# Energy-Optimal Circuit Design

**Scott Hanson, Bo Zhai, David Blaauw, and Dennis Sylvester**
Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI, U.S.A.
{dmcs,hansons,bzhai,blaauw}@umich.edu

***Abstract -*** *Energy efficiency is an emerging metric for the quality of integrated circuit designs. Applications ranging from wireless sensor networks to RFID tags to embedded microprocessors require extremely low power consumption to maintain good battery life. We advocate the use of aggressively scaled supply voltages in such applications to maximize energy efficiency. This paper reviews our recent progress in mapping out the low energy design space including the presence of an energy-optimal supply voltage, and also touches on complications arising from variability at low supply voltages. We conclude with a survey of open research directions in the ultra-low voltage design space.*

**Keywords:** Integrated circuits, CMOS, low-power

## 1 Introduction

New classes of applications such as wireless sensor networks have dramatically different requirements from traditional integrated circuits (ICs) where performance (i.e., frequency) is the primary metric of interest. In this example, each node in the sensor network system must consume very little power to achieve a battery life of months or years. In addition, each node in the system must be fabricated at a negligible cost. Nodes should ideally be disposable and networks with thousands or millions of nodes should be achieved at a moderate cost. The cost of a single node can be measured in a number of ways, including the cost per unit area of silicon, design cost, and the cost of maintenance. It is therefore important that a design minimizes area and maximizes simplicity. For most such applications, performance (i.e., the time required to complete a given computation) is a secondary concern. For example, a sensor network deployed on an island off the coast of Maine monitored habitat conditions by measuring all sensors every 70 seconds [15].

The above discussion motivates the growing need for ultra low power (or more frequently, low energy, since this translates to battery life) ICs. In this paper we analyze the low energy design space and make conclusions about how such circuits should be designed. The models and guidelines developed from this work have been successfully used in the design of a robust ultra-low energy microprocessor [16] and a sub-200mV 6T SRAM [19], which are described briefly. We also highlight a number of open research directions that will enable ubiquitous reliable and energy-efficient systems.

## 2 Theoretical Foundations

Voltage scaling has emerged as one of the most effective techniques for meeting the increasingly stringent power demands in modern chip designs. A number of industrial designs have ventured as low as half of the nominal supply voltage. This voltage scaling has been the source of dramatic energy reductions but has left several lingering questions: How far can the supply voltage scale in conventional CMOS logic? Even if CMOS logic functions properly at very low supply voltages, is it worthwhile to operate at these voltages? We find in the subsequent discussion that aggressive voltage scaling into the subthreshold regime (i.e., $V_{dd} < V_{th}$) yields energy optimal circuits but is complicated by a high sensitivity to process-induced $V_{th}$ variability.

### 2.1 Energy Minimization

In [1], it was shown that CMOS gates composed of ideal transistors with a subthreshold swing of 60 mV/decade should function properly with a supply voltage as low as 36 mV. Despite a non-ideal subthreshold swing, measurements of an inverter show that functionality can be achieved with a supply voltage of just 65 mV [2]. It is clear that CMOS logic functions at extremely low voltages, but we must still consider the question of whether operation at these voltages is worthwhile. Figure 1(a) shows how the power consumed by an inverter chain scales with supply voltage. Total power consumption is broken into dynamic power (the power consumed by switching gates) and leakage power (the power consumed by idle gates). Minimum power is achieved by choosing the minimum functional supply voltage.

However, power is not always the most appropriate metric. For many applications, especially those in which battery life is the primary concern, energy per instruction may be a more sensible metric. There is a subtle but important difference between energy and power that is highlighted in Figure 1(b), which shows the energy consumed per switching event (which we call an operation) for the inverter chain from Figure 1(a).

Although Figure 1(a) shows that minimizing supply voltage will minimize power, the energy inflection point in Figure 1(b) shows that minimum energy is achieved at some voltage that is greater than the minimum functional supply voltage. This energy minimum is due to a rapid

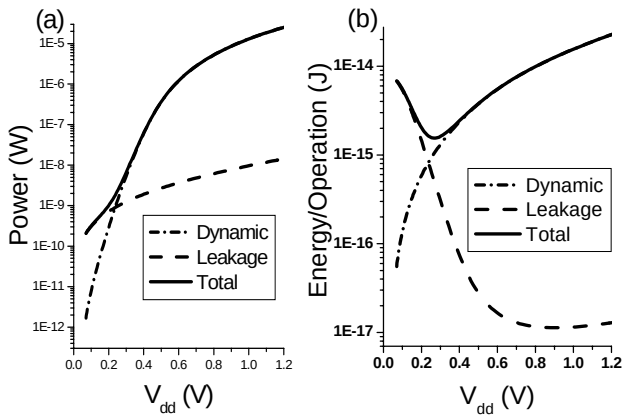


**Figure 1. (a) Power consumed by an inverter chain as a function of supply voltage ($V_{dd}$) (b) Energy consumed per switching operation by the same inverter chain as a function of $V_{dd}$.**

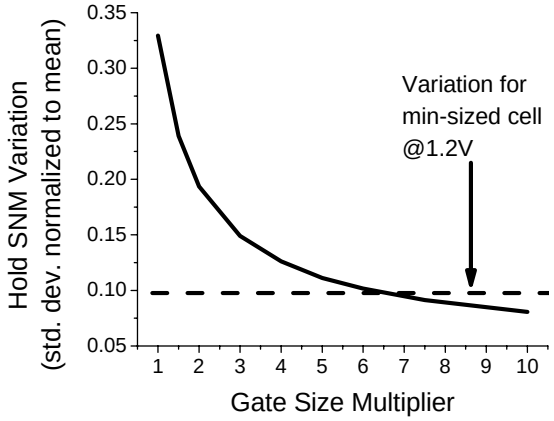


**Figure 2. Hold noise margin variation for a 6T SRAM cell at 0.3V. All transistor sizes in the cell are multiplied by an identical scaling factor.**

increase in gate delay as the supply voltage scales below the threshold voltage. As gate delay increases, the amount of time that each gate spends leaking also increases. As a result, the total leakage energy (the product of leakage current, supply voltage, and total leakage time) increases quickly and creates the minimum apparent in Figure 1(b). The location of this minimum energy supply voltage ($V_{min}$) is a strong function of both switching activity and logic depth (the number of gates between an input and an output) and was derived in [3] as:

$$V_{\min} = \left[1.587 \cdot \ln\left(\eta \cdot \frac{n}{\alpha}\right) - 2.355\right] \cdot m \cdot v_T \quad (1)$$

where $\alpha$ is the switching activity, $n$ is the logic depth, $m$ is the subthreshold slope factor, $v_T$ is the thermal voltage, and $\eta$ is a delay-related technology parameter.

## 2.2 Variability Considerations

While the energy benefits of low voltage operation are evident, it is not immediately clear that those benefits will exist after variability is considered. In addition to a reduction in absolute noise margins, low voltage operation suffers from a heightened sensitivity to $V_{th}$ variations. There are two fundamental concerns about variability in subthreshold circuits: 1) the robustness of subthreshold circuits, specifically SRAM, under variability and 2) the energy optimality of subthreshold circuits after accounting for delay variations.

The robustness of subthreshold SRAM has been the subject of intense research recently [7,19]. Random variations, which are largely caused by random dopant fluctuations, are most concerning for subthreshold SRAM. Mismatch between the cross-coupled inverters in the traditional 6T SRAM cell can lead to significant reductions in hold, read, and write noise margins. Monte Carlo simulations of a 6T SRAM cell in a 0.13µm process show that the variation in hold noise margins (as measured by $\sigma_{SNM}/\mu_{SNM}$) increases by 3.4X between $V_{dd}$=1.2V and $V_{dd}$=0.3V, greatly increasing the likelihood of a failure. The simplest solution to the random variability problem is to use larger gate sizes since random dopant fluctuations have an inverse dependence on the square root of gate area [4]. Figure 2 shows hold noise margin variation for a 6T SRAM cell as a function of gate size at $V_{dd}$=0.3V. For every point along the horizontal axis, the size of each transistor in the SRAM cell is multiplied by the same factor. At 0.3V, the transistor sizes must be increased by ~6.5X in order to achieve noise margin variation equal to that observed at 1.2V (relative to $V_{dd}$). Recent research has shown that robustness can also be recovered using alternative memory cell designs [7,19]. One such design will be described in the next section.

While robustness is the most immediate concern for subthreshold circuits, it is also important to investigate the energy implications of variation. In particular, we are interested in how the minimum energy point ($E_{min}$) and the minimum energy supply voltage ($V_{min}$, described by Equation 1) are affected by variation. Figure 3 plots $E_{min}$ and $V_{min}$ for an inverter chain of length $n$ under gate length and $V_{th}$ variability. Global $V_{th}$ variation is not included in this analysis, though should be considered in future work. For an inverter chain length of 16, $V_{min}$ increases by 58% and $E_{min}$ increases by 85%. We again note that random variability can be mitigated by increasing gate area, or alternatively, by "averaging" the variations over a longer inverter chain. In Figure 3, the $V_{min}$ penalty is reduced to 20% and the $E_{min}$ penalty is reduced to 46% when the inverter chain length is increased to 60, a vast improvement over the chain of length 16. A careful choice of $V_{min}$ in combination with judicious selection of logic depth, $n$, can help reduce energy under variability, but further work is undoubtedly necessary to maintain high yield under the dramatic variability expected in low voltage operation.

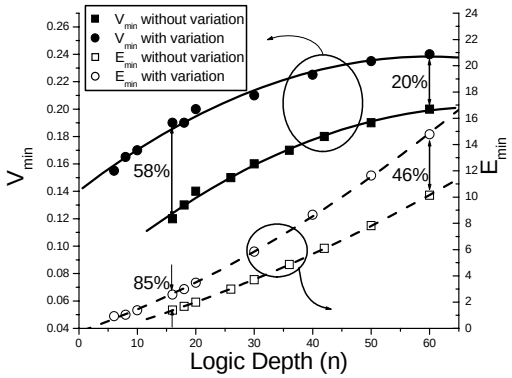


**Figure 3. Minimum energy, $E_{min}$, and the energy optimal supply voltage, $V_{min}$, for inverter chain of length *n*. Variation numbers are worst case (μ+3σ) [4].**

# 3 Design Examples

We now explore the topics of the last section further within the context of two test-chips: an 8-bit subthreshold processor for sensor applications and a 6T subthreshold SRAM.

## 3.1 A Subthreshold Sensor Processor

An 8-bit processor with 1.5kb instruction memory and 1kb data memory was fabricated in a 0.13μm process with $V_{th}$~400mV [16]. The processor targets mobile sensor applications where energy consumption is the primary metric and performance is a secondary concern. Memories were implemented using a robust latch-based memory [6]. Figure 4(a) shows energy and frequency measurements of the processor as functions of $V_{dd}$. At $V_{dd}$=350mV, energy consumption reaches a minimum of 3.52pJ/instruction at a frequency of 354kHz, giving an ~8X energy reduction as compared to normal superthreshold operation.

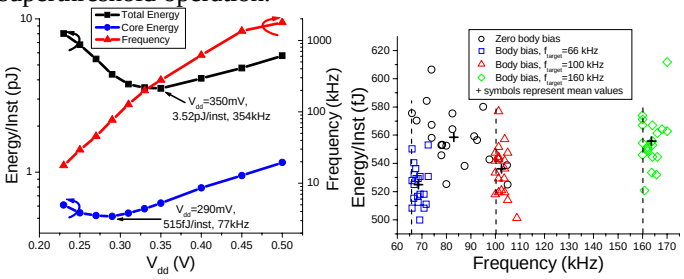


**Figure 4. (a) Energy and frequency for an 8-bit subthreshold processor [16]. (b) Energy and frequency for 20 dies under different body bias configurations.**

The 8-bit processor was also used to test the effectiveness of body biasing in addressing variability in subthreshold circuits. Due to the exponential sensitivity of subthreshold current to body bias, global current variability can be eliminated by selecting the proper body bias. Figure 4(b) shows energy and frequency for 20 measured dies under four body bias configurations. In the first configuration, no body bias is applied. In the remaining three configurations, body bias is applied uniquely to each die to target frequencies of 66kHz (the worst case frequency when no body bias is applied), 100kHz, and 160kHz. Frequency variations are eliminated by applying a body bias while energy variations are reduced significantly. Also note that the frequency can be tuned from 66kHz-160kHz with minimal energy penalties.

## 3.2 A Sub-200mV 6T SRAM

Body biasing can be used to compensate for global $V_{th}$ variability but is ineffective against random $V_{th}$ variability. As shown in the last section, robust low voltage SRAM design is particularly challenging under random variability. To investigate this, a 2kb SRAM array was fabricated in a 0.13μm process using the modified 6T SRAM cell shown in Figure 5 [19]. The single-ended design is inherently more robust to read upsets since noise is isolated to the single bitline. Additionally, the transmission gate can drive the bitline from rail-to-rail, eliminating the need for area-intensive sense-amplifiers. To decouple read and write operations and regain lost write margins, the power supply of the feedback inverter is gated during write operations. Transistor sizes were also increased to combat random dopant fluctuations. Measurements of the SRAM show that it remains functional below 200mV. Figure 5 shows that the proposed cell is far more energy efficient than a latch-based memory similar to the one described in Section 3.1.

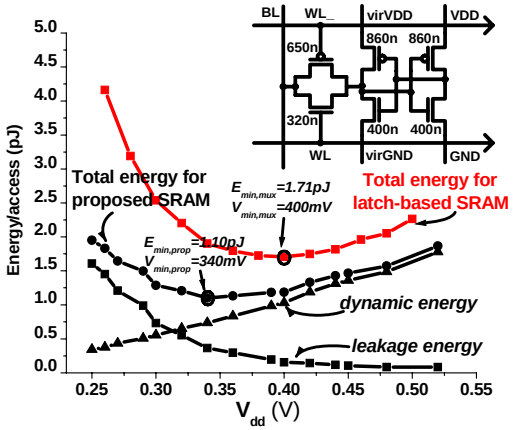


**Figure 5. Energy measurements for a modified 6T SRAM [19] and a latch-based memory are compared.**

# 4 Open Research Directions

Below we suggest several avenues for future exploration in the general area of ultra low-voltage design. In addition to dedicated low-power applications, subthreshold processing elements could form the basis of a heterogeneous massively multi-core processor architecture [17]. Each core would emphasize a different power-performance point and memory arrays would natively run at higher voltages since they are more energy-efficient in that space as well as more robust. This would allow for very high memory bandwidth relative to the processor speeds as well, effectively reversing the memory bottleneck commonly seen on large processors today [2].

Variability due to all sources, including process, voltage, and temperature (PVT) are all magnified in subthreshold as discussed earlier. There is a great need for a range of effective techniques to combat this variability. In particular, temperature-insensitive design approaches become an interesting topic of study; this may employ adaptive body bias [5]. Logic families other than CMOS may offer greater resiliency to certain variation sources such as voltage or process. This is another open area for exploration. Architectural approaches to variability are also vital – these may include a proliferation of traditional schemes such as error correction and redundancy, although the power costs of added hardware need to be weighed carefully against the improvements in robustness achieved.

Very low-power wireless communication schemes are needed; else the low energy budget of the digital processing component of a system will simply be swamped out by the communication requirements. This is highly application-specific; there may be cases where proximity communication schemes [9,10] are suitable and others where more traditional radios are used but with new architectures to address strict power budgets [8].

Extremely low duty cycle applications are very common – for instance environmental monitoring such as that mentioned in the introduction, or the monitoring of cracks along oil pipelines. All of these applications require a sensor to be read and data to be processed on a relatively infrequent basis (on the order of minutes). For such systems, standby mode power will completely dominate and current techniques to reduce standby mode leakage such as power gating are insufficient to provide required battery lifetimes. In these cases transistors should be employed as frugally as possible since added devices will inherently leak during standby. Novel architectures and low-power modes will need to be invented to enable sleep mode current levels below 1nA.

Subthreshold circuits can greatly benefit from a rethinking of device [11,18] and interconnect architectures. Since it may not be feasible to have a specific process variant dedicated to subthreshold, work is needed to find ways to tailor device and interconnect behavior to the unique needs of subthreshold design. Interconnect is interesting since parameters such as wire resistance become inconsequential compared to channel resistances. This may lead to a complete re-thinking of how clock and power routing are done, as well as how back-end stacks are manufactured (e.g., thick/wide wires are counterproductive since capacitance is the only important parasitic) [2].

Finally, given the possibility of nW processing there is an opportunity for on-die power sources that are CMOS-compatible. Most simply this could take the form of on-die thin-film batteries [12,13] that in the past have been impractical due to their limited energy capacities per unit area. These would be low cost and process compatible. Another option is energy scavenging through solar power, ambient vibration, or other techniques [14]. These power sources could also be used to recharge the primary battery, rather than supply power to the entire system.

# References


[1] J. Meindl and J. Davis, "The Fundamental Limit on Binary Switching Energy for Terascale Integration (TSI)," *IEEE J. Solid-State Circuits*, pp. 1515-1516 October 2000.

[2] S. Hanson, *et al*., "Ultralow-voltage minimum-energy CMOS," *IBM Journal of Research and Development*, pp. 469-490, July/Sept 2006.

[3] B. Zhai, D. Blaauw, D. Sylvester, K. Flautner, "Theoretical and Practical Limits of Dynamic Voltage Scaling," *Design Automation Conf.*, pp. 868-873, 2004.

[4] B. Zhai, S. Hanson, D. Blaauw, D. Sylvester, "Analysis and Mitigation of Variability in Subthreshold Design," *Intl. Symp. Low Power Electronics and Design*, pp. 20-25, 2005.

[5] S. V. Kumar, C. H. Kim, S. S. Sapatnekar, "Mathematically-Assisted Adaptive Body Bias (ABB) for Temperature Compensation in Gigascale LSI Systems," *Asia-South Pacific Design Automation Conf.*, pp. 559 – 564, 2006.

[6] A. Wang, A. Chandrakasan, "A 180mV FFT processor using subthreshold circuit techniques," *IEEE Intl. Solid-State Circuits Conf.,* pp. 292-293, 2004.

[7] B. Calhoun, A. Chandrakasan, "A 256kb subthreshold SRAM in 65nm CMOS," *IEEE Intl. Solid-State Circuits Conf.,* pp. 628-629, 2006.

[8] J. Chen, M.P. Flynn, J. Hayes, "A Fully Integrated Auto-Calibrated Super-Regenerative Receiver," *IEEE Intl. Solid State Circuits Conf.*, 2006.

[9] R.J. Drost, *et al.,* "Proximity communication," *IEEE J. Solid-State Circuits*, pp. 1529-1535, Sept. 2004.

[10] N. Miura, *et al*. ,"Analysis and design of inductive coupling and transceiver circuit for inductive inter-chip wireless superconnect," *IEEE J. Solid-State Circuits*, pp. 829-837, April 2005.

[11] A. Raychowdhury *et al*., "Computing with subthreshold leakage: device/circuit/ architecture co-design for ultralow-power subthreshold operation," *IEEE Trans. VLSI Systems,* pp. 1213-1224, Nov. 2005.

[12] J Harb, R LaFollette, R Selfridge, L Howell, "Microbatteries for self-sustained hybrid micropower supplies," *Journal of Power Sources*, 104 (1) (2002), 46 – 51

[13] C Dewan, D Teeters, "Vanadia xerogel nanocathodes used in lithium microbatteries," *Journal of Power Sources*, 119 (2003), 310 – 315

[14] J.A. Paradiso, T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, pp. 18-27, Jan-Mar 2005.

[15] R. Szewczyk, *et al*., "Lessons from a sensor network expedition," *European Workshop on Sensor Networks,* 2004.

[16] S. Hanson, *et al.*, "Performance and Variability Optimization Strategies in a Sub-200mV, 3.5pJ/inst, 11nW Subthreshold Processor," *IEEE Symp. on VLSI Circuits*, 2007.

[17] R. Kumar, *et al.*, "Single-ISA Heterogeneous Multi-Core Architectures: The Potential for Processor Power Reduction,'' *Int. Sym. Microarchitecture*, pp. 81–92, 2003.

[18] S. Hanson, M. Seok, D. Sylvester, and D. Blaauw, "Nanometer device scaling in subthreshold circuits," *ACM/IEEE Design Automation Conference*, pp. 700-705, 2007.

[19] B. Zhai, D. Blaauw, D. Sylvester, S. Hanson, "A sub-200mV 6T SRAM in 130nm CMOS, " *Int. Solid-State Circuits Conf.*, pp 332-333,2007.