1-1-1986

Reliable Computation in the Presence of Noise

Nicholas Pippenger
Harvey Mudd College

Proceedings of the International Congress of Mathematicians
Berkeley, California, USA, 1986

Reliable Computation in the Presence of Noise

NICHOLAS PIPPENGER

1. Introduction. This talk concerns computation by systems whose components exhibit noise (that is, errors committed at random according to certain probabilistic laws). If we aspire to construct a theory of computation in the presence of noise, we must possess at the outset a satisfactory theory of computation in the absence of noise. A theory that has received considerable attention in this context is that of the computation of Boolean functions by networks (with perhaps the strongest competition coming from the theory of cellular automata; see [**G**] and [**GR**]).

The theory of computation by networks associates with any two sets Q and R of Boolean functions a number $L_Q(R)$ (the "size" of R with respect to Q), defined as the minimum number of "gates," each computing a function from the basis Q, that can be interconnected to form a "network" that computes all of the functions in R. This theory has many pleasant properties, among which is the fact that if Q and Q' are finite and "complete," then

$$L_Q(R) \leq C_{Q,Q'} L_{Q'}(R), \tag{1.1}$$

for some constant $C_{Q,Q'}$ independent of R (see [**M**]). Thus, if one is unconcerned with constant factors, one may drop the subscript Q and consider $L(R)$ as a measure of the complexity of computing the functions in R. Another pleasant property, however, is the existence of an exquisitely precise theory of the complexity of "generic" functions. Thus for "almost all" functions f of degree n (that is, depending on n arguments), one has

$$L_Q(f) \sim C_Q 2^n/n \tag{1.2}$$

as $n \to \infty$, where C_Q is a constant independent of n (see [**L**]).

The theory of computation by networks in the presence of noise was founded by von Neumann [**N**]. Firstly, von Neumann showed that reliable computation in the presence of noise is possible. If a network N contains L gates, each of which fails with probability at most ε, then N fails with probability at most $L\varepsilon$. This crude bound becomes uninformative, however, if L grows while $\varepsilon > 0$ remains fixed. It was proved by von Neumann that N can be replaced by a network N', with a larger number L' of gates, so that N' fails with probability at most δ,

where $\delta < 1/2$ is fixed (independent of L and L') when ε is sufficiently small and the gates of N' fail independently with probability ε.

Let $L'_{Q,\varepsilon,\delta}(R)$ denote the counterpart to $L_Q(R)$ when the network must fail with probability at most δ, given that each gate fails independently with probability ε. A heuristic argument to the effect that

$$L'_{Q,\varepsilon,\delta}(R) = O(L_Q(R) \log L_Q(R)) \tag{1.3}$$

was given by von Neumann; this was proved rigorously by Dobrushin and Ortyukov [**DO1**]. They also gave, in [**DO2**], a sequence f_n of functions such that

$$L_Q(f_n) = O(n),$$

but

$$L'_{Q,\varepsilon,\delta}(f_n) = \Omega(n \log n),$$

so that the estimate (1.3) is, in general, the best possible. On the other hand, I have shown in [**P**] that the estimate

$$L'_{Q,\varepsilon,\delta}(f) = O(L_Q(f)) \tag{1.4}$$

holds not only for many specific functions, but also for "almost all" functions in the sense of (1.2). Results such as (1.3) and (1.4), and others not mentioned here, form the core of a theory with many of the properties typified by (1.1) and characterized by a lack of concern for constant factors. A theory with results like (1.2), however, seems far beyond our grasp at this time.

My goal in this talk is to sketch a theory in which results like (1.2) may be within reach, though they have not yet been obtained. My proposal is to consider formulae, which behave rather more simply than networks, and to consider depth, which behaves rather more simply than size.

Let $\mathbf{B}$ denote the Boolean algebra with 2 elements. These elements will be denoted 0 ("false") and 1 ("true"); the operations will be denoted $(x, y) \mapsto x \wedge y$ ("and," or conjunction), $(x, y) \mapsto x \vee y$ ("or," or disjunction) and $x \mapsto \bar{x}$ ("not," or negation).

By a *Boolean function* we shall mean a map $f\colon \mathbf{B}^n \to \mathbf{B}$, for some n which is called the *degree* of f. Let $x_1, \ldots, x_n$ be indeterminates, and let $\mathbf{B}(x_1, \ldots, x_n)$ denote the extension of $\mathbf{B}$ by $x_1, \ldots, x_n$. The Boolean functions of degree n are in an obvious one-to-one correspondence with the elements of $\mathbf{B}(x_1, \ldots, x_n)$, which will therefore also be called Boolean functions. Boolean functions of various degrees are thereby identified in accordance with the filtration $\mathbf{B} \subseteq \mathbf{B}(x_1) \subseteq \mathbf{B}(x_1, x_2) \subseteq \cdots \subseteq \mathbf{B}(x_1, \ldots, x_n) \subseteq \cdots$.

By a *formula* on $x_1, \ldots, x_n$ over Q we shall mean an expression of one of three kinds. The first kind, a *source*, is an expression c, where $c \in \mathbf{B}$; it has *depth* 0 and computes the constant function $c \in \mathbf{B}(x_1, \ldots, x_n)$. The second kind, an *input*, is an expression x_m, where $1 \leq m \leq n$; it has depth 0 and computes the projection function $x_m \in \mathbf{B}(x_1, \ldots, x_n)$. The third kind, a *gate*, is an expression $g(N_1, \ldots, N_k)$, where $g \in Q$ and $N_1, \ldots, N_k$ are formulae on

$x_1, \ldots, x_n$ over Q; if $N_1, \ldots, N_k$ have depths $d_1, \ldots, d_k$, respectively, and compute the functions $f_1(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n)$, respectively, then it has depth $1 + \max\{d_1, \ldots, d_k\}$ and computes the function

$$g(f_1(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n)) \in \mathbf{B}(x_1, \ldots, x_n).$$

A set Q is *complete* if every Boolean function is computed by some formula over Q. If Q is complete, define $D_Q(f)$ to be the minimum possible depth of a formula over Q that computes f. If R is finite, define $D_Q(R)$ to be the maximum of $D_Q(f)$ over $f \in R$.

It is easy to see that

$$D_Q(RS) \leq D_Q(R) + D_Q(S), \tag{1.5}$$

where RS denotes the set of functions obtained by substituting functions from S for the arguments of functions from R. We also have

$$D_Q(S) \leq D_Q(R) D_R(S), \tag{1.6}$$

which is the counterpart to (1.1) for depth.

To discuss computation by formulae in the presence of noise, we must adopt probabilistic assumptions about the errors, then reconsider what it means for a formula to "compute" a function. For technical reasons it is convenient to work not with probabilities of incorrect behavior, ε and δ, but with probabilities of correct behavior, $\rho = 1 - \varepsilon$ and $\sigma = 1 - \delta$. The assumptions we shall make are not the simplest ones, but they have the merit that they yield counterparts to (1.5) and (1.6).

Consider the evaluation of a function $f(x_1, \ldots, x_n)$ at a point $c_1, \ldots, c_n \in \mathbf{B}^n$ by a formula N. We shall say that $f(c_1, \ldots, c_n)$ is the *correct* value for N. Let M be a subformula of N. If M is a source c, it produces the correct value, c. If M is an input x_m, it produces the correct value, c_m, if M is *proper*; otherwise it produces $\overline{c_m}$. If $M = g(M_1, \ldots, M_k)$ is a gate, and if the subformulae $M_1, \ldots, M_k$ produce the values $m_1, \ldots, m_k$ (correct or not), then it produces $g(m_1, \ldots, m_k)$ (correct or not) if M is proper; otherwise it produces $\overline{g(m_1, \ldots, m_k)}$. We shall assume that each input is proper with probability at least α and each gate is proper with probability at least ρ, even when these probabilities are conditioned on other inputs or gates being proper or improper; these probabilities may also depend on $c_1, \ldots, c_n$. If in this situation N produces the correct value with probability at least β for all $c_1, \ldots, c_n$, we shall say that N (ρ, α, β)-*computes* f.

Let $D^*_{Q,\rho,\alpha,\beta}(f)$ denote the minimum possible depth of a formula over Q that (ρ, α, β)-computes f, and let $D^*_{Q,\rho,\alpha,\beta}(R)$ denote the maximum of $D^*_{Q,\rho,\alpha,\beta}(f)$ over $f \in R$.

It is clear that $D^*_{Q,\rho,\alpha,\beta}(R)$ is decreasing in Q, ρ, and α, and increasing in R and β, and that $D^*_{Q,\rho,\alpha,\beta}(R) \geq D_Q(R)$. We have

$$D^*_{Q,\rho,\alpha,\gamma}(RS) \leq D^*_{Q,\rho,\alpha,\beta}(R) + D^*_{Q,\rho,\beta,\gamma}(S), \tag{1.7}$$

which is the counterpart to (1.5). This inequality suggests that $D^*_{Q,\rho,\sigma,\sigma}(R)$ behaves particularly simply. Indeed,

$$D^*_{Q,\rho,\tau,\tau}(S) \leq D^*_{Q,\rho,\sigma,\sigma}(R) D^*_{R,\sigma,\tau,\tau}(S),$$

which is the counterpart to (1.6).

Let $D^*_{Q,\rho,\sigma}(d)$ denote the maximum of $D^*_{Q,\rho,\sigma,\sigma}(f)$ over all functions f such that $D_Q(f) \leq d$. A subadditivity argument based on (1.7) shows that

$$\lim_{d\to\infty} D^*_{Q,\rho,\sigma}(d)/d$$

exists; the limit represents the factor by which computations take longer in the presence of noise than in its absence.

2. An upper bound. We shall start with an exemplary theorem, due in essence to von Neumann [N]. All formulae will be over the complete basis {minor} (where

$$\text{minor}\{x, y, z\} = \overline{(x \wedge y) \vee (x \wedge z) \vee (y \wedge z)}$$

denotes the minority of its three arguments), so we shall drop subscripts indicating the basis.

LEMMA 2.1. *Let N be a formula that (ρ, α, ξ)-computes f. Then the formula* $\text{minor}\{N, N, N\}$ *$(\rho, \alpha, F(\xi))$-computes $\bar{f}$, where $F(\xi) = \rho(3\xi^2 - 2\xi^3)$.*

PROOF. The formula $\text{minor}\{N, N, N\}$ produces the correct value if at least two of its immediate subformulae produce the correct value and if the gate is proper. □

LEMMA 2.2. *Let N_1, N_2, and N_3 be formulae that (ρ, α, ξ)-compute f_1, f_2, and f_3, respectively. Then* $\text{minor}\{N_1, N_2, N_3\}$ *is a formula that $(\rho, \alpha, G(\xi))$-computes* $\text{minor}\{f_1, f_2, f_3\}$, *where $G(\xi) = \rho\xi^2$.*

PROOF. When f_1, f_2, and f_3 all assume the same value, we are in the situation of Lemma 2.1, and $F(\xi) \geq G(\xi)$. Otherwise, two of these functions assume a common value and the third assumes the complementary value. The formula $\text{minor}\{N_1, N_2, N_3\}$ produces the correct value provided that the corresponding two immediate subformulae produce the correct value and the gate is proper. □

If $\rho = (10/9)(5/6)^{2/3} = 0.9839\ldots$, $\sigma = (9/10)(6/5)^{1/3} = 0.9563\ldots$, and $\tau = 9/10$, then $\sigma = F(\tau)$ and $\tau = G(\sigma)$. (This value of ρ is the smallest for which such values of σ and τ can be found; it is a root of the discriminant of $F(G(\xi)) = \xi$.)

THEOREM 2.3. *Let $\rho = (10/9)(5/6)^{2/3}$ and $\sigma = (9/10)(6/5)^{1/3}$. For any Boolean function f,*

$$D^*_{\rho,\sigma,\sigma}(f, \bar{f}) \leq 2D(f) + 1.$$

PROOF. We proceed by induction on $D(f)$. If $D(f) = 0$, then f is a constant or a projection, and the claim follows from Lemma 2.1. Otherwise, f and $\bar{f}$ are each of the form $\text{minor}\{f_1, f_2, f_3\}$ where $D(f_1, f_2, f_3) \leq D(f) - 1$. The claim

follows by applying the inductive hypothesis to f_1, f_2, and f_3, then applying Lemma 2.2, and finally applying Lemma 2.1. □

The foregoing theorem shows that reliable computation in the presence of noise is possible, at least if $\rho \geq \rho_2 = (10/9)(5/6)^{2/3}$ and if we are willing to spend about twice the time. This is done by alternating "correcting steps" (Lemma 2.1) with "computing steps" (Lemma 2.2). If $\rho > \rho_2$, we might hope to perform more than one computing step per correcting step, and thus to obtain

$$D^*_{\rho,\sigma,\sigma}(f) \leq C_\rho D(f) + o(D(f)), \qquad (*)$$

with $C_\rho \to 1$ as $\rho \to 1$. If $\rho < \rho_2$, we still might hope to compute reliably by performing more than one correcting step per computing step, and thus to obtain $(*)$ for some $C_\rho > 2$, at least if ρ is not too small. When $\rho \leq 1/2$, the value produced by a gate can be statistically independent of the values computed by its immediate subformulae, and reliable computation will certainly not be possible. Thus we must expect $C_\rho \to \infty$ as $\rho \to \rho_1$ for some $\rho_1 \leq 1/2$. In the remainder of this section we shall indicate how these hopes may be fulfilled.

Let us consider the action of the maps F and G on the interval $(0, 1]$. If $\rho < 1$, G is deflationary: $G(\xi) < \xi$. Thus if the damage done by a computation step is to be ameliorated by a correction step, there must be values $\xi \in (0, 1]$ for which F is inflationary: $F(\xi) > \xi$. This happens precisely when $\rho > \rho_0 = 8/9$ (this value is a root of the discriminant of $F(\xi) = \xi$). When $\rho > \rho_0$, the equation $F(\xi) = \xi$ has two roots:

$$\xi^{\pm} = \frac{3 \pm \sqrt{9 - 8/\rho}}{4}.$$

Under iteration of F, ξ^- is a repulsive fixed point and ξ^+ is an attractive one. Thus if the damage done by a computation step is to be undone by a finite number of correction steps, we must in fact have $G(\xi^+) > \xi^-$. This happens precisely when $\rho > \rho_1 = (10 + 4\sqrt{13})/27 = 0.904\ldots$. This is the lower limit to the reliability for which the scheme we are describing works.

Suppose then that $\rho_1 < \rho < 1$. Suppose further that $\xi^- < \sigma < \xi^+$.

Let $\{F, G\}^*$ be the free monoid generated by the symbols F and G, and let $\langle F(\xi), G(\xi)\rangle$ be the monoid of polynomials under composition generated by $F(\xi)$ and $G(\xi)$. For every $W \in \{F, G\}^*$, let $P_W(\xi) \in \langle F(\xi), G(\xi)\rangle$ be the image of W under the homomorphism that sends $F \mapsto F(\xi)$ and $G \mapsto G(\xi)$. Given d, let $M(d)$ denote the minimum possible number of symbols in a word $W \in \{F, G\}^*$ that contains d occurrences of the symbol G and satisfies $P_W(\sigma) \geq \sigma$. A subadditivity argument shows that $\lim_{d\to\infty} M(d)/d$ exists. This is the ratio C_ρ by which computations are slowed down by the scheme we are describing.

To determine the behavior of $M(d)$, it is helpful to transform the problem. Given c, d, and ξ, let $T(c, d, \xi)$ denote the maximum of $P_W(\xi)$ over all words $W \in \{F, G\}^*$ that contain c occurrences of the symbol F and d occurrences of the symbol G. Then $M(d) = \min\{m \geq d\colon T(m - d, d, \sigma) \geq \sigma\}$. It is clear that $T(0, 0, \xi) = \xi$, $T(c, 0, \xi) = T(c - 1, 0, F(\xi))$ for $c \geq 1$, and $T(0, d, \xi) = T(0, d - 1, G(\xi))$ for $d \geq 1$. The only problem arises when $c, d \geq 1$ and one must

decide whether to apply $F(\xi)$ or $G(\xi)$ first. This problem can be resolved by considering the Poisson bracket: $[F,G](\xi) = F(G(\xi)) - G(F(\xi))$. If $[F,G](\xi) > 0$, it is more advantageous to apply $G(\xi)$ before $F(\xi)$. This happens precisely when $\xi > \xi_0$, where

$$\xi_0 = \frac{1}{1+\sqrt{(1-\rho)/3}}$$

(this value is a root of $F(G(\xi)) = G(F(\xi))$). A monotonicity argument shows that if $c, d \geq 1$, then $T(c,d,\xi) = T(c-1,d,F(\xi))$ if $\xi \leq \xi_0$ and $T(c,d,\xi) = T(c,d-1,G(\xi))$ if $\xi > \xi_0$. This recurrence, together with the boundary conditions given above, determines $T(c,d,\xi)$ and therefore $M(d)$.

For $\xi^- < \xi < \xi^+$, define $H(\xi)$ to be $F(\xi)$ if $\xi \leq \xi_0$ and to be $G(\xi)$ if $\xi > \xi_0$. The iteration of the map H generates the sequence of values of ξ that governs the recurrence for $T(c,d,\xi)$. Let H^* be the restriction of H to the interval $[G(\xi_0), F(\xi_0)]$ with the identification $F(\xi_0) = G(\xi_0)$. Then H^* is an orientation-preserving homeomorphism of the circle. Let θ be the rotation number of H^*; θ is the average number of cycles per step in the iteration of H^*. Since H^* has no fixed point, $\theta > 0$. Since a cycle must contain at least one application of $F(\xi)$ and one of $G(\xi)$, $\theta \leq 1/2$. If $\rho \leq \rho_2$, there is exactly one computation step per cycle; thus $C_\rho = 1/\theta$. If $\rho \geq \rho_2$, there is exactly one correction step per cycle; thus $C_\rho = 1/(1-\theta)$.

The foregoing analysis describes the factor C_ρ in terms of the rotation number θ of a certain homeomorphism of the circle. Some further analysis yields the following asymptotic formulae:

$$C_\rho - 1 \sim \frac{2\log 2}{\log \frac{1}{(1-\rho)}},$$

as $\rho \to 1$, and

$$C_\rho \sim \frac{\left(\log \frac{7+\sqrt{13}}{6}\right)\left(\log \frac{1}{(\rho-\rho_1)}\right)}{\left(\log \frac{11-\sqrt{13}}{6}\right)\left(\log \frac{5+\sqrt{13}}{6}\right)},$$

as $\rho \to \rho_1$.

3. A lower bound. I conjecture that the method described above is essentially optimal, in the sense that reliable computation is impossible if $\rho \leq \rho_1$ and takes C_ρ times as long if $\rho > \rho_1$. I have only succeeded in proving, however, that it is impossible if $\rho \leq 2/3$ and takes $1/(1+\log_3(2\rho-1))$ times as long if $\rho > 2/3$. We shall continue to confine our attention to formulae over the basis {minor}. An advantage of the argument we shall present is that it applies to formulae over any basis, with $2/3$ replaced by $(k+1)/2k$ and $\log_3(2\rho-1)$ replaced by $\log_k(2\rho-1)$, where k is the largest of the degrees of the functions in the basis. The corresponding disadvantage is that it is unable to predict the threshold ρ_1 and the factor C_ρ, which undoubtedly depend on the particular functions present in the basis, and not merely on their degrees.

Let us say that f is a *subfunction* of g if f can be obtained from g by evaluation (substituting constants for indeterminates). Let $d \geq 2$ be even, let $n = 3^d$, and

let f_d denote a function of degree n such that $D(f_d) = d$ and all n projections are subfunctions of f_d.

THEOREM 3.1. *Suppose that $\sigma > 1/2$ and N is a formula on $x_1, \ldots, x_n$ that (ρ, σ, σ)-computes f_d. Then $\rho > 2/3$ and $D(N) \geq d/(1 + \log_3(2\rho - 1))$.*

PROOF. For each input M of N, let $\Delta(M)$ denote the number of gates on the unique path from M to the root of N. Let $\Phi(\xi)$ denote the sum of $\xi^{\Delta(M)}$ over all inputs M of N.

We shall prove below that

$$\Phi(1/3) \leq 1/3 \tag{3.1}$$

and

$$\Phi(2\rho - 1) \geq n. \tag{3.2}$$

For now, let us see how these inequalities imply the theorem. Suppose first that $\rho > 2/3$. Let $r = 1 + \log_3(2\rho - 1)$, so that $0 < r \leq 1$, and recall Hölder's inequality:

$$\sum_M a_M^{1-r} b_M^r \leq \left(\sum_M a_M\right)^{1-r} \left(\sum_M b_M\right)^r.$$

By (3.2), Hölder's inequality (with $a_M = 1/3^{\Delta(M)}$ and $b_M = 1$) and (3.1) we have

$$n \leq \Phi(2\rho - 1) \leq \Phi(1/3)^{1-r} \Phi(1)^r \leq \Phi(1)^r.$$

Since $\Phi(1)$ is the number of inputs of N, and is thus at most $3^{D(N)}$, and since $n = 3^d$, taking logarithms yields $D(N) \geq d/r$, as claimed. Since this lower bound diverges as $\rho \to 2/3$, we conclude that $\rho > 2/3$ is necessary as well. □

It remains to prove (3.1) and (3.2). To do this we shall write $\Phi_N(\xi)$ rather than $\Phi(\xi)$, to indicate the dependence on the formula N. If M is a source, then $\Phi_M(\xi) = 0$. If M is an input, then $\Phi_M(\xi) = 1$. If $M = \text{minor}\{M_1, M_2, M_3\}$, then $\Phi_M(\xi) = \xi(\Phi_{M_1}(\xi) + \Phi_{M_2}(\xi) + \Phi_{M_3}(\xi))$. Inequality (3.1) now follows immediately by induction on the structure of N.

To prove (3.2), let $\Phi^{(m)}(\xi)$ denote the sum of $\xi^{\Delta(M)}$ over all inputs M in N that compute the projection x_m. Since $\Phi(\xi) = \sum_{1 \leq m \leq n} \Phi^{(m)}(\xi)$, it will suffice to prove that

$$\Phi^{(m)}(2\rho - 1) \geq 1$$

for all $1 \leq m \leq n$. Since f_d contains all n projections as subfunctions, we can substitute sources for inputs in N to obtain a formula $N^{(m)}$ that (ρ, σ, σ)-computes the projection x_m and such that $\Phi^{(m)}(\xi) = \Phi_{N^{(m)}}(\xi)$. Thus it will suffice to show that if N is a formula on x that (ρ, σ, σ)-computes the projection x, then

$$\Phi_N(2\rho - 1) \geq 1. \tag{3.3}$$

Let $K = 1 + \sigma \log_2 \sigma + (1 - \sigma) \log_2(1 - \sigma)$. Since $\sigma < 1/2$, $K > 0$. With each subformula M of N we shall associate a number Ψ_M with the following properties. If M is a source, then

$$\Psi_M = 0. \tag{3.4}$$

If M is an input computing the projection x, then

$$\Psi_M = K. \tag{3.5}$$

If $M = \text{minor}\{M_1, M_2, M_3\}$, then

$$\Psi_M \leq (2\rho - 1)(\Psi_{M_1} + \Psi_{M_2} + \Psi_{M_3}). \tag{3.6}$$

These properties imply

$$\Psi_N \leq K\Phi_N(2\rho - 1),$$

by induction on the structure of N. We shall also prove that if N is a formula on x that (ρ, σ, σ)-computes the projection x, then

$$\Psi_N \geq K. \tag{3.7}$$

This will complete the proof of (3.3).

To define Ψ_M with the desired properties, we shall use Shannon's information theory. If X is a random variable assuming t distinct values with probabilities $p_1, \ldots, p_t$, define the *entropy* $H(X)$ by

$$H(X) = -\sum_{1 \leq s \leq t} p_s \log_2 p_s.$$

If X and Y are jointly distributed random variables, we shall write $H(X, Y)$ for $H((X, Y))$. The entropy satisfies the following properties: (A) $H(X) \geq 0$, and $H(X) = 0$ if and only if X is constant with probability one; (B) $H(X, Y) \geq H(X)$; and (C) $H(X, Y, Z) + H(X) \leq H(X, Y) + H(X, Z)$. These properties are immediate consequences of the fact that the logarithm is increasing, concave, and vanishes at unity. Define the *mutual information* $I(X; Y)$ by $I(X; Y) = H(X) + H(Y) - H(X, Y)$.

Let X be a random variable assuming values 0 and 1 with equal probability. Let N be a formula on x that (ρ, σ, σ)-computes the projection x, and let the random variable Y_M assume the value produced by the subformula M of N when x is assigned the value X, inputs have reliability σ (independently of X and of each other), and gates have reliability ρ (independently of X, of the inputs and of each other). Set $\Psi_M = I(X; Y_M)$.

With this definition, it is straightforward to verify properties (3.4)—(3.7); the proof of (3.6) is best broken into three parts: if $M = \text{minor}\{M_1, M_2, M_3\}$, then

$$I(X; (Y_{M_1}, Y_{M_2}, Y_{M_3})) \leq I(X; Y_{M_1}) + I(X; Y_{M_2}) + I(X; Y_{M_3});$$

$$I(X; \text{minor}\{Y_{M_1}, Y_{M_2}, Y_{M_3}\}) \leq I(X; (Y_{M_1}, Y_{M_2}, Y_{M_3}));$$

and

$$I(X; Y_M) \leq (2\rho - 1) I(X; \text{minor}\{Y_{M_1}, Y_{M_2}, Y_{M_3}\}).$$

These inequalities, and the other properties of Ψ_M, are easy consequences of (A), (B), and (C).

REFERENCES

[DO1] R. L. Dobrushin and S. I. Ortyukov, *Upper bound for the redundancy of self-correcting arrangements of unreliable functional elements*, Problems Inform. Transmission **13** (1977), 203–218.

[DO2] ——, *Lower bound for the redundancy of self-correcting arrangements of unreliable functional elements*, Problems Inform. Transmission **13** (1977), 59–65.

[G] P. Gács, *Reliable computation with cellular automata*, J. Comput. System. Sci. **32** (1986), 15–78.

[GR] P. Gács and J. H. Reif, *A simple three-dimensional real-time reliable cellular array*, ACM Sympos. Theory of Comp. **17** (1985), 288–395.

[L] O. B. Lupanov, *Ob Odnom Metode Sinteza Skhem*, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **1** (1958), 120–140.

[M] D. E. Muller, *Complexity in electronic switching circuits*, IRE Trans. Electr. Comp. **5** (1956), 15–19.

[N] J. von Neumann, *Probabilistic logics and the synthesis of reliable organisms from unreliable components*, Automata Studies, C. E. Shannon and J. McCarthy, editors, Princeton Univ. Press, Princeton, N.J., 1956, pp. 43–98.

[P] N. Pippenger, *On networks of noisy gates*, IEEE Sympos. Found. of Comp. Sci. **26** (1985), 30–38.

IBM ALMADEN RESEARCH LABORATORY, SAN JOSE, CALIFORNIA 95120, USA