Gate Leakage and Delay Analysis in Logic Gates of Non-Classical High$-K$ Nano-CMOS Transistors

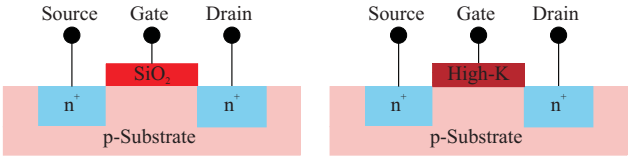
Saraju P. Mohanty
Computer Science and Engineering
University of North Texas, Denton, TX 76203.
Email: smohanty@cse.unt.edu

Elias Kougianos
Electrical Engineering Technology
University of North Texas, Denton, TX 76203.
Email: eliask@unt.edu

***Abstract*— Gate leakage (direct tunneling current) has emerged as the major component of power dissipation due to the use of ultra-thin SiO_2 layer. Replacement of SiO_2 as gate dielectric with alternative high-K dielectric materials, such as SiON, Si_3N_4, Al_2O_3, etc. is considered as a method to contain the gate leakage; thus constructing a type of non-classical nano-CMOS transistor. This paper provides novel attempts to evaluate the gate leakage and delay of basic logic gates comprising of such transistors. The results presented will guide a design engineer in selecting a non-classical transistor for leakage-performance tradeoffs.**

I. Introduction and Contributions

Transistor feature sizes have dramatically shrunk with technology scaling and the value of oxide thickness has reached the range of $12-16\mathring{A}$ which is just a few monolayers of SiO_2. This has led to a drastic change in the leakage components of the device both in its active and inactive states. Consequently, gate oxide leakage has emerged as the most prominent form of leakage in nano-CMOS devices, particularly in the $65nm$ and below regime. This has led to the construction of non-classical transistors demonstrated in. Fig. 1.



(a) with SiO_2 gate dielectric (b) with High$-K$ gate dielectric

Fig. 1. Nanoscale Planer CMOS Tansistors : Classical Versus Nonclassical

The use of high$-K$ serves the dual purpose of scaling of the device as well as reduction of gate leakage [1], [2]. A proper characterization high$-K$ nano-CMOS devices and circuits is necessary to give the designer a complete idea of the efficacy of the material as a gate dielectric and to evaluate its potential in replacing SiO_2. It further necessitates the development of an analytical method for on-the-fly calculation of electrical characteristics of logic cells to facilitate automatic synthesis.

Few research works exist in the literature for analysis and estimation of gate leakage in nano-CMOS circuits. However, none of them address non-classical high$-K$ nano-CMOS logic gates. In [3], the authors presented mechanisms for estimation of leakage current. The authors in [4] have formulated "state-dependent" gate leakage data. The authors in [5] have developed several methods of modeling, estimation and analysis of total leakage and its components especially, gate leakage.

Contributions of this paper: We introduce a new approach for the characterization of non-classical high-K nano-CMOS transistors and corresponding logic gates. We give a complete view of the design and process parameter variations of the devices when high-K dielectric materials are used. We present novel analytical functions for on-the-fly calculations of the gate leakage and delay, which would provide a useful mechanism for the design engineer for design space and technology exploration during the automatic synthesis of circuits.

II. High-K Dielectrics in Nano-CMOS Technology

Recently several materials have been investigated for use in nano-CMOS technology, such as ZrO_2, TiO_2, BST, HfO_2, Al_2O_3, SiON, and Si_3N_4 [1], [6]. There is some progress in the development of various technologies for high$-K$ gate dielectric deposition [7]. This includes the extension of chemical vapor deposition (CVD), as well as single wafer methodologies, namely rapid thermal chemical vapor deposition (RTCVD), rapid plasma-enhanced chemical vapor deposition (RPECVD), and liquid source misted chemical vapor deposition (LSCVD). Other techniques include physical vapor deposition (PVD) [8], jet vapor deposition (JVD) [9], oxidation of metallic films [10], and molecular beam epitaxy [11].

While the materials research is in full swing, there is no research addressing automatic design and synthesis of systems using high-K nano-CMOS transistors. For compact modeling based study of high-K non-classical transistors using BSIM4/5, *two possible options* can be considered: (i) varying the model parameter in the model card that denotes relative permittivity (EPSROX) and/or (ii) finding the equivalent oxide thickness (EOT) for a dielectric under consideration. Approach (i) may not be sufficient to model the behavior of non-classic nano-CMOS with non-SiO_2 dielectrics as it is does not correctly account for the barrier height of non-SiO_2 dielectrics. Using method (ii) EOT will be calculated so as to keep the ratio of relative permittivity over dielectric thickness constant.

Both of these approaches ignore several aspects of the physics behind non-SiO_2 dielectrics, particularly in the Si/dielectric interface. However, in the absence of published device data, the methodology presented will provide meaningful information of the various materials under consideration to match EDA development with material science trends.

III. Transient Study of Logic Gates

In our analysis we use the Berkeley Predictive Technology Model (BPTM) [12] since it is well established and is able to predict the general trend of device attributes. In the absence of published data and device models, the BPTM provides the means for timely and effective analysis. Since the BPTM is physics based the simulations results are accurate and the calculated data are of comparable accuracy to TCAD simulations which are typically time and computation intensive.

We performed a complete transistor level characterization of several logic gates (NOT, NAND, NOR, AND and OR) with respect to gate leakage and delay using analog simulator, but present only the results for NAND for brevity. A capacitive load of value of 10 times the total gate capacitance C_{gg} of the PMOS device has been used [13]. For fair comparison of the performance of the various dielectrics, this value was calculated for a SiO_2 device and kept fixed throughout.

A. Transient Response

The gate leakage of a logic gate can be calculated from the direct tunneling current of each PMOS and NMOS device in it. There are several components of direct tunneling current within each device, such as I_{gs} and I_{gd} (components due to the overlap of gate and diffusions), I_{gcs} and I_{gcd} (components due to tunneling from the gate to the diffusions via the channel) and I_{gb}, the component due to tunneling from the gate to the bulk via the channel. These components are quantitatively and flow direction wise different for different states (ON or OFF) and types (NMOS or PMOS) of a device. We calculate the total gate leakage current for each device as:

$$I_{gate}[i] \quad = \quad I_{gs}[i] + I_{gd}[i] + I_{gcs}[i] + I_{gcd}[i] + I_{gb}[i], \quad (1)$$

where the index i identifies the device within a logic gate. A total gate leakage for the logic gate (I_{gate}) can then calculated by summing the gate leakage of all the devices:

$$I_{gate} \quad = \quad \textstyle\sum_i I_{gate}[i]. \quad (2)$$

During its various states of operation, a logic gate presents different dominant leakage paths, depending on the combination of inputs. For a 2-input logic gate, for each of the four possible states (00, 01, 10 and 11), the overall gate leakage current (I_{00}, I_{01}, I_{10}, and I_{11}, respectively) is calculated from Eqn. 1 and 2. Assuming that all states are to occur with equal probability, an "average gate leakage" ($\overline{I_{gate}}$) is calculated as:

$$\overline{I_{gate}} \quad = \quad \left(\tfrac{I_{00}+I_{01}+I_{10}+I_{11}}{4}\right). \quad (3)$$

Following standard approaches we define the delay as the time difference between the 50% level of the input and output waveforms. For worst-case scenario, we chose the maximum delay time regardless of whether this was due to a low-to-high or a high-to-low transition.

B. Process and Design Parameters Variation

The effect of varying oxide thickness (T_{ox}) is incorporated by varying $TOXE$ in the SPICE model deck directly. The effect of varying dielectric material K is modeled by calculating an equivalent oxide thickness (T^*_{ox}) according to the formula:

$$T^*_{ox} \quad = \quad \left(\frac{^{K}\mathrm{SiO_2}}{K_{gate}}\right) \times T_{gate}, \quad (4)$$

where, K_{gate} is the relative permittivity and T_{gate} is the thickness of the gate dielectric material other than SiO_2, while $K_{\mathrm{SiO_2}}$ is the dielectric constant of SiO_2 (=3.9). The length of the device is proportionately changed to minimize the impact of higher dielectric thickness on device performance and to maintain the per width gate capacitance constant as per CMOS fabrication requirements [14]. Thus, the scaling ratio of channel length to the gate thickness is maintained constant, i.e. $\left(\frac{L}{T}\right)$ = Constant. Moreover, the length and width of the transistors are chosen to maintain a $\left(\frac{W}{L}\right)$ ratio of $4:1$ for NMOS and $8:1$ for PMOS. The variation of V_{DD} is achieved by running a parameter sweep.

1) Effect of Variation on Gate Leakage: It is observed that the gate leakage current show a uniform trend when the gate dielectric (K) is varied assuming a fixed T_{gate} and V_{DD} at their nominal values. It can be seen that with an increase in the value of K there is a steady decrease in the value of tunneling leakage, I_{gate}. This trend continues up to a knee region of $K = 12$, after which it becomes almost constant and there is not much of a decrease in the value of I_{gate} with an increase in K. This indicates the presence of a saturation zone in the gate leakage versus K characteristics which implies that there is *not much benefit in deploying gate dielectrics of very high-K* (i. e. $K > 12$) *for the gate tunneling leakage reduction*. There is a uniform decrease in the gate leakage with an increase in the gate thickness T_{gate}. It is seen that an increase in the supply voltage leads to a corresponding increase in the gate leakage.

2) Effect of Variation on Propagation Delay: We observe that there is a sharp increase in the value of T_{pd} with an increase in the dielectric constant; this increase continues until a value of around $K = 6$ after which the slope is much lower. The increase in propagation delay can be attributed to the increase in capacitance per unit area ($C^{'}_{gate}$) of gate oxide with dielectric constant i.e.

$$C^{'}_{gate} \quad = \quad \left(\tfrac{K_{gate}}{T_{gate}}\right). \quad (5)$$

The effect of variation of propagation delay with respect to the process parameter T_{gate} indicates that there is an increase in the propagation delay with increase in dielectric thickness. This happens due to the increase in gate capacitance (C_{gate}) with thickness T_{gate} for a particular dielectric, say SiO_2 with K or $\epsilon_{ox} = 3.9$ as evident from the following discussion. The gate capacitance (C_{gate}) is given by [15]:

$$C_{gate} \quad = \epsilon_{ox}\left(\tfrac{L}{T_{ox}}\right) W \quad = \epsilon_{ox} * k_{\mathrm{size}} * L\left(\tfrac{L}{T_{ox}}\right), \quad (6)$$

for constant $\left(\frac{W}{L}\right) = k_{\mathrm{size}}$. Thus, as per the suggestion in [14], with increase in T_{ox} when we maintain $\left(\frac{L}{T_{ox}}\right)$ constant by increasing L, C_{gate} increases, so also the propagation delay.

The propagation delay shows a decreasing trend with an increase in the value of V_{DD} when K_{gate} and T_{gate} are kept fixed. This is due to the increase in the drive current resulting from the increase in V_{DD}. However, a better insight of the situation can be obtained from the following. For a technology parameter α the propagation delay is given by [15],

$$\begin{aligned} T_{pd} &= \tfrac{\alpha C_{load} V_{DD}}{V_{DD} - V^*_{Th}}, \text{ (where } V^*_{Th} = V_{Th} + 0.5V_{DSAT}), \\ &= \alpha C_{load}\left(\tfrac{1}{1 - V^*_{Th}/V_{DD}}\right), \text{ (dividing by } V_{DD}). \end{aligned} \quad (7)$$

Since, $-1 < \frac{V^*_{Th}}{V_{DD}} < 1$, using McLaurin series we get:

$$\begin{aligned} T_{pd} \quad &= \quad \alpha C_{load}\left(1 + \tfrac{V^*_{Th}}{V_{DD}} + \left(\tfrac{V^*_{Th}}{V_{DD}}\right)^2 + ...\right), \\ &\approx \quad \alpha C_{load}\left(1 + \tfrac{V^*_{Th}}{V_{DD}}\right). \end{aligned} \quad (8)$$

This clearly suggest that for fixed load and threshold voltage, as V_{DD} increases, T_{pd} decreases. However, as scaling continues, the trend is to scale down the supply along with other device features and that is also compatible with the objective of decreasing the gate leakage current in the nanometer regime.

IV. Logic Gates with Specific High-K Nano-CMOS

A number of dielectrics are being investigated as potential alternatives to SiO_2. Among them the more prominent ones are SiON, Si_3N_4, Al_2O_3, $ZrSiO_4$, $HfSiO_4$, and HfO_2 [1], [6].

A. Behavior of High-K Dielectrics on Variation of T_{gate}

The effect of a change in the gate dielectric thickness (process parameter) with different dielectrics are presented in Fig. 2(a) for gate leakage. Among the dielectrics considered, HfO_2 has the highest dielectric constant while SiO_2 has the lowest. It can be clearly seen that in all cases, the gate tunneling current for SiO_2, having the least K, is the highest across all thicknesses. As the value of K increases the gate tunneling current reaches a constant level for all values of thickness. In other words, the tunneling current can be obviated to a great extent with the use of high-K dielectrics.

The behavior of different dielectrics for a change in the process parameter T_{gate} for propagation delay is demonstrated in Fig. 2(b). In this case the dielectric with the least value of K, i.e. SiO_2 has the least propagation delay across all thicknesses. However, as the thickness increases the propagation delay increases too. In the case of dielectrics with a medium value of K the propagation delay stays constant over the range of gate thickness considered.

B. Behavior of High-K dielectrics on variation of V_{DD}

The effect of variation of design parameters on a number of gate dielectrics is also studied. As can be seen from Fig. 2(c), the gate tunneling current increases with an increase in the value of the supply voltage for all dielectrics. In the case of the default gate dielectric i.e. SiO_2, the gate tunneling current is maximum among all the other dielectrics across the range of supply voltages from $0.5V$ to $1.0V$. On the other hand, for HfO_2, having the highest value of K, the gate tunneling current is the smallest throughout the range of supply voltages. For any given supply voltage, it can be seen that the gate tunneling current is lower for dielectrics with higher K.

The plot in Fig. 2(d) shows the behavior of propagation delay for various dielectrics when the design parameter supply voltage is changed. When the supply voltage is kept fixed at $0.5V$, the propagation delay is larger for gate dielectrics with higher values of K. As the supply voltage increases it can be seen from the same figure that there is a corresponding decrease in the propagation delay.

V. Characterization for Circuit Synthesis

We perform curve-fitting of the data for each parameter viz. dielectric constant (K), gate dielectric thickness (T_{gate}) and supply voltage (V_{DD}) for their effect on gate tunneling current (I_{gate}) and propagation delay (T_{pd}). The results of this modeling can be used in the back-end tools of design and automatic synthesis frameworks for on-the-fly calculation of the gate tunneling current and propagation delay under the influence of various design and process parameters.

We present curve-fitting functions for a 2-input NAND logic gate in Table I. Derivations for other gates are done in a similar manner, but we do not present the results for brevity. For the 2-input NAND gate, the gate tunneling current decreases exponentially with an increase in the dielectric constant. The propagation delay varies in a more complex fashion as can be seen from the table: it is represented as a combination of two functions over the range of dielectric constant considered.

Across the range of variation of T_{gate} the tunneling shows a trend of exponential decrease with increase in the value of dielectric thickness. In the case of propagation delay, the fitting corresponds to a Langmuir function and shows an excellent fit over the range in consideration with correlation factor of 0.99913. This shows that the modeling functions are highly reliable for use in design. The advantages of our approach over experimentally derived relations (however, no published data are available as of now) is the predictiveness for process and design space exploration for a design engineer.

In the case of the variation of V_{DD}, both I_{gate}, while the gate leakage shows a uniform exponential increase with an increase in the value of the supply, the propagation delay shows an exponential decrease and the nature remains the same for all logic gates. All the functions yield a perfect fit and have a correlation factor of the order of 0.999 in all cases.

VI. Summary and Conclusions

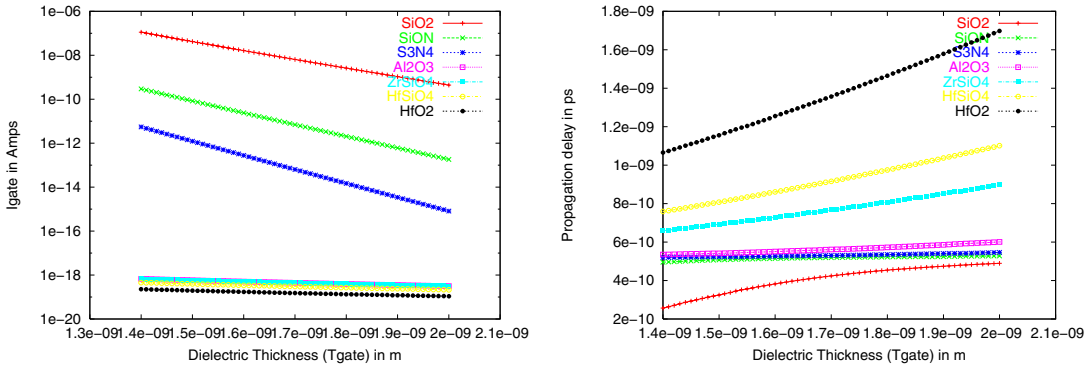
We presented a comprehensive analysis of the transient behavior of a nano-CMOS NAND gate, for a realistic $45nm$ BSIM4 model. We used this information for the characterization of two crucial process (K_{gate}, T_{gate}) and one design (V_{DD}) parameter of the tunneling effect and propagation delay of the logic gates made of high-K nano-CMOS transistors. It was observed that the *NAND had minimal gate leakage* among all the logic gates under consideration for same experimental conditions. We performed a further analysis of the interdependence of these parameters and used the data for curve-fitting. We propose that the data from the characterization and modeling equations provides an excellent means for the design of novel and complex low-leakage nano-CMOS devices involving high-K dielectrics and provides valuable information to estimate the effect of gate leakage and propagation delay which can then be used to characterize entire cells and libraries leading ultimately to optimized synthesis algorithms for nanoscale CMOS circuit design.

References

[1] L. Manchanda et al., "High K gate Dielectrics for the Silicon Industry," in *Proc. of International Workshop on Gate Insulator*, 2001, pp. 56–60.

[2] E. M. Vogel, K. Z. Ahmed, B. Hornung, P. K. McLarty, G. Lucovsky, J. R. Hauser, and J. J. Wortman, "Modeled Tunnel Currents for High Dielectric Constant Dielectrics," *IEEE Transactions on Electron Devices*, vol. 45, no. 6, pp. 1350–1355, June 1998.

[3] R. M. Rao, J. L. Burns, A. Devgan, and R. B. Brown, "Efficient Techniques for Gate Leakage Estimation," in *Proceedings of the International Symposium on Low Power Devices*, 2003, pp. 100–103.



(a) $\overline{I_{gate}}$ Vs T_{gate} (b) T_{pd} Vs T_{gate}

(c) $\overline{I_{gate}}$ Vs V_{DD} (d) T_{pd} Vs V_{DD}

Fig. 2. Evaluation of selected gate dielectrics (SiO_2, SiON, Si_3N_4, Al_2O_3, $ZrSiO_4$, $HfSiO_4$, and HfO_2 are listed from top to bottom in legend) for effect of variation of process parameter (T_{gate}) and design parameter (V_{DD}) on gate leakage and propagation delay using BSIM4 $45nm$ CMOS.

TABLE I

CURVE FITTING FOR EFFECT OF PROCESS AND DESIGN VARIATION IN 2-INPUT NAND GATE

Parameters	Attributes	Fitting Functions	Function Parameters	Correlations
	$\overline{I_{gate}}$	$\overline{I_{gate}} = A * exp\left(\frac{-K}{\alpha}\right) + \overline{I_{gate}}_0$	$\overline{I_{gate}}_0 = -4.6 * 10^{-9}$, $A = 0.00966$, $\alpha = 0.36115$	0.99655
K	T_{pd}	$T_{pd} = A_2 + \frac{A_1 - A_2}{1+exp\left(\frac{K-K_0}{\alpha}\right)}$, $2.5 \leq K < 6$ $T_{pd} = A_3 * exp\left(\frac{-K}{\beta}\right) + T_{pd_0}$, $6 \leq K < 30$	$A_1 = 4.12 * 10^{-11}$, $A_2 = 5.14 * 10^{-10}$ $K_0 = 4.02176$, $\alpha = 0.47847$ $A_3 = 6.94 * 10^{-11}$, $\beta = -10.63$, $T_{pd_0} = 3.75 * 10^{-10}$	0.99725
T_{gate}	$\overline{I_{gate}}$	$\overline{I_{gate}} = A * exp\left(\frac{-K}{\alpha}\right) + \overline{I_{gate}}_0$	$\overline{I_{gate}}_0 = 3.13 * 10^{-10}$ $A = 0.09475$, $\alpha = 1.03 * 10^{-10}$	0.99564
	T_{pd}	$T_{pd} = \frac{A * B * T_{gate}{}^{1-\alpha}}{1+B * T_{gate}{}^{1-\alpha}}$	$A = 5.10110^{10}$, $B = 1.50110^{75}$, $\alpha = -7.49133$	0.99913
V_{DD}	$\overline{I_{gate}}$	$\overline{I_{gate}} = A * exp\left(\frac{V_{DD}}{\alpha}\right) + \overline{I_{gate}}_0$	$\overline{I_{gate}}_0 = -2.47 * 10^{-08}$, $A = 4.76 * 10^{-9}$, $\alpha = 0.20795$	0.99996
	T_{pd}	$T_{pd} = A * exp\left(\frac{-V_{DD}}{\beta}\right) + T_{pd_0}$	$T_{pd_0} = 1.92 * 10^{-10}$, $A = 4.21 * 10^{-9}$, $\beta = 0.16781$	0.99903

[4] R. S. Guindi and F. N. Najm, "Design Techniques for Gate-Leakage Reduction in CMOS Circuits," in *Proceedings of the International Symposium On Quality Electronic Design*, 2003, pp. 61–65.

[5] A. Agarwala, S. Mukhopadhyay, C. H. Kim, and K. Roy, "Leakage Power Analysis and Reduction: Models, Estimation and Tools," *IEE Proceedings in Computers and Digital Techniques*, vol. 152, no. 3, pp. 353 – 368, 2005.

[6] M. Yang et al., "Performance Dependence of CMOS on Silicon Substrate Orientation for Ultrathin and HfO_2 Gate Dielectrics," *IEEE Electron Device Letters*, vol. 24, no. 5, pp. 339–341, May 2003.

[7] H. R. Huff and et. al., "Integration of High-K Gate Stack Systems into Planar CMOS Process Flows," in *International Workshop on Gate Insulator*, 2001, pp. 2–11.

[8] W. J. Qi et al., "Ultrathin Zirconium Silicate Film With Good Thermal Stability for Alternative Gate Dielectric Application," *Applied Physics Letters*, vol. 77, pp. 1704–1706, 2000.

[9] T. P. Ma, "Making Silicon Nitride a Viable Gate Dielecrtric," *IEEE Transaction on Electron Devices*, vol. 45, pp. 680–690, 1999.

[10] B. H. Lee et al., "Thermal Stability and Electrical Characteristics of Ultrathin Hafnium Oxide Gate Dielectric Reoxidized with Rapid Thermal Annealing," *Applied Physics Letters*, vol. 77, pp. 1926–1928, 2000.

[11] A. I. Kingon, J. P. Maria, and S. K. Streifferr, "Alternative Dielectrics to Silicon Dioxide for memory and Logic Devices," *Nature*, vol. 406, pp. 1021–1038, 2000.

[12] "Predictive Technology Model Website," `http://www.eas.asu.edu/~ptm`.

[13] V. Mukherjee, S. P. Mohanty, and E. Kougianos, "A Dual Dielectric Approach for Performance Aware Gate Tunneling Reduction in Combinational Circuits," in *Proceedings of the 23rd IEEE International Conference of Computer Design (ICCD)*, 2005, pp. 431–436.

[14] Y. Taur, "CMOS Design Near the Limits of Scaling," *IBM Journal on Research and Development*, vol. 46, no. 2/3, pp. 131–139, Mar./May 2002.

[15] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits*, Prentice Hall, 2003.