ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **11**, Number 1, 2008, 3–15

The New Generation of SOI MOSFETs

J.-P. COLINGE

Tyndall National Institute, Lee Maltings,
Prospect Row, Cork, Ireland
E-mail: `jean-pierre.colinge@tyndall.ie`

Abstract. The classical MOSFET is reaching its scaling limits and "end-of-roadmap" alternative devices are being investigated. Amongst the different types of SOI devices proposed, one clearly stands out: the multigate field-effect transistor (multigate FET). This device has a general "wire-like" shape. Multigate FETs are commonly referred to as "multi(ple)-gate transistors", "FinFETs", "tri(ple)-gate transistors", "GAA transistors", etc. This Paper describes the reasons for evolving from single-gate to multi-gate structures. It also describes some issues in ultra-small devices, such as doping fluctuation effects and quantum confinement effect.

1. Introduction

The adoption of Silicon-on-Insulator (SOI) substrates for the manufacturing of mainstream semiconductor products such as microprocessors has given SOI research an unprecedented impetus. In the past, novel transistor structures proposed by SOI scientists were often considered exotic and unpractical, but the recent success of SOI in the field of microprocessor manufacturing has finally given this technology the credibility and respect it deserves. The classical CMOS structure is reaching its scaling limits and "end-of-roadmap" alternative devices are being investigated. Amongst the different types of SOI devices proposed to remedy the problem, one clearly stands out: the multigate field-effect transistor (multigate FET). This device has a general "wire-like" shape with a gate electrode that controls the flow of current between source and drain. Multigate FETs are commonly referred to as "multi(ple)-gate transistors", "FinFETs", "tri(ple)-gate transistors", "Gate-all-Around transistors", etc. The International Technology Roadmap for Semiconductors (ITRS) recognizes the importance of these devices and calls them "advanced non-classical CMOS devices".

2. Historical perspective

The first article on the double-gate MOS (DGMOS) transistor was published by T. Sekigawa and Y. Hayashi 1984 [1]. That paper shows that one can obtain significant reduction of short-channel effects by sandwiching a fully depleted SOI device between two gate electrodes connected together. The device was called XMOS because its cross section looks like the Greek letter Ξ (*Xi*). Using this configuration, a better control of the channel depletion region is obtained than in a "regular" SOI MOSFET, and, in particular, the influence of the drain electric field on the channel is reduced, which reduces short-channel. In 1987, Balestra *et al.* [2] discovered another important property of double-gate MOSFETs: Volume inversion. Volume inversion is a phenomenon that appears in very thin (or narrow) film multigate SOI MOSFETs due to the fact that inversion carriers are not confined near the Si/SiO_2 interface, as predicted by classical device physics, but rather at the centre of the film. Volume inversion was first observed in Gate-all-Around (GAA) MOSFETs in 1990 [3]. The original GAA MOSFET had a polysilicon gate electrode wrapped around the entire channel region. Because the width of the device was much larger than the silicon film thickness, the original GAA device was really a double-gate device, and the contribution of the sidewall gates to electrostatic control of the channel was negligible (Figs. 1 and 2). Other double-gate MOSFETs implementations include the DELTA FET [4], the FinFET [5], the Silicon-on-Nothing) MOSFET [6], the Multi-Fin XMOS (MFXMOS) [7], the triangular-wire SOI MOSFET [8] and the Δ-channel SOI MOSFET [9].

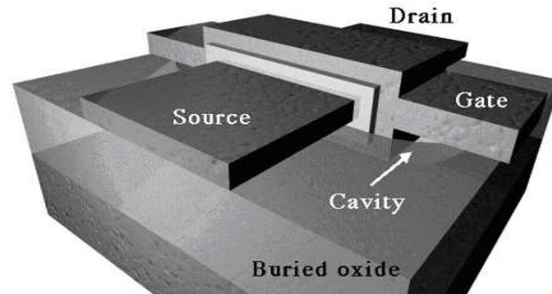

Fig. 1. Bird's eye view of the original GAA device.

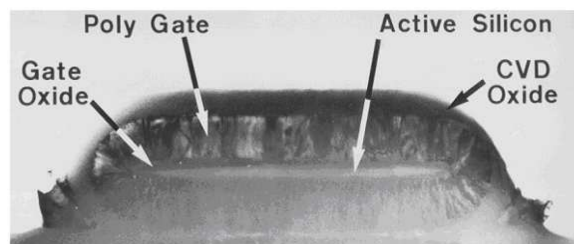

Fig. 2. TEM cross section of the original GAA device.

3. Electrostatic control

Short-channel effects arise when control of the channel region by the gate is affected by electric field lines from source and drain.

In a bulk device, the electric field lines propagate through the depletion regions associated with the junctions. Their influence on the channel can be reduced by increasing the doping concentration in the channel region. In very small devices, the doping concentration becomes too high (10^{19} cm^{-3}) for proper device operation, unfortunately. In a fully depleted SOI (FDSOI) device, most of the field lines propagate through the buried oxide (BOX) before reaching the channel region. Short-channel effects can be reduced in FDSOI MOSFETs by using a thin buried oxide and an underlying ground plane. This approach, however, has the inconvenience of increased junction capacitance and body effect. A much more efficient device configuration is obtained by using the double-gate transistor structure. The electric field lines from source and drain underneath the device terminate on the bottom gate electrode and cannot, therefore, reach the channel region. The two main manifestations of the short-channel effect are the Drain-Induced Barrier Lowering (DIBL) and the threshold voltage roll-off Short-Channel Effect (SCE). There is a parameter called the "Electrostatic Integrity" (EI) that can be related to both DIBL and SCE, and which describes the quality of the electrostatic control of the channel by the gate [10]:

$$DIBL = 0.80\frac{\varepsilon_{Si}}{\varepsilon_{ox}}EI\ V_{DS} \text{ and } SCE = 0.64\frac{\varepsilon_{Si}}{\varepsilon_{ox}}EI\ V_{bi},$$

where V_{bi} is the source or drain built-in potential. The values for EI in a bulk, FDSOI and double-gate device are shown in Table 1.

Table 1. Electrostatic integrity for different devices

Device	EI
A: Bulk	$EI = \left[1 + \frac{x_j^2}{L_{el}^2}\right]\frac{t_{ox}}{L_{el}}\frac{t_{dep}}{L_{el}}$
B: FDSOI	$EI = \left[1 + \frac{t_{Si}^2}{L_{el}^2}\right]\frac{t_{ox}}{L_{el}}\frac{t_{Si} + \lambda t_{BOX}}{L_{el}}$
C: Double Gate	$EI = \frac{1}{2}\left[1 + \frac{t_{Si}^2/4}{L_{el}^2}\right]\frac{t_{ox}}{L_{el}}\frac{t_{Si}/2}{L_{el}}$

Based on short-channel and DIBL considerations, the minimum gate length that can be used with the different technologies has been calculated. The result of these calculations is shown in Fig. 3 for three different types of CMOS circuits: high-performance (HP), low operating power (LOP), and low standby power (LSTP) digital circuits. An important conclusion can be derived from the data presented in Fig. 3: bulk transistors run out of steam once they reach a gate length of 15–20 nm. FDSOI can be used until 10 nm, but smaller gate lengths can be only achieved by the double-gate structure.

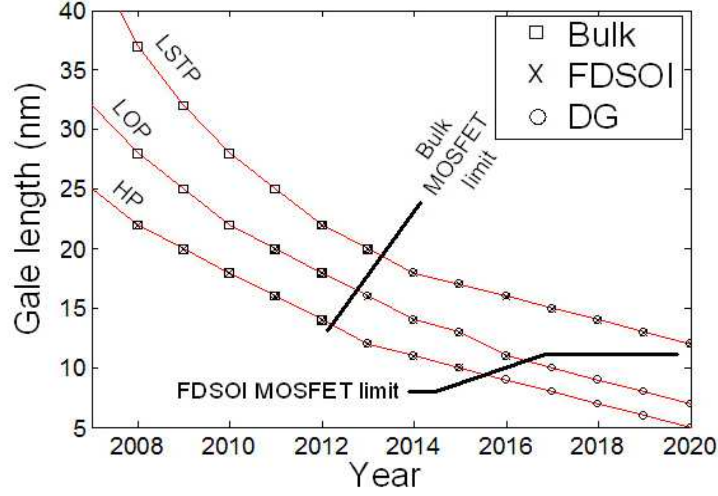

Fig. 3. Evolution of gate length predicted by the 2005 ITRS for high-performance (HP), low operating power (LOP), and low standby power (LSTP) digital circuits.

The concept of "*natural length*" can be derived from Poisson's equation. The natural length of a device basically represents the length of the region of the channel that is controlled by the drain. It is related to the Electrostatic Integrity. A device is free of short-channel effects if the effective gate length of a MOS device is larger than 5 to 10 times the natural length [11]. Table 2 shows the natural length for different gate configurations.

Table 2. Natural length, λs in devices with different geometries

Single gate	$\lambda_1 = \sqrt{\frac{\varepsilon_{si}}{\varepsilon_{ox}} t_{si}\, t_{ox}}$
Double gate	$\lambda_2 = \sqrt{\frac{\varepsilon_{si}}{2\varepsilon_{ox}} t_{si}\, t_{ox}}$
Quadruple gate (square section)	$\lambda_4 \cong \sqrt{\frac{\varepsilon_{si}}{4\varepsilon_{ox}} t_{si}\, t_{ox}}$
Surrounding gate (circular section)	$\lambda_o = \sqrt{\frac{2\varepsilon_{si}\, t_{si}^2\, \ln\left(1+\frac{2\, t_{ox}}{t_{si}}\right) + \varepsilon_{ox}\, t_{si}^2}{16\, \varepsilon_{ox}}}$

4. Multigate device design guidelines

The following observations can be made regarding the natural length (and regarding short-channel effects). The natural length can be reduced by decreasing the gate oxide thickness, the silicon film thickness and by using a high-k gate dielectric instead of SiO_2. In addition, the natural length is reduced when the number of gates is increased. In very small devices, the reduction of oxide thickness below 1.5 nm causes gate tunneling current problems. Using multi-gate devices, it is possible to trade a thin gate oxide for thin silicon film/fin thinning since λ is proportional to the product $t_{si} \times t_{ox}$.

The "*equivalent number of gates*" (*ENG*) is basically equal to the number of gates (a square cross section is assumed) but is also equal to the number that divides $\varepsilon/\varepsilon_{ox} \times t_{si} t_{ox}$ in the equations defining the natural length. Thus we have *ENG*=1 for a single-gate FDSOI MOSFET, *ENG*=2 for a double-gate device and *ENG*=4 for a quadruple-gate MOSFET. *ENG*=3 for a triple-gate device and, by some strange coincidence, *ENG* is close to π in a Π-gate device. In the Ω-gate device the value of *ENG* ranges between 3 and 4 depending on the extension of the gate under the fin [12].

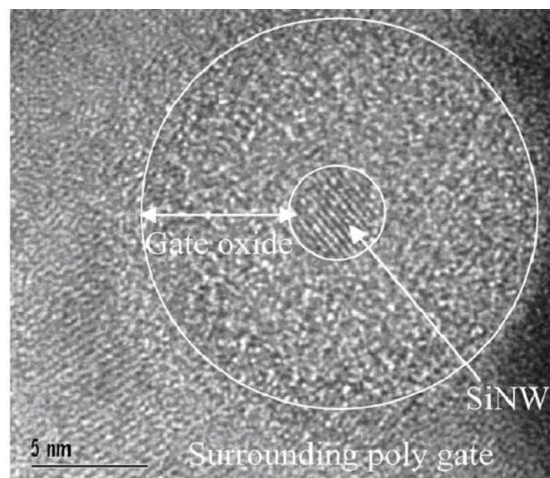

Fig. 4. TEM section of silicon GAA nanowire [19].

The electrostatic control of the channel can be improved if the gate electrode is wrapped around three or four sides of the device. The triple-gate MOSFET is a thin-film, narrow silicon island with a gate on three of its sides.[13] Implementations include the quantum-wire SOI MOSFET [14] and the tri-gate MOSFET [15]. The Electrostatic Integrity of triple-gate MOSFETs can be improved by extending the sidewall portions of the gate electrode to some depth in the buried oxide and underneath the channel region (Π-gate device [16] and Ω-gate device [17]). From an electrostatic point of view, the Π-gate and Ω-gate MOSFETs have an effective number of gates between three and four. The use of strained silicon, a metal gate and/or

high-k dielectric as gate insulator can further enhance the current drive of the device [18]. Ultimate electrostatic control performance is obtained using a cylindrical GAA nanowire structure (Fig. 4) [19].

5. Doping fluctuation effects

Random doping fluctuation effects are known to cause variations of the electrical parameters of small MOSFETs. A trigate device with a channel length, width and height of 10 nm and a doping concentration of 10^{18} cm^{-3} has a single doping atom in its channel, statistically. This means some devices will be undoped, while some will have one or two doping atoms. Even "undoped" devices will contain doping atoms arising from the original SOI film doping, from random contamination effects or from the S/D implantation process.

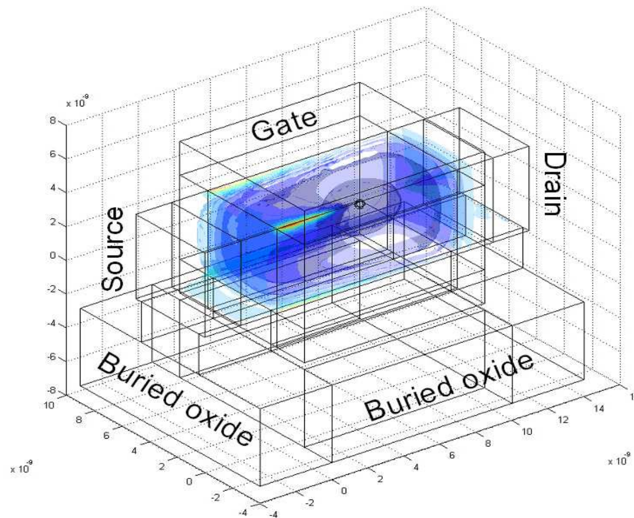

Fig. 5. Electron concentration in an n-channel trigate FET with one acceptor atom in the channel.

Figures 5 and 6 show the electron concentration in a trigate FET with $W_{si} = t_{si} = 5$ nm and L=10 nm if the channel contains a single acceptor or donor atom in the channel, respectively. In the simulations, the doping atom can be moved around and placed at any location (x, y, z) in the silicon. The coordinate system in the channel is such that$0 < x < L$, $0 < y < W_{si}$ and $0 < z < t_{si}$. The variation of threshold voltage is higher when the doping atom is located near the centre of the device and the highest variation is observed in the middle of the slice where $x = L/2$, which is the centre of the channel. Near the four corners, the impact of the doping atom is very limited. The presence of a P-type impurity in the transistor creates around the doping atom a region that is depleted of electrons. While volume inversion is observed in most of the device, the electron concentration drops in the vicinity of the

doping atom (Fig. 5). Thus a higher gate voltage is required to reach the same current level as in an undoped device, which increases the threshold voltage. The diameter of the region where the electron concentration is reduced is around 2.5 nm at threshold. Thus a doping atom in the centre of a slice ($z = y = 2.5$ nm) will influence the current flow on each side de of the device and in each corner. On the other hand, an impurity atom located in one corner will influence conduction in that corner, but not in the 3 other corners. This explains why atoms located at the centre of the device have the largest influence on the threshold voltage.

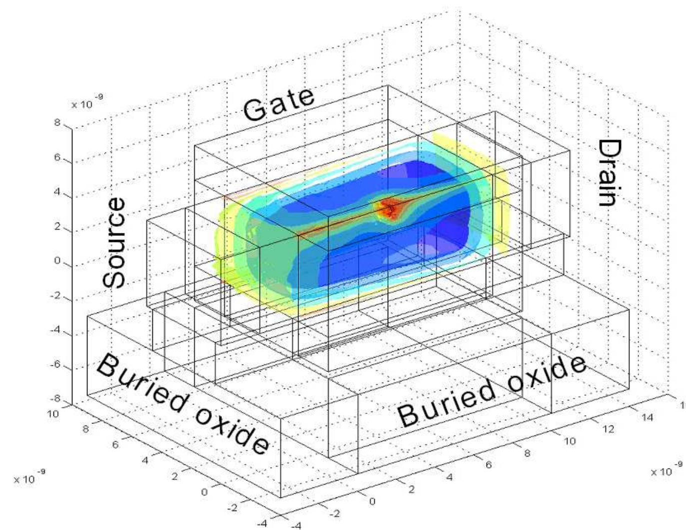

Fig. 6. Electron concentration in an n-channel trigate FET with one donor atom in the channel.

If the doping atom is a donor impurity or N-type doping atom, the transistor becomes an N^+NN^+ transistor. The FET is now an accumulation-mode device. In such a device, the entire channel region is n-type near threshold. The addition of an n-type impurity atom to the channel region increases the electron concentration locally, but its influence does not “radiate out” like in the case of the p-type impurity atom (Fig. 6). A donor atom slightly increases the electron concentration and, thus, reduces the threshold voltage, but the magnitude of the threshold shift is significantly less than when an acceptor atom is introduced, especially in the device with a smaller cross section. As expected, this atom renders V_{TH} more negative.

Figure 7 shows the variation of threshold voltage in the entire 5 nm $\times$5 nm$\times$20 nm trigate device, averaged for each slice. For both N-type and P-type doping impurities, the greatest shifts of threshold voltage occurs when the doping atom is placed in the centre of the channel. The variation tends to be zero when the impurity moves close to the source or drain. The variation is greater for an inversion-mode (IM) than an accumulation-mode (AM) device.

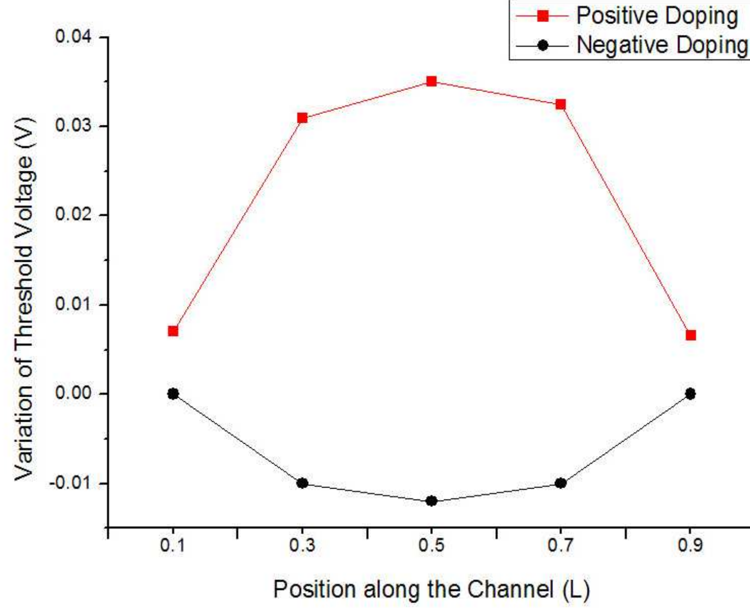

Fig. 7. The average variations of threshold voltage in the entire 5 nm ×5 nm×20 nm trigate device.

6. Quantum confinement effects

The thickness and/or width of multi-gate FETs is reaching values that are less than 10 nanometers. Under these conditions the electrons in the "channel" (if we take the example of an n-channel device) form either a two-Dimensional Electron Gas (2DEG) if we consider a double-gate device or a one-Dimensional Electron Gas (1DEG) if we consider a triple or quadruple-gate MOSFET. This confinement is at the origin of the "volume inversion" effect and yields an increase of threshold voltage when the width/thickness of the devices is reduced [20, 21].

The calculation of the electron concentration in the presence of confinement effect requires to solve the Poisson and Schrödinger equation self-consistently [22]. An example of electron concentration calculated in the cross section of a FinFET, a trigate MOSFET and a gate-all-around (GAA) transistor for with a gate voltage higher than the threshold voltage ($V_G > V_{TH}$) is presented in Fig. 8. The fin width of the devices is 5 nm. In each device one can observe the relatively large electron concentration in the centre of the silicon film or fin, corresponding to volume inversion. Another interesting feature of nanowire devices is the shape of the Density of States (DoS) distribution. It is composed of a succession of spikes. An infinite DoS is found at the bottom of each energy subband, and the then the DoS decreases as a function of $1/\sqrt{E}$in each subband. Figure 9 shows the product of the DoS by the

Fermi-Dirac function at room temperature for trigate devices with a cross section of 5 nm×5 nm and 10 nm×10 nm. The energy separation between the spikes increases as the dimensions of the device are reduced.

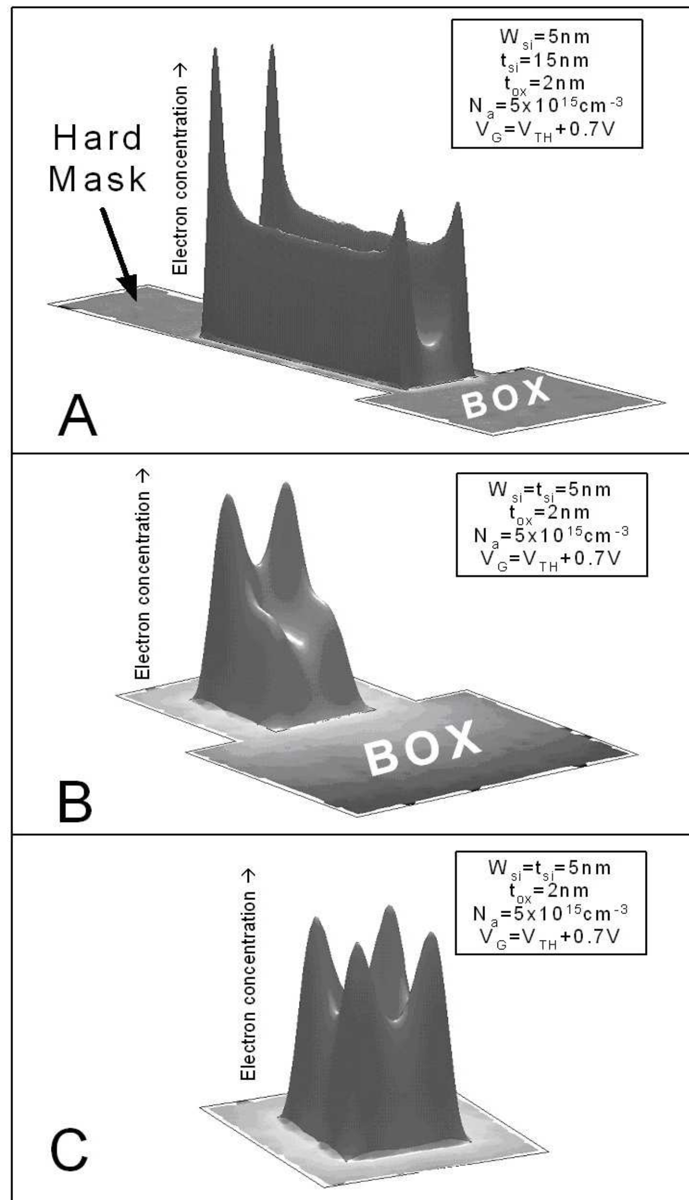

Fig. 8. Electron concentration profile in A: a FinFET; B: a trigate FET; C: a GAA FET

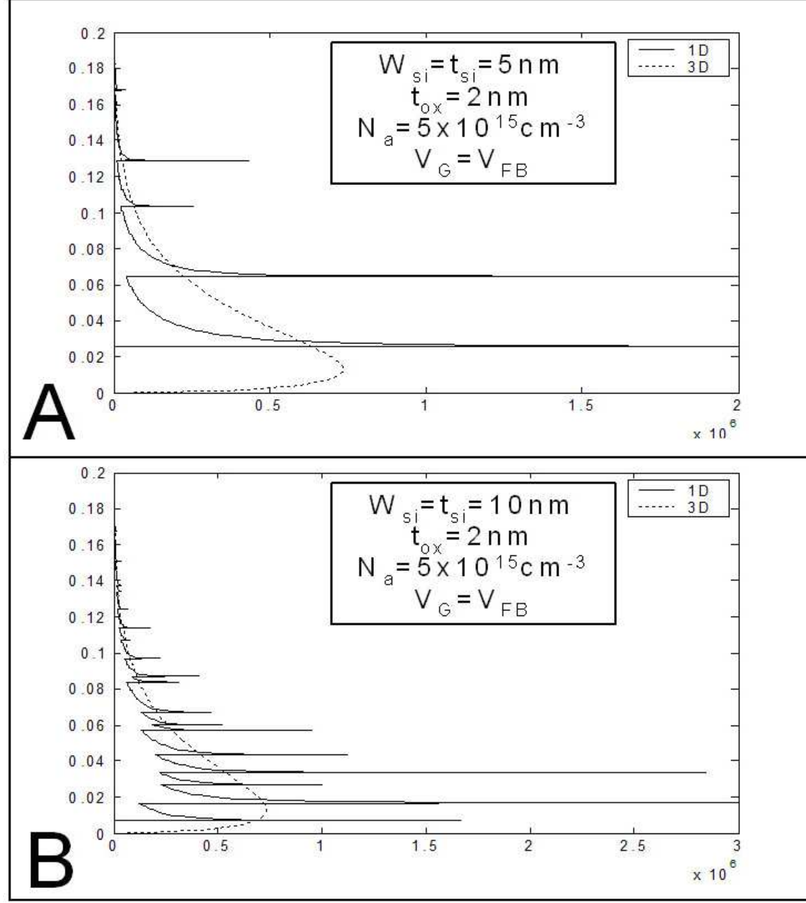

Fig. 9. Product of Density of States (DoS) by the Fermi-Dirac function in a trigate device with A: a 5 nm section and B: a 10 nm section. The equivalent distribution for a 3D system is show for comparison (dotted line).

Inter-subband scattering occurs between electrons that belong to different energy subbands. These scattering events reduce the electron mobility. By definition, there is no intersubband scattering if only one subband is occupied, which occurs right above threshold. As the gate voltage and the electron concentration are increased, however, a larger number of subbands become populated, and scattering occurs between electrons belonging to different subbands. If the temperature is not too high (compared to $\Delta E/k$, where ΔE is the energy separation between two subbands, and k is Boltzmann's constant) and if the drain voltage not much larger than $\Delta E/q$, inter-subband scattering phenomena can be directly observed in the form of oscillations of drain current amplitude when the gate voltage is increased. This effect can be seen in Fig. 10, in which each "dip" of the curve corresponds to a reduction of mobility caused by scattering due to the population of each new subband.

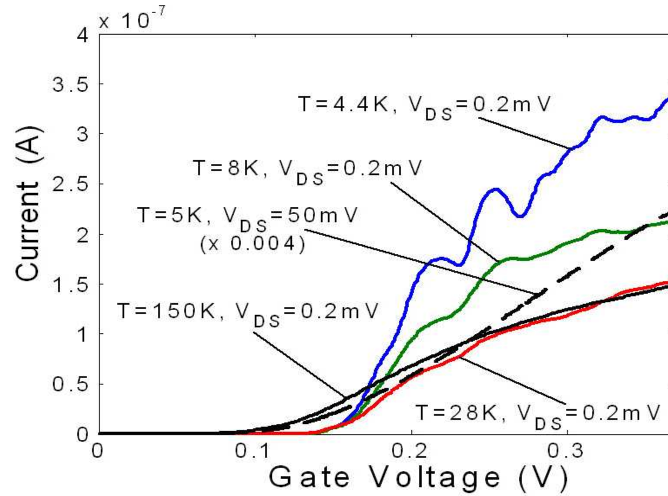

Fig. 10. Drain current vs. gate voltage for different temperature and drain voltage values. The amplitude of the curve for $V_{DS} = 50$ mV is multiplied by a factor 200 µV/50 mV = 0.004 to fit in the same graph as the curves measured at $V_{DS} = 200$µV.

This effect is more pronounced at low temperature, but can be observed at room temperature provided the cross-section of the devices is small enough. Table 3 shows the maximum temperature and maximum drain voltage at which current oscillations due to inter-subband scattering have been experimentally observed in MuGFETs.

Table 3. Subband energy separation, ΔE, maximum temperature and drain voltage at which current oscillations due to intersubband scattering have been observed

$W_{si} \times t_{si}$	ΔE	Temperature	Drain voltage	Reference
45 nm×82 nm	0.15 meV	28 K	0.2 mV	23
11 nm×58 nm	1–2 meV	300 K	1 mV	24
6 nm×6 nm	35 meV	5 K	100 mV	25
6 nm×6 nm	35 meV	200 K	50 mV	*ibidem*

7. Conclusions

This paper describes the reasons for evolving from single-gate to multi-gate structures. It also describe some issues in ultra-small devices, such as doping fluctuation effects and quantum confinement effect.

Acknowledgements. This material is based upon works supported by Science Foundation Ireland under Grant 05/IN/I888.

References

[1] SEKIGAWA T., HAYASHI Y., *Calculated threshold-voltage characteristics of an XMOS transistor having an additional bottom gate*, Solid-State Electronics, vol. **27**, p. 827, 1984.

[2] BALESTRA F., CRISTOLOVEANU S., BENACHIR M., BRINI J., ELEWA T., *Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance*, IEEE Electron Device Letters, vol. **8**, no. 9, p. 410, 1987.

[3] COLINGE J.P., GAO M.H., ROMANO A., MAES H., CLAEYS C., *Silicon-on-insulator 'gate-all-around' MOS device*, Technical Digest of IEDM, p. 595, 1990.

[4] HISAMOTO D., KAGA T., KAWAMOTO Y., TAKEDA E., *A fully depleted lean-channel transistor (DELTA)-a novel vertical ultra thin SOI MOSFET*, Technical Digest of IEDM, p. 833, 1989.

[5] HISAMOTO E., LEE W.C., KEDZIERSKI J., ANDERSON E., TAKEUCHI H., ASANO K., KING T.J., BOKOR J., HU C., *A folded-channel MOSFET for deep-sub-tenth micron era*, Technical Digest of IEDM, p. 1032, 1998.

[6] PRETET J., MONFRAY S., CRISTOLOVEANU S., SKOTNICKI T., *Silicon-on-nothing MOSFETs: performance, short-channel effects, and backgate coupling*, IEEE Transactions on Electron Devices, vol. **51**, p. 240, 2002.

[7] LIU Y., ISHII K., TSUTSUMI T., MASAHARA M., SUZUKI E., *Ideal rectangular cross-section Si-Fin channel double-gate MOSFETs fabricated using orientation-dependent wet etching*, IEEE Electron Device Letters, vol. **24**, no. 7, p. 484, 2003.

[8] HIRAMOTO T., *Nano-scale silicon MOSFET: towards non-traditional and quantum devices*, *IEEE International SOI Conference Proceedings*, p. 8, 2001.

[9] JIAO Z., SALAMA A.T., *A Fully Depleted Delta Channel SOI NMOSFET*, *Electrochem. Society Proceedings*, vol. **2001-3**, p. 403, 2001.

[10] SKOTNICKI T., *Heading for decananometer CMOS - is navigation among icebergs still a viable strategy?*, *Proceedings of ESSDERC*, p. 19, 2000.

[11] YAN R.H., OURMAZD A., LEE K.F., *Scaling the Si MOSFET: from bulk to SOI to bulk*, IEEE Trans. on Electron Devices, vol. **39**, no. 7, p. 1704, 1992.

[12] LEE C.-W., YUN S., YU C.-G., PARK J.-T., COLINGE J.P., *Device design guidelines for nano-scale MuGFETs*, Solid-State Electronics, vol. **51**, no. 3, p. 505, 2007.

[13] LEMME M.C., MOLLENHAUER T., HENSCHEL W., WAHLBRINK T., BAUS M., WINKLER O., GRANZNER R., SCHWIERZ F., SPANGENBERG B., KURZ H., *Subthreshold behavior of triple-gate MOSFETs on SOI Material*, Solid State Electronics, vol. **48**, p. 529, 2004.

[14] COLINGE J.P., BAIE X., BAYOT V., GRIVEI E., *A silicon-on-insulator quantum wire*, Solid-State Electronics, vol. **39**, p. 49, 1996.

[15] DOYLE B.S., DATTA S., DOCZY M., JIN B., KAVALIEROS J., LINTON T., MURTHY A., RIOS R., CHAU R., *High performance fully-depleted tri-gate CMOS transistors*, IEEE EDL, vol. **24**, no. 4, p. 263, 2003.

[16] PARK J.T., COLINGE J.P., *Multiple-gate SOI MOSFETs: device design guidelines*, IEEE Trans. on Elect. Dev., vol. **49**, no. 12, p. 2222, 2002.

[17] YANG F.L., CHEN H.Y., CHENG F.C., HUANG C.C., CHANG C.Y., CHIU H.K., LEE C.C., CHEN C.C., HUANG H.T., CHEN C.J. *et al.*, *25 nm CMOS Omega FETs*, Technical Digest of IEDM, p. 255, 2002.

[18] ANDRIEU F., DUPRÉ C., ROCHETTE F., FAYNOT O., TOSTI L., BUJ C., ROUCHOUZE E., CASSÉ M., GHYSELEN B., CAYREFOURCQ I. *et al.*, *25 nm Short and Narrow Strained FDSOI with* TiN/HfO_2 *Gate Stack*, *Symposium on VLSI Technology*, paper 16.4, 2006.

[19] SINGH N., AGARWAL A., BERA L.K., LIOW T.Y., YANG R., RUSTAGI S.C., TUNG C.H., KUMAR R., LO G.Q., BALASUBRAMANIAN N., KWONG D.L., *High-performance fully depleted silicon nanowire (diameter < 5 nm) gate-all-around CMOS devices*, IEEE Electron Device Letters vol. **27**, no. 5, p. 383, 2006.

[20] OMURA Y., HORIGUCHI S., TABE M., KISHI K., *Quantum-mechanical effects on the threshold voltage of ultrathin SOI nMOSFETs*, IEEE Electron Device Letters, vol. **14**, no. 12, p. 569, 1993.

[21] POIROUX T., VINET M., FAYNOT O., WIDIEZ J., LOLIVIER J., ERNST T., PREVITALI B., DELEONIBUS S., *Multiple gate devices: advantages and challenges*, Microelectronic Engineering, vol. **80**, p. 378, 2005.

[22] COLINGE J.P., ALDERMAN J.C., XIONG W., CLEAVELIN C.R., *Quantum-Mechanical Effects in Trigate SOI MOSFETs*, IEEE TED, vol. **53**, no. 5, p. 1131, 2006.

[23] COLINGE J.P., QUINN A.J., FLOYD L., REDMOND G., ALDERMAN J.C., XIONG W., CLEAVELIN C.R., SCHULZ T., SCHRUEFER K., KNOBLINGER G., PATRUNO P., *Low-Temperature Electron Mobility in Trigate SOI MOSFETs*, IEEE EDL, vol. **27**, no. 2, p. 120, 2006.

[24] COLINGE J.P., XIONG W., CLEAVELIN C.R., SCHULZ T., SCHRÜFER K., MATTHEWS K., PATRUNO P., *Room-Temperature Low-Dimensional Effects in Pi-Gate SOI MOSFETs*, IEEE EDL, vol. **27**, no. 9, p. 775, 2006.

[25] SINGH N., LIM F.Y., FANG W.W., RUSTAGI S.C., BERA L.K., AGARWAL A., TUNG C.H., HOE K.M., OMAMPULIYUR S.R., TRIPATHI D., ADEYEYE A.O., LO G.Q., BALASUBRAMANIAN N., KWONG D.L., *Ultra-narrow silicon nanowire gate-all-around CMOS device: impact of diameter, channel-orientation and low temperature on device performance*, Technical Digest of IEDM, paper 20.4, 2006.