

# Energy-Efficiency Bounds for Deep Submicron VLSI Systems in the Presence of Noise

Lei Wang, *Member, IEEE*, and Naresh R. Shanbhag, *Senior Member, IEEE*

**Abstract**—In this paper, we present an algorithm for computing the bounds on energy-efficiency of digital very large scale integration (VLSI) systems in the presence of deep submicron noise. The proposed algorithm is based on a *soft-decision channel* model of noisy VLSI systems and employs information-theoretic arguments. Bounds on energy-efficiency are computed for multimodule systems, static gates, dynamic circuits and noise-tolerant dynamic circuits in 0.25- $\mu\text{m}$  CMOS technology. As the complexity of the proposed algorithm grows linearly with the size of the system, it is suitable for computing the bounds on energy-efficiency for complex VLSI systems. A key result presented is that noise-tolerant dynamic circuits offer the best trade off between energy-efficiency and noise-immunity when compared to static and domino circuits. Furthermore, employing a 16-bit noise-tolerant Manchester adder in a CDMA receiver, we demonstrate a 31.2%–51.4% energy reduction over conventional systems when operating in the presence of noise. In addition, we compute the lower bounds on energy dissipation for this CDMA receiver and show that these lower bounds are  $2.8\times$  below the actual energy consumed, and that noise-tolerance reduces the gap between the lower bounds and actual energy dissipation by a factor of  $1.9\times$ .

**Index Terms**—CDMA communications, deep submicron noise, energy-efficiency bounds, low power, noise-tolerance.

## I. INTRODUCTION

THE ABILITY to scale CMOS technology [1]–[3] has been a key driver for the development of low-cost broadband communication and computing systems. However, with feature sizes being scaled into the deep submicron (DSM) regime, DSM noise [4], [5] consisting of ground bounce, crosstalk,  $IR$  drops, clock jitter, charge sharing, process variations, etc., has emerged as a critical factor that may ultimately determine the performance achievable in the future at an affordable cost. Compounding the problem further is the adoption of aggressive design practices such as dynamic logic, low-supply voltage, use of low- $V_t$  and hence, high-leakage devices, and high clock-frequencies. As a result, the semiconductor industry is facing a reliability problem that challenges the very foundation of the cost and performance benefits of very large scale integration (VLSI).

In this paper, we address the issue of determining achievable bounds on energy-efficiency for digital VLSI systems in

the presence of noise. The issue of lower bounds on power dissipation was addressed in [6] by considering thermal noise as the limiting factor for energy reduction. In [7], the lower bounds on power dissipation per pole for analog circuits and empirical lower bounds for digital circuits were derived from the desired signal-to-noise ratio (SNR) by assuming noise-free elements.

In the past, we have proposed an *information-theoretic framework* [8] that enables us to determine the bounds on energy-efficiency in a rigorous manner. The central thesis of this framework is to establish an energy-related correspondence between an algorithm and its architectural implementation. Specifically, we view a DSM VLSI system as a network of communication channels; a view also echoed recently in the ITRS2001 [1]. The algorithmic complexity of the system is quantified by the *information transfer rate*  $R$ , while each implementation is said to have an *information transfer capacity*  $C$ . Employing the information-theoretic constraint  $C > R$  on reliability [9], we presented [8] a common basis for various power reduction techniques such as voltage scaling, pipelining, parallel processing, adiabatic logic, etc. The same constraint was utilized in [10] to obtain the fundamental limit on signal energy transfer during a binary switching transition. In [11], we proposed a *binary symmetric channel* (BSC) model to determine the lower bounds on energy dissipation for single-output static modules. An information-theoretic approach was also employed in [12] to determine the maximum achievable energy-reduction in high-speed busses.

While the proposed information-theoretic approach enables the computation of energy-efficiency bounds, it does not directly enable the construction of low-power design techniques. Fortunately, the proposed approach does indicate that design techniques based on *noise-tolerance*, in contrast to *noise mitigation*, can enable the design of systems that operate at or close to their energy-efficiency bounds. Indeed, our past work [13]–[16] has shown that design techniques based on noise-tolerance at the algorithmic level [or algorithmic noise-tolerance (ANT)] [13], [14] and at the circuit level [15], [16] are effective in achieving energy-efficiency in the presence of noise. In the long run, we see algorithmic and circuit level noise-tolerance techniques being applied concurrently to design systems with energy consumption that approach these bounds. Indeed, *error-tolerance* has been referred to as one of the difficult design challenges in the next decade [1].

Ideally, we would like to be able to compute energy-efficiency bounds for complex systems and then employ a combination of algorithmic and circuit level noise-tolerance techniques to approach these bounds. Unfortunately, computing the bounds for complex systems is quite difficult at present. Thus, we are

Manuscript received August 14, 2001; revised February 8, 2002. This work was supported by the National Science Foundation under Grant CCR-0000987 and Grant CCR-9979381.

L. Wang is with Microprocessor Technology Laboratories, Hewlett-Packard Company, Fort Collins, CO 80521 USA (e-mail: lei.wang@onebox.com).

N. R. Shanbhag is with the Coordinated Science Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: shanbhag@uiuc.edu).

Digital Object Identifier 10.1109/TVLSI.2003.810783

faced with a “gap” between our ability to compute bounds on energy-efficiency for a specific system and our ability to devise design techniques for approaching these bounds. Fortunately, the *soft-decision channel* (SDC) model presented in this paper holds the promise that lower bounds on energy-efficiency can be computed for complex systems. An evidence of this promise is presented in this paper where we compute the bounds on energy-efficiency for multi-input multi-output systems and for a CDMA receiver. Specific examples where the above mentioned gap is bridged are the chip I/O signaling [11] and the CDMA receiver example in this paper.

Employing the proposed SDC model, we obtain the lower bounds on energy-efficiency for static, domino [19], and noise-tolerant domino [15] gates. We show that noise-tolerant gates have a lower bound that is smaller than the corresponding bound for a domino gate when operating in the presence of noise. This is an interesting result because noise-tolerance implies an energy overhead and thus it is not obvious that it is an energy-efficient design technique. Note that while the proposed SDC model enables us to compute the lower bounds on energy-efficiency, it does not indicate the design technique to employ. Nevertheless, this paper presents simple examples that do suggest that design techniques based on noise-tolerance can be quite effective in achieving energy-efficiency in the presence of noise.

In Section II, we briefly review our past work on the information-theoretic framework. In Section III, we propose the soft-decision channel model for noisy VLSI systems and develop the associated information-theoretic measures. In Section IV, we determine the lower bounds on energy-efficiency by solving an energy optimization problem. In Section V, we employ noise-tolerance to design an adder circuit that approaches the lower bounds on energy-efficiency in the presence of noise.

## II. INFORMATION-THEORETIC FRAMEWORK: A REVIEW

In this section, we review our past work on the information-theoretic framework for deriving the lower bounds on energy dissipation of noisy logic gates.

From Shannon’s joint source-channel coding theory [9], the information content of a continuous source  $\mathcal{X}$  is given by

$$H(X) = - \int_{\mathcal{X}} f_X(x) \log_2 f_X(x) dx \quad (1)$$

where  $f_X(x)$  is the probability density function (PDF) of a continuous-valued output  $X \in \mathcal{X}$ .

Assume that the output of the source  $\mathcal{X}$  is passed through a noisy transformation  $\mathcal{F}: \mathcal{X} \rightarrow \mathcal{Y}$ , i.e.,

$$Y = \mathcal{F}(X) + N \quad (2)$$

where  $\mathcal{F}$  is a deterministic mapping function from  $X \in \mathcal{X}$  to  $Y \in \mathcal{Y}$ , and  $N$  denotes the noise, which is typically assumed to be white with a Gaussian distribution but could also have an arbitrary distribution. The maximum information content that the noisy transformation  $\mathcal{F}$  can transfer with arbitrarily low probability of error is given by its capacity as

$$C = C_u f_c = \max_{\forall f_X(x)} I(X; Y) f_c \quad (3)$$

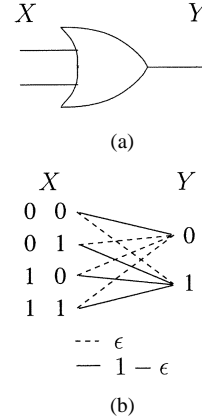


Fig. 1. A simple information transfer system. (a) A 2-input OR gate. (b) The corresponding BSC model.

where  $C_u$  is the information transfer capacity per use,  $f_c$  is the rate at which the system is being operated, and  $I(X; Y)$  is the mutual information, which is defined as

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (4)$$

and  $H(X|Y)$  is the conditional entropy of  $X$  conditioned on  $Y$ .

In [8], we have shown that any transformation  $\mathcal{F}$  with input  $X$  and output  $Y$  has an information transfer rate  $R$  given by  $R = f_s H(Y)$ , where  $H(Y)$  is the entropy of the output  $Y$  and  $f_s$  is the rate at which the original input data are being generated by a source. Note that  $f_s$  can be different from  $f_c$  due to input coding. Information theory [9] indicates that it is possible to achieve an information transfer rate  $R$  with arbitrarily low probability of error (by properly coding the input) as long as  $C > R$ .

The lower bounds on energy dissipation for single-output static gates have been studied previously [11] by modeling noisy gates as a binary symmetric communication channel (BSC) and employing information-theoretic concepts. Consider a 2-input OR gate as shown in Fig. 1(a), where the input  $X$  is generated from  $\mathcal{X} = \{00, 01, 10, 11\}$ . Due to noise, the output of the gate will deviate from its nominal values. Assume that the output is passed through a hard-decision device such as a latch. The latched output, denoted by  $Y$ , can be regarded as a binary signal that contains errors with a probability  $\epsilon$ . This can be represented by the BSC model as shown in Fig. 1(b).

Employing the BSC model, the lower bounds on energy dissipation of single-output static gates can be determined under the information-theoretic constraint  $C > R$ . A key advantage of the BSC model is that it can be used to model logic errors quite easily. However, the BSC model has two disadvantages. First, the complexity of a BSC model for a complex VLSI system is very high. Second, the energy-efficiency bounds obtained using the BSC model will not be as accurate as those obtained from models that do not quantize the output and noise. In this paper, we propose a model that eliminates these drawbacks.

## III. THE SDC MODEL

In this section, we develop a SDC model for DSM VLSI systems. We first provide a physical basis for the proposed model and then develop the channel capacity formula to compute the bounds on energy-efficiency.

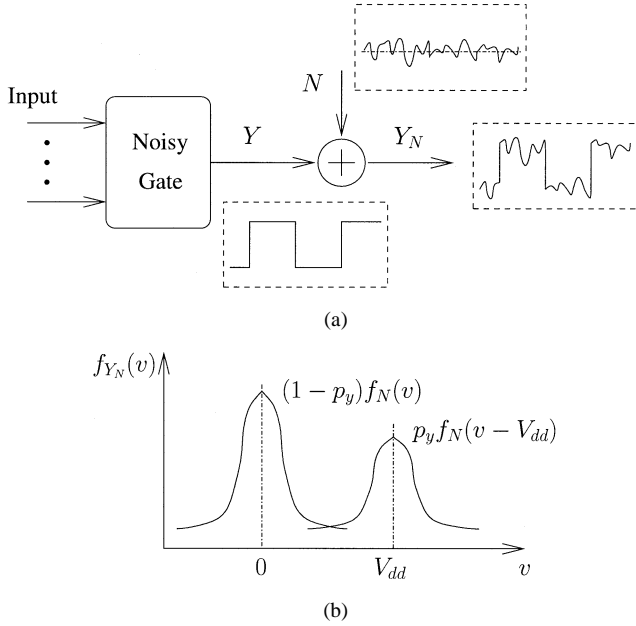


Fig. 2. The proposed SDC model for noisy gates. (a) Voltage waveforms and (b) output distribution.

### A. SDC Model

The proposed SDC model for a single-output noisy gate is illustrated in Fig. 2(a). In the presence of noise, the voltage waveform at the output  $Y_N$  is composed of an ideal output voltage  $Y$  and additive noise voltage  $N$  that is assumed to be independent of the input and output  $Y$ , i.e.,

$$Y_N = Y + N. \quad (5)$$

The ideal output  $Y$  is a binary signal with a statistical distribution given by  $P(Y = V_{dd}) = p_y$  and  $P(Y = 0) = 1 - p_y$ , where  $p_y$  is determined by the input statistics and the logic function of the gate. The noise voltage  $N$  represents a composite effect due to thermal noise and other DSM phenomena such as ground bounce, crosstalk, charge-sharing, leakage, and process variations. In this paper, we assume that the noise voltage  $N$  has a PDF denoted by  $f_N(v)$ . From (5), the statistical distribution  $f_{Y_N}(v)$  of the noisy output  $Y_N$  can be expressed as

$$f_{Y_N}(v) = \begin{cases} f_N(v - V_{dd}), & \text{if } Y = V_{dd} \\ f_N(v) & \text{if } Y = 0. \end{cases} \quad (6)$$

Note that  $Y_N$  can be considered as a bimodal continuous random signal whose value is either  $V_{dd} + N$  or  $N$ . A typical distribution of  $Y_N$  is depicted in Fig. 2(b), from which we can rewrite  $f_{Y_N}(v)$  as

$$f_{Y_N}(v) = p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v). \quad (7)$$

It can be shown that the previously proposed BSC model [11] is a special case of the proposed SDC model when the input and output of a noisy gate are latched synchronously. Assume that the latch being employed has a logic threshold denoted by

$V_{\text{th-latch}}$ . Thus, noise voltage  $N$  can introduce logic errors at the output of the latch with a probability  $\epsilon$  given by

$$\epsilon = (1 - p_y) \int_{V_{\text{th-latch}}}^{\infty} f_N(v) dv + p_y \int_{-\infty}^{V_{\text{th-latch}}} f_N(v - V_{dd}) dv. \quad (8)$$

There are two fundamental differences between the BSC model and the proposed SDC model. First, the BSC model quantizes noise contributions by employing error probability  $\epsilon$  in the computation of lower bounds. This requires noiseless hard-decision devices (e.g., latches) employed at the inputs and outputs of noisy gates. The proposed SDC model relaxes this constraint by modeling noiseless logic signals as binary signals with voltage levels  $V_{dd}$  and  $0$  V and signaling probability  $p_y$ , and DSM noise as a continuous random signal with a specific statistical distribution. Thus, all the signals and noise are captured in such a way that reflects their inherent physical nature.

Second, the proposed SDC model leads to an efficient algorithm to compute the lower bounds on energy dissipation. The computational complexity of the algorithm increases linearly with system complexity, making the SDC model applicable to complex digital systems. We will discuss this point in detail in Section IV.

### B. Information Transfer Capacity

We now derive the information transfer capacity for DSM VLSI systems using the proposed SDC model. We start with a simple  $n$ -input, single-output logic gate and then extend the framework to complex systems. Lemma 1 presented below provides the formula of mutual information for noisy gates. It is then employed in deriving the information transfer capacity in Theorem 1.

*Lemma 1:* Consider a noisy digital gate with  $n$  binary inputs  $X_0, X_1, \dots, X_{n-1}$  and single output  $Y_N = Y + N$ , where  $Y$  is the noiseless output and all the noise sources contribute a noise voltage  $N$  with a distribution denoted by  $f_N(v)$ . The mutual information  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  for this gate is given by

$$I(Y_N; X_0, X_1, \dots, X_{n-1}) = - \int_{-\infty}^{\infty} [f_{Y_N}(v) \log_2 f_{Y_N}(v) - f_N(v) \log_2 f_N(v)] dv, \quad (9)$$

where  $f_{Y_N}(v)$  is the PDF of  $Y_N$ , given by

$$f_{Y_N}(v) = p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v). \quad (10)$$

The proof of Lemma 1 is straightforward. From the definition of mutual information (4), we have

$$\begin{aligned} I(Y_N; X_0, X_1, \dots, X_{n-1}) &= H(Y_N) - H(Y_N | X_0, X_1, \dots, X_{n-1}) \\ &= H(Y_N) - H(N) \end{aligned} \quad (11)$$

where both  $Y_N$  and  $N$  are continuous signals. Substituting  $f_{Y_N}(v)$  and  $f_N(v)$  into (11) and using (1), we get (9).

Note that  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  given by (9) is a function of supply voltage  $V_{dd}$ , noise distribution  $f_N(v)$ , and output

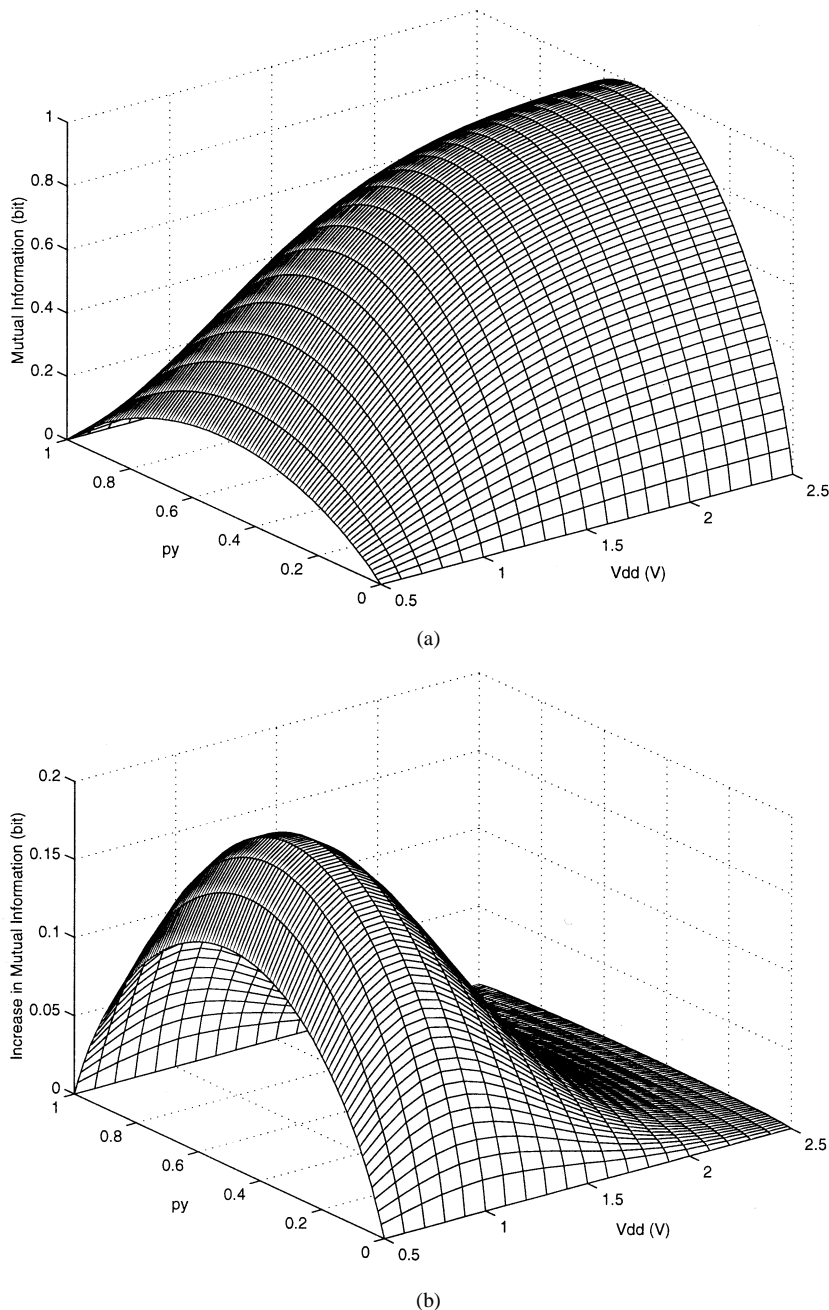


Fig. 3. Mutual information with respect to  $V_{dd}$  and  $p_y$ . (a)  $\sigma_N = 0.4$  V. (b) The increase for  $\sigma_N = 0.3$  V.

probability  $p_y$ . The value of  $p_y$  is determined by the input statistics and the logic function of the gate. In general, it is difficult to obtain an analytical solution for (9); however, numerical solutions are easy to obtain [20].

From Lemma 1, we obtain the information transfer capacity per use as summarized below.

**Theorem 1:** The information transfer capacity per use  $C_u$  of an  $n$ -input, single-output noisy gate is given by

$$C_u = \max_{p_y} I(Y_N; X_0, X_1, \dots, X_{n-1}) \\ = I(Y_N; X_0, X_1, \dots, X_{n-1})|_{p_y=0.5}. \quad (12)$$

The proof of Theorem 1 is provided in Appendix A. Theorem 1 indicates that the capacity per use  $C_u$  is achieved when the

ideal output  $Y$  has a uniform distribution [i.e.,  $P(Y = V_{dd}) = P(Y = 0) = 0.5$ ]. This is consistent with the observation that  $p_y = 0.5$  implies maximum uncertainty in  $Y$  and hence, the largest information content being transferred.

Fig. 3(a) plots the mutual information  $I(Y_N; X)$  using Lemma 1 for a 2-input OR gate in a  $0.25 \mu\text{m}$  CMOS process. The noise voltage  $N$  is assumed to be a zero-mean Gaussian distribution with a variance  $\sigma_N = 0.4$  V. As indicated, for every supply voltage  $V_{dd}$ ,  $I(Y_N; X)$  reaches the maximum when  $p_y = 0.5$ . In addition,  $I(Y_N; X)$  is symmetric around  $p_y = 0.5$ . This is consistent with the BSC model. The capacity per use  $C_u$  approaches 1 b/use as  $V_{dd}$  increases. This can be interpreted as the ability of the 2-input OR gate to transfer 1 bit of information for every use provided the supply voltage

$V_{dd}$  is sufficiently high with respect to noise. This is to be expected given the fact that for high enough  $V_{dd}$ , noise becomes negligible and the gate can be approximated as being noiseless. On the other hand,  $C_u$  decreases toward zero with the reduction in  $V_{dd}$ . This reflects the practical scenario where the gate stops functioning on being powered down.

The impact of noise on  $C_u$  is illustrated in Fig. 3(b), where we observe an increase in  $I(Y_N; X)$  (and hence  $C_u$ ) as noise variance  $\sigma_N$  being reduced from 0.4 V to 0.3 V. Obviously, a gate operating in a less noisy medium is more robust and hence can transfer more information.

We now compute the information transfer capacity. Assume that the NMOS and PMOS transistors being used are balanced implying the same low-to-high and high-to-low delays. The maximum signaling rate  $f_c$  at a supply voltage  $V_{dd}$  can be approximated by [21]

$$f_c = \frac{k_m(V_{dd} - V_t)^\alpha}{C_L V_{dd}} \quad (13)$$

where  $k_m$  is the transconductance of the balanced NMOS and PMOS transistors,  $C_L$  is the load capacitance,  $V_t$  is the transistor threshold voltage, and  $\alpha$  is the velocity saturation index ranging from 1 (velocity saturated) to 2 (without velocity saturation). The information transfer capacity can be obtained as  $C = C_u f_c$ , where  $C_u$  and  $f_c$  are given by (12) and (13), respectively.

As an example, we consider the 2-input OR gate in a 0.25  $\mu\text{m}$  CMOS process with the following parameters: 1)  $k_m = 80 \mu\text{A}/\text{V}^2$ , 2)  $V_{tn} = |V_{tp}| = 0.4 \text{ V}$ , 3)  $C_L = 30 \text{ fF}$ , 4)  $\alpha = 1.2$ , and 5) information transfer rate  $R = 800 \text{ MB/s}$ . The noise voltage  $N$  is assumed to be a zero-mean Gaussian distribution with a variance  $\sigma_N = 0.4 \text{ V}$ . With  $V_{dd} = 2.5 \text{ V}$ , we obtain  $C_u \approx 1 \text{ b/use}$  and  $f_c = 2.598 \text{ GHz}$ . Therefore, the information transfer capacity of the gate is given by  $C = C_u f_c = 2.598 \text{ Gb/s}$ . Similarly, with  $V_{dd} = 1.0 \text{ V}$ , we obtain  $C_u \approx 0.6 \text{ b/use}$  and  $f_c = 1.445 \text{ GHz}$ . This gives  $C = 867 \text{ MB/s}$ , which is still larger than the information transfer requirement  $R$ , implying that through appropriate coding, the gate can be operated reliably at a supply voltage as low as 1.0 V. The degradation in  $C$  is primarily due to noise impact becoming increasingly significant as  $V_{dd}$  reduces. Present-day digital circuits operate at sufficiently high voltages so that  $C_u = 1 \text{ b/use}$  and hence  $R$  can be as high as  $f_c$ .

It is worth mentioning that (12) and (13) provide a direct correspondence between the information transfer capacity  $C$  and implementation details such as supply voltage, load capacitance, circuit style, CMOS process, and noise parameters. While in this paper we employ rather general models for speed and noise, any specific assumptions or modifications can be easily incorporated into the framework. One such example is the computation of bounds on energy dissipation for noise-tolerant circuit techniques (see Section IV-D), where noise contributions are modeled as an input noise voltage rather than at the output as in the SDC model. Assuming the gate with a voltage transfer characteristic (VTC) function  $\mathcal{T}$  [22], the equivalent output noise voltage can be expressed as

$$N_{\text{out}} = \mathcal{T}(N_{\text{in}}) \quad (14)$$

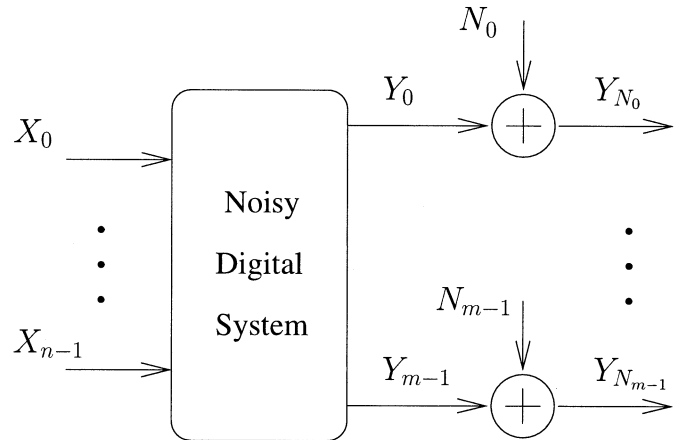


Fig. 4. The SDC model for noisy digital systems.

and has a distribution given by [23]

$$f_{N_{\text{out}}}(v) = \sum_{i=1}^P \frac{f_{N_{\text{in}}}(v_i)}{\mathcal{T}'(v_i)} \quad (15)$$

where  $f_{N_{\text{in}}}(v)$  and  $f_{N_{\text{out}}}(v)$  are the PDFs for the input noise voltage  $N_{\text{in}}$  and the equivalent output noise voltage  $N_{\text{out}}$ , respectively, and  $v_i$  is the root of (14). Thus, we can compute the information transfer capacity for gates subject to input noise by substituting  $f_N(v) = f_{N_{\text{out}}}(v)$  into (9), (10), and (12).

#### IV. LOWER BOUNDS ON ENERGY DISSIPATION

In this section, we determine the lower bounds on energy dissipation using the proposed SDC model. These bounds are obtained by solving an energy optimization problem while being subject to the information-theoretic constraint  $C > R$  on reliability. In Section IV-A, we formulate the constrained optimization problem and develop an analytical solution by employing the *Lagrange multiplier* method [24]. In Section IV-B, we present an algorithm to compute the lower bounds and then apply it to multimodule systems, dynamic circuits and noise-tolerant dynamic circuits in Section IV-C to IV-D, respectively.

##### A. Problem Formulation and Solution

Consider a generic digital system as depicted in Fig. 4. We assume, without loss of generality, that the system consists of  $m$  noisy modules, each generating an ideal (noiseless) output denoted by  $Y_i$  for  $i = 0, 1, \dots, m-1$ . The input of the system is given by  $X = \{X_0, X_1, \dots, X_{n-1}\}$ , where  $X_i$  is a binary signal. From the SDC model, the total noise contribution can be represented by  $m$  output noise voltages  $N_0, N_1, \dots, N_{m-1}$ . In addition, we assume these noise voltages are independent of each other. The outputs of the system can be expressed as

$$Y_{N_i} = Y_i + N_i \quad i = 0, 1, \dots, m-1 \quad (16)$$

where  $Y_{N_i}$  denotes the actual voltage waveform at the  $i$ th output and  $N_i$  is the corresponding noise voltage with a distribution given by  $f_{N_i}(v)$ . Note that the assumption on independent noise sources is pessimistic, i.e., the resulting bounds would be greater

than those obtained if the noise sources were correlated. This is due to the fact that independent noise sources incur the largest information loss [9] and thus require higher output probabilities  $p_{y_i}$ s and transition probabilities  $t_i$ s to compensate. Furthermore, making the independence assumption simplifies the mathematical development so that the key ideas in the paper can be illustrated clearly.

We consider a silicon implementation where all the NMOS and PMOS transistors share a common power supply and ground. In addition, the NMOS and PMOS transistors are properly sized so that the modules can be operated at the same speed at a nominal supply voltage. For the sake of simplicity, we assume that all the capacitances including the parasitic capacitance, interconnect capacitance, and input capacitance from the following stage are lumped into the load capacitance at the output of the system.

In what follows, we consider the total power dissipation  $P_{\text{tot}}$  to consist primarily of the capacitive component of power dissipation, also referred to as dynamic power dissipation  $P_{\text{dyn}}$ , which is given by [22]

$$P_{\text{dyn}} = \sum_{i=0}^{m-1} t_i C_{L_i} V_{dd}^2 f_c \quad (17)$$

where  $t_i$  and  $C_{L_i}$  are the average transition probability and the equivalent load capacitance, respectively, for the  $i$ th output, and  $f_c$  is given by (13). In Section IV-D, we will include other power components (e.g., short-circuit and static power) to determine the lower bounds for domino and noise-tolerant circuits.

We note that for dynamic circuits such as conventional domino,  $t_i$  will be equal to the probability of output  $Y_i$  being  $V_{dd}$  (or a logic “1”), i.e.,  $t_i = p_{y_i}$ . This is valid for static circuits as well provided transition signaling [11] (i.e., a logic “1” is represented with a transition and a logic “0” is represented with no transition) is employed at the output. Therefore, in the rest of the paper we will replace  $p_y$  in the information capacity expressions (9)–(12) with the transition probability  $t$  as these two measures are equivalent.

We now determine the lower bounds on energy dissipation for DSM VLSI systems as shown in Fig. 4, using the SDC model. As discussed in Section III, the mutual information  $I(Y_{N_0}, \dots, Y_{N_{m-1}}; X_0, \dots, X_{n-1})$  for such systems is determined by the supply voltage  $V_{dd}$ , output transition probabilities  $t_0, t_1, \dots, t_{m-1}$  and noise parameters. To simplify notation, we rewrite  $I(Y_{N_0}, \dots, Y_{N_{m-1}}; X_0, \dots, X_{n-1})$  as an explicit function of these parameters as follows

$$I(Y_{N_0}, \dots, Y_{N_{m-1}}; X_0, \dots, X_{n-1}) \triangleq I_{Y_N; X}(V_{dd}, t_0, \dots, t_{m-1}, \sigma_{N_0}^2, \dots, \sigma_{N_{m-1}}^2) \quad (18)$$

where  $\sigma_{N_i}^2$  is the variance (energy) of noise voltage  $N_i$ , which is determined by the distribution function  $f_{N_i}(v)$ . Employing similar arguments as in [11], the lower bounds on energy dissipation can be obtained by solving the following optimization problem

$$\text{minimize: } E_b = \frac{P_{\text{tot}}}{R} \quad (19)$$

$$\text{subject to: } I_{Y_N; X}(V_{dd}, t_0, \dots, t_{m-1}, \sigma_{N_0}^2, \dots, \sigma_{N_{m-1}}^2) f_c = R \quad (20)$$

where  $E_b$  is the energy per information bit.

Note that for any given supply voltage  $V_{dd}$  and noise parameters, the information-theoretic constraint (20) and power dissipation are functions of transition probabilities  $t_0, t_1, \dots, t_{m-1}$ . Hence, the optimum solution to (19) and (20) is a set of  $t_i$ s ( $i = 0, 1, \dots, m-1$ ) that minimizes the power dissipation while satisfying the information-theoretic constraint on reliability. Employing the Lagrange multiplier method [24], we obtain the optimum solution to (19) and (20) as summarized in Theorem 2.

**Theorem 2:** The lower bound on energy dissipation for a digital system consisting of  $m$  noisy modules is achieved with transition probabilities  $t_0^{\text{opt}}, t_1^{\text{opt}}, \dots, t_{m-1}^{\text{opt}}$  satisfying

$$\left. \frac{\partial I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)}{\partial t_i} \right|_{(t_i=t_i^{\text{opt}})} = \beta_{\text{max}} C_{L_i} V_{dd}^2, \quad i = 0, 1, \dots, m-1 \quad (21)$$

where  $C_{L_i}$ ,  $t_i$ , and  $\sigma_{N_i}^2$  are the load capacitance, transition probability, and noise variance, respectively, at the  $i$ th output,  $\beta_{\text{max}}$  is a constant determined by the information-theoretic constraint on reliability, and  $I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)$  is the mutual information for the  $i$ th output, which is given by

$$\begin{aligned} I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2) &= - \int_{-\infty}^{\infty} [(t_i f_{N_i}(v - V_{dd}) + (1 - t_i) f_{N_i}(v)) \\ &\quad \cdot \log_2(t_i f_{N_i}(v - V_{dd}) + (1 - t_i) f_{N_i}(v)) \\ &\quad - f_{N_i}(v) \log_2 f_{N_i}(v)] dv \end{aligned} \quad (22)$$

where  $f_{N_i}(v)$  is the distribution function for the output noise  $N_i$ .

The proof of Theorem 2 is provided in Appendix B. In Section IV-B to IV-D, we will demonstrate the use of Theorem 2 to compute the lower bounds on energy dissipation for various digital systems.

### B. Computation of Lower Bounds

We assume that the NMOS and PMOS transistors being used are properly sized to operate at the same speed at a nominal supply voltage. Thus, the parameters  $k_m$  and  $C_{L_i}$ s are fixed making the lower bounds a function of the supply voltage  $V_{dd}$  and transition probabilities  $t_i$ s. From Theorem 2, the objective is to find an optimum combination  $(t_0^{\text{opt}}, t_1^{\text{opt}}, \dots, t_{m-1}^{\text{opt}})$  at each  $V_{dd}$  such that the power dissipation is minimized subject to the information-theoretic requirement  $R$ . Fig. 5 shows the algorithm for computing the lower bounds. For each  $V_{dd}$  we start with a sufficiently small  $\beta$  and compute the values of  $t_i$ s using (21) and (22). These  $t_i$ s are then employed to compute the information transfer metric  $I_{Y_N; X}(V_{dd}, t_0, \dots, t_{m-1}, \sigma_{N_0}^2, \dots, \sigma_{N_{m-1}}^2) f_c$  and the result is compared to the information transfer rate  $R$ . If  $I_{Y_N; X}(V_{dd}, t_0, \dots, t_{m-1}, \sigma_{N_0}^2, \dots, \sigma_{N_{m-1}}^2) f_c > R$ , the value of  $\beta$  will be increased in small steps until (20) is just

```

inputs:  $f_{N_0}(v), \dots, f_{N_{m-1}}(v), C_{L_0}, \dots, C_{L_{m-1}}, V_{dd}, V_L, k_m, R;$ 
output:  $E_b;$ 
initial:
  compute  $I_{Y_{N_i};X}(V_{dd}, t_i, \sigma_{N_i}^2)$  using (9);
  compute  $\partial I_{Y_{N_i};X}(V_{dd}, t_i, \sigma_{N_i}^2)/\partial t_i;$ 
  let  $\beta = \beta_{init};$ 

begin
  do
    compute  $f_c$  using (13);
    compute  $t_i^o : \partial I_{Y_{N_i};X}(V_{dd}, t_i, \sigma_{N_i}^2)/\partial t_i = \beta C_{L_i} V_{dd}^2;$ 
    compute  $\sum_{i=0}^{m-1} I_{Y_{N_i};X}(V_{dd}, t_i^o, \sigma_{N_i}^2) f_c;$ 
    if  $\sum_{i=0}^{m-1} I_{Y_{N_i};X}(V_{dd}, t_i^o, \sigma_{N_i}^2) f_c < R$ 
       $\beta = \beta - \Delta\beta;$ 
    else if  $\sum_{i=0}^{m-1} I_{Y_{N_i};X}(V_{dd}, t_i^o, \sigma_{N_i}^2) f_c > R$ 
       $\beta = \beta + \Delta\beta;$ 
    else if  $\sum_{i=0}^{m-1} I_{Y_{N_i};X}(V_{dd}, t_i^o, \sigma_{N_i}^2) f_c = R$ 
       $\beta = \beta_{max};$ 
       $t_i^{opt} = t_i^o, i = 0, \dots, m-1;$ 
      compute  $E_b$  using (19);
    end if
     $V_{dd} = V_{dd} - \Delta V_{dd};$ 
  until  $V_{dd} = V_{dd,min}$ 
end

```

Fig. 5. The algorithm to compute the lower bounds on energy dissipation.

satisfied. The corresponding transition probabilities  $t_i$ s are then employed with  $V_{dd}$  to obtain the lower bounds. Note that the value of  $\Delta\beta$  in Fig. 5 is determined by the precision to which the bounds need to be computed. A smaller value of  $\Delta\beta$  results in higher precision. Therefore, one can start with a fixed step-size and then reduce it as the solution converges.

Note that the proposed algorithm determines the lower bounds by joint optimization of power components from all the modules in the system. The associated computational complexity increases linearly with the number of modules in the system, making the proposed algorithm suitable for determining the energy-efficiency bounds of complex digital systems. In addition, it can be shown that the lower bounds derived via the SDC model are smaller than those obtained via the BSC model.

In Section IV-C and IV-D, we will demonstrate the computation of the lower bounds via the proposed SDC model for multimodule systems, dynamic circuits, and noise-tolerant circuit techniques.

### C. Lower Bounds for Multimodule Systems

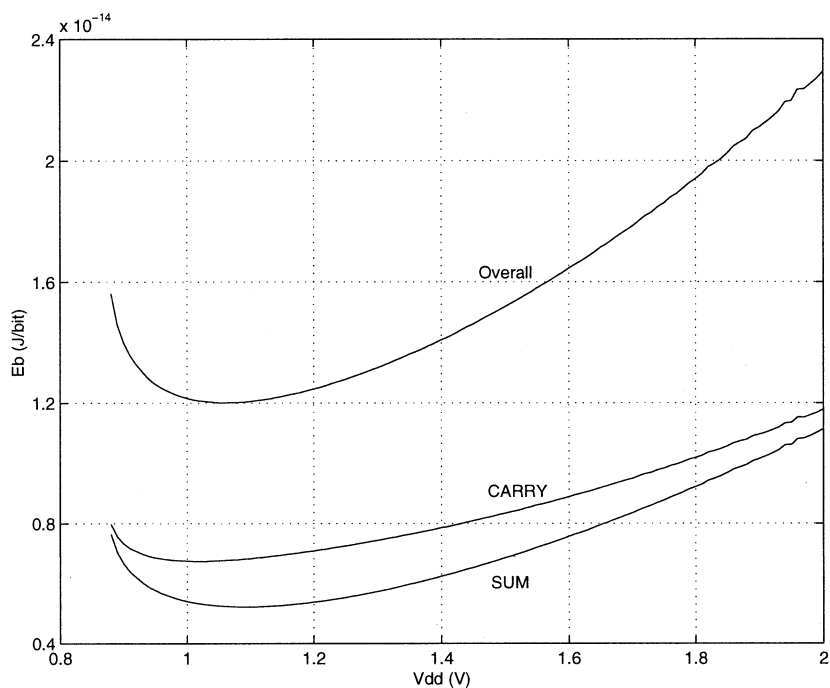
We now determine the lower bounds on energy dissipation for digital systems consisting of multiple logic modules, each of which generates a noisy output while consuming a certain amount of power. For the purpose of demonstration, we will consider a full-adder which has a SUM module and a CARRY module. The lower bounds for more complex digital systems can be determined in a similar manner.

We assume the following design parameters for the full-adder.

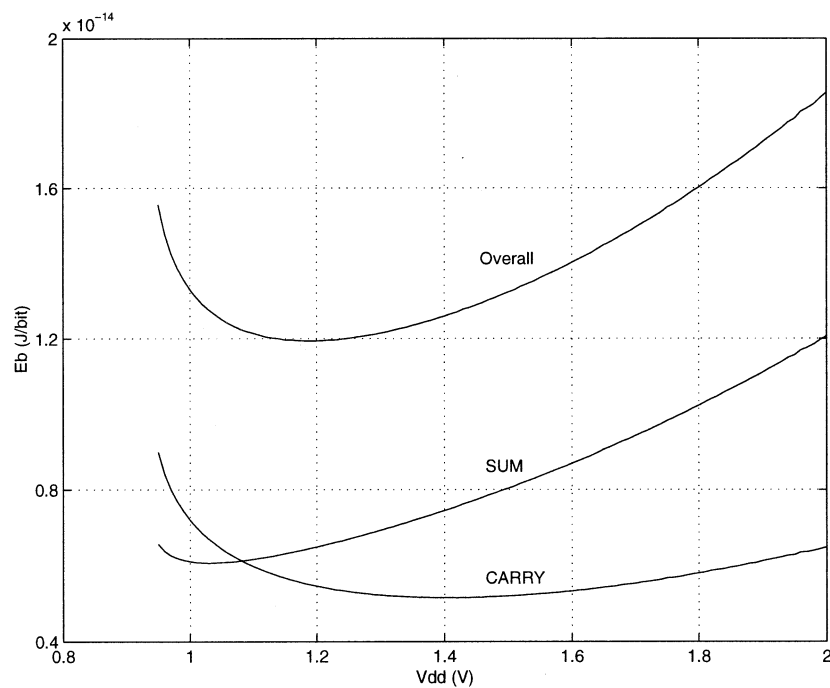
- 1) The gate is implemented in a 0.25  $\mu\text{m}$  CMOS process in static CMOS logic style with dual NMOS and PMOS networks.
- 2) The NMOS and PMOS transistors are balanced with the same propagation delay. The speed of the gate is given by (13), where  $C_L = 20 \text{ fF} - 30 \text{ fF}$ ,  $k_m = 80 \mu\text{A/V}^2$ ,  $V_{tn} = |V_{tp}| = 0.4 \text{ V}$ , and  $\alpha = 1.2$ .
- 3) The noise voltage  $N$  has a zero-mean Gaussian distribution with  $\sigma_N = 0.3 \text{ V} - 0.4 \text{ V}$ . The noise is uncorrelated to the desired input and output signals.
- 4) The gate has an information transfer rate requirement  $R = 800 \text{ Mb/s}$ .

Note that in practice the CARRY and SUM modules are possibly subject to noise with different amplitudes or driving different load capacitances. This results in different energy bounds and, hence, we will evaluate them separately.

We first consider the case where the two modules are subject to different noise amplitudes but with the same load capacitance of 30 fF. We assume that the noise voltages  $N_{\text{sum}}$  and  $N_{\text{carry}}$  are zero-mean Gaussian distributions with  $\sigma_{N_{\text{sum}}} = 0.4 \text{ V}$  and  $\sigma_{N_{\text{carry}}} = 0.3 \text{ V}$ , respectively. This implies a more reliable CARRY module. Fig. 6(a) illustrates the lower bounds for the SUM module, CARRY module and the full-adder. Note that the lower bound for the full-adder is obtained by jointly optimizing the power components of different modules under the information-theoretic constraint on reliability. The minimum values of energy per information bit  $E_b$  for the SUM module, CARRY module and the full-adder equal 5.1 fJ/b, 6.7 fJ/b and 12 fJ/b at  $V_{dd} = 1.1 \text{ V}$ , 1.0 V and 1.08 V, respectively. Also shown in



(a)



(b)

Fig. 6. Lower bounds on energy dissipation for multimodule systems subject to (a) different noise and (b) different load capacitances.

Fig. 6(a) is that the lower bound for the full-adder is achieved when the CARRY module consumes more energy than the SUM module does.

Fig. 6(b) illustrates the lower bounds on energy dissipation when the SUM and CARRY modules drive different load capacitances of 20 fF and 30 fF, respectively, but subject to the same noise amplitude ( $\sigma_N = 0.4$  V). It indicates that the lower bound for the full-adder is achieved when the output driving a larger capacitance consumes less energy. This is to be expected

because such output will have a smaller transition probability [see (21)] which offsets the energy overhead due to the large capacitance.

#### D. Energy-Efficiency Bounds for Noise-Tolerant Circuits

Noise-tolerant circuit techniques [15]–[18] improve noise-immunity by employing additional elements to prevent logic errors from occurring in the presence of noise. Thus, one would expect noise-tolerant circuits to be less energy-efficient



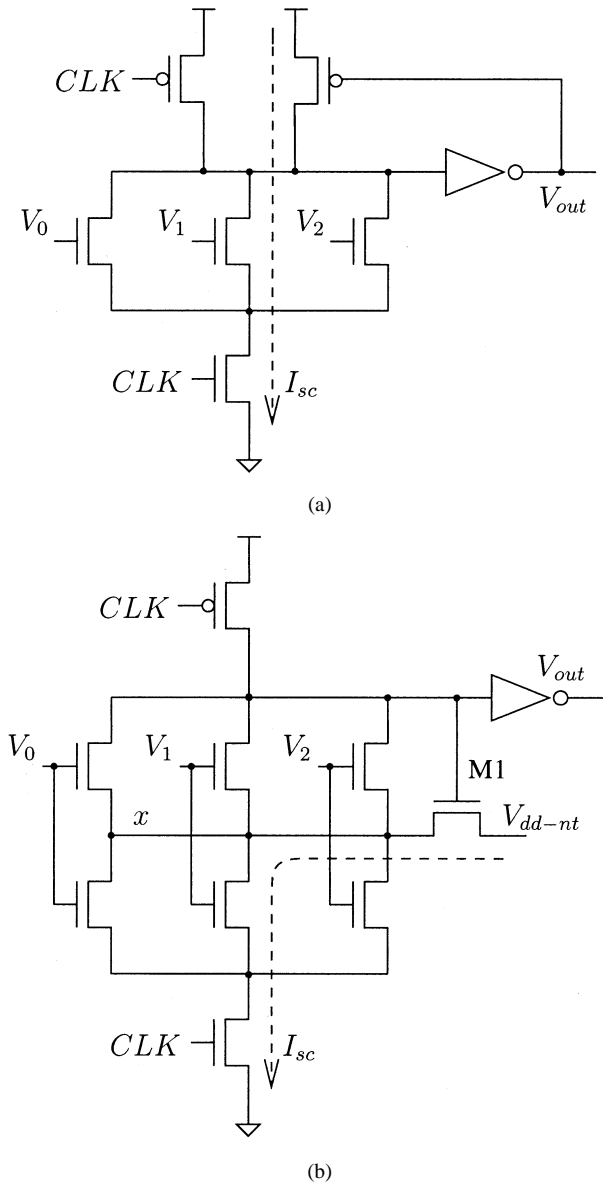


Fig. 7. Dynamic style 3-input OR gates. (a) Conventional domino. (b) Mirror technique.

than conventional circuits. In this subsection, we determine the lower bounds on energy-efficiency for noise-tolerant circuit techniques such as the mirror technique [15]. It will be shown that noise-tolerance improves the energy-efficiency when operating at the lower bound.

Fig. 7 depicts two 3-input OR gates implemented by the conventional domino (with a keeper) and the mirror technique in a 0.25  $\mu\text{m}$  CMOS technology. It is known that domino circuits are inherently susceptible to noise [5] due to their low switching threshold voltage  $V_{th}$ , defined as the input voltage at which the output changes state. For the domino OR gate shown in Fig. 7(a),  $V_{th} = V_{tn}$ , where  $V_{tn}$  is the threshold voltage of an NMOS transistor. The previously proposed mirror technique [15] improves noise-immunity via employing two identical NMOS evaluation nets. One additional NMOS transistor M1, whose gate voltage is controlled by the dynamic node voltage, provides a conduction path between the common node  $x$  of the two evaluation nets and  $V_{dd-nt}$ . During the precharge phase, transistor M1 is turned on

and the common node voltage  $V_x$  is charged up to  $(V_{dd-nt} - V_{tn})$ . Due to body-effect, the switching threshold voltage of the upper NMOS net is increased, thereby improving the noise-immunity. Note that the noise-immunity of the gate can be tuned by either changing the voltage  $V_{dd-nt}$  or resizing the transistor M1.

The total power dissipation of digital gates consists of dynamic power  $P_{dyn}$ , short-circuit power  $P_{sc}$  and static power  $P_{stat}$ , where  $P_{stat}$  has two components  $P_{sub}$  (due to sub-threshold leakage) and  $P_{DC}$  (due to DC power dissipation). Considering the fact that the two gates in Fig. 7 have  $P_{DC} = 0$  and  $P_{sub}$  is relatively small (this is because  $V_{tn} = 0.4$  V), we express the total power dissipation as

$$P_{tot} = tC_L V_{dd}^2 f_c + tI_{sc} V_{dd} \quad (23)$$

where  $I_{sc}$  is the average short-circuit current evaluated over each signaling period when the gate switches. Note that the two gates consume nontrivial short-circuit power as indicated in Fig. 7.

From (23), the problem of deriving the energy-efficiency bounds for domino and mirror circuits is stated as follows:

$$\begin{aligned} \text{minimize: } E_b &= \frac{tC_L V_{dd}^2 f_c + tI_{sc} V_{dd}}{R} \\ \text{subject to: } I_{Y_N; X}(V_{dd}, t, \sigma_N^2) f_c &= R. \end{aligned} \quad (24)$$

From Theorem 2, the solution to (24) can be obtained as

$$\left. \frac{\partial I_{Y_N; X}(V_{dd}, t, \sigma_N^2)}{\partial t} \right|_{(t=t^{opt})} = \beta_{max} (C_L V_{dd}^2 + I_{sc} V_{dd} T_c) \quad (25)$$

where  $T_c = 1/f_c$  is the signaling period.

Fig. 8 illustrates the lower bounds derived from the proposed SDC model for a 3-input static, domino, and mirror OR gates. The mirror gate was designed to have  $V_{dd-nt} = V_{dd}$  with M1 being eight times the minimum size. For consistency, we sized the transistors in all three gates to operate at the same speed at nominal  $V_{dd} = 2.5$  V while driving a 30 fF load. This implies that the pull-down NMOS transistors in Fig. 7(b) are sized up resulting in larger parasitic capacitances. We account for this design overhead by extracting the capacitances from the layout and adding them to the 30 fF load capacitance. Please note that during the computation of the lower bounds we relax the speed constraint and instead keep only the constraint on the information transfer rate  $R = 800$  Mb/s. We assume all three gates to be subject to an input noise voltage  $N_{in}$  which is zero-mean Gaussian with  $\sigma_{N_{in}} = 0.4$  V. The equivalent output noise voltages are obtained from (14) and (15). As shown, the minimum  $E_b$  of the conventional domino gate is found to be 20 fJ/b, whereas that of the mirror gate is 13 fJ/b, which is 35% lower. Furthermore, we also see that the static gate, while being inherently noise-tolerant, is also energy-inefficient when operating at the lower bound.

Thus, we find that from an information-theoretic perspective, noise-tolerant circuits provide the best trade off between noise-immunity and energy-efficiency than either domino or static circuits.

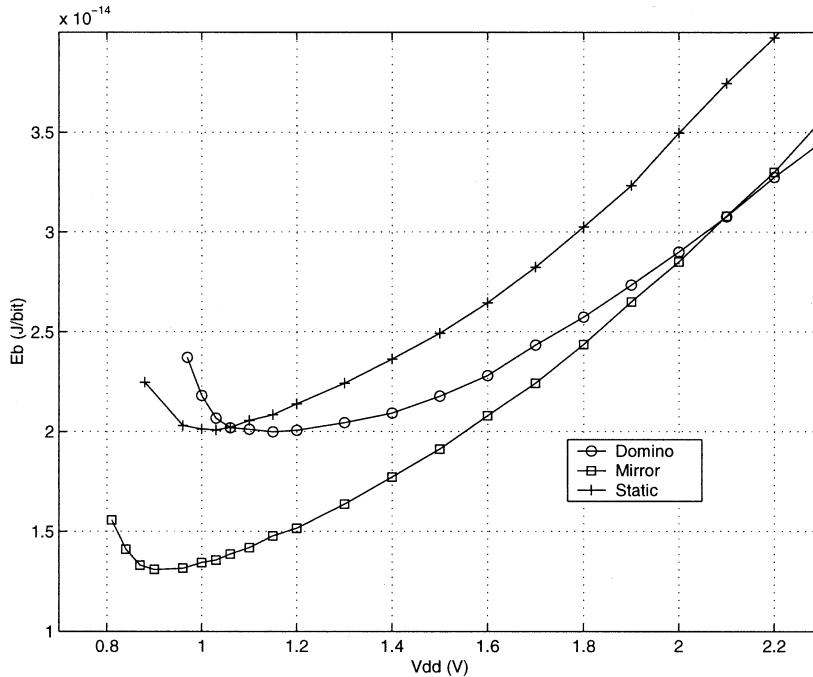


Fig. 8. Lower bounds on energy dissipation via the proposed SDC model.

## V. APPLICATION TO HIGH-SPEED, LOW-POWER ARITHMETIC CIRCUITS

In this section, we employ noise-tolerance in high-speed arithmetic circuits to reduce power dissipation subject to a specified level of algorithmic performance, thereby reducing the gap between the bounds on energy-efficiency and the actual energy dissipation of practical systems. In particular, we consider adders with large word sizes which are a key datapath element in the design of high-performance microprocessors and digital signal processing systems. The problem of improving reliability while maintaining energy-efficiency is of great importance given the design challenge of achieving high data rate in noisy media. This requires joint optimization of noise-immunity, energy dissipation, speed as well as other design parameters.

Wide adders are typically constructed by combining identical smaller adder modules. One commonly used module is the *Manchester* domino adder [22] which employs the concept of *carry lookahead* for speed improvement. In Section V-A, we propose a noise-tolerant scheme based on the mirror technique to improve the noise-immunity of conventional Manchester adders. In Section V-B, we define performance measures that will be employed to evaluate the usefulness of the proposed noise-tolerant Manchester adder in a digital signal processing system. In Section V-C, we apply this noise-tolerant Manchester adder in a CDMA receiver and demonstrate a 31.2%–51.4% energy reduction. In addition, we compute the lower bounds on energy dissipation for this CDMA receiver and compare it with the actual energy dissipated via the use of a noise-tolerant design. We show that the lower bounds on energy dissipation are 2.8× below the actual energy consumed. This example shows very clearly that noise-tolerance is effective in reducing the gap between the actual energy consumed and the lower bounds.

### A. Noise-Tolerant, High-Speed Adder Design

The speed of wide adders is limited by the speed of carry signals propagating through the carry chain. Carry lookahead technique [22] computes carry signals to each stage in parallel, thereby improving the speed. For the  $i$ th stage, the carry signal  $C_i$  and sum signal  $S_i$  are obtained as

$$C_i = G_i + P_i C_{i-1} \quad (26)$$

$$S_i = P_i \oplus C_{i-1} \quad (27)$$

where  $\oplus$  denotes the XOR operation,  $G_i$  and  $P_i$  are the generate and propagate signals, respectively, which are given by

$$G_i = A_i B_i \quad (28)$$

$$P_i = A_i + B_i \quad (29)$$

where  $A_i$  and  $B_i$  are the  $i$ th bit of two input operands  $A$  and  $B$ , respectively. Expanding (26), we get

$$C_i = G_i + P_i G_{i-1} + P_i P_{i-1} G_{i-2} + \dots + P_i \dots P_1 C_0. \quad (30)$$

From (30), the complexity of computing  $C_i$  becomes large very quickly as the bit-width increases. Hence, carry lookahead addition typically spans no more than four stages.

Manchester adders employ a domino style of carry lookahead addition for high-speed and low-complexity. Fig. 9(a) illustrates the circuit schematic of a conventional Manchester adder, where the carry signals  $\overline{C_0} \text{--} \overline{C_4}$  are generated in parallel from internal nodes. A carry-bypass scheme is employed to reduce the worst-case delay when all  $P_i$ s are “1.” It is known that Manchester adders improve speed by approximately 4× over ripple-carry adders.

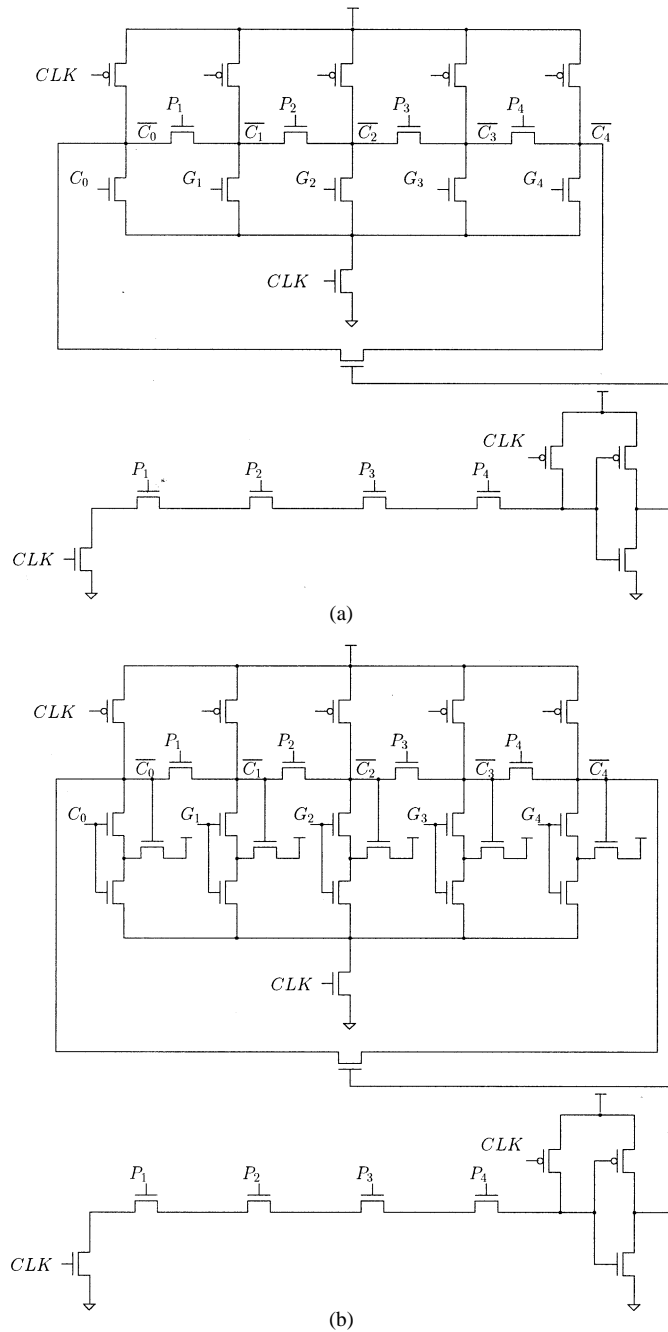


Fig. 9. Manchester carry chain: (a) domino and (b) noise-tolerant design.

While Manchester adders are fast, the inherent presence of domino style makes it susceptible to noise, thereby putting a tight requirement on supply voltage for reliable operation. Such adders are, therefore, power hungry compared to ripple-carry adders. In this section, we apply the mirror technique to the design of a noise-tolerant Manchester adder. The result we expect to demonstrate is that a noise-tolerant Manchester adder will be much more energy-efficient than a conventional Manchester adder while delivering the same algorithmic performance, i.e., the SNR, when employed in a CDMA receiver.

As shown in Fig. 9(b), the proposed scheme protects the error-prone dynamic nodes  $\overline{C_0}$ – $\overline{C_4}$  by employing mirror transistors for short pull-down paths, i.e., paths consisting of NMOS transistors with  $G_i$  and  $CLK$  as their gate inputs. Note that this ap-

proach is effective because a longer pull-down path with more stacked NMOS transistors is more robust to noise. In addition, short pull-down paths are not on the critical delay paths and, hence, do not affect the overall speed when applying noise-tolerance.

### B. Performance Measures

We assume that the magnitude and duration of intermittent noise pulses are sufficient to cause logic errors. The error probability  $\epsilon$  for dynamic gates can be obtained from the *noise-immunity curves* (NICs) [15], [25]. As shown in Fig. 10, a point on the NIC indicates the duration and amplitude of an input noise pulse  $V_n$  that will erroneously discharge dynamic nodes and cause an output error. Thus, noise pulses corresponding to the points that lie above the NIC will cause output errors. Obviously, the more noise-immune a circuit technique is, the higher its NIC will be.

Assume that the evaluation time equals  $T_c/2 = 1/(2f_c)$  and that the corresponding point on the NIC is denoted by  $V_N$ . Given a noise model that consists of a distribution  $f_{V_n, T_n}(v, t)$  on the amplitude and duration as shown in Fig. 10, the error probability  $\epsilon$  can be obtained as

$$\epsilon = \int_{V_N}^{\infty} \int_{t_v}^{T_c/2} f_{V_n, T_n}(v, t) dt dv \quad (31)$$

where the second integral starts from  $t_v$  (determined by the variable  $v$  of the first integral and the given NIC) and ends at  $T_c/2$ . From (31),  $\epsilon$  is a function of supply voltage  $V_{dd}$  because both  $T_c$  and  $V_N$  are functions of  $V_{dd}$ . Note that noise-tolerant circuits will have a smaller probability of error  $\epsilon$  as compared to conventional domino. Also, a higher  $V_{dd}$  reduces  $\epsilon$ .

It is more convenient to use the measure of *mean-squared error* (MSE) for arithmetic circuits because errors at different output bits have different weights (e.g., an error occurring at the  $i$ th bit has a value of  $\pm 2^i$ ). The MSE, denoted by  $\sigma_e^2$ , for an  $n$ -bit adder is defined as

$$\sigma_e^2 = \sum_{i=0}^{n-1} \epsilon_i (2^i)^2 \quad (32)$$

where  $\epsilon_i$  is the error probability at the  $i$ th output bit and can be computed via (31).

### C. Performance Comparison

We now present the results of algorithmic performance and energy dissipation for the proposed noise-tolerant adder scheme in the context of a CDMA wireless communication system [26]. The basic principle of CDMA is to spread the spectrum of a narrowband message signal by multiplying it with a wideband binary pseudo-noise (PN) sequence whose rate is  $N_c$  times that of the original signal, where  $N_c$  is the length of the PN sequence. It can be shown that this type of modulation has the property of suppressing jamming, interference from other users, and self-interference due to multipath propagation. Due to this, CDMA techniques have been widely employed in multiuser wireless communications.

In the receiver, the incoming signal needs to be despread by the same PN sequence to recover the transmitted symbols.

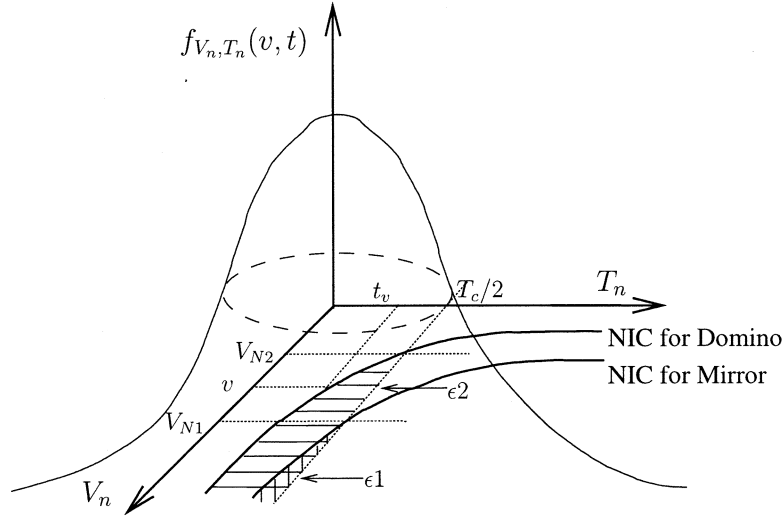
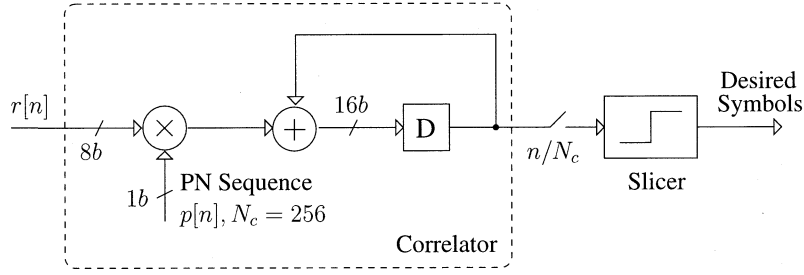

 Fig. 10. Determining  $\epsilon$  from noise-immunity curves.


Fig. 11. A correlator for CDMA communications.

This can be achieved by a correlation operation as illustrated in Fig. 11, where a multiplication involves computing the absolute value of the received signal. We assume that the received signal  $r[n]$  has 8-bit precision and the length of the binary PN sequence  $p[n]$  is  $N_c = 256$ . The accumulator has 16-bit precision, which requires four 4-bit Manchester adders (see Fig. 9) for high-speed addition. The noise from the underlying circuits is assumed to have an amplitude  $V_n$  which is zero-mean Gaussian with  $\sigma_n = 0.4$  V, and a duration  $T_n$  which is uniformly distributed between 0 and  $T_c/2$ . The magnitude and duration of intermittent noise pulses are sufficient to cause logic errors during accumulation. The error probability is computed from (31) and then employed to flip the adder outputs in order to emulate a noisy hardware. The final output has a  $SNR$  requirement of 20 dB as recommended in [27] and expressed as

$$SNR = 10 \log_{10} \left( \frac{\sigma_s^2}{\sigma_n^2 + \sigma_e^2} \right) \geq 20 \text{ dB} \quad (33)$$

where  $\sigma_s^2$  and  $\sigma_n^2$  are the variances of the desired signal and signal noise, respectively, and  $\sigma_e^2$  is given by (32).

Fig. 12(a) shows the plot of energy dissipation versus SNR for different designs. The curve denoted by “NT( $i$ )” refers to the 16-bit adder where the top  $i$  MSBs are implemented via the mirror technique [see Fig. 9(b)]. To achieve the specified SNR, the domino Manchester adder consumes the maximum energy due to its low noise-immunity requiring a high  $V_{dd}$  for

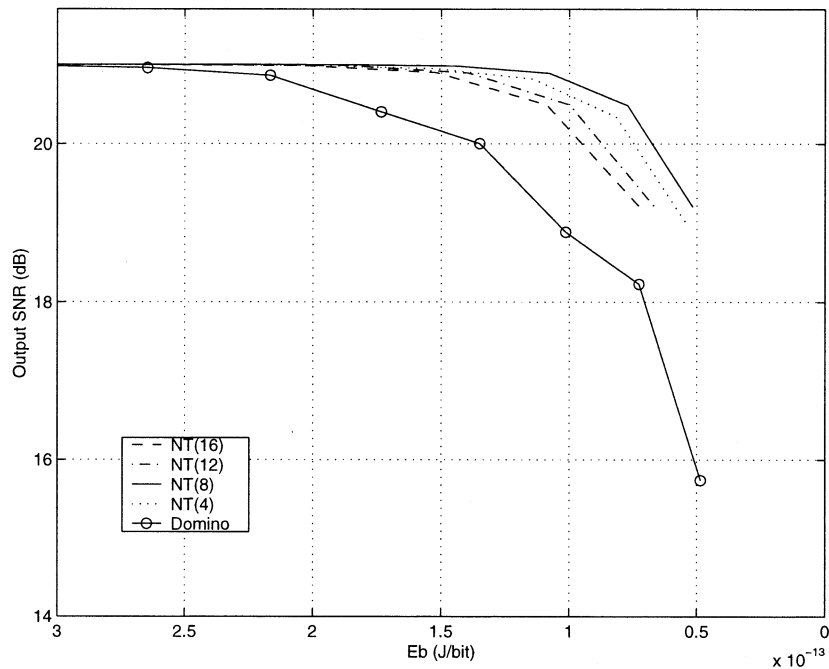
reliable operation. For the proposed technique, NT(8) consumes the minimum amount of energy indicating that it is the optimum in terms of energy-efficiency.

Fig. 12(b) plots the energy per information bit  $E_b$  at the specified SNR of 20 dB for different implementations along with the lower bounds. The lower bounds were computed by modeling Manchester adders as a multi-input multi-output SDC model and employing Theorem 2. First, we observe that noise-tolerant designs reduce energy dissipation by 31.2%–51.4% over conventional systems. This is due to the fact that any improvement in noise-immunity makes it easy to achieve reliable operation at low supply voltages, thereby, improving the energy-efficiency in the presence of noise. Second, the actual energy dissipation for conventional domino system is  $5.3\times$  above its lower bound while that for NT(8) is only  $2.8\times$  above the bound. This is an improvement by a factor of  $1.9\times$ . Finally, we observe that the overhead due to noise-tolerance starts to dominate in NT(12) and NT(16), which offsets the improvement in noise-tolerance and the resulting improvement in energy-efficiency.

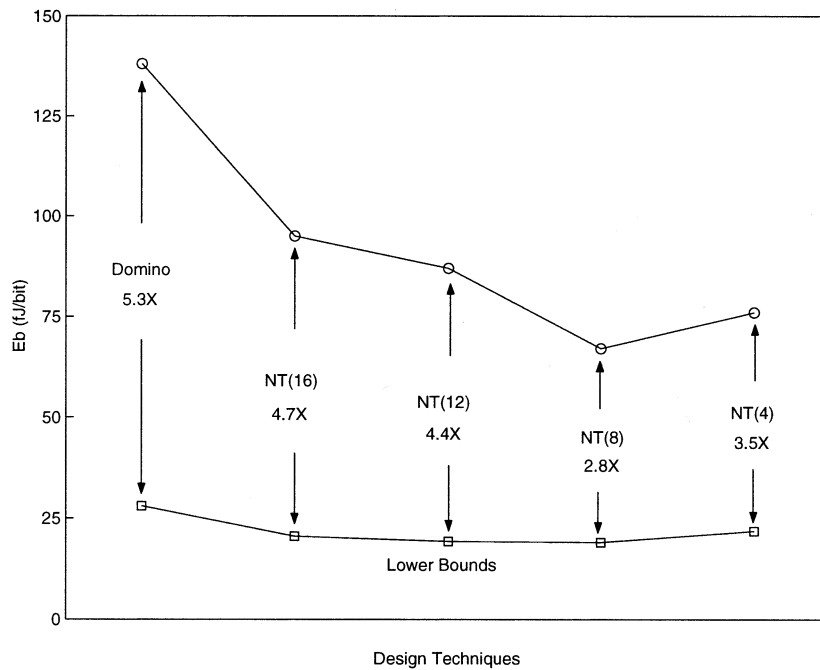
It is worth mentioning that algorithmic noise-tolerance (ANT) techniques [13], [14] can be employed concurrently with circuit-level noise-tolerant techniques [NT(8) in this case] in order to further improve the energy-efficiency.

## VI. CONCLUSION

An algorithm has been presented in this paper for deriving the lower bounds on energy dissipation of noisy digital systems.



(a)



(b)

Fig. 12. Performance of the proposed noise-tolerant scheme. (a) Energy dissipation versus algorithmic performance and (b) in comparison with the lower bounds.

These bounds are obtained by modeling digital systems as a SDC and employing information-theoretic considerations. We have shown that noise-tolerant dynamic circuits offer the best trade off between energy-efficiency and reliability when operating in the presence of noise. Employing a 16-bit noise-tolerant Manchester adder in a CDMA receiver, we demonstrate a 31.2%–51.4% energy reduction, and also show that the lower bounds on energy for this receiver are 2.8 $\times$  below the actual energy consumed. Further, we show that noise-tolerance reduces the gap between the lower bounds and actual energy dissipation

by a factor of 1.9 $\times$ . The results presented in this paper are a continuation of our past work [8], [11] on developing an information-theoretic framework for deep submicron VLSI systems. The elements of this framework are consistent with the recommendation in [1] to view DSM VLSI systems as communication networks and to develop noise-tolerance techniques at the circuit, architectural and algorithmic levels.

Future work needs to be directed toward reducing the gap between the lower bounds and the actual power dissipation. One approach for achieving this goal is to employ the SDC

model for computing the bounds on energy-efficiency of complex VLSI systems, and to develop design methodologies based on a concurrent application of circuit [15], [16] and algorithmic noise-tolerance [13], [14] design techniques to approach these bounds.

#### APPENDIX A PROOF OF THEOREM 1

In this appendix, we prove that the mutual information  $I(Y_N = Y + N; X_0, X_1, \dots, X_{n-1})$  of an  $n$ -input, single-output noisy gate achieves the maximum when  $p_y \triangleq P(Y = 1) = 0.5$ .

From Lemma 1, we rewrite  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  as

$$\begin{aligned} I(Y_N; X_0, X_1, \dots, X_{n-1}) &= - \int_{-\infty}^{\infty} [(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) \\ &\quad \cdot \log_2(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) \\ &\quad - f_N(v) \log_2 f_N(v)] dv. \end{aligned} \quad (\text{A1})$$

Taking partial derivative of  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  with respect to  $p_y$ , we get

$$\begin{aligned} \frac{\partial I(Y_N; X_0, X_1, \dots, X_{n-1})}{\partial p_y} &= - \int_{-\infty}^{\infty} [(f_N(v - V_{dd}) - f_N(v)) \\ &\quad \cdot \log_2(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) \\ &\quad + \frac{1}{\ln 2} (f_N(v - V_{dd}) - f_N(v))] dv \\ &= - \int_{-\infty}^{\infty} (f_N(v - V_{dd}) - f_N(v)) \\ &\quad \cdot \log_2(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) dv \end{aligned} \quad (\text{A2})$$

where we utilize the fact that

$$\int_{-\infty}^{\infty} \frac{1}{\ln 2} [f_N(v - V_{dd}) - f_N(v)] dv = 0. \quad (\text{A3})$$

The maximum value of  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  is obtained at the point where

$$\frac{\partial I(Y_N; X_0, X_1, \dots, X_{n-1})}{\partial p_y} = 0. \quad (\text{A4})$$

From (A2), this implies

$$\begin{aligned} &\int_{-\infty}^{\infty} f_N(v - V_{dd}) \\ &\quad \cdot \log_2(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) dv \\ &= \int_{-\infty}^{\infty} f_N(v) \log_2(p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)) dv. \end{aligned} \quad (\text{A5})$$

Obviously, (A5) is satisfied if and only if  $p_y = 0.5$ . Furthermore, we have

$$\begin{aligned} &\frac{\partial^2 I(Y_N; X_0, X_1, \dots, X_{n-1})}{\partial p_y^2} \\ &= - \frac{1}{\ln 2} \int_{-\infty}^{\infty} \frac{(f_N(v - V_{dd}) - f_N(v))^2}{p_y f_N(v - V_{dd}) + (1 - p_y) f_N(v)} dv \leq 0. \end{aligned} \quad (\text{A6})$$

Thus,  $I(Y_N; X_0, X_1, \dots, X_{n-1})$  at  $p_y = 0.5$  is indeed a global maximum.

#### APPENDIX B PROOF OF THEOREM 2

In this appendix, we derive the optimum solution  $(t_0^{\text{opt}}, t_1^{\text{opt}}, \dots, t_{m-1}^{\text{opt}})$  for the energy optimization problem (19) and (20).

Consider the mutual information for the noisy system shown in Fig. 4

$$\begin{aligned} I_{Y_N; X}(V_{dd}, t_0, \dots, t_{m-1}, \sigma_{N_0}^2, \dots, \sigma_{N_{m-1}}^2) &\triangleq I(Y_{N_0}, \dots, Y_{N_{m-1}}; X_0, \dots, X_{n-1}) \\ &= H(Y_{N_0}, \dots, Y_{N_{m-1}}) \\ &\quad - H(Y_{N_0}, \dots, Y_{N_{m-1}} | X_0, \dots, X_{n-1}) \\ &= H(Y_{N_0}, \dots, Y_{N_{m-1}}) - H(N_0, \dots, N_{m-1}) \\ &\leq \sum_{i=0}^{m-1} (H(Y_{N_i}) - H(N_i)) \end{aligned} \quad (\text{B1})$$

where the equality is achieved if the  $m$  outputs  $Y_{N_0}, Y_{N_1}, \dots, Y_{N_{m-1}}$  are statistically independent. Note that for typical digital systems where  $n \geq m$ , there always exists certain input probabilities such that the outputs are independent of each other [23]. As will be shown, the lower bound on energy dissipation is obtained when the outputs are statistically independent.

We denote the mutual information  $(H(Y_{N_i}) - H(N_i))$  for the  $i$ th output as  $I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)$ . From Lemma 1,  $I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)$  is given by

$$\begin{aligned} I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2) &= - \int_{-\infty}^{\infty} [(t_i f_{N_i}(v - V_{dd}) + (1 - t_i) f_{N_i}(v)) \\ &\quad \cdot \log_2(t_i f_{N_i}(v - V_{dd}) + (1 - t_i) f_{N_i}(v)) \\ &\quad - f_{N_i}(v) \log_2 f_{N_i}(v)] dv \end{aligned} \quad (\text{B2})$$

where  $f_{N_i}(v)$  is the distribution function for the noise voltage  $N_i$ .

Consider the power dissipation  $P_{\text{tot}}$  as primarily consisting of dynamic power dissipation. Employing (B1) and (B2), we rewrite the optimization problem (19) and (20) as

$$\text{minimize: } \sum_{i=0}^{m-1} t_i C_{L_i} V_{dd}^2 f_c \quad (\text{B3})$$

$$\text{subject to: } \sum_{i=0}^{m-1} I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2) f_c \geq R. \quad (\text{B4})$$

This is a standard optimization problem that can be solved using the Lagrange multiplier method [24]. Define function  $\mathcal{J}(t_0, t_1, \dots, t_{m-1}; \lambda)$  as

$$\begin{aligned} \mathcal{J}(t_0, t_1, \dots, t_{m-1}; \lambda) &= \sum_{i=0}^{m-1} t_i C_{L_i} V_{dd}^2 f_c \\ &+ \lambda \left( R - \sum_{i=0}^{m-1} I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2) f_c \right) \end{aligned} \quad (\text{B5})$$

where  $\lambda$  is the a real-valued Lagrange multiplier. Differentiating  $\mathcal{J}(t_0, t_1, \dots, t_{m-1}; \lambda)$  with respect to  $t_i$ , we get

$$\begin{aligned} \frac{\partial \mathcal{J}(t_0, t_1, \dots, t_{m-1}; \lambda)}{\partial t_i} &= C_{L_i} V_{dd}^2 f_c - \lambda f_c \frac{\partial I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)}{\partial t_i}. \end{aligned} \quad (\text{B6})$$

Let (B6) be zero, we have

$$\begin{aligned} \frac{\partial I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)}{\partial t_i} &= \beta C_{L_i} V_{dd}^2, \quad i = 0, 1, \dots, m-1 \end{aligned} \quad (\text{B7})$$

where  $\beta = 1/\lambda$ . The optimum solution  $(t_0^{\text{opt}}, t_1^{\text{opt}}, \dots, t_{m-1}^{\text{opt}})$  to (B3) and (B4) is thus obtained from (B7) at  $\lambda = \lambda_{\min}$ , or equivalently  $\beta = \beta_{\max}$ , where the information-theoretic constraint (B4) is just met, i.e.,

$$\sum_{i=0}^{m-1} I_{Y_{N_i}; X}(V_{dd}, t_i^{\text{opt}}, \sigma_{N_i}^2) f_c = R. \quad (\text{B8})$$

This is because that  $\partial I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2)/\partial t_i$  is a monotonically decreasing function with respect to  $t_i$  [see (A6) with  $p_y$  replaced by  $t_i$ ]. From (B7), for any  $\beta < \beta_{\max}$  (or  $\lambda > \lambda_{\min}$ ),  $t_i > t_i^{\text{opt}}$  making  $\sum_{i=0}^{m-1} I_{Y_{N_i}; X}(V_{dd}, t_i, \sigma_{N_i}^2) f_c > R$ , which also leads to a larger power dissipation due to the increase in  $t_i$ .

## REFERENCES

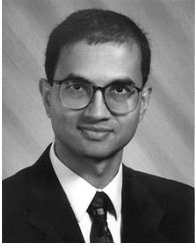
- [1] The 2001 International Technology Roadmap for Semiconductors [Online]. Available: <http://public.itrs.net/Files/2001ITRS/Home.htm>.
- [2] B. Davari, R. H. Dennard, and G. G. Shahidi, "CMOS scaling for high-performance and low power—The next ten years," *Proc. IEEE*, vol. 83, pp. 595–606, Apr. 1995.
- [3] R. Gonzalez, B. M. Gordon, and M. A. Horowitz, "Supply and threshold voltage scaling for low power CMOS," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1210–1216, Aug. 1997.
- [4] K. L. Shepard and V. Narayanan, "Noise in deep submicron digital design," in *Proc. '96 Int. Conf. Computer-Aided Design*, San Jose, CA, Nov. 1996, pp. 524–531.

- [5] P. Larsson and C. Svensson, "Noise in digital dynamic CMOS circuits," *IEEE J. Solid-State Circuits*, vol. 29, pp. 655–662, June 1994.
- [6] J. D. Meindl, "Low power microelectronics: Retrospect and prospect," *Proc. IEEE*, vol. 83, pp. 619–635, Apr. 1995.
- [7] E. A. Vittoz, "Low-power design: Ways to approach the limits," in *Proc. '94 IEEE Int. Solid-State Conf.*, San Francisco, CA, Feb. 1994, pp. 14–18.
- [8] N. R. Shanbhag, "A mathematical basis for power-reduction in digital VLSI systems," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 935–951, Nov. 1997.
- [9] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, pt. I, vol. 27, pp. 379–423, part II, pp. 623–656, 1948.
- [10] J. D. Meindl and J. A. Davis, "The fundamental limit on binary switching energy for terascale integration (TSI)," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1515–1516, Oct. 2000.
- [11] R. Hedge and N. R. Shanbhag, "Toward achieving energy-efficiency in presence of deep submicron noise," *IEEE Trans. VLSI Syst.*, vol. 8, pp. 379–391, Aug. 2000.
- [12] P. P. Sotiriadis, A. Chandrakasan, and V. Tarokh, "Maximum achievable energy reduction using coding with applications to deep sub-micron buses," in *Proc. of Intl. Symp. on Circuits and Syst.*, 2002, p. 85.
- [13] R. Hegde and N. R. Shanbhag, "Soft digital signal processing," *IEEE Trans. VLSI Syst.*, vol. 9, pp. 813–823, Dec. 2001.
- [14] L. Wang and N. R. Shanbhag, "Low-power AEC-based MIMO signal processing for Gigabit Ethernet 1000 Base-Transceivers," in *Proc. of Int. Symp. Low-Power Electronics Design (ISLPED)*, Huntington Beach, CA, Aug. 2001, pp. 334–339.
- [15] —, "An energy-efficient, noise-tolerant dynamic circuit technique," *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 1300–1306, Nov. 2000.
- [16] G. Balamurugan and N. R. Shanbhag, "The twin-transistor noise-tolerant dynamic circuit technique," *IEEE J. Solid-State Circuits*, vol. 36, pp. 273–280, Feb. 2001.
- [17] J. J. Covino, "Dynamic CMOS circuits with noise immunity," U.S. Patent 5 650 733, 1997.
- [18] G. P. D'Souza, "Dynamic logic circuit with reduced charge leakage," U.S. Patent 5 483 181, 1996.
- [19] R. H. Krambeck, C. M. Lee, and H.-F. S. Law, "High-speed compact circuits with CMOS," *IEEE J. Solid-State Circuits*, vol. 17, pp. 614–619, June 1982.
- [20] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*. New York: Academic, 1984.
- [21] J. M. Daga and D. Auvergne, "A comprehensive delay macro modeling for submicrometer CMOS logics," *IEEE J. Solid-State Circuits*, vol. 34, pp. 42–55, Jan. 1999.
- [22] S. M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*. New York: McGraw-Hill, 1996.
- [23] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [24] D. P. Bertsekas, *Nonlinear Programming*. Boston, MA: Athena Scientific, 1995.
- [25] G. A. Katopis, "Delta-I noise specification for a high-performance computing machine," *Proc. IEEE*, vol. 73, pp. 1405–1415, Sept. 1985.
- [26] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [27] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.
- [28] S. Ramprasad, N. R. Shanbhag, and I. N. Hajj, "Signal coding for low-power: Fundamental limits and practical realizations," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 923–929, July 1999.



**Lei Wang** (M'01) received the B.Eng. and M.Eng. degree from Tsinghua University, Beijing, China, in 1992 and 1996, respectively, and the Ph.D. degree from the University of Illinois at Urbana-Champaign, Urbana, IL, in 2001.

In summer 1999, he was with Microprocessor Research Labs, Intel Corporation, Hillsboro, OR, where his work involved the development of high-speed and noise-tolerant VLSI design techniques. In 2001, he joined Hewlett-Packard Microprocessor Design Labs, Fort Collins, CO. His current research interests include design and implementation of low-power, high-speed, and noise-tolerance VLSI systems.



**Naresh R. Shanbhag** (M'93–SM'93) received the B.Tech. degree from the Indian Institute of Technology, New Delhi, India, the M.S. degree from Wright State University, Dayton, OH, and the Ph.D. degree from the University of Minnesota, Minneapolis, all in electrical engineering in 1988, 1990, and 1993, respectively.

From July 1993 to August 1995, he was with AT&T Bell Laboratories, Murray Hill, NJ, where he was responsible for the development of VLSI algorithms, architectures, and implementation of broadband data communications transceivers. In particular, he was the lead chip architect for AT&T's 51.84 MB/s transceiver chips over twisted-pair wiring for asynchronous transfer mode (ATM)-LAN and broadband access chip-sets. Since August 1995, he has been with the Department of Electrical and Computer Engineering, and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign (UIUC), where he is presently an Associate Professor and the Director of the Illinois Center for Integrated Microsystems. At UIUC, he founded the VLSI Information Processing Systems (ViPS) Group, whose charter is to explore issues related to low-power, high-performance, and reliable integrated circuit implementations of broadband communications and digital signal processing systems spanning the algorithmic, architectural and circuit domains. He has published more than 90 journal articles, book chapters, and conference publications in this area and holds three U.S. patents. He is also a coauthor of the research monograph *Pipelined Adaptive Digital Filters* (Norwell, MA: Kluwer, 1994).

Dr. Shanbhag received the 2001 IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS Best Paper Award, the 1999 IEEE Leon K. Kirchmayer Best Paper Award, the 1999 Xerox Faculty Award, the National Science Foundation CAREER Award in 1996, and the 1994 Darlington Best Paper Award from the IEEE Circuits and Systems Society. From July 1997 to 2001, he was a Distinguished Lecturer for the IEEE Circuits and Systems Society. From 1997 to 1999, he served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: ANALOG AND DIGITAL SIGNAL PROCESSING. He has served on the technical program committees of various international conferences.