Amplification and Percolation

Moshe Dubiner *

Dept. of Applied Math.
Tel Aviv university
Tel Aviv 69978, Israel

Uri Zwick *

Dept. of Computer Science
Tel Aviv university
Tel Aviv 69978, Israel

Abstract

Moore and Shannon had shown that relays with arbitrarily high reliability can be built from relays with arbitrarily poor reliability. Valiant used similar methods to construct monotone read-once formulae of size $O(n^{\alpha+2})$ (where $\alpha = \log_{\sqrt{5}-1} 2 \simeq 3.27$) that amplify $(\psi - \frac{1}{n}, \psi + \frac{1}{n})$ (where $\psi = \frac{\sqrt{5}-1}{2} \simeq 0.62$) to $(2^{-n}, 1-2^{-n})$ and deduced as a consequence the existence of monotone formulae of the same size that compute the majority of n bits.

Boppana had shown that any monotone read-once formula that amplifies $(p - \frac{1}{n}, p + \frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$ (where $0 < p < 1$ is constant) has size of at least $\Omega(n^{\alpha})$ and that any monotone, not necessarily read-once, contact network (and in particular any monotone formula) that amplifies $(\frac{1}{4}, \frac{3}{4})$ to $(2^{-n}, 1-2^{-n})$ has size of at least $\Omega(n^2)$.

We extend Boppana's results in two ways. We first show that his two lower bounds hold for general read-once formulae, not necessarily monotone, that may even include exclusive-or gates. We are then able to join his two lower bounds together and show that any read-once, not necessarily monotone, formula that amplifies $(p-\frac{1}{n}, p+\frac{1}{n})$ to $(2^{-n}, 1-2^{-n})$ has size of at least $\Omega(n^{\alpha+2})$. This result does not follow from Boppana's arguments and it shows that the amount of amplification achieved by Valiant is the maximal achievable using read-once formulae.

Next, extending the results of Moore and Shannon and Valiant, we construct monotone read-once networks of size $O(n^{2.99})$ that amplify $(\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$. This construction improves the reliability that can be obtained using a given number of relays, and shows that the first lower bound of Boppana, in contrast to his second, does not apply to contact networks, even if they are required to be both monotone and read-once.

Finally, we establish a connection between amplification and percolation and show that certain conjectures in percolation theory imply the existence of monotone read-once networks of size $O(n^{8/3+o(1)})$, and perhaps even less, that amplify $(\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$.

1 Introduction

In a classical paper, Moore and Shannon [13] use what is now called the *amplification method* to show that relays with arbitrarily high reliability can be built from (so-called *crummy*) relays with arbitrarily poor reliability.

The amplification method was next used by several researchers to show that particular Boolean functions have small Boolean circuits and formulae. Benett and Gill [3], extending a result of Adelman [1] used it to show that every language in the complexity class BPP has polynomial-size circuits. Ajtai and Ben-Or [2] used the amplification method to show that probabilistic constant-depth circuits can be simulated by deterministic constant-depth circuits, with only a polynomial increase in size.

The main focus of attention in this paper is the elegant application of the amplification method by Valiant [19] to the construction of monotone formulae of size $O(n^{\alpha+2})$ (where $\alpha = \log_{\sqrt{5}-1} 2 \simeq 3.27$) for the majority function and its extension by Boppana [4] to the construction of $O(k^{\alpha+1} n \log n)$ size monotone formulae for the k-th threshold function of n variables.

To show the existence of $O(n^{\alpha+2})$ size monotone formulae for the majority function Valiant first con-

*Full affiliations : Departments of Applied Mathematics and Computer Science, School of Mathematical Sciences, The Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv university, Tel Aviv 69978, Israel. The second author was supported in part by THE BASIC RESEARCH FOUNDATION administrated by THE ISRAEL ACADEMY OF SCIENCES AND HUMANITIES. Contact author: Uri Zwick, e-mail: zwick@math.tau.ac.il.

structs monotone read-once formulae that amplify $(\psi - \frac{1}{n}, \psi + \frac{1}{n})$ (where $\psi = \frac{\sqrt{5}-1}{2}$) to $(2^{-n}, 1-2^{-n})$. Formal definitions of all these terms will appear in the next section. The existence of monotone formulae of the same size for majority follows using a simple probabilistic argument.

Boppana [4], considered the question of whether Valiant had obtained an optimal amount of amplification in his construction and came very close to answering it positively. He observed that Valiant had actually constructed $O(n^{\alpha})$ size monotone read-once formulae that amplify $(\psi - \frac{1}{n}, \psi + \frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$ and $O(n^2)$ size monotone read-once formulae that amplify $(\frac{1}{4}, \frac{3}{4})$ to $(2^{-n}, 1-2^{-n})$. The $O(n^{\alpha+2})$ size monotone read-once formulae that amplify $(\psi - \frac{1}{n}, \psi + \frac{1}{n})$ to $(2^{-n}, 1-2^{-n})$ are easily obtained by combining these two sub-constructions. Boppana was able to show that each one of these sub-constructions achieved an optimal amount of amplification. It does not seem to follow however from Boppana's arguments that the combined construction is also optimal.

We are able to join together the two lower bounds of Boppana and show that any monotone read-once formula that amplifies $(p-\frac{1}{n}, p+\frac{1}{n})$ to $(2^{-n}, 1-2^{-n})$ (where $0 < p < 1$ is fixed) does indeed have to be of size $\Omega(n^{\alpha+2})$. This gives a complete positive answer to the question of whether Valiant's construction obtains an optimal amount of amplification.

We are also able to strengthen Boppana's results in another respect. We show that the combined $\Omega(n^{\alpha+2})$ lower bound applies even if the read-once formulae are allowed to use negations and exclusive-or (XOR) gates (as well as the monotone AND and OR gates). To that end we show that for every formula with XOR gates there exists a probabilistic formula without XOR gates of the same size that achieves the same amount of amplification.

Boppana had used two different (we are tempted to say incompatible) methods to get his two lower bounds for the two amplification stages. To obtain the combined lower bound we have to slightly strengthen his first lower bound and exhibit an alternative proof for his second. In particular, we obtain a slightly stronger version of, what we call, Boppana's inequality (see next section). We present an analytical proof of this strengthened inequality (Boppana had resorted to numerical experimentation in the proof of his inequality). We also generalize Boppana's bound on the derivatives of univariate amplification functions to bounds on the partial derivatives of multivariate amplification functions.

As mentioned, Boppana's second lower bound applies to general monotone networks. Using a result of Lupanov [12], it can be shown that it does not apply to non-monotone contact networks. Thus, in the contact networks model negations do help amplify. We do not elaborate on it in this abstract. We also note that negations and XOR gates seem to help in the construction of formulae for the majority function. In [14],[15], non-monotone formulae of size $O(n^{4.57})$, without XOR gates, and of size $O(n^{3.13})$, with XOR gates were constructed.

Next, we construct monotone read-once undirected contact networks of size $O(n^{2.99})$ that amplify $(\frac{1}{2} - \frac{1}{n}, \frac{1}{2}+\frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$. This improves the results of Moore and Shannon [13] and those of Valiant [19] and shows that Boppana's first lower bound does not apply to contact networks, even if they are required to be undirected, monotone and read-once. It also implies the existence of *undirected* monotone contact networks of size $O(n^{4.99})$ that compute the majority of n bits. Smaller *directed* monotone networks for majority had been recently constructed by Radhakrishnan and Subrahmanyam [17].

Finally, we exhibit a connection between amplification and percolation. We show that certain conjectures in percolation theory imply the existence of monotone read-once (undirected) networks of size $O(n^{8/3+o(1)})$, and perhaps even less, that amplify $(\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$ and consequently the existence of monotone (undirected) networks of size $O(n^{14/3+o(1)})$ that compute the majority function of n bits.

Other works relevant to the subjects considered in this paper are [7],[8],[9] and [16].

2 Preliminaries

Following Moore and Shannon [13] and Boppana [4] we introduce the following two definitions

Definition 2.1 (Amplification Functions) Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, we define its *multivariate amplification function* $f : [0,1]^n \to [0,1]$ as follows: $f(p_1, p_2, \ldots, p_n) = \Pr[f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = 1]$, where $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are independent random variables, $\mathbf{x}_i$ assumes the value 1 with probability p_i and the value 0 with probability $1 - p_i$. We define the *univariate amplification function* of f to be $f(p) = f(p, p, \ldots, p)$. As the (multivariate) amplification function is an extension of the original Boolean function, we use the same notation for both.

Definition 2.2 (Amplification from/to) A Boolean function function $f : \{0,1\}^n \to \{0,1\}$ *amplifies* (p,q) to (p',q') if $f(p) = p'$ and $f(q) = q'$. It amplifies (p,q) to *at least* (p',q') if $f(p) \leq p'$ and $f(q) \geq q'$.

The main computational models studied in this work are formulae and contact networks.

Definition 2.3 (Formulae) A *formula* of n variables is defined recursively as follows: (i) for $1 \leq i \leq n$, the variables x_i and their negations $\overline{x_i}$ are formulae; (ii) if f and g are formulae then so are $(\neg f)$, $(f \wedge g)$, $(f \vee g)$ and $(f \oplus g)$. A formula in which no XOR gates are used is called *unate* or *de Morgan*. A unate formula in which no negations are used is called *monotone*. Formulae define Boolean functions in the obvious way. The *size* of a formula f, denoted by $size(f)$, is the number of occurrences of variables in it.

Definition 2.4 (Probabilistic Formulae) A *probabilistic* (monotone, unate) formula is simply a discrete probability distribution over conventional deterministic (monotone, unate) formulae. The size of a probabilistic formula F is defined to be the maximum size of a formula f whose probability according to the distribution induced by F is positive. The amplification function $F(p_1,\ldots,p_n)$ of a probabilistic formula F is simply $F(p_1,\ldots,p_n) = E_{f \in F} f(p_1,\ldots,p_n)$ where E denotes expectation relative to the distribution induced by F.

Definition 2.5 (Contact Networks) A *contact network* of n variables is an undirected graph with two distinguished vertices s (the source) and t (the sink) in which each edge is labeled by a variable x_i (where $1 \leq i \leq n$) or its negation. The network defines a function in the following way; for any given assignment of Boolean values to the variables $x_1,\ldots,x_n$ we check whether there is a path from s to t in which all the edges are labeled by 1. If there is such a path then the value of the function for this assignment is defined to be 1, otherwise 0. The *size* of a network is defined to be the number of edges in it. A network is said to be *monotone* if no edges in it are labeled by negations of variables.

It is easy to see that monotone and unate formulae correspond to monotone and non-monotone *series-parallel* contact networks. For more details the reader is referred to [4].

Most of the formulae and networks encountered in this work will be read-once.

Definition 2.6 (Read-Once) A formula or a network are said to be *read-once* if every variable appears in them at most once.

Definition 2.7 (Amplification Complexity) If $0 < p_0, q_0, p_1, q_1 < 1$, we denote by $N(p_1, q_1|p_0, q_0)$ the minimum size of a unate read-once formula that amplifies (p_0, q_0) to (p_1, q_1). We denote by $N_\oplus(p_1, q_1|p_0, q_0)$ the corresponding measure for general read-once formulae.

In the sequel we refer to the following inequalities

Theorem 2.8 (Hölder's inequality) *If $\alpha, \beta > 1$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ then for any two real vectors x and y we have*

$$\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_\alpha \cdot \|y\|_\beta$$

where $\|x\|_\alpha = \left(\sum_{i=1}^{n} |x_i|^\alpha\right)^{1/\alpha}$.

Theorem 2.9 (Boppana's inequality) *Let $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ be the usual binary entropy function and let $\beta = \log_{\frac{\sqrt{5}-1}{2}} 2 \simeq 1.44$. Then, for every $0 < x, y \leq 1$ we have*

$$\left(\frac{H(x)}{x}\right)^\beta + \left(\frac{H(y)}{y}\right)^\beta \leq \left(\frac{H(xy)}{xy}\right)^\beta \quad .$$

In the sequel we would say that a function I satisfies Boppana's inequality iff $I(x) = I(1-x) \geq 0$ for every $0 \leq x \leq 1$ and if I satisfies the inequality in the above lemma.

3 Strengthening Boppana's inequality

Boppana's inequality, stated in the previous section forms the basis of his lower bound for the first amplification stage. To get some of our extended lower bounds we need a slightly stronger version of his inequality. This version is obtained by replacing the entropy function $H(x)$ by a function $G(x)$ with the tightest asymptotic behaviour near $x = 0$ and $x = 1$.

To prove his inequality Boppana had to make many numerical checks involving functions of two variables (in one point of the argument, for example, he has to estimate the third derivatives of the function $\widetilde{H}(x,y) = (H(xy)/xy)^\beta - (H(x)/x)^\beta - (H(y)/y)^\beta$ and show that the Hessian matrix of $\widetilde{H}$ is positive definite on the entire region $[0.55, 0.65] \times [0.55, 0.65]$). We are able to exhibit a much simpler proof to our strengthened inequality. We show that it follows from some relatively simple one variable inequalities.

Our new function $G(x)$ is defined as follows

$$G(x) = \begin{cases} x\left(\ln \frac{\gamma}{x}\right)^{1/\beta} & \text{if } 0 \le x \le \frac{1}{2} \\ (1-x)\left(\ln \frac{\gamma}{1-x}\right)^{1/\beta} & \text{if } \frac{1}{2} \le x \le 1 \end{cases}$$

where

$$\ln \gamma = \frac{2 - \psi - \ln 4}{\beta} \simeq -0.003005 \quad .$$

The constant γ was chosen so that $\widehat{G}(\psi) = \widehat{G}(1-\psi)$ where $\widehat{G}(x) = x \cdot \frac{d}{dx}\left(\frac{G(x)}{x}\right)^\beta$. This is a necessary condition that must be satisfied by any symmetric function wishing to satisfy Boppana's inequality. The function $G(x)$ has a small cusp at $x = \frac{1}{2}$ but this will not cause us any trouble. We now claim

Lemma 3.1 *For every $0 < x, y \le 1$ we have*

$$\left(\frac{G(x)}{x}\right)^\beta + \left(\frac{G(y)}{y}\right)^\beta \le \left(\frac{G(xy)}{xy}\right)^\beta \quad .$$

Proof : See the full version of the paper. □

4 Bounds on derivatives of amplification functions

The basic theorem from which all the results of this section follow is

Theorem 4.1 *If f is a unate read-once formula (or a read-once series-parallel contact network) that depends on n variables and if I satisfies Boppana's inequality then*

$$\left(\sum_{i=1}^{n} \left| I(x_i) \cdot \frac{\partial f}{\partial x_i} \right|^\beta \right)^{1/\beta} \le I(f) \quad .$$

Proof : The proof is by induction on the structure of f. If $f = x_i$ or $f = \overline{x_i}$ then the inequality is easily verified as $\left|\frac{\partial f}{\partial x_i}\right| = 1$ and $I(x_i) = I(1 - x_i) = I(f)$. If $f = \neg g$ then the inequality follows easily from the induction hypothesis as $f = 1 - g$ and therefore $\left|\frac{\partial f}{\partial x_i}\right| = \left|\frac{\partial g}{\partial x_i}\right|$ and $I(f) = I(1 - f) = I(g)$. If $f = f_1 \wedge f_2$ and if $f_1(x_1, \ldots, x_n)$ and $f_2(y_1, \ldots, y_m)$ are the amplification functions of f_1 and f_2 then the amplification function of f is $f(x_1, \ldots, x_n, y_1, \ldots, y_m) = f_1(x_1, \ldots, x_n) f_2(y_1, \ldots, y_m)$. As a consequence $\frac{\partial f}{\partial x_i} = \frac{\partial f_1}{\partial x_i} \cdot f_2$ and $\frac{\partial f}{\partial y_j} = f_1 \cdot \frac{\partial f_2}{\partial y_j}$. For convenience we let $x_{n+1} = y_1, \ldots, x_{n+m} = y_m$. Using the induction hypothesis and then Boppana's inequality we therefore get

$$\begin{aligned} &\sum_{i=1}^{n+m} \left| I(x_i) \cdot \frac{\partial f}{\partial x_i} \right|^\beta \\ \le\ & f_2^\beta \cdot \sum_{i=1}^{n} \left| I(x_i) \cdot \frac{\partial f}{\partial x_i} \right|^\beta + f_1^\beta \cdot \sum_{i=n+1}^{n+m} \left| I(x_i) \cdot \frac{\partial f}{\partial x_i} \right|^\beta \\ \le\ & (f_2 \cdot I(f_1))^\beta + (f_1 \cdot I(f_2))^\beta \\ =\ & (f_1 f_2)^\beta \cdot \left[\left(\frac{I(f_1)}{f_1}\right)^\beta + \left(\frac{I(f_2)}{f_2}\right)^\beta \right] \\ \le\ & (f_1 f_2)^\beta \cdot \left(\frac{I(f_1 f_2)}{f_1 f_2}\right)^\beta = (I(f))^\beta \quad . \end{aligned}$$

This completes the proof as any unate gate can be obtained by combining an $\wedge$-gate with negations. □

If, in addition to satisfying Boppana's inequality, the function I also satisfies the following inequality

$$\left(\frac{I\left(\frac{1-x}{2}\right)}{x}\right)^\beta + \left(\frac{I\left(\frac{1-y}{2}\right)}{y}\right)^\beta \le \left(\frac{I\left(\frac{1-xy}{2}\right)}{xy}\right)^\beta$$

for every $0 < x, y \le 1$ then we can show directly that the inequality in the above lemma holds for general, not necessarily unate, read-once formulae. Both the entropy function $H(x)$ and the function $G(x)$ introduced in the previous section satisfy this inequality. We omit the details here as the same amplification bounds will follow from the more general arguments of Section 8.

We denote by $\nabla f(x)$ the *gradient* $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$ of f evaluated at the point $(x, \ldots, x)$. As an immediate corollary to the previous lemma we get

Corollary 4.2 *If f is a unate read-once formula then*

$$\|\nabla f(x)\|_\beta \le \frac{I(f(x))}{I(x)} \quad .$$

In particular if p_c is the *critical probability* for which $f(p_c) = p_c$, we get that $\|\nabla f(p_c)\|_\beta \le 1$. We will encounter a similar relation in Section 9 that presents constructions of small amplifying networks.

As a second corollary to Lemma 4.1 we get the following result obtained by Boppana for monotone read-once formulae.

Corollary 4.3 *If f is a unate read-once formula that depends on n variables then*

$$|f'(x)| \le n^{1/\alpha} \cdot \frac{I(f(x))}{I(x)} \quad .$$

Proof : Applying Hölder's inequality we get

$$\begin{aligned} |f'(x)| &\le \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(x) \right| \\ &\le \left(\sum_{i=1}^{n} 1 \right)^{1/\alpha} \cdot \left(\sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(x) \right|^{\beta} \right)^{1/\beta} \\ &= n^{1/\alpha} \cdot \|\nabla f(x)\|_\beta \ \le \ n^{1/\alpha} \cdot \frac{I(f(x))}{I(x)} \quad . \end{aligned}$$

□

Again, if I satisfies the additional inequality mentioned above, as G and H do, then the above corollary holds for general, not necessarily unate, read-once formulae. We claim that the inequality in Corollary 4.3 with $G(x)$ plugged into it is, up to constant factors, the strongest valid inequality of its kind. We do not elaborate on it here.

5 Local amplification bounds

Boppana [4] used the above corollary, together with the mean value theorem to get a lower bound on the size of formulae required to amplify (p_0, q_0) to (p_1, q_1). A stronger lower bound is obtained by integrating the upper bound obtained for $|f'(x)|$. Let

$$M(x,y) = \left| \int_x^y \frac{du}{G(u)} \right|^{\alpha} \quad .$$

Theorem 5.1 *Any read-once formula that amplifies (p_0, q_0) to (p_1, q_1) has size of at least*

$$N(p_1, q_1 | p_0, q_0) \ge \frac{M(p_1, q_1)}{M(p_0, q_0)} \quad .$$

Proof : Let f be a read-once formula that depends on n variables and amplifies (p_0, q_0) to (p_1, q_1). Using Corollary 4.3 and a simple change of variables we get that

$$\begin{aligned} \left| \int_{p_1}^{q_1} \frac{dx}{G(x)} \right| &\le \int_{p_0}^{q_0} \left| \frac{f'(x)dx}{G(f(x))} \right| \\ \le \left| \int_{p_0}^{q_0} \frac{n^{1/\alpha}dx}{G(x)} \right| &= n^{1/\alpha} \left| \int_{p_0}^{q_0} \frac{dx}{G(x)} \right| \quad . \ \Box \end{aligned}$$

Since

$$\int \frac{dx}{x\left(\ln \frac{\gamma}{x}\right)^{1/\beta}} = -\alpha \left(\ln \frac{\gamma}{x} \right)^{1/\alpha} \quad ,$$

it is easily verified that

$$M(x,y) = \left| \int_x^y \frac{du}{G(u)} \right|^{\alpha} =$$

$$C \cdot \begin{cases} m(x,y) & \text{if } x \le y \le \frac{1}{2} \\ (m(x,\frac{1}{2})^{1/\alpha} + m(1-y,\frac{1}{2})^{1/\alpha})^{\alpha} & \text{if } x \le \frac{1}{2} \le y \\ m(1-y, 1-x) & \text{if } \frac{1}{2} \le x \le y \\ M(y,x) & \text{if } y < x \end{cases}$$

where $C = \alpha^\alpha$ and

$$m(x,y) = \left[\left(\ln \frac{\gamma}{x} \right)^{1/\alpha} - \left(\ln \frac{\gamma}{y} \right)^{1/\alpha} \right]^{\alpha} \quad .$$

Clearly $M(x,y) = M(y,x) = M(1-x, 1-y)$. In section 7, the function $M(x,y)$ will be compared with some other functions. It will be convenient to scale $M(x,y)$ before these comparisons and assume that $C = 10^6$. This of course will not affect the validity of Theorem 5.1.

As an immediate corollary of Theorem 5.1 we get

Corollary 5.2 *Any unate read-once formula that amplifies $(p - \frac{1}{n}, p + \frac{1}{n})$ to at least $(\frac{1}{4}, \frac{3}{4})$ (where $0 < p < 1$ is fixed) is of size $\Omega(n^\alpha)$ and any unate read-once formula that amplifies $(p - \frac{1}{n}, p + \frac{1}{n})$ to at least $(2^{-n}, \frac{1}{2})$ is of size $\Omega(n^{\alpha+1})$.*

Proof : Follows easily as $M((p - \frac{1}{n}, p + \frac{1}{n}) = \theta(n^{-\alpha})$ and $M(2^{-n}, \frac{1}{2}), M(\frac{1}{2}, 1 - 2^{-n}) = \theta(n)$. □

It can be checked (see [4]) that both these bounds are optimal. As $M(2^{-n}, 1-2^{-n})$ is still $\theta(n)$, the methods used so far imply only an $\Omega(n^{\alpha+1})$ lower bound on amplification from $(p - \frac{1}{n}, p + \frac{1}{n})$ to $(2^{-n}, 1 - 2^{-n})$. This will be improved in section 7 to a tight $\Omega(n^{\alpha+2})$ lower bound.

We note that the use of the function $G(x)$ in place of $H(x)$ is essential in obtaining the $\Omega(n^{\alpha+1})$ lower bounds of Corollary 5.2. Had we used $H(x)$, we would only have obtained $\Omega(n^\alpha \log n)$ lower bounds.

6 Global amplification bounds

The main result of this section is the following simple, yet powerful theorem

Theorem 6.1 *If $L(x,y)$ is defined on $(0,1)^2$ and satisfies the following two conditions*

$$\begin{aligned} L(x_1x_2, y_1y_2) &\leq L(x_1,y_1) + L(x_2,y_2) \\ L(x,y) &= L(1-x,1-y) \end{aligned}$$

for every $0 < x_1,x_2,y_1,y_2,x,y < 1$ then for every $0 < p_0,p_1,q_0,q_1 < 1$ we have

$$N(p_1,q_1|p_0,q_0) \geq \frac{L(p_1,q_1)}{L(p_0,q_0)} \quad .$$

Proof : We prove by induction on the structure of f that if f amplifies (p_0,q_0) to (p_1,q_1) then $size(f) \geq L(p_1,q_1)/L(p_0,q_0)$. The basis of the induction is easily established.

If $f = \neg g$ then g amplifies (p_0,q_0) to $(1-p_1,1-q_1)$. As $size(f) = size(g)$ and $L(1-p_1,1-q_1) = L(p_1,q_1)$, the required inequality follows immediately from the induction hypothesis.

If $f = f_1 \wedge f_2$ and if f_1, f_2 amplify (p_0,q_0) to (x_1,y_1) and (x_2,y_2) respectively, then $p_1 = x_1x_2$ and $q_1 = y_1y_2$. By the induction hypothesis and the first condition of L we have

$$size(f) = size(f_1) + size(f_2) \geq$$
$$\frac{L(x_1,y_1)+L(x_2,y_2)}{L(x_0,y_0)} \geq \frac{L(x_1x_2,y_1y_2)}{L(p_0,q_0)} = \frac{L(p_1,q_1)}{L(p_0,q_0)} \quad .$$

□

Note that in both Theorem 5.1 and Theorem 6.1 the lower bounds obtained were of a similar form. Both involved a quotient of some two-variable *potential* function evaluated at the pre and post amplification probabilities.

The approach of Theorem 6.1 seems to be more general than the approach of the previous section. Numerical tests seem to suggest that the function $M(x,y)$ of the previous section satisfies the condition $M(x_1x_2,y_1y_2) \leq M(x_1,y_1) + M(x_2,y_2)$ of Theorem 6.1. If this is indeed the case, then Theorem 5.1 follows immediately from Theorem 6.1.

With some additional work, the numerical tests that we had performed can probably be turned into a proof, similar in spirit to the original proof that Boppana had given to his inequality, that the function $M(x,y)$ satisfies the inequality $M(x_1x_2,y_1y_2) \leq M(x_1,y_1) + M(x_2,y_2)$. This should be tedious however, as four, and not just two, variables are involved this time. We did not carry out this extra work as the direct proof given to Theorem 5.1 seems to be more informative.

We now use Theorem 6.1 to obtain an alternative proof to the lower bounds for the second amplification stage obtained by Moore and Shannon [13] and Boppana [4]. The new proof works only for read-once formulae and not for general read-once networks. It does work however for general, not necessarily monotone, read-once formulae.

Let

$$L_1(x,y) = \ln\frac{1}{x}\cdot\ln\frac{1}{1-y} \quad , \quad L_2(x,y) = \ln\frac{1}{x}\cdot\ln\frac{\ln\frac{1}{x}}{\ln\frac{1}{y}}$$

for $x \leq y \leq 1-x$ and extend their definition using the symmetry condition $L(x,y) = L(y,x) = L(1-x,1-y)$.

Lemma 6.2 *The functions $L_1(x,y)$ and $L_2(x,y)$ satisfy the conditions of Theorem 6.1.*

Proof : The proof for $L_1(x,y)$ is a simple exercise. For proof for $L_2(x,y)$ is given in the full version of the paper. □

Using $L_2(x,y)$ we extend Boppana's second lower bound to unate formulae.

Corollary 6.3 *Every unate read-once formula that amplifies $(\frac{1}{4},\frac{3}{4})$ to at least $(2^{-n_1}, 1-2^{-n_2})$ is of size $\Omega(n_1n_2 + n_1\log n_1 + n_2\log n_2)$.*

7 A unified lower bound for amplification

Corollary 5.2 gives in particular an $\Omega(n^{\alpha+1})$ lower bound on the size of unate read-once formulae amplifying $(p-\frac{1}{n}, p+\frac{1}{n})$ to at least $(2^{-n}, 1-2^{-n})$ (for a fixed $0<p<1$). In this section we combine the proof techniques of the two preceding sections and improve this to a tight $\Omega(n^{\alpha+2})$ lower bound. We begin with the following generalization of Theorem 6.1

Theorem 7.1 *If $K(x,y)$ is defined on $(0,1)^2$ and satisfies the following three conditions*

$$\begin{aligned} K(x_1x_2,y_1y_2) &\leq \max\left\{ \begin{matrix} K(x_1,y_1)+K(x_2,y_2) \\ M(x_1x_2,y_1y_2) \end{matrix} \right\} \\ K(x,y) &\geq M(x,y) \\ K(x,y) &= K(1-x,1-y) \end{aligned}$$

and if it is already known that

$$N(p_1,q_1|p_0,q_0) \geq \frac{M(p_1,q_1)}{M(p_0,q_0)} \quad ,$$

for every $0<p_0,p_1,q_0,q_1<1$, then we also have

$$N(p_1,q_1|p_0,q_0) \geq \frac{K(p_1,q_1)}{K(p_0,q_0)} \quad .$$

Proof : Similar to the proof of Theorem 6.1. □

To get our unified lower bound we apply Theorem 7.1 to a function $K(x,y)$ obtained by stitching together the definitions of $M(x,y)$ from section 5 (with $C = 10^6$), that did well for the first amplification stage, and $L(x,y) = L_2(x,y)$ of section 6, that did well for the second amplification stage. The function $K(x,y)$ will equal $M(x,y)$ when x and y are very close to one another and it will equal $L(x,y)$ when they are far apart.

The function $K(x,y)$ is defined as follows

$$K(x,y) = \begin{cases} M(x,y) & \text{if } (x,y) \in \mathcal{A} \\ \max\{\ M(x,y), L(x,y)\ \} & \text{if } (x,y) \notin \mathcal{A} \end{cases}$$

where

$$\mathcal{A} = \mathcal{A}' \cup \mathcal{A}'' \quad , \quad \mathcal{A}' = \mathcal{R}'_2 \quad , \quad \mathcal{A}'' = \mathcal{R}''_2$$

and

$$\mathcal{R}'_d = \{\ (x,y) : \frac{1}{d} \le \left| \frac{\ln \frac{1}{x}}{\ln \frac{1}{y}} \right| \le d\ \} \quad ,$$

$$\mathcal{R}''_d = \{\ (x,y) : \frac{1}{d} \le \left| \frac{\ln \frac{1}{1-x}}{\ln \frac{1}{1-y}} \right| \le d\ \} \quad .$$

It is immediate that $K(x,y)$ satisfies the last two conditions of Theorem 7.1. To show that it also satisfies the first we define the following additional regions

$$\mathcal{C} = \mathcal{C}' \cup \mathcal{C}'' \quad , \quad \mathcal{C}' = \mathcal{R}'_6 \quad , \quad \mathcal{C}'' = \mathcal{R}''_6$$

Clearly $\mathcal{A} \subseteq \mathcal{C}$. Note that if $(x_1x_2, y_1y_2) \in \mathcal{A}$ then $K(x_1x_2, y_1y_2) = M(x_1x_2, y_1y_2)$. The fact that the function $K(x,y)$ satisfies the first condition of Theorem 7.1 will therefore follow from the following lemma

Lemma 7.2

$$L(x_1x_2, y_1y_2) \le$$

$$\begin{cases} K(x_1,y_1) + K(x_2,y_2) & \text{if } (x_1x_2, y_1y_2) \notin \mathcal{C} \\ M(x_1x_2, y_1y_2) & \text{if } (x_1x_2, y_1y_2) \in \mathcal{C} \setminus \mathcal{A} \end{cases} \quad .$$

The proof of this lemma is very technical. It will be presented at the full version of the paper. As an immediate corollary of Theorem 7.1 and Lemma 7.2 we get

Corollary 7.3 *Any unate read-once formula that amplifies $(p - \frac{1}{n_1}, p + \frac{1}{n_1})$ to at least $(2^{-n_2}, 1 - 2^{-n_3})$ (where $0 < p < 1$ is fixed) is of size $\Omega(n_1^\alpha(n_2n_3 + n_2 \log n_2 + n_3 \log n_3)$.*

8 Exclusive-Or gates as convex combinations of unate gates

Theorem 8.1 *If f is a read-once formula that amplifies (p,q) to (p',q'), then there exists a probabilistic unate read-once formula F with $size(F) \le size(f)$ that also amplifies (p,q) to (p',q').*

Proof : The proof is by induction on the structure of f. If f is a variable then the result is clear. If $f = \neg f_1$ or $f = f_1 \wedge f_2$ then the result follows immediately from the induction hypothesis. The interesting case is if $f = f_1 \oplus f_2$. Assume that f_1, f_2 amplify (p,q) to (x_1,y_1) and (x_2,y_2) respectively. Let F_1, F_2 be two probabilistic unate formulae with $size(F_1) \le size(f_1)$ and $size(F_2) \le size(f_2)$ that also amplify (p,q) to (x_1,y_1) and (x_2,y_2) respectively. The existence of F_1, F_2 follows from the induction hypothesis. Using the next lemma, we obtain a unate connective $\circ$ and three constant $0 \le \alpha, \beta, \gamma$ with $\alpha + \beta + \gamma = 1$, depending on F_1 and F_2, such that the probabilistic formula

$$F = \begin{cases} 0 & \text{with prob. } \alpha \\ F_1 \circ F_2 & \text{with prob. } \beta \\ 1 & \text{with prob. } \gamma \end{cases}$$

achieves the same amplification as f. Clearly $size(F) \le size(F_1) + size(F_2) \le size(f_1) + size(f_2) = size(f)$. This completes the proof of the Theorem. □

A unate connective here is one of the twelve connectives $(x^a \wedge y^b)^c$, x^a and y^b where $x^a = x \oplus a$, that may be obtained using one AND gate and some negations.

Lemma 8.2 *For every $0 \le x_1, x_2, y_1, y_2 \le 1$ there exists a unate connective $\circ$ such that $(x_1,y_1) \oplus (x_2,y_2)$ is in the convex hull of the points $(0,0), (x_1,y_1) \circ (x_2,y_2), (1,1)$.*

Proof : The proof uses case analysis and it will be given in the full version of the paper. □

It is easy to check that the lower bounds of the previous sections apply to probabilistic and not only to deterministic read-once formulae. All the lower bounds claimed for unate read-once formulae are valid therefore for general read-once formulae. We also have the following connection between probabilistic and non-probabilistic formulae

Theorem 8.3 *If F is a probabilistic formula amplifying (p,q) to (p',q') then there exists a deterministic formula f that amplifies (p,q) to (p'',q'') where $p'' \le 2p'$ and $1 - q'' \le 2(1-q')$. Furthermore, if F is read-once, unate or monotone then so is f.*

Proof : Recall that $p' = F(p) = E_{f\in F} f(p)$ where the expectation is according to the distribution that F induces on the deterministic formulae that it assumes. Since for every non-negative random variable X we have $\Pr[X > 2E(X)] < \frac{1}{2}$ we get that $\Pr[f(p) > 2p'] < \frac{1}{2}$, where again f is chosen according to the distribution induced by F. Applying a similar argument to $q' = F(q)$ we get that $\Pr[(1-f(q)) > 2(1-q')] < \frac{1}{2}$. As a consequence

$$Pr[\, f(p) > 2p' \;\vee\; (1-f(q)) > 2(1-q')\,] < 1$$

and therefore a deterministic formula chosen according to the distribution of F will satisfy the required conditions with a positive probability. □

Corollary 8.4

$$N_{\oplus}(p_1, q_1|p_0, q_0) \;\geq \min_{\substack{p_1' \leq 2p_1 \\ 1-q_1' \leq 2(1-q_1)}} N(p_1', q_1'|p_0, q_0) \quad .$$

9 Amplifying networks

Valiant's construction of $O(n^\alpha)$ size monotone formulae that amplify $(\psi-\frac{1}{n}, \psi+\frac{1}{n})$ to $(\frac{1}{4}, \frac{3}{4})$ uses a repeated composition of the basic series-parallel unit A shown in Fig. 1. Better amplification can be obtained using other units (e.g., units B, C, D of Fig. 1).

Theorem 9.1 *Let $f(x_1, \ldots, x_k)$ be the amplification function of a monotone read-once contact network with k edges. Let p_c be the critical probability for which $f(p_c) = p_c$. Let $\delta > 1$ be the unique number for which $\|\nabla f(p_c)\|_\delta = 1$. Then, by a repeated composition of f with composition ratios proportional to $(\nabla f(p_c))^\delta$ we get monotone read-once networks of size $O(n^\epsilon)$, where $\epsilon = \frac{\delta}{\delta-1}$ is the conjugate of δ, that amplify $(p_c - \frac{1}{n}, p_c + \frac{1}{n})$ to at least $(\frac{1}{4}, \frac{3}{4})$.*

Proof : (Sketch) Let $c_1, \ldots, c_k$ be positive constants that satisfy $\sum_{i=1}^k c_i = 1$. We define a sequence f_n of networks in the following way: $f_k = f$ and for $n > k$, f_n is obtained from f by replacing its i-th edge by a copy of the network $f_{c_i n}$. We ignore integrallity problems here. Clearly, we have the relation $f_n(p) = f(f_{c_1 n}(p), \ldots, f_{c_k n}(p))$ and therefore $f_n'(p_c) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p_c) \cdot f_{c_i n}'(p_c)$. It is easy to check that asymptotically $f_n'(p_c) \sim a n^\gamma$ where $\gamma = \gamma(c)$ is the unique real satisfying $\sum_{i=1}^k \frac{\partial f}{\partial x_i}(p_c) \cdot c_i^\gamma = 1$ and $a > 0$ is some constant. It can be checked then that $f_{bn^{1/\gamma}}$ for some fixed $b > 0$, which is a network of size $O(n^{1/\gamma})$, performs the required amplification.

We want to find the composition ratios $c_1, \ldots, c_k$ for which $\gamma = \gamma(c)$ is maximized. By Hölder's inequality we get that if $\delta' = \frac{1}{1-\gamma}$ is the conjugate of $1/\gamma$ then $1 = \sum_{i=1}^k \frac{\partial f}{\partial x_i}(p) \cdot c_i^\gamma \leq \|\nabla f(p)\|_{\delta'}$. Thus, $\delta' \leq \delta$ and $1/\gamma = \frac{\delta'}{\delta'-1} \geq \frac{\delta}{\delta-1}$. On the other hand, if we choose $c_i \propto \left(\frac{\partial f}{\partial x_i}(p)\right)^\delta$ and $\gamma = 1 - \frac{1}{\delta}$ we get the required equality. Thus, the choice specified is indeed the optimal one. □

Plugging unit C of Fig. 1 into the above theorem we get networks for the first amplification stage of size $O(n^{3.12})$. The resulting exponents for the other basic units are specified in the figure. As we shall see in the next section further improvements are possible.

10 Amplification and percolation

Blocks B and C are simply 2×3 and 3×4 rectangular grids in which all the vertices on the leftmost and rightmost columns were glued together to form the source and the sink, respectively. Using a computer program we were able to compute (exactly) the partial derivatives the the amplification functions of such $m \times (m+1)$ rectangular grids, for $m = 4, \ldots, 10$, at their critical point (this took about 30 minutes on a Sun workstation) . As all these networks are self-dual, their critical probability is exactly $p_c = \frac{1}{2}$. Using a 10×11 grid as a basic building block we get amplifying networks of size $O(n^{2.99})$ for the first stage of amplification stage. Using larger and larger building blocks we expect to get better and better results. The exact computation of the derivatives of much larger grids seems impractical. Using Monte Carlo methods we estimate that by using a 50×51 grid as the basic building block we can get networks of size below $O(n^{2.9})$.

It follows from certain conjectures in *percolation theory* that constructions of the form described, using larger and larger two dimensional grids as basic building blocks will yield amplifying networks of size $O(n^{8/3+o(1)})$. There is no need in fact to use the composition process, $n^{4/3+o(1)} \times n^{4/3+o(1)}$ grids, as conjectured, will achieve the required amplification!

Percolation theory is the study of the connectivity properties of random subgraphs of d-dimensional grids and other regular graphs. Good references to the mathematical theory of percolation are Grimmett [6], and Kesten [10],[11]. We will briefly describe its fea-

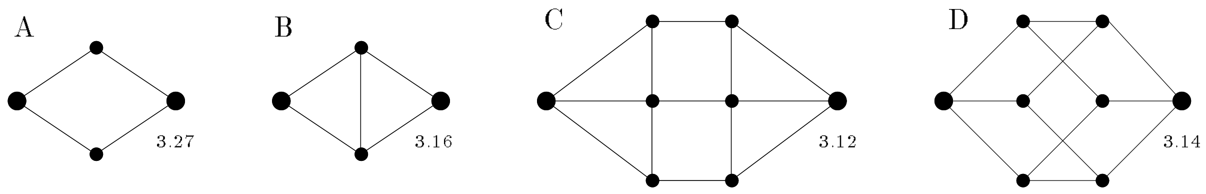

Figure 1: Basic contact networks

tures relevant to our investigation.

Consider an $m \times (m+1)$ 2-dimensional grid where each edge is independently open (i.e., present) with probability p (see Figure 2 for an example). Let $f_m(p)$ be its (univariate) amplification function and $p_c = \frac{1}{2}$ be its critical probability. Its *correlation length* $\xi(p)$ is defined for $p < p_c$ to be

$$\xi(p) = \lim_{m \to \infty} \frac{m}{\ln \frac{1}{f_m(p)}} \quad .$$

There are other definitions of $\xi(p)$ but they all agree up to a uniformly bounded factor. This quantity can be shown to exist and moreover, when p is bounded away from 0, the convergence is uniform in some sense. In particular for every $\epsilon > 0$ there exists $c > 0$ such that for each $m > \xi(p)^{1+\epsilon}, c$ we have

$$e^{-\frac{m}{\xi(p)^{1-\epsilon}}} \leq f_m(p) \leq e^{-\frac{m}{\xi(p)^{1+\epsilon}}} \quad .$$

The correlation length $\xi(p)$ is finite for $p < p_c$ and tends to $+\infty$ when $p \to p_c$. It is conjectured that

$$\ln \xi(p) \sim \nu \cdot \ln \frac{1}{p_c - p} \qquad \text{where} \quad \nu = \frac{4}{3} \quad .$$

The constant ν is one of the *critical exponents* encountered in percolation theory. Den Nijs [5] has given plausible but not rigorous arguments in support of this conjecture.

Assuming that this conjecture is true, we get that an $m \times (m+1)$ grid, where $m = n^{\frac{4}{3}(1+2\epsilon)}$ amplifies $(p_c - \frac{1}{n}, p_c + \frac{1}{n})$ to about $(\frac{1}{4}, \frac{3}{4})$. It is interesting to note that by taking $m = n^{\frac{7}{3}(1+2\epsilon)}$ we get networks that amplify $(p_c - \frac{1}{n}, p_c + \frac{1}{n})$ to about $(2^{-n}, 1 - 2^{-n})$. The bounds on $f_m(p_c - \frac{1}{n})$ follow easily from the above estimates. The bounds on $f_m(p_c + \frac{1}{n})$ follow using the same arguments for the dual grid.

An amazing consequence of the *scaling* or *renormalization* theory of percolation is the conjecture that critical exponents, such as ν, depend only on the dimension of the grid used and not on it exact structure (note that the critical probability p_c for example does depend on the exact structure). If this conjecture is true, the same asymptotic results could be obtained using triangular, hexagonal or any other regular two dimensional construct such as the *hammock* structure favoured by Moore and Shannon [13]. We have experimented with many small size networks and found however that of all networks with the same (small) size, rectangular grids seem to give the best results.

How about using higher dimensional grids? (note that unit D of Fig. 1 is a cube and it had quite a good behaviour). Experimental evidence (see Stauffer and Aharony [18], p. 52), gives the following approximations to the critical exponent ν in dimensions 3,4 and 5: $\nu_3 \simeq 0.88$, $\nu_4 \simeq 0.68$ and $\nu_5 \simeq 0.57$. As $3\nu_3 \simeq 2.64 < \frac{8}{3}$ it seems that we could do a bit better using 3-dimensional grids. In fact, it might be possible to use fractional dimensions between 2 and 4 to get further improvements. A d-dimensional grid is just an (infinite) graph in which each vertex has $\theta(k^d)$ neighbours at distance k.

11 Open problems

1. Do the lower bounds on amplifying formulae given in this paper apply to general, not necessarily read-once formulae? Do they apply to general monotone formulae? Do they apply, say, to read-twice formulae?

2. What is asymptotically the minimal size of a contact network that amplifies $(p - \frac{1}{n}, p + \frac{1}{n})$ (for some fixed $0 < p < 1$) to $(\frac{1}{4}, \frac{3}{4})$? Of all networks with the same number of edges, which ones have the largest derivative at their critical probability?

3. Are the optimal monotone formulae and undirected monotone contact networks for the majority obtained using the amplification method?

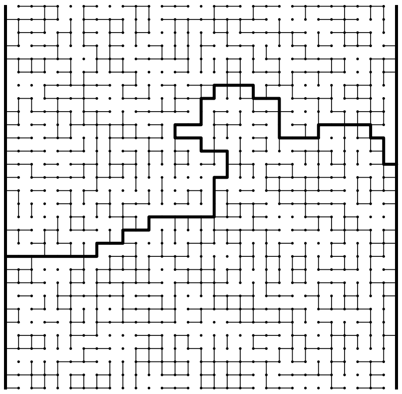

Figure 2: Percolation on a rectangular grid.

Acknowledgement

The authors would like to thank Noga Alon for many helpful discussions and to Edith Magen and Dalya Jacobsen for their continued interest and encouragement.

References

[1] L. Adelman, "Two theorems on random polynomial time", *Proceedings of the 19th IEEE Symposium on Foundations of Computer Science, 1978, pp. 75–83.*

[2] M. Ajtai and M. Ben-Or, "A theorem on probabilistic constant depth computations", *Proceedings of the 16th ACM Symposium on Theory of Computing, 1984, pp. 471–474.*

[3] C.G. Bennett and J. Gill, "Relative to a random oracle A, $P^A \neq NP^A \neq co-NP^A$ with probability 1", *SIAM Journal on Computing, Vol. 10 (1981), pp. 96–113.*

[4] R.B. Boppana, "Amplification of probabilistic Boolean formulas", in *Advances in Computer Research, Vol.5: Randomness and Computation*, JAI Press, Greenwich, Conn., 1989.

[5] M.P.M. den Nijs, "A relation between the temperature exponents of the eight-vertex and q-state Potts model", *Journal of Physics A: Mathematical and General, Vol. 12 (1979), pp. 1857-1868.*

[6] G. Grimmett, *Percolation*, Springer-Verlag, New York, 1989.

[7] Q.P., Gu and A. Maruoka, "Amplification of Bounded Depth Monotone Read-Once Boolean Formulae", *SIAM J. Computing, Vol. 20 (1991), pp. 41-55.*

[8] A. Gupta and S. Mahajan, "Using Amplification to Compute Majority with Majority", *preprint.*

[9] J. Håstad, A. Razborov and A. Yao, "On the Shrinkage Exponent for Read-Once Formulae", *preprint.*

[10] H. Kesten, *Percolation Theory for Mathematicians*, Birkhäuser, Boston, 1982.

[11] H. Kesten, "Percolation Theory and First-Passage Percolation", *Annals of Probability, Vol. 15 (1987), pp. 1231-1271.*

[12] O.B. Lupanov, "On Computing symmetric functions of the propositional calculus by switching networks" (in Russian), *Problemy Kibernet. Vol. 15 (1965), pp. 85-100.*

[13] E.F. Moore and C.E. Shannon, "Reliable circuits using less reliable relays", *Journal of the Franklin Institute Vol. 262 (1956), pp. 191-208 and pp. 281-297.*

[14] M.S. Paterson, N. Pippenger and U. Zwick, "Faster circuits and shorter formulae for multiple addition, multiplication and symmetric Boolean functions," *Proceedings of the 31st Ann. IEEE Symp. on Found. of Comp. Sci.*, St. Louis 1990, pp. 642-650.

[15] M.S. Paterson, N. Pippenger and U. Zwick, "Optimal carry save networks," *Boolean function complexity: Selected papers from the LMS symposium, Durham 1990.* To appear, Cambridge Univ. Press, 1992.

[16] M.S. Paterson, U. Zwick, "Shrinkage of de Morgan formulae under restriction", *Proceedings of 32st FOCS, San Juan, Puerto-Rico, 1991, pp. 324-333.* Journal version to appear in *Random Structures & algorithms.*

[17] J. Radhakrishnan and K.V. Subrahmanyam, *personal communication.*

[18] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, Taylor & Francis, London, 1992.

[19] L.G. Valiant, "Short monotone formulae for the majority function", *J. Algorithms 5 (1984) 363-366.*