# *Quantum confinement in silicon-germanium electron waveguides*

**Gilberto Curatola**
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni, Università di Pisa
**Giuseppe Iannaccone**
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni, Università di Pisa

G. Curatola, G. Iannaccone, "*Quantum confinement in silicon-germanium electron waveguides*", Nanotechnology, **13** , 3, pp.267-273 (2002).

INSTITUTE OF PHYSICS PUBLISHING NANOTECHNOLOGY
Nanotechnology **13** (2002) 267–273 PII: S0957-4484(02)31451-X

# Quantum confinement in silicon–germanium electron waveguides

**G Curatola and G Iannaccone**

Dipartimento di Ingegneria dell'Informazione, Universitá degli studi di Pisa, Via Diotisalvi 2, I-56122 Pisa, Italy

E-mail: g.curatola@iet.unipi.it and g.iannaccone@iet.unipi.it

Received 3 December 2001, in final form 18 March 2002
Published 9 May 2002
Online at stacks.iop.org/Nano/13/267

**Abstract**
We have simulated the electronic properties of silicon–germanium electron waveguides defined by selective etching on a SiGe heterostructure. In particular, we have investigated the dependence of quantum confinement and of one-dimensional subband separation on the waveguide width. Indeed, a larger subband separation means a larger dephasing length and larger electron mobility in the waveguide, and therefore increased possibilities of detecting mesoscopic transport effects. Accurate modelling of SiGe electron waveguides requires us to take into account the effect of strain in the SiGe heterostructure and of the interface states at the exposed SiGe surface, and to solve the Poisson–Schrödinger equation in two dimensions. Results are also shown for a structure in which a gate electrode is evaporated onto the SiGe waveguide, realizing a three-terminal device in which the gate voltage is used to control the number of propagating modes, and therefore the conductance of the channel.

## 1. Introduction

Silicon–germanium technology holds great promise to reduce the cost of consumer products and help make new applications possible. It would allow device designers to simultaneously increase speed, reduce noise and reduce power supply requirements. Although high-speed devices have been fabricated with the $Al_xGa_{1-x}As$/GaAs material system, a silicon-based technology would allow integration with the CMOS process. By using SiGe heterostructures it is possible to realize conduction and valence band offsets for obtaining enhanced-mobility devices.

SiGe heterojunction bipolar transistors (HBTs) have recently reached the marketplace for high-frequency applications: mixers, low-power amplifiers and global position systems (GPSs). HBTs can be integrated with CMOS by inserting a narrow $Si_xGe_{1-x}$ ($x < 0.3$) base layer into a BiCMOS fabrication process. The $Si_xGe_{1-x}$ base may be doped with higher densities with respect to a normal bipolar transistor and a graded Ge base may be grown in order to accelerate the carriers across the base. In research centres, HBTs with $f_T$ up to 120 GHz [1] and $f_{max}$ up to 150 GHz [2] have been demonstrated, while more modest values have been obtained for HBTs introduced into the market place ($f_T = 50$–60 GHz) [3, 4].

In contrast, SiGe field-effect transistors (HFETs) still require significant research before marketable products may come to fruition, even if recently the interest in high-mobility silicon–germanium devices has been greatly increased.

As an example, silicon–germanium field effect transistors cannot withstand the high temperature of standard CMOS processing, which can cause diffusion of dopants and of Ge, and the relaxation of the strained layer. A solution could be to use a modified low-thermal-budget CMOS process with deposited gate oxides, as indicated by Paul *et al* [5]. A proper choice of the thermal budget, which would allow the activation of impurities implanted into the ohmic contacts and in the poly gate and avoid segregation effects, would allow us to realize HFETs with higher transconductance than comparable Si MOSFETs. The higher current drive at low $V_{DS}$ can be attributed to the higher mobility of the strained Si channel, as simulated by Sadek *et al* [6] in the case of 0.2 $\mu$m devices.

These encouraging results could be used to improve the electrical behaviour of p-MOSFETs, which limit performance of CMOS circuits, because the hole mobility is 2.5 times lower than the electron mobility. Unfortunately, several problems must be solved to fabricate silicon–germanium

MOSFETs. The mobility improvement due to the splitting of the conduction and valence bands is undermined by alloy scattering and the small valence band discontinuity can produce parallel conducting channel at small $V_{GS}$, which causes a mobility degradation for both carriers.

SiGe modulation-doped field effect transistors (MODFETs) are not affected by parallel conducting channels and have shown switching times comparable to GaAs/AlGaAs n-MODFETs and lower power requirements with respect to the silicon counterpart [7, 8]. These facts make the investigation of the practical performance limits of SiGe MODFETs important. At present, novel techniques of fabrication [9] allow us to realize silicon–germanium quantum wires with widths down to 15 nm.

As the dimension of the wire is shrunk, the effects of the interface state become more and more important, simply because they are responsible for depletion of the wire [10]. For this reason, we need an accurate model for interface states.

In this paper we focus on the effects of interface states and of quantization on the carrier concentration in silicon–germanium quantum wires of different dimensions. Furthermore, acting on the germanium concentration and on the wire dimensions, it is possible to vary the separation between occupied modes in the waveguide and hence to select the number of propagating modes.

Because of the reduced dimensions of such structures, quantum effects are very important and, in order to obtain accurate analysis of the problems investigated, we have used a two-dimensional simulation code based on the nested solution of the Poisson–Schrödinger equation. The equations are solved numerically with the Newton–Raphson algorithm, after a discretization based on the box-integration method.

## 2. Model

### *2.1. Band alignment at the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ interface*

Let us consider a strained $Si_\alpha Ge_{1-\alpha}$ epitaxial layer deposited onto a relaxed $Si_\beta Ge_{1-\beta}$ substrate, which imposes the lattice constant in the plane perpendicular to the growth direction. Let $\alpha$ represent the germanium mole fraction in the epi-layer and $\beta$ represent the germanium mole fraction in the substrate. The strain in the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ heterostructure, as a consequence of the thermal and lattice mismatch between the substrate and the epitaxial layer, affects the energy band structure, the energy gap, the curvature at the conduction and valence band minima and the degeneracy. Changes in the curvature affect the carrier effective masses, mobility and effective channel velocity. Stress causes a splitting of the sixfold degenerate conduction band into two- and fourfold degenerate valleys and hence leads to a preferential occupation of conduction band minima and to reduced scattering. In addition, the valence band degeneracy at the $\Gamma$ point is lifted. Heavy holes are at the valence band edge under biaxial compressive strain ($\alpha > \beta$); light holes are at the valence band edge in the case of tensile strain ($\alpha < \beta$). The splitting of the conduction band minima and of valence band edges has been computed in [11] by means of self-consistent simulations based on the local density functional and *ab initio* pseudopotential methods. Spin–orbit splitting effects in the valence band are included *a posteriori*. The effective masses for electron and holes are calculated in [12] by means of nonlocal empirical pseudopotential calculations with spin–orbit interactions, while the energy gap in the strained layer has been derived by adding the band offsets, previously calculated, to the gap of the relaxed layer given by [13].

### *2.2. Band parameters in strained $Si_\alpha Ge_{1-\alpha}$ alloys on $Si_\beta Ge_{1-\beta}$ substrates*

Here the procedure is described for calculating the band alignment at the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ interface. In figure 1 the band alignment at the interface between a strained layer deposited onto a relaxed substrate is shown and the strain-induced splitting of the conduction band minima and valence band edges is highlighted. It is important to notice that the procedure implemented is still valid even if the substrate is itself strained and the lattice constant is imposed by a third unstrained layer, but fails if the germanium mole fraction in the substrate or in the epi-layer is larger than 0.85, because the position of conduction band minima in the $k$-space is changed.

The lattice constant of the substrate, as a function of the germanium concentration, can be expressed as [12]

$$a_o(\beta) = a_o(\mathrm{Si}) + 0.200\,326\alpha(1-\beta) + [a_o(\mathrm{Ge}) - a_o(\mathrm{Si})]\beta^2. \tag{1}$$

The epi-layer lattice constant $a_\perp$ in the growth direction is given by [12]

$$a_\perp(\alpha) = a_o(\alpha)\left[1 - 2\frac{c_{12}(\alpha)}{c_{11}(\alpha)}\frac{a_\parallel(\alpha) - a_o(\alpha)}{a_o(\alpha)}\right] \tag{2}$$

where $c_{11}$ and $c_{12}$ are the elastic constants.

Starting from the difference, $\Delta E_{vav}$, between the weighted averages of valence band edges at $\Gamma$, we have calculated the splitting of the three valence band edges $\Delta E_{v1}$, $\Delta E_{v2}$ and $\Delta E_{v3}$ as a function of $\alpha$ and $\beta$ [11, 12]:

$$\Delta E_{vav} = (0.047 - 0.06\beta)(\alpha - \beta) \tag{3}$$

$$\Delta E_{v2} = \tfrac{1}{3}\Delta_0 - \tfrac{1}{2}\delta E_{001} \tag{4}$$

$$\Delta E_{v1} = -\tfrac{1}{6}\Delta_0 + \tfrac{1}{4}\delta E_{001} + \tfrac{1}{2}[(\Delta_0)^2 + \Delta_0\delta E_{001} + \tfrac{9}{4}(\delta E_{001})^2]^{\frac{1}{2}} \tag{5}$$

$$\Delta E_{v3} = -\tfrac{1}{6}\Delta_0 + \tfrac{1}{4}\delta E_{001} - \tfrac{1}{2}[(\Delta_0)^2 + \Delta_0\delta E_{001} + \tfrac{9}{4}(\delta E_{001})^2]^{\frac{1}{2}} \tag{6}$$

where $\Delta_0$ [11] is the experimental spin–orbit splitting in the unstrained material, and $\delta E_{001}$ [11] is the linear splitting of the multiplet.

Similarly, the splitting of the two- and fourfold conduction band minima [11, 12] can be expressed as follows:

$$\Delta E_{c2} = \tfrac{2}{3}E_u(\varepsilon_{zz} - \varepsilon_{xx}) \tag{7}$$

$$\Delta E_{c4} = -\tfrac{1}{3}E_u(\varepsilon_{zz} - \varepsilon_{xx}) \tag{8}$$

where $\varepsilon_{ii}$, with $i = x, y, z$, are the components of the symmetric strain tensor and $E_u$ is a deformation potential. Hence we can calculate the discontinuity between the valence band edges, the gap in the strained layer and finally the discontinuity between the conduction band minima:

$$\Delta E_v = \Delta E_{vav} - \Delta E_{vi}^{sub} + \Delta E_{vi} \tag{9}$$

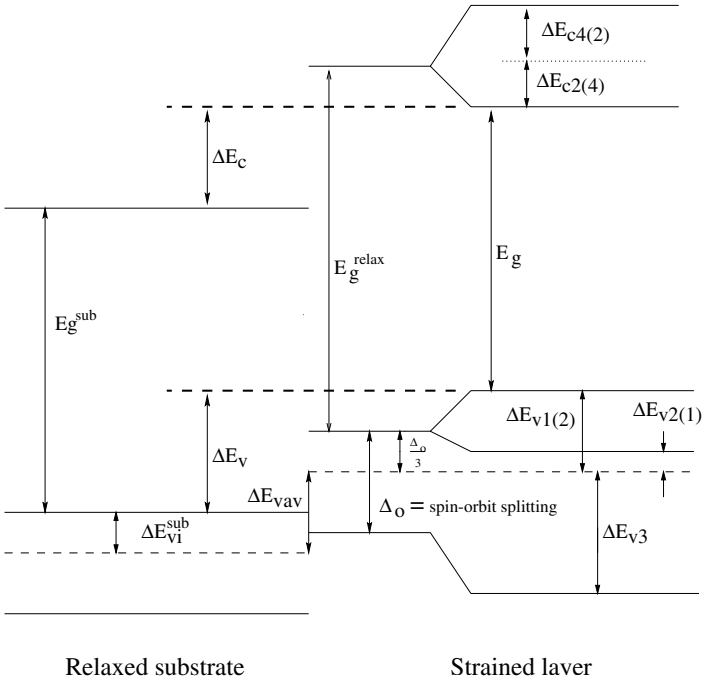



**Figure 1.** Band alignment at the interface between a $Si_\alpha Ge_{1-\alpha}$ layer and a $Si_\beta Ge_{1-\beta}$ substrate. The figure represents the case $\alpha < \beta$, while the numbers in parentheses refer to the case $\alpha > \beta$. Note that the conduction band minimum splittings are referred to the energy of the unstrained sixfold degenerate valleys and hence $\Delta E_{c4(2)}$ and $\Delta E_{c2(4)}$ have opposite sign.

where $\Delta E_{vi}^{sub}$ and $\Delta E_{vi}$ represent the difference between the maximum and the average valence band edge in the substrate and in the epitaxial layer, respectively.

$i = 1$ (2) in the case of tensile (compressive) strain. We have

$$E_g = E_g^{relax} + \Delta E_{ci} - \Delta E_{vj} + \Delta E_{go} - \frac{\Delta_0}{3} \tag{10}$$

where $\Delta E_{go}$ [13] is the gap variation as a consequence of an uniaxial strain in the growth direction and $j = 2, 4$ in the case of tensile or compressive strain, respectively. $E_g^{relax}$ is the value of the energy gap known for the unstrained layer [13]. Finally

$$\Delta E_c = E_g + \Delta E_v - E_g^{sub}. \tag{11}$$

The electron effective masses are calculated as a function of $\alpha$ and $\beta$ [12], while we have used for holes the values known for silicon. In particular, electron effective masses are calculated in the following way:

$$m_s(\alpha, \beta) = [1, (\alpha - \beta), (\alpha - \beta)^2] \cdot \boldsymbol{W} \cdot \begin{bmatrix} 1 \\ (\alpha + \beta) \end{bmatrix} \tag{12}$$

where $s = l, t1, t2$ identifies the longitudinal and transverse mass and $\boldsymbol{W}$ is a $3 \times 2$ matrix [12].

### *2.3. Self-consistent solution of the Poisson–Schrödinger equation*

The nonlinear Poisson equation for the electrostatic potential $\Phi$ is

$$\begin{aligned} \nabla \cdot (\epsilon \nabla \Phi) &= -\rho[\Phi] \\ &= -q[-n[\Phi] + p[\Phi] + N_D^+[\Phi] - N_A^-[\Phi]] \end{aligned} \tag{13}$$

where $\epsilon$ is the dielectric constant, $q$ the electron charge, $n$ and $p$ the electron and hole concentrations, respectively, and $N_D^+$ and $N_A^-$ the ionized donor and acceptor concentrations.

Consider an electron in a region of dimensions $L_x$, $L_y$, $L_z$ with $L_x, L_y \ll L_z$ and let us assume that the structure has translational symmetry along the $z$-axis, so that all quantities depend only on $x$ and $y$.

The two-dimensional Schrödinger equation, in the case of time-independent potential $V(x, y)$, reads

$$-\frac{\hbar^2}{2} \nabla \cdot (m^{-1} \nabla \Psi_i) + V(x, y) \Psi_i = E_i \Psi_i \tag{14}$$

where $\Psi_i$ represents the $i$th eigenfunction, $E_{i,j}$ is the $i$th eigenenergy and $m$ is the electron effective mass tensor in the plane perpendicular to the direction of propagation:

$$m = \begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix}. \tag{15}$$

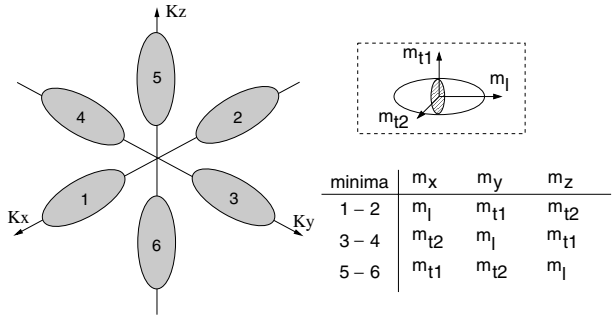
Note that we enumerate the eigenvalues $E_i$ in ascending order ($E_1$ has the lowest energy, $E_n$ has the highest energy, for $i = 1, \ldots, n$), and $\Psi_i$ is the corresponding eigenfunction. The electron energy can be written as

$$E(i, n_z) = E_i + \frac{\hbar^2 k_{nz}^2}{2m_z} \tag{16}$$

where $E_i$ represents the eigenvalue of (14) and $\hbar^2 k_{nz}^2 / 2m_z$ is the energy along the $z$-axis.

The density of states per unit of volume and energy near a conduction band minimum is given by

$$N(E) = \frac{\sqrt{2m_z}}{\pi \hbar} \sum_i |\Psi_i|^2 (E - E_i)^{-\frac{1}{2}} u(E - E_i). \tag{17}$$




| minima | $m_x$ | $m_y$ | $m_z$ |
|---|---|---|---|
| 1 – 2 | $m_l$ | $m_{t1}$ | $m_{t2}$ |
| 3 – 4 | $m_{t2}$ | $m_l$ | $m_{t1}$ |
| 5 – 6 | $m_{t1}$ | $m_{t2}$ | $m_l$ |

**Figure 2.** Sixfold degenerate conduction band minima in Si(100) and associated effective masses.

The Schrödinger equation must be solved three times, for each of the three pairs of conduction band minima with identical energy and symmetrical wavevector. In fact, as indicated in figure 2, for each direction we have different effective masses to consider in (14) and different masses for the calculation of the two-dimensional density of states.

The electron concentration in the conduction band is therefore

$$n = g_e \frac{\sqrt{2m_z^{1-2}kT}}{\pi\hbar} \sum_i |\Psi_i^{1-2}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{1-2}}{kT}\right) + g_e \frac{\sqrt{2m_z^{3-4}kT}}{\pi\hbar} \sum_i |\Psi_i^{3-4}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{3-4}}{kT}\right) + g_e \frac{\sqrt{2m_z^{5-6}kT}}{\pi\hbar} \sum_i |\Psi_i^{5-6}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{5-6}}{kT}\right) \quad (18)$$

where $g_e$ represents the degeneracy of each valley ($g_e = 2$ in our case), apexes of $m_z$, $\Psi_i$ and $E_i$ indicate the associated pair of conduction band minima and $F_{-1/2}$ denotes the Fermi–Dirac integral of order $-1/2$.

To calculate the hole concentration we solve the Schrödinger equation for heavy holes and for light holes. Therefore we have

$$p = g_h \frac{\sqrt{2m_z^{\mathrm{lh}}kT}}{\pi\hbar} \sum_i |\Psi_i^{\mathrm{lh}}|^2 F_{-\frac{1}{2}}\left(\frac{E_i^{\mathrm{lh}} - E_f}{kT}\right) + g_h \frac{\sqrt{2m_z^{\mathrm{hh}}kT}}{\pi\hbar} \sum_i |\Psi_i^{\mathrm{hh}}|^2 F_{-\frac{1}{2}}\left(\frac{E_i^{\mathrm{hh}} - E_f}{kT}\right) \quad (19)$$

where $g_h$ is valley degeneracy ($g_h = 1$ in our case) and lh and hh refer to light and heavy holes, respectively.

The flow diagram of the algorithm implemented is shown in figure 3. The two-dimensional self-consistent Poisson–Schrödinger equation is discretized on a rectangular grid with the box-integration method. The energy bands and the electron and hole concentrations are calculated using the semiclassical approximation except where differently specified.

The algorithm implemented starts from an initial guess of the unknown potential and solves the Poisson equation using the Newton–Raphson method. The semiclassical solution obtained is used as the initial guess in the case of quantum simulation. The quantum simulation requires large memory occupancy and computing time, because of

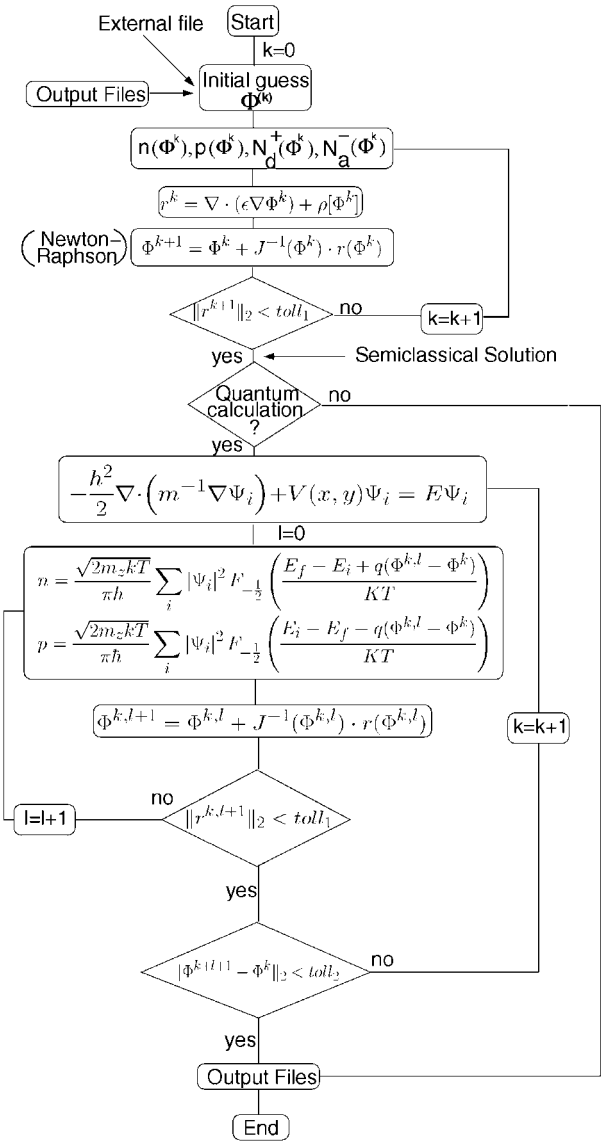



**Figure 3.** Flow diagram of the algorithm implemented.

the nested Poisson/Schrödinger solution within a Newton–Raphson cycle. In order to obtain a fast and converging algorithm we used an approximation [14] that allowed us to solve the Schrödinger equation as few times as possible and to reach the solutions in a reduced number of steps.

Within a Newton–Raphson cycle eigenfunctions are affected by small variations and hence it is possible to consider them as constant and to solve only the nonlinear Poisson equation. At each step, each eigenvalue is corrected by the difference between the actual guess of the potential and the potential with which we have solved the previous Schrödinger equation. Hence the term $E_i$ of equations (18) and (19) becomes $E_i - q(\Phi - \Phi^{old})$.

The algorithm stops when the two-norm of the difference between the value of $\Phi$ at the end of two successive Newton–Raphson cycles is smaller than a fixed value.

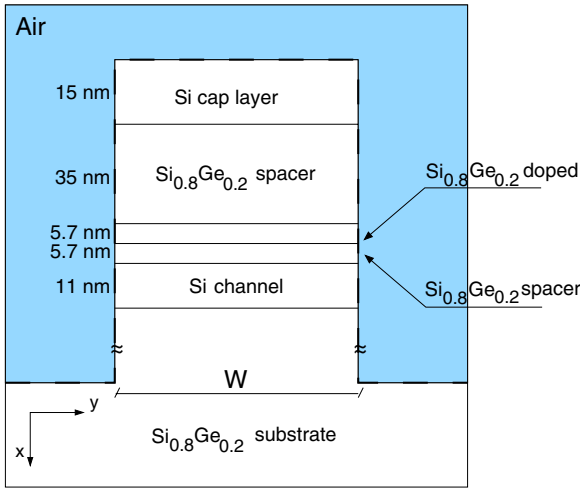



**Figure 4.** Structure of the etched silicon–germanium quantum wire considered in the simulation.

(This figure is in colour only in the electronic version)

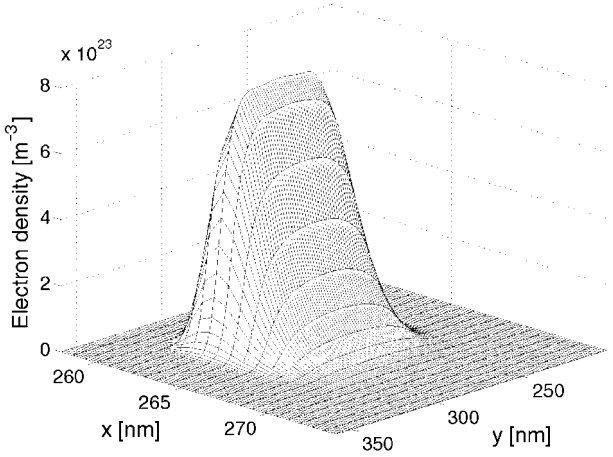



**Figure 5.** Electron density in the Si channel in the case $D_s = 2.7 \times 10^{13}$ cm$^{-2}$ eV$^{-1}$, $W = 160$ nm and $\Phi^* = 4.4$ V. The electron concentration is calculated by solving the Schrödinger equation in the strained Si layer.

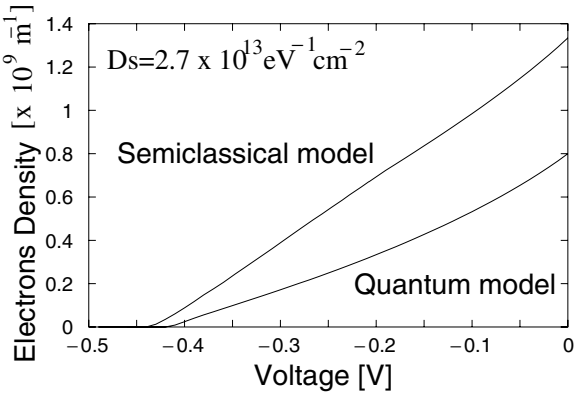



**Figure 6.** Electron density in the strained silicon channel calculated with a semiclassical model and with a quantum model as a function of the voltage applied to the external gate.

### *2.4. Model for interface states*

The states at the air–semiconductor interface are described by a model based on two parameters, $\Phi^*$ and $D_s$. The 'effective work function' at the surface $q\Phi^*$ is the energy difference between the vacuum energy level $E_0$ and the Fermi level $E_F$ at the surface when the surface charge density is zero. We make the assumption that all surface states below $\Phi^*$ behave as donors and all surface states above $\Phi^*$ behave as acceptors, and assume a uniform concentration of states per unit energy per unit area $D_s$. Occupied acceptor states at the exposed surface deplete the semiconductor in the vicinity of the surface. The charge density $Q_s$ at the exposed semiconductor surface can be expressed as

$$Q_s = -qD_s[E_F - (E_0 - q\Phi^*)]. \tag{20}$$

In the case of $D_s \to \infty$ we have Fermi-level pinning at the surface, that is $E_F$ in the semiconductor is pinned at the value imposed by surface states and therefore $E_F = (E_0 - q\Phi^*)$. We have to make the additional assumption that surface charge is very effective in screening the electric field, and that therefore we have vanishing electric field in the air above the semiconductor. Such an assumption has been verified in a few cases, and allows us to exclude from the simulation domain the region above the semiconductor.

## 3. Results and discussion

First, we have studied the effects of interface states in narrow silicon–germanium quantum wires. In particular, the considered structure is schematically represented in figure 4. Here, a Si layer is grown on a $Si_{0.8}Ge_{0.2}$ virtual substrate and, because of the tensile strain imposed by the substrate, it represents the channel in which the inversion layer will form. The $Si_{0.8}Ge_{0.2}$ doped layer provides carriers to the channel formation in the strained silicon layer, which can be undoped in order to reduce impurity scattering.

The self-consistent Poisson/Schrödinger equation is discretized onto a rectangular grid of 108 × 137 points and the electron concentration has been calculated by solving the Schrödinger equation inside the tensile strained silicon channel. We have computed, at a temperature of 4.2 K, the band alignment for a width of 160 nm and observed that electrons are confined by the discontinuity in the conduction band $\Delta E_c$, of about 122 meV, between the strained silicon layer and the virtual substrate. The quantum electron density is represented in figure 5, where it is shown that the electrical waveguide width is about 95 nm, instead of 160 nm, because of the electron depletion induced by interface states at the exposed surfaces. Finally, we can note that, as a consequence of the strain-induced splitting and quantum confinement, there are only two propagating modes in the waveguide.

In order to test the screening effects of interface states and the sensitivity of the electron channel to an external voltage, we have surrounded the heterostructure with a metallic gate and calculated the electron density in the channel as a function of the applied voltage, by using first a semiclassical model and then a quantum model. Results are shown in figure 6.

Electron density obtained with the semiclassical model is always larger than the quantum electron density because in the first case electrons occupy levels starting from the bottom of

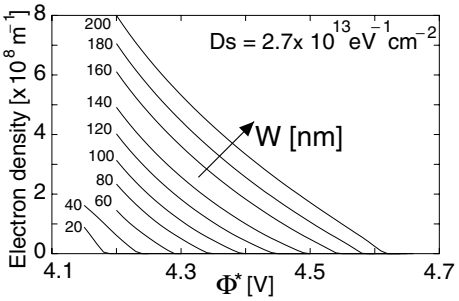



**Figure 7.** Integrated electron density in the channel as a function of the effective work function at the exposed surface, for different widths of the quantum wire.

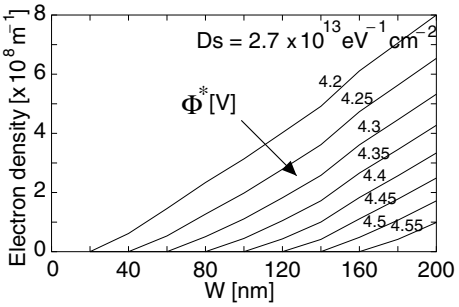



**Figure 8.** Integrated electron density in the strained Si layer as a function of the etched wire width $W$ for different values of the effective work function. The value of $W$ for which the electron density is zero is a sort of threshold width for the wire.

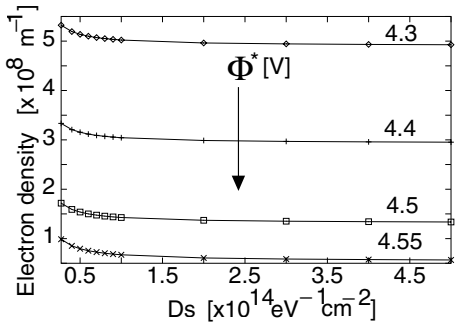



**Figure 9.** For $D_s > 10^{14}$ cm$^{-2}$ eV$^{-1}$, the Fermi level in the semiconductor is pinned at the value imposed by the interface states. Hence the electron concentration in the Si channel becomes independent of the exact value of $D_s$.

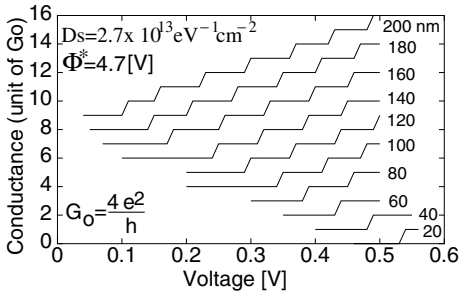



**Figure 10.** Number of propagating modes as a function of the gate voltage for different widths of the etched wire. For the purpose of presentation each curve is shifted by one conductance quantum.

the conduction band, while in the second case available levels start from the first eigenvalue of the Schrödinger equation.

We have calculated the quantum electron density in waveguides of different widths and observed that the considered value of interface state density $D_s$ [15] induces a quasi-Fermi-level pinning: range of variation of the Fermi level within the gap is very small for finite density of interface states and, as $\Phi^*$ is increased, the distance between the conduction band and the Fermi level in the semiconductor increases and the silicon channel is progressively depleted. Results are shown in figure 7, where the electron density is represented as a function of $\Phi^*$ and we can see that the larger the electron waveguide, the higher the value of $\Phi^*$ at which the channel is completely depleted.

At this point we have chosen a fixed value for $D_s$ and $\Phi^*$ and calculated the electron density in the strained Si channel as a function of the wire width. It is possible to calculate, as a function of $\Phi^*$, a sort of threshold width below which there is no electron density in the channel, as represented in figure 8. Indeed, when the wire is particularly thin, the depletion induced by surface states is of primary importance and causes the complete depletion of the channel.

In order to verify the Fermi-level pinning due to surface states at the exposed surface we have calculated the electron concentration in the Si channel as a function of the interface state density. As shown in figure 9, for $D_s > 10^{14}$ eV$^{-1}$ cm$^{-2}$ the electron density saturates because the Fermi level in the semiconductor is pinned at a value imposed by the surface conditions.

The use of silicon–germanium to implement high-mobility electron devices is suggested by the consideration that strain causes a splitting of the sixfold degenerate conduction band into two- and fourfold degenerate valleys. In our particular case, the twofold degenerate valleys are 120 meV lower in energy than the other fourfold degenerate valleys.

In the case of tensile strain we can consider that only the twofold valleys are occupied by electrons while the other valleys are empty. Hence we have several inter-band transitions forbidden, so intervalley scattering is reduced. In the case of silicon–germanium quantum wires it is possible to obtain an additional mobility improvement by increasing the separation between one-dimensional subbands to suppress inter-subband scattering. For this reason, we have investigated the dependence of subband splitting on the waveguide width.

In figure 10 the conductance of the wire is plotted as a function of the voltage applied to the external gate. In the absence of magnetic field the conductance is quantized in units of $N\frac{4e^2}{h}$ as the gate voltage is tuned, where $N$ represents the number of occupied subbands. The number 4 in the previous formula is due to the contribution of both the spin degeneracy ($g_s = 2$) and valley degeneracy ($g_e = 2$) [16].

As the gate voltage is tuned, it is possible to select the number of propagating modes in the waveguide and hence to vary the quantized conductance in the channel.

In figure 11 the one-dimensional subband separation is plotted as a function of the waveguide width. The one-dimensional subband separation can also be varied as a function of the germanium concentration in the substrate: the discontinuity $\Delta E_c$ between adjacent layers is responsible for the electron confinement in the waveguide. For a waveguide 20 nm wide, we obtain a separation between the first and the

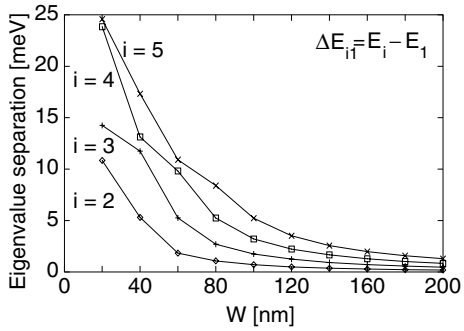



**Figure 11.** In the figure, the distance between the first four occupied modes and the first is represented as a function of the waveguide width. In a waveguide 100 nm wide a quasi-parabolic potential provides equally separated eigenvalues.

second subband of 10 meV, that means population mainly in the first subband. However, it is difficult to evaluate the improvement of mobility at room temperature, because subband separation is still smaller than thermal energy at 300 K (25.9 meV).

## 4. Conclusion

We have developed a program for the simulation of etched silicon–germanium quantum wires, which takes into account strain in the SiGe material system and interface states at the exposed semiconductor surface. The program in based on the self-consistent solution of Poisson and Schrödinger equations in two dimensions. We have investigated the effect of surface states, described by a simple model based on two parameters, on the electron density in the quantum wire. Comparison with experimental results would allow us to determine the value of the relevant parameters. In addition, we have investigated the subband splitting in the quantum wire, as a function of the etched wire width. Such splitting is important to evaluate possible mobility improvements in SiGe quantum wires with respect to the silicon bulk, due to the strain-induced lift of the degeneracy of conduction band minima, and to the suppression of inter-subband scattering. Results from this model need to be validated with experiments.

## References

[1] Crabbé E F, Meyerson B S, Stork J M C and Harame D L 1993 *Tech. Dig. IEDM* **93** 83–6
[2] Gruhle A and Schuppen A 1997 *Thin Solid Films* **294** 246–9
[3] Harame D L, Comfort J H, Cressler J D, Crabbé E F, Sun J Y-C, Meyerson B S and Tice T 1995 *IEEE Trans. Electron Devices* **42** 455–68
[4] Harame D L, Comfort J H, Cressler J D, Crabbé E F, Sun J Y-C, Meyerson B S and Tice T 1995 *IEEE Trans. Electron Devices* **42** 469–82
[5] Dunford R B, Paul D J, Pepper M, Coonan B, Griffin N, Redmon G, Crean G M, Hollander B and Mantl S 2000 *Microelectron. Eng.* **53** 209–12
[6] Sadek A, Ismail K, Armstrong A, Antoniadis A and Stern F 1996 *IEEE Trans. Electron Devices* **43** 1224–32
[7] Ismail K, Nelson S F, Chu J O and Meyerson B S 1993 *Appl. Phys. Lett.* **63** 660–2
[8] König U and Schäffler F 1993 *IEEE Electron Device Lett.* **14** 205–7
[9] Giovine E, Cianci E, Foglietti V, Notargiacomo A and Evangelisti F 2000 *Microelectron. Eng.* **53** 217–19
[10] Giovine E, Notargiacomo A, Di Gaspare L, Palange E, Evangelisti F, Leoni R, Castellano G, Torrioli G and Foglietti V 2001 *Nanotechnology* **12** 132–5
[11] Van de Walle C G and Martin R M 1986 *Phys. Rev.* B **34** 5621–33
[12] Rieger M and Vogl P 1993 *Phys. Rev.* B **48** 14 276–86
[13] People R and Bean J C 1986 *Appl. Phys. Lett.* **48** 538–40
[14] Trellakis A and Galick T 1997 *J. Appl. Phys.* **81** 7880–4
[15] Sze S M 1981 *Physics of Semiconductor Devices* (New York: Wiley)
[16] Többen D, Wharam D A, Abstreiter G, Kotthaus J P and Schäffler F 1995 *Semicond. Sci. Technol.* **10** 711–4

INSTITUTE OF PHYSICS PUBLISHING NANOTECHNOLOGY
Nanotechnology **13** (2002) 267–273 PII: S0957-4484(02)31451-X

# Quantum confinement in silicon–germanium electron waveguides

**G Curatola and G Iannaccone**

Dipartimento di Ingegneria dell'Informazione, Universitá degli studi di Pisa, Via Diotisalvi 2, I-56122 Pisa, Italy

E-mail: g.curatola@iet.unipi.it and g.iannaccone@iet.unipi.it

Received 3 December 2001, in final form 18 March 2002
Published 9 May 2002
Online at stacks.iop.org/Nano/13/267

**Abstract**
We have simulated the electronic properties of silicon–germanium electron waveguides defined by selective etching on a SiGe heterostructure. In particular, we have investigated the dependence of quantum confinement and of one-dimensional subband separation on the waveguide width. Indeed, a larger subband separation means a larger dephasing length and larger electron mobility in the waveguide, and therefore increased possibilities of detecting mesoscopic transport effects. Accurate modelling of SiGe electron waveguides requires us to take into account the effect of strain in the SiGe heterostructure and of the interface states at the exposed SiGe surface, and to solve the Poisson–Schrödinger equation in two dimensions. Results are also shown for a structure in which a gate electrode is evaporated onto the SiGe waveguide, realizing a three-terminal device in which the gate voltage is used to control the number of propagating modes, and therefore the conductance of the channel.

## 1. Introduction

Silicon–germanium technology holds great promise to reduce the cost of consumer products and help make new applications possible. It would allow device designers to simultaneously increase speed, reduce noise and reduce power supply requirements. Although high-speed devices have been fabricated with the $Al_xGa_{1-x}As$/GaAs material system, a silicon-based technology would allow integration with the CMOS process. By using SiGe heterostructures it is possible to realize conduction and valence band offsets for obtaining enhanced-mobility devices.

SiGe heterojunction bipolar transistors (HBTs) have recently reached the marketplace for high-frequency applications: mixers, low-power amplifiers and global position systems (GPSs). HBTs can be integrated with CMOS by inserting a narrow $Si_xGe_{1-x}$ ($x < 0.3$) base layer into a BiCMOS fabrication process. The $Si_xGe_{1-x}$ base may be doped with higher densities with respect to a normal bipolar transistor and a graded Ge base may be grown in order to accelerate the carriers across the base. In research centres, HBTs with $f_T$ up to 120 GHz [1] and $f_{max}$ up to 150 GHz [2] have been demonstrated, while more modest values have been obtained for HBTs introduced into the market place ($f_T = 50$–60 GHz) [3, 4].

In contrast, SiGe field-effect transistors (HFETs) still require significant research before marketable products may come to fruition, even if recently the interest in high-mobility silicon–germanium devices has been greatly increased.

As an example, silicon–germanium field effect transistors cannot withstand the high temperature of standard CMOS processing, which can cause diffusion of dopants and of Ge, and the relaxation of the strained layer. A solution could be to use a modified low-thermal-budget CMOS process with deposited gate oxides, as indicated by Paul *et al* [5]. A proper choice of the thermal budget, which would allow the activation of impurities implanted into the ohmic contacts and in the poly gate and avoid segregation effects, would allow us to realize HFETs with higher transconductance than comparable Si MOSFETs. The higher current drive at low $V_{DS}$ can be attributed to the higher mobility of the strained Si channel, as simulated by Sadek *et al* [6] in the case of 0.2 $\mu$m devices.

These encouraging results could be used to improve the electrical behaviour of p-MOSFETs, which limit performance of CMOS circuits, because the hole mobility is 2.5 times lower than the electron mobility. Unfortunately, several problems must be solved to fabricate silicon–germanium

MOSFETs. The mobility improvement due to the splitting of the conduction and valence bands is undermined by alloy scattering and the small valence band discontinuity can produce parallel conducting channel at small $V_{GS}$, which causes a mobility degradation for both carriers.

SiGe modulation-doped field effect transistors (MODFETs) are not affected by parallel conducting channels and have shown switching times comparable to GaAs/AlGaAs n-MODFETs and lower power requirements with respect to the silicon counterpart [7, 8]. These facts make the investigation of the practical performance limits of SiGe MODFETs important. At present, novel techniques of fabrication [9] allow us to realize silicon–germanium quantum wires with widths down to 15 nm.

As the dimension of the wire is shrunk, the effects of the interface state become more and more important, simply because they are responsible for depletion of the wire [10]. For this reason, we need an accurate model for interface states.

In this paper we focus on the effects of interface states and of quantization on the carrier concentration in silicon–germanium quantum wires of different dimensions. Furthermore, acting on the germanium concentration and on the wire dimensions, it is possible to vary the separation between occupied modes in the waveguide and hence to select the number of propagating modes.

Because of the reduced dimensions of such structures, quantum effects are very important and, in order to obtain accurate analysis of the problems investigated, we have used a two-dimensional simulation code based on the nested solution of the Poisson–Schrödinger equation. The equations are solved numerically with the Newton–Raphson algorithm, after a discretization based on the box-integration method.

## 2. Model

### *2.1. Band alignment at the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ interface*

Let us consider a strained $Si_\alpha Ge_{1-\alpha}$ epitaxial layer deposited onto a relaxed $Si_\beta Ge_{1-\beta}$ substrate, which imposes the lattice constant in the plane perpendicular to the growth direction. Let $\alpha$ represent the germanium mole fraction in the epi-layer and $\beta$ represent the germanium mole fraction in the substrate. The strain in the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ heterostructure, as a consequence of the thermal and lattice mismatch between the substrate and the epitaxial layer, affects the energy band structure, the energy gap, the curvature at the conduction and valence band minima and the degeneracy. Changes in the curvature affect the carrier effective masses, mobility and effective channel velocity. Stress causes a splitting of the sixfold degenerate conduction band into two- and fourfold degenerate valleys and hence leads to a preferential occupation of conduction band minima and to reduced scattering. In addition, the valence band degeneracy at the $\Gamma$ point is lifted. Heavy holes are at the valence band edge under biaxial compressive strain ($\alpha > \beta$); light holes are at the valence band edge in the case of tensile strain ($\alpha < \beta$). The splitting of the conduction band minima and of valence band edges has been computed in [11] by means of self-consistent simulations based on the local density functional and *ab initio* pseudopotential methods. Spin–orbit splitting effects in the valence band are included *a posteriori*. The effective masses for electron and holes are calculated in [12] by means of nonlocal empirical pseudopotential calculations with spin–orbit interactions, while the energy gap in the strained layer has been derived by adding the band offsets, previously calculated, to the gap of the relaxed layer given by [13].

### *2.2. Band parameters in strained $Si_\alpha Ge_{1-\alpha}$ alloys on $Si_\beta Ge_{1-\beta}$ substrates*

Here the procedure is described for calculating the band alignment at the $Si_\alpha Ge_{1-\alpha}/Si_\beta Ge_{1-\beta}$ interface. In figure 1 the band alignment at the interface between a strained layer deposited onto a relaxed substrate is shown and the strain-induced splitting of the conduction band minima and valence band edges is highlighted. It is important to notice that the procedure implemented is still valid even if the substrate is itself strained and the lattice constant is imposed by a third unstrained layer, but fails if the germanium mole fraction in the substrate or in the epi-layer is larger than 0.85, because the position of conduction band minima in the $k$-space is changed.

The lattice constant of the substrate, as a function of the germanium concentration, can be expressed as [12]

$$a_o(\beta) = a_o(\mathrm{Si}) + 0.200\,326\alpha(1-\beta) + [a_o(\mathrm{Ge}) - a_o(\mathrm{Si})]\beta^2. \tag{1}$$

The epi-layer lattice constant $a_\perp$ in the growth direction is given by [12]

$$a_\perp(\alpha) = a_o(\alpha)\left[1 - 2\frac{c_{12}(\alpha)}{c_{11}(\alpha)}\frac{a_\parallel(\alpha) - a_o(\alpha)}{a_o(\alpha)}\right] \tag{2}$$

where $c_{11}$ and $c_{12}$ are the elastic constants.

Starting from the difference, $\Delta E_{vav}$, between the weighted averages of valence band edges at $\Gamma$, we have calculated the splitting of the three valence band edges $\Delta E_{v1}$, $\Delta E_{v2}$ and $\Delta E_{v3}$ as a function of $\alpha$ and $\beta$ [11, 12]:

$$\Delta E_{vav} = (0.047 - 0.06\beta)(\alpha - \beta) \tag{3}$$

$$\Delta E_{v2} = \tfrac{1}{3}\Delta_0 - \tfrac{1}{2}\delta E_{001} \tag{4}$$

$$\Delta E_{v1} = -\tfrac{1}{6}\Delta_0 + \tfrac{1}{4}\delta E_{001} + \tfrac{1}{2}[(\Delta_0)^2 + \Delta_0\delta E_{001} + \tfrac{9}{4}(\delta E_{001})^2]^{\frac{1}{2}} \tag{5}$$

$$\Delta E_{v3} = -\tfrac{1}{6}\Delta_0 + \tfrac{1}{4}\delta E_{001} - \tfrac{1}{2}[(\Delta_0)^2 + \Delta_0\delta E_{001} + \tfrac{9}{4}(\delta E_{001})^2]^{\frac{1}{2}} \tag{6}$$

where $\Delta_0$ [11] is the experimental spin–orbit splitting in the unstrained material, and $\delta E_{001}$ [11] is the linear splitting of the multiplet.

Similarly, the splitting of the two- and fourfold conduction band minima [11, 12] can be expressed as follows:

$$\Delta E_{c2} = \tfrac{2}{3}E_u(\varepsilon_{zz} - \varepsilon_{xx}) \tag{7}$$

$$\Delta E_{c4} = -\tfrac{1}{3}E_u(\varepsilon_{zz} - \varepsilon_{xx}) \tag{8}$$

where $\varepsilon_{ii}$, with $i = x, y, z$, are the components of the symmetric strain tensor and $E_u$ is a deformation potential. Hence we can calculate the discontinuity between the valence band edges, the gap in the strained layer and finally the discontinuity between the conduction band minima:

$$\Delta E_v = \Delta E_{vav} - \Delta E_{vi}^{sub} + \Delta E_{vi} \tag{9}$$

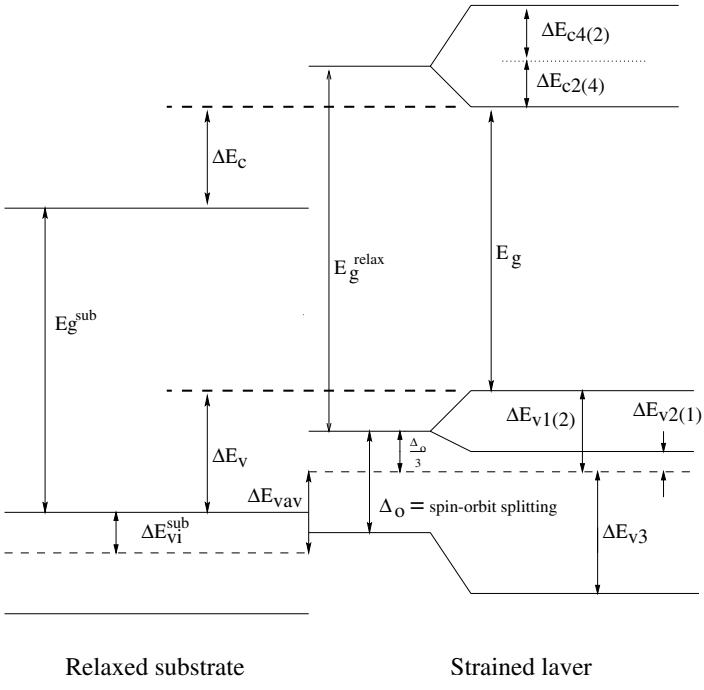



**Figure 1.** Band alignment at the interface between a $Si_\alpha Ge_{1-\alpha}$ layer and a $Si_\beta Ge_{1-\beta}$ substrate. The figure represents the case $\alpha < \beta$, while the numbers in parentheses refer to the case $\alpha > \beta$. Note that the conduction band minimum splittings are referred to the energy of the unstrained sixfold degenerate valleys and hence $\Delta E_{c4(2)}$ and $\Delta E_{c2(4)}$ have opposite sign.

where $\Delta E_{vi}^{sub}$ and $\Delta E_{vi}$ represent the difference between the maximum and the average valence band edge in the substrate and in the epitaxial layer, respectively.

$i = 1$ (2) in the case of tensile (compressive) strain. We have

$$E_g = E_g^{relax} + \Delta E_{ci} - \Delta E_{vj} + \Delta E_{go} - \frac{\Delta_0}{3} \tag{10}$$

where $\Delta E_{go}$ [13] is the gap variation as a consequence of an uniaxial strain in the growth direction and $j = 2, 4$ in the case of tensile or compressive strain, respectively. $E_g^{relax}$ is the value of the energy gap known for the unstrained layer [13]. Finally

$$\Delta E_c = E_g + \Delta E_v - E_g^{sub}. \tag{11}$$

The electron effective masses are calculated as a function of $\alpha$ and $\beta$ [12], while we have used for holes the values known for silicon. In particular, electron effective masses are calculated in the following way:

$$m_s(\alpha, \beta) = [1, (\alpha - \beta), (\alpha - \beta)^2] \cdot \boldsymbol{W} \cdot \begin{bmatrix} 1 \\ (\alpha + \beta) \end{bmatrix} \tag{12}$$

where $s = l, t1, t2$ identifies the longitudinal and transverse mass and $\boldsymbol{W}$ is a $3 \times 2$ matrix [12].

### *2.3. Self-consistent solution of the Poisson–Schrödinger equation*

The nonlinear Poisson equation for the electrostatic potential $\Phi$ is

$$\begin{aligned} \nabla \cdot (\epsilon \nabla \Phi) &= -\rho[\Phi] \\ &= -q[-n[\Phi] + p[\Phi] + N_D^+[\Phi] - N_A^-[\Phi]] \end{aligned} \tag{13}$$

where $\epsilon$ is the dielectric constant, $q$ the electron charge, $n$ and $p$ the electron and hole concentrations, respectively, and $N_D^+$ and $N_A^-$ the ionized donor and acceptor concentrations.

Consider an electron in a region of dimensions $L_x$, $L_y$, $L_z$ with $L_x, L_y \ll L_z$ and let us assume that the structure has translational symmetry along the $z$-axis, so that all quantities depend only on $x$ and $y$.

The two-dimensional Schrödinger equation, in the case of time-independent potential $V(x, y)$, reads

$$-\frac{\hbar^2}{2} \nabla \cdot (m^{-1} \nabla \Psi_i) + V(x, y) \Psi_i = E_i \Psi_i \tag{14}$$

where $\Psi_i$ represents the $i$th eigenfunction, $E_{i,j}$ is the $i$th eigenenergy and $m$ is the electron effective mass tensor in the plane perpendicular to the direction of propagation:

$$m = \begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix}. \tag{15}$$

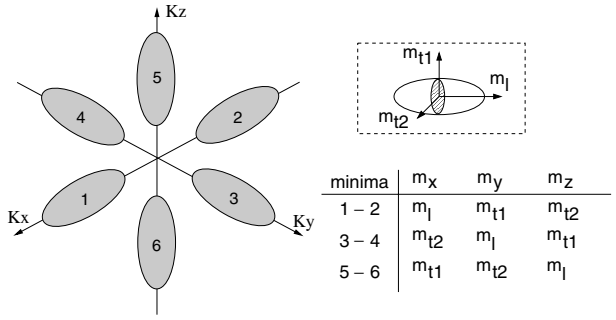
Note that we enumerate the eigenvalues $E_i$ in ascending order ($E_1$ has the lowest energy, $E_n$ has the highest energy, for $i = 1, \ldots, n$), and $\Psi_i$ is the corresponding eigenfunction. The electron energy can be written as

$$E(i, n_z) = E_i + \frac{\hbar^2 k_{nz}^2}{2m_z} \tag{16}$$

where $E_i$ represents the eigenvalue of (14) and $\hbar^2 k_{nz}^2 / 2m_z$ is the energy along the $z$-axis.

The density of states per unit of volume and energy near a conduction band minimum is given by

$$N(E) = \frac{\sqrt{2m_z}}{\pi \hbar} \sum_i |\Psi_i|^2 (E - E_i)^{-\frac{1}{2}} u(E - E_i). \tag{17}$$




| minima | $m_x$ | $m_y$ | $m_z$ |
|---|---|---|---|
| 1 – 2 | $m_l$ | $m_{t1}$ | $m_{t2}$ |
| 3 – 4 | $m_{t2}$ | $m_l$ | $m_{t1}$ |
| 5 – 6 | $m_{t1}$ | $m_{t2}$ | $m_l$ |

**Figure 2.** Sixfold degenerate conduction band minima in Si(100) and associated effective masses.

The Schrödinger equation must be solved three times, for each of the three pairs of conduction band minima with identical energy and symmetrical wavevector. In fact, as indicated in figure 2, for each direction we have different effective masses to consider in (14) and different masses for the calculation of the two-dimensional density of states.

The electron concentration in the conduction band is therefore

$$n = g_e \frac{\sqrt{2m_z^{1-2}kT}}{\pi\hbar} \sum_i |\Psi_i^{1-2}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{1-2}}{kT}\right) + g_e \frac{\sqrt{2m_z^{3-4}kT}}{\pi\hbar} \sum_i |\Psi_i^{3-4}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{3-4}}{kT}\right) + g_e \frac{\sqrt{2m_z^{5-6}kT}}{\pi\hbar} \sum_i |\Psi_i^{5-6}|^2 F_{-\frac{1}{2}}\left(\frac{E_f - E_i^{5-6}}{kT}\right) \quad (18)$$

where $g_e$ represents the degeneracy of each valley ($g_e = 2$ in our case), apexes of $m_z$, $\Psi_i$ and $E_i$ indicate the associated pair of conduction band minima and $F_{-1/2}$ denotes the Fermi–Dirac integral of order $-1/2$.

To calculate the hole concentration we solve the Schrödinger equation for heavy holes and for light holes. Therefore we have

$$p = g_h \frac{\sqrt{2m_z^{\mathrm{lh}}kT}}{\pi\hbar} \sum_i |\Psi_i^{\mathrm{lh}}|^2 F_{-\frac{1}{2}}\left(\frac{E_i^{\mathrm{lh}} - E_f}{kT}\right) + g_h \frac{\sqrt{2m_z^{\mathrm{hh}}kT}}{\pi\hbar} \sum_i |\Psi_i^{\mathrm{hh}}|^2 F_{-\frac{1}{2}}\left(\frac{E_i^{\mathrm{hh}} - E_f}{kT}\right) \quad (19)$$

where $g_h$ is valley degeneracy ($g_h = 1$ in our case) and lh and hh refer to light and heavy holes, respectively.

The flow diagram of the algorithm implemented is shown in figure 3. The two-dimensional self-consistent Poisson–Schrödinger equation is discretized on a rectangular grid with the box-integration method. The energy bands and the electron and hole concentrations are calculated using the semiclassical approximation except where differently specified.

The algorithm implemented starts from an initial guess of the unknown potential and solves the Poisson equation using the Newton–Raphson method. The semiclassical solution obtained is used as the initial guess in the case of quantum simulation. The quantum simulation requires large memory occupancy and computing time, because of

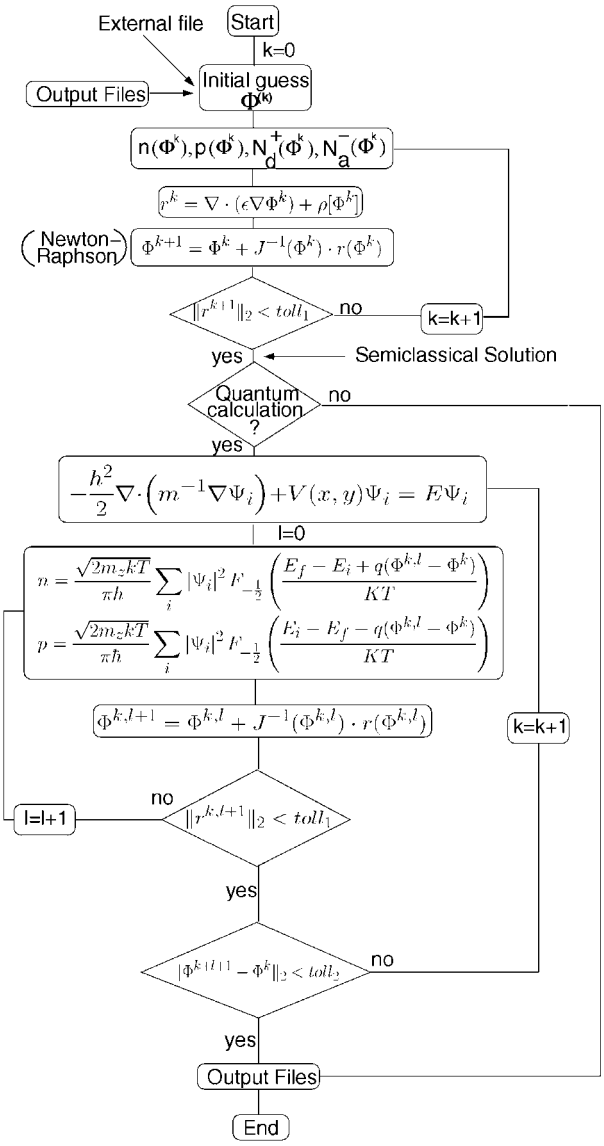



**Figure 3.** Flow diagram of the algorithm implemented.

the nested Poisson/Schrödinger solution within a Newton–Raphson cycle. In order to obtain a fast and converging algorithm we used an approximation [14] that allowed us to solve the Schrödinger equation as few times as possible and to reach the solutions in a reduced number of steps.

Within a Newton–Raphson cycle eigenfunctions are affected by small variations and hence it is possible to consider them as constant and to solve only the nonlinear Poisson equation. At each step, each eigenvalue is corrected by the difference between the actual guess of the potential and the potential with which we have solved the previous Schrödinger equation. Hence the term $E_i$ of equations (18) and (19) becomes $E_i - q(\Phi - \Phi^{old})$.

The algorithm stops when the two-norm of the difference between the value of $\Phi$ at the end of two successive Newton–Raphson cycles is smaller than a fixed value.

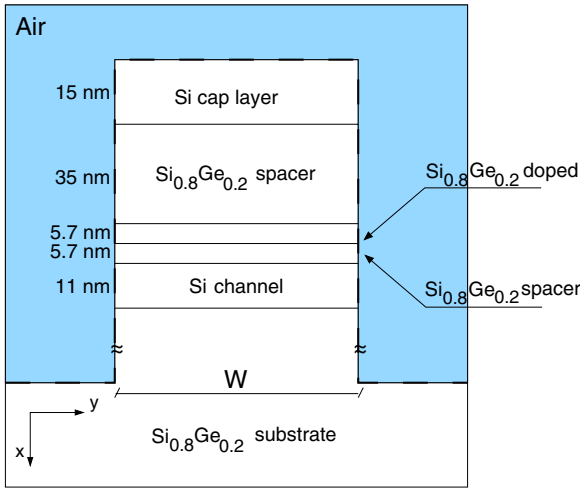



**Figure 4.** Structure of the etched silicon–germanium quantum wire considered in the simulation.

(This figure is in colour only in the electronic version)

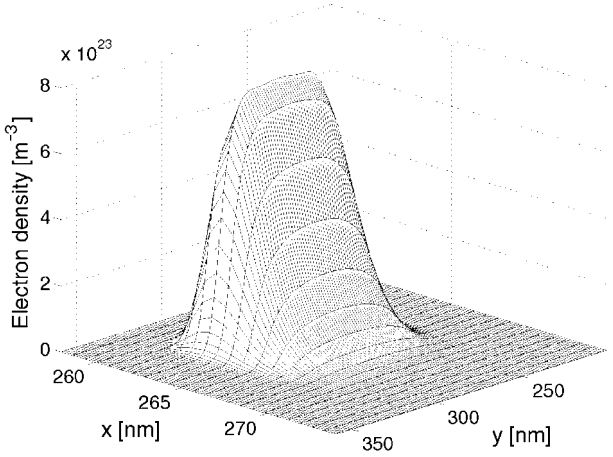



**Figure 5.** Electron density in the Si channel in the case $D_s = 2.7 \times 10^{13}$ cm$^{-2}$ eV$^{-1}$, $W = 160$ nm and $\Phi^* = 4.4$ V. The electron concentration is calculated by solving the Schrödinger equation in the strained Si layer.

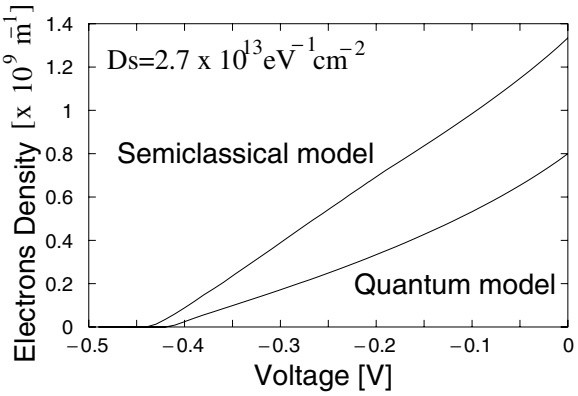



**Figure 6.** Electron density in the strained silicon channel calculated with a semiclassical model and with a quantum model as a function of the voltage applied to the external gate.

### *2.4. Model for interface states*

The states at the air–semiconductor interface are described by a model based on two parameters, $\Phi^*$ and $D_s$. The 'effective work function' at the surface $q\Phi^*$ is the energy difference between the vacuum energy level $E_0$ and the Fermi level $E_F$ at the surface when the surface charge density is zero. We make the assumption that all surface states below $\Phi^*$ behave as donors and all surface states above $\Phi^*$ behave as acceptors, and assume a uniform concentration of states per unit energy per unit area $D_s$. Occupied acceptor states at the exposed surface deplete the semiconductor in the vicinity of the surface. The charge density $Q_s$ at the exposed semiconductor surface can be expressed as

$$Q_s = -qD_s[E_F - (E_0 - q\Phi^*)]. \tag{20}$$

In the case of $D_s \to \infty$ we have Fermi-level pinning at the surface, that is $E_F$ in the semiconductor is pinned at the value imposed by surface states and therefore $E_F = (E_0 - q\Phi^*)$. We have to make the additional assumption that surface charge is very effective in screening the electric field, and that therefore we have vanishing electric field in the air above the semiconductor. Such an assumption has been verified in a few cases, and allows us to exclude from the simulation domain the region above the semiconductor.

## 3. Results and discussion

First, we have studied the effects of interface states in narrow silicon–germanium quantum wires. In particular, the considered structure is schematically represented in figure 4. Here, a Si layer is grown on a $Si_{0.8}Ge_{0.2}$ virtual substrate and, because of the tensile strain imposed by the substrate, it represents the channel in which the inversion layer will form. The $Si_{0.8}Ge_{0.2}$ doped layer provides carriers to the channel formation in the strained silicon layer, which can be undoped in order to reduce impurity scattering.

The self-consistent Poisson/Schrödinger equation is discretized onto a rectangular grid of 108 × 137 points and the electron concentration has been calculated by solving the Schrödinger equation inside the tensile strained silicon channel. We have computed, at a temperature of 4.2 K, the band alignment for a width of 160 nm and observed that electrons are confined by the discontinuity in the conduction band $\Delta E_c$, of about 122 meV, between the strained silicon layer and the virtual substrate. The quantum electron density is represented in figure 5, where it is shown that the electrical waveguide width is about 95 nm, instead of 160 nm, because of the electron depletion induced by interface states at the exposed surfaces. Finally, we can note that, as a consequence of the strain-induced splitting and quantum confinement, there are only two propagating modes in the waveguide.

In order to test the screening effects of interface states and the sensitivity of the electron channel to an external voltage, we have surrounded the heterostructure with a metallic gate and calculated the electron density in the channel as a function of the applied voltage, by using first a semiclassical model and then a quantum model. Results are shown in figure 6.

Electron density obtained with the semiclassical model is always larger than the quantum electron density because in the first case electrons occupy levels starting from the bottom of

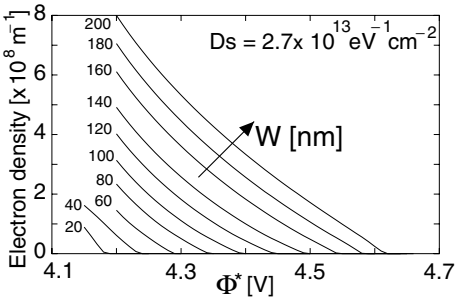



**Figure 7.** Integrated electron density in the channel as a function of the effective work function at the exposed surface, for different widths of the quantum wire.

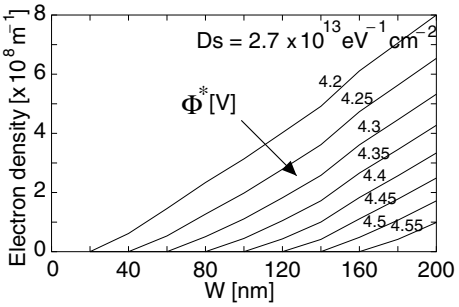



**Figure 8.** Integrated electron density in the strained Si layer as a function of the etched wire width $W$ for different values of the effective work function. The value of $W$ for which the electron density is zero is a sort of threshold width for the wire.

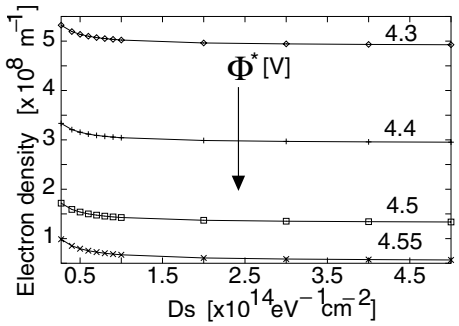



**Figure 9.** For $D_s > 10^{14}$ cm$^{-2}$ eV$^{-1}$, the Fermi level in the semiconductor is pinned at the value imposed by the interface states. Hence the electron concentration in the Si channel becomes independent of the exact value of $D_s$.

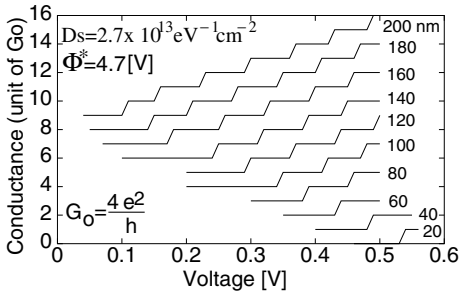



**Figure 10.** Number of propagating modes as a function of the gate voltage for different widths of the etched wire. For the purpose of presentation each curve is shifted by one conductance quantum.

the conduction band, while in the second case available levels start from the first eigenvalue of the Schrödinger equation.

We have calculated the quantum electron density in waveguides of different widths and observed that the considered value of interface state density $D_s$ [15] induces a quasi-Fermi-level pinning: range of variation of the Fermi level within the gap is very small for finite density of interface states and, as $\Phi^*$ is increased, the distance between the conduction band and the Fermi level in the semiconductor increases and the silicon channel is progressively depleted. Results are shown in figure 7, where the electron density is represented as a function of $\Phi^*$ and we can see that the larger the electron waveguide, the higher the value of $\Phi^*$ at which the channel is completely depleted.

At this point we have chosen a fixed value for $D_s$ and $\Phi^*$ and calculated the electron density in the strained Si channel as a function of the wire width. It is possible to calculate, as a function of $\Phi^*$, a sort of threshold width below which there is no electron density in the channel, as represented in figure 8. Indeed, when the wire is particularly thin, the depletion induced by surface states is of primary importance and causes the complete depletion of the channel.

In order to verify the Fermi-level pinning due to surface states at the exposed surface we have calculated the electron concentration in the Si channel as a function of the interface state density. As shown in figure 9, for $D_s > 10^{14}$ eV$^{-1}$ cm$^{-2}$ the electron density saturates because the Fermi level in the semiconductor is pinned at a value imposed by the surface conditions.

The use of silicon–germanium to implement high-mobility electron devices is suggested by the consideration that strain causes a splitting of the sixfold degenerate conduction band into two- and fourfold degenerate valleys. In our particular case, the twofold degenerate valleys are 120 meV lower in energy than the other fourfold degenerate valleys.

In the case of tensile strain we can consider that only the twofold valleys are occupied by electrons while the other valleys are empty. Hence we have several inter-band transitions forbidden, so intervalley scattering is reduced. In the case of silicon–germanium quantum wires it is possible to obtain an additional mobility improvement by increasing the separation between one-dimensional subbands to suppress inter-subband scattering. For this reason, we have investigated the dependence of subband splitting on the waveguide width.

In figure 10 the conductance of the wire is plotted as a function of the voltage applied to the external gate. In the absence of magnetic field the conductance is quantized in units of $N\frac{4e^2}{h}$ as the gate voltage is tuned, where $N$ represents the number of occupied subbands. The number 4 in the previous formula is due to the contribution of both the spin degeneracy ($g_s = 2$) and valley degeneracy ($g_e = 2$) [16].

As the gate voltage is tuned, it is possible to select the number of propagating modes in the waveguide and hence to vary the quantized conductance in the channel.

In figure 11 the one-dimensional subband separation is plotted as a function of the waveguide width. The one-dimensional subband separation can also be varied as a function of the germanium concentration in the substrate: the discontinuity $\Delta E_c$ between adjacent layers is responsible for the electron confinement in the waveguide. For a waveguide 20 nm wide, we obtain a separation between the first and the

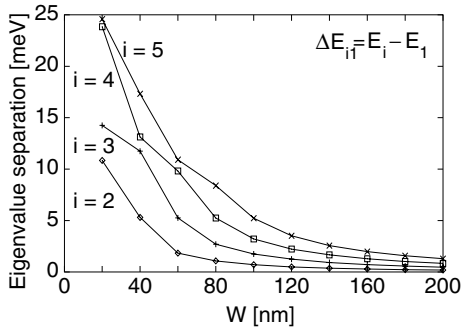



**Figure 11.** In the figure, the distance between the first four occupied modes and the first is represented as a function of the waveguide width. In a waveguide 100 nm wide a quasi-parabolic potential provides equally separated eigenvalues.

second subband of 10 meV, that means population mainly in the first subband. However, it is difficult to evaluate the improvement of mobility at room temperature, because subband separation is still smaller than thermal energy at 300 K (25.9 meV).

## 4. Conclusion

We have developed a program for the simulation of etched silicon–germanium quantum wires, which takes into account strain in the SiGe material system and interface states at the exposed semiconductor surface. The program in based on the self-consistent solution of Poisson and Schrödinger equations in two dimensions. We have investigated the effect of surface states, described by a simple model based on two parameters, on the electron density in the quantum wire. Comparison with experimental results would allow us to determine the value of the relevant parameters. In addition, we have investigated the subband splitting in the quantum wire, as a function of the etched wire width. Such splitting is important to evaluate possible mobility improvements in SiGe quantum wires with respect to the silicon bulk, due to the strain-induced lift of the degeneracy of conduction band minima, and to the suppression of inter-subband scattering. Results from this model need to be validated with experiments.

## References

[1] Crabbé E F, Meyerson B S, Stork J M C and Harame D L 1993 *Tech. Dig. IEDM* **93** 83–6
[2] Gruhle A and Schuppen A 1997 *Thin Solid Films* **294** 246–9
[3] Harame D L, Comfort J H, Cressler J D, Crabbé E F, Sun J Y-C, Meyerson B S and Tice T 1995 *IEEE Trans. Electron Devices* **42** 455–68
[4] Harame D L, Comfort J H, Cressler J D, Crabbé E F, Sun J Y-C, Meyerson B S and Tice T 1995 *IEEE Trans. Electron Devices* **42** 469–82
[5] Dunford R B, Paul D J, Pepper M, Coonan B, Griffin N, Redmon G, Crean G M, Hollander B and Mantl S 2000 *Microelectron. Eng.* **53** 209–12
[6] Sadek A, Ismail K, Armstrong A, Antoniadis A and Stern F 1996 *IEEE Trans. Electron Devices* **43** 1224–32
[7] Ismail K, Nelson S F, Chu J O and Meyerson B S 1993 *Appl. Phys. Lett.* **63** 660–2
[8] König U and Schäffler F 1993 *IEEE Electron Device Lett.* **14** 205–7
[9] Giovine E, Cianci E, Foglietti V, Notargiacomo A and Evangelisti F 2000 *Microelectron. Eng.* **53** 217–19
[10] Giovine E, Notargiacomo A, Di Gaspare L, Palange E, Evangelisti F, Leoni R, Castellano G, Torrioli G and Foglietti V 2001 *Nanotechnology* **12** 132–5
[11] Van de Walle C G and Martin R M 1986 *Phys. Rev.* B **34** 5621–33
[12] Rieger M and Vogl P 1993 *Phys. Rev.* B **48** 14 276–86
[13] People R and Bean J C 1986 *Appl. Phys. Lett.* **48** 538–40
[14] Trellakis A and Galick T 1997 *J. Appl. Phys.* **81** 7880–4
[15] Sze S M 1981 *Physics of Semiconductor Devices* (New York: Wiley)
[16] Többen D, Wharam D A, Abstreiter G, Kotthaus J P and Schäffler F 1995 *Semicond. Sci. Technol.* **10** 711–4