

Lower Bounds on Power-Dissipation for DSP Algorithms

Naresh R. Shanbhag

ECE Dept./Coordinated Science Laboratory
University of Illinois at Urbana-Champaign,
Urbana, IL 61801. *shanbhag@uivlsi.csl.uiuc.edu*

Abstract

Presented in this paper is a fundamental mathematical basis for determining the lower bounds on power dissipation in digital signal processing (DSP) algorithms. This basis is derived from information-theoretic arguments. In particular, a digital signal processing algorithm is viewed as a process of information transfer with an inherent information transfer rate requirement of R bits/sec. Different architectures implementing a given algorithm are equivalent to different communication networks each with a certain capacity C (also in bits/sec). The absolute lower bound on the power dissipation for any given architecture is then obtained by minimizing the signal power such that its channel capacity C is equal to the desired information transfer rate R . The proposed framework is employed to determine the lower-bounds for simple digital filters. Furthermore, lower bounds on the power dissipation achievable via adiabatic logic is also presented thus demonstrating the versatility of the proposed approach.

1 Introduction

In recent years, the need for implementing highly complex algorithms within a limited power budget has forced researchers to explore various methods for power-reduction [1]. These methods are applicable at various levels of the VLSI design methodology including architectural, logic and circuit level. However, the interaction between these techniques is highly complex and there exists a need for a unifying basis for power-reduction. A desirable characteristic of such a basis is that it should be applicable at all levels of the design methodology including the algorithmic level.

Most of the work done in the area of low-power design can be placed into the following two broad categories: 1.) Development of power reduction techniques and 2.) investigating the lower bounds on power dissipation. Work done in Category 1 can be viewed as a collection of techniques to reduce power at all possible levels of the VLSI design methodology. These techniques include pipelining and parallel processing [1-2] (architectural level),

logic minimization [3], adiabatic computation [4-5], optimal coding [6] and device threshold voltage (V_t) reduction (device level) [7].

Some related work includes determining the lower bounds on the achievable power dissipation [8-9]. In [8], the lower bound on power dissipation per pole for analog circuits was presented. This bound is based upon the desired signal-to-noise ratio (SNR). Empirical lower bound estimates for digital circuits were also presented in [8], which was based upon a desired SNR , estimates of gate complexity and energy per operation. The theory presented in this paper can provide the desired SNR for a given algorithm. Employing thermodynamic arguments [9], the thermal noise power spectrum was determined to be kT and the minimum energy required for a logic change was shown to be lower bounded by approximately $4kT$. The lower bounds calculated via the proposed theory is also a function of the noise power spectrum of the implementation media.

In this paper, we employ an information-theoretic approach to develop a mathematical basis for determining the lower-bounds on power-dissipation for DSP algorithms. Information-theoretic measures have been employed in the past to define a measure of *computational work* of a boolean transformation [10-11]. This measure of computational work is closely related to the area [12] occupied by an implementation of the boolean function. Elegant approaches to power estimation at the register-transfer level (RTL) have been proposed [13-14], which incorporate concepts such as *entropy*.

Our work differs significantly in that we begin with an *implementation-independent* view of an algorithm and then proceed to determine the lower-bounds on power dissipation for a given architecture. This is done via the introduction of the information transfer rate R (an implementation-independent quantity) and the channel capacity C (an architecture and technology-dependent quantity). The proposed basis has two advantages: 1.) it allows us to derive lower bounds on the power dissipation in DSP algorithms and 2.) it enables us to unify existing power-reduction techniques under a common framework. We will illustrate 1.) via examples of simple digital filters and 2.), by determining the lower bounds on power dissipation for adiabatic logic [4-5]. However, it must be mentioned that with respect to 2.), the proposed theory has been employed to provide a common basis for architectural techniques such as pipelining and parallel processing [15].

2 Preliminaries

In order to provide the necessary background, we will review some basic information theoretic concepts in this section.

2.1 Entropy and Mutual Information

Consider a discrete source generating symbols from the set $S_X = X_0, X_1, \dots, X_{L-1}$ according to a probability distribution $Pr(X)$. A measure of the information content of this source is given by its *entropy* $H(X)$, which is defined as follows

$$H(X) = - \sum_{i=0}^{L-1} P_i \log_2(P_i), \quad (2.1)$$

where $P_i \stackrel{\text{def}}{=} Pr(X = X_i)$ for $i = 0, \dots, L-1$ and $H(X)$ is in bits. The *mutual information* $I(X; Y)$ between the input X and the output Y (of a channel) is defined as

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \quad (2.2)$$

where $H(X|Y)$ is the *conditional entropy* of X conditioned on Y . The mutual information $I(X; Y)$ can be viewed as the *reduction in uncertainty* in X due to the knowledge of Y .

2.2 Information Transfer Rate

The reduction in uncertainty (by an amount $I(X; Y)$) in X is due to the information transferred from the input of the channel to its output. Thus, the *information transfer rate* R is defined as

$$R = f_{op} I(X; Y), \quad (2.3)$$

where f_{op} is the rate at which the symbols are generated by the source.

2.3 Channel Capacity

In his seminal work [16], Shannon showed that the capacity (C) of a channel band-limited to frequency W , is given by

$$C = \int_0^W \log_2[1 + SNR(f)] df, \quad (2.4)$$

where C is in bps. From (2.4), it is clear that the capacity C depends upon the SNR and the transmission bandwidth W .

3 A Fundamental Basis for Power-Reduction

In this section, we will first show that all logic transformations have an inherent information transfer rate requirements R associated with them. Next, we will present a theorem which allows one to determine the lower bounds on the power dissipation for a given architectural implementation.

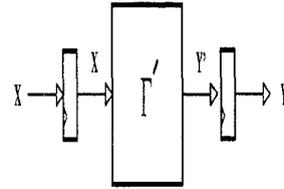


Figure 1: A noisy transformation.

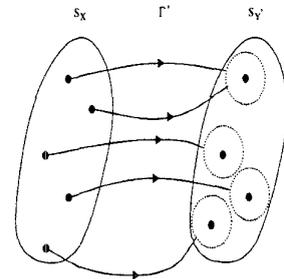


Figure 2: Input-output mapping for Γ' in Figure 1.

3.1 Logic Transformation

We can represent any noisy logic transformation as shown in Fig. 1, where noise could have many sources including the implementation media itself. Without any loss of generality, we assume that the inputs and the outputs are latched synchronously.

The definition of Γ' is shown in Fig. 2, where the input space S_X is mapped onto the output space S_Y . The dark dots in the set S_X represent the discrete values that the input X can assume, while the ones in the set S_Y denote the values that the output can assume if the noise power were zero.

Assuming that the noise probability density function is identical for all possible noiseless outputs \bar{Y} , then we can represent the system in Fig. 1 as shown in Fig. 3 with the corresponding mapping for Γ as shown in Fig. 4. In this figure, all the noise in Γ' has been referred to the output and we now have a noiseless transformation Γ mapping the input space S_X to a noiseless output space S_Y . In this paper, we will assume that the Fig. 3 can be employed to represent any noisy transformation Γ' . Note that the computation model in Fig. 3 is quite similar to a generic digital communications system. The input and output latches can be viewed as transceivers (transmitter-receiver) and the clock as providing the correct sampling epoch (output of a timing-recovery block). In a digital communications system, the transformation Γ is usually an identity transformation. Thus, the model in Fig. 3 allows us to apply results from digital communications theory [16] to determine the information transfer rate requirements for a given DSP algorithm.

Theorem 1: If $\Gamma: \mathbf{Z}^L \rightarrow \mathbf{Z}^M$ is a deterministic mapping, where \mathbf{Z} is the set of integers, then the information transfer rate is given by

$$R = f_{op} H(Y). \quad (3.1)$$

Theorem 1 can be easily proved by observing that in a noiseless case $I(X; Y') = I(X; Y) = H(Y)$ and substi-

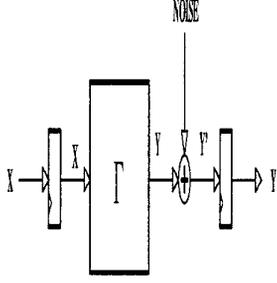


Figure 3: Another representation of a noisy transformation with noise referred to the output.

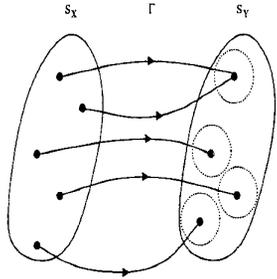


Figure 4: Input-output mapping for Γ in Figure 3.

tuting the result in (2.3). As an example, for a two-input AND gate with equiprobable inputs, (2.2) implies that $I(X; Y) = 0.8112$ bits. If such a system operates at 100 Mhz then $R = 81.12$ Mb/s.

Thus, all digital transformations, in particular linear finite-precision DSP algorithms have an inherent information transfer rate requirement R given by (3.1). This requirement is an inherent property of the transformation and is *independent of the implementation media or the architecture*.

3.2 Lower Bounds on Power Dissipation

It is well-known that there can be many different digital architectures which achieve the same functionality. In the present context, we view each of these architectures as a communication network with a certain capacity C . The capacity C for a point-to-point link is given by (2.4), where it can be seen that C is a function of: 1.) the transmission bandwidth W , 2.) the signal-to-noise ratio $SNR(f)$. In the context of a digital system, a connection between two logic gates can be viewed as a point-to-point link. This connection would have a certain resistance R_L and capacitance C_L (inductance L can also be considered if necessary). Hence to a first order approximation, the transmission bandwidth W is given by $1/(R_L C_L)$. The signal-to-noise ratio $SNR(f)$ is a function of the supply-voltage V_{dd} and the predominant sources of noise in the system.

A general digital network consists of numerous point-to-point links each connecting either combinational logic gates or synchronously clocked latches. We will not attempt to compute the capacity of such a general network as this is currently an open problem in information-theoretic literature. However, employing the abstraction indicated in Fig. 3, we can obtain an

equivalent noise source, and an equivalent resistance R_L and C_L , which are lumped at the output of the digital system. For example, the equivalent C_L could be the sum of the capacitances of the critical path of the architecture. Therefore, the transmission bandwidth W , $SNR(f)$ and therefore the capacity C can be determined.

Clearly, from [16], the capacity C should be greater than or equal to R for a meaningful computation to take place. We exploit this result to formally present the following theorem.

Theorem 2: For a given channel with the following properties

- *bandlimited to W Hz.*
- *with a noise power spectral density of $S_{NN}(f)$.*
- *a desired information transfer rate R bps.*
- *power dissipation at frequency f ($P_D(f)$) being related to the signal power spectrum $S_{XX}(f)$ as*

$$P_D(f) = F[S_{XX}(f)], \quad (3.2)$$

where $F()$ is a linear monotonically increasing function of its argument.

The lower bound on the power dissipation for such a channel is given by

$$P_D > F\left[\int_0^W (\nu_{min} - S_{NN}(f))^+ df\right], \quad (3.3)$$

where $(arg)^+$ denotes the positive part of arg and ν_{min} is a unique constant which can be obtained as a solution to the following equation

$$R = \int_0^W \log_2\left[1 + \frac{(\nu_{min} - S_{NN}(f))^+}{S_{NN}(f)}\right] df. \quad (3.4)$$

The proof of this Theorem 2 follows along similar lines as that of [16] and is omitted here for the sake of brevity.

Thus, for a given DSP transformation Γ , Theorem 1 allows us to calculate the information transfer rate R and Theorem 2 enables us to determine the lower bound on the power dissipation of a *particular architectural implementation* of Γ . In particular, Theorem 2 relates an algorithm level parameter R to a circuit level parameter SNR . This enables us to view the problem of power-reduction in a unified manner. It should be clear that the results obtained from the application of Theorem 1 and Theorem 2 would be a function of the algorithm (Γ), the technology ($S_{NN}(f)$ and $F(.)$) and the architecture (C).

Note that Theorem 2 does not provide us with the technique to achieve the lower bound. This is not surprising given the fact that Theorem 2 is derived from Shannon's joint source-channel coding theorem [16], which in turn provides a proof of achievability but not the method.

4 Lower Bound Calculations

In this section, we present four examples of the application of Theorem 1 and Theorem 2. The noise

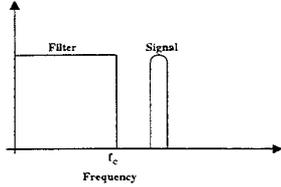


Figure 5: A degenerate case.

source in Fig. 3 needs to be determined for the system. However, for conventional digital systems the noise is mainly due to the phenomenon of ground bounce. Just for the sake of demonstration, and without any loss of generality, we assume that the implementation technology is CMOS with a flat noise spectrum with average power $\sigma_N^2 = 10^{-2} \text{ V}^2$ over a usable bandwidth of $W = 100 \text{ MHz}$. This is consistent with the value of ground bounce in a typical sub-micron CMOS technology. Furthermore, we can approximate the transmit pulse as a square-wave with an amplitude of $\pm V_{dd}/2$ volts for transmitting a '1' and '0', respectively. The signal power σ_X^2 (or the variance) is therefore equal to $V_{dd}^2/4$. The function F for static CMOS is defined as follows

$$F = C_L V_{dd}^2 2W, \quad (4.1)$$

where we assume a maximum of $2W$ channel uses per second. It can be shown that for a given value of R , a maximum of $2W$ channel uses would result in the lowest power dissipation.

4.1 Degenerate Case

Consider an ideal low-pass filter with a cutoff frequency f_c lower than the lower edge of the input signal spectrum (see Fig. 5). From Theorem 1 and (3.1), we get $H(Y) = 0$ and $R = 0$. Next, employing Theorem 2, we find that the lower bound on power dissipation $P_{D,min}$ is equal to zero. Clearly, the output of the low-pass filter is equal to zero. One particular implementation of such a system is an output line connected to ground. Other than non-idealities, such a system will not consume any power.

4.2 Filter 1

Let us consider an FIR filter with following transfer-function

$$H(z^{-1}) = z^{-1} + 0.5z^{-2}. \quad (4.2)$$

Further, assume that the sampling rate $f_s = 100 \text{ MHz}$ and that the input $x(n)$ to the filter is a 1-bit word, which is equally likely to be a '1' or a '0'. From (4.2), the output $y(n)$ depends upon the input vector $X(n) = [x(n-1), x(n-2)]$, which can take the values $[0,0]$, $[0,1]$, $[1,0]$ and $[1,1]$. The state sequence of shift-register is a Markov process and hence we can determine the steady-state value of the state probabilities $P(X(n))$. Assuming the initial state to be equally likely in any of the four possible values, we can show that the state probabilities are equal to 0.25 for all four states. This implies that the input entropy $H(X) = 2$ bits. From (4.2), we can see that every input state results in a unique output and hence the output entropy is also

equal to 2 bits. Hence, the information transfer rate $R = 200 \text{ Mb/s}$. A direct-form implementation of (4.2) would require one multiplier, one adder and two latches. Furthermore, we assume that the critical path delay for the multiply-add combination allows a 100 MHz operation (i.e., $W = 100 \text{ Mhz}$). For such an architecture, Theorem 2 can be employed to determine the lower-bound on the V_{dd} as follows

$$\begin{aligned} C &= R \\ W \log_2 \left[1 + \frac{V_{dd1,min}^2}{4\sigma_N^2} \right] &= R \\ V_{dd1,min} &= \left[\left(2^{\frac{R}{W}} - 1 \right) 4\sigma_N^2 \right]^{\frac{1}{2}}, \\ V_{dd1,min} &= 0.3464V, \end{aligned} \quad (4.3)$$

where $R = 200 \text{ Mb/s}$, $W = 100 \text{ MHz}$, and $\sigma_N^2 = 10^{-2} \text{ V}^2$. Note that the value of $V_{dd1,min}$ obtained via the proposed theory is in the same range as the 20 MHz encoder-decoder circuit in [7], where $V_{dd} = 0.2 \text{ V}$ was employed.

Substituting the value of (4.3) into (4.1) and with $C_L = 0.5pF$, we obtain the corresponding lower bounds on the power-dissipation as follows

$$P_{D,min} = 12\mu W. \quad (4.4)$$

4.3 Filter 2

In this example, we consider the following transfer function

$$H(z^{-1}) = 0.5z^{-1} - 0.5z^{-2}. \quad (4.5)$$

Assume all the other parameters to be the same as in Filter 1. Hence, the input entropy is $H(X) = 2$ bits. However, (4.5) is no longer a one-to-one mapping. It can be seen that the states $[0,0]$ and $[1,1]$ result in the same output $y(n) = 0$, while the other two states result in a unique output. This implies that the output entropy is $H(Y) = 1.5$ bits and from (3.1), the information transfer rate is $R = 150 \text{ Mb/s}$. Proceeding in a manner similar to that for Filter 1, we find that the lower bound on the supply voltage for a direct-form implementation is given by

$$V_{dd2,min} = 0.2704V. \quad (4.6)$$

Comparing (4.3) and (4.6), we find that the lower-bound on the supply voltage for Filter 2 is lower than that of Filter 1. This is due to the fact that Filter 1 has an information transfer rate requirements that is higher than that of Filter 2. The lower bound on the power dissipation is given by

$$P_{D,min} = 7.32\mu W. \quad (4.7)$$

A comparison of (4.4) and (4.7) shows that Filter 2 has a smaller lower bound than that of Filter 1. Note that this is a counter-intuitive result given the fact that there are two non-zero coefficients in a direct-form implementation of (4.5) as compared to one non-zero coefficient in the corresponding direct-form implementation of (4.2). However, by applying the *associativity* transformation [17], (4.5) can be implemented with one multiplier and one adder. The addition would be done first followed

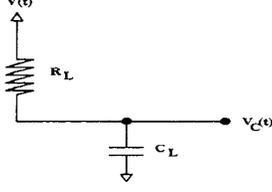


Figure 6: Adiabatic computation.

by one multiplication. Hence, the number of hardware units for a direct-form implementation of both filters would be identical. This example shows that the proposed theory does indicate the possibility of finding an architecture that would dissipate lesser power than the existing one. Hence it can be employed to compare the power dissipation characteristics of two algorithms under similar implementation constraints.

4.4 Adiabatic Logic

Adiabatic computation [4-5] and pulsed power-supply CMOS logic [18] work in a unique fashion. In particular, these techniques are based upon the fact that power dissipation can be minimized by ensuring that the voltage across any resistor is kept as small as possible. This is shown in Fig. 6, where the capacitor C_L is charged up by applying a voltage source $V(t)$, whose rise time $T_r \gg R_L C_L$. Under this condition, the power dissipation can *approach* zero. A lower bound on the energy dissipation for adiabatic logic has been calculated in [4-5] for a given value of T_r . In this subsection, we show that T_r itself is a function of the supply-voltage V_{dd} and the information transfer rate R . Hence, there is an upper limit for T_r and therefore a lower bound on the power dissipation. In practice, R_L is the source to drain resistance of a MOSFET, which is a function of the current. However, for the sake of simplicity and without any loss of generality, we will assume that R_L is a constant.

The power reduction property of adiabatic computation [4-5] and pulsed power-supply CMOS logic [8] can be explained via the information-theoretic framework presented in this paper. By applying a time-varying voltage source $V(t)$, adiabatic computation involves reducing the bandwidth of the transmit pulse and hence the transmission bandwidth W . From (2.4), we see that reducing W also reduces the capacity C of the channel. Furthermore, it can be shown that the power dissipated in an adiabatic computation is a function of W . For a desired information transfer rate R and supply voltage V_{dd} , there is a lower limit on W and hence on the power dissipation. The limit to which the power dissipation can be reduced is computable from Theorem 2. First assume a sinusoidal supply waveform $V(t) = V_{dd} \sin(2\pi Wt)$. The power in $V(t)$ is known to be $V_{dd}^2/2$. Employing (2.4) with a constant $SNR(f)$, we get

$$R < W \log_2 \left[1 + \frac{V_{dd}^2}{2\sigma_N^2} \right] \quad (4.8)$$

$$\frac{R}{\log_2 \left[1 + \frac{V_{dd}^2}{2\sigma_N^2} \right]} < W. \quad (4.9)$$

Next, we assume that $W \ll 1/(R_L C_L)$, so that the voltage across the capacitor $V_C(t)$ is given by

$$V_C(t) = V_{dd} \sin(2\pi Wt + \tan^{-1}(2\pi W R_L C_L)). \quad (4.10)$$

The power dissipation in Fig. 6 is given by

$$P_D(t) = \frac{[V(t) - V_C(t)]^2}{R_L}. \quad (4.11)$$

Substituting for $V(t)$ and $V_C(t)$ (from (4.10)) into (4.11), we obtain

$$P_D(t) = \frac{V_{dd}^2 \cos^2(2\pi Wt) [\tan^{-1}(2\pi W R_L C_L)]^2}{R_L}. \quad (4.12)$$

The lower bound on power dissipation for adiabatic computation is then obtained by averaging $P_D(t)$ over one period of $V(t)$ and is given by

$$P_{D,min,adb} = \frac{V_{dd}^2}{2R_L} [\tan^{-1}(2\pi W R_L C_L)]^2, \quad (4.13)$$

where W is the required transmission bandwidth, which is given by

$$W_{min,adb} = \frac{R}{\log_2 \left(1 + \frac{V_{dd}^2}{2\sigma_N^2} \right)}. \quad (4.14)$$

From (4.14), we see that for a given supply voltage V_{dd} and R , there is a lower limit on W . In fact, the relationship between V_{dd} and $W_{min,adb}$ is in inverse proportion. A lower bound on W implies a lower bound on P_D , which is given by (4.13). Unlike other methods of power reduction, the lower bound given by (4.13) is easily approachable. This is due to the fact that we have included all implementation constraints in the derivation of (4.13).

Note that (4.13) is not an absolute lower bound as it is a function of the supply voltage V_{dd} . In order to compute the absolute lower bound, we first substitute (4.14) into (4.13) to obtain

$$P_{D,min,adb} = \frac{V_{dd}^2}{2R_L} \left[\tan^{-1} \left(2\pi \frac{R}{\log_2 \left(1 + \frac{V_{dd}^2}{2\sigma_N^2} \right)} R_L C_L \right) \right]^2. \quad (4.15)$$

From (4.15), we notice that $P_{D,min,adb}$ is a monotonically decreasing function of V_{dd} for sufficiently small V_{dd} . This is because the function $\tan^{-1}(x)$ attains a constant value as x approaches infinity. Clearly, the fundamental lower bound would be obtained if the minimum value of V_{dd} could be determined. This minimum value is obtained by recognizing the fact that the channel is bandlimited to $1/(R_L C_L)$ Hz. Hence, the maximum value of $W_{min,adb}$ in (4.14) is equal to $1/(R_L C_L)$. Substituting $W_{min,adb} = 1/(R_L C_L)$ in (4.14) and solving for V_{dd} , we obtain the following lower bound on the supply voltage

$$V_{dd,min,adb} = [(2^{RR_L C_L} - 1) 2\sigma_N^2]^{1/2}. \quad (4.16)$$

Substituting (4.16) back into (4.15) we obtain the fundamental lower bound on P_D as follows

$$P_{D,min,adb} = \frac{(2^{RR_L C_L} - 1) 2\sigma_N^2 [\tan^{-1}(2\pi)]^2}{R_L}. \quad (4.17)$$

Clearly, from (4.17) the power dissipation for adiabatic logic will be zero only if the desired information transfer rate $R = 0$.

5 Conclusions

A mathematical basis for power-reduction in digital VLSI circuits was presented in this paper. The proposed basis was applied to determine the lower-bounds on power dissipation for DSP filtering algorithms and also adiabatic logic. This clearly indicates that the proposed approach can traverse a wide range of design abstraction from algorithms to circuits. For the sake of simplicity, we have assumed in this paper that a constant value of ground bounce voltage is the major noise contributor. Clearly, ground bounce is a function of signal power and the architecture. Hence, future work is being directed towards an accurate characterization of this and other noise sources in digital circuits. Approaching the lower bounds presented here requires a cohesive application of power-reduction techniques at all levels of the design hierarchy starting from algorithms, architectures, logic and circuits. We are currently in the processes of developing such a strategy employing the framework proposed in this paper.

References

- [1] M. Horowitz, T. Indermaur and R. Gonzalez, "Low-power digital design", *1994 IEEE Symposium on Low Power Electronics*, pp. 8-11, San Diego, Oct. 10-12.
- [2] A. P. Chandrakasan, and R. W. Brodersen, "Minimizing power consumption in digital CMOS circuits", *Proceedings of IEEE*, vol. 83, no. 4, pp. 498-523, April 1995.
- [3] J. Monteiro, S. Devdas, and A. Ghosh, "Retiming sequential circuits for low power," *IEEE Int. Conf. on CAD*, pp. 398-402.
- [4] W. C. Athas, L. J. Svensson, J. G. Koller, N. Tzartzanis, and E. Y.-C. Chou, "Low-power digital systems based on adiabatic-switching principles", *IEEE Trans. on VLSI Systems*, vol. 2, no. 4, pp. 398-407, Dec. 1994.
- [5] A. G. Dickinson and J. S. Denker, "Adiabatic dynamic logic", *Proceedings of the Custom Integrated Circuits Conference*, 1994.
- [6] A. P. Chandrakasan, S. Sheng and R. W. Brodersen, "Low power CMOS digital design", *IEEE J. of Solid-State Circuits*, vol. 27, pp. 473-484, April 1992.
- [7] J. B. Burr and J. Schott, "A 200mv self-test encoder/decoder using Stanford ultra-low power CMOS", *ISSCC '94*, pp. 84-85, San Francisco, CA.
- [8] E. A. Vittoz, "Low-power design: Ways to approach the limits", *ISSCC '94*, pp. 14-18, San Francisco, CA.
- [9] J. D. Meindl, "Low power microelectronics: Retrospect and prospect", *Proceedings of IEEE*, vol. 83, no. 4, pp. 619-635, April 1995.
- [10] L. Hellerman, "A measure of computational work", *IEEE Trans. on Computers*, vol. C-21, no. 5, pp. 439-446, May 1972.
- [11] R. W. Cook and M. J. Flynn, "Logical network cost and entropy", *IEEE Trans. on Computers*, vol. C-22, no. 9, pp. 823-826, Sept. 1973.
- [12] K-T Cheng and V. Agrawal, "An entropy measure for the complexity of multi-output Boolean functions", *27th ACM/IEEE Design Automation Conference*, Orlando, FL, pp. 302-305, June 24-28, 1990.
- [13] F. N. Najm, "Towards a high-level power estimation capability", *ACM/IEEE International Symposium on Low Power Design*, pp. 87-92, 1995.
- [14] D. Marculescu, R. Marculescu, and M. Pedram, "Information theoretic measures of energy consumption at register transfer level", *ACM/IEEE International Symposium on Low Power Design*, pp. 81-86, 1995.
- [15] N. R. Shanbhag, "A fundamental basis for power-reduction in VLSI circuits", *IEEE Intl. Symposium on Circuits and Systems*, Atlanta, GA, May 1996.
- [16] C. E. Shannon, "A mathematical theory of communications," *Bell System Technical Journal*, vol. 27, Part I, pp. 379-423, Part II, pp. 623-656.
- [17] K. K. Parhi, "High-level transformations for DSP synthesis", *Microsystems for Multimedia Applications*, ISCAS '95 Tutorial, Seattle, WA.
- [18] T. Gabara, "Pulsed power supply CMOS - PPS CMOS", *1994 IEEE Symposium on Low Power Electronics*, pp. 98-99, San Francisco, CA.