Three-Dimensional Simulation of Charge-Trap Memory Programming—Part I: Average Behavior

Salvatore Maria Amoroso, Alessandro Maconi, Aurelio Mauri, Christian Monzio Compagnoni, *Member, IEEE*, Alessandro S. Spinelli, *Senior Member, IEEE*, and Andrea L. Lacaita, *Fellow, IEEE*

***Abstract*—This paper presents a detailed investigation of charge-trap memory programming by means of 3-D TCAD simulations accounting both for the discrete and localized nature of traps and for the statistical process ruling granular electron injection from the substrate into the storage layer. In addition, for a correct evaluation of the threshold-voltage dynamics, cell electrostatics and drain current are calculated in presence of atomistic doping, largely contributing to percolative substrate conduction. Results show that the low average programming efficiency commonly encountered in nanoscaled charge-trap memory devices mainly results from the low impact of locally stored electrons on cell threshold voltage in presence of fringing fields at the cell edges. Programming variability arising from the discreteness of charge and matter will be addressed in Part II of this paper.**

***Index Terms*—Atomistic doping, charge-trap memory devices, Monte Carlo simulations, semiconductor device modeling.**

I. Introduction

THE PROGRAMMING efficiency of NAND-type charge-trap memory devices has been shown to largely decrease when device dimensions are reduced to the deca–nanometer scale [1]–[4]. By referring to the incremental step-pulse programming (ISPP) algorithm [5], [6], in fact, significant differences have been shown to appear in the ratio between the threshold-voltage increase per step $(\Delta V_{T,s})$ and the step amplitude V_s for large-area capacitors and nanoscaled cells. While the ratio is only slightly below 1 for the former, at least far from the saturation of the available traps in the storage layer [7], quite lower values in the range of 0.5–0.65 are typically reported for the latter [1], [2], [4]. Despite the decrease in the ISPP slope for small-area cells has been clearly correlated not only to cell dimensions but also to cell geometry [3], its origin has not been well assessed so far.

Manuscript received January 27, 2011; revised March 25, 2011; accepted March 27, 2011. Date of publication May 12, 2011; date of current version June 22, 2011. This work was supported in part by the European Commission under the Seventh Framework Programme under Contract 214431 “GOSSAMER.” The review of this paper was arranged by Editor H. Jaouen.

S. M. Amoroso, A. Maconi, and C. Monzio Compagnoni are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milano, Italy, and also with the Italian University Nano-Electronics Team (IU.NET), 20133 Milano, Italy.

A. Mauri is with the Numonyx (now Micron) R&D Technology Development, 20041 Agrate Brianza, Italy.

A. S. Spinelli and A. L. Lacaita are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milano, Italy, with the Italian University Nano-Electronics Team (IU.NET), 20133 Milano, Italy, and also with the Istituto di Fotonica e Nanotecnologie–Consiglio Nazionale delle Ricerche, 20133 Milano, Italy.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2011.2138708

In this paper, extending our previous work [8], we present a detailed simulation analysis of the ISPP dynamics on nanoscaled charge-trap memory cells, highlighting the basic features of the electron storage process and its impact on cell threshold voltage V_T. In order to account in detail for all the effects of discreteness, 3-D technology computer-aided design (TCAD) simulations have been performed accounting for the discrete nature both of traps in the storage layer and of the electron flow charging it [9]–[13]. Moreover, to carefully evaluate the impact of each single stored electron on cell V_T, substrate doping was treated as atomistic when solving for cell electrostatics and source-to-drain-current conduction [14]–[20]. Statistical results were collected following a Monte Carlo approach, randomly changing the number and the position of both the nitride traps and subtrate dopants and reproducing the stochastic process ruling the discrete electron injection into the storage layer during programming. The average results from the Monte Carlo simulations show that the low programming efficiency of nanoscaled charge-trap cells mainly results from the low impact of locally stored electrons on V_T due to fringing fields at the cell edges. This is also supported by results from a continuous 3-D model for the program operation, including only the localized nature of charge storage but neglecting the discreteness of traps, dopants, and electron flow. The statistical variability of the ISPP transients will be addressed by our Monte Carlo simulations in the companion paper [21].

II. Physics-Based Numerical Model

Fig. 1 shows the TCAD structure of the charge-trap memory cell investigated in this paper, featuring an aluminum metal gate, a bottom-oxide/nitride/top-oxide dielectric stack with thicknesses equal to 4/4.5/5 nm, and shallow trench isolations (STI) at the active-area edges. Note that the nitride layer is patterned over the channel area, having width W and length L equal to 18 nm. Instead, the gate extends beyond the active area in the W direction, remaining completely planar. A uniform substrate doping $N_a = 4.2 \times 10^{18}$ cm^{-3} was assumed, discretizing the acceptor atoms in the channel region down to 25 nm from the bottom-oxide interface. A trap density equal to $N_t = 6 \times 10^{19}$ cm^{-3} was assumed for the nitride. Details on doping and nitride-trap discretization can be found in [9].

Fig. 2 shows the simulation procedure adopted for the statistical analysis of the ISPP dynamics. After the definition of the deterministic features of the device, a first Monte Carlo loop is used to gather information on cells having different stochastic configurations of atomistic dopants and nitride traps,

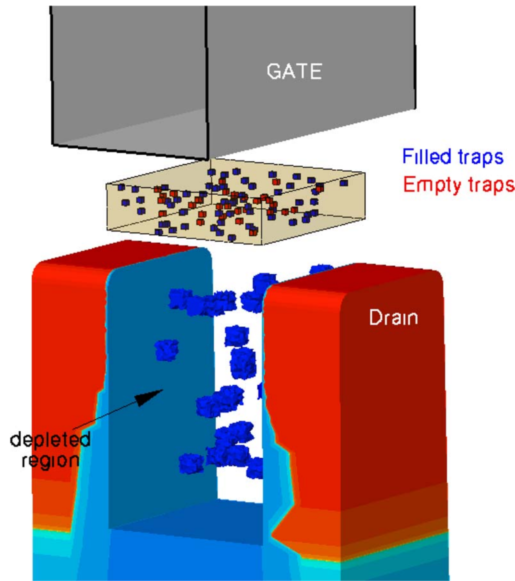

Fig. 1. TCAD structure for the simulated 18-nm metal–oxide–nitride–oxide–semiconductor cell, highlighting discrete dopants in the substrate and discrete traps in the nitride. Red regions: n-type implants. Blue regions: p-type substrate. The oxide regions are not shown for clarity.

which were obtained by drawing their number from a Poisson statistics and their spatial position from a uniform distribution in their discretization region. Assuming empty traps in the nitride, the neutral cell V_T is extracted from the drain-current-versus-gate-voltage (I_D–V_G) transcharacteristics, which was calculated by solving the drift–diffusion equations, as discussed in [9], keeping the source and the bulk grounded and the drain at $V_D = 0.7$ V. An inner Monte Carlo loop is then used to simulate the electron-injection process during each ISPP step. Once cell electrostatics is solved for $V_G = V_{G,\mathrm{prog}}$ (with source, drain and bulk grounded), the tunneling current density $J(x, y)$ is calculated over the channel area in the Wentzel–Kramers–Brillouin approximation taking into account the local electron supply in each point of the channel, and the average electron-injection time from the substrate to each trap is computed as

$$\overline{\tau}_{\mathrm{inj}} = \frac{q}{\int_\sigma J(x, y)dxdy} \tag{1}$$

where q is the electron charge, and the integral is evaluated on an area equal to the trap capture cross section σ (assumed equal to 10^{-14} cm^2 throughout this paper [22]) centered at the trap position. Then, for each trap, the stochastic electron-injection time τ_{inj} is drawn from an exponential distribution with average value equal to its corresponding $\overline{\tau}_{\mathrm{inj}}$, and a single electron is placed in the nitride trap having the smallest τ_{inj} value. This value is added to the total time t elapsed since the beginning of the program operation, then going back to the solution of cell electrostatics with the new electron in the nitride and repeating the calculations for the stochastic injection of the next electron until the end of the ISPP step. When this happens, the I_D–V_G transcharacteristics is calculated again to extract cell V_T after the programming step, then reentering the Monte Carlo loop for electron injection after adding V_s to $V_{G,\mathrm{prog}}$. The simulation flow reaches its end when $V_{G,\mathrm{prog}}$ is equal to the maximum value selected for the ISPP algorithm.

As a final remark, note that the model does not include the possibility for electron emission from filled nitride traps. This process is surely very important in the late stages of the program operation, when the number of electrons stored in the nitride is large and when the electric field within the top oxide is high [22]. The analysis is therefore expected to hold for V_T values far from the saturation of the available traps in the nitride. Moreover, no trapping was included in the bottom- and top-oxide layers, which were considered as ideal trap-free dielectrics. Despite this choice allows the investigation of the ultimate programming performance of the charge-trap technology, the trapping in the top dielectric should be carefully taken into account when a high-k material is adopted. Similar to electron emission, results on TANOS memory devices revealed, in fact, that trapping in the top alumina layer has a nonnegligible impact on the late stages of the programming transient [23]–[25].

III. Electron Injection

In order to highlight the basic features of the statistical process leading to a discrete electron injection into the nitride traps, ISPP was first investigated on a single stochastic cell, having a random configuration of substrate dopants and nitride traps, whose number was set nearly equal to the corresponding average value. Fig. 3 shows $J(x, y)$ over the channel area in the case of $V_{G,\mathrm{prog}} = 13$ V and empty nitride, representing the initial condition for the simulated ISPP. A largely nonuniform tunneling current density clearly appears, with higher $J(x, y)$ at the STI corners where field peaks occur [9]. As a result, $\overline{\tau}_{\mathrm{inj}}$ strongly depends on the trap position over the channel, as shown in Fig. 4. Traps placed near the active-area edges along W display a smaller average capture time than traps located near the channel center, with only a relatively small statistical dispersion of the scatter plot. This dispersion is due, first of all, to a weak $\overline{\tau}_{\mathrm{inj}}$ dependence on the trap position along L, resulting into a slightly smaller capture time when the trap is placed halfway between the source and the drain than near the junctions. Moreover, $\overline{\tau}_{\mathrm{inj}}$ also depends on the trap position along the nitride thickness when the trap is vertically aligned to other traps. In this case, in fact, the injected electron is assumed to be captured by the lowest trap in the stack, and $\overline{\tau}_{\mathrm{inj}}$ for the higher traps was calculated by limiting the integral in (1) only to the fraction of the σ projection not shadowed by the lowest traps.

From the results in Figs. 3 and 4, electron storage in the nitride layer is not uniform during programming. This is shown in Fig. 5, where the average distance of stored electrons from the channel center is reported along the L and W directions. If electrons were uniformly stored in the nitride, then their average distance from the channel center in our 18-nm cell should be 4.5 nm (i.e., the dashed line in the figure). Fig. 5 instead shows that the real distance in the W direction is larger than this value during the first steps of ISPP, due to a larger electron injection/capture near the cell STI corners. However, such a trapped charge reduces the tunneling

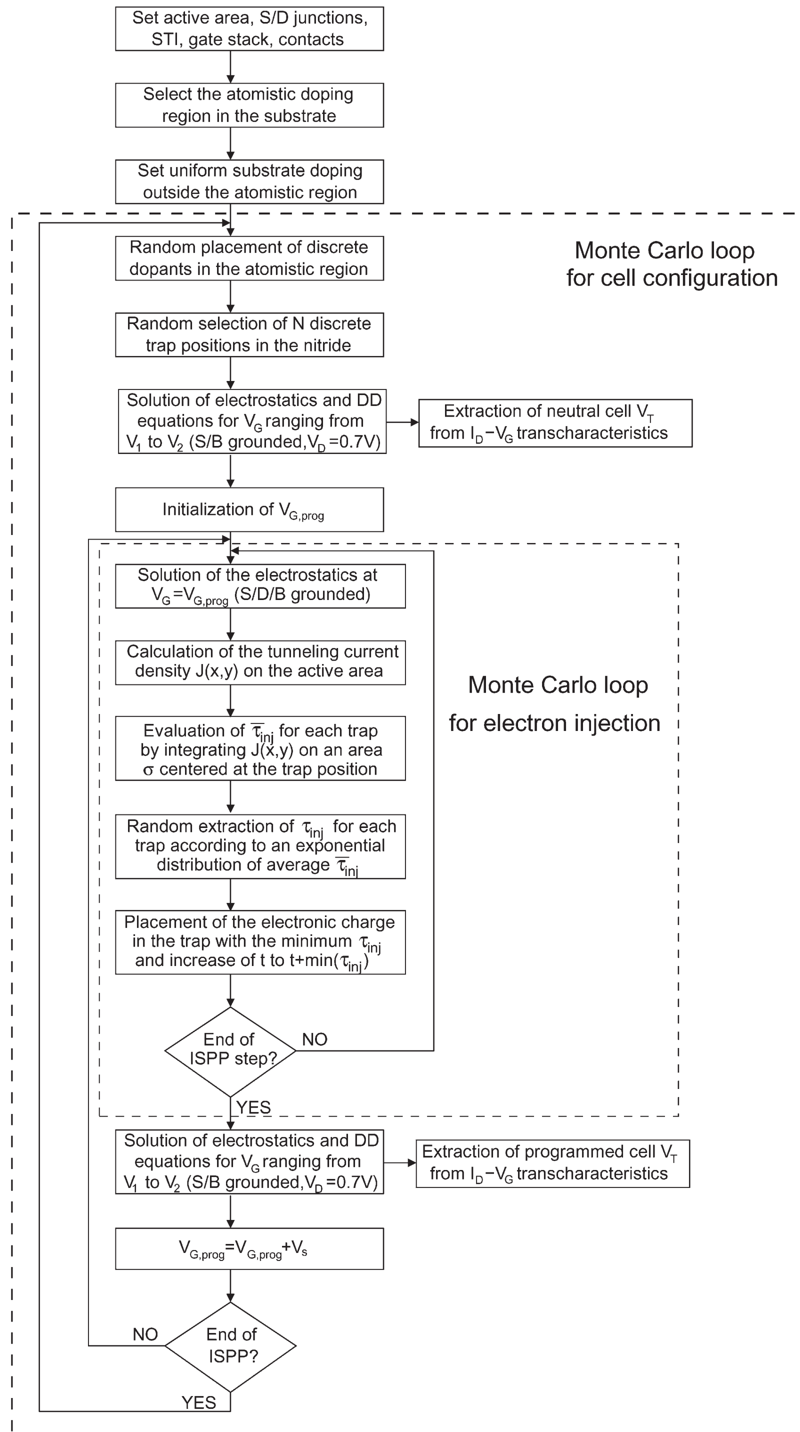

Fig. 2. Block diagram for the simulation procedure used to statistically investigate ISPP on nanoscaled charge-trap memory devices.

probability in the same regions, making the electron trapping in the central part of the channel more and more important as programming proceeds. As a result, the average electron distance from the channel center decreases and approaches 4.5 nm. Fig. 5 also shows a weak increase in the stored charge position along L. This trend can be explained considering the results in Figs. 3 and 4, showing that, during the initial ISPP steps, electrons are nearly stored halfway between the source and the drain. However, note that disuniformities in the electron distribution along L are recovered as ISPP proceeds, with the mean stored charge position approaching 4.5 nm. Finally, Fig. 5 also shows that the average distance of stored electrons from the nitride/bottom-oxide interface slightly increases as programming proceeds.

IV. ΔV_T Transients and ISPP Efficiency

The impact of electrons stored in the nitride layer on cell V_T is strongly affected by fringing fields and atomistic substrate doping [9], [10]. In fact, both these effects result into a

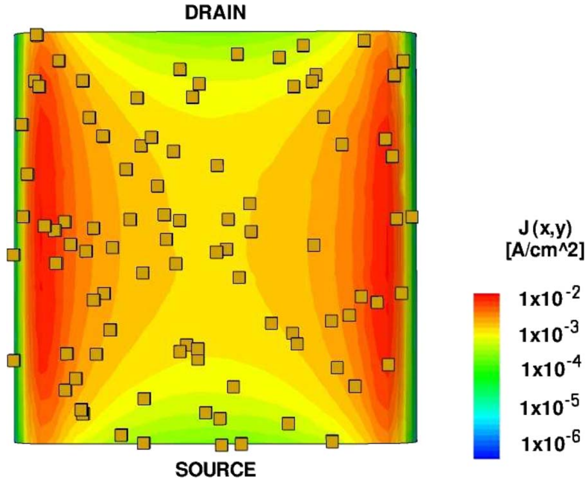

Fig. 3. Tunneling current density $J(x, y)$ at the channel surface of the stochastic cell investigated in Sections III and IV, for $V_G = 13$ V and empty nitride (i.e., the beginning of ISPP). The squares are trap positions in the nitride.

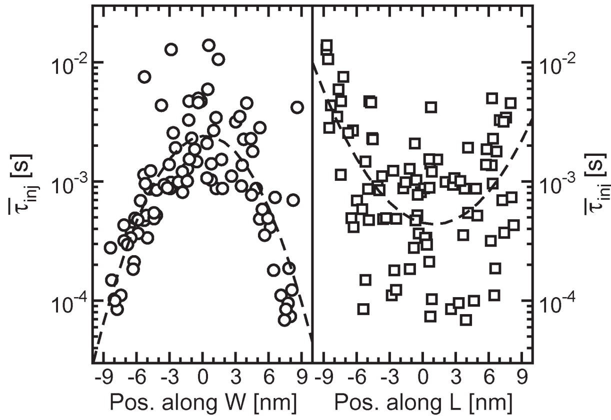

Fig. 4. Average injection time into the nitride traps $\overline{\tau}_{\rm inj}$ as a function of trap position along W (left) and L (right), for $V_G = 13$ V and empty nitride. The channel center is at (0, 0). The dashed lines are the average trend guidelines.

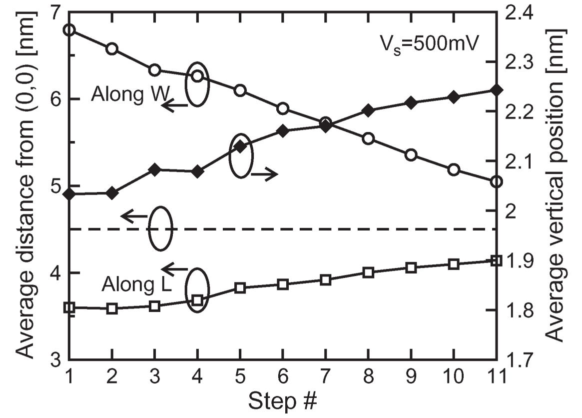

Fig. 5. Average distance of trapped electrons from the channel center [i.e., (0, 0)] in the W and L directions as ISPP proceeds (the dashed line represents the expected average distance in the case electrons were uniformly distributed over the channel). The average vertical position of the trapped electrons in the nitride is also shown (0 is the bottom-oxide/nitride interface).

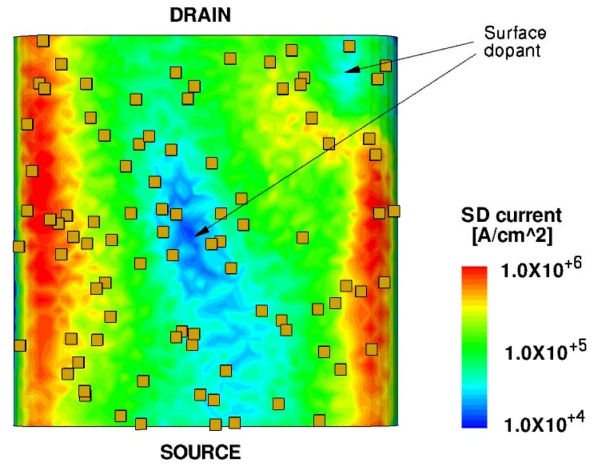

Fig. 6. Source-to-drain-current density at the channel surface of the stochastic cell investigated in Sections III and IV, during read at $V_G = V_T$, with empty nitride. The squares are the trap positions in the nitride.

nonuniform source-to-drain-current density on the active area during read, as shown in Fig. 6, where percolative substrate conduction clearly appears. Note that the current density profile is not symmetrical in the W direction, as expected in presence only of field intensifications at the cell STI corners, due to the variability contribution given by atomistic doping on substrate inversion. In these conditions, the V_T shift ΔV_T obtained by a single electron stored in the nitride depends on the electron position over the active area, with electrons placed above a current percolation path having a larger impact on V_T than the others, due to their larger possibility to stop source-to-drain conduction. This is shown in Fig. 7, where the ΔV_T produced by a single electron placed in the different nitride traps is reported as a function of the trap position along W and L. In addition to the dispersion of the scatter plot, matching the current density profile in Fig. 6, the resulting ΔV_T values in the graph are all below the 1-D prediction $\Delta V_T^{1-D} = q/C_{\rm NG}^{1-D} \simeq$ 88 mV, where $C_{\rm NG}^{1-D} = 1.81$ aF is the 1-D capacitance from the nitride center to the gate, which is given by

$$C_{\rm NG}^{1-D} = \epsilon_{\rm ox} \frac{WL}{t_{\rm eq}} \qquad (2)$$

with $t_{\rm eq} = t_{\rm top} + t_N \epsilon_{\rm ox}/2\epsilon_N$ representing the equivalent oxide thickness of the dielectric materials from the nitride central position to the gate ($\epsilon_{\rm ox}$ and ϵ_N and $t_{\rm top}$ and t_N are the oxide and nitride dielectric constants and thicknesses, respectively). The smaller ΔV_T in Fig. 7 with respect to the 1-D prediction is the result of fringing fields in the 3-D electrostatics, increasing the gate coupling both with the nitride stored charge and the channel, as discussed in [10].

In order to correctly quantify the stored charge effect on cell V_T, Fig. 8 shows the average number of electrons in the nitride ($\overline{n}_t$) as a function of the average V_T shift $\overline{\Delta V_T}$ at the end of the ISPP steps ($V_s = 500$ mV and step duration $\tau_s = 10$ μs), resulting from more than 100 Monte Carlo simulations on the investigated stochastic cell. From the slope of this graph, an effective 3-D electrostatic capacitance value $C_{\rm NG}^{3-D} = 4.2$ aF can be extracted, allowing a correct evaluation of $\overline{\Delta V_T}$ as $q\overline{n}_t/C_{\rm NG}^{3-D}$. Note that the effective $C_{\rm NG}^{3-D}$ value is larger than $C_{\rm NG}^{1-D}$, as required to give rise to ΔV_T values lower than the 1-D predictions in Fig. 7. The variability of the scatter plot cannot, however, be explored by C_{NG}^{3D}, as this is defined from Fig. 8 using the integral of n_t and ΔV_T as programming

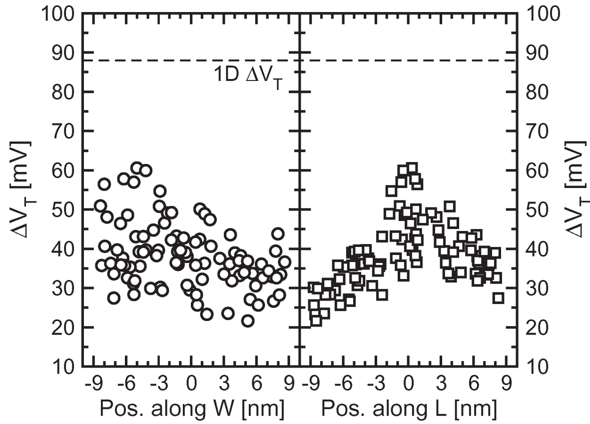

Fig. 7. ΔV_T given by a single electron stored in the nitride layer as a function of the trap position along W (left) and L (right). The channel center is at (0, 0).

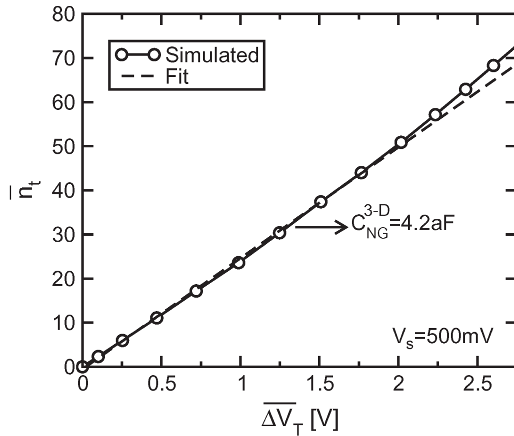

Fig. 8. $\overline{n}_t$ versus $\overline{\Delta V_T}$ relation during ISPP with $V_s = 500$ mV and $\tau_s = 10\ \mu$s. The linear fit of the curve allows the evaluation of $C_{\rm NG}^{3-D}$.

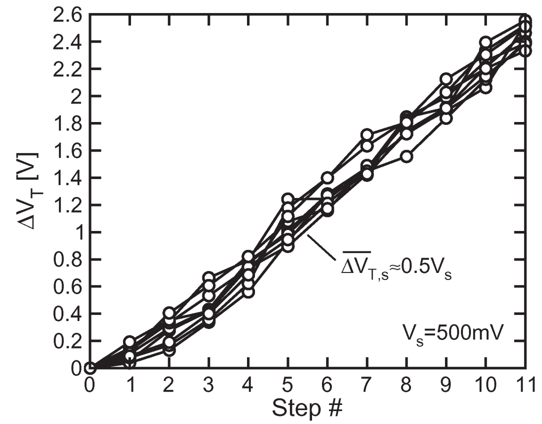

Fig. 9. Monte Carlo simulation results for the ΔV_T transients of the stochastic cell investigated in Sections III and IV.

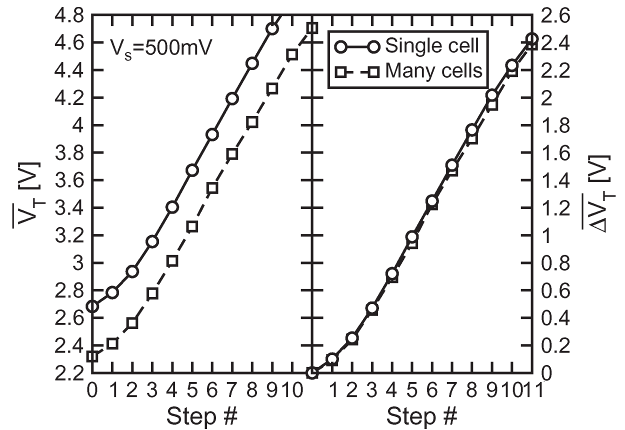

Fig. 10. Average V_T and ΔV_T transients during ISPP simulated from many Monte Carlo runs on the same single cell or from a single Monte Carlo run on many stochastic cells.

proceeds, therefore averaging the different effects of stored electrons on cell V_T. The displacement of the $\overline{n}_t$ versus $\overline{\Delta V_T}$ relation from the linear behavior for large numbers of stored electrons in Fig. 8 clearly highlights, however, a change in the average impact of the stored charge on V_T. This reflects the displacement of the storage position from the STI cell corners to the central channel regions and from half-way between source and drain to the junctions as ISPP proceeds, in agreement with the results in Fig. 5 and determining a consequent reduction of the stored electron control on V_T.

Fig. 9 shows some Monte Carlo simulations for the ΔV_T transients during ISPP on the same stochastic cell. In addition to the statistical dispersion of the curves, which will be addressed in detail in Part II of this paper, an average increase per step $\overline{\Delta V_{T,s}} \simeq 0.5 V_s$ appears, confirming the low ISPP efficiency commonly observed on charge-trap memory devices [1], [2], [4]. The low $\overline{\Delta V_{T,s}}/V_s$ is mainly the result of the low impact exerted by stored electrons on V_T and, in turn, of the large effective $C_{\rm NG}^{3-D}$. Note, in fact, that, by assuming for the electrons an electrostatic control as in the 1-D case, the resulting ISPP slope would be increased by ratio $C_{\rm NG}^{3-D}/C_{\rm NG}^{1-D} \simeq 2.3$, i.e., it would be even a little bit larger than 1. This means that the nonuniform $J(x,y)$ and $\overline{\tau}_{\rm inj}$ profiles over the active area not only do not degrade the electron-injection process but also enhance indeed the process with respect to the 1-D case.

In order to exclude that the previous results are a specific feature of the single stochastic cell investigated, we simulated the ISPP transient on a large number of different cells, following the complete Monte Carlo procedure in Fig. 2. Fig. 10 shows the comparison between the average results from many ISPP transients on the same cell and from a single ISPP transient on many stochastically different cells, in terms of $\overline{V_T}$ and $\overline{\Delta V_T}$. Despite the number of atomistic dopants in the substrate of the previously considered single cell was selected nearly equal to the average value expected from N_a, a higher neutral V_T value (i.e., V_T at step 0) appears for this cell with respect to the average value of the cells statistics. This is due to a significant impact of dopants position on the cell neutral V_T value [9]. However, no significant difference appears between the single- and multicell $\overline{\Delta V_T}$ transients in Fig. 10. This confirms in more general terms the low programming efficiency of the investigate charge-trap cell, leading to an average $\overline{\Delta V_{T,s}}/V_s \simeq 0.5$.

V. Scaling Analysis

The results in Fig. 10 revealed that the low ISPP efficiency is a general feature of nanoscaled charge-trap memory devices,

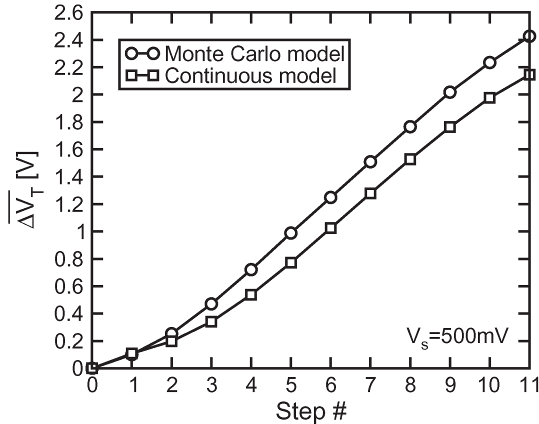

Fig. 11. Comparison between the average ΔV_T transient during ISPP from our Monte Carlo model and from a 3-D numerical tool treating substrate doping, nitride traps, and the programming electron flow as continuous.

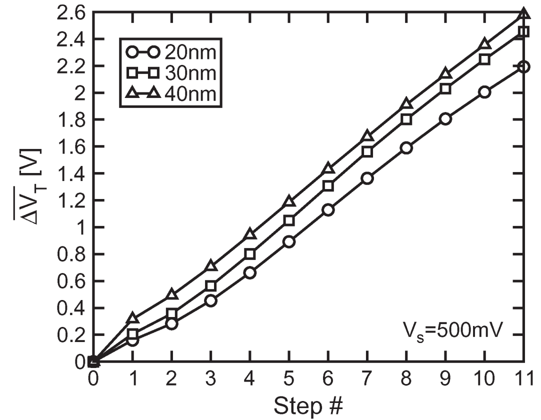

Fig. 12. Simulated average ΔV_T transients during ISPP for different cell sizes.

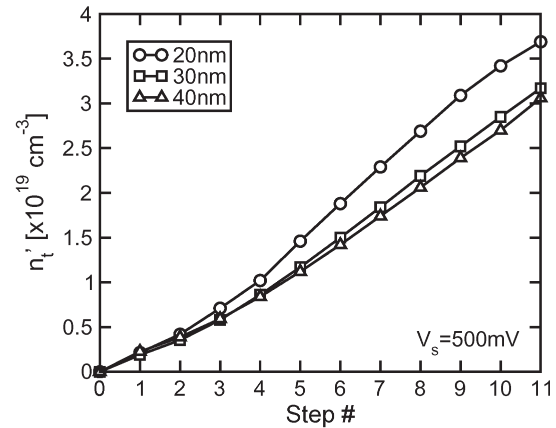

Fig. 13. Simulated average trapped electron density n_t' in the nitride during ISPP for different cell sizes.

clearly appearing in the average programming behavior extracted from single- and multi-cell statistics. Fig. 11 shows that this average behavior can be reproduced, as reasonably expected, even neglecting the discrete nature of charge and matter, i.e., using 3-D simulations with continuous substrate doping, continuous trap density in the nitride, and continuous electron flow from the substrate to the nitride during programming. These simulations were obtained by a numerical tool extending to 3-D geometries the model for charge-trap memory programming that we presented in [22], implementing electron trapping as [26], [27]

$$\frac{dn_t'(x,y)}{dt} = \frac{J(x,y)}{q}\sigma\left[N_t - n_t'(x,y)\right] \tag{3}$$

where $n_t'(x,y)$ is the trapped electron density in the nitride. Note that, for a correct comparison of the Monte Carlo and continuous models, no emissivity from filled electron traps was included in (3). Moreover, in the continuous tool cell, V_T was obtained from the simulation of cell I_D–V_G transcharacteristics, as discussed in Fig. 2. Despite a small displacement of the ΔV_T curves, which are mainly attributed to numerical differences between the two simulation codes, Fig. 11 confirms, first of all, the correctness of the Monte Carlo approach for the program operation that we presented in Fig. 2. Moreover, the agreement between the results in Fig. 11 makes possible the use of the continuous 3-D model to investigate the programming performance when only the average results are of interest. However, note that the Monte Carlo model allows a more complete analysis of the program operation, including variability effects representing the topic of Part II of this paper [21].

Fig. 12 shows a scaling analysis of the ISPP $\overline{\Delta V_T}$ transient obtained using the continuous 3-D model for cell programming and assuming a reduction of the cell area while keeping the same gate stack investigated in the previous sections. A reduction of the ISPP efficiency with cell scaling clearly appears, in terms both of the slope and the horizontal delay of the curves. This is due to a decreasing impact of the electrons stored in the nitride on V_T as cell dimensions are reduced, as discussed in Section IV. Fig. 13 shows, in fact, that the average n_t' curves display faster electron injection and storage in the nitride as scaling proceeds, due to a larger field enhancement at the corners of the cell area.

VI. Conclusion

This paper has presented a detailed simulation analysis of charge-trap memory programming, carefully reproducing the discrete and localized nature of storage traps and the statistical process ruling the granular electron injection into the storage layer. The average results for ISPP on single- and multi–cell statistics revealed that the low programming efficiency of nanoscaled charge-trap cells mainly results from the low impact of locally stored electrons on cell V_T in presence of fringing fields and 3-D electrostatics. Programming variability arising from the discreteness of charge and matter will be addressed in Part II of this paper.

Acknowledgment

The authors would like to thank P. Cappelletti, E. Camerlenghi, R. Bez, L. Baldi, P. Tessariol, A. Ghetti,

E. Greco, E. Camozzi, S. Viganò of Numonyx (now Micron Technology Inc.) for the discussions and support.

References

[1] H.-T. Lue, T.-H. Hsu, S.-Y. Wang, Y.-H. Hsiao, E.-K. Lai, L.-W. Yang, T. Yang, K.-C. Chen, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, "Study of local trapping and STI edge effects on charge-trapping NAND Flash," in *IEDM Tech. Dig.*, 2007, pp. 161–164.

[2] H.-T. Lue, T.-H. Hsu, S.-Y. Wang, E.-K. Lai, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, "Study of incremental step pulse programming ISPP and STI edge effect of BE-SONOS NAND Flash," in *Proc. IRPS*, 2008, pp. 693–694.

[3] H.-T. Lue, T.-H. Hsu, Y.-H. Hsiao, S.-C. Lai, E.-K. Lai, S.-P. Hong, M.-T. Wu, F. H. Hsu, N. Z. Lien, C.-P. Lu, S.-Y. Wang, J.-Y. Hsieh, L.-W. Yang, T. Yang, K.-C. Chen, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, "Understanding STI edge fringing field effect on the scaling of charge-trapping (CT) NAND Flash and modeling of incremental step pulse programming (ISPP)," in *IEDM Tech. Dig.*, 2009, pp. 839–842.

[4] M. F. Beug, T. Melde, M. Czernohorsky, R. Hoffmann, J. Paul, R. Knoefler, and A. T. Tilke, "Analysis of TANOS memory cells with sealing oxide containing blocking dielectric," *IEEE Trans. Electron Devices*, vol. 57, no. 7, pp. 1590–1596, Jul. 2010.

[5] G. J. Hemink, T. Tanaka, T. Endoh, S. Aritome, and R. Shirota, "Fast and accurate programming method for multi-level NAND EEPROMs," in *VLSI Symp. Tech. Dig.*, 1995, pp. 129–130.

[6] C. Calligaro, A. Manstretta, A. Modelli, and G. Torelli, "Technological and design constraints for multilevel Flash memories," in *Proc. 3rd IEEE Int. Conf. Electron., Circuits Syst.*, 1996, pp. 1005–1008.

[7] A. Maconi, C. Monzio Compagnoni, S. M. Amoroso, E. Mascellino, M. Ghidotti, G. Padovini, A. S. Spinelli, A. L. Lacaita, A. Mauri, G. Ghidini, N. Galbiati, A. Sebastiani, C. Scozzari, E. Greco, E. Camozzi, and P. Tessariol, "Investigation of the ISPP dynamics and of the programming efficiency of charge-trap memories," in *Proc. ESSDERC*, 2010, pp. 444–447.

[8] S. M. Amoroso, A. Maconi, A. Mauri, C. Monzio Compagnoni, E. Greco, E. Camozzi, S. Viganò, P. Tessariol, A. Ghetti, A. S. Spinelli, and A. L. Lacaita, "3D Monte Carlo simulation of the programming dynamics and their statistical variability in nanoscale charge-trap memories," in *IEDM Tech. Dig.*, 2010, pp. 540–543.

[9] A. Mauri, C. Monzio Compagnoni, S. M. Amoroso, A. Maconi, A. Ghetti, A. S. Spinelli, and A. L. Lacaita, "Comprehensive investigation of statistical effects in nitride memories—Part I: Physics-based modeling," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2116–2123, Sep. 2010.

[10] C. Monzio Compagnoni, A. Mauri, S. M. Amoroso, A. Maconi, E. Greco, A. S. Spinelli, and A. L. Lacaita, "Comprehensive investigation of statistical effects in nitride memories—Part II: Scaling analysis and impact on device performance," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2124–2131, Sep. 2010.

[11] C. Monzio Compagnoni, A. S. Spinelli, R. Gusmeroli, A. L. Lacaita, S. Beltrami, A. Ghetti, and A. Visconti, "First evidence for injection statistics accuracy limitations in NAND Flash constant-current Fowler–Nordheim programming," in *IEDM Tech. Dig.*, 2007, pp. 165–168.

[12] C. Monzio Compagnoni, A. S. Spinelli, R. Gusmeroli, S. Beltrami, A. Ghetti, and A. Visconti, "Ultimate accuracy for the NAND Flash program algorithm due to the electron injection statistics," *IEEE Trans. Electron Devices*, vol. 55, no. 10, pp. 2695–2702, Oct. 2008.

[13] C. Monzio Compagnoni, R. Gusmeroli, A. S. Spinelli, and A. Visconti, "Analytical model for the electron-injection statistics during programming of nanoscale NAND Flash memories," *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 3192–3199, Nov. 2008.

[14] H.-S. Wong and Y. Taur, "Three-dimensional 'atomistic' simulation of discrete random dopant distribution effects in sub-0.1 μm MOSFET's," in *IEDM Tech. Dig.*, 1993, pp. 705–708.

[15] A. Asenov, A. R. Brown, J. H. Davies, and S. Saini, "Hierarchical approach to 'atomistic' 3-D MOSFET simulation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 18, no. 11, pp. 1558–1565, Nov. 1999.

[16] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1837–1852, Sep. 2003.

[17] N. Sano, K. Matsuzawa, M. Mukai, and N. Nakayama, "On discrete random dopant modelling in drift-diffusion simulations: Physical meaning of 'atomistic' dopants," *Microelectron. Reliab.*, vol. 42, no. 2, pp. 189–199, Feb. 2002.

[18] G. Roy, A. R. Brown, F. Adamu-Lema, S. Roy, and A. Asenov, "Simulation study of individual and combined sources of intrinsic parameter fluctuations in conventional Nano-MOSFETs," *IEEE Trans. Electron Devices*, vol. 53, no. 12, pp. 3063–3070, Dec. 2006.

[19] M. F. Bukhori, S. Roy, and A. Asenov, "Statistical aspects of reliability in bulk MOSFETs with multiple defect states and random discrete dopants," *Microelectron. Reliab.*, vol. 48, no. 8/9, pp. 1549–1552, Aug./Sep. 2008.

[20] A. Ghetti, C. Monzio Compagnoni, A. S. Spinelli, and A. Visconti, "Comprehensive analysis of random telegraph noise instability and its scaling in deca–nanometer Flash memories," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1746–1752, Aug. 2009.

[21] A. Maconi, S. M. Amoroso, C. Monzio Compagnoni, A. Mauri, A. S. Spinelli, and A. L. Lacaita, "Three-dimensional simulation of charge-trap memory programming—Part II: Variability," *IEEE Trans. Electron Devices*, 2011, to be published.

[22] C. Monzio Compagnoni, A. Mauri, S. M. Amoroso, A. Maconi, and A. S. Spinelli, "Physical modeling for programming of TANOS memories in the Fowler–Nordheim regime," *IEEE Trans. Electron Devices*, vol. 56, no. 9, pp. 2008–2015, Sep. 2009.

[23] S. M. Amoroso, A. Mauri, N. Galbiati, C. Scozzari, E. Mascellino, E. Camozzi, A. Rangoni, T. Ghilardi, A. Grossi, P. Tessariol, C. Monzio Compagnoni, A. Maconi, A. L. Lacaita, A. S. Spinelli, and G. Ghidini, "Reliability constraints for TANOS memories due to alumina trapping and leakage," in *Proc. IRPS*, 2010, pp. 966–969.

[24] L. Larcher, A. Padovani, V. della Marca, P. Pavan, and A. Bertacchini, "Investigation of trapping/detrapping mechanisms in Al_2O_3 electron/hole traps and their influence on TANOS memory operation," in *Proc. VLSI-TSA*, 2010, pp. 52–53.

[25] G. Molas, L. Masoero, P. Blaise, A. Padovani, J. P. Colonna, E. Vianello, M. Bocquet, E. Nowak, M. Gasulla, O. Cueto, H. Grampeix, F. Martin, R. Kies, P. Brianceau, M. Gely, A. M. Papon, D. Lafond, J. P. Barnes, C. Licitra, G. Ghibaudo, L. L. S. Deleonibus, and B. De Salvo, "Investigation of the role of H-related defects in Al_2O_3 blocking layer on charge-trap memory retention by atomistic simulations and device physical modelling," in *IEDM Tech. Dig.*, 2010, pp. 536–539.

[26] P. C. Arnett, "Transient conduction in insulators at high fields," *J. Appl. Phys.*, vol. 46, no. 12, pp. 5236–5243, Dec. 1975.

[27] E.-S. Choi, H.-S. Yoo, K.-H. Park, S.-J. Kim, J.-R. Ahn, M.-S. Lee, Y.-O. Hong, S.-G. Kim, J.-C. Om, M.-S. Joo, S.-H. Pyi, S.-S. Lee, S.-K. Lee, and G.-H. Bae, "Modeling and characterization of program/erasure speed and retention of TiN-gate MANOS (Si-Oxide-SiN_x-Al_2O_3-Metal gate) cells for NAND Flash memory," in *Proc. Non-Volatile Semicond. Memory Workshop*, 2007, pp. 83–84.

Salvatore Maria Amoroso was born in Catania, Italy, in 1983. He received the B.S. and M.S. degrees in physics engineering, in 2005 and 2008, respectively, from the Politecnico di Milano, Milan, Italy, where he is currently working toward the Ph.D. degree in information technology in the Dipartimento di Elettronica e Informazione.

His research interests include modeling and numerical simulation of semiconductor devices, with particular interest on innovative nonvolatile memories.

Alessandro Maconi was born in Carate Brianza, Italy, in 1983. He received the Laurea degree in electronics engineering, in 2008, from the Politecnico di Milano, Milan, Italy, where he is currently working toward the Ph.D. degree in the Dipartimento di Elettronica e Informazione.

His research interests include characterization and modeling of advanced nonvolatile memories, with particular interests to TANOS memories.

Aurelio Mauri was born in1969. He received the M.S. degree in plasma physics (*cum laude*) from the University of Milano, Milan, Italy, in 1995.

Then, he spent a research period at Rutherford Appleton Laboratory, U.K., and ENEA, Frascati, Italy, working on characterization and modeling of the highly dense plasmas produced by laser confinement. In 1996, he started to work for a semiconductor company focused on chemistry treatment of the silicon surfaces. From 1998 to 2000, he spent a "nonworking" time helping young generation. In 2000, he jointed ST Microelectronics, Catania, Italy, and since 2004, he has been with the Nonvolatile Technology Development TCAD Group. Then, he moved to Numonyx and finally to Micron R&D, Agrate Brianza, Italy. His research interests include electronics, particularly modeling of advanced nonvolatile memories and tunneling in dielectrics media, and physics, focusing on advanced modeling of impurity diffusion in silicon with the Kinetic Mone Carlo approach and of materials properties with the ab-initio approach.

Christian Monzio Compagnoni (M'08) was born in Busto Arsizio, Italy, in 1976. He received the Laurea (*cum laude*) degree in electronics engineering and the Ph.D. degree in information technology from the Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively.

Since 2002, he has been with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, where he became an Assistant Professor in 2006. He is also currently with the Italian University Nanoelectronics Team (IU.NET), Milano, Italy. His research interests include characterization and modeling of advanced nonvolatile memories and metal–oxide–semiconductor devices.

Dr. Monzio Compagnoni was a member of the memory committee of the International Reliability Physics Symposium (IRPS) in 2009 and 2010. He was the recipient of the Outstanding Paper Award at the IRPS in 2008.

Alessandro S. Spinelli (M'99–SM'07) was born in Bergamo, Italy, in 1966. He received the Laurea (*cum laude*) and Ph.D. degrees in electronics engineering from the Politecnico di Milano, Milano, Italy.

In 1995, he was a Visiting Scholar with the University of Tennessee Space Institute, Tullahoma, where he worked on single-molecule detection in solution, and in 1996, he worked as a Consultant with the Central Department of Research and Development, STMicroelectronics, Agrate Brianza, Italy. In 1997, he became an Assistant Professor with the Politecnico di Milano and was with the Università degli Studi dell'Insubria, Como, Italy, as an Associate Professor of electronics in 1998. In 2001, he was a Visiting Professor with the Institute National Polytechnique de Grenoble, Grenoble, France. Since 2006, he has been a Full Professor of electronics with the Dipartimento di Elettronica e Informazione, Politecnico di Milano. He has conducted experimental and theoretical research in electronics instrumentation and microelectronics, coauthoring more than 140 papers published in international journals or presented at international conferences. His current research interests include experimental characterization and modeling of nonvolatile memory cell reliability, development of innovative nonvolatile memories, and circuit design for neural signal readout.

Dr. Spinelli has served in the technical committees of the IEDM and IRPS conferences.

Andrea L. Lacaita (F'09) received the Laurea degree in nuclear engineering from Politecnico di Milano, Milano, Italy, in 1985.

From 1987 to 1992, he was a Researcher with the CNR (Italian National Research Council). In1992, he joined the Politecnico di Milano, Milano, Italy, as an EE Professor, became a Full Professor in 2000, and is currently a Full Professor of electronics, teaching "Electron Devices" and "Electronic Circuit Design." Since 1992, he has been Head of the Micro and Nanoelectronics Laboratory, Dipartimento di Elettronica e Informazione (DEI). He has been a Visiting Scientist/Professor with AT&T Bell Laboratories, Murray Hill, NJ (1989–1990), and the IBM T.J. Watson Research Center, Yorktown Heights, NY (1999). He is coauthor of more than 200 papers published in international journals or presented to international conferences, patents, and several educational books in electronics. In the field of device physics he has contributed to the study of quantum effects, as well as experimental characterization techniques and numerical models of nonvolatile memories [both Flash and emerging (phase-change memory and resistive random-access memory)].

Prof. Lacaita has been serving in several scientific committees, i.e., IEEE-IEDM (2001–2002), IEDM European Chair (2003–2004), ESSDERC (2005 and 2009), and IEEE VLSI Symposium (2005–2008). From 2006 to 2008, he has served as Department Chair of the Dipartimento di Elettronica ad Informazione and member of the Academic Senate (2007–2008).