Blaze Simulation of SiGe:Si Heterostructure p-MOSFETs

P. A. Clifton, and A. G. O'Neill*

*Department of Electrical and Electronic Engineering, University of Newcastle, UK.

Introduction

The 2-D heterostructure device simulator Blaze is used here to investigate the design criteria for sub-micron p-channel SiGe:Si heterostructure MOSFETs. Thin layers of compressively strained SiGe grown 'pseudomorphically' on silicon substrates exhibit improved in-plane hole drift mobility relative to Si (800 cm^2/V.s [1] c.f. 450 cm^2/V.s) and may be applied in future enhanced p-channel devices (as shown schematically in Figure 1) for high performance CMOS. The use of a buried channel is also expected to improve carrier mobility and noise performance by reducing the interaction of carriers with the oxide interface. A major constraint on HMOSFET performance is the onset of parasitic inversion at the cap-oxide interface where carriers have degraded mobility. This limits the degree of inversion in the strained channel layer by electrostatic screening and hence degrades the small signal transconductance.

Figure 1. Schematic of the strained SiGe p-channel heterostructure MOSFET.

Device Simulation

Numerical simulation based on ATLAS-Blaze is used here to explore the design parameter space for enhancement (inversion) mode SiGe p-HMOSFETs. Fermi-Dirac statist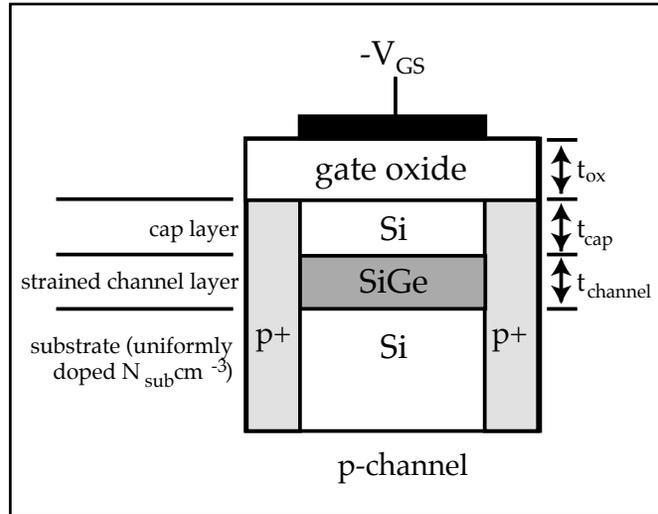ics for carrier populations and a dense mesh specification for the thin epitaxial layers are required for accurate modeling of charge distributions and drift-diffusion based current formulations have been found to be sufficient for the range of channel lengths investigated (down to 0.25µm).

The epitaxial cap and channel layers in the simple device structure are assumed to have low background doping (5×10^{13} cm^{-3}) and the underlying substrate (or n-well) uniformly doped to varying degrees (N_{sub} cm^{-3}) in the range 1×10^{16} cm^{-3} to 5×10^{17} cm^{-3}. Interface states are neglected, a p+ polysilicon gate is used and the threshold voltage is allowed to shift freely according to channel doping and layer thicknesses.

D.C. output characteristics and small signal transconductance have been obtained (the latter by AC analysis [2]) and the respective inversion layer carrier populations in the cap and channel have been extracted by integrating the carrier profiles across the total depths of these layers. The region of integration was defined midway between source and drain at zero drain bias so as to avoid the complicating effect of junction depletion regions.

Density of States in Strained Layers

A reduced effective density of states (DOS) in the valence band, N_V, is inherent in the use of compressively strained SiGe channels on Si [3], being intimately linked to the enhanced hole mobility [4,5]. The lower density of states effective hole mass and the reduced carrier scattering due to the lifting of the valence band degeneracy are both thought to contribute to higher μ_h. As the Ge fraction in a strained SiGe layer, x, is increased, N_V is predicted to fall monotonically [5] - by a factor of 5.6 at x=0.3, an effect that should not be ignored in modeling HMOSFETs.

Benchmarking

An experimental 0.7 micron channel length SiGe p-channel MOSFET reported by Nayak et. al. [6] has been used as a benchmark to support the validity of the simulation results. In modeling this device, which had a SiGe composition factor x=0.2, the effective N_V in the SiGe channel layer was reduced from 3.09×10^{19} cm^{-3} to 7.09×10^{18} cm^{-3} according to [5] and the band offset was taken as 0.15eV based on the mean of the reported values collated in [7]. The carrier mobility in the surface channel was assumed to be degraded with increasing transverse and longitudinal fields using the 'CVT' model [8] and the mobility in the SiGe channel was assumed to be insensitive to the transverse field. By using only the layer mobility values as fitting parameters (μ_h =185 cm^2/V.s in the SiGe and μ_h(max)=255 cm^2/V.s in the Si cap), a good fit to the experimental transconductance curves has been obtained, as shown in Figure 2(a). Extracted inversion layer carrier concentrations, shown in figure 2(b), indicate that the level of inversion in the SiGe channel is in fact very low in these devices and that the main contribution to the drain current comes from the undesirable surface inversion layer in the silicon cap.

The effect of a reduced density of states on the carrier concentration in the strained channel is marked, as shown in Figure 2(b), where the saturated carrier concentration falls from 1.19×10^{12} cm^{-2} to 0.83×10^{12} cm^{-}

^{2} when N_V in the SiGe is reduced by strain from 3.09×10^{19} cm^{-3} to 7.09×10^{18} cm^{-3}. Neglecting the reduction in N_V is liable to lead to falsely low extracted hole mobilities in experimental devices due to overestimation of carrier concentrations in the SiGe channel.

Response Surface Methodology.

In order to thoroughly explore the large parameter space which determines the performance of a HMOS-FET, the experimental $0.7\mu m$ channel length device [6] has been taken as the basis for an investigation by response surface methodology (RSM), a feature of the VWF Automation Tools. The dependencies of the SiGe: Si valence band offset and the hole mobility on SiGe layer composition (x) were respectively modeled as $\Delta E_V = 0.75x$ eV [11] and $\mu_h(SiGe) = 750x$ cm^2/V.s, a conservative assessment of hole mobility based on early literature values [6] [9] [10] for experimental devices. The drawn channel length was maintained at $0.7\mu m$ and N_V in the SiGe was fitted to the data given in [5].
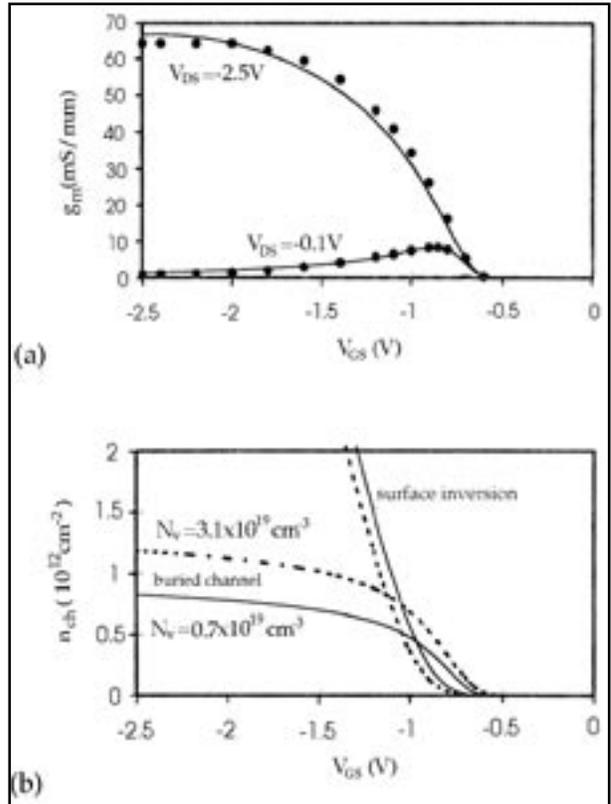
Figure 2.

(a) Transconductance curves obtained by simulation using the fitted model for a $0.7\mu m$ gate length p-HMOSFET. Upper curve is for device in saturation (V_{DS}=-2.5V) and lower curve is linear (V_{DS}=-0.1V). The dots represent the experimental data reported by Nayak et. al. [6].

(b) Surface and channel inversion layer hole concentrations obtained for the same device, both with full (3.09×10^{19} cm^{-3}, dotted lines) and reduced (7.09×10^{18} cm^{-3}, full lines) DOS in the strained SiGe.

The same calibrated device model was used throughout and a 2^4 full factorial design based upon a body centered cubic structure was implemented with the four input factors; oxide thickness (t_{ox}), cap layer thickness (t_{cap}), substrate doping (N_{ox}) and Ge composition in the channel (x). Previous (unreported) studies have shown that neither the channel thickness nor the presence of a spacer layer between the channel and the doped substrate play a significant part in the device performance and hence these factors are neglected here.

Three key output metrics have been modeled as functions of the four input factors: n_{ch}(max) - the maximum hole concentration achieved in the SiGe channel with V_{GS} = -2.5V, ΔV_{GS} - the gate voltage range above the threshold voltage over which the integrated inversion charge in the SiGe exceeds that in the cap and g_m(max) - the maximum transconductance obtained with V_{DS} = -2.5V.

A wide range of functional relationships can be examined between each of the output factors and the four input factors, usually by the use of response surface plots. Selected cuts through the 4-dimensional factor space of the experiment are presented in figures 3 to 5 to illustrate major trends.

Discussion

It may be seen in Figures 3-5 that all the metrics of device performance; n_{ch}(max), g_m(max) and ΔV_{GS} generally deteriorate as the thickness of the cap layer is increased.
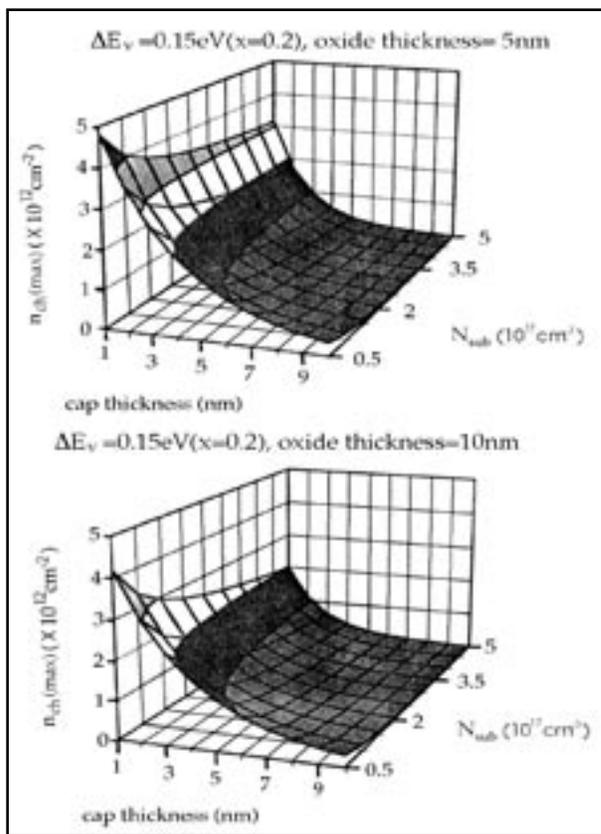

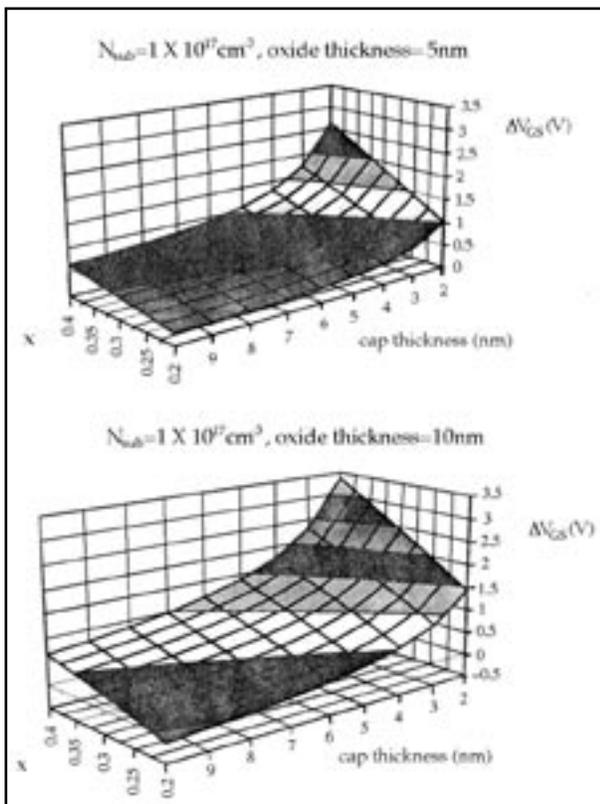Figure 3. RSM results obtained for the device represented in Figure 2.

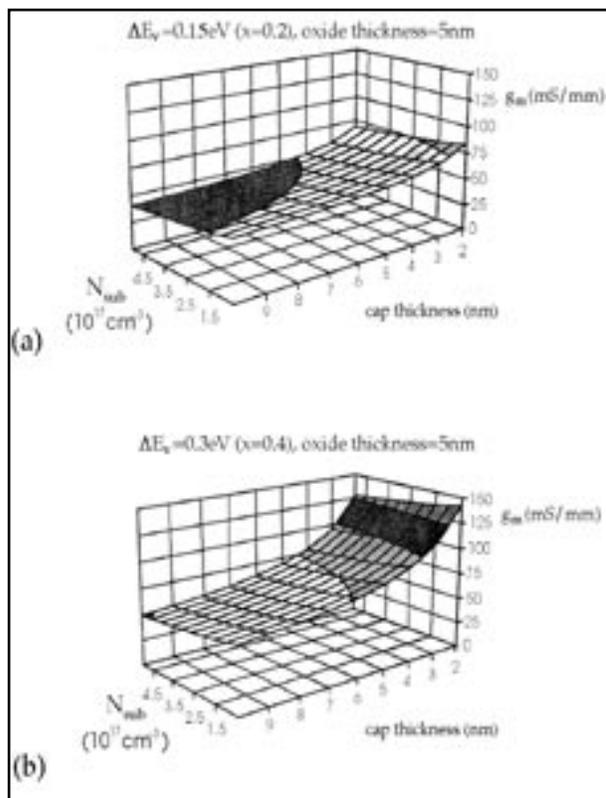Figure 4. RSM results obtained for the device represented in Figure 2.

Figure 5. RSM results obtained for the device represented in Figure 2.

This is particularly important at high substrate doping levels (Figure 3) which lead to strong band bending (high transverse field) in the channel region under gate bias. The threshold for parasitic surface inversion is consequently reduced relative to that for SiGe channel inversion, so causing ΔV_{GS} to decrease. A smaller useful range ΔV_{GS} is concomitant with reduced SiGe inversion charge and hence smaller transconductance. In Figure 4, g_m(max) can be seen to fall as t_{cap} and N_{sub} increase. The use of a large value of band offset (high x) improves ΔV_{GS} markedly, as expected (Figure 5), since the threshold voltage for inversion of the buried channel is reduced relative to that of the surface channel. It is important to note that the role of the band offset is not so much in providing carrier confinement in the strained SiGe as in allowing an inversion layer to be induced in the SiGe rather than at the oxide interface. That is, a smaller offset does not degrade performance by allowing carriers to leak out of the SiGe potential well but rather it reduces the critical bias range ΔV_{GS}.

A thinner gate oxide also results in a higher n_{ch}(max) (not shown) and hence improves current drive and g_m(max) due to the improved capacitive coupling between gate and channel charges. These improvements in performance will always overcome the disadvantage of the small reduction in ΔV_{GS} arising with the thinner oxide which is evident in comparing Figures 5(a) and (b).

Deep Sub-micron HMOSFETs.

It has been seen that the main limitations on HMOSFET operation arise at high levels of sub-channel doping, i.e. under the conditions associated with deep sub-micron devices. A great benefit of RSM is that any combination of input factors (within the modeled parameter space) may be inserted into the derived empirical RS model to determine a good estimation of the device performance. For example, taking a typical 0.35μm MOSFET with $N_{sub} = 4.5 \times 10^{17}$ cm^{-3} and a 9nm thick oxide, negligible conduction in the SiGe channel is achieved for $x=0.2$ and the channel population is still rather limited for $x=0.4$ unless very thin cap layers are employed. Going a stage further to a 0.25μm channel length p-HMOSFET with $N_{sub} \cong 5.5 \times 10^{17}$ cm^{-3}, ΔV_{GS} may diminish to zero (the surface inverts before the channel) even for a cap thickness as small as 7nm. This is illustrated by the simulation results shown in Figure 6 ($\delta=0$) for a device based on the one reported by Kesan et. al. [10] ($\Delta E_v=0.15$eV). This device was simulated using the same numerical model as for the Nayak [6] HMOSFET but fitting values of 350 and 300cm^2/V.s for the maximum channel and cap layer mobilities respectively. The graph of integrated sheet hole concentrations, Figure 6(b) ($\delta=0$), shows that this structure supports very little inversion charge in the buried SiGe channel. It is possible that the reported improvement over an equivalent conventional Si MOSFET may be largely attributable to the use of undoped silicon in the cap.

Clearly, based on the foregoing results, the cap layer should be made as thin as possible. However, a minimum value may be determined by two primary limitations; avoidance of high interface state densities (a minimum thickness of Ge-free silicon of order 6nm may be required [11]) and the avoidance of remote carrier scattering (by the insulator-semiconductor interface). Some experimental evidence suggests the latter may require a cap layer thickness of the order of 10nm [12] [13]. To enable significant benefit to be gained from the use of buried strained layer channels in sub-micron MOSFETs, two options exist - increase the offset potential between the cap and channel layers or reduce the peak field in the semiconductor. Growth of the HMOSFETs on silicon-on-insulator (SIMOX) substrates is one approach to field reduction [14], but, because of the heightened technological demands, does not appear to be a near term solution. Modulation doping provides one alternative.

Pulsed Modulation Doping.

It has been shown experimentally that a 'pulse' or 'delta' doped acceptor layer in the spacer below (but in close proximity to) the SiGe channel allows the inversion layer carrier concentration in the SiGe to be increased [11] [15]. In addition, the pulse doping layer reduces
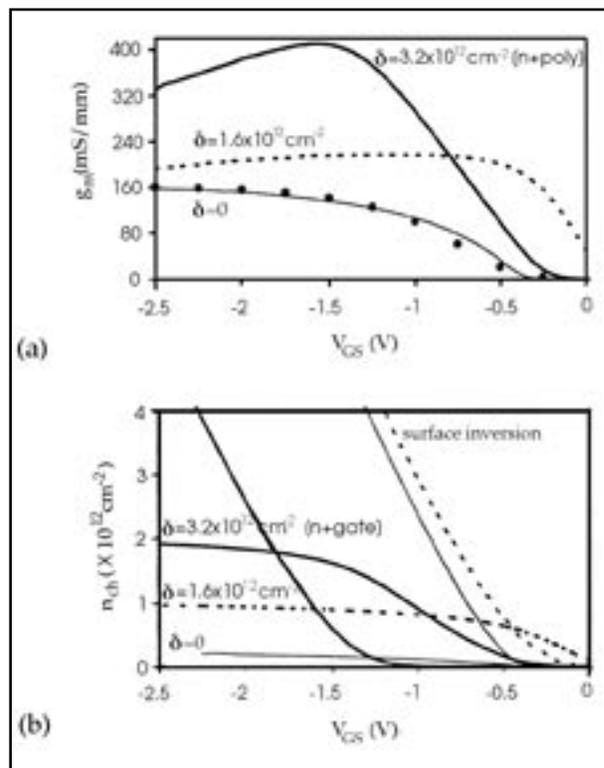

Figure 6.

(a) Transconductance curves obtained by simulation for an experimental 0.25μm gate length p-HMOSFET [9] with and without delta doping (V_{DS}=-2.5V). Dots represent experimental results.

(b) The inversion layer hole concentrations without delta (□) doping layer (full line) and with a □-dose of 0.8×10^{12} cm^{-2} (dotted).

the threshold voltage for inversion of the channel and increases ΔV_{GS}. Figure 6(a), shows that the inefficient device with a 7nm thick cap and effective gate length of 0.25μm (N_{sub} of 5 X 10^{17} cm^{-3}) is enhanced by the addition of 4nm thick pulse doping layers (of varying dose) set back 4nm below the channel. Peak transconductance improves from 144 mS/mm to 268 mS/mm and ΔV_{GS} from 0V to 0.4V as the sheet dose is raised from 0 to 2.0 X 10^{12} cm^{-2}. This very significant increase in gm(max) demonstrates the improvement in performance possible through epitaxial growth capabilities, such as in-situ modulation doping, quite apart from gains achieved by further increasing mobility. A δ-dose of $3.2x10^{12}$cm^{-2} increases g_m to a value in excess of 400 mS/mm, remarkably high for any type of p-channel FET. Note that in this case, an n+ polysilicon gate is required to ensure enhancement mode operation (negative V_T), in the same manner as for a conventional buried channel p-MOSFET.

The increase in ΔV_{GS} is largely due to the reduction in transverse field achieved by the presence of the fully depleted pulse doped layer. This effect is the same as that proposed in the use of a 'back junction' [16] to modify the transverse field but in this case, excessive source-drain leakage in the off state (not shown) is avoided as long as the pulse doped layer is fully depleted.

Summary

As MOSFET gate lengths are shrunk to deep sub-micron dimensions, dopant concentrations in the channel region are forced higher to suppress short channel effects (~$6x10^{17}$ cm^{-3} and $1x10^{18}$ cm^{-3} for 0.25μm and 0.12μm channel lengths respectively). 2d device simulation shows that it becomes ever more difficult to ensure inversion of a buried channel layer in HMOSFET devices with these levels of (sub channel) doping. This trend must be offset by employing maximum offset potentials at the cap-channel interface and minimum cap layer thicknesses. The same argument applies to n-channel HMOSFETs based on strained Si layers on SiGe. In the limit, this would suggest the use of Si (n-) and Ge (p-) channels for CMOS based on a relaxed Si$_{1-x}$Ge$_x$ (x≈0.5) virtual substrate. Modulation doping using pulse doping layers below the strained channel is a useful approach to extending the range of gate voltage over which conduction along the strained channel dominates. Indeed, this may prove essential if improved current drive and transconductance are to be realized in ultra-small geometry devices.

References.

[1]   K. Ismail, J.O. Chu and B.S. Meyerson, Appl. Phys. Lett., 64 (1994) 3124.

[2]   Silvaco International, User Manual Edition 2, (1994).

[3]   R. People, Physics and applications of GeSi/Si strained-layer heterostructures, IEEE J. Quantum Electron., QE-22 (1986) 1696.

[4]   J. M. Hinckley and J Singh, Hole transport theory in pseudomorphic SiGe alloys grown on Si(001) substrates, Phys. Rev. B, 41(5) (1990) 2912.

[5]   T. Manku, J. M. McGregor, A. Nathan, D.J. Roulston, J-P. Noel and D.C. Houghton, Drift hole mobility in strained and unstrained doped SiGe alloys, IEEE Trans. Electron Devices, ED-40 (1993) 1990.

[6]   D.K. Nayak, J.C.S. Woo, J.S. Park, K-L. Wang and K.P. MacWilliams, Enhancement-mode quantum-well GeSi PMOS, Electron Device Lett., 12 (1991) 154.

[7]   C. H. Gan, J. A. DelAlamo, B. R. Bennett, B. S. Meyerson, E. F. Crabbe, C. G. Sodini and L. R. Reif, Si/SiGe valence band discontinuity measurements using a semiconductor-insulator-semiconductor (SIS) heterostructure, IEEE Trans. Electron Devices, ED-41 (12) (1994) 2430.

[8]   C. Lombardi, S. Manzini, A. Saporito and M. Vanzi, A physically based mobility model for numerical simulation of non-planar devices, IEEE Trans. Computer-Aided Design, CAD-7 (11) (1988) 1164.

[9]   V.P. Kesan, S. Subbanna, P.J. Restle, M.J Tejwani, J.M. Aitken, S.S. Iyer and J.A. Ott, High performance 0.25?m p-MOSFETs with silicon-germanium channels for 300K and 77K operation, IEDM Tech. Dig., 91 (1991) 25.

[10]   S. Verdonckt-Vandebroek, E.F. Crabbe, B.S. Meyerson, D.L. Harame, P.J. Restle, J.M.C. Stork, A.C. Megdanis, C.L. Stanis, A.A. Bright, G.M.W. Kroesen and A.C. Warren, SiGe-channel heterojunction p-MOSFETs, IEEE Trans. Electron Devices, ED-41(1) (1994) 90.

[11]   S.S. Iyer, P.M. Solomon, V.P. Kesan, A.A. Bright, J.L. Freeouf, T.N. Nguyen and A.C. Warren, A gate-quality dielectric system for SiGe metal-oxide-semiconductor devices, IEEE Electron Device Lett., 12 (1991) 246.

[12]   J. Ruolian, L. Jianlin, Z. Youdou, L. Haifeng and Z. Houzh, Hole transport properties of Si/SiGe modulation doped heterostructures, Superlattices and Microstructures, 16(4) (1994) 375.

[13]   P.M. Garone, V. Venkataraman and J.C. Sturm, Mobility enhancement and quantum mechanical modeling in GeSi channel MOSFETs from 90 to 300 K, IEDM Tech. Dig. 91 (1991) 29.

[14]   D.K. Nayak, J.C.S. Woo, G.K. Yabiku, K.P. MacWilliams, J.S. Park and K.L. Wang, High mobility GeSi PMOS on SIMOX, IEEE Electron Device Lett., EDL-14 (1993) 520.

[15]   S.P. Voinigescu, P.B. Rabkin, C.A.T. Salama and P.A. Blakey, 2D numerical investigation of the impact of compositional grading on the performance of submicrometer Si-SiGe MOSFETs, IEEE Trans. Electron Devices, ED-42(6) (1995) 1039.

16]   G.F. Niu, G. Ruan and T.A. Tang, Solid-State Electron., 38 (1995) 323.