



High-performance low-leakage regions of nano-scaled CMOS digital gates under variations of threshold voltage and mobility

Hossein AGHABABA[†], Behjat FOROUZANDEH, Ali AFZALI-KUSHA

(Nanoelectronics Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran 14395, Iran)

[†]E-mail: h.aghababa@ece.ut.ac.ir

Received Sept. 13, 2011; Revision accepted Feb. 17, 2012; Crosschecked Apr. 9, 2012

Abstract: We propose a modeling methodology for both leakage power consumption and delay of basic CMOS digital gates in the presence of threshold voltage and mobility variations. The key parameters in determining the leakage and delay are OFF and ON currents, respectively, which are both affected by the variation of the threshold voltage. Additionally, the current is a strong function of mobility. The proposed methodology relies on a proper modeling of the threshold voltage and mobility variations, which may be induced by any source. Using this model, in the plane of threshold voltage and mobility, we determine regions for different combinations of performance (speed) and leakage. Based on these regions, we discuss the trade-off between leakage and delay where the leakage-delay-product is the optimization objective. To assess the accuracy of the proposed model, we compare its predictions with those of HSPICE simulations for both basic digital gates and ISCAS85 benchmark circuits in 45-, 65-, and 90-nm technologies.

Key words: High-performance circuit, Low-leakage circuit, Manufacturing process variation, CMOS integrated circuit
doi:10.1631/jzus.C1100273 **Document code:** A **CLC number:** TN4

1 Introduction

In the design of complementary metal–oxide–semiconductor (CMOS) integrated circuits, optimizing leakage power consumption and delay is challenging. The difficulty in simultaneous minimization of these two parameters leads the designers to make trade-off between the leakage consumption and delay for different applications. In addition, variations present in scaled technologies have made the design and manufacturing process more complicated. The variations which affect both leakage and performance have made the design even more challenging. The scaling of CMOS has resulted in increasing magnitude of variability. The large variation in CMOS technology could be due to several reasons. Three reasons among the others are widely accepted as the main factors of variability in the manufacturing process (Orshansky *et al.*, 2008).

The first reason is the rise of systematic variations at the front/back end of line (F/B EOL) phase of manufacturing. For instance, optical proximity effects at photolithography and interconnect variation at chemical mechanical polishing (CMP) are good examples of front end and back end of line, respectively (Srivastava *et al.*, 2005; Orshansky *et al.*, 2008). The second reason is that technology scaling and manufacturing tolerances are not correspondingly moving at the same time. For example, the pace at which the effective channel length is reduced is faster than the improvement of the mask fabrication error and mask overlay control. The third reason is that technology is approaching the fundamental randomness in the behavior of silicon structure, leading to dopant fluctuation in the channel of the transistor. This phenomenon is known to be the main contributor for random variation.

Process variations, whether systematic or random, can drastically affect the delay and leakage power consumption of the circuits. Leakage as a

non-negligible component of power dissipation has proved to increase from 18% at 130 nm to 54% at the 65 nm node (Narendra *et al.*, 2003). The researchers are constantly trying to improve the classical methods of statistical leakage estimation, such as Wilkinson's approach, see, e.g., Cheng *et al.* (2009), who proposed an efficient additive statistical method to estimate the moments of a lognormal distribution. They did not propose explicit expressions, however, for leakage power of standard cells versus the variations of device parameters. D'Agostino *et al.* (2009) and Hao *et al.* (2011) proposed approaches to model the leakage power under statistical process variations. However, these approaches are useful only to estimate the statistical distribution of leakage power without proposing any closed-form expression for leakage power under variations. Ye and Yu (2009) proposed an efficient technique to model the leakage power in the presence of process variations. They, however, proposed algorithms only to reduce the computation time. In this work, also, no expression was suggested for clearly separating the variation part and the nominal parts of the leakage consumption.

Variation-aware delay has been always a significant issue in modeling and analysis of circuits as reflected in recent works (Liu and Sapatnekar, 2009; 2010; Guerra *et al.*, 2010). However, these works are mainly algorithm-based methodologies, which try to improve the modeling time through a statistical framework without proposing closed-form expressions for the delay of gates and circuits. There have been works, such as Ramalingam *et al.* (2006), which have tried to improve gate delay modeling. However, the variability was not considered. Okada *et al.* (2003) investigated the delay model under variability but there have been no efforts to differentiate the effects of variability from the nominal part. Miryala *et al.* (2011) proposed a linear model to estimate the delay in the presence of variations. Although this model is easy to use, it may not be accurate for diverse types of variation sources. Chen *et al.* (2011) proposed a technique to capture the delay variations. Their method was based on modeling threshold voltage variability. It, however, did not incorporate other variation parameters such as mobility. Also, it did not propose a closed-form expression for delay under variations. Joshi *et al.* (2010) and Xu *et al.* (2011) also suggested models for the delay of a single transistor

under variations. No model was presented for the delay of a gate.

Our contribution in this paper is to propose a model for the delay and leakage power consumption of digital circuits in terms of the product of the nominal part and variable part. Using this model, in the plane of threshold voltage and mobility, regions for different combinations of performance (speed) and leakage are specified. Based on these regions, the trade-off between leakage and delay is discussed.

2 Modeling the effects of variations on leakage consumption

Static power consumption is a major component of the total power consumption and continues to grow in future technologies. The main contributor to leakage power consumption is the sub-threshold current (Christian, 2007). The equation describing sub-threshold current is given by (Sakurai and Newton, 1990; Sze and Ng, 2007)

$$I_{\text{sub}} = I_0 e^{-V_{\text{th}}/(mV_T)} (1 - e^{-V_{\text{ds}}/V_T}) \approx I_0 e^{-V_{\text{th}}/(mV_T)}, \quad I_0 = \Omega \mu_{\text{eff0}}, \quad (1)$$

where I_0 is the reference static current, V_{th} is the threshold voltage, and m , V_T , and V_{ds} stand for the sub-threshold slope coefficient, thermal potential, and drain-source voltage, respectively. Note that I_0 is a function of the diffusion coefficient (Sze and Ng, 2007), which is related to the mobility via the Einstein relation. In Eq. (1), Ω is a constant and μ_{eff0} is the low-field mobility. Eq. (1) can justify the strong (exponential) dependency of sub-threshold current on the threshold voltage. Since V_{ds} is much larger than V_T in digital circuits, Eq. (1) is simplified to $I_0 \exp(-V_{\text{th}}/(mV_T))$. To calculate the static power, we assume that the static power is the average of two power consumptions for all inputs equal to 1 and all inputs equal to 0. Thus,

$$P_{\text{Static Power}} = \frac{P(\text{all inputs} = 1) + P(\text{all inputs} = 0)}{2}. \quad (2)$$

As shown in Eq. (1), the sub-threshold current depends on I_0 and V_{th} . The reference static current, I_0 , is a function of mobility which could be subject to

variation as a result of systematic sources such as mechanical stress (Scott *et al.*, 1999; Mistry *et al.*, 2004; Andrieu *et al.*, 2005). The threshold voltage is also very susceptible to variations coming from various systematic sources, including layout-dependent systematic variations originating from photolithography effects. Examples include non-rectangular gates (Singhal *et al.*, 2007; Tsai *et al.*, 2008), line-end extension, and tapering (Gupta *et al.*, 2008), which can lead to a threshold voltage shift. In addition, random process variations, such as dopant fluctuation, result in threshold voltage variation (Roy and Asenov, 2005).

The above discussion implies that the leakage consumption is affected by the threshold voltage and mobility in two different ways. To model the impact of the threshold voltage and mobility variations on the leakage of a circuit, we include the threshold voltage shift and low-field mobility multiplier in our leakage expression. These two parameters are denoted by DELVT0 and MULU0, respectively, and defined by

$$V_{th} = V_{th0} + \text{DELVT0}, \quad \mu_{eff} = \text{MULU0} \cdot \mu_{eff0}, \quad (3)$$

where V_{th0} and DELVT0 are zero-bias threshold voltage and zero-bias threshold voltage shift, respectively. Also, μ_{eff0} and MULU0 are low-field mobility and low-field mobility multiplier, respectively. These parameters are also defined as back-annotated instance parameters used in the model of a transistor in HSPICE (Synopsys Corporation, 2008), explained in Pramanik *et al.* (2006), Ma (2009), and Morshed (2009). In fact, HSPICE simulations of digital gates under variations could be performed using these parameters.

Based on the above discussion, we can express the standby leakage consumption under variation as

$$P = P_{nom} \cdot f_p(\text{DELVT0}) \cdot g_p(\text{MULU0}), \quad (4)$$

where $f_p(\text{DELVT0})$ and $g_p(\text{MULU0})$ are the coefficients for modeling the change of the threshold voltage and low-field mobility (under variation), and the nominal standby leakage consumption is given by

$$P_{nom} = V_{DD} I_0 e^{-V_{th}/(mV_T)}. \quad (5)$$

Here, V_{DD} is the supply voltage.

2.1 Modeling $f_p(\text{DELVT0})$

To consider only the variability of the threshold voltage, one may assume that the mobility is not affected by variations, and hence, Eq. (4) is reduced to

$$\hat{P}_1 = P_{nom} f_p(\text{DELVT0}), \quad (6)$$

where

$$f_p(\text{DELVT0}) = e^{\alpha \cdot \text{DELVT0}}, \quad (7)$$

and α is $-1/(mV_T)$ from Eq. (1). Eq. (7) is valid for both states when all input signals are 0 or all input signals are 1. Since in the low-input (high-input) state only the NMOS (PMOS) transistors determine the leakage current, Eq. (7) behaves differently depending on the state of the gate input. Fig. 1 shows the leakage consumption of some primitive gates as a function of DELVT0 (from -50 to $+50$ mV) for both input states. The results, which have been obtained for a 45-nm predictive technology model (PTM) (Arizona State University, 2006) at $V_{DD}=1$ V, include both the predictions of the model and those of HSPICE. The model parameter was extracted by fitting the results to those of HSPICE.

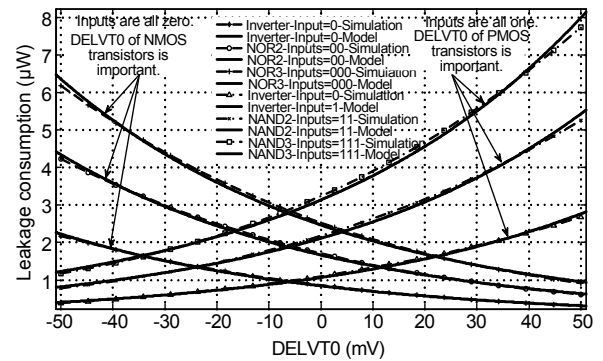


Fig. 1 Standby leakage consumption of primitive digital gates in 45-nm technology versus DELVT0 at different states of input patterns

2.2 Modeling $g_p(\text{MULU0})$

Similarly, one can consider solely the variability of the mobility by assuming that the threshold voltage is not affected by variations, and write Eq. (4) as

$$\hat{P}_2 = P_{nom} g_p(\text{MULU0}), \quad (8)$$

where

$$g_p(\text{MULU0}) = \beta(\text{MULU0} - 1) + 1, \quad (9)$$

where β is a fitting parameter. In Fig. 2, we have plotted the change of the leakage power versus MULU0 for both input states. The results reveal a roughly linear behavior for the dependence of the leakage power consumption on the low-field mobility multiplier. Linear approximation is more accurate when PMOS transistors are involved in determining the leakage current, while in the case of NMOS transistors a second-order polynomial may be more accurate. For simplicity, however, we keep the linear model given by Eq. (9) for both input states.

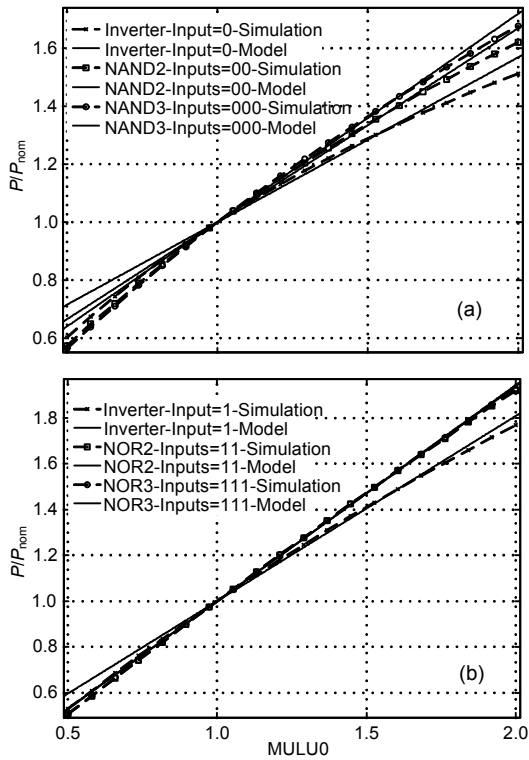


Fig. 2 The ratio of standby leakage to nominal leakage consumption of primitive digital gates in 45-nm technology versus MULU0

(a) All input signals are 0; (b) All input signals are 1

3 Modeling the effects of variations on delay

One approach in accurately modeling delay is the use of lookup tables, which are not computationally cost-effective compared to the closed-form delay models (Alpert *et al.*, 2001). The modeling techniques based on closed-form expressions for the delay analysis, however, provide simplicity. In this work, similar to the leakage power consumption, we pro-

pose a closed-form expression for delay using DELVT0 and MULU0. The delay expression is developed based on a classic yet reasonably accurate delay model, known as the Sakurai-Newton delay metric (Sakurai and Newton, 1990), given by

$$t_{\text{sn}} \propto C_L V_{\text{DD}} / (V_{\text{DD}} - V_{\text{th}})^\alpha, \quad (10)$$

where C_L is the load capacitance and α is the alpha-power law parameter. Using Eq. (10) and including the threshold voltage roll-off effect described by the drain-induced barrier lowering (DIBL) phenomenon, the nominal delay of a digital gate can be expressed as (Christian, 2007)

$$\begin{cases} \text{Delay}_{\text{nom}} = \lambda V_{\text{DD}} / [V_{\text{DD}}(1 + \eta) - V_{\text{th0}}]^\alpha, \\ \lambda = \gamma C_L / (\mu_{\text{eff}} C_{\text{ox}} W / L). \end{cases} \quad (11)$$

Here, η is the DIBL coefficient, and λ contains the mobility coefficient (μ_{eff}), the gate capacitance per unit area (C_{ox}), the load capacitance, the aspect ratio of the transistor (W/L), and a constant (γ) accounting for this fact that driving current is not constant during the capacitance charge (Christian, 2007). Note that MULU0 affects the parameter λ in a complicated way as the driving current is not constant during the capacitance charge. Hence, a function of MULU0 affects the parameter λ as a multiplier. Conversely, DELVT0 appears in the denominator of Eq. (11). To model the delay under variations, we separate the effects of DELVT0 and MULU0 by defining two independent multiplying functions of $f_D(\text{DELVT0})$ and $g_D(\text{MULU0})$. Thus, we propose the following form for delay of digital gates under variations:

$$\text{Delay} = \text{Delay}_{\text{nom}} \cdot f_D(\text{DELVT0}) \cdot g_D(\text{MULU0}). \quad (12)$$

Here, $f_D(\text{DELVT0})$ and $g_D(\text{MULU0})$ are functions of zero-bias threshold voltage shift and low-field mobility multiplier, respectively. We propose two independent models for both functions and verify their accuracy using HSPICE simulations.

3.1 Modeling $f_D(\text{DELVT0})$

Assuming that the variations affect only the threshold voltage, Eq. (12) is reduced to

$$\widehat{\text{Delay}}_1 = \text{Delay}_{\text{nom}} \cdot f_D(\text{DELVT0}). \quad (13)$$

We need to find the expression for f_D . By explicitly expressing the variation of the threshold voltage in Eq. (11), one may express delay as

$$\text{Delay} = KV_{DD} / [V_{DD}(1+\eta) - (V_{th0} + \text{DELVT0})]^\alpha. \quad (14)$$

Thus,

$$\frac{\text{Delay}_{\text{nom}}}{\widehat{\text{Delay}}_1} = \left[\frac{V_{DD}(1+\eta) - (V_{th0} + \text{DELVT0})}{V_{DD}(1+\eta) - V_{th0}} \right]^\alpha. \quad (15)$$

Eq. (15) may be rewritten as

$$\frac{\text{Delay}_{\text{nom}}}{\widehat{\text{Delay}}_1} = \left[1 - \frac{\text{DELVT0}}{V_{DD}(1+\eta) - V_{th0}} \right]^\alpha. \quad (16)$$

The following lemma will help to simplify Eq. (16):

$$\lim_{x \rightarrow 0} (1+x)^{1/x} = e, \quad \lim_{x \rightarrow 0} (1-x)^{1/x} = e^{-1}. \quad (17)$$

Eqs. (16) and (17) lead to the following simplified approximation:

$$\widehat{\text{Delay}}_1 \cong \text{Delay}_{\text{nom}} \cdot \exp\left(\frac{\alpha \cdot \text{DELVT0}}{V_{DD}(1+\eta) - V_{th0}}\right). \quad (18)$$

The beauty of Eq. (18) is that the effect of the threshold voltage shift has been modeled as a multiplying function. Using Eqs. (13) and (18), we obtain f_D as

$$f_D(\text{DELVT0}) \cong \exp\left(\frac{\alpha \cdot \text{DELVT0}}{V_{DD}(1+\eta) - V_{th0}}\right). \quad (19)$$

Note that this function depends on the supply voltage.

To determine the accuracy of the model, we have compared the results of model and HSPICE simulation for the ratio of delay under variation to nominal delay versus V_{DD} for different zero-bias threshold voltage shifts (± 50 and ± 20 mV), as shown in Fig. 3. The comparison shows a reasonable accuracy over a relatively wide range of threshold voltage variation. Even for the shift in the threshold voltage of ± 50 mV ($\pm 20\%$ of the nominal value of the threshold voltage $V_{th0}=250$ mV), the model acceptably tracks the data points obtained from HSPICE simulations. Another important issue is that the parameter α in Eq. (19) may serve as a fitting parameter for every single digital

gate in the technology of interest. In Table 1, we have presented this parameter for different digital gates in the 45-, 65-, and 90-nm technologies. The accuracy of the model has been further investigated using the ‘goodness of fit’ parameters (R-square and RMSE). These parameters are given in Table 2. The R-square indicates the accuracy of the model for the variance of the data, while the root mean squared error (RMSE) represents the fit standard error. In our delay analysis for both threshold voltage shift and mobility variation, we measure the pull-down network delay assuming that pull-up and pull-down networks are sized such that they exhibit similar delays.

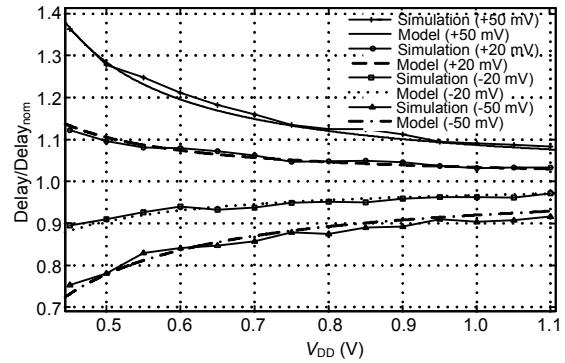


Fig. 3 The ratio of delay under variation to nominal delay of a NAND3 gate in 45-nm technology versus V_{DD} for different zero-bias threshold voltage shifts (± 50 and ± 20 mV)

Table 1 Fitting parameter for the threshold voltage shift

Technology	Fitting parameter α in Eq. (19)				
	Inverter	NAND2	NAND3	NOR2	NOR3
45 nm	0.9	1.1	1.25	0.95	1
65 nm	1	1.25	1.45	1.1	1.2
90 nm	1.1	1.4	1.6	1.2	1.4

3.2 Modeling $g_D(\text{MULU0})$

It is assumed that variations affect only mobility, and hence Eq. (12) is reduced to

$$\widehat{\text{Delay}}_2 = \text{Delay}_{\text{nom}} \cdot g_D(\text{MULU0}). \quad (20)$$

By inspecting the behavior of delay versus MULU0 obtained by HSPICE simulations and the functional dependence of delay on MULU0 defined by Eq. (11), we propose the following relation for $g_D(\text{MULU0})$:

$$g_D(\text{MULU0}) = \frac{\text{MULU0}}{\beta(\text{MULU0}^\gamma - 1) + 1}, \quad (21)$$

Table 2 ‘Goodness of fit’ parameters of modeling of delay under different threshold voltage shifts

Technology (nm)	DELVT0 (mV)	R-square					RMSE				
		Inverter	NAND2	NAND3	NOR2	NOR3	Inverter	NAND2	NAND3	NOR2	NOR3
45	5	0.9748	0.9760	0.9851	0.9799	0.9748	0.010	0.012	0.010	0.009	0.010
45	2	0.9817	0.9873	0.9479	0.9476	0.9269	0.003	0.003	0.006	0.004	0.005
45	-2	0.9420	0.9528	0.8977	0.9246	0.8415	0.004	0.004	0.006	0.005	0.006
45	-5	0.9248	0.9571	0.9217	0.8800	0.7631	0.010	0.010	0.013	0.013	0.018
65	5	0.9813	0.9783	0.992	0.9900	0.9676	0.008	0.011	0.008	0.006	0.013
65	2	0.9787	0.9591	0.9746	0.9824	0.9590	0.003	0.006	0.005	0.003	0.005
65	-2	0.9892	0.9654	0.9596	0.9827	0.9756	0.002	0.004	0.005	0.003	0.003
65	-5	0.9929	0.9862	0.9825	0.9781	0.9574	0.004	0.006	0.007	0.007	0.011
90	5	0.9704	0.9847	0.9744	0.9800	0.9089	0.010	0.008	0.011	0.009	0.025
90	2	0.9697	0.9032	0.9504	0.9746	0.9556	0.003	0.007	0.006	0.004	0.005
90	-2	0.9439	0.9768	0.9374	0.9707	0.9852	0.004	0.000	0.006	0.003	0.003
90	-5	0.9846	0.9816	0.9684	0.9708	0.9535	0.005	0.007	0.010	0.007	0.011

where β and γ are fitting parameters. Note that this function is independent of the supply voltage. Fig. 4 shows a comparison between the prediction of the model and the results of HSPICE simulations for the ratio of delay under variation to nominal delay versus MULU0 for V_{DD} in the range 0.5–1 V in the 45-nm technology. The comparison which is for a three-input NAND gate reveals a very good accuracy for the model. Table 3 shows the fitting parameters β and γ as well as R-square and RMSE of fitting. As can be seen, R-square is larger than 0.99 for all the gates in different technologies. That is, the proposed model (21) tracks the simulation results obtained from HSPICE with a very good accuracy.

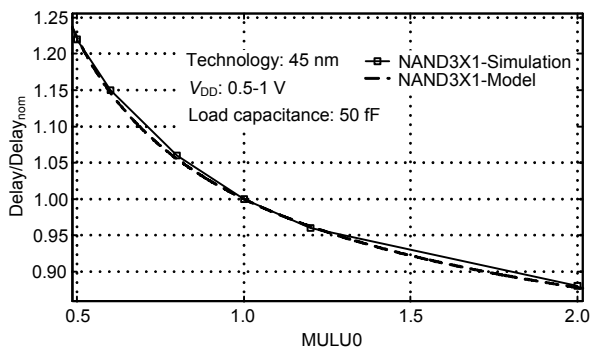


Fig. 4 The ratio of delay under variation to nominal delay of a NAND3 in 45-nm technology against MULU0 (low-field mobility multiplier)

As discussed above, the effects of variation on delay and leakage can be taken into account by multiplying coefficients. One can use Eqs. (4), (7), and (9) to express the leakage power consumption of a CMOS digital gate under variation:

Table 3 ‘Goodness of fit’ parameters of modeling of delay against the low-field mobility multiplier

Gate	Technology	β	γ	R-square	RMSE
Inverter	45 nm	1.088	1.027	0.9999	0.0008
	65 nm	1.105	1.016	0.9989	0.0026
	90 nm	1.112	1.025	0.9988	0.0031
NAND2	45 nm	1.110	1.059	0.9990	0.0035
	65 nm	1.118	1.067	0.9986	0.0046
	90 nm	1.109	1.084	0.9991	0.0038
NAND3	45 nm	1.099	1.113	0.9993	0.0036
	65 nm	1.134	1.086	0.9999	0.0011
	90 nm	1.117	1.115	0.9996	0.0033
NOR2	45 nm	1.101	1.016	0.9995	0.0018
	65 nm	1.084	1.043	0.9991	0.0024
	90 nm	1.108	1.042	0.9984	0.0039
NOR3	45 nm	1.097	1.030	1.0000	0.0004
	65 nm	1.085	1.048	0.9999	0.0007
	90 nm	1.104	1.048	0.9990	0.0032

$$P = P_{nom} e^{\alpha \cdot DELVT0} [\beta(MULU0 - 1) + 1]. \quad (22)$$

Similarly, using Eqs. (12), (19), and (21), the delay under variation may be expressed as

$$\text{Delay} = \text{Delay}_{nom} \cdot \exp\left(\frac{\alpha \cdot DELVT0}{V_{DD}(1 + \eta) - V_{th0}}\right) \cdot \frac{MULU0}{\beta(MULU0^\gamma - 1) + 1}. \quad (23)$$

Using Eq. (22), the total leakage consumption of a circuit composed of several standard cells can be obtained from

$$P_{\text{total}} = \sum_{i=1}^n P(i), \quad (24)$$

where $P(i)$ is the leakage power consumption of the i th cell in the circuit obtained from Eq. (22). Similarly, the total delay of a circuit can be obtained from

$$\text{Delay}_{\text{total}} = \sum_{i=1}^k \text{Delay}(i), \quad (25)$$

where $\text{Delay}(i)$ is the delay of the i th standard cell in the critical path of the circuit obtained by Eq. (23).

4 High-performance low-leakage regions

Eqs. (22) and (23) provide useful information on the domain of DELVT0 and MULU0 to identify the high-performance and low-leakage regions. To identify these regions, first we obtain the contours of $f_p(\text{DELVT0}) \times g_p(\text{MULU0})=1$ for both gate input states of all zeros and all ones (Fig. 5). The contours are the borders of high-leakage from low-leakage regions. Moving on these contours does not change the leakage power consumption.

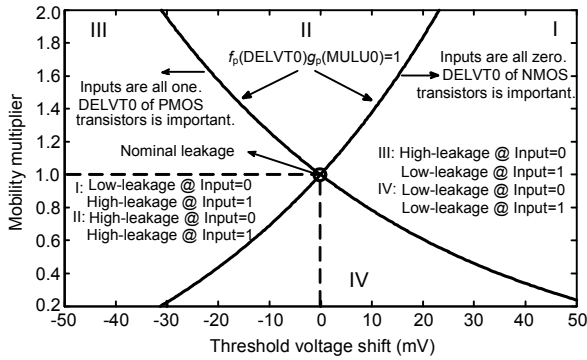


Fig. 5 Contours of $f_p(\text{DELVT0}) \times g_p(\text{MULU0})=1$ of an inverter in 45-nm technology at two states of input patterns

Note that the leakage current is determined by NMOS transistors when the inputs are low and by PMOS transistors when the inputs are high. Thus, the contours of the leakage power consumption move in different directions (Fig. 5). Low- and high-leakage power consumption regions are determined based on the contours that correspond to the nominal values of

leakage for a gate. As shown in Fig. 5, when the threshold voltage shift and mobility multiplier equal 1 and 0, respectively, the contours intersect each other.

Fig. 6 shows the contours of delay for $f_D(\text{DELVT0}) \times g_D(\text{MULU0})=1$ at two supply voltages of 0.45 and 1 V. Similarly, for each V_{DD} , we can distinguish low- and high-performance regions using the contours that correspond to the nominal delay. The intersection region of the low-leakage and high-performance regions is the ideal region where both the leakage power consumption and performance have better values. To find such a region, the delay and leakage contours should be plotted on the same graph. We have drawn the delay and leakage contours in Fig. 7, where regions with different combinations of high/low performance and high/low leakage are indicated.

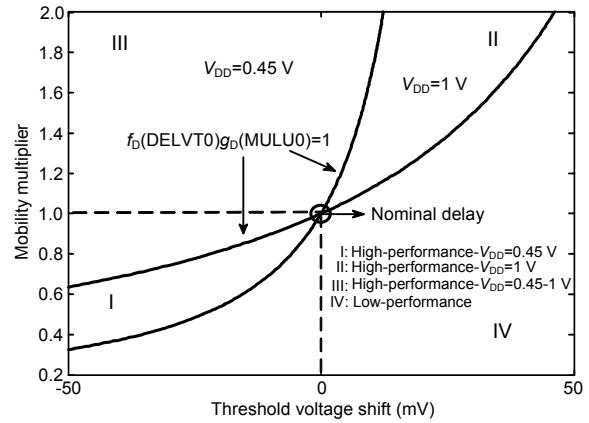


Fig. 6 Contours of $f_D(\text{DELVT0}) \times g_D(\text{MULU0})=1$ of an inverter in 45-nm technology for V_{DD} at two supply voltages of 0.45 V and 1 V

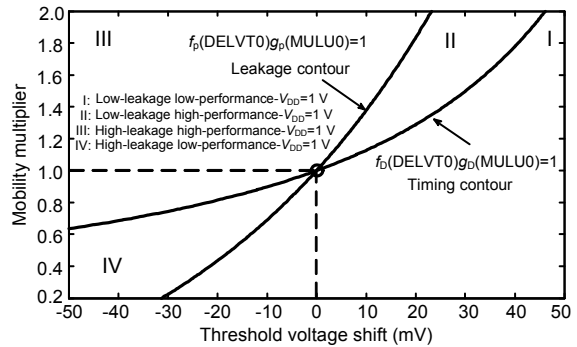


Fig. 7 Contours of $f_p(\text{DELVT0}) \times g_p(\text{MULU0})=1$ and $f_D(\text{DELVT0}) \times g_D(\text{MULU0})=1$ of an inverter in 45-nm technology and the high-performance low-power region

Based on the above analysis, one may use techniques such as mechanical stress (Scott *et al.*, 1999; Mistry *et al.*, 2004; Andrieu *et al.*, 2005) to modify the nominal threshold voltage and mobility such that the operation region moves to the desired region. These regions had not been studied in previous works (Agarwal *et al.*, 2004; Rao *et al.*, 2004; Chen *et al.*, 2008; da Silva *et al.*, 2009) on modeling of delay and leakage and the effects of process variations on them.

As shown in Fig. 7, the points that lie in region II make the digital gate work in high-performance low-leakage state. One can make the design based on a desired point in region II. Note that after fabrication, this point may be changed because of process variations. The variations may change the operation region. To study this effect, we selected a point inside region II (leakage=0.725 μW and delay=240 ps, which correspond to DELVT0=25 mV and MULU0=0.5), and performed Monte Carlo simulations. Fig. 8 shows the results of the analysis, including two cases: (1) sigma(DELVT0)=20 mV and sigma(MULU0)=0.5; (2) sigma(DELVT0)=10 mV and sigma(MULU0)=0.25. Obviously, the variabilities of leakage and delay are lower in case (2) where a larger number of points remain inside the low-leakage high-performance region in the presence of process variations. Finding the optimum point inside the region whose distances from the contours ensure the maximum margin of safety under process variations is out of the scope of our work.

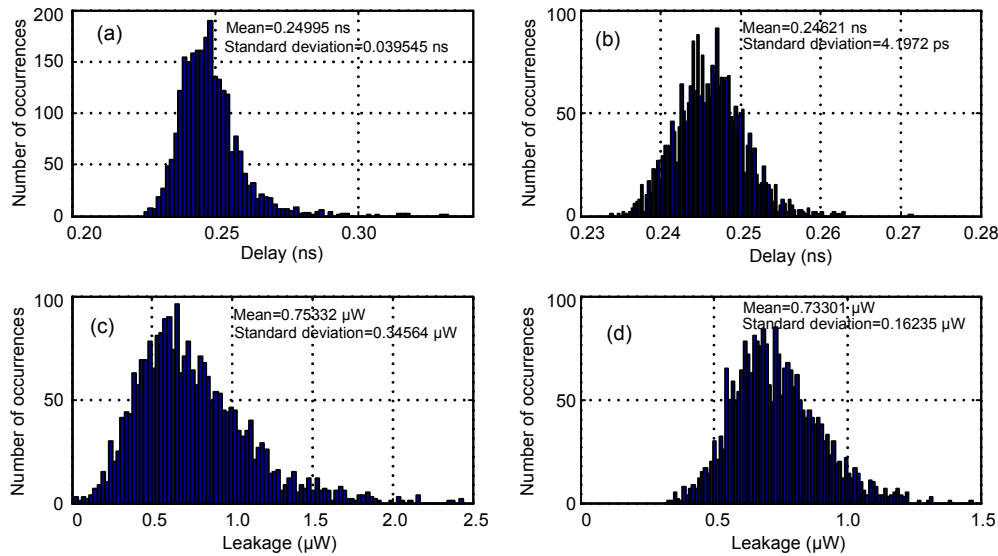


Fig. 8 Probability density functions of delay and leakage of an inverter in 45-nm technology for different values of the standard deviation of DELVT0 and MULU0 in the high-performance low-leakage region

(a) and (c) are delay and leakage histograms, respectively, with sigma(DELVT0)=20 mV and sigma(MULU0)=0.5; (b) and (d) are delay and leakage histograms, respectively, with sigma(DELVT0)=10 mV and sigma(MULU0)=0.25

Finally, note that other figures of merit, such as leakage-delay², which describes the performance of a gate by considering both leakage and delay, may also be analyzed using the above technique. For this figure of merit, for example, the contour may be obtained by setting $f_p(\text{DELVT0})g_p(\text{MULU0})[f_D(\text{DELVT0})g_D(\text{MULU0})]^2=1$. The contour in Fig. 9 separates the area in two regions of low and high values for the figure of merit.

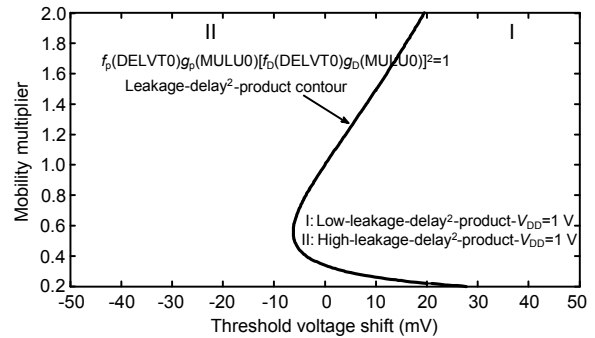


Fig. 9 Contour of $f_p(\text{DELVT0})g_p(\text{MULU0})[f_D(\text{DELVT0})g_D(\text{MULU0})]^2=1$ for an inverter in 45-nm technology

5 Results and discussion

In this section, we investigate the accuracy of the proposed modeling technique for both the leakage and delay. For this purpose, the model predictions and the results of the HSPICE simulations were compared

for primitive digital gates (inverter, NAND2, NAND3, NOR2, and NOR3) and ISCAS85 benchmark circuits in 45-, 65-, and 90-nm technologies. The ranges of the threshold voltage variations were assumed to be $\pm 20\%$ of the nominal value while the low-field mobility multiplier was varied from 0.5 to 2.

The leakage power consumption and delay of an Inverter in the 45-nm technology are illustrated in Figs. 10 and 11, respectively. The accuracy of the model was quantified using the parameters R-square and mean percentage error (MPE). MPE was used to measure the mean error between corresponding points of the model and simulation results, obtained from

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{\text{Model}(i) - \text{SPICE}(i)}{\text{SPICE}(i)} \quad (26)$$

The values of R-square and MPE for the leakage and delay are given in Tables 4 and 5, respectively. The average value of MPE for the leakage (delay) was 4.75% (2.54%) while its R-square was higher than 97% (95%). These figures show a very good accuracy for the model. In the case of ISCAS85 benchmark circuits, we performed 10000 Monte Carlo simulations (assuming $3\sigma/\mu=20\%$) for the threshold voltage and low-field mobility. Note that the leakage power and delay of the primitive gates were calculated from

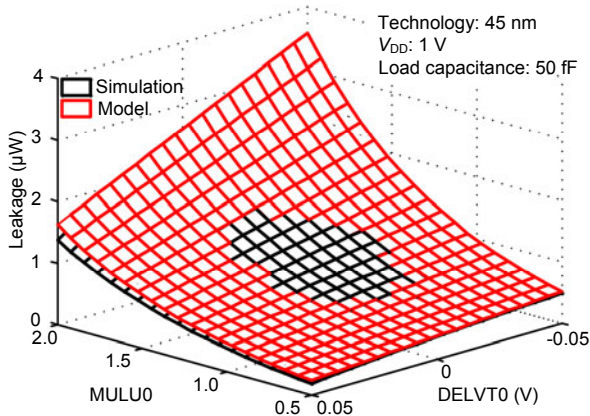


Fig. 10 Leakage consumption of an inverter in 45-nm technology versus DELVT0 and MULU0 at $V_{DD}=1$ V

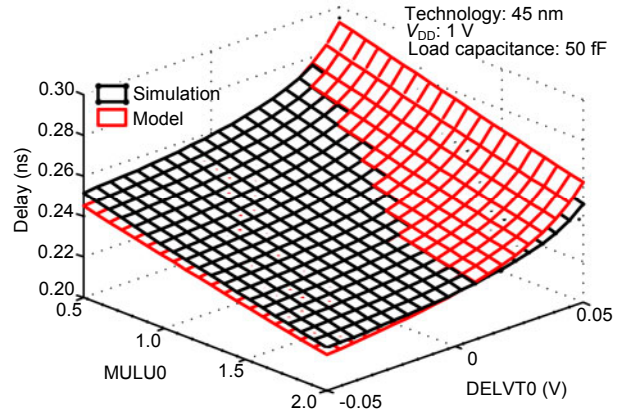


Fig. 11 Delay of an inverter in 45-nm technology versus DELVT0 and MULU0 at $V_{DD}=1$ V

Table 4 ‘Goodness of fit’ and MPE of the leakage model, and CPU time for predicting the leakage power consumption and delay by HSPICE simulations and our proposed model

Circuit	R-square			MPE (%)			CPU time* (s)	
	45 nm	65 nm	90 nm	45 nm	65 nm	90 nm	Simulation	Our model
Inverter	0.991	0.993	0.993	5.11	4.93	5.13	22	0.81
NAND2	0.998	0.998	0.998	3.13	3.32	3.79	38	0.83
NAND3	0.999	0.999	0.998	2.48	2.71	3.58	89	0.82
NOR2	0.990	0.993	0.992	5.35	4.98	5.14	43	0.79
NOR3	0.989	0.992	0.991	5.40	5.02	5.46	97	0.81
C17	0.998	0.998	0.998	3.17	3.24	3.12	420	5.04
C432	0.987	0.991	0.993	4.25	4.29	4.19	4.12e3	23.4
C499	0.974	0.978	0.981	4.89	4.56	4.32	4.88e3	33.5
C880	0.969	0.975	0.972	5.21	5.35	5.17	7.22e3	57.8
C1355	0.958	0.963	0.969	5.37	5.86	5.91	1.23e4	99.4
C1908	0.951	0.959	0.961	4.48	4.12	4.33	2.59e4	156.8
C2670	0.946	0.951	0.959	5.12	5.02	5.14	6.23e4	212.3
C5315	0.927	0.931	0.928	5.57	5.15	5.45	13.38e4	423.6
C6288	0.915	0.919	0.922	6.48	6.53	6.32	16.72e4	508.2
C7552	0.921	0.928	0.934	5.23	5.29	5.31	21.34e4	567.8

MPE: mean percentage error. * The measurements were performed on a 2.66 GHz Intel Core i7 CPU for 10 000 Monte Carlo points. The circuit simulations were performed by HSPICE and modeling was done using MATLAB

Table 5 ‘Goodness of fit’ and MPE of the delay model , and CPU time for predicting the leakage power consumption and delay by HSPICE simulations and our proposed model

Circuit	R-square			MPE (%)			CPU time* (s)	
	45 nm	65 nm	90 nm	45 nm	65 nm	90 nm	Simulation	Our model
Inverter	0.959	0.948	0.938	1.31	1.38	1.58	87	0.71
NAND2	0.945	0.941	0.942	1.91	2.09	2.26	166	0.72
NAND3	0.946	0.946	0.934	2.34	2.50	2.54	314	0.74
NOR2	0.966	0.953	0.926	1.21	1.48	1.88	197	0.69
NOR3	0.956	0.961	0.949	1.57	1.46	1.68	378	0.73
C17	0.947	0.943	0.951	1.93	1.87	1.82	983	4.56
C432	0.949	0.968	0.971	2.21	2.31	2.35	1.45e3	7.32
C499	0.952	0.941	0.945	2.37	2.25	2.19	1.69e3	10.21
C880	0.938	0.952	0.957	2.58	2.62	2.68	4.23e3	22.89
C1355	0.942	0.976	0.982	3.35	3.39	3.21	9.43e3	48.34
C1908	0.928	0.937	0.942	2.45	2.57	2.62	1.26e4	64.53
C2670	0.936	0.944	0.952	2.89	2.79	2.68	1.83e4	89.30
C5315	0.943	0.949	0.953	3.69	3.45	3.38	2.94e4	124.67
C6288	0.919	0.928	0.932	4.23	4.12	3.98	3.77e4	189.70
C7552	0.927	0.922	0.932	3.76	3.68	3.54	3.35e4	165.80

MPE: mean percentage error. * The measurements were performed on a 2.66 GHz Intel Core i7 CPU for 10 000 Monte Carlo points. The circuit simulations were performed by HSPICE and modeling was done using MATLAB

Eqs. (22) and (23), while those for the ISCAS85 benchmark circuits were obtained using Eqs. (24) and (25), respectively. All the simulations were performed at room temperature (300 K) and with $V_{DD}=1$ V.

We have also reported the CPU times for predicting the leakage power consumption and delay by HSPICE simulations and our proposed model (evaluated using MATLAB) in Tables 4 and 5. The model CPU time will be even smaller if it is implemented in a programming language.

6 Conclusions

In this paper we propose a methodology to extract the effects of variations from the nominal values of leakage consumption and delay. The proposed models for leakage and delay were verified against the simulation results obtained from primitive digital gates and ISCAS85 benchmark circuits in 45-, 65-, and 90-nm PTM. The average MPE of leakage and delay models were 4.75% and 2.54%, respectively. Also, the R-square of leakage and delay models were higher than 97% and 95%, respectively. We also illustrated the contours of leakage and delay, which could be useful for identifying the domains of

DELVT0 and MULU0 where the circuit relies on high-performance low-leakage area of operation. Furthermore, we discussed how these contours could be used as margins of design when it comes to process variation effects on circuit performance. In-depth mathematical discussion, however, remains an open research topic. We also discussed figures of merit such as leakage-delay²-product and how process variations could deviate the performance from the desired region.

References

- Agarwal, A., Dartu, F., Blaauw, D., 2004. Statistical Gate Delay Model Considering Multiple Input Switching. Proc. 41st Annual Design Automation Conf., p.658-663. [doi:10.1145/996566.996746]
- Alpert, C.J., Devgan, A., Kashyap, C.V., 2001. RC delay metrics for performance optimization. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.*, **20**(5):571-582. [doi:10.1109/43.920682]
- Andrieu, F., Ernst, T., Lime, F., Rochette, F., Romanjek, K., Barraud, S., Ravit, C., Boeuf, F., Jurczak, M., Casse, M., et al., 2005. Experimental and Comparative Investigation of Low and High Field Transport in Substrate- and Process-Induced Strained Nanoscale MOSFETs. Symp. on VLSI Technology, Digest of Technical Papers, p.176-177. [doi:10.1109/2005.1469257]
- Arizona State University, 2006. Predictive Technology Models. Available from <http://ptm.asu.edu/> [Accessed on Sept. 13,

- 2011].
- Chen, H., Neely, S., Xiong, J., Zolotov, V., Visweswariah, C., 2008. Statistical modeling and analysis of static leakage and dynamic switching power. *LNCS*, **5349**:178-187. [doi:10.1007/978-3-540-95948-9_18]
- Chen, M., Yi, Y., Zhao, W., Ma, D., 2011. Variation-Aware Deep Nanometer Gate Performance Modeling: an Analytical Approach. Int. Symp. on VLSI Design, Automation, and Test, p.1-4. [doi:10.1109/VDAT.2011.5783560]
- Cheng, L., Gupta, P., He, L., 2009. Efficient additive statistical leakage estimation. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.*, **28**(11):1777-1781. [doi:10.1109/TCAD.2009.2030433]
- Christian, S., 2007. Leakage Aware Digital Design Optimization for Minimal Total Power Consumption in Nanometer CMOS Technologies. PhD Thesis, University of Neuchâtel, Switzerland.
- D'Agostino, C., le Coz, J., Flatresse, P., Beigne, E., Belleville, M., 2009. An Accurate Approach for Statistical Estimation of Leakage Current Considering Multi-parameter Process Variations in Nanometer CMOS Technologies. Proc. European Solid State Device Research Conf., p.427-430. [doi:10.1109/ESSDERC.2009.5331488]
- da Silva, D.N., Reis, A.I., Ribas, R.P., 2009. CMOS logic gate performance variability related to transistor network arrangements. *Microelectron. Reliab.*, **49**(9-11):977-981. [doi:10.1016/j.microrel.2009.07.023]
- Guerra, L., Phillips, J., Silveira, L.M., 2010. Effective corner-based techniques for variation-aware IC delay verification. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.*, **29**(1):157-162. [doi:10.1109/TCAD.2009.2034343]
- Gupta, P., Jeong, K., Kahng, A.B., Park, C., 2008. Electrical metrics for lithographic line-end tapering. *SPIE*, **7028**(2): 70283A.1-70283A.12. [doi:10.1117/12.793117]
- Hao, Z., Tan, S., Shi, G., 2011. An Efficient Statistical Chip-Level Total Power Estimation Method Considering Process Variations with Spatial Correlation. Int. Symp. on Quality Electronic Design, p.1-6. [doi:10.1109/ISQED.2011.5770801]
- Joshi, V., Agarwal, K., Sylvester, D., 2010. Simultaneous Extraction of Effective Gate Length and Low-Field Mobility in Non-uniform Devices. Int. Symp. on Quality Electronic Design, p.158-162. [doi:10.1109/ISQED.2010.5450409]
- Liu, Q., Sapatnekar, S., 2009. A framework for scalable post-silicon statistical delay prediction under process variations. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.*, **28**(8):1201-1212. [doi:10.1109/TCAD.2009.2021732]
- Liu, Q., Sapatnekar, S., 2010. Capturing post-silicon variations using a representative critical path. *IEEE Trans. Comput.-Aided Des. Integr. Circ. Syst.*, **29**(2):211-222. [doi:10.1109/TCAD.2009.2035552]
- Ma, T., 2009. Improving Design Quality by Managing Process Variability. Int. Symp. on Quality Electronic Design (Presentation), p.18-23.
- Miryala, S., Kaur, B., Anand, B., Manhas, S., 2011. Efficient Nanoscale VLSI Standard Cell Library Characterization Using a Novel Delay Model. Int. Symp. on Quality Electronic Design, p.1-6. [doi:10.1109/ISQED.2011.5770767]
- Mistry, K., Armstrong, M., Auth, C., Cea, S., Coan, T., Ghani, T., Hoffmann, T., Murthy, A., Sandford, J., Shaheed, R., et al., 2004. Delaying Forever: Uniaxial Strained Silicon Transistors in a 90nm CMOS Technology. Symp. on VLSI Technology, Digest of Technical Papers, p.50-51. [doi:10.1109/VLSIT.2004.1345387]
- Morshed, T.H., 2009. BSIM4.6.4 MOSFET Model. User's Manual. Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA.
- Narendra, S., Blaauw, D., Devgan, A., Najm, F., 2003. Leakage Issues in IC Design: Trends, Estimation and Avoidance. Int. Conf. on Computer Aided Design (tutorial), p.11. [doi:10.1109/ICCAD.2003.145]
- Okada, K., Yamaoka, K., Onodera, H., 2003. A Statistical Gate-Delay Model Considering Intra-Gate Variability. Int. Conf. on Computer Aided Design, p.908-913. [doi:10.1109/ICCAD.2003.20]
- Orshansky, M., Nassif, S.R., Boning, D., 2008. Design for Manufacturability and Statistical Design: a Constructive Approach. Springer, New York, NY.
- Pramanik, D., Moroz, V., Lin, X.W., 2006. Process Induced Layout Variability for sub 90nm Technologies. Int. Conf. on Circuit Technology, p.1849-1852. [doi:10.1109/ICSICT.2006.306464]
- Ramalingam, A., Kodakara, S.V., Devgan, A., Pan, D., 2006. Robust Analytical Gate Delay Modeling for Low Voltage Circuits. Asia and South Pacific Design Automation Conf., p.61-66. [doi:10.1145/1118299.1118315]
- Rao, R., Srivastava, A., Blaauw, D., Sylvester, D., 2004. Statistical analysis of subthreshold leakage current for VLSI circuits. *IEEE Trans. VLSI Syst.*, **12**(2):131-139. [doi:10.1109/TVLSI.2003.821549]
- Roy, S., Asenov, A., 2005. Where do the dopants go? *Science*, **309**(5733):388-390. [doi:10.1126/science.1111104]
- Sakurai, T., Newton, A.R., 1990. Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. *IEEE J. Sol.-State Circ.*, **25**(2):584-594. [doi:10.1109/4.52187]
- Scott, G., Lutze, J., Rubin, M., Nouri, F., Manley, M., 1999. NMOS Drive Current Reduction Caused by Transistor Layout and Trench Isolation Induced Stress. Int. Electron Device Meeting, p.827-830. [doi:10.1109/IEDM.1999.824277]
- Singhal, R., Balijepalli, A., Subramaniam, A., Liu, F., Nassif, S., Cao, Y., 2007. Modeling and Analysis of Non-rectangular Gate for Post-Lithography Circuit Simulation. Proc. 44th Annual Design Automation Conf., p.823-828. [doi:10.1145/1278480.1278685]
- Srivastava, A., Sylvester, D., Blaauw, D., 2005. Statistical Analysis and Optimization for VLSI: Delay and Power. Springer, New York, NY.
- Synopsys Corporation, 2008. Synopsys HSPICE. Available from <http://www.synopsys.com/Tools/Verification/AMS>

Verification/CircuitSimulation/HSPICE/Pages/default.aspx
[Accessed on Sept. 13, 2011].

Sze, S.M., Ng, K.K., 2007. Physics of Semiconductor Devices. John Wiley & Sons, Hoboken, NJ.

Tsai, K.Y., You, M.F., Lu, Y.C., Ng, P.C.W., 2008. A New Method to Improve Accuracy of Leakage Current Estimation for Transistors with Non-rectangular Gates due to Sub-wavelength Lithography Effects. Int. Conf. on Computer-Aided Design, p.286-291. [doi:10.1109/ICCAD.

2008.4681587]

Xu, N., Wang, L., Neureuther, A., Liu, T., 2011. Physically based modeling of stress-induced variation in nanoscale transistor performance. *IEEE Trans. Dev. Mater. Reliab.*, **11**(3):378-386. [doi:10.1109/TDMR.2011.2144598]

Ye, Z., Yu, Z., 2009. An Efficient Algorithm for Modeling Spatially-Correlated Process Variation in Statistical Full-Chip Leakage Analysis. Proc. Int. Conf. on Computer-Aided Design, p.295-301. [doi:10.1145/1687399.1687455]

Accepted manuscript available online (unedited version)

<http://www.zju.edu.cn/jzus/inpress.htm>



JZUS-A
(Applied Physics & Engineering)



JZUS-B
(Biomedicine & Biotechnology)



JZUS-C
(Computers & Electronics)

- As a service to our readers and authors, we are providing the unedited version of accepted manuscripts.
- The section "Articles in Press" contains peer-reviewed, accepted articles to be published in *JZUS (A/B/C)*. When the article is published in *JZUS (A/B/C)*, it will be removed from this section and appear in the published journal issue.
- Please note that although "Articles in Press" do not have all bibliographic details available yet, they can already be cited as follows: Author(s), Article Title, Journal (Year), **DOI**. For example:
ZHANG, S.Y., WANG, Q.F., WAN, R., XIE, S.G. Changes in bacterial community of anthrance bioremediation in municipal solid waste composting soil. *J. Zhejiang Univ.-Sci. B (Biomed. & Biotechnol.)*, in press (2011). [doi:10.1631/jzus.B1000440]
- Readers can also give comments (Debate/Discuss/Question/Opinion) on their interested articles in press.