Università di Pisa

Three-dimensional simulation of the dependence of the programming window of SOI nanocrystal memories on the channel width

Gianluca Fiori

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. Molas

CEA-LETI, Grenoble

Barbara De Salvo

CEA-LETI, Grenoble

Three-Dimensional Simulation of the Dependence of the Programming Window of SOI Nanocrystal Memories on the Channel Width

Gianluca Fiori, Giuseppe Iannaccone, Gabriel Molas, and Barbara De Salvo

Abstract—In this paper, we present an approach based on three-dimensional simulations for the investigation of the dependence of the programming window of silicon-on-insulator (SOI) nanocrystal memories on the width of the silicon channel. Recent experiments show that the threshold voltage shift after programming increases with decreasing channel width. By evaluating the consequences of possible assumptions on the charge stored in the nanocrystals, we show that such behavior can be consistently explained by the preferential injection of electrons, during the program operation, through the oxide near the edges of the channel. As a consequence, charge is mostly stored in the dots close to the edges and, therefore, is more and more effective as the channel width is decreased. Experiments and simulations on SOI nanocrystal memories show support of our interpretation with respect to different proposed mechanisms.

Index Terms—Nonvolatile memories (NVMs), programming window, silicon-on-insulator (SOI) technology, silicon nanocrystals.

I. INTRODUCTION

NANOCRYSTAL memories [1] are promising candidates for future nonvolatile memories (NVMs) for two main reasons: on the one hand, their fabrication requires only few additional process steps with respect to conventional CMOS technology; on the other hand, they offer the perspective of better scalability than conventional double-poly Flash EEPROMs that would translate into reduced program/erase times, lower power consumption, and higher integration density [2]–[6].

Recently, for silicon-on-insulator (SOI) nanocrystal floating gate memories with ultra-narrow channels [7], [8], an increasing threshold voltage shift has been observed with decreasing channel width, at a constant program time and write voltage. Such behavior might suggest that wider programming window, which is defined as the difference between the threshold voltages measured when dots are charged (after programming) and when dots are discharged (after erasing), is

achievable with the use of ultra-narrow channels and, therefore, deserves detailed investigation.

A possible mechanism responsible for such an effect has been recently proposed that is based on the role played by percolating paths in the channel due to the random distribution of charged dots above the silicon wire [7], [8]. However, we will show experiments in which the behavior under investigation is also observed when the threshold voltage shift is measured in strong inversion in a situation in which percolating paths cannot possibly be present.

In this paper, we present a simulation approach that allows us to propose an alternative physical explanation based on experiments and on detailed simulations of very narrow SOI nanocrystal memories. In particular, with the use of a self consistent Poisson/Schrödinger solver in three-dimensional (3-D) domains, we show that the above-mentioned effect can be consistently explained by the preferential injection of electrons near the edges of the channel, which makes the dots in the oxide region surrounding the edges much more charged than those over the center of the channel.

II. DEVICE FABRICATION AND SIMULATED STRUCTURE

In Fig. 1, we show the longitudinal and transversal sections of the fabricated devices (see Fig. 1(a) and (b), respectively). The 15-nm-thick (100)-oriented channel with varying width W and length L of 50 nm is defined by electron beam lithography (EBL) and reactive ion etching (RIE). The SiO_2 tunneling layer has a thickness of 2.5 nm and is grown by rapid thermal oxidation.

The layer of silicon dots, with a density of approximately 5×10^{11} cm^{-2} and average dot diameter of 6 nm, is obtained by the phase separation of a silicon rich oxide $(SiO_{0.5})$ layer through an annealing step at 1050 °C for approximately 3 min under N_2 [9]. The dot layer is then covered with a deposited 20-nm control oxide and, finally, polysilicon gates are defined by EBL and RIE.

In Fig. 2, we show the threshold voltage shifts as a function of the channel width, obtained in the strong inversion and in the subthreshold regime, for a writing voltage of 10 V and a programming time of 10 s.

As can be seen, the effect is also present in strong inversion, when the Debye length [10] can be estimated to approximately 3 nm, for an electron concentration $n = 2 \times 10^{18}$ cm^{-3}; percolating paths in this condition cannot be a relevant cause of the

G. Fiori and G. Iannaccone are with the Dipartimento di Ingegneria dell'Informazione, Universitá di Pisa, 56122 Pisa, Italy (e-mail: g.fiori@iet.unipi.it).

G. Molas and B. De Salvo are with the Commissariat à l'energie atomique–Laboratory of Electronics, Technology, and Instrumentation, 38054 Grenoble, France.
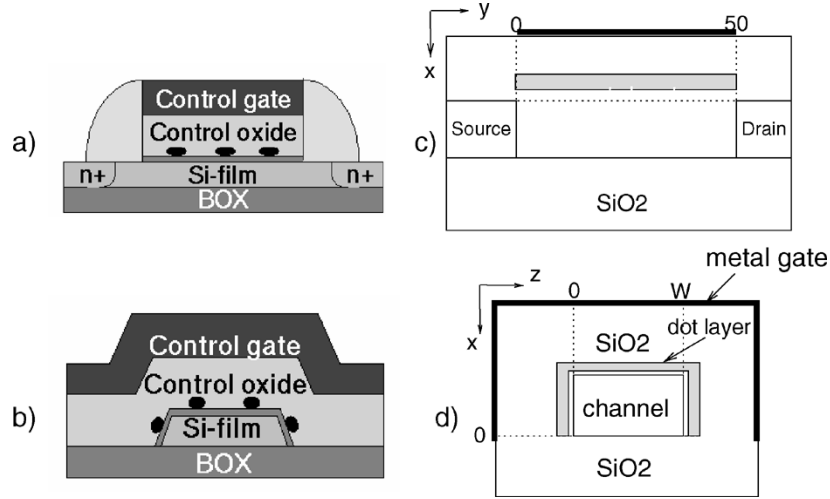
Fig. 1. (a) and (b) Sketch of the fabricated nanomemory (longitudinal and transversal views, respectively). (c) and (d) Simulated 3-D simplified structure (longitudinal and transversal sections, respectively).
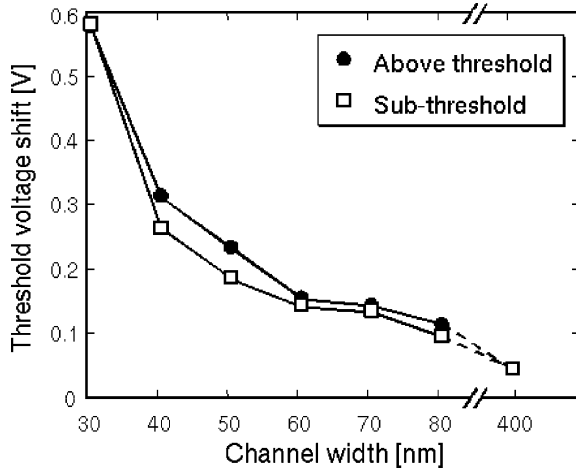

Fig. 2. Measured threshold voltage shift as function of the channel width for a device, in the strong inversion and subthreshold regime, as obtained with a writing voltage of 10 V and a programming time of 10 s.

observed effect since, for such electron density, charged dots are not able to deplete the channel underneath.

In order to investigate the reduction of the threshold voltage shift for increasing channel width, we have performed detailed simulations of six different devices with the same L (50 nm) and W varying from 30 to 80 nm by means of an in-house developed 3-D Poisson/Schrödinger solver.

III. APPROACH

A. Physical Model

The Poisson equation reads

$$\nabla\left[\epsilon(\vec{r})\nabla\phi(\vec{r})\right] = -q\left[p(\vec{r}) - n(\vec{r}) + N_D^+(\vec{r}) - N_A^-(\vec{r}) + \rho_{\text{fix}}\right] \quad (1)$$

where ϕ is the electrostatic potential, q is the electron charge, ϵ is the dielectric constant, p and n are the hole and electron densities, respectively, N_D^+ is the concentration of ionized donors, N_A^- is the concentration of ionized acceptors, and ρ_{fix} is the fixed charge density. Hole, acceptor, and donor densities are

computed in the whole domain with the semiclassical approximation, while the electron concentration in strongly confined regions needs to be computed by solving the Schrödinger equation with density functional theory in the local density approximation [11].

Preliminary two-dimensional (2-D) investigation stated that 2-D confinement is not relevant so we have solved the one-dimensional (1-D) Schrödinger equation in the direction perpendicular to the channel since the solution of the 2-D Schrödinger equation would have been too computational demanding because of the high number of required transversal modes without adding significant information.

The Poisson/Schrödinger equation is solved self-consistently with the Newton–Raphson scheme with a predictor/corrector algorithm similar to that proposed in [12]. In particular, to simplify the Jacobian and to achieve faster convergence, the Schrödinger equation is solved at the beginning of the Newton–Raphson iterative cycle: the eigenfunctions are considered fixed within the cycle, while eigenvalues are varied by a quantity equal to $q(\phi - \tilde{\phi})$, where $\tilde{\phi}$ is the potential used to solve the Schrödinger equation and ϕ is the potential at the current iteration. The electron density then becomes

$$n(x) = \frac{2k_B T m_t}{\pi \hbar^2}$$
$$\times \sum_i |\psi_{li}|^2 \ln\left[1 + \exp\left(\frac{E_F - E_{li} + q(\tilde{\phi} - \phi)}{k_B T}\right)\right]$$
$$+ \frac{4k_B T \sqrt{m_l m_t}}{\pi \hbar^2}$$
$$\times \sum_i |\psi_{ti}|^2 \ln\left[1 + \exp\left(\frac{E_F - E_{ti} + q(\tilde{\phi} - \phi)}{k_B T}\right)\right] \quad (2)$$

where anisotropy of the mass has been taken into account and $\psi_{li}, E_{li}, \psi_{ti}$, and E_{ti} are the eigenfunctions and eigenvalues obtained from the 1-D Schrödinger equation using the longitudinal effective mass m_l and the transverse effective mass m_t, respectively. The algorithm is then repeated cyclically until the norm of $(\phi - \tilde{\phi})$ is smaller than a predetermined value.

B. Threshold Voltage Computation

The computation of the programming window first requires the definition of the threshold voltage. In particular, in order to make comparisons with experiments, we have defined two threshold voltages: one in the subthreshold and one in the above threshold regime. From the transfer characteristics, we have defined the threshold voltage in the subthreshold region as the gate voltage (V_{GS}) at which the drain-to-source (I_{DS}) current is equal to 1 nA, while in the above threshold region, as the intercept with the V_{GS} axis of the line that best fits the transfer characteristics.

We compute I_{DS} by assuming that for very small V_{DS}, the electron density is equal to the value computed for $V_{DS} = 0$ [13]. The current density can be expressed as

$$\vec{J}_n = -q n \mu_n \nabla \phi_n \tag{3}$$

where μ_n is the electron mobility and ϕ_n is the quasi-Fermi level for the electrons. If we now suppose to be very close to the equilibrium condition, (3) can be written as

$$\vec{J}_n = -q n_0 \mu_n \nabla \phi_n \tag{4}$$

where n_0 is the charge density computed for drain-to-source voltage equal to zero. Considering a constant electron mobility and negligible generation-recombination processes, the continuity equation for electrons reads

$$\nabla \cdot J_n = 0 \tag{5}$$

and substituting (4) in (5), we obtain

$$\nabla \cdot (n_0 \nabla \phi) = 0. \tag{6}$$

Null Neumann boundary conditions are considered in the lateral surfaces delimiting the channel, while Dirichlet boundary conditions are assumed in correspondence of the source and drain contacts.

IV. SIMULATION

Since we are interested in the programming window, the key point of our approach is the modeling of the charged dots. We are not interested now in the issue of confinement in the silicon dots and of single electron charging of the dots since we want to focus mainly on the electrostatics of the device. We have then decided to model the dot layer with a layer of negative fixed charge in the silicon oxide. Since we do not know *a priori* how the charge is distributed in the dot layer, we have made different hypotheses that will be critically assessed in Section V.

As a first attempt, we can assume that dots are uniformly charged: once the dot density and the number of electrons per dot are fixed, the simplest choice is to model the dot layer as a region with fixed uniform charge density. The discrete nature of dot and charge distribution is an aspect that might have a role, and that can be considered by assuming that charged dots obey a Poisson distribution in space. In particular, for each grid point, we have considered the associated volume element and multiplied its volume ΔV by the nominal fixed charge concentration. A random number N' has then been extracted with Poisson distribution and divided by ΔV in order to have the "actual" charge concentration in the volume element. In this way, it has been
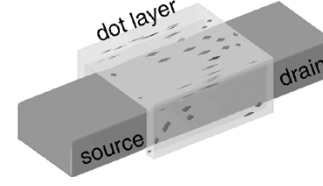

Fig. 3. Isosurface of the charge density ($n = 2 \times 10^{18}$ cm^{-3}) computed for a $V_{GS} = 1.6$ V and of the discrete fixed charge density in the storage layer (average charge in the dot layer 2×10^{18} cm^{-3}).
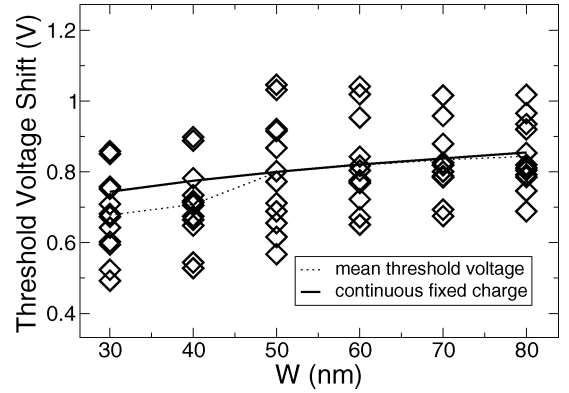

Fig. 4. Simulated threshold voltage shift as function of the channel width for a continuous distribution of fixed charge in the dot layer (solid line) and for 12 different discrete distribution of charged dots in the dot layer (symbols) computed in the subthreshold regime.
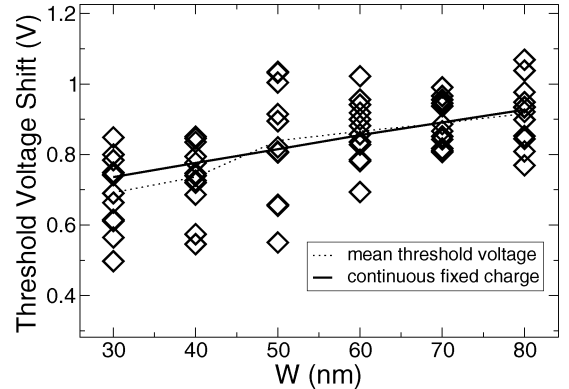

Fig. 5. Simulated threshold voltage shift as function of the channel width for a continuous distribution of fixed charge in the dot layer (solid line) and for 12 different discrete distribution of charged dots in the dot layer (symbols) computed in the strong inversion regime.

possible to take into account the threshold voltage dispersion due to the random distribution of charged dots.

In Fig. 3, we show the isosurface of the charge density ($n = 2 \times 10^{18}$ cm^{-3}) computed for $V_{GS} = 1.6$ V: the spots around the channel are the isosurfaces of the discrete fixed charge density ($\rho_{fix} = 2 \times 10^{18}$ cm^{-3}) and represent each individual charged dot. The shape of each spot is instead not particularly meaningful, as it depends both on the number of electrons per dot and on the grid.

In Figs. 4 and 5, we show results in the subthreshold and the above threshold regime, respectively, for a continuous distribution of fixed charge in the dot layer, corresponding to two electrons per dot and to an experimental dot density of 5 $\times$
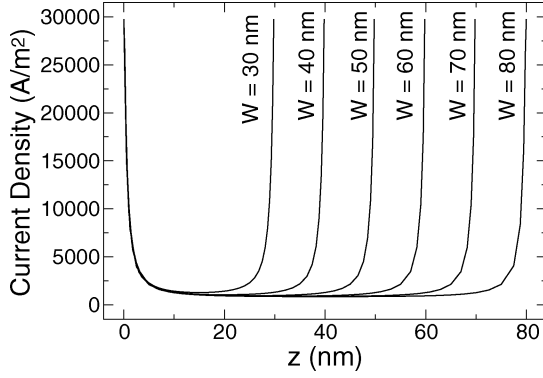
Fig. 6. Local direct tunneling current density in the correspondence of the dot layer as a function of the z-coordinate.
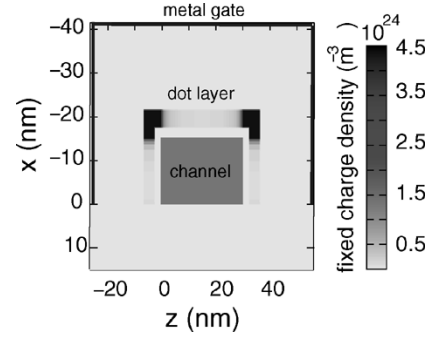

Fig. 7. Transversal cross section of the nanocrystal SOI memory with charge density in the dot layer proportional to the direct tunneling current.
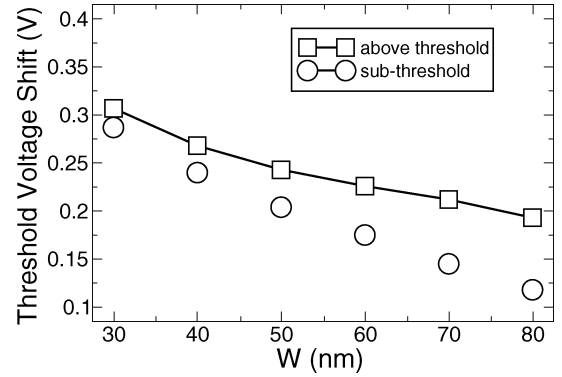

Fig. 8. Threshold voltage shift as a function of the channel width for a charge density in the dot layer proportional to the local tunnel current computed in the subthreshold and in the strong inversion regime.

10^{11} cm^{-2}. The threshold voltage shift is plotted as a function of W. As can be noticed, simulation results are in evident contrast with experiments since the computed shift slightly increases with increasing W. In Figs. 4 and 5, we also show results for a discrete distribution of fixed charge in the dot layer. We have computed the threshold voltage shift over a sample of 12 devices (as in the experiments) with the same nominal dot density, but with a different discrete distribution of charged dots. Even in this case, the threshold voltage shift is almost independent of channel width.

We have, therefore, removed the hypothesis of uniformly charged dots, and have considered the possibility of nonhomogeneous charging of the dots.

Indeed, in the devices considered, for the applied program voltage (10 V), electrons traverse the oxide via direct tunneling. An analytical formula for the direct tunneling current reads [14]

$$J_{DT}(\psi_b, E_{\text{ox}}) = \frac{C_1}{\left[1 - \left(\dfrac{\psi_b + qE_{\text{ox}}t_{\text{ox}}}{\psi_s}\right)^{0.5}\right]^2}$$
$$\cdot \exp\left[-\frac{C_2}{E_{\text{ox}}}\left[1 - \left(1 - \frac{qt_{\text{ox}}}{\psi_b}E_{\text{ox}}\right)^{1.5}\right]\right] \quad (7)$$

with

$$C_1 = \frac{q^3}{16\pi^2\hbar\psi_s} \quad (8)$$

and

$$C_2 = \frac{4}{3}\frac{(2m_{\text{ox}})^{0.5}}{q\hbar}\psi_s^{1.5} \quad (9)$$

where ψ_b is the difference between the top of the conduction band and the Fermi level in the bulk, $\hbar$ is the reduced Plank's constant, t_{ox} is the oxide thickness, m_{ox} is the electron effective mass in the oxide, and E_{ox} is the electric field in the oxide.

By extracting for each point at the Si/SiO$_2$ interface both ψ_b and the local E_{ox} in the tunnel oxide, we have been able to compute the local tunneling current. We then make the assumption that the charge locally stored in the dot layer is proportional to the local direct tunneling current, as in (7).

Indeed, the electrostatics of the structure is such that close to the channel edges the electric field in the oxide is higher and the
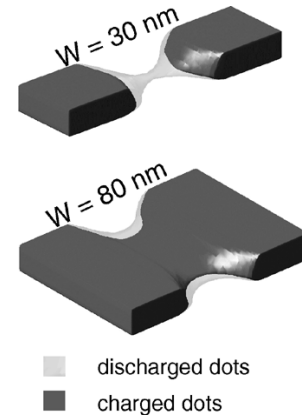
Fig. 9. Isosurface of the electron density ($n = 7.5 \times 10^{17}$ cm^{-3}) computed for a $V_{\text{GS}} = -0.4$ V in case of fixed charge proportional to the tunneling current for devices with channel width equal to 30 and 80 nm.

tunnel barrier lower and, therefore, the direct tunneling current [14] through the oxide is larger at the edges than at the center of wire (Fig. 6) and, since storage nodes are discrete, charge is preferentially stored in dots close to the edges of the channel, as shown in Fig. 7. The narrower the channel, the more effective the charge stored near the edges on device behavior. In the limit of an infinitely large channel, electrons stored in dots near the edges have no effect, and the threshold voltage shift is only due to the smaller amount of charge injected in the middle of the channel.

Results based on the hypothesis of locally stored charge proportional to the local tunnel current are shown in Fig. 8. As can

be noticed, simulations are now able to reproduce the experimental behavior.

In Fig. 9, we show the charge density computed for a $V_{GS} = -0.4$ V and $n = 7.5 \times 10^{17}$ cm^{-3} for devices with channel width equal to 30 and 80 nm: semitransparent surfaces correspond to simulation results when dots are discharged, while red surfaces are those when dots are charged. As it appears, charged dots at the corner are more effective in depleting the channel in the case where the width is smaller.

V. CONCLUSION

In conclusion, we have described a 3-D approach that has allowed us to investigate the dependence of the programming window of SOI nanocrystal memories on the channel width. As confirmed by the experiments, the observed effect can be explained by preferential injection of charge in correspondence of channel edges due to the reduced barrier height and increased electric field in the tunnel oxide near the channel edges. Percolating transport might represent an additional mechanism reinforcing such an effect in subthreshold regions. Let us also stress the fact that since charge is mostly stored in dots close to the channel edges, the dispersion of the threshold voltage shift can be significant and even larger than that shown in Figs. 4 and 5. This imposes the choice of a larger acceptable programming window.

REFERENCES

[1] S. Tiwari, F. Rana, K. Chan, H. Hassafi, W. Chan, and D. Buchanan, "Volatile and nonvolatile memories in silicon with nano-crystal storage," in *Int. Electron Devices Meeting Tech. Dig.*, 1995, pp. 521–524.

[2] J. J. Welser, S. Tiwari, S. Rishton, K. Y. Lee, and Y. Lee, "Room temperature operation of a quantum dot flash memory," *IEEE Trans. Electron Devices*, vol. 18, no. 6, pp. 278–280, Jun. 1997.

[3] K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, "Room-temperature single electron memory," *IEEE Trans. Electron Devices*, vol. 41, no. 9, pp. 1628–1638, Sep. 1994.

[4] B. J. Hinds, T. Yamanaka, and S. Oda, "Emission lifetime of polarizable charge stored in nano-crystalline Si based single-electron memory," *J. Appl. Phys.*, vol. 90, pp. 6402–6408, 2001.

[5] L. Guo, E. Leobandung, and S. Y. Chou, "A single-electron transistor memory operating at room temperature," *Science*, vol. 275, p. 649, 1997.

[6] J. D. Blauwe, "Nanocrystal nonvolatile memory devices," *IEEE Trans. Nanotechnol.*, vol. 1, no. 1, pp. 72–77, Mar. 2002.

[7] M. Saitoh, E. Nagata, and T. Hiramoto, "Effects of ultra-narrow channel on characteristic of MOSFET memory with silicon nanocrystal floating gates," in *Int. Electron Devices Meeting Tech. Dig.*, 2002, pp. 181–184.

[8] C. Monzio Compagnoni, D. Ielmini, A. S. Spinelli, A. Lacaita, C. Gerardi, B. De Salvo, L. Perniola, and S. Lombardo, "Program/erase dynamics and channel conduction in nanocrystal memories," in *Int. Electron Devices Tech. Dig.*, 2003, pp. 549–552.

[9] G. Molas, B. De Salvo, G. Ghibaudo, D. Mariolle, A. Toffoli, N. Buffet, R. Puglisi, S. Lombardo, and S. Deleonibus, "Single electron effects and structural effects in ultrascaled silicon nanocrystal floating-gate memories," *IEEE Trans. Nanotechnol.*, vol. 3, no. 1, pp. 42–48, Mar. 2004.

[10] Y. Taur and T. H. Ning, *Fundamental of Modern VLSI Devices*. Cambridge, U.K.: Cambridge Univ. Press, 1998, pp. 26–28.

[11] J. C. Inkson, *Many Body Theory of Solids—An Introduction*. New York: Plenum, 1984, pp. 265–286.

[12] A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli, "Iteration Scheme for the solution of the two-dimensional Schrödinger Poisson equations in quantum structures," *J. Appl. Phys*, vol. 81, pp. 7800–7804, 1997.

[13] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFET's: A 3-D 'atomistic' simulation study," *IEEE Trans. Electron Devices*, vol. 45, p. 2505, 1998.

[14] M. Depas, B. Vermeire, P. W. Mertens, R. L. Van Meirhaeghe, and M. M. Heyns, "Determination of tunneling parameter in ultrathin oxide layer Poly — Si/SiO$_2$/Si structure," *Solid State Electron.*, vol. 38, pp. 1465–1471, 1995.

Gianluca Fiori received the Electronic Engineering degree and Ph.D. degree in electronic engineering from the Universitá di Pisa, Pisa, Italy, in 2001 and 2005, respectively. His Ph.D. thesis concerned approaches for the simulation of 3-D nanoelectronic devices.

His main field of activity is the development of models and codes for the simulations of semiconductor nanoscale devices.

Giuseppe Iannaccone was born on April 28, 1968. He received the Laurea degree (*cum laude*) in electrical engineering and Ph.D. degree from the Universitá di Pisa, Pisa, Italy, in 1992 and 1996, respectively. His Ph.D. thesis concerned transport and noise phenomena in nanoelectronic devices.

Since January 2001, he has been an Associate Professor with the Dipartimento di Ingegneria dell'Informazione, Universitá di Pisa. He is also the principal investigator in other national and international projects. His interests include transport and noise modeling in nanoscale devices, devices and architectures for nanoelectronics, the design of passive RF identification (RFID) transponders, and the exploitation of quantum effects in conventional electron devices. He has participated in a series of European and National research projects as consortium coordinator or principal investigator. He has authored or coauthored over 80 papers in peer-reviewed journals and 50 papers in proceedings of international conferences.

Gabriel Molas received the B.S. and M.S. degrees in physics engineering with a microelectronics specialization and Ph.D. degree in microelectronics and nanoelectronics from the Polytechnics Institute of Grenoble, Grenoble, France, in 2001 and 2004, respectively.

In 2004, he joined the Laboratory of Electronics, Technology and Instrumentation (LETI), Grenoble, France, as a Research Engineer. His research activities cover the engineering and physics of advanced memory devices, in particular silicon dots based, for ULSI technologies.

Barbara De Salvo received the B.S. and M.S. degrees in electronics engineering from the University of Parma, Parma, Italy, in 1996, and the Ph.D. degree in microelectronics from the Polytechnics Institute of Grenoble, Grenoble, France, in 1999.

From 1995 to 1996, she was with the Department of Information Technology, University of Parma, where she was involved with electrical characterization of A/D converters. From 1996 to 1999, she was with the Laboratory of Physics of Semiconductor Devices (LPCS), Centre National de la Recherche Scientifique/Institut National Polytechnique de Grenoble (CNRS/INPG), Grenoble, France, where she was involved with research on reliability of NVMs, gate oxide transport mechanisms, and degradation phenomena. In 1999, she joined the Laboratory of Electronics, Technology and Instrumentation (LETI), Commissariat à l'energie atomique (CEA), Grenoble, France, as a Research Engineer. Her current interests and activities cover the engineering and physics of advanced memory devices (in particular, silicon dots based) for ULSI technologies. From March 1 2002 to March 1 2004, she was the coordinator of the ADAMANT Project (IST-2001-34234) funded by the European Commission within the Fifth Framework Programme. She currently manages the Advanced Memories Group, LETI, covering several projects with industrial partners and advanced topics such as 3-D memory devices and molecular memories. She has authored or coauthored over 100 papers appearing in international refereed journals and conferences. She has supervised several Master and Ph.D. students.