Voltage- and Temperature-Dependent Gate Capacitance and Current Model: Application to ZrO_2 n-channel MOS Capacitor

Yang-Yu Fan, Renee E. Nieh, Jack C. Lee, *Fellow, IEEE*, Gerry Lucovsky, George A. Brown, *Member, IEEE*, Leonard Frank Register, and Sanjay K. Banerjee, *Fellow, IEEE*

***Abstract*—Based on the energy-dispersion relation in each region of the gate-dielectric-silicon system, a tunneling model is developed to understand the gate current as a function of voltage and temperature. The gate capacitance is self-consistently calculated from Schrödinger and Poisson equations subject to the Fermi–Dirac statistics, using the same band structure in the silicon as used for tunneling injection. Franz two-band dispersion is assumed in the dielectric bandgap. Using a Wentzel–Kramer–Brillouin (WKB)-based approach, direct and Fowler–Nordheim (FN) tunneling and thermionic emission are considered simultaneously. The model is implemented for both the silicon conduction and valence bands and both gate- and substrate-injected currents. ZrO_2 NMOSFETs were studied through temperature-dependent C_g-V_g and I_g-V_g simulations. The extracted band gaps and band offsets of the ZrO_2- and interfacial-Zr-silicate-layer are found to be comparable with the reported values. The gate currents in ZrO_2-NMOSCAPs are found to be primarily contributed from the silicon conduction band and tunneling appears to be the most probable primary mechanism through the dielectric. Oscillations of gate currents and kinks of gate capacitance were observed near flat-band in the experiments. These phenomena might be caused by the interface states.**

***Index Terms*—High-K gate dielectric, leakage currents, MIS devices, MOSFETs, semiconductor device modeling, tunneling.**

I. Introduction

SiO_2 has been used as the gate dielectric because of its process advantages and good interface properties with silicon. As gate lengths have been reduced, the SiO_2 thickness also has been reduced to maintain gate control over the channel. However, in the deep submicron regime, the direct tunneling current limits this scaling process [1]–[3]. In order to overcome this limitation, high-K (or high dielectric constant) materials are being introduced to achieve a greater physical thickness and thus reduce the direct tunneling current while retaining a low equivalent oxide thickness (EOT).

Different materials such as ZrO_2 and HfO_2 have been shown [4]–[6] promise for transistor application. Owing to the complex fabrication process of transistors, most studies have been conducted on MOSCAPs. The uncertainty of the material properties, especially for the thin films, imposes another difficulty in terms of understanding the experimental results. In this work, through C_g–V_g and I_g–V_g simulations, the material properties of ZrO_2-NMOSCAPs are extracted in the accumulation region.

High-K materials, not coincidentally, generally have smaller band gaps and smaller band offsets with silicon than SiO_2 [7]. In addition to direct tunneling, Fowler-Nordheim (FN) tunneling can be important because of the smaller band offsets. The thermionic emission also needs to be considered for materials with extremely small band offsets and for transistors in which hot carriers are of concern. Gate currents from the silicon conduction band and valence band may be nonnegligible at the same time. Furthermore, an interfacial layer is generally found between the high-K layer and the silicon substrate. These all impose difficulties in terms of understanding the experimental gate currents.

With increasing normal fields as a result of high doping concentrations and reduced EOTs, quantum confinement effects can no longer be neglected [8]–[11]. In this paper, we report a Wentzel–Kramer–Brillouin (WKB)-based gate current model that considers quantum confinement effects and is applied to the inversion and accumulation regions and both the conduction and valence band components. Direct, FN tunneling and thermionic emission are considered simultaneously.

The gate capacitance has been used for extracting the EOTs [8]–[10], beyond which the capacitance-voltage behavior should be modeled concurrently with current-voltage. The charge redistribution in the gate-dielectric-silicon structure must be determined self-consistently by solving the Poisson-Schrödinger equations. The gate capacitance provides an experimental probe of the charge distribution at different gate voltages and this is key to modeling the gate current. Temperature-dependence study of gate currents and gate capacitance of MOS capacitors with SiO_2 shows this correlation: the temperature-dependent region of the gate capacitance, which is near the flat-band, is the same as that of the gate current [12]. Such temperature-dependence is attributed to temperature-dependence of the charge distribution or "supply function" instead of the temperature-dependence of the carrier transport in the SiO_2. In this work, the simultaneous agreement of gate capacitance and gate current simulation with experiment has been obtained.

Manuscript received April 2, 2002. This work was supported in part by Semiconductor Research Corporation. The review of this paper was arranged by Editor C. C. McAndrew.

Y.-Y. Fan, R. E. Nieh, J. C. Lee, L. F. Register, and S. K. Banerjee are with the Microelectronics Research Center, University of Texas at Austin, Austin, TX 78758 USA (e-mail: yyfan@mail.utexas.edu).

G. Lucovsky is with the Department of Physics, North Carolina State University, Raleigh, NC 27695 USA.

G. A. Brown is with the International SEMATECH, Austin, TX 78741 USA.

Digital Object Identifier 10.1109/TED.2002.804713

0018-9383/02$17.00 © 2002 IEEE

In Section II, the key features of the model and the model validation using SiO_2 will be presented. The experimental analysis of ZrO_2-NMOSCAPs and parameter extraction will be discussed in Section III. Finally, we will summarize this work in Section IV.

II. Modeling

The independent particle approach [13] is adopted to model the tunneling currents through the gate dielectric. If the energy dispersion relation (E verus k) in each region is available, the tunneling probability is obtained by a WKB method. In the originally proposed approach, the wave number vector parallel to region interfaces, $k_{||}$, is required to be conserved. However, well-known inconsistencies between the assumption and experiment are found between gate currents in SiO_2-MOS devices fabricated on Si(110) and Si(111) [14] and this issue has not been resolved. In this work, simultaneous agreement of gate current simulation with experiment in inversion and accumulation is achieved for SiO_2 devices fabricated on Si(100) with $k_{||}$ conserved. The primary effect of using this assumption or not seems to be primarily to decrease the calculated band-edge masses.

A physical picture of this model can be described in terms of electrons tunneling from the silicon substrate to the gate (see Fig. 1). The incident component of the electron wave function, subject to the band structure (E, k_x, k_y, k_z) tunnels through the barrier and contributes to the transmitted currents. The energy-dispersion relation in the dielectric bandgap and available states in the gate determine the transmission probability. The continuity of electron flux allows the gate-injected electron current to be calculated by the transmission of electrons into the silicon at silicon-dielectric interface.

The carrier states in the silicon channel are roughly divided into bound states and unbound states. Bound states are quantized states inside the quantum well formed by the silicon band-edge. Subband structures and envelope functions are obtained by solving the Schrödinger equation in k-space using a full-band formalism [15], [25]. Those states outside the quantum well are approximated as free and unbound and the energy-dispersion at the silicon-dielectric interface is approximated by the bulk-silicon band structure.

The apparent energy dispersion "seen" by the tunneling electron in the dielectric band gap is approximated by Franz's two-band model [16]

$$\frac{1}{k^2} = \frac{\hbar^2}{2m_c(E-E_c)} + \frac{\hbar^2}{2m_v(E_v-E)} \tag{1}$$

where

E total energy of the electron;

m_c and m_v *band-edge* effective masses for the conduction and valence bands, respectively, of the dielectric;

E_c and E_v energies of the respective band-edges.

Such dispersion relation has previously been shown to be effective approximation for conventional oxides [17], [18]. m_c and m_v are unknown parameters to be extracted by comparing gate current simulation with experiment. Table I summarizes the difference between this work and others of tunneling through SiO_2.

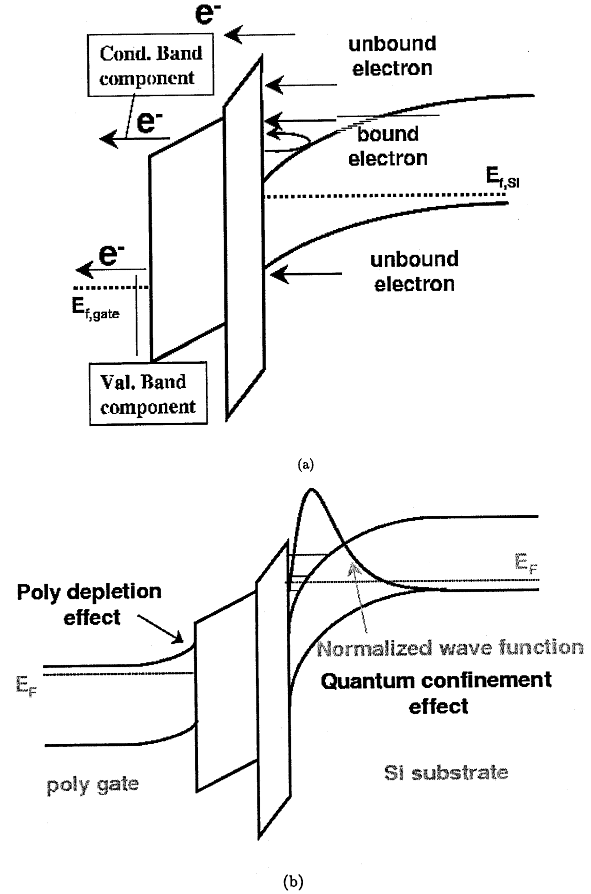


Fig. 1. Band diagrams of a gate-dielectric-silicon system being biased in the inversion region for (a) gate current model and (b) gate capacitance model.

The tunneling probability $T(E, k_{||})$ through the dielectric is calculated by the WKB method [22] when $k_z^2 < 0$

$$-\ln T(E, k_{||}) = \int_{\text{dielectric}} |k_z| \, dz \tag{2}$$

with

$$k_z^2 = k^2 - k_{||}^2 \tag{3}$$

where k^2 is calculated from Franz dispersion and $k_{||}$-conservation is applied.

The total energy E is used, thus direct, FN tunneling and thermionic emission are considered simultaneously:

- direct tunneling: $k_z^2 < 0$ across the whole dielectric;
- FN tunneling: k_z^2 passes through 0 within the dielectric;
- thermionic emission: $k_z^2 > 0$ across the whole dielectric.

These mechanisms are all considered elastic as the total energy and $k_{||}$ are conserved during transport through the dielectric. The image force is lower in higher-K materials and reduced by the quantum mechanical repulsion of both bound and unbound carriers from the interface. Thus, its effect on barrier lowering is not considered.

Finally, the total gate current density J_g is calculated as

$$J_g = J_{\text{Si-conduction}} + J_{\text{Si-valence}} \tag{4}$$

TABLE I
COMPARISON BETWEEN THIS WORK AND OTHER MODELS OF TUNNELING THROUGH GATE SILICON DIOXIDE. EMA STANDS FOR EFFECTIVE MASS APPROXIMATION

	This Work	Parabolic EMA	IBM's EMA	NC state's EMA
Reference	N/A	Ref. 20,21	Ref. 19	Ref. 9
Numerical approach	WKB, self-consistent	Resonance, self-consistent	Transfer Matrix, self-consistent	WKB, non-self-consistent
$k_{\parallel}$-conserved	yes	no	no	no
Dielectric Modeled	Franz-type like	parabolic	Franz-type	Franz-type
Dielectric treated	SiO_2, High-K	SiO_2	SiO_2	SiO_2
Transport modeled	Direct, FN tunneling Thermionic emission	Direct tunneling	Direct tunneling	Direct tunneling

where $J_{\text{Si-conduction}}$ is from the conduction-band longitudinal valleys. With $k_{\parallel}$-conservation, the effective barrier for transverse valleys is higher and tunneling from them is negligible.

Each component J_i, where i designates the band, is calculated separately for unbound states and bound states. For unbound states in the conduction band

$$J_{i,\text{unbound}} = \frac{2q}{(2\pi)^3}\int_{E_{z,\min,c}}^{\infty} dE_z \cdot \int_{k_{\parallel}=(0,0)}^{\infty} dk_{\parallel}^2 T\left(E,k_{\parallel}\right)\left(f_{FD,\text{gate}} - f_{FD,\text{Si}}\right) \quad (5)$$

and for unbound states in the valence band

$$J_{i,\text{unbound}} = \frac{2q}{(2\pi)^3}\int_{-\infty}^{E_{z,\max,v}} dE_z \cdot \int_{k_{\parallel}=(0,0)}^{\infty} dk_{\parallel}^2 T\left(E,k_{\parallel}\right)\left(f_{FD,\text{gate}} - f_{FD,\text{Si}}\right) \quad (6)$$

where

$f_{FD,\text{gate}}$ and $f_{FD,\text{Si}}$ Fermi–Dirac distribution functions for the gate and silicon substrate, respectively;

$E_{z,\min,c}(E_{z,\max,v})$ lowest (highest) unbound state in the conduction (valence) band;

q charge of electron.

While for bound states

$$J_{i,\text{bound}} = \frac{2q}{(2\pi)^2}\sum_{\mu}\tau_{\mu}^{-1} \cdot \int_{k_{\parallel}=k_{\parallel,\min}}^{\infty} dk_{\parallel}^2 T\left(E,k_{\parallel}\right)\left(f_{FD,\text{gate}} - f_{FD,\text{Si}}\right) \quad (7)$$

where τ_{μ}^{-1} is approximated by the surface impact frequency for a two-dimensional (2-D) subband μ [21]

$$\tau_{\mu}^{-1} = \left(\frac{\partial I}{\partial E}\right)^{-1}_{\text{subband}\,\mu} \quad (8)$$

with the action integral between the classical turning points

$$I = \oint k_z^{(\mu)} dz \quad (9)$$

and

$$k_z^{(\mu)}(z) = \frac{1}{\hbar}\sqrt{2m_z\left(E_{\mu} - E_{\text{band min.}}(z)\right)} \quad (10)$$

where the contour integral is around the quantum well formed in the channel; E_{μ} is the minimum energy of μth subband; $E_{\text{band min.}}$ is the band-edge in which the subbands formed; and m_z is the effective mass along the tunneling direction z.

$m_z = 0.2m_e$ and $m_z = 0.9m_e$ are used for the conduction-band longitudinal and transverse valleys, respectively, where m_e is the free electron mass. Because the effective mass approximation is not very accurate for the valence band, m_z then is treated as a subband-independent adjustable parameter. From comparison of gate current simulation with experiment for a SiO_2-PMOSFET in inversion, it is found that $m_z = 0.10 \sim 0.25m_e$ gives good agreement (see Fig. 2). The SiO_2 thickness is extracted by comparison of C_g–V_g simulation with experiment in accumulation (see Fig. 2). Similar agreement is also achieved for a SiO_2-NMOSFET, which was fabricated on Si(100) (see Fig. 3). The gate current simulation agrees well with experiment both in inversion and accumulation. The band gap and conduction band offset with silicon of SiO_2 are 9.0 eV and 3.15 eV, respectively. The band-edge effective masses are adjustable parameters but remain the same for the silicon conduction band and valence band components, respectively, whether for substrate-injected or gate-injected currents. More details about gate currents in SiO_2 devices will be discussed in a future paper.

Near the flat-band region, $V_g \sim -0.8$ V in Fig. 3, the quantum well is wide and shallow. Quantization effects due to the quantum well formed by the band-edge become less important. Thus, a clear division between the bound and unbound states can cause significant errors in gate current calculation, which result in abnormal I_g–V_g behavior in simulation. The error may originate from either the impact frequency or the supply of carriers. Below flat-band, only energetic carries can tunnel elastically, greatly reducing the calculated current. In practice, other mechanisms may dominate. A possible mechanism might be related to the interface states, which is not considered in the model. Near flat-band, the carrier density is much lower than in accumulation or inversion. The trapped charges can contribute appreciably to the total gate currents; empty traps may assist the carrier tunneling into/out of the silicon substrate. Contribution from the overlap regions with the source and drain may also be important in that region.

Gate current characteristics at different temperatures are shown in Fig. 4. NMOSFETs with single-layer dielectric of 3.9 with different thicknesses and barrier heights are simulated. The flat-band voltages of these devices are ~ -0.8 V. When

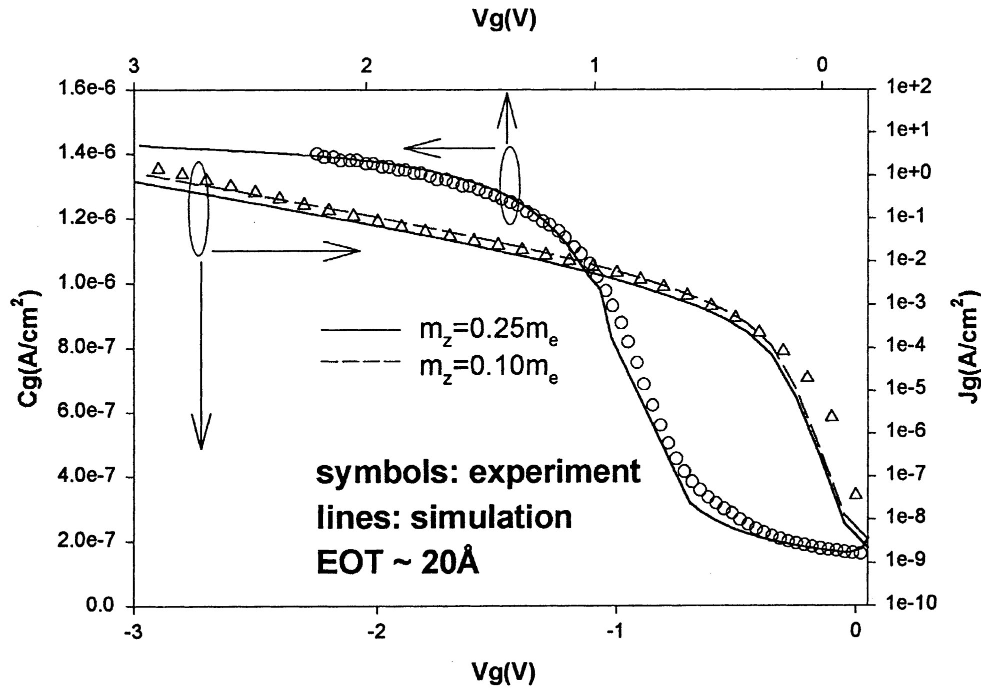


Fig. 2. Comparison of gate current and gate capacitance simulation with the experimental data for a SiO_2-PMOSFET. EOT $\sim$ 20 Å is found from C_g–V_g simulation. Area = 10^{-4} cm^{-2} with $W/L = 1$.

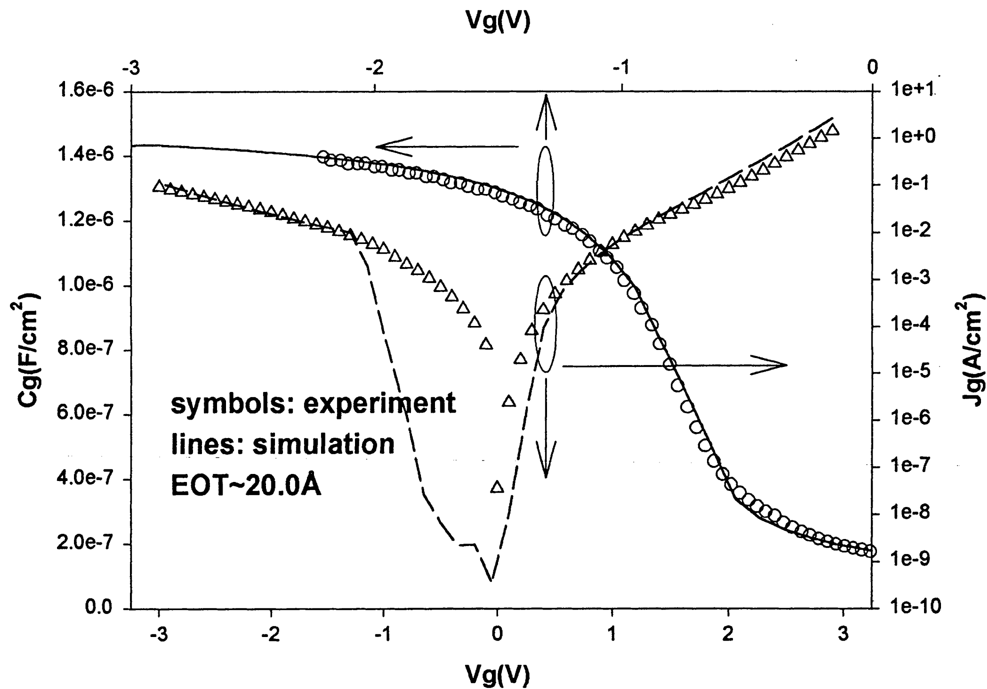


Fig. 3. Comparison of gate current and gate capacitance simulation with the experimental data for a SiO_2-NMOSFET. EOT $\sim$ 20 Å is found from C_g–V_g simulation. Area = 10^{-4} cm^{-2} with $W/L = 1$.

the potential barrier of dielectric is high and/or thin, direct tunneling prevails. However, when the potential barrier is low, transition from direct to FN tunneling is seen. With the potential barrier of 1.0 eV high and 40 Å wide, such transition occurs around $V_g = -2$ V in accumulation and $V_g = 0.8$ V in inversion.

More energetic carriers see a relatively small barrier. Therefore, the hotter the carrier distribution, the greater the tunneling. This field-assisted thermionic effect is particularly significant when the barrier is low and/or thick (see Fig. 4). Polarity effects are also seen in the temperature-dependence of gate current, which is stronger in accumulation than in inversion. Such

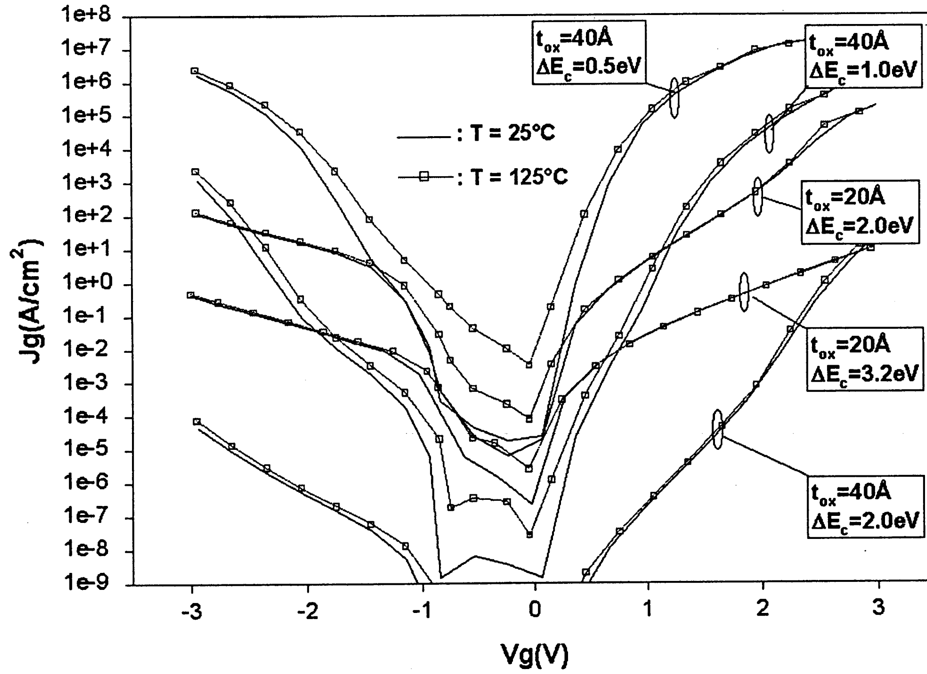


Fig. 4. Simulated I_g–V_g at 25 °C and 125 °C for n-channel MOSFETs with different dielectric thicknesses and conduction band offsets with silicon. Single layer of dielectric of dielectric constant of 3.9 is assumed in the devices.

effects are not due to the transport in the dielectric but the charge supply and asymmetry of the gate stack.

III. Experimental Analysis

TaN/ZrO_2/p-Si NMOSCAPs of area of 5×10^{-5} cm^2 were fabricated on Si(100) wafers with resistivity of 5–25 Ω-cm. The Zr was deposited via dc magnetron sputtering in an argon ambient at room temperature. After deposition, the wafers were rapid thermal annealed (RTA) in N_2 at temperatures ranging from 500 °C to 800 °C at atmospheric pressure. TaN was dc magnetron sputter-deposited in Ar and N_2 at room temperature and dry etched to form the top electrode. Finally, aluminum was deposited on the backside of the wafers for ohmic contacts prior to measurements.

Two samples of different thicknesses of ZrO_2 were fabricated, labeled as "A" and "B." According to TEM analysis, an interfacial Zr-silicate layer exists between the ZrO_2-layer and the silicon substrate as reported in [23]. The physical thickness of each layer is found as: for sample A, $t_{ZrO_2} = 31 \sim 39$ Å and $t_{int} = 6 \sim 10$ Å for sample B, $t_{ZrO_2} = 36 \sim 40$ Å, and $t_{int} = 7 \sim 9$ Å, where t_{ZrO_2} and t_{int} are physical thicknesses of the ZrO_2 and interfacial silicate layers, respectively. Because of the thickness variation, the experimental data being analyzed for either sample were measured from the same device. The data shown is representative of the entire wafer.

To analyze the devices, C_g–V_g simulation is first compared with the experimental data to extract the related device structural parameters (see Fig. 5). At this step, EOT is the only physical parameter that can be extracted and is required for the dielectric stack. The interface state effects are neglected. Effects of fixed charges are compensated by adjusting the metal work function to fit the experimental data. For the fabrication process described above, it was found that the substrate doping concentration is 5×10^{15} cm^{-3}, the metal work function is 4.50 eV for sample A and 4.45 eV for sample B, the EOT of sample A is 10.8 Å and 13.6 Å for sample B. The difference of metal work functions indicates slightly different fixed charge densities at the dielectric-silicon interfaces for A and B. The flat-band voltages are ~ -0.55 V and ~ -0.60 V for samples A and B, respectively.

If abrupt interfaces are assumed, the EOT of the dielectric stack can be calculated by

$$\mathrm{EOT} = t_{\mathrm{ZrO_2}} \times \frac{\epsilon_{\mathrm{SiO_2}}}{\epsilon_{\mathrm{ZrO_2}}} + t_{\mathrm{int}} \times \frac{\epsilon_{\mathrm{SiO_2}}}{\epsilon_{\mathrm{int}}} \tag{11}$$

where ϵ_{SiO_2}, ϵ_{ZrO_2} and ϵ_{int} are dielectric constants of SiO_2, ZrO_2, and Zr-silicate, respectively; $\epsilon_{SiO_2} = 3.9$.

From ranges of dielectric-layer thicknesses, EOTs from C_g–V_g simulation, $\epsilon_{ZrO_2} = 20$ and to satisfy (11) for both samples A and B, $\epsilon_{int} \sim 5.5$ is found. The physical thickness of each layer is determined as: $t_{ZrO_2} = 33.6$ Å and $t_{int} = 6.0$ Å for sample A and $t_{ZrO_2} = 38.5$ Å and $t_{int} = 8.6$ Å for sample B. These values are within ranges obtained from TEM but are not the medium values. To satisfy (11) with the medium values will result in negative ε_{int} and small ε_{ZrO2} ($\sim$ 4.18). With limited experimental information available, the above approximation is the most reasonable. These parameters, along with the ones obtained from C_g–V_g simulation, are used later for gate current simulation.

Temperature-dependence of gate current is shown in Fig. 6. Sample A, with $t_{ZrO_2} + t_{int} = 39.6$ Å, shows temperature-independence in accumulation but temperature-dependence in

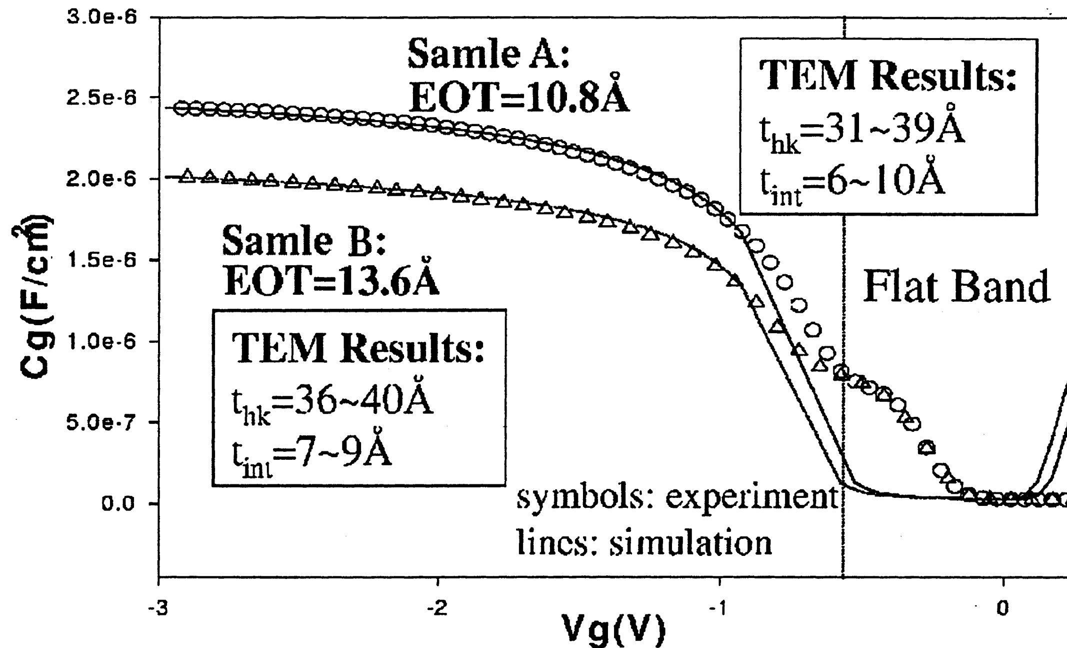


Fig. 5. Comparison of C_g–V_g simulation with the experimental data in accumulation for ZrO_2-NMOSCAPs with different dielectric thicknesses, samples A and B.

inversion. Sample B, with $t_{ZrO_2} + t_{int} = 47.1$ Å, shows weak temperature-dependence in accumulation.

In principle, the temperature-dependence could be contributed from either the transport mechanism, or by the charge distribution in the system, or both. In an NMOSCAP, electrons in the conduction band are minority carriers in the silicon channel. Without a source/drain as external supply, its concentration cannot be maintained at the thermodynamic-equilibrium value. As the temperature is increased, substantially more minority carriers are available and, as a result, the gate currents show strong temperature-dependence in inversion. This strong temperature-dependence contrasts markedly to the much weak temperature-dependence for the field-assisted thermionic current for accumulation shown in Fig. 4, thus is attributed to the temperature-dependence of the charge distribution. The strong temperature-dependence in inversion also indicates that the gate current is primarily from the conduction band.

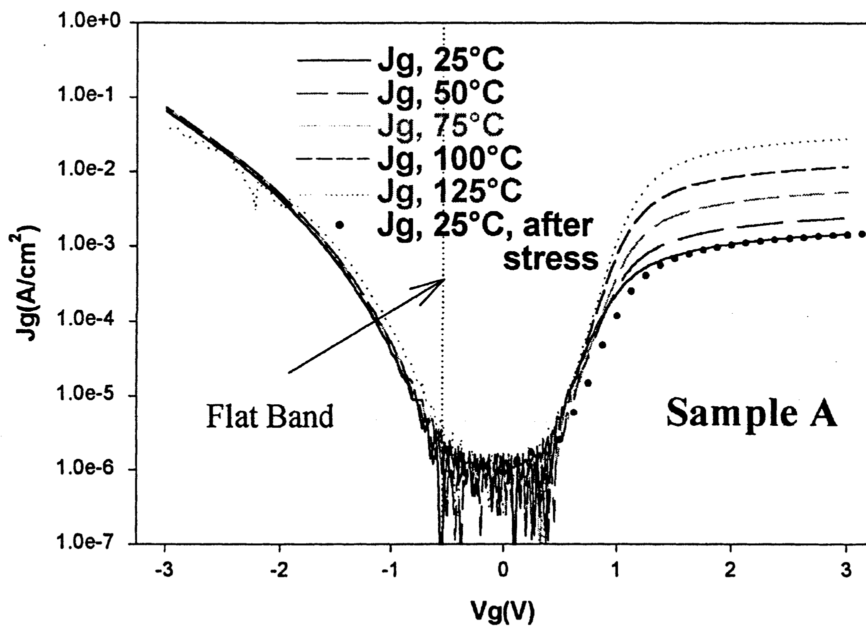
Yamaguchi *et al.* reported that band gaps of ZrO_2 and Zr-silicate layers are 5.7 eV and 4.5 eV, respectively, from XPS analysis for sputter-deposited films [23]. Using these values in Franz dispersion, the gate current simulation agrees well with experiment for both samples (see Fig. 7). The conduction band offsets of ZrO_2 and Zr-silicate are found as 1.45 eV and 1.0 eV, respectively. The band-edge effective masses are $m_c = m_v = 0.35m_e$ for both ZrO_2 and Zr-silicate.

Yamaguchi assumed Frenkel-Poole conduction through the dielectric stack and reported the conduction band offsets with silicon of the ZrO_2 and Zr-silicate layers as 1.5 eV and 1.0 eV, respectively. For Frenkel-Poole conduction [24], the conducting electrons are thermally emitted out of the traps in the dielectric. The trapping potential barrier will be reduced if an electric field is applied. As a result, the current will be increased by a factor proportional to $\exp(\Delta E_i/(k_B T))$, where $k_B T$ is the thermal energy and ΔE_i is the decrease of the trapping potential barrier due to the electric field. The barrier is further decreased if the electric field is increased. Thus, stronger field/voltage-dependence of current is to be expected at low temperatures. Such characteristic was not observed in the gate currents of ZrO_2-NMOSCAPs in our experiments. We believe, therefore, the temperature-dependence observed in sample B is not primarily attributed by the the Frenkel-Poole mechanism.

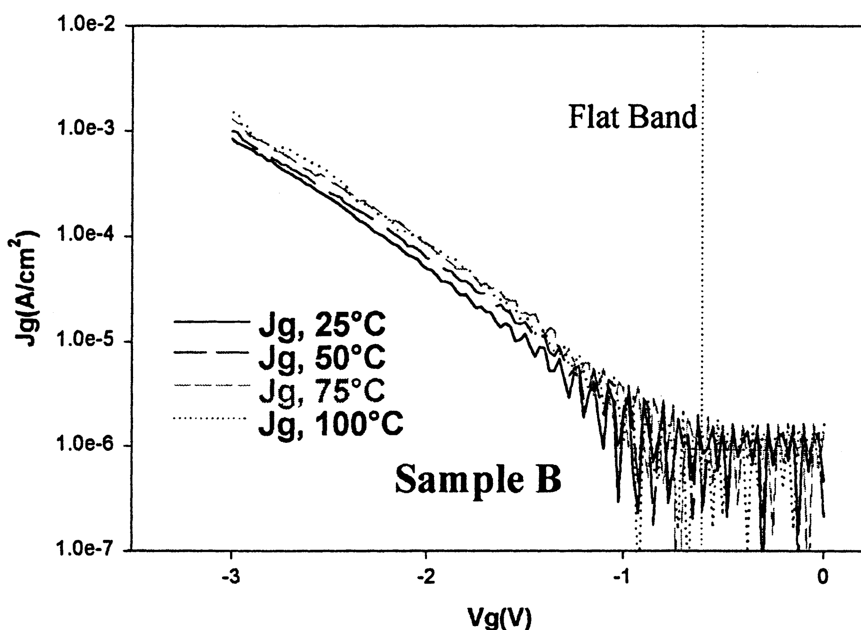
Although simulation and experiment agree quite well, it is not conclusively clear that direct tunneling or FN tunneling, which are the temperature-independent conductions being modeled in this work, is the primary transport mechanism. This is due to the uncertainties about the device structural information, especially those of the dielectric stack. Other temperature-insensitive or weakly dependent transport mechanisms such as trap-assisted tunneling remain possibilities.

Comparison between simulation and experiment of gate currents at different temperatures is shown in Fig. 8. In simulation, the gate current of sample B has stronger temperature-dependence than sample A, consistent with the model for thicker barriers (see Fig. 4). However, the temperature-dependence predicted by the model is weaker than experiment. Thus, some thermally assisting process(es) should be involved, but its influence is weak. Such thermal processes can either be related to the transport through the dielectric or the supply of carriers. If it is related to the transport, it should be weakly dependent on the electric field in the dielectric.

At low gate voltages, oscillations of gate currents were observed near the flat-band and depletion region. Kinks in C_g–V_g were also observed. These might be caused by the interface states between regions (gate/dielectric or dielectric/silicon). The



(a)



(b)

Fig. 6. Gate currents of ZrO_2-NMOSCAPs at different temperatures: (a) sample A, with $t_{\mathrm{ZrO_2}} + t_{\mathrm{int}} = 39.6$ Å and (b) sample B, with $t_{\mathrm{ZrO_2}} + t_{\mathrm{int}} = 47.1$ Å.

high-density of interface states might significantly affect the gate capacitance and gate current by trapping or detrapping the electrons in the absence of inversion or accumulation layers.

The I_g –V_g characteristics from sample A were measured at 25 °C before and after thermal and electrical stresses. Gate voltages from −3 V to 3 V were applied at temperatures of 25 °C–125 °C. The measured gate current is similar to that of the fresh device [see Fig. 6(a)]. This implies good quality of the thin films and few charges are trapped or fresh traps are created within the dielectric stack. This is yet another piece of circumstantial evidence of the unimportance of Frenkel-Poole transport in these samples, at least in accumulation.

The comparison of gate current simulation with experiment for a TaN/HfO_2/p-Si capacitor is shown in Fig. 9. The simulation agrees well with experiment in accumulation. The EOT is ~ 10.5 Å obtained from C_g–V_g simulation. $E_g = 5.7$ eV,

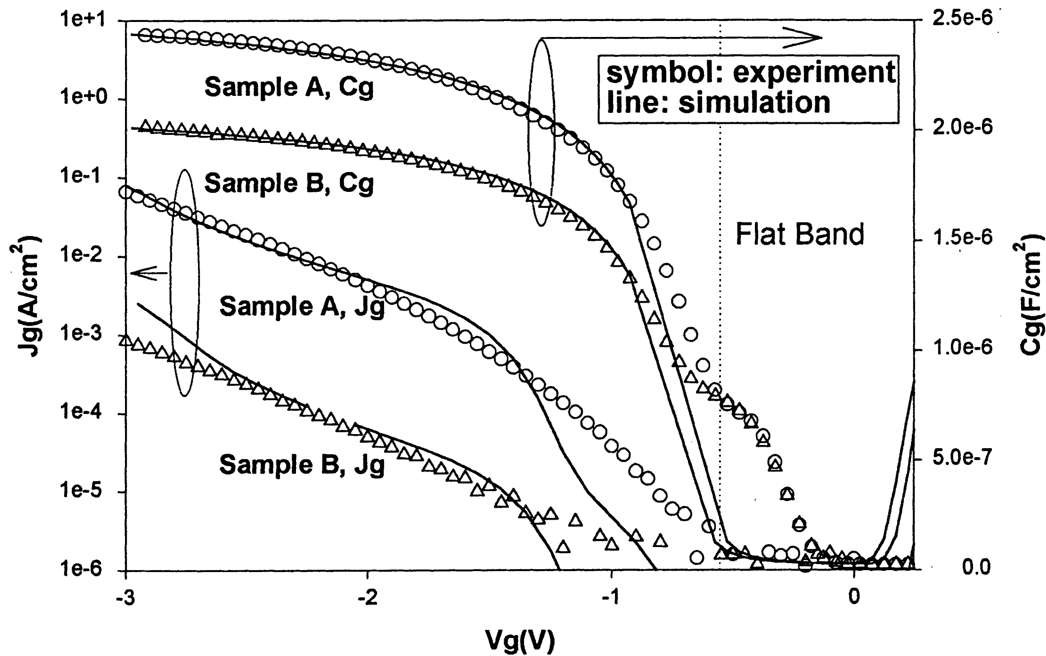


Fig. 7. Comparison of gate current and gate capacitance simulations with the experimental data for samples A and B.

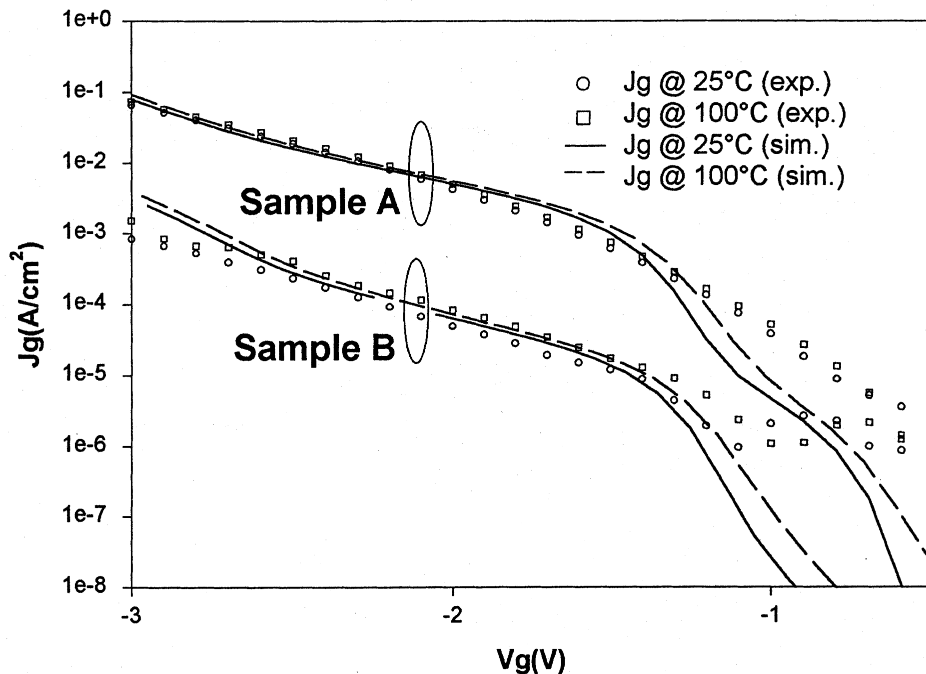


Fig. 8. Comparison of simulation and experiment for gate currents at different temperatures for samples A and B.

$\Delta E_c = 1.6$ eV and $m_c = m_v = 0.32 m_e$ are used in the Franz dispersion for the HfO_2 layer and $E_g = 4.5$ eV, $\Delta E_c = 1.2$ eV and $m_c = m_v = 0.5 m_e$ for the interfacial layer.

IV. Summary and Conclusions

A gate current model has been developed taking into account the quantum confinement effects in the silicon channel, direct and FN tunneling and thermionic emission transport through the gate dielectric. Both gate-injected or substrate-injected currents are modeled for the silicon conduction-band and valence-band components. Subject to the energy dispersion relation in each region of the gate-dielectric-silicon system and available carriers and empty states in either side of the dielectric, the gate current is determined. The energy dispersion in the dielectric band gap is approximated by Franz dispersion. The subband structures and carrier distribution in energy and position in the silicon channel are obtained by solving the Schrödinger and Poisson

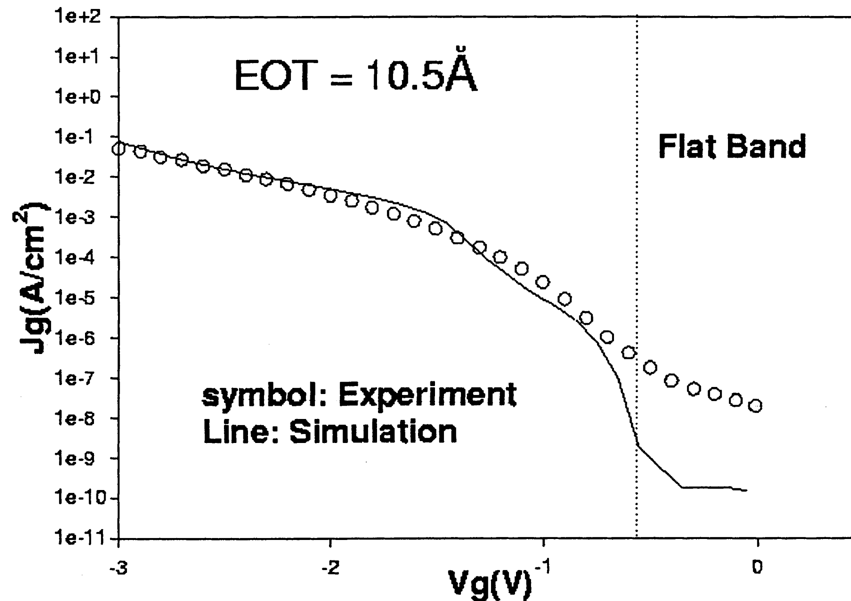


Fig. 9. Comparison of gate current simulation with experimental data for a TaN/HfO_2/p-Si capacitor.

equations self-consistently, both for gate capacitance and gate current calculations. A self-consistent C_g, I_g –V_g model thus is established. This model was validated using SiO_2 dielectric devices and good agreement with experiment is achieved.

Device structure information of ZrO_2 NMOSCAPs is extracted in accumulation using C_g –V_g and I_g –V_g simulation. The simulation agrees well with the experiment for different thicknesses of dielectric stack, using consistent parameters. The extracted band gaps and band offsets with silicon are comparable with those that have been reported. The gate current simulation of a HfO_2 NMOSCAP also agrees well with the experiment.

Characteristics of gate current transport mechanism of ZrO_2-NMOSCAPs were studied. The temperature-dependence study shows that the gate current is primarily contributed from the silicon conduction band and tunneling is the most likely primary transport mechanism. However, other tunneling processes such as trap-assisted tunneling may be possible. More temperature-dependent processes are also not completely excluded, but their effects are weak. Interface states between different regions might significantly affect the gate capacitance and gate current at low voltages. Kinks in gate capacitance-voltage and oscillations of gate currents were observed. The gate current does not change much after the device is stressed electrically and thermally. This indicates good quality of the film and few charges or traps created in the dielectric stack as a result of the stress.

ACKNOWLEDGMENT

The authors would like to thank Prof. D.-L. Kwong and his students at the University of Texas at Austin for the discussion about their high-K work.

REFERENCES

[1] *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Association (SIA), Austin, TX, 1999.
[2] S. Thompson, P. Packan, and M. Bohr, "MOS scaling: Transistor challenges for the 21st century," *Intel Technol. J.*, vol. Q398, 1998.
[3] C. Hu, "Gate oxide scaling limits and projection," in *IEDM Tech. Dig.*, 1996, pp. 319–322.
[4] W. Qi, R. Nieh, B. H. Lee, L. Kang, Y. Jeon, K. Onishi, T. Ngai, S. Banerjee, and J. C. Lee, "MOSCAP and MOSFET characteristics using ZrO_2 gate dielectric deposited directly on Si," in *IEDM Tech. Dig.*, 1999, pp. 145–148.
[5] C. H. Lee, Y. H. Kim, H. F. Luan, S. J. Lee, T. S. Jeon, W. P. Bai, and D. L. Kwong, "MOS devices with high-quality ultrathin CVD ZrO_2 gate dielectrics and self-aligned TaN and TaN/poly-Si gate electrodes," in *Proc. VLSI*, 2001, pp. 137–138.
[6] C. Hobbs, H. Tseng, K. Reid, B. Taylor, L. Dip, L. Hebert, R. Garcia, R. Hedge, J. Grant, D. Gilmer, A. Franks, V. Dhandapani, M. Azrak, L. Prabhu, R. Rai, S. Bagchi, J. Conner, S. Backer, F. Dumbuya, B. Nguyen, and P. Tobin, "80 nm poly-Si gate CMOS with HfO_2 gate dielectric," in *IEDM Tech. Dig.*, 2001, pp. 30.1_1–30.1_4.
[7] J. A. Duffy, *Bonding Energy Levels and Bands in Inorganic Solids*. New York: Wiley, 1990.
[8] S.-H. Lo, A. Buchanan, Y. Taur, and W. Wang, "Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFETs," *IEEE Electron Device Lett.*, vol. 18, pp. 209–211, May 1997.
[9] N. Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, "Modeling study of ultrathin gate oxides using direct tunneling current and capacitance-voltage measurements in MOS devices," *IEEE Trans. Electron Devices*, vol. 46, pp. 1464–1471, July 1999.
[10] F. Rana, S. Tiwari, and D. A. Buchanan, "Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides," *Appl. Phys. Lett.*, vol. 69, pp. 1104–1106, Aug. 1996.
[11] M. V. Fischetti and S. E. Laux, "Monte Carlo study of electron transport in silicon inversion layers," *Phys. Rev. B, Condens. Matter*, vol. 48, pp. 2244–2274, July 1993.
[12] "SRC/SEMATECH FEP Center Review," North Carolina State Univ., Raleigh, N.C., 2001.
[13] W. Harrison, "Tunneling from an independent-particle point of view," *Phys. Rev.*, vol. 123, pp. 85–89, July 1961.
[14] Z. A. Weinberg, "On modeling oxide tunneling current," *J. Appl. Phys.*, vol. 53, pp. 5052–5056, July 1982.
[15] S. Jallepalli, J. Bude, W.-K. Shih, M. R. Pinto, C. M. Maziar, and A. F. Tasch, "Electron and hole quantization and their impact on deep submicron silicon p- and n-MOSFET characteristics," *IEEE Trans. Electron Devices*, vol. 44, pp. 297–303, Feb. 1997.

[16] W. Franz, *Handbych der Physik*, 1956, vol. 17, p. 155.
[17] G. Lewicki and C. A. Mead, "Experimental determination of E-k relationship in electron tunneling," *Phys. Rev. Lett.*, vol. 16, pp. 939–941, May 1966.
[18] J. Maserjian and G. P. Petersson, "Tunneling through thin MOS structures: Dependence on energy (E-κ)," *Appl. Phys. Lett.*, vol. 25, pp. 50–52, July 1974.
[19] S.-H. Lo, D. A. Buchanan, and Y. Taur, "Modeling and characterization of quantization, polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides," *IBM J. Res. Develop.*, vol. 43, pp. 327–337, 1999.
[20] E. Cassan, P. Dollfus, S. Galdin, and P. Hesto, "Semiclassical and wave-mechanical modeling of charge control and direct tunneling leakage in MOS and H-MOS devices with ultrathin oxides," *IEEE Trans. Electron Devices*, vol. 48, pp. 715–721, Apr. 2001.
[21] W.-K. Shih, E. X. Wang, S. Jallepalli, F. Leon, C. M. Maziar, and A. F. Tasch, "Modeling gate leakage current in nMOS structures due to tunneling through an ultrathin oxide," *Solid State Electron.*, vol. 42, no. 6, pp. 997–1006, 1998.
[22] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*. New York: Pergamon, 1977.
[23] T. Yamaguchi, H. Satake, N. Fukushima, and A. Toriumi, "Band diagram and carrier conduction mechanism in ZrO_2/Zr-silicate/Si MIS structure fabricated by pulsed-laser-ablation deposition," in *IEDM Tech. Dig.*, 2000, pp. 19–22.
[24] J. Frenkel, "On pre-breakdown phenomena in insulators and electronic semi-conductors," *Phys. Rev.*, vol. 43, pp. 647–648, 1938.
[25] *UTQUANT Manual*, Univ. of Texas at Austin.

Yang-Yu Fan received the B.S. degree in physics from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1994, and the M.S. degree in electrical and computer engineering from the University of Texas at Austin in 1999. He is currently pursuing the Ph.D. degree in electrical and computer engineering at the Microelectronics Research Center, University of Texas at Austin. His research interest is in the area of submicron device modeling and simulation. He is currently conducting research on high-K gate stack and nonvolatile memory modeling.

Renee E. Nieh received the B.S. degree in engineering science from Trinity University, San Antonio, TX, in 1997, and the M.S. degree in electrical engineering from the University of Texas at Austin in 1999. She is currently pursuing the Ph.D. degree in electrical engineering at the Microelectronics Research Center, University of Texas at Austin. Her current research interest is in the area of high-K gate dielectrics, particularly the electrical and material evaluation and process intergration of ZrO2- and Zr-based dielectric stacks.

Jack C. Lee (S'85–M'88–SM'93–F'01) received the B.S. and M.S. degrees in electrical engineering from the University of California, Los Angeles, in 1980 and 1981, respectively, and the Ph.D. degree in electrical engineering from the University of California, Berkeley, in 1988.

Currently, he is Professor with the Electrical and Computer Engineering Department and holds the Cullen Trust For Higher Education Endowed Professorship in Engineering at The University of Texas at Austin. From 1981 to 1984, he was a Member of the technical staff at the TRW Microelectronics Center in the High-Speed Bipolar Device Program. He worked on bipolar circuit design, fabrication, and testing. In 1988, he joined the faculty of The University of Texas at Austin. His current research interests include thin dielectric breakdown and reliability, high-K gate dielectrics and gate electrode, high-K thin films for semiconductor memory applications, electronic materials and semiconductor device fabrication processes, characterization, and modeling. He has published over 250 journal publications and conference proceedings.

Dr. Lee has been awarded two Best Paper Awards, numerous teaching awards, and several patents.

Gerry Lucovsky, photograph and biography not available at the time of publication.

George A. Brown (M'64) received the B.S.E.E. degree from the University of Pennsylvania, Philadelphia, in 1959, and the M.S.E. degree from Princeton University, Princeton, NJ, in 1961.

Currently, he is with International SEMATECH, Inc., Austin, TX, as a Member of the Front End Processes Research Center. He retired from Texas Instruments, Inc., in April 2001 as a Distinguished Member, Technical Staff, with 37 years of service. At SEMATECH, he is involved in the electrical and reliability characterization of high-K gate stack materials and as an Industrial Resident for the University of Texas at Austin FEP research projects. Prior to joining Texas Instruments, he was with RCA Laboratories, Princeton, NJ, where he was involved in early development of the MOS transistor device. His primary interests lie in the areas of processing, characterization and reliability of insulator-silicon interface structures and devices.

Prof. Brown is a Member of the Electrochemical Society and is a Section Chairman of the ASTM.

Leonard Frank Register received the B.S. degrees (summa cum laude) in electrical engineering and in physics and the Ph.D. degree in electrical and computer engineering, all from North Carolina State University, Raleigh.

He was a Visiting Assistant Research Professor and then a Research Scientist in the Beckman Institutes Computational Electronics Group at the University of Illinois at Champaign-Urbana before recently joining The University of Texas at Austin as an Assistant Professor in the Microelectronics Research Center. He is a Device Theorist and has published on diverse subjects including quantum transport, high-energy transport, transport through oxides, optical fields of vertical-cavity surface-emitting lasers, single electronics, and device degradation.

Sanjay K. Banerjee (S'80–M'83–SM'89–F'96) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, and the M.S. and Ph.D. degress from the University of Illinois at Urbana-Champaign in 1979, 1981, and 1983, respectively, all in electrical engineering.

He is the Cockrell Family Regents Chair Professor of Electrical and Computer Engineering and Director, Microelectronics Research Center, at The University of Texas at Austin. As a Member of the Technical Staff, Corporate Research, Development, and Engineering of Texas Instruments, Inc., from 1983 to 1987, he worked on polysilicon transistors and dynamic random access trench memory cells used by Texas Instruments in the world's first 4–Mb DRAM, for which he received the Best Paper Award, IEEE International Solid State Circuits Conference, 1986. He has been Assistant Professor (1987–1990), Associate Professor (1990–1993), and Professor (1993-present). He has more than 400 archival refereed publications/talks, five books/chapters, and over 20 U.S. patents. He has supervised over 25 Ph.D. and 45 M.S. students. He is currently active in the areas of ultrahigh vacuum and remote plasma-enhanced chemical vapor deposition for silicon-germanium-carbon heterostructure MOSFETs and nanostructures. He is also interested in the areas of ultrashallow junction technology and semiconductor device modeling.

Dr. Banerjee received the Engineering Foundation Advisory Council Halliburton Award (1991), the Texas Atomic Energy Fellowship (1990–1997), Cullen Professorship (1997–2001), and the NSF Presidential Young Investigator Award (1988). His recent awards include the IEEE Millennium Medal (2000) and the SRC Inventor Recognition Award (2000). He is a Distinguished Lecturer for IEEE Electron Devices Society and was the General Chair of the IEEE Device Research Conference, 2002.