Accurate Estimation of Total Leakage Current in Scaled CMOS Logic Circuits Based on Compact Current Modeling*

Saibal Mukhopadhyay, Arijit Raychowdhury, and Kaushik Roy
Dept. of Electrical & Computer Engineering, Purdue University, West Lafayette, IN-47907-1285, USA
<sm, araycho, kaushik>@ecn.purdue.edu

ABSTRACT
Dramatic increase of subthreshold, gate and reverse biased junction band-to-band-tunneling (BTBT) leakage in scaled devices, results in the drastic increase of total leakage power in a logic circuit. In this paper a methodology for accurate estimation of the total leakage in a logic circuit based on the compact modeling of the different leakage current in scaled devices has been developed. Current models have been developed based on the exact device geometry, 2-D doping profile and operating temperature. A circuit level model of junction BTBT leakage (which is unprecedented) has been developed. Simple models of the subthreshold current and the gate current have been presented. Here, for the first time, the impact of quantum mechanical behavior of substrate electrons, on the circuit leakage has been analyzed. Using the compact current model, a transistor has been modeled as a Sum of Current Sources (SCS). The SCS transistor model has been used to estimate the total leakage in simple logic gates and complex logic circuits (designed with transistors of 25nm effective length) at the room and at the elevated temperatures.

Categories & Subject Descriptor:
B.6.3 [Logic Design]: Design Aids – *Simulation, Estimation*
B.7.2 [Integrated Circuits]: Design Aids – *Simulation, Estimation*

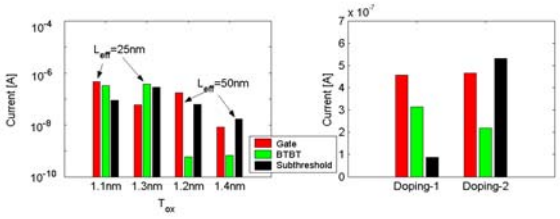
General Terms: Design, Experimentation, Theory.

Keywords: doping profiles, leakage, tunneling, threshold voltage.

1. INTRODUCTION

Aggressive scaling of CMOS devices in each technology generation has resulted in higher integration density and performance. Simultaneously, supply voltage scaling has reduced the switching energy per device. However, the leakage current (i.e. the current flowing through the device in its "off" state) has increased drastically with technology scaling [1] and leakage power has become a major contributor to the total power. Hence, the estimation of the total leakage is absolutely necessary for estimating total power and designing low power logic circuits.

Among different leakage mechanisms in scaled devices [1], three major ones can be identified as: Subthreshold leakage, Gate leakage and reverse biased drain-substrate and source-substrate junction Band-To-Band-Tunneling (BTBT) leakage [1]. The threshold voltage (V_{th}) scaling and the V_{th} reduction due to Short Channel Effects (SCE) [1], result in an exponential increase in the subthreshold current. The oxide thickness scaling, required to maintain reasonable SCE immunity, results in a considerable direct tunneling current through the gate insulator of the transistor [1], [2]. In scaled devices, the higher substrate doping density and the application of the "halo" profiles (used to reduce SCE) [2] cause significantly large BTBT current through the reverse biased drain-substrate and source-substrate junctions. In the small devices each of the different leakage components increases resulting in a dramatic increase of the overall leakage. The magnitudes of each of these components depend strongly on the device geometry (namely, channel length, oxide thickness and transistor width) and the doping profiles as shown in Fig. 1.



(a) (b)

Figure 1: Variation of different leakage components with (a) technology generation and oxide thickness; (b) doping profile. "Doping-1" has a different halo profile than "Doping-2". (simulation)

Different leakage current components in the devices vary differently with varying temperature. Subthreshold and BTBT leakage show a strong dependence of temperature, whereas gate leakage is relatively insensitive to temperature variations. Since digital VLSI circuits usually operate at elevated temperatures, estimation of the various leakage components and the total leakage in devices and circuits is necessary both at room and elevated temperatures.

In this paper we have developed a methodology for accurately estimating the total leakage of a logic circuit for different primary input vectors, based on the knowledge of, (a) the device geometry, (b) the exact 2-D doping profile of the device and (c) the operating temperature. Although, a number of previous work are reported on the estimation of leakage in logic circuits [3], [3], [4] but they have only considered the subthreshold leakage. However, as shown in Fig. 1, gate and BTBT leakage are also becoming extremely important and thus cannot be neglected for estimation of total leakage. We have developed a compact circuit level model of BTBT leakage in a MOSFET with halo [2] and retrograde doping [2]. To the best of our knowledge it is unprecedented. A simple and reasonably accurate model of the subthreshold current has been developed based on the exact 2-D doping profile. Here, for the first time, we have evaluated the direct impact of quantization of the electron energy in the substrate [2], on the leakage in logic circuits. We have used the gate leakage model presented in [5], [6]. Finally, the compact models of the leakage components have been used to model a transistor as a Sum of Current Sources (SCS) for accurate leakage estimation. A numerical solver has been developed to evaluate leakage in simple logic gates by solving the Kirchoff's Current Law (KCL) at intermediate nodes, using SCS transistor model. A method for calculating the total leakage of a logic circuit by adding the individual leakage contribution of its constituent gates is also proposed. We have verified the leakage estimation technique on simple logic gates, such as INVERTER, NAND and

*This research is supported in part by SRC, GSRC, Intel and IBM Corporations.

DAC2003, June 2-6, Anaheim, California, USA

NOR gate, and on complex logic circuits, such as, an adder and a multiplier.

2. LEAKAGE ESTIMATION STEPS:

In scaled devices leakage strongly depends on transistor geometry, doping profile (Fig. 1) and temperature. Hence, accurate estimation of total leakage of a logic circuit starts with the accurate description (device geometry, doping profile) of the transistor used to fabricate the circuit and the operating temperature. The steps followed to estimate the total leakage are shown in Fig. 2.

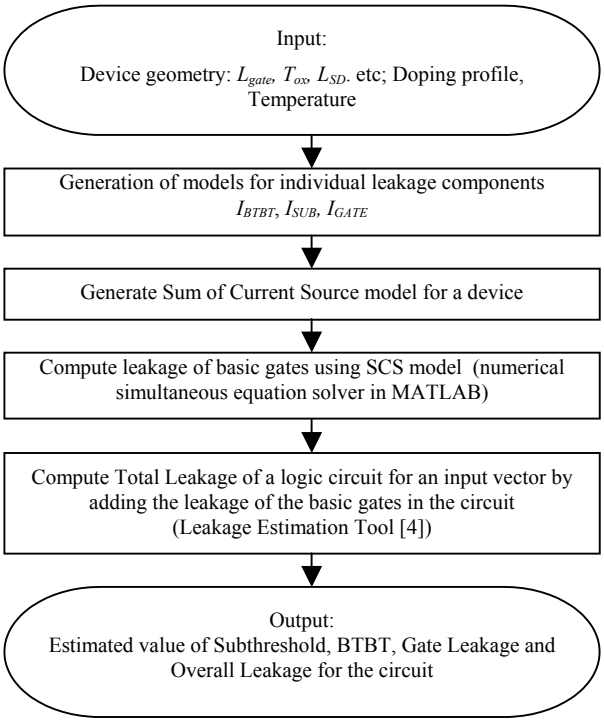


Figure 2: Leakage estimation steps.

3. MODELING LEAKAGE COMPONENTS

This section represents the general approach used to formulate the model for the BTBT, subthreshold and gate leakage, in a MOSFET. The formulation, developed for NMOS transistors, can be easily extended to PMOS transistors. Device structures with Gaussian-shaped channel ("super halo" channel doping) and source/drain (S/D) doping profiles have been considered while deriving these models. A schematic of the device structure (symmetric about the middle of the channel) is shown in Fig. 3 [7]. The 2-D Gaussian doping profile in the channel ($N_a(x,y)$) and S/D ($N_{sd}(x,y)$) can be represented as [7],[8]:

$$x > 0,$$
$$N_{(a/sd)}(x,y) = A_{(p/sd)}\Gamma_{x(a/sd)}(x)\mathrm{K}_{y(a/sd)}(y) + N_{SUB}$$
$$\text{where,}\quad \mathrm{K}_{y(a/sd)}(y) = exp\left(\frac{-(y-\alpha_{a/sd})^2}{\quad 2}\right) \qquad (1)$$
$$\text{ar} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad sd)\Big]$$

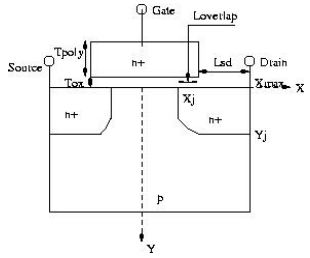


Figure 3: Architecture of the device

where, suffix *a* and *sd* represents channel and S/D region respectively. A_p and A_{sd} represent the peak "halo" and S/D doping respectively. N_{SUB} is the constant uniform doping in the bulk and is much less compared to contributions from Gaussian profiles at and near the channel and S/D regions. Parameters α_a, α_{sd} (=0), β_a and β_{sd} control the positions and σ_{ya} , σ_{xa} and σ_{ysd} , σ_{xsd} control the variances of the Gaussian profiles in channel and S/D regions [7], [8]. Unless otherwise specified in this paper we have used NMOS (N_{ref}) and PMOS (P_{ref}) transistors with L_{eff}=25nm, W_{eff}=1μm and channel doping profile α_a=0.018μm, σ_{ya}=0.016μm β_a=0.016μm, σ_{xa}=0.020μm and S/D profile from [8].

3.1. Modeling Band-to-band leakage current (I_{BTBT}):

A high electric field across a reverse biased p-n junction causes significant current to flow through the junction due to tunneling of electrons from the valence band of the p-region to the conduction band of the n-region (Fig. 4(a)) [2]. Tunneling occurs when the total voltage drop across the junction (applied reverse bias (V_{app}) + built-in voltage (ψ_{bi})) is more than the band-gap (Σ_g). The tunneling current density through a silicon p-n junction is given by [2]:

$$J_{b-b} = A\frac{EV_{app}}{\Sigma_g^{1/2}}exp\left(-B\frac{\Sigma_g^{3/2}}{E}\right)$$
$$A = \frac{\sqrt{2m^*}\,q^3}{4\pi^3\hbar^2}, \text{ and } B = \frac{4\sqrt{2m^*}}{3q\hbar} \qquad (2)$$

where, m^* is the effective mass of electron, E is the electric field at the junction, q is the electronic charge and $\hbar$ is the reduced Plank's constant.

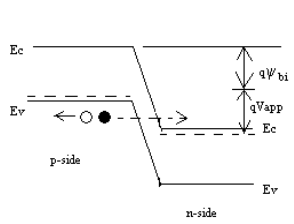
In an NMOSFET when the drain and/or the source is biased at a potential higher than that of the substrate, a significant BTBT current flows through the drain-substrate and/or the source-substrate junctions. The total BTBT current in the MOSFET is the sum of the currents flowing through two junctions and is given by:

$$I_{BTBT} = w_{eff}\int_l J_{b-b}(x,y)dl\Big|_{drain} + w_{eff}\int_l J_{b-b}(x,y)dl\Big|_{source} \qquad (3)$$

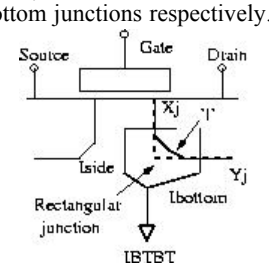
where, $J_{b-b}(x,y)$ is the current density at a point (x,y) at the junction. For a symmetric device the current expressions for the drain and the source junctions will be identical. Hence, here we have considered only the drain junction. The integration in (3) has to be done along the junction line 'l' (Fig. 4(b)) within the tunneling region. This integration cannot be solved analytically. To obtain an accurate analytical estimate of the total current, we can apply the "rectangular junction" approximation (Fig. 4(b)). Using this approximation the total current through the drain junction is given by:

$$I_{BTBTdrain} = I_{side} + I_{bottom}$$
$$= \left[w_{eff}\int_{y_1}^{y_2} J_{b-b}(X_j,y)dy\right]_{side} + \left[w_{eff}\int_{x_1}^{x_2} J_{b-b}(x,Y_j)dx\right]_{bottom} \qquad (4)$$

where, X_j and Y_j are the position of the side and the bottom junctions respectively (Fig. (4(b)). y_1 to y_2 and x_1 to x_2 are the tunneling regions in the side and the bottom junctions respectively.



(a)



(b)

Figure 4: Band-to-band tunneling in MOSFET. (a) physical picture of valence band electron tunneling in a reversed bias p-n junction, (b)Circuit modeling of tunneling current in drain-substrate junction of MOSFET with rectangular junction approximation.

Here, we present the derivation of the side junction current. The bottom junction current can be derived following a similar procedure.

Due to the non-uniform doping in the substrate and the drain region, the integration in (4) can not be solved analytically. Hence, we approximate the integral using an average tunneling current density ($J_{b\text{-}bside}$) which is determined by the average electric field (E_{side}) across the junction. Hence, using (4) and (2), I_{side} is given by:

$$I_{side} = w_{eff}\left|y_2 - y_1\right| J_{b-bside} = w_{eff}\left|y_2 - y_1\right| \frac{A E_{side} V_{app}}{\Sigma_g^{1/2}} exp\left(-\frac{B\Sigma_g^{3/2}}{E_{side}}\right) \quad (5)$$

To evaluate E_{side} we simplify the p-n junction as a step junction with doping at the p and n side equal to N_{aside} and N_{dside}, respectively. The detail argument leading to that simplification can be found in our technical report [9]. Using the expression for the electric field at a step junction [10], E_{side} is given by:

$$E_{side} = \sqrt{\frac{2qN_{aside}N_{dside}(V_{app} + \psi_{biside})}{\varepsilon_{si}(N_{aside} + N_{dside})}} \quad (6)$$

where, N_{aside} and N_{dside} are given by:

$$N_{aside} = \frac{1}{\left|y_2 - y_1\right|}\int_{y_1}^{y_2} N_a(X_j, y)dy$$
$$N_{dside} = \frac{1}{\left|y_2 - y_1\right|}\int_{y_1}^{y_2} \left[N_{sd}(x = \beta_{sd}, y) - N_a(x = \beta_a, y)\right]dy \quad (7)$$

ψ_{biside} (the built-in potential for a step junction [10]) is given by :

$$\psi_{biside} = \frac{kT}{q} ln\left(\frac{N_{aside}N_{dside}}{n_i^2}\right) \quad (8)$$

X_j and Y_j are found by solving following equations:

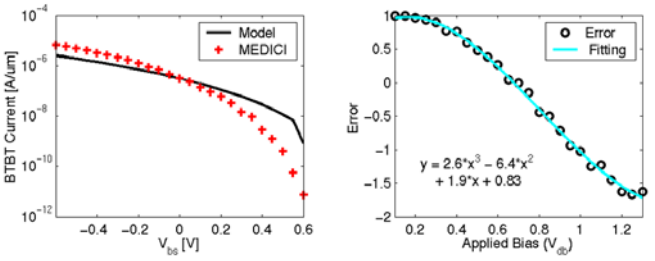
$$N_{sd}(X_j, y = 0) = N_a(X_j, y = 0)$$
$$N_{sd}(x = x_{\max}, Y_j) = N_a(x_{\max}, Y_j) \quad (9)$$

For simplicity the whole side junction is assumed to be tunneling (i.e y_1=0 and $y_2 = Y_j$). For bottom junction x_1=X_j and $x_2 = x_{max}$.

Using expressions from (7)-(9), into (13) E_{side} (and similarly E_{bottom}) can be obtained. E_{side} (and E_{bottom}) can be used in (5) to obtain $J_{b\text{-}bside}$ (similarly $J_{b\text{-}bbottom}$). If $(V_{app}+\psi_{biside}) < \Sigma_g/q$, then no tunneling occurs and $J_{b\text{-}bside}$ is zero (similar argument holds for $J_{b\text{-}bbottom}$). Hence, the total BTBT current in the drain junction is given by:

$$I_{BTBTdrain} = w_{eff}(Y_j J_{b-bside}) + w_{eff}\left(\left|x_{\max} - X_j\right| J_{b-bbottom}\right) \quad (10)$$

For a 25nm transistor, the comparison of the analytical model given in (10) and the simulated data from MEDICI [11] shows close match for small reverse and forward substrate bias (Fig. 5(a)). However, deviations are observed at high forward (i.e. low V_{app}) and reverse (i.e. even higher V_{app}) substrate bias. At high V_{app} , the average electric field (calculated using average doping density) used in the model is considerably less than the peak field (at the peak doping region). Since the tunneling is dominated by the peak field, the analytical current is less than the simulated one at high V_{app}. In the low bias region, reduction of V_{app} considerably reduces the tunneling volume. The model does not consider the reduction of the tunneling volume. Moreover, the derived field is based on the abrupt junction approximation which also predicts a higher field. Hence, the evaluated current is higher than the simulated current at low V_{app} (i.e. high forward substrate bias). Also, at a high gate voltage, (a) small increase in the potential near the substrate side of the side junction and (b) non-negligible voltage drop at the S/D series resistance caused by the high "on" current flowing through the transistor reduce the effective applied reverse bias across the junction. Hence, the BTBT current reduces by a small amount at a high gate bias. Exact modeling of these effects requires calculation of the tunneling rate at each point, which makes the formulation of a compact circuit model of the current extremely difficult. To take care of these effects an empirical parameter *(a_0)*, function



(a) (b)

Figure 5: Variation of BTBT current with substrate bias. (a) Comparison of analytical and simulated data for N_{ref}. (b) Variation of error

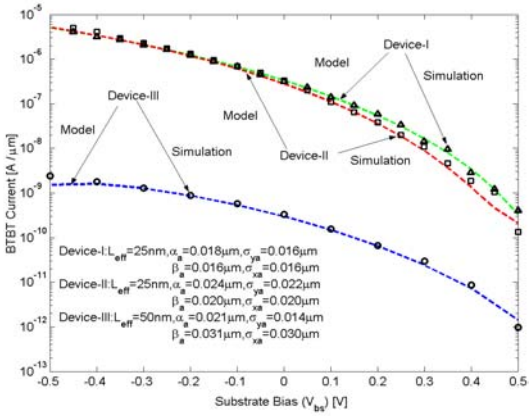


Figure 6: Variation of BTBT current with substrate bias for different devices.

($\lambda_{BTBT}(V_{app})$) and an empirical gate correction factor (δ_g) have been introduced in the model. With these corrections the current due to the drain junction (or source) is given by:

$$I_{BTBTcorrected} = a_0 I_{BTBTdrain}\left(1 - \lambda_{BTBT}(V_{app})\right)\left(1 - \delta_g V_G\right) \quad (11)$$

where a_0 is the zero substrate bias multiplication factor defined as the ratio of the actual BTBT (measured/simulated) current and the analytical value at zero substrate bias and $\lambda_{BTBT}(V_{app})$ is an empirical function (for drain-substrate junction V_{app}=V_{db} and for source-substrate junction V_{app}=V_{sb}). From experiments it was found that a cubic function gives a good fit of the simulated result (Fig. 5(b)). Gate correction factor (δ_g) can be calculated from the measured/simulated BTBT current at a low and a high gate bias.

The final expression for the total BTBT current is given by:

$$I_{BTBT} = \sum_{i=drain,source}\left[\left(\sum_{k=side,bottom} I_{k\_i}\right)(1 - \delta_g V_G)\right](1 - \lambda_{BTBT}(V_{ib}))$$
$$I_{k\_i} = w_{eff} A \frac{E_{k\_i}}{\Sigma_g^{1/2}} V_{ib} exp\left(-\frac{B\Sigma_g^{3/2}}{E_{k\_i}}\right) \quad (12)$$

The parameters, namely, E_{side_i}, E_{bot_i} can be evaluated following the procedure discussed above. Fig. 6 shows a comparison plot of the analytical results with the simulated results from MEDICI for devices with L_{eff}=25nm (V_{dd}=*0.7V*) and 50nm (V_{dd}=*0.9V*) and different doping profiles. It shows that, the analytical results follow very closely the simulated results.

3.2. Modeling subthreshold current (I_{ds}):

In the "off" state of a device ($V_{gs} < V_{th}$), diffusion of the minority carriers through the channel cause current to flow from the drain to the source of a transistor. This is known as the subthreshold current. The subthreshold current is given by [2]:

$$I_{sub} = \frac{w_{eff}}{L_{eff}}\mu\sqrt{\frac{q\varepsilon_{si}N_{cheff}}{2\Phi_s}}v_T^2 exp\left(\frac{V_{gs} - V_{th}}{nv_T}\right)\left(1 - exp\left(\frac{-V_{ds}}{v_T}\right)\right) \quad (13)$$

where, N_{cheff} is the effective channel doping, Φ_s is surface potential, n is subthreshold swing and v_T is thermal voltage given by kT/q. Using charge sharing model the threshold voltage can be expressed as [12], [13]:

$$V_{th} = V_{FB} + \Phi_s + \gamma\sqrt{\Phi_{s0} - V_{bs}}\left(1 - \lambda\frac{X_d}{L_{eff}}\right) \quad (14)$$

where, V_{FB} is the flat-band voltage, Φ_{s0} is the zero bias surface potential, γ is the body factor, C_{ox} is the oxide capacitance, X_d is the depletion layer thickness and λ is a fitting parameter (≈ 1). The expressions for each of the parameters are given in [12], [13]. The surface potential (Φ_s) of short channel devices is reduced from its zero bias value due to short channel effects like DIBL and V_{th} roll-off as given in [12]. In scaled devices, due to high electric field at the surface (E_s) and high substrate doping, the quantization of inversion-layer electron energy modulates V_{th}. Quantum-mechanical behavior of the electrons increases V_{th} , thereby reducing the subthreshold current, since more band bending is required to populate the lowest subband, which is at a energy higher than the bottom of the conduction band. When E_s is larger than 10^6 V/cm, electrons occupy only the lowest subband. In that case, the quantization effect can be modeled as an increase in threshold voltage by an amount ΔV_{QM}, given by [2]:

$$\Delta V_{QM} = \left(1 + \frac{3t_{ox}}{X_d}\right)\left(\frac{\Sigma_0}{q} - \frac{kT}{q}\ln\left(\frac{8\pi q m_d E_s}{h^2 N_C}\right)\right) \quad (15)$$

where, Σ_0 is the lowest subband energy given by [2] $\Sigma_o = \left[\frac{3hq_s E_s}{4\sqrt{2m_x}}\frac{3}{4}\right]^{2/3}$, N_C is the effective conduction band density of states, m_x is the quantization effective mass of electron and m_d is the density of states effective mass of electron.

The effect of 2D Gaussian profile is used to calculate the effective channel (N_{cheff}) and S/D doping (N_{sdeff}) as shown below:

$$N_{cheff} = \frac{1}{\Delta_{ch}}\iint_{\Delta_{ch}} N_a(x,y)dxdy + N_{sub} \quad (16)$$

$\Delta_{ch}=L_{eff}\alpha_a$ is the channel area under the influence of gate.

$$N_{sdeff} = \frac{1}{\Delta_{SD}}\iint_{\Delta_{SD}} N_d(x,y)dxdy \quad (17)$$

where, $\Delta_{SD}=(L_{overlap}+L_{sd})Y_j$ is the S/D area, $L_{overlap}$ is the gate and the S/D overlap length and L_{sd} is the S/D length as shown in Fig. 3. The simplified model shows reasonable match with the simulated result from MEDICI under substrate and drain bias variation (Fig. 7) with and without quantum correction. Substantial reduction in the subthreshold current is observed using the quantum correction.

3.3. Modeling Gate Direct Tunneling Current (I_{gate})

Gate direct tunneling current is due to the tunneling of electrons (or holes) from the bulk silicon and source/drain (S/D) overlap region through the gate oxide potential barrier into the gate [2]. The tunneling current increases exponentially with decrease in the oxide

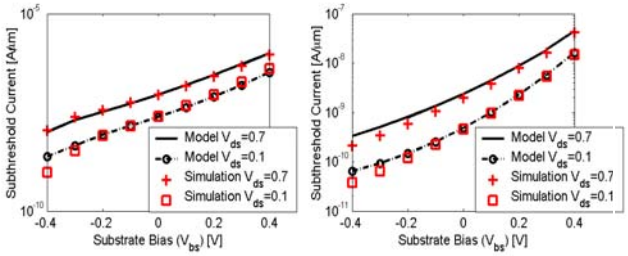


Figure 7: Variation of subthreshold leakage with substrate bias (V_{bs}) and drain bias (V_{ds}) for NMOS transistor N_{ref} (a) Without and (b) With quantum correction.

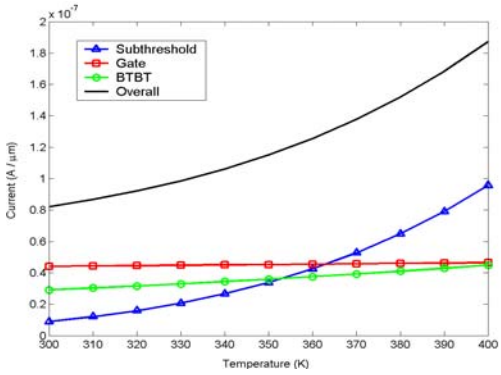


Figure 8: Variation of leakage current components with temperature.

thickness and increase in the potential drop across oxide. Major components of gate tunneling in a scaled MOSFET device are [5]: (1) Gate to S/D overlap region current (Edge Direct Tunneling (EDT)) components (I_{gso} & I_{gdo}),(2) Gate to channel current (I_{gc}), part of which goes to source (I_{gcs}) and rest goes to drain (I_{gcd}), (3) Gate to substrate leakage current (I_{gb}). Accurate modeling of each of the components is based on the following equation [5], [6]:

$$J_{DT} = A_g\left(\frac{T_{oxref}}{t_{ox}}\right)^{ntox}\left(\frac{V_g V_{aux}}{t_{ox}^2}\right)exp\left(-B_g t_{ox}\left(\alpha_g - \beta_g|V_{ox}|\right)\left(1+\gamma_g|V_{ox}|\right)\right) \quad (18)$$

where, T_{oxref} is the reference oxide thickness at which all parameters are extracted, $ntox$, α_g, β_g, γ_g are fitting parameters and V_{aux} is an auxiliary function. A_g and B_g are physical parameters given in [6]. We have used the current models from [5], [6] with the effective channel and the S/D doping density obtained from (16) and (17).

3.4: Temperature Dependence of Leakage Components

The different leakage components show different temperature dependence. Subthreshold current, which is governed by the carrier diffusion, increases exponentially with temperature due to (a)

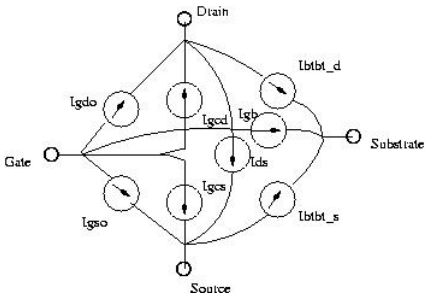


Figure 9: Sum of Current Source model of a transistor.

reduction in threshold voltage and (b) increase in thermal voltage (v_T). The gate tunneling current is almost insensitive to temperature since the tunneling probability and the electric field across the oxide does not strongly depend on temperature. Band-gap of Silicon reduces with an increase in the operating temperature [2], [10]. Due to the band-gap narrowing, BTBT increases with temperature. Fig. 8 shows the variation of each leakage component with temperature in an NMOS transistor (N_{ref}) using the models introduced in the last three sub-sections. It is observed that, at room temperature (T=300K) gate leakage and BTBT dominates over subthreshold current, while at elevated temperature, subthreshold leakage is the dominant component of overall leakage.

4. MODELING OVERALL LEAKAGE

The overall leakage in a device is the summation of the three major leakage components. We can model the overall leakage ($I_{overall}$) as:

$$I_{overall} = I_{BTBT} + I_{sub} + I_{gate} \quad (19)$$

Hence, for leakage estimation we have modeled the device as a combination of voltage controlled current sources as shown in Fig.

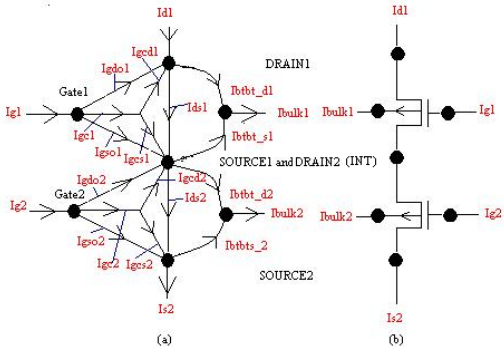


Fig. 10: Circuit configuration with SCS model for a 2-transistor stack, (a) SCS model, (b) transistor-circuit diagram.

9. Based on (12) the BTBT current is modeled as two current sources, one between the drain and the substrate (I_{btbt_d}) controlled by V_{db} and another between the source and the substrate (I_{btbt_s}) controlled by V_{sb}. Each component of gate leakage described in 3.3 is modeled as a current source. I_{ds} models the subthreshold current. The SCS model of the transistor can be effectively used to calculate the overall leakage in a circuit. This model can also be effectively used to describe the SPICE model of a transistor.

5. MODELING OF LEAKAGE IN LOGIC GATES

The SCS model of the transistor can be effectively used to calculate the overall leakage in a circuit. Fig. 10 shows the circuit containing two series connected NMOS transistors and the equivalent SCS model. To calculate the overall leakage, we have to solve the KCL at the intermediate node INT. From Fig. 10 the node equation at INT is given by:

$$I_{ds1} + I_{gcs1} + I_{gso1} - I_{BTBT\_s1} = I_{ds2} - I_{gdo2} - I_{gcd2} + I_{BTBT\_d2} \qquad (20)$$

In circuits involving more than one such node, we will have a set of simultaneous equations that needs to be solved. The overall leakage in the circuit can be defined as the sum of all currents collected at the ground node. Hence, the overall leakage in a CMOS circuit is given by (assuming $V_{bulk}=0$ for all NMOS and $V_{bulk}=V_{dd}$ for all PMOS):

$$I_{leakage} = \sum_{\substack{NMOS+PMOS \\ with\ \ Vg='0'}} I_{gk} + \sum_{NMOS} I_{BTBTk} + \sum_{\substack{NMOS\ with\ source \\ connected\ \ to\ ground}} I_{sourcek} \qquad (21)$$

A numerical equation solver (SCS solver) is written in MATLAB to solve the set of simultaneous equations in a circuit and to determine the overall leakage under a specific input condition. Fig. 11 shows the comparison of the evaluated result and simulated result in MEDICI for a stack of 2 NMOS transistors (N_{ref}), at normal temperature (without the quantum correction). The evaluated results closely match the simulated results.

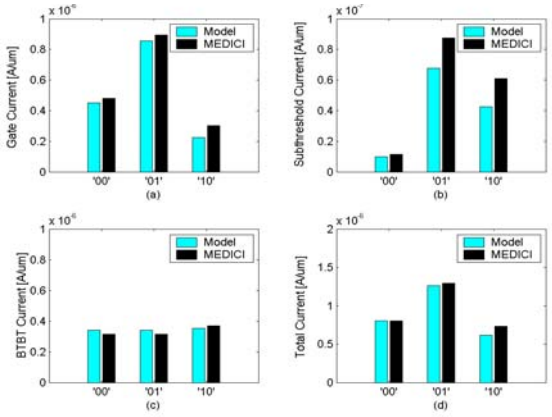


Figure 11: Comparison of simulator (MEDICI) and model current values for a 2-transitor stack for different input vectors: (a) gate, (b) subthreshold, (c) BTBT and (d) total leakage

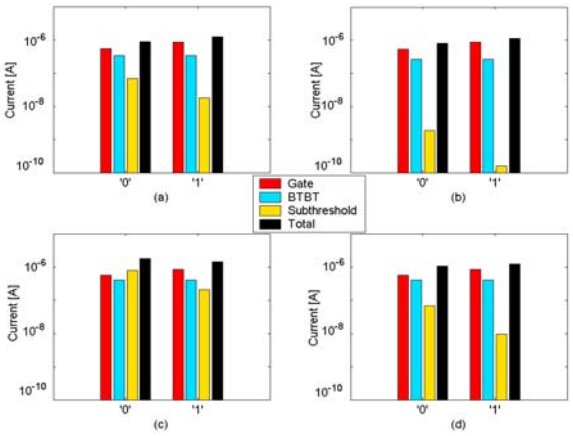


Figure 12: Leakage current of an INVERTER with input '0' and '1'. (a) T=300K and no quantum correction, (b) T=300K and including the quantum correction, (c) T=400K and no quantum correction, (d) T=400K and including the quantum correction.

SCS solver can be used to evaluate the leakage components of basic gates. Fig. 12 and 13 show the different leakage components of INVERTER, NAND and NOR gates (designed with N_{ref} and P_{ref}) at normal (T=300K) and high temperature (T=400K) (with and without the quantum correction). It is observed that, the overall leakage increases considerably with the temperature. At normal temperature the gate leakage dominates the subthreshold leakage and the BTBT leakage, whereas later two are high at higher temperatures. Also, application of the quantum correction reduces the subthreshold current considerably. The solver can easily be extended to handle other logic gates.

5.1. Stacking Effect

Turning "off" more than one transistor in a stack of transistors forces the intermediate node (say INT in Fig. 10) voltage to go to a value higher than zero [1], [4]. This causes a negative V_{gs}, negative V_{bs} (more body effect) and reduced V_{ds} (less DIBL) in the top transistor, thereby considerably reducing the subthreshold current flowing through the stack [1], [4]. This effect, known as the "stacking effect", has been used to reduce the subthreshold leakage in logic circuits in stand-by mode [1], [4]. The estimation tool described here, effectively models the stacking effect for subthreshold, gate and BTBT leakage. Fig. 11 shows that, the input

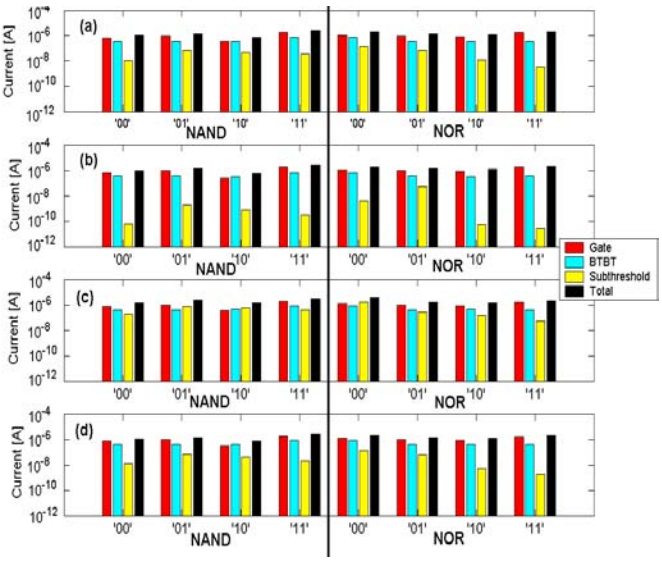


Figure 13: Leakage current of a 2-input NAND and NOR gate with different input. (a) T=300K and no quantum correction, (b) T=300K and including the quantum correction, (c) T=400K and no quantum correction, (d) T=400K and including the quantum correction.

'00' (turning "off" both transistors) produces the minimum subthreshold and BTBT leakage (BTBT leakage in fact does not depend much on stacking (Fig.11)), however, '10' produces the minimum gate leakage condition (reducing V_{gd} of the top transistor to 0, thereby reducing its I_{gdo} to nearly zero). Hence, *the input condition that minimizes the total leakage depends on the relative magnitude of the different components. In devices where gate leakage is the dominant component the input '10' minimizes the total leakage in a stack of two NMOS transistors* (Fig. 11).

6. ESTIMATION OF TOTAL CIRCUIT LEAKAGE

Evaluation of the leakage components of basic logic gates is used to estimate the total leakage in a gate level logic circuit. To evaluate the different leakage components in a logic circuit we have modified the leakage estimation tool described in [4]. Leakage of a logic circuit depends on the primary input vector. The primary input vector is propagated by simulating the circuit level by level. The subthreshold (I_{Tsub}), the gate (I_{Tgate}), the BTBT (I_{Tbtbt}) and the overall leakage ($I_{Toverall}$) through the circuit is defined as the sum of the leakage through each of the basic gates present in the circuit and is given by:

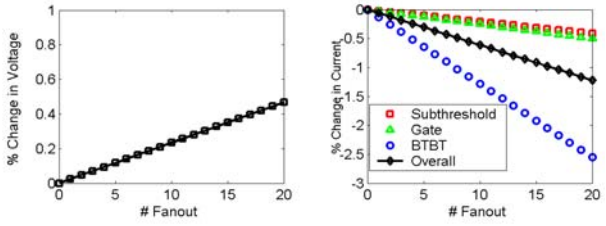
$$I_{Tsub} = \sum_{k=all\ gate} I_{ksub}\ ;\quad I_{Tgate} = \sum_{k=all\ gate} I_{kgate}\ ;\ I_{Tbtbt} = \sum_{k=all\ gate} I_{kbtbt}\ ;$$
$$I_{Toverall} = \sum_{k=all\ gate} I_{koverall} = I_{Tsub} + I_{Tgate} + I_{Tbtbt}\ ; \qquad (22)$$

The total leakage power (P_{leak}) is obtained by:

$$P_{Leak} = I_{Toverall} \times V_{dd} \qquad (23)$$

6.1: Loading effect

The estimation method using (19) neglects the change of the leakage currents of a gate due to the loading by its fanout gates. When, output node (say OUT1) of a gate (say G1) is connected to the input of the other gates, the gate leakage from these other gates get added to the current at OUT1. Hence, the voltage at OUT1 changes, thereby modifying the leakage of the gate G1. This effect can be defined as loading effect. To understand how the leakage of a gate varies with its loading, we studied the variation of the leakage of an inverter (input at '1') with loading (Fig. 14). It is observed that, even for a fanout of 20 the leakage of the inverter remains almost constant. Hence, we conclude that the summation of the leakage of individual gates give a reasonably accurate estimate of the total leakage of a circuit.



(a) (b)

Figure 14: Percentage change in the (a) output voltage and (b) leakage components in an inverter due to loading.

6.2: Results

The leakage estimation tool is used to estimate the total leakage in complex logic circuits, under different primary input vectors. Fig. 15 shows the different leakage components along with the total leakage of an 8-bit ripple carry adder and a 2-bit array multiplier circuit (designed using NAND, NOR and INVERTER) averaged over a large number of primary input vectors. The leakage is evaluated at both normal (T=300K) and high (T=400K) temperatures and with and without quantum correction. The result shows that on the average the gate leakage is the dominant component of the total leakage. However, at higher temperature contributions of the subthreshold and the BTBT are increased.

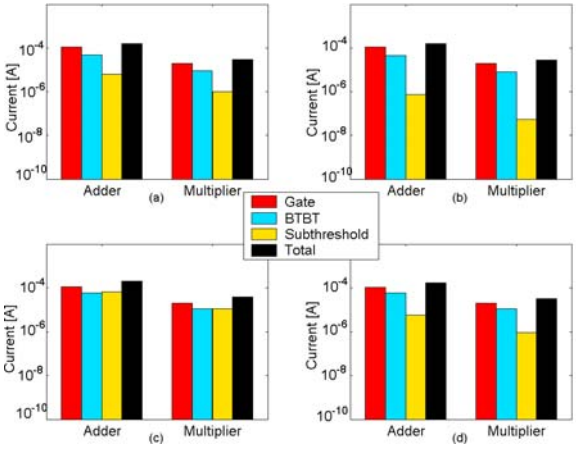


Figure 15: Average leakage current of an 8-bit adder and a 2-bit array multiplier. (a) T=300K and no quantum correction, (b) T=300K and including the quantum correction, (c) T=400K and no quantum correction, (d) T=400K and including the quantum correction

7. SUMMARY and CONCLUSION

In this paper we have developed a compact model for the total leakage in a transistor as the summation of the subthreshold, the BTBT and the gate leakage. It has been shown that for leakage estimation the transistor can be modeled as a Sum of Current Sources, where, each current source describes a leakage mechanism. SCS model can be used to describe a transistor in SPICE circuit simulator. We have developed a CAD tool to estimate the total leakage in CMOS circuits based on the SCS model. The described method for leakage estimation is based on the knowledge of the transistor geometry, 2-D doping profile and operating temperature and can be effectively used to accurately estimate leakage in a scaled CMOS logic circuit.

8. REFERENCE

[1] K. Roy. et. al. "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicron CMOS Circuits", Proceeding of IEEE, Feb, 2003.

[2] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices, Cambridge University Press, 1998.

[3] W. Shiue, "Leakage power estimation and minimization in VLSI circuits", ISCAS, 2001, pp.178-181

[4] Z. Chen, et. al., "Estimation of Standby Leakage Power in CMOS Circuits Considering Accurate Modeling of Transistor Stacks". Int. ISLPED, 1998, pp.239-244.

[5] K. Cao, et.al. "BSIM4 Gate Leakage Model Including Source Drain Partition", IEDM, pp. 815-818, 2000.

[6] *BSIM4.2.1 MOSFET Model,* BSIM Group, UC Berkeley, http://www-device.eecs.berkeley.edu/~bsim3/

[7] Z. Lee, et.al., "Two-dimensional doping profile characterization of MOSFET's by inverse modeling using characteristics in the subthreshold Region", IEEE Trans. on Elec. Dev., Aug. 1999, pp. 1640 –1649.

[8] http://www-mtl.mit.edu/Well/

[9] S. Mukhopadhyay, et.al, "Modeling and estimation of total leakage in scaled CMOS logic circuits", Technical report, TR-ECE 03-01, School of ECE, Purdue University, USA.

[10] R. Pierret, Semiconductor Device Fundamentals, Addison-Wesley Publishing Company,1996.

[11] MEDICI: Two-dimensional semiconductor device simulation program, Synopsys, Inc. CA 94043

[12] Z. Liu, et.al.,"Threshold voltage model for deep-submicrometer MOSFET's", IEEE Trans. On Elec. Dev., Jan. 1993

[13] X. Zhou, et.al., "A general approach to compact threshold voltage formulation based on 2-D numerical simulation and experimental correlation for deep-submicron ULSI technology development" , IEEE Trans. On Elec. Dev., Jan. 2000