

Analysis of Leakage Currents and Impact on Off-State Power Consumption for CMOS Technology in the 100-nm Regime

W. Kirklen Henson, Nian Yang, *Member, IEEE*, Stefan Kubicek, Eric M. Vogel, Jimmie J. Wortman, *Senior Member, IEEE*, Kristen De Meyer, *Member, IEEE*, and Abdalla Naem

Abstract—Off-state leakage currents have been investigated for sub-100 nm CMOS technology. The two leakage mechanisms investigated in this work include conventional off-state leakage due to short channel effects and gate leakage through ultrathin gate oxides. The conventional off-state leakage due to short channel effects exhibited the similar characteristics as previously published; however, gate leakage introduces two significant consequences with respect to off-state power consumption: 1) an increase in the number of transistors contributing to the total off-state power consumption of the chip and 2) an increase in the conventional off-state current due to gate leakage near the drain region of the device. Using experimentally measured data, it is estimated that gate leakage does not exceed the off-state specifications of the National Technology Roadmap for Semiconductors for gate oxides as thin as 1.4 to 1.5 nm for high performance CMOS. Low power and memory applications may be limited to an oxide thickness of 1.8 to 2.0 nm in order to minimize the off-state power consumption and maintain an acceptable level of charge retention. The analysis in this work suggests that reliability will probably limit silicon oxide scaling for high performance applications whereas gate leakage will limit gate oxide scaling for low power and memory applications.

Index Terms—MOS devices, off-state leakage, power consumption, ultrathin gate oxide.

I. INTRODUCTION

ONE of the most challenging CMOS device design criteria is maintaining a required off-state leakage specification while maximizing the device drive current. For conventional CMOS devices, the dominant leakage mechanism is due to short channel effects (SCE) where the drain potential lowers the source junction barrier to minority carriers. This barrier lowering is typically referred to as drain induced barrier lowering (DIBL) and has been investigated extensively [1]–[5]. As CMOS gate lengths approach the 100 nm regime other

leakage mechanisms will contribute to the off-state power consumption. The physical gate oxide thickness for the 100-nm technology regime is projected to be less than 2.5 nm [6], which introduces a nonnegligible gate current due to direct tunneling [7]–[16]. The direct tunneling current increases exponentially with decreasing oxide thickness, which introduces a potential limitation in aggressively scaling the gate oxide for advanced CMOS technology. Although several reports have demonstrated the advantages of using direct tunnel oxides [7]–[9] and the corresponding consequences on power consumption [10], the scaling limitations due to off-state leakage for advanced CMOS devices have not been critically addressed.

The analysis presented in this paper combines experimental data from two different device experiments to address leakage current limitations of CMOS scaling. One set of experimental data was collected on devices fabricated at North Carolina State University (NCSU) and a second set was collected on devices fabricated at the Interuniversity Micro-Electronics Center (IMEC). The NCSU device experiment consists of NMOS devices with gate oxides ranging in thickness from 1.4 to 2.0 nm; however the minimum gate length was only 0.32 μm . The IMEC device experiment consists of CMOS devices with various halo implant doses, source/drain extension doses, thermal budgets and has gate lengths down to 100 nm; however, the gate oxide thickness is fixed at 2.5 nm. Using the experimental data from both device lots, the impact of gate leakage and conventional channel leakage on off-state power consumption can be analyzed.

The device fabrication details are described in Section II. The gate leakage current and channel leakage current are individually analyzed in Section III and then used in Section IV to estimate the off-state leakage levels for sub-100 nm CMOS technology. The results are summarized in Section V.

II. EXPERIMENT

The NCSU devices were fabricated on p-type wafers with a nominal uniform doping of $5 \times 10^{17} \text{ cm}^{-3}$. *N*-channel transistors were fabricated with a four mask process (active, poly, contact, and metal). The pre-gate oxidation clean consisted of a standard RCA clean followed by a 25-nm sacrificial oxide grown via wet oxidation. The sacrificial oxide was removed one wafer at a time with a 2% HF solution then immediately loaded into the rapid thermal processor (RTP). The gate oxides were grown by rapid thermal oxidation (RTO) at 800 °C where

Manuscript received August 23, 1999. This work was supported in part by the National Science Foundation through the Advanced Electronic Materials Processing Research Center and by the Semiconductor Research Corporation through Contract BP-132. The review of this paper was arranged by Editor J. M. Vasi.

W. K. Henson, S. Kubicek, and K. De Meyer are with the Interuniversity Micro-Electronics Center, 3001 Leuven, Belgium (e-mail henson@imec.be).

N. Yang is with Advanced Micro Devices, Sunnyvale, CA 94088 USA. E. M. Vogel is with the National Institute of Standards and Technology, Semiconductor Electronics Division, Gaithersburg, MD 20899 USA.

J. Wortman is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695 USA.

A. Naem is with National Semiconductor Corporation, Santa Clara, CA 95052 USA.

Publisher Item Identifier S 0018-9383(00)05198-4.

oxidation pressure and cycle time were used to control thickness. A 150-nm *in-situ* phosphorous doped polysilicon gate was deposited before removing the wafer from the RTP. The gate was patterned by reactive ion etching, a single arsenic implant formed the source/drain junctions and titanium/aluminum was evaporated for front side metal.

The IMEC devices were fabricated in a full CMOS process with a 2.5-nm NO based furnace gate oxide. The nitrogen concentration incorporated into the oxide is estimated to be 7 to 8 atomic % from SIMS analysis. Further analysis of the nitrogen concentration and the impact on electrical characteristics can be found in [17]. Advanced 248 nm deep-UV lithography with alternating phase shift mask technique was used for the gate patterning. Optimization of the lithography process resulted in low across the chip linewidth variations for the 130-nm isolated lines and the smallest poly lines printed were 100 nm. As⁺ and B⁺ implants were employed to form a retrograde channel profile prior to gate stack formation while a second large tilt angle or halo implant with As⁺ and BF₂⁺ was employed after gate stack lithography. Low energy BF₂⁺ and As⁺ implants formed the shallow extensions followed by nitride spacer formation and deep junction formation with B⁺ and As⁺ implants. An advanced Co/Ti silicidation scheme was used and the transistor fabrication was completed conventionally up to level 1 metallization. The electrical characteristics of the NCSU devices were measured with an HP4145B parameter analyzer and an HP4284A LCR meter. The IMEC devices were measured with an HP4156A parameter analyzer and an HP4284A LCR meter.

III. ANALYSIS OF DIBL AND GATE LEAKAGE

A. Definition of Leakage Components with Device Bias

Fig. 1(a) illustrates the various leakage mechanisms for CMOS technology in the 100-nm gate length regime for an *n*-channel device biased in the off-state ($V_{\text{gate}} = 0.0$, $V_{\text{drain}} = V_{\text{DD}}$). The three possibilities for leakage are the channel current (I_{channel}), gate oxide tunneling current (I_{gate}) and the junction tunneling current (I_{junction}). The channel current represents the conventional DIBL leakage mechanism and is a function of channel doping, temperature, applied bias and junction architecture [1]–[5]. The gate current represents leakage through the oxide and is exponentially dependent on oxide thickness and applied bias [11]–[16]. The junction current represents the band to band tunneling or leakage mechanisms due to defects [18], [19].

Fig. 1(b) illustrates a second bias configuration that becomes significant for devices with direct tunnel gate oxides. Consider for example a device with $V_{\text{gate}} = V_{\text{DD}}$ and $V_{\text{drain}} = 0$ (e.g. NMOS of an inverter). In this case the gate leakage occurs uniformly across the device area which represents an additional leakage mechanism that is not present in more conventional CMOS technology with thicker gate oxides ($t_{\text{ox}} > 3.0$ nm). The leakage in Fig. 1(b) must also be considered when analyzing the off-state power consumption for devices with direct tunnel oxides.

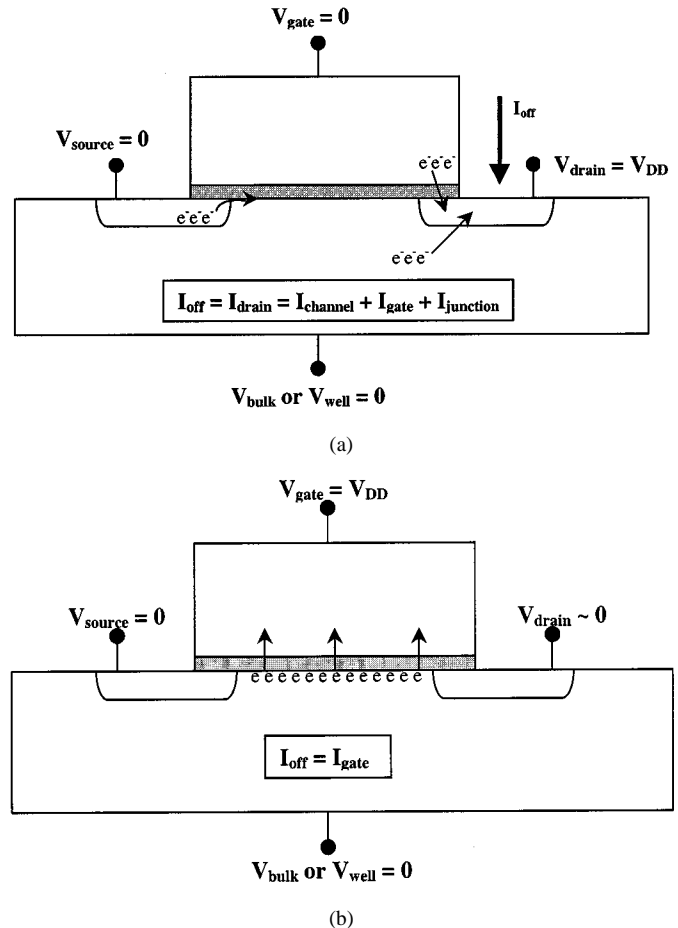


Fig. 1. Definition of leakage components for advanced CMOS devices. (a) Applied bias to drain and all other terminals grounded. Note that the arrows indicate the direction of electron flow in an NMOS device. (b) Gate leakage component for a device with a direct tunnel gate oxide. Applied bias to gate and all other terminals grounded. Note that the arrows indicate the direction of electron flow in an NMOS device.

B. Gate Current Leakage Characteristics

Fig. 2 illustrates the typical gate leakage characteristics versus gate bias with several different oxide thicknesses. The data in Fig. 2 were measured on *n*-channel MOSFET's with the source, drain and substrate connected to ground. The values of oxide thickness are estimates of the physical thickness using conventional capacitance-voltage measurements and modeling to account for the quantum confinement and poly-depletion effects [20], [21]. The polysilicon depletion is accounted for by modeling the polysilicon gate as a uniformly doped semiconductor with doping N_{poly} . The polysilicon doping is extracted from capacitance-voltage (*C*-*V*) measurements on transistors with gate oxides thicker than 2.5 nm to avoid attenuation of inversion capacitance due to leakage [22]. The quantum mechanical confinement effects of the semiconductor charge are modeled in a similar fashion as described in [23]. A nonlinear least squares technique is used to fit the theoretical model to the measured data where the oxide thickness, substrate doping and flatband voltage are used as fitting parameters. More details of the modeling are described in [20] while the impact of leakage current and comparison with high resolution

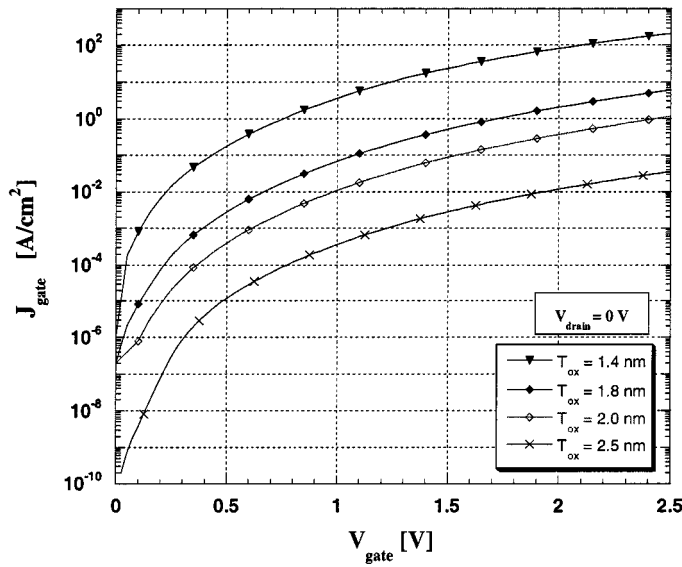


Fig. 2. Gate current density versus gate bias measured on NMOS transistors with the drain, source and substrate grounded. Measurements for 1.4 nm, 1.8 nm and 2.0 nm taken from NCSU devices and 2.5 nm taken from IMEC devices.

tunneling electron microscopy can be found in [21]. The data in Fig. 2 are in good agreement with recent results published in [16] but exhibit lower leakage at a given oxide thickness and bias than in other published results [14]. We believe this discrepancy is due to the modeling and not necessarily a real improvement in oxide quality.

To illustrate how dramatically the gate current increases power consumption, the data in Fig. 2 have been converted to a power density [W/cm^2] and plotted versus oxide thickness in Fig. 3. The inverse slopes of the data in Fig. 3 are -0.291 nm/decade, -0.288 nm/decade and -0.284 nm/decade for a gate bias of 1.3 V, 1.0 V and 0.7 V, respectively. The values of these slopes are not encouraging since a 0.3 nm reduction in oxide thickness increases the power density approximately one order of magnitude.

The off-state bias configuration of $V_{\text{gate}} = 0.0$ V and $V_{\text{drain}} = V_{\text{DD}}$ [Fig. 1(a)] also exhibits a nonnegligible gate current but not as significant as the case when $V_{\text{gate}} = V_{\text{DD}}$ and $V_{\text{drain}} = 0$. This is intuitively expected since the tunneling current is localized in the drain region for the off-state condition. Fig. 4 shows the gate current per unit width versus drain bias characteristics with $V_{\text{gate}} = 0$ for different oxide thicknesses and channel lengths. The detailed physics of describing the gate leakage current due to drain bias requires two-dimensional analysis of tunneling which is beyond the scope of this work, however, several features can be observed.

The first obvious trend to note is the gate current exhibits a strong exponential dependence with oxide thickness as expected from a tunneling process. A second trend to note is the gate current weakly depends on channel length suggesting that the tunneling may be concentrated at the gate to drain overlap region. Note that the magnitude of the gate current for the 1.8 nm gate oxide is less than 1 nA/ μm for drain biases up to 2 V. However, the magnitude of the gate current for the 1.4-nm

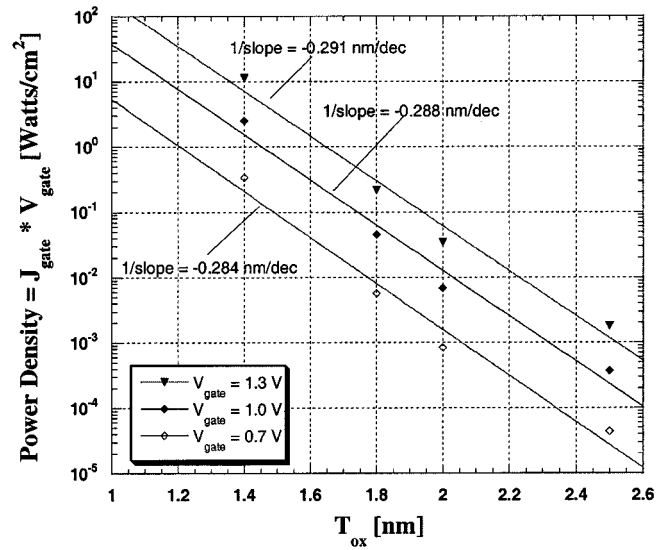


Fig. 3. Gate current density data converted to power density and plotted versus oxide thickness. Measurements for 1.4 nm, 1.8 nm and 2.0 nm taken from NCSU devices and 2.5 nm taken from IMEC devices.

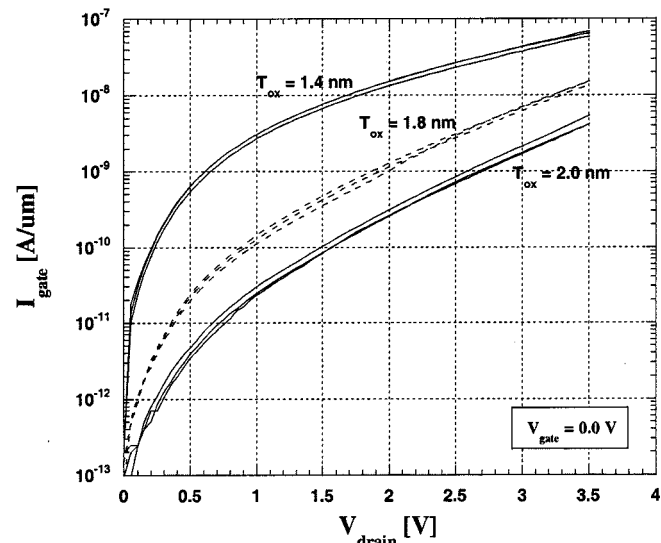


Fig. 4. Gate leakage current as a function of drain bias for three different oxide thicknesses. For each oxide thickness the gate leakage is measured for a 4.82 μm , 0.82 μm and 0.32 μm channel length. No trends are observed with channel length.

gate oxide is in the range of 1 to 10 nA/ μm for applied drain biases between 0.6–1.7 V. Therefore, as the oxide thickness scales below 1.8 nm, a significant fraction of the conventional off-state leakage current will be due to gate leakage near the drain portion of the device. This places more stringent requirements on channel engineering to reduce the leakage due to short channel effects in order to compensate for the extra leakage from the gate current.

C. DIBL Leakage Current Characteristics

The off-state leakage current that flows between the source and drain can be attributed to surface potential through and bulk

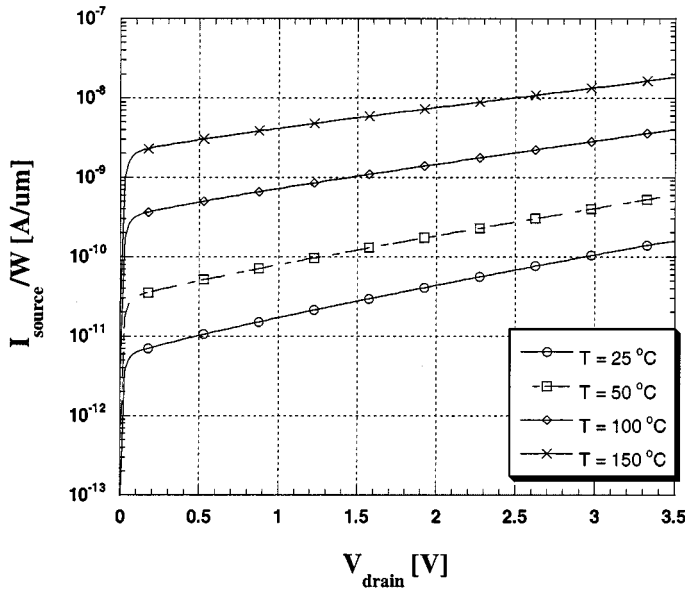


Fig. 5. Channel leakage current versus drain bias for an NMOS device with nominal L_{poly} of 130 nm and different temperatures.

punchthrough. Bulk punchthrough leakage exhibits characteristics that are weakly dependent on temperature whereas the surface punchthrough, or DIBL leakage, is strongly dependent on temperature [3]–[5]. Leakage measurements were performed as a function of temperature to determine the dominant leakage mechanism. Fig. 5 illustrates the measured source current versus drain bias for an NMOS transistor with nominal L_{poly} of 130 nm and temperature ranging from 25 °C to 150 °C. The drastic increase in source current with temperature demonstrates that surface leakage is the dominant mechanism and can be described by the following equation [3], [5]:

$$I_{source/drain} = I_o \exp\left(\frac{qm|V_{ds}|}{kT}\right) \\ I_o \propto \exp\left(\frac{-E_g + q\phi_s}{kT}\right). \quad (1)$$

The parameter m is a dimensionless constant that represents the DIBL effect and depends on device architecture. q represents the electronic charge, V_{ds} represents the applied drain-source voltage, k represents Boltzmann's constant, T represents temperature, E_g is the bandgap, and ϕ_s represents the potential at the surface [5].

To analyze the DIBL leakage as a function of device parameters, the leakage data are modeled by (1) for three halo implant conditions ($1.8 \times 10^{13} \text{ cm}^{-2}$, $2.4 \times 10^{13} \text{ cm}^{-2}$, $3.0 \times 10^{13} \text{ cm}^{-2}$), four gate lengths (250 nm, 150 nm, 130 nm, and 100 nm) and four temperatures (25 °C, 50 °C, 100 °C, and 150 °C). Note that the mask length (L_{mask}) is used instead of the effective channel length (L_{eff}). The reason for using L_{mask} is due to the inability of extracting L_{eff} from devices with a strong reverse short channel effect (RSCE) [24]. Fig. 6 illustrates the RSCE with the linear threshold voltage versus L_{mask} for three different halo implant conditions. The increase in threshold voltage is due

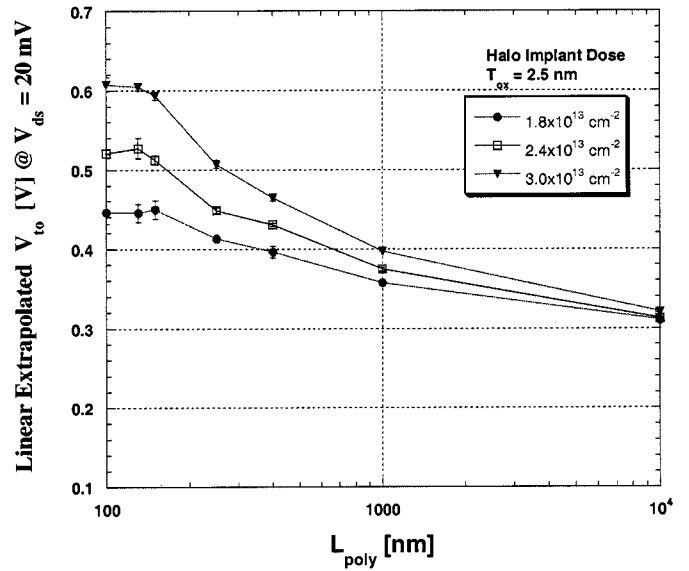


Fig. 6. Threshold voltage versus gate length for three different halo implant doses with all other fabrication steps identical. Error bars represent the maximum and minimum extracted values from four measurements.

to an increase in channel doping when the halo implants begin to overlap. The increasing doping means the effective mobility is also a function of channel length and currently there is not a proven technique that can self-consistently extract L_{eff} for devices with different effective mobility. The second detail to note is only nMOS devices are shown in this analysis. The reason for using nMOS devices is that the L_{eff} is much closer to L_{mask} compared to pMOS devices due to less dopant diffusion from the junctions.

The results are summarized in Fig. 7(a) and (b) with the extracted I_o and m parameter, respectively. The I_o parameter exhibits an exponential dependence on $1/kT$ which verifies the functional relationship of (1). Increasing the channel doping via the halo implant dose substantially reduces the DIBL leakage as seen in Fig. 7(a); however, as a consequence the threshold voltage is drastically increased which leads to a lower current drive. The m parameter increases significantly with decreasing gate length due to the stronger influence of the drain potential on the source junction. The higher channel doping only lowers the m parameter marginally as seen in Fig. 7(b).

IV. DISCUSSION

As CMOS technology approaches the 100-nm technology node an additional component of power consumption is introduced due to the gate leakage current. As described in Section III-B, the magnitude of the gate leakage current is a strong function of the applied bias [e.g., leakage described in Fig. 1(a) versus (b)]. A precise analysis of power consumption is complicated due to the different tunneling properties of electrons and holes (e.g., NMOS devices versus PMOS devices). Also, the gate leakage contributing to the conventional off-state leakage is expected to be dependent on processing and device architecture (poly oxidation, junction-gate overlap, etc.). A general expres-

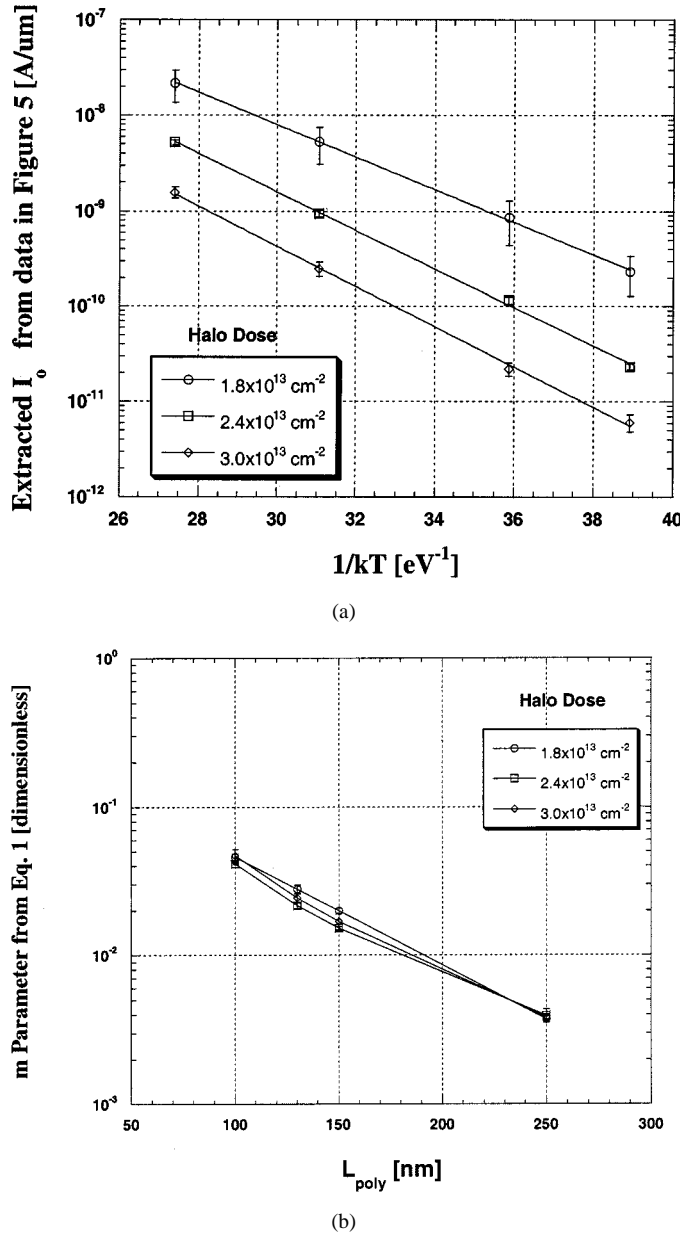


Fig. 7. Extracted I_o and m parameters from fitting equation (1) to experimental data in Fig. 5. (a) Extracted I_o parameter in (1) for $L_{poly} = 100$ nm NMOS devices with different halo implant doses. The error bars represent the maximum and minimum from three measurements. (b) Extracted m parameter in (1) with different halo implant doses.

sion for off-state power can be written as

$$P_{off} = \left[\begin{array}{l} \text{All devices with} \\ \text{"conventional" leakage} \\ \text{[see Fig. 1(a)]} \end{array} \right] + \left[\begin{array}{l} \text{All devices with} \\ \text{uniform gate leakage} \\ \text{[see Fig. 1(b)]} \end{array} \right] \quad (2a)$$

$$P_{off} = (n_1 W V_{DD} I_{off} + n_2 J_{gate-e} V_{DD} W L_{poly})_{NMOS} + (p_1 W V_{DD} I_{off} + p_2 J_{gate-h} V_{DD} W L_{poly})_{PMOS} \quad (2b)$$

where V_{DD} is the supply voltage, W and L_{poly} are the transistor width and length, respectively. The I_{off} term represents the leakage illustrated in Fig. 1(a) while J_{gate} represents the leakage illustrated in Fig. 1(b). The indices n_1, n_2, p_1 and p_2 represent the number of transistors at the corresponding bias defined in Fig. 1.

There are at least two important consequences to consider regarding the impact of gate leakage on off-state power consumption. First, the number of transistors contributing to P_{off} significantly increases. For conventional CMOS technology only devices biased such as n_1 and p_1 contribute to P_{off} . The addition of the gate leakage term as defined by n_2 and p_2 introduces more devices contributing to P_{off} . As a consequence, the I_{off} specification for a given technology or application may need to be decreased in magnitude in order to compensate for the contributions from gate leakage. The second consequence to consider is the impact of gate current on conventional off-state leakage. As illustrated by Fig. 4, the magnitude of the gate current becomes a significant fraction of the off-state drain current for gate oxides thinner than 1.8 nm; therefore, the constraints on the leakage due to DIBL become more severe in order to achieve the desired total I_{off} . This second consequence suggests that higher demands will be placed on channel engineering if oxides thinner than 1.8 nm are to be used.

To further examine the impact of uniform gate leakage [Fig. 1(b)], the data in Fig. 2 has been converted to $I_{leakage}$ per unit width as described in (3) and plotted in Fig. 8 for biases of 1.3 V, 1.0 V and 0.7 V.

$$\frac{I_{leakage}}{W} = J_{gate} L_{poly} \quad (3)$$

where $I_{leakage}$ is the leakage current illustrated in Fig. 1(b), W and L_{poly} represent the channel width and poly length, respectively, and J_{gate} represents the measured gate leakage plotted in Fig. 2. Equation (3) normalizes the uniform gate leakage to amperes per unit width, which is consistent to compare with the conventional I_{off} per unit width. Conventional DIBL leakage data from IMEC's sub 250 nm NMOS devices are included for comparison only. Note that values of DIBL leakage for $L_{poly} = 50$ nm is not available since the smallest gates from the IMEC devices are 100 nm.

Note that the electrical data in Fig. 8 are not optimized for gate lengths of 100 nm and sub 100-nm CMOS technology; however, the trends of the data provide some insight into the limitations of oxide scaling due to gate leakage. For example, assuming a high performance specification for off-state leakage of 10 nA/μm, a 1.4–1.5 nm oxide exhibits leakage less than 10 nA/μm for a gate length of 50 nm and a supply voltage of 1.3 V. Assuming a portable or low power application, the estimated gate leakage for a gate oxide of 1.8 to 1.9 nm, a gate length of 50 nm and supply voltage of 1.3 V is approximately 0.1 nA/μm. Further reduction of the applied voltage reduces the gate leakage substantially. For an applied bias of 1 V and gate lengths between 50 nm to 100 nm, the estimated gate leakage for the 1.4 nm oxide is between 1–3 nA/μm. The corresponding gate leakage for an applied bias of 0.7 V is between 0.2–0.6 nA/μm. These examples are simplistic, however, the estimates show that the gate leakage

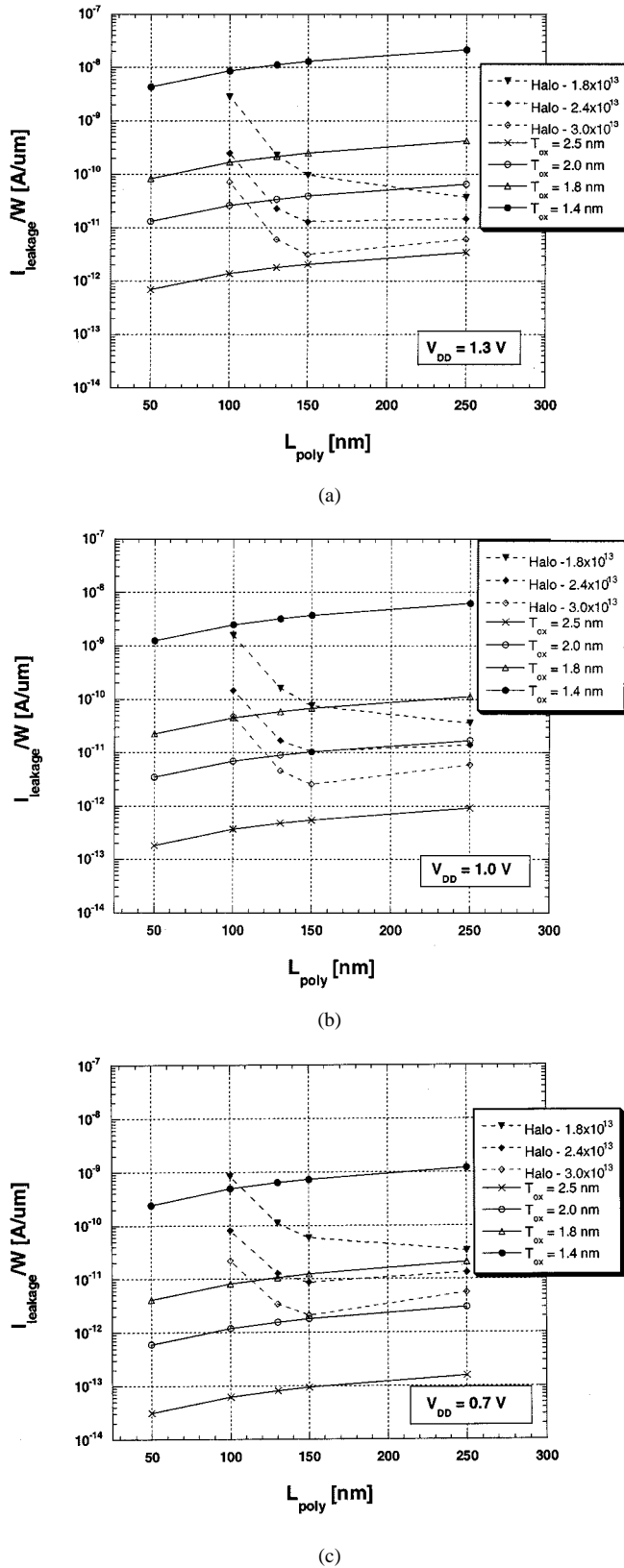


Fig. 8. Off-state leakage per unit width for a single transistor. Values are calculated using experimental data. The dashed lines represent DIBL leakage and the solid lines represent the gate leakage. (a) $V_{\text{DD}} = 1.3$ V. (b) $V_{\text{DD}} = 1.0$ V. (c) $V_{\text{DD}} = 0.7$ V.

illustrated in Fig. 1(b) is within NTRS specifications for oxides down to 1.4 nm.

A detailed analysis of the conventional off-state bias [Fig. 1(a)] where gate leakage and DIBL leakage are present in the same device is beyond the scope of the available data. Although the gate leakage from Fig. 4 are also within the NTRS specifications for off-state leakage, the addition of gate leakage and DIBL leakage may require less aggressive scaling of oxide thickness or more aggressive scaling of supply voltage. The combination of the sharp increase of DIBL leakage with decreasing channel length observed in Fig. 8 plus the gate leakage component illustrated in Fig. 4 certainly requires serious consideration in sub-100-nm CMOS device design.

Note that the analysis of gate leakage in this work did not include any data from PMOS devices; however, the trends are not expected to become worse since the magnitude of electron tunneling is larger than hole tunneling for a given oxide thickness and bias [25]. Also, the impact of high channel doping on gate leakage was not considered. Again the trends are not expected to become worse since for a given bias and oxide thickness, the gate leakage is expected to decrease with increasing doping [15]. With respect to gate leakage, the trends of the data in Fig. 8 are slightly pessimistic since the channel doping is expected to be significantly higher than the doping of the experimental devices.

Although gate leakage does not appear to limit the oxide scaling down to 1.4 to 1.5 nm or 1.8 to 2.0 nm for high performance and low power applications, respectively, there are other issues to consider. The experimental data in this work primarily consist of pure silicon oxide, whereas the real technologies will use some form of nitrided oxide to control boron penetration. A more realistic analysis of gate leakage should study the oxynitrides [26], [27] intended for use in these advanced CMOS technologies since tunneling is very sensitive to dielectric and interface properties. More important, reliability constraints have been reported that propose a more severe limit on oxide scaling [28]–[30] while other reports indicate that reliability constraints are not as severe [31]–[33].

V. SUMMARY

Leakage currents for advanced CMOS technology include conventional DIBL leakage and gate leakage. The DIBL leakage characteristics exhibit surface punchthrough conduction mechanisms that are exponentially dependent on channel length, drain bias and temperature. The gate leakage current depends exponentially on oxide thickness and gate bias. The most critical bias for gate leakage corresponds to $V_{\text{gate}} = V_{\text{DD}}$ and $V_{\text{drain}} = 0$ V where the leakage current occurs across the entire device area. The conventional off-state bias of $V_{\text{gate}} = 0$ V and $V_{\text{drain}} = V_{\text{DD}}$ will have an increasingly significant contribution from the gate due to tunneling; however, the magnitude of gate leakage is very small since the effective tunneling area is concentrated at the gate to drain overlap region.

Gate leakage current is expected to have two major consequences on off-state power consumption. First, the total number of devices contributing to the off-state power consumption increases. This may result in a lower off-state current specification to compensate for the additional contribution from gate leakage. Second, the conventional off-state bias (e.g., $V_{\text{gate}} =$

0, $V_{\text{drain}} = V_{\text{DD}}$) will have an increasing contribution from the gate current. This increases the demands on channel engineering to lower DIBL leakage in order to compensate for the gate leakage component.

Although gate leakage is expected to introduce new challenges in CMOS device design, it is estimated that the silicon oxide can scale down to the 1.4 to 1.5 nm regime for high performance applications and the off-state leakage will remain within the target specifications ($I_{\text{off}}/W < 10 \text{ nA}/\mu\text{m}$). However, reliability of silicon oxide remains a controversial issue and it will most likely be the limiting factor for scaling silicon oxide thickness in high performance applications. For low power applications with lower operating voltages, reliability of the gate oxide becomes less of an issue and gate leakage becomes the limiting factor with respect to scaling silicon oxides. Using the NTRS specifications, the gate oxide thickness is estimated to scale down to the 1.8 to 2.0 nm regime for low power and memory applications.

ACKNOWLEDGMENT

The authors would like to thank D. Blackburn and H. Maes for their helpful comments and suggestions on the manuscript.

REFERENCES

- [1] R. R. Troutman and A. G. Fortino, "Simple model for threshold voltage in short-channel IGFETs," *IEEE Trans. Electron Devices*, vol. ED-24, p. 1266, 1977.
- [2] R. R. Troutman, "VLSI limitations from drain-induced barrier lowering," *Solid-State Circuits*, vol. 14, p. 383, 1979.
- [3] K.-Y. Fu and Y. L. Tsang, "On the punchthrough phenomenon in submicron MOS transistors," *IEEE Trans. Electron Devices*, vol. 44, p. 847, 1997.
- [4] B. Eitan and D. Frohman-Bentchkowsky, "Surface conduction in short-channel MOS devices as a limitation to VLSI scaling," *IEEE Trans. Electron Devices*, vol. ED-29, p. 254, 1982.
- [5] M. B. Barron, "Low-level currents in insulated gate field effect transistors," *Solid-State Electron.*, vol. 15, p. 293, 1972.
- [6] *National Technology Roadmap for Semiconductors*: Semiconductor Indust. Assoc., 1997.
- [7] H. Iwai *et al.*, "The future of ultra-small geometry MOSFET's beyond 0.1 micron," *Microelectron. Eng.*, vol. 28, p. 147, 1995.
- [8] G. Timp *et al.*, "Low leakage, ultra-thin gate oxides for extremely high performance sub-100 nm nMOSFETs," in *IEDM Tech. Dig.*, 1997, p. 930.
- [9] H. S. Momose *et al.*, "1.5 nm direct-tunneling gate oxide Si MOSFETs," *IEEE Trans. Electron Devices*, vol. 43, p. 1233, 1996.
- [10] H. Iwai and H. S. Momose, "Technology toward low power/low voltage and scaling of MOSFETs," *Microelectron. Eng.*, vol. 39, p. 7, 1997.
- [11] E. O. Kane, "Theory of tunneling," *J. Appl. Phys.*, vol. 32, p. 83, 1961.
- [12] S. Nagano *et al.*, "Mechanism of leakage current through nanoscale SiO_2 layer," *J. Appl. Phys.*, vol. 7, p. 3530, 1994.
- [13] W. A. Harrison, "Tunneling from an independent particle point of view," *Phys. Rev.*, vol. 123, p. 85, 1961.
- [14] S. Lo *et al.*, "Quantum mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFETs," *IEEE Electron Device Lett.*, vol. 18, p. 209, 1997.
- [15] N. Yang *et al.*, "Modeling study of ultrathin gate oxides using direct tunneling current and capacitance-voltage measurements in MOS devices," *IEEE Trans. Electron Devices*, vol. 46, p. 1464, 1999.
- [16] C.-H. Choi *et al.*, "Direct tunneling current model for circuit simulation," in *IEDM Tech. Dig.*, 1999, p. 735.
- [17] S. Kubicek *et al.*, "0.13 μm CMOS technology with optimized poly-Si/NO-oxide gate stack," in *Proc. 1st Int. Symp. ULSI Process Integration*. Honolulu, HI, 1999, p. 193.
- [18] J. L. Moll, *Physics of Semiconductors*, New York: McGraw-Hill, 1964.
- [19] S. M. Sze, *Physics of Semiconductor Devices*, New York: Wiley, 1981.

- [20] J. R. Hauser and K. Ahmed, "Characterization of ultrathin oxides using electrical C-V and I-V measurements," in *Characterization and Metrology for ULSI Technology: 1998 Int. Conf.*, p. 235.
- [21] W. K. Henson *et al.*, "Estimating oxide thickness of tunnel oxides down to 1.4 nm using conventional capacitance-voltage measurements on MOS capacitors," *IEEE Electron Device Lett.*, vol. 20, p. 179, 1999.
- [22] K. Ahmed *et al.*, "Impact of tunnel currents and channel resistance on the characterization of channel inversion layer charge and polysilicon-gate depletion of sub-20 Å gate oxide MOSFET's," *IEEE Trans. Electron Devices*, vol. 46, p. 1650, 1999.
- [23] M. J. Van Dort, "A simple model for quantization effects in heavily doped silicon MOSFET's at inversion conditions," *Solid-State Electron.*, vol. 37, p. 411, 1994.
- [24] H. Van Meer *et al.*, "Limitations of shift-and-ratio based I_{eff} extraction techniques for MOS transistors with halo or pocket implants," *IEEE Electron Device Lett.*, vol. 21, p. 133, Mar. 2000.
- [25] Y. Shi *et al.*, "Polarity dependent gate tunneling currents in dual-gate CMOSFET's," *IEEE Trans. Electron Devices*, vol. 45, p. 2355, 1998.
- [26] E. Ibong *et al.*, "Gate quality ultrathin (2.5 nm) PECVD deposited oxynitrides and nitrided oxide dielectrics," *IEEE Electron Device Lett.*, vol. 20, p. 442, 1999.
- [27] X. Guo and T. P. Ma, "Tunneling leakage current in oxynitride: Dependence on oxygen/nitrogen content," *IEEE Electron Device Lett.*, vol. 19, p. 207, 1998.
- [28] J. H. Stathis and D. J. DiMaria, "Reliability projection for ultra-thin oxides at low voltage," in *IEDM Tech. Dig.*, 1998, p. 167.
- [29] R. Degraeve *et al.*, "Temperature acceleration of oxide breakdown and its impact on ultra-thin gate oxide reliability," in *VLSI Tech. Dig.*, 1999, p. 59.
- [30] R. Degraeve *et al.*, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Trans. Electron Devices*, vol. 45, p. 904, 1998.
- [31] Y. Wu *et al.*, "Time dependent dielectric wearout (TDDW) technique for reliability of ultra-thin gate oxides," *IEEE Electron Device Lett.*, vol. 20, p. 262, 1999.
- [32] B. E. Weir *et al.*, "Ultra-thin gate dielectrics: They break down but do they fail?," in *IEDM Tech. Dig.*, 1997, p. 73.
- [33] T. Ghani *et al.*, "100 nm gate length high performance/low power CMOS transistor structure," in *IEDM Tech. Dig.*, 1999, p. 415.

W. Kirklen Henson was born in Shelby, NC, in 1972. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from North Carolina State University, Raleigh, in 1994, 1995, and 1998, respectively. For his Ph.D. dissertation he worked on direct tunnel oxides for MOSFET applications comparing gate oxides formed by rapid thermal chemical vapor deposition and rapid thermal oxidation.

He is currently with the Interuniversity Micro-Electronics Center, Leuven, Belgium. His research interests involve CMOS process integration, device physics and electrical characterization.

Nian Yang (S'96-M'99) received the B.S. degree in electronics from Peking University, Beijing, China, in 1991, the M.S. degree in physics from Purdue University, West Lafayette, IN, in 1993, the M.S. degree in biophysics from the University of Illinois, Urbana-Champaign (UIUC), in 1996, and the Ph.D. degree in electrical engineering from North Carolina State University (NCSU), Raleigh, in 1999. His Ph.D. research at NCSU involved MOS device physics and modeling, gate dielectrics processing, reliability and interface study, and rapid thermal techniques.

From 1991 to 1996, he conducted research in nuclear magnetic resonance and its biomedical applications at both Purdue and UIUC. Currently, he is with Advanced Micro Devices, Sunnyvale, CA, working on the nonvolatile memory technology.

Stefan Kubicek received the M.S. degree in electrical engineering from Slovak Technical University, Bratislava, in 1987, and M.S. degree in physics of microelectronics and materials science in 1992 from Katholieke Universiteit Leuven, Belgium.

Since August 1993, he works in Advanced Silicon Devices group, Interuniversity Micro-Electronics Center, Leuven, Belgium. The topic of his research is scaling the CMOS technology to sub-0.1 μm region with the focus on device integration issues.

Eric M. Vogel was born in Charleroi, PA, in 1972. He received the B.S. degree in electrical engineering (with honors) from the Pennsylvania State University, University Park, in 1994, and the M.S. and Ph.D. degrees in electrical engineering, in 1996 and 1998, respectively, both from North Carolina State University, Raleigh. His dissertation was on the use of RTCVD oxynitrides for advanced MOS gate dielectrics.

In June 1998, he joined the Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD, working in the areas of electrical characterization, gate dielectric reliability, and thin-film metrology for CMOS and beyond devices.

Jimmie J. Wortman (M'72–SM'88) was born in North Carolina on February 23, 1936. He received the B.S. degree (with honors) in nuclear engineering from North Carolina State University (NCSU), Raleigh, in 1960, and the M.S. and Ph.D. degrees from Duke University, Durham, NC, in 1962 and 1965, respectively.

He was with Bell Telephone Laboratories as Member of Technical Staff from 1960 to 1962. From 1962 to 1980, he worked at the Research Triangle Institute, Research Triangle Park, Durham. During that period, he held the positions of Research Engineering, Senior Engineering, Senior Scientist, manager of several departments, and the Director of the Engineering Division. He is currently a professor in the Electrical and Computer Engineering Department at NCSU. His current research interests include rapid thermal processing, rapid thermal oxides, and stacked gate dielectric structures.

Kristin De Meyer (M'85) received the degree in electrical engineering in 1974 and the Ph.D. degree in 1979, both from the Katholieke Universiteit Leuven (KUL), Belgium.

From 1974 to 1975, she was a Fellow of the Institute for Scientific Research in Industry and Agriculture at the Laboratory for Physics and Electronics of Semiconductors, KUL, working on CCD filter design. From October 1975 through September 1979, she worked as a Research Assistant of the National Fund for Scientific Research (NFWO) at the same laboratory on the design, development, and characterization of two-phase buried channel CCD's. From November 1979 through November 1980, she was granted an IBM World Trade Post Doctoral Fellowship and worked at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, on the development of EARAM's using Si-rich oxides. At the end of 1980, she returned to KUL as a Senior Research Assistant of the NFWO, and from October 1982 as a Research Associate of the NFWO. With the founding of the Interuniversity Micro-Electronics Center (IMEC), her activities moved from the KUL to IMEC, where she initially remained Research Associate of the NFWO. Since October 1989, she has been with IMEC. From 1985 through 1992, she was Head of the Group on Process and Device Modeling and Simulation, working on process optimization techniques, parameter extraction, analytical and numerical device modeling, and computational physics. Her group also monitored Statistical Process Control Activities. Since January 1993, she has been in charge of deep submicron MOS technology and exploratory field effect devices. Since October 1986, she is also a part-time Professor at KUL. Dr. De Meyer is the coordinator for IMEC in the ESPRIT 962 EVEREST, STORM, NOVA, ULTRA, VAHMOS2000, FASEM, ULIS projects. She was also involved in the ESPRIT ACCESS program, ADEQUAT, ACE. She also managed the HCM-SUSTAIN network on deep submicron silicon technology with 21 partners from academia and research institutes. She authored or co-authored over 100 publications and organized summercourses and the SISPAD 98 conference.



Abdalla Naem received the Ph.D. degree in electrical engineering from Carleton University, Ottawa, Ont., Canada, in 1984, in the area of VLSI microelectronics. His doctoral research dealt with narrow channel device analysis and modeling.

From 1982 to 1984, he was a full-time consultant for Advanced Processing Division, Bell-Northern Research, Ottawa, where he carried out several research projects including: the invention and application of selective laser processing technique in VLSI circuits, semiconductor/metal interfaces,

and 3-D integration. He joined Northern Telecom Electronics, Ottawa, as Member of Scientific Staff in 1984. In 1987, he was a Senior Member of Scientific Staff. His activities at Northern included formation of ultrashallow junctions, titanium/platinum/cobalt salicides process development/integration for submicron CMOS/BiCMOS technologies, rapid thermal processing, and high precision resistors for telecom applications. He joined National Semiconductor Corporation, Santa Clara, CA, in 1995, as a Principal Engineer engaged in research and development/integration of 0.35 and 0.25 μm CMOS technologies. In 1997, he became a senior engineering manager involved in the research and development of advanced CMOS and BiCMOS technologies down to 0.13 μm devices. He has published over 40 papers in refereed journals or conferences, and holds 12 patents. He has also been engaged in supervising M.Sc. and Ph.D. students from different Canadian universities, where several joint research programs were carried out. He has also collaborated with the National Research Council of Canada and IMEC of Belgium.

Dr. Naem chaired many sessions at international conferences and served on the committees of international conferences and workshops in the area of microelectronics technology.