Leakage Power Reduction by Dual-Vth Designs Under Probabilistic Analysis of Vth Variation

Michael Liu, Wei-Shen Wang, and Michael Orshansky

Department of Electrical and Computer Engineering
University of Texas at Austin

1. Abstract

Low-power circuits are especially sensitive to the increasing levels of process variability and uncertainty. In this paper we study the problem of leakage power minimization through dual Vth design techniques in the presence of significant Vth variation. For the first time we consider the optimal selection of Vth under a statistical model of threshold variation. Probabilistic analytical models are introduced to account for the impact of Vth uncertainty on leakage power and timing slack. Using this model, we show that the non-probabilistic analysis significantly (by 3x) underestimates the leakage power. We also show that in the presence of variability the optimal value of the second Vth must be about 30mV higher compared to the variation-free scenario. In addition, this model provides a way to compute the optimal value of the second Vth for a variety of process conditions.

Categories & Subject Descriptors:
C.4 **[Performance of Systems]** – *modeling techniques*
General Terms:
Reliability, experimentation
Keywords:
Power minimization, variability, yield

2. Introduction

The minimization of leakage power becomes the dominant concern of the nanometer scale CMOS designs, and is part of the general struggle to contain the increase in the overall circuit power consumption. At the 90nm technology node, leakage power may make up 42% of total power [1][2]. The primary reason for this increase in leakage power is the reduction of threshold voltage (Vth) of devices, which is causing an exponential increase in leakage current.

Dual or multiple threshold voltage processes are becoming more popular due to the rising leakage current levels (Ioff) of ultra-small MOSFETs [3][6][7]. In order to maintain sufficient MOS current drive, a simultaneous reduction of Vth and Vdd is needed. This leads to an exponential increase in leakage current. Introducing a second Vth allows maintaining the overall high performance, while reducing leakage current by setting some transistors to high Vth. While performance difference between the high and low Vth transistors is roughly 2X, the leakage current differs by almost 30X [3]. The possibility of setting some transistors to high Vth is a potent way to reduce leakage.

The major contribution of this paper is the analysis of the dual Vth design methodology in the presence of large variation in the threshold voltage. Random variation in Vth is caused by the fluctuation in the number and location of the dopant atoms in the channel of the MOS transistor. The magnitude of the variation in Vth is growing as devices shrink [4][11][12]. Even small fluctuations in dopants will cause a large change in Vth. In order to improve the effectiveness of the dual Vth method in reducing leakage power, we must treat Vth probabilistically.

In this paper we derive a set of analytical models that allow the probabilistic analysis of leakage power within the dual-Vth design methodology. The equations that we derive can also be used to probabilistically describe the leakage power, in general. A specific issue that we seek to answer is the way in which the values of the two Vth can be selected in the presence of variability, which is, we seek to find the optimal separation between the nominal values of low and high Vth such that the overall power savings are optimized. With large Vth variability, the separation must be big enough to overcome the statistical noise. Without the loss of generality in our work, we assume that the value of lower Vth is fixed by the timing requirements [8], and thus are focusing on the optimal selection of the high Vth. Some prior work [5][7] considered the problem non-statistically. We are the first to introduce a rigorous probabilistic description of the problem.

Our model demonstrates that the true average leakage power is three times as large as that predicted by a non-probabilistic model. The models shows that a significantly higher Vth, compared to previous models, is needed to achieve a substantial reduction in leakage power: the optimal value of the second Vth must be about 30mV higher compared to the variation-free scenario. The rest of the paper is organized as follows. Section 3.1 introduces the probabilistic model of static power, and section 3.2 describes the probabilistic timing model. In section 4, the results are presented and analyzed.

ISLPED'04, August 9–11, 2004, Newport Beach, California, USA.

3. Minimizing Static Power Under a Probabilistic Vth Model

In this section, we develop a probabilistic model to describe static power reduction using a dual Vth design under Vth variability. We seek a model and an optimization strategy that will allow us to minimize leakage current (power), while ensuring that the clock frequency satisfies certain requirements. The traditional way is to formulate the power reduction problem as a constrained optimization problem:

$$\min P_{leak},\ s.t.\ f_{clock} \geq f^*$$

However, in contrast to the above model, we will treat *both* P_{leak} and delay (frequency) probabilistically. Also, in our actual formulation, we normalize the power of a dual Vth design to that of the single Vth design, and use it as a measure of static power optimization. The dual-to-single leakage power ratio (R) is

$$R = P_{dual} / P_{single}$$

Now, considering the probabilistic nature of our formulation, we will seek to minimize the expected value of the power ratio, R, under the constraint that with a certain probability the frequency target will be met:

$$\min E\{R\},\ s.t.\ \mathrm{P}\{f_{clock} \geq f^*\} \geq \alpha$$

In the following sections we derive the appropriate probabilistic models for power ratio and delay.

3.1 Model of Static Power Optimization

To calculate the minimum static power achievable in a dual Vth design, we first model the ratio of static power with respect to the two threshold voltages. The amount of static power optimization is defined as the ratio of the static power of a dual Vth design compared to the static power in a single Vth design so this value is less than one. First, we follow [1] and [2] to formulate the static power improvement without the variation in Vth, and then modify the formulation to include a probabilistic Vth model. In the model proposed by [1] and [2], the circuit is modeled as a collection of non-crossing combinational logic paths. Leakage power is reduced by assigning a subset of gate stages to higher Vth. To model the leakage power, we assume that an arbitrary fraction of the total paths gate stages can be assigned to higher Vth. Since the total transistor width and gate stages are largely proportional, we can model the power as a ratio of gate stages. Then the dual-to-single leakage power ratio (R) can be modeled by:

$$R = 1 - \frac{N_h}{N}\left(1 - 10^{-\frac{V_{th}^h - V_{th}^l}{s}}\right) \qquad (1)$$

where s is the subthreshold swing, V_{th}^l and V_{th}^h are the low and high threshold voltages, N_h is the number of gate stages in the entire circuit set to the high threshold voltage, and N is the total number of gates stages in the circuit.

We now modify Eq. 1 to allow the probabilistic treatment of threshold voltage variation. We model threshold voltage as a random, Gaussian variable. Empirical evidence suggests that variation in threshold voltage can be modeled by normal (Gaussian) distribution [1]. Under the assumption that V_{th} is a Gaussian random variable, the power ratio R follows a lognormal distribution. Although the fraction of gates assigned to the high threshold voltage, N_h/N , is also a function of the threshold voltage, a simulation suggests that Eq. 2 is a reasonable approximation of $E[R]$, which allows us to simplify the derivation. The expected value of a log-normally distributed function can be found in closed form from Eq. 1[14]:

$$E[R] = 1 - \frac{N_h}{N}\left(1 - g(V_{os})\right) \qquad (2)$$

where $V_{os} = V_{th}^h - V_{th}^l$, and

$$g(V_{os}) = \exp\left(-\bar{V}_{os} \cdot \frac{\ln(10)}{s} + \left(\frac{\sigma_{V_{th}}^2}{2}\right) \cdot \left(\frac{\ln(10)}{s}\right)^2\right) \qquad (3)$$

where $\bar{V}_{os} = E[V_{os}]$ and $\sigma_{V_{th}}^2 = \sigma_{V_{th}^h}^2 + \sigma_{V_{th}^l}^2$.

Furthermore, the variance of R is given by

$$\sigma_R^2 = \left(\frac{N_h}{N}\right)^2 \left(g(V_{os})\right)^2 \left(e^{\sigma_{V_{th}}^2 \left(\ln(10)/s\right)^2} - 1\right) \qquad (4)$$

Now the power ratio R represents the ratio of the static power of a dual Vth design and that of a single Vth design taking into account the statistical nature of the threshold voltage. In our work we would analyze the impact of V_{th}^h and σ_{Vth} on the power ratio.

In the next section we derive a probabilistic model of path delay that allows us to express the ratio of gates stages set to high Vth and thereby allow us to calculate the expected power ratio.

3.2 Probabilistic Circuit Delay Modeling

In a dual Vth design methodology, the static power is minimized by setting the fraction of gates on paths with timing slack to a higher Vth. This makes the paths slower but less leaky. In this section, we provide a probabilistic description of path delay degradation as a result of setting a portion of the gate stages on a path to a higher Vth.

Let $\gamma = d_h / d_l$ be the degradation of delay per gate stage, with d_h and d_l corresponding to the gate delay at high and low Vth. Using the alpha-power law model [15], we can show that this ratio can be described as:

$$\gamma = \left(\frac{V_{dd} - V_{th}^l}{V_{dd} - V_{th}^h}\right)^{\alpha} \qquad (5)$$

where V_{dd} is the power supply voltage, and V_{th}^{l} and V_{th}^{h} are the low and high threshold voltages. We now describe the degradation of delay of a path when a certain fraction of its gates has been set to high Vth. Let

$$\Delta D = D_{lh} - D_{l} \tag{6}$$

be the change in delay of a path after a fraction of gates along the path is set to high Vth. Here, D_l is the delay of a path with all its gates at low Vth, and D_{lh} is the delay of a path with a mixture of gates at low and high Vth. We also rely on the observation that, to a good approximation, path delay is proportional to the total path capacitance [1] and [2], which, we in turn approximate by gate stages. Combining (5) and (6), we can show that:

$$\Delta D = \sum_{i}^{n_h} \left(d^h - d^l\right) = \sum_{i}^{n_h} d^l\left(\gamma_i - 1\right) = d^l \sum_{i}^{n_h} \left(\gamma_i - 1\right) \tag{7}$$

where n is the number of gate stages along a path, and n_h is the number of gate stages assigned to high Vth.

Due to the uncertainty in threshold voltage, path delay is needed to be modeled probabilistically. To do this we first derive the variance of path delay (Eq. 7). In order to simplify the derivation, it is convenient to include all Vth-induced variability into γ, that is, we assume that only the difference between the low and high threshold voltages is a random quantity. This reduction can be easily arrived, and does not jeopardize the generality of our framework. We further assume that the individual threshold voltage variations are uncorrelated. Then, the path delay variance is given by

$$\begin{aligned} Var\{\Delta D\} &= \sigma_{\Delta D}^2 = Var\left\{d^l \sum_{i}^{n_h}(\gamma_i - 1)\right\} \\ &= d^{l^2} Var\left\{\sum_{i}^{n_h}(\gamma_i - 1)\right\} = d^{l^2} \cdot n_h \cdot Var\{\gamma\} \end{aligned} \tag{8}$$

Defining $V_T = V_{dd} - V_{th}^h$ and $V_{os} = V_{th}^h - V_{th}^l$ we can re-write Eq. 5 as

$$\gamma = \left(1 + \frac{V_{os}}{V_T}\right)^{\alpha} \tag{9}$$

We now account for the statistical variation of Vth by lumping this variability into Vos: $V_{os} = \bar{V}_{os} + \Delta V_{os} \sim N(\bar{V}_{os}, \sigma_{V_{th}}^2)$. To finish the derivation, we use the statistical delta-method to find the variance and mean of γ [8]:

$$\gamma = \left(1 + \frac{V_{os}}{V_T}\right)^{\alpha} \approx 1 + \frac{\alpha V_{os}}{V_T} = 1 + \frac{\alpha(\bar{V}_{os} + \Delta V_{os})}{V_T} \tag{10}$$

The variance of γ, which follows from the statistical delta method, is given by

$$Var\{\gamma\} = \frac{\alpha^2}{V_T^2} Var\left(\Delta V_{os}\right) = \frac{\alpha^2}{V_T^2}\sigma_{Vth}^2 \tag{11}$$

and the mean of γ is given by

$$\gamma_o = E[\gamma] = 1 + \alpha\left(\frac{\bar{V}_{os}}{V_T}\right) \tag{12}$$

By modeling the variability of Vth in γ, the mean path delay is

$$E\left[\Delta D\right] = n_h d^l\left(\gamma_o - 1\right) \tag{13}$$

and the variance of the path delay is

$$Var\{\Delta D\} = \sigma_{\Delta D}^2 = d^{l^2} \cdot n_h \cdot \frac{\alpha^2}{V_T^2}\sigma_{Vth}^2 \tag{14}$$

Having derived the mean and variance of the path delay, we go on to specify the delay constraints in the next section.

3.3 Finding Optimal Vth Separation Under the Probabilistic Models

In this section we derive the analytical framework for finding the optimal value of the high Vth under the statistical description of path delays and power consumption. The objective is to minimize the expected power ratio, $E[R]$, under the timing constraint that no path delay exceeds the critical path delay. In order to make the analysis tractable, we use a simplified model in which the circuit consists of a collection of M non-crossing paths. In contrast to earlier approaches, this constraint is formulated probabilistically, in terms of the probability of meeting the timing constraint:

$$P\{\text{ckt delay } \le T\} = \{\max\{D_1, \ldots D_M\} \le T\} = \alpha_d$$

Because we assume that the paths are non-crossing,

$$P\{\text{ckt delay } \le T\} = \prod_{i=1}^{M}\{D_i \le T\} = \alpha_d$$

For a specified confidence level α_d, we can compute the probability that every path does not exceed the timing constraint. This probability is given by

$$P\left\{(D_i \le T)\right\} = \alpha_d^{1/M} \tag{15}$$

The timing constraint that we enforce is that no path delay can be greater than the critical path delay, which is normalized to one. We can express this constraint by first letting S_i be the amount of slack in a single path.

$$S_i = \left(1 - D_L^i\right) \tag{16}$$

To insure that no path delay is greater than the critical path delay, we must satisfy:

$$P\{\Delta D_i \le S_i\} = \alpha_c \tag{17}$$

Since we assume the path delays to be described by the Gaussian distribution, we can re-write (17) as

$$\Delta D_i^{\alpha} = E[\Delta D_i] + \phi^{-1}(\alpha_c) \cdot \sigma_{\Delta D_i} \tag{18}$$

where ϕ is the cumulative distribution function of the standard normal distribution. Any value of α_c can be used, for example, a convenient value is $\alpha_c = 99.7\%$, such that $\phi^{-1}(\alpha_c) = 3$.

With the constraint on delay given in Eq. 17 we can now find the number of gates along a path to set to the high Vth. As follows from Eq. 1, the power ratio R is linearly proportional to the number of gate stages that we can set to high Vth in the entire circuit. Note that

$$N_h = N \cdot N_{ave}^h \tag{19}$$

where N_{ave}^h is the average ratio of gates set to high Vth per path, which can be found by evaluating:

$$N_{ave}^h = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{n_h}{n}\right)_i \tag{20}$$

where M is the number of paths in the circuit.

We can compute the value of $(n_h/n)_i$ per path by considering the amount of timing slack per path. Using Eq. 13, 14, and 17 we may re-write Eq. 18 as:

$$\Delta D_i^{\alpha} = n_h d^l(\gamma_o - 1) + \phi^{-1}(\alpha_c)\cdot d^l \cdot \sqrt{n_h}\cdot\frac{\alpha}{V_T}\sigma_{V_{th}} \le S_i \tag{21}$$

We can now set an analytical quadratic equation to find n_h per path as a function of S_i. To quantitatively evaluate the average ratio of gates set to the higher Vth, we assume a specific shape of the initial path delay distribution, $p(D_L)$. First we use the triangular distribution that has been shown to be characteristic of many circuits [10]. To make the analysis simpler, the model normalizes all path delays to critical path delay. Then, integrating $n_h(S_i)$ for individual paths over the entire path delay distribution we get:

$$N_{ave}^h = \frac{\int_0^1 n_h(S_i)\,p(D_L)\partial D}{\int_0^1 n_i\,p(D_L)\partial D} \tag{22}$$

where n_i is the number of gate stages per path. The analytical integration of Eq. 22 is possible; however, it does not provide any particular insight into the nature of the problem. In our implementation, we used numerical integration to compute the value of N_{ave}^h for a distribution of specific form. Then, we can link it back to the original equation for the expected power ratio, Eq. 2, such that the result is only a function of the Vth separation:

$$E[R] = 1 - N_{ave}^h \cdot (1 - g(V_{os})) \tag{23}$$

This equation can now be used to explore the dependence of the expected power ratio on Vth separation, and for finding the optimal value of the higher Vth.

4. Analysis and Results

We have implemented the analytical results developed in the earlier sections in a Matlab-based analysis environment. This is a fast implementation that allows us to rapidly consider multiple scenarios with respect to the magnitude and character of Vth variation, other circuit parameters, such as the value of Vdd, and the confidence level (α).

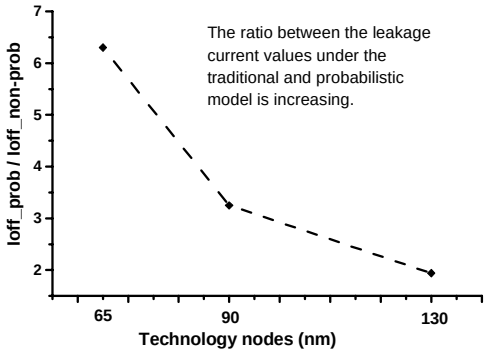

Figure 1. The ratio between the expected leakage current and non-probabilistic leakage current estimates for each technology node. We use σ_{Vth} equal to 30, 40, and 50 mV for each technology node.

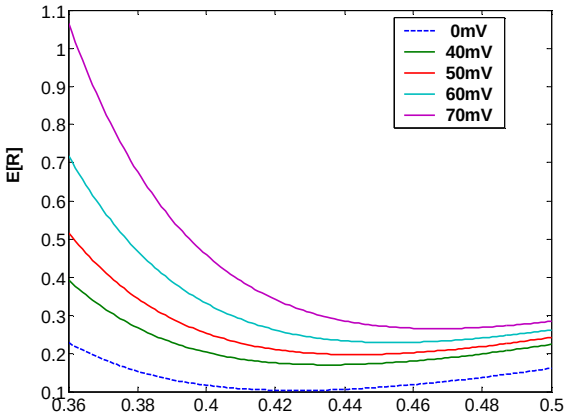

Figure 2. The $E[R]$ vs. the value of the higher Vth for different values of σ_{Vth}. (Vdd=1V, Vth^l=0.3V).

First, we use the developed formulation to demonstrate the importance of accounting for the statistical nature of Vth variation, in the leakage power analysis. Figure 1 shows the difference between the leakage current values under the non-probabilistic description and the new model. Thus, in the probabilistic case we are minimizing from a 3x larger static current in the probabilistic model compared to the non-probabilistic model. This graph was made by setting the high Vth equal to the low Vth and computing the relative leakage current.

Figure 2 shows the power ratio of non-probabilistic and probabilistic models. The minimum point for R and $E[R]$, is 0.10 and 0.26 for the non-probabilistic and probabilistic case with σ_{Vth}=60mV, respectively. High variation of threshold voltage results in substantial leakage power dissipation, which is in agreement with Figure 1.

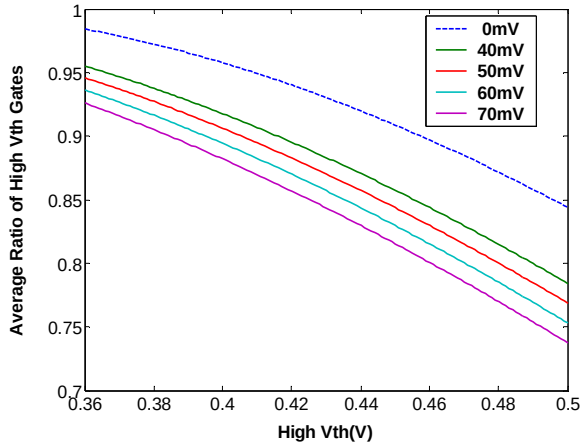

Figure 3. Average ratio of high Vth gates vs. higher Vth for different values of σ_{Vth}. (Vdd=1V, Vth^l=0.3V).

Besides, the non-probabilistic model of static power optimization under a dual-Vth approach skews the actual gains, and does not allow one to pick the truly optimal values of Vth. The Vth value that minimizes the expected static power is approximately $Vth^h=Vth^l+0.15Vdd$ with σ_{Vth}=60mV. The non-probabilistic model, on the other hand, predicts that the minimum static power occurs at a high Vth value of about 0.12·Vdd greater than low Vth.

The greater the variation in Vth the higher the value of high Vth has to be to minimize static power. Hence, the probabilistic model for Vth becomes more important as the variation in Vth gets larger, which is predicted to occur given current scaling trends.

Figure 3 shows the average ratio of high Vth gates N_{ave}^{h} under different Vth variations for the triangular path delay distribution. As expected, high Vth variation leads to low N_{ave}^{h}, thus results in high static power dissipation.

Figure 4 shows the dependence of $E[R]$ on the quantile point of the probabilistic path distribution. Our basic approach is to take the 3-sigma point of ΔD^{α}. Clearly, the higher the confidence level that the timing constraints are not violated, the fewer the gates that can be assigned to high Vth, and the higher is the power dissipation. However, the optimal high Vth value appears to be weakly dependent on the confidence level. Comparing the use of the 50^{th} percentile (mean delay) to the 3-sigma point, we find that the optimal Vth value changes by 9mV.

Not all circuits can be approximated by the triangular path delay distribution and here we also include the results for a sloped path distribution, where most of the paths are near the critical path, and a flat path distribution. As expected, the achievable power savings in a dual Vth approach for both distributions is smaller due to the greater number of paths with the delay near the critical path. Table 1 gives a summary of the optimal Vth value for each case.

We find from the analysis of different distributions that although the amount of power savings is different, the optimal value of the higher Vth changes only slightly. In addition, since the dependence of $E[R]$ on Vth is rather weak for a range of Vth values near the optimum, any Vth value in this range provides about the same amount of power savings. That is, there is a range in which raising Vth provides diminishing returns in terms of power savings. As a rule of thumb, for most typical path distributions, a second Vth voltage of 0.14·Vdd provides the highest amount of savings.

Figure 4 would seem to suggest that any value of high Vth from Vth^l+0.11·Vdd to Vth^l+0.18·Vdd would be a suitable choice, as all result in approximately the same amount of power dissipation. However, there is a cost of further Vth increase: removing slack from the sub-critical path prevents using other circuit optimization techniques for power reduction, such as, transistor sizing. The key is to use the technique that best trades slack for lower power, similar to work done in [9] and [10]. If the high Vth value is calculated on a pre-optimized path delay distribution, then a value for high Vth can be chosen that represents the best power-delay tradeoff.

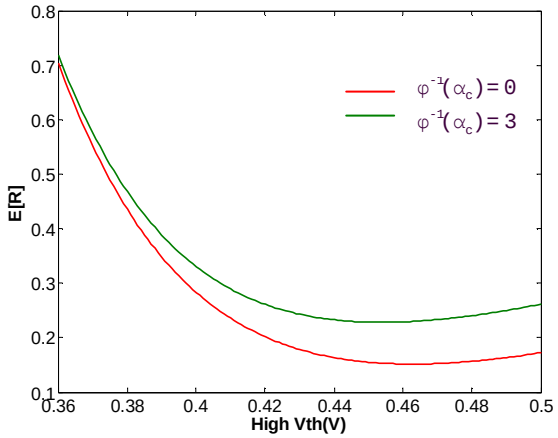

Figure 4. $E[R]$ vs. the value of higher Vth for the mean delay and 3-sigma point (Vdd=1V, Vthl=0.3V, and σ_{Vth}=60mV).

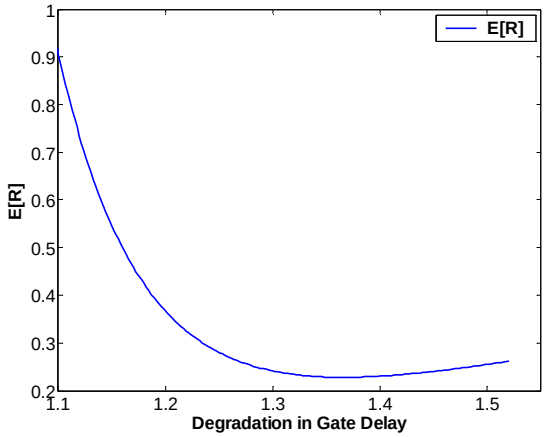

Figure 5. Degradation in gate delay, γ, versus $E[R]$ (σ_{Vth}= 60mV). Minimum point of $E[R]$ is at γ=1.37.

Table 1. The value of optimal higher Vth for different initial path delay distributions (σ_{Vth}= 60mV).

Distribution Type	Optimal Higher Vth	Minimum $E[R]$	N^h_{ave}
Triangle	0.15·Vdd	0.23	0.82
Flat	0.14·Vdd	0.41	0.64
Sloped	0.13·Vdd	0.52	0.54

To show this in Figure 5, we plot the degradation of $E[R]$ as a function of the increased gate stage delay (γ). The graph is similar in shape to Figure 3, because the degradation in delay is nearly linear for small values of the high Vth. It is clear that the value of Vth that minimizes static power is situated at a point of very unfavorable power-delay tradeoff. Comparing Figure 4 and 5, we can see that the minimum power is achieved at a point at which $\gamma = 1.37$. A better value for the high Vth would instead be at a slightly lower value for the high Vth. For example, at a high Vth of Vth^l+0.09·Vdd there is still a 61% savings in power with a gate stage delay of only $\gamma = 1.19$, which would leave considerably more slack for other circuit optimization techniques.

5. Conclusions

In this paper we derive a probabilistic analytical framework to minimize expected static power using a dual Vth design technique. From this analysis we find that the non-probabilistic model severely underestimates the expected leakage current. We also observe that under variability a larger separation between the lower and higher Vth is required to achieve optimal leakage power reduction. This further highlights the increased importance of using a fully probabilistic approach as fundamental variability continues to increase.

6. References

[1] S. Narendra, D. Blaauw, A. Devgan, F. Najm, "Leakage issues in IC design: trends, estimation, and avoidance," *Proc. of ICCAD*, 2003.

[2] J. Kao, S. Narendra, A. Chandrakasan, "Subthreshold Leakage Modeling and Reduction Techniques," ICCAD, 2002, pp. 141-149.

[3] S. Sirichotiyakul, et al, "Stand-by power minimization through simultaneous threshold voltage selection and circuit sizing," *DAC*, 1999, pp. 436-441.

[4] A. Asenov, S. Kaya, JH. Davies, "Intrinsic threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations," IEEE Transactions on Electron Devices, vol.49, no.1, Jan. 2002, pp.112-19.

[5] M. Hamada, Y. Ootaguro, and T. Kuroda, "Utilizing surplus timing for power reduction," *Proc. IEEE Custom Integrated Circuits Conference*, pp. 89-92, 2001.

[6] P. Pant, R. Roy and A. Chatterjee, "Dual-Threshold Voltage Assignment with Transistor Sizing for Low Power CMOS Circuits," *Trans. on VLSI*, Vol. 9, No. 2, 4, 2001, 390-394.

[7] A. Srivastava, D. Sylvester, "Minimizing Total Power by Simultaneous Vdd/Vth Assignment," Design Automation Conference, 2003. *Proceedings of the ASP-DAC 2003.,* pp. 400-403, 21-24 Jan. 2003.

[8] M. Hirabayashi, K. Nose, T. Sakurai, "Design methodology and optimization strategy for dual-V_{TH} scheme using commercially available tools," *ISLPED*, pp. 283 – 286, Aug. 2001.

[9] R. Gonzalez, B.M. Gordon, M.A. Horowitz, **"**Supply and threshold voltage scaling for low power CMOS," *IEEE Journal of Solid-State Circuits*, Volume: 32 Issue: 8 , Aug. 1997 pp. 1210 –1216

[10] R.W. Brodersen, M.A. Horowitz, D. Markovic, B. Nikolic and V. Stojanovic. "Method for True Power Optimization," *IEEE/ACM International Conference on Computer Aided Design*, November 2002, pp. 35-42.

[11] 5. Vivek K. De, Xinhhai Tang, and James D. Meindl, "Random MOSFET Parameter Fluctuation Limits to Gigascale Integration(GSI)," *Symposium of VLSI Technology Digest*, pp. 198-199, 1996.

[12] H.-S. Wong, D. J. Frank, P. Solomon, C. Wann, and J. Wesler, "Nanoscale CMOS, " in *Proc. of the IEEE*, Vol. 87, No. 4, pg. 537-570, April 1999.

[13] M. Orshansky, L. Milor, P. Chen, K. Keutzer, C. Hu, "Impact of spatial intrachip gate length variability on the performance of high-speed digital circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Volume: 21 Issue: 5 , pp. 544-553, May 2002.

[14] W. Feller, "Probability Theory and its Applications," *John Wiley & Sons*, second ed., vol. II. 1971.

[15] T. Sakurai, A.R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journal of Solid-State Circuits*, Volume: 25 Issue: 2, pp. 584-594, April 1990.

[16] J.D. Warnock, et al., "The circuit and physical design of the POWER4 microprocessor," *IBM J. of R&D*, Vol. 46, pp. 27-52, Jan. 2002.