Subthreshold Leakage Modeling and Estimation of General CMOS Complex Gates

Author1, Author2, Author3

Local 1
e-mail 1
Local 2
e-mail 2
Local 3
e-mail 3

Abstract. A new subthreshold leakage model is proposed in order to improve the static power estimation in general CMOS complex gates. Series-parallel transistor arrangements with more than two logic depth, as well as non-series-parallel off-switch networks are covered by such analytical modeling. The occurrence of on-switches in off-networks, also ignored by previous works, is considered in the proposed analysis. The model has been validated through electrical simulations, taking into account transistor sizing, operating temperature, supply voltage and threshold voltage variations.

1 Introduction

Power consumption represents an emerging issue nowadays due to recent mobile products and the challenges from new advanced technologies. The leakage currents responsible for static power dissipation during idle mode are increasing significantly in sub-100nm CMOS processes, where the device threshold voltage and the gate oxide thickness tend to reduce. Great effort has been concentrated to understand the leakage current mechanisms, to model their behavior and to develop design techniques for power saving in standby operation [1]-[9].

The main contributors to the total leakage dissipation in CMOS circuitry are the subthreshold and the gate oxide currents. The subthreshold current, negligible in old processes when compared to dynamic charging and short-circuit currents, starts to be taken into account in bellow 180nm technologies. Gate leakage, on the other hand, tends to become the main factor in the static consumption from sub-100nm CMOS processes [3]. Design techniques for leakage reduction have been proposed considering for instance the input state dependency, multi-V_{th} devices and bulk biasing when subthreshold current is addressed, while future high-K gate dielectric and metal gate electrodes are expected to solve static power consumption due to gate oxide leakage current [1]. The most considered design strategies are based on the already well known stack effect, which suggests the use of more than one off-device in series to reduce the subthreshold current. Note that such effect presents distinct influence on different leakage mechanisms, e.g. gate oxide currents [3].

The modeling of the stack effect has been treated in the literature, where only basic NOR and NAND gates, i.e. purely series and parallel arrangements of transistors have been actually targeted. It should be quite enough for standard cell libraries applied to the technology mapping performed by commercial EDA tools.

However, in most recent EDA technologies, library-free mapping technique has been explored [10]. In this sense, CMOS complex gates, composed by a great variety of mixed series-parallel pull-up/pull-down logic networks, i.e. arrangements with more than two levels of logic depth, are frequently applied.

On the other hand, transistor networks derived from BDD graphs provide interesting improvements in device count and cell performance [11-13]. However, unlike standard CMOS gates, non-series-parallel transistor arrangements are usually obtained. Such a kind of topology has not been treated in literature for leakage current prediction.

Finally, the occurrence of on-transistors in off-networks for subthreshold current analysis is another factor largely ignored by previous works, which considers it as ideal short-circuits or on-switches, as reviewed in Section 2. In this paper, the subthreshold leakage modeling has been improved to be applied in general transistor networks, as described in Section 3. Section 4 presents experimental results demonstrating the model accuracy and validation. The conclusions are given in Section 5.

2 Related Works

Subthreshold leakage models have been recently reported in the literature. In [4], Gu and Elmasry presented the subthreshold current model for purely series and parallel off-transistors arrangements (Inverter, NAND and NOR gates), assuming equivalent standby current in both NMOS and PMOS devices, and considering up to three series devices. The two XOR topologies, presented by Gu [4], result, in fact, in single off-transistor configurations during the steady state analysis, as verified in pass-transistor logic (PTL) design style. Transistor stacks composed by single transistors in chain are also addressed in [5]-[8], while on-devices are considered as ideal short-circuits.

AOI and OAI gates are mentioned as examples for analysis of series-parallel off-networks by Lee *et al.* in [9]. Multiple parallel transistor stacks are computed, taking into account each branch separately, and the values are then summed to obtain the total leakage in the logic gate. In case of parallel transistors within a stack, such transistors are initially collapsed and replaced by a single equivalent device with transistor size equal to the sum of their original sizes. Such strategy is also applied by other authors, such as Rosseló *et al.* in [8], and it seems to be quite accurate. However, series-parallel networks with more than two levels of logic depth, as illustrated in Fig.1, cannot be treated by such a kind of network reduction technique.

Yang *et al.* [3], in turn, presents a detailed analysis to estimate the total leakage, including and interacting subthreshold and gate leakage currents, for NAND and NOR gates. Different AOI and OAI gates are also discussed emphasizing the strong influence of the electrical topology in the leakage value. Although not explicitly mentioned, probably the same strategy of transistor width equivalence is applied to treat parallel devices, as suggested in [8] and [9].

The major contributions of this paper are the subthreshold current modeling extension for general off-networks (series-parallel and non-series-parallel arrangements), and the analysis and modeling of subthreshold leakage considering the presence of on-devices in off-networks. In terms of gate oxide leakage, the proposed method can be combined with the estimation method proposed Yang *et al.* [3], improving the subthreshold part in the total leakage calculation.

3 Subthreshold Leakage Model

The total leakage dissipation results from the sum of the current in each branch of off-transistors between the supply voltage (V_{dd}) and ground. To present the proposed method, the off-network illustrated in Fig. 1 can be considered as the entire NMOS pull-down arrangement. The same analysis is applicable to PMOS pull-up trees.

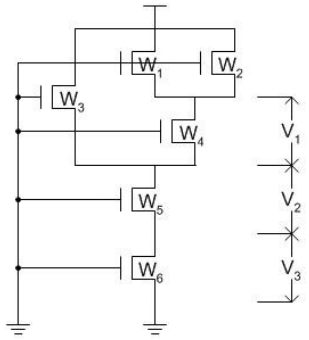


Fig. 1. NMOS series-parallel network.

According to BSIM transistor model [13], the subthreshold current for a MOSFET can be modeled as

$$I_S = I_0 We^{\frac{V_{gs}-(V_{t0}-\eta V_{ds}-\gamma V_{bs})}{nV_T}}\left[1-e^{-\frac{V_{ds}}{V_T}}\right] \quad (1)$$

where $I_0 = \frac{\mu_0 C_{ox} V_T^2 e^{1.8}}{L}$ and $V_T = \frac{kT}{q}$. V_{gs}, V_{ds} and V_{bs} are the gate, drain and bulk voltage of the transistor, respectively. V_{t0} is the zero bias threshold voltage. W and L are the effective transistor width and length, respectively. γ is the body effect coefficient and η is the DIBL coefficient. C_{ox} is the gate oxide capacitance, μ_0 is the mobility and n is the subthreshold swing coefficient.

The subthreshold current through the top devices (I_{Si}), i.e. transistors connected to V_{dd}, can be expressed by equation (2). This equation considers the variable V_j as the voltage over every transistor placed under the transistor in evaluation in the stack.

$$I_{Si} = I_0 W_i e^{\frac{-\sum V_j - \left[V_{t0} - \eta\left(V_{dd} - \sum V_j\right) + \gamma \sum V_j\right]}{nV_T}} \tag{2}$$

The subthreshold current through other transistors in the network is expressed by equation (3). The differences between both equations are observed in the η term (DIBL effect), where V_i represents the voltage over the specific transistor in analysis, and also in the last term, which can be eliminated when $V_i >> V_T$.

$$I_{Si} = I_0 W_i e^{\frac{-\sum V_j - \left[V_{t0} - \eta V_i + \gamma \sum V_j\right]}{nV_T}} \left[1 - e^{-\frac{V_i}{V_T}}\right] \tag{3}$$

The transistor voltages can be evaluated in three different situations. The analysis assumes that $V_{dd} >> V_i$, which drop out all the V_i terms, and it also considers the fact that $V_i >> V_T$, so that the $\left[1 - e\left(-V_i / V_T\right)\right]$ term can be ignored.

The first situation is represented by the voltage V_1 in Fig. 1. In this case, it is possible to associate every transistor connected in that node by series-parallel association. The terms W_{above} and W_{below}, in equation (4), represent the effective width of the transistor over and under the node V_i, respectively. For this condition, V_i is given by

$$V_i = \frac{\eta V_{dd} + nV_T \ln\left(\frac{W_{above}}{W_{below}}\right)}{1 + 2\eta + \gamma} \tag{4}$$

The second situation is presented by the voltage V_2 in Fig. 1. In such condition, it is not possible to make series-parallel associations between the transistors connected at *i*-index node. The term V_{above} means the voltage on the transistors above the node V_i. For this state, the voltage V_i is given by

$$V_i = \frac{nV_T \ln\left(\frac{\sum W_{above} e^{\frac{\eta V_{above}}{nV_T}}}{W_{below}}\right)}{1 + \eta + \gamma} \tag{5}$$

Finally, the third situation is illustrated by the voltage V_3 in the same arrangement. This case only happens at the bottom transistors and the analysis cannot assume $V_i >> V_T$, so that the term (V_i/V_T) should not be ignored. To simplify the mathematic calculation, the term $e\left(-V_i / V_T\right)$ in (1) can be expressed by $\left(1 - V_i / V_T\right)$. Then, *Vi* is obtained by equation (6), where $C = 1 + \eta + \gamma$.

$$\frac{C}{n}\left(\frac{V_i}{V_T}\right) + \ln\left(\frac{V_i}{V_T}\right) = \frac{\eta V_{above}}{nV_T} + \ln\left(\frac{W_{above}}{W_i}\right) + \ln\left(\frac{V_{above}}{V_T}\right) \tag{6}$$

3.1 Subthreshold Leakage in Non-Series-Parallel Networks

Standard CMOS gates derived from logic equations are usually composed by series-parallel (SP) device arrangements. When a Wheatstone bridge configuration is observed in the transistor schematic, as observed in some BDD-based networks [11-13], a non-series-parallel topology is identified, as depicted in Fig. 2b.

The proposed model, discussed above, can be used to calculate the voltage across each single transistor, estimating accurately the leakage current. When the model is applied in non-series-parallel configuration sometimes is somewhat difficult to calculate the voltage across certain transistors, as occur in Fig. 2b for the transistor controlled by 'c'. In this case, the transistor receiving 'c' may be ignored at a first moment, until the voltage at one of its terminals is evaluated. For evaluating the other terminal, such device is then included. Similar procedure is suitable for any kind of non-SP networks.

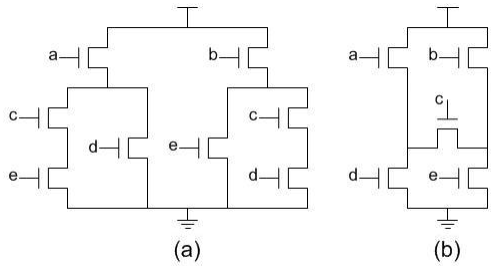


Fig. 2. SP (a) and non-SP (b) transistor arrangements of the same logic function.

3.2 On-Transistors in Off-Networks

Previous analysis considers only off-networks composed exclusively by transistors that are turned off. Usually, in the most cases, the transistors that are turned on could be treated as ideal short-circuits, because the drop voltage across such devices is some orders of magnitude smaller than the drop voltage across the off-transistors.

However, in the case of NMOS transistors switched on and connected to power supply V_{dd}, the drop voltage V_{drop} across them should be taken into account. In the leakage current analysis, this voltage drop is somewhat important when the transistor stacking presents only one off-device at the bottom of the stack, as depicted in Fig. 3.

In the proposed model, the term $V_{dd} - \sum V_j$ in equation (2) is replaced by $V_{dd} - V_{drop} - \sum V_j$.

In arrangements with more than one off-transistor in series, the on-devices could be accurately considered as zero drop voltage (i.e. ideal short-circuit). Similar analysis is valid for PMOS transistors in off-networks when they are connected to the ground node.

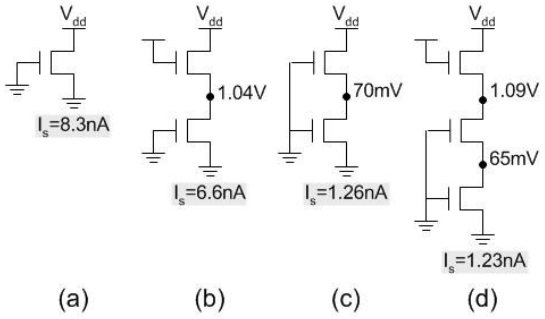


Fig. 3. Influence of on-transistor in off-stack current.

4 Results

In order to validate this model, the results obtained analytically were compared to Hspice simulation results, considering commercial 130nm CMOS process parameters and operating temperature at 100°C. It is known that down to 100nm processes, the subthreshold currents represent the main leakage mechanism, while for sub-100nm technologies, the model presented herein should be combined with gate oxide leakage estimation, already proposed by other authors [3][9].

The parameters used in the analytical evaluation were: V_{dd} = 1.2V, V_{drop} = 0.14, I_0 = 20.56mA, η = 0.078. γ = 0.17, n = 1.45, and W = 0.4μA. In a first moment, transistors with equal sizing were applied to simplify the analysis.

The leakage current was calculated and correlated with Hspice results for several pull-down NMOS off-networks, depicted in Fig 4. The results presented in Table I show a good agreement between the model and the simulation data, showing an absolute average error of less than 10%. It is interesting to note that the static current in networks (f), (g), and (h) from Fig. 4, not treated by previous models, are accurately predicted. The main difference is observed for structures (c) and (d), when the model assumes $V_i < V_T$ and the term $e^{(-V_i / V_T)}$ is changed by $(1 - V_i / V_T)$.

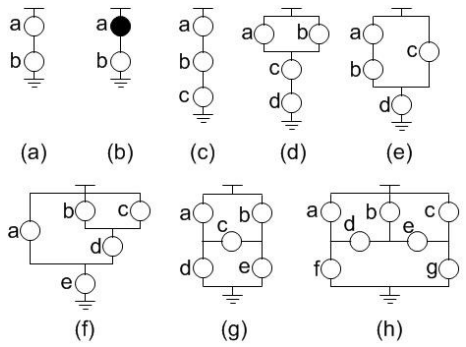


Fig. 4. Pull-down NMOS networks ('●' and 'o' represent on- and off-device, respectively).

Table II and III correspond to the results related to both NMOS trees in Fig. 4(f) and Fig. 4(g), respectively. In these tables, the input dependence leakage current is evaluated for all input combinations. In some cases, different input vectors result in equivalent off-device arrangements. For that, the Hspice values are presented for minimum and maximum values obtained applying the set of equivalent input states. Moreover, the previous model presented in [5] was also calculated for such logic states. Note that, the first input vector in both cases, which represents the entire network composed by off-switches, is not treated by the previous model proposed in [5]. Basically, different values from both methods are obtained when on-transistors are considered in the off-networks, providing more accurate correlation with the electrical simulation results.

Table I. Subthreshold leakage current (nA) related to the off-networks depicted in Fig. 4.

Network	HSPICE	Model	Diff.(%)
(a)	1.26	1.26	-
(b)	6.58	6.60	0.30
(c)	0.69	0.75	8.70
(d)	0.72	0.77	6.94
(e)	1.29	1.28	0.78
(f)	1.29	1.28	0.78
(g)	2.52	2.53	0.40
(h)	2.56	2.54	0.79

Table II. Input dependence leakage estimation (nA) in the logic network from Fig. 5(f).

Input-state (abcde)	HSPICE results*	Proposed model	Previous model [5]
00000	1.29	1.28	-
00001	9.71	9.65	9.65
00010	1.43	1.34	1.34
00011	25.00	25.02	25.02
00100 [a]	1.36/1.37	1.30	1.31
00101 [b]	14.91/15.14	14.94	16.69
00110 [c]	6.30/6.73	6.60	8.34

* The HSPICE value is given for min./max. currents related to such equivalent vectors.
[a,b,c] Equivalent vectors – a.[01000, 01100]; b.[01001, 01101]; c.[01010, 01110, 10000, 10010, 10100, 10110, 11000, 11010, 11100, 11110].

Table III. Input dependence leakage estimation (nA) in the logic network from Fig. 4(g).

Input-state (abcd)	HSPICE results*	Proposed model	Previous model [5]
00000	2.52	2.53	-
00001[a]	10.54/10.54	10.76	10.76
00011[b]	16.68/16.68	16.68	16.68
00100	2.52	2.52	2.52
01000	7.90	7.90	7.91
01010[c]	21.05/21.05	21.54	24.02
01100[d]	12.55/13.15	13.20	16.68
10000	7.90	7.90	9.65

* The HSPICE value is given for min./max. currents related to such equivalent vectors.
[a,b,c,d] Equivalent vectors – a.[00010]; b.[00101, 00110, 00111]; c.[10001]; d.[10100, 11000, 11100].

Fig. 5 shows the subthreshold average leakage current related to the NMOS networks illustrated in Fig. 4(f) and Fig. 4(g). As discussed before, the previous model from [5] cannot estimate the subthreshold current for the first input vector in both cases, and it is not considered in the average static current calculation. Unlike the previous model, the proposed one presents results close to Hspice simulations. The main reason for that is the influence of on-transistors in off-networks.

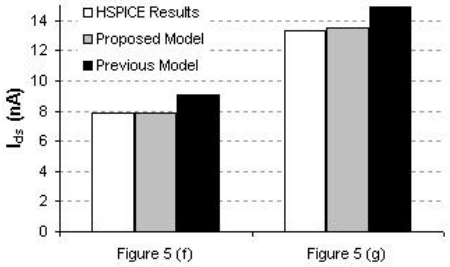


Fig. 5. Average subthreshold leakage current for pull-down networks in Fig. 4(f) and Fig. 4(g).

In terms of combinational circuit static dissipation analysis, the technology mapping task divides the entire circuit in multiple logic gates or cells. Thus, they can be treated separately for the leakage estimation, since the input state of each cell is known according to the primary input vector of the circuit. A complex CMOS logic gate, whose transistor sizes were determined by considering the Logical Effort method [14], is depicted in Fig. 6. Table IV presents the comparison between electrical simulation data and the model calculation, proving the method in sized gates.

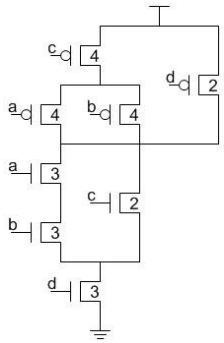
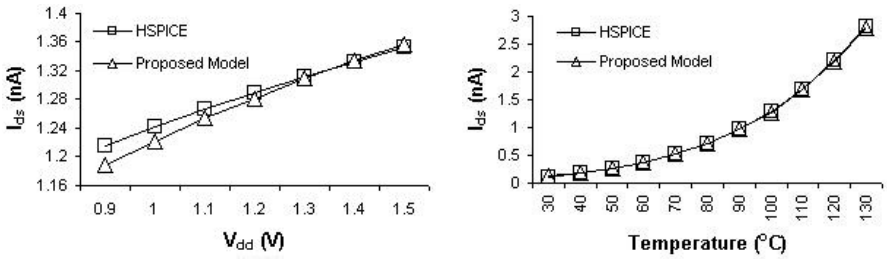


Fig. 6. CMOS gate, with different transistor sizing, according to the Logical Effort [14].

Finally, the proposed model has been verified to the variation of power supply voltage and operating temperature, as depicted in Fig. 7. The influence of temperature variation in the predicted current shows good agreement with Hspice results. On the other hand, the difference between the subthreshold currents obtained from electrical simulation and analytical modeling to voltage variation can be justified by eventual inaccuracy in the parameter extraction. Fig. 8 shows the leakage current analysis in respect to the threshold voltage variation, validating the proposed method for this factor, critical in the most advanced CMOS processes.

Table IV. Subthreshold leakage current related to the CMOS complex gate depicted in Fig 6.

Input state (abcd)	HSPICE results	Proposed model	Diff(%)
0000	4.01	4.13	3.0
0001	20.67	20.68	0.0
0010	19.93	19.99	0.3
0011	44.52	43.27	2.8
0100	4.44	4.29	3.4
0101	42.34	42.37	0.1
0110	19.93	19.99	0.3
0111	43.38	43.27	0.3
1000	4.40	4.26	3.2
1001	36.81	36.50	0.8
1010	19.93	19.99	0.3
1011	43.38	43.27	0.3
1100	19.50	19.99	2.5
1101	96.67	96.92	0.3
1110	20.43	19.99	2.2
1111	20.48	20.21	1.3



(a) (b)

Fig. 7. Subthreshold current in (a) supply voltage and (b) operating temperature variations.

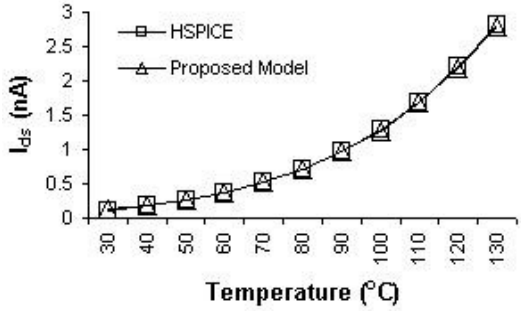


Fig. 8. Subthreshold current according to the operating temperature variation.

5 Conclusions

A new subthreshold leakage current model has been presented for application in general CMOS networks, including series-parallel and non-series-parallel transistor arrangements. The occurrence of on-devices in off-networks is also taken into account. These features make the model useful for different logic styles, including BDD-derived networks, improving previous works not suitable to such a kind of structure. The proposed model has been validated considering a 130nm CMOS technology, in which the subthreshold current is the most relevant leakage mechanism. In the case of sub-100nm processes where gate leakage becomes more significant, the present work should be combined with already published works which address the interaction between the subthreshold and the gate oxide leakage currents, such as proposed by Yang *et al.* [3], in order to improve the accuracy of the total leakage estimation.

6 References

1. K. Roy, S. Mukhopadhyay and H. M.-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits", Proceedings of the IEEE, vol. 91, no. 2, Feb. 2003, pp. 302-327.
2. J. H. Anderson and F. N. Najm, "Active Leakage Power Optimization for FPGAs", IEEE Trans. on CAD, vol. 25, no. 3, Mar. 2006, pp. 423-437.
3. S. Yang, W. Wolf, N. Vijaykrishnan, Y. Xie and W. Wang, "Accurate Stacking Effect Macro-modeling of Leakage Power in Sub-100nm Circuits", Proc. Int. Conference on VLSI Design, Jan. 2005, pp. 165-170.
4. R. X. Gu and M. I. Elmasry, "Power Distribution Analysis and Optimization of Deep Submicron CMOS Digital Circuit", IEEE J. Solid-State Circuits, vol.31, May 1996.
5. Z. Cheng, M. Johnson, L. Wei and K. Roy, "Estimation of Standby Leakage Power in CMOS Circuits Considering Accurate Modeling of Transistor Stacks", Proc. Int. Symposium Low Power Electronics and Design, Aug. 1998, pp. 239-244.
6. K. Roy, S. Prasad, "Low-Power CMOS VLSI Circuit Design", John Wiley & Sons, 2000.
7. S. G. Narendra and A. Chandrakasan, "Leakage in Nanometer CMOS Technologies", Springer, 2006.
8. J. L. Rosseló, S. Bota and J. Segura, "Compact Static Power Model of Complex CMOS Gates", PATMOS, 2005.
9. D. Lee, W. Kwong, D. Blaauw and D. Sylvester, "Analysis and Minimization Techniques for Total Leakage Considering Gate Oxide Leakage", Proc. DAC, June 2003, pp.175-180.
10. S. Gavrilov *et al*., "Library-less Synthesis for Static CMOS Combinational Logic Circuits", Proc. ICCAD, Nov. 1997, pp. 658-662.
11. C. Yang and M. Ciesielski. "BDS: a BDD-based logic optimization system". IEEE Trans. CAD, vol.21, no.7, July 2002, pp. 866-876.
12. P. Lindgren, M. Kerttu, M. Thornton and R. Drechsler. "Low power optimization technique for BDD mapped circuits". ASP-DAC 2001, pp. 615-621.
13. R. S. Shelar and S. Sapatnekar. "BDD decomposition for delay oriented pass transistor logic synthesis". IEEE Trans. VLSI, vol.13, no.8, Aug. 2005, pp. 957-970.
14. B. J. Sheu, D. L. Scharfetter, P. K. Ko, M. C. Jeng, "BSIM: Berkeley Short-Channel IGFET Model for MOS Transistors", IEEE J.Solid-State Circuits, vol. SC-22, Aug. 1987.
15. I. Sutherland, B. Sproull, and D. Harris, "Logical Effort: Designing Fast CMOS Circuits", San Francisco, CA: Morgan Kaufmann, Jan. 1999.