Modeling of Gate Current and Capacitance in Nanoscale-MOS Structures

J. P. Sun, Wei Wang, Toru Toyabe, *Fellow, IEEE*, Ning Gu, and Pinaki Mazumder, *Fellow, IEEE*

***Abstract*—By applying a fully self-consistent solution of the Schrödinger–Poisson equations, a simple unified approach has been developed in order to study the gate current and gate capacitance of nanoscale-MOS structures with ultrathin dielectric layer. In this paper, the model has been employed to investigate various gate structure and material combinations, thereby demonstrating wide applicability of the present model in the design of nanoscale-MOSFET devices. The results obtained by applying the proposed model are in good agreement with experimental data and previous models in the literature. A new result concerning optimum nitrogen content in HfSiON high-k gate-dielectric structure reported in this paper requires experimental verification through device fabrication.**

***Index Terms*—Gate current, high-k dielectric, nanoscale MOSFETs, quantum modeling.**

I. Introduction

THE NEED for research in material, device structure, and assembly method for electronic devices with feature sizes smaller than 20 nm is crucial for Moore's law of exponential growth of semiconductor integrated circuits to continue for at least another two decades [1]. To understand the behavior of sub-100-nm devices, numerous studies have been conducted in the literature on modeling of gate current and gate capacitance of nanoscale MOSFETs. The evaluation of gate current due to tunneling through the thin-gate oxide between polysilicon gate and substrate has attracted special attention, since it represents one of the major leakage-current components in nanoscale MOSFETs. Control of gate-leakage current is paramount in low-power CMOS-circuit design.

The gate current transport across the oxide–substrate interface under a positive gate bias can be divided into the following five mechanisms, as shown in Fig. 1.

1) Thermionic emission of hot electrons which gain sufficient kinetic energies from the gate and drain bias fields.
2) Fowler–Nodheim (FN) tunneling through a triangular potential barrier at the top of the oxide-conduction band.
3) Direct tunneling due to electrons tunneling through an approximate trapezoidal potential barrier.

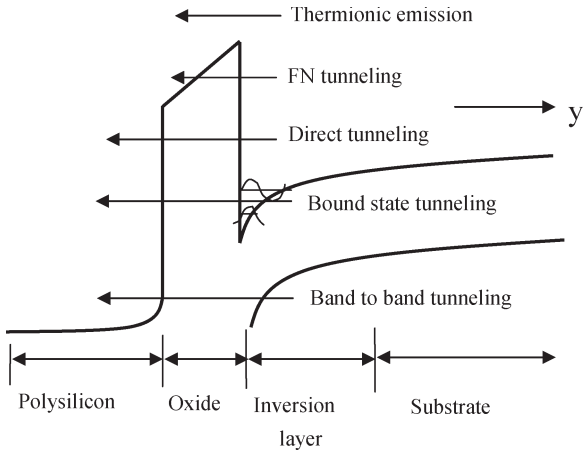


Fig. 1. Schematic gate-current components between the polysilicon gate and semiconductor substrate in nanoscale MOSFETs.

4) Tunneling of electrons residing in the subbands of strongly inverted surface of the P-type substrate into or through the gate oxide.
5) Band-to-band tunneling of electrons from the valence band of substrate to conduction band of polysilicon gate at high gate voltages.

Many models for the gate current have been developed in the past, with various degrees of sophistication and success in their agreement with the experimental results. For the gate-current evaluation, either the Wentzel–Kramer–Brillouin (WKB) approach or the Schrödinger equation was used to calculate the direct-tunneling current components coupled with Poisson's equation [2]–[12]. Efforts have been devoted to calculating the electron distribution available for tunneling such as a displaced Maxwellian, or from a solution of the Boltzmann equation, or based on a Monte Carlo simulation to account for the hot-electron effects. However, most of these models involve complicated formulas and, in many cases, fitting parameters.

We consider that the separation of the current-transport components is a reflection of the historical development of the MOS-scaling efforts toward nanoscale dimensions. In principle, tunneling depends on the overall potential profile that electrons encounter rather than a particular part of the potential profile. Therefore, the separation and delineation of various gate-current components may involve inconsistencies in the interfacing and integration of these components, thereby making the adjustments and fittings necessary to achieve agreement between analytical models and experimental results. In addition, the gradual potential-barrier approximation inherent in the WKB approach may become improper for the sharp

Manuscript received January 24, 2006; revised August 31, 2006. This work was supported in part by the Office of Naval Research Grant and in part by the National Science Foundation Grant. The review of this paper was arranged by Editor S. Datta.

J. P. Sun and P. Mazumder are with the University of Michigan, Ann Arbor, MI 48109-2122 USA.

W. Wang and N. Gu are with Southeast University, Nanjing 210096, China.

T. Toyabe is with the Bio-Nano Electronics Research Centre, Toyo University, Kawagoe 350-8585, Japan.

Digital Object Identifier 10.1109/TED.2006.885637

potential interfaces between the polysilicon and substrate for nanoscale MOSFETs. In our opinion, as the feature sizes approach nanoscale dimensions, the device structure in the direction of interest may well be treated as an integrated open-quantum system. A fully self-consistent solution to the Schrödinger–Poisson equations is sufficient to model all the gate current transport components and obtain all the device-terminal properties of interest.

In this paper, a simple unified model has been developed for modeling gate current and capacitance of nanoscale-MOS structures based only on the first principles without having to use a number of approximations and fitting parameters. We solve the effective mass Schrödinger equation and Poisson's equations in full self-consistency under various gate-bias conditions, using the Fermi–Dirac distribution for the carrier-supply function at the device electrodes to model all the gate-current components listed above. The self-consistent charge and potential distributions also give gate capacitance under various gate voltages by simply applying the definition of differential capacitance. The present approach is computationally efficient and takes only a couple of hours to obtain the necessary solutions on a regular PC. Our gate-current and capacitance models agree with the experimental data and results of other approximate models. We have also calculated the gate current using the WKB approach to establish that it is in good agreement with the present model.

An additional merit of the present model lies in its capability of modeling various combinations of high-k dielectric materials and structures, since the material and structural parameters of individual layers in the high-k stacks can readily be specified accurately. This feature will help carry out the tasks for significant engineering of high-k stack structures and silicon–dielectric interface in designing nanoscale MOSFETs.

In Section II, we briefly list the main formulations of the present model and then discuss physical considerations that make present model useful. The calculated results of the gate current and capacitance under various bias conditions for various high-k dielectric structures are presented in Section III. A wide applicability of our model and its comparisons with experimental results and other approximate models are also discussed.

II. Description of Model

The present model is based on a self-consistent solution of the Schrödinger–Poisson equations previously developed for modeling of quantum-tunneling devices such as the resonant-tunneling diodes and transistors [13], [14], with the modifications made for nanoscale-MOS structures in the y direction, as shown in Fig. 1. The five gate-current components are calculated by sending a spectrum of electrons from the device electrodes with their energy distributions governed by the Fermi statistics. Sufficiently high-energy electrons are included to ensure that the hot-electron effects are modeled especially at larger gate-bias voltages. The quasi-bound states in the inversion-layer quantum well on the substrate side are fully included, and a proper interfacing mechanism is implemented between the three-dimensional (3-D) current components originated from the electrodes and two-dimensional (2-D) component originated from the inversion layer.

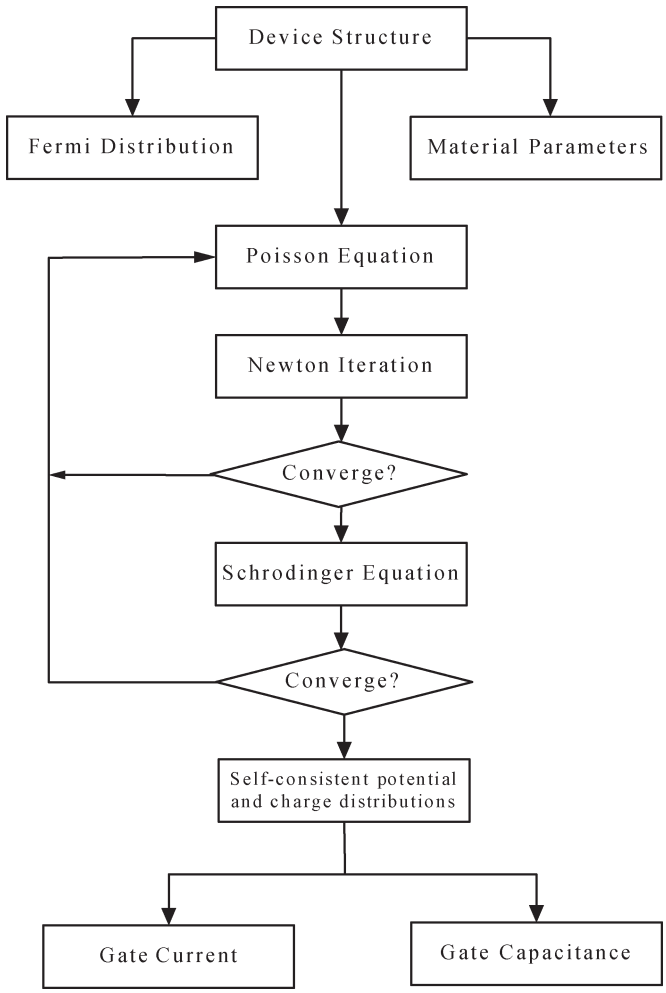


Fig. 2. Flowchart of modeling nanoscale-MOS structures.

The 2-D current transport is treated by calculating the transmission of electrons residing in the inversion quantum well at all possible quasi-bound states found in the quantum well under various bias voltages. Open-boundary conditions, based on the quantum-transmitting-boundary method, are employed for various current evaluations. The overall computational aspects are summarized in a flowchart, as shown in Fig. 2. The main formulas used in the calculation are

$$\frac{\partial^2 V(y)}{\partial y^2} = -\frac{q}{\varepsilon(y)}\left[N_D(y) - N_A(y) - n(y) + p(y)\right] \quad (1)$$

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial y}\left(\frac{1}{m^*(y)}\frac{\partial \psi(y)}{\partial y}\right) + E_C(y)\psi(y) = E\psi(y). \quad (2)$$

The self-consistent charge at the jth energy subband in the ith valley in the inversion layer is given by

$$n_{ij}(y) = \frac{m_i^* k_B T}{\pi \hbar^2} \ln\left[1+\exp\left(-\frac{E_{ij}-E_F}{k_B T}\right)\right] \times |\psi_{ij}(E_{ij}, y)|^2 . \quad (3)$$

The gate-current components of thermionic emission, FN tunneling, and direct tunneling through the oxide barrier are

calculated as a whole by using traveling-wave calculation, hereafter referred to as J_{3D}. The tunneling component from the inversion-layer quantum well is calculated using a transmission calculation, hereafter referred to as J_{2D}. All currents are calculated by solving the Schrödinger equation, self-consistently with its potential term determined from the Poisson equation.The gate current is then given by

$$J_{3D} = -q\hbar \sum_k W(k) \mathrm{Im}\left(\psi_k^*(y) \frac{1}{m^*(y)} \frac{\partial \psi_k(y)}{\partial y}\right). \quad (4)$$

We consider the inclusion of the thermioinc emission, FN tunneling, and direct-tunneling components in a unified treatment viable and convenient, although it may not be as accurate as those approaches where special efforts are made to model the hot-electron distribution, as in [7] and [8]. However, by including a large number of electron wave vectors in the Fermi distribution, which grows with increasing positive gate-bias voltages, it gives the corresponding portions of the electrons available for these components relative to all the parts of the oxide barrier above the conduction-band edge. This approximation allows us to treat those transport components in a unified fashion, thereby avoiding the separation and delimitation among calculations of those components, and the use of other approximations and fitting parameters as in many other approximate models. In nanoscale MOSFETs, the limit of the oxide thickness is considered to be set by direct tunneling rather than by hot-carrier injection. When oxide thickness decreases further, the power-supply voltage is scaled down; hence, hot-carrier effects become insignificant by decreasing in an exponential fashion in comparison to direct-tunneling components.

For the electron-tunneling current from the inversion layer into the oxide and gate electrode, a transmission calculation is performed. Based on the wave functions calculated in the inversion layer, the transmission is evaluated by the square ratio of the transmitted electron flux to the incident flux:

$$T_{ij} = \frac{|\mathrm{C}_{tr,ij}|^2}{|\mathrm{A}_{in,ij}|^2} \frac{k_{tr,ij}}{k_{in,ij}} \frac{m^*_{in,i}}{m^*_{tr,i}} \quad (5)$$

where $\mathrm{C}_{tr,ij}$ is the transmitted wave amplitude in the ith valley and the jth subband, and $\mathrm{A}_{in,ij}$ the incident wave amplitude in the corresponding subband and valley, respectively. The 2-D gate-current component originating from the subbands in the inversion layers is given by

$$J_{2D} = \sum_{i,j} J_{i,j} = q \sum_{i,j} n_{e,ij} \mathrm{T}_{ij} f_{ij} \quad (6)$$

where T_{ij} is the electron-transmission probability, $f_{ij} = E_{ij}/j\pi\hbar$ is the interface-impact frequency, $n_{e,ij}$ is the electron concentration, and E_{ij} is the quasi-bound state energies, in the ith valley and the jth subband, respectively. The total gate-current density is then the sum of the 2-D and 3-D components.

In all the calculations for the electron-wave functions, the quantum-transmitting-boundary conditions are employed, as in [15] and [16]. We also note that electron inelastic scatterings are not modeled in the present approach since (2) is a pure-state Schrödinger equation in the single-electron picture, where the electron energy must be conserved. Fig. 3 shows an example of the calculated self-consistent potential and electron-wave functions, and it illustrates that the current-transport components originated from the 3-D electrodes and 2-D quantum well in the inversion layer.

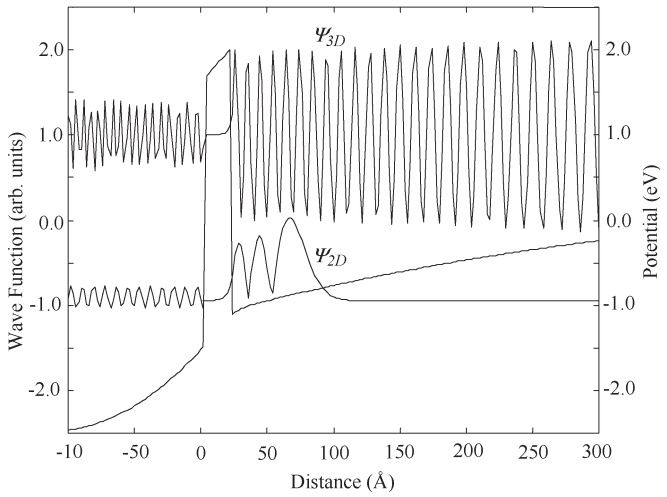


Fig. 3. Electron wave functions from the 3-D and 2-D electron states. The magnitude of the wave functions on the gate side is scaled suitably to illustrate the wave behavior of the transmitting electrons. Only the third quasi-bound state in the inversion layer is shown.

For the purpose of comparison, we have also performed a gate-current calculation based on a WKB approach similar to that in [17] and [18]. However, in our transmission calculations, the field and energy dependencies on the distance are directly obtained from the solution of Possion's equation, instead of the empirical formulas used by other authors.

Previous work on the gate capacitance of submicrometer- and deep-submicrometer-MOS devices has developed numerical and analytic models based on the oxide and inversion-layer components. The analytical results have been found to be in good agreement with the experimental data [19]–[22]. Influences of quantum effects such as the carrier depletion in the polysilicon and inversion-layer charge have been quantified to make a reduction of the gate capacitance. In this paper, as an integrated part of the self-consistent solution to the Schrödinger–Poisson equations, the charge distribution in the y direction over the entire device structure is available under different bias voltages. Therefore, the gate capacitance is calculated taking into account of the quantum effects. The gate capacitance is obtained by taking the derivative of the gate charge with respect to the gate-bias voltages. The readers are referred to [20] and [22] for more details.

III. Results and Discussions

The present model was used to perform studies of gate-current behavior for various MOS structures of nanometer dimensions, including different high-k gate-dielectric materials and their combinations. Effects of dielectric thickness, interfacial and transition layers (IL), as well as that of nitrogen incorporation are included in this paper to demonstrate wide applications of the present approach.

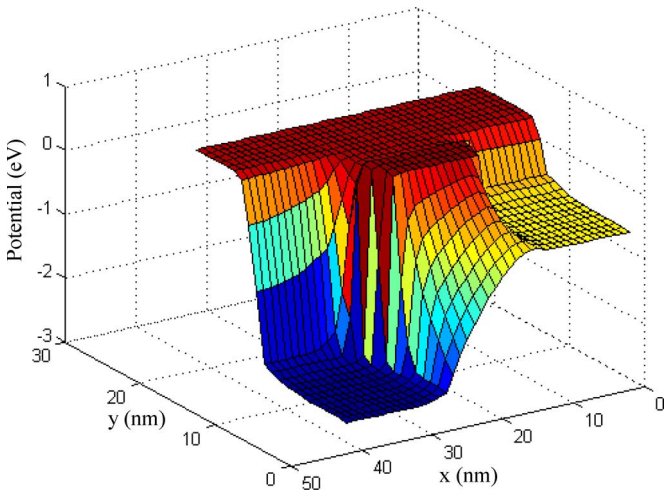


Fig. 4. Electron potential, in 2-D, of the MOS structure under study at a gate-bias voltage of 2.0 V. The x direction is along the channel, and the y direction is vertical to the gate dielectric. (Color version available online at http://ieeexplore.ieee.org.)

A 2-D MOSFET simulator has been developed and used to find a suitable doping profile that will suppress subthreshold leakage current for the nanoscale MOSFETs under this study. Subthreshold leakage current is calculated for a combined retrograde and halo-doping profile as compared to the case of the simple doping. It is found that the realistic doping profile will result in lower subthreshold current. We will not discuss details of subthreshold leakage in this paper. For the gate-leakage-current evaluation, Fig. 4 shows a calculated 2-D realistic potential profile used for the gate-current calculations. However, under the present one-dimensional (1-D) model, our results show that the effects of doping profiles on the gate-leakage current are insignificant as compared to that using a uniform doping of 1.0×10^{17} cm^{-3}. Therefore, the features of the realistic doping profile are either not in 1-D tunneling path or have insignificant influence in comparison to other tunneling parameters. In this paper, we will not discuss the doping effects in further detail. A 2-D or 3-D leakage model would be desired for correct assessment of effects of doping profiles on the gate-leakage current.

The device structure under study has the oxide layer thickness running from 16 Å to 36 Å for various calculations, with an oxide barrier height of 3.15 eV. The electron effective mass is taken to be 0.5 m_0 in the oxide layer, 0.98 m_0 for the twofold longitudinal subband valley, and 0.19 m_0 for the fourfold transverse valley. All the calculations are performed for the device structure at room temperature (300 K) with the above material parameters without adjustment or fittings for a particular calculation.

A. Effects of Dielectric Thickness and Comparison With Experiment

The calculated gate-tunneling-current densities as a function of gate-bias voltage for various gate oxide thickness ranging from 16 Å to 36 Å are shown in Fig. 5. They are in overall good agreement with the experimental data taken from [6] and other modeled results [4], [6], [9].

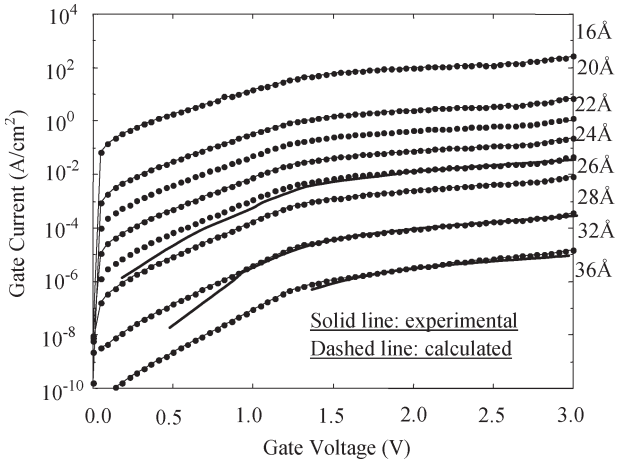


Fig. 5. Gate current versus gate voltage with the oxide thickness of 16 Å to 36 Å. The dotted lines are calculated using the present model, and the solid lines are experimental data taken from [6].

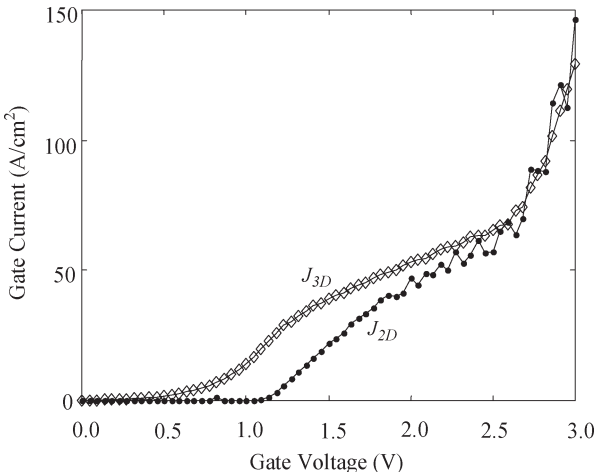


Fig. 6. Comparison of the 3-D and 2-D gate-current components. The oxide thickness of the device structure is 16 Å. The current oscillations occur at about 2.0 V gate bias and above, where the incident electrons encounter strong multiple quantum reflections at and above the top of the oxide barrier.

The following observations can be made from the modeled and measured gate current.

1) The two gate-current components, J_{3D} and J_{2D}, as plotted in Fig. 6, are combined into the total gate-tunneling current. It is shown that the overall agreement between our model and the experiment is achieved except at the low gate-bias voltage below about 0.2 V. We note that many second-order effects and nonidealities are not included in our model, such as the interface charges, oxide charges, and other barrier lowering effects and some 3-D effects.
2) The 2-D current component originated in the inversion quantum well is much smaller than the 3-D tunneling current at low gate bias. This result may not be valid for realistic MOSFETs where the inversion electrons come from drain and source electrodes. In the present approach, they can only be assumed from the bulk, as effects of electron transport in other (x and z) dimensions are not modeled in the 1-D approach. We note that both the 2-D and 3-D gate-current components depend mainly on the

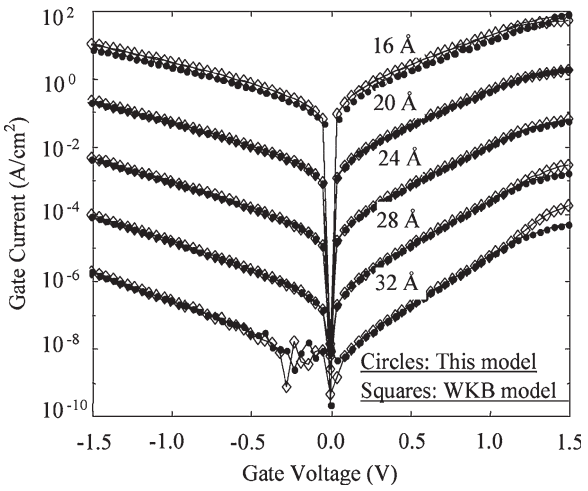


Fig. 7. Comparison of gate-current calculations using present model and the WKB approximation. The circles are from the present model, whereas the squares are calculated using our WKB model.

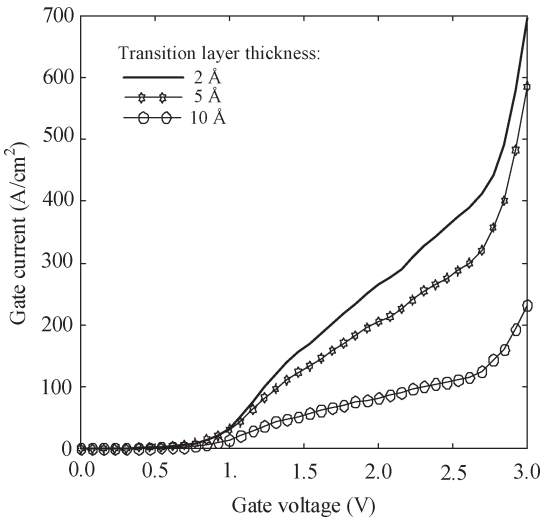


Fig. 8. Calculated gate current versus gate voltage for high-k structures with transition layer between the high-k dielectric and silicon-dioxide layers.

barrier height the 2-D and 3-D electrons encounter and the available 2-D and 3-D electrons at a particular bias. At low gate bias, the inversion electrons may be much more in number than the higher energy 3-D electrons as the 2-D electrons fill up first to the lower energy quasi-bound states. On the other hand, the 2-D electrons encounter a much higher barrier than the fewer 3-D electrons. The result depends on the interplay between the above factors and may not be conclusive for realistic MOSFETs for which a 3-D open-quantum model would be required to model these current components more accurately. At increased gate bias, more inversion electrons will tunnel through the lower barrier. The 3-D and 2-D current components become comparable in quantity when the strong inversion occurs. Eventually, the 2-D current component can become significantly larger than the 3-D tunneling current as shown in Fig. 6, where the structure under study has an oxide thickness of 16 Å. Note that current oscillations occur at about 2.0 V gate bias and above, where the incident electrons encounter multiple quantum reflections at and above the top of the oxide barrier.

3) In our results, it is also observed that a maximum tolerable gate-current limit of approximately 1 A/cm^2 is related to the oxide thickness of 20 Å due to the standby-power requirement. When the oxide thickness is below that limit, the gate-tunneling-current density increases more rapidly as the exponential dependence of the tunneling current on the barrier thickness becomes more pronounced.

The calculated gate current using the WKB approach is shown in Fig. 7 together with the corresponding results obtained from our model. The comparison of the two approaches presents very good agreement. However, for a thicker oxide barrier, both models show some oscillations and kinks in the flatband region. It is believed that at small gate bias and for thick barrier, the fewer tunneling electrons experience multiple reflections at and within the barrier, thereby resulting in the current fluctuations. Note that the WKB model itself does not generate the capacitance model, while the present model provides not only the gate current but also the gate capacitance, simultaneously.

B. Effects of IL in High-k Structures

High-k gate-stack structures as possible candidates to replace silicon-dioxide layer for nanoscale MOSFETs have been of great interest very recently due to their promise in reduction of gate current. A number of high-k structures have been proposed and studied. However, significant engineering of high-k stack structures and silicon-dielectric interface are required due to the complications associated with the compatibility of the materials and processing of the high-k stacks with silicon technologies. It has been found that, in order to maintain good interface quality between the high-k layer and silicon substrate, an interfacial oxide layer is present, and a transition layer between high-k dielectric and silicon substrate may exist [23]. Furthermore, combinations of various high-k dielectric materials with different dielectric constant and other material properties need to be investigated for their effects on the gate current and capacitance. The present approach is particularly suitable for analysis of high-k structures consisting of multiple layers of different materials with interface layers, since all the material parameters in the individual layers can readily be specified with accuracy.

In most of the previous studies of high-k stack structures, an abrupt transition between the high-k dielectric layer and IL was assumed. However, in practice, a gradual transition may exist. Effects of IL region on gate-current performance have been exploited theoretically using the WKB approximation [23]. In this paper, the present modeling tool is applied to a high-k stack structure with different transition thickness. Fig. 8 shows the calculated gate current versus gate voltage of Si_3N_4–transition–SiO_2 structures with an effective oxide thickness (EOT) of 15 Å and an oxide thickness of 5 Å.

The comparison between the calculated gate current for different transition thickness is illustrated in Fig. 8 to show

the effects of transition layer. In this case, the gate current is reduced by the addition of a transition layer and with increasing thickness of the transition layer for the same EOT. However, it is noted that the effect of gate-current reduction with transition layer cannot be generalized. The magnitude of gate current in high-k stack structures with transition layer depends, in general, on the interplay between various factors of the structure and material parameters like the barrier height, electron effective masses, dielectric constant, as well as proportion of the individual layers. Therefore, numerical simulations may be required to achieve an optimum design for a practical case. These effects have been discussed more extensively in [23] and will not be elaborated further here.

C. Effects of Nitrogen Incorporation

Incorporation of nitrogen at the dielectric interface suppresses dopant diffusion from gate polysilicon into the channel, which can cause a shift of threshold voltage. Since it is desirable to use highly doped p-type polysilicon gate for p-MOS transistors in the scaled-CMOS technology to minimize short-channel effects, the problem is more serious as boron diffuses more readily into oxide film. It is reported that incorporation of nitrogen into hafnium silicate (Hf/SiO), hafnium aluminate (Hf/AlO), and HfO_2 greatly enhances the dielectric constant of silicates, suppresses dopant diffusion from gate polysilicon into the channel during high-temperature annealing process, and increases crystallization temperature of the high-k stacks [24]. These properties are ascribed to the homogeneity of the bond structure in the film containing nitrogen through high-temperature annealing. These improvements make these Hf-based materials more suitable for the CMOS process. The increase of nitrogen content leads to an increase of the dielectric constant but a decrease of the offset in both conduction and valence bands. It is interesting to observe that as the only exception among most high-k dielectrics, the conduction-band offset with nitrogen incorporation is larger than the valence-band offset as in the case of Hf/SiO/N. It is an advantage for the reduction of gate current in n-MOSFETs. For the p-MOSFETs, however, such property would not be useful since the hole tunneling is dominant for the gate current.

In Fig. 9, the effects of nitrogen composition on the gate current are shown for tunneling current through Hf/SiO/N dielectric film. Three different electron effective masses in between that of SiO_2, HfO_2, and Si_3N_4 were used in the calculations as the values of the electron effective masses are not available, and also, such modeling calculations have not been found in the open literature to date. On the other hand, the values of dielectric constant and barrier height were taken from the experimental work with the nitrogen content up to 40%, for the curves of Hf/(Hf + Si) being 60%, and up to 35% for the rest of the curves [25]. The results here show that, for both cases, there is an optimum value of the nitrogen content (approximately 10%), where a minimum gate current may be achieved. This is attributed to the interplay between the dielectric constant and barrier height. The incorporation of nitrogen leads to an increase of the dielectric constant while maintaining essentially the same barrier height until the nitrogen content reaches 10%, where the barrier height starts decreasing significantly. This expected optimum value of the nitrogen content is a new finding from this study, and it needs to be verified experimentally. It is also noted that the gate current in the p-MOSFET is larger than that of the n-MOSFET with the same EOT and nitrogen content. This is because the valence-band offset in the p-MOSFET is smaller than the conduction-band offset in the n-MOSFET when the dielectric constants are virtually the same.

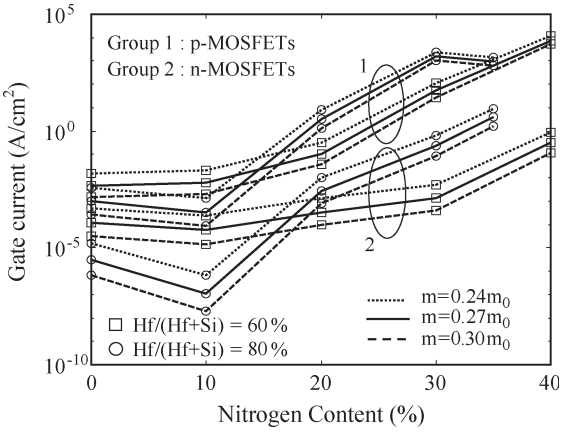


Fig. 9. Gate-tunneling current through Hf/SiO/N dielectric film of different N content. The EOT is 1.1 nm. The effect of $\pm 10\%$ variation of Hf/SiO/N electron and hole effective-mass value $(0.27\ m_0)$ on gate current is demonstrated in this figure. We use $m = 0.30\ m_0$ for all the dashed lines, $m = 0.27\ m_0$ for all the solid lines, and $m = 0.24\ m_0$ for all the dot-dashed lines. Group 1 curves are for p-MOSFETs, and group 2 curves for n-MOSFETs. The circle symbols indicate a ratio of Hf/SiO/N to Hf/Hf + Si of 60%; the square symbols indicate a ratio of Hf/SiO/N to Hf/Hf + Si of 80%. The values of dielectric constant and barrier height were taken from [19].

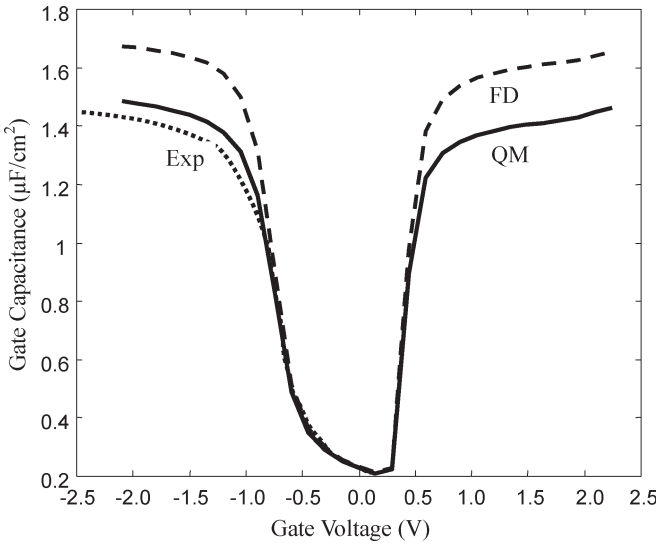


Fig. 10. Comparison of calculated gate capacitance with experimental result for a MOSFET of 20-Å oxide layer. The experimental data were taken from [5].

D. Gate Capacitance

The calculated results of the gate capacitances of the MOS structure with the gate oxide thickness of 24 Å as a function of the gate-bias voltage are shown in Fig. 10. The comparison

with the experimental data shows very good agreement with the well-known variations with respect to the bias voltages ranging from negative gate bias to positive gate bias, corresponding to the hole accumulation to the electron inversion in the inversion layer. Both the classical capacitance and the quantum capacitance are shown. The classical capacitance is calculated using the Fermi statistics, whereas the quantum capacitance is modeled using the present approach described previously in Section II. The measured experimental capacitance is slightly smaller than the quantum capacitance for larger bias voltages. The results obtained from the present model show very similar behavior as compared to the other approximate models [4], [5].

IV. Conclusion

A simple unified approach of gate and capacitance models for nanoscale-MOS structures with ultrathin oxide thickness is presented in this paper. This approach uses only a fully self-consistent solution to the Schrödinger–Poisson equations, and it has the features of being conceptually simple and computationally efficient. In the present model, the commonly used WKB approximation for the gate-current evaluation and other empirical approximations to model the gate capacitance are not made. Our modeling results are in good agreement with experimental data and other approximation models. The present model has demonstrated its wide applicability to various material and structural combinations of high-k dielectric stacks. Therefore, it provides a design aid for nanoscale-MOS structures with high-k stack-gate structures. As an example, an optimum nitrogen content in a Hf/SiO/N gate dielectric for minimizing gate-leakage current is found. This requires further experimental verification.

Acknowledgment

The authors would like to thank Dr. Hiroo Masuda of Device Research Center, Hitachi Corporation, Japan, for his contributions to this paper.

References

[1] [Online]. Available: http://www.sia-online.org/pre-release

[2] C. Huang, T. Wang, C. N. Chen, M. C. Chang, and J. Fu, "Modeling hot electron gate current in Si MOSFETs using a coupled drift-diffusion and Monte Carlo method," *IEEE Trans. Electron Devices*, vol. 39, no. 11, pp. 2562–2568, Nov. 1992.

[3] E. Cassan, S. Galdin, P. Dollfus, and P. Hesto, "Analysis of electron energy distribution function in ultrathin gate oxide n-MOSFETs using Monte Carlo simulation for direct tunneling gate current calculation," in *Proc. 11th International Conference on Nonequilibrium Carrier Dynamics in Semiconductors (HCIS-11)*, Kyoto, Japan, Jul. 1999, vol. 272, pp. 1464–1471.

[4] 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, no. 7, pp. 1464–1471, Jul. 1999.

[5] Y. Y. Fan, R. E. Nieh, J. C. Lee, G. Lucovsky, G. A. Brown, L. F. Register, and S. K. Banerjee, "Voltage and temperature dependent gate capacitance and current model: Applications to ZrO2 n-channel MOS capacitor," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 1969–1978, Nov. 2002.

[6] S. H. Lo, D. A. Buchanan, and Y. Taur, "Modeling characterization of quantization, polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides," *IBM J. Res. Develop.*, vol. 43, no. 3, pp. 327–337, 1999.

[7] C. K. Huang and N. Goldsman, "Modeling the limits of gate oxide scaling with a Schrödinger based method of direct tunneling gate currents of nanoscale MOSFETs," in *Proc. IEEE NANO—Nanoelectronics and Giga-Scale Systems*, 2001, pp. 335–339.

[8] H. Lin, J. Z. Peng, N. Goldsman, and I. D. Mayergoyz, "An efficient physics-based gate current calculation by solving space-dependent Boltzmann transport equation," in *Proc. 11th Biennial Univ./Government/Ind. Microelectron. Symp.*, Austin, TX, May 16–17, 1995, pp. 193–196.

[9] B. Govoreanu, W. Magnus, W. Schoenmaker, J. Van Houdt, and K. De Meyer, "On the calculation of the quasi-bound-state energies and lifetimes in inverted MOS structures with ultra-thin oxides and its application to the direct tunneling current," *IEEE Trans. Electron Devices*, vol. 51, no. 5, pp. 764–773, May 2004.

[10] T. Janik and B. Majkusiak, "Analysis of the MOS transistor based on the self-consistent solution to the Schrödinger and Poisson equations and on the local mobility model," *IEEE Trans. Electron Devices*, vol. 45, no. 6, pp. 1263–1271, Jun. 1998.

[11] K. Hasnat, C.-F. Yeap, S. Jallepalli, S. A. Hareland, W.-K. Shih, V. M. Agostinelli, Jr., A. F. Tasch, and C. M. Maziar, "Thermionic emission model of electron gate current in submicrometer nMOSFETs," *IEEE Trans. Electron Devices*, vol. 44, no. 1, pp. 129–138, Jan. 1997.

[12] Y. Ma, L. Liu, Z. Yu, and Z. Li, "Validity and applicability of triangular potential well approximation in modeling of MOS structure inversion and accumulation layer," *IEEE Trans. Electron Devices*, vol. 47, no. 9, pp. 1764–1767, Sep. 2000.

[13] J. P. Sun, R. K. Mains, W. L. Chen, J. R. East, and G. I. Haddad, "C–V and I–V characteristics of quantum well varactors," *J. Appl. Phys.*, vol. 72, no. 6, pp. 2340–2346, Sep. 1992.

[14] J. P. Sun, G. I. Haddad, P. Mazumder, and J. Schulman, "Resonant tunneling diodes: Models and properties," *Proc. IEEE*, vol. 86, no. 4, pp. 641–661, Apr. 1998.

[15] C. S. Lent and D. J. Kirkner, "The quantum transmitting boundary method," *J. Appl. Phys.*, vol. 67, no. 10, pp. 6353–6355, May 1990.

[16] J. P. Sun, R. K. Mains, K. Yang, and G. I. Haddad, "A self-consistent model of Γ-X mixing in GaAs/AlAs/GaAs quantum well structures using the quantum transmitting boundary method," *J. Appl. Phys.*, vol. 74, no. 8, pp. 5053–5057, Oct. 1993.

[17] Y. T. Hou, M. F. Li, Y. Jin, and W. H. Lai, "Direct tunneling hole currents through ultrathin gate oxides in metal-oxide-semiconductor devices," *J. Appl. Phys.*, vol. 91, no. 1, pp. 258–264, Jan. 2002.

[18] X. Liu, J. Kang, and R. Han, "Direct tunneling current model for MOS devices with ultra-thin gate oxide including quantization effect and polysilicon depletion effect," *Solid State Commun.*, vol. 125, no. 3/4, pp. 219–223, Jan. 2000.

[19] Y. Ma, L. Liu, L. Tian, Z. Yu, and Z. Li, "Analytical charge control and I–V model for submicrometer and deep-submicrometer MOSFETs fully comprising quantum mechanical effects," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 20, no. 4, pp. 495–502, Apr. 2001.

[20] S. Takagi and A. Toriumi, "Quantitative understanding of inversion-layer capacitance in Si MOSFETs," *IEEE Trans. Electron Devices*, vol. 42, no. 12, pp. 2125–2130, Dec. 1995.

[21] A. S. Spinelli, A. Pacelli, and A. L. Lacaita, "Polysilicon quantization effects on the electrical properties of MOS transistors," *IEEE Trans. Electron Devices*, vol. 47, no. 12, pp. 2366–2371, Dec. 2000.

[22] D. Vasileska, D. K. Schroder, and D. K. Ferry, "Scaled silicon MOSFETs: Degradation of the total gate capacitance," *IEEE Trans. Electron Devices*, vol. 44, no. 4, pp. 584–587, Apr. 1997.

[23] Y. Y. Fan, Q. Xiang, J. An, L. F. Register, and S. K. Banerjee, "Impact of interfacial layer and transition region on gate current performance for high-k gate dielectric stack: Its tradeoff with gate capacitance," *IEEE Trans. Electron Devices*, vol. 50, no. 2, pp. 433–439, Feb. 2003.

[24] M. Koyama, A. Kaneko, T. Ino, M. Koike, Y. Kamata, R. Iijima, Y. Kamimuta, A. Takashima, M. Suzuki, C. Hongo, S. Inumiya, M. Takayanagi, and A. Nishiyama, "Effects of nitrogen in HfSiON gate dielectric on the electrical and thermal characteristics," in *IEDM Tech. Dig.*, 2002, pp. 849–852.

[25] M. Koike, T. Ino, Y. Kamimuta, M. Koyama, Y. Kamata, M. Suzuki, Y. Mitani, A. Nishiyama, and Y. Tsunashima, "Effect of Hf–N bond on properties of thermally stable amorphous HfSiON and applicability of this material to Sub-50 nm technology node LSIs," in *IEDM Tech. Dig.*, 2003, pp. 107–110.

J. P. Sun received the B.S. degree from Southeast University, Nanjing, China, and the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, all in electrical engineering.

He was with Nanjing Solid State Devices Research Institute, China, the Electronics Research Laboratory, University of California at Berkeley, and Solid State Electronics Laboratory, the University of Michigan. He is also currently with the Department of Electrical Engineering and Computer Science, the University of Michigan, Ann Arbor, where he is a Visiting Associate Research Professor. His research interests include modeling and design of nanoelectronic devices and microwave circuits.

Wei Wang is currently working toward the Ph.D. degree at the Department of Biomedical Engineering, Southeast University, Nanjing, China.

He is also with the College of Optoelectronic Engineering, Nanjing University of Posts and Telecommunications, as a Lecturer. His research interests include nanoelectronics, focusing primarily on quantum effects in nanoscale device modeling and analog-circuit design.

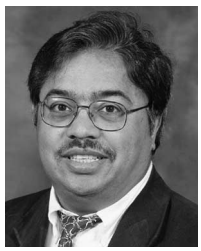
Toru Toyabe (M'76–SM'91–F'96) received the B.S. and M.S. degrees in physics and the Ph.D. degree in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1964, 1966, and 1981, respectively.

He was with Hitachi Central Research Laboratory, Hitachi Ltd., Tokyo, where he worked on GaAs devices and Si devices with two- and three-dimensional device simulations. He joined the Faculty of Engineering, Toyo University, Kawagoe, Japan, in 1994, where he is currently a Professor of the Department of Computational Science and Engineering. His research interests include SOI MOSFETs modeling with particular emphasis in quantum effects and trap models.

Dr. Toyabe is the recipient of the IEEE Paul Rappaport Award in 1984 and Hattori Hoko Award in 1998.

Ning Gu received the Bachelor's degree in radio engineering from Nanjing Institute of Technology, Nanjing, China, the Master's degree in bioelectronics, and the Doctorate's degree in biomedical engineering from Southeast University (SEU), Nanjing, China, in 1986, 1990, and 1996, respectively.

After completing his doctorate, he was with Tokyo University as a Visiting Scholar. In 1997, he became a Professor of SEU. He was with the National Institute of Materials and Chemical Research, Yamanashi University, as a Postdoctoral Fellow in 1998–1999, and the Institute for Molecular Science as a Visiting Professor in 2002. Currently, he is the Chairman of the Department of Biological Science and Medical Engineering, and Director of Research Center for Nano-scale Science and Technology at SEU. He has coauthored more than 120 journal papers on organic materials, thin films, and their properties, characterization of biomaterials, nanoparticles, and nanofabrication for biomedical application, etc. His current research interests include functional nanomaterials, molecular assembly, nanofabrication, nanodevices, and their applications in biomedicine.



Pinaki Mazumder (M'88–SM'95–F'99) received the Ph.D. degree from the University of Illinois at Urbana-Champaign, Urbana-Champaign, in 1988.

He is currently a Professor of electrical engineering and computer science with the University of Michigan, Ann Arbor. For six years, he had been with industrial R&D centers including AT&T Bell Laboratories, where he started the CONES project, which is the first C-modeling-based very large scale integration (VLSI) synthesis tool, in 1985 and India's premiere electronics company Bharat Electronics Ltd., where he had developed several high-speed and high-voltage analog integrated circuits intended for consumer electronics products. He has published more than 200 technical papers and four books on various aspects of VLSI research works. His research interests include current problems in very deep submicron CMOS VLSI design, computer-aided design tools, and circuit designs for emerging technologies including Quantum MOS and resonant tunneling devices, semiconductor memory systems, and physical synthesis of VLSI chips.

Dr. Mazumder was a recipient of the Digital's Incentives for Excellence Award, the BF Goodrich National Collegiate Invention Award, and the DARPA Research Excellence Award.