UNIVERSITY
of
GLASGOW

Watling, J.R. and Brown, A.R. and Asenov, A. and Svizhenko, A. and Anantram, M.P. (2002) Simulation of direct source-to-drain tunnelling using the density gradient formalism: Non-Equilibrium Greens Function calibration. In, *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD 2002), 4-6 September 2002*, pages pp. 267-270, Kobe, Japan.

http://eprints.gla.ac.uk/3009/

Glasgow ePrints Service
http://eprints.gla.ac.uk

Simulation of direct source-to-drain tunnelling using the density gradient formalism: Non-equilibrium Green's function calibration

J. R. Watling*†, A. R. Brown†, A. Asenov†, A. Svizhenko‡ and M. P. Anantram‡
†Device Modelling Group, Dept. of Electronics and Electrical Engineering
University of Glasgow, Glasgow G12 8LT, Scotland, UK
‡NASA Ames Research Center, Moffett Field, CA 90345-1000, U.S.A.
*E-mail: J.Watling@elec.gla.ac.uk, Tel: +44 141 330 4792, Fax: +44 141 330 4907

Abstract – **Quantum mechanical confinement effects, gate, band-to-band and source-to-drain tunnelling will dramatically affect the characteristics of future generation nanometre scaled devices. It has been demonstrated already that first-order quantum corrections, which satisfactorily describe quantum confinement effects, can be introduced into efficient TCAD orientated drift-diffusion simulators using the density gradient approach. In this paper we refer to Non-Equilibrium Green's Function simulations in order to calibrate the density gradient formalism in respect of both confinement and source-to-drain tunnelling using different effective masses in directions normal and parallel to the conducting channel. We demonstrate that the density gradient formalism can describe accurately the current characteristics in sub 20nm double gate MOSFETs.**

I. INTRODUCTION

As MOSFETs are scaled to decananometre/nanometre dimensions quantum mechanical (QM) effects start to dramatically affect their characteristics. The combination of thin gate oxides and heavy doping in conventional MOSFETs and the thin silicon body of the double-gate structures result in substantial quantum mechanical threshold voltage shifts, gate tunnelling leakage and transconductance degradation [1]. Additionally, below gate-lengths of 10 nm direct source-to-drain tunnelling may became one of the major scaling limitation factors [2,3].

Complete self-consistent treatment of QM effects in the simulation of nano-scale devices necessitates the use of sophisticated quantum simulation techniques such as Non-Equilibrium Green's functions (NEGF) [4], Wigner Functions [5] or a direct solution of Schrödinger's equation [6], all of which are computationally expensive. It has therefore become necessary to find accurate ways of introducing QM corrections in Computer Aided Design tools without compromising their computational efficiency. Recently it has been demonstrated that economical first-order quantum corrections describing quantum confinement effects can be introduced into conventional drift-diffusion simulators using the density gradient (DG) approach [7] or effective potential (EP) formalism [8]. However, both the EP and DG approach must be carefully calibrated in order to represent accurately the quantum confinement effects [9,10]. In previous work we have been able to demonstrate that the DG simulations in the subthreshold regime qualitatively reproduce features in MOSFET characteristics attributable to source-to-drain tunnelling [3,8].

In this paper, using careful comparison and calibration with respect to comprehensive 2D Non-Equilibrium Green's Function (NEGF) simulations [4], we demonstrate that the DG formalism can quantitatively describe the degradation in the subthreshold characteristics in nano-scale double gate MOSFETs introduced by direct source-to-drain tunnelling.

The next section describes the double gate MOSFET structure considered in this work. Section 3 briefly outlines the density gradient formalism. Our results are presented in section 4, while section 5 contains the conclusion and further discussions.

II. THE DOUBLE GATE MOSFET

Double gate MOSFETs are at present the winning candidates for scaling below the 10 nm barrier. Here we consider a planar double gate MOSFET architecture, similar in design to that proposed by Ren *et al* [2]. We have simulated a family of double-gate MOSFETs illustrated schematically in Figure 1, with channel lengths, L_{chan}, ranging from 20nm down to 4nm and with a channel width, W_{chan}, of 10nm. The silicon channel thickness, T_{chan}, is 2 nm and oxide thickness, t_{ox}, is 1nm. The source and drain junctions are 10nm long and n-doped at 1×10^{20}cm^{-3}. The channel is lightly p-doped at 1×10^{14}cm^{-3}.

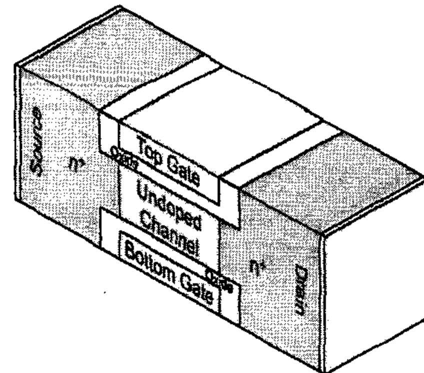

Figure 1: Schematic illustration of the simulated double-gate MOSFET structure.

III. THE DENSITY GRADIENT APPROACH

The density gradient approach follows from the transport equation for the Wigner distribution function given in the form of a modified Bolzmann Transport Equation (BTE) [11]:

$$\frac{\partial f}{\partial t}+\mathbf{v}\cdot\nabla_r f-\frac{1}{\hbar}\nabla_r V\cdot\nabla_k f+$$
$$+\sum_{\alpha=1}^{\infty}\frac{\hbar^{(2\alpha-1)}(-1)^{\alpha+1}}{4^{\alpha}(2\alpha+1)!}(\nabla_r V\cdot\nabla_k f)^{2\alpha+1}=\left(\frac{\partial f}{\partial t}\right)_{Coll} \quad (1)$$

Quantum effects are included through the inherently non-local driving potential in the fourth term on the left-hand side. The modified Drift Diffusion (DD) transport equation with density gradient corrections may be derived in a manner similar to that for deriving the drift diffusion approximation from the semi-classical Boltzmann Transport Equation with special attention paid to the quantum potential terms in (1) [8,11]:

$$J_n = qn\mu_n\nabla\psi + qD_n\nabla n + 2q\mu_n b_n\nabla\left(\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right) \quad (2)$$

where $b_n = \hbar^2/4qm_n^* r$, all other symbols have their usual meaning and r is a dimensionless parameter. In situations with strong quantum confinement, (when only a single sub-band is occupied) the parameter r is considered equal to 1. However, as more subbands become filled, for example due to increase in temperature, statistical averaging causes r to change, approaching 3 asymptotically [12,13]. Although the modified BTE (1) undoubtedly includes tunnelling, it is still debatable to what extent the modified DD equation (2) retains the tunnelling handling capabilities [3].

To avoid the discretisation of fourth order derivatives in multidimensional numerical simulations a generalised electron quasi-Fermi potential ϕ_n is usually introduced by $F_n = n\mu_n\nabla\phi_n$ [14]. Thus the unipolar DD system of equations with DG quantum corrections, which in most of the cases is sufficient for MOSFET simulations, becomes:

$$\nabla\cdot(\varepsilon\nabla\psi) = -q\left(p-n+N_D^+-N_A^-\right) \quad (3)$$

$$2b_n\frac{\nabla^2\sqrt{n}}{\sqrt{n}} = \phi_n - \psi + \frac{k_BT}{q}\ln\left(\frac{n}{n_i}\right) \quad (4)$$

$$\nabla\cdot(n\mu_n\nabla\phi_n) = 0 \quad (5)$$

Equation (4) which is the DG analogue of the Schrödinger equation can be interpreted as an electron state equation with an additional term that is dependent on the gradient of the carrier density. The effect of including a dependence on the gradient of the density means that the equation of state contains an aspect of *nonlocality*, a primary characteristic of quantum mechanics [7].

IV. CALIBRATION

It is commonly accepted [8,9,14] that the effective mass and/or the parameter r in the DG formalism should be considered as an adjustable parameter and have to be properly calibrated in respect of more comprehensive quantum mechanical simulations in order to obtain quantitatively accurate results. For the temperatures and gate biases considered here, an appreciable number of sub-bands are likely to be occupied so that the parameter r may be assumed to equal 3. To achieve satisfactory calibration in respect of both quantum confinement effects and source-to-drain tunnelling we have split the scalar effective mass in the b_n parameter of (4) into two components, $m_\perp$ and $m_\parallel$, normal and parallel to the current flow respectively. The effective mass component, $m_\perp$, perpendicular to the Si/SiO_2 interface, is responsible for the accurate description of the confinement effects and determines for example the threshold voltage shift. The lateral effective mass, $m_\parallel$ is held responsible for the accurate description of source-to-drain tunnelling. The introduction of these two different effective masses results in a modified density gradient equation (4) into:

$$\frac{2\hbar^2}{12q\sqrt{n}}\left[\frac{1}{m_\parallel}\left(\frac{\partial^2\sqrt{n}}{\partial x^2}+\frac{\partial^2\sqrt{n}}{\partial y^2}\right)+\frac{1}{m_\perp}\frac{\partial^2\sqrt{n}}{\partial z^2}\right] = \phi_n - \psi + \frac{k_BT}{q}\ln\left(\frac{n}{n_i}\right) \quad (6)$$

where z is the direction normal to the Si/SiO_2 interface.

We have adjusted the vertical effective mass $m_\perp$ in (4) in order to match the vertical charge distribution in the channel of the double gate device. The results of this calibration are illustrated in figure 2, which compares the vertical electron density distribution through the middle of the channel, obtained from NEGF simulations and from the calibrated DG simulations. Due to quantum confinement the electron density peaks at the centre of the channel. In the DG simulations we have neglected charge density penetration into the SiO_2 barrier, which results in a slightly more bended electron distribution. This, however, does not affect significantly the threshold voltage shift associated with the quantum confinement.

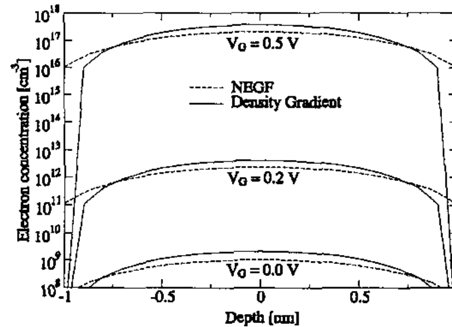

Figure 2: Quantum electron concentration profiles through the centre of the double-gate MOSFET.

Following the above procedure we have found that an effective mass of $0.19m_0$, where m_0 is the free electron rest mass, is appropriate for accurate calibration in respect of the vertical confinement effects. The lateral effective mass $m_\parallel$, which in our case is parallel to the gates of the MOSFET,

has been calibrated in respect of the degradation of the subthreshold slope partially associated in the very short devices with direct source-to-drain tunnelling [8]. The effect of changing the lateral effective mass in the DG simulations on the I_D-V_G characteristic of the 4 nm device can be seen in figure 3. The reduction in lateral effective mass results in degradation of the subthreshold slope which may be attributed to source-drain tunnelling. The lateral effective mass can thus be used as a fitting parameter for calibrating our DG simulations in respect of source-to-drain tunnelling.

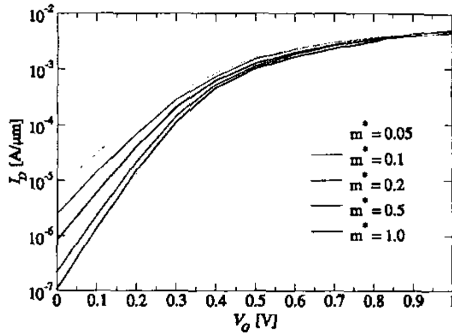

Figure 3: I_D-V_G characteristics for a 4nm channel length double gate structure obtained using DG simulations with different lateral effective masses at 300K.

Figure 4 shows the results from NEGF simulations of a double gate MOSFET, for a variety of gate lengths ranging from 20 nm down to 4nm, along with our calibrated DG simulations. By using a single value of the lateral effective mass ($m_{\parallel} \sim 0.5m_0$) over the entire range of channel lengths and gate voltages we have found good agreement between the DG and NEGF simulations. This is an indication for the physically meaningful nature of the calibration parameters.

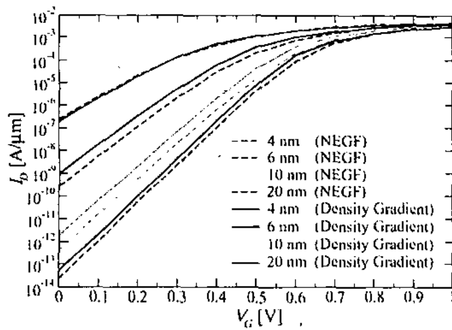

Figure 4: I_D-V_G characteristics obtained from NEGF and calibrated DG simulations for double gate MOSFETs with gate lengths ranging from 20nm down to 4nm. V_D= 1V and V_G is applied to both top and bottom gate contacts.

The large difference between the calibrated lateral and vertical effective masses means that the predictive power of DG simulations utilising a single isotropic effective mass is likely to be somewhat doubtful. Indeed, the sensitively of the source-drain current to the lateral effective mass means that such simulations may significantly over-estimate the tunnelling current. However this splitting of the masses offers the opportunity for calibration in respect of source-to-drain tunnelling.

For the simulated set of double gate MOSFETs with relatively thin silicon bodies the quantum mechanical confinement results in a threshold voltage shift of approximately 0.16V, between the classical and the quantum (NEGF and DG) simulations. We also observe a noticeable degradation in the subthreshold slope with decreasing gate length. A significant fraction of this effect at room temperature is associated with the drain induced barrier lowering (DIBL). The displacement of the charge centroid from the Si/SiO_2 interface in the quantum simulations, shown in figure 2, results in an increase in the *effective* oxide thickness and reduced *effective* gate-to-channel capacitance. Thus the confinement effects contribute to the DIBL and the degradation of the subthreshold slope. It becomes rather difficult, particularly at room temperature, to disentangle the contribution of the source-to-drain tunnelling from the degradation of the subthreshold slope from the conventional DIBL and its enhancement due to vertical confinement.

We further examine the contribution of the source-to-drain tunnelling in subthreshold regime by studying the temperature dependence of the subthreshold slope, since any tunnelling current will exhibit a temperature dependence weaker than the direct temperature proportionality (S~kT/q) of the thermionic current limited subthreshold slope S. The simulation results at 300K and 200K for the two extreme gate lengths, 20 and 4 nm are shown in figures 5 and 6 respectively.

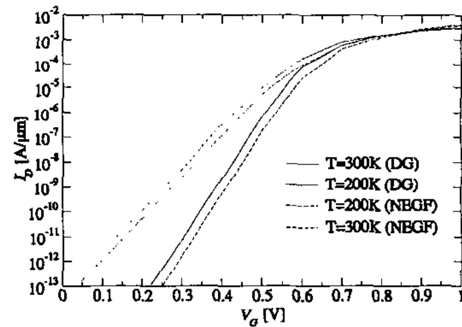

Figure 5: I_D-V_G characteristics for a 20nm channel length double gate structure from density gradient and NEGF simulations, at 200K and 300K.

The subthreshold current in the 20nm channel length device exhibits the expected thermionic current behaviour with subthreshold slope increasing by a factor of 1.5, which exactly corresponds to the temperature ratio. As Figure 6 shows, the subthreshold current for the 4nm channel length device has, however, a much weaker dependence on temperature, with increase in the subthreshold slope of approximately 1.2, smaller than the expected 1.5 ratio for

thermionic current. This is indicative of the existence of a second current transport mechanism in the subthreshold region in addition to the classical over-barrier (thermionic) current. This is consistent with the expectation of a source-drain tunnelling current contribution at 4 nm channel length, since the tunnelling current exhibits a much weaker dependence on temperature [15] than a classical thermionic emission current.

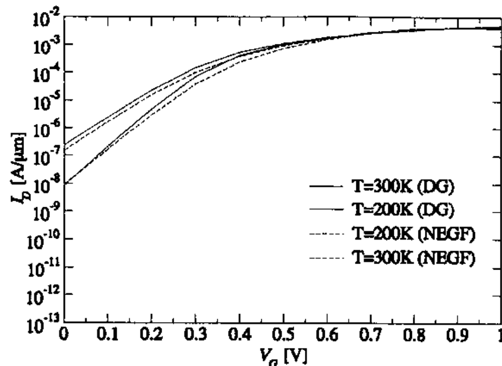

Figure 6: I_D-V_G characteristics for a 4nm channel length double gate structure from density gradient and NEGF simulations, at 200K and 300K.

The ability to fit the current characteristics obtained from detailed NEGF simulations with only two adjustable parameters, over a wide range of biases and temperatures, gives us confidence that the density gradient approach can be used to predict the quantum behaviour of devices scaled to 10 nm and below. The results also clearly indicate that it is important to use two different parameters (effective masses) to calibrate the device current characteristics in respect of tunnelling and confinement independently.

It is likely that the double gate devices will still be able to function satisfactorily below the 10nm barrier, with tunnelling not being such a dramatic scaling limitation as previously postulated. This is in agreement with recently published results [16], where they considered channel-lengths down to a nominal 0 nm.

V. CONCLUSIONS

Using comprehensive NEGF simulation results we have calibrated our 3D DD simulator with DG quantum corrections in respect of confinement effects and source-to-drain tunnelling using different effective masses in directions normal and parallel to current flow. The results presented here provide strong evidence that the density gradient approach can, at least to a first approximation, reproduce the important aspects of the I_D-V_G characteristics that are consistent with the presence of source-drain tunnelling. Both Density Gradient and Non-Equilibrium Greens Functions show that the source-drain tunnelling in sub-10nm double gate MOSFETs is expected to play an important role especially in the subthreshold regime. However, source drain tunnelling current may not be as large as has been previously speculated.

ACKNOWLEDGMENTS

JRW would like to acknowledge the support of EPSRC under grant No: GR/L53755. SHEFC Research Development Grant VIDEOS provided support for ARB. We also acknowledge support from IBM through a Shared University Research (SUR) grant.

REFERENCES

[1]. S. Jallepalli, J. Bude, W. K. Shih, M. R. Pinto, C. M. Maziar, and A. F. Tasch Jr., IEEE Trans. Electron Devices, **44**: 297-303, 1997.

[2] Z. Ren, R. Venugopal, S. Datta, M. Lundstrom, D. Jovanovic, and J. Fossum, IEDM Tech. Digest, 107-110, 2001.

[3] J. R. Watling, A. R. Brown and A. Asenov, Proc. IWCE-8, Journal of Computational Electronics, in press

[4] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, Journal of Applied Physics, **91**: 2343-2354, 2002.

[5] W. R. Frensley, Physical Review B, **36**: 1570-1580,1997

[6] S. E. Laux, Proc. IWCE-8, Journal of Computational Electronics, 2002, in press

[7] M. G. Ancona and G. J. Iafrate, Physical Review B, **39**: 9536-9540, 1989

[8] A. Asenov, J. R. Watling, A. R. Brown and D. K. Ferry, submitted to Journal of Computational Electronics.

[9] A. Asenov, *et al*, IEEE Trans. Electron Devices, **48**: 722-729, 2001

[10] J. R. Watling, A. R. Brown, A. Asenov and D. K. Ferry, Proc. SISPAD 2001, 82-85 2001

[11] H. Tsuchia, B. Winsted and U. Ravioli, VLSI Design, 13: 335-340, 2001

[12] F. Perrot, Physical Review B, **A20**: 586, 1979

[13] M. G. Ancona, IEEE Transactions on Electron Devices, **47**: 1449-1456, 2000

[14] S. Rafferty, B. Biegel, Z. Yu, M. G. Ancona, J. Bude, and R. W. Dutton, SISPAD'98, Eds De Meyer, and Biesemans, 137-140, 1998.

[15] H. Kawaura, T. Sakamoto, T. Baba, Y. Ochiai, J. Fujita and J. Sone, IEEE Trans. Electron Devices, **47**: 856-860, 2000

[16] A. Schenk, A. Wettstein, Modeling and Simulation of Microsystems 2002, Puerto Rico, 552-555.