

Claremont Colleges Scholarship @ Claremont

All HMC Faculty Publications and Research

HMC Faculty Scholarship

1-1-1991

On a Lower Bound for the Redundancy of Reliable Networks with Noisy Gates

Nicholas Pippenger
Harvey Mudd College

Recommended Citation

Pippenger, N.; Stamoulis, G.D.; Tsitsiklis, J.N., "On a lower bound for the redundancy of reliable networks with noisy gates," *Information Theory, IEEE Transactions on*, vol.37, no.3, pp.639,643, May 1991. doi: 10.1109/18.79921

This Article is brought to you for free and open access by the HMC Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in All HMC Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

than the new bound [1] for the case of a channel with finite impulse response and with equalization over the full response length.

The notations and definitions of [1] are used here. The attention there is restricted to the case where the Z_k satisfy [1, (2.1)],

$$Z_k = a_k + \sum_{l=-M}^{-1} h_l e_{k+l} + \sum_{l=-\infty}^{-M-1} h_l a_{k+l} + n_k. \quad (1)$$

We note that this case (1) does not include precursors. The bound of [1], P_K , for general PAM systems where a_k can assume one of $2m$ values is given by

$$P_e \leq \left[\frac{1 - \left(\frac{1}{2m}\right)^M}{\left(1 - \frac{1}{2m}\right) \left(\frac{1}{2m}\right)^M} - (M-1) \right] \cdot \Pr\{Q(y_k) \neq a_k\} = P_K. \quad (2)$$

We consider an important subset of the set of channels modelled by (1); those having nonzero response only for a finite time duration with equalization over the full response length [2], [3]–[5]. Then the Z_k satisfy

$$Z_k = a_k + \sum_{l=-M}^{-1} h_l e_{k+l} + n_k. \quad (3)$$

To obtain an upper bound using the method of [2], one starts with the following expression for the probability of error [2, (14)–(16)]

$$P_e = \left[1 + \sum_{i=0}^{M-2} \prod_{l=0}^i \alpha_l + (1 - \alpha_M)^{-1} \prod_{l=0}^{M-1} \alpha_l \right]^{-1}. \quad (4)$$

One has that [1], [3]

$$\alpha_l \geq \frac{1}{2m}, \quad l = 0, \dots, M-1, \quad (5a)$$

and

$$\alpha_M = \Pr\{e_k = 0 | e_{k-1} = 0, \dots, e_{k-M} = 0\}. \quad (5b)$$

Note that $\epsilon = 1 - \alpha_M = \Pr\{Q(y_k) \neq a_k\}$ when the noise samples n_k are independent and (3) is satisfied. Using (5) with (4) yields the bound, P_{DMM} , given by

$$P_e \leq \frac{1 - \frac{1}{2m}}{1 - \left(\frac{1}{2m}\right)^M \left[1 - \frac{1}{\epsilon} \left(1 - \frac{1}{2m}\right) \right]} = P_{DMM}. \quad (6)$$

We will show that $P_{DMM} < P_K$ whenever (3) is satisfied and $\epsilon \neq 0$. Note that $\epsilon = 0$ implies that $P_e = 0$, $P_K = 0$ and $P_{DMM} = 0$.

One has from (2) and (6) that

$$P_K = \left[\frac{(2m)^{M+1} - 2m}{2m - 1} - (M-1) \right] \epsilon \quad (7)$$

and

$$P_{DMM} = \frac{\epsilon(2m)^M(2m-1)}{2m\epsilon[(2m)^M - 1] + 2m - 1} < \epsilon(2m)^M, \quad (8)$$

since $M \geq 1$, $m \geq 1$, and $\epsilon > 0$. Let $f(M) = P_K - \epsilon(2m)^M$ where M is interpreted as continuous and real and $M \geq 1$. Then

$P_{DMM} < P_K$ if $f(M) \geq 0$ for all $M \geq 1$. Differentiation gives

$$f'(M) = \epsilon \left[(2m)^M \left[\frac{1}{2m-1} \right] \ln(2m) - 1 \right] > \epsilon [\ln(2m) - 1] > 0, \quad (9)$$

since $\epsilon > 0$ and $\ln(\cdot)$ is increasing. Since $f(M)$ is increasing for $M \geq 1$, $f(M) \geq f(1) = 0$ for $M \geq 1$. Thus, $P_{DMM} < P_K$ is established.

REFERENCES

- [1] P. Kabaila, "On the probability of error for decision-feedback equalizers," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 5, pp. 735–738, Sept. 1977.
- [2] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 4, p. 490–497, July 1974.
- [3] A. Cantoni and P. Butler, "Stability of decision-feedback inverses," *IEEE Trans. Commun.*, vol. COM-24, no. 9, pp. 970–977, Sept. 1976.
- [4] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Trans. Commun.*, vol. COM-35, no. 10, pp. 1022–1028, Oct. 1987.
- [5] P. Mosen, "Adaptive equalization of the slow fading channel," *IEEE Trans. Commun.*, vol. COM-22, no. 8, pp. 1064–1075, Aug. 1974.

On a Lower Bound for the Redundancy of Reliable Networks with Noisy Gates

Nicholas Pippenger, George D. Stamoulis,
and John N. Tsitsiklis

Abstract—A proof is provided that a logarithmic redundancy factor is necessary for the reliable computation of the parity function by means of a network with noisy gates. This is the same as the main result of Dobrushin and Ortyukov except that the analysis therein seems to be not entirely correct.

Index Terms—Reliable computation, noisy gates, parity function.

I. INTRODUCTION

Computation of Boolean functions by means of noisy gates is a topic that started attracting the attention of researchers in the early '50s. The first related work was that of von Neumann [4] in 1952. The problem defined there is as follows: Suppose that the gates available for the computation of a Boolean function are *not completely reliable*; in particular, each one of them *fails* with probability $\epsilon < 1/2$, independently of the other gates. Given the values of its input bits, a gate is said to "fail" if it produces the complement of the output bit that it would have produced if it were completely reliable and its *inputs were the same*. Is it

Manuscript received February 21, 1990; revised September 10, 1990. N. Pippenger was supported by the NSERC under Grant OGP-0041640. G. D. Stamoulis and J. N. Tsitsiklis were supported by the NSF under Grant ECS-8552419 with matching funds from Bellcore, Inc. and DuPont, and by the ARO under Grant DAAL03-86-K-0171.

N. Pippenger is with the Department of Computer Science, University of British Columbia, Vancouver, BC, V6T 1W5, Canada.

G. D. Stamoulis and J. N. Tsitsiklis are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9041808.

possible under these assumptions to build, for any given function f , a network that computes $f(x)$ correctly with *high probability* for every input vector x ?

Von Neumann proved in [4] that computations may be done reliably for all sufficiently small $\epsilon > 0$. In his construction, each intermediate result is computed several times and its value is determined by majority voting. One then obtains a probability of error $\eta(\epsilon)$ for the final result, where $\eta(\epsilon) < 1/2$ for all sufficiently small $\epsilon > 0$. Unfortunately, this procedure for constructing reliable networks results in an unacceptably *large* number of gates.

Almost 25 years after von Neumann introduced the problem, Dobrushin and Ortyukov [1] claimed that there are cases where a considerable increase in complexity is *necessary* for reliable computation. Indeed, let $L(f)$ be the number of gates of a *minimal* noise-free network that computes some Boolean function f ; these authors stated the following result: There exists some function f^* (namely, the parity function) such that any network that computes f^* with probability of error $p < 1/3$ must contain $\Omega(L(f^*) \ln L(f^*))$ gates; i.e., the *order of magnitude* of the number of gates in any such reliable network is at least $L(f^*) \ln L(f^*)$. Thus, reliable computation of f^* requires at least *logarithmic redundancy*. The proof of this claim in [1] contains two questionable arguments; moreover, there does not seem to be any obvious modification that could result in a correct proof. In this correspondence, we present a new proof of this result. Our analysis follows steps similar to those in [1]; however, our approach to the questionable points in [1] is completely different. Moreover, our proof extends the validity of the claim in [1] to all $p \in (0, 1/2)$, which is the broadest acceptable range for the probability of error.

It is worth noting that for all Boolean functions there exist reliable networks with logarithmic redundancy; this result was proved by Dobrushin and Ortyukov in [2]. Moreover, as was proved by Pippenger [5], a rather broad class of Boolean functions may be computed reliably by networks that involve only constant redundancy. Thus, the logarithmic lower bound for the redundancy factor is *tight* only in the *worst* case.

To the best of our knowledge, our proof is the first correct one to be published. After writing this correspondence, it was communicated to us that Peter Gács had obtained another correct proof, which was never published.

The remainder of this correspondence is organized as follows: In Section II, we present an outline of the analysis in [1] and we state a result that implies the logarithmic lower bound on the redundancy factor. In Section III, we give our proof of this auxiliary result. Finally, in Section IV, we present some concluding remarks.

II. AN OUTLINE OF THE ANALYSIS IN [1]

In this section, we use a notation similar to that of [1]. First, we give some of the definitions therein.

We consider a *finite and complete* basis Φ ; that is, a set of gates that allows for the computation of *any* Boolean function by means of a network consisting exclusively of *finitely* many such gates (e.g., the AND-NOT basis). The maximum fan-in (i.e., number of input wires) of the gates in Φ is denoted by $n(\Phi)$. All networks considered in the analysis are assumed to consist only of gates belonging to this basis Φ . In the presence of noise, the gates available are assumed to fail according to the model presented in Section I; the probability ϵ of failure is taken to be fixed. Let f be a Boolean function and \mathcal{N} be a network over Φ .

Moreover, let $\xi(x, \epsilon)$ be the output of \mathcal{N} , where x is some assignment of the values of the input bits of \mathcal{N} ; of course, $\xi(x, \epsilon)$ is a random variable. The network \mathcal{N} is said to compute the function f with probability of error p if the following holds:

$$\Pr[\xi(x, \epsilon) \neq f(x)] \leq p, \quad \text{for all } x; \quad (1)$$

$p \in (0, 1/2)$ is a given scalar. Let $L_{p,\epsilon}(f, \Phi)$ be the *minimum* number of gates in a reliable network that computes the function f in such a way that (1) is satisfied. Similarly, $L_{0,0}(f, \Phi)$ denotes the number of gates in the *minimal* network that computes f in the absence of noise.

The *redundancy factor* $R_{p,\epsilon}(N, \Phi)$ for the basis Φ is defined as follows:

$$R_{p,\epsilon}(N, \Phi) = \max_{(f: L_{0,0}(f, \Phi) = N)} \frac{L_{p,\epsilon}(f, \Phi)}{L_{0,0}(f, \Phi)},$$

i.e., it equals the maximum of the required redundancy factor over all functions f that are computable in the absence of noise with the same minimum complexity N . The main result in [1] is given in Theorem 2.1 of that article; we repeat it next, in simplified notation.

Proposition 1: For any $p \in (0, 1/2)$, the redundancy factor $R_{p,\epsilon}(N, \Phi)$ is $\Omega(\ln N)$; that is, there exists some function $h(N)$ such that $R_{p,\epsilon}(N, \Phi) \geq h(N)$ and $\lim_{N \rightarrow \infty} h(N)/\ln N = h^* > 0$.

The expression for the function $h(N)$ mentioned in Proposition 1 is of no particular importance; what is important is that $h(N)$ is asymptotically linear in $\ln N$. Henceforth, we mainly focus on arguments involving orders of magnitude rather than giving detailed expressions.

Proposition 1 may be established by proving that some specific function f^* satisfies

$$L_{p,\epsilon}(f^*, \Phi) = \Omega(L_{0,0}(f^*, \Phi) \cdot \ln(L_{0,0}(f^*, \Phi))). \quad (2)$$

In particular, the authors of [1] considered the parity function $f^*(x) = x_1 \oplus \cdots \oplus x_n$, i.e., the sum modulo 2 of x_1, \dots, x_n . (Note that \oplus is the symbol for the XOR operation.) The choice of this function makes intuitive sense, because, when the value of one of the x_i 's is reversed, the value of $f^*(x)$ changes; in some sense, $f^*(x)$ is a "sensitive" function.

For this sensitivity of the parity function f^* to be exploited, a new model for noise is introduced in [1]. Under the new model, each of the *wires* fails with probability δ , independently of all other wires and gates; failure of a wire results in transmission of the complement of the input bit-signal. Consider now some gate that receives j binary input bits τ_1, \dots, τ_j and computes the function $\phi(\tau)$. Due to failures of the input wires, the vector $\tau = (\tau_1, \dots, \tau_j)$ may be different than the vector $t = (t_1, \dots, t_j)$ of the bits that the gate should have received. Moreover, given the distorted input vector τ , the gate may not produce $\phi(\tau)$; this is assumed to occur with probability $P(\tau, \delta)$, independently of all other gates. However, since the output of the gate in the absence of noise would have been $\phi(t)$, the gate is considered to fail if it does not produce $\phi(t)$. It is established in Lemma 3.1 of [1] that, given some $\delta \in [0, \epsilon/j]$, there exists a unique vector of malfunction probabilities $(P(\tau, \delta))_{\tau \in \{0,1\}^j}$ such that the overall probability that the gate does not produce $\phi(t)$ is equal to ϵ (for all t), as was the case in the original model. Though technically complicated, the underlying idea is clear: failures of gates may be visualized as not caused only by noisy computation, but also by noisy reception of the inputs. The parameters of this new model for noise can be selected in such a way that each gate still fails with probability ϵ . In this case, the state-vector of the

network has the *same* statistical properties as originally, which is intuitively clear. This is established in Lemma 3.2 of [1], by using induction on the depth of the network; this result holds for all $\delta \in [0, \epsilon/n(\Phi)]$. Thus, as far as reliability is concerned, the two types of networks are *equivalent*. On the other hand, under the new model for failures, wires also are unreliable, which suggests that the number of wires plays a key role in reliability; this was not that clear under the original model for noise. Since the parity function f^* is the most sensitive in the noisy transmission of inputs, it is expected that the redundancy involved in its reliable computation is of the worst possible order of magnitude.

So far, we have discussed the preliminary part of the analysis in [1], where the original problem was transformed into an equivalent one. Henceforth, we are only dealing with the newly introduced problem.

It is well known that, in the absence of noise, f^* may be calculated by using a tree of XOR gates. Thus, if the basis Φ includes the gate for $x_1 \oplus x_2$, then we have $L_{0,0}(f^*, \Phi) \leq n-1$; if not, then we have $L_{0,0}(f^*, \Phi) \leq C(\Phi)(n-1)$, where $C(\Phi)$ is the complexity of the noise-free network over Φ that computes $x_1 \oplus x_2$. (Notice that $C(\Phi)$ is finite because Φ is a finite and complete basis.) On the other hand, it is straightforward that $L_{0,0}(f^*, \Phi) \geq n/n(\Phi)$. Therefore, proving (2) is equivalent to proving that $L_{p,\epsilon}(f^*, \Phi)$ is $\Omega(n \cdot \ln n)$. (Recall that n is the number of input bits.)

We consider a reliable *minimal complexity* noisy network \mathcal{N} for the function f^* . We denote by m_i the number of wires of \mathcal{N} over which the input bit x_i is transmitted, for $i=1, \dots, n$. Thus, \mathcal{N} has at least $\sum_{i=1}^n m_i$ wires, which implies that

$$L_{p,\epsilon}(f^*, \Phi) \geq \frac{\sum_{i=1}^n m_i}{n(\Phi)}.$$

It follows from the above discussion that to prove (2) (which implies Proposition 1), it suffices to prove the following result.

Proposition 2: The total number $\sum_{i=1}^n m_i$ of input wires in any reliable network that computes f^* with probability of error p is $\Omega(n \cdot \ln n)$ for all $p \in (0, 1/2)$.

In [1], this result is dealt with in Theorem 3.1 and in its auxiliary Lemma 3.3. This part of the analysis in [1] seems not to be correct; we comment on this in the Appendix. In the next section, we present our proof of Proposition 2.

It is worth noting that Theorem 3.1 of [1] would hold for several Boolean functions that are "sensitive" under some particular assignment of the input bits (e.g., the AND function, which is "sensitive" for $x = (1, \dots, 1)$). On the contrary, Proposition 2 holds only for the parity function.

III. PROOF OF PROPOSITION 2

We fix some $p \in (0, 1/2)$. Moreover, we fix some $\delta \in (0, \epsilon/n(\Phi))$; note that such a δ satisfies $\delta < \epsilon < 1/2$. Henceforth, we assume that the input bits X_1, \dots, X_n are independent *random variables* and that $\Pr[X_i = 0] = 1/2$ for $i=1, \dots, n$. We use the notation (x_1, \dots, x_n) to denote some particular *value* of the random vector (X_1, \dots, X_n) . Under this assumption, we shall prove that the *average* (over all possible inputs) probability of an erroneous output for the noisy network for the parity function f^* must be greater than p unless $\sum_{i=1}^n m_i$ is $\Omega(n \cdot \ln n)$. This implies that if $\sum_{i=1}^n m_i$ is not $\Omega(n \cdot \ln n)$, then there exists at least one input assignment for which the probability of an

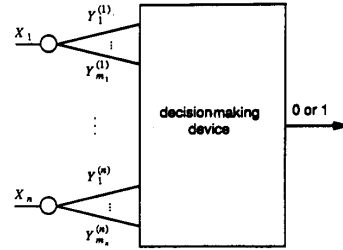


Fig. 1.

erroneous output exceeds p ; this statement is equivalent to Proposition 2.

After introducing the assumption of equally likely input assignments, any noisy network for f^* may be visualized as a device for estimating the binary parameter $f^*(X) = X_1 \oplus \dots \oplus X_n$. The decision is to be based on the values of the signals communicated by the input wires. Notice that such a decision-making device employs randomization due to the presence of noise. We denote by \vec{Y} the random vector $(Y^{(1)}, \dots, Y^{(n)})$, where $Y^{(i)} = (Y_1^{(i)}, \dots, Y_{m_i}^{(i)})$ is the vector of binary random variables corresponding to the output signals of the input wires for X_i (see Fig. 1). The *value* $y^{(i)} = (y_1^{(i)}, \dots, y_{m_i}^{(i)})$ of $Y^{(i)}$ is a vector of *distorted* copies of the i th input bit X_i , for $i=1, \dots, n$. Thus, the data on which estimation is based is contained in the vector $\vec{Y} = (Y^{(1)}, \dots, Y^{(n)})$. Clearly, we have

$$\Pr[f^*(X) = 0] = \Pr[f^*(X) = 1] = 1/2.$$

Therefore, the decisionmaking device that has the minimum average probability of error is the one based on the *maximum likelihood* (ML) test. Hence, to prove Proposition 2, it suffices to prove the following result.

Proposition 3: If the average probability of error for the device based on the maximum likelihood rule does not exceed p , then $\sum_{i=1}^n m_i$ is $\Omega(n \cdot \ln n)$.

Proof: We fix some index $i \in \{1, \dots, n\}$. We denote by w_i the number of entries of the observed vector $y^{(i)}$ that equal 0. Recalling that wires fail with probability δ and independently of each other, it follows that the ML rule for estimating X_i is equivalent to the majority-voting test:

$$\hat{X}_i = \begin{cases} 0 & \text{if } w_i \geq \frac{m_i}{2} \\ 1 & \text{otherwise} \end{cases}$$

If $w_i = m_i/2$, then the tie may be broken arbitrarily.

We denote by Z_i the Boolean random variable indicating whether the ML estimate of X_i is correct or not; that is, we have $Z_i = 1$ if and only if $\hat{X}_i \neq X_i$. If all copies of the input bit X_i are communicated erroneously, then we have $Z_i = 1$; this implies that

$$\Pr[Z_i = 1] \geq \delta^{m_i}. \quad (3)$$

Because wires fail independently of each other and because different input bits are independent, the random vectors $Y^{(1)}, \dots, Y^{(n)}$ are independent conditioned on any given X . The parameter to be estimated is the parity among the input bits; thus, it is intuitively clear that the ML estimate of $f^*(X)$ should be equal to $\hat{X}_1 \oplus \dots \oplus \hat{X}_n$, where \hat{X}_i is the ML estimate of X_i . This is proved formally in Lemma 5; first, we state the following auxiliary result (also see [3]), whose proof we include for completeness. \square

Lemma 4: There holds

$$1 - 2\Pr[Z_1 \oplus \cdots \oplus Z_n = 1] = \prod_{i=1}^n (1 - 2\Pr[Z_i = 1]).$$

Proof: We denote by $\phi_i(\cdot)$ the probability generating function of the Boolean random variable Z_i for $i = 1, \dots, n$. We have $\phi_i(t) = E[t^{Z_i}] = 1 - \Pr[Z_i = 1] + t\Pr[Z_i = 1]$. Since Z_1, \dots, Z_n are independent, the probability generating function $\varphi(\cdot)$ of the random variable $\sum_{i=1}^n Z_i$ can be expressed in the following product form

$$\varphi(t) = \prod_{i=1}^n \phi_i(t). \quad (4)$$

(Note that $\sum_{i=1}^n Z_i$ denotes the ordinary sum of Z_1, \dots, Z_n ; this should not be confused with the modulo 2 sum $Z_1 \oplus \cdots \oplus Z_n$.) Clearly, we have

$$\frac{\varphi(1) - \varphi(-1)}{2} = \sum_{\text{odd } k} \Pr\left[\sum_{i=1}^n Z_i = k\right] = \Pr[Z_1 \oplus \cdots \oplus Z_n = 1].$$

This together with (4) and the fact that $\varphi(1) = 1$ establishes the lemma. \square

Next, we prove the result on the ML estimate of $f^*(X)$.

Lemma 5: The ML estimate of $X_1 \oplus \cdots \oplus X_n$ is $\hat{X}_1 \oplus \cdots \oplus \hat{X}_n$, where \hat{X}_i is the ML estimate of X_i .

Proof: Let V be some Boolean random variable with $\Pr[V = 0] = \Pr[V = 1] = 1/2$. Assume that V is to be estimated based on the observation of some data vector \vec{U} . Then, the Boolean random variable \hat{V} is the ML estimate of V given \vec{U} if and only if the following is true:

$$\Pr[\hat{V} \neq V | \vec{U} = \vec{u}] \leq \frac{1}{2}, \quad \forall \vec{u}. \quad (5)$$

Thus, in order to prove the lemma, it suffices to show that

$$\Pr[\hat{X}_1 \oplus \cdots \oplus \hat{X}_n \neq X_1 \oplus \cdots \oplus X_n | \vec{Y} = \vec{y}] \leq \frac{1}{2}, \quad \forall \vec{y}. \quad (6)$$

(Recall that $X_1 \oplus \cdots \oplus X_n$ takes the values 0 and 1 with probability 1/2.)

Since Z_i takes the value 1 if and only if $\hat{X}_i \neq X_i$, we have $\hat{X}_1 \oplus \cdots \oplus \hat{X}_n \neq X_1 \oplus \cdots \oplus X_n$ if and only if $Z_1 \oplus \cdots \oplus Z_n = 1$, that is, if and only if an odd number of the input bits are estimated erroneously. Therefore, we have

$$\begin{aligned} \Pr[\hat{X}_1 \oplus \cdots \oplus \hat{X}_n \neq X_1 \oplus \cdots \oplus X_n | \vec{Y} = \vec{y}] \\ = \Pr[Z_1 \oplus \cdots \oplus Z_n = 1 | \vec{Y} = \vec{y}]. \end{aligned} \quad (7)$$

Reasoning similarly as in proving Lemma 4, we obtain

$$\begin{aligned} 1 - 2\Pr[Z_1 \oplus \cdots \oplus Z_n = 1 | \vec{Y} = \vec{y}] \\ = \prod_{i=1}^n (1 - 2\Pr[Z_i = 1 | \vec{Y} = \vec{y}]). \end{aligned} \quad (8)$$

Since \hat{X}_i is the ML estimate of X_i given \vec{Y} , it follows from (5) that

$$\Pr[Z_i = 1 | \vec{Y} = \vec{y}] = \Pr[\hat{X}_i \neq X_i | \vec{Y} = \vec{y}] \leq \frac{1}{2}, \quad \forall \vec{y}.$$

Combining this with (8), we obtain

$$\Pr[Z_1 \oplus \cdots \oplus Z_n = 1 | \vec{Y} = \vec{y}] \leq \frac{1}{2}, \quad \forall \vec{y}.$$

This together with (7) proves (6). \square

We now complete the proof of Proposition 3.

We have already argued that the event $Z_1 \oplus \cdots \oplus Z_n = 1$ coincides with the event that the ML estimate of $f^*(X)$ is erroneous. Thus, we have

$$\Pr[Z_1 \oplus \cdots \oplus Z_n = 1] \leq p$$

by assumption. This together with (3) and Lemma 4 implies that

$$1 - 2p \leq \prod_{i=1}^n (1 - 2\delta^{m_i}).$$

Using the inequality between arithmetic and geometric means, we obtain

$$1 - 2p \leq \left(\frac{1}{n} \sum_{i=1}^n (1 - 2\delta^{m_i})\right)^n = \left(1 - \frac{2}{n} \sum_{i=1}^n \delta^{m_i}\right)^n.$$

Again, using the inequality between arithmetic and geometric means, followed by the inequality $1 - \alpha \leq \exp(-\alpha)$, we obtain

$$1 - 2p \leq (1 - 2\delta^{1/n \sum_{i=1}^n m_i})^n \leq \exp(-2n\delta^{1/n \sum_{i=1}^n m_i}),$$

which implies that

$$\sum_{i=1}^n m_i \geq n \frac{\ln(2n) - \ln \ln(1/(1-2p))}{\ln(1/\delta)}.$$

This result holds for all $\delta \in (0, \epsilon/n(\Phi)]$; taking $\delta = \epsilon/n(\Phi)$ we obtain

$$\sum_{i=1}^n m_i \geq n \frac{\ln(2n) - \ln \ln(1/(1-2p))}{\ln(n(\Phi)/\epsilon)};$$

clearly, this proves that $\sum_{i=1}^n m_i$ is $\Omega(n \cdot \ln n)$. \square

IV. CONCLUSION

In this correspondence we have proved a lower bound for the redundancy involved in constructing reliable networks by means of noisy gates. In particular, we have established that a redundancy factor logarithmic in n is necessary for reliable computation of the parity (i.e., the sum modulo 2) of n bits. This result was first stated by Dobrushin and Ortyukov in [1]. As we have argued in the Appendix, we believe that the proof given in [1] is not entirely correct. We have established the result by following the same steps as Dobrushin and Ortyukov and by replacing the questionable part of their analysis with entirely new arguments. Nevertheless, formulating the lower bound problem and introducing a suitable problem transformation has proved to be a valuable contribution of [1].

APPENDIX

In this appendix we discuss the proofs of Theorem 3.1 and of Lemma 3.3 in [1]. First we consider the latter.

Lemma 3.3 of [1]: Let $p \in (0, 1/3)$ and $\delta \in (0, 1/2)$. Moreover, let $(H_i)_{i \in Q}$ be independent events satisfying the following:

$$\Pr[H_i] \geq \delta^{m_i}, \quad \forall i \in Q$$

and

$$p \geq (1-p) \Pr \left[\bigcup_{l \in Q} \tilde{H}_l \right],$$

where $\bigcup_{l \in Q} \tilde{H}_l$ is the event that exactly one of the events $(H_l)_{l \in Q}$ has occurred. Then,

$$\sum_{l \in Q} m_l \geq \frac{|Q|}{\ln(1/\delta)} \ln \left(\frac{|Q|(1-3p)}{p} \right). \quad \square$$

This lemma seems to be incorrect. Indeed, consider the simple case where $\Pr[H_l] = \delta^m$ for all $l \in Q$. We fix some $m > 0$ and some $p \in (0, 1/3)$. Notice that

$$\Pr \left[\bigcup_{l \in Q} \tilde{H}_l \right] = |Q| \delta^m (1 - \delta^m)^{|Q|-1}.$$

Thus, according to the lemma, the inequality

$$p \geq (1-p) |Q| \delta^m (1 - \delta^m)^{|Q|-1} \quad (\text{A.1})$$

implies that

$$m \geq \frac{1}{\ln(1/\delta)} \ln \left(\frac{|Q|(1-3p)}{p} \right). \quad (\text{A.2})$$

However, this is seen to be false, because (A.1) holds for all sufficiently large $|Q|$ whereas (A.2) fails to hold for all sufficiently large $|Q|$. (Notice that the right-hand quantity in (A.1) tends to 0, as $|Q| \rightarrow \infty$, whereas the right-hand quantity in (A.2) tends to ∞ , as $|Q| \rightarrow \infty$.)

Lemma 3.3 is crucial for the proof of Theorem 3.1 of [1]. Thus, it does not seem that the proof of Theorem 3.1 can be fixed. Also note that at some point of that proof (namely, (3.30) of [1]), the authors seem to use the property that the inequalities $\Pr[\Gamma \Delta_1] \geq 1-p$ and $\Pr[\Gamma \Delta_2] \geq 1-p$ imply

$$\Pr \left[\Gamma \bigcup_{l=1,2} \tilde{\Delta}_l \right] \geq 1-p.$$

However, this property is not generally valid. Indeed, taking $\Gamma = \Delta_1 \cap \Delta_2$, we have

$$\Pr \left[\Gamma \bigcup_{l=1,2} \tilde{\Delta}_l \right] = 0.$$

These observations lead us to doubt that the analysis in [1] can be corrected by local modifications.

REFERENCES

- [1] R. L. Dobrushin and S. I. Ortyukov, "Lower bound for the redundancy of self-correcting arrangements of unreliable functional elements," *Prob. Inf. Trans.*, vol. 13, pp. 59-65, 1977.
- [2] —, "Upper bound for the redundancy of self-correcting arrangements of unreliable functional elements," *Prob. Inf. Trans.*, vol. 13, pp. 203-218, 1977.
- [3] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
- [4] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," in *Automata Studies*, C. E. Shannon and J. McCarthy, Eds. Princeton, NJ: Princeton University Press, 1956, pp. 329-378.
- [5] N. Pippenger, "On networks of noisy gates," in *Proc. 26th Ann. Symp. Foundations Comput. Sci.*, 1985, pp. 30-36.

Exponential Error Bounds for Random Codes on Gaussian Arbitrarily Varying Channels

Tony G. Thomas and Brian Hughes

Abstract—The Gaussian arbitrarily varying channel (GAVC) models a channel corrupted by thermal noise and by an unknown interfering signal of bounded power. The upper and lower bounds are presented to the best error probability achievable on this channel with random coding. The asymptotic exponents of these bounds agree in a range of rates near capacity. The exponents are *universally larger* than the corresponding exponents for the discrete-time Gaussian channel with the same capacity. It is further shown that the decoder can be taken to be the minimum Euclidean distance rule at all rates less than capacity.

Index Terms—Error exponents, arbitrarily varying channels, Gaussian channels, random codes.

I. INTRODUCTION

The Gaussian arbitrarily varying channel (GAVC), introduced in [1], models a communication channel corrupted by thermal noise and by an unknown and arbitrary interfering signal of bounded power. The channel is described as follows (see Fig. 1): Once in each unit of time, the transmitter sends a real number to the receiver, say x_i at time i , so that the sequence $\{x_i\}$ has power (defined in the next section) at most b . This number is received as $y_i = x_i + n_i + s_i$ where $\{n_i\}$ is an independent and identically distributed (i.i.d.), zero-mean Gaussian sequence with variance a , and $\{s_i\}$ is an unknown and arbitrary sequence with power at most c . The goal of the transmitter is to construct a communication system that is *robust* in the sense that is performed reliably for all $\{s_i\}$ that satisfy the power constraint.

In [1], Hughes and Narayan investigated the random coding capacity of the GAVC for a variety of power constraints on transmitter and interference. For peak time-averaged power constraints on both transmitter and interference, the authors showed that the GAVC has random coding capacity

$$C = \frac{1}{2} \ln \left(1 + \frac{b}{a+c} \right) \text{ nats/channel use.} \quad (1)$$

It is interesting to note that (1) is identical to the capacity formula for the discrete-time Gaussian channel that would be formed if $\{s_i\}$ were an i.i.d. sequence of zero-mean, Gaussian random variables with variance c . For ensemble-averaged constraints on transmitter or interference, the GAVC has no capacity in the usual sense; rather, the best achievable error probability depends continuously on the rate of transmission.

The main objective of this paper is to develop exponential bounds to the best error probability achievable with random coding on the GAVC in the one case where a (strong) capacity exists (i.e., with peak time-averaged power constraints on both the transmitter and interference). Stiglitz [2] has derived a random-coding exponential upper bound for the discrete arbitrarily varying channel (AVC) (with correlated random encoding

Manuscript received February 7, 1989; revised October 3, 1990. This work was supported by the National Science Foundation under Grant NSF-NCR-8804257 and by the Army Research Office under Grant DAAL03-89-K-0130.

The authors are with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218. IEEE Log Number 9142963.