Nanometer MOSFET Effects on the Minimum-Energy Point of 45nm Subthreshold Logic

David Bol, Dina Kamel, Denis Flandre and Jean-Didier Legat
Microelectronics laboratory, Université catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
{david.bol,dina.kamel,denis.flandre,jean-didier.legat}@uclouvain.be

ABSTRACT

In this paper, we observe that minimum energy $E_{min}$ of subthreshold logic dramatically increases when reaching 45 nm node. We demonstrate by circuit simulation and analytical modeling that this increase comes from the combined effects of variability, gate leakage and DIBL. We then investigate the new impact of MOSFET parameters on $E_{min}$ in nanometer technologies. We finally propose an optimum MOSFET selection intended for subthreshold circuit designers, which favors low-$V_t$ mid-$L_g$ devices in standard 45nm GP technology. The use of such optimum MOSFETs yields 35% $E_{min}$ reduction for a benchmark multiplier with good speed performances and negligible area overhead.

Categories and Subject Descriptors
B.7.1 [Hardware]: Advanced technologies;
D.8.2 [Hardware]: Performance Analysis and Design Aids.

General Terms
Design, performance.

Keywords
CMOS digital integrated circuits, gate leakage, short-channel effects, subthreshold logic, ultra-low power, variability.

1. INTRODUCTION

Subthreshold logic is an efficient technique to drastically reduce the energy per operation in applications with moderate demand in speed performances [1]. Lowering the supply voltage $V_{dd}$ yields a quadratic reduction in switching energy $E_{sw}$, at the expense of increased delay and thus execution time of the operation. This makes leakage energy $E_{leak}$ increase, as it results from the integration of leakage power over the execution time. There is thus an optimum supply voltage $V_{min}$ that yields minimum energy per operation $E_{min}$ [2]. Over the last decade, minimum energy has become a popular research direction for ultra-low-power applications [3]. Moreover, operating at this minimum-energy point has recently been proposed in medium-performance applications relying on massive parallelization [4].

Beyond Silicon area savings, CMOS technology scaling yields $E_{leak}$ reduction through capacitance reduction at the expense of increased leakage currents, short-channel effects and variability. In [5], Hanson et al. report that minimum energy level $E_{min}$ is reduced when migrating from 90 nm to 32 nm node, when considering LSTP technology trend under typical conditions and neglecting gate leakage. They also show that $E_{min}$ is proportional to the switched capacitance multiplied by the square of the subthreshold swing $C_{LS}$, a factor that decreases with technology scaling. In nanometer technologies, variability cannot be neglected, especially uncorrelated $V_t$ variations due to random dopant fluctuations (RDF) as they cannot be compensated by adaptive body biasing. In [6], we showed using PTM models [7] that variability strongly affects $E_{min}$ scaling trend. Additionally, as depicted in Fig. 1, when considering RDF-induced 3σ worst-case delay and mean leakage through Monte-Carlo Spice simulations of an 8-bit RCA multiplier\(^1\), we observe a dramatic $E_{leak}$ increase in general-purpose (GP) technology $E_{min}$.

\(^1\)Statistical delay extraction on the critical path (logig depth is 23) and $E_{tot}$ extraction on the full circuit with random inputs at the maximum data rate, given by the obtained worst-case delay. The input activity factor is 0.5, the duty cycle is 1 (no stand-by period) and room temperature is considered. Although these parameters do affect $E_{min}$ and $V_{min}$ values [2], they do not change the scaling trend.
gies from an industrial foundry. Although technology scaling reduces $E_{sw}$ contribution ($E_{tot}$ curves above 0.4V) and $C_L S^2$ factor as shown in the insert of Fig. 1, new effects make $E_{min}$ deviate from $C_L S^2$ factor. $E_{min}$ thus increases from 90nm to 45nm node, which suggests a major challenge to keep energy consumption reasonable in nanometer subthreshold circuits.

In this paper, we therefore first analyze the effects that make $E_{min}$ rise in nanometer technologies. We then address the issue of selecting the optimum MOSFETs for minimum energy in 45nm technology, from a circuit designer point of view i.e. within a practical set of available devices. Our contributions are:

- a detailed analysis of minimum-energy point in nanometer subthreshold logic showing the new combined effects of variability, gate leakage and drain-induced barrier lowering (DIBL), that clearly outweigh the previously-reported $C_L S^2$ figure of merit;
- an investigation of the impact of $V_t$, $T_{ox}$ and $L_g$ on minimum-energy point, given these new effects;
- a study intended for circuit designers, to select the optimum MOSFETs for nanometer minimum-energy subthreshold logic circuits, in a standard CMOS technology without any process modification.

This article is organized as follows. Section 2 briefly reviews the concept of minimum-energy point and the basic impact of device parameters in submicron technologies. In Section 3, we propose a pre-Silicon MOSFET compact modeling approach for realistic nanometer subthreshold circuit simulations. We then use the generated models in Section 4 to show the impact of nanometer MOSFET effects and device parameters on minimum energy. Optimum nanometer MOSFET selection is finally addressed in Section 5.

2. MINIMUM ENERGY

Total energy comes from $E_{sw}$ and $E_{leak}$ contributions:

$$E_{tot} = \frac{1}{2} N_{sw} C_L V_{dd}^2 + V_{dd} I_{leak} T_{del}$$

where $N_{sw}$ is the number of switched nodes to perform the operation, $C_L$ the average node capacitance, $I_{leak}$ the total leakage current and $T_{del}$ the circuit delay, corresponding to the execution time of the operation. Minimum energy is often achieved when operating the circuit in subthreshold region [2]. Subthreshold drain current is expressed as:

$$I_{sub} = I_0 \times 10^{\frac{V_{gs}+V_{th}}{S}} \times \left(1 - e^{-\frac{V_{dd}}{\eta V_{th}}}ight)$$

where $I_0$ is a reference current proportional to $W/L$, which exponentially depends on the threshold voltage $V_t$. $S$ is the subthreshold swing, $\eta$ the DIBL coefficient and $U_{th}$ the thermal voltage. In previous works on minimum-energy point [2, 5], leakage currents are assumed to be dominated by subthreshold leakage and leakage energy is expressed as:

$$E_{leak} = V_{dd} \times I_{leak} \times T_{del}$$

$$\propto V_{dd} \times I_{sub,off} \times C_L V_{dd}$$

$$\propto V_{dd} \times I_0 \times 10^{\frac{nV_{dd}}{S}} \times \frac{C_L V_{dd}}{I_0 10^{\frac{nV_{dd}}{S}} (1+\eta) V_{th}}$$

$$\propto C_L \times \frac{V_{dd}}{V_{th}} V_{dd}^2.$$ (3)

Several works investigate the interests of new MOSFET architectures for subthreshold circuits such as double-gate SOI [8] and underlap devices [9] or post-Si devices such as super cut-off transistors [10]. In particular, we showed that fully-depleted SOI planar technology with ultra-thin body can significantly improve $E_{min}$ at 45nm node [11]. Nevertheless, in this article we rather focus on planar bulk CMOS devices as it is today’s mainstream for circuit designers. For optimum subthreshold operation of such standard devices, it has been shown that halo doping can be reduced [5], [12] and that a high-to-low vertical channel doping profile is preferable [12]. However, to the authors’ knowledge no industrial foundry offers a technology that specifically targets minimum-energy subthreshold logic. Moreover, process modifications such as doping profile optimizations are hardly available to circuit designers. In this article, we therefore focus on the 3 main device parameters that circuit designers can usually choose within a set in a versatile yet standard technology menu: $V_t$, $T_{ox}$ and $L_g$.

Under typical conditions, minimum energy $E_{min}$ is shown in [5] to be proportional to $C_L S^2$. In [13], variability is shown to worsen $E_{min}$ because worst-case delay has to be considered in order for all manufactured chips to correctly operate with the same clock frequency under a given $V_{dd}$.

In [2], it is suggested that $V_t$ has no impact on $E_{min}$, provided that the devices actually remain in subthreshold regime. This is confirmed by $E_{leak}$ expression from Eq. (3) where the $V_t$ impact through $I_0$ parameter is canceled when multiplying $I_{leak}$ by $T_{del}$. However, if $V_t$ is too low, $I_{on}$ is no longer a subthreshold current and Eq. (2) does not hold. The delay is thus no longer exponentially-dependent on $V_{dd}$ and the $I_{leak} \times T_{del}$ product increases.

In [5], it is shown that there is an optimum $L_g$ to minimize $E_{min}$, which results from a trade-off between $S$ improvement from short-channel behavior mitigation and $C_L$ increase at longer $L_g$. Notice that the optimum $L_g$ is quite long as the impact of intrinsic gate capacitance $C_g$ on $C_L$ is small. This comes from the reduced $C_g$ in subthreshold regime as compared to parasitic capacitances because subthreshold $C_g$ is dominated by the channel depletion capacitance [12].

Finally, it is shown in [14] that an optimum $T_{ox}$ also results from a trade-off between $S$ improvement from improved channel control and $C_g$ increase (intrinsic, fringing and overlap), at thinner $T_{ox}$. In Section 4, we will show that these trade-offs remain valid in nanometer technologies but are outweighed by DIBL, gate leakage and variability.

3. PRE-SILICON MODEL CARDS FOR SUBTHRESHOLD SIMULATION

Throughout this paper, we investigate circuit-level implications of MOSFET characteristics with an 8-bit multiplier as benchmark circuit. For Monte-Carlo simulation time issues, we thus use SPICE simulator based on BSIM4 compact models. In order to investigate the impact of device parameters, we need a generic yet accurate model card to define BSIM4 parameters. The methodology we use for generating these model cards is illustrated in Fig. 2 and main resulting subthreshold characteristics of baseline devices ($L_g=35nm$, $V_{t}=0.35V$, $T_{ox}=1.1nm$) are given in Table 1.

**Basic parameters:** we start from 45nm Predictive Technology Models (PTM) from Arizona State University [7] as

2Models are available on-line at [www.eas.asu.edu/~ptm](http://www.eas.asu.edu/~ptm).
When considering devices with different nominal relations \[18\] while the benchmark multiplier is quite small. Devices because CD variations exhibit a strong spatial correlation \[18\] while the benchmark multiplier is quite small. We consider a single \[18\] and 1.7 nm BSIM4 model considered for the simulations of Fig. 1 at 1.1 nm technology, with BSIM4 models generated according to the methodology presented in Section 3. The circuit is first simulated with ideal devices, i.e. without variability, gate leakage nor DIBL. The same simulation is then carried out by successively adding DIBL, gate leakage and variability. The resulting \(E_{min}\) values are plotted in Fig. 3 showing the high \(E_{min}\) overhead of these effects. Let us analyze the reasons of this overhead.

4. **NANOMETER MOSFET EFFECTS ON MINIMUM ENERGY**

As shown in Fig. 1, new effects in nanometer technologies make \(E_{min}\) increase and deviate from \(C_L S^2\) trend. In order to investigate these effects, we consider the benchmark multiplier with baseline devices \((L_g=35nm, V_t=0.35V, T_{ox}=1.1nm)\) in 45 nm technology, with BSIM4 models generated according to the methodology presented in Section 3. The circuit is first simulated with ideal devices, i.e. without variability, gate leakage nor DIBL. The same simulation is then carried out by successively adding DIBL, gate leakage and variability. The resulting \(E_{min}\) values are plotted in Fig. 3 showing the high \(E_{min}\) overhead of these effects. Let us analyze the reasons of this overhead.

4.1 **Drain-Induced Barrier Lowering**

The DIBL effect, which is important in nanometer technologies, implies an exponential dependence of \(I_{sub}\) on \(V_{dd}\) as shown in Eq. (2). According to Eq. (3), DIBL should not impact \(E_{leak}\) as the \(I_{leak}\) increase from DIBL \((\eta V_{dd}^3)\) is compensated by an equal delay reduction. In this equation, the delay is assumed to be inversely proportional to \(I_{leak}\) and \(I_{on}\), i.e with \(V_{dd} = V_{dd}\). However, during a transition, the current to charge/discharge the load capacitance is not constant. In particular as \(V_{dd}\) varies, the delay depends on the integral of \(I_{sub}\) current over the transition \(S_{tr}\), that we model as:

\[
T_{del} \propto \frac{C_L V_{dd}}{\int_{S_{tr}} I_{sub}} \approx \frac{C_L V_{dd}}{I_{0} 10^{k_{DIBL} \eta V_{dd}} \times \int_{S_{tr}} 10^{k_{DIBL} \eta V_{dd}}} = \frac{C_L V_{dd}}{I_{sub,dd}} \times 10^{k_{DIBL} \eta V_{dd}},
\]

where \(k_{DIBL}\) is a fitting parameter, whose value depends on \(V_{dd}\) with \(0 < k_{DIBL} < 1\). It accounts for the DIBL-induced \(I_{sub}\) reduction during a transition, the ideal case being \(k_{DIBL} = 0\). The value of \(k_{DIBL}\) can be empirically extracted from simulation of the delay with and without the DIBL effect. Simulations of the benchmark circuit have been carried out with a large set of device parameters \((L_g, V_t, T_{ox})\) and show that the value of \(k_{DIBL}\) is pretty much independent on these parameters, provided that the devices actually stay in subthreshold regime, i.e. \(V_{ds} \leq V_t\). The value of \(k_{DIBL}\) is 0.65 at 0.2V and 0.75 at 0.5V.
This shows that the DIBL effect has a delay overhead. When injecting the DIBL-aware delay expression from Eq. (5) into $E_{\text{leak}}$ formula from Eq. (1), the impact of DIBL effect on $E_{\text{leak}}$ clearly appears:

$$E_{\text{leak}} = V_{dd} \times I_{\text{leak}} \times T_{\text{del}}$$

$$\propto V_{dd} \times I_{o} 10^{-\eta \frac{V_{dd}}{S}} \times \frac{C_{L} V_{dd}}{I_{o} 10^{-\eta \frac{V_{dd}}{S}}} \times 10^{\frac{k_{\text{DIBL}} V_{dd}}{S}}$$

$$\propto C_{L} 10^{-\eta \frac{V_{dd}}{S}} V_{dd} \times (1 + \frac{I_{g} \eta \frac{V_{dd}}{I_{\text{sub,off}}}}{10^{-\eta \frac{V_{dd}}{S}}}).$$

This clearly shows that $I_{g}$ worsens $E_{\text{leak}}$ unless it is much lower than subthreshold leakage $I_{\text{sub}} = I_{o} 10^{-\eta \frac{V_{dd}}{S}}$. Rather than having low absolute $I_{g}$, the important target to minimize $E_{\text{min}}$ is to keep $I_{g}/I_{\text{sub}}$ ratio lower than 1.

### 4.2 Gate leakage

When shrinking $T_{ox}$, gate leakage exponentially increases and becomes comparable to subthreshold leakage in nanometer technologies. Therefore, $E_{\text{leak}}$ expression must include $I_{gate}$ contribution to total $E_{\text{leak}}$, i.e. back from Eq. (3):

$$E_{\text{leak}} = V_{dd} \times I_{\text{leak}} \times T_{\text{del}}$$

$$\propto V_{dd} \times (I_{o} 10^{-\eta \frac{V_{dd}}{S}} + I_{g}) \times \frac{C_{L} V_{dd}}{I_{o} 10^{-\eta \frac{V_{dd}}{S}}}$$

$$\propto C_{L} 10^{-\eta \frac{V_{dd}}{S}} V_{dd} \times (1 + \frac{I_{g} \eta \frac{V_{dd}}{I_{\text{sub,off}}}}{10^{-\eta \frac{V_{dd}}{S}}}),$$

which are plotted vs. $V_{t}$ for ideal devices.

### 4.3 Variability

Device variability has a strong impact on subthreshold circuit delay as $I_{\text{sub}}$ exponentially depends on $V_{t}$ [13]. Amongst variability sources, uncorrelated $V_{t}$ variations are especially important as they cannot be compensated by adaptive body biasing. For computing the average $E_{\text{leak}}$, statistically-extracted $95\%$ worst-case delay of the small benchmark multiplier has to be used as the execution time for enabling 99.99% timing yield of a full chip. This leads to 20% $E_{\text{min}}$ increase. Moreover, RDF lead to a mean leakage current for large circuits much higher than the typical value because subthreshold leakage is a lognormal distribution (exponential dependence on $V_{t}$ normal distribution). The total $E_{\text{min}}$ overhead of variability is 40%, as shown in Fig. 3.

This discussion shows that there are new device targets in nanometer technologies to design/select optimum MOSFETs for minimum-energy subthreshold logic:

- low $C_{L} S^{2}$ factor [5],
- low DIBL effect,
- $I_{g}/I_{\text{sub}}$ ratio significantly lower than 1,
- low variability [13].

## 5. OPTIMUM MOSFET SELECTION FOR MINIMUM-ENERGY CIRCUITS

As shown in Fig. 1, direct porting of a subthreshold circuit from 90 nm to 45 nm general-purpose technologies results in 50% $E_{\text{min}}$ overhead. However at 45 nm node, circuit designers have new opportunities as technologies are versatile. They offer a menu of several flavors, with various speed/power trade-offs resulting from different device configurations ($L_{g}$, $T_{ox}$ and $V_{t}$). Moreover, each technology flavor often features dual or triple-$V_{t}$ devices and one can choose to use minimum or longer $L_{g}$. Circuit designers thus face the complex problem of optimum technology/device selection with multiple degrees of freedom. In this section, we address this selection problem in two steps: first by analyzing the impact of basic $L_{g}$, $V_{t}$, and $T_{ox}$ parameters on $E_{\text{min}}$, given the nanometer MOSFET effects and second, by investigating the optimum MOSFET, to get the lowest $E_{\text{min}}$. Notice that for each device parameters, a new model card is generated according to the methodology from Section 3 to get realistic MOSFET subthreshold characteristics.

### 5.1 Gate length impact

Fig. 4(a) shows $E_{\text{min}}$ vs. $L_{g}$ for ideal to real devices by successively adding DIBL, gate leakage and variability (worst-case $T_{\text{del}}$ and mean $I_{\text{leak}}$). The $C_{L} S^{2}$ factor is plotted too for comparison purpose. For ideal devices, $E_{\text{min}}$ exhibits the same dependence vs. $L_{g}$ as $C_{L} S^{2}$; it decreases until 55 nm $L_{g}$ thanks to $S$ improvement, whereas $C_{L}$ increases slowly because it is dominated by parasitic capacitances [5].

High DIBL effect of short devices implies an important energy overhead through large $\eta$ coefficient ($k_{\text{DIBL}}$ does not vary). Long devices are affected by gate leakage due to a high $I_{g}/I_{\text{sub}}$ ratio because $I_{g}$ increases with $L_{g}$ while $I_{\text{sub}}$ decreases from DIBL mitigation. Finally, variability worsens the picture for short devices. It comes from higher RDF due to smaller channel area, as well as magnified current sensitivity against $L_{g}$ variations because of high DIBL. The optimum $L_{g}$ in nanometer technologies thus results from a trade-off between low variability, $S$ and DIBL for long devices, and low $I_{g}/I_{\text{sub}}$ for short devices. This trade-off clearly outweighs the $C_{L} S^{2}$ trade-off.

### 5.2 Threshold voltage impact

Minimum energy level and $C_{L} S^{2}$ factor are plotted vs. $V_{t}$ in Fig. 4(b). Lowering $V_{t}$ implies reducing the channel doping $N_{ch}$, which results in a lower channel depletion capacitance $C_{dp}$ and in turn a better subthreshold swing at low $V_{ds}$. Low-$V_{t}$ devices thus feature a low $C_{L} S^{2}$ factor. Nevertheless, lowering $V_{t}$ does not improve $E_{\text{min}}$ of ideal devices because the $S$ improvement is compensated by the fact the devices leave the subthreshold regime, which results in lower $I_{on}/I_{off}$ ratio at a given $V_{dd}$ and in turn $E_{\text{leak}}$ overhead.

With a reduced $N_{ch}$ for low-$V_{t}$ devices, the short-channel effects are increased and DIBL thus degrades $E_{\text{min}}$. On the other hand, higher $V_{t}$ lowers $I_{\text{sub}}$, thereby degrading $I_{gate}/I_{\text{sub}}$ ratio. Variability has an important impact on $E_{\text{min}}$ for both high- and low-$V_{t}$ devices. High-$V_{t}$ devices have high $\sigma_{V_{t}}$ because of high channel doping, as shown in Eq. (4). Low-$V_{t}$ devices suffer from an important current sensitivity against $\sigma_{V_{t}}$ due to their high DIBL. Optimum $V_{t}$ selection thus results from a trade-off between DIBL mitigation and $I_{gate}/I_{\text{sub}}$ ratio reduction.

### 5.3 Oxide thickness impact

Finally, the impact of $T_{ox}$ on $E_{\text{min}}$ is shown in Fig. 4(c). Notice that, for illustration purpose, we keep $N_{ch}$ constant rather than $V_{t}$ when varying $T_{ox}$ because $S$, $\eta$ and $\sigma_{V_{t}}$ primarily depend on $N_{ch}$ rather than on exact $V_{t}$ value.
Fig. 5 shows that, although $C_L$ is a surprising observation because, whereas the DIBL-induced energy overhead is lower for thin-oxide devices, the outer fringing capacitance from gate to source/drain electrodes is very important [6, 16] as the gate electrode thickness is high and the spacer width is small. This capacitance component hardly depends on $T_{ox}$ and thus sets, together with junction capacitance, a lower bound on the achievable $C_L$ reduction.

For ideal devices, a thinner $T_{ox}$ implies a lower $V_t$, which translates into an $E_{min}$ increase as the devices leave the subthreshold regime, as explained in previous section. The DIBL-induced energy overhead is lower for thin-oxide devices, whereas the $I_{gate}$ overhead is only important for thin-oxide devices, which feature a high $I_{gate}/I_{sub}$ ratio despite their low $V_t$. Finally, variability makes $E_{min}$ dramatically rise for thick $T_{ox}$ because of low channel control and thus important RDF-induced $v_{dd}$, according to Eq. (4). Again, the trade-off between variability/short-channel behavior mitigation and low $I_{gate}/I_{sub}$ ratio implies an optimum $T_{ox}$ value.

5.4 Optimum MOSFET selection

Let us now consider the selection of optimum MOSFETs.

In this paper, we restrict the discussion to GP technology flavors (LOP and LSTP). Within a given flavor, circuit designers have 2 degrees of freedom for device selection: $V_t$ and $L_g$. $V_t$ can be chosen between 2 or 3 discrete values, while $L_g$ can take any value higher than the minimum $L_g$. In 45 nm technologies, there are often restrictive design rules that prevent circuits designers from using any $L_g$ for regularity issues. We therefore consider $L_g$ values that are multiples of 5 nm. Fig. 5 shows $E_{min}$ for the devices with 3 considered $V_t$ vs. $L_g$, within the GP technology flavor.

A $T_{ox}$ reduction yields on one hand a better channel control by the gate, which results in an improved $S$. On the other hand, it increases the load capacitance [14]. The resulting $C_L S^2$ factor is slightly improved for thin-oxide devices. No minimum $C_L S^2$ is seen for the considered $T_{ox}$ range because $C_L$ increases slowly with $T_{ox}$ reduction. This is a surprising observation because, although $C_L$ is dominated by parasitic capacitances, one could expect a strong $C_L$ reduction from fringing and overlap capacitance mitigation when increasing $T_{ox}$. Nevertheless, in nanometer technologies the outer fringing capacitance from gate to source/drain electrodes is very important [6, 16] as the gate electrode thickness is high and the spacer width is small. This capacitance component hardly depends on $T_{ox}$ and thus sets, together with junction capacitance, a lower bound on the achievable $C_L$ reduction.

For ideal devices, a thinner $T_{ox}$ implies a lower $V_t$, which translates into an $E_{min}$ increase as the devices leave the subthreshold regime, as explained in previous section. The DIBL-induced energy overhead is lower for thin-oxide devices, whereas the $I_{gate}$ overhead is only important for thin-oxide devices, which feature a high $I_{gate}/I_{sub}$ ratio despite their low $V_t$. Finally, variability makes $E_{min}$ dramatically rise for thick $T_{ox}$ because of low channel control and thus important RDF-induced $v_{dd}$, according to Eq. (4). Again, the trade-off between variability/short-channel behavior mitigation and low $I_{gate}/I_{sub}$ ratio implies an optimum $T_{ox}$ value.

5.4 Optimum MOSFET selection

Let us now consider the selection of optimum MOSFETs.

In this paper, we restrict the discussion to GP technology flavors (LOP and LSTP). Within a given flavor, circuit designers have 2 degrees of freedom for device selection: $V_t$ and $L_g$. $V_t$ can be chosen between 2 or 3 discrete values, while $L_g$ can take any value higher than the minimum $L_g$. In 45 nm technologies, there are often restrictive design rules that prevent circuits designers from using any $L_g$ for regularity issues. We therefore consider $L_g$ values that are multiples of 5 nm. Fig. 5 shows $E_{min}$ for the devices with 3 considered $V_t$ vs. $L_g$, within the GP technology flavor.

Figure 4: Impact of device parameters on $E_{min}$ ($V_{dd}$ implicitly adapted to $V_{min}$). Solid lines represent $E_{min}$ from ideal devices (square markers) to real devices (circle markers) by successively enabling DIBL, gate leakage and variability. The previously-reported $C_L S^2$ figure of merit [5] is plotted with dashed lines for comparison purpose. $E_{min}$ significantly deviates from $C_L S^2$ trend because of DIBL, variability and gate leakage.

Figure 5: $E_{min}$ vs. gate length in multi-$V_t$ technology (considered $V_t$ values are 0.27, 0.35 and 0.44V).

For high-$V_t$ devices, $E_{min}$ is the highest because of its high $I_{gate}/I_{sub}$. Therefore, an $L_g$ upsize further degrades $E_{min}$ by worsening this ratio. For low- and std-$V_t$ devices, $E_{min}$ is comparable at minimum $L_g$. When upsizing $L_g$, $E_{min}$ is first improved thanks to variability and short-channel behavior mitigation but is then degraded by higher $I_{gate}/I_{sub}$ ratio, as detailed in Section 5.1. This ratio is lower for low-$V_t$ devices and $L_g$ can thus further be upszied to mitigate more efficiently variability and short-channel behavior, while keeping $I_{gate}/I_{sub}$ low. Low-$V_t$ devices with 45nm $L_g$ yield 35% $E_{min}$ reduction, as compared to baseline 35nm-Lg std-$V_t$ devices, thereby making low-$V_t$ mid-Lg devices the optimum choice for minimum-energy subthreshold logic.

Table 1 shows the subthreshold features of the optimum MOSFET. It features improved subthreshold characteristics at the expense of higher $I_{gate}/I_{sub}$ ratio. This indicates that, for energy concern, it is worth tolerating higher $I_{gate}$ to limit $S$, DIBL and variability. Table 2 summarizes the subthreshold circuit performances with baseline and optimum MOSFETs. At minimum energy point, optimum MOSFET selection increases $T_{del}$ because of lower $V_{min}$. However, at iso-$V_{dd}$ ($V_{min}$ of circuit with baseline devices), $T_{del}$ is reduced by 50% with still 20% $E_{min}$ saving while $E_{min}$ at iso-$T_{del}$ is reduced by 30%. In any cases, optimum MOSFETs yield important energy reduction with a mitigation of $T_{del}$ variability as an extra benefit.
Table 1: Subthreshold MOSFET characteristics

<table>
<thead>
<tr>
<th>Device type</th>
<th>$S$ [mV/dec]</th>
<th>$\eta$ [mV/V]</th>
<th>$I_{on}$ var.</th>
<th>$I_{gate}/I_{sub}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>92.5</td>
<td>183</td>
<td>29.7×</td>
<td>0.09</td>
</tr>
<tr>
<td>Optimum</td>
<td>81.9</td>
<td>83</td>
<td>13.1×</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 2: Subthreshold 8-bit multiplier performances

<table>
<thead>
<tr>
<th>Device type</th>
<th>$E_{tot}$ [μJ]</th>
<th>$V_{id}$ [V]</th>
<th>$3\sigma$ WC $T_{del}$ [μs]</th>
<th>$T_{del}$ var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline @$V_{min}$</td>
<td>28.8</td>
<td>0.39</td>
<td>0.14</td>
<td>1.76×</td>
</tr>
<tr>
<td>Optimum @$V_{min}$</td>
<td>18.6</td>
<td>0.31</td>
<td>0.39</td>
<td>1.52×</td>
</tr>
<tr>
<td>Optimum @iso-$V_{id}$</td>
<td>22.9</td>
<td>0.39</td>
<td>0.07</td>
<td>1.36×</td>
</tr>
<tr>
<td>Optimum @iso-$T_{del}$</td>
<td>20.1</td>
<td>0.35</td>
<td>0.14</td>
<td>1.43×</td>
</tr>
</tbody>
</table>

We carried out the same optimum device selection in the industrial general-purpose 45 nm technology considered in Fig. 1. As this technology is a dual-$V_t$ technology, the lowest-$V_t$ device was already considered as the standard device in Fig. 1. Nevertheless, simulations show that an $L_g$ upsine from 35 to 50 nm alone leads to 35% $E_{min}$ improvement with only 5% MOSFET area overhead (at constant gate-to-contact distance). This overhead is assumed to weakly impact the Silicon area of the full circuit because interconnection area remains unchanged. As shown in Fig. 6, this optimum device selection yields an $E_{min}$ level at 45 nm node lower than at 90 nm node with a reduced Silicon area.

6. CONCLUSION

In this article, we analyzed the minimum-energy point of subthreshold circuits in nanometer CMOS technologies. Simulations of a benchmark multiplier show that direct porting to 45 nm technology leads to $E_{min}$ overhead. We confirmed that this overhead partly comes from high variability and subthreshold swing $S$. We reported and demonstrated by circuit simulation and analytical modeling that DIBL and gate leakage contribute to this overhead as much as variability. We then investigated the impact of nanometer MOSFET parameters $L_g$, $V_t$ and $T_{off}$ on $E_{min}$ and we showed that improving $E_{min}$ results from a trade-off between variability and DIBL mitigation vs. reduction of the $I_{gate}/I_{sub}$ ratio. This new trade-off in nanometer technologies completely outweighs previously-reported $C_L$ vs. $S$ trade-off.

We showed that selecting low-$V_t$ mid-$L_g$ devices in a 45 nm technology leads to 35% saving in minimum energy, thereby bringing $E_{min}$ back to the level of a 90 nm technology. It delivers good speed performances, reduced delay variability and less than 5% area overhead, as compared to baseline devices in 45 nm technology. If the application features timing constraints (e.g. in real-time applications), the proposed optimum MOSFET selection can further be combined with the technology flavor selection proposed in [19] to meet the constraints at the minimum-energy point.

This study draws a new route for device optimization towards ultimate minimum-energy subthreshold logic. It indicates that efforts should be devoted to minimizing subthreshold swing, DIBL and variability, while gate leakage increase can be tolerated provided that it remains below the subthreshold leakage level.

7. ACKNOWLEDGMENTS

This work was funded by the FNRS and the Walloon Region of Belgium under TABLOID and E.USER projects.

8. REFERENCES