Threshold Phenomena and Influence with Some Perspectives from Mathematics, Computer Science, and Economics

Gil Kalai
Shmuel Safra

1 INTRODUCTION

Threshold phenomena refer to settings in which the probability for an event to occur changes rapidly as some underlying parameter varies. Threshold phenomena play an important role in probability theory and statistics, physics, and computer science, and are related to issues studied in economics and political science. Quite a few questions that come up naturally in those fields translate to proving that some event indeed exhibits a threshold phenomenon, and then finding the location of the transition and how rapid the change is. The notions of sharp thresholds and phase transitions originated in physics, and many of the mathematical ideas for their study came from mathematical physics. In this chapter, however, we will mainly discuss connections to other fields.

A simple yet illuminating example that demonstrates the sharp threshold phenomenon is Condorcet's Jury Theorem (CJT), which can be described as follows. Say one is running an election process, where the results are determined by simple majority, between two candidates, Alice and Bob. If every voter votes

Computational Complexity and Statistical Physics,
edited by Allon Percus, Gabriel Istrate, and Cristopher Moore, Oxford University Press.

for Alice with probability $p > 1/2$ and for Bob with probability $1 - p$, and if the probabilities for each voter to vote either way are independent of the other votes, then as the number of voters tends to infinity the probability of Alice getting elected tends to 1. The probability of Alice getting elected is a monotone function of p, and when there are many voters it rapidly changes from being very close to 0 when $p < 1/2$ to being very close to 1 when $p > 1/2$.

The reason usually given for the interest of CJT to economics and political science is that it can be interpreted as saying that even if agents receive very poor (yet independent) signals, indicating which of two choices is correct, majority voting nevertheless results in the correct decision being taken with high probability, as long as there are enough agents, and the agents vote according to their signal. This is referred to in economics as *asymptotically complete aggregation of information.*

Condorcet's jury theorem is a simple consequence of the weak law of large numbers. The central limit theorem implies that the "threshold interval" is of length proportional to $1/\sqrt{n}$. Some extensions, however, are much more difficult. When we consider general economic or political situations, the aggregation of agents' votes may be much more complicated than a simple majority. The individual signal (or signals) may be more complicated than a single bit of information, the distribution of signals among agents can be more general and, in particular, agents' signals may depend on each other. On top of that, voters may vote strategically by taking into account the possible actions of others in addition to their own signal, and distinct voters may have different goals and interests, not only different information. In addition, the number of candidates may be larger than two, resulting in a whole set of new phenomena.

Let us now briefly mention two other areas in which threshold behavior emerges. The study of random graphs as a separate area of research was initiated in the seminal paper of Erdös and Rényi [29] from 1959. Consider a random graph $G(n, p)$ on n vertices where every edge among the $\binom{n}{2}$ possible edges appears with probability p. Erdös and Rényi proved a sharp threshold property for various graph properties. For example, for every $\epsilon > 0$, if $p = (1 + \epsilon) \log n/n$ the graph is connected with probability tending to 1 (as n tends to infinity) while for $p = (1 - \epsilon) \log n/n$ the probability that the graph will be connected tends to zero. Since the time of their work, extensive studies of specific random graph properties have been carried out and, in recent years, results concerning the threshold behavior of general graph properties have been found. For a general understanding of the threshold properties of graphs, *symmetry* plays a crucial role: when we talk about properties of graphs we implicitly assume that those properties depend only on the isomorphism type of the graphs, and not on the labeling of vertices. This fact introduces substantial symmetry to the model. We will discuss how to exploit this symmetry.

Next, we mention complexity theory. Threshold phenomena play a role, both conceptual and technical, in various aspects of computational complexity theory. One of the major developments in complexity theory in the last two decades

is the emerging understanding of the complexity of approximating optimization problems. Here is an important example: for a graph G let $m(G)$ be the maximum number of edges between two disjoint sets of vertices of G. MAX-CUT, the problem of dividing the vertices of a given input graph into two parts so as to maximize the number of edges between the parts, is known to be NP-hard. However, simply finding a partition such that the number of edges between the two parts is at least $m(G)/2$ is easy. The emerging yet unproven picture for this problem is that if we wish to find a partition of the vertices with at least $cm(G)$ edges between the parts then there is a critical value c_0 such that the problem is easy (there is a randomized polynomial time algorithm to solve it) for $c < c_0$ and hard (likely NP-hard) for $c > c_0$. For MAX-CUT, the critical value $c_0 = 0.878567\ldots$ is reached by the famous Goemans-Williamson algorithm [39] based on semidefinite programming. More generally, for many other problems we can expect a sharp threshold between the region where approximation is easy and the region where approximation is hard. In addition, the study of threshold phenomena and other related properties of Boolean functions is an important technical tool in understanding the hardness of approximation.

Another connection with complexity theory occurs in the area of circuit complexity. It turns out that Boolean functions in very "low" complexity classes necessarily exhibit coarse threshold behavior. For example, the majority function that exhibits a very sharp threshold behavior cannot be represented by a bounded-depth Boolean circuit of small size. This insight is related to another major success of complexity theory: lower bounds for the size of bounded-depth circuits.

Let us now explicitly define the basic mathematical object that is the subject of our considerations. A *Boolean function* is a function $f(x_1, x_2, \ldots, x_n)$ where each variable x_i is a Boolean variable, taking the value 0 or 1. The value of f is also 0 or 1. A Boolean function f is monotone if $f(y_1, y_2, \ldots, y_n) \geq f(x_1, x_2, \ldots, x_n)$ when $y_i \geq x_i$ for every i. Some basic examples of Boolean functions are named after the voting method they describe. For an odd integer n, the *majority* function $M(x_1, x_2, \ldots, x_n)$ equals 1 if and only if $x_1+x_2+\ldots+x_n > n/2$. The *dictatorship* function is $f(x_1, x_2, \ldots, x_n) = x_i$. *Juntas* refer to the class of Boolean functions that depend on a bounded number of variables, namely functions that disregard the value of almost all variables except for a few, whose number is independent of n.

Now consider the probability $\mu_p(f)$ that $f(x_1, x_2, \ldots, x_n) = 1$, when the probability that $x_i = 1$ is p, independently for $i = 1, 2, \ldots, n$, just as we had earlier for the election between Alice and Bob. When f is a monotone Boolean function, the function $\mu_p(f)$ is a monotone real function of p. Given a real number $1/2 > \epsilon > 0$, the *threshold interval* depending on ϵ is the interval $[p_1, p_2]$ where $\mu_{p_1}(f) = \epsilon$ and $\mu_{p_2}(f) = 1-\epsilon$. Understanding the length of this threshold interval is one of our central objectives.

Before we describe this chapter's sections it is worth noting that the notion of a sharp threshold is an asymptotic property and therefore it applies to a sequence

of Boolean functions when the number of variables becomes large. Giving explicit, realistic, and useful estimates is an important goal. In the election example above, the central limit theorem provides explicit, realistic, and useful estimates. In more involved settings, however, this task can be quite difficult.

The main messages of this chapter can be summarized as follows:

- The threshold behavior of a system is intimately related to combinatorial notions of "influence" and "pivotality" (section 2).
- Sharp thresholds are common. We can expect a sharp threshold unless there are good reasons not to (section 3 and 5.3).
- A basic mathematical tool in understanding threshold behavior is Fourier analysis of Boolean functions (section 4).
- Higher symmetry leads (in a subtle way) to sharper threshold behavior (section 5.2).
- Sharp thresholds occur unless the property can be described "locally" (section 5.3).
- Systems whose description belongs to a very low complexity class have rather coarse (not sharp) threshold behavior (section 6.1).
- In various optimization problems, when we seek approximate solutions, there is a sharp transition between goals that are algorithmically easy and those that are computationally intractable (section 6.3).

In section 2 we introduce the notions of pivotality and influence and discuss *Russo's lemma*, which relates these notions to threshold behavior. In section 3 we describe basic results concerning influences and threshold behavior of Boolean functions. In section 4 we discuss a major mathematical tool required for the study of threshold phenomena and influences: Fourier analysis of Boolean functions. In section 5 we discuss the connection to random graphs and hypergraphs and to the k-SAT problem. In section 6 we discuss the connections to computational complexity. Section 7 is devoted to the related phenomenon of noise sensitivity. Section 8 discusses connections with the model of percolation. Section 9 discusses an example from social science: a result by Feddersen and Pesendorfer that exhibits a situation of self-organized criticality. Section 10 concludes with some of the main open problems and challenges.

2 PIVOTALITY, INFLUENCE, POWER, AND THE THRESHOLD INTERVAL

In this section we describe the n-dimensional hypercube, and define the notions of "pivotal" variables and influence for Boolean functions. We state Russo's fundamental lemma connecting influences and thresholds.

2.1 THE DISCRETE CUBE

Let $\Omega_n = \{0,1\}^n$ denote the discrete n-dimensional cube, namely, the set of 0-1 vectors with n entries. A Boolean function is a map from Ω_n to $\{0,1\}$. Boolean functions on Ω_n are of course in 1-1 correspondence with subsets of Ω_n. Elements in Ω_n are themselves in 1-1 correspondence with subsets of $[n] = \{1,2,\ldots,n\}$. Boolean functions appear under different names in many areas of science. We will equip Ω_n with a metric, namely a distance function, and a probability measure. For $x, y \in \Omega_n$ the Hamming distance $d(x,y)$ is defined by

$$d(x,y) = |\{i : x_i \neq y_i\}| . \tag{1}$$

Denote by $\Omega_n(p)$ the discrete cube endowed with the product probability measure μ_p, where $\mu_p(\{x : x_j = 1\}) = p$. In other words,

$$\mu_p(x_1, x_2, \ldots, x_n) = p^k (1-p)^{n-k} , \tag{2}$$

where $k = x_1 + x_2 + \ldots + x_n$.

2.2 PIVOTALITY AND INFLUENCE OF VARIABLES

Consider a Boolean function $f(x_1, x_2, \ldots, x_n)$ and the associated event $A \subset \Omega_n(p)$, such that $f = \chi_A$, namely that f is the indicator function of A. For $x = (x_1, x_2, \ldots, x_n) \in \Omega_n$ we say that the kth variable is *pivotal* if flipping the value of x_k changes the value of f. Formally, let

$$\sigma_k(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_{k-1}, 1 - x_k, x_{k+1}, \ldots, x_n) \tag{3}$$

and define the kth variable to be pivotal at x if

$$f(\sigma_k(x)) \neq f(x) . \tag{4}$$

The *influence* of the kth variable on a Boolean function f, denoted by $I_k^p(f)$, is the probability that the kth variable is pivotal, that is,

$$I_k^p(f) = \mu_p(\{x : f(\sigma_k(x)) \neq f(x)\}) \tag{5}$$

The influence of a variable in a Boolean function and more general notions of influences were introduced by Ben-Or and Linial [11] in the context of "collective coin-flipping."

The total influence $I^p(f)$ is the sum of the individual influences.

$$I^p(f) = \sum_{k=1}^{n} I_k^p(f) . \tag{6}$$

We omit the superscript p for $p = 1/2$. For a monotone Boolean function thought of as an election method, $I_k(f)$ $(= I_k^{1/2}(f))$ is referred to as the Banzhaf

power index of voter k. The quantity

$$\phi_k(f) = \int_0^1 I_k^p(f)dp\,, \tag{7}$$

is called the Shapley-Shubik power index of voter k.

The mathematical study (under different names) of pivotal agents and influences is quite basic in percolation theory and statistical physics, as well as in probability theory and statistics, reliability theory, distributed computing, complexity theory, game theory, mechanism design and auction theory, other areas of theoretical economics, and political science.

2.3 RUSSO'S LEMMA AND THRESHOLD INTERVALS

A Boolean function f is monotone if its value does not decrease when we flip the value of any variable from 0 to 1. For a monotone Boolean function $f \subset \Omega_n$, let $\mu_p(f)$ be the probability that $f(x_1, \ldots, x_n) = 1$ with respect to the product measure μ_p. Note that $\mu_p(f)$ is a monotone function of p. Russo's fundamental lemma [41, 82] asserts that

$$\frac{d\mu_p(f)}{dp} = I^p(f)\,. \tag{8}$$

Suppose now that f is a non-constant monotone Boolean function. Given a small real number $\epsilon > 0$, let p_1 be the unique real number in $[0,1]$ such that $\mu_{p_1}(f) = \epsilon$ and let p_2 be the unique real number such that $\mu_{p_2}(f) = 1 - \epsilon$. The interval $[p_1, p_2]$ is called a *threshold interval* and its length $p_2 - p_1$ is denoted by $t_\epsilon(f)$. Denote by p_c the value satisfying $\mu_{p_c}(f) = 1/2$, and call it the critical probability of the event A.

By Russo's lemma, a large total influence around the critical probability implies a short threshold interval.

Remark: Let us now exhibit the notions introduced here using a simple example. We will return to this example to demonstrate several issues discussed in the chapter. Let M_3 represent the majority function on three variables. Thus, $M_3(x_1, x_2, x_3) = 1$ if $x_1 + x_2 + x_3 \geq 2$ and $M_3(x_1, x_2, x_3) = 0$ otherwise. Clearly, $\mu(M_3) = 1/2$. This follows from the fact that M_3 is an *odd* Boolean function, namely one that satisfies the relation

$$f(1 - x_1, 1 - x_2, \ldots, 1 - x_n) = 1 - f(x_1, x_2, \ldots, x_n)\,. \tag{9}$$

A simple calculation gives, for general p,

$$\mu_p(M_3) = p^3 + 3p^2(1 - p)\,. \tag{10}$$

As for the influence of the variables, we obtain $I_k(M_3) = 1/2$ and $I_k^p(M_3) = 2p(1-p)^2 + 2p^2(1-p)$ for $k = 1, 2, 3$. Therefore, $I(M_3) = 3/2$ and $I^p(M_3) = 6(p(1-p))$, which is indeed equal to $d\mu_p(M_3)/dp$.

3 BASIC RESULTS ON INFLUENCES AND THRESHOLD BEHAVIOR OF BOOLEAN FUNCTIONS

Some basic facts on influences and the corresponding results on threshold intervals are as follows. Dictatorships and juntas have small total influence, and thus coarse thresholds. Conversely, when the critical probability is 1/2, a coarse threshold implies that the function "looks like" a junta. These results are formalized as follows.

3.1 THE TOTAL INFLUENCE CANNOT BE OVERLY SMALL

Theorem 3.1. *For every Boolean function f,*

$$I(f) \geq 2\mu(f)\log_2(1/\mu(f))\,. \tag{11}$$

In particular, if $\mu_{1/2}(f) = 1/2$ then $I(f) \geq 1$ and equality holds if and only if f is a dictatorship, namely $f(x_1, \ldots, x_n) = x_i$ for some i, or an "antidictatorship," $f(x_1, \ldots, x_n) = 1 - x_i$ for some i. Inequality (11) has its origins in the works of Whitney and Loomis, Harper, Bernstein, Hart, and others. It is of great importance in many mathematical contexts. Inequality (11) is often referred to as the *edge-isoperimetric inequality*. It can be regarded as an isoperimetric relation for subsets of the discrete cube, analogous to the famous Euclidean isoperimetric relations. This analogy goes a long way, and we will return to it in section 5.4. Ledoux's book [67] is an excellent source for the related phenomenon of "measure concentration."

An upper bound for the length of the threshold interval can be derived from the bounds on the sum of influences combined with Russo's lemma.

Theorem 3.2 (Bollobás and Thomason[16]). *For every monotone Boolean function f,*

$$t_\epsilon(f) = O(\min(p_c, 1 - p_c))\,. \tag{12}$$

Two brief remarks are in order. First, note that for a function $f(x_1, x_2, \ldots, x_n)$ we can consider the "dual" function defined by

$$g(x_1, x_2, \ldots, x_n) = f(1 - x_1, 1 - x_2, \ldots, 1 - x_n)\,. \tag{13}$$

Then it is easily seen that

$$\mu_p(g) = 1 - \mu_{1-p}(f)\,. \tag{14}$$

Due to this duality we may, without loss of generality, restrict ourselves to the case where $p_c(f) \leq 1/2$, which will simplify several of the statements below. Second, note that another way to state the Bollobás-Thomason result is that for

every Boolean function f and every $\epsilon > 0$ there exists a value $c(\epsilon)$ such that $t_\epsilon(f)/p_c(f) \leq c(\epsilon)$.

Theorem 3.2 is the basis for the following definition: we say that a sequence (f_n) of Boolean functions has a *sharp threshold* if for every $\epsilon > 0$,

$$t_\epsilon(f_n) = o(\min(p_c, 1 - p_c)) \,. \tag{15}$$

Otherwise, we say that the sequence demonstrates a *coarse threshold* behavior. When the critical probabilities for the functions f_n are bounded away from 0 and 1 then having a sharp threshold simply means that for every $\epsilon > 0$, $t_\epsilon(f_n) = o(1)$.

3.2 SIMPLE MAJORITY MAXIMIZES THE TOTAL INFLUENCE OF MONOTONE BOOLEAN FUNCTIONS

Let n be an odd integer. Denote by M_n a simple majority function on n variables.

Proposition 3.3. *Let f be a monotone Boolean function over n variables, n odd, and with $p_c(f) = 1/2$. Then for every p, $0 < p < 1$,*

$$I^p(f) \leq I^p(M_n) \,. \tag{16}$$

See, for example, lemma 6.1 of Friedgut and Kalai [34] and Chayes et al. [24]. By Russo's lemma it follows that:

Proposition 3.4. *Let f be a monotone Boolean function over n variables, n odd, and with $p_c(f) = 1/2$. Then, for every $p > 1/2$, $\mu_p(M_n) \geq \mu_p(f)$.*

3.3 NOT ALL INDIVIDUAL INFLUENCES CAN BE SMALL

Theorem 3.5 (Kahn-Kalai-Linial [53]). *There exists a universal constant K such that for every Boolean function f,*

$$\max_k I_k(f) \geq K \min(\mu(f), 1 - \mu(f)) \log n/n \,. \tag{17}$$

This theorem answered a question posed by Ben-Or and Linial [11], who gave an example of a Boolean function f with $\mu(f) = 1/2$ and $I_k(f) = \Theta(\log n/n)$. Note that theorem 3.5 implies that when all individual influences are the same, that is, when A is invariant under the induced action from a transitive permutation group on $[n]$, then the total influence is at least $K \min(\mu(f), 1 - \mu(f)) \log n$. An extension for arbitrary product probability spaces was found by Bourgain, Kahn, Kalai, Katznelson, and Linial [23]. Talagrand [92] extended the result of Kahn, Kalai, and Linial in various directions and applied these results for studying threshold behavior. Talagrand also presented a very useful extension for arbitrary real functions on the discrete cube. Talagrand's extension for the product measure μ_p is stated as follows:

Theorem 3.6 (Talagrand [92]). *There exists a universal constant K such that for every Boolean function f,*

$$\sum_{k=1}^{n} \frac{I_k^p(f)}{\log 1/I_k^p(f)} \geq K \frac{\mu_p(f)(1-\mu_p(f))}{\log 2/(p(1-p))}\,. \tag{18}$$

Our next result Friedgut [35] describes Boolean functions with a small total influence.

Theorem 3.7 (Friedgut). *Let f be a monotone Boolean function. For every $0 < z \leq 1/2$, $a \geq 1$ and $\gamma > 0$, there exists a value $C = C(z, a, \gamma)$ such that if $z \leq p \leq 1 - z$ and $I^p(f) \leq a$, then there is a monotone Boolean function g depending on at most C variables, such that*

$$\mu_p(\{x \in \Omega_n : f(x) \neq g(x)\}) \leq \gamma\,. \tag{19}$$

Theorem 3.7 asserts that if the critical probability is bounded away from 0 and 1 and the threshold is coarse, then for most values of p in the threshold interval, f can be approximated by a junta with respect to the probability measure μ_p. Note that when p tends to zero with increasing n, the size of the junta is no longer bounded; when p tends to zero as a fractional power of $1/n$, the theorem carries no information. We will return to this important range of parameters later.

Likewise, if no one influence is unduly large then the threshold is sharp, as demonstrated by the following.

Theorem 3.8 (Russo-Talagrand-Friedgut-Kalai). *Let f be a Boolean function. For every $0 < z \leq 1/2$, $\epsilon > 0$ and $\gamma > 0$, there exist values $\delta_i = \delta_i(z, \epsilon, \gamma) > 0$, $i = 1, 2, 3$ such that if $z \leq p_c(f) \leq 1 - z$, then any of the following conditions implies that*

$$t_\epsilon(f) < \gamma.$$

1. *For every k, $1 \leq k \leq n$, and for every p, $0 < p < 1$, $I_k^p(f) \leq \delta_1$ [34, 81, 92].*
2. *For every k, $1 \leq k \leq n$, and for p such that $\epsilon < \mu_p(f) < 1-\epsilon$ (e.g., $p = p_c(f)$), $I_k^p(f) < \delta_2$ [57].*
3. *For every k, $1 \leq k \leq n$, the Shapley-Shubik power index $\phi_k(f) \leq \delta_3$ [57] .*

Part (1) of the theorem was proved by Russo [81]. A sharp version was proved by Talagrand [92] and Friedgut and Kalai [34] based on the Kahn-Kalai-Linial theorem and its extensions.

Parts (2) and (3) are based on Friedgut's result and some additional observations, and are derived in Kalai [57], but the values of δ_2, δ_3 are rather weak (doubly logarithmic in γ). It would be interesting to find better bounds. Part (3) in the theorem above is, in fact, a characterization:

Theorem 3.9. *Let (f_n) be a sequence of monotone Boolean functions. For every $\epsilon > 0$,*

$$\lim_{n\to\infty} t_\epsilon(f_n) = 0$$

if and only if the maximal Shapley-Shubik power index for f_n tends to zero [57].

4 FOURIER ANALYSIS OF BOOLEAN FUNCTIONS

In this section we describe an important mathematical tool in the study of threshold phenomena and in various related areas. The material described here is not essential for reading most of the remaining sections, and so the reader who wishes to skip this section may safely do so. But as the topic is central to many of the mathematical results presented in this chapter, we feel it is important familiarize the reader with it at this early stage.

4.1 ALL THE WAY TO PARSEVAL

Let Ω_n denote the set of 0-1 vectors $(x_1, \ldots, x_n)$ of length n. Let $L_2(\Omega_n)$ denote the space of real functions on Ω_n, endowed with the inner product

$$\langle f, g\rangle = \sum_{(x_1,x_2,\ldots,x_n)\in\Omega_n} 2^{-n} f(x_1, \ldots, x_n) g(x_1, \ldots, x_n)\,. \tag{20}$$

The inner product space $L_2(\Omega_n)$ is 2^n-dimensional. The L_2-norm of f is defined by

$$\|f\|_2^2 = \langle f, f\rangle = \sum_{(x_1,x_2,\ldots,x_n)\in\Omega_n} 2^{-n} f^2(x_1, x_2, \ldots, x_n)\,. \tag{21}$$

Note that if f is a Boolean function, then $f^2(x)$ is either 0 or 1 and therefore $\|f\|_2^2 = \sum_{(x_1,\ldots,x_n)\in\Omega_n} 2^{-n} f^2(x)$ is simply the probability $\mu(f)$ that $f = 1$ (with respect to the uniform probability distribution on Ω_n). If the Boolean function f is odd (i.e., satisfying relation (9)) then $\|f\|_2^2 = 1/2$.

For a subset S of $[n]$ consider the function

$$u_S(x_1, x_2, \ldots, x_n) = (-1)^{\sum_{i\in S} x_i}\,. \tag{22}$$

It is not difficult to verify that the 2^n functions u_S for all subsets S form an orthonormal basis for the space of real functions on Ω_n.

For a function $f \in L_2(\Omega_n)$, the Fourier-Walsh coefficient $\hat{f}(S)$ of f is

$$\hat{f}(S) = \langle f, u_S\rangle\,. \tag{23}$$

Since the functions u_S form an orthogonal basis, it follows that

$$\langle f, g\rangle = \sum_{S\subset[n]} \hat{f}(S)\hat{g}(S)\,. \tag{24}$$

In particular,

$$\|f\|_2^2 = \sum_{S\subset[n]} \hat{f}^2(S)\,. \tag{25}$$

This last relation is called Parseval's formula.

Remark: To demonstrate the notions introduced here we return to our example. Let M_3 represent the majority function on three variables. The Fourier coefficients of M_3 are easy to compute: $\hat{M_3}(\varnothing) = \sum(1/8)M_3(x) = 1/2$. In general, if f is a Boolean function then $\hat{f}(\varnothing)$ is the probability that $f(x) = 1$ and when f is an odd Boolean function, $\hat{f}(\varnothing) = 1/2$. Next, $\hat{M_3}(\{1\}) = 1/8(M_3(0,1,1) - M_3(1,0,1) - M_3(1,1,0) - M_3(1,1,1)) = (1-3)/8$ and thus $\hat{M_3}(\{j\}) = -1/4$, for $j = 1,2,3$. Next, $\hat{M_3}(S) = 0$ when $|S| = 2$ and finally $\hat{M_3}(\{1,2,3\}) = 1/8(M_3(1,1,0) + M_3(1,0,1) + M_3(0,1,1) - f(1,1,1)) = 1/4$.

4.2 THE RELATION WITH INFLUENCES

It is surprising how far one can get with the simple base-change of the Fourier-Walsh transform and Parseval's formula. The relation between influences and Fourier coefficients is given by the following expressions, whose proof is elementary:

$$I_k(f) = 4\sum_{S:k\in S} \hat{f}^2(S)\,. \tag{26}$$

$$I(f) = 4\sum_{S\subset[n]} \hat{f}^2(S)|S|\,. \tag{27}$$

If f is monotone we also have $I_k(f) = -2\hat{f}(\{k\})$.

The following notation is useful:

$$W_k(f) = \sum_{S:|S|=k} \hat{f}^2(S)\,, \tag{28}$$

allowing us to rewrite relation (27) as $I(f) = 4\sum_{k\geq 0} kW_k(f)$.

To practice these notions, observe that $\hat{f}(\varnothing) = \|f\|_2^2 = \mu(f)$, so from Parseval's formula, $\sum_{S\subset[n],S\neq\varnothing} \hat{f}^2(S) = \mu(f)(1-\mu(f))$. It follows from eq. (27) that

$$I(f) \geq 4\mu(f)(1-\mu(f))\,. \tag{29}$$

If one considers a Boolean function f where $\mu(f) = 1/2$, $I(f) \geq 1$. This is an important special case of the edge-isoperimetric inequality (11).

4.2.1 Remark: Indeed, for our example M_3 we have

$$3/2 = I(M_3) = 4 \sum_{S \subset [n]} \hat{M}_3^2(S)|S| = 4(3(1/16) + (1/16)3)\,.$$

4.3 BERNOULLI MEASURES

When we consider the probability distribution μ_p, we have to define the inner product by

$$\langle f, g \rangle = \sum_{(x_1,x_2,\ldots,x_n)\in\Omega_n} f(x_1,\ldots,x_n)g(x_1,\ldots,x_n)\mu_p(x_1,\ldots,x_n). \tag{30}$$

We need an appropriate generalization for the Walsh-Fourier orthonormal basis for general Bernoulli probability measures μ_p. Those are given by

$$u_S^p(x_1,x_2,\ldots,x_n) = \left(-\sqrt{\frac{1-p}{p}}\right)^{\sum_{i\in S} x_i} \left(\sqrt{\frac{p}{1-p}}\right)^{n-\sum_{i\in S} x_i}. \tag{31}$$

Let p be a fixed real number, $0 < p < 1$. Every real function f on Ω_n can be expanded to

$$f = \sum_{S\subset[n]} \hat{f}(S;p)u_S^p\,,$$

where

$$\hat{f}(S;p) = \sum_{x\in\Omega_n} f(x)u_S^p(x)\mu_p(x)\,.$$

The relations with influences also extend as follows:

$$p(1-p)I_k^p(f) = \sum_{S:k\in S} \hat{f}^2(S;p)\,, \tag{32}$$

$$I^p(f) = \frac{1}{p}\frac{1}{1-p}\sum_{S\subset[n]} \hat{f}^2(S)|S|\,. \tag{33}$$

Exercise: Compute the coefficients $\hat{M}_3(S,p)$ and verify eq. (33) for the case of M_3.

4.4 THE BONAMIE-GROSS-BECKNER RELATION

The reader who did not skip this whole section may still wish to skip this subsection. We will consider here a technical inequality that will not be explicitly mentioned again in the chapter, but nevertheless underlies many of the proofs

and results. There are many ways of viewing the inequality, and its remarkable effectiveness remains somewhat mysterious. We will present the "simplest" application of it that we know.

For a real function $f : \Omega_n \to \mathcal{R}$, $f = \sum \hat{f}(S)u_S$, define the L_w-norm of a function f to be

$$\|f\|_w = \left(\sum_{x \in \Omega_n} 2^{-n} |f(x)|^w \right)^{1/w} . \tag{34}$$

Note that, due to the normalization coefficient 2^{-n} in the definition, if $1 \leq v < w$ then

$$\|f\|_v \leq \|f\|_w . \tag{35}$$

Next define the operator

$$T_\rho(f) = \sum_{S \subset [n]} \hat{f}(S)\rho^{|S|}u_S , \tag{36}$$

so that

$$\|T_\rho(f)\|_2^2 = \sum_{S \subset [n]} \hat{f}^2(S)\rho^{2|S|} . \tag{37}$$

The Bonamie-Gross-Beckner (BGB) inequality [7, 18, 43] asserts that for every real function f on Ω_n,

$$\|T_\rho(f)\|_2 \leq \|f\|_{1+\rho^2} . \tag{38}$$

Because this inequality involves two different norms, it is referred to as *hypercontractive* [42]. The inequality can be regarded as an extension of the Khintchine inequality [62], which states that the different L_w-norms of functions of the form $\sum_k \alpha_k u_{\{k\}}$ differ only by absolute multiplicative constants. Beckner used this inequality in the early 1970s to handle classical problems in harmonic analysis. The work was influenced by earlier hypercontractive inequalities by Nelson and others, originating in the mathematical study of quantum field theory [43, 74].

Here is a quick and sketchy argument giving a flavor of the use of the Bonamie-Gross-Beckner inequality. Note that for a Boolean function f and every $w \geq 1$,

$$\|f\|_w^w = \mu(f) . \tag{39}$$

Let $0 < \rho < 1$. Now, if a large portion of the L_2-norm of f is concentrated at "low frequencies" $|S|$, then $\|T_\rho(f)\|_2$ will not be too much smaller than $\|f\|_2$. The BGB inequality implies that in this case, $\|f\|_{1+\rho^2}$ cannot be too much smaller than $\|f\|_2$ either. This fact, however, cannot coexist with eq. (39) if $\mu(f)$ is sufficiently small.

More formally, suppose that $\mu(f) = s \leq 1/2$, and we will try to give lower bounds for $I(f)$. In section 4.2 we derived from Parseval's formula that $I(f) \geq 4(s - s^2)$. The edge-isoperimetric inequality (eq. (11)) asserts that $I(f) \geq$

$2s\log_2(1/s)$. Let us try to understand the appearance of $\log(1/s)$. Take $\rho = 1/2$ and thus $1 + \rho^2 = 5/4$. The BGB inequality and eq. (39) give

$$\sum \frac{\hat{f}^2(S)}{2^{2|S|}} \leq \|f\|_{5/4}^2 = s^{1+3/5}.$$

Noting that $2^{2|S|} < 1/\sqrt{s}$ for $0 < |S| < \log_2(1/s)/4$,

$$\sum_{0<|S|<\log(1/s)/4} \hat{f}^2(S) \leq \sqrt{s}s^{3/5} \leq K\sqrt{s(1-s)}$$

for some constant $K < 1$, since $s \leq 1/2$. This implies that a finite fraction of the L_2 norm of f is concentrated at Fourier coefficients $\hat{f}(S)$ where $|S| \geq K' \log(1/s)$. It then follows from the discussion in section 4.2 that $I(f) \geq K''(\mu(f)(1 - \mu(f))\log(1/\mu(f))$. Up to a multiplicative constant this gives the fundamental edge-isoperimetric relation (eq. (11)), but the information on Fourier coefficients, while not sharp, is even stronger.

An extension of the BGB inequality for general p can be found in Talagrand [92].

REMARKS

- The Fourier coefficients of Boolean functions are tailor-made to deal with the total influence that by Russo's lemma gives the "local" threshold behavior. However, to understand the behavior in the entire threshold interval, a further understanding of the relation between the behavior at different points is required. For a global understanding of influences over the entire threshold interval, the quantities $\int_0^1 \hat{f}(S,p)dp$ may play a role: it would be interesting to study them.
- This section is only a taste of a rather young field of Fourier analysis of Boolean functions which has many connections, extensions, applications, and problems. We hope to be able to give a fuller treatment elsewhere.

5 FROM ERDÖS AND RÉNYI TO FRIEDGUT: RANDOM GRAPHS AND THE K-SAT PROBLEM

5.1 GRAPH PROPERTIES AND BOOLEAN FUNCTIONS

Another origin for the study of threshold phenomena in mathematics is random graph theory and, particularly the seminal works by Erdös and Rényi [29]. Some good references on random graphs are Alon and Spencer [2], Bollobás [13], and Janson et al. [51].

Consider a graph $G = (V, E)$, where V is the set of vertices and E is the set of edges. Let $x_1, x_2, \ldots, x_{|E|}$ be Boolean variables corresponding to the edges

of G. An assignment of the values 0 and 1 to the variables x_i corresponds to a subgraph $H \subseteq G$, where $H = (V, E')$ and $e \in E'$ if and only if $x_e = 1$. We will mostly consider the case where G is the complete graph, namely, $E = \binom{V}{2}$.

This basic Boolean representation of subgraphs (or substructures for other structures) is very important. A graph property P is a property of graphs that does not depend on the labeling of the vertices. In other words, P depends only on the isomorphism type of G. The property is monotone if when a graph H satisfies it, every graph G on the same vertex set obtained by adding edges to H also satisfies the property. Examples include: "the graph is connected," "the graph is not planar" (a graph is planar if it can be drawn in the plane without crossings), "the graph contains a triangle," and "the graph contains a Hamiltonian cycle." Understanding the threshold behavior of monotone graph properties for random graphs was the main motivation behind the theorem of Bollobás and Thomason ([16], theorem 3.2). Their result applies to arbitrary monotone Boolean functions, so it does not rely on the symmetry that Boolean functions representing graph properties have.

Theorem 5.1 (Friedgut and Kalai [34]). *For every monotone property P of graphs, there exists a constant C such that*

$$t_\epsilon(P) \leq C \log(1/\epsilon)/\log n\,. \tag{40}$$

Theorem 5.1, which answered a question suggested by Nati Linial, is a simple consequence of the Kahn-Kalai-Linial theorem and its extensions combined with Russo's lemma. The crucial observation is that all influences of variables are equal for Boolean properties defined by graph properties. As a matter of fact, this continues to be true for Boolean functions f describing random subgraphs of an arbitrary edge-transitive graph.[1] All influences being equal implies that the total influence $I^p(f)$ is at least as large as $K \min(\mu_p(f), 1-\mu_p(f)) \log n$. By Russo's lemma, this gives the required result.

Friedgut and Kalai [34] raised several questions that were addressed in later works:

- What is the relation between the group of symmetries of a Boolean function and its threshold behavior?
- What would guarantee a sharp threshold when the critical probability p_c tends to zero with increasing n?
- What is the relation between influences, the threshold behavior, and other isoperimetric properties of f?

We will describe in some detail the work of Bourgain and Kalai [22] on the first question and the works of Friedgut [36] and Bourgain [21] on the second.

[1] A graph is edge-transitive if for every two edges e and e' there is an automorphism of the graph that maps e to e'.

The last question was addressed by several papers of Talagrand [90, 91] and also Benjamini et al. [10], but we will not elaborate on it here.

Let us make one comment at this point. When we consider the Fourier coefficients $\hat{f}(S)$ of a Boolean function representing a graph property then the set S, which can be regarded as a subset of the variables, also represents a graph. As mentioned above, being a graph property implies large symmetry for the original Boolean function: it is invariant under permutations of the variables that correspond to permutations of the vertices of the graph. The same is true for the Fourier coefficients: the Fourier coefficient $\hat{f}(S)$ depends only on the isomorphism type of the graph described by the set S. This is a crucial observation for the results that follow.

5.2 THRESHOLD UNDER SYMMETRY

We now describe a measure of symmetry that is related to the threshold behavior. The key intuition is that the more symmetry we have, the sharper the threshold behavior we observe. The measure of symmetry is based on the size of orbits.

A graph property for graphs with n' vertices is described by a Boolean function on $n = \binom{n'}{2}$ variables. Such Boolean functions are invariant under the induced action of the symmetric group $S_{n'}$ on the vertices, namely the group of all permutations of the vertices, acting on the edges. (Note that the variables of f correspond to the n edges of the complete graph on n' vertices.) In the previous section we used this symmetry to argue that all individual influences are the same. Here we would like to exploit further the specific symmetry in the situation at hand.

Bourgain and Kalai [22] studied the effect of symmetry on the threshold interval, leading to the following result:

Theorem 5.2 (Bourgain and Kalai). *For every monotone property P of graphs with n' vertices, and every $\tau > 0$, there exists a value $C(\tau)$ such that*

$$t_\epsilon(P) \leq C(\tau) \log(1/\epsilon)/(\log n')^{2-\tau}. \tag{41}$$

It is conjectured that the theorem continues to hold for $\tau = 0$. Let Γ be a group of permutations of $[n]$. Thus Γ is a subgroup of the group of all $n!$ permutations of $[n]$. The group Γ acts on Ω_n as follows:

$$\pi(x_1, x_2, \ldots, x_n) = (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}),$$

for $\pi \in \Gamma$. A Boolean function is Γ-invariant if $f(\pi(x)) = f(x)$ for every $x \in \Omega_n$ and every $\pi \in \Gamma$. We would like to understand the influences and threshold behavior of Boolean functions that are Γ-invariant.

We now describe certain parameters of Γ that depend on the size of the orbits in the action of Γ on subsets of $[n]$. Divide the discrete hypercube Ω_n

into layers: write Ω_n^m for the vectors in Ω_n with exactly m 1's. For a group Γ of permutations of $[n]$, let $T(m)$ denote the number of orbits in the induced action of Γ on Ω_n^m and let $B(m)$ be the smallest size of an orbit of Γ acting on Ω_n^m. For graph properties, $T(m)$ is the number of isomorphism types of graphs with n' vertices and m edges, and $B(m)$ is the minimum number of (labeled) graphs with n' vertices and m edges that are isomorphic to a specific graph H. The number of graphs isomorphic to H is $n'!/|\mathrm{Aut}(H)|$, where $\mathrm{Aut}(H)$ denotes the automorphism group of H.

When we consider graph properties for graphs with n' vertices, $B(m)$ grows as $\binom{n'}{\sqrt{m}}$. To see this, note that when $m = \binom{s}{2}$ for some $s \leq n'$, graphs H with the fewest isomorphic copies (hence with the largest automorphism groups) are complete graphs on s vertices, leading to $B(m) = \binom{n'}{s}$.

Define the parameter $\kappa(\Gamma)$ as follows:

$$\kappa(\Gamma) = \min\{m : B(m) < 2^m\}\,. \tag{42}$$

Since greater symmetry leads to smaller $B(m)$, $\kappa(\Gamma)$ measures the "size" of the group of symmetries.

Define also for $\tau > 0$:

$$\kappa_\tau(\Gamma) = \min\{m : B(m) < 2^{m^\tau}\}\,. \tag{43}$$

Bourgain and Kalai showed that for every $\tau > 0$ the total influence $I^p(f)$ of a Γ-invariant Boolean function f satisfies the inequality

$$I^p(f) \geq K(\tau)\kappa_\tau(\Gamma)\min(\mu_p(f), 1-\mu_p(f))\,, \tag{44}$$

where $K(\tau)$ is a positive function of τ. It can be shown that this reduces to Theorem 5.2 when we specialize to graph properties, emphasizing that the symmetry implied by Γ-invariance leads directly to a sharp threshold.

Bourgain and Kalai also gave examples of Γ-invariant functions f_n such that $\mu(f_n)$ is bounded away from 0 and 1 and $I(f_n) = \Theta(\kappa(f_n))$. Based on this result and results on primitive permutation groups (that require the classification of finite simple groups), it is possible to classify the coarsest threshold behavior for Γ-invariant Boolean functions, when Γ is a primitive permutation group. Welcome results here would include sharper lower bounds for the influences and, for example, proving a lower bound of $K\log^2 n\mu(f)(1-\mu(f))$ on the influence of Boolean functions that describe graph properties. See Bourgain and Kalai [22] for further details.

5.3 THRESHOLD BEHAVIOR FOR SMALL CRITICAL PROBABILITIES

Theorem 3.7 addressed the consequences of a coarse threshold when p is bounded away from 0. In this section we state theorems by Friedgut [36] and by Bourgain [21] on the sharpness of thresholds (as defined by eq. (15)), that apply when

the critical probability p_c tends to zero. These theorems yield sharp threshold results for graph properties when p_c tends to zero. Recall that theorem 5.2 asserts that a sharp threshold is guaranteed for graph properties when the critical probability is bounded away from 0 and 1.

Given a family $\mathcal{G}$ of graphs, let $g_{\mathcal{G}}$ be the Boolean function describing the graph property: "The graph contains a subgraph H, where $H \in \mathcal{G}$." For a graph H, $e(H)$ denotes the number of edges in H.

Theorem 5.3 (Friedgut [36]). *Let f represent a monotone graph property. For every $a \geq 1$ and $\gamma > 0$, there exists a value $C = C(a, \gamma)$ such that if $I^p(f) < a$, then there is a family $\mathcal{G}$ of graphs such that*

$$e(H) \leq C \text{ for every } H \in \mathcal{G}$$

and

$$\mu_p(\{x : f(x) \neq g_{\mathcal{G}}(x)\}) \leq \gamma\,. \tag{45}$$

The interpretation of the theorem is that a coarse threshold implies that the function has "local" behavior.

Friedgut's proof relies on symmetry and the statement extends to hypergraphs and similar structures. The crucial property appears to be that the number of orbits of sets of a given size, or $T(m)$ in the notation of the previous section, has a uniform upper bound. (For graphs this reads: For a fixed nonnegative integer m the number of isomorphism types of graphs with n' vertices and m edges is uniformly bounded.)

Friedgut conjectured that his theorem can be extended to arbitrary Boolean functions. For a collection $\mathcal{G}$ of subsets of $[n]$ (which without loss of generality we assume to be an antichain of sets, so it does not contain two sets Q and R with $Q \subset R$) let $g_{\mathcal{G}}(x_1, x_2, \ldots, x_n)$ be defined as follows: $g_{\mathcal{G}}(x_1, x_2, \ldots, x_n) = 1$ if and only if for some $S \in \mathcal{G}$, $x_i = 1$ for every $i \in S$. The sets S in $\mathcal{G}$ are called minterms for the function $g_{\mathcal{G}}$. Of course, every Boolean function can be represented in such a way.

Conjecture 5.4 (Friedgut). *Let f be a monotone Boolean function. For every $a \geq 1$ and $\gamma > 0$, there is a value $C = C(a, \gamma)$ such that if $I^p(f) < a$, then there is a family $\mathcal{G}$ of subsets of $[n]$ such that*

$$|S| \leq C \text{ for every } S \in \mathcal{G}$$

and

$$\mu_p(\{x : f(x) \neq g_{\mathcal{G}}(x)\}) \leq \gamma\,.$$

In other words, Friedgut's conjecture asserts that a Boolean function with low influence can be approximated by a Boolean function with small minterms.

A theorem of Bourgain [21] towards this conjecture which is very useful for applications is:

Theorem 5.5 (Bourgain). *Let f be a monotone Boolean function. For every $a \geq 1$, there is a value $\delta = \delta(a) > 0$ such that if $I^p(f) < a$ then there is a set S of variables, $|S| < 10a$, such that*

$$\mu_p(f(x)|x_i = 1 \text{ for every } i \in S) \geq (1+\delta)\mu_p(f)\,.$$

Both Friedgut's and Bourgain's theorems are very useful for proving sharp threshold behavior in many cases. We will mention one example that was studied in Friedgut's original paper, and is central to this volume. We refer the reader to Friedgut's recent survey article [37] for many other examples. This survey article also describes various handy formulations of theorems 5.3 and 5.5.

The 3-SAT problem. This problem has been discussed at length in Percus et al. [77]. Consider n Boolean variables, $x_1, \ldots, x_n$. A "literal" z_i is either x_i or $\overline{x_i}$. A clause c is an expression of the form $(z_i \vee z_j \vee z_k)$ where the symbol $\vee$ represents the logical OR and $1 \leq i < j < k \leq n$. A 3-CNF formula with m clauses is a formula of the form $(c_1 \wedge c_2 \wedge \cdots \wedge c_m)$, where the symbol $\wedge$ represents the logical AND. A random formula of length m is obtained by choosing c_i uniformly at random among the possible $8\binom{n}{3}$ possible clauses. A closely related model is obtained by choosing each one of the possible $8\binom{n}{3}$ clauses at random with probability p. (See Kirousis et al. [66] for further discussion of the differences between these ensembles.) A formula is satisfiable if we can assign truth values to the variables so that the Boolean value of the entire formula is TRUE. The larger m is, the more difficult it is. Using a slight extension of Theorem 5.3, Friedgut proved that there is a threshold $\alpha_c(n)$ such that for every $\epsilon > 0$, a random formula with $(\alpha_c(n) + \epsilon)n$ clauses is satisfiable with probability tending to 0 (as n tends to infinity) while a random formula with $(\alpha_c(n) - \epsilon)n$ clauses is satisfiable with probability tending to 1. It is still an outstanding problem to show that $\alpha_c(n)$ can be replaced by a constant α_c in the large n limit, meaning that the location of the critical probability does not oscillate.

Recent advances concerning the location of the critical value for the k-SAT problem are discussed in Kirousis et al. [66].

5.4 MARGULIS' THEOREM

Margulis [71] found in 1974 a remarkable condition guaranteeing a sharp threshold for Boolean functions, and applied it to study random subgraphs of highly connected graphs. His paper also contains an earlier proof of Russo's lemma. The theorem later improved by Talagrand [89] gives another general method for proving threshold behavior.

Let f be a monotone Boolean function. For $x \in \Omega_n$ let

$$h(x) = |\{y \in \Omega_n : d(x,y) = 1, f(y) \neq f(x)\}|\,, \tag{46}$$

with the Hamming distance $d(x, y)$ as defined in eq. (1). Thus, $h(x)$ counts the number of neighbors of x for which the value of f changes, which is the number of pivotal variables at x. Note that the total influence is then given by

$$I^p(f) = \sum_{x \in \Omega_n} \mu_p(x) h(x) \,. \tag{47}$$

Define $h_+(x) = h(x)$ if $f(x) = 1$ and $h_+(x) = 0$ if $f(x) = 0$. Since every pair x, y with $f(x) \neq f(y)$ has precisely one element where f attains the value one, one finds

$$pI^p(f) = \sum_{x \in \Omega_n} \mu_p(x) h_+(x) \,.$$

Theorem 5.6 (Talagrand [89]).

$$\sum_{x \in \Omega_n} \mu_p(x) \sqrt{h_+(x)} \geq \mu_p(f)(1 - \mu_p(f)) \frac{\sqrt{2} \min(p, 1-p)}{\sqrt{p(1-p)}}. \tag{48}$$

Suppose (for simplicity) that $p_c(f)$ is bounded away from 0 and 1. Suppose also that if $h_+(x) > 0$ then $h_+(x) \geq k$. This implies that

$$I^p(f) = (1/p)\mu_p(x) \sum_{x \in \Omega_n} h_+(x) \geq \sqrt{k} \sum_{x \in \Omega_n} \mu_p(x) \sqrt{h_+(x)} \,.$$

It then follows from eq. (48) that

$$I^p(f) \geq C\sqrt{k}.$$

By Russo's lemma the length of the threshold interval is $O(1/\sqrt{k})$.

Here is Margulis' original application. Let G be a k-connected graph, that is, at least k vertices must be deleted from G for it to no longer be connected. Consider a random spanning subgraph H where an edge of G is taken to be absent from H with probability p. We assume that H has n edges and let f be the Boolean function that represents the property: "H is not connected." Margulis proved that the threshold interval for connectivity is of length $O(1/\sqrt{k})$. The reason is that if H is not connected, but it is possible to make H connected by adding back a single edge of G (so that $h_+(x) > 0$), then H must have precisely two connected components. Since G is k-connected, there are at least k edges in $G \backslash H$ such that adding any of them to H yields a connected graph. It thus follows that if $h_+(x) > 0$ then $h_+(x) \geq k$.

5.5 FURTHER CONNECTIONS AND PROBLEMS

- **The giant component.** Both Talagrand's strengthening of Margulis' theorem and Friedgut's theorem give the sharp threshold of graph connectivity as

a special case. This is nice, but a serious criticism would be that the more interesting phase transition relating to connectivity occurs earlier, when p is around $1/n$. The value $1/n$ is the critical probability of the emergence of the "giant component" [2, 51]. It would be desirable to understand even the basic facts concerning the giant component in the context of general threshold phenomena, discrete isoperimetry, and Fourier analysis.

- **Graph invariants.** We have discussed a monotone graph property, or more generally a monotone Boolean function, and varied the parameter p. A different scenario would be to consider a parameter of graphs or a function defined on the discrete cube and study its distribution for a fixed p. We can consider, for example, the chromatic number, the clique number, the size of the maximal component, etc. The probabilistic properties of monotone functions on the discrete cube, and especially those which come from interesting graph parameters are of great interest. Discrete isoperimetric relations play a central role in this study. But direct relations with threshold results and with Fourier analysis are sparse.
- **Hereditary properties.** We could also consider non-monotone properties. A property of graphs (on n vertices) described by a Boolean function f is *hereditary* if there is a collection $\mathcal{H}$ of graphs such that $f = 1$ if the graph contains a subgraph H from $\mathcal{H}$ as an *induced* subgraph. Alon and Kalai asked for which hereditary properties is it the case that the measure of the set of p's for which $\epsilon < \mu_p(f) < 1 - \epsilon$ tends to 0 as n tends to infinity. Since f need not be monotone, this set will not necessarily be an interval. Of course, monotone properties are hereditary.
- **Influence of Boolean functions with tiny measure.** Another criticism would be that we concentrate on the secondary problem of threshold behavior while neglecting the primary problem of finding the location of the critical probability. Indeed, finding the critical probability of particular properties of random structures is a large and beautiful field, and is the subject of later chapters of this book. We comment that there are a very few cases where knowing that the threshold is sharp helps in estimating its location, since it is sufficient to show that the property is satisfied with a probability that is small but bounded away from zero. The analogy with physical models suggests that the threshold behavior, like certain critical exponents for models of statistical physics, may exhibit more "universal" behavior than the location of the critical probability.

Finally, recent work of Kahn and Kalai [52] suggests that for a large class of problems, good estimates on the location of critical probabilities can follow from understanding the behavior of the function $t_\epsilon(f)$ when ϵ itself is a function that tends to zero with increasing n. Such an understanding can be derived from some conjectures, quite similar to theorems 5.3, 5.5 and conjecture 5.4, about influences of Boolean functions when $\mu_p(f)$ tends to zero with increasing n.

6 THRESHOLD BEHAVIOR AND COMPLEXITY

In this section we will discuss two areas where threshold phenomena and complexity theory are related. First we will describe results on bounded depth circuits, a very basic notion in computational complexity. Second we will describe the connection to the area of "hardness of approximation."

6.1 BOUNDED DEPTH BOOLEAN CIRCUIT

The important complexity class AC0 of Boolean functions consists of those that can be expressed by Boolean circuits of polynomial size (in the number of variables) and bounded depth. Although functions belonging to AC0 are of very low complexity, the class is an important one. Here we show that such functions must have a coarse threshold behavior.

A Boolean circuit is a directed acyclic graph with $2n$ sources, each corresponding to a variable x_i or its negation $\overline{x_i}$, and one sink representing the output of the computation. The intermediate vertices are called gates and can represent the Boolean operations AND and OR. The size of a Boolean circuit is the number of vertices including all sources, gates, and sink. The depth is the maximum length of a directed path.

Boppana [19] proved that if a Boolean function f is expressed by a depth-c circuit of size N, then

$$I(f) \leq C_1 \log^{c-1} N\,. \tag{49}$$

Earlier, Linial, Mansour, and Nisan [69] proved that for Boolean functions that can be expressed by Boolean circuits of polynomial (or quasi-polynomial) size and bounded depth the Fourier coefficient sum $W_k(f)$ defined in equation (28) decays exponentially with k when k is larger than poly-logarithmic in the number of variables. This result relies on the fundamental Håstad switching lemma [46, 2], and a more precise result was recently given by Håstad [48]. It appears that all these results and their proofs apply to the probability measure $\mu_p(f)$ when p is bounded away from 0 and 1.

Remark: A *monotone circuit* is one where all the gates are monotone increasing in the inputs, that is, there are no NOT gates. The Håstad lemma for monotone Boolean circuits is easier, and was already proved much earlier by Boppana [20].

It can be conjectured that the "only" reason for a small total influence, and hence for a coarse threshold behavior, comes from bounded depth small circuits. Here, *small* means a slowly growing function of n. For that to be the case, an inequality that is roughly the reverse of eq. (49) must also hold. The following conjecture is a particularly bold version of the statement:

Conjecture 6.1 (Reverse Håstad). *Let f be a monotone Boolean function. For every $\epsilon > 0$ there is a value $K = K(\epsilon) > 0$ and another function g expressible as*

a Boolean circuit of size N and depth c, such that

$$\log^{c-1} N < KI(f)\,,$$

and

$$\mu\{x : f(x) \neq g(x)\} < \epsilon\,.$$

Remarks:

- As discussed in the previous chapter, a large number of papers in recent years have suggested a bold and far-reaching statistical physics approach to fundamental questions in complexity. These papers regard classical optimization problems as zero-temperature cases of statistical physics systems. The approach further proposes that the complexity of problems may be related to the type of phase transition of the physical system. In addition, statistical physics suggests both a way of thinking and heuristic mathematical machinery for dealing with these problems. This approach has met with some skepticism within the complexity theory community, and evidence for its usefulness is still tentative. The results by Håstad, Linial-Mansour-Nisan, and Boppana can be interpreted as going in the direction suggested by physicists. Of course, when we deal with complexity classes beyond AC0, caution is still advised.
- Connections between influences and the model of decision trees can be found in Friedgut et al. [38] and O'Donnell et al. [75].

6.2 HARDNESS OF APPROXIMATION AND PCP

Given an optimization problem, what is the complexity of finding an approximation to an optimal solution? Sometimes approximation is intractable and sometimes it is easy. The theory of probabilistically checkable proofs (PCP) is a powerful tool for studying approximation. Technical results pertaining to sharp threshold phenomena are important for showing that certain approximation problems are difficult.

The PCP theorem concerns *constraint satisfaction problems* (sometimes referred to as label-cover problems) of various types, and is the main tool in proving NP-hardness for approximation problems. As examples, consider the following two computational problems:

Vertex Cover: Given a graph G, find the smallest set of vertices whose complement is an independent set.

MAX-CUT: Given a graph G, find a partition of its vertices that maximizes the number of edges between the two sets of the partition.

Coming up with the optimal solution for these problems is known to be NP-hard [58]. The next best option is to approximate the optimal solution. In the case

of Vertex Cover, that means coming up with an appropriate set that may not be the smallest, but whose size is larger by at most some fixed *approximation factor*. Approximating MAX-CUT requires coming up with a partition that may not maximize the cut size, but gives a cut whose size is within a fixed approximation factor of the maximum.

Proving that such problems are NP-hard requires extending the Cook-Levin [25, 68] characterization of NP, which in simple terms states that SAT is NP-complete. One has to show that even approximating SAT is NP-hard, in the following sense.

A *Constraint Satisfaction Problem* (CSP) involves a set of variables and constraints over the assignment to those variables. Let X and Y be two sets of (not necessarily Boolean) variables, whose range is R_X and R_Y respectively. R_X and R_Y are two fixed sets independent of the sizes of X and Y. For some pairs of variables (x, y) where $x \in X$ and $y \in Y$, there is a *constraint* $\phi_{x,y} \subset R_X \times R_Y$, specifying the values of x and y that satisfy it. The constraints imposed on the variables are local, in the sense that they only involve one variable in X and one in Y. Let us further assume that all constraints have the *projection* property: for each constraint $\phi_{x,y}$, for every $a \in R_X$ there is only one $b \in R_Y$ so that both satisfy $\phi_{x,y}$. Our objective is to find an assignment for all variables $x \in X$ and $y \in Y$ such that no constraint will be violated.

A very general version of the PCP theorem is as follows:

Theorem 6.2 (PCP [5, 4, 79]). *Given a CSP Φ as defined above, there exists a constant $\delta > 0$ such that it is NP-hard to exclude either of the following alternatives:*

- *There is a variable assignment satisfying all the constraints $\phi \in \Phi$.*
- *There is no variable assignment satisfying even a fraction $\epsilon = |R_X|^{-\delta}$ of the constraints $\phi \in \Phi$.*

Note that if we had an approximation algorithm determining whether or not there is an assignment satisfying at least an ϵ fraction of the constraints, this algorithm would necessarily rule out one of the two alternatives. Namely, given a CSP instance, if the algorithm satisfies an ϵ fraction of the entire set of constraints, the second alternative is ruled out, while if it satisfies less than an ϵ fraction of the constraints, the first alternative is ruled out. Therefore, the corresponding approximation problem is NP-hard.

A general scheme for proving hardness of approximation was developed in Arora et al. [5], Arora and Safra [4], Bellare et al. [8], Dinur and Safra [26], and Håstad [47, 49]. Let us demonstrate this scheme on the Vertex Cover problem from above. We consider a basic combinatorial construction in which sufficiently large independent sets—or alternatively, small vertex covers—are represented by juntas. We then sketch a reduction of CSP to vertex cover, such that juntas lead

to variable assignments satisfying an ϵ fraction of the constraints. By the PCP theorem, this implies that approximating the Vertex Cover problem is NP-hard.

We proceed as follows. First, consider the graph $G_I^{[n]}$, whose vertex set Ω_n is the set of all binary vectors $\{0,1\}^n$ of length n. One may think of these vertices as all possible input vectors to a function over n Boolean variables. In $G_I^{[n]}$, two vertices v and u are adjacent if there is no $i \in [n]$ so that $v_i = u_i = 1$. This is referred to as the *non-intersection* graph, and it is the complement of the *intersection* graph (where two vectors are adjacent if the sets of indices where they are 1 have non-empty intersection), which has been investigated extensively. It is easy to see that no independent set in $G_I^{[n]}$ contains more than half of the vertices. This upper bound corresponds to an independent set that for some index i takes all vectors whose ith entry is 1. Such an independent set is the pre-image of a dictatorship Boolean function. What other large independent sets can one find in $G_I^{[n]}$?

The pre-image of the majority function (or any other odd monotone Boolean function) is also an independent set in the non-intersection graph, as any two vectors with more than half of their indices being 1 must have an index in which both are 1. For odd n that independent set matches the upper bound. To apply the PCP theorem we will need to "eliminate" independent sets, such as the majority function, that are not close to juntas.

For this purpose, one may impose a different distribution on the vertices of $G_I^{[n]}$ that will rule out such examples. One can assign weights to the vertices of $G_I^{[n]}$ according to μ_p for some p smaller than $1/2$, weighting independent sets as the sum of their vertices' weight. In that case, dictatorships' weights are p, while majority's weight tends to 0 as n tends to infinity.

What about independent sets that are smaller than those corresponding to dictatorships, but still within some constant factor of that size? It turns out that for $p < 1/2$ any independent set of non-negligible weight must correspond in some sense to a junta. The following result relies on Friedgut's theorem 3.7 and Russo's lemma.

Theorem 6.3 (Dinur and Safra [26]). *Let W be a locally maximal independent set in $G_I^{[n]}$ (thus, every vertex $x \in G_I^{[n]}$ is either in W or is adjacent to a vertex in W), and let f be a Boolean function where $f(x) = 1$ if $x \in W$ and $f(x) = 0$ if $x \notin W$. For every $0 < p < 1/2$, $\gamma > 0$ and $\epsilon > 0$, there exists a value $q \in [p, p+\gamma]$, a value $C(\gamma, \epsilon) \leq 2^{O(1/\gamma\epsilon)}$ and another Boolean function g depending on at most C variables, such that*

$$\mu_q(\{x \in \Omega_n : f(x) \neq g(x)\}) \leq \epsilon\,.$$

Note that if we let $J \subseteq [n]$ denote the C variables that g depends on, the pre-image $g^{-1}(1)$ represents a set of vectors over J that constitutes an independent set over G_I^J.

We now sketch the reduction from the CSP instance Φ above to the Vertex Cover problem. One constructs a graph G_Φ as follows. G_Φ consists of one copy of $G_I^{R_X}$ for every variable $x \in X$, and one copy of $G_I^{R_Y}$ for every variable $y \in Y$. Additional edges, representing constraints, are then added to connect the copies. The effect of these edges is that large independent sets reflect consistent assignments of Φ: in particular, if there is an assignment satisfying all constraints, then the set of vertices made up of the dictatorships in each copy forms an independent set in G_Φ. Theorem 6.3 guarantees that any independent set in G_Φ corresponds to juntas in many of the copies of G_I in G_Φ, so a sufficiently large independent set allows one to design an assignment that satisfies at least an ϵ fraction of Φ. This excludes the second alternative in the PCP theorem. Consequently, finding whether or not such a large independent set exists must be NP-hard.

We now describe another powerful form of PCP. Consider a further restricted CSP variant. Above we required the constraints to satisfy the projection property, meaning that for any constraint $\phi_{x,y}$, the value for x, $a \in R_X$, determines a unique value for y so that both satisfy $\phi_{x,y}$. What if we require in addition that the value for y uniquely determines the value of x?

Given a CSP instance satisfying this *uniqueness* property, one can efficiently figure out whether there is an assignment satisfying all constraints. Nevertheless, one may consider the following problem:

Khot's Unique Game [63] Given a CSP instance Φ that conforms to the uniqueness property, decide whether one of the following alternatives can be exlcuded:

- There exists an assignment satisfying at least a fraction $1-\epsilon$ of the constraints $\phi \in \Phi$
- No assignment satisfies even a fraction ϵ of Φ.

For $\epsilon > 0$, the complexity of this problem is still wide open. No polynomial algorithm is known for it; neither is it known to be NP-hard. (Khot himself conjectures that the problem is NP-hard.) Placing this problem within the known complexity classes is an exciting open question. The motivation for this problem, and the reason it is so interesting, is that it is often possible to relate the hardness of approximation problems to that of the Unique Game problem. We will give examples in the next section.

6.3 THE SHARP THRESHOLD BETWEEN EASY AND HARD PROBLEMS

In the previous section we briefly discussed PCP and indicated how technical results for threshold phenomena are used. There is another threshold aspect to the story. It turns out that for various optimization problems, when we try to approximate the solution, there is a sharp threshold between cases that are very easy to solve and cases in which the problem is NP-hard. This insight and the

methodology for observing such phenomena are fairly recent, and a deeper understanding of the issues involved may lead both to improved approximation algorithms and to tighter hardness results. (We do not see a clear connection between the two appearances of sharp thresholds in this story.) Harmonic analysis of Boolean functions has already proved to be a powerful tool for such considerations.

Here are some results concerning sharp transitions between easy and hard computational problems:

- MAX-3-LIN(2): Given a set of linear equations over $\mathbb{Z}_2$ (integers modulo 2), assign variables in such a way as to satisfy as many of them as possible. Satisfying half of the equations is easy—by just taking a random assignment—and this "algorithm" can be derandomized easily. However, for all $\epsilon > 0$, it is NP-hard to distinguish instances where $1/2 + \epsilon$ of the equations are satisfied and instances where $1 - \epsilon$ of the equations are satisfied [49].
- MAX-3-SAT: A similar problem—only instead of equations one has ORs over three literals each. A fraction $7/8$ of the constraints are expected to be satisfied by a random assignment, yet distinguishing between $7/8 + \epsilon$ and 1 is NP-hard [49].
- SET-COVER: Given a collection of subsets of $[n]$, find the smallest number of sets from the collection such that their union is $[n]$. A $\log n$ approximation (one that uses at most $\log n$ times as many sets as actually necessary) is simple to obtain, but nothing better can be achieved unless NP-complete problems with input size n have a deterministic algorithm with running time $n^{O(\log \log n)}$ [32, 80].

When we consider reductions to Khot's Unique Game problem, further results can be proved.

- MIN-2-SAT-DELETION: The instance is a formula in 2-CNF form, that is, a conjunction of clauses, each one consisting of 2 literals connected by OR. The goal is to delete as few of the clauses as possible, such that the remaining instance is completely satisfiable. Approximation within any constant factor (finding a solution that deletes at most a constant times as many clauses as actually necessary) is as hard as Khot's Unique Game problem [63].
- VERTEX COVER: Given an undirected graph, find the minimal number of nodes that touch all edges. A 2-approximation, namely covering the edges by at most twice the number of nodes needed, is quite easy—for example, by taking "both" ends of each as yet uncovered edge. Any better approximation is as hard as Khot's Unique Game problem [64].
- MAX-CUT: Find a 2-partition of the nodes of a given graph such that there are as many edges as possible between the two parts. We will return to this problem in the next section.

Remark: Other interesting cases of threshold behavior in complexity theory concern fault-tolerant computations, both for classical notions of computation and for "quantum computers."

7 NOISE SENSITIVITY

Motivated by mathematical physics, Benjamini, Kalai, and Schramm [10] have studied the sensitivity of an election's outcome to low levels of noise in the signals—or viewed differently, to small errors in the counting of votes. Their assumption is that there is a probability $\epsilon > 0$ of a mistake in counting a given vote and these probabilities are independent. Simple majority tends to be quite stable in the presence of noise. Two-level majority like the U.S. electoral system is less stable and multi-tier council democracy is quite sensitive to noise. This study is also closely related to works by Tsirelson, Vershik and Schramm [93, 94, 85]. For an attempt to apply the notion of noise sensitivity in finance, see [1].

For a Boolean function f and $\omega > 0$, consider the following scenario. First choose voter signals $x_1, x_2, \ldots, x_n$ randomly such that $x_i = 1$ with probability p, independently for $i = 1, 2, \ldots, n$. Let $S = f(x_1, x_2, \ldots, x_n)$. Next let $y_i = x_i$ with probability $1 - \omega$ and $y_i = 1 - x_i$ with probability ω, independently for $i = 1, 2, \ldots, n$. Let $T = f(y_1, y_2, \ldots, y_n)$. Define $C_\omega(f)$ to be the correlation between S and T.

Let $p, 0 < p < 1$, be fixed. A sequence $(f_n)_{n=1,2,\ldots}$ of Boolean functions such that $\mu_p(f_n)$ is bounded away from 0 and 1 is called *asymptotically noise-sensitive* if, for every $t > 0$,

$$\lim_{n\to\infty} C_\omega(f_n) = 0\,. \tag{50}$$

We will now define the complementary notion of noise stability. A class $\mathcal{F}$ of Boolean functions is *uniformly noise-stable* if for every $f \in \mathcal{F}$ and every $s > 0$ there exists a value $\omega = \omega(s) > 0$ such that $C_\omega(f) \geq 1 - s$.

A basic result concerning noise sensitivity is that the class of simple and weighted majority functions f such that $\mu_p(f)$ is bounded away from 0 and 1 is noise-stable. A sharp version was recently demonstrated by Peres [78]. Note that when the individual influences tend to 0, the property is a consequence of the central limit theorem.

The main result of Benjamini et al. [10] is a sort of converse of this. It asserts the following:

Theorem 7.1. *For every sequence (f_n) of monotone Boolean functions such that $\mu_p(f_n)$ is bounded away from 0 and 1 and (f_n) is not asymptotically noise-sensitive, there exists a weighted majority function g such that the correlation between (f_n) and g is bounded away from zero.*

The basic relation between noise sensitivity and influences is that for a sequence (f_n) of asymptotically noise-sensitive monotone Boolean functions,

$\lim I^p(f_n) = \infty$. Therefore, if f is noise-sensitive in its threshold interval, it must have a sharp threshold behavior. On the other hand, in this case the threshold interval is of length $\Omega(1/\sqrt{n})$.

In this chapter, we have described several results where, in order to demonstrate a sharp threshold behavior, we exhibited a large total influence. In some of these results the proofs actually give the stronger property of noise sensitivity.

The following four remarks will further demonstrate the relevance of noise sensitivity:

1. **The connection with Fourier coefficients.** A simple but important result from Benjamini et al. [10] asserts

   Theorem 7.2. *For every sequence (f_n) of Boolean functions such that $\mu(f_n)$ is bounded away from 0 and 1, (f_n) is asymptotically noise-sensitive if and only if for every $k > 0$*

   $$\lim_{n\to\infty} \sum_{i=1}^{k} W_i(f_n) = 0\,. \tag{51}$$

   Thus, f is noise-sensitive if and only if most of the L_2-norm of f is concentrated at "high frequencies." By the same token, noise stability is equivalent to the statement that most of the L_2-norm of f is concentrated at "low" frequencies.

   Theorem 7.3. *A class $\mathcal{F}$ of Boolean functions is uniformly noise-stable if and only if for every $f \in \mathcal{F}$ and every $\epsilon > 0$ there exists a value k such that*

   $$\sum_{i\geq k} W_i(f) < \epsilon\,. \tag{52}$$

2. **The majority-is-stablest conjecture.** What are the Boolean functions most stable under noise? It was conjectured by several authors that under several conditions that exclude individual variables having a large influence, majority is (asymptotically) most stable to noise. This conjecture has recently been proved by Mossel, O'Donnell, and Oleszkiewicz [73].
   We define a sequence (f_n) of Boolean functions to have a *diminishing individual influence* if

   $$\lim_{n\to\infty} \max\{I_k(f_n) : 1 \leq k \leq n\} = 0\,. \tag{53}$$

   Theorem 7.4 (Mossel, O'Donnell, and Oleszkiewicz [73]). *For every sequence (f_n) of Boolean functions with diminishing individual influence,*

   $$C_\omega(f_n) \leq (1 - o(1)) \left(1 - \frac{2}{\pi} \arccos(1 - 2\omega)\right)\,. \tag{54}$$

The fact that the right-hand side gives the precise asymptotic description of the noise stability of the majority function is a nineteenth-century result by Sheppard.

3. **MAX-CUT.** Khot, Kindler, Mossel, and O'Donnell [65] showed that the majority-is-stablest theorem (which at the time was a conjecture that they posed) implies a sharp threshold for approximating MAX-CUT based on Khot's unique game problem. The famous Goemans-Williamson algorithm based on semidefinite programming achieves the ratio $\alpha = .878567\ldots$ Khot, Kindler, Mossel, and O'Donnell showed that assuming the majority-is-stablest theorem, anything better is as hard as Khot's Unique Game problem.
4. **Monotone threshold circuits.** Threshold circuits form an important class of circuits that are more general than Boolean circuits, since they allow weighted majority gates. Contrary to the situation for Boolean circuits, it is not the case that functions expressible by constant depth threshold circuits have coarse threshold behavior, as is evident from the majority of such circuits. But there is a far-reaching conjecture [10] regarding their stability to noise that is analogous to the theorems by Boppana, Linial-Mansour-Nisan, and Håstad mentioned in the previous section:

Conjecture 7.5. *Consider the class $\mathcal{F}$ of monotone Boolean functions f that are expressed by monotone depth-c threshold circuits (of size $N(f)$). Then, for every $f \in \mathcal{F}$ and every $\epsilon > 0$ there is a value $K = K(\epsilon)$ such that*

$$\sum_{k > K \log^{c-1} N(f)} W_k(f) < \epsilon\,. \tag{55}$$

Equation (52) shows that a noise-stable Boolean function can be well approximated by a low depth threshold circuit, but we do not know whether, when the function is monotone, this can be achieved by a monotone threshold circuit.

Finally, let us note an important criticism arising from works by Tsirelson [85, 93]. These demonstrate that Boolean functions are too restricted for various problems and applications concerning noise sensitivity, and indicate that "binary trees" (in the form used in basic probability theory) rather than "cubes" are the correct mathematical framework. Tsirelson's more general setting allows him to study, for example, "correlated" random walks and Brownian motions. It suggests that the extensive investigation of Boolean functions, based on the discrete cube, may be complemented by investigations based on binary trees. This point of view may reflect on other topics studied in this chapter.

8 PERCOLATION

We have mentioned in the introduction that the area in which threshold behavior was originally studied is Physics. In this section we will discuss the model of percolation.

Consider the graph G of an m by $m+1$ planar rectangular grid. The vertices of G are thus points of the form $(i, j) : 1 \leq i \leq m, 1 \leq j \leq m+1$, and two vertices are adjacent in the graph G if they agree in one coordinate and differ by one in the other coordinate. Questions concerning percolation in the plane (usually on the infinite grid) are very important. Russo's lemma was proved in the context of percolation, and Kesten proved a sharp threshold result on the way to proving his famous result [59] on critical probabilities for planar percolation. (For a simple proof of Kesten's theorem and an extension to Voronoi percolation, see the recent papers by Bollobás and Riordan [15, 14].)

Choose every edge in G to be "open" with probability p. What is the probability of an open path from the left side of the rectangle to the right side? Is there a sharp threshold? We can ask and immediately answer the analogous question on the torus when we identify the left and right sides of the rectangle and the top and bottom sides, or even just for a cylinder when we identify only the left and right sides. When we look for a path homotopic to the horizontal path from (0,0) to $(0, m + 1)$, a sharp threshold follows from the proof of theorem 5.1.

The total influence of the Boolean function f described by "left-right" percolation on the $m + 1$ by m grid is a basic notion in percolation theory. It is conjectured that $I(f) \approx m^{3/4} \approx n^{3/8}$, where n is the number of variables. This conjecture was recently verified for one of the variants of planar percolation (site percolation on the triangular grid) based on the works of Smirnov, Lawler, Schramm, and Werner.

Basic Problem: For a Boolean function f with $\mu(f)$ bounded away from 0 and 1, find sufficient conditions to guarantee that for some $\alpha, \beta > 0$, $n^{\alpha} < I(f) < n^{1/2-\beta}$.

It was shown by Kesten [60, 61] that this property holds for the crossing event for planar percolation. Why does the total influence for percolation behave as a power of n? We can expect that the reason lies in some symmetry like the one considered in theorem 5.2 of Bourgain and Kalai. However, two facts are worth noting. The first is that the present formulation of Theorem 5.2 is not sufficiently strong to yield lower bounds of the form $I(f) > n^{\alpha}$. The second is that the Boolean function we described does not admit many symmetries. What it does seem to have is "approximate" symmetries. We expect that as the grid becomes finer, there will be some "limit object" (the scaling limit) reflecting an approximate symmetry of our functions under continuous maps of the square to itself. Such a symmetry is expected in any dimension. In two dimensions, it is expected that the limit object is symmetric under conformal maps. This was proved by Smirnov for another variant of planar percolation, namely site

percolation on the triangular grid. Noise sensitivity for the crossing event was proved in Benjamini et al. [10] and Schramm and Steif [84] recently proved a very strong form of it.

We now briefly discuss several related issues:

1. **First passage percolation.** Let f be a Boolean function. Consider a real function g defined on the discrete cube. Let $y_1, y_2, \ldots, y_n$ be independent, identically distributed random variables. Define

$$g(x_1, x_2, \ldots, x_n) = \min\{\sum x_1 y_1 + x_2 y_2 + + x_n y_n : f(x_1, x_2, \ldots, x_n) = 1\}. \tag{56}$$

   Understanding the behavior of the function g is of interest in percolation theory. In this context f is the Boolean function that describes the existence of a path of open edges between two points on the grid. Curiously, the same model is related to questions raised in mechanism design in economics theory. Influences and methods used to study them apply very nicely to the study of first passage percolation [9].
2. **Models with dependence.** One of the major research challenges is to extend the results described in this chapter to models where the probability distribution is not a product distribution. Important cases are the Ising and the more general Potts and random cluster models, as well as models based on random walks of various types. The random cluster model is a model of random subgraphs of a graph G with n edges, where one has a real parameter $q > 0$. The probability of a spanning graph H with k edges is proportional to

$$p^k (1-p)^{n-k} q^c,$$

   where c is the number of connected components of H. This model thus defines a two-parameter probability distribution on random subgraphs. The challenge is to find useful discrete isoperimetric theory and useful harmonic analysis for these probability distributions that will allow us to extend some of the general theorems described in this chapter.

   Very recently, Graham and Grimmett [40] have made a breakthrough in this area, extending the Kahn-Kalai-Linial theorem and deducing sharp threshold theorems for measures of the random-cluster type.
3. **The Fourier coefficients.** The Fourier coefficients of the crossing (and other) events for percolation are indexed by subgraphs of the grid. The Fourier transform gives a distribution on such subgraphs which is very interesting.

9 ECONOMICS AND VOTING: AN EXAMPLE OF SELF-ORGANIZED CRITICALITY

Let us now return to the Condorcet Jury theorem from the section 1. A key assumption in Condorcet Jury theorem is that each agent votes according to his or her signals. There is recent interesting literature on the case where voters vote *strategically* based on their signal. Suppose that every voter wishes to minimize the probability of mistakes, where we may assign different weights to mistakes in the two directions. Feddersen and Pesendorfer [31] considered the example of juries, where a much larger weight is typically given to an innocent person being convicted than to a guilty one being acquitted. Suppose that in order to convict, one needs two thirds of the votes. Suppose furthermore that each juror k receives a Boolean signal s_k such that if the defendant is guilty then $s_k = 1$ with probability $p > 1/2$ and if the defendant is innocent then $s_k = 1$ with probability $1 - p$. (We assume these signals are independent.) Now, if jurors vote according to their signals, then when $p = 0.51$ and the number of jurors is large, they will hardly ever convict.

Feddersen and Pesendorfer considered the case where jurors vote strategically, observing how their peers are voting, and use mixed (randomized) strategies. The surprising conclusion is that in such a situation, ever with a high threshold for conviction and a weak signal, the probability of either convicting an innocent defendant or acquitting a guilty one tends to zero as the number of jurors grows, even if the signal is weak. The one case where this does not hold is where unanimity among all jurors is required. Feddersen and Pesendorfer's result and analysis is based on the notion of Nash equilibrium. Nash equilibrium in this case gives us a nice example of "self-organized criticality." The behavior at the critical point is significant even when the voting method is biased from the beginning.

For the reader who is not familiar with game theory, some explanation is in order. To start with, every member of the jury has four pure strategies for how to act, given the signal he or she receives: act according to the signal, act opposite to the signal, acquit regardless of the signal, and convict regardless of the signal. A mixed strategy means a strategy involving randomization, so the outcome is probabilistic. In our case, a mixed strategy for juror k would be: upon receiving a signal to acquit, acquit with probability α_k and convict with probability $1 - \alpha_k$; upon receiving a signal to convict, acquit with probability β_k and convict with probability $1 - \beta_k$. We assume that each juror knows the signal s_k he or she has received, but not the signals or strategies of the other voters, and the jurors vote in a secret ballot. Furthermore, we assume that the signal strength p is known to all.

Each juror now votes in such a way as to maximize his or her own perceived "payoff," defined as follows. Jurors want to minimize the probability of a wrong decision, and it is considered worse to convict an innocent defendant than to

acquit a guilty defendant. So if the jury reaches the right decision, the payoff for each juror is zero. If the jury acquits a guilty defendant, the payoff for each juror is $-q$, where $q \in (0, 0.5)$. If the jury convicts an innocent defendant, the payoff for each juror is $q-1$. Note that the payoff function is the same for all jurors, and depends only on the collective decision of the jury. Given a sequence of mixed strategies, one for each juror, and based on an equal prior probability of innocence and guilt, a juror can estimate the posterior probability that the defendant is guilty as well as the expected payoff. In game theory, the Nash equilibrium point is a sequence of mixed strategies such that no player can expect a gain in payoff by deviating from his or her strategy as long as none of the other players deviates.

When we consider general voting methods and not only majority rules, it can be shown that "asymptotically complete aggregation of information" is intimately related to having a sharp threshold [83]. In particular, if there is a sharp threshold, then there is always a Nash equilibrium point for which the probability of mistakes tends to zero as the number of voters grows.

Fedderson and Pesendorfer's result is related to the question of why we care about critical behavior to start with. Why is it so often the case that shortly before an election between two candidates, both of them appear to have a significant chance of being elected? How come the probabilities we can assign to the choices of each individual voter do not "sum up" to a decisive collective outcome? This seems especially surprising in view of the sharp threshold phenomena. Fedderson and Pesendorfer's result suggests that the strategic behavior of voters can push the situation towards criticality. Another explanation would challenge the independence of the signals received by the voters.

There are other relations between threshold phenomena and economics and social choice theory. We have already seen in theorem 3.9 that having a sharp threshold for a sequence of monotone Boolean functions is equivalent to having a diminishing Shapley-Shubik power index. A famous result in social choice theory is Arrow's impossibility theorem concerning election methods when there are three or more candidates. Condorcet's famous "paradox" demonstrates that given three candidates A, B, and C, the majority rule may result in the society preferring A to B , B to C, and C to A. Arrow's Impossibility Theorem is an extension of Condorcet's paradox, and states that under certain general conditions such non-transitive social preferences cannot be avoided under *any* non-dictatorial voting method. Relations between threshold phenomena and Arrow type theorems are described in Kalai [54, 55].

As in the percolation discussion in section 8, a further problem in the context of economics is to understand matters under more realistic probabilistic assumptions, moving away from product distributions. This poses interesting conceptual and technical problems. Haggstrom, Kalai, and Mossel [44] studied aggregation of information in models with dependence. Another challenge in the economic arena is to study threshold phenomena (aggregation of information) and related notions such as noise sensitivity for more complex models.

10 CONCLUSIONS AND OPEN PROBLEMS

Threshold phenomena and related concepts such as pivotality, influence, and noise sensitivity are important in many areas of mathematics, science, and engineering. We have described some mathematical advances in the understanding of threshold behavior and related phenomena, as well as various applications and connections, and some open problems. The underlying mathematical concepts are similar in different disciplines. However, bridging the different points of view, methodologies, and interpretations is a major challenge. The subequent chapters of this book address this challenge from the perspectives of physics and computer science.

Over the course of this chapter, we have highlighted some important open problems. These include proving Friedgut's conjecture 5.4 and finding sharper versions of Bourgain and Kalai's theorem 5.2.[2] A less explicit but nevertheless important problem is to explain the emergence of power laws in the threshold interval, where the width of the interval behaves as $n^{-\beta}$ where $\beta > 0$ is a real number.

A fundamental challenge is to relate the threshold behavior to the threshold's location, and to find methods to exclude the possibility of oscillating critical probabilities. We mentioned this issue in the context of the k-SAT problem. It is equally of interest for many other problems as well.

Another important challenge is to find methods to deal with the influence of events of small probability. This is related to a detailed understanding of how the function $\mu_p(f)$ behaves, and especially to the analysis of large deviations of the threshold behavior. In this chapter we have dealt mainly with $t_\epsilon(f)$ when ϵ is fixed. It is of great interest to understand dependence on ϵ. The precise behavior of the function $\mu_p(f)$ in the threshold interval and the situation when ϵ itself is very small and expressed as a function of n are both very interesting topics. Kahn and Kalai [52] have proposed far-reaching conjectures concerning the influence $I^p(f)$ of Boolean functions f when $\mu_p(f)$ is a function of n and tends to 0 with increasing n. They also studied possible applications towards finding the location of the critical probability.

It would also be interesting to study threshold behavior and influences when we replace the Boolean cube $\{0,1\}^n$ by Σ^n when Σ is a finite alphabet with more than two letters. We expect, in that case, that for symmetric monotone functions the transition will occur in small "membranes" [56]. There is interesting related work concerning powers of arbitrary graphs by Alon, Dinur, Friedgut, and Sudakov [3]. There are various other generalizations of Boolean functions. Some can be found in Ben-Or and Linial's original paper [11] on collective coin flipping and are waiting to be explored further. Another important generalization is to

[2] Falik and Samorodnitsky [30] have very recently found a new proof of the Kahn-Kalai-Linial theorem based on an extension of the edge-isoperimetric inequality. Their methods may be relevant to some of the problems that we have mentioned.

functions of the form

$$f : \{0,1\}^n \to \{0,1\}^m . \tag{57}$$

These are of great importance in mathematics (e.g., error-correcting codes) and computer science (e.g., extractors).

Finally, it is worth repeating a problem already mentioned in several contexts: study threshold behavior and related notions of noise sensitivity and Fourier analysis for various models, with non-product probability distributions. These are examples of some of the many open problems suggested by our discussion of threshold phenomena. We challenge the reader to explore other applications.

ACKNOWLEDGMENTS

Research supported in part by an NSF grant and by an ISF bikura grant. We thank Allon Percus without whose encouragement and help this chapter would not have been written, and are thankful to many friends and colleagues including Noga Alon, Itai Benjamini, Irit Dinur, Ehud Friedgut, Jeff Kahn, Guy Kindler, Nati Linial, Elchanan Mossel, Ryan O'Donnell, Yuval Peres, Oded Schramm, and Boris Tsirelson for inspiring discussions and helpful remarks.

REFERENCES

[1] Akahori, J. "Asymptotics of Hedging Errors in A Slightly Incomplete Discrete Market: A Noise-Sensitive Example." Preprint (2002). http://www.ritsumei.ac.jp/se/∼akahori/papers/pp/slightly_incomplete.pdf (accessed July 28, 2005).

[2] Alon, N., and J. Spencer. *The Probabilistic Method*. New York: Wiley, 1992.

[3] Alon, N., I. Dinur, E. Friedgut, and B. Sudakov. "Graph Products, Fourier Analysis and Spectral Techniques." *Geom. Funct. Anal.* 14 (2004): 913–940.

[4] Arora, S., and S. Safra, "Probabilistic Checking of Proofs: A New Characterization of NP." *J. ACM* 45 (1998): 70–122.

[5] Arora, S., C. Lund, R. Motwani, M. Sudan, and M. Szegedy. "Proof Verification and Intractability of Approximation Problems." *J. ACM* 45 (1998): 501–555.

[6] Arrow, K. "A Difficulty in the Theory of Social Welfare." *J. Pol. Econ.* 58 (1950): 328–346.

[7] Beckner, W. "Inequalities in Fourier Analysis." *Ann. Math.* 102 (1975): 159–182.

[8] Bellare, M., O. Goldreich, and M. Sudan. "Free Bits, PCPs, and Nonapproximability—Towards Tight Results." *SIAM J. Comp.* 27 (1998): 804–915.

[9] Benjamini, I., G. Kalai, and O. Schramm. "First Passage Percolation has Sublinear Distance Variance." *Ann. Probab.* 31 (2003): 1970–1978.

[10] Benjamini, I., G. Kalai, and O. Schramm. "Noise Sensitivity of Boolean Functions and Applications to Percolation." *Publ. I.H.E.S.* 90 (1999): 5–43.

[11] Ben-Or, M., and N. Linial. "Collective Coin Flipping." In *Randomness and Computation*, ed. S. Micali, 91–115. New York: Academic Press, 1990.

[12] Bobkov, S., and F. Götze. "Discrete Isoperimetric and Poincaré-Type Inequalities." *Prob. Theor. Related Fields* 114 (1999): 245–277.

[13] Bollobás, B. *Random Graphs*. London, New York: Academic Press, 1985.

[14] Bollobás, B., and O. Riordan. "The Critical Probability for Random Voronoi Percolation in the Plane is 1/2." Preprint (2004). http://citebase.eprints.org/cgi-bin/citations?id=oai.arXiv.org:math/0410336 (accessed on July 28, 2005).

[15] Bollobás, B., and O. Riordan. "A Short Proof of the Harris-Kesten Theorem." Preprint (2004). http://citebase.eprints.org/cgi-bin/citations?id=oai.arXiv.org:math/0410359 (accessed on August 2, 2005).

[16] Bollobás, B., and A. Thomason. "Threshold Functions." *Combinatorica* 7 (1987): 35–38.

[17] Bollobás, B., C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson. "The Scaling Window of the 2-SAT Transition." *Rand. Struct. & Algorithms* 18 (2001): 201–256.

[18] Bonami, A. "Etude des Coefficients Fourier des Fonctiones de $L^p(G)$." *Ann. Inst. Fourier* 20 (1970): 335–402.

[19] Boppana, R. "The Average Sensitivity of Bounded Depth Circuits." *Infor. Process. Lett.* 63 (1997): 257–261.

[20] Boppana, R. "Threshold Functions and Bounded Depth Monotone Circuits." In *Proceedings of the 16th Annual ACM Symposium on the Theory of Computing (STOC '84)*, 475–479.

[21] Bourgain, J. "On Sharp Thresholds of Monotone Properties, Appendix to E. Friedgut. "Sharp Thresholds of Graphs Properties, and the k-SAT Problem."" *J. Amer. Math. Soc.* 12 (1999): 1017–1054.

[22] Bourgain, J., and G. Kalai. "Influences of Variables and Threshold Intervals under Group Symmetries." *Geom. Funct. Anal.* 7 (1997): 438–461.

[23] Bourgain, J., J. Kahn, G. Kalai, Y. Katznelson, and N. Linial. "The Influence of Variables in Product Spaces." *Israel J. Math.* 77 (1992): 55–64.

[24] Chayes, J. T., L. Chayes, D. S. Fisher, and T. Spencer. "Finite-Size Scaling and Correlation Length for Disordered Systems." *Phys. Rev. Lett.* 57 (1986): 2999–3002.

[25] Cook, S. "The Complexity of Theorem-Proving Procedures." In *Proceedings of the 3rd ACM Symposium on the Theory of Computing (STOC '71)*, 151–158.

[26] Dinur, I., and S. Safra. "On The Importance of Being Biased (1.36 Hardness of Approximating Vertex-Cover)." In *Annals of Math, Proceedings of the 34th Annual Symposium on the Theory of Computing (STOC '02)*. http://www.cs.huji.ac.il/∼dinuri/online-papers.htm (accessed on August 1, 2005).

[27] Dinur, I., V. Guruswami, S. Khot, and O. Regev. "A New Multilayered PCP and the Hardness of Hypergraph Vertex Cover." In *Proceedings of the 35th ACM Symposium on Theory on the Theory of Computing (STOC '03)*, 595–601.

[28] Erdös, P., and A. Rényi. "On the Evolution of Random Graphs." *Magyar Tud. Akad. Mat. Kut. Int.* 5 (1960): 17–61.

[29] Erdös, P., and A. Rényi. "On Random Graphs I." *Publ. Math. Debrecen* 6 (1959): 290–297.

[30] Falik, D., and A. Samorodnitsky. "A Combinatorial Proof for A Theorem of Kahn, Kalai, and Linial and Some Applications." Preprint (2005). http://www.google.com/search?q=cache:iG7zXxuVhXkJ:www.ma.huji.ac.il/∼kalai/ML.pdf+falik+samorodnitsky+combinatorial+theorem+kahn+linial&hl=en&ie=UTF-8 (accessed on August 2, 2005).

[31] Feddersen, T., and W. Pesendorfer. "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *Amer. Pol. Sci. Rev.* 92 (1998): 23–35.

[32] Feige, U. "A Threshold of $\ln n$ for Approximating Set Cover." *J. ACM* 45(1998): 634–652.

[33] Feige, U., S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. "Approximating Clique is Almost NP-Complete." *J. ACM* 43 (1996): 268–292.

[34] Friedgut, E., and G. Kalai. "Every Monotone Graph Property Has a Sharp Threshold." *Proc. Amer. Math. Soc.* 124 (1996): 2993–3002.

[35] Friedgut, E. "Boolean Functions with Low Average Sensitivity Depend on Few Coordinates." *Combinatorica* 18 (1998): 27–35.

[36] Friedgut, E. "Sharp Thresholds of Graph Properties, and the k-SAT Problem." *J. Amer. Math. Soc.* 12 (1999): 1017–1054.

[37] Friedgut, E. "Hunting for Sharp Thresholds." In *Random Structures and Algorithms*. 26(1-2) (2005): 37–51. http://citeseer.ist.psu.edu/650951.html (accessed on August 1, 2005).

[38] Friedgut, E., J. Kahn, and A. Wigderson. "Computing Graph Properties by Randomized Subcube Partitions." In *Randomization and Approximation Techniques in Computer Science, 6th International Workshop*, RANDOM 2002, 105–113. 2002.

[39] Goemans, M. X., and D. P. Williamson. "Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems using Semidefinite Programming." *J. Assoc. Comput. Mach.* 42 (1995): 1115–1145.

[40] Graham, B. T., and G. R. Grimmett. "Influence and Sharp Threshold Theorems for Monotonic Measures." Preprint (2005).

http://www.arxiv.org/abs/math.PR/0505057 (accessed on August 2, 2005).
[41] Grimmett, G. *Percolation*. Berlin: Springer-Verlag, 1989.
[42] Gross, L. "Hypercontractivity, Logarithmic Sobolev Inequalities and Applications: A Survey of Surveys." Preprint (2005). http://www.ma.huji.ac.il/∼kalai/ML.pdf (accessed on August 2, 2005).
[43] Gross, L. "Logarithmic Sobolev Inequalities." *Amer. J. Math.* 97 (1975): 1061–1083.
[44] Haggstrom, O., G. Kalai, and E. Mossel. "A Law of Large Numbers for Weighted Majority." Discussion Paper Series No. dp363. Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem. http://www.ratio.huji.ac.il/dp.asp (accessed on August 1, 2005).
[45] Hart, S. "A Note on the Edges of the n-Cube." *Discrete Math.* 14 (1976): 157–163.
[46] Håstad, J. "Almost Optimal Lower Bounds for Small Depth Circuits." In *Randomness and Computation*, ed. S. Micali, 143–170. 1989.
[47] Håstad, J. "Clique is Hard to Approximate within n to the Power $1 - \epsilon$." *Acta Mathematica* 182 (1999): 105–142.
[48] Håstad, J. "A Slight Sharpening of LMN." *J. Comp. Sys. Sci.* 63 (2001): 498–508.
[49] Håstad, J. "Some Optimal Inapproximability Results." *J. ACM* 48 (2001): 798–859.
[50] Holzman, R., E. Lehrer, and N. Linial. "Some Bounds for the Banzhaf Index and Other Semivalues." *Math. Oper. Res.* 13 (1988): 358–363.
[51] Janson, S., T. Luczak, and A. Rucinski. *Random Graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. New York: Wiley-Interscience, 2000.
[52] Kahn, J., and G. Kalai. "A Discrete Isoperimetric Conjecture with Probabilistic Applications." Preprint (2005). http://www.ma.huji.ac.il/∼kalai/ML.pdf (accessed on August 2, 2005).
[53] Kahn, J., G. Kalai, and N. Linial. "The Influence of Variables on Boolean Functions. In *Proceedings of the 29th Annual Symposium on the Foundations of Computer Science (FOCS '88)*, 68–80.
[54] Kalai, G. "A Fourier-Theoretic Perspective for the Condorcet Paradox and Arrow's Theorem." *Adv. in Appl. Math.* 29 (2002): 412–426.
[55] Kalai, G. "Noise Sensitivity and Chaos in Social Choice Theory." http://www.nyu.edu/gsas/dept/politics/seminar_pe_f04.shtml (accessed on August 1, 2005).
[56] Kalai, G. "Social Choice and Threshold Phenomena." Discussion Paper Series No. dp279. Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem. http://ratio.huji.ac.il/dp/dp279.pdf (accessed onAugust 2, 2005).
[57] Kalai, G. "Social Indeterminacy." *Econometrica* 72 (2004): 1565–1581.

[58] Karp, R. M. *Reducibility Among Combinatorial Problems*, 85–103. New York: Plenum Press, 1972.

[59] Kesten, H. "The Critical Probability of Bond Percolation on the Square Lattice Equals 1/2." *Comm. Math. Phys.* 74 (1980): 41–59.

[60] Kesten, H. "Scaling Relations for 2D-Percolation." *Comm. Math. Phys.* 109 (1987): 109–156.

[61] Kesten, H., and Y. Zhang. "Strict Inequalites for Some Critical Exponents in 2D-Percolation." *J. Stat. Phys.* 46 (1987): 1031–1055.

[62] Khintchine, A. "Über dyadische Brüche." *Math. Z.* 18 (1923): 109–116.

[63] Khot, S. "On the Power of Unique 2-Prover 1-Round Games." In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC '02)*, 767–775. New York: ACM Press, 2002.

[64] Khot, S., and O. Regev. "Vertex Cover Might be Hard to Approximate to Within $2-\varepsilon$." In *Proceedings of the 18th IEEE Annual Conference on Computational Complexity (CCC '03)*, 379–386.

[65] Khot, S., G. Kindler, E. Mossel, and R. O'Donnell. "Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs?" In *Proceedings of the 45th Symposium on the Foundations of Computer Science (FOCS 04)*.

[66] Kirousis, L. M., Y. C. Stamatiou, and M. Zito. "The Unsatisfiability Threshold Conjecture: Techniques Behind Upper Bound Improvements." This volume.

[67] Ledoux, M. *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001. http://www.lsp.ups-tlse.fr/Ledoux/ (accessed on August 2, 2005).

[68] Levin, L. "Universal'nyĭe perebornyĭe zadachi, universal search problems, (in Russian)." *Problemy Peredachi Informatsii* 9(3) (1973): 265–266.

[69] Linial, N., Y. Mansour, and N. Nisan. "Constant Depth Circuits, Fourier Transform, and Learnability." *J. Assoc. Comp. Mach.* 40 (1993): 607–620.

[70] Loomis, L., and H. Whitney. "An Inequality Related to the Isoperimetric Inequality." *Bull. Amer. Math. Soc.* 55 (1949): 961–962.

[71] Margulis, G. "Probabilistic Characteristics of Graphs with Large Connectivity (in Russian)." *Probl. Pered. Inform.* 10 (1974): 101–108.

[72] Mezard, M., and R. Zecchina. "The Random K-Satisfiability Problem: From an Analytic Solution to an Efficient Algorithm." *Phys. Rev. E* 66 (2002).

[73] Mossel, E., R. O'Donnell, and F. Oleszkiewicz. "Noise Stability of Functions with Low Influences: Invariance and Optimality." Preprint (2005). http://citebase.eprints.org/cgi-bin/citations?id=oai:arXiv.org:math/0503503 (accessed on August 2, 2005).

[74] Nelson, E. "The Free Markov Field." *J. Functional Analysis* 12 (1973): 211–227.

[75] O'Donnell, R., M. Saks, O. Schramm, and R. Servedio. "Every Decision Tree has an Influential Variable." In *The 46th Annual Symposium on the Foundations of Computer Science (FOCS '05)*, to appear.

[76] Owen, G. "Multilinear Extensions of Games." In *The Shapley Value*, ed. A. E. Roth, 139–151. Cambridge, UK: Cambridge University Press, 1988.

[77] Percus, A., G. Istrate, and C. Moore. "Combinatorial Problems in Computer Science and Their Relation to Statistical Physics Models: Random K-SAT, Replica Method, Overview of Main Concepts." This volume.

[78] Peres, Y. "Noise Stability of Weighted Majority." http://front.math.ucdavis.edu/math.PR/ (accessed on August 2, 2005).

[79] Raz, R. "A Parallel Repetition Theorem." *SIAM J. Comp.* 27 (1998): 763–803.

[80] Raz, R., and S. Safra. "A Sub-Constant Error-Probability Low-Degree Test, and A Sub-Constant Error-Probability PCP Characterization of NP." In *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing (STOC '97)*, 475–484.

[81] Russo, L. "An Approximate Zero-One Law." *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 61 (1982): 129–139.

[82] Russo, L. "A Note on Percolation." *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete.* 43 (1978): 39–48.

[83] Samet, Y. "Equilibria with Information Aggregation in Sharp Threshold Voting Rules." M.Sc. thesis, Hebrew University of Jerusalem, 2004.

[84] Schramm, O., and J. Steif. "Quantitative Noise Sensitivity and Exceptional Times for Percolation." http://arxiv.org/abs/math/0504586 (accessed on August 2, 2005).

[85] Schramm, O., and B. Tsirelson. "Trees, Not Cubes: Hypercontractivity, Cosiness, and Noise Stability." *Elec. Comm. Prob.* 4 (1999): 39–49.

[86] Steele, M. "Probability Theory and Combinatorial Optimization." In *CBMS-NSF Regional Conference Series in Applied Mathematics, 69, Society for Industrial and Applied Mathematics (SIAM)*, Philadelphia, PA, 1997.

[87] Talagrand, M. "Concentration of Measure and Isoperimetric Inequalities in Product Spaces." *Publ. I.H.E.S.* 81 (1995): 73–205.

[88] Talagrand, M. "How Much are Increasing Sets Positively Correlated?" *Combinatorica* 16 (1996): 243–258.

[89] Talagrand, M. "Isoperimetry, Logarithmic Sobolev Inequalities on the Discrete Cube, and Margulis' Graph Connectivity Theorem." *Geom. & Funct. Anal.* 3 (1993): 295–314.

[90] Talagrand, M. "On Boundaries and Influences." *Combinatorica* 17 (1997): 275–285.

[91] Talagrand, M. "On Influence and Concentration." *Israel J. Math.* 111 (1999): 275–284.

[92] Talagrand, M. "On Russo's Approximate Zero-One Law." *Ann. Prob.* 22 (1994): 1576–1587.

[93] Tsirelson, B. "Scaling Limit, Noise, Stability." In *Lectures on Probability Theory and Statistics*, 1–106. Lecture Notes in Mathematics, 1840. Berlin: Springer, 2004.

[94] Tsirelson, B., and A. Vershik. "Examples of Nonlinear Continuous Tensor Products of Measure Spaces and Non-Fock Factorizations." *Rev. Math. Phys.* 10 (1998): 81–145.

[95] Young, H. P. "Condorcet's Theory of Voting." *Amer. Econ. Rev.* 82 (1988): 1231–1244.