

Thresholds and expectation-thresholds of monotone properties with small minterms

Ehud Friedgut*

Faculty of Mathematics and Computer Science
Weizmann Institute of Science, Rehovot, Israel

ehud.friedgut@weizmann.ac.il

Jeff Kahn†

Department of Mathematics
Rutgers University, New Brunswick NJ 08854, U.S.A.

jkahn@math.rutgers.edu

Clara Shikhelman

School of Mathematical Sciences, Tel Aviv University, Israel and
Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel

clara.shikhelman@gmail.com

Submitted: Oct 20, 2014; Accepted: May 12, 2015; Published: Jun 3, 2015

Mathematics Subject Classifications: 05C88, 05C89

Abstract

Let N be a finite set, let $p \in (0, 1)$, and let N_p denote a random binomial subset of N where every element of N is taken to belong to the subset independently with probability p . This defines a product measure μ_p on the power set of N , where $\mu_p(\mathcal{A}) := \Pr[N_p \in \mathcal{A}]$ for $\mathcal{A} \subseteq 2^N$.

In this paper we study monotone (upward-closed) families \mathcal{A} for which all minimal sets in \mathcal{A} have size at most k , for some positive integer k . We prove that for such a family $\mu_p(\mathcal{A})/p^k$ is a decreasing function, which implies a uniform bound on the coarseness of the thresholds of such families.

We also prove a structure theorem which enables one to identify in \mathcal{A} either a substantial subfamily \mathcal{A}_0 for which the first moment method gives a good approximation of its measure, or a subfamily which can be well approximated by a family with all minimal sets of size strictly smaller than k .

Finally, we relate the (fractional) expectation threshold and the probability threshold of such a family, using linear programming duality. This is related to the threshold conjecture of Kahn and Kalai.

*Research supported in part by I.S.F. grant 0398246, BSF grant 2010247 and MINERVA grant 712023.

†Research supported in part by BSF grant 2010247.

1 Introduction

A family of sets \mathcal{A} is called *monotone* if $A \in \mathcal{A}$ and $A \subset B$ implies $B \in \mathcal{A}$. For such a family, a set which is minimal with respect to inclusion is called a *minterm*. One of the fundamental phenomena in random graph theory is that of thresholds of monotone properties. This dates back to the seminal papers of Erdős and Rényi [5, 6] who defined the notion of thresholds and discovered that for many interesting graph properties the probability of the property appearing in the binomial random graph $G(n, p)$, for large n , behaves much like a step function of the edge probability p , increasing from close to 0 to close to 1 abruptly as p is varied slightly. The study of thresholds of random structures in general, and in random graphs specifically, has been a thriving area ever since, and thousands of papers have covered related problems. Bollobás and Thomason [4] showed that every monotone property of sets has a threshold function and, using the Kruskal-Katona theorem, gave optimal quantification of such thresholds. In [8] it was observed that the KKL theorem [11], and its extension in [3] imply sharp thresholds for properties which are symmetric under the action of a group on the elements of the ground set, in particular for graph properties.

For most interesting families of graph properties the threshold function $p(n)$ tends to zero as n tends to infinity. In this case it is of interest to study the sharpness of the threshold. Fixing a graph property \mathcal{A} , and a parameter ϵ , one may ask for the width of the interval of values of p in which the probability of $G(n, p)$ having property \mathcal{A} climbs from ϵ to $1 - \epsilon$. This width is measured with respect to the value of p for which the probability of $G(n, p) \in \mathcal{A}$ is, say, $1/2$, henceforth *the critical p* . For a series of properties \mathcal{A}_n , of graphs on n vertices, we will say that the threshold is sharp if the ratio between the width of the threshold interval and the critical p tends to 0. We will shortly give a more precise definition of sharp thresholds in a more general setting.

In his Ph.D. thesis, the first author [7] gave a necessary condition for monotone graph properties to have a sharp threshold. Roughly speaking, if a property does not have a sharp threshold it must be well approximable by a *local property* (e.g. containing a triangle), as opposed to properties that are *global* (e.g. connectivity) and cannot be well approximated by the property of containing a subgraph from a fixed list. In the appendix to [7] Bourgain proved a similar statement, with a slightly weaker conclusion, in a much more general setting, without the assumption of symmetry. In a recent paper Hatami, [9] gives a common generalization of these two results.

Returning to the question of thresholds of local properties, the appearance of any fixed subgraph in $G(n, p)$ has a coarse threshold, and this is well understood, see Bollobás' paper [1] for a complete description. Roughly speaking, if a fixed graph H is strictly balanced then the number of copies of H in $G(n, p)$ will be approximately Poisson, and the governing parameter, the expectation of the random variable, will be of order $p^{|E(H)|}n^{|V(H)|}$, which varies smoothly with p : when p is multiplied by a constant c , the expectation changes by a factor of $c^{|E(H)|}$. When H is not balanced the situation is only slightly more complicated, and appearance of copies of H in $G(n, p)$ can be understood by studying the appearances of the densest subgraphs of H . Intuitively, if H has a subgraph H' which is much denser

than H , then every copy of H' that appears in $G(n, p)$ is extremely likely to be contained in many copies of H . Kahn and Kalai, [10] have a far-reaching conjecture as to what this looks like for general monotone properties. In the graph setting a variant of their conjecture (see [10, Conj. 2.1]) says that for a family generated by copies of a single *nonfixed* H (e.g. a Hamiltonian cycle) there is at most a logarithmic gap between the actual threshold and the probability that guarantees that the expected number of copies of each subgraph of H is at least (say) 1. (It is an open question—see [10, Question 2.2]—whether this variant is implied by the main conjecture of [10].)

The basic question which led to the writing of this paper was: how specific is this behavior to graphs? The proofs of this behavior use the symmetry of graphs very strongly, yet it seemed possible that something similar should hold also for properties of random binomial subsets of a ground set without any symmetry assumptions. This would imply a converse to the main theorems of [7] and its appendix: not only does a non-sharp threshold imply that the property in question has local nature, but also any property determined by small minimal sets has a non-sharp threshold.

We will prove in this paper that this indeed is the case.

2 Setting and main results

Let $[n]$ denote the set $\{0, 1, \dots, n\}$, and let $[n]_p$ denote a random subset of $[n]$, where each element is chosen independently with probability p . Recall that a family of sets \mathcal{A} is called *monotone* if $A \in \mathcal{A}$ and $A \subset B$ implies $B \in \mathcal{A}$. For \mathcal{A} , a family of subsets of $[n]$ and $p \in [0, 1]$ we define $\mu(\mathcal{A}, p)$ to be the probability that $[n]_p \in \mathcal{A}$. Note that if \mathcal{A} is monotone and nontrivial (i.e. neither \mathcal{A} nor its complement are empty) then this function is strictly increasing in p . We will also use the notation μ_p to denote the measure $\mu(\cdot, p)$.

For a fixed non-trivial monotone family \mathcal{A} and any $x \in [0, 1]$ we define p_x to be the unique number such that $\mu(\mathcal{A}, p_x) = x$. For $0 < \epsilon < 1/2$ define $\delta_\epsilon(\mathcal{A}) = \frac{p_{1/2} - p_\epsilon}{p_{1/2}}$. The numerator is the length of the *threshold interval*, in which the probability of \mathcal{A} climbs from ϵ to $1/2$. The denominator, $p_{1/2}$, supplies the correct yardstick with which to measure this length. The slower $\delta_\epsilon(\mathcal{A})$ tends to 1 as ϵ tends to 0, the sharper the threshold is (in other words, the smaller the threshold interval is.) Note it would also make sense to study the interval $[p_\epsilon, p_{1-\epsilon}]$, but our choice gives a neater normalization, bounding $\delta_\epsilon(\mathcal{A})$ between 0 and 1.

Theorem 1. *Let \mathcal{A} be a monotone family of subsets of $[n]$, with all minterms of size at most k . Then the function $\frac{\mu(\mathcal{A}, p)}{p^k}$ is monotone decreasing. Consequently*

$$\delta_\epsilon(\mathcal{A}) \geq 1 - (2\epsilon)^{\frac{1}{k}}.$$

The simple derivation of this theorem from the Margulis-Russo lemma was pointed out to us by Oliver Riordan. Note that the theorem is tight, e.g., for a family with a single minterm. We present the proof of Theorem 1 in Section 3 below. An upper bound on δ_ϵ will follow from a different approach which we present in Section 5:

Theorem 2. *Let \mathcal{A} be a monotone family of subsets of $[n]$, with all minterms of size at most k , then one has:*

$$1 - 2\epsilon \frac{(k-1)^{k-1}}{k^k} \geq \delta_\epsilon(\mathcal{A}).$$

To present the results of Section 5 we first need a definition. As usual, let $P([n])$ denote the power set of $\{1, 2, \dots, n\}$. For a monotone family $\mathcal{A} \subseteq P([n])$ with a set of minterms \mathcal{M} we define $\mathbb{E}_p(\mathcal{A})$ to be the expected number of minterms of \mathcal{A} that are contained in the random set $[n]_p$. $\mathbb{E}_p(\mathcal{A})$ is the expectation of the function $f(A) = \sum_{M \in \mathcal{M}} \mathbf{1}_{A \subseteq M}$, which takes positive integer values on all A in \mathcal{A} , and consequently gives an upper bound on $\mu(p, \mathcal{A})$. This upper bound can be tightened by a fractional version, which we call the *fractional expectation* of \mathcal{A} .

Definition 3. For a monotone family $\mathcal{A} \subseteq P([n])$, with a set of minterms \mathcal{M} , we define the fractional expectation of \mathcal{A} with respect to μ_p to be

$$\mathbb{E}_p^*(\mathcal{A}) = \min \sum \beta(B) p^{|B|},$$

where the minimum is taken over all functions $\beta : P([n]) \rightarrow \mathbf{R}^{\geq 0}$ such that

$$\sum_{B \subseteq A} \beta(B) \geq 1 \text{ for all } A \in \mathcal{M}.$$

If β is a function for which the minimum in the definition of the fractional expectation is achieved, then $\mathbb{E}^*(\mathcal{A})$ is the expectation of the function $g(A) = \sum_B \beta(B) \mathbf{1}_{B \subseteq A}$, which is non negative and assumes values at least 1 on all A in \mathcal{A} , hence this too gives a (better) upper bound on $\mu(\mathcal{A})$, namely

$$\mathbb{E}_p(\mathcal{A}) \geq \mathbb{E}_p^*(\mathcal{A}) \geq \mu_p(\mathcal{A}).$$

The main result of Section 5 is that for monotone families with minterms of bounded size this bound is not too far off mark.

Theorem 4. *Let \mathcal{A} be a monotone family with all minterms of size at most k . Then for any $\alpha > 0$*

$$\mathbb{E}_p^*(\mathcal{A}) \geq \mu_p(\mathcal{A}) \geq \mathbb{E}_{\alpha p}^*(\mathcal{A})(1 + \alpha)^{-k}.$$

As a corollary we deduce a special case of the expectation-threshold conjecture of [10].

Corollary 5. *Let \mathcal{A} be a monotone family of subsets of $[n]$, with all minterms of size at most k . If $\mathbb{E}_p^*(\mathcal{A}) = 1$ then $\mu_{kp}(\mathcal{A}) > 1/e$.*

Note that for a family \mathcal{A} with minterms of size at most k the expectation threshold (p for which $\mathbb{E}_p(\mathcal{A}) = 1$), and the fractional expectation threshold (p for which $\mathbb{E}_p^*(\mathcal{A}) = 1$), differ only by a multiplicative factor of 2^k . Talagrand conjectures in [16] that the gap between the thresholds for families with minterms of size at most k , is of order at most $\log(k)$.

In Section 4 we approach the question of understanding the threshold behavior of a monotone family $\mathcal{A} \subseteq P([n])$ via the parameter $\mathbb{E}_p(\mathcal{A})$, the expected number of minterms of \mathcal{A} in a random set $[n]_p$. If we have good control over the second moment of this random variable, the expectation gives us a good indication as to the probability $\mu_p(\mathcal{A})$. An example of this setting is, say, the family of all subgraphs of K_m that contain a copy of K_4 , when $p = \Theta(m^{-2/3})$. It is easy to verify that in this case the expected number of minterms (i.e. the expected number of copies of K_4 in $G(m, p)$) is $\Theta(1)$, whereas the variance is also of this order of magnitude, which enables us to get an effective lower bound on the measure of the family using the Payley-Zygmund bound

$$Pr[Z > 0] \geq \frac{E[Z^2]}{(E[Z])^2},$$

which holds for any non-negative random variable Z . On the other hand, consider the example where the minterms are all subgraphs of K_m containing a copy of “ K_4 with a tail”, a graph consisting of K_4 with a fifth vertex connected to precisely one of the four. We again set $p = m^{-2/3}$ and a moment of thought shows that although this family is properly contained in the previous one, the measure of their symmetric difference is negligible, as any copy of K_4 that appears in $G(m, m^{-2/3})$ is overwhelmingly likely to have many tails. This is reflected by the fact that the expectation now is huge rather than constant. In this case one has to realize that the tail connected to K_4 is a red herring, and proper analysis can, and should, focus on the previous family.

Such examples are almost canonical in any introductory course on random graphs. Our main theorem in Section 4 guarantees that something similar to one of these two case should hold in any family \mathcal{A} defined by minterms of bounded size k . Either there is a substantial subfamily \mathcal{B} for which the first and second moments are well behaved, or a substantial subfamily \mathcal{B} that may be approximated by a different family with minterms of size strictly smaller than k . This structure theorem then allows us to deduce a theorem quite similar to Theorem 1, with a slightly worse rate of decay (e.g. $\mu_{p/2}(\mathcal{A}) \geq \mu_p(\mathcal{A})/(k8^k)$), as opposed to the truth which is $\mu_{p/2}(\mathcal{A}) \geq \mu_p(\mathcal{A})/2^k$.)

3 The Margulis-Russo lemma, and proof of Theorem 1

For a monotone family \mathcal{A} , the Margulis-Russo lemma, ([14], [15]) relates the derivative of $\mu(\mathcal{A}, p)$ with respect to p with the *edge boundary* of \mathcal{A} . If $A \in \mathcal{A}$, but $(A \setminus a)$ is not in \mathcal{A} we say that a is a pivotal element of A and that there is a boundary edge “leaving A in the direction of a ”. Let $\text{Piv}(A)$ denote the number of pivotal elements in A (which is necessarily 0 if $A \notin \mathcal{A}$). This quantity is sometimes called the sensitivity of A . Let A be a random set chosen according to $\mu(p)$, then $\text{Piv}(A)$ is a random variable, and its expectation is a measure of the size of the boundary of \mathcal{A} . As the lemma below shows, this is a parameter intimately correlated with the threshold behavior of \mathcal{A} .

Lemma 6. [Margulis, Russo]

$$p \frac{d\mu(\mathcal{A}, p)}{dp} = \mathbb{E}[\text{Piv}(A)].$$

This lemma, which is so simple to state, and, as we shall see shortly, very easy to prove, is extremely useful. See, e.g., [8], [7], [2]. It is not surprising that this lemma is relevant when studying thresholds, as the expression on the right hand side is clearly related to the ratio between the width of the threshold interval and the value of p within the interval (it is an approximation of its reciprocal). Both Russo and Margulis proved this lemma by induction on n , the size of the ground set from which A is chosen. For the sake of being self contained we present a different proof, which is well known folklore, perhaps due to Gil Kalai.

Proof of Lemma 6: Let \mathcal{A} be a monotone family of subsets of $[n]$, and for some fixed (p_1, p_2, \dots, p_n) consider the following product measure $\mu_{(p_1, p_2, \dots, p_n)}$ on $P([n])$. The measure of a set A is $\prod_{i \in A} p_i \prod_{j \notin A} (1 - p_j)$. For $i \in [n]$ and a random set A chosen according to $\mu_{(p_1, p_2, \dots, p_n)}$ let a_i denote the probability that $(A \in \mathcal{A} \text{ and } (A \setminus \{i\}) \in \mathcal{A})$, and let b_i denote the probability that i is pivotal in $A \cup \{i\}$, i.e. $((A \cup \{i\}) \in \mathcal{A} \text{ and } (A \setminus \{i\}) \notin \mathcal{A})$. This means that the probability that i is pivotal is $p_i b_i$. Recalling that \mathcal{A} is monotone, for all i we have $\mu_{(p_1, p_2, \dots, p_n)}(A) = a_i + p_i b_i$, and $\mathbb{E}[\text{Piv}(A)] = \sum_i p_i b_i$.

We now let all of the p_i depend on a common parameter p in the following trivial way: $p_i(p) = p$, and note that the resulting measure is μ_p . The Margulis-Russo formula now follows from a simple application of the chain rule.

$$\frac{d\mu(\mathcal{A}, p)}{dp} = \sum_i \frac{\partial \mu_{(p_1, p_2, \dots, p_n)}(\mathcal{A})}{\partial p_i} \cdot \frac{dp_i}{dp} = \sum_i b_i = \frac{\mathbb{E}[\text{Piv}(A)]}{p}.$$

□

The fact that Theorem 1 follows immediately from the Margulis-Russo formula, as pointed out to us by Oliver Riordan, is yet another example of how useful this result is.

Proof of Theorem 1: Let \mathcal{A} be a monotone family with all minterms of size at most k . Let A be a random set chosen according to μ_p . Note that A can never have more than k pivotal elements, and that if $A \notin \mathcal{A}$ then, by definition there are no pivotal elements. Therefore

$$\mathbb{E}[\text{Piv}(A)] \leq k \cdot \mu_p(A).$$

Using this in conjunction with the Margulis-Russo formula, and deriving with respect to p gives

$$\left(\frac{\mu_p(\mathcal{A})}{p^k} \right)' = \frac{\mu' p^k - k p^{k-1} \mu}{p^{2k}} \leq 0.$$

□

4 Structure theorem for monotone families with small minterms

We begin with some notation. Let \mathcal{A} be a monotone family and let $\mathcal{M}(A)$ be the set of its minterms. Throughout this section we will assume that all minterms are of size at most k . Let $X^{\mathcal{A}}$ be the random variable that counts the number of minterms of \mathcal{A} in $[n]_p$. For a set $V \subset [n]$ we will define the family of its m -supplements with respect to \mathcal{A} to

be the family of sets W of size m whose disjoint union with V form a minterm, namely $N_{\mathcal{A}}^m(V) = \{W \subset [n] \text{ s.t. } |W| = m \text{ and } \exists M \in \mathcal{M}(\mathcal{A}) \text{ s.t. } W \uplus V = M\}$.

We say that \mathcal{A} is *tame* with respect to p if for any $1 \leq m \leq k-1$ and for any $V \subset [n]$ one has $|N_{\mathcal{A}}^m(V)| < p^{-m}$. We say that \mathcal{B} is a *tame m -approximation* of \mathcal{A} at p , if there is a subfamily $\mathcal{A}' \subset \mathcal{A}$, such that for all minterms $B \in \mathcal{M}(\mathcal{B})$ it holds that the set $N_{\mathcal{A}'}^m(B)$ has size at least p^{-m} and is tame.

The definition of a tame family is useful because it implies that for such a family the first moment bound on the measure is not too far from the truth. This is captured by the following lemma.

Lemma 7. *Let \mathcal{A} be a tame family with respect to p with minterms of size at most k . Then*

$$\mu_p(\mathcal{A}) \geq \frac{\min\{\mathbb{E}_p[X^{\mathcal{A}}], 1\}}{k2^k}$$

Proof. We would like to use the Paley-Zygmund inequality to bound the probability from below:

$$\mu_p(\mathcal{A}) \geq \frac{\mathbb{E}^2[X^{\mathcal{A}}]}{\mathbb{E}[(X^{\mathcal{A}})^2]} \tag{1}$$

Let us calculate the numerator and denominator separately. Denoting $\mathcal{M}(m) = |\{M_i \in \mathcal{M} \text{ s.t. } |M_i| = m\}|$ it's easy to see that:

$$\mathbb{E}[X^{\mathcal{A}}]^2 = \sum_{m,\ell=1}^k p^{m+\ell} \mathcal{M}(m)\mathcal{M}(\ell)$$

The denominator needs a bit more careful work. Remembering that $\mathcal{M}(\mathcal{A}) = \{M_i\}$ and that $X^{\mathcal{A}}$ is the random variable counting the number of minterms, we can define X_i to be the indicator of M_i and write $X^{\mathcal{A}} = \sum_i X_i$. With this we have:

$$\mathbb{E}[X^2] = \mathbb{E}\left[\sum_{i,j} X_i X_j\right] = \mathbb{E}\left[\sum_i X_i^2\right] + \mathbb{E}\left[\sum_{\substack{i \neq j \\ M_i \cap M_j \neq \emptyset}} X_i X_j\right] + \mathbb{E}\left[\sum_{M_i \cap M_j = \emptyset} X_i X_j\right].$$

It's easy to see that:

$$\mathbb{E}\left[\sum_{M_i \cap M_j = \emptyset} X_i X_j\right] \leq \sum_{m,\ell=1}^k p^{m+\ell} \mathcal{M}(m)\mathcal{M}(\ell) = \mathbb{E}[X^{\mathcal{A}}]^2$$

and as the X_i s are indicators we get

$$\mathbb{E}\left[\sum_i X_i^2\right] = \mathbb{E}\left[\sum_i X_i\right] = \mathbb{E}[X^{\mathcal{A}}]$$

and so we are left with taking care of the second summand.

Note that \mathcal{A} is a tame family, thus for any M_i and $1 \leq m \leq k-1$ one has:

$$|\{M_j \text{ s.t. } M_i \cap M_j \neq \emptyset, |M_j \setminus M_i| = l\}| = \sum_{V \subset M_i} N_{\mathcal{A}}^l(V) \leq 2^k p^{-l}.$$

With this we are ready to do the calculation. We break the sum into sums corresponding to the different sizes of minterms and supplements.

$$\begin{aligned} \mathbb{E}\left[\sum_{\substack{i \neq j \\ M_i \cap M_j \neq \emptyset}} X_i X_j\right] &= \sum_{m=1}^k \sum_{i: |M_i|=m} \left(\sum_{j=1}^{k-1} \sum_{\substack{j: M_j \cap M_i \neq \emptyset \\ |M_j \setminus M_i|=l}} \mathbb{E}[X_i X_j]\right) \\ &= \sum_{m=1}^k \sum_{i: |M_i|=m} \left(\sum_{j=1}^{k-1} |\{M_j \text{ s.t. } M_i \cap M_j \neq \emptyset, |M_j \setminus M_i| = l\}| p^{m+l}\right) \\ &\leq \sum_{m=1}^k \sum_{i: |M_i|=m} \left(\sum_{j=1}^{k-1} 2^k p^m\right) \\ &= (k-1)2^k \sum_{m=1}^k \sum_{i: |M_i|=m} p^m \\ &= (k-1)2^k \sum_{m=1}^k \mathcal{M}(m) p^m \\ &= (k-1)2^k \mathbb{E}[X^{\mathcal{A}}] \end{aligned}$$

Summing everything together we get that:

$$\mathbb{E}[(X^{\mathcal{A}})^2] \leq ((k-1)2^k + 1) \mathbb{E}[X^{\mathcal{A}}] + \mathbb{E}[X^{\mathcal{A}}]^2$$

And now plugging this in to Paley-Zygmund we get:

$$\mu_p(\mathcal{A}) \geq \frac{\mathbb{E}[X^{\mathcal{A}}]^2}{\mathbb{E}[(X^{\mathcal{A}})^2]} \geq \frac{\mathbb{E}[X^{\mathcal{A}}]^2}{((k-1)2^k + 1) \mathbb{E}[X^{\mathcal{A}}] + \mathbb{E}[X^{\mathcal{A}}]^2} \geq \frac{\min\{\mathbb{E}[X^{\mathcal{A}}], 1\}}{(k-1)2^k + 2}$$

For the simplicity of further calculations we can obviously write:

$$\mu_p(\mathcal{A}) \geq \frac{\min\{\mathbb{E}[X^{\mathcal{A}}], 1\}}{k2^k}$$

as needed. □

Corollary 8. *Let \mathcal{A} be a monotone family with minterms of size at most k . Then*

(i) *If \mathcal{A} is tame with respect to $p/2$ then $\mu_{p/2}(\mathcal{A}) \geq \mu_p(\mathcal{A})/(k2^{2k})$.*

(ii) *If \mathcal{B} is a tame m -approximation of \mathcal{A} at p then $\mu_p(\mathcal{A}) \geq \mu_p(\mathcal{B})/(m2^m)$*

Note that (i) is weaker than what can be deduced using Theorem 1, nonetheless we include it with its proof in order to demonstrate the information one can deduce from the structural approach.

Proof. To see (i) we only need to note that as the minterms of \mathcal{A} are of size at most k one has $\mathbb{E}_{p/2}[X^{\mathcal{A}}] \geq \frac{1}{2^k} \mathbb{E}_p[X^{\mathcal{A}}]$, together with the inequality on the first moment $\mathbb{E}_p[X^{\mathcal{A}}] \geq \mu_p(\mathcal{A})$ and Lemma 7 we get

$$\mu_{p/2}(\mathcal{A}) \geq \frac{\min\{\mathbb{E}_{p/2}[X^{\mathcal{A}}], 1\}}{k2^k} \geq \frac{\min\{\mathbb{E}_p[X^{\mathcal{A}}], 1\}}{k2^{2k}} \geq \frac{\mu_p(\mathcal{A})}{k2^{2k}}$$

as needed.

For (ii) note that as any $B \in \mathcal{M}(\mathcal{B})$ is a subset of a minterm of \mathcal{A} the following inequality holds for any p :

$$\mu_p(\mathcal{A}) \geq \mu_p(\mathcal{B}) \min_{B \in \mathcal{B}} \mu_p(N_{\mathcal{A}}^m(B))$$

It is left to show that for any $B \in \mathcal{B}$ one has $\mu_p(N_{\mathcal{A}}^m(B)) \geq \frac{1}{m2^m}$. \mathcal{B} is an m -approximation so there is a family $\mathcal{A}' \subset \mathcal{A}$ for which $|N_{\mathcal{A}'}^m(B)| > p^{-m}$, furthermore each minterm in $N_{\mathcal{A}'}^m(B)$ is of size m , so $\mathbb{E}[X^{N_{\mathcal{A}'}^m(B)}] \geq 1$. As $N_{\mathcal{A}'}^m(B)$ is tame we can apply Lemma 7 and get:

$$\mu_p(N_{\mathcal{A}'}^m(B)) \geq \frac{\min\{\mathbb{E}_{p/2}[X^{N_{\mathcal{A}'}^m(B)}], 1\}}{m2^m} \geq \frac{1}{m2^m}$$

Finally as $\mathcal{A}' \subset \mathcal{A}$ we get $\mu_p(N_{\mathcal{A}}^m(B)) \geq \mu_p(N_{\mathcal{A}'}^m(B)) \geq \frac{1}{m2^m}$ as needed. □

Now we are ready to present the structural result and its corollaries.

Theorem 9. *Let \mathcal{A} be a monotone with minterms of size at most k , and let $p \in [0, 1]$. Then at least one of the following two possibilities holds.*

1. *There exists a subfamily $\mathcal{B} \subseteq \mathcal{A}$, with $\mu_p(\mathcal{B}) \geq \mu_p(\mathcal{A})/2$, which is tame with respect to $p/2$.*
2. *There exists m between 1 and $k-1$, and a family \mathcal{B} which is a tame m -approximation of \mathcal{A} at $p/2$ with $\mu_p(\mathcal{B}) \geq \mu_p(\mathcal{A})/2^{m+1}$.*

By induction on the size of the minterms and application of the Theorem 9 and corollary 8 we deduce the following corollary, whose proof we defer until after the proof of the theorem.

Corollary 10. *Let \mathcal{A} be a monotone family of subsets of $[n]$ with all minterms of size at most k . Then*

$$\mu_{p/2}(\mathcal{A}) \geq \frac{1}{k2^{3k-1}} \mu_p(\mathcal{A}),$$

By repeated application of the corollary above one gets:

Corollary 11. *Let \mathcal{A} be a monotone family of subsets of $[n]$ with minterms of size at most k . Then:*

$$\delta_\epsilon(\mathcal{A}) \geq 1 - (2\epsilon)^{\frac{1}{3k + \log k - 1}}$$

Note, again, that this is weaker than what follows from Theorem 1.

Proof of Theorem 9. Let us iteratively define new families, one of which will be either a tame family or a tame approximation. Let $\mathcal{A}_1 := \mathcal{A}$, for each $1 \leq m \leq k - 1$:

$$\begin{aligned} \mathcal{B}_m &= \{V \subset [n] \text{ s.t. } |N_{\mathcal{A}_m}^m(V)| \geq (p/2)^{-m}\} \\ \mathcal{M}(\mathcal{A}_{m+1}) &= \mathcal{M}(\mathcal{A}_m) \setminus \{M \mid \exists V \in \mathcal{B}_m, V \subset M\} \end{aligned}$$

and let \mathcal{A}_{m+1} be the family spanned by $\mathcal{M}(\mathcal{A}_{m+1})$.

As $\mathcal{A} = \mathcal{A}_1 \supseteq \mathcal{A}_2 \cdots \supseteq \mathcal{A}_{k-1} \supseteq \mathcal{A}_k$ there are two possible options. Either there is some m for which $\mu_p(\mathcal{A}_m \setminus \mathcal{A}_{m+1})$ is large, or if for all m we have that $\mu_p(\mathcal{A}_m \setminus \mathcal{A}_{m+1})$ is small then $\mu_p(\mathcal{A}_k)$ is large.

Note that \mathcal{A}_k is tame with respect to $p/2$ as we removed all subsets $V \subset [n]$ that have many supplements (of any size.) If $\mu_p(\mathcal{A}_k) \geq \frac{1}{2}\mu_p(\mathcal{A})$ this gives us immediately the first case of the theorem.

If $\mu_p(\mathcal{A}_k) < \frac{1}{2}\mu_p(\mathcal{A})$ then there must be some m for which $\mu_p(\mathcal{A}_m) - \mu_p(\mathcal{A}_{m+1}) \geq \frac{1}{2^{m+1}}\mu_p(\mathcal{A})$. Let us show that in this case \mathcal{B}_m is a tame m -approximation as guaranteed in the second case of the theorem.

Taking \mathcal{A}' to be $\mathcal{A}_m \subset \mathcal{A}$ we see that from the definition of \mathcal{B}_m one has that for any $B \in \mathcal{B}_m$ $|N_{\mathcal{A}_m}^m(B)| > (p/2)^{-m}$ so we only need to show that $N_{\mathcal{A}_m}^m(B)$ is tame. Indeed, denote \mathcal{N} to be the family spanned by $N_{\mathcal{A}_m}^m(B)$ and assume there is some $U \subset [n]$ and some $l < m$ such that $|N_{\mathcal{N}}^l(U)| \geq (p/2)^{-l}$, then $|N_{\mathcal{A}_m}^l(U \cup B)| \geq |N_{\mathcal{N}}^l(U)| \geq (p/2)^{-l}$ and thus $B \cup U \in \mathcal{B}_l$ for $l < m$ in contradiction to the definition of \mathcal{A}_m .

Finally note that $\mu_p(\mathcal{B}_m) = \mu_p(\mathcal{A}_m) - \mu_p(\mathcal{A}_{m+1}) \geq \frac{1}{2^{m+1}}\mu_p(\mathcal{A})$. From this and the above \mathcal{B} is as guaranteed in the second case of the theorem. □

Proof of Corollary 10. We will use induction on the size of the minterms. For $k = 1$ we note that \mathcal{A} is tame by definition, so we can directly apply Lemma 7 and together with the first moment we get the required inequality:

$$\mu_{p/2}(\mathcal{A}) \geq \frac{\min\{\mathbb{E}_{p/2}[X^{\mathcal{A}}], 1\}}{2} \geq \frac{\min\{\mathbb{E}_p[X^{\mathcal{A}}], 1\}}{4} \geq \frac{1}{4}\mu_p(\mathcal{A})$$

Now assume we have proved for any $\ell < k$ and let us proof for k . Theorem 9 gives us two options. If we have the first one, then there is a tame family \mathcal{B} which is a subfamily of \mathcal{A} and $\mu_p(\mathcal{B}) \geq \frac{1}{2}\mu_p(\mathcal{A})$. Then together with Corollary 8 we have:

$$\mu_{p/2}(\mathcal{A}) \geq \mu_{p/2}(\mathcal{B}) \geq \frac{\mu_p(\mathcal{B})}{k2^k} \geq \frac{\mu_p(\mathcal{A})}{k2^{k+1}}$$

which is stronger than the required inequality.

If we are in the second case of the theorem note that any $B \in \mathcal{M}(\mathcal{B})$ has supplements of size m and so the size of each minterm in \mathcal{B} is at most $k - m$. Thus we can use the induction assumption on \mathcal{B} and get that $\mu_{p/2}(\mathcal{B}) \geq \frac{1}{(k-m)2^{3(k-m)-1}} \mu_p(B)$.

Recalling that from Theorem 9 $\mu_p \mathcal{B} \geq \frac{1}{2^{m+1}} \mu_p(\mathcal{A})$, it is left to apply Corollary 8 and get:

$$\mu_{p/2}(\mathcal{A}) \geq \frac{1}{m2^m} \mu_{p/2}(\mathcal{B}) \geq \frac{1}{m(k-m)2^{3k-2m-1}} \mu_p(B) \geq \mu_p(\mathcal{A}) \frac{1}{m(k-m)2^{3k-m}}$$

a simple calculation will give us the fact that $\frac{1}{m(k-m)2^{3k-m}} \geq \frac{1}{k2^{3k-1}}$ and so we get :

$$\mu_{p/2}(\mathcal{A}) \geq \frac{1}{k2^{3k-1}} \mu_p(\mathcal{A})$$

as required □

5 Fractional expectation and the expectation threshold

The technique we apply in this section, using duality of linear programming, follows an idea presented by Talagrand in the same context (see “weakly p -small” vs. “ p -spread” in [16]). This leads to a calculation of a weighted-second-moment, as done by Lyons in [12].

We begin by proving Theorem 4.

Proof: Let $\mathcal{A} \subseteq P([n])$ be a monotone family with $\mathcal{M} = \mathcal{M}(\mathcal{A})$ as a set of minterms, all of which have size at most k . Note that whenever $f : P([n]) \rightarrow \mathbb{R}^{\geq 0}$ assumes values greater than 1 on all $A \in \mathcal{A}$ then $\mathbb{E}(f) \geq \mu(\mathcal{A})$, and whenever $g : P([n]) \rightarrow \mathbb{R}^{\geq 0}$ has its support contained in \mathcal{A} then by Payley-Zygmund,

$$\mu(\mathcal{A}) \geq \frac{\mathbb{E}(g)^2}{\mathbb{E}(g^2)}. \tag{2}$$

To get good control on $\mu(\mathcal{A})$ it makes sense to try and find such a function f which is as small as possible, and a function g for which the second moment is well behaved (say, not too much weight on the upset generated by any single set, a quantity that arises naturally when calculating the second moment.) The trick will be to relate these two functions via LP duality. First, for $q \in [0, 1]$, define

$$\mathbb{E}_q^*(\mathcal{A}) = \min \sum \beta(B) q^{|B|},$$

where the minimum is taken over all functions $\beta : P([n]) \rightarrow \mathbf{R}^{\geq 0}$ such that

$$\sum_{B \subseteq A} \beta(B) \geq 1 \text{ for all } A \in \mathcal{M}. \tag{3}$$

By LP duality $\mathbb{E}_q^*(\mathcal{A}) = L_q^*(\mathcal{A})$ where

$$L_q^*(\mathcal{A}) = \max \sum_{A \in \mathcal{M}} \nu(A),$$

where the maximum is taken over all functions $\nu : \mathcal{M}(\mathcal{A}) \rightarrow \mathbf{R}^{\geq 0}$ such that

$$\sum_{B \subseteq A} \nu(A) \leq q^{|B|} \text{ for all } B. \tag{4}$$

Now, for any $p, q \in (0, 1)$ we let $\alpha = q/p$ and proceed to relate μ_p and \mathbb{E}_q^* . Let ν be a function achieving the maximum in the definition of $L_q^*(\mathcal{A})$ and define

$$g(X) = \sum_{A \in \mathcal{M}} \nu(A) p^{-|A|} \mathbf{1}_{A \subseteq X}$$

and note that

$$\mathbb{E}_p(g) = L_q^*(\mathcal{A}). \tag{5}$$

So to complete the calculation of a Payley-Zygmund type lower bound on $\mu_p(A)$ what is left is to calculate $\mathbb{E}_p(g^2)$. (This is the ‘‘weighted second moment’’ calculation, as in [12].) With A, B running over $\mathcal{M}(\mathcal{A})$ we have

$$\begin{aligned} \mathbb{E}_p(g^2) &= \sum \sum \nu(A) \nu(B) p^{-|A \cap B|} \\ &\leq \sum_{I \subseteq [n]} p^{-|I|} \left(\sum_{A \supseteq I} \nu(A) \right)^2 \\ &\leq \sum_i p^{-i} \left(\max_{|I|=i} \sum_{A \supseteq I} \nu(A) \right) \left(\sum_{|I|=i} \sum_{A \supseteq I} \nu(A) \right) \\ &\leq \sum_i p^{-i} (\alpha p)^i \sum_A \nu(A) \binom{|A|}{i} \end{aligned} \tag{6}$$

$$\begin{aligned} &\leq \sum_A \nu(A) \sum_i \alpha^i \binom{|A|}{i} \\ &\leq \sum_A \nu(A) (1 + \alpha)^{|A|} \\ &\leq L_q^*(\mathcal{A}) (1 + \alpha)^k \end{aligned} \tag{7}$$

where (6) follows from (4), and (7) follows from the definition of ν and the fact that all minterms are of size at most k . We now use (7) and (5) in (2), together with the fact that $L^* = \mathbb{E}^*$:

$$\mu_p(\mathcal{A}) \geq \frac{\mathbb{E}(g)^2}{\mathbb{E}(g^2)} \geq \frac{(L_q^*(\mathcal{A}))^2}{L_q^*(\mathcal{A})(1 + \alpha)^k} = \mathbb{E}_{\alpha p}^*(\mathcal{A})(1 + \alpha)^{-k}.$$

□

A nice feature of Theorem 4 is that it gives sufficient control over the rate of change of μ_p (as a function of p) to give both lower and upper bounds. This is embodied in the following corollary which also implies theorem 2.

Corollary 12. *Let $b < a$ and for a fixed non-trivial family \mathcal{A} and any $x \in [0, 1]$ let p_x be the unique probability that $\mu(\mathcal{A}, p_x) = x$. Then*

$$\left(\frac{a}{b}\right)^{1/k} - 1 \leq \frac{p_a}{p_b} \leq \frac{k^k}{(k-1)^{k-1}} \frac{a}{b}.$$

Note that setting $b := \epsilon, a := 1/2$ implies Theorem 2. Also, note that Theorem 1 yields the bound $\left(\frac{a}{b}\right)^{1/k} \leq \frac{p_a}{p_b}$, so that when a/b is large this is almost as good.

Proof: For the lower bound observe that Theorem 4 implies for any p and α

$$\frac{\mu_{\alpha p}}{(1+\alpha)^k} \leq \mu_p.$$

Setting $p := p_b$ and $\alpha := p_a/p_b$ gives the required result.

For the upper bound it is useful to use the inverse function to \mathbb{E}^* . Let q_x be the value q for which $\mathbb{E}_q^*(\mathcal{A}) = x$. Theorem 4 implies for any x and α

$$p_x \leq \frac{q_x(1+\alpha)^k}{\alpha}.$$

Also, it is easy to see that for $y \leq x$ it holds that $q_x \leq q_y \frac{x}{y}$. Furthermore, for every $x \leq 1$ we have $q_x \leq p_x$. Putting these together gives

$$p_a \leq \frac{q_a(1+\alpha)^k}{\alpha} \leq q_b \frac{(1+\alpha)^k a}{\alpha b} \leq p_b \frac{(1+\alpha)^k a}{\alpha b}.$$

The function $\frac{(1+\alpha)^k}{\alpha}$ is minimized at $\alpha = 1/(k-1)$. Plugging this value of α into the above expression yields the result. \square

Acknowledgements

The authors wish to thank Oliver Riordan for pointing out to us Theorem 1 and its proof. We would also like to thank Gil Kalai for useful conversations.

References

- [1] B. Bollobás, Threshold functions for small subgraphs. *Math. Proc. Cambridge Philos. Soc.* 90 (1981), no. 2, 197–206.
- [2] B. Bollobás, O. Riordan, A short proof of the Harris-Kesten Theorem, *Bulletin of the London Mathematical Society* 38 (2006), 470-484
- [3] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson and N. Linial, The influence of variables in product spaces, *Israel J. Math.* 77(1992), 55-64.
- [4] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* 7 (1987), 35–38.
- [5] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959), 290–297.

- [6] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* 5 (1960), 17-61.
- [7] E. Friedgut, Sharp thresholds of graph properties, and the k -sat problem. *Journal of the American Math. Soc.*, 12 (1999), no. 4, 1017–1054.
- [8] E. Friedgut, G. Kalai, Every monotone graph property has a sharp threshold, *Proc. Amer. Math. Soc.* 124 (1996), pp. 2993-3002.
- [9] H. Hatami, A structure theorem for Boolean functions with small total influences, To appear in *Annals of Math.*
- [10] J. Kahn, G. Kalai, Thresholds and expectation thresholds., *Combinatorics Probability and Computing* 16 (2007), no. 3, pp. 495–502
- [11] J. Kahn, G. Kalai, and N. Linial, The influence of variables on Boolean functions, *Proc. 29-th Ann. Symp. on Foundations of Comp. Sci.*, 68-80, Computer Society Press, 1988.
- [12] R. Lyons, Random walks and percolation on trees, *Ann. Probab.* **18** (1990), 931-958
- [13] M. Karonski, A. Rucinski, On the number of strictly balanced subgraphs of a random graph. *Graph theory (Lagow, 1981)*, 79–83, *Lecture Notes in Math.*, 1018, Springer, Berlin, 1983.
- [14] G. Margulis, Probabilistic characteristics of graphs with large connectivity, *Prob. Peredachi Inform.* 10(1974), 101-108.
- [15] L. Russo, An approximate zero-one law, *Z. Wahrsch. verw. Gebiete*, 61 (1982), 129-139.
- [16] M. Talagrand, Are many small sets explicitly small?, *Proceedings of the forty-second ACM symposium on Theory of computing*, 13-36, 2010.