Subthreshold FIR Filter Architecture for Ultra Low Power Applications

Biswajit Mishra and Bashir M. Al-Hashimi *

Electronic Systems and Devices Group,
School of Electronics and Computer Science,
University of Southampton, UK SO17 1BJ
{bm2,bmah}@ecs.soton.ac.uk
http://www.esd.ecs.soton.ac.uk

Abstract. Subthreshold design has been proposed as an effective technique for designing signal processing circuits needed in wireless sensor nodes powered by sources with limited energy. In this paper we propose a subthreshold FIR architecture which brings the benefits of reduced leakage energy, reduced minimum energy point, reduced operating voltage and increased operating frequency when compared with recently reported subthreshold designs. We shall demonstrate this through the design of a 9-tap FIR filter operating at *220mV* with operational frequency of *126kHz/sample* consuming *168.3nW* or *1.33pJoules/sample*. Furthermore, the area overhead of the proposed method is less than that of the transverse structure often employed in subthreshold filter designs. For example, a 9-tap filter based on transverse structure has $5\times$ higher area than the filter designed using our proposed method.

Key words: Subthreshold design, FIR, Minimum Energy Point, Ultra Low Power Design, Leakage

1 Introduction and Related Work

In wireless sensor nodes there is limited energy and therefore careful usage of the available energy is required. Subthreshold approach has been demonstrated to be effective in designing circuits with limited energy supply and therefore is receiving continuing attention from researchers interested in ultra low power design in particular wireless sensor networks and ubiquitous computing. The key to subthreshold design is the recent work reported by several authors which has already established the importance of leakage current contribution to the total power in subthreshold designs.

In [1] the authors have demonstrated that an optimal supply voltage $V_{optimal}$ exist below the threshold voltage V_T for maximum energy efficiency in subthreshold circuits. This occurs when the dynamic energy and leakage energy is comparable and is often referred to as the '*minimum energy point*'. Scaling the supply

voltage further below $V_{optimal}$ may result in correct circuit operation but doesn't necessarily improve energy efficiency because the leakage energy starts to dominate. Therefore the dynamic and leakage energy have two opposing trend in this region that gives rise to a minimum energy point at $V_{optimal}$. The subthreshold FFT design in [2] showed that the circuit can operate down to $V_{dd} = 180mV$ with very low operating frequency of $64Hz$, but the minimum energy point voltage is much higher than this minimal voltage and is reported at $350mV$ having operational frequency of $10kHz$. Transistor sizing that affects the energy consumption and the minimum energy point voltage is considered in the FIR design in [3]. The minimum energy point analysis through an analytical model for the delay and energy of an inverter chain in subthreshold circuits is discussed in [4]. The study showed that the minimum supply voltage $V_{optimal}$ for obtaining minimum energy point is dependent on several circuit parameters including transistor sizing, dynamic voltage scaling, threshold voltage scaling, body biasing and size of logic depth. The adaptive filter design in [5] proposed dynamic threshold voltage scaling approach to reduce leakage energy through substrate biasing. In [6] the improvement of leakage energy in subthreshold circuits was investigated by simultaneously scaling the supply voltage and threshold voltage.

One key application in wireless sensor nodes with limited energy supply is filtering and therefore the design of filter function has been considered in the recently reported subthreshold designs including [3] and [5]. In this paper we propose a subthreshold FIR architecture which brings the benefits of reduced leakage energy, reduced minimum energy point, reduced operating voltage and increased performance when compared with recently reported subthreshold designs. Our approach is based on reducing the number of transistors needed to implement a particular filter order. We will demonstrate the proposed architecture in the design of a 9 tap FIR filter. To the best of our knowledge this is the first study that shows improvement in leakage energy in the context of subthreshold design through reduced transistor count.

2   Minimum Energy Operation in Subthreshold Design

The total energy of CMOS a circuit is [4]:

$$E_{total} = N\alpha C_s V_{dd}^2 + \frac{N(1-\alpha)V_{dd}I_{off}}{f} \tag{1}$$

where N is the number of gates in the circuit, α is the average circuit switching activity, C_s is the switch capacitance of a single inverter, V_{dd} is the supply voltage, $I_{off} = I_o e^{-\frac{V_T}{mV_{th}}}$ is the off current, m is the subthreshold slope factor, V_{th} is the thermal voltage and f is the frequency of operation. The frequency of operation is $f = \frac{1}{L_{slow} \times t_{delay}}$ and depends on the number of inverters in the critical path (L_{slow}) and the delay of a single inverter (t_{delay}). In the above equation, V_{dd} can be scaled down to obtain the $V_{optimal}$ for the minimum energy point but is bound by a certain limit for the subthreshold operation [4]. The

$V_{optimal}$ can be found by expanding the terms and differentiating equation 1:

$$E_{total} = N\alpha C_s V_{dd}^2 + N(1-\alpha)V_{dd}I_o e^{\frac{-V_T}{mV_{th}}} t_{delay} L_{slow}$$
$$= N\alpha C_s V_{dd}^2 + N(1-\alpha)KC_s L_{slow} V_{dd}^2 e^{\frac{-V_{dd}}{mV_{th}}}$$

where K is a process dependent parameter. Differentiating with respect to V_{dd} gives:

$$\frac{\partial E_{total}}{\partial V_{dd}} = 2N\alpha C_s V_{dd} + 2N(1-\alpha)KC_s L_{slow} V_{dd} e^{\frac{-V_{dd}}{mV_{th}}}$$
$$-\frac{1}{mV_{th}}N(1-\alpha)KC_s L_{slow} V_{dd}^2 e^{\frac{-V_{dd}}{mV_{th}}} = 0 \qquad (2)$$

From equation 2 the first term contributes to the dynamic energy while the second and third term contributes to the leakage energy. Equating the above non linear equation 2 to zero and solving for the V_{dd} would provide the optimal supply voltage at $V_{dd} = V_{optimal}$ for the minimum energy point. A solution of this can be obtained by a curve fitting method. Our approach to obtain the minimum energy point is to reduce the number(N) of minimum sized ($W \times L$) transistors through the elimination of multipliers. In [1], it has already been established that the minimum energy point is dependent on α. It is shown that the $V_{optimal}$ occurs at a higher voltage when α is low because a low α gives a circuit more time to leak and the effective critical path becomes longer. A longer chain of gates in the critical path (L_{slow}) is also detrimental to the overall energy performance of the circuit as more gates are leaking relative to the dynamic energy. Reducing the transistor count will increase the switching activity (α or transistor utilization), hence the increased α can be used to reduce V_{dd} which leads to reduced overall energy. In the proposed filter, a short critical path (L_{slow}) is achieved through the elimination of multipliers. We will illustrate the effects of the above parameters (N, α, L_{slow}, V_{dd}) in our proposed FIR filter in sec 5.

3   Filtering

A key application for subthreshold wireless sensor node is physiological monitoring application where filtering and convolution is required. In [3], [5] and [7] the authors have reported how such functions can be implemented using subthreshold designs. A standard FIR realization often employed in subthreshold designs is the transversal structure depicted in Fig.1. The filter input $x(n)$ and output $y(n)$ is:

$$y(n) = \sum_{m=0}^{M-1} h(m)x(n-m) \qquad (3)$$

In the figure, the symbol z^{-1} is a delay of one sample or unit of time and is implemented using shift registers. The output sample $y(n)$ is the weighted sum

of the current input $x(n)$ and $(M-1)$ previous samples. The calculation of each of the output sample requires $(M-1)$ shift registers to store the $(M-1)$ input samples, M registers to store M coefficients, M multiplications and $(M-1)$ additions. Therefore, the critical path or delay of an M-tap filter would consist of one single multiplier and $[ceiling(\log_2 M)]$ number of adder delays. An example is the critical path of an 9-tap filter that consist of one multiplier and $[ceiling(\log_2 9)] = 4$ adder delays shown as dashed lines in Fig.1. It should be noted that the critical path of the multiplier consists of 15 full adder stage (tiny square boxes) as shown in Fig.1.
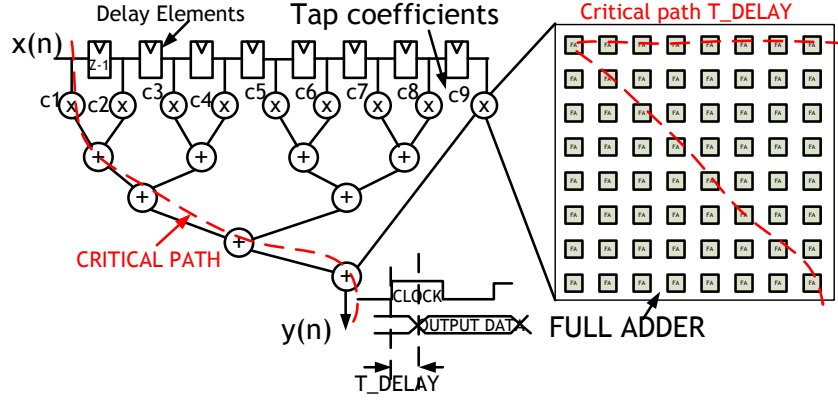

Fig. 1. FIR transverse architecture

3.1   Minimum Energy Point Analysis of Adders

As our method eliminates multipliers which consume significant power and also this leads to reduced critical path or delay. Since we discuss the derivation of the minimum energy point, both delay and power is important. We will demonstrate in Section 4 and in Section 5 that removing multipliers from the data path will have significant energy savings. As a result of eliminating the multipliers, the only key building block left in the proposed FIR structure (Fig.5) is the adders. We investigate the minimum energy point for different adders. To the best of our knowledge, no explicit investigation of obtaining $V_{optimal}$ and the minimum energy point for different adder topologies in the context of subthreshold design has been reported. We examine four adder circuits: Carry Look Ahead(CLA), Ripple Carry(RC), Carry Select(CS) and Carry Skip(CSK), for which minimum energy point is determined, using 0.13μ Berkeley Predictive Technology Models [8]. Fig.2 shows *hspice* simulation of the minimum energy point analysis of the adders as a function of V_{dd}. As it can be seen all adders have the minimum energy point within a $\pm5\%$ range of 200mV, and the CS adder has the minimum energy point (*i.e.* lowest energy consumption). This is explained as follows,

The carry select adder has the shortest critical path when compared with the other adders and its critical path consists of 4 full adders (one RCA-4) and 2 gates (AND, OR) as shown in Fig.3. For comparison the critical path of the
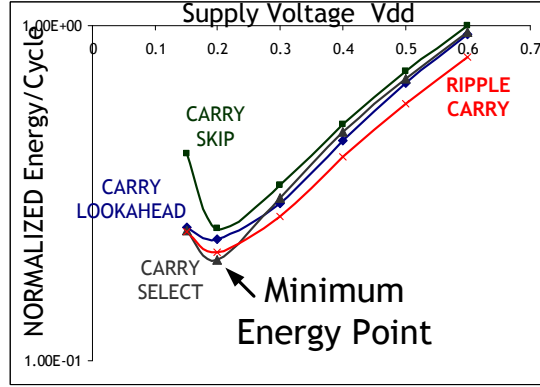
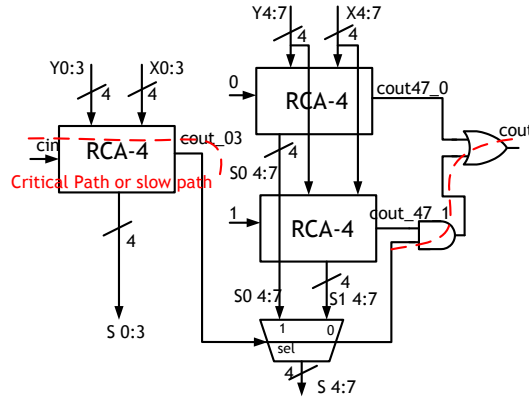**Fig. 2.** Minimum Energy Point of Adders

Fig. 3. Carry Select Adder

carry skip adder is shown in Fig.4. As shown in the dashed lines, the critical path of the carry skip adder is longer than the carry select adder since it consists of 2 full adder delays (one RCA-2) and 12 stages of 2-input gates (AND, OR). The overall delay or the critical path of the carry select adder contains 10, 2-input gates, whereas the carry skip adder has 16, 2-input gates. So, the carry select adder has a lower delay than the carry skip adder. The carry select adder also ensures that for any inputs most of the gates are switching during the circuit operation due to the two 4-bit ripple carry adder stage (RCA-4) for the most significant bit that has two carry inputs tied to '0' and '1'. From the simulations we observe that for the same set of inputs the average switching activity of the carry select adder is 1.3× more than that of the carry skip adder. Due to the higher switching of the gates the optimal voltage occurs at a much low voltage for the CS adder because the leakage energy is reduced and an improvement in overall energy is achieved. It should be noted that in designing the adders only two input gates with fan-out limited to three and minimum sized transistors were employed in order to reduce leakage energy and to avoid circuit failure [9].
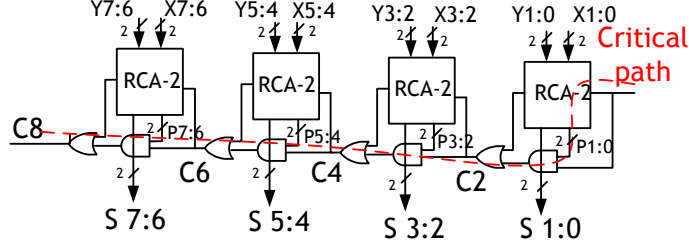
Fig. 4. Carry Skip Adder

4   Proposed FIR Architecture

The proposed FIR architecture is shown in Fig.1. As it can be seen, it consist of functional units (FU), adder stage and only one shift and accumulate stage without any multipliers. We have implemented a 9-tap filter and included the multiplexors after three delay stages (shaded region) in the FIR to show the added benefit of this FIR to be configured as a convolution filter often used in physiological monitoring applications [10]. Assuming the tap coefficient to be 8-bit wide, a standard M-tap transverse FIR filter equation 3 can be modified to:

$$y(n) = \sum_{m=0}^{M-1} h(m)x(n-m) = \sum_{m=0}^{M-1} \left[\sum_{k=0}^{7} x(n-m)h_k(m)2^k \right] \qquad (4)$$

The square term in equation 4 can be implemented by using shift registers and adders. The term h_k is a one bit data '0' or '1', and is the weight of the coefficient. The resulting architecture based on equation 4 contains same M number of shift-add-accumulate blocks as multipliers in conventional FIR (Fig.1). This can be simplified further [11] to equation 5 for area critical implementation resulting in the following:

$$y(n) = \sum_{k=0}^{7} \left[\sum_{m=0}^{M-1} x(n-m)h_k(m) \right] 2^k \qquad (5)$$

This results in area efficient architecture because the term inside the square bracket reduces from 16-bits to 8-bits. For a M-tap filter, a transverse filter with multipliers will contain $2 \times M$ shift registers, M multipliers and $(M-1)$ adders, while the proposed filter will contain $8 \times M$ AND gates, $16 \times M$ shift registers and $(M-1)$ adders. As shown in Fig.5, the 9-tap filter consists of nine functional units, an adder stage and one add-accumulate block.

As shown in Fig.5 the functional unit(FU) is the core of the architecture and is defined in the square bracket term in equation 5. Each FU is capable of one partial product. In every clock cycle, one 8-bit partial product is calculated. So a complete 8-bit sample would be delivered once in every eight clock cycle. The nine functional unit outputs 72-bits of partial product every clock cycle which is one eighth of the sample. The partial product of each of the functional unit is
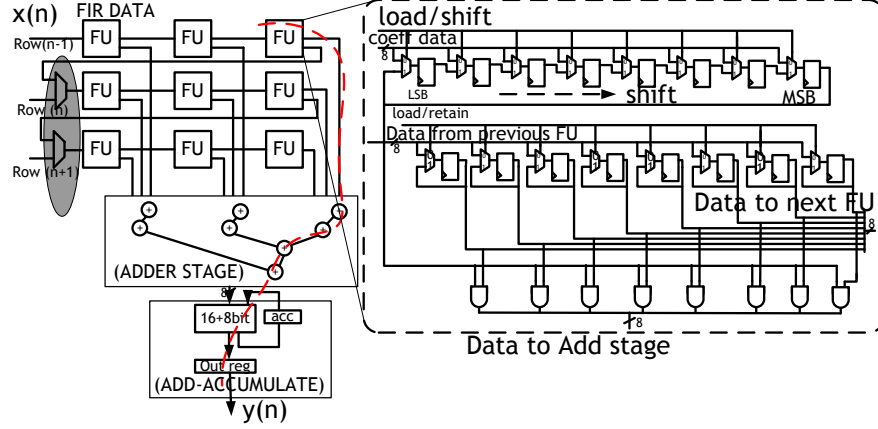
Fig. 5. Proposed FIR architecture

fed to the adder stages that sums up the nine partial product. The adder stages are 8-bit wide instead of 16-bit, which again reduces area. Coefficient bits are shifted left in each clock cycle so that the partial product is ANDed from most significant bit to least significant bit as shown in Fig.5. To avoid overflow, a 16-bit wider adder structure (with 8-bit half adder and a 8-bit full adder) in the shift, add-accumulate stage is implemented. The left shift in the accumulator and the add takes care of the weight associated with the left shift of the coefficient data. A shift operation is done in the accumulator by tying the least significant bit to '0' to adjust the weight of the coefficients. This process is continued 8 times till one filtered sample or convolved data is obtained. The new data is loaded after every eight clock cycles. A simple 8-bit shift register is implemented to generate the control signal once every 8 clock cycles for loading or shifting of the input data. The critical path (or longest path) of the design is the dotted line marked in the Fig.5 which is clearly shorter than the transverse structure. We assume that the data input is done directly and completely avoids any buffering stages in the FIR.

5 Results and Discussion

To validate the efficiency of the proposed architecture, we have designed two 9-tap filters; one is based on the proposed architecture (Fig.5), and is denoted as Design 1 and the other denoted as Design 2 based on the transverse structure with multipliers (Fig.1) which has also been employed in recently reported subthreshold filters [3]. Both designs were simulated using *hspice* with realistic transistor models from [8]. Apart from the minimum sized two input gates, the use of shift registers and associated flip flops for data buffering presents a significant problem because the flip flops fail to function below the threshold voltage. To mitigate this problem we have used the flip flop design discussed in [3]. In both designs, 8-bit wide input data and 8-bit coefficients were used. Fig.6

shows the minimum energy point analysis of both filters. As it can be seen, both filters can operate down to $V_{dd} = 150mV$ (points Ⓒ and Ⓓ). From the spice simulations the power obtained for Design 1 is $168.3nW$ and for Design 2 is $816.0nW$. Design 1 has lower minimum energy point at Ⓐ and happen at lower supply voltage ($220mV$) than Design 2 (Ⓑ, $275mV$). The reason why Design 1 outperforms Design 2 in terms of energy consumption is because of the following reasons: From the simulations we observe that the operating voltage increases as the switching activity decreases as expected [4]. This is because the ratio of the dynamic and leakage energy is proportional to the switching activity (α). A higher α will have a lower operating voltage V_{dd}, because the influence of leakage energy on the total energy will be minimal. From the spice simulations we observe that Design 1 has a higher utilization of the transistors and therefore has a higher average switching activity, $6\times$ than that of Design 2. This allows for a lower V_{dd} for the circuit to be operated resulting in lower dynamic energy. Also, due to a higher utilization of the transistor and due to the smaller critical path, fewer transistors are leaking and hence the leakage energy is low. The critical path of Design 1 has 60 gate delays whilst the Design 2 consists of 98 gate delays.
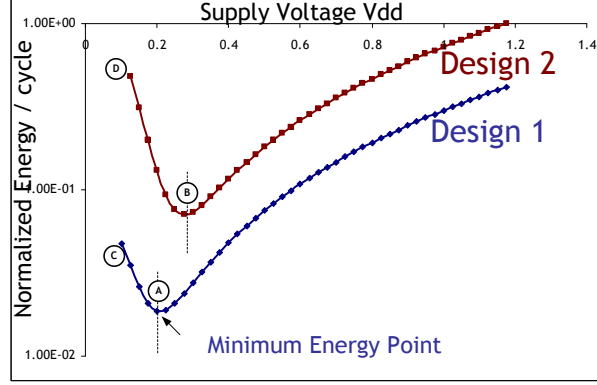

Fig. 6. Minimum Energy Point of two Filters

Fig.7 gives insight into the leakage and dynamic power consumption of both filter designs as function of V_{dd}. Again, as expected Design 1 have lower dynamic and leakage power components than Design 2. Fig.8 shows the delay performance of both filter designs as a function of V_{dd}. Design 1 has an operating frequency at $126kHz$ and Design 2 has an operating frequency at $100kHz$. As it can seen the filter designed using the proposed architecture exhibits better performance than Design 2. This is because Design 1 filter has much smaller critical path than that of Design 2 and is illustrated in Fig.1 and Fig.5 respectively. In summary, Fig.6, 7 and 8 clearly demonstrates that the proposed architecture produce filters with lower energy consumption ($1.33\frac{pjoule}{sample}$ at $220mV$) and better delay performance ($126kHz$) than Design 2 using the transverse structure with multipliers ($8.16\frac{pjoule}{sample}$ at $275mV$, $100kHz$). An 8-tap subthreshold filter reported in [3] operates at $V_{dd} = 250mV$ and $30kHz$ using the transverse structure shows further evidence.
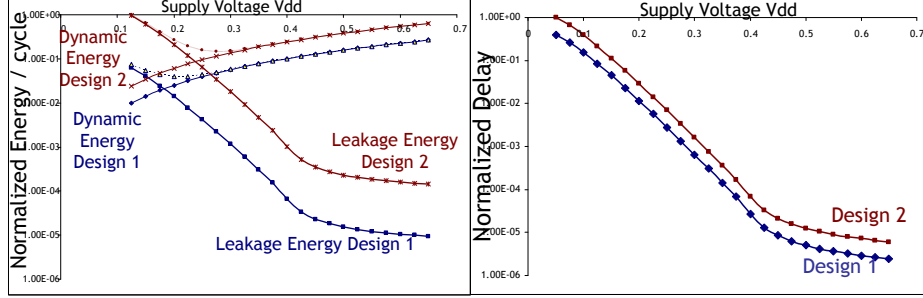
Fig. 7. Dynamic and Leakage Energy   **Fig. 8.** Delay Comparison of two Filters

It should be noted that the multipliers takes up considerable area and hence to the overall transistor count in an FIR up to 30-40% of the total and therefore reducing the multipliers will reduce the transistor count. As indicated earlier that the better energy and delay performance of the filters designed using the proposed architecture is achieved through the removal of multipliers from the filter architecture. This leads to significant reduction in transistor count. As it can be observed Design 1 has 144 shift registers (16 reg×9 FU) and 72 AND gates (8 gates×9 FU) whilst Design 2 has 16 shift registers and 9 multipliers. Table 1 gives the block count and the transistor count of the 9-tap filter (Design1). For example, 8, 8-bit carry select adders were needed, each has 91 gates, and a total of 362 transistors. The total transistor count of the 8 adders is 2896. Due to space limitations it is not possible to conclude the area overhead details of Design 2. But it can be stated that the overall transistor count is roughly $50k$ nearly $5\times$ higher than the proposed filter which consist of 9 multipliers, adder stage and the registers. The area cost of the proposed architecture is low compared with that of filters based on the transverse structure consisting of multipliers. For example, it was reported in [3] that the 8-tap subthreshold filter has $200k$ transistors, which is nearly $20\times$ higher than the proposed filter (Table 1).

Block	Circuit Blocks	Transistors
9 FU(2×8b Reg+8 AND)	9 × 640	5760
Add Stage(8×8b CSA)	8 × 362	2896
Control(1×8b Reg)	1 × 304	304
Adder(8b CSA+8×HA)	1 × 362 + 8 × 12	458
Accumulator (2× 16b Reg)	2 × 608	1216
2 MUX	2×14	28
Total Count of FIR	FIR	10,662

Table 1. Design 1 Filter Area Overhead.

6   Conclusions and Future Work

We have proposed an FIR filter architecture based on subthreshold transistor operation. The architecture generates filters with lower minimum energy points, and operates with lower V_{dd} and exhibits better delay performance than designs

obtained using the transverse structure that has been employed in previously reported subthreshold FIR filters. These energy and performance benefits have been achieved as a result of reducing the number of transistor count needed to implement the filtering function. This reduction in area overhead brings another benefit of the proposed filter architecture. We envisage a potential application for the proposed FIR filter architecture is to be part of DSP architectures aimed at wireless sensor nodes powered by limited energy sources.

The performance and stability of the subthreshold designs are greatly affected by Process, Voltage and Temperature variations. The effect on the circuit performance due to these variations will be studied further and is left as a future work.

References

1. B. Zhai, D. Blaauw, D. Sylvester, and K. Flautner, "Theoretical and Practical Limits of Dynamic Voltage Scaling," in *DAC '04: Proceedings of the 41st annual conference on Design automation.* New York, USA: ACM, 2004, pp. 868–873.
2. A. Wang and A. Chandrakasan, "A 180-mV Subthreshold FFT Processor Using a Minimum Energy Design Methodology," *IEEE Journal of Solid State Circuits*, vol. 40, no. 1, pp. 310–319, 2001.
3. B. Calhoun, A. Wang, and A. Chandrakasan, "Modeling and sizing for minimum energy operation in subthreshold circuits," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 9, pp. 1778–1786, Sept. 2005.
4. B. Zhai, S. Hanson, D. Blaauw, and D. Sylvester, "Analysis and Mitigation of Variability in Subthreshold Design," in *ISLPED '05: Proceedings of the 2005 international symposium on Low power electronics and design.* New York, USA: ACM, 2005, pp. 20–25.
5. H. Kim and K. Roy, "Ultra-Low Power DLMS Adaptive Filter for Hearing Aid Applications," in *ISLPED '01: Proceedings of the 2001 international symposium on Low power electronics and design.* New York, USA: ACM, 2001, pp. 352–357.
6. A. Wang, A. Chandrakasan, and S. Kosonocky, "Optimal Supply and Threshold Scaling for Subthreshold CMOS Circuits," *VLSI, 2002. Proceedings. IEEE Computer Society Annual Symposium on*, pp. 5–9, 2002.
7. R. Amirtharajah, J. Wenck, J. Collier, J. Siebert, and B. Zhou, "Circuits for Energy Harvesting Sensor Signal Processing," *Design Automation Conference, 2006 43rd ACM/IEEE*, pp. 639–644, 24-28 July 2006.
8. Y. Cao, T. Sato, M. Orshansky, D. Sylvester, and C. Hu, "New Paradigm of Predictive MOSFET and Interconnect Modeling for Early Circuit Simulation," *Conference 2006, IEEE Custom Integrated Circuits*, pp. 201–204, Jun 2000.
9. J. Kwong and A. P. Chandrakasan, "Variation-driven Device Sizing for Minimum Energy Sub-threshold Circuits," in *ISLPED '06: Proceedings of the 2006 international symposium on Low power electronics and design.* New York, NY, USA: ACM, 2006, pp. 8–13.
10. R. Amirtharajah, J. Collier, J. Siebert, B. Zhou, and A. Chandrakasan, "DSPs for Energy Harvesting Sensors: Applications and Architectures," *IEEE Pervasive Computing*, vol. 4, no. 3, pp. 72–79, 2005.
11. M. H. Sunwoo and S. K. Oh, "A Multiplierless 2-D Convolver Chip for Real-Time Image Processing," *Journal VLSI Signal Processing Syst.*, vol. 38, no. 1, pp. 63–71, 2004.