Modeling of High-k-Metal-Gate-Stacks Using the Non-Equilibrium Green's Function Formalism

O. Baumgartner, M. Karner, and H. Kosina

Institute for Microelectronics, TU Wien
Gußhausstraße 27–29, A–1040 Wien, Austria
E-mail: {baumgartner|karner|kosina}@iue.tuwien.ac.at

***Abstract*—A high-k-Metal-Gate stack has been investigated using an open boundary model based on the non-equilibrium Green's function formalism. The numerical energy integration, which is crucial because of the very narrow resonant states, is pointed out in detail. The model has been benchmarked against the established classical and closed boundary Schrödinger-Poisson model. In contrast to the established models, the solution covers distinct resonant states with a realistic broadening and results in a major difference in the current density spectrum.**

I. Introduction

The recent introduction of high-k-metal-gate transistors [1] draws the attention to a more accurate modeling of gate leakage current. Two different models are commonly used, namely the Tsu-Esaki formula [2] and the quasi-bound state (QBS) tunneling formalism [3]. The current expressions are given by

$$J_{\mathrm{Tsu}} = N_{\mathrm{Tsu}} \int TC(\mathcal{E}) SF(\mathcal{E})\, \mathrm{d}\mathcal{E} \tag{1}$$

according to Tsu and Esaki, and

$$J_{\mathrm{QBS}} = N_{\mathrm{QBS}} \sum_i \frac{n_i}{\tau_i} \tag{2}$$

for the QBS case. Expression (1) relies on a the transmission coefficient TC of the barrier and a supply function SF, determined by the carrier distributions in the gate and channel regions. The QBS method is based on the electron populations n_i of the discrete subbands in the MOS inversion layer and a finite lifetime τ_i. Both approaches neglect the carrier density in the dielectric due to the hard wall boundary conditions assumed and are thus inconsistent with the non-vanishing current density.

II. Non-Equilibrium Green's Functions

A more rigorous description by the non-equilibrium Green's functions (NEGF) formalism [4] overcomes the aforementioned problem. It allows for a full quantum mechanical treatment and yields the current density consistently with the carrier density. The influence of level broadening due to scattering processes was modeled by means of an optical potential [5]. Using this model a high-k gate-stack has been analyzed.

The gate and the bulk regions have been assumed to be in thermal equilibrium and are characterized by the Fermi energies $\mathcal{E}_{\mathrm{FG}}$ and $\mathcal{E}_{\mathrm{FB}}$, respectively (c.f. Fig.1). The leakage current through the gate dielectric, which separates the equilibrium regions, has been calculated assuming ballistic transport between the two reservoirs [6], [7]. The retarded and advanced Green's functions are determined by

$$\begin{aligned} G^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E}) &= G^{\mathrm{A}\dagger}(\mathbf{r},\mathbf{r}',\mathcal{E}) \\ &= [\mathcal{E}I - H(\mathbf{r},\mathbf{r}',\mathcal{E}) - \Sigma^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})]^{-1}, \end{aligned} \tag{3}$$

where $H(\mathbf{r},\mathbf{r}',\mathcal{E})$ is the Hamiltonian of the system and $\Sigma^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})$ is the retarded self-energy. The optical potential determined by the carrier lifetime τ is added to the diagonal elements of the Hamiltonian:

$$H(\mathbf{r},\mathbf{r},\mathcal{E}) = H_0(\mathbf{r},\mathbf{r},\mathcal{E}) + \mathrm{i}\hbar/(2\tau). \tag{4}$$

Assuming Fermi Dirac statistics, the occupation is given by $f_{\mathrm{G,B}}(\mathcal{E}) = N_{\mathrm{C,2D}}\mathcal{F}_0\left(\beta\left(\mathcal{E}_{\mathrm{F_{G,B}}} - \mathcal{E}\right)\right)$ with $\beta = 1/\mathrm{k_B}T$. Within the equilibrium regions, the lesser Green's function is calculated as $G^{<}(\mathbf{r},\mathbf{r}',\mathcal{E}) = G^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E}) f_{\mathrm{G,B}}(\mathcal{E})$. The lesser Green's function in the dielectric is determined by $G^{<}(\mathbf{r},\mathbf{r}',\mathcal{E}) = G^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})\Sigma^{<}(\mathbf{r},\mathbf{r}',\mathcal{E})G^{\mathrm{A}}(\mathbf{r},\mathbf{r}',\mathcal{E})$. The lesser self energy of the left and right contact is given as $\Sigma^{<}_{\mathrm{G,B}}(\mathcal{E}) = \mathrm{i}\Im\left\{\Sigma^{\mathrm{R}}_{\mathrm{G,B}}(\mathcal{E})\right\} f_{\mathrm{G,B}}(\mathcal{E})$. The electron density and

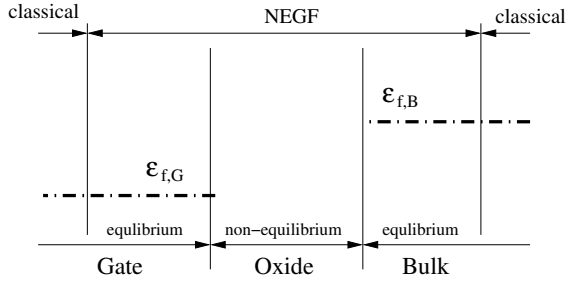

Fig. 1: The simulation domain is split into a classical lead region and a quantum mechanical device region. The gate and bulk contacts are assumed to be in thermal equilibrium.

978-1-4244-1753-7/08/$25.00 ©2008 IEEE

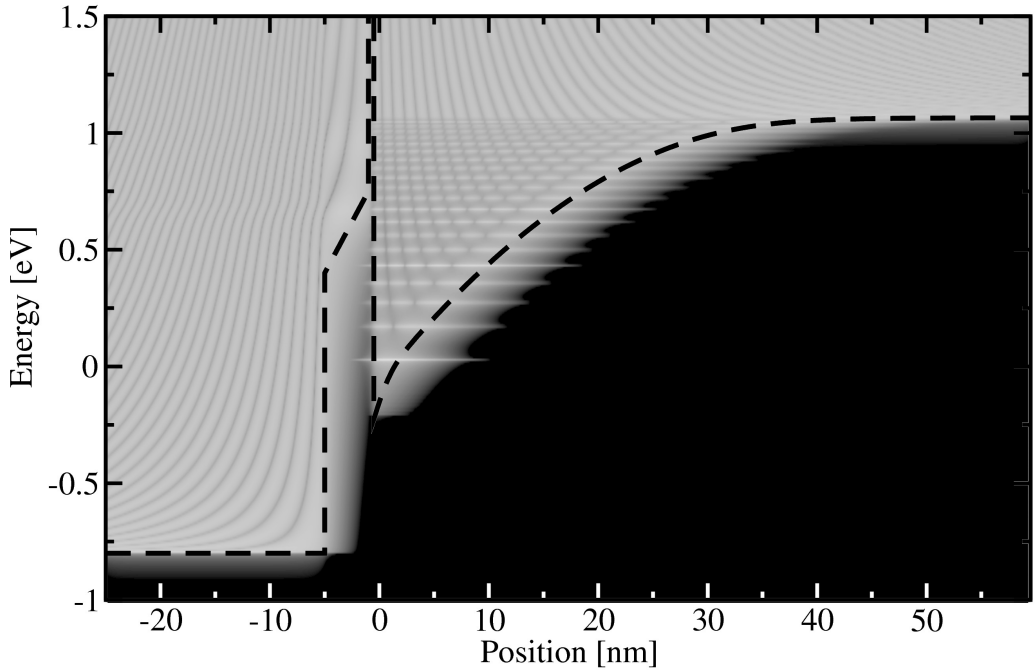

Fig. 2: Self-consistent band edge and the local density of states. Quantum mechanical effects like the penetration of the wavefunctions into classically forbidden regions and reflections at the barrier are clearly seen. Furthermore, in the channel the formation of quasi-bound states and the transition to the continuum states is observed.

the leakage current are given by the integrals

$$n(\mathbf{r}) = -2\mathrm{i} \int G^{<}(\mathbf{r},\mathbf{r},\mathcal{E}) \frac{\mathrm{d}\mathcal{E}}{2\pi} \,, \tag{5}$$

$$j(\mathbf{r}) = -\frac{\hbar q}{m^*} \int \left[(\nabla - \nabla') G^{<}(\mathbf{r},\mathbf{r}',\mathcal{E})\right]\Big|_{\mathbf{r}'=\mathbf{r}} \frac{\mathrm{d}\mathcal{E}}{2\pi} \,. \tag{6}$$

III. Numerical Methods

In inversion, numerous quasi-bound states arise in the channel of a MOS transistor as displayed in Fig. 2. These states correspond to narrow resonances in the energy spectrum. To correctly calculate the integrals (5) and (6) these resonances need to be accurately resolved. Using a fixed, equidistant energy grid does not necessarily yield higher numerical accuracy but greatly increases the computational cost, since the Green's functions need to be solved for every energy grid point. Therefore, an adaptive energy integration method has been implemented [8]. One of the realized algorithms, which is based on the doubly adaptive quadrature routine reported in [9], is depicted in Fig. 3. The method utilizes Newton-Cotes quadrature of the order five, nine, 17 and 33.

Starting from an initial grid, for example, provided by a resonance finder to further increase numerical accuracy, the Green's functions are calculated for the given energies. Then the integral and the error are computed for all subintervals. The interval with the biggest contribution to the aggregated global error is then extracted from the datastructure and subdivided. If the integration error is reduced hereby, the two new subintervals are reinserted into the datastructure. Otherwise, the Green's functions on additional energy grid points are calculated and the next higher order Newton-Cotes rule is applied to the processed interval. This procedure is repeated until the previously chosen global error criterion is fulfilled. By this means, the algorithm generates an energy grid, automatically refined in critical ranges of the energy spectrum, namely near the potential of the contact regions and at the energies of the resonant states.

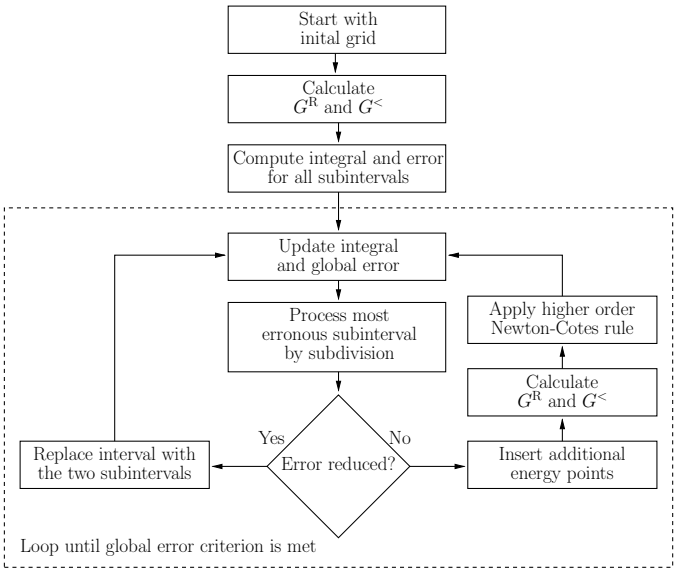

Fig. 3: Illustration of the doubly adaptive global quadrature routine.