Copyright © 2008 American Scientific Publishers
All rights reserved
Printed in the United States of America

Journal of
Computational and Theoretical Nanoscience
Vol. 5, 1–22, 2008

REVIEW

Ballistic Quantum Simulators for Studying Variability in Nanotransistors

A. Martinez[1,*], J. R. Barker[1], A. Svizhenko[2], A. Anantram[3], M. Bescond[4], and A. Asenov[1]

[1] *Department of Electronic and Electrical Engineering, University of Glasgow, Glasgow G12 8LT, UK*
[2] *SILVACO International, Santa Clara, CA 95054, USA*
[3] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada*
[4] *IM2NP UMR 6242 Bât. IRPHE Technopole de Château-Gombert 13384 Marseille Cedex 13, France*

This work will review the development of 2D/3D device simulators based on the Non-Equilibrium Green Function (NEGF) formalism, which has been developed in order to study variability issues. We have concentrated on the variability deriving from the particular arrangement of dopants, unintentional defects and roughness of the Si/SiO_2 interface. The device simulators developed have the capability to address very strong local potential inhomogeneities that stem from discrete dopants. We have implemented volume discretisation techniques that allow us a more precise representation of interfaces and dopant localization. The simulators have been parallelized and recursive algorithms have been implemented in order to reduce the computational time, which is a killer effect in the Poisson-NEGF loop in a 3D quantum simulator. The simulator is designed for an effective mass tensor approximation. Finally, we have carried out studies of the performance of nanotransistors due to the effect of the discreteness of dopant and interface roughness.

Keywords: Double Gate MOSFET, Nanowire MOSFET, Non Equilibrium Green Function, Surface Roughness, Stray Charge.

CONTENTS

1. Introduction 1
2. The Non Equilibrium Green Function Approach 4
3. The 2D NEGF Device Simulator and Applications 6
4. Impurity in Channel 6
5. Three Impurities in the Channel 9
6. Doping Inhomogeneity 11
7. Interface Roughness 14
8. The 3D NEGF Device Simulator and Applications 17
9. Conclusions 21
References 21

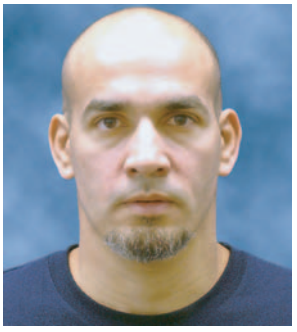
1. INTRODUCTION

Currently the semiconductor industry works at the 45 nm technology node, which has approximately 1 billion transistors per chip area (1.5 cm^2). The channel length of these transistors is around 30 nm. In the next few years the semiconductor industry will be targeting to double or triple the number of transistors per chip. The dimensions of the transistor active region will shrink to under 10 nm. Considering that the density of silicon is approximately 64 atoms per cubic nanometre, the number of atoms involved in the transport (along the channel) will be less than 40 atoms. At these dimensions transistor behaviour is mainly governed by the quantum statistical mechanics of several atoms coupled to a relatively large system, forming the contacts. It is crucial that the development of a new set of simulation tools will incorporate the relevant physics, in order to predict device behaviour in a reliable and accurate manner.

When a positive charge is located in the channel of a small nanotransistor, the repulsive potential of the charge acts to block the current flow, but some electrons are able to tunnel through the potential barrier changing the screening potential and the current density. It is clear that a treatment of these phenomena requires a quantum mechanical approach, and similarly for an accurate treatment of discrete dopants in source and drain. In the latter case the impurity potential is attractive but the concentration of electrons around the acceptor follows the wave nature of electrons dictated by their wavelength. Besides, at nanometre scale the interaction of the electron with the dielectric interface is more like to be dependent on the electron wavelength and the size of the spatial variation of the interface. The reflection at the interface is more like a scattering of waves than a particle bouncing on a stair and for small channel lengths the electron can reach

*Author to whom correspondence should be addressed.

the drain end even without scattering with phonons, keeping the transport coherent throughout the entire channel. In this case, the probability amplitude that an electron moves from source to drain is given as a superposition of the amplitudes that describe all the possible ways that the electron can move ballistically from the source to the drain. This picture shows clearly the difference between the quantum and the Monte Carlo approach to transport.



Antonio Martinez received in 1987 a M.Sc. degree in physics from the University of Havana, Cuba. In 1997, he received another M.Sc. degree in physics from the University of Umea, Sweden and a Ph.D. degree from the Royal Institute of Technology, Stokholm, Sweden, in 2003. His Ph.D. thesis deals with electron transport in anisotropic materials (SiC, GaN) at high electric field using various simulation techniques as drift-diffusion, ensemble Monte Carlo and quantum Monte Carlo. He is currently a member of Device Modelling Group at Department of Electronics and Electrical Engineering, University of Glasgow. He experience include Monte Carlo simulation technique for electron and hole transport, Nanoscale quantum electron transport using full 3D Non-Equilibrium Green's Functions approach, specifically simulation of nanoscale MOSFET and nanowire transistors.



John R. Barker has been Professor of Electronics at the University of Glasgow, Glasgow, Scotland UK since 1985 where he was a co-founder of the Nanoelectronics Research Centre with responsibility for the Theory and Modelling activities. He is a graduate of the University of Edinburgh and the University of Durham. He obtained his Ph.D. in Theoretical Solid state Physics from the University of Warwick in 1969 where he held a Personal SRC Fellowship. From 1970 to 1985 he was a lecturer then senior lecturer in Theoretical Physics at the University of Warwick, Coventry, UK. He has been a visiting Professor at Colorado State University, North Texas State University, and visiting scientist at The IBM T. J. Watson Laboratories. He is a Fellow of the Royal Society of Edinburgh and a Fellow of the Royal Astronomical Society. He has a broad range of research interests that currently include Green function modelling of nano-devices, silicon-germanium hetero-MOS, interface roughness modelling, smart dust and highly parallel computer architectures. He has over 300 publications in semiconductor physics, hot electron theory, quantum transport theory, Monte Carlo simulation, device physics and modelling, quantum theory of atomistic effects in ultra-small devices, parallel computing, olfactory sensors, molecular electronics.



Alexei Svizhenko received the Ph.D. degree in electrical engineering from the University of Nebraska, Lincoln, NE, in 1999, and M.S. and B.S. degrees from Moscow Institute of Physics and Technology, Moscow, Russia, in 1995 and 1993, respectively. In 1999–2005 he was a Staff Scientist National Aeronautics and Space Administration (NASA) Ames Center for Nanotechnology, Moffett Field, CA. In 2005–2006, he was a research scientist at Stanford University. In 2006, he joined SILVACO International as a Senior TCAD engineer. His research interests are in the modeling and computation of quantum transport in nanotransistors, silicon nanowires and carbon nanotubes. Dr. Svizhenko was the recipient of the 2000 NASA Group Achievment Award and the 2003 and 2004 Computer Sciences Corporation (CSC) Best Paper Award.



M. P. Anantram received the B.Sc. degree in applied sciences from the P. S. G. College of technology, India, in 1986, the M.Sc. Degree in physics from the University of Pune, India, in 1989, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1995. He was a postdoctoral fellow working on physical models of computing at Purdue University and the University of California, Los Angeles. He has been scientist at the NASA Ames Research Center, Mottfet Field, CA from 1997–2006, and he is currently Professor at the University of Waterloo, Ontario, Canada. His main research focus is modeling and theory of semiconductor and molecular nanostructures, such as nanotransistors, carbon nanotubes, and DNA. Dr. Anantram received the S. S. Joshi Award for academic exellence in 1989, the Veridian Technical Award for the best Paper, in 2000 for work on the electromechanical properties of carbon nanotubes, the NASA group achievement Award to the Nanotechnology Team, the CSC best paper award in 2002, for work on the electromechanical properties of carbon nanotubes, and the 2003 CSC technical excellence Award for work on modeling and theory of nanostructure.



Marc Bescond was born in Aubagne, France in 1978. He received the Master Degree in Material Engineering from INSA (National Institute of Applied Sciences) Lyon, France, in 2001, and the Ph.D. degree from the University of Provence, Marseille, France, in 2004 (thesis supervisor: Michel Lannoo). Since 2005, he is Associate Professor of Electrical Engineering and Physics at the INPG (Grenoble Institute of Technology). Before attending Grenoble, he was post-doc in the Device Modelling Group of Professor A. Asenov at the University of Glasgow. His research interest is in the modelling of nanoelectronic devices, three dimensional simulations of quantum transport and band-structure calculations.



Asen Asenov (M'96) received his M.Sc. degree in Solid State Physics from Sofia University, Bulgaria in 1979 and Ph.D. degree in Physics from The Bulgarian Academy of Science, Sofia, in 1989. He had 10 years industrial experience as a head of the Process and Device Modelling Group in IME-Sofia, developing one of the first integrated process and device CMOS simulators IMPEDANCE. In 1989–1991 he was a visiting professor at the Physics Department of TU Munich. He joined the Department of Electronics and Electrical Engineering at the University of Glasgow in 1991 and served as a Head of Department in 1999–2003. As a professor of Device Modelling, leader of the Glasgow Device Modelling Group and academic director of the Glasgow Process and Device Simulation Centre he coordinates the development of 2D and 3D quantum mechanical, Monte Carlo and classical device simulators and their application in the design of advanced and novel CMOS devices. He has pioneered the simulations and the study of various sources of intrinsic parameter fluctuations in decanano and nano CMOS devices including random dopants, interface roughness and line edge roughness. Professor Asenov is a fellow of the Royal Academy of Scotland and a member of the IEEE EDS TCAD Committee. He is currently a member of the programme committees for IEDM, ESSDERC, IWCE, SNW, HCIS and IEEE Nano Conf, and SNW2006. He has over 330 publications in semiconductor device physics on process and device modelling and simulation. He is an expert on atomistic effects in ultra-small devices and impact of variations on circuits and systems including more than 15 papers in IEEE Trans. Electron Devices in the last 5 years. He has also given more than 65 invited talks at prestigious international conferences and meetings in Europe, USA and Japan.

In the Monte Carlo approach the particle takes only one of the possible paths (minimum action).

Previous quantum transport simulations have been carried out, but most of these are using simulators, which simplify the transport to 1D,[1–4] or are using some kind of phenomenological models. Recently a full atomistic based simulator has appeared but the simulator is very time consuming and the implementation of dissipative physics in this simulator will increases the computational times considerably unless further physical approximations are made. Our approach uses a simple effective mass model Hamiltonian which folds the physics of the band structure into the effective masses. Based on this Hamiltonian we may calculate the Green function of the system under open boundary conditions to allow electron to be injected into the system (this method when applied to non-equilibrium conditions is called the Non Equilibrium Green Function (NEGF) approach and will be explained later). This approach is very similar to the atomistic or tight-binding one but uses essentially one effective orbital, instead of the 10 orbitals usually used in the atomistic case. The property of this orbital can be obtained from tight-binding band structure calculations or from first principles. This simulator is therefore fast and can incorporate the correct physics of confined system through the mass renormalization. We can incorporate full band effects due to the fact that the low-lying energy dispersion can be fitted by energy-dependent masses in our Hamiltonian. The future implementation of dissipative process such as phonon scattering is very natural except for the extra computational load.

As devices become ever smaller, two identical technologically fabricated devices, which differ in the particular microscopic atomic configurations, may perform in a completely different way, due to atomistic effects, which do not self-average away at such scales. In this paper we will concentrate on the development of a quantum simulator to study specifically two sources of variability: the variability that arises from the discreteness of the dopants and defect charges; and variability coming from the atomic scale roughness of the interface between the dielectric and the semiconductor.

In the case of the discrete dopants and stray charges, we have studied the effect of the position of a stray charge in the channel of a Double Gate (DG) MOSFET. We have also examined the effect on the performance of a DG MOSFET due to different spatial configurations of

3 unintentional charges in the channel of a nano-transistor. The variability of I_D–V_G characteristics has also been studied for the case of dopant fluctuations in source/drain.

The influence of the interface roughness between oxide and semiconductor on carrier transport increases particularly in thin-body double gate devices where the roughness is correlated to body thickness fluctuations.[5] The effect of interface roughness on the device performance was studied in the past, predominantly using Monte Carlo techniques with various surface roughness scattering models.[6–8]

When the body thickness of double gate transistors is on the order of a few nanometres the current flow is 2D due to the irregularity of *both* Si/SiO_2 interfaces. A careful treatment is therefore required in order to resolve the sub-nanometre scale features of the interface.

In this work, we have used a full (real space) 2D self-consistent NEGF simulator,[5, 9] while implementing a new discretisation scheme that allows us to describe precisely the interface profile. The same simulator has been used to simulate the effect of the dopants and stray charges in the device performance. The 2D work is concentrated on DG MOSFET structures.

We have also developed a full 3D real space simulator based on the NEGF formalism and we have made preliminary studies of the effect of interface roughness and a stray charge in a wrap-around nanowire transistor. The simulator uses a recursive algorithm[5, 9, 10] and the energy cycle has been parallelized. The code is written in Fortran.

Following the introduction, this paper is divided into four chapters; first we will give a glimpse into the NEGF technique as applied to nano-transistors including the recursive algorithm. The second and third chapters describe the 2D and 3D simulators respectively. Those chapters also present the applications of the simulator to study the various sources of fluctuation mentioned before. Conclusions are drown in the final chapter.

2. THE NON EQUILIBRIUM GREEN FUNCTION APPROACH

The NEGF approach to carrier transport in mesoscopic systems, comes from the marriage of two old ideas due to Schwinger-Keldysh (SK) and Landauer-Buttiker (LB).[11–14] The first idea, (SK), was the use of propagators and correlation functions (Green's functions) to describe the motion of electrons in non-equilibrium quantum statistical systems or systems coupled to non-coherent systems such as random oscillators or reservoirs. The second idea, LB, is based on the observation that when a difference in potential is applied to a mesoscopic system, which essentially consists of two or more reservoirs connected by (short channel length and cross section) quantum channels, the same difference in the chemical potentials between the respective reservoirs is induced and this determines the current, while the electrostatic potential is allowed to float (Neumann boundary conditions) in order to fulfil charge neutrality. The current flowing between two reservoirs is controlled by (quantum) transmission in the channel connecting the reservoirs, so implicitly the reservoirs are kept with a fixed difference in chemical potential between them. Under certain simplifying assumptions[15] the current flow between two reservoirs is proportional to the transmission function and the difference of chemical potentials between the reservoirs.[13, 15] Essentially, in this approach resistance is represented by transmission.[16]

In the modern NEGF formalism[17] the current and electron density in a mesoscopic system, when a difference in potential is applied, is computed with the help of the electron propagator G^R and correlation function $G^<$. The boundary conditions are included in the self-energy Σ^R (dynamic boundary condition) and statistical self-energy matrix $\Sigma^<$ (statistical boundary condition). The matrix equations (the Green functions are matrices in the coordinate representation) for the electron propagator and the correlation function may be written as:[17, 18]

$$\left[G_0^{-1} - \Sigma^R\right] G^R = I \tag{1}$$

$$G^< = G^R \Sigma^< G^A \tag{2}$$

$$n(E;\mathbf{r}) = -2iG^<(E;\mathbf{r},\mathbf{r})/\pi \tag{3}$$

$$\mathbf{J}(E;\mathbf{r}) = -\frac{e\hbar}{4\pi}\mathbf{m}^{*-1}\cdot(\nabla - \nabla')G^<(E;\mathbf{r},\mathbf{r}')\,|_{\mathbf{r}=\mathbf{r}'} \tag{4}$$

Where $G_0^{-1} = H - i\eta$ and H is the Hamiltonian of the finite device; Σ^R contains the effect of the injecting and absorbing (IA) boundaries on the open system. The statistical self-energy matrix $\Sigma^<$ is given by the imaginary part of the self-energy for each IA boundary multiplied by the distribution function at the boundary value of the chemical potential (including the external potential). Scattering is not considered. $iG^<(E;\mathbf{r},\mathbf{r})$ or $n(E;\mathbf{r})$ are the density of electrons at the position $\mathbf{r}$ with energy E; $\mathbf{J}$ is the energy-resolved current density associated with electrons of energy E. The total electron density is obtained by integrating equation (3) over all energies. Expression (4) also should be integrated for all energies to give the total current density. The electron density must satisfy Poisson's equation; therefore a self-consistent iterative cycle is set up to compute the electron density and potential throughout the device.

As can be seen from Eqs. (3) and (4), we only need to calculate the diagonal and off-diagonal elements of the $G^<$ function in order to compute the current and electron density. For this purpose we deploy a recursive algorithm technique based on the symmetry of the Dyson equation, which allows us to compute the diagonal, and off-diagonal blocks of $G^<$ without the need to invert the Hamiltonian matrix of the whole device. A detailed presentation and proof of the recursive algorithm used in this work may be found in Ref. [9]. In the following, we present the equations in a

simplified fashion, but they can directly apply for computational purposes, these formulas are essentially in Ref. [9], but a few misprints[19] may cause difficulty for implementation. If we call A the inverse matrix of the retarded Green function G^{r}, we may write the equations which allow the computation of the diagonal (density) and first off-diagonal block of $G^{<}$; these diagonal blocks represent the interaction inside the device cross section and the first off-diagonal represents the coupling of the cross section with the neighbouring cross section. The crucial point is the partition of the A matrix in blocks, which represent the interaction inside the cross section (A_{qq}) and the cross section–cross section interactions (A_{qq+1}). To calculate the G_{qq} and G_{qq+1} blocks of the retarded Green function, a series of auxiliary matrices need to be calculated recursively; there is a forward and a backward recursion, the steps are as follows:

$$g^{rLq+1}_{q+1,\,q+1} = \left[A_{q+1,\,q+1} - A_{q+1,\,q} g^{rLq}_{q,\,q} A_{q,\,q+1}\right]^{-1} \tag{5}$$

$$g^{rL1}_{11} = A^{-1}_{11} \tag{6}$$

$$q = 1, \ldots, N-1$$

Following these steps we can calculate the diagonal and first off diagonal blocks of G^{r} by backward recursion. They are used to compute the electron density

$$G^{r}_{N,\,N} = g^{rLN}_{N,\,N} \tag{7}$$

$$G^{r}_{q,\,q} = g^{rLq}_{q,\,q} + g^{rLq}_{q,\,q}\left(A_{q,\,q+1} G^{r}_{q+1,\,q+1} A_{q+1,\,q}\right) g^{rLq}_{q,\,q} \tag{8}$$

$$G^{r}_{q+1,\,q} = -G^{r}_{q+1,\,q+1} A_{q+1,\,q} g^{rLq}_{q,\,q} \tag{9}$$

$$q = N-1, \ldots, 1$$

The $G^{<}_{q,\,q+1}$ and $G^{<}_{q,\,q}$ are calculated from the above matrices, using the following forward and backward recursion schemes; the forward recursion computes auxiliary matrices using a forward algorithm:

$$g^{<Lq+1}_{q+1,\,q+1} = g^{rLq+1}_{q+1,\,q+1}\Bigg(\sum^{<}_{q+1,\,q+1} + \sigma^{<}_{q+1} - \sum^{<}_{q+1,\,q} g^{aLq}_{q,\,q} A^{+}_{q,\,q+1} - A_{q+1,\,q} g^{rLq}_{q,\,q} \sum^{<}_{q,\,q+1}\Bigg) g^{aLq+1}_{q+1,\,q+1} \tag{10}$$

$$\sigma^{<}_{q+1} = A_{q+1,\,q} g^{<Lq}_{q,\,q} A^{+}_{q,\,q+1} \tag{11}$$

$$g^{<L1}_{11} = g^{r0}_{11} \sum^{<}_{11} g^{a0}_{11} \tag{12}$$

$$g^{r0}_{11} = A^{-1}_{11} \tag{13}$$

$$q = 1, \ldots, N-1$$

A backward recursion is used to finally computed the $G^{<}_{q,\,q}$ and $G^{<}_{q,\,q+1}$:

$$G^{<}_{q,\,q} = g^{<Lq}_{q,\,q} + g^{rLq}_{q,\,q} A_{q,\,q+1} G^{<}_{q+1,\,q+1} A^{+}_{q+1,\,q} g^{aLq}_{q,\,q} - g^{<Lq}_{q,\,q} A^{+}_{q,\,q+1} G^{a}_{q+1,\,q} - G^{r}_{q,\,q+1} A_{q+1,\,q} g^{<Lq}_{q,\,q} - g^{<0}_{q,\,q+1} A^{+}_{q+1,\,q} G^{a}_{q,\,q} - G^{r}_{q,\,q} A_{q,\,q+1} g^{<0}_{q+1,\,q} \tag{14}$$

$$g^{<0}_{q,\,q+1} = g^{r0}_{q,\,q} \sum^{<}_{q,\,q+1} g^{a0}_{q+1,\,q} \tag{15}$$

$$g^{<0}_{q+1,\,q} = g^{r0}_{q+1,\,q+1} \sum^{<}_{q+1,\,q} g^{a0}_{q,\,q} \tag{16}$$

$$G^{<}_{N,\,N} = g^{<LN}_{N,\,N} \tag{17}$$

$$g^{r0}_{q,\,q} = A^{-1}_{q,\,q} \tag{18}$$

$$G^{<}_{q+1,\,q} = g^{<0}_{q+1,\,q} - G^{r}_{q+1,\,q} A_{q,\,q+1} g^{<0}_{q+1,\,q} - G^{<}_{q+1,\,q+1} A^{+}_{q+1,\,q} g^{gLq}_{q,\,q} - G^{r}_{q+1,\,q+1} A_{q+1,\,q} g^{<Lq}_{q,\,q} \tag{19}$$

$$q = N-1, \ldots, 1$$

The above equations form the generalized recursive algorithm for calculating the diagonal and off diagonal block of Eq. 2.

We are using a simplified version of the recursive algorithm due to the fact that the $\Sigma^{<}$ is diagonal. Equations (20) and (22) are solved in a forward recursion and Equations (23)–(25) are solved in a backward recursion. Those equations are solved following the order in their equation number.

$$g^{rLq+1}_{q+1,\,q+1} = \left[A_{q+1,\,q+1} - A_{q+1,\,q} g^{rLq}_{q,\,q} A_{q,\,q+1}\right]^{-1} \tag{20}$$

$$g^{<Lq+1}_{q+1,\,q+1} = g^{rLq+1}_{q+1,\,q+1}\left(\sum^{<}_{q+1,\,q+1} + \sigma^{<}_{q+1}\right) g^{aLq+1}_{q+1,\,q+1} \tag{21}$$

$$G^{r}_{q+1,\,q} = -G^{r}_{q+1,\,q+1} A_{q+1,\,q} g^{rLq}_{q,\,q} \tag{22}$$

$$G^{<}_{q+1,\,q} = -G^{<}_{q+1,\,q+1} A^{+}_{q+1,\,q} g^{aLq}_{q,\,q} - G^{r}_{q+1,\,q+1} A_{q+1,\,q} g^{<Lq}_{q,\,q} \tag{23}$$

$$G^{<}_{q,\,q} = g^{<Lq}_{q,\,q} + g^{rLq}_{q,\,q} A_{q,\,q+1} G^{<}_{q+1,\,q} - g^{<Lq}_{q,\,q} A^{+}_{q,\,q+1} G^{a}_{q+1,\,q} \tag{24}$$

$$G^{r}_{q,\,q} = g^{rLq}_{q,\,q} - g^{rLq}_{q,\,q} A_{q,\,q+1} G^{r}_{q+1,\,q} \tag{25}$$

Note that Eqs. (23) and (24) compute the blocks of $G^{<}$, which are necessary in order to calculate the density and current (Equations (3) and (4)). The simplified recursive algorithm is used in the 2D and 3D simulator. Essentially the same algorithm can be used for the implementation of scattering if the corresponding self-energies are diagonal.

As we mention before we used a control volume element discretisation to allow a more precise treatment of the interface, the variation of the dielectric constant and the variation of the effective mass from element to element. An element in 2D is the rectangle formed by the four nodes (for a rectangular mesh). In 3D this is the region inside the 8 nodes forming a cuboid. An elliptical second order partial differential equation can be put in a canonical form (not mixed derivatives). Using Gauss's law in the node bounded for elements we can integrate the Laplace operator in the Poisson and in the Schrödinger equation producing a discretised equation which takes into account the difference in dielectric constants of neighbouring elements. In the following we present the explicit steps to

REVIEW

get the element discretisation equation for the case of the Poisson equation:

$$\nabla \cdot (\varepsilon(r)\nabla\varphi) = \rho \tag{26}$$

$$\iiint_V \nabla \cdot (\varepsilon(r)\nabla\varphi)\, dV = \oiint_{\partial V} \varepsilon(r)\nabla\varphi \cdot ds \tag{27}$$

$$\oiint_{\partial V} \varepsilon(r)\vec{E} \cdot ds = \iiint_V \rho\, dV \tag{28}$$

Here ε is the dielectric constant, ρ the total charge density, φ the electrostatic potential and E the electric field. Equation (28) is discretised around the nodes and used in the simulators. This equation contains the continuity of the displacement field and therefore the polarization charge.

3. THE 2D NEGF DEVICE SIMULATOR AND APPLICATIONS

This section describes simulation studies of variability mechanisms in Double Gate MOSFET using the 2D NEGF simulator. As it has been mentioned before, the 2D simulator uses control volume discretisation. The equation which follows, represent explicit the basis transformation step of the volume discretisation for the Laplacian operator and the retarded Green function.

$$\int HG\,ds = \frac{\hbar^2}{2}\int \nabla\left(\frac{1}{m}\nabla G\right)ds = \frac{\hbar^2}{2}\int\left(\frac{1}{m}\nabla G\right)\cdot n\, dl \tag{29}$$

Figure 1 illustrates the control volume discretisation (in this case the area) for the retarded Green function equation around the node (i, k). The four parameters m1, m2, m3 and m4 represent the effective masses of the neighbouring elements to that node. The dashed line encloses the area used in the integration to eliminate one of the derivatives. The expression on the far right in the Eq. (29) is discretised, but now in the dashed line, which is always inside an element with constant mass. In the case of the Poisson equation, instead of different masses we will have different dielectric constants for each element. Figure 2 shows a section of the Si/SiO_2 interface. Six nodes are highlighted in red showing the correspondence with the 6 nodes of Figure 1. The yellow or light colour represents the SiO_2 region, and the grey or dark colour represents the Silicon region. The central node has two elements in Si and two elements in the SiO_2. The effective masses in Si, and the dielectric constants are different from those in SiO_2. In our code, properties such as doping are associated with elements and not with nodes, this makes clearer the characterisation of the interface. In the 2D case the matrix that represent the cross section Eq. (9d) is tri-diagonal, making the inversion straightforward.

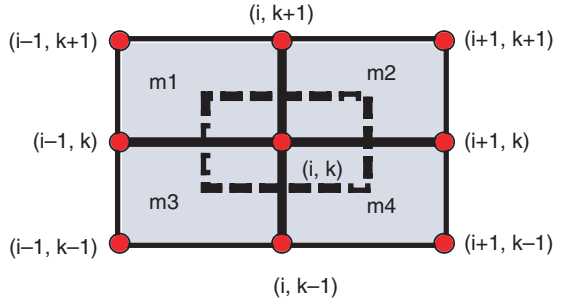


Fig. 1. Control volume discretisation scheme for the Green function operator showing the average on the node (i, k) of the gradient of G.

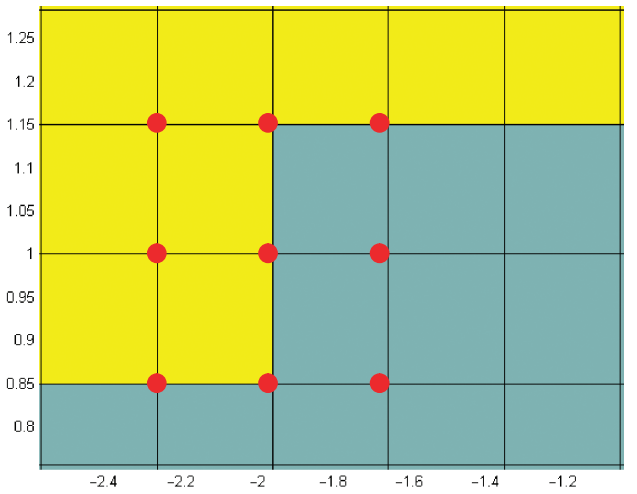


Fig. 2. A piece of the SiO_2/Si rough interface, yellow and green stand for SiO_2 and Si regions respectively. Red dots stand for the discretisation nodes.

4. IMPURITY IN CHANNEL

We will start with the effect of a stray charge in the channel. The device considered is a Silicon DG MOSFET with SiO_2 gate oxide. The transistor structure is shown in Figure 3. The channel length is 10 nm, which justifies the use of a ballistic approach. The channel thickness is 3 nm; the doping in the source and drain regions, N_D,

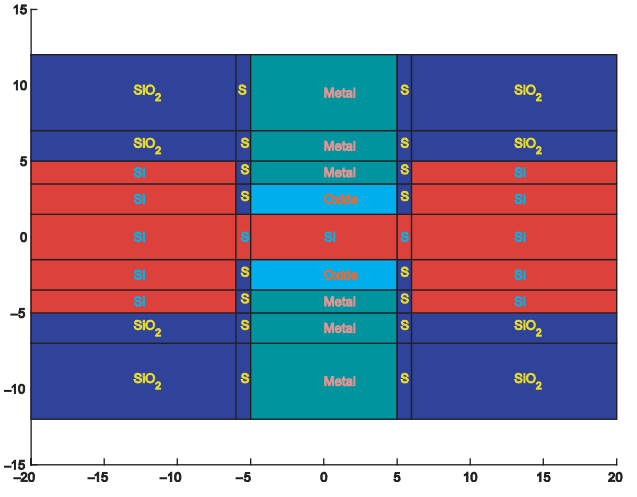


Fig. 3. Structure of the double gate MOSFET.

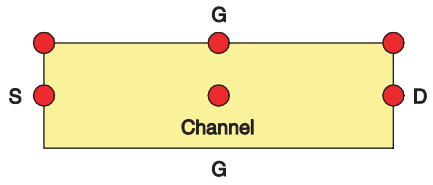


Fig. 4. Location of charges investigated, S, D and G stand for source, drain and gate.

is 10^{20} cm^{-3}; and the background channel doping, N_A, is 10^{14} cm^{-3}. In general, the location of the single stray charge in the device is completely random. However, in the case of our 2-dimensional simulations we consider six different key cases of charge location illustrated in Figure 4. The circles in this figure represent the different positions of the stray charge considered in this work. Those cases are chosen because they represent worst-case scenarios or because they are a representative case. Cases not considered have the same qualitative behaviour as the cases considered or an intermediate behaviour between the cases considered.

The resulting I_D–V_G characteristics for the six cases studied are shown in Figure 5. Through most of the range of applied gate voltages the largest change in threshold voltage (relative to the control device with no stray charge) is produced by a charge placed in the middle of the channel. The reduction in current due to this charge is very high at low gate voltages but as the gate voltage increases the screening effect from the charge in the channel and increasing tunnelling reduce the blocking effect. Due to the applied drain bias, when the impurity is located in the middle of the source boundary the reduction in current is larger than the case where the charge is at the middle of the drain interface.

As the gate voltage is increased, and the potential barrier lowered, the influence of the charge located in the very centre of the channel is reduced. As a consequence, at very high gate voltage the location with the largest impact is at the source end of the channel.

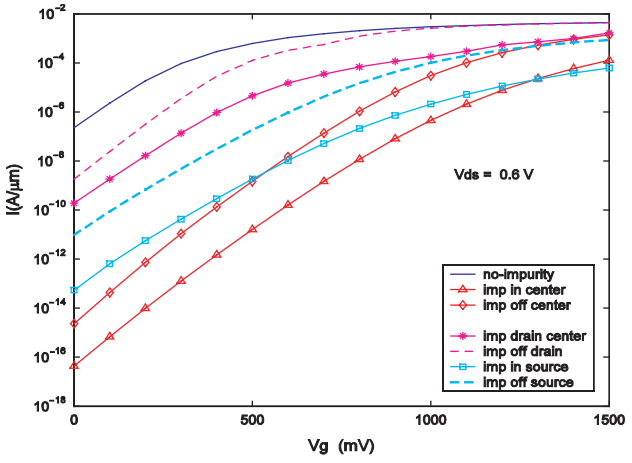


Fig. 5. I_D–V_G curves showing the effect of a fixed charge at different points within the channel. $V_D = 0.6$ V.

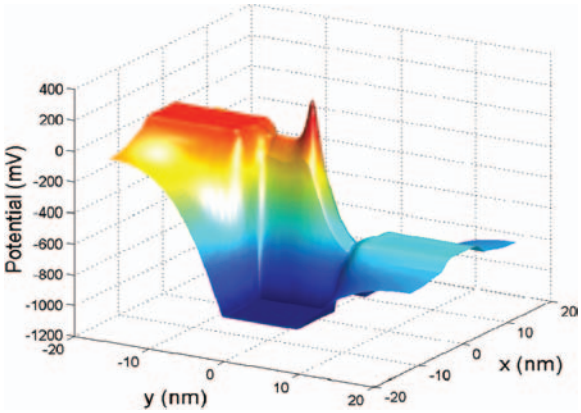


Fig. 6. 2D potential distribution with a charge in the centre of the channel. $V_G = 1.5$ V and $V_D = 0.6$ V.

The 2D potential profile for the case of a centrally located charge is shown in Figure 6, for $V_G = 1.5$ V and $V_D = 0.6$ V. The potential profile along the centre line of the transistor, midway between the two gates is shown in Figure 7. The height of the potential barrier of the charge located close to the oxide interface is very small in the middle of the channel at high gate voltage as shown here. This is due to the screening effect of the mobile charge in the channel. In the sub-threshold regime, where screening does not play a part, the effect of the charge at the interface is felt in the middle of the channel and the current reduction is substantial.

Figure 8 shows the density for one electron with energy of 0.7 eV coming from the source showing the penetration through the potential from a fixed charge in the middle of the channel. This high energy electron distribution shows few oscillations in the channel corresponding to the short wavelengths. But even the high energy electrons show their wave nature by reflecting from the stray charge potential barrier. Low energy electrons are expected to show even stronger wave behaviour. This shows to some extent the quantum mechanical effects taking place at microscopic levels and for a single carrier. More severe effects come

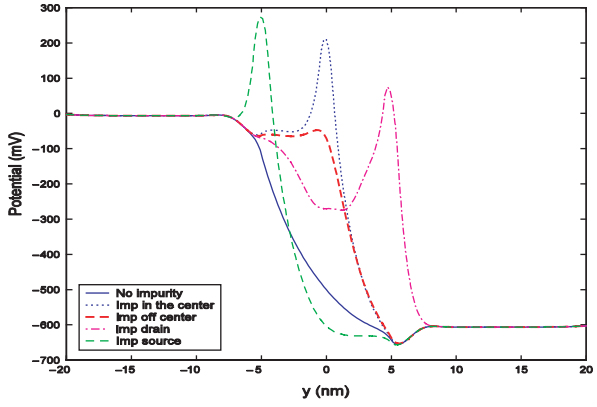


Fig. 7. Potential profile through the middle of the channel with different locations of fixed charge.

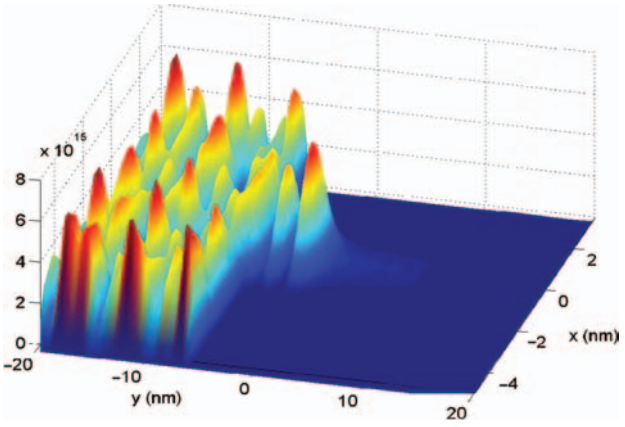


Fig. 8. Density for one electron with energy of 0.7 eV coming from the source and reflecting from the potential barrier resulting from a stray charge located in the middle of the channel.

with the addition of several charges in the channel, or inhomogeneity of the interface caused by interface roughness. It is shown in the figure that there is a tunnelling penetration into the potential resulting from the charge at the point $x = 0.0$, $y = 0.0$. However, the probability to find an electron drops off rapidly for positive y coordinates.

Figure 9 shows the electron concentration through the channel, at the middle of the channel, and at 0.6 and 1.2 nm off the middle of the channel. As is expected the density decreases away from the centre of the channel. The picture shows clearly a substantial electron density in the region where the stray charge is located, and where classically a significant reduction in electron concentration would be expected. These electrons will contribute to the current and will affect the I–V curves dramatically when compared to classical simulations.

The current density for this case, with the stray charge in the middle of the channel, is shown in Figure 10. In the central portion of the channel the current partially flows around the potential "spike" resulting from the charge, however there is also a tunnelling current flowing through the centre of the channel.

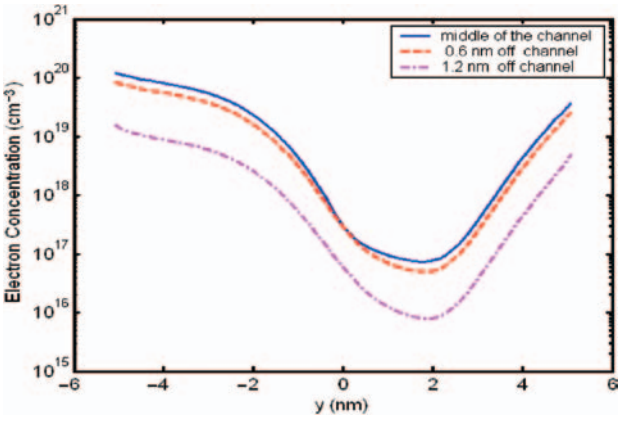


Fig. 9. The electron concentration along the channel for the case of the stray charge in the middle at $V_G = 1.2$ V.

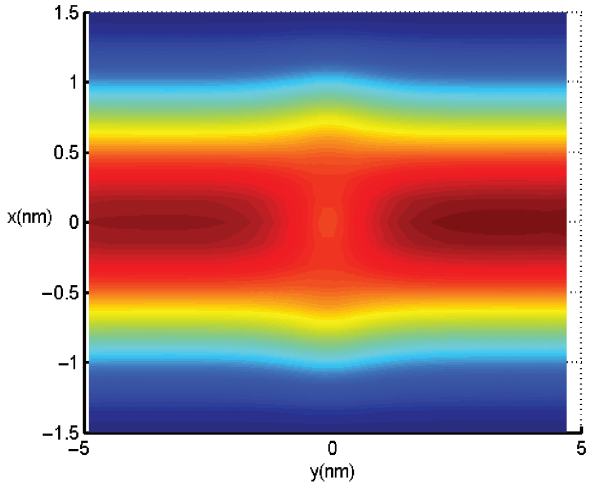


Fig. 10. The 2D-surface plot of the current density in the transport direction for the case of Figure 6.

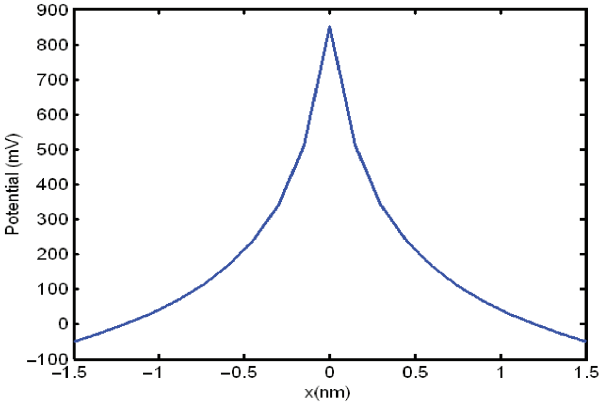


Fig. 11. The potential profile through the channel cross section at the middle of the channel for the case of Figure 7.

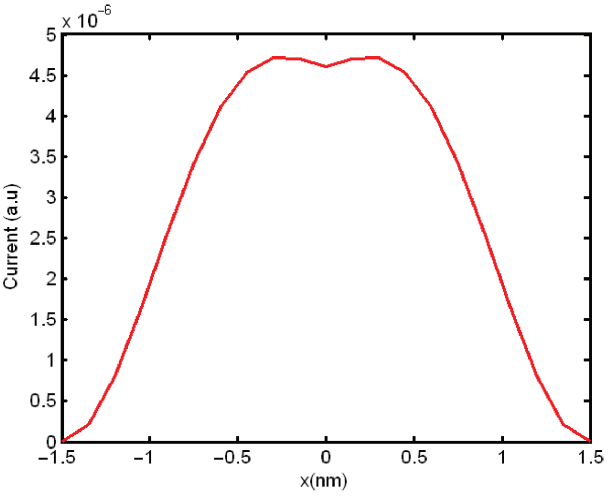


Fig. 12. The current density in the transport direction (y) plotted along a cross section in the channel halfway between source and drain.

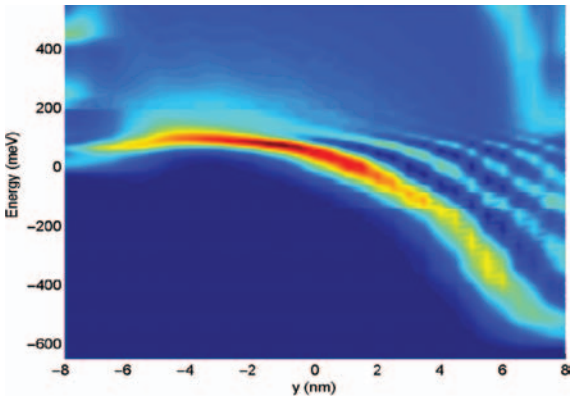


Fig. 13. Density of states through the middle of the channel at $V_G = 1$ V without stray charge.

The potential and current density through the cross-section between the two gates are shown in Figures 11 and 12. Figure 12 shows that approximately 30% of the total current flows through a high potential barrier at $x = 0$ in Figure 11 which indicates that the tunnelling current cannot be ignored.

Figures 13 and 14 show the density of states along the centre line of the transistor, midway between the two gates for a gate potential of 1.0 V without and with the fixed charge present, respectively. In the case of Figure 13 we can see the interference effect at the drain and the density of electrons in the first sub-band. In Figure 14 the energy of the conduction band edge in the middle of the device has increased because of the confinement and the potential barrier introduced by the stray charge. The figure shows interference patterns at the source and drain. A high density of states is located between the source and the stray charge because of the formation of quasi-bound states as can be inferred from the potential profile in Figure 7.

As conclusions, threshold voltage and the sub-threshold slope of the I_D–V_G characteristics are strongly affected by the position of the charge in the channel. The sub-threshold slope in the I_D–V_G curve depends more on the distance of the charge from the source than the in the direction perpendicular to the gates. The tunnelling of carriers through the stray charge potential plays an important role in the I–V characteristics. There is a substantial tunnelling current through the stray charge at high gate voltage. The current density is pushed to the oxide interface by the repulsive potential of the stray charge.

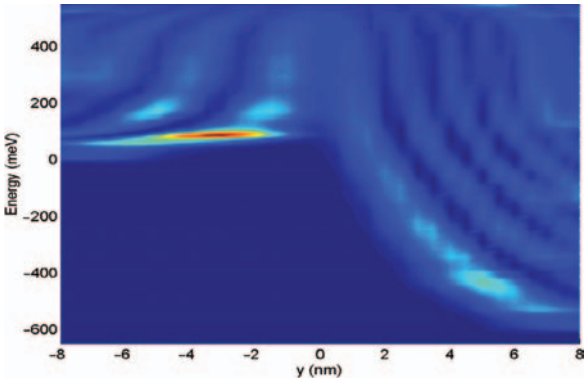


Fig. 14. Same as Figure 7 but with stray charge placed at the centre of the channel.

5. THREE IMPURITIES IN THE CHANNEL

We have also studied the effect on the I_D–V_G characteristics of various spatial configurations of 3 acceptor dopants in the channel of a nano-transistor.

The simulated device is a DG MOSFET at 300 K with metal gate and 1 nm oxide thickness. The channel length of the transistor is 12 nm that justifies the use of a ballistic approach. The channel body thickness is 5 nm, the doping in the source and drain regions, N_D, is 10^{20} cm^{-3}, and the channel doping, N_A, is 10^{14} cm^{-3}. Four devices which differ in the arrangement of three point charges in the transistor channel have been considered in the calculation. We have used both multi-valley anisotropic and isotropic effective mass models and compared the results. Figure 15 shows the potential for four different configurations of three dopants in the channel of the transistor at $V_G = 1.5$ V. The source and drain are located at the left and right of every panel. The two gates are located at the top and the bottom of the panels.

Figure 16 shows the I_D–V_G characteristics of the four devices using isotropic and anisotropic mass models at $V_D = 0.6$ V. The device with the vertical arrangement produces the lowest current because the channel is completely blocked by the potential barrier created by the impurities.

In the horizontal arrangement of dopants the blocking of the current is very effective at low gate voltages but at high bias there is a substantial lowering of the barrier and the current flow is concentrated close to the gates. The blocking by the triangle configuration is very effective at low gate voltages but as the gate voltage increases the potential barrier decreases substantially because the charges are mainly located in the middle of the channel. The symmetric triangle configuration has two point charges very close to the source that block the current effectively and are less sensitive to the gate voltage. As a result of these effects the I_D–V_G curve for the symmetric triangle is crossed by the I_D–V_G curve for the case of the asymmetric triangle as can be seen in Figure 16. The current for the anisotropic mass device models are larger than for the isotropic mass models due to a smaller threshold voltage shift associated with the large effective mass components in the confining direction and the increase of tunnelling due to small effective mass components in the transport direction. The horizontal arrangement of dopants displays the largest difference in currents between the two effective mass models. This follows from the high degree of confinement in the channel created by the alignment of the impurities along

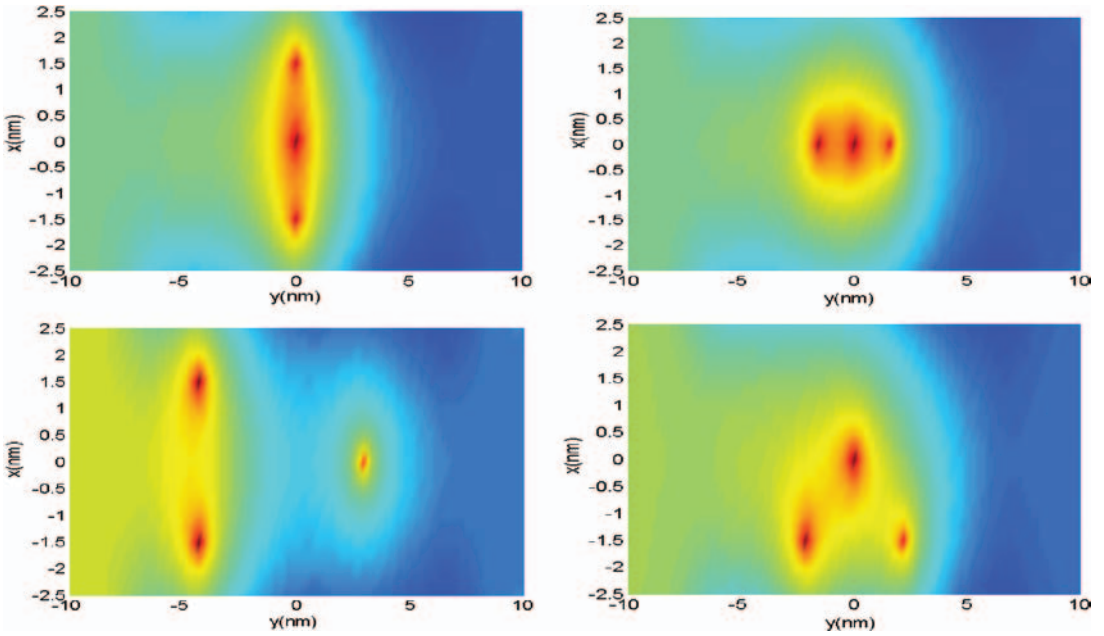


Fig. 15. The potential landscape (at $V_G = 1.5$ eV) for the four arrangements of the three charges in the channel: upper left panel (vertical configuration), upper right panel (horizontal configuration), lower left panel (symmetric triangle configuration), and lower right panel (triangle configuration).

the channel. For the device with the symmetric triangle configuration Figure 17 shows a comparison of the current density vector fields between the isotropic and the anisotropic effective mass models. It is clear that for the anisotropic mass model the current flow is more horizontal compared to the isotropic mass model. This is mainly due to the low effective mass in the transport direction and the high effective mass in the confined direction for the anisotropic model that act to reduce the meandering of the current flow through the potential landscape.

The current density profile $J_y(x, y)$ at low gate voltage ($V_G = 1.5$ V) is illustrated in the right panel of Figure 17 for the symmetric triangle configuration. It shows that the majority of the current flow is between the two foremost dopants followed by bifurcation of the flow around the farthest dopant. At very high gate voltage ($V_G = 2.4$ V) the current flow concentrates close to the gates in two parallel streams that avoid the dopants. This high voltage behaviour does not occur for the isotropic mass model. The left panel of Figure 18 shows the electron density in the channel. A charge density of approximately 10^{19} cm^{-3} is present in the dopant locations indicating electron tunnelling.

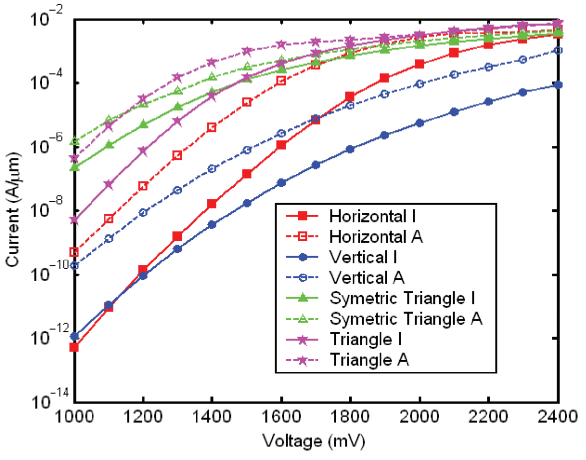


Fig. 16. I_D–V_G characteristics of the four devices (see Fig. 1) using the isotropic (I) and the anisotropic (A) mass models at $V_D = 0.6$ V.

As a conclusion, the threshold voltage and the subthreshold I_D–V_G slope are strongly dependent on the particular arrangement of a small number of point charges inside the channel of a short channel double gate MOSFET. Distortion of the ideal uniform quasi 1D flow through the channel occurs due to tunnelling, diffraction and interference from the potential landscape of the discrete dopants. The distortion is significantly smaller for the realistic multi-valley anisotropic mass model than in the isotropic mass model. The differences between the isotropic and anisotropic models are strongly dependent on the specific charge configurations. The current flow through the device shows a non-classical non-uniform 2D character: the current meanders through the channel potential landscape created by the point charges. There is significant tunnelling through the potential barrier created by the unintentional doping charge at high gate voltage. For the perfect channel the underlying transmission function is a multi-stepped function of carrier energy. For the imperfect device the transmission functions have reduced slope and the steps are smoothed out due to scattering and tunnelling processes in the potential landscape induced by the dopants.

These results are for a 2D model: a worst-case scenario for a wide 3D device. A full 3D model would be expected

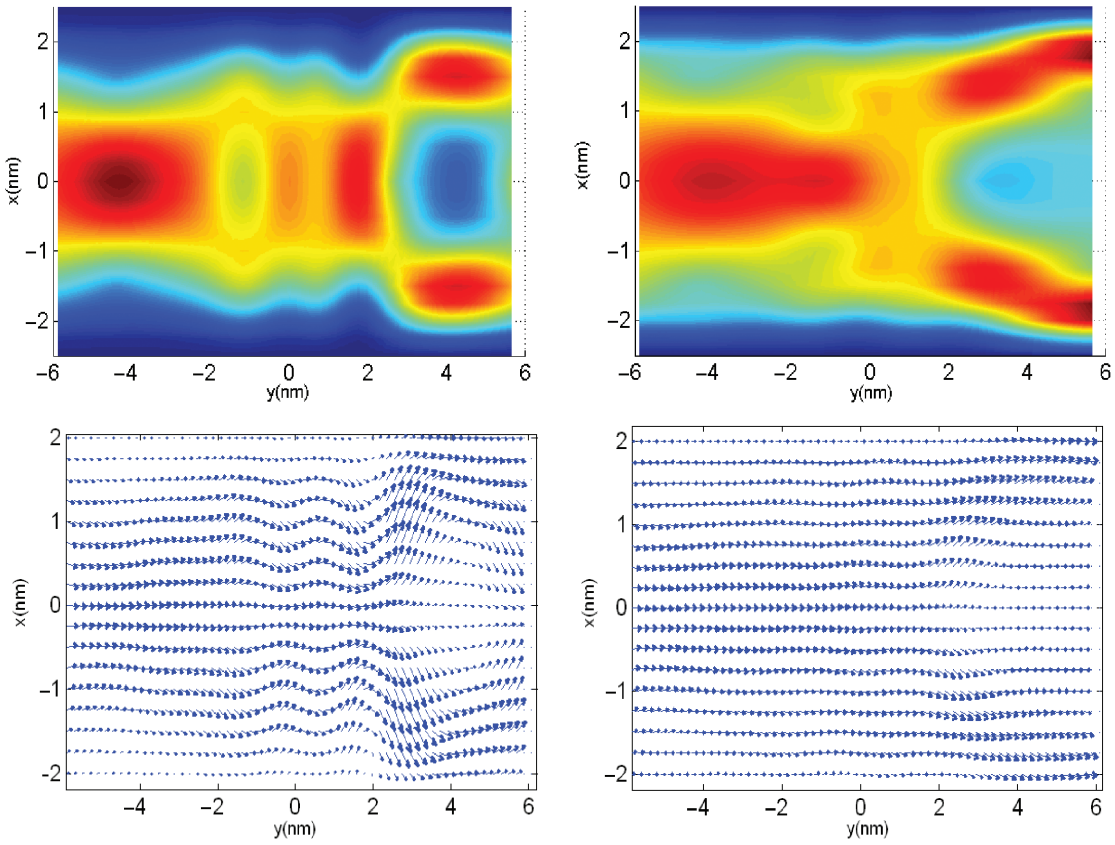


Fig. 17. The current density J_y in the channel of the symmetric triangle configuration device at $V_G = 1.5$ V. The upper left and right panels represent the current landscape for the isotropic and anisotropic effective mass model; the two lower panels show the current density vector field J respectively.

to exhibit more scope for the current flow to avoid the discrete point dopants.

6. DOPING INHOMOGENEITY

In this section we have assumed that the doping in source and drain is inhomogeneous and/or forms clusters. The device studied is a DG MOSFET shows in Figure 19, S/D dopings is 2e20 cm^{-3}, and the channel is un-doped. Channel length and width are 12 nm and 5 nm respectively. We have studied several devices that differ in the spatial configuration of dopant clusters in source and drain, and in the number of clusters.

To set the background for the device characteristics we first present the potential distribution of one of the devices with a random distribution of clusters (upper panel of Fig. 20). The corresponding electron distribution in the channel and the discrete dopant aggregation regime

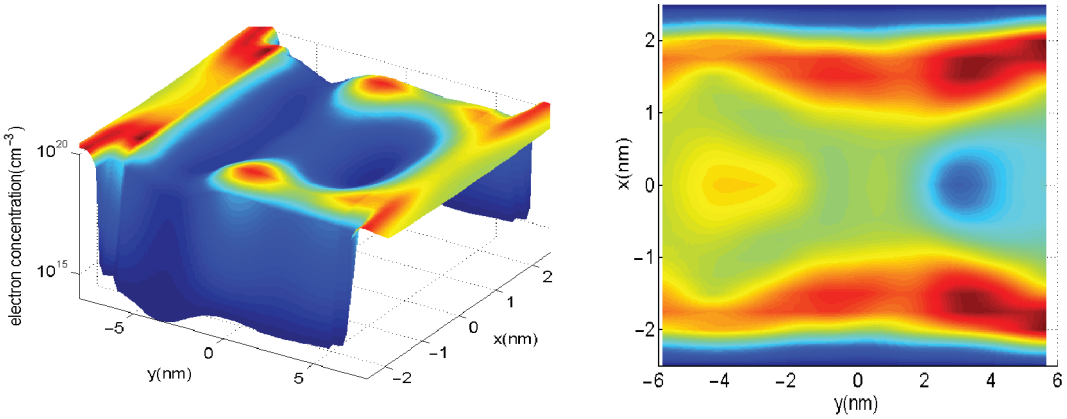


Fig. 18. The left and right panel show the electron density and the current density profile J_y respectively for the symmetric triangle device using the anisotropic model at $V_G = 2.4$ V.

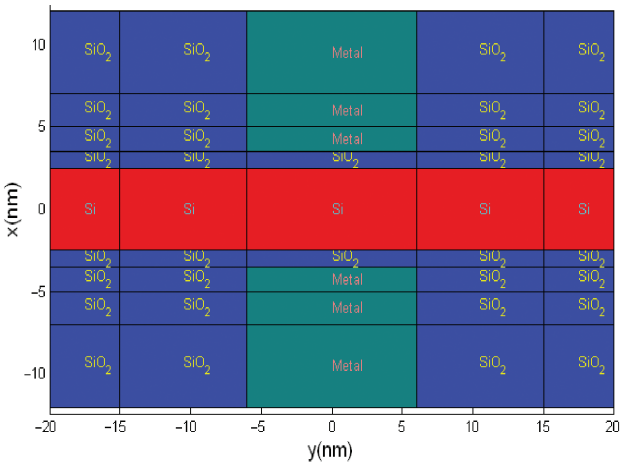


Fig. 19. Double gate MOSFET structure (DG-MOSFET).

(clustering) of source and drain is shown in the lower panel of Figure 20. Similar data is shown in Figure 21 but for fewer clusters (i.e., an effective lower average doping). The corresponding configurations of the clusters are presented in Figure 22 (the upper panel corresponds to Fig. 20; the lower panel corresponds to Fig. 21). Figure 23 shows the I_D–V_G characteristics for 5 devices that differ only in the arrangement of clusters ($\triangle N = 0$). Data for the case of uniform continuum doping (smooth doping distribution) is included for comparison. Figure 24 shows two cluster configurations with the same number of dopants corresponding to Figure 23. Figure 25 shows the I_D–V_G characteristics for a different number of total clusters ($\triangle N \neq 0$) ranging from 13 to 35 discrete cluster locations. Again, data for the case of continuum doping (smooth doping distribution) is included for comparison. In Figure 26 we represent the

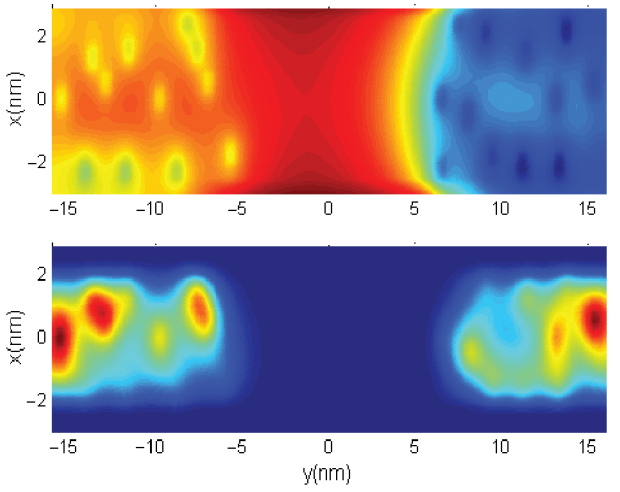


Fig. 20. Spatial profile of potential (upper panel) and electron concentration (lower panel) in the source, channel and drain regions of a 12 nm channel length DG-MOSFET for the dopant aggregation configuration shown in Figure 6 (upper panel). Average doping concentration in S/D = 10^{20} cm^{-3}, number of high doping clusters $\langle N \rangle = 10$; actual number of high doping clusters $N = 12$.

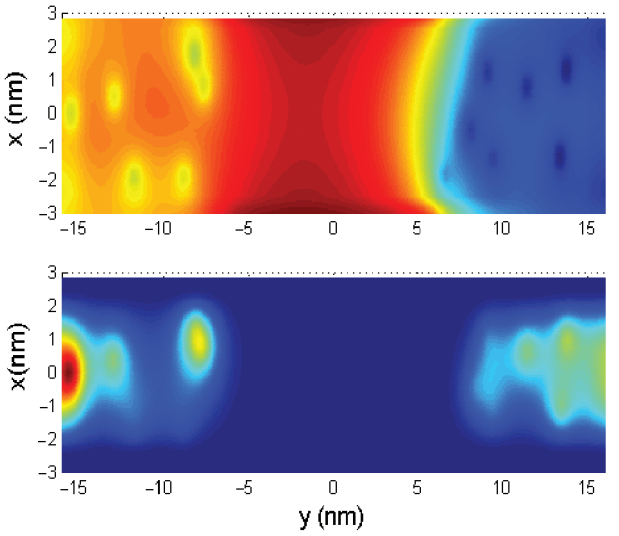


Fig. 21. Spatial profile of potential (upper panel) and electron concentration (lower panel) in the source, channel and drain regions of a 12 nm channel length DG-MOSFET for the dopant aggregation configuration shown in Figure 6 (lower panel). Average doping concentration in S/D = 10^{20} cm^{-3}, number of high doping clusters $\langle N \rangle = 10$; actual number of high doping clusters $N = 6$.

current density J_x (lower panel) through the self-consistent potential landscape (upper panel) corresponding to the device with the doping configuration shown in the lower panel of Figure 24.

From Figures 20, 21 and 22 it becomes clear that the self-consistent potential landscape effectively obscures the bare cluster potentials. It should be stressed that this result is a consequence of the thermal averaging (at 300 K) over the quasi-continuum of energy states inherent in the Green function simulation. Sharper features are found in the spectrally-resolved Green functions $G(E)$ which lead to complex features in the energy-resolved electron concentration $n(E; \mathbf{r})$. The total electron concentration $n(\mathbf{r})$ and hence the total self-consistent potential distribution results from the thermal average of $n(E; \mathbf{r})$ which obscures many of the obvious quantum effects in $n(E; \mathbf{r})$. However, full self-averaging does not take place and the source and drain regions, for example, exhibit inhomogeneous

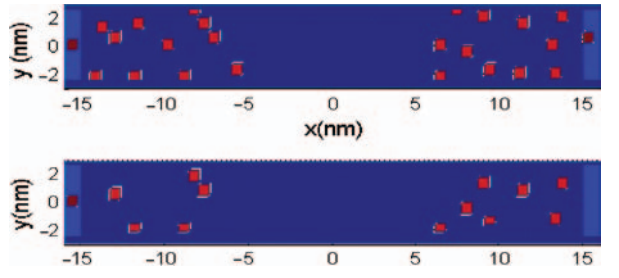


Fig. 22. Doping cluster aggregation map for DG MOSFET source and drain regions. Upper panel: the high doping configuration corresponding to Figure 4. Lower panel: high doping configuration corresponding to Figure 5.

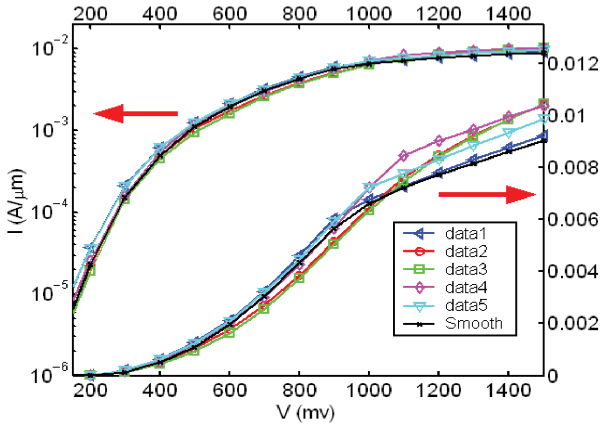


Fig. 23. Computed I_D–V_G characteristics of devices that differ only in the arrangement of high doping clusters ($\Delta N = 0$).

potential distributions and carrier distributions. Figure 26 shows the typical result that, despite the uniform undoped channel, the current meanders through the channel; the flow is non-uniform. This is a direct result of the inhomogeneous injection through the random potential landscape of the source; it is characteristic of spatially inhomogeneous doping resolved structures similar to atomistically resolved structures.[20–22] The resulting injected current flow is reflected from the walls of the channel. The spatial characteristics of the flow are different from classical modelling: the transport is thermodynamic quantum transport. The spectrally-resolved current density along the channel $J_x(E;\mathbf{r})$ shows more complex behaviour, consistent with channel quantization, but the total current density is smoothed by the thermal average over the energy states, E. No trace of the characteristic quantized current response found at low temperatures is found at electron temperatures $T_e \geq 300$ K. However, the monotonic current–voltage characteristics display quite different values for threshold voltage, sub-threshold response and on-current response from the classical results.[23]

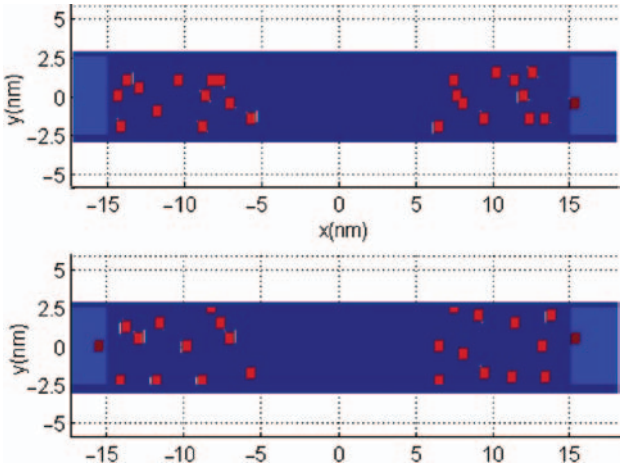


Fig. 24. Two high doping configurations with the same number of clusters corresponding to Figure 6.

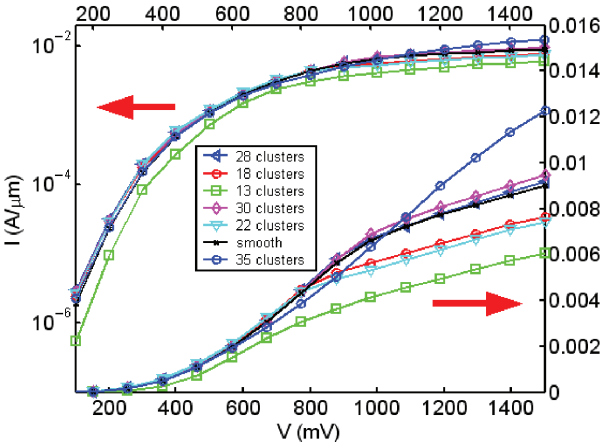


Fig. 25. Computed I_D–V_G characteristics of devices that differ in the arrangement of high doping clusters and in fluctuations of doping cluster number ($\Delta N \neq 0$).

Comparing Figures 23 and 25, we find that the fluctuation in total number of dopants has a stronger effect (50% fluctuation in on-current) than purely positional fluctuations of the clusters (10%). In both cases the impact on the on-current is stronger than the effect on the sub-threshold response. The computed on-current fluctuations are significantly larger than the classical predictions.[23] We conclude that for practical device design it is crucial to control precisely the total number of dopants ($\triangle N \to 0$) in the source and drain to avoid large performance fluctuations.

This section has explored the effects of the doping inhomogeneity in the source/drain regions confirmed recently in experiments[24] using a 2D NEGF formalism in a DG MOSFET. As the STM experiments show,[24] the existence of clustering and aggregation are likely to appear around the end terminations. This clearly affects the injection of carriers in the channel of the device, and creates fluctuations from device to device as has been show in our calculations.

In this work, symmetry in the perpendicular direction is assumed and all the physical magnitudes have

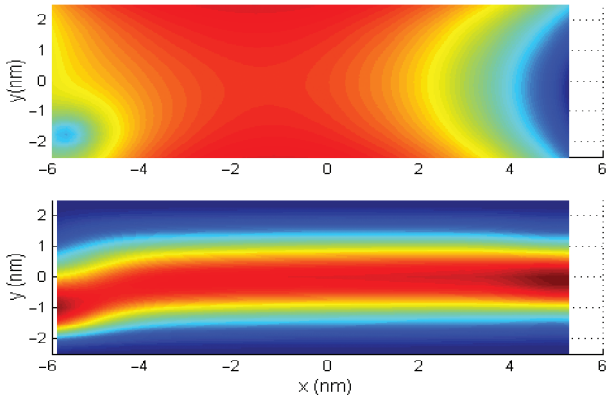


Fig. 26. Potential and current density landscapes corresponding to the device with the high doping configuration of the lower panel in Figure 7. Upper panel: potential landscape. Lower panel: current density J_x along the channel.

REVIEW

lateral symmetry reflected in the 2D electrostatics as a logarithmic dependence of the potential as a function of coordinates. This is the case in every 2D simulation. An analysis in three dimensions of three-dimensional doping inhomogeneities is beyond the scope of this work. Even so, a 3D analysis could be more relevant for bulk MOSFETs (three spatial dimensions of the device in the same order of magnitude) with intrinsically 3D character. This type of analysis for DG MOSFET will require a very large three-dimensional model in order to avoid the effect of the lateral boundaries and to be able to restore the global lateral symmetry. A 3D NEGF description of this device is computationally complex and will be addressed in a future work.

7. INTERFACE ROUGHNESS

The effect of the microscopic roughness profile of the interface on the I_D–V_G characteristic of a DG MOSFET has been studied. The current–voltage characteristics of eight devices that differ in their specific interface roughness profiles have been compared in order to understand the impact of the roughness on the device performances.

In our simulations the electron density can penetrate into the oxide in order to describe properly the quantum mechanical boundary condition when small interface irregularities occur. Several randomly generated interfaces have been simulated and a detailed study of the 2D current density has been carried out. The interfaces are randomly generated assuming an exponential auto-correlation function with correlation length, $\lambda = 1.8$, and assuming a digitised interatomic layer step height of 0.3 nm.[25] The grid at the interface is designed to match the one-interatomic-layer interface steps.[26] Figure 27 shows the simulated double gate (DG) MOSFET with a channel length of 10 nm and a channel thickness of 2 nm. The oxide thickness is fixed to 1.6 nm. An impression of a realistic device composed using TEM pictures of real Si/SiO_2 interfaces is shown in Figure 28. The picture shows atomic layer steps at the top and bottom interfaces. Figure 29 shows the interface configurations for one of the simulated transistors (denoted *rough3* later in Fig. 37). The body thickness variation correlated to the random steps at the interfaces results in confinement variations and produces a meandering flux of the current (see Fig. 30). The current is flowing through an effective cross-section, perpendicular to the current flow, which could be wider than the vertical channel cross section but the current path is longer compared to the smooth interface case. The 2D map of the horizontal (y-direction) current vector component is plotted in Figure 31 showing reduction in the corner regions of the channel.

The body thickness variation associated with the rough interfaces, and the corresponding variation in quantum confinement throughout the channel, produces strong variations in the electron concentration in the channel as

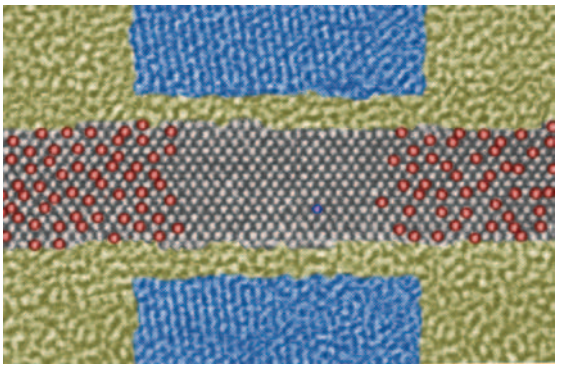


Fig. 28. Photoshop™ impression of a 10 nm DG MOSFET constructed from TEM images of real interfaces.

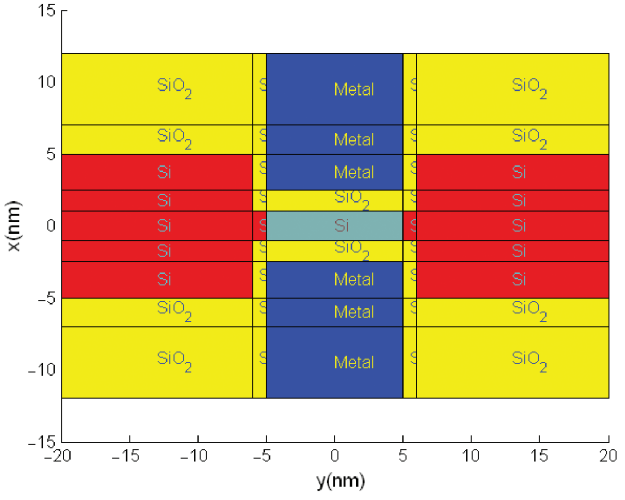


Fig. 27. Structure of the DG MOSFET. The channel thickness is 2 nm, the doping in the source and drain regions is 10^{20} cm^{-3} and the background channel doping is 10^{14} cm^{-3}.

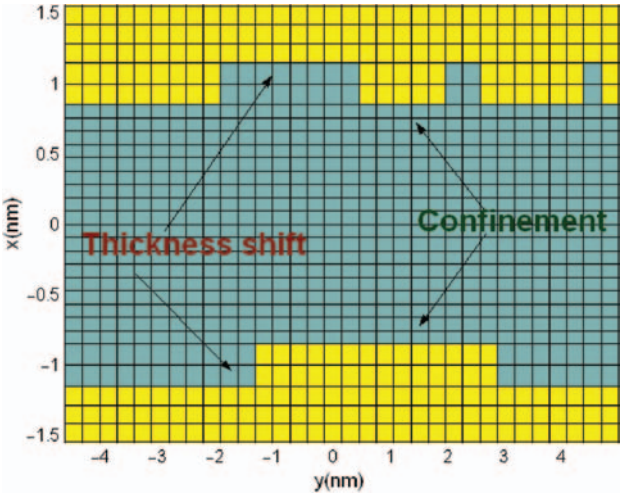


Fig. 29. Channel discretisation of one of the simulated cases showing various effects created by the interface thickness.

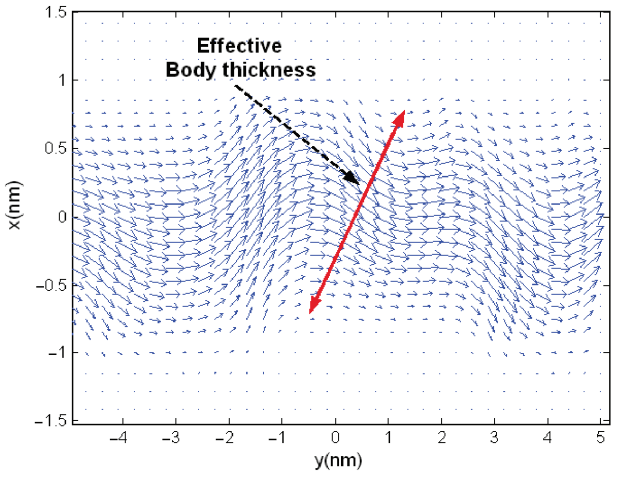


Fig. 30. The 2D current density vector field and the effective cross-section created as a result of the interface roughness shown in Figure 4.

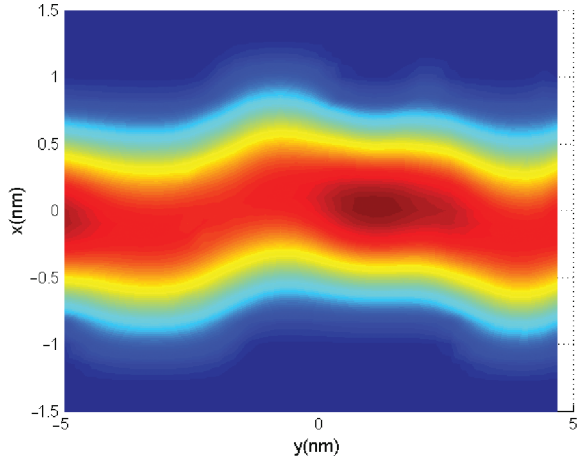


Fig. 31. 2D profile of the current in the transport direction showing the non homogeneity along the channel.

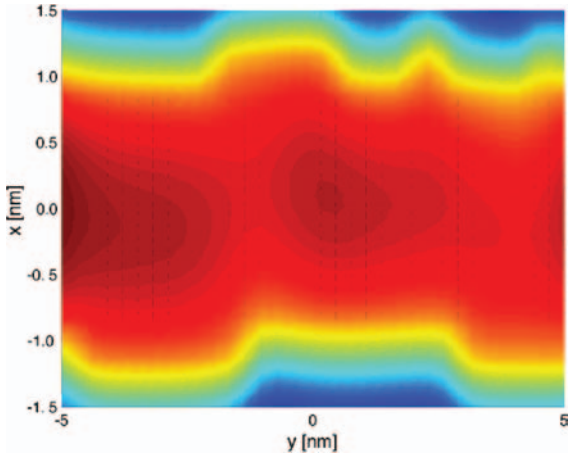


Fig. 32. 2D electron density profile showing the inhomogeneity along the channel.

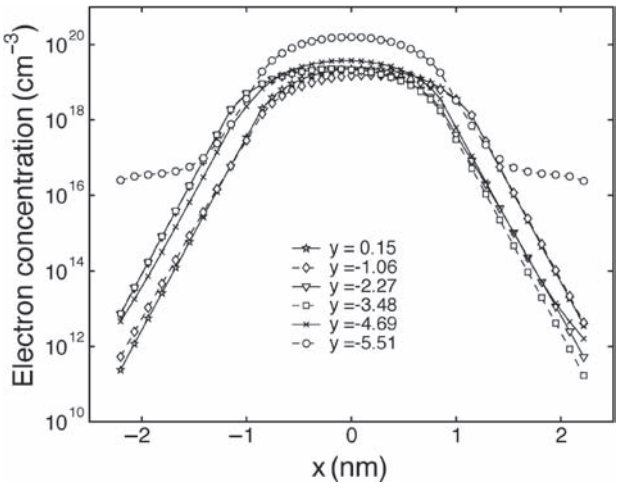


Fig. 33. Electron concentration on a vertical cross-section at different positions from the middle of the channel ($y = 0.15$) to the source ($y = -5.0$).

illustrated in Figure 32. In Figure 33, the electron concentration along cut lines normal to the channel are plotted for different positions along the channel ($y = -5.51$ nm is in the source region and $y = 0.15$ nm is close to the middle of the channel). There is a significant carrier penetration into the oxide at the interface in the regions with interface steps. There the wave function is penetrating the oxide from two different directions—perpendicular to the channel and parallel to the channel. The fluctuation in the electron concentration along the channel at different distances from the middle of the channel ($x = 0.0$ denotes the middle of the channel) is presented in Figure 34. It shows that the variation in the electron density near to the interface may be more than one order of magnitude.

Figures 35 and 36 show the density of states for the device of Figure 31 and the device without roughness respectively at identical gate voltages. The density of states of the device with roughness exhibits localised peaks due

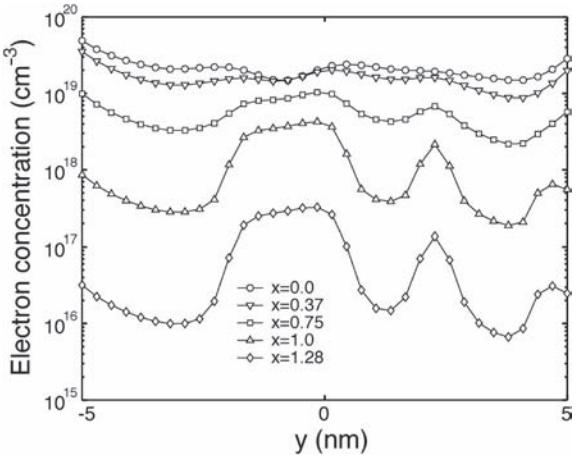


Fig. 34. Electron concentration along the channel at different distances from the middle of the channel (see Fig. 33).

to the body thickness fluctuations. The I_D–V_G characteristics at 300 K for 8 different devices with randomly generated surface roughness are shown in Figure 37 for a drain voltage of 0.5 V. The current for the device without the surface roughness is also shown for a comparison. The device with highest current has an averaged body thickness of almost 2.3 nm (*rough6*) as the patterns of the upper and lower interfaces combine to maximise the channel

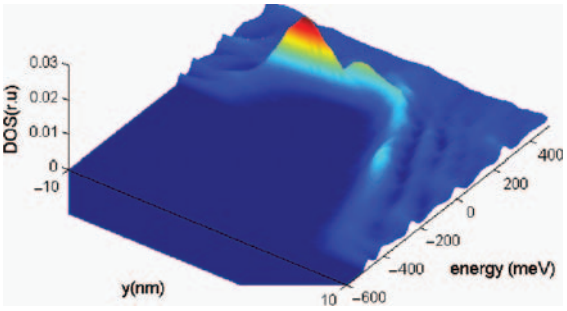


Fig. 35. The density of states at $V_G = 0.4$ V for the surface roughness of Figure 4.

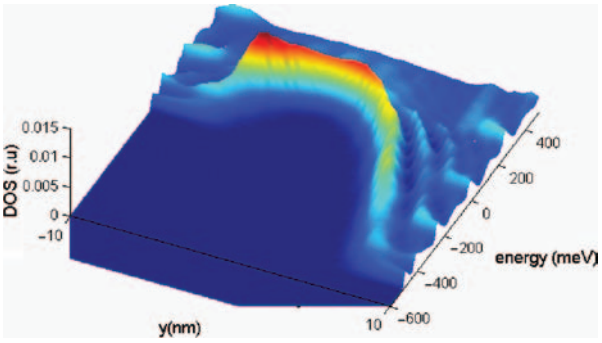


Fig. 36. The density of states at $V_G = 0.4$ V for the device without surface roughness.

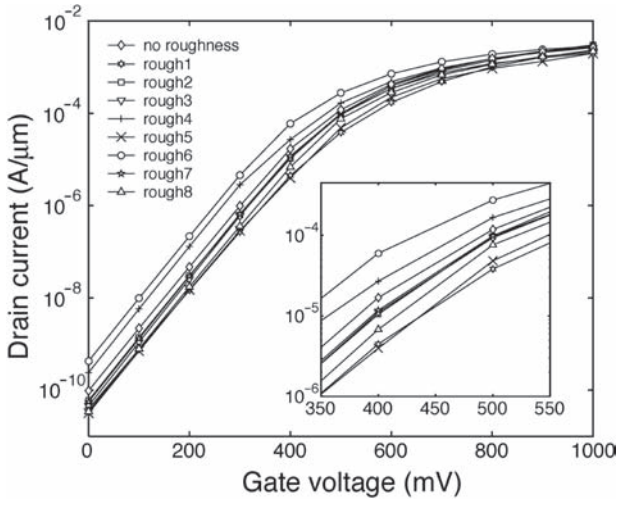


Fig. 37. I_D–V_G characteristic for different randomly generated interfaces. The I–V characteristic of the device without interface roughness is showing for a comparison.

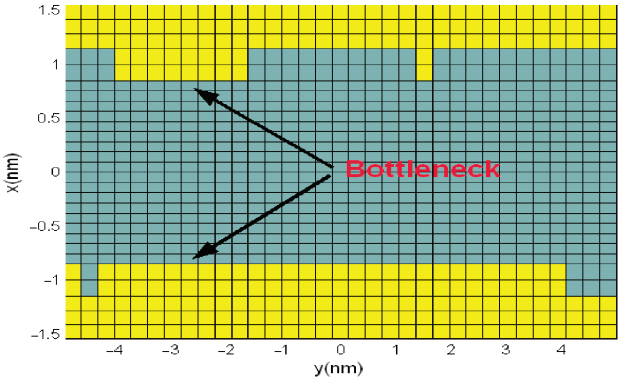


Fig. 38. The interface roughness corresponding to the case *rough1* (see the Fig. 12). The bottleneck is located at the beginning of the channel.

thickness. The lowest current comes from interface roughness patterns, which result in an effective body thickness approaching the minimum limit of 1.7 nm. In general if the body thickness is reduced at the beginning of the channel (close to the source) instead of in the middle, or the end, the current will be more drastically reduced, as will be shown later.

The sub-threshold slope is similar for all of the simulated cases, although the devices with larger average oxide thickness do exhibit minor degradation in sub-threshold slope due to increased short channel effects. However, the fluctuations in oxide thickness induced by the roughness are a small perturbation compared to the total oxide thickness, which controls the electrostatics. The magnitude of the current is affected by the interface because the constrains in the body result in stronger carrier confinement and local quantum mechanical shift of the threshold voltage.

Figures 38 and 39 illustrate the interfaces/body pattern for two transistors denoted *rough1* and *rough8*, respectively. As indicated with the arrows in the figures, the confinement for the device *rough1* is located at the beginning of the channel while in the device *rough8* the confinement is located at the middle of the channel. Figures 40 and 41 represent the electron density for the two cases. We can

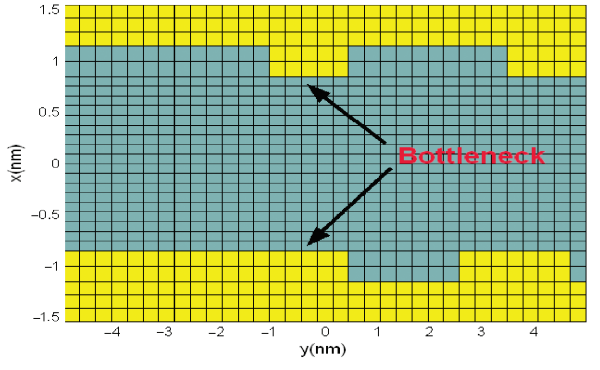


Fig. 39. The interface roughness corresponding to the case *rough8* (see Fig. 12). The bottleneck is located at the middle of the channel.

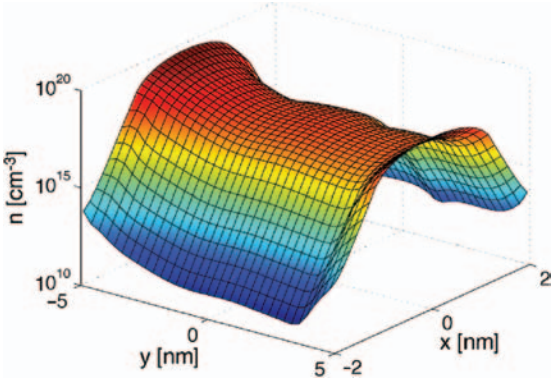


Fig. 40. 2D electron density for the case of Figure 39.

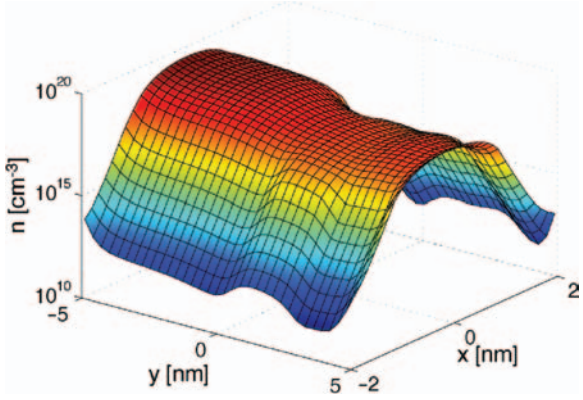


Fig. 41. 2D electron density for the case of Figure 40.

see that the electron injection in the channel is greater in the case *rough8* producing a slightly higher current than in the case of *rough1* as shown in Figure 39.

As a result of this work, we can conclude that the detailed unique pattern of the interfaces is reflected in a 2D current and density profile of the electrons in the channel which a 1D transport model is unable to describe properly. The current is changing direction inside the channel; therefore the effective current path is longer than the metallurgical channel length. This, in combination with the average increase or reduction of channel thickness, causes an intrinsic current variation as illustrated in Figure 37. The injection of electrons into the channel is affected by the surface roughness. The density exhibits an inhomogeneous pattern in the channel and strong fluctuation (one order of magnitude) close to the interface. Fluctuations in the threshold voltage produced by the interface roughness are on the order of 100 mV but the sub-threshold slope is only slightly affected.

8. THE 3D NEGF DEVICE SIMULATOR AND APPLICATIONS

The newly developed 3D quantum simulator[10] is based on the NEGF equations and uses the recursive algorithm in order to compute the current and electron density assuming that the Hamiltonian is expressed in the effective mass approximation. The random rough interface between Si and SiO_2 is introduced using the approach described in Ref. [5, 27] based on results reported in Ref. [28]. The model describes the random interface as a randomly generated surface with one interatomic layer step (0.28 nm) normal to the interface, and mesh space resolution, separating the Silicon and the SiO_2, and the corresponding values of the dielectric constant. The correlation length of 2 nm along the interface plane used to generate the surface is essentially identical to the one used in Monte Carlo simulations to reproduce the Si/SiO_2 universal mobility curve.[29]

A 3D Cartesian mesh with variable mesh spacing is used to discretise the solution domain. In order to include, in a simple way, interface roughness in the simulations we have adopted a control volume discretisation[5] approach for both the Poisson and the NEGF equations. Following the interface roughness generation approach the complex shape of the $Si\text{-}SiO_2$ interface is easily approximated by a collection of nodal planes. Hence a unique dielectric constant can be assigned to every volume element (brick) of the grid. In this case the Poisson equation can be easily integrated around a node (control volume discretisation), conserving the electric displacement vector around every node and naturally taking into account the dielectric constants of the eight bricks condensing at the node. Bearing in mind that each of the eight bricks could have 2 different dielectric constants (Si or SiO_2) each node may have 256 essentially different neighbourhood configurations. A finite difference scheme would need to treat all those different cases but in the control volume discretisation scheme these cases are automatically taken into consideration.

In very small nanowire transistors, surface roughness and discrete defects introduce non-perturbative effects, which greatly distort the transverse wave function from cross-section to cross-section along the transport direction. Such effects lead to a current density vector that is not necessarily parallel to the transport direction: The flow direction alters along the channel of the transistor leading to a meandering current flow.

This work intends only to show the capability of the 3D simulator and therefore presents only a qualitative study. We have applied the simulator also to study the effect of a stray charge in the channel of a nanowire transistor.

In order to demonstrate the capabilities of our 3D NEGF simulator we have simulated a gate-all-around silicon nanowire MOSFET with 2.2×2.2 nm^2 cross-section, 6 nm channel length and 1 nm gate oxide thickness, as illustrated in Figure 42. The doping in the source and drain region is 1×10^{20} cm^{-3} and the channel is undoped. The Poisson equation is solved in the external box of Figure 42 and the Green function equations are solved in the internal box.

Initially, we focus on the effects associated with the atomic scale roughness at the channel/gate dielectric

REVIEW

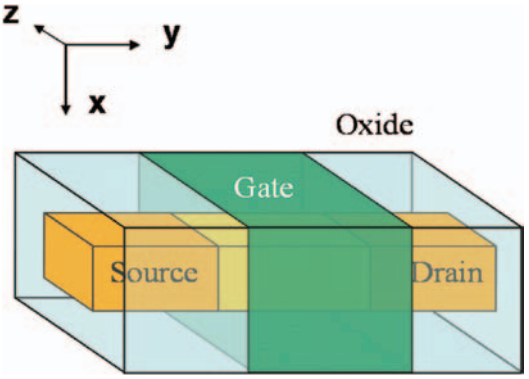


Fig. 42. The nanowire transistor, and the orientation of the coordinated system; y is the transport direction.

interface. In all simulations of this work, the surface roughness is only considered in the channel region. Figure 43(a) shows the 3D surface potential distribution associated with one randomly generated atomic scale interface roughness configuration at $V_G = 0.4$ V and $V_D = 0.2$ V. In this case, the interface roughness produces a constriction in the middle of the channel. Figure 43(b) shows the 3D plot of iso-surfaces of the y component of the current density. The tubular iso-surfaces became wider in the middle of the channel reflecting an increase in current density. The colour code of the figures associates maximum values of the mapped field to the dark red colour and minimum values to the dark blue colour.

Figure 44(a) maps in the zy plane the electrostatic potential and the current density for the constricted nanowire from Figure 43 at $V_G = 0.4$ V and $V_D = 0.2$ V. The current density increases in the middle of the channel due to the constraint. Figure 44(b) shows the potential and current in the same plane but for a nanowire with a synchronised step in both the top and bottom interfaces, which does not reduce the cross section but increases the current path. In this case the current changes direction in the vicinity of the step, following a meandering path.

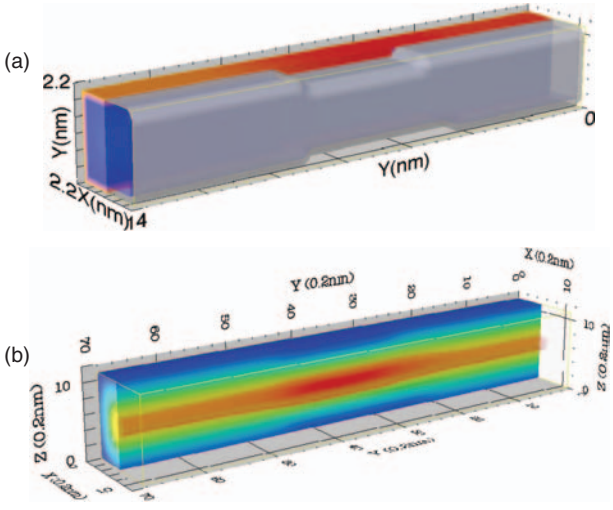


Fig. 43. (a) 3D view of the rough surface potential through a constrained nanowire. (b) 3D view of the current of the constricted nanowire.

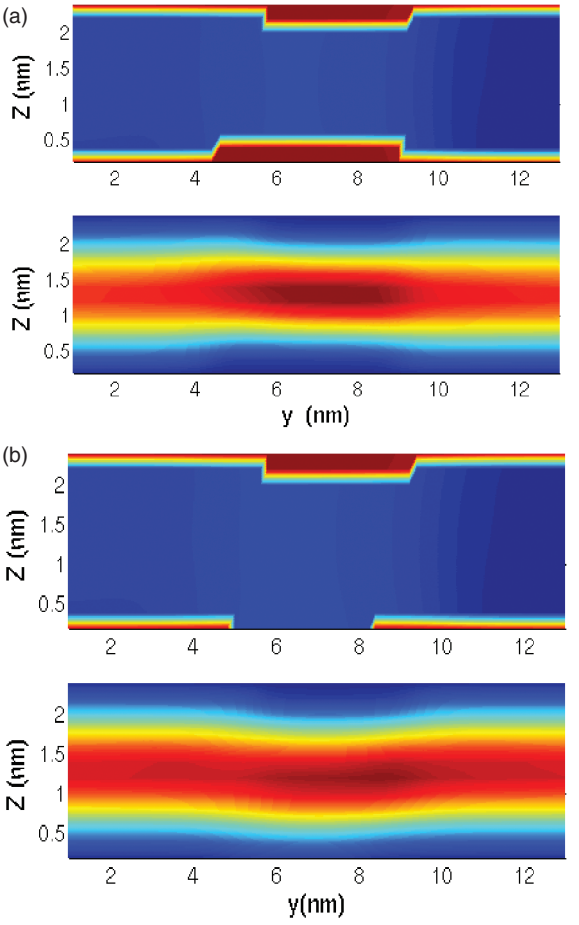


Fig. 44. (a) A Nanowire with constriction. Upper: Potential profile. Lower: Current density profile (both at $x = 1.1$ nm). (b) The potential and current density in a stepped nanowire ($x = 1.1$ nm).

Figure 45(a) and (b) show the electrostatic potential along the middle of the channel, in both the uniform and the constricted devices respectively, for gate voltages in the range 0.1 V to 0.5 V and a drain voltage of 0.2 V. The electrostatic potential distribution at low gate bias is similar for both devices. At high gate bias ($V_G = 0.4 - 0.5$ V), the gate control of the electrostatic potential in the constriction region is stronger resulting in lower potential compared to the uniform case. The electron density distributions along the channel of the uniform and the constricted devices, under the same bias conditions, are illustrated in Figure 46(a) and (b) respectively. Despite the better electrostatic control the electron density in the constrained device at equivalent gate biases is lower compared to the uniform device. This is due to an energy shift in the transverse states in the constriction. For the same reason, the injection of electrons in the channel of the constrained device is lower, when compared to the smooth device at all gate voltages.

The I_D–V_G characteristics of the uniform device and the two devices from Figure 44 are shown in Figure 47. The

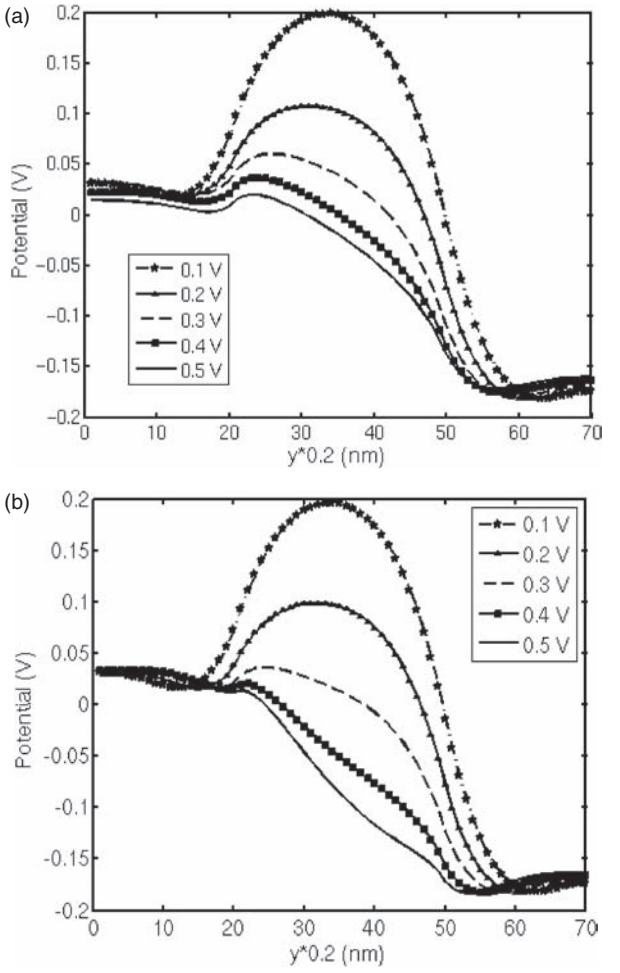


Fig. 45. (a) The potential along the centre of the uniform nanowire at different gate biases. (b) The potential along the constricted nanowire for different gate bias.

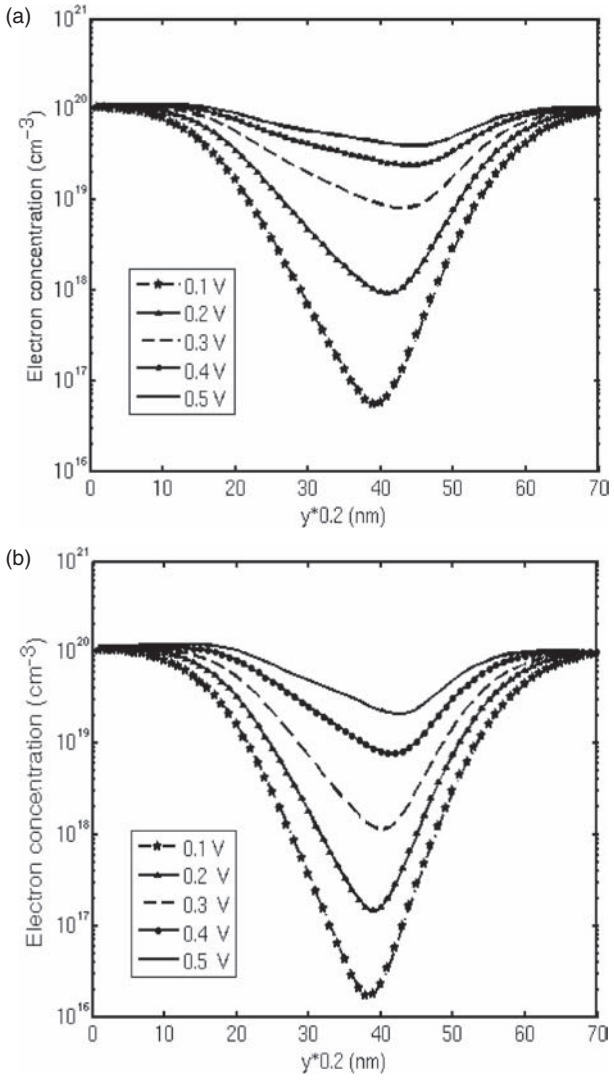


Fig. 46. (a) The electron concentration the along the uniform nanowire. (b) The electron concentration along the constrained nanowire.

effect of the constriction is dramatic and produces a threshold voltage shift of 150 mV when compared to the uniform device. The threshold voltage shift of the device with synchronised step (∼50 mV) is significantly smaller. Additionally, the on current at $V_G = 0.5$ V in the constrained device is 3 times smaller than that in the uniform device. The constriction introduces a change in body thickness, which is around 30% of the original thickness. The effect on the carrier concentration and the current is not perturbative and the corresponding change in current is large. The difference between the I_D–V_G curves of the nanowire with the synchronised steps and the uniform one reflects the effect of the surface roughness scattering only. Clearly, for the nanowire studied here, the confinement effect has a much greater impact on the current compared with the scattering coming from the interface roughness. However with the increase in the nanowire cross-section the confinement effects associated with atomic scale interface roughness will play a less important role.

Figure 48 shows the iso-surfaces of the y-component of the current density in a nanowire with a single electronic

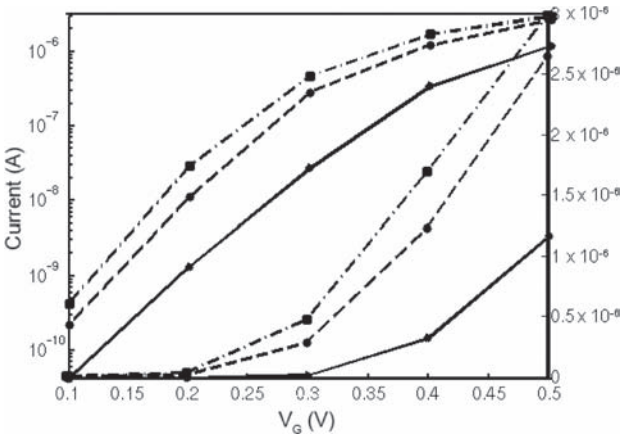


Fig. 47. The I_D–V_G characteristics of the uniform nanowire (dot-dash), the stepped nanowire (dashed) and the constrained nanowire (solid).

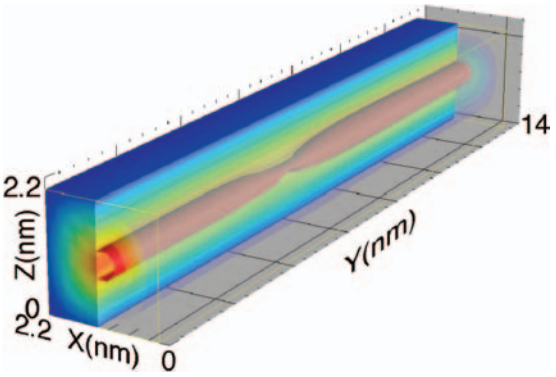


Fig. 48. 3D plot of the y current component in a nanowire with a single impurity in the centre of the channel.

charge associated with an ionised acceptor or a trapped electron located almost at the middle of the channel ($x =$ 1.1 nm, $y = 6.0$ nm, $z = 1.1$ nm) $V_G = 0.4$ V and $V_D =$ 0.2 V. The electrons are able to propagate across the middle of the channel mainly through the ground state in the potential created by the impurity and the channel interface, which requires higher injection energy. There is also a small contribution associated with a direct tunnelling through the coulomb potential. This is reflected in the narrowing of the tubular iso-surfaces at the middle of the channel. The situation is clarified further in Figure 49(a) and (b) which maps the electrostatic potential and the current density through the impurity in the zy plane. Note the crowding of the current between the impurity and the interface.

The potential distribution along the centre of the channel for different positions of the impurity in the xz plane at $y = 6$ nm and at low gate bias $V_G = 0.2$ V and $V_D = 0.2$ V is illustrated in Figure 50. The impact of the Coulomb tail on the potential barrier in the middle of the channel is still strong even when the impurity is 1 nm off the centre of the channel. In order to study the current sensitivity with respect to the position of the impurity in a particular cross-section of the channel we have carried out simulation

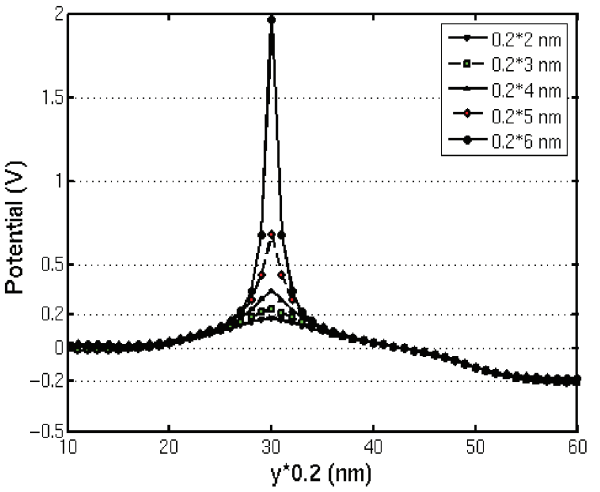


Fig. 50. The potential profile along the channel at different distances from the interface (at low bias).

for two impurity positions ($x = 1.1$ nm, $z = 1.1$ nm) and ($x = 1.1$ nm, $z = 1.5$ nm) in the xz plane at $y = 6$ nm. The iso-surfaces of the electrostatic potential at $V_G = 0.4$ V and $V_D = 0.2$ V are plotted in Figure 51(a) and (b) for the two impurity configurations. The electrostatic perturbation caused by the impurities in the channel is rather strong for both impurity positions and should have a strong impact on the I_D–V_G characteristic.

Figure 52 illustrates the iso-surface plot of electron concentrations for the impurity in the middle of the channel for low ($V_G = 0.2$ V) and high ($V_G = 0.4$ V) gate bias. The high doping regions in the source and drain are clearly seen. At high gate bias the electron concentration at the source extends further in the channel as expected.

Figure 53 compares the I_D–V_G characteristics of the uniform device and the two devices with a single negative charge in the middle and 1nm off the middle of the channel at $y = 6$ nm. The characteristics of the transistors with the two different positions of the impurities are very

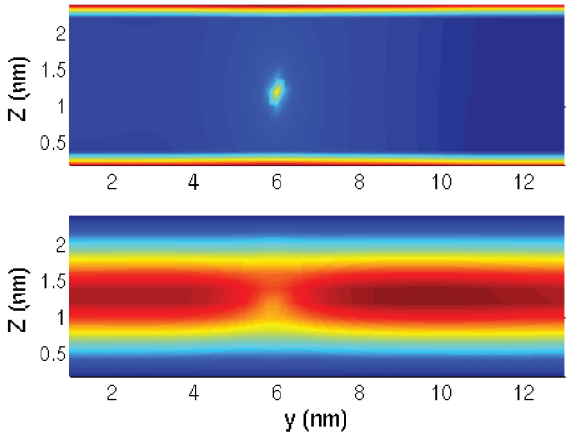


Fig. 49. The potential in a nanowire with an impurity in the middle of the channel (upper), and current density (lower), at $x = 1.1$ nm

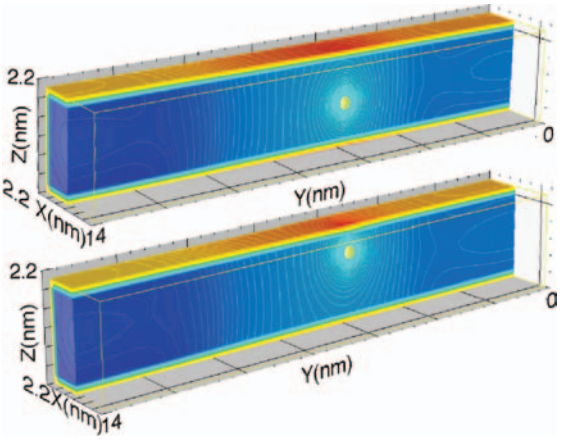


Fig. 51. Potential contour plots for the impurity located at (1.1, 6.0, 1.1) nm and (1.1, 6.0, 1.5) nm.

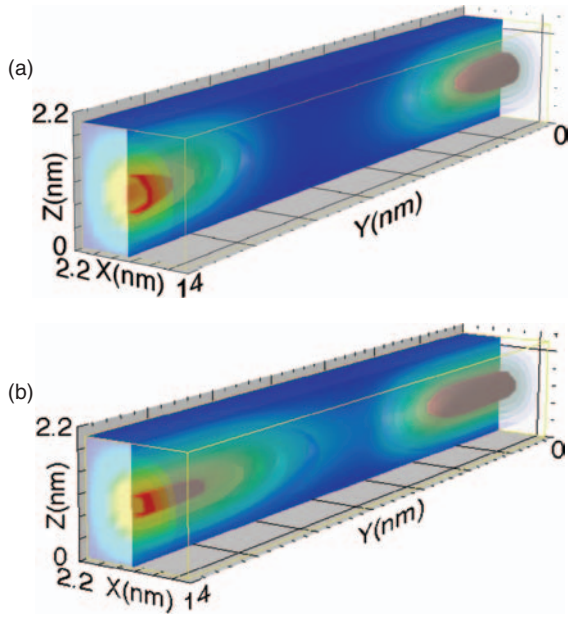


Fig. 52. (a) Electron density isosurfaces at low gate bias with the impurity in the middle of the nanowire. (b) Electron density isosurfaces at high bias with the impurity in the middle of the nanowire.

similar, both yielding a threshold voltage shift of approximately 100 mV. This similarity is associated with the small (∼2.2 nm) cross-section of the nanowire which is comparable with the characteristic radius of the screened Coulomb potential which is of the order of 1 nm, as is shown in Figures 50 and 51.

The full 3D NEGF simulations at room temperature produce smooth monotonic I_D–V_G characteristics, despite size-quantisation and contributing sub-band quantum transport. Surface roughness has a dramatic effect at nano-scale dimensions due to the fact that both the energy shifts, surface roughness scattering and three-dimensional meandering current flows associated with the random surface roughness configurations. Constrains of the channel

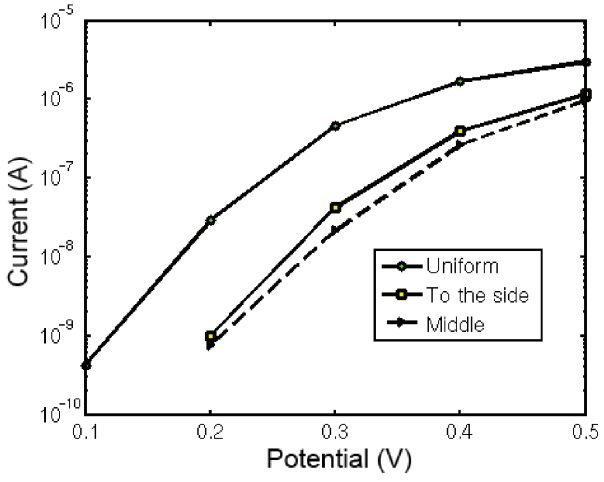


Fig. 53. The I_D–V_G characteristics of the uniform and non-uniform nanowires.

have a more dramatic effect on the current compare to synchronised roughness that does not change the channel width. The influence of a single dopant on device behaviour is dramatic, independent of its spatial location due to the comparable characteristic coulomb radius and the cross-sectional dimensions of the nanowire channel.

9. CONCLUSIONS

In this work we have presented the development of self-consistent 2D/3D ballistic quantum device simulators; we have explained in detail the recursive algorithm and the principles used for the discretisation of the NEGF and Poisson equations. The simulators solve the electron transport problem in 2 and 3 dimensions calculating the full 3D and 2D electron and current density without any approximation; this allows us to address very strongly varying potentials such as surface roughness and discrete dopants modelled by Coulomb potentials. Finally, we demonstrate the capability of these simulators by addressing specific problems: interface roughness and discrete dopants (stray charges are consider as acceptors in the channels) in two and three dimensions. These simulators use an effective mass approximation, but the masses can be extracted from various atomistic techniques such as Tight binding and first principles methods, which makes the simulators powerful fast tools for predicting the device performance of very small devices.

REVIEW

References

1. Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. Lundstrom, *IEEE Trans. Electron Devices* 50, 1914 **(2003)**.
2. M. Luisier, A. Shenk, and W. Fichtner, *J. Appl. Phys.* 100, 043713 **(2006)**.
3. J. Wang, A. Raman, G. Klimeck, and M. Lundstrom, *IEDM Tech. Dig.* 537 **(2005)**.
4. A. Raman, G. Klimeck, and M. Lundstrom, *IEDM Tech Dig.* 615 **(2005)**.
5. A. Martinez, A. Svizhenko, M. P. Anantram, J. R. Barker, and A. Asenov, *IEDM Tech Dig.* 626 **(2005)**.
6. A. R. Brown, F. Adamu-Lema, and A. Asenov, *Superlattices and Microstructures* 34, 283 **(2003)**.
7. J. R. Watling, L. Yang, M. Boriçi, R. C. W. Wilkins, A. Asenov, J. R. Barker, and S. Roy, *Solid-State Electronics* 48, 1337 **(2004)**.
8. J. Fonseca and S. Kaya, *Solid-State Electronics* 48, 1843 **(2004)**.
9. A. Svizhenko, M. P. Anatram, T. R. Govindan, B. Biegel, and R. Venugopal, *J. App. Phys.* 91, 2343 **(2002)**.
10. A. Martinez, M. Bescond, J. Barker, A. Svizhenko, M. P. Anantram, and C. Millar, *IEEE Trans. Electron Devices* 54, 2213 **(2007)**.
11. J. Schwinger, *J. Math. Phys.* 2, 407 **(1961)**.
12. L. V. Keldysh, *Sov. Phys.-JEPT* 20, 1018 **(1965)**.
13. R. Landauer, *Phil. Mag.* 21, 863 **(1970)**.
14. M. Buttiker, *IBM J. Res. Dev.* 32, 317 **(1988)**.
15. H. L. Engquist and P. L. Anderson, *Phys. Rev. B* 24, 1151 **(1981)**.
16. Y. Imry and R. Landauer, *Rev. Mod. Phys.* 71, 306 **(1999)**.
17. R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, *J. App. Phys.* 81, 7845 **(1997)**.
18. S. Datta, Electronic Transport in Mesoscopic Systems, Cambridge University Press **(1995)**.

19. M. P. Anantram, A. Svizhenko, and A. Martinez, *J. App. Phys.* 100, 119903 **(2006)**.
20. J. R. Barker, A. Martinez, A. Svizhenko, A. Anantram, and A. Asenov, *J. Phys.: Conf. Ser.* 35, 233 **(2006)**.
21. J. R. Barker, *Physica* E19, 62 **(2003)**.
22. J. R. Barker, *Semiconductor Science and Technology* 19S, 56 **(2004)**.
23. A. R. Brown, A. Asenov, and J. R. Watling, *IEEE Transactions on Nanotechnology* 1, 195 **(2002)**.
24. H. Fukotome, Y. Momiyama, E. Yoshida, M. Okuno, T.Itakura, and T. Aoyama, *IEDM Technical Dig.* 57 **(2005)**.
25. A. Asenov, S. Kaya, and J. H. Davies, *IEEE Trans. Electron Dev.* 49, 112 **(2002)**.
26. A. R. Brown, F. Adamu-Lema, and A. Asenov, *Superlattices and Microstructures* 34, 283 **(2003)**.
27. A. Asenov, S. Kaya, and J. H. Davies, *IEEE Trans. Electron Dev.* 49, 112 **(2002)**.
28. S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Lilienthal, D. Fathy, and O. L. Krivanek, *Phys. Rev. B* 32, 8171 **(1985)**.
29. S. Takagi, A. Toriumi, M. Iwase, and H. Tango, *IEEE Trans. Electron Devices* 41, 2357 **(1994)**.

Received: 5 February 2008. Accepted: 12 February 2008.

REVIEW