Journal of Computational Electronics 2: 219–223, 2003
© 2003 Kluwer Academic Publishers. *Manufactured in The Netherlands.*

Analysis of Gate Dielectric Stacks Using the Transmitting Boundary Method

A. GEHRING*, H. KOSINA AND S. SELBERHERR
Institut für Mikroelektronik, TU Wien, Gusshausstrasse 27–29, A-1040 Vienna, Austria
gehring@iue.tuwien.ac.at

Abstract. We present simulations of tunneling current through layers of high-κ dielectrics using the Tsu-Esaki model and the quantum transmitting boundary method (QTBM) to estimate the transmission coefficient. In contrast to transfer-matrix based methods, which suffer from numerical instabilities due to rounding errors, the QTBM is numerically stable even for large stacks and is suitable for implementation in device simulators. The method is used to investigate the tradeoff between conduction band offset and permittivity in alternative high-κ materials.

Keywords: device simulation, tunneling, high-k dielectrics, transfer-matrix, transmitting-boundary

1. Introduction

To enable further device scaling into the sub-100 nm channel length regime it is necessary to reduce MOSFET effective oxide thicknesses (EOT) below 2 nm. This is not possible using SiO_2 due to an exponential increase in gate leakage current: a gate current density of 1 A/cm^2 is usually regarded as upper limit for proper functioning CMOS circuits [1], while low-power devices may require even stricter limitations. Gate dielectric stacks consisting of high-κ dielectric layers such as Si_3N_4, Al_2O_3, Ta_2O_5, HfO_2, or ZrO_2 have been suggested to act as alternative dielectrics. Apart from interface quality and reliability, the dielectric permittivity and the conduction band offset to silicon are of utmost importance as they determine the gate current density through the layer. Unfortunately, materials with a high permittivity have a low band offset and vice versa, so a trade off between these parameters has to be found. Furthermore, at the interface to the underlying silicon substrate, an interface layer exists which is either created unintentionally during processing or intentionally deposited to improve the interface quality. Due to this layer, the band edge in the dielectric is not linear but shows a kink. Simulation of gate leakage thus requires the calculation of tunneling current through piece-wise linear energy barriers, for which the applicability of the commonly used Wentzel-Kramers-Brillouin method is questionable. To get a clear picture of the transmission through the stack, a rigorous solution of Schrödinger's equation in the dielectric region is necessary.

*To whom correspondence should be addressed.

2. Simulation Method

The most prominent and almost exclusively used expression to describe tunneling through insulating layers has been developed by Duke [2] and used by Tsu and Esaki [3] for the simulation of tunneling through a one-dimensional superlattice. The gate current density is computed as

$$J_g = \frac{4\pi m_{\text{eff}} q}{h^3} \int_0^\infty TC(\mathcal{E}_t) N(\mathcal{E}_t)\, d\mathcal{E}_t \qquad (1)$$

where $TC(\mathcal{E}_t)$ is the transmission coefficient, which only depends on the energy perpendicular to the interface, and $N(\mathcal{E}_t)$ the supply function which is defined as

$$N(\mathcal{E}_t) = \int_0^\infty [f_1(\mathcal{E}_t + \mathcal{E}_l) - f_2(\mathcal{E}_t + \mathcal{E}_l + \Delta\mathcal{E}_C)]\, d\mathcal{E}_l \qquad (2)$$

where $f_1(\mathcal{E})$ and $f_2(\mathcal{E})$ denote the distribution of carriers next to the oxide. For the case of a Fermi-Dirac

distribution the gate current density evaluates to

$$J_g = \frac{4\pi m_{\mathrm{eff}} \mathrm{q} \mathrm{k_B} T}{h^3} \int_0^\infty TC(\mathcal{E}_t) \times \ln\left\{\frac{1+\exp[(\mathcal{E}_f - \mathcal{E}_t)/\mathrm{k_B}T]}{1+\exp[(\mathcal{E}_f' - \mathcal{E}_t)/\mathrm{k_B}T]}\right\} \mathrm{d}\mathcal{E}_t \quad (3)$$

where $\mathcal{E}_f$ and $\mathcal{E}_f'$ denote the Fermi energies at the semiconductor-oxide interfaces. This expression is frequently used in the literature and implemented in various device simulators.

Several methods have been proposed to calculate the transmission coefficient $TC(\mathcal{E}_t)$ of an energy barrier as shown in Fig. 1. Using the Wentzel-Kramers-Brillouin (WKB) approximation, the transmission coefficient is found by integration of the wave number within the classical turning points, i.e. the region where $\mathcal{E} \leq V(x)$. Gundlach's method [4] provides an analytically exact expression for the transmission coefficient of a linear potential barrier using Airy functions. For arbitrary shaped barriers the transfer-matrix method [5] can be used. It is based on the approximation of an energy barrier by a sequence of constant or linear energy barriers as shown in Fig. 1. Since the wave function in constant or linear barriers can easily be calculated, the total transfer matrix and the transmission coefficient can be derived by a number of subsequent matrix computations. The main shortcoming of the method is that it becomes numerically instable for low energies [6]. During the matrix multiplications exponentially growing and decaying states have to be multiplied, leading to rounding errors which eventually exceed the amplitude of the wave function itself. The problems have been overcome by the use of scattering matrices [7], iterative methods [8], or by simply setting the transfer matrix entries to zero if the decay factor $\sum k_j(\mathcal{E})\Delta$ exceeds a given value [9].

Figure 1. The energy barrier for the linear and constant potential transfer-matrix methods.

An alternative method to solve Schrödinger's equation with open boundary conditions has been proposed by Frensley [10] and is based on the quantum transmitting boundary method. Applying a finite-difference approximation of Schrödinger's equation on an equidistant grid with an effective mass m_j and a grid spacing Δ, the discretized one-dimensional Hamiltonian reads

$$s_{j-1}\Psi_{j-1} + d_j\Psi_j + s_{j+1}\Psi_{j+1} = \mathcal{E}\Psi, \quad (4)$$

where $s_j = -\hbar^2/(2m_j\Delta^2)$ and $d_j = \hbar^2/(m_j\Delta^2) + V_j$. Open boundary conditions are introduced by writing the wave functions at the boundaries of the simulation domain as

$$\begin{aligned} \Psi_1 &= a_1 + b_1 \\ \Psi_n &= a_n + b_n \end{aligned} \quad (5)$$

and the wave functions outside of the simulation domain as

$$\begin{aligned} \Psi_0 &= a_1 \exp(-ik_1\Delta) + b_1 \exp(ik_1\Delta) \\ \Psi_{n+1} &= a_n \exp(-ik_n\Delta) + b_n \exp(ik_n\Delta). \end{aligned} \quad (6)$$

Setting $a_1 = \alpha_1\Psi_0 + \beta_1\Psi_1$ and $a_n = \alpha_n\Psi_{n+1} + \beta_n\Psi_n$ eliminates the unknown values of b_1 and b_n and gives a linear system for the $n+2$ complex values Ψ_j

$$\begin{pmatrix} \alpha_1 & \beta_1 & & & & \\ s_1 & d_1' & s_2 & & & \\ & s_2 & d_2' & s_3 & & \\ & & \cdots & & & \\ & & & s_n & d_n' & s_n \\ & & & & \beta_n & \alpha_n \end{pmatrix} \begin{pmatrix} \Psi_0 \\ \Psi_1 \\ \Psi_2 \\ \cdots \\ \Psi_n \\ \Psi_{n+1} \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ 0 \\ \cdots \\ 0 \\ a_n \end{pmatrix} \quad (7)$$

where $d_i' = d_i - \mathcal{E}$. Setting $a_1 = 1$ and $a_n = 0$ yields the values of the wave function in the whole simulation domain for an incident wave from the left side. The method is easy to implement, fast, and more robust than the transfer-matrix method.

Figure 2 shows that the methods are equivalent for a simple triangular barrier. The QTBM does not suffer from numerical problems like the linear- or constant-potential transfer-matrix methods. Furthermore, it directly yields the wave function in the insulator, which

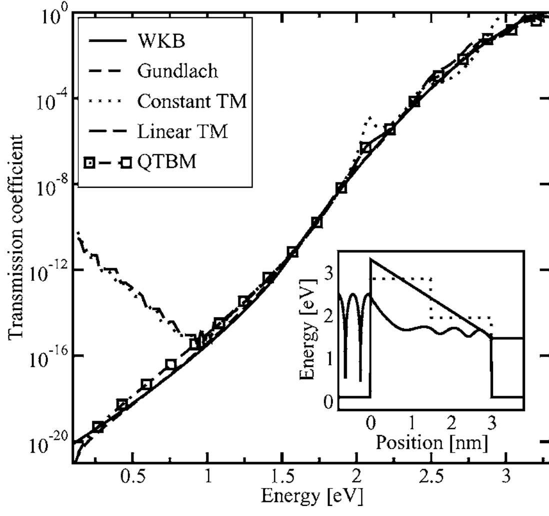

Figure 2. Transmission coefficient as a function of energy for a 3 nm thick layer of SiO_2 computed by different methods. The transfer-matrix methods fail for $\mathcal{E} < 1$ eV.

allows to account for the charge of the tunneling carriers via

$$n(x) = \frac{2m\mathrm{k_B}T}{h^2}\int_0^\infty |\overleftarrow{\Psi}(x)|^2 \times \ln\left[1+\exp\left(\frac{E_{\mathrm{f,g}}-E}{\mathrm{k_B}T}\right)\right]\mathrm{d}k + \frac{2m\mathrm{k_B}T}{h^2} \times \int_0^\infty |\overrightarrow{\Psi}(x)|^2 \ln\left[1+\exp\left(\frac{E_{\mathrm{f,s}}-E}{\mathrm{k_B}T}\right)\right]\mathrm{d}k$$

where $\overleftarrow{\Psi}(x)$ and $\overrightarrow{\Psi}(x)$ denote wave functions impinging from the gate and the substrate, respectively, with $E_{\mathrm{f,g}}$ and $E_{\mathrm{f,s}}$ being the respective Fermi energies.

3. Implementation and Results

The QTBM was implemented in the general-purpose device simulator MINIMOS-NT. The tunneling current is calculated between two specified boundaries of insulator or semiconductor segments with N interface nodes. The total current is found by numerical integration along the boundary, where the local tunneling current density is calculated from (3). To assure self-consistency with the transport model in the underlying silicon substrate, the local tunneling current density is considered in the continuity equation of the neighboring segments. Since trap-assisted tunneling processes have been neglected, this model gives a lower bound for the leakage current.

Table 1. Dielectric permittivity, band gap, and conduction band offset of dielectric materials [11–17].

	Permittivity κ/κ_0 [1]	Band gap $\mathcal{E}_\mathrm{g}$ [eV]	Band offset $\Delta\mathcal{E}_\mathrm{C}$ [eV]
SiO_2	3.9	8.9–9.0	3.0–3.5
Si_3N_4	7.0–7.9	5.0–5.3	2.0–2.4
Ta_2O_5	23.0–26.0	4.4–4.5	0.3–1.5
TiO_2	39.0–170.0	3.0–3.5	0.0–1.2
Al_2O_3	7.9–12.0	5.6–9.0	2.78–3.5
ZrO_2	12.0–25.0	5.0–7.8	1.4–2.5
HfO_2	16.0–40.0	4.5–6.0	1.5

When comparing different dielectric materials, one encounters a large discrepancy in material parameters, as seen in Table 1. Choosing some intermediate material parameters from Table 1 the gate current density can be computed as a function of the gate bias for different materials with an underlying 0.5 nm layer of SiO_2 and a fixed EOT of 1 nm as shown in Fig. 3. Both SiO_2 and Si_3N_4 show a much too high leakage, while Ta_2O_5, ZrO_2, and HfO_2 stay below 1 A/cm^2 at $V_{\mathrm{GS}} = 1$ V. Due to the low conduction band offset, TiO_2 shows an especially pronounced current increase.

To assess the material parameters necessary to reach a specific maximum gate current density we calculated the gate current density as a function of the

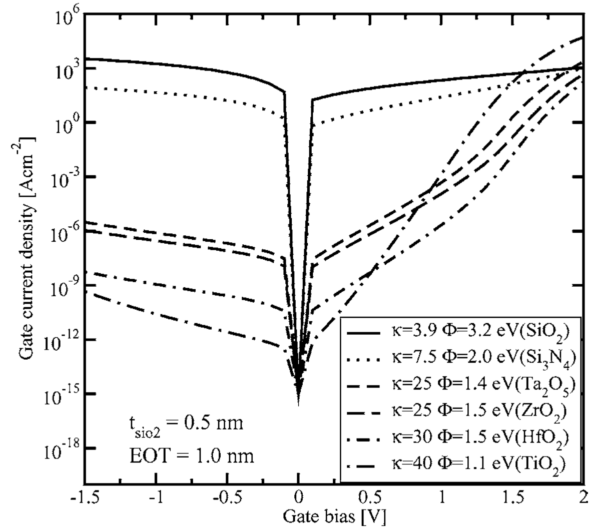

Figure 3. Gate current density as a function of the gate voltage for different materials. The dielectric stack consists of a 0.5 nm SiO_2 layer and a high-κ layer with a total EOT of 1.0 nm.

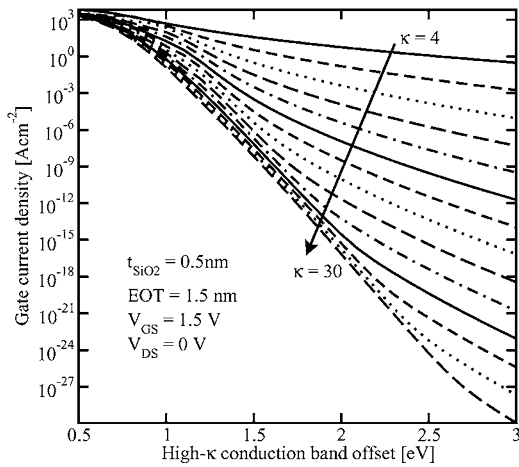

Figure 4. Dependence of the gate current on the high-κ conduction band offset and permittivity of a stack with EOT = 1.5 nm, a 0.5 nm SiO_2 interface layer at $V_{DS} = 0$ V and $V_{GS} = 1.5$ V.

conduction band offset and dielectric permittivity as shown in Fig. 4. Since it is often not possible to vary the thickness of the underlying SiO_2 layer we fixed it at 0.5 nm and adjusted the high-κ thickness to reach an EOT of 1.5 nm. The simulations were performed at $V_{GS} = 1.5$ V and $V_{DS} = 0$ V. Increasing the value of κ strongly reduces the leakage current due to the higher physical stack thickness. However, materials with Φ below 1 eV never give acceptable gate current densities. Fig. 5 shows the gate current density for an EOT

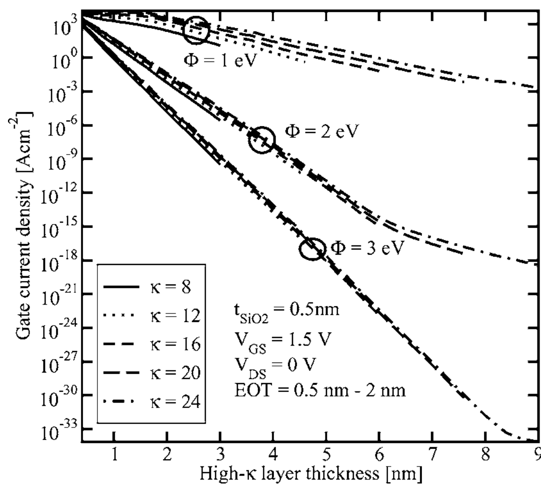

Figure 5. Dependence of the gate current on the high-κ layer thickness, conduction band offset, and permittivity of a stack with EOT = 2.0 nm, a 0.5 nm SiO_2 interface layer at $V_{DS} = 0$ V and $V_{GS} = 1.5$ V.

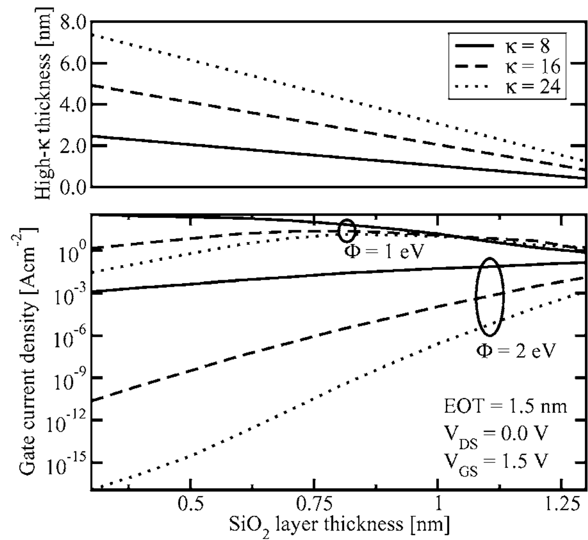

Figure 6. Dependence of the gate current on the interface layer thickness, conduction band offset, and permittivity of a stack with EOT = 1.5 nm at $V_{DS} = 0$ V and $V_{GS} = 1.5$ V.

ranging from 0.5 nm to 2.0 nm as a function of the high-κ layer thickness. Again, the stack consists of an underlying 0.5 nm layer of SiO_2 and the bias point is $V_{GS} = 1.5$ V and $V_{DS} = 0$ V. For $\Phi = 1$ eV, large high-κ thicknesses are necessary to reduce the leakage. Such large stacks may pose problems due to fringing fields from the drain contact which reduce the threshold voltage. The tradeoff between permittivity and conduction band offset gives rise to further effects as shown in Fig. 6: If the EOT is fixed, an increase of the SiO_2 layer thickness causes a reduced thickness of the high-κ layer. This is shown in the upper part of Fig. 6 for different values of the permittivity. That reduction of the total stack thickness may cause the gate current density at a specific bias point to stay constant, increase, or even decrease depending on the material parameters. So, a clear statement about the optimum thickness of the interface layer depends on the material parameters.

4. Conclusions

We presented simulations of high-κ dielectric materials using the device simulator MINIMOS-NT. The transmitting boundary method has been used to compute the transmission coefficient of the piecewise linear energy barrier within the framework of the Tsu-Esaki tunneling model. In contrast to transfer-matrix based methods, the transmitting-boundary method does not suffer

from numerical problems and directly yields the values of the wave function in the insulator. This allows to take the charge of tunneling carriers into account. We used the method to study the influence of permittivity and conduction band offset on the leakage current of gate dielectrics. To ensure a leakage current below 1A/cm^2 at an effective oxide thickness of 1.5 nm, a band offset of at least 1.5 eV and a permittivity higher than 8 are necessary.

References

1. S. Song *et al.*, *Journal of Semiconductor Technology and Science*, **1**, 103 (2001).
2. C.B. Duke, Tunneling in Solids (Academic Press New York, 1969).
3. R. Tsu and L. Esaki, *Appl. Phys. Lett.*, **22**, 562 (1973).
4. K.H. Gundlach, *Solid-State Electron.*, **9**, 949 (1966).
5. Y. Ando and T. Itoh, *J. Appl. Phys.*, **61**, 1497 (1987).
6. D.Z.Y. Ting *et al.*, *Physical Review B*, **45**, 3583 (1992).
7. D.Y.K. Ko and J.C. Inkson, *Physical Review B*, **38**, 9945 (1988).
8. T. Usuki *et al.*, *Physical Review B*, **52**, 8244 (1995).
9. B.A. Biegel, *Dissertation* (Stanford University, 1997).
10. W.R. Frensley, *Superlattices & Microstructures*, **11**, 347 (1992).
11. J. Zhang *et al.*, *Solid-State Electron.*, **44**, 2165 (2000).
12. J.D. Casperson *et al.*, *J. Appl. Phys.*, **92**, 261 (2002).
13. M.L. Roy *et al.*, *J. Appl. Phys.*, **93**, 2966 (2003).
14. C.M. Osburn *et al.*, *IBM J. Res. Dev.*, **46**, 299 (2002).
15. G.D. Wilk *et al.*, *J. Appl. Phys.*, **89**, 5243 (2001).
16. J. Robertson, *J. Vac. Sci. Technol.*, **18**, 1785 (2000).
17. H.-S. P. Wong, *IBM J. Res. Dev.*, **46**, 133 (2002).