Impact of Gate Tunneling Current in Scaled MOS on Circuit Performance: A Simulation Study

Chang-Hoon Choi, Zhiping Yu, and Robert W. Dutton.
Center for Integrated Systems, Stanford University, Stanford, CA 94305, USA
phone: (650) 723-9484, fax: (650) 725-7731, e-mail: chchoi@stanford.edu

Abstract—**The influence of gate direct tunneling current on ultra-thin gate oxide MOS (1.1 nm $\leq t_{ox} \leq$ 1.5 nm, L_g = 50 – 70 nm) circuits has been studied based on detailed simulations. For the gate oxide thickness down to 1.1 nm, gate direct tunneling currents, including the edge direct tunneling, show only a minor impact on low V_{dd} static-logic circuits. However, some dynamic logic and analog circuits are more significantly influenced by the off-state leakage current for oxide thickness below 1.5 nm, under low voltage operation.**

I. Introduction

According to the International Technological Roadmap for Semiconductors (ITRS), gate oxide thicknesses of 1.2–1.5 nm will be required by 2004 for sub-100 nm CMOS [1]. In this thin gate oxide regime, direct tunneling current increases exponentially with decreasing oxide thickness [2][3], which is of primary concern for CMOS scaling.

For conventional CMOS devices, the dominant leakage mechanism is mainly due to short channel effects owing to drain induced barrier lowering (i.e. DIBL). In the ultra-thin gate oxide regime, however, the gate leakage current can contribute significantly to off-state leakage, which may result in faulty circuit operation since designers may assume that there is no appreciable gate current. A recent study has shown that direct tunneling current appearing between the Source-Drain Extension (SDE) and the gate overlap, so-called the Edge Direct Tunneling (EDT), dominates off-state drive current, especially in very short channel devices [4][5]. This results from the fact that the ratio of the gate overlap to the total channel length becomes large in the short channel device compared to that of the long channel device. Thus, the gate current effect is expected to become appreciable in ultra-thin oxide, sub-100 nm MOS circuits. Even though many researchers have discussed the effects of gate leakage current, scaling limitations due to gate tunneling current from the viewpoint of circuit operation have not been critically addressed. Assessment of circuit immunity against the gate tunneling current, depending on various device structures and bias conditions, is of great importance in determining directions for future gate oxide scaling.

II. Gate Current Modeling

A. Edge Direct Tunneling (EDT)

Gate direct tunneling current is produced by the quantum-mechanical wavefunction of a charged carrier through the gate oxide potential barrier into the gate, which depends not only on the device structure but also bias conditions. Fig. 1 illustrates various gate tunneling components in a scaled NMOSFET; the gate-to-channel current (I_{gc}) and the EDT currents (I_{gso} and I_{gdo}), respectively.

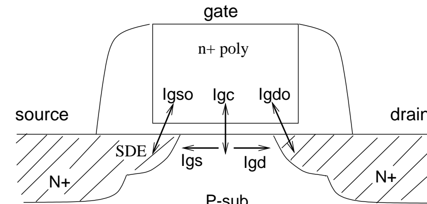


Fig. 1. Illustration of gate direct tunneling components of a very short-channel NMOSFET. I_{gso} and I_{gdo} are edge direct tunneling (EDT) currents.

In long channel devices, I_{gso} and I_{gdo} are less important than I_{gc} because the gate overlap length is small compared to the channel length. In very short channel devices, the portion of the gate overlap compared to the total gate length becomes a large fraction. For example, the typical gate overlap for a physical gate length of a 50 nm NMOSFET, estimated by the two-dimensional process simulation, is around 20 nm which corresponds to 40 % of the total. To reduce the direct tunneling current and Miller capacitance, a smaller overlap length is desirable. However, the overlap cannot be scaled easily in advanced MOS devices due to difficulties in controlling the doping profiles. Moreover, even if devices can be scaled successfully, too short an overlap region may cause unacceptably high external resistance in shallow source-drain junctions.

Fig. 2 illustrates the band diagrams and electron tunneling directions along the gate-to-channel and gate-to-SDE directions for a highly doped drain (HDD) NMOSFET. For $V_g > 0$ V, the gate-to-channel tunneling current (I_{gc}) is the dominant current component, since higher gate oxide voltage (V_{ox}) appears between the gate and the channel, as shown in Fig. 2(a). Namely, the V_{fb} of an NMOSFET with an n-type polysilicon gate (i.e. n^+-poly/SiO_2/p-substrate) is approximately -1 V, while the V_{fb} along the gate-to-SDE (i.e. n^+-poly/SiO_2/n^+ SDE) is approximately 0 V. On the contrary, assuming that the overlap length is comparable to the channel length, the EDT currents (I_{gso} and I_{gdo}) can become dominant for bias conditions of $V_{fb} < V_g < 0$ V. For the gate-to-SDE case, electrons accumulated in the n^+-poly gate tunnel to the SDE region can lead to an appreciable off-state current. Meanwhile, operating in the depletion mode along the n^+-poly/SiO_2/p-substrate surface, few electrons are present in the channel that could in turn tunnel into the gate, as shown in

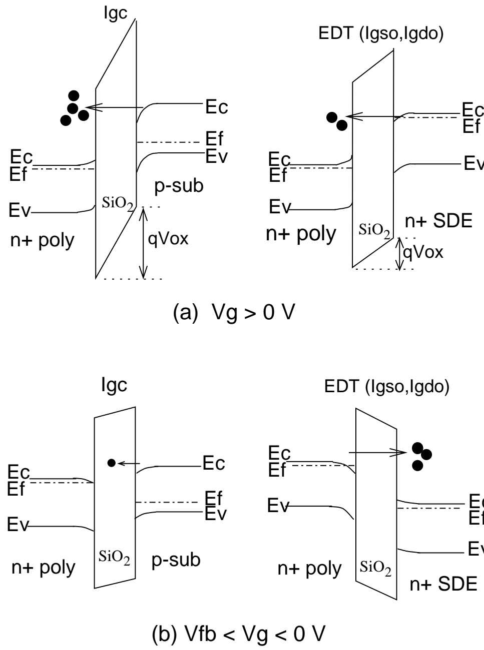


Fig. 2. Gate bias dependent band diagrams and electron tunneling in the channel (I_{gc}) and the gate edge (I_{gso} and I_{gdo}). (a) $V_g > 0$ V (inversion mode). (b) $V_{fb} < V_g < 0$ V (depletion mode).

Fig. 2(b).

B. Direct Tunneling Device Simulation

The edge direct tunneling in the gate-to-SDE region must be treated as a two-dimensional problem in very short channel devices [5], owing to the laterally finite doping gradient in the SDE region and the drain electric field effects. The compact direct tunneling current model for circuit simulation has shown good agreement with experimental data for long channel devices [6]; however, it cannot accurately represent the direct tunneling current for highly non-uniform SDE and channel regions in sub-100 nm MOSFETs. The exact gate direct tunneling should be modeled based on solution of the two-dimensional Schrödinger equation, coupled with the semiconductor transport equations.

In order to model the edge direct tunneling behavior, MEDICI [7] was used. The electron direct tunneling is calculated for two sources: conduction band electron tunneling (CBET) and valence band electron tunneling (VBET). CBET is the tunneling of electrons from the conduction band of the silicon substrate to the polysilicon, while VBET is the electron tunneling from the valence band of the silicon substrate due to generation of free holes.

To validate the direct tunneling model, gate currents were simulated and compared to the experimental data of long channel NMOSFETs (i.e. W/L = 100 μm/100 μm and t_{ox} = 1.3 and 1.5 nm). Though agreement is not perfect, simulated gate currents from MEDICI show reasonable correspondence to the measurements, as reflected in Fig. 3.

Device simulations were also performed for a very-short

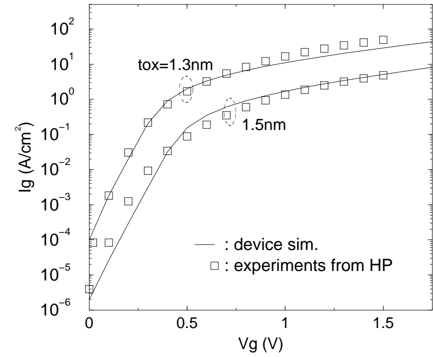


Fig. 3. Simulated gate currents from MEDICI and comparison with the measured data for a long channel (L_g = 100 μm) NMOSFET.

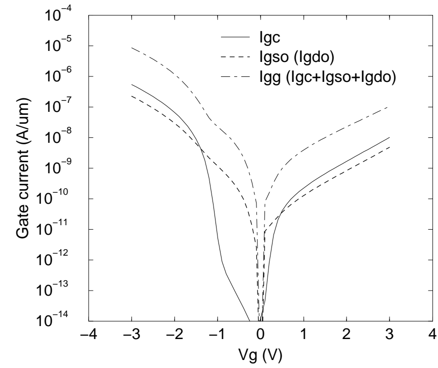


Fig. 4. Simulated I_{gc}, I_{gso} and I_{gg} (= I_{gc}+I_{gso}+I_{gdo}) for an NMOSFET with t_{ox} = 1.5 nm and L_g = 50 nm.

channel NMOSFET with 50 nm gate length; Fig. 4 illustrates resulting gate currents for an NMOSFET with t_{ox} = 1.5 nm. The source and drain are tied to ground and the gate bias is forced from negative to positive values. Note that the EDT current (I_{gso}, I_{gdo}) is higher than the gate-to-channel current (I_{gc}) for gate biases of -1.5 V $< V_g <$ 0 V, implying that the EDT is the dominant leakage source for the off-state current in the low voltage range of operation for MOS circuits. Fig. 5 shows the total simulated gate current ($I_{gg} = I_{gc} + I_{gso}$ + I_{gdo}) for different gate oxide thicknesses, ranging from 1.1 – 1.8 nm; gate current increases exponentially as the gate oxide thicknesses are scaled down.

In order to observe the transient behavior of thin oxide MOSFETs with significant gate tunneling current, mixed mode circuit-device simulation is performed for a single off-state transistor with a loading capacitor initially charged to 2.5 V. The discharge of the output node connected to the NMOS drain is determined by the off-state current (I_{OFF}), which is the sum of the direct tunneling leakage (I_{DT}) and the DIBL leakage (I_{DIBL}) currents, as illustrated in Fig. 6(a). The gate length and the oxide thickness are 70 nm and 1.3 nm, respectively. Even though the NMOS is in the off-state, it acts more like a resistor due to the off-state current of the transistor, therefore V_{out} decays over time, as shown in Fig. 6(b). With the direct tunneling model, V_{out} drops more sharply than the

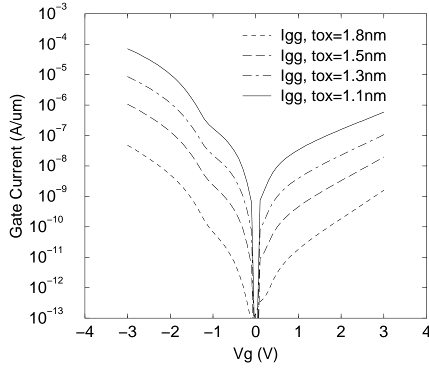


Fig. 5. Simulated I_{gg} (= I_{gc}+I_{gso}+I_{gdo}) for different t_{ox}'s ranging 1.1 – 1.8 nm and L_g = 50 nm.

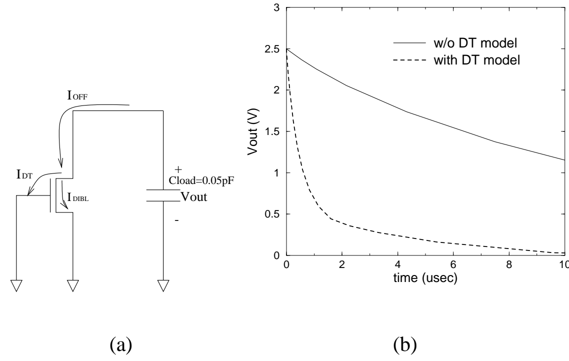


Fig. 6. Transient simulation by using MEDICI for a single off-state NMOS transistor with t_{ox} = 1.3 nm, L_g = 70 nm and W = 10μm. (a) discharge through a single off-transistor, V_{out} (t = 0) = 2.5 V. (b) comparison of V_{out}'s between with and without the direct tunneling model.

case without considering the direct tunneling model. Defining an equivalent time constant (τ) as, $\tau = V_{dd}C_{load}/I_{OFF} = V_{dd}C_{load}/(I_{DT} + I_{DIBL})$ provides a useful estimate. Over-estimating τ will be predictable without considering the I_{DT} term.

III. Circuit Application

In order to evaluate circuit performance by considering gate direct tunneling effects, a macro-circuit model has been constructed in the circuit simulator, HSPICE [9]. Gate direct tunneling currents, obtained from the device simulation for the gate oxide thicknesses of 1.1, 1.3 and 1.5 nm, are described using voltage-dependent current sources as a function of the terminal voltage, as shown in Fig. 7. The partitioning of I_{gc} into I_{gs} and I_{gd} is modeled by using variable resistances, R_{cs} and R_{cd}, respectively, in each part of the channel. R_{cs} and R_{cd} are the channel resistance corresponding to $0.5L_{ch}$, where $R_{cs} = V_{cs}/I_{cs}$ and $R_{cd} = V_{cd}/I_{cd}$. I_{cs} and I_{cd} are channel currents of each region; they have been obtained by adjusting the BSIM3-model parameters to fit the I-V curves generated from device simulation.

The macro-circuit model has been applied to several MOS

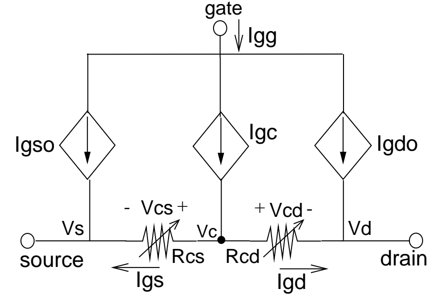


Fig. 7. A macro-circuit model for direct tunneling current combined circuit simulation.

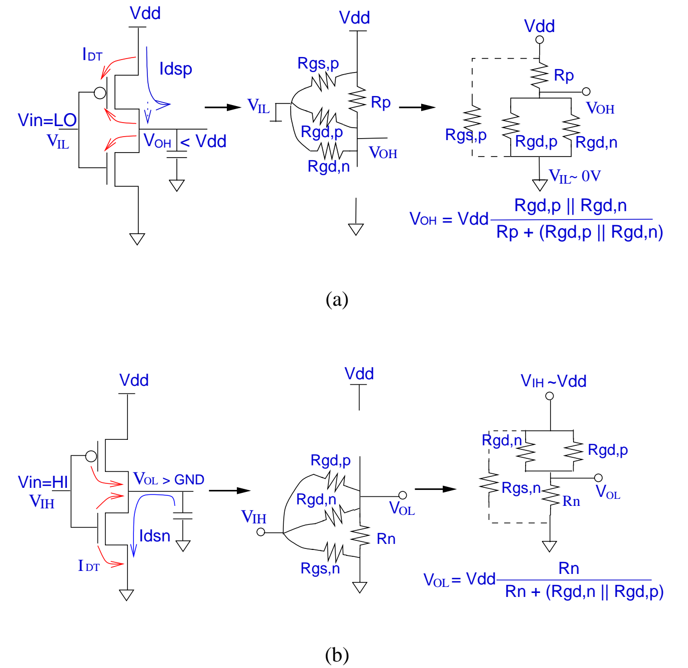


Fig. 8. V_{OH} and V_{OL} modeling considering gate tunneling currents of a CMOS inverter. (a) V_{in} = logic-low = 0 V. (b) V_{in} = high-high = V_{dd}.

circuits – CMOS inverter, dynamic AND gate, sample and hold (S/H) and bootstrapping circuits.

A. Static CMOS Inverter

For the CMOS inverter application we assumed the amount of hole direct tunneling of the PMOSFET is the same as that of the NMOSFET. The magnitude of the channel current is assumed to be identical, regardless of the gate oxide thickness, in order to focus on the circuit performance difference based on the oxide thickness dependent gate tunneling current contributions. Estimated gate current paths during the operation are shown schematically in Fig. 8.

When the input is *low* and the gate tunneling current is significant, V_{out} (i.e. V_{OH}) will not reach V_{dd} due to the leakage current that flows from the output node, as shown in Fig. 8(a). Here, direct tunneling current components can be modeled as

resistors and V_{OH} is approximated by the voltage divider:

$$V_{OH} \approx V_{dd} \times \frac{R_{gd,p}||R_{gd,n}}{R_p + (R_{gd,p}||R_{gd,n})} \quad (1)$$

where, R_p is the on-state channel resistance of the PMOS. $R_{gd,n}$ and $R_{gd,p}$ are gate-to-drain resistances of the N- and PMOSFET, respectively, modeling gate direct tunneling effects. The resistor values are approximately $R_{gd,p} \approx V_{dd}/I_{gd,p}$ and $R_{gd,n} \approx V_{dd}/I_{gd,n}$. As an example, if the ratio of $R_{gd,n}$ (or $R_{gd,p}$) to R_p is 100, then V_{OH} will drop by 2 % from the V_{dd} level.

As shown in Fig. 8(b), V_{OL} is again approximated using a voltage divider:

$$V_{OL} \approx V_{dd} \times \frac{R_n}{R_n + (R_{gd,n}||R_{gd,p})} \quad (2)$$

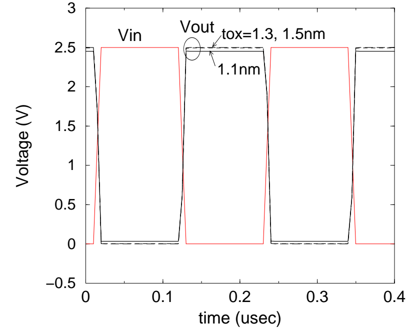
where, R_n is the on-state channel resistance of the NMOS.

As a result, when tunneling current is significant V_{OL} will not fall to the GND level and the V_{out} swing (GND $< V_{out} <$ V_{dd}) is reduced for very leaky, thin gate oxide CMOS inverters.

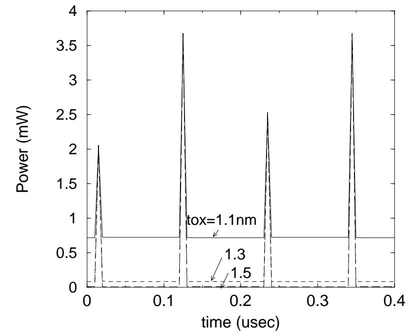
Again considering an example, assuming that V_{dd} = 2.5 V, $I_{ds,n}$ = 0.50 mA/μm, $I_{ds,p}$ = 0.25 mA/μm and $I_{gd,p} = I_{gd,n}$ = 5.0$\times$ 10^{-6} A/μm (i.e. $W_p = 2W_n$ = 20 μm), based on the simulations for t_{ox} = 1.1 nm and L_g = 50 nm, then $R_{gd,n}/R_n$ or $R_{gd,p}/R_p$ is about 100. In this case, the estimated V_{OH} and V_{OL} values are V_{OH} = 2.42 V and V_{OL} = 0.08 V, respectively, from Eqs. (1) and (2), or a total reduced logic swing of 160 mV.

Fig. 9 shows simulated input/output waveforms and power consumptions for a CMOS inverter using the macro-circuit model parameterized with t_{ox} = 1.1, 1.3, and 1.5 nm. The output capacitance is assumed the value when fanout is 4 (FO4). For V_{dd} = 2.5 V and t_{ox} = 1.1 nm, simulated V_{OH} and V_{OL} are 2.45 V and 0.04 V, respectively. The full logic-high (V_{dd}) and logic-low (GND) levels are achieved for t_{ox} = 1.3 and 1.5 nm. The DC power consumption of the CMOS inverter at V_{dd} = 2.5 V can be substantial for t_{ox} = 1.1 nm, as shown in Fig. 9(b); the average power consumption is 0.75, 0.16, and 0.09 mW for t_{ox} = 1.1, 1.3, and 1.5 nm, respectively. The power consumption of 0.75 mW for t_{ox} = 1.1 nm, is about 10 times larger than the case when the gate tunneling current is negligible – 0.08 mW baseline.

For V_{dd} = 1.5 V, the output node of the inverter swings between full logic-high and logic-low (i.e. only 0.1 % of V_{OH} reduction even for t_{ox} = 1.1 nm). Since the channel current becomes much higher than the gate tunneling current, gate tunneling current effects can be minimized when using lower voltage static-logic circuits. The power consumption for V_{dd} = 1.5 V is exponentially reduced compared to the case for V_{dd} = 2.5 V, due to the exponential decrease in gate current; basically a reduction of 20 $\sim$ 100 times is realized, compared to the V_{dd} = 2.5 V case. According to the ITRS roadmap, V_{dd} of



(a)



(b)

Fig. 9. CMOS inverter simulation results including gate direct tunneling for three gate oxide thicknesses, V_{dd} = 2.5 V (C_{out} = 0.08 pF (FO4) and L_n = L_p = 50 nm). (a) input and output waveforms. (b) power consumptions.

1.0 – 1.5 V is required for 70 nm CMOS technology. For the low V_{dd}, gate direct tunneling current effects on static-logic circuits will be less serious for oxide thicknesses down to 1.1 nm.

B. *Dynamic AND Gate*

Consider the domino CMOS AND-2 gate shown in Fig. 10. The circuit operation relies on first *precharging* the output node capacitance and subsequently, *evaluating* the output level according to the applied inputs, V_A and V_B. These operations are scheduled by a single clock signal, CK, which drives one NMOS and one PMOS transistor in each dynamic stage. Assume all inputs are low initially and the intermediate node voltage across C_2 has an initial value of 0 V. During the precharge phase, the output node capacitance (C_1) is charged up to its logic-high level of V_{dd} through the PMOS transistor. In the next phase, CK switches logic-high and the evaluation begins. If the input V_A switches from low to high during the evaluation phase, charge stored on C_1 will be shared with C_2, and the node voltage V_X drops after the charge-sharing. If V_X is less than $V_{dd}/2$, the output of the inverter V_{out} erroneously switches to logic-high during the evaluation phase.

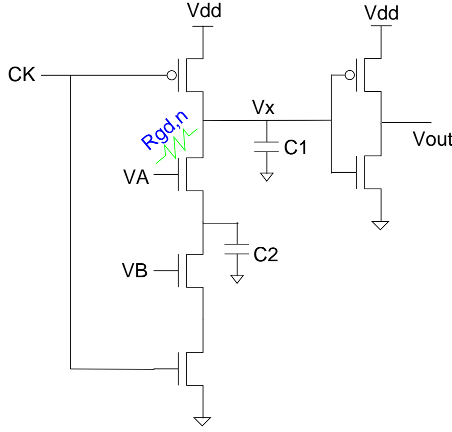
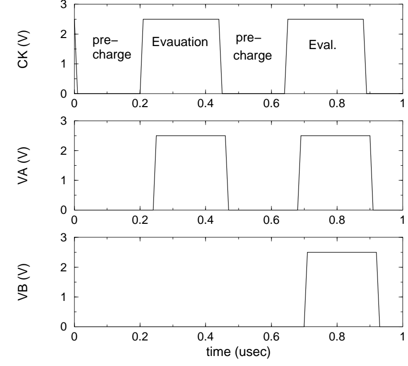
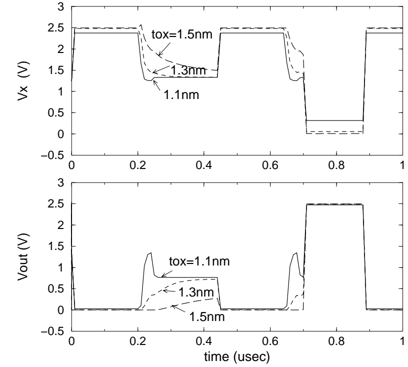


Fig. 10. Domino CMOS AND-2 gate.



(a)



(b)

Fig. 11. Simulated waveforms of the domino AND-2 gate. (a) clock and input signals. (b) output waveforms for V_{dd} = 2.5 V.

When gate tunneling current becomes significant, V_X may be less than the logic-high level during the precharge phase, so that V_X can even drop to less than $V_{dd}/2$ during the first evaluation period. As a result, V_{out} will inadvertently switch to a logic-high, resulting in a logic error.

Fig. 11 shows the simulated input and output waveforms of a domino AND-2 gate for V_{dd} = 2.5. When t_{ox} = 1.1 nm and V_{dd} = 2.5 V, V_X drops to about 1.2 V during the evaluation phase due to gate tunneling current effects. As a result, V_{out} erroneously switches during the first evaluation period and a glitch appears prior to the second evaluation phase.

These phenomena can be simply modeled as RC circuits, as shown in Fig. 12. Initially, V_X has a value of V_{dd} during the precharge phase. When the evaluation begins at t = 0.2 μs by switching of the CK signal to logic-high (i.e. V_A remains logic-low until t = 0.23 μs), charge stored in C_1 flows to ground through the tunneling resistance ($R_{gd,n}$), as illustrated in Fig. 12(a). During the discharge process the V_X level drops as a function of time according to the following:

$$V_X(t) = V_X(t_0)\, e^{-\frac{t-t_0}{R_{gd,n}C_1}} \approx V_{dd}\, e^{-\frac{I_{gd,n}\ (t-t_0)}{V_{dd}C_1}} \qquad (3)$$

where, t_0 = 0.20 μs in the case.

With a tunneling current of $I_{gd,n} = 4\times10^{-6}$ A, V_X drops to 1.60 V from its initial voltage of V_{dd} = 2.5 V during 0.03 μs (t = 0.20 – 0.23 μs), according to Eq. (3).

When V_A switches to logic-high at t = 0.23 μs, direct tunneling current is reduced due to a smaller voltage difference between V_X and V_A. Thus, instead of the discharge process, charge sharing begins; charge in C_1 is shared with C_2, as shown in Fig. 12(b). The final V_X after the charge sharing is approximated as follows:

$$V_X \approx \frac{V_X(t = 0.23\mu s)}{1 + C_2/C_1} \qquad (4)$$

When $V_X\ (t = 0.23\mu s)$ = 1.60 V, C_1 = 0.15 pF and C_2 = 0.04 pF, V_X after charge sharing estimated by Eq. (3) is about 1.26 V, which corresponds to $\sim V_{dd}/2$ level. Hence, V_{out} may erroneously switch during the evaluation phases, as shown in the simulation results of Fig. 11(b). Even though these spurious results can be reduced by lowering V_{dd}, the dynamic logic circuit may have potential problems due to gate tunneling-induced off-state current during the precharge-and-evaluation phases of operation.

C. Sample and Hold Circuit

The sample and hold (S/H) circuit is an important analog building block in data-converter systems used to acquire analog signals and to store the value for some length of time. A simple S/H circuit is formed by a sampling CMOS switch followed by a hold capacitor, as shown in Fig. 13(a). When the clock (*Phi*) is high, V_{out} follows V_{in}; when *Phi* goes low, V_{out} will ideally remain at a constant level. However, V_{out} will not hold this sampled value if leakage paths exist. This tunneling current-induced decay in V_{out} during the hold period can be modeled using the RC circuit shown in Fig. 13(a). As for the previous dynamic AND gate, V_{out} decays as a function of time, again using the expression given in Eq. (3).

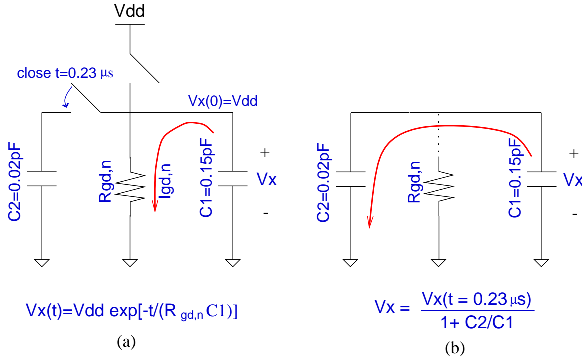


Fig. 12. Modeling of discharge and charge sharing behaviors during first evaluation period of domino AND-2 gate. (a) discharge through tunneling resistance before V_A switches to logic-high (t = 0.20–0.23 μs). (b) charge sharing of C_1 with C_3 after the V_A switches to logic-high ($t >$ 0.23 μs).

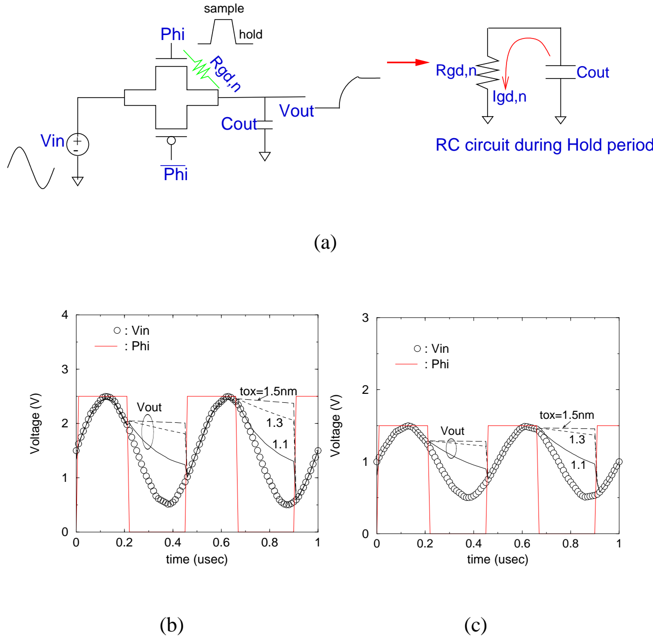


Fig. 13. CMOS sample and hold circuit and simulated waveforms for different t_{ox}'s. (a) CMOS S/H circuit schematic and RC circuit model during hold period. (b) waveforms for V_{dd} = 2.5 V. (c) waveforms for V_{dd} = 1.5 V.

Fig. 13(b) and (c) show simulation results of a S/H switch for three gate oxide thicknesses. During the holding period, the output node does not maintain the sampled value due to gate leakage current, and degradation becomes increasingly severe as the oxide thickness is scaled down. These gate tunneling effects appear even at low V_{dd}, as shown in Fig. 13(c), implying that the S/H circuit has poor robustness in the face of gate leakage current, limiting its operation to oxide scaling only to the 1.5 nm regime.

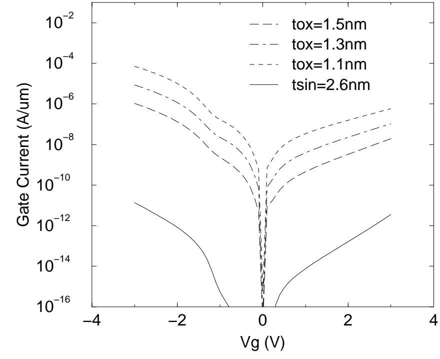


Fig. 14. Simulated gate tunneling current of L_g = 50 nm NMOSFET with an alternative gate dielectric of Si_3N_4, $\kappa = 2\epsilon_{ox}$ and thickness is 2.6 nm.

D. Impact of Alternative Gate Dielectrics

Alternate insulating materials with a dielectric constant larger than that of SiO_2 are under evaluation to replace the conventional SiO_2 gate stack, thereby overcoming the tunneling problems cited above; remote-plasma nitrided oxide (RPNO) [10], JVD-Si_3N_4 and nitride/oxide (N/O) [11] are all possible replacement materials. Using such gate dielectrics, devices with lower gate-leakage current can be achieved as a result of the increased film thickness resulting from the increased dielectric constant of nitride-based layers ($\epsilon_{nitride} \sim$ 7.8).

Fig. 14 illustrates simulated gate tunneling current for an L_g = 50 nm NMOSFET with alternative gate dielectric of Si_3N_4 shown as a solid line, assuming a dielectric constant of $\kappa = 2\epsilon_{ox}$ and thickness of 2.6 nm ($t_{ox,eq}$ = 1.3 nm). Much lower gate current is produced by using the Si_3N_4 device, compared to the pure oxide device with the same equivalent oxide thickness of t_{ox} = 1.3 nm. Secondary effects such as surface roughness and interface traps are not considered.

Voltage bootstrapping is used to overcome threshold voltage drops in digital circuits. Fig. 15(a) shows a schematic of the bootstrapping circuit, including the bootstrap MOS capacitor; the voltage V_X is increased during the V_{in} switching event. As a result, the threshold voltage drop can be compensated for the output node, V_{out}.

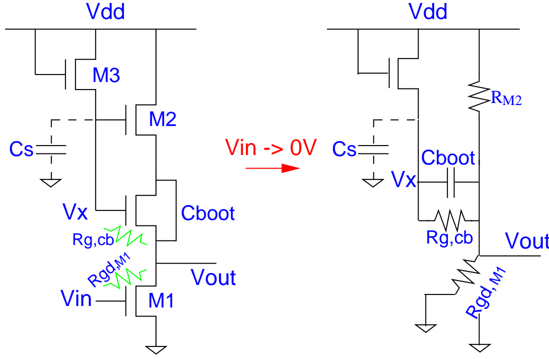
When V_{in} switches to logic-low, V_{out} and V_X are approximated as follows [12]:

$$V_{out} \approx V_X - V_{th,M2} \tag{5}$$

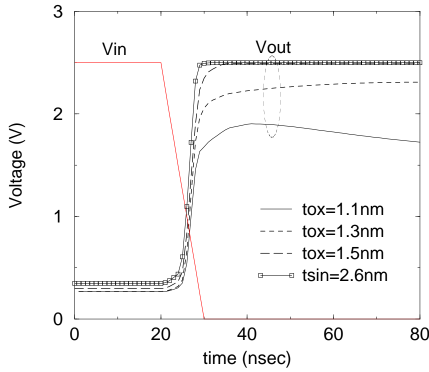
where

$$V_X \approx (V_{dd} - V_{th,M3}) + V_{dd}\frac{C_{boot}}{C_s + C_{boot}} \tag{6}$$

(V_{dd}-$V_{th,M3}$) is the initial condition of V_X and the second term of Eq. (6) represents the increase in V_X after the V_{in} switches to 0 V. However, this V_X expression should be modified to account for the gate tunneling current. First, the initial V_X is reduced due to the discharge via the gate leakage resistor (i.e. $R_{g,cboot}$), such that



(a)



(b)

Fig. 15. Voltage bootstrapping circuit and simulation results. (a) circuit schematic and its equivalent circuit when V_{in} switches to 0 V. (b) simulated waveforms with an alternative gate dielectric (Si_3N_4) and pure oxides.

$$(V_{dd} - V_{th,M3}) \longrightarrow (V_{dd} - V_{th,M3})\, e^{-\frac{t}{R_{g,cboot} C_{boot}}} \qquad (7)$$

In addition, the second term of Eq. (6) is modified due to the gate-to-drain resistance effects of M1 such that,

$$V_{dd} \frac{C_{boot}}{C_s + C_{boot}} \longrightarrow \left(V_{dd} \frac{R_{gd,M1}}{R_{M2} + R_{gd,M1}} \right) \cdot \frac{C_{boot}}{C_s + C_{boot}} \qquad (8)$$

Thus, both V_X and V_{out} will be reduced in the presence of substantial gate tunneling current.

Fig. 15(b) illustrates the simulated input and output waveforms of the bootstrapping circuit. For t_{ox} = 1.1 and 1.3 nm, the final V_{out} does not reach V_{dd} because of the leakage in both the MOS capacitor (C_{boot}) and the driving transistor (M1). In contrast, a full V_{dd} output voltage can be achieved by adopting a Si_3N_4 gate dielectric due to the substantial reduction in leakage current. Note that the equivalent oxide thickness of $t_{Si_3N_4}$ = 2.6 nm is t_{ox} = 1.3 nm. Alternative gate dielectrics will be necessary to replace leaky gate oxides in MOS circuits, especially where charge conservation or charge bootstrapping techniques are required.

IV. Conclusions

CMOS circuit robustness in the presence of gate tunneling currents has been studied using circuit simulation, combined with a macro-circuit model of gate tunneling current and analytic estimation of the effects. CMOS static inverters at V_{dd} = 1.5 V show acceptable noise margins with low power consumption for the oxide thicknesses down to 1.1 nm, while dynamic AND gates have a potential weakness in the presence of gate current during the precharge and evaluation phases. For circuits that require charge-conservation or charge-bootstrapping, including the S/H circuit, significant performance degradation can be expected for $t_{ox} < 1.5$ nm, even considering low voltage operation. A dual-gate oxide process or use of high-κ dielectric will be necessary on these circuits to continue device scaling.

Acknowledgement

The authors wish to thank the National Science Foundation (NSF) for support of this work though the Distributed Center of Advanced Electronics Simulations (DesCArtES).

References

[1] "International Technology Roadmap for Semiconductors 2000 Edition," *http://public.itrs.net/, Semiconductor Industry Assoc.*, Austin, TX, 2000.
[2] H.S. Momose, M. Ono, T. Yoshitomo, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, "1.5 nm Direct-Tunneling Gate Oxide Si MOSFET's," *IEEE Trans. Electron Devices*, vol. 43, pp. 1233–1242, Aug. 1996.
[3] H. Iwai and H.S. Momose, "Ultra-Thin Gate Oxides – Performance and Reliability," in *IEDM Tech. Dig.*, pp. 163–166, 1998.
[4] K.N. Yang, H.T. Huang, M.J. Chen, Y.M. Lin, M.C. Yu, S.M. Jang, C.H. Yu, and M.S. Liang, "Edge Hole Direct Tunneling in Off-State Ultrathin Gate Oxide p-Channel MOSFETs," in *IEDM Tech. Dig.*, pp. 679–682, 2000.
[5] N. Yang, W.K. Henson, and J. Wortman, "A Comparative Study of Gate Direct Tunneling and Drain Leakage Currents in N-MOSFET's with Sub-2 nm Gate Oxides," *IEEE Trans. Electron Devices*, vol. 47, no. 8, Aug. 2000.
[6] C-H. Choi, K-H. Oh, J-S. Goo, Z. Yu, and R.W. Dutton, "Direct Tunneling Current Model for Circuit Simulation," in *IEDM Tech. Dig.*, pp. 735–738, 1999.
[7] *MEDICI : Two-Dimensional Semiconductor Device Simulation.* AVANT! Corp., 2000.
[8] A. Shanware, J.P. Shiely, and H.Z. Massoud, "Extraction of the Gate Oxide Thickness of N- and P-Channel MOSFETs Below 20Å from Substrate Current Resulting from Valence-Band Electron Tunneling," in *IEDM Tech. Dig.*, pp. 815–818, 1999.
[9] *HSPICE User's Manual,* AVANT! Corp., 1999.
[10] M. Rodder, S. Hattangady, N. Yu, W. Shiau, P. Nicollian, T. Laaksonen, M. Mehrotra, C. Lee, and S. Aur, "A 1.2V, 0.1μm Gate Length CMOS Technology : Design and Process Issues," in *IEDM Tech. Dig.*, pp. 623–626, 1998.
[11] Y. Wu, Y-M. Lee, and G. Lucovsky, "1.6 nm Oxide Equivalent Gate Dielectrics Using Nitride/Oxide (N/O) Composites Prepared by RPECVD/Oxidation Process," *IEEE Electron Device Lett.*, vol. 21, pp. 116–118, Mar. 2000.
[12] S.M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits.* The McGraw-Hill Companies, Inc., 1996.