Very Weak Zero One Law for Random Graphs with Order and Random Binary Functions

Saharon Shelah*
Institute of Mathematics, The Hebrew University, Jerusalem, Israel
Department of Mathematics, Rutgers University, New Brunswick, NJ
Department of Mathematics, University of Wisconsin, Madison, WI

Received 10 November 1994; revised 8 December 1995

ABSTRACT: Natural languages and random structures are given for which there are sentences A with no limit probability, yet for every A the difference between the probabilities that A holds on random structures of sizes n and $n+1$ approaches zero with n. © 1996 John Wiley & Sons, Inc. *Random Struct. Alg.*, **9**, 351–358 (1996)

INTRODUCTION

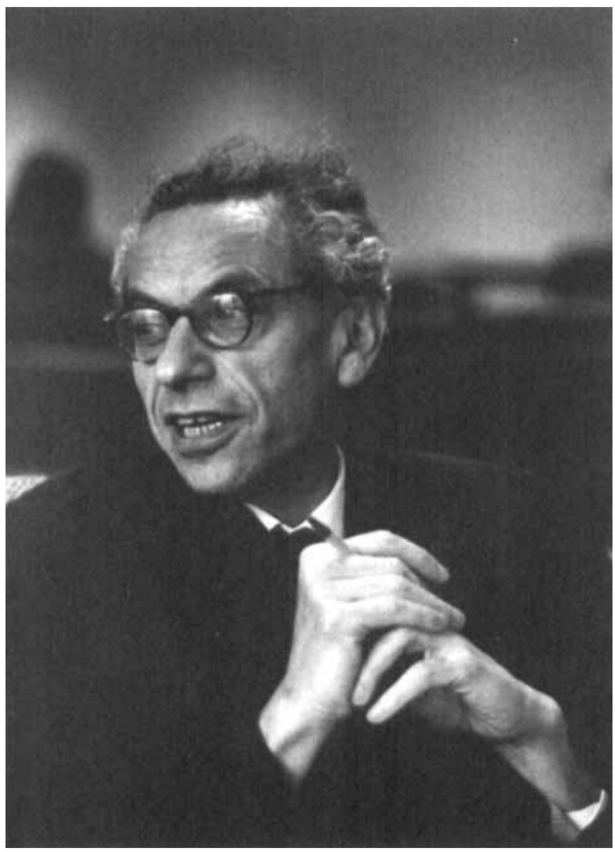
Let $G_<(n, p)$ denote the usual random graph $G(n, p)$ on a totally ordered set of n vertices. (We naturally think of the vertex set as $1, \ldots, n$ with the usual $<$). We will fix $p = \frac{1}{2}$ for definiteness. Let $L^<$ denote the first-order language with predicates equality $(x = y)$, adjacency $(x \sim y)$ and less than $(x < y)$. For any sentence A in $L^<$ let $f(n) = f_A(n)$ denote the probability that the random $G_<(n, p)$ has property A. It is known (Compton, Henson, and Shelah [4]) that there are A for which $f(n)$ does not converge. Here we show what is called a *very weak zero–one law* (from [8]):

---

*Correspondence to: Dr. Saharan Shelah.
Contract grant sponsor: United States–Israel Binational Science Foundation and NSF
Contract grant number: 144-EF67.

Sh:548

"The search for truth is more precious than its possession."—Albert Einstein



Paul Erdős

1913–1996

Theorem 0.1. *For every A in language $L^{<}$*

$$\lim_{n\to\infty} (f_A(n+1) - f_A(n)) = 0 .$$

Note, as an extreme example, that this implies the nonexistence of a sentence A holding with probability $1 - o(1)$ when n is even and with probability $o(1)$ when n is odd (as in Kaufman and Shelah [6]).

In Section 2 we give the proof, based on a circuit complexity result. In Section 3 we prove that result, which is very close to the now classic theorem, that parity cannot be given by an AC^0 circuit. In Section 4 we give a very weak zero–one law for random two-place functions. The proof is very similar, the random function theorem being perhaps of more interest to logicians, the random graph theorem to discrete mathematicians.

On a work of Boppana and Spencer continuing this, see Section 3.2(**5**).

1. THE PROOF

Let G be a fixed graph on the ordered set $1, \dots, 2n+1$. For a property A and for $i = n, n+1$, let $g(i) = g_{G,A}(i)$ denote the probability that $G\restriction_S$ satisfies A, where S is chosen uniformly from all subsets of $1, \dots, 2n+1$ of size precisely i. We shall show

Theorem 1.1. $g(n+1) - g(n) = o(1)$.

More precisely, given A and $\epsilon > 0$ there exists n_0 so that for any G as above with $n \ge n_0$ we have $|g(n+1) - g(n)| < \epsilon$.

We first show that Theorem 0.1 follows from Theorem 1.1. The idea is that a random $G_<(i, p)$ on $i = n$ or $n+1$ vertices is created by first taking a random $G_<(2n+1, p)$ and then restricting to a random set S of size i. Thus (fixing A) $f(n)$ and $f(n+1)$ are the averages of $g_G(n)$ and $g_G(n+1)$ over all G. By Theorem 1.1 we have $g_{G,A}(n) - g_{G,A}(n+1) = o(1)$ for all G, and therefore their averages are only $o(1)$ apart.

Now we show Theorem 1.1. Fix G and A as above. Let $P(S)$ be the Boolean value of the statement that $G|_S$ satisfies A. For $1 \le x \le 2n+1$ let z_x denote the Boolean value of "$x \in S$" so that $P(S)$ is a Boolean function of $z_1, \dots, z_{2n+1}$. We claim this function has a particularly simple form. Any A can be built up from primitives $x = y$, $x < y$, $x \sim y$ by $\wedge$, $\neg$ and, critically, $\exists_x$. As G is fixed the primitives have values true or false. Let $\wedge$ and $\neg$ be themselves. Consider $\exists_x W(x)$ where, for each $1 \le x \le 2n+1$, we let $W(x)$ on $G|_S$ is given by $W^*(x)$. Then $\exists_x W(x)$ has the interpretation $\exists_{x\in S} W(x)$ which is expressed as $\vee_{x=1}^{2n+1} (z_x \wedge W^*(x))$. For convenience we can be redundant and replace $\forall_x W(x)$ by $\wedge_{x=1}^{2n+1} (z_x \Rightarrow W^*(x))$. For example $\forall_x \exists_y x \sim y$ becomes

$$\wedge_x [z_x \Rightarrow \vee_{y\sim x} z_x] .$$

Thus $P(S)$ can be built up from $z_1, \dots, z_{2n+1}$ by means of the standard $\neg$, $\wedge$, $\vee$, and $\wedge$, $\vee$ over (at most) $2n+1$ inputs. That is (see Section 3), $P(S)$ can be expressed by an AC^0 circuit over $z_1, \dots, z_{2n+1}$ [of course, with the number of

Sh:548

levels bounded by the length d_A of the sequence A (can get less) and the number of nodes bounded by $d_A n^{d_A}$]. Now $g(i)$, for $i = n, n+1$, is the probability P holds when a randomly chosen set of precisely i of the z's are set to True. From Theorem 2.1 below $g(n+1) - g(n) = o(1)$ giving Theorem 1.1 and hence Theorem 0.1.

2. AC^0 FUNCTIONS

We consider Boolean functions of $z_1, \ldots, z_m$. (In our application $m = 2n+1$.) The functions z_i and $\neg z_i$, called literals, are the level 0 functions. A level $i+1$ function is the $\wedge$ or $\vee$ of polynomially many level i functions. An AC^0 function is a level d function for any constant d. By standard technical means we can express any AC^0 function in a "leveled" form so that the level $i+1$ functions used are either all $\wedge$'s of level i functions or all $\vee$'s of level i functions and the choice alternates with i (at most doubling the number of levels). It is a classic result of circuit complexity that parity is not an AC^0 function. Let C be an AC^0 function. For $0 \le i \le m$ let $f(i) = f_C(i)$ denote the probability C holds when precisely i of the z_j are set to True and these i are chosen randomly.

Theorem 2.1. $f(n+1) - f(n) = o(1)$.

Called a restriction ρ balanced if $|\{i: \rho(i) = 0\}| = |\{i: \rho(i) = 1\}|$. Now more fully the theorem says:

(*) For every ϵ, d, t there is $n_{\epsilon,d,t}$ satisfying: If $n \ge n_{\epsilon,d,t}$ and C is an AC^0 Boolean circuit of $z_1, \ldots, z_{2n+1}$ of level $\le d$ with $\le n^t$ notes, then $|f_C(n+1) - f_C(n)| < \epsilon$.

This statement is proved by induction on d.

We choose the following:

(i) $c_0 = (\ln 4)t > 0$,
(ii) $\epsilon = \frac{1}{2}$, $\epsilon_\ell = \frac{1}{2^{1+\ell}}$,
(iii) k is such that $\epsilon \cdot k \ge t$,
(iv) we choose k_ℓ inductively on $\ell \le k$ such that k_ℓ large enough,
(v) c_1 is large enough to real,
(vi) n_0 is large enough.

For a node x of the circuit C let $\mathcal{Y}_x$ be the set of nodes which fans into it; (without loss of generality in the level 1 we have only OR).

First we assume $d > 2$. Note:

$\otimes_1$ Drawing as below a balanced restriction ρ with domain with $\le n$ elements, with probability $\ge 1 - \epsilon/3$, we have that, in $C^1 = C \restriction_\rho$, every node of the level 1 (i.e., for which $\mathcal{Y}_x$ is a set of atoms) satisfies $|\mathcal{Y}_x| \le c_0(\ln n)$.

Sh:548

[Why? Choose randomly a set $\mathbf{u}_0$ of $[n/2]$ pairwise disjoint pairs of numbers among $\{1, \dots, 2n+1\}$, and then for each $\{i, j\} \in \mathbf{u}$ decide with probability half that $\rho(i) = 0$, $\rho(j) = 1$ and with probability half that $\rho(i) = 1$, $\rho(j) = 0$ (independently for disjoint pairs). This certainly gives a balanced ρ.

Now if x is a node of C of the level 1, the probability that ρ does not decide the truth value which the node compute is $\leq (\frac{1}{4})^{|\mathcal{Y}_x|}$. Note: after drawing $\mathbf{u}$, if $\mathcal{Y}_x$ contains a pair from $\mathbf{u}$ the probability is zero, we only increase compared to drawing just a restriction. So the probability that, for *some* x of the level 1 of C, $|\mathcal{Y}_x| \geq (\ln 4)t(\ln n) + 1$ and the truth value is not computed, is $\leq |C| \times (\frac{1}{4})^{(\ln 4)t(\ln n)+1} \leq 1/2$, so there is ρ_0 for which for any such x the truth value is computed.]

Next, we say that a restriction ρ' extends a restriction ρ if $\rho'(i) \neq \rho(i) \Rightarrow \rho(i) = *$. Now:

$\otimes_2$ Choosing randomly a restriction ρ_1 as below we have: ρ_1 is a balanced restriction extending ρ such that $|\{i : i \in \{1, \dots, 2n+1\},\ \rho(i) = *\}| \geq 2[n^\epsilon] + 1$ and with probability $\geq 1 - \epsilon/3$ for every node y of C of the level 1 we have, $|\mathcal{Y}_y| \leq k$.

[Why? We draw a set $\mathbf{u}_1$ of $(2n + 1 - |\mathrm{dom}(\rho_0)| - (2[n^{\epsilon_0}] + 1))/2$ pairs from $\{i : \rho_0(i) = *\}$ pairwise disjoint and for each $\{i, j\} \in \mathbf{u}$, decide with probability $\frac{1}{2}$ that $\rho_1(i) = 0$, $\rho_1(j) = 1$ and with probability half that $\rho_1(i) = 1$, $\rho_1(j) = 0$.

For each node $y \in C^1$ of the level 1 the probability that "the number of $y' \in \mathcal{Y}_y$ not assigned a truth value by ρ_1 is $\geq k+1$" is at most $\binom{|\mathcal{Y}_y|}{k+1} \times (\frac{1}{2n^{\epsilon_0}+1})^{k+1} \leq (c_0 \ln n)^{k+1} \cdot n^{-\epsilon_0(k+1)} < n^{-t}$.]

We now choose by induction on $\ell \leq k$ a restriction $\rho_{2,\ell}$ such that

$\otimes_3$ **(a)** $\rho_{2,\ell_0} = \rho_1$, $\rho_{2,\ell} \subseteq \rho_{2,\ell+1}$, $2n + 1 - (2[n^{\epsilon_\ell}] + 1) = |\mathrm{dom}\ \rho_{2,\ell}|$,

(b) for every $y \in C$ of the level 2 there is a set $w_{y,\ell}$ of $\leq k_\ell$ atoms such that if $z \in \mathcal{Y}_y$, then $|\mathcal{Y}_z \backslash w_{y,\ell}| \leq k - \ell$.

Now for $C_{\restriction \rho_{2,k}}$ we can invert AND and OR (multiplying the size by a constant $\leq c_1$) decreasing d by 1, thus carrying the induction step.

For $\ell = 0$ let $\rho_{2,0} = \rho_1$. For $\ell + 1$, for each $y \in C$ of level 2 let $\Xi = \{\nu : \nu$ a restriction with domain $w_{y,\ell}\}$, let

$$\mathcal{Y}_y^\nu = \{z \in \mathcal{Y}_y : \text{the truth value at } z \text{ is still not computed under } \rho_{2,\ell} \cup \nu\},$$

and try to choose by induction on i an atom $z_{y,\ell,\nu,i} \in \mathcal{Y}_y^\nu \backslash \{z_{y,\ell,\nu,j} : j < i\}$, such that $\mathrm{dom}(z_{y,\ell,\nu,i})$ is disjoint to $\bigcup_{j<i} \mathrm{dom}(z_{y,\ell,\nu,j}) \backslash w_{y,\ell}$. Let it be defined if $i < i_{y,\ell}$.

Now $\rho_{2,\ell+1}$ will for each $\nu \in \Xi$ decide that ν makes the truth value computed in y true, or will leave only $\leq (k_{\ell+1} - k_\ell)/2^{k_\ell}$ of the atoms in $\bigcup \mathrm{dom}\ z_{y,\ell,\nu,i} \backslash w_{y,\ell}$ undetermined (this is done as in the previous two stages).

But now by $\otimes_1 + \otimes_2$, $C_{\restriction \rho_{2,k}}$ can be considered having $d - 1$ levels (because, as said above we can invert the AND and OR in level 1 and 2).

We have translate our problem to one with $[n^{\epsilon_k}]$, $d - 1$, $\epsilon_k(t + \epsilon_1)$, $\frac{\epsilon}{3}$ instead n, d, t, and ϵ (the $t + \epsilon$ is just for $n^{t+\epsilon} > c_1 n^t$).

Also note that ϵ and c_1 do not depend on n. So we can use the induction hypothesis. We still have to check the case $d \leq 2$, we still are assuming level 1 consist of cases of OR, and for almost all random ρ_1 (as in $\otimes_1$) for every x of level 1 we have $|\mathcal{Y}_x| \leq \mathbf{c}_0 \ln n$ (so again changing n).

So as above we can add this assumption. Choose randomly a complete restriction ρ^0 with $|\{i: \rho^0(i) = 1\}| = n$, and let ρ^1 be obtained from ρ^0 by changing one zero to 1, so $|\{i: \rho^1(i) = 1\}| = n + 1$.

Now the probability that $C_{\restriction \rho^0} = 0$ but $C_{\restriction \rho^1} = 1$ is small: It requires, that, for some node, x of level 1 is made false in $C_{\restriction \rho^0}$ while there is no such x for $C_{\restriction \rho^1}$, but if $x(*)$ is such for $C_{\restriction \rho^0}$ that is made true, then with probability $\geq 1 - \frac{|\mathcal{Y}_x|}{2n+1} \geq 1 - \frac{\mathbf{c}_0 \ln n}{n}$ the z_i changed is not in $\mathcal{Y}_{x(*)}$. Contradiction, thus finishing the proof.

3. TWO PLACE FUNCTIONS

Here we consider the random structure $([n], F_n)$, where $F_n(x, y)$ is a random function from $[n] \times [n]$ to $[n]$. (We no longer have order. A typical sentence would be $\forall_x \exists_y F(x, y) = x$.): Again for any sentence A we define $f(n) = f_A(n)$ to be the probability A holds in the space of structures on $[n]$ with uniform distribution. Again it is known [4] that convergence fails; there are A's for which $f(n)$ does not converge. Again our result is a very weak zero–one law.

Theorem 3.1 *For every A*

$$\lim_{n\to\infty} f_A(n+1) - f_A(n) = 0\,.$$

Again let $m = 2n + 1$. Let $F^*(x, y, z)$ be a *three*-place function from $[m] \times [m] \times [m]$ to $[m]$. For $S \subset [m]$ of cardinality $i = n$ or $n + 1$ we define F^*_S, a partial function from $[S] \times [S]$ to $[S]$ by setting $F^*_S(x, y) = F^*(x, y, z)$, where z is the minimal value for which $F^*(x, y, z) \in S$. If there is no such z then $F^*_S(x, y)$ is not defined. This occurs with probability $(\frac{m-i}{m})^m$ for any particular x, y so the probability f^*_S is not always defined is at most $i^2(\frac{m-i}{m})^m = o(1)$.

We generate a random three-place F^* and then consider F^*_S with S a random set of size $i = n$ or $n + 1$. Conditioning on F^*_S being always defined it then has the distribution of a random two-place function on i points. Thus $\Pr[A]$ over $[n]$, F_n is within $o(1)$ of $\Pr[A]$ when $F_n = F^*_S$ is chosen in this manner. Thus, as in Section 2, it suffices to show for any F^* and A that, letting $g(i)$ denote the probability that F^*_S satisfies A with S a uniformly chosen i-set, $g(n+1) - g(n) = o(1)$. Again fix F^* and A and let z_x be the Boolean value of $x \in S$ for $1 \leq x \leq 2n + 1$. In A replace the ternary relation $F(a, b) = c$ by $\wedge_{y<d} \neg z_{F(a,b,y)}$ & $z_{F(a,b,d)}$. (For $d = 1$ this is simply True.) As in Section 2 replace $\exists_x P(x)$ by $\vee_x (z_x \wedge P^*(x))$, where $P^*(x)$ has been inductively defined as the replacement of $P(x)$. Then the statement that F^*_S satisfies A becomes a Boolean function of the $z_1, \ldots, z_m$, as before it is an AC^0 function, and by Section 2 we have $g(n+1) - g(n) = o(1)$.

* * *

The following discussion is directed mainly for logicians but may be of interest for CS-oriented readers as well.

Discussion 3.2.

(1) Note that the results of [8] cannot be obtained in this way as the proof here uses high symmetry. The problem there was: let $\bar{p} = \langle p_i : i \in \mathbb{N} \rangle$ be a sequence of probabilities such that $\Sigma_i\, p_i < \infty$. Let $G(n, \bar{p})$ be the random graph with set of nodes $[n] = \{1, \ldots, n\}$ and the edges drawn independently, and for $i \neq j$ the probability of $\{i, j\}$ being an edge is $p_{|i-j|}$.

The very weak 0–1 law was proved for this context in [8] (earlier on this context (probability depending on distance) was introduced and investigated in Łuczak and Shelah [7]. Now drawing $G(2n+1, \bar{p})$ and then restricting ourselves to a random $S \subseteq \{1, \ldots, 2n+1\}$ with n, and with $n+1$ elements, fail as $G(2n+1, \bar{p})_{\restriction S}$ does not have the same distribution as $G(|S|, \bar{p})$.

(2) We may want to phrase the result generally;
One way: just say that M_n, M_{n+1} can be obtained as above: Draw a model on $[2n+1] = \{1, \ldots, 2n+1\}$ (i.e., with this universe), and then choose randomly subsets P_n^{ℓ} with $n + \ell$ elements and restrict yourself to it.

(3) Two random linear order satisfies the very weak 0–1 zero law (meaning take two random functions from $[n]$ to $[0, 1]_{\mathbb{R}}$). The proof should be clear.

(4) All this is for fixed probabilities; we then allow probabilities depending on n, e.g., we may consider $G_<(n, p_n)$ as the model with set of elements $\{1, \ldots, n\}$, the order relation, and we draw edges with edge probability p_n depending on n. This call for estimating two number (for φ first-order sentence):

$$\alpha_n = |\mathrm{Prob}(G_<(n, p_{n+1}) \models \varphi) - \mathrm{Prob}(G_<(n, p_n) \models \varphi)|\,,$$

$$\beta_n = |\mathrm{Prob}(G_<(n, p_{n+1}) \models \varphi) - \mathrm{Prob}(G_<(n+1, p_{n+1}) \models \varphi)|\,.$$

As for β_n the question is how much does the proof here depend on having the probability $\frac{1}{2}$. Direct inspection on the proof shows that it does not at all (this just influence on determining the specific Boolean function with $2n+1$ variables) so we know that the β_n converge to zero.

As for α_n, clearly the question is how fast the p_n change.

(5) As stated in [8], we can also consider $\lim(\mathrm{Prob}_{n+h(n)}(M_{n+h(n)} \models \psi) - \mathrm{Prob}_n(M_n \models \psi)) = 0$, i.e., characterize the function h for which this holds, but this was not dealt with there. Hopefully there is a threshold phenomena. Probably this family of problems will appeal to mathematicians with an analytic background.

Another problem, closer to my heart, is to understand the model theory:

In some sense first-order formulas cannot express too much, but can we find a more direct statement fulfilling this?

* * *

Another way to present the first problem for our case is: Close (or at least narrow) the analytic gap between [4] and the present paper.

After this work, Boppana and Spencer [3], continuing the present paper and [8], address the problem and completely solve it. More specifically they proved the following.

For every sentence A there exists a number t so that $m(n) = O(n \ln)^{-t} n)$ implies

$$\lim_{n\to\infty} f_A(n + m(n)) - f(A)(n) = 0 .$$

Also,

For every number t there exists a sentence A and a function $m(n) = O(n \ln^{-t} n)$ so that $f_A(n + m(n) - f_A(n)$ does not approach zero.

Together we could say: A function $m(n)$ has the property that for all A and all $m'(n) \leq m(n)$ we have $f_A(n + m'(n)) - f_A(n) \to 0$ if and only if $m(n) = o(n \ln^{-t} n)$ for all t.

For improving the bound from this side they have used Hastad switching lemma [Hastad] (see [1], Section 11.2, Lemma 2.1).

(6) If we use logic stronger than first order it cannot be too strong (on monadic logic see [6]), but we may allow quantification over subsets of size k_n; e.g., $\log(n)$ there are two issues:

(A) When for both n and $n+1$ we quantify over subsets of size k_n, we should just increase M by having the set $[n]^{k_n}$ as a set of extra elements, so in $(*)$, P is chosen as a random subset of $\{1, 2, 3, \ldots, 2n, 2n+1\}$ with n or $n+1$ elements but the model has about $(2n+1)^{k_n}$ elements; this requires a stronger theorem, still true (up to very near to exponentiation).

(B) If $k_n \neq k_{n+1}$, we need to show that it does not matter; we may choose to round $k_n = \log_2(n)$ so only for rare n do the values change. Thus we weaken a little the theorem *or* we may look at sentences for which this does not matter.

Maybe more naturally, together with randomly choosing $\mathcal{M}_n$ we choose a number $\underline{k}_n$, and the probability of $\underline{k}_n = k_n + i$ if $i \in [-k_n/2, k_n/2]$ being $1/k_n$.

And we ask for $p_n^{\varphi} =: \mathrm{Prob}(\mathcal{M}_n \models \varphi$, where the monadic quantifier is interpreted as varying on set with $\leq \underline{k}_n$ elements) for sentence φ (the point of the distribution of $\underline{k}_n$ is just that for n, $n+1$ they differ a little). For example, if for a random graph on n (probability 0.5) we ask on the property that "the size of maximal clique of size at most $[\log_2 n]^2$ is even," it satisfies the very weak zero–one law.

Of course, we know much more on this, still it shows that this old result (more exactly, a weakened version) can be put in our framework.

ACKNOWLEDGMENTS

The reader should thank Joel Spencer who totally rewrote the paper (using the computer science jargon rather than that of logicians), and with some revisions up to the restatement in the proof of 2.1 but with 3.1, this is the version presented here. We thank the referee for comments on the exposition, and we thank Tomasz Łuczak and Joel Spencer for reminding me this problem on $G_<(n, p)$ in summer 1993.

REFERENCES

1. M. Ajtai, A Σ_1^1 formula on finite structures, *Ann. Pure Appl. Logic*, **24**, 1–49 (1983).
2. N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York, 1992.
3. R. B. Boppana and J. H. Spencer, Smoothness laws for random ordered graphs, *Logic and Random Structures: DIMACS Workshop*, November 5–7, 1995, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI, to appear.
4. K. Compton, C. S. Henson, and S. Shelah, Non-convergence, undecidability and intractability in asymptotic problems, *Ann. Pure Appl. Logic*, **36**, 207–224 (1987).
5. J. Hastad, Almost optimal lower bounds for small depth circuits, in *Advances in Computer Research*, S. Micali, Ed., JAI Press, Greenwich, CT, Vol. 5, pp. 143–170.
6. M. Kaufman and S. Shelah, On random models of finite power and monadic logic, *J. Symb. Logic*, **54**, 285–293 (1985).
7. T. Łuczak and S. Shelah, Convergence in homogeneous random graphs, *Random Struct. Alg.*, **6**, 371–391 (1965).
8. S. Shelah, On the very weak 0–1 law for random graphs with order, *J. Logic and Comput.*, **6**, 137–159 (1966).