Ultra-Low-Power Digital Design with Body Biasing for Low Area and Performance-Efficient Operation

Maurice Meijer1,∗, Jose Pineda de Gyvez1,2, and Ajay Kapoor1
1 NXP Semiconductors, Eindhoven, The Netherlands
2 Department of Electrical Engineering, Technical University of Eindhoven, Eindhoven, The Netherlands

(Received: 15 June 2010; Accepted: 11 October 2010)

We present a design methodology towards minimum-area maximum-performance designs in sub-/near-threshold operation. Our methodology is based on a new metric called performance-per-area. Unlike conventional gate sizing, we use forward body biasing at synthesis time to render faster, smaller and more energy-efficient circuits. Our theory introduces body biasing into delay and energy models in the form of nonlinear derating functions that can easily be fitted to a technology node. The methodology is validated using an industrial microprocessor consisting of approximately 31 K gates and 3.7 K flip-flops in CMOS 90 nm. We obtain 4.2× better EDP, 3.8× higher speed and 9% smaller area than the non-body-biased counterpart.

Keywords: Sub-Threshold Circuits, Near-Threshold Design, Synthesis, Body Biasing.

1. INTRODUCTION

The quest for more energy-efficient integrated circuit solutions opened avenues for ultra-low-power enabling design styles. Reducing supply voltage (V_DD) is the most favored approach to achieve ultra-low-power. For very low speed and energy starving applications such as autonomous sensor nodes, RF-ID tags, and biomedical applications, sub-threshold operation became extremely popular. In sub-threshold operation, the V_DD is set below the threshold voltage (Vth), thereby minimizing energy consumption. The operation at minimum energy is achieved in sub-threshold regime, as illustrated in Figure 1.4–5 However, operating in sub-threshold comes at a large performance penalty: clock frequencies are limited to a few hundred kilohertz only. This has led to the fact the sub-threshold design is only for niche applications where performance is a second priority.

Many digital circuit applications require clock frequencies beyond a few hundred kilohertz. This motivated the introduction of near-threshold design, where V_DD is raised to a level just above the threshold voltage of transistors. In this way, circuit performance can be increased into the MHz-range while sacrificing minimum energy operation. Near-threshold design can achieve up to 10× higher energy-efficiency as compared to nominal V_DD operation. Typical applications benefiting from near-threshold operation are portable applications. To satisfy more performance-demanding applications, near-threshold design has been combined with circuit-level parallelism and wide-range voltage scaling. Although this approach demonstrated its effectiveness to achieve the required performance levels, the disadvantages are the increase in circuit area and the fact that not every circuit is suited to be parallelized.

A key issue for ultra-low-power circuits is the increased sensitivity of circuit performance to process parameter variations when operating at reduced V_DD supplies. Circuit design with minimum size devices was suggested to operate at the theoretically minimum energy point (MEP). However, device upsizing may be required to achieve operational robustness against process variability at the expense of a higher energy consumption and larger area occupation. Achieving process balance, e.g., an equal NMOS and PMOS on-current, has been shown to be essential for minimum energy operation. Moreover, industrial integrated circuits must be dimensioned such that specifications are met across the entire process window as defined by the foundry to ensure high production yield; this motivates a worst-case design approach. Researchers have proposed statistical circuit sizing as an alternative to worst-case design. Yet such approach has not totally found its way in industrial practices, because of, among other reasons, the moving average of process parameters, the fabrication of the same chip in multiple foundries,

∗Author to whom correspondence should be addressed. Email: maurice.meijer@nxp.com
Ultra-Low-Power Digital Design with Body Biasing for Low Area and Performance-Efficient Operation

Meijer et al.

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2
10^0 10^-1 10^-2 10^-3 10^-4

Fig. 1. Energy per cycle versus supply voltage for a generic digital circuit in 90 nm HVT LP-CMOS. MEP refers to the minimum energy point $V_{th} = 0.6$ V.

and the lack of appropriate EDA tools for statistical logic synthesis.

Multi-$V_{th}$ assignment during the design synthesis phase is a known problem. In practice, there is a small set of discrete $V_{th}$ values available in a given process technology. For example, the difference between a low-$V_{th}$ (LVT) and high-$V_{th}$ (HVT) in CMOS 90 nm is about 100 mV. Figure 2 illustrates the trade-off between energy and performance for a generic digital circuit design in 90 nm Low-Power (LP) CMOS. The design has been optimized for minimum energy while meeting a given clock frequency constraint under variable $V_{DD}$ and in case of two discrete $V_{th}$ values. As expected, there exists a MEP for both design cases. Observe that both $V_{th}$'s are suboptimal for a narrow range of performance. The MEP of the LVT design is located at a higher operating frequency than in case of the HVT design. This implies that LVT is the preferred choice for performance-constrained ultra-low-power designs but at a larger leakage overhead. The impact of process parameter variations on circuit performance remains an issue in multi-$V_{th}$ designs.

Finally, $V_{DD}$ scaling and body biasing have been proposed to counteract process variation in both strong-inversion and sub-threshold operation. Such approaches have been applied in the form of post-silicon circuit tuning. Body biasing offers several advantages over multi-$V_{th}$ design:

(i) it offers a continuum of $V_{th}$ values,
(ii) it can be used on top of any $V_{th}$ assignment,
(iii) it can be used for energy-performance tuning, and
(iv) it can be dynamically controlled.

Body biased designs can effectively reach LVT performance in active operation and HVT leakage in standby.

1.1. Contribution of This Work

It follows then that the state of the art either uses post silicon tuning for performance compensation, or multi-$V_{th}$ assignments for low energy operation. Still, these two approaches are mutually exclusive. In this work we propose a new design strategy that utilizes body biasing during the design synthesis of ultra-low-power digital circuits. We prove that our approach renders better circuit performance, smaller area, lower energy consumption, and on-line energy-delay product (EDP) control in the presence of process parameter variations. Thus, we cope with the shortcomings of speed and energy mentioned above. Moreover, we introduce a new metric that constrains area over-dimensioning in the search of an “optimum” energy-efficient design. Our focus is on low-energy performance-efficient design strategies applicable to consumer portable applications with target operating frequencies in the range of 1 MHz to 50 MHz.

1.2. Outline of the Paper

The rest of this paper is organized as follows. In Section 2 we introduce body biased low-power design. Section 3 presents the theoretical background and modeling. In Section 4 we explore the area, performance and power trends. Section 5 discusses the physical design aspects, body bias selection and generation. Section 6 presents synthesis results for an industrial microprocessor design. Finally, Section 7 presents our conclusions.
2. BODY-BIASED ULTRA-LOW-POWER DESIGN

In this section we introduce the concept of body bias driven design. Moreover, we present the effectiveness of forward body biasing (FBB) in 90 nm LP-CMOS. But first, the process variability influence on ultra-low-power circuit design will be discussed.

2.1. Circuit Design Under Process Variation

The energy consumption and the maximum performance of a circuit are sensitive to variations in process parameters. Figure 3 illustrates this sensitivity for minimum energy designs under a $V_{TH}$ scaling scenario. The analysis concerns a generic digital circuit design in 90 nm LP-CMOS. Each curve corresponds to a distinct process corner, for example slow, nominal, and fast process. A point on each curve corresponds to the lowest energy design that meets a given clock frequency constraint under variable $V_{DD}$. The energy and performance values have been normalized to the nominal $V_{DD}$ design at 1.2 V for a slow-process condition. Observe the increasing variation in clock frequency at reduced $V_{DD}$'s. While a frequency spread of 20% has been observed at $V_{DD} = 1.2$ V between slow and nominal process conditions, this spread increases to about 4x at $V_{DD} = 0.5$ V. Eventually, the spread saturates at very low $V_{DD}$'s. A MEP can be reached irrespective of process condition; however, it is located at different supply voltages and clock frequencies. The MEP is located at a $V_{TH}$opt of 0.36 V, 0.32 V and 0.31 V for slow, nominal and fast process conditions, respectively. The clock frequency at which the MEP occurs can differ by about one order of magnitude depending on process condition. At a given energy target, the clock frequency of the design is limited by the slow process corner. In this case, all transistors in the logic gates of the critical path are in the slow-process corner. Since practical systems are implemented such that the dynamic energy is much larger than the leakage energy, the slow process corner also gives the largest energy design at a given clock frequency due to a higher required $V_{TH}$. This motivates design optimization under slow process conditions for improving both performance and energy-efficiency of the design.

Figure 4 shows the conventional behavior of today's synthesis tools under a fixed $V_{DD}$ and $V_{TH}$ assignment. Essentially, conventional synthesizers upsize the circuit area to meet speed constraints. The area versus performance curves under different process conditions are plotted in Figure 4 when operating at $V_{DD} = 0.5$ V. The curves are constructed from a multitude of points, each corresponding to a design with unique speed requirements. As indicated before, the worst circuit performance relates to the slow process condition. Observe that circuit sizing is a weak parameter to increase performance for sub-threshold circuits. Instead, $V_{DD}$ scaling is more effective for performance increase. Therefore, the amount of circuit sizing for achieving higher performance or operational robustness against process parameter variations should be limited to avoid spending excessive area and energy.

2.2. Body Biased Driven Design Concept

Traditionally, digital CMOS circuits are implemented to meet timing specifications for slow process conditions. The circuit is dimensioned such that it can operate at a minimum possible $V_{DD}$ at which timing can just be
Ultra-Low-Power Digital Design with Body Biasing for Low Area and Performance-Efficient Operation

Meijer et al.

In this way, the total energy of the circuit is minimized. Observing that FBB enhances circuit speed, one can account for FBB at design-time to reduce circuit area and energy consumption at a given clock frequency. This creates opportunities for more cost-effective and lower energy solutions without sacrificing performance specifications and parametric yield, provided that the IC has FBB tuning capabilities. Figure 5 illustrates how FBB is utilized under body bias driven (BBD) design. In contrast to conventional design, the performance of the BBD design is optimized such that most chip samples operate at, or close to, the target clock frequency. FBB needs to be applied to slow die samples only for reducing the process-dependent performance spread. At design-time, the amount of FBB is chosen such that slow-corner samples achieve nominal-corner performance. Notice that a performance increase by FBB can also be traded-off against a performance decrease due to smaller area or lower energy circuit solutions. The amount of FBB required depends on circuit performance increase needed to achieve the target clock frequency. FBB is applied only to the speed-critical part of the chip design, which could consist of one or more digital functional (sub-)blocks. At silicon-time, the actual FBB for a given chip sample can be determined during production test, a chip calibration phase, or alternatively with a closed-loop approach using speed monitors. Like in conventional design styles, the impact of $V_{DD}$ variations and temperature changes on circuit behavior can be accounted for through guard-banding.

Figure 6 shows the energy and performance trade-off for BBD designs under $V_{DD}$ scaling in case of a generic digital circuit design 90 nm LP-CMOS. The indicated trend lines relate to slow-process corner conditions. The same normalization has been used as before. As way of example, two trade-off curves have been illustrated for BBD design, each with a different FBB value. Observe that BBD designs achieve lower energy per cycle than the conventional design over a wide $V_{DD}$ range to meet a given frequency target. This is because FBB enables lower $V_{DD}$ operation than the non body biased designs. Since BBD design makes use of FBB, one could expect a leakage energy penalty for BBD design. However, the impact of BBD design on leakage energy is rather limited primarily because of the increased operating frequency with FBB. Moreover, BBD design operates at similar or better circuit timing than the conventional design, resulting in better EDP figures. Finally, BBD design can effectively speed-up slow-process corner samples to achieve the operating performance associated to a nominal process-corner, as can be deduced by comparing Figures 3 and 6. The amount of FBB required to achieve such performance compensation is $V_{DD}$ dependent.

2.3. Body Bias Tuning Capabilities

The effectiveness of BBD design depends on the performance tuning range available with FBB. This tuning range depends on process technology, transistor threshold voltage option, and power supply voltage of the design. To quantify the effectiveness of FBB in 90 nm LP-CMOS, circuit simulations have been carried out on HVT ring-oscillator circuits. The circuit descriptions contain all layout parasitics including the bulk diffusion diodes. Slow process conditions and 25 °C operating temperature were applied. At these conditions, the NMOS and PMOS $V_{th}$ are at 0.6 V and 0.55 V, respectively. Our simulations used a $V_{DD}$ of 0.5 V and 0.7 V as a reference for sub-threshold and near-threshold designs, respectively. Using FBB, the $V_{th}$ is lowered and consequently moves the transistor’s biasing point towards the strong-inversion regime. The applied FBB voltage is limited to 0.5 V to avoid turning on the junction diodes. The FBB is applied symmetrically to PMOS and NMOS transistors, e.g., $V_{pwell} = V_{DD} - V_{nwell}$.

Fig. 5. Forward body bias utilization under body bias driven design.

Fig. 6. Energy per cycle versus clock frequency for a generic digital circuit in 90 nm HVT LP-CMOS under conventional and body bias driven design.

$V_{DD}$ operation than the non body biased designs. Since BBD design makes use of FBB, one could expect a leakage energy penalty for BBD design. However, the impact of BBD design on leakage energy is rather limited primarily because of the increased operating frequency with FBB. Moreover, BBD design operates at similar or better circuit timing than the conventional design, resulting in better EDP figures. Finally, BBD design can effectively speed-up slow-process corner samples to achieve the operating performance associated to a nominal process-corner, as can be deduced by comparing Figures 3 and 6. The amount of FBB required to achieve such performance compensation is $V_{DD}$ dependent.

2.3. Body Bias Tuning Capabilities

The effectiveness of BBD design depends on the performance tuning range available with FBB. This tuning range depends on process technology, transistor threshold voltage option, and power supply voltage of the design. To quantify the effectiveness of FBB in 90 nm LP-CMOS, circuit simulations have been carried out on HVT ring-oscillator circuits. The circuit descriptions contain all layout parasitics including the bulk diffusion diodes. Slow process conditions and 25 °C operating temperature were applied. At these conditions, the NMOS and PMOS $V_{th}$ are at 0.6 V and 0.55 V, respectively. Our simulations used a $V_{DD}$ of 0.5 V and 0.7 V as a reference for sub-threshold and near-threshold designs, respectively. Using FBB, the $V_{th}$ is lowered and consequently moves the transistor’s biasing point towards the strong-inversion regime. The applied FBB voltage is limited to 0.5 V to avoid turning on the junction diodes. The FBB is applied symmetrically to PMOS and NMOS transistors, e.g., $V_{pwell} = V_{DD} - V_{nwell}$.
From Table I, notice the large impact of FBB to increase circuit performance at the applied VDD levels. With 0.5 V FBB, the performance increases by ~16x for VDD = 0.5 V, and ~3x for VDD = 0.7 V. Contrarily, the respective leakage increase is ~436x and ~237x. At such large FBB, the leakage increase is because the transistor junctions become forward-biased. Obviously a judicious choice of FBB is needed. One advantage of using FBB is that it can be turned off in standby mode operation to avoid the large leakage penalty.

### 3. ANALYTICAL CIRCUIT MODELS

In this section we present analytical circuit models for energy, delay, and area, which account for the impact of FBB. Moreover, we introduce the performance-per-area metric for binding silicon area to circuit speed.

#### 3.1. Energy, Delay and Area Models

The total energy consumption of a digital gate can be modeled as the sum of the active (or dynamic) energy, and the leakage energy consumed during the clock period:

\[ E_{\text{gate}} = E_{\text{d}} + E_{\text{leak}} \]

\[ = (x C_{\text{inv}} + C_{\text{load}}) V_{\text{DD}}^2 + T_{\text{ak}} \cdot x \cdot I_{\text{leak}} \cdot s_{\text{i}} \cdot V_{\text{DD}} \]  

(1)

where \( x \) is the gate sizing factor (\( x \geq 1 \)), \( C_{\text{inv}} \) and \( C_{\text{load}} \) are the switching intrinsic and load capacitance of a gate, respectively, and \( T_{\text{ak}} \) is the operating clock period. The leakage current of the gate depends on gate size, \( V_{\text{DD}} \) and \( V_{\text{BB}} \); \( I_{\text{leak}} \propto e^{-V_{\text{BB}}/V_{\text{T}}/} \), where \( m \) is the sub-threshold slope factor and \( U_{\text{T}} \) is the thermal voltage. Parameters \( s_{\text{i}} \) and \( s_{\text{l}} \) are non-linear derating functions for modeling the impact of FBB on \( C_{\text{inv}} \) and \( I_{\text{leak}} \), as will be explained later. The delay of a digital logic gate can be modeled as:

\[ d_{\text{gate}} = \frac{(x C_{\text{inv}} + C_{\text{load}}) V_{\text{DD}} I_{\text{drive}}}{x I_{\text{leak}}} \]  

(2)

where \( I_{\text{drive}} \) is the current drive of a gate. In sub-threshold, the drive current is exponentially dependent on \( V_{\text{DD}} \) and \( V_{\text{BB}} \): \( I_{\text{drive}} \propto e^{(V_{\text{DD}} - V_{\text{BB}})/V_{\text{T}}/} \). In near- or super-threshold, \( I_{\text{drive}} \propto (V_{\text{DD}} - V_{\text{BB}})^{a} \) where \( a \) is a parameter that models velocity saturation. Parameter \( s_{\text{i}} \) is a non-linear derating function for modeling the FBB dependence of both \( C_{\text{inv}} \) and \( I_{\text{leak}} \), as will be explained later. In sub-/near-threshold, the circuit speed is reduced substantially due to a low gate drive voltage, \( V_{\text{DD}} - V_{\text{BB}} \). Now, reducing \( V_{\text{DD}} \) with FBB has a large impact on gate delay because \( V_{\text{BB}} \) is a larger portion of the gate drive voltage. This makes FBB attractive to use in ultra-low-power circuits. In fact, FBB is more energy-efficient than \( V_{\text{DD}} \) upscaling to increase the gate drive voltage when dynamic energy dominates the overall energy. The contrary holds true for leakage dominant situations.

At a given \( V_{\text{DD}} \), we use non-linear derating functions \( s_{\text{i}}, s_{\text{l}}, s_{\text{c}}, s_{\text{b}} \) to model the energy and delay dependence on FBB. Such functions can be easily calibrated for a given process technology by means of circuit simulations. For instance, the intrinsic capacitance, leakage, and delay of a reference circuit or logic gate can be plotted as function of FBB. We introduce the following derating functions:

\[ s_{\text{c}} = 1 + \frac{1}{m} \sum_{i=1}^{n} V_{\text{BB}}^{i} \]  

(3)

\[ s_{\text{i}} = \frac{1}{k} \left( e^{I_{\text{leak}}/V_{\text{T}}/} + I_{\text{leak}} \left( e^{I_{\text{leak}}/V_{\text{T}}/} - 1 \right) \right) \]  

(4)

\[ s_{\text{l}} = e^{I_{\text{leak}}/V_{\text{T}}/} \]  

(5)

where \( k, I, m \) are fitting parameters. \( V_{\text{BB}} \) represents the symmetrical FBB value: \( V_{\text{BB}} = V_{\text{DD}} - V_{\text{BB}} \). The left term in (4) models the sub-threshold leakage increase due to FBB, while the right term models the current due to forward-biased junctions. Simulation results for a CMOS 90 nm ring-oscillator revealed the accuracy of expressions (3)–(5). A maximum error of 3%, 15% and 6% was observed for (3), (4), and (5), respectively, when applying up to 0.5 V FBB for a \( V_{\text{DD}} \) range from 0.5 V to 0.7 V in our experiments.

Based on the models presented above, the total energy, and delay of a CMOS digital circuit design can be modeled as:

\[ E_{\text{total}} = V_{\text{DD}} \sum_{i=1}^{n} \left( s_{\text{i}} \cdot C_{\text{inv}}, \left( 1 + \frac{1}{m} \sum_{i=1}^{n} m_{\text{inv}}, V_{\text{BB}}^{i} \right) \right) C_{\text{load}} V_{\text{DD}} \]

\[ + T_{\text{ak}} \cdot x \cdot I_{\text{leak}} \cdot (e^{I_{\text{leak}}/V_{\text{T}}/} + I_{\text{leak}} \left( e^{I_{\text{leak}}/V_{\text{T}}/} - 1 \right) ) \]  

(6)

\[ D_{j} = V_{\text{DD}} \sum_{i=1}^{n} \left( C_{\text{inv}}, + x \cdot C_{\text{load}} \cdot I_{\text{drive}} \cdot e^{I_{\text{leak}}/V_{\text{T}}/} \right) \]  

\[ \forall j \in \Psi \]  

(7)

where \( i \) is an index that runs over all gates in the circuit, \( q \) is the total number of gates in the circuit, \( j \) is an index that runs over all circuit paths, \( D_{j} \) is the delay of path \( j \), and \( \Psi \) is the collection of all paths in the circuit. Expression (7) constrains the delay of each circuit.
path to be less than the targeted clock period, $T_{ck}$. In (7), we accommodate the $C_{load}$ dependence on FBB in fitting parameter $k_i$. At a given $V_{dd}$, the lowest energy design is obtained when no gates are up-sized, e.g., $x_i = 1 \forall$ gates $i$. However, this also leads to the slowest design, as can be inferred from (7).

The area of a CMOS circuit design is approximately equal to the sum of all gate areas. This gives

$$A_{total} = \sum_{i=1}^{n} x_i A_i$$

(8)

where $A_i$ represents the minimum area of gate $i$. The actual circuit area depends on the targeted operating frequency. The smallest design is obtained when no gates are up-sized. Both dynamic and leakage energy consumed by a circuit is proportional to the circuit area. Notice that $C_{load,i} \propto A_i$ and $I_{load,i} \propto A_i$. By combining (6) and (8), we obtain the following expression

$$E_{total} = (\lambda_c V^2 + \lambda_d V_{dd} T_{ck}) A_{total} + C_{load,total} V_{dd}^2$$

(9)

where $\lambda_c$ is the switching capacitance-area proportionality factor, $\lambda_d$ is the leakage-area proportionality factor, and $C_{load,total}$ is the summed switching load capacitance of each gate. Observe from (9) the linear dependence between energy and circuit area.

The dependence between circuit area and clock period can be expressed by a rational function, as proposed in Ref. [24]. For a given supply voltage and body bias assignment we obtain:

$$A_{total} = \frac{x}{\delta + T_{ck} + \eta}$$

(10)

where $\chi$, $\delta$, and $\eta$ are independent fitting parameters. Parameter $\chi$ models gate-up-sizing and logic re-structuring associated to a timing-constrained design. The minimum clock period of the design that is theoretically possible is achieved at $T_{ck} = -\delta$. Parameter $\eta$ models the circuit area in case of unconstrained or very relaxed timing. High-performance circuits consume more area than low-performance circuits by up-sizing gates to speed-up critical circuit paths. For a given clock constraint for both circuits, high-performance circuits can operate at a lower $V_{dd}$ than low-performance circuits to achieve the required performance while saving energy.

3.2. Performance-Per-Area Figure-of-Merit

We introduce the performance-per-area (PPA) metric to qualify how effectively silicon area is utilized to achieve circuit speed. The PPA is defined as

$$PPA = \frac{1}{T_{ck} A_{total}} = \frac{\delta + T_{ck}}{(\chi \delta \eta + \eta T_{ck}) T_{ck}}$$

(11)

The clock period at which the maximum PPA occurs ($T_{max}$), can be determined by making the derivative of PPA with respect to $T_{ck}$ equal to zero. By solving the equation for $T_{ck}$, we obtain

$$T_{max} = -\delta + \sqrt{\frac{\delta}{\eta}} \forall T_{ck} \geq T_{max} \land \delta \eta \leq 0$$

(12)

$T_{ck} > T_{max}$ yields circuits without area over-dimensioning, and the contrary holds true for $T_{ck} < T_{max}$. Therefore, $T_{max}$ identifies the minimum possible clock period without circuit over-dimensioning. In its turn, the circuit area at $T_{ck} = T_{max}$ defines the upper bound for circuit area of the design. Beyond this point, the area of the circuit design is seriously penalized to achieve performance, and the same holds for energy. The PPA metric depends on the technology node, the technology’s $V_{th}$ option, and the standard cells available for circuit synthesis. A higher PPA indicates that the circuit design utilizes silicon area more effectively to achieve performance. The maximum PPA design identifies the maximum performance design acceptable under gate sizing, while avoiding area over-dimensioning.

4. BODY BIAS DRIVEN DESIGN

SPACE EXPLORATION

In this section we explore the design space of energy, area, and timing for conventional and BBD design styles. The analysis has been performed on a generic digital circuit design in 90 nm LP-CMOS. In particular and without loss of generality, we focus on an ultra-low-power design at $V_{dd} = 0.5$ V using HVT devices only. All results relate to slow-process corner, and 25 °C operating temperature. This implies that the design operates in sub-threshold since $V_{th} = 0.6$ V for the applied conditions.

Figure 7 shows the results from today’s synthesis tools under a fixed $V_{dd}$ and $V_{th}$ assignment. The area versus clock-period curves of the conventional design (solid line) and the BBD design (dashed line) are plotted. The curves are constructed from a multitude of points, each corresponding to a design with unique speed requirements. Iso-energy curves are plotted as overlay. As way of example, all values in Figure 7 have been normalized to the conventional design for which a maximum PPA occurs (maximum PPA point or MPPAP at $T_{ck} = T_{max}$), as highlighted by the circle symbol. The MPPAP indicates the upper area bound for design without area over-dimensioning for achieving the higher circuit speed. Consider now a constant circuit area, then one can see that BBD design achieves a higher performance than the conventional design (arrow A in Fig. 7). The curve for a BBD design with a higher FBB value will be located at smaller clock period values. Notice also that the circuit’s energy consumption is nearly constant over a large clock period range, which implies the dominance of dynamic energy. However, energy increases as well for large FBB values. This is because of the increased junction capacitance and leakage with FBB. Alternatively, BBD design enables smaller circuit designs.
for a fixed clock period (arrow B). In this case BBD design can lower the circuit’s energy significantly; the amount of energy savings is design dependent. In general, we can say that for a given energy budget, BBD design offers better performance and area figures.

Figure 8 presents the total energy versus the amount of FBB. The solid-line curve indicates a multitude of BBD designs optimized for distinct FBB values and with the smallest possible design area. The dashed line relates to the BBD designs with a maximum PPA. The same normalization as before has been used. Notice that energy can be reduced by more than 25% by walking along the iso-clock “1” line when using ~40 mV FBB for BBD design (star symbol). This operating condition renders as well the smallest area design. The minimum area BBD design with 0.4 V FBB achieves about 7.7x higher operating speed for the same energy consumption as the conventional design at maximum PPA. Alternatively, the BBD design with 0.4 V FBB in case of maximum PPA achieves about 9.2x higher operating speed at a energy penalty of about 28% w.r.t. as the conventional design at maximum PPA. One can also observe that the increase in energy is rather limited up to 0.4 V FBB at the considered operating temperature of 25 °C. For larger FBB values, the leakage current due to forward-biased junctions significantly increases and consequently so does the energy.

Figure 9 plots the EDP versus circuit area for the sub-threshold design. The solid line relates to the conventional design, while the dashed line relates to various BBD design cases. The circle symbol highlights the reference non body biased design which is designed at its maximum PPA point. The star symbol indicates the minimum area design with small FBB (~40 mV) that has the same operating speed as the reference design. The iso-clock period lines have been plotted as overlay. In general, the smallest designs achieve the best (or minimum) EDP. BBD design consistently achieves a better EDP than the conventional design, irrespective of the FBB applied. Without loss of generality, consider now the particular example of BBD design with a 0.4 V FBB. Observe that the BBD design can run at 0.13x the clock period of the conventional design, and that the corresponding EDP is 7.7x lower (as indicated by the triangle symbol in Fig. 9). In other
words, the minimum area BBD design can improve EDP by almost one order of magnitude with respect to the conventional design. Also notice that sensitivity of EDP to circuit sizing reduces under BBD design. This implies that circuit sizing increases its effectiveness for trading-off circuit speed against energy consumption under BBD design, without seriously affecting energy efficient operation.

Figure 10 highlights the PPA and total energy trends. The iso-clock period lines have been plotted as overlay.

As before, the red circle indicates the MPPAP of the conventional design, the star symbol points to the minimum area BBD design with the same performance, and the triangle symbol relates to the minimum area BBD design with the same energy consumption as the maximum PPA conventional design. We can observe that BBD design exhibits a noteworthy higher PPA over the conventional design under all circumstances. This higher PPA comes primarily from the increased circuit performance and reduced circuit area. FBB is simply a more energy-efficient approach for speeding-up the circuit than the conventional gate area up-sizing. Observe also that for each BBD design, there exists a MPPAP at which the design utilizes circuit area most effectively, as indicated by the circle symbols. In other words, it is the area-speed point for which performance is the highest before the circuit is over-dimensioned. Over-dimensioning comes at the price of a significant energy penalty as compared to an insignificant performance increase. Consequently, circuit sizing should be constrained by the MPPAP of the design. When a higher performance is needed, the threshold voltage of the transistors should be lowered, or alternatively the $V_{th}$ of the circuit can be increased.

Table II illustrates the impact of technology scaling on maximum PPA designs for the same generic digital circuit design in different process technology nodes. All values have been normalized to the maximum PPA design under conventional design in 90 nm LP-CMOS. Observe that BBD design consistently outperforms conventional design; A maximum PPA design is smaller in a next-generation technology. Consequently, the PPA increases with technology scaling. The total energy consumption reduces in a next-generation technology due to reduced circuit capacitances. The leakage energy is more pronounced in 45 nm LP-CMOS, which can be observed by the energy increase of the BBD design case. In a next-generation technology, the speed increase with BBD design is somewhat lower, mainly because a lower device threshold voltage. In general, we can state that BBD design remains effective with scaling in the analyzed process technology nodes.

### 5. BODY BIAS IMPLEMENTATION

Body biasing requires deep N-well isolation, power delivery networks and voltage generation. In this section we discuss the physical design aspects as well as body bias selection and generation.

#### 5.1. Physical Design Aspects

Under BBD design, FBB is utilized at design-time to speed-up speed-critical circuit parts of the design. All body-biased PMOS transistors experience the same amount of FBB, and the same holds true for the NMOS devices. In this way, body bias digital cells can share the same physical N-well and P-well. Deep N-well isolation is added to (part of) the design for separating the P-well of the body-biased NMOS devices from the P-substrate. Only 2 μm extra is needed for the Deep N-well at each side of the body biased circuit part in 90 nm LP-CMOS.

Figure 11 illustrates a layout implementation example of a body-biased circuit block. The circuit is place-and-routed in standard-cell rows. The N-well and P-well connections are made through dedicated well tap cells; the tap cells have been inserted in columns at a maximum pitch of 60 μm. A two-layer routing grid has been utilized for connecting the tap cells to the body bias supply.
5.2. Body Bias Selection and Generation

It is possible to use a fully programmable body bias generator for providing the required FBB values for post-silicon tuning. Alternatively, and without loss of generality, we illustrate the implementation of a fixed FBB voltage. A single FBB voltage of \( V_{DD}/2 \) has been selected for joint N-well and P-well biasing of the body biased circuit when operating at reduced supply voltages \( (V_{DD} \leq 1 \text{ V}) \). This FBB voltage is found sufficient for slow-to-nominal process corner performance compensation at \( V_{DD} = 0.5 \text{ V or 0.7 V} \). More FBB can further increase performance but at leakage penalty, as shown in Table I. By using a single FBB voltage, we limit both the area and energy overhead associated the use and generation of multiple body bias voltages. Figure 12 shows a conceptual circuit diagram of the proposed body bias power delivery. The FBB generator consists of a \( V_{DD}/2 \) voltage generator and a switch box to support dynamic FBB. The voltage generator can be implemented by, for example a resistor divider, a low-power LDO or switched-capacitor converter unit. The FBB generator only needs to supply the difference between well currents, because the N-well current sources the P-well current. Therefore, the generator only needs a limited output current drive capability.

6. SYNTHESIZED DESIGN EXAMPLE

Commercial synthesis tools target area optimization subject to delay constraints. In this section we present synthesis results for an industrial microprocessor design in 90 nm LP-CMOS. Design synthesis has been performed for slow-process and 25 °C temperature conditions. We have utilized BBD design with a FBB voltage of \( V_{DD}/2 \). Digital cell timing and power have been characterized using Altos’ Liberate library characterizer. Logic synthesis has been done with Cadence RTL Compiler. Synthesis has been directed towards the smallest design with minimum leakage and dynamic power that meets timing constraints.

The circuit area accounts for layout effects such as row utilization, FBB power delivery networks and Deep N-well isolation.

The microprocessor contains 3764 flip-flops and about 31 K combinational gates. It makes use of HVT devices only. For both conventional and BBD design styles, the design synthesis has been done at \( V_{DD} = 0.5 \text{ V and 0.7 V} \). This relates to a sub-threshold and a near-threshold design, respectively. Figure 13 presents the area and clock period trends for the sub-threshold design. The results obtained from synthesis, have been indicated by circles and triangles for the conventional and BBD designs, respectively. The trend lines are calculated from expression (10) by using the fitting parameters shown.

---

**Fig. 11.** Layout implementation example of a body biased digital circuit.

**Fig. 12.** Conceptual circuit diagram of the proposed body bias power delivery.

**Fig. 13.** Area versus clock period for the microprocessor design in CMOS 90 nm at \( V_{DD} = 0.5 \text{ V} \). Lines: expression (10), symbols: synthesis results.

---

*J. Low Power Electron. 6, 1–12, 2011*
in Table III. The minimum energy and maximum PPA designs have been indicated in Figure 13 by diamond and circle symbols, respectively. The key results have been highlighted in Table IV.

Notice in Figure 13 that the maximum PPA conventional design has a significantly higher performance than its minimum energy design. The minimum energy design is located at $T_{ck} = 4 \mu s$ ($f_{ck} = 250$ kHz), while the maximum PPA design is located at $T_{ck} = 1.33 \mu s$ ($f_{ck} = 751$ kHz). Observe in Figure 13 large performance increase that can be achieved with BBD design. The maximum PPA BBD design can be found at $T_{ck} = \sim 350$ ns ($f_{ck} = \sim 2.9$ MHz). By comparing the maximum PPA designs, BBD design achieves up to 3.8x higher performance w.r.t. the conventional design. In addition, it achieves a 9% smaller circuit area. This gives a PPA improvement of 4.2x for the BBD design over the conventional design. Moreover, the overall energy consumption of the BBD design is about 10% lower, and the EDP is improved by 4.2x. At the same clock frequency, BBD design reduces circuit area by 29% at the same clock frequency as the maximum PPA conventional design, which results in a PPA improvement of 39%, an overall energy reduction of 36%, and an EDP improvement of 36%. By comparing the minimum energy designs, BBD design achieves up to 3x higher performance w.r.t. the conventional design at the same energy consumption. Moreover, the BBD design achieves a 3% smaller circuit area, and PPA and EDP improvements of 3.1x and 3.3x, respectively.

Figure 14 illustrates the PPA versus energy trends for this sub-threshold design. Notice that the same trends as shown before in Figure 10 are observed from this experimental synthesis. BBD design improves the maximum PPA value by about 4.2x w.r.t. the maximum PPA conventional design. If operation at a maximum PPA is not selected, one could also trade-off the energy per cycle of the design against a smaller circuit area but at a significant performance penalty. For the conventional design, this would result into operation at 85 pJ/cycle for 26% lower area and 3x lower frequency. Typically, in conventional design methods, the $V_{DD}$ is raised to compensate this performance loss, but this comes at the price of worsening energy consumption.

The previous analysis has also been performed for a near-threshold design ($V_{DD} = 0.7$ V). Table V summarizes the results. Notice the 20x higher performance than the sub-threshold design for maximum PPA ($f_{ck, max} = 15.7$ MHz vs. $f_{ck, rel} = 751$ kHz). For the near-threshold maximum PPA designs, the BBD design achieves almost 2x higher performance, up to 8% smaller circuit area, 2.1x higher PPA, and 2x improved EDP w.r.t. the

Table III. Parameters of expression (10) for the microprocessor design in 90 nm HVT LP-CMOS at $V_{DD} = 0.5$ V

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Conventional design</th>
<th>BBD (0.25 V FBB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$ [\mu m$^3$/ms]</td>
<td>3.94 $\times 10^{-6}$</td>
<td>9.8 $\times 10^{-7}$</td>
</tr>
<tr>
<td>$\delta$ [\mu s]</td>
<td>-1.04</td>
<td>-0.26</td>
</tr>
<tr>
<td>$\eta$ [\mu m]</td>
<td>2.47 $\times 10^{-6}$</td>
<td>2.72 $\times 10^{-6}$</td>
</tr>
</tbody>
</table>

Table IV. Microprocessor design characteristics in 90 nm HVT LP-CMOS at $V_{DD} = 0.5$ V. BBD utilizes an FBB value of 0.25 V.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum energy design</th>
<th>Maximum PPA design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit area [\mu m$^2$]</td>
<td>298951</td>
<td>288202</td>
</tr>
<tr>
<td>Performance [kHz]</td>
<td>250</td>
<td>1333</td>
</tr>
<tr>
<td>PPA [MHz/\mu m$^2$]</td>
<td>8.6 $\times 10^{-3}$</td>
<td>2.7 $\times 10^{-6}$</td>
</tr>
<tr>
<td>Energy [pJ/cycle]</td>
<td>85</td>
<td>77</td>
</tr>
<tr>
<td>EDP [pJ/\mu s]</td>
<td>339</td>
<td>103</td>
</tr>
</tbody>
</table>
conventional design. At the same clock period, BBD design reduces circuit area by 23%, increases PPA by 30%, consumes 29% less energy, and improves EDP by 29%. By comparing the minimum energy designs, BBD design achieves up to 50% higher performance w.r.t. the conventional design at the same energy consumption. In this case, the BBD design achieves a comparable circuit area, a PPA improvement of 53%, and a 33% improved EDP.

7. CONCLUSION

We presented a design synthesis strategy for (near-) sub-threshold designs that is supported by forward body biasing. Furthermore, we include in our strategy silicon area utilization which is based on a performance per area metric. Our approach renders low-area energy-efficient designs without the major performance penalties associated to conventional sub-threshold design. We demonstrated its effectiveness on a microprocessor designed associated to conventional sub-threshold design. We demonstrated its effectiveness on a microprocessor designed.

Acknowledgments: The authors would like thank Aatish Kumar from NXP Semiconductors, Central Research and Development, for his support during the library characterization stage.

References


16. T. Chen and S. Naffziger, Comparison of adaptive body bias (ABB) and adaptive supply voltage (ASV) for improving delay and leakage under the presence of process variation. IEEE Transactions on VLSI Systems 11, 888 (2003).

Table V. Microprocessor design characteristics in 90 nm HVT LP-CMOS at $V_{DD} = 0.7$ V. BBD utilizes an FBB value of 0.35 V.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>286401</td>
<td>6.7</td>
<td>3.3·10⁻⁴</td>
<td>162</td>
<td>24.3</td>
<td>378921</td>
<td>347325</td>
<td>8%</td>
</tr>
<tr>
<td>283028</td>
<td>10.0</td>
<td>3.5·10⁻⁴</td>
<td>162</td>
<td>16.2</td>
<td>15.7</td>
<td>30.3</td>
<td>1.9x</td>
</tr>
</tbody>
</table>

J. Low Power Electron. 6, 1–12, 2011
Ultra-Low-Power Digital Design with Body Biasing for Low Area and Performance-Efficient Operation


Maurice Meijer
Maurice Meijer received the B.Eng. degree in electrical engineering from Eindhoven Polytechnic, Eindhoven, The Netherlands, in 1999 and the M.Sc. degree in electrical engineering from Eindhoven University of Technology, The Netherlands, in 2004. In 1999, he joined Philips Research Laboratories, The Netherlands, where he has been a Research Scientist in the Digital Design and Test group. He is currently a Senior Scientist at NXP Semiconductors, The Netherlands. His research interests are in the areas of low power and variation tolerant digital integrated circuit design.

José Pineda de Gyvez
José Pineda de Gyvez (F’09) received the Ph.D. degree from the Eindhoven University of Technology, The Netherlands, in 1991. From 1991 until 1999 he was a Faculty member in the Department of Electrical Engineering at Texas A&M University, USA. He is currently a Senior Principal at NXP Semiconductors in The Netherlands. Since 2006 he also holds the professorship “Deep Submicron Integration” in the Department of Electrical Engineering at the Eindhoven University of Technology. Dr. Pineda de Gyvez has been Associate Editor in IEEE Transactions on Circuits and Systems Part I and Part II, and also Associate Editor for Technology in IEEE Transactions on Semiconductor Manufacturing. He is also a member of the editorial board of the Journal of Low Power Electronics. Pineda has co-authored more than 100 publications in the fields of testing, nonlinear circuits, and low power design. He is (co)-author of three books, and has a number of granted patents. His work has been acknowledged in academic environments as well as in patent portfolios of many companies. Pineda’s research has been funded by the Dutch Ministry of Science, US Office of Naval Research, US National Science Foundation, among others.

Ajay Kapoor
Ajay Kapoor received the B.Tech. degree (1995–1999) in electrical engineering from the Indian Institute of Technology (IIT), Delhi, India, in 1999, and the M.Sc. degree (2003–2005) with cum laude in embedded systems from the University of Twente, Enschede, The Netherlands, in 2005. From 1999 to 2005, he was with Philips (now NXP) Semiconductors as a CMOS/BiCMOS logic circuit designer. From 2005 to 2008, he was working with Philips (now NXP) as a scientific researcher for low power digital system and algorithms for wireless communication. From 2008 to 2009, he was working at ST-Ericsson as a scientific researcher for systems and algorithms for low power signal processing and baseband architectures for wireless radio communication systems. Since Jan. 2010, he is working as a scientific researcher for low power circuit, systems and architectures at NXP–research. His current interests are low power circuits, architectures and signal processing.