Scaling limits for the threshold window: When does a monotone Boolean function flip its outcome?

Daniel Ahlberg and Jeffrey E. Steif

May 29, 2014

Abstract

Consider a monotone Boolean function \( f : \{0, 1\}^n \to \{0, 1\} \) and the canonical monotone coupling \( \{\eta_p : p \in [0, 1]\} \) of an element in \( \{0, 1\}^n \) chosen according to product measure with intensity \( p \in [0, 1] \). The random point \( p \in [0, 1] \) where \( f(\eta_p) \) flips from 0 to 1 is often concentrated near a particular point, thus exhibiting a threshold phenomenon. For a sequence of such Boolean functions, we peer closely into this threshold window and consider, for large \( n \), the limiting distribution (properly normalized to be nondegenerate) of this random point where the Boolean function switches from being 0 to 1. We determine this distribution for a number of the Boolean functions which are typically studied; it turns out that these limiting distributions have quite varying behavior. In fact, we show that any nondegenerate probability measure on \( \mathbb{R} \) arises in this way for some sequence of Boolean functions.

Keywords. Boolean functions; sharp thresholds; influences; iterated majority function

1 Introduction

It has been known for quite some time that typical events involving many independent random variables exhibit “thresholds” in the sense that the probability of the given event changes sharply as the parameter of the independent random variables changes. Observations of this kind were first made in the context of random graphs by Erdős and Rényi [4]. A more general understanding of the existence of threshold phenomena has since then been obtained through a series of papers. For instance, Russo [13] showed that a monotone event defined in terms of a family of independent Bernoulli variables exhibits a threshold if its dependence on each variable is small. Russo’s result was later refined by Talagrand [15]. The first estimates on the “sharpness” of the threshold were obtained by Friedgut and Kalai [6], critically building on work originating from Kahn, Kalai and Linial [10]. Related results also appeared in [2], [5] and elsewhere.

Less is known when it comes to closer inspections of the “threshold window”. Although the windows corresponding to certain graph properties are well understood, there is to our knowledge no general study of this transition. We aim with the present paper to offer a unified perspective on threshold transitions, and show that these transitions present quite varying behavior.

Let \( f : \{0, 1\}^n \to \{0, 1\} \) be a monotone (increasing) Boolean function and assign \([0, 1]\)-uniform random variables \( \xi_1, \xi_2, \ldots, \xi_n \) to the elements of \([n] := \{1, 2, \ldots, n\}\). For \( p \in [0, 1] \), let \( \eta_p = \{i \in [n] : \xi_i \leq p\} \), and note that \( a) \) each \( i \in [n] \) is present in \( \eta_p \) with probability \( p \) independently for different \( i \); and \( b) \) \( i \in \eta_p \) implies \( i \in \eta_{p'} \) for \( p' > p \). Thus, \( \eta_p \) corresponds to an element \( \omega \in \{0, 1\}^n \) chosen according to product measure with intensity \( p \), in the sequel.
denoted by $\mathbb{P}_p$, and $(\eta_n)_{n \in [0,1]}$ constitutes the standard monotone coupling of elements in $\{0,1\}^n$ chosen according to $\mathbb{P}_p$, as $p$ varies between 0 and 1. We study the random point $p$ at which $f(\eta_n)$ changes from 0 to 1. For a given sequence of monotone Boolean functions $(f_n)_{n \geq 1}$, our goal will be to find the (nondegenerate) limiting distribution of this random point after proper normalization should it exist. In the most interesting examples, one has a threshold phenomenon where, for large $n$, $\mathbb{P}_p(f_n)$ goes from 0 to 1 within a very small interval, which results in this transition point having a degenerate limit; one then needs to renormalize (scale in) in order to obtain a nondegenerate limiting distribution. (If there is no threshold phenomenon, then this point will have a nondegenerate limit and no further analysis is made.) One of our goals is to describe the distribution of this random point for some commonly studied Boolean functions.

To be more precise, given a monotone Boolean function $f : \{0,1\}^n \to \{0,1\}$, we define the random variable

$$T(f) := \min\{p \in [0,1] : f(\eta_p) = 1\};$$

this is the point where $f$ switches from 0 to 1 in the canonical coupling. Given a sequence $(f_n)_{n \geq 1}$ of monotone Boolean functions $f_n : \{0,1\}^n \to \{0,1\}$, we want to find, if possible, normalizing constants $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ with the $a_n$’s nonnegative such that $a_n(T(f_n) - b_n)$ converges, as $n \to \infty$, to a nondegenerate limiting distribution, and to determine what that limit may be. Observe that for $x \in \mathbb{R}$,

$$\mathbb{P}(a_n(T(f_n) - b_n) \leq x) = \mathbb{P}(T(f_n) \leq b_n + x/a_n) = \mathbb{P}_{p_n}(f_n(\omega) = 1), \quad (1)$$

where $p_n = b_n + x/a_n$. Recall the theorem of types (see [3, Theorem 3.7.5]) which tells us that there is essentially only one way to normalize a sequence of random variables and that there is essentially at most one possible nondegenerate limiting distribution. (The word “essentially” in the latter part of the statement means “up to a change of variables of the form $x \mapsto ax + b$”.

**Notation.** When understood from the context, we will write $T_n$ for $T(f_n)$.

The dictatorship function, which for every $n$ outputs the value of the first coordinate of the input, clearly has that $T_n$ is uniformly distributed on the interval [0,1] for each $n$ and so no scaling is needed. A simple example where scaling is needed is the OR function which is 1 if and only if at least one bit is 1. In this case, it is immediate to check that $nT_n$ converges in distribution to a unit exponential random variable. The cases when nontrivial normalization is needed are exactly those covered in the next definition.

**Definition.** We say that $(f_n)_{n \geq 1}$ has a threshold if there exists a sequence $(p_n)_{n \geq 1}$ such that for all $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}_{p_n+\varepsilon}(f_n = 1) = 1 \quad \text{and} \quad \lim_{n \to \infty} \mathbb{P}_{p_n-\varepsilon}(f_n = 1) = 0.$$ 

This is equivalent to $T_n - p_n$ approaching 0 in distribution. (Often $p_n$ will not depend on $n$.)

We give an alternative description of $T(f)$ which will be useful to have in mind, in particular when we study the Boolean function known as “tribes”. Recall that a 1-witness for $f$ is a minimal set $W \subseteq [n]$ such that $\{\omega_i = 1 \text{ for all } i \in W\}$ implies $f(\omega) = 1$. Similarly, a 0-witness for $f$ is a minimal set $W \subseteq [n]$ such that $\{\omega_i = 0 \text{ for all } i \in W\}$ implies $f(\omega) = 0$. Witnesses may be used to characterize $T(f)$. Writing $\mathcal{W}^1$ for the set of 1-witnesses and $\mathcal{W}^0$ for the set of 0-witnesses for $f$, it is immediate to check that for any monotone Boolean function one has

$$T(f) = \min_{W \in \mathcal{W}^1} \max_{i \in W} \xi_i = \max_{W \in \mathcal{W}^0} \min_{i \in W} \xi_i. \quad (2)$$
We next briefly discuss what one should expect the normalizing constants \((a_n)_{n \geq 1}\) and \((b_n)_{n \geq 1}\) to be in typical situations. Certainly it is reasonable that \(b_n\) should be close to \(E[T_n]\). In cases where we have a threshold, heuristically, the size of the “threshold interval” around \(p_n\) where \(\mathbb{P}_p(f_n = 1)\) moves from being near 0 to being near 1 should be governed by \(\frac{d}{dp} \mathbb{P}_p(f_n = 1)\) evaluated at \(p = p_n\). The Margulis-Russo formula (see e.g. [9, Theorem III.1]) tells us that this is equal to the total influence at \(p_n\) (defined below). Therefore the total influence dictates what the scaling factor \(a_n\) should be. We mention that while this heuristic for the scaling works in most natural examples, it is certainly not true in general. For example, if \(f_n\) is the Boolean function which is the AND of majority (to be defined later) on \(n\) bits and dictator, then the total influence will be of order \(\sqrt{n}\) but no scaling is used to obtain a limit for \(T_n\).

**Definition.** Given a Boolean function \(f\) of \(n\) variables and a variable \(i \in [n]\), we say that \(i\) is pivotal for \(f\) for \(\omega\) if \(f(\omega) \neq f(\omega^i)\) where \(\omega^i\) is \(\omega\) but flipped in the \(i\)th coordinate. The influence of the \(i\)th bit with respect to \(p\), denoted by \(\text{Inf}_i^p(f)\), is defined by

\[
\text{Inf}_i^p(f) := \mathbb{P}_p(i \text{ is pivotal for } f)
\]

and the total influence with respect to \(p\) is defined to be \(\sum_{i \in [n]} \text{Inf}_i^p(f)\).

We now summarize some of the results that we will obtain. The paper will begin by analyzing the limiting distribution of \(T_n\) for the majority function (which will be normal), the tribes function (which will be a reverse Gumbel distribution) and certain properties associated to graphs, such as connectivity and clique containment. A connection between the tribes function and the coupon collector problem is discussed. These results, which are not difficult, are presented in Section 2.

Next, for the class of so-called iterated majority functions, the analysis of the limiting distribution of \(T_n\) (both its existence and its properties) requires somewhat more work and involves dynamical systems. This will be done in Section 3. The precise definition and result are as follows. Given an odd integer \(m \geq 3\), the iterated \(m\)-majority function is defined recursively on \(m^n\) bits as follows. One constructs an \(m\)-ary tree of height \(n\) and places 0’s and 1’s at the leaves. One takes the majority of the bits in each family of \(m\) leaves and thus obtains 0 and 1 values for the nodes at height \(n - 1\). One then continues iteratively until the root is assigned a value. This is defined to be the output of the function, which we denote by \(f_n\) (where \(m\) is implicit).

To state the result of the analysis for iterated majority, we let, for odd integers \(m \geq 3\),

\[
\gamma(m) := m \left(\frac{m - 1}{m - 1}\right) 2^{-(m-1)} \quad \text{and} \quad \beta(m) := \frac{\log m + 1}{\log \gamma(m)}.
\]

**Theorem 1.** Consider, for each odd integer \(m \geq 3\), iterated \(m\)-ary majority on \(m^n\) bits.

a) Then \(\gamma(m)^n(T_n - \frac{1}{2})\) converges in distribution, as \(n\) tends to infinity, to a random variable whose distribution \(F_m\) is absolutely continuous and fully supported on \(\mathbb{R}\). Moreover, for \(m \neq m'\), \(F_m\) and \(F_{m'}\) are not related by a linear change of variables.

b) There exist constants \(c_1 = c_1(m)\) and \(c_2 = c_2(m)\) in \((0, 1)\) so that for all \(k \geq 1\),

\[
c_1^{k \beta(m)} \leq \mathbb{P}(W_m \geq k) \leq c_2^{k \beta(m)},
\]

where \(W_m\) has distribution \(F_m\).
c) $\beta(m)$ is strictly increasing, taking values in the interval $(1, 2)$ and approaches 2 as $m \to \infty$.
d) The sequence $(F_m)_{m \geq 1}$ approaches, as $m \to \infty$, a centered Gaussian with variance $(2\pi)^{-1}$.

Remark. Note that parts b) and c) together state that the tails of $F_m$ are between those of an exponential and a Gaussian. The fact that $\beta(m)$ approaches 2 is consistent with part d).

One of the most interesting and studied sequence of Boolean functions correspond to percolation crossings of a square. The coming book [8] would not exist if it were not for this example. In Section 4, we state a result for this example. However, unlike Section 3 which is completely self-contained, Section 4 will appeal to recent highly nontrivial results.

In Section 4, we state a result for this example. However, unlike Section 3 which is completely self-contained, Section 4 will appeal to recent highly nontrivial results.

The coming book [8] would not exist if it were not for this example.

Theorem 2. Given any probability measure $\mu$ on $\mathbb{R}$, and any sequence $(a_n)_{n \geq 1}$ satisfying $1 \ll a_n \ll \sqrt{n}$, there exists a sequence $(f_n)_{n \geq 1}$ of monotone Boolean functions $f_n : \{0, 1\}^n \to \{0, 1\}$ for which $a_n(T_n - \frac{n}{2})$ approaches $\mu$ in distribution.

We remark that we cannot in general choose the sequence of scaling coefficients to be growing faster than $\sqrt{n}$. Indeed, it is well known that for monotone Boolean functions the total influence, for $p$ bounded away from 0 and 1, is of order at most $\sqrt{n}$. The Margulis-Russo formula then implies that the scaling coefficient $a_n$ cannot be growing faster than this, assuming that we are centering around a value $b_n$ which is bounded away from 0 and 1. Also in this vein, we remark that the centralizing coefficient of Theorem 2 could in greater generality be replaced by any sequence $(b_n)_{n \geq 1}$ bounded away from 0 and 1, although this may be of less interest.

The fact that there is scaling at all involved is essential. Not all distributions on $[0, 1]$ can arise without any scaling, as the following proposition, also proved in Section 5, shows.

Proposition 3. There exist probability measures $\mu$ on $[0, 1]$ for which there is no sequence $(f_n)_{n \geq 1}$ of monotone Boolean functions so that $T_n$ approaches $\mu$.

Interestingly, there are sequences of nondegenerate random variables $(X_n)_{n \geq 1}$ which are not renormalizable in the sense that for no subsequence $(X_{nk})_{k \geq 1}$ are there normalizing constants $(a_k)_{k \geq 1}$ and $(b_k)_{k \geq 1}$ with the $a_k$’s nonnegative so that $a_k(X_{nk} - b_k)$ converges, as $k \to \infty$, to a nondegenerate limiting distribution. A typical example of such a sequence is given by $X_n = e^{nZ}$ where $Z$ is a standard normal random variable. The vague idea is that when we try to scale down to keep mass from going to infinity, then the result will be that all the mass is accumulating at 0. Another example, which we will exploit, is when $X_n$ is uniformly distributed on $\{\pm 2^k : k = 1, 2, \ldots, n\}$.

The following proposition, also proved in Section 5, shows that one cannot necessarily extract a subsequence of $(T_n)_{n \geq 1}$ which after normalization converges to a nondegenerate limit.

Proposition 4. There exists a (nondegenerate) sequence $(f_n)_{n \geq 1}$ of monotone Boolean functions $f_n : \{0, 1\}^n \to \{0, 1\}$ so that no subsequence of $(T_n)_{n \geq 1}$ can be renormalized to have a nondegenerate limiting distribution.
2 Some elementary examples

2.1 Majority and the standard normal

Majority is an example providing nontrivial, although classical, scaling behavior. The majority function on \( n \) bits is defined to output the value 1 if there are at least \( n/2 \) bits with the value 1. More generality, we consider biased majority, which is the function with output 1 if and only if there are at least \( pn \) bits valued 1, where \( p \in (0, 1) \) is a fixed parameter. The correct scaling factor will be of order \( \sqrt{n} \) and the limit will be Gaussian, as stated in the following proposition.

**Proposition 5.** For every \( p \in (0, 1) \) we have for the \( p \)-biased majority function on \( n \) bits that \( \sqrt{n} p (1-p)(T_n - p) \) converges in distribution to a standard normal.

Note that the multiplicative scaling is of order \( \sqrt{n} \) and coincides with the order of the total influence at the relevant parameter \( p \); hence it is consistent with the heuristic described above.

**Proof.** Let \( x \in \mathbb{R} \) and set \( p_n = p + x \sqrt{p(1-p)/n} \). For large \( n \) we have \( p_n \in [0, 1] \), and

\[
P \left( \sqrt{n} \frac{p}{p(1-p)}(T_n - p) \leq x \right) = P(T_n \leq p_n) = P_{p_n} \left( \sum_{i=1}^{n} \omega_i \geq np \right).
\]

We of course have a sum of \( n \) Bernoulli variables with success probabilities \( p_n \).

A consequence of the Lindeberg-Feller central limit theorem (see e.g. [3, Theorem 3.4.5]) is that if \( \{X_{i,n} : 1 \leq i \leq n, n \geq 1\} \) is a family of bounded random variables, such that for each \( n \), \( \{X_{i,n} : 1 \leq i \leq n\} \) are i.i.d. with zero mean and variance that tends to 1 as \( n \) increases, then \( \sum_{i=1}^{n} X_{i,n}/\sqrt{n} \) converges in distribution to a standard normal.

Since \( \text{Var}_{p_n}(\omega_i) = p_n(1-p_n) \), which tends to \( p(1-p) \), and \( (np - np_n)/\sqrt{np(1-p)} = -x \), the above consequence of the Lindeberg-Feller theorem implies that, as \( n \to \infty \),

\[
P_{p_n} \left( \sum_{i=1}^{n} \omega_i \geq np \right) = P_{p_n} \left( \sum_{i=1}^{n} \frac{\omega_i - p_n}{\sqrt{np(1-p)}} \geq \frac{np - np_n}{\sqrt{np(1-p)}} \right) \to \Phi(x),
\]

the distribution function of a standard normal distribution. \( \square \)

2.2 Tribes, Gumbel and coupon collectors

The tribes function on \( n \) bits is defined as follows. Given \( \ell_n \), partition \( [n] \) into \( \lfloor n/\ell_n \rfloor \) sets (‘tribes’) of length \( \ell_n \) (plus some residual bits). Then \( f_n(\omega) = 1 \) if and only if \( \omega \) is all 1’s for at least one tribe. The correct choice for \( \ell_n \), in order for the distribution of to be nondegenerate for the uniform measure, is of order \( \log_2 n - \log_2 \log_2 n \).

**Proposition 6.** Consider tribes with \( \ell_n = \lfloor \log_2 n - \log_2 \log_2 n \rfloor \), set \( \alpha_n = (\log_2 n - \log_2 \log_2 n)/\ell_n \). Then for all \( x \in \mathbb{R} \) we have

\[
\lim_{n \to \infty} P \left( 2(\log_2 n)(T_n - (\frac{1}{2})^{\alpha_n}) \leq x \right) \to 1 - \exp(-e^x).
\]
Note that the multiplicative scaling is of order \( \log_2 n \), which can be checked to be the order of the total influence at the relevant parameter \( 1/2 \); again, this is consistent with the heuristic described in the introduction. Note also that the upper tail of this limiting distribution decays super-exponentially, whereas the lower tail just decays exponentially.

**Proof.** Note that the \( \lceil n/\ell_n \rceil \) tribes of length \( \ell_n \) corresponds to the 1-witnesses for the tribes function. Let \( X_n \) denote the number of tribes (1-witnesses) for which \( \omega \) is all 1. For \( \omega \sim \mathbb{P}_p \), we see that \( X_n \) is binomially distributed with parameters \( \lceil n/\ell_n \rceil \) and \( p^\ell_n \). Clearly, \( \lceil n/\ell_n \rceil \to \infty \) as \( n \to \infty \). Also, for all \( x \in \mathbb{R} \) we have

\[
\lceil n/\ell_n \rceil \left( \frac{1}{2} \right)^{\alpha_n \ell_n} \frac{x}{2 \log_2 n} = \lceil n/\ell_n \rceil \left( 1 + \frac{x}{2^{1-\alpha_n \log_2 n}} \right)^{\ell_n} \to e^x
\]

as \( n \to \infty \). Given \( x \in \mathbb{R} \) and letting \( p_n = (\frac{1}{2})^{\alpha_n \ell_n} + \frac{x}{2 \log_2 n} \), we therefore have, by the Poisson convergence theorem (see e.g. [3, Theorem 3.6.1]), that for \( \omega \sim \mathbb{P}_{p_n} \), \( X_n(\omega) \) converges in law to a Poisson distribution with parameter \( e^x \). Since for each \( n \),

\[
\mathbb{P}\left( T_n \leq \left( \frac{1}{2} \right)^{\alpha_n \ell_n} + \frac{x}{2 \log_2 n} \right) = \mathbb{P}_{p_n}(X_n \geq 1),
\]

we thus conclude that

\[
\lim_{n \to \infty} \mathbb{P}\left( T_n \leq \left( \frac{1}{2} \right)^{\alpha_n \ell_n} + \frac{x}{2 \log_2 n} \right) = 1 - \exp(-e^x),
\]

as we needed to show. \( \square \)

**Remark.** The unfortunate term \( \alpha_n \) arises due to the fact that \( \log_2 n - \log_2 \log_2 n \) is not an integer. A related fact is that if \( p = 1/2 \), then the number of tribes which are identically 1 has all Poisson distributions with parameter in \([1, 2]\) as subsequential limits. If we were to restrict ourselves to \( n \)'s of the form \( k2^k \), then \( \log_2 n - \log_2 \log_2 n \) would be an integer and we would have the simpler form that

\[
\lim_{n \to \infty} \mathbb{P}\left( 2(\log_2 n)(T_n - \frac{1}{2}) \leq x \right) \to 1 - \exp(-e^x)
\]

along this thin subsequence of \( n \).

The reader might recognize the limiting distribution obtained in Proposition 6. In general if \( X \) has distribution \( F(x) \), then \( -X \) has distribution \( 1 - F(-x) \). If \( Y \) is distributed according to the above limiting distribution, then \( -Y \) has distribution \( \exp(-e^{-x}) \) which is known as the standard Gumbel distribution. This distribution often arises in extreme value theory and in particular is the limiting distribution after proper normalization of \( a) \) the maximum of \( n \) independent unit exponential random variables (where one subtracts \( \log n \) but uses no scaling factor to normalize); and \( b) \) the number of picks needed to collect \( n \) coupons when each pick is uniform (where one subtracts \( n \log n \) and divides by \( n \) to normalize). Heuristically, the reason that one gets the same limiting distribution in these two models is that in the latter case, we have the maximum of \( n \) weakly dependent geometric random variables with parameters \( 1/n \). When dividing by \( n \) (which explains the difference of a factor of \( n \) in the two normalizations), the geometric random variables become unit exponentials in the limit.
While we will not give an alternative proof of Proposition 6 based on these ideas, we want to explain why we obtained the limiting distribution there that we did.

Given a Boolean function \( f \), define its reversal \( \hat{f} \) by \( \hat{f}(\omega) = 1 - f(1 - \omega) \) and observe that \( \hat{f} \) is also a monotone Boolean function. One immediately checks that \( T(\hat{f}) \) and \( 1 - T(f) \) have the same distribution. If \( f_n \) is our tribes function, then this distributional relationship and Proposition 6 easily yields that

\[
2(\log_2 n)(T(\hat{f}_n) - 1 + \left(\frac{1}{2}\right)^{\alpha n})
\]

converges to the standard Gumbel distribution. We now give a heuristic for this. Clearly, \( \hat{f}_n \) is the function which is 0 if and only if there is a tribe which is all 0’s. (The tribes are 0-witnesses for \( \hat{f}_n \).) One easily checks that \( T(\hat{f}_n) \) is the smallest \( p \) such that each tribe has a 1 in it with respect to \( \eta_p \); compare with (2). The distribution of the time at which a given tribe gets its first 1 is equal to the distribution of the minimum of \( \ell_n \) uniform random variables. The minimum of \( k \) uniform random variables after multiplying by \( k \) converges to a unit exponential. Therefore, since different tribes are disjoint (and hence their corresponding uniform random variables are independent) and have size \( \ell_n \), it follows that \( \ell_n T(\hat{f}_n) \) is approximately the maximum of \( [n/\ell_n] \) unit exponential random variables. Therefore one should have that

\[
\ell_n T(\hat{f}_n) - \log([n/\ell_n]) = \ell_n \left(T(\hat{f}_n) - \frac{\log([n/\ell_n])}{\ell_n}\right)
\]

converges to the Gumbel distribution. This is certainly close to (3) and heuristically explains the reverse Gumbel distributional limit.

Remark. The so-called circular tribes function is a more symmetric version of tribes and perhaps more natural. It is defined as follows. We place the \( n \) bits in a circle and define \( f_n(\omega) \) to be 1 if \( \omega \) contains an interval of 1’s of length \( \lfloor \log_2 n \rfloor \). One can prove in a similar manner that the corresponding sequence \( T_n \) also has the reverse Gumbel distribution as a limit. The situation is however slightly different than for tribes since the number of such intervals containing all 1’s is no longer Poisson but rather compound Poisson, where the summands are mean 2 geometric random variables.

2.3 Random graph properties

In this subsection we cover a few monotone functions related to random graphs. We remind the reader that a random graph on \( n \) vertices is obtained by declaring each of the possible \( \binom{n}{2} \) edges open with probability \( p \in (0, 1) \). Equivalently, this amounts to determining an element \( \omega \in \{0, 1\}^{\binom{n}{2}} \) according to \( p_p \). We first discuss two functions whose critical values occur near 0. The proof of the following proposition is very straightforward (when using well known results) and hence we only sketch the proof.

**Proposition 7.**

a) Let \( f_n \) be the function corresponding to containing a triangle in a graph with \( n \) vertices. Then, for all \( x \geq 0 \), we have that

\[
\lim_{n \to \infty} \mathbb{P}(nT_n \leq x) = 1 - \exp(-x^3/6).
\]
b) Let \( f_n \) be the function corresponding to a graph with \( n \) vertices being connected. Then, for all \( x \in \mathbb{R} \), we have that
\[
\lim_{n \to \infty} \mathbb{P}(nT_n - \log n \leq x) = \exp(-e^{-x}).
\]

The multiplicative scaling \( n \) can in both cases be checked to be the order of the total influence at the relevant parameter.

Proof. a) It is well known (see e.g. [1, Theorem 4.1]) that if \( p = x/n \), then the number of triangles contained in the random graph converges to a Poisson distribution with parameter \( x^3/6 \). The result follows immediately using (1).

b) It is well known (see e.g. [1, Theorem 7.3]) that for any \( x \in \mathbb{R} \), if \( p = (\log n + x)/n \), the probability that the random graph is connected approaches \( \exp(-e^{-x}) \). The result follows.

A clique is a maximal complete subgraph of a graph. At a given parameter \( p \in (0, 1) \), the expected number of complete subgraphs of size \( \ell \) of a random graph on \( n \) vertices falls abruptly from being very large to being very small, as \( \ell \) increases. As a consequence, the maximal clique size of a random graph is highly concentrated, with high probability equal to either of two consecutive values \( \ell_n - 1 \) or \( \ell_n \), where \( \ell_n = \ell_n(p) \). Using Stirling’s approximation one sees that this sequence must satisfy \( \ell_n \sim 2 \log_{1/p} n \). This is well known; see e.g. [1, Chapter 4].

We will be interested in the function encoding the existence of a clique of size \( \ell_n \). For most values of \( n \) the maximal sized clique consists of \( \ell_n \) vertices with probability close to 1. However, along certain subsequences this probability remains bounded away from 1. Instead of restricting to subsequences we may allow \( p \) to vary, similar to the case of tribes. We simply state this result without proof since the argument follows more or less the argument for tribes. One obtains the result by proving Poisson approximation for the number of complete graphs of a given size. While this is more involved than for tribes, it is proved in [1, Theorems 11.7 and 11.9].

**Proposition 8.** Let \( p \in (0, 1) \) and \( \ell_n = \ell_n(p) \) be the above mentioned sequence. Let \( p_1, p_2, \ldots \) be any sequence bounded away from 0 and 1 such that the limit
\[
\lambda := \lim_{n \to \infty} \left( \frac{n}{\ell_n} \right)^{p_n} \text{ exists in } (0, \infty).
\]

Then, for the Boolean function encoding the existence of a complete graph of size \( \ell_n \), we have
\[
\lim_{n \to \infty} \mathbb{P}\left( \frac{\ell_n^2}{2p_n} (T_n - p_n) \leq x \right) = 1 - \exp(-\lambda e^x) \quad \text{for } x \in \mathbb{R}.
\]

### 3 Iterated majority

In this section, we will analyze iterated majority and prove Theorem [1]. In order to understand the asymptotic behavior of iterated majority, one is led to study its recursive structure. The limiting distribution will be described through the iterates of some function \( g : [0, 1] \to [0, 1] \), and the appropriate scaling will be determined by the derivative of \( g \) at \( 1/2 \).
We begin by describing what the limiting distribution $F_m$ will be. Define $g : [0, 1] \to [0, 1]$ as the probability at parameter $x \in [0, 1]$ that the majority on $m$ bits equals 1. Formally, $g$ is given by

$$g(x) = \sum_{k=(m+1)/2}^{m} \binom{m}{k} x^k (1-x)^{m-k}.$$  

Observe that $\gamma(m)$, which will be our scaling coefficient satisfies

$$\gamma(m) := \frac{m}{m-1} \gamma(m-2). \tag{4}$$

It is clear that $\gamma(m)$ is increasing in $m$, and Stirling’s approximation says that $\gamma(m) \sim \sqrt{2m/\pi}$ as $m$ tends to infinity. It turns out that the total influence for $f_n$ is $\gamma(n)$ and we will below see that $\gamma$ coincides with the derivative of $g$ at $\frac{1}{2}$. The recursive structure of $\gamma(m)$ stated in (4) easily yields that $\gamma(m) < \frac{m}{2}$, implying in turn that $\beta(m) > 1$ for all $m \geq 3$. Also, using $\gamma(m) \sim \sqrt{2m/\pi}$, we find that $\beta(m) \to 2$ as $m \to \infty$.

It turns out to be convenient to consider the translate

$$h(x) = g(\frac{1}{2} + x) - \frac{1}{2}$$

of $g$. The scaling limit of iterated $m$-majority will be described in terms of the limit as $n \to \infty$ of $h^{(n)}(\alpha \gamma^{-n})$, where $\alpha \in \mathbb{R}$ and $h^{(n)}$ denotes the composition of $h$ with itself $n$ times.

We will break up the proof of the four parts of Theorem 1 into subsections.

### 3.1 Proof of part a)

We begin with the following proposition which will be central for our analysis.

**Proposition 9.** For every odd integer $m \geq 3$ and $\alpha \in \mathbb{R}$, the limit $L(\alpha) := \lim_{n \to \infty} h^{(n)}(\alpha \gamma^{-n})$ exists and the resulting function $L : \mathbb{R} \to (-\frac{1}{2}, \frac{1}{2})$ is odd, onto, 1-Lipschitz continuous, strictly increasing and continuously differentiable.

We first need the following lemma.

**Lemma 10.** The function $h : [-\frac{1}{2}, \frac{1}{2}] \to [-\frac{1}{2}, \frac{1}{2}]$ is odd, onto, strictly increasing, strictly convex on $[-\frac{1}{2}, 0]$ and strictly concave on $[0, \frac{1}{2}]$. In particular $h'(x) \leq h'(0) = \gamma(m)$ for $x \in [-\frac{1}{2}, \frac{1}{2}]$.

**Proof.** It suffices to demonstrate the corresponding characteristics for $g$. From the interpretation of $g$ as a probability, it is clear that $g$ is strictly increasing, maps $0, \frac{1}{2}$ and 1 to themselves, and that

$$g(x) = \mathbb{P}(\text{Bin}(m,x) \geq m/2) = 1 - \mathbb{P}(\text{Bin}(m,x) < m/2) = 1 - g(1-x).$$

Thus $h$ is odd, strictly increasing, has fixed points at $-\frac{1}{2}, 0$ and $\frac{1}{2}$, and is therefore also onto.

We know that $g$ is differentiable and we aim to determine its derivative. Note that

$$g'(1) = g'(0) = \lim_{x \to 0} \sum_{k \geq m/2} \binom{m}{k} x^{k-1} (1-x)^{m-k} = 0.$$
Next, pick δ > 0 and let ξ₁, ξ₂, . . . , ξₘ be independent and [0, 1]-uniformly distributed. Using the monotone coupling we find that

\[ g(x + δ(1 - x)) - g(x) = P\left(\#\{ξ_i ≤ x\} < m/2, \#\{ξ_i ≤ x + δ(1 - x)\} ≥ m/2\right). \]

Conditioning on the number of ξᵢ’s whose value is at most x we arrive at

\[ \sum_{k≤m/2} \binom{m}{k} x^k (1 - x)^{m-k} P\left(\#\{ξ_i ≤ x + δ(1 - x)\} ≥ m/2 \mid \#\{ξ_i ≤ x\} = k\right). \]

The above conditional probability coincides with the probability that a binomial random variable with parameters m − k and δ is at least m/2 − k, and is thus independent of x. In addition,

\[ δ^{-1} P(\text{Bin}(m, k, δ) ≥ m/2 − k) = \sum_{ℓ≥m/2−k} \binom{m−k}{ℓ} δ^{ℓ−1}(1 − δ)^{m−k−ℓ}, \]

and sending δ to 0 leaves us with m − k = (m + 1)/2 in case k = (m − 1)/2, and 0 for all smaller values of k. In conclusion, for x ∈ (0, 1),

\[ g'(x) = \lim_{δ→0} \frac{g(x + δ(1 - x)) - g(x)}{δ(1 - x)} = \frac{m + 1}{2} \left(\frac{m}{m - 1}\right) \left(x(1 - x)\right)^{m-1}. \]  

Differentiating once more gives

\[ g''(x) = \frac{m + 1}{2} \frac{m - 1}{2} \left(\frac{m}{m - 1}\right) \left(x(1 - x)\right)^{m-3}(1 - 2x). \]

In conclusion, the derivative of g is strictly positive on (0, 1), and the second derivative is strictly positive on (0, 1/2) and strictly negative on (1/2, 1). So h possesses the claimed properties and h′ reaches its maximum at the origin, which is easily seen to equal γ(m). \( \square \)

The proof of Proposition \[ \square \] will make repeated use of the properties of h displayed in Lemma \[ \square \]. For instance, we note that h cannot have any fixed points other than \( −1/2, 0 \) and \( 1/2 \).

**Proof of Proposition.** Since h(0) = 0 we also have L(0) = 0, and since h is odd the limit L(α), if it exists, has to be odd as well. In particular, it will be sufficient to consider α ≥ 0 for the rest of this proof.

**Existence.** Given α ≥ 0, choose n₀ such that \( αγ^{-n} ≤ 1/2 \) for all n ≥ n₀. Note that we may obtain \( h^{(n)}(αγ^{-n}) \) from \( αγ^{-(n+1)} \) by first multiplying by γ, and then applying h n times. \( h^{(n+1)}(αγ^{-(n+1)}) \) is similarly obtained from \( αγ^{-(n+1)} \) by first applying h once, and then another n times. Lemma \[ \square \] shows that the derivative of \( h \) is bounded by γ. Hence \( γx ≥ h(x) \) for all \( x ∈ [0, 1/2] \), and it follows that \( h^{(n)}(αγ^{-n}) \) is decreasing in n for n ≥ n₀. Since the sequence is bounded below by 0, the limit L(α) necessarily exists for all α ≥ 0.

**1-Lipschitz Continuity.** Using again that \( |h'| ≤ γ \), together with iterated use of the mean value theorem, we find for α, α′ ∈ \( \mathbb{R} \) that

\[ |L(α) - L(α')| = \lim_{n→∞} |h^{(n)}(αγ^{-n}) - h^{(n)}(α'γ^{-n})| ≤ \liminf_{n→∞} γ^n |αγ^{-n} - α'γ^{-n}| = |α - α'|, \]

10
where we also have used that $\alpha \gamma^{-n}$ and $\alpha' \gamma^{-n}$ are contained in $[-\frac{1}{2}, \frac{1}{2}]$ for large $n$.

An observation that will be important for the rest of this proof is that, by continuity of $h$, for all $\alpha \in \mathbb{R}$

$$h(L(\alpha)) = \lim_{n \to \infty} h^{(n+1)}(\alpha \gamma^{-n}) = L(\alpha \gamma).$$

(6)

Iterating this yields that

$$L(\alpha) = h(L(\alpha \gamma^{-1})) = h^{(n)}(L(\alpha \gamma^{-n})).$$

(7)

**Strict Monotonicity.** Note that weak monotonicity of course follows from $h$ being increasing. We will next aim to show that for $\alpha \geq \alpha' \geq 0$ sufficiently small, we have

$$L(\alpha) - L(\alpha') \geq (\alpha - \alpha') \prod_{k=1}^{\infty} (1 - \alpha \left(\frac{3}{4}\right)^k).$$

(8)

Apart from showing that $L$ is strictly increasing in a neighborhood around the origin, (8) will be an important step in the proof of differentiability of $L$. Note that strict monotonicity of $L$ would follow for all $\alpha \in \mathbb{R}$ by (7) and (8), since $h$ is strictly increasing.

We now deduce (8). Using concavity of $h$ on $[0, \frac{1}{2}]$, we observe that $\frac{3}{4} \leq \gamma - \frac{1}{4} \leq h'(x) \leq \gamma$ on some interval $[0, c]$, where $c = c(m) > 0$. So, by the mean value theorem we conclude that

$$\frac{3}{4} x \leq h(x) \leq \gamma x \text{ on } [0, c].$$

Consequently, for all $\alpha \in [0, c]$ and $1 \leq k \leq n$, we have $h^{(k)}(\alpha \gamma^{-n}) \leq \alpha$, and therefore

$$h^{(k)}(\alpha \gamma^{-n}) \leq \frac{3}{4} h^{(k+1)}(\alpha \gamma^{-n}) \leq \left(\frac{3}{4}\right)^{n-k} h^{(n)}(\alpha \gamma^{-n}) \leq \alpha \left(\frac{3}{4}\right)^{n-k}.$$

Now, for any $\alpha' \leq \alpha$ in $[0, c]$ and given $n$, we obtain, from iterated use of the mean value theorem, the existence of constants $\{s_k^n\}_{1 \leq k \leq n}$ with $s_k^n \in [h^{(k-1)}(\alpha' \gamma^{-n}), h^{(k-1)}(\alpha \gamma^{-n})]$ and such that

$$h^{(n)}(\alpha \gamma^{-n}) - h^{(n)}(\alpha' \gamma^{-n}) = (\alpha \gamma^{-n} - \alpha' \gamma^{-n}) \prod_{k=1}^{n} h'(s_k^n).$$

(9)

Since $h'$ is decreasing on $[0, \frac{1}{2}]$ and $h''(0) = 0$, we have that $h'(x)$ is bounded below by $\gamma(1 - x)$ on some, possibly smaller, interval $[0, c']$. As a consequence we obtain the lower bound on (9),

$$\left(\alpha \gamma^{-n} - \alpha' \gamma^{-n}\right) \prod_{k=1}^{n} \gamma(1 - h^{(k-1)}(\alpha \gamma^{-n})) \geq (\alpha - \alpha') \prod_{k=1}^{n} \left(1 - \alpha \left(\frac{3}{4}\right)^{n-k+1}\right)$$

$$= (\alpha - \alpha') \prod_{k=1}^{n} \left(1 - \alpha \left(\frac{3}{4}\right)^{n-k}\right) \geq (\alpha - \alpha') \prod_{k=1}^{\infty} \left(1 - \alpha \left(\frac{3}{4}\right)^{k}\right).$$

(10)

Combining (8) and (10) and letting $n \to \infty$ yields (8) for every $\alpha' \leq \alpha$ in $[0, c']$.

**Continuous Differentiability.** Using (7) we have for any $\alpha \geq 0$ and $\delta \in \mathbb{R}$ that

$$L(\alpha + \delta) - L(\alpha) = h^{(n)}(L((\alpha + \delta) \gamma^{-n})) - h^{(n)}(L(\alpha \gamma^{-n}))$$

$$= \left[L((\alpha + \delta) \gamma^{-n}) - L(\alpha \gamma^{-n})\right] \prod_{k=1}^{n} h'(h^{(k-1)}(L(\alpha_k \gamma^{-n}))).$$

11
Proposition 9. Using $h$ probability $a)$ to prove part of $h$ bounded between and where we in the last step have used the mean value theorem iteratively; the $\alpha_k$’s are bounded. By continuity and monotonicity of $h$ and $L$, these $\alpha_k$’s exist. Using $h^{(k-1)}(L(\alpha)) = L(\alpha \gamma^{k-1})$, and reindexing the terms of the product, we arrive at

$$
\frac{L(\alpha + \delta) - L(\alpha)}{\delta} = \frac{L((\alpha + \delta)\gamma^{-n}) - L(\alpha \gamma^{-n})}{\delta \gamma^{-n}} \prod_{k=1}^{n} \gamma^{-1} h'(L(\alpha_{n-k+1} \gamma^{-k})).
$$
(11)

We now want to take limits. First, since $L$ is 1-Lipschitz continuous,

$$
\limsup_{\delta \to 0} \frac{L(\alpha + \delta) - L(\alpha)}{\delta} \leq \prod_{k=1}^{n} \gamma^{-1} h'(L(\alpha \gamma^{-k})),
$$

which is decreasing in $n$. Second, we note that the infinite product in (8) tends to 1 as $\alpha \to 0$. Applying this to the first term in (11), we conclude that for every $\varepsilon > 0$, if $n$ is sufficiently large, then

$$
\liminf_{\delta \to 0} \frac{L(\alpha + \delta) - L(\alpha)}{\delta} \geq (1 - \varepsilon) \prod_{k=1}^{n} \gamma^{-1} h'(L(\alpha \gamma^{-k})).
$$

Sending $n$ to infinity, and then $\varepsilon$ to zero, we conclude that the inferior and superior limits coincide and that

$$
L'(\alpha) = \lim_{\delta \to 0} \frac{L(\alpha + \delta) - L(\alpha)}{\delta} = \prod_{k=1}^{\infty} \gamma^{-1} h'(L(\alpha \gamma^{-k})).
$$
(12)

Since $h' \leq \gamma$ the limit is finite, and since $L(\alpha \gamma^{-k}) \leq \alpha \gamma^{-k}$ and $h'(x) \geq \gamma(1 - x)$ for small $x \geq 0$, the limit is strictly positive for all $\alpha \in \mathbb{R}$. This, again, shows that $L$ is strictly monotone on $\mathbb{R}$.

We need to show that $L'$ is continuous, and note, based on (12), that $L'$ is decreasing on $[0, \infty)$ since $L$ is increasing. Since also $L' > 0$ on $\mathbb{R}$ it follows that

$$
\prod_{k=\ell}^{\infty} \gamma^{-1} h'(L(\alpha \gamma^{-k})) \to 1 \quad \text{as } \ell \to \infty
$$

uniformly on compact sets. Thus, for every $\varepsilon > 0$

$$
\left| \lim_{x \to \alpha} L'(x) - \prod_{k=1}^{\ell} \gamma^{-1} h'(L(\alpha \gamma^{-k})) \right| < \varepsilon
$$

for large enough $\ell$, showing that $\lim_{x \to \alpha} L'(x) = L'(\alpha)$.

Surjectivity. Since $L(0) = 0$ and $L$ is continuous, it remains to show that $L(\alpha) \to \frac{1}{2}$ as $\alpha \to \infty$. For any $k \in \mathbb{N}$ we have from (7) that $L(\alpha \gamma^k) = h^{(k)}(L(\alpha))$. Since $L(\alpha) > 0$ for $\alpha > 0$, the properties of $h$ imply that $h^{(k)}(L(\alpha)) \to \frac{1}{2}$ as $k \to \infty$. Together with the proven continuity of $L$ and its weak monotonicity, this shows that $L$ maps $[0, \infty)$ onto $[0, \frac{1}{2})$.

We now use the above to analyze the asymptotics of $T_n$ for iterated $m$-majority on $m^n$ bits to prove part a) of Theorem 1. With $x \in \mathbb{R}$ fixed, the goal will be to relate, for large $n$, the probability $\mathbb{P}(T_n \leq p_n)$, where $p_n = \frac{1}{2} + x \gamma^{-n}$, with the function $L : \mathbb{R} \to [-\frac{1}{2}, \frac{1}{2}]$ defined in Proposition 9.
Given $p \in (0, 1)$ and $n \in \mathbb{N}$, let
\[ q_n(p) := \mathbb{P}_p(f_n(\omega) = 1). \]
Using the iterated majority structure, it is easy to express $q_n(p)$ in terms of $q_{n-1}(p)$. Specifically, $q_n(p)$ is the probability that a binomial random variable with parameters $m$ and $q_{n-1}(p)$ is at least $m/2$. That is,
\[ q_n(p) = \mathbb{P}(\text{Bin}(m, q_{n-1}(p)) \geq m/2) = g(q_{n-1}(p)), \]
where $g$ is defined as above. Of course, the base case $q_0(p)$ equals the probability of the majority function on one bit being 1, which equals $p$. Iterating the above procedure we get $q_n(p) = g^{(n)}(p)$.
Replacing $p$ by $p_n = \frac{1}{2} + x\gamma^{-n}$ leaves
\[ \mathbb{P}(T_n \leq p_n) = \mathbb{P}_{p_n}(f_n(\omega) = 1) = q_n(p_n) = g^{(n)}(\frac{1}{2} + x\gamma^{-n}) = \frac{1}{2} + h^{(n)}(x\gamma^{-n}), \]
which, according to Proposition [9], converges as $n \to \infty$ to $\frac{1}{2} + L(x)$.
That the limiting distribution is absolutely continuous and fully supported on $\mathbb{R}$ is a consequence of the properties of $L$ established in Proposition [9]. The final sentence of part a) follows from parts b) and c).

Remark. Observe that the first equality in [7] has a very nice interpretation. It says that the two homeomorphisms $x \mapsto \gamma x$ from $\mathbb{R}$ to itself and $x \mapsto h(x)$ from $[-1/2, 1/2]$ to itself are conjugate and that $L$ provides a conjugation between these homeomorphisms (showing that they are conjugate). We know that $L$ is also a homeomorphism. Since the two conjugate mappings are analytic, one might guess that one can show that $L$ has good properties (such as analyticity or being continuously differentiable) by virtue of the fact it is a conjugacy between these systems. This is unfortunately not true. One can construct self-conjugacies $f$ of $x \mapsto \gamma x$ which, while being homeomorphisms, are not continuously differentiable. Being a self-conjugacy amounts to saying that $f(\gamma x) = \gamma f(x)$ and so in particular we are saying that $f(\gamma x) = \gamma f(x)$ does not imply that $f$ is linear even for homeomorphisms. Since $L \circ f$ would also be a conjugacy between the two systems, one cannot conclude good properties of $L$ using only the fact that it is a conjugacy.

### 3.2 Proof of part b)

We first need the following lemma.

**Lemma 11.** For every odd integer $m \geq 3$ there exists $\varepsilon_0 = \varepsilon_0(m) > 0$ such that for all $\ell \geq 1$ and $\varepsilon \in (0, \varepsilon_0)$,
\[ \varepsilon^{(m+1)/\ell} \leq \left| h^{(\ell)}(\frac{1}{2} - \varepsilon) - \frac{1}{2} \right| \leq (9\varepsilon)^{(m+1)/\ell}. \]

**Proof.** First note that
\[ \frac{1}{2} - h(\frac{1}{2} - \varepsilon) = 1 - g(1 - \varepsilon) = g(\varepsilon), \]
which after iteration leaves us with $\frac{1}{2} - h^{(\ell)}(\frac{1}{2} - \varepsilon) = \gamma h^{(\ell)}(\varepsilon)$. In addition, each term in $g$ with non-zero coefficient has degree at least $(m + 1)/2$. Thus,
\[ g(x) = \left( \frac{m}{m+1} \right)^{m+1} \left( 1 + r(x) \right) \]
for some polynomial \( r(x) \to 0 \) as \( x \to 0 \). Since

\[
\left( \frac{m}{m+1} \right)^{n_k} = \left( \frac{m}{m-1} \right) \left( \frac{2cm}{m-1} \right)^{m-1} \leq 9^{m-1},
\]

we see that \( g(\varepsilon) \) lies in \( [\varepsilon^{m+1}, (9\varepsilon)^{m+1}/9] \) for all sufficiently small \( \varepsilon \). Using the fact that \( g \) is increasing and \( g(x) \leq x \) on \( (0, \frac{1}{2}) \), it then follows by two inductions that \( \varepsilon^{(\alpha+1)/2} \) and \( (9\varepsilon)^{(m+1)/2}/9 \) are lower respectively upper bounds for \( g(\varepsilon) \).

Now, fix \( m \geq 3 \), and let \( \varepsilon_0 = \varepsilon_0(m) > 0 \) be given as in Lemma 11. We first show the second inequality. Since \( L \) approaches \( \frac{1}{2} \) continuously, we can choose \( a_0 > 0 \) so that \( L(a_0) = \frac{1}{2} - \frac{\varepsilon_0}{9} \).

Given an integer \( k \geq 1 \), let \( n_k := \lceil \log_{\gamma} \frac{k}{a_0} \rceil \). We restrict to \( k \)'s which are sufficiently large so that \( n_k \geq 1 \). This immediately yields

\[
a_0 \leq k\gamma^{-n_k}.
\]

Using (6) and monotonicity of \( L \) and \( h \), we have

\[
L(k) = h^{(n_k)}(L(k\gamma^{-n_k})) \geq h^{(n_k)}(L(a_0)).
\]

By Lemma 11 and the definition of \( a_0 \), we have that

\[
\mathbb{P}(W_m \geq k) = \frac{1}{2} - L(k) \leq \frac{1}{2} - h^{(n_k)}(L(a_0)) = \frac{1}{2} - h^{(n_k)}\left( \frac{1}{2} - \frac{\varepsilon_0}{9} \right) \leq \varepsilon_0^{(m+1)/2k}.
\]

An easy computation shows that, for all \( k \) for which \( n_k \geq 1 \)

\[
\left( \frac{m+1}{2} \right)^{n_k} \geq \frac{2}{m+1} \left( \frac{k}{a_0} \right)^{\beta(m)}.
\]

From this, the upper bound follows with \( c_2 = \varepsilon_0^0 \) for all large \( k \) and \( \rho = \frac{2}{m+1}a_0^{-\beta(m)} \). By increasing \( c_2 \) if necessary, one can of course get the desired inequality for all \( k \geq 1 \).

We now move to the lower bound. This time, choose \( a_0 > 0 \) so that \( L(a_0) = \frac{1}{2} - \frac{\varepsilon_0}{2} \), and given an integer \( k \geq 1 \), let \( n_k := \lfloor \log_{\gamma} \frac{k}{a_0} \rfloor \). We again restrict to \( k \)'s which are large so that \( n_k \geq 1 \). This immediately yields

\[
a_0 \geq k\gamma^{-n_k}.
\]

Using (7) and monotonicity of \( L \) and \( h \), we have

\[
L(k) = h^{(n_k)}(L(k\gamma^{-n_k})) \leq h^{(n_k)}(L(a_0)).
\]

By Lemma 11 we have that

\[
\mathbb{P}(W_m \geq k) = \frac{1}{2} - L(k) \geq \frac{1}{2} - h^{(n_k)}(L(a_0)) = \frac{1}{2} - h^{(n_k)}\left( \frac{1}{2} - \frac{\varepsilon_0}{2} \right) \geq \left( \frac{a_0}{2} \right)^{(m+1)/2k}.
\]

An easy computation shows that one has, for all \( k \) for which \( n_k \geq 1 \), that

\[
\left( \frac{m+1}{2} \right)^{n_k} \leq m+1 \left( \frac{k}{a_0} \right)^{\beta(m)}.
\]

From this, the lower bound follows for some \( c_1 \) for all large \( k \), and by decreasing \( c_1 \) if necessary, one can of course get the desired inequality for all \( k \geq 1 \). This proves part b).
3.3 Proof of part c)

We set out to show that \( \beta(m) \) is strictly increasing. Since

\[
\beta(m + 2) - \beta(m) = \frac{\log \frac{m+3}{2} \log \gamma(m) - \log \frac{m+1}{2} \log \gamma(m + 2)}{\log \gamma(m) \log \gamma(m + 2)},
\]

it will suffice to show that the numerator in the right-hand side of (13) is strictly positive. Using the recursive structure of \( \gamma(m) \), i.e., that \( \gamma(m + 2) = \frac{m+2}{m+1} \gamma(m) \), we aim to show that

\[
\log \gamma(m) \log \frac{m+3}{m+1} - \log \frac{m+1}{2} \log \frac{m+2}{m+1} > 0.
\]

For \( x \geq 0 \) a Taylor estimate for \( \log(1 + x) \) gives the lower and upper bounds \( x - \frac{x^2}{2} \) and \( x \), respectively. A lower bound on the numerator in the right-hand side of (13) is thus given by

\[
\log \gamma(m) \left[ \frac{2}{m+1} - \frac{2}{(m+1)^2} \right] - \frac{1}{m+1} \log \frac{m+1}{2}.
\]

A lower bound on \( \gamma(m) \) can be obtained from known bound on the central binomial coefficient. For instance, Wallis’ product formula states that \( a_n := \prod_{k=1}^{n} \frac{2k-1}{2k} \frac{2k}{2k+1} \) converges to \( \frac{\pi}{2} \) as \( n \to \infty \) (see e.g. [16] for an elementary proof). Since \( a_n \) is increasing we have \( a_n \leq \frac{\pi}{2} \) for all \( n \geq 1 \), leading to the bound

\[
\left( \frac{2n}{n} \right) \geq 4^n \sqrt{\frac{2}{\pi (2n+1)}}.
\]

Consequently \( \gamma(m) \geq \sqrt{2m/\pi} \) for all \( m \geq 3 \). After multiplication by \( m + 1 \), a lower bound on the expression in (14) is given by

\[
\log \frac{2m}{\pi} \left[ 1 - \frac{1}{m+1} \right] - \log \frac{m+1}{2} = \log \frac{4}{\pi} - \log \frac{m+1}{m} - \frac{1}{m+1} \log \frac{2m}{\pi}.
\]

One may check that the latter expression is increasing in \( m \) and positive for \( m = 13 \). Using the slightly sharper lower bound \( \frac{4^n}{\sqrt{\pi n}} (1 - \frac{1}{8n}) \) on the central binomial coefficient, obtained from the Stirling series, one may arrive at an alternative lower bound on the difference in (13), which is positive for all \( m \geq 5 \). In either case, one further checks that \( \beta(m) < \beta(m+2) \) for the remaining values of \( m \) by hand, so that \( \beta(m) \) is strictly increasing for all \( m \geq 3 \).

3.4 Proof of part d)

Since \( m \) will now be changing, it is natural to now write \( L_m \) instead of \( L \). The distribution given by \( F_m(x) = \frac{1}{2} + L_m(x) \) has density \( L_m'(x) \). We will show that \( \lim_{m \to \infty} L_m'(x) = e^{-\pi x^2} \), which we recognize as the density of a centered normal distribution with variance \( 1/(2\pi) \). By virtue of Scheffé’s theorem (see e.g. [3]), pointwise convergence of densities implies the desired weak convergence of \( F_m \) to a normal distribution. By symmetry, it suffices to prove this for \( x \geq 0 \) which we now assume to be the case.
We first show that $L'_m(x^{\gamma-1}) \to 1$ as $m \to \infty$, for every $x \geq 0$. First recall that by Proposition 9, we have that $L'_m(x) \leq 1$ for all $x \geq 0$. Using (12) and (5) and then that $L'_m \leq 1$, we find that
\[
L'_m(x^{\gamma-1}) = \prod_{k=1}^{\infty} \left[ 1 - 4[L_m(x^{\gamma-k-1})]^2 \right]^{m-1} \geq \prod_{k=1}^{\infty} \left[ 1 - 4x^{2\gamma-2(k+1)} \right]^{m-1}.
\]
Since $e^{-2y} \leq 1 - y$ for small positive $y$, replacing the terms of the product by an exponential, we obtain that for all large $m$
\[
L'_m(x^{\gamma-1}) \geq \prod_{k=1}^{\infty} \exp \left( -4x^2(m-1)\gamma^{-2(k+1)} \right) = \exp \left( -4x^2(m-1) \sum_{k=1}^{\infty} \gamma^{-2(k+1)} \right)
\geq \exp \left( -8x^2m\gamma^{-4} \right),
\]
where we in the last step have used that $\gamma^2 \geq 2$. Since $\gamma = \gamma(m)$ increases at the rate of $\sqrt{m}$, we may conclude that $L'_m(x^{\gamma-1}) \to 1$ as $m \to \infty$. Moreover, the convergence is uniform in $x$ over compact sets. Consequently, for every $\varepsilon > 0$ and $x \geq 0$ there is $m_0$ such that $L_m(x^{\gamma-1}) \geq (1 - \varepsilon)x^{\gamma-1}$ for all $m \geq m_0$.

Second, again using (12) and (5), or differentiating (7), we arrive at
\[
L'_m(x^{\gamma-1}) \geq \prod_{k=1}^{\infty} \exp \left( -4x^2(m-1)\gamma^{-2(k+1)} \right) = \exp \left( -4x^2(m-1) \sum_{k=1}^{\infty} \gamma^{-2(k+1)} \right)
\geq \exp \left( -8x^2m\gamma^{-4} \right),
\]
Together with our previous conclusions we find that
\[
L'_m(x^{\gamma-1}) \left[ 1 - 4x^2\gamma^{-2} \right]^{m-1} \leq L'_m(x) \leq L'_m(x^{\gamma-1}) \left[ 1 - (1 - \varepsilon)^24x^2\gamma^{-2} \right]^{m-1}.
\]
Taking limits, first as $m \to \infty$ and then as $\varepsilon \to 0$, leaves us with
\[
\lim_{m \to \infty} L'_m(x) = e^{-\pi x^2},
\]
as required.

4 Percolation Crossings

We now consider percolation crossings in the hexagonal lattice. Unlike every other result in this paper, this will be based on recent highly nontrivial developments in percolation and in so called near-critical percolation; see e.g. [11] and [12].

To even begin this, we need to introduce a number of different concepts. However, we will be very brief and refer to [17] and [9] for background and explanation of terms which are not clear. We consider percolation on the hexagonal lattice embedded into $\mathbb{R}^2$. Given $n$, we will consider the set of hexagons contained inside of $[0,n] \times [0,n]$, denoted by $B_n$, and we will think of these hexagons as indexing our underlying i.i.d. random variables of which we will then have approximately $n^2$. We let $f_n$ be the indicator function of the event that there is a path of hexagons from the left side of this box to the right side all of whose values are 1. It is well known that there is a threshold at $p = 1/2$ which is the critical value for percolation on the (full) hexagonal lattice. For critical percolation on the hexagonal lattice, we let $\alpha_4(R)$ be the
probability that there are four paths of alternating value from a neighbor of the origin 0 to
distance $R$ away; this event is usually called the four-arm event. See Figure 1 for a realization
of this event (where 1 is replaced by black and 0 is replaced by white).

Using SLE and conformal invariance, it was proved by Smirnov and Werner \[14\] that
\[
\alpha_4(R) = R^{-\frac{5}{4} + o(1)} \quad \text{as } R \to \infty.
\]

A little bit of thought shows that if we have a hexagon $H$ in $B_n$, not too close to the boundary,
which is pivotal for this crossing event, then the four-arm event to distance approximately $n$
centered at $H$ occurs. From here, it is possible to argue (see \[9\]) that the expected number of
pivotal hexagons for $f_n$ is, up to constants, $n^2 \alpha_4(n)$. This suggests what the proper scaling of
$T_n$ is and this again turns out to be correct. The result for percolation on the hexagonal lattice
turns out to be the following.

**Proposition 12.** There is a distribution function $F$ which is continuous and has full support
so that for all $x \in \mathbb{R}$ we have
\[
\lim_{n \to \infty} \mathbb{P}\left(n^2 \alpha_4(n) \left( T_n - \frac{1}{2} \right) \leq x \right) = F(x).
\]

Proposition 12 follows from the results due to Garban, Pete and Schramm \[7, Theorem 1.5\]. That the
limit, should it exist, would have to satisfy the stated properties essentially follows from the work of Kesten \[11\]. However, proving that the limit indeed exists is more involved
and requires further arguments. By diving deeply into \[7\], one would be able to obtain other
properties of the limiting distribution but this would bring us too far afield.

## 5 All measures are distributional limits

In this section, we prove Theorem 2 and Propositions 3 and 4 first showing that all probability
measures appear as distributional limits.
Proof of Theorem 3. The main part of the proof is to prove the result for a restricted class of probability measures, namely those $\mu$ of the form $\sum_{i=1}^k p_i \delta_{x_i}$, i.e., having finite support. Assume such a $\mu$ is given and let $(a_n)_{n \geq 1}$ satisfy $1 \ll a_n \ll \sqrt{n}$. We may assume that $x_1 < x_2 < \ldots < x_k$ and that the $p_i$'s are all positive. Express $n = k(m + \lfloor a_n \rfloor) + \ell$ for integers $m$ and $\ell$, where $m$ is as large as possible. Partition $[n]$ into sets $A_1, A_2, \ldots, A_k$ of size $m$, $B_1, B_2, \ldots, B_k$ of size $\lfloor a_n \rfloor$, and let $C$ contain the residual $\ell$ bits. For each $i = 1, 2, \ldots, k$ fix $y_i \in \mathbb{R}$ such that

$$1 - \Phi(y_i) = \frac{p_i}{1 - p_1 - p_2 - \ldots - p_{k-1}},$$

where $\Phi(\cdot)$ denotes the distribution function of the standard Gaussian. (Of course, this defines $y_i$ to be $-\infty$, but we allow this slight abuse of notation.)

Now let $E_i$ denote the event that the proportion of 1’s among the bits in $A_i$ is at least $\frac{1}{2} + x_i/a_n$, and let $F_i$ denote the event that the proportion of 1’s among the bits in $B_i$ is at least $\frac{1}{2} + y_i/(2\sqrt{a_n})$. Although not explicit in the notation, these events depend on $n$. Finally, we define $f_n$ as the indicator function of the event $\bigcup_{i=1}^k (E_i \cap F_i)$.

Let $p_n = \frac{1}{2} + x/a_n$. To complete the first part of the proof we need to verify that

$$\mathbb{P}(a_n(T_n - \frac{1}{2}) \leq x) = \mathbb{P}_{p_n}(E_i \cap F_i \text{ for some } i)$$

tends to 0, $\sum_{i=1}^j p_i$ or 1, depending on whether $x < x_1$, $x \in (x_j, x_{j+1})$ and $j = 1, 2, \ldots, k - 1$, or $x > x_k$. We first examine the events $E_i$ and $F_i$. Appealing to the Lindeberg-Feller central limit theorem, or Chernoff’s inequality, we find that

$$\mathbb{P}_{p_n}(E_i) = \mathbb{P}_{p_n}\left( \frac{1}{m} \sum_{j \in A_i} \omega_j \geq \frac{1}{2} + \frac{x_i}{a_n} \right) \rightarrow \begin{cases} 0 & x < x_i, \\ 1 & x > x_i, \end{cases} \quad (15)$$

as by assumption $a_n \ll \sqrt{n}$. On the other hand, using Lindeberg-Feller, for any $x \in \mathbb{R}$

$$\mathbb{P}_{p_n}(F_i) = \mathbb{P}_{p_n}\left( \frac{1}{\lfloor a_n \rfloor} \sum_{j \in B_i} \omega_j \geq \frac{1}{2} + \frac{y_i}{2\sqrt{a_n}} \right) \rightarrow 1 - \Phi(y_i). \quad (16)$$

(The above abuse of notation is here manifested in that $F_k$ equals the whole sample space.)

Since $\mathbb{P}_{p_n}(E_i \cap F_i \text{ for some } i)$ is at most $\mathbb{P}_{p_n}(E_i \text{ for some } i)$, the case $x < x_1$ is immediate from (15). Similarly for $x \in (x_j, x_{j+1})$, using (15), assuming the limits exists, gives

$$\lim_{n \to \infty} \mathbb{P}_{p_n}(E_i \cap F_i \text{ for some } i) = \lim_{n \to \infty} \mathbb{P}_{p_n}(F_i \text{ for some } i \leq j),$$

and for $x > x_k$ the $j$ is replaced by $k$ in the right-hand side. By independence of the $F_i$’s and (16) we conclude that the above limit exists, and that for every $j = 1, 2, \ldots, k$,

$$\mathbb{P}_{p_n}(F_i \text{ for some } i \leq j) = 1 - \prod_{i=1}^j (1 - \mathbb{P}_{p_n}(F_i)),$$

which due to the definition of the $y_i$’s approaches $p_1 + p_2 + \ldots + p_j$ as $n \to \infty$, as required.
To now obtain the general result, we first state a simple lemma, whose proof is left to the reader, concerning metric spaces. Assume, in a metric space, we are given $x_m$ converging to $x_\infty$ and for each $m$, we have $x_{m,n}$ converging to $x_m$ as $n \to \infty$. Then there is a sequence $m_n$ (not necessarily strictly increasing) so that we have that $x_{m,n}$ converges to $x_\infty$ as $n \to \infty$.

We note that it is well known that convergence in distribution is metrizable. Assume now that we are given an arbitrary probability measure $\mu$ and a sequence $a_n$ satisfying the stated properties. It is clear we can find a sequence $(\mu_m)_{m \geq 1}$, each with finite support as above, converging to $\mu$. By the case already proved, for each $m$, we can find a sequence of monotone Boolean functions $(f_{m,n})_{n \geq 1}$ such that $f_{m,n}$ is defined on $n$ bits and $a_n(T(f_{m,n}) - \frac{1}{2})$ approaches, as $n \to \infty$, $\mu_m$ in distribution. By the above general metric space result, there exists a sequence $(m_n)_{n \geq 1}$ (not necessarily strictly increasing) so that $a_n(T(f_{m_n,n}) - \frac{1}{2})$ approaches $\mu$ in distribution, as $n \to \infty$. This proves the claim.

We continue proving that not all measures appear without scaling.

**Proof of Proposition 3.** We will show that there is no sequence of monotone Boolean functions for which $T_n$ approaches the measure giving equal weight to $1/3$ and $2/3$. Recall the discrete Poincaré inequality which states that for any Boolean function $f : \{0, 1\}^n \to \{0, 1\}$

$$\sum_{i=1}^{n} \inf_{x} p_i^x(f) \geq \text{Var}_p(f).$$

From this, the Margulis-Russo formula allows us to conclude that

$$\frac{d}{dp} \mathbb{P}_p(f = 1) \geq \text{Var}_p(f).$$

This easily implies that there is no Boolean function $f$ for which $\mathbb{P}_p(f = 1)$ is larger than .499 at .334 and less than .501 at .66. The statement in the first sentence of the proof follows.

**Remark.** While the above proof shows that no monotone Boolean function $f$ may have $\mathbb{P}_p(f = 1)$ making a finite number of “jumps” but otherwise remaining more or less constant, it is still possible to have sudden jumps between which $\mathbb{P}_p(f = 1)$ grows linearly. An example would be the event $A$ consisting of all configurations for which either $\omega_1 = 1$ and the proportion of 1’s is at least 1/3 or the proportion of 1’s is at least 2/3.

As mentioned in the introduction, one easily shows that no subsequence of the probability measures giving equal weight to the points in $\Omega_m = \{ \pm 2^k : k = 1, 2, \ldots, m \}$ can be normalized in order to give a nondegenerate limit. We base the proof of Proposition 3 on this example.

**Proof of Proposition 3.** We prefer to work with continuous distributions. Let $\mu_m$ be the measure whose density function equals $\frac{1}{2m}$ for $x \in [k - \frac{1}{2}, k + \frac{1}{2}]$ and $k \in \Omega_m$, and 0 otherwise; let $F_m$ denote the corresponding distribution function. Then $F_m$ is continuous and $\mu_m$ effectively has the same properties as the uniform measure on $\Omega_m$. According to Theorem 2, we may choose $a_n = n^{1/4}$, say, and monotone Boolean functions $f_{m,n} : \{0, 1\}^n \to \{0, 1\}$ such that $a_n(T(f_{m,n}) - \frac{1}{2})$ tends to $\mu_m$ in distribution, as $n \to \infty$. Writing $F_{m,n}$ for the distribution function of $a_n(T(f_{m,n}) - \frac{1}{2})$ and using that $F_m$ is continuous, we find for each $m$ an integer $n_m$ such that

$$\sup_{x \in \mathbb{R}} |F_{m,n}(x) - F_m(x)| \leq \frac{1}{m} \quad \text{for all } n \geq n_m.$$

(17)
Define $f_n := f_{m,n}$ for $n \in [m, n_{m+1})$, and note that $f_n$ is a monotone function on $n$ variables.

Let $m_n := \max\{m \in \mathbb{N} : n_m \leq n\}$. Now, assume there are nonnegative sequences $(b_n)_{n \geq 1}$ and $(c_n)_{n \geq 1}$, and a nondegenerate probability measure with distribution function $F$ such that, along some subsequence, $F_{m_n,n}(c_n x + b_n) \to F(x)$ for all continuity points of $F$. Then,

$$|F_{m_n}(c_n x + b_n) - F(x)| \leq |F_{m_n}(c_n x + b_n) - F_{m_n,n}(c_n x + b_n)| + |F_{m_n,n}(c_n x + b_n) - F(x)|,$$

which, for continuity points of $F$, would tend to zero along this subsequence, in virtue of (17). This would contradict the fact that no subsequence of $\mu_m$ can be normalized to obtain a non-degenerate limit, and therefore shows that no subsequence of $(T(f_n))_{n \geq 1}$ can be normalized to obtain a nondegenerate limit.

Acknowledgements

This work was begun while DA was visiting the Department for Mathematical Sciences, Gothenburg, to which he is grateful for their hospitality. DA was during this work supported by postdoctoral grants from the Brazilian CNPq and the Swedish Research Council. JES also acknowledges the support of the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The authors thank Bo Berndtsson, Christophe Garban and Benjy Weiss for some discussions.

References

[1] B. Bollobás. Random graphs, volume 73 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition, 2001.

[2] B. Bollobás and A. Thomason. Threshold functions. Combinatorica, 7(1):35–38, 1987.

[3] R. Durrett. Probability: theory and examples. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[4] P. Erdős and A. Rényi. On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl., 5:17–61, 1960.

[5] E. Friedgut. Sharp thresholds of graph properties, and the $k$-sat problem. J. Amer. Math. Soc., 12(4):1017–1054, 1999. With an appendix by Jean Bourgain.

[6] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. Proc. Amer. Math. Soc., 124(10):2993–3002, 1996.

[7] C. Garban, G. Pete, and O. Schramm. The scaling limits of near-critical and dynamical percolation. Preprint, see arXiv:1305.5526.

[8] C. Garban and J. E. Steif. Noise sensitivity of Boolean functions and percolation. Cambridge University Press, to appear.

[9] C. Garban and J. E. Steif. Noise sensitivity and percolation. In Probability and statistical physics in two and more dimensions, volume 15 of Clay Math. Proc., pages 49–154. Amer. Math. Soc., Providence, RI, 2012.
[10] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In 29th Annual Symposium on Foundations of Computer Science, pages 68–80, 1988.

[11] H. Kesten. Scaling relations for 2D-percolation. Comm. Math. Phys., 109(1):109–156, 1987.

[12] P. Nolin. Near-critical percolation in two dimensions. Electron. J. Probab., 13:no. 55, 1562–1623, 2008.

[13] L. Russo. An approximate zero-one law. Z. Wahrsch. Verw. Gebiete, 61(1):129–139, 1982.

[14] S. Smirnov and W. Werner. Critical exponents for two-dimensional percolation. Math. Res. Lett., 8(5-6):729–744, 2001.

[15] M. Talagrand. On Russo’s approximate zero-one law. Ann. Probab., 22(3):1576–1587, 1994.

[16] J. Wästlund. An elementary proof of the Wallis product formula for pi. Amer. Math. Monthly, 114(10):914–917, 2007.

[17] W. Werner. Lectures on two-dimensional critical percolation. In Statistical mechanics, volume 16 of IAS/Park City Math. Ser., pages 297–360. Amer. Math. Soc., Providence, RI, 2009.

DANIEL AHLBERG
INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA
Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brasil
DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY
www.impa.br/~ahlberg
ahlberg@impa.br

JEFFREY E. STEIF
MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY
MATHEMATICAL SCIENCES, UNIVERSITY OF GOTHENBURG
SE-41296 GOthenburg, Sweden
http://www.math.chalmers.se/~steif
steif@chalmers.se