Modeling Strategies for Flash Memory Devices

Andrea Padovani*, Luca Larcher* and Paolo Pavan**

*DISMI, Università di Modena e Reggio Emilia, Via Amendola 2, 42122 Reggio Emilia (RE), Italy, and IU.NET – Italian University Nano-Electronics Team
**DII, Università di Modena e Reggio Emilia, Strada Vignolese 905, 41125 Modena (MO), Italy, and IU.NET – Italian University Nano-Electronics Team

ABSTRACT

In this paper, we will review the modeling strategies for standard and advanced Flash memory devices based on Floating Gate devices developed by our research group in the last ten years. We will show a complete compact model that includes program/erase and leakage currents that can be used to simulate memory cells in both DC (read operation) and transient conditions (Program/Erase). The same model can be used also for reliability simulations by providing good descriptions of the degradation mechanisms. We will also show the extended model for circuit simulation of NAND strings, modified to account for capacitive coupling effects. Finally, we will show how the same framework can be used to develop a compact model for operations of advanced planar charge-trapping memory devices.

Keywords: nonvolatile memory, floating gate, compact model, Flash, Nand, TANOS

1 INTRODUCTION

It was exactly ten years ago, at the first Workshop on Compact Modeling, that we started reporting results of our activity on Flash Memory Compact Modeling [1]. Flash memories in those years were not as pervaisve as today, they were mostly arranged in NOR architecture, for both EPROM and EEPROM applications [2]. Today, Flash memories are one of the leading technologies, the market boomed and they are eventullay challenging HDD for storage. They are mostly arranged in NAND architecture and capacity is doubling every year. The Floating Gate transistor is still the building block of a full array of memory cells and a memory chip. New devices emerged and tried to enter the market, but the main actor is still the Floating Gate (FG) transistor [3].

In the literature, one of the important modeling issues of Flash memory cell is calculation of the FG voltage. Several models have been introduced, which are [4] "classified into two kinds of models. One is capacitive coupling model and the other is Larcher's model."

The model we developed is simple and it exploits the research effort in MOS modeling. It is a general framework where we can add a specific current generator to model: the physics of Program/Erase (P/E) mechanisms, the reliability of the single device and of new architectures, statistical effecs [5, 6]; we can also use this model framework as an inspiration to model new devices.

This paper is organized as follows: in Section 2 we will show the basics of the Compact Model (CM) and its capability to simulate a single FG device. In Section 3 we will show how it can be used to simulate a circuit actually used in a memory device. In Section 4 we will show how this model can be adapted for reliability simulation and finally, in Section 5, we will show our more recent results, using this model as an inspiration for new nonvolatile memory CMs based on Charge Trapping devices.

2 COMPACT MODEL

The FG device is the building block of a nonvolatile memory cell and the FG device CM is the basic building block to model a single memory cell, a full array, a memory chip. The simple idea underneath is to model the FG device as a circuit with a MOS transistor and a capacitor between the control gate and the FG node [1, 7] (which is the gate of the MOS transistor), Fig 1. This CM fully exploits the MOS transistor modeling research efforts. Many MOS models have been developed over the last decade (BSIM4 [8], EKV [9], PSP [10], HiSIM [11]). The basic concepts and the functionality of this kind of device are easily understood after knowing the FG potential; it allows modeling of programming and erasing operations by simply adding a set of suitable current generators between the various electrodes.

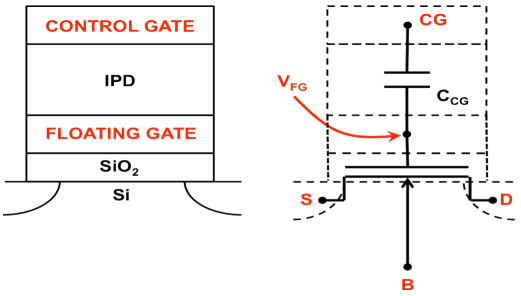

Figure 1. Cross section of a FG device and basic schematic of the CM subcircuit.

This approach gives many advantages compared to standard models [7]: 1. *scalability*: scaling rules are already included in the compact MOS model adopted and they do

not affect directly the V_{FG} calculation routine; 2. *implementation*: it uses standard circuit elements whose parameters can be determined by applying the MOS parameter extraction procedure to the *dummy cell* (which is the cell where FG and CG are short-circuited), and the few additional parameters can be easily estimated from cell layout and cross section; 3. *accuracy*: it depends mainly on the compact MOS model adopted, taking advantage of the many efforts to improve and scale MOS CMs; 4. *computation time*: comparable to a MOS transistor; 5. *modularity*: it can be easily extended to simulate transient behaviors of FG memories by adding a suitable set of voltage controlled current sources to its basic structure.

2.1 DC Operation: Read

In this CM, the FG node is biased to its correct value by an external source: the voltage-controlled voltage source, V_{FG}, (not shown in Fig. 1) which constitutes the core of the model in DC conditions [7] (moreover, it is necessary since there is no general solution in circuit simulators to the calculation in DC conditions of the potential of a floating node). This generator implements the routine to solve the charge balance equation at the FG node: the charge on the MOS gate is equal to the charge of the capacitor between FG and CG, plus the charge forced in/out the FG during cell operation (constant in DC conditions).

The procedure to extract the parameters of device CMs is not a "push-button" task. For FG devices, this task is even more complex than for standard MOS transistors. Reasonable results are obtained paying attention to the differences of the *dummy cell* and a standard MOS transistor: narrow and short geometry, different process steps for junctions. Care has to be devoted to extract the overlap capacitance values (overlap capacitances are very small: their evaluation is particularly critical) [12].

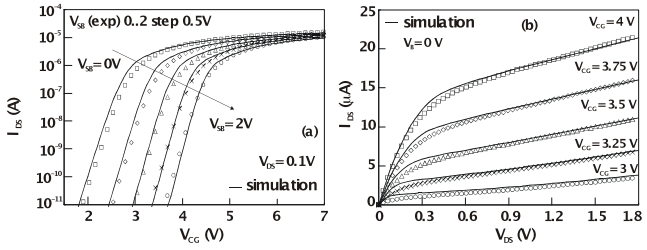

Figure 2. DC characteristics: experimental (symbols) and model (solid lines) for a 0.25μm Flash memory cell (W=0.25μm, L=0.375μm, C_{CG}=0.8fF) [7].

Except the CG-FG capacitance (C_{CG}), additional parameters depend on the kind of FG memory considered. Generally, these parameters are either directly evaluated from the layout of the cell (C_{CG}, tunnel and overlap region areas), or straightly derived from the process recipe (dopings). Sometimes, dedicated measurements performed on MOS capacitor test structures.

Results are shown in Fig. 2, where sub- and above-threshold behavior of a FG device are correctly simulated [7].

2.2 Transient Operations: Program and Erase

This model can simulate the program/erase operations of FG devices by a suitable set of voltage controlled current sources to implement compact formulae of program/erase currents. The number and position of current generators depend on the FG memory considered (EEPROM, Flash NOR, NAND, …) and the writing mechanisms used to transfer charge to and from the FG. For example, we used three voltage controlled current sources to extend the DC model of Flash memory devices in an old technology to reproduce program/erase currents, see Fig. 3 [13]: 1. a voltage controlled current source between FG and S, $\mathbf{I_{W1}}$, which models the Fowler Nordheim (FN) current flowing at the source side (needed when modeling Flash memories erased by FN tunnel at the source side); 2. a voltage controlled current source connected between FG and B, $\mathbf{I_{W2}}$, which models the FN tunnel current flowing toward the substrate; 3. a voltage controlled current source connected between FG and D, $\mathbf{I_{W3}}$, which models Channel Hot Electron (CHE) via approximated compact formulae.

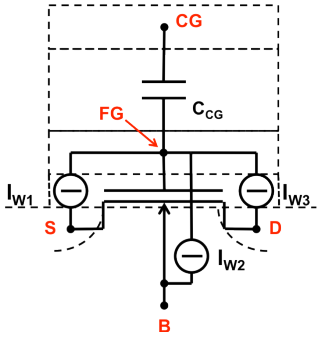

Figure 3. The complete CM of a Flash memory cell: basic framework plus three voltage controlled current sources, I_{W1}, I_{W2} and I_{W3}, that model P/E currents.

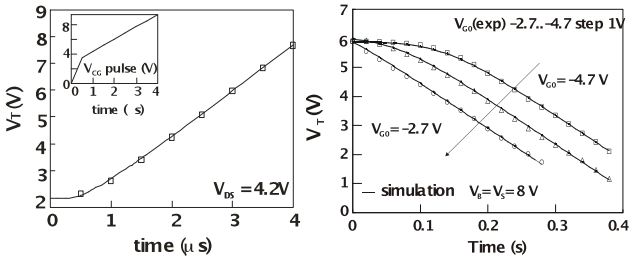

Figure 4. Threshold voltage (V_T) shifts measured (symbols) and simulated (solid lines) during program and erase of a Flash memory cell [14].

Results are shown in Fig. 4. When using this model, the simulation accuracy of FG device program/erase operations depends strictly on the precision of CMs developed to describe write/erase currents. This modular CM can be easily adapted to new generation devices by developing effective CMs of the currents mechanisms used for their functioning. Again, we can exploit the research effort in other fields.

3 CIRCUIT SIMULATIONS

The CM of FG devices we developed can be used within standard circuit simulator software: it can describe single-cell operations and the interaction with the rest of the device, and hence it can be used to check the design of the circuitry around the memory array: algorithms for cell addressing, charge pump sizing taking into account current consumption and voltage drops, etc.

We simulated the sense amplifier circuit shown in Fig. 5 [13]. It is a classic scheme where Mn1 and Mn3, Mn2 and Mn4 provide the current/voltage conversion to bias the reference cell and the cell to be read in the matrix. V_{CELL} and V_{REF} are voltages deriving from the I-V conversion of currents driven by the cell in the memory array and the reference cell, that are compared to generate the V_{SENSE_OUT} digital level.

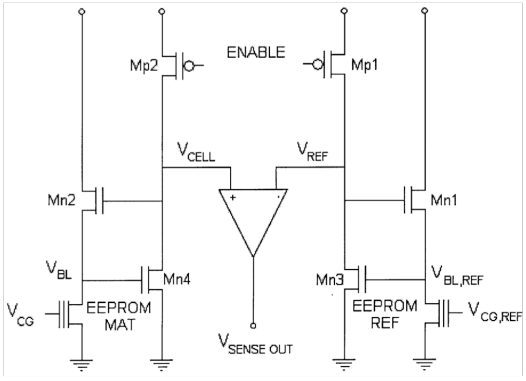

Figure 5. Schematic of the sense amplifier and the direct I-V conversion circuits of an EEPROM memory [13].

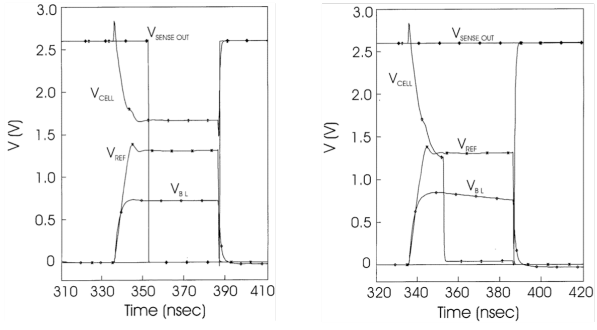

Figure 6. Control signals, and sense amplifier output obtained from read-path circuit simulations in the two cases of a programmed (left) and erased (right) EEPROM memory cell [13].

Simulation results in Figure 6 [13] shows that the output signal of the sense amplifier switches correctly according to the programmed/erased state of the EEPROM memory cell.

This model is effective to simulate FG-based memory cells also in complex circuits, and therefore it can be used to simulate any circuit including a FG memory cell: read paths, non-volatile latches, X and Y decoders, voltage pumps.

4 RELIABILITY SIMULATIONS

Usually, the reliability of FG memory devices is investigated through experimental techniques and the use of suitable ad-hoc models to describe leakage currents through their oxide layers. In fact, leakage currents through gate and interpoly oxides are the most serious concern for the reliability of FG memory devices. In this scenario, we will show that this CM of FG devices (extended to include leakage current effects) can be a versatile and powerful tool for reliability predictions [15]. CMs allow also to bridge the gap between the oxide quality characterization activity performed traditionally on MOS transistors and capacitors, and the actual impact of Stress Induced Leakage Current (SILC) on FG memory reliability. This CM is an effective tool to predict FG memory reliability degradation, the influence on data retention of P/E cycles, P/E bias conditions, thickness and quality of tunnel oxide, and storage field. To this purposes, this CM can be extended by including a voltage-controlled current source implementing the empirical SILC expression proposed in [16].

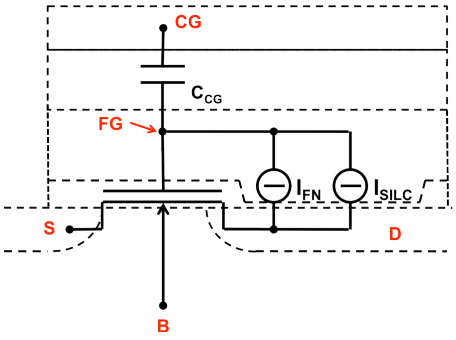

Figure 7. The CM of an EEPROM memory cell extended to simulate SILC-induced EEPROM reliability degradation by including the current generator, I_{SILC}, implementing an analytical SILC formula.

As shown in Fig. 7 for an EEPROM cell, the SILC current generator is connected between the drain and the floating gate since in this region SILC is much larger due to the thinner thickness (~7nm) of tunnel oxide compared to the gate oxide one (~20nm). To correlate the tunnel oxide degradation induced by high-field stress to the P/E cycles, N_C, and the P/E bias conditions, the charge exchanged during a P/E cycle and the P/E current density flowing through the tunnel oxide have been evaluated using the model. Read disturb simulations have demonstrated that SILC is not a concern for EEPROM cells considered, since

the oxide field and the time involved in read operations are too low to induce significant FG charge variations, i.e. V_T modification regardless the SILC magnitude. On the contrary, data retention losses are strongly affected by Stress Induced Leakage Current, as predicted by reliability simulations and also confirmed by experimental data. Results are shown in Fig. 8 [16].

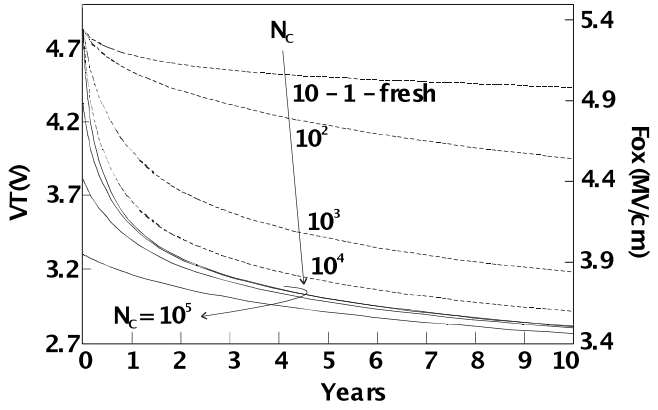

Figure 8. V_T decay for a single EEPROM cell left unbiased at room temperature on increasing the cycle number N_C (dashed lines). The oxide field is also indicated. All dashed lines start from the same initial storage field. For $N_C=10^5$, different initial storage fields have been assumed, and the decay curves are shown by solid lines [16].

This CM has been extended to simulate NAND Flash memories. NAND Flash memories are organized in strings of 32 cells, connected by two selector transistors to drain (DSL) and source (SSL) lines. The building block of the macromodel we proposed in order to reproduce real device behavior is shown in Fig. 9 [17], which shows three NAND cell strings. The central string is the one to be analyzed, whereas the two lateral ones are inserted to account for coupling capacitive effects. Each cell is composed of a stacked oxide/nitride/oxide (ONO) capacitor and an equivalent transistor, as shown in Fig. 9. Capacitive coupling effects are taken into account by an array of capacitors modeling the floating gate (FG) couplings to adjacent cells on the same WLs and BLs (Cx and Cy, respectively, in Fig. 9). To model P/E operations, voltage-controlled current sources are inserted between FG and S, B, and D nodes to model tunneling currents flowing through the bottom oxide during program and erase operations [7]. To demonstrate the model application to future Flash memory generations, we tested its capabilities to reproduce the crosstalk between adjacent cells. Fig. 10 shows the measured and simulated threshold voltage shift ΔV_{TH} of a cell in the string, due to capacitive coupling between adjacent FGs on the same BL (y-direction) and WL (x-direction). As shown, the model accurately reproduces both TCAD simulations and measurements, demonstrating that it is a valuable tool for circuit simulations of present and future NAND Flash memories, essential to optimize the program algorithms of multilevel design where sensing margins become a strong concern.

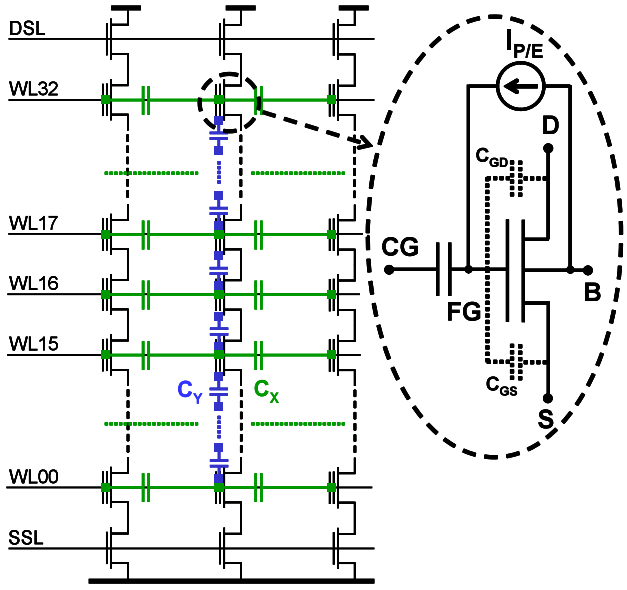

Figure 9. Schematic of basic structure of the NAND Flash string macro-model. Coupling capacitance between adjacent cells in the same word line (Cx) and bit line (Cy) are shown. The explosion shows the description of a single cell in terms of the CM of the FG device [17].

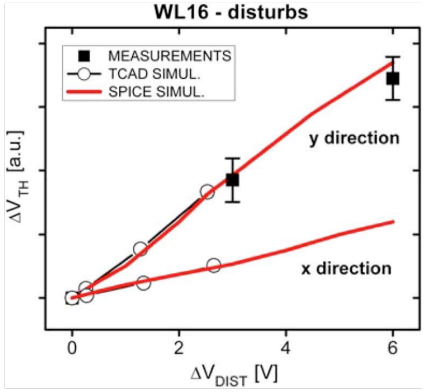

Figure 10. (Full symbols) V_T shift measured on the sixteenth cell in the string manufactured in 5-nm technology as a function of the threshold voltage shift of the adjacent disturbing cell (ΔV_{DIST}) in the same BL (x-direction) and WL (y-direction). Empty symbols represent TCAD simulation results, whereas the solid line represents the Spice simulation results obtained with the model we have developed [17].

5 EXTENSION TO CT DEVICES

This modeling approach can be adopted also to invesitgate the operations of advanced CT memory devices, like TANOS [18]. A TANOS SPICE-like compact model allows designers to investigate ancillary circuit solution and

optimized algorithms for read and program/erase operations, especially when considering multi-level memories, having very tight margins between different threshold voltage levels [19].

Although the charge storage mechanisms exploited by FG and TANOS memory cells are different, the main physical mechanisms governing their operation are similar and can be described using the same formalism. Thus, a SPICE-like TANOS model can be developed adopting the same circuital approach. We have developed an analytical model of Program/Erase operations in TANOS devices [20], [21] which can be used to implement a compact SPICE-like model of these memory devices [19].

The model is schematically illustrated in Fig. 11. In this case the Gate of the MOS transistor corresponds to the charge centroid of the TANOS device and the capacitor C_{EQ} is connected from there to the metal gate.

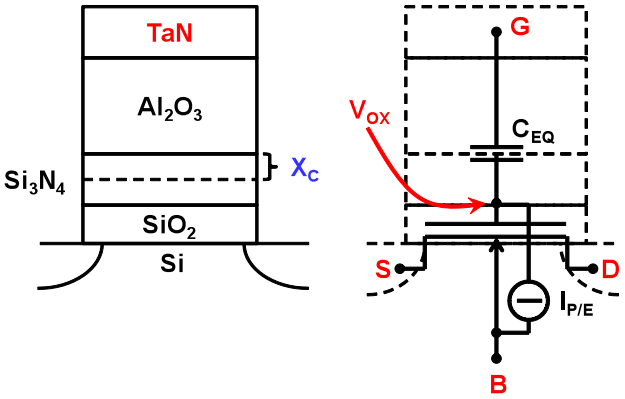

Figure 11. Schematic representation of the compact TANOS model [19].

The model can be used to simulate the TANOS devices in DC conditions (e.g. read operation), provided that the gate of the MOSFET reproducing the I_D characteristics of the TANOS transistor is initialized to its correct voltage.

Fig. 12 shows the threshold voltage shift during program measured (symbols) and simulated (lines) using the compact TANOS model. Simulations are in a reasonable agreement with the experimental data. The threshold voltage shift measured during erase is compared to the simulations performed using the compact model in Fig. 13. The agreement with the experimental data is excellent over the whole time and erase voltage range considered, proving that the assumptions adopted model correctly the physical mechanisms involved in the TANOS operation.

As expected, accuracy is low at high program times (since the electron emission current has been purposely neglected during program, see Fig. 12 [19]) and the saturation occurring at high erase times is not reproduced (since the gate current has been purposely neglected during erase, see Fig. 13 [19]).

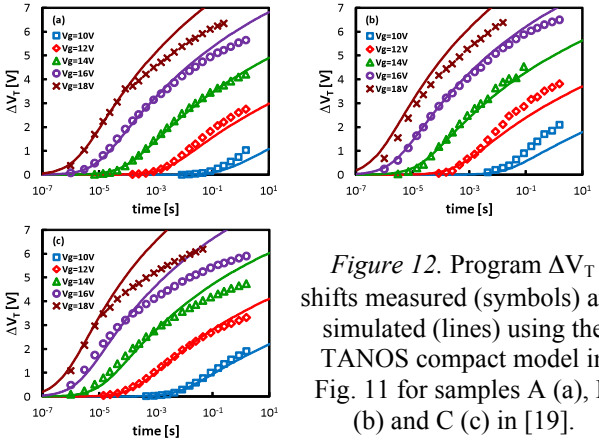

Figure 12. Program ΔV_T shifts measured (symbols) and simulated (lines) using the TANOS compact model in Fig. 11 for samples A (a), B (b) and C (c) in [19].

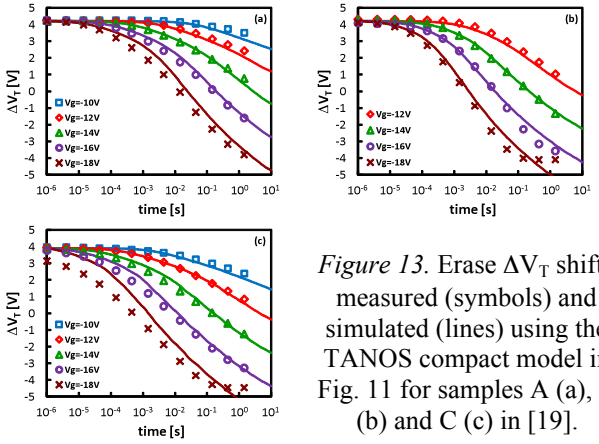

Figure 13. Erase ΔV_T shifts measured (symbols) and simulated (lines) using the TANOS compact model in Fig. 11 for samples A (a), B (b) and C (c) in [19].

6 CONCLUSIONS

We have developed a new FG memory cell model suitable for circuit simulation that is capable to correctly simulate a FG-based memory cell in every bias condition. By adding some appropriate voltage/controlled current source, we were able to describe correctly transient behaviors during program and erase operations. This model can be easily included in circuit simulators and it does not increase simulation time. If degradation mechanisms can be described analytically, they can be included and thus reliability simulations and predictions are also possible.

The same framework can be adopted to simulate TANOS devices, provided an analytical description of program and erase currents.

Results obtained in these last ten years confirm the validity of this approach. This simple SPICE-like model framework can be updated by adding appropriate current generators to describe many different effects, linking the fundamental physical mechanism description to the electrical behavior of devices in various operating conditions.

REFERENCES

[1] L. Larcher, P. Pavan, F. Gattel, L. Albani, A. Marmiroli, "A New Compact Model of Floating Gate Non-Volatile Memory Cells", Proc. International Conference on Modeling and Simulations of Microsystems, MSM 2001, March 19/21 2001, Hilton Oceanfront Resort, Hilton Head, S.C. (USA)

[2] P. Pavan, R. Bez, P. Olivo, E. Zanoni, "Flash Memory Cells - An Overview", Proceedings of IEEE 85, pp. 1248-1271, 1997.

[3] C. Lee, et al, "A 32-Gb MLC NAND Flash Memory With Vth Endurance Enhancing Schemes in 32 nm CMOS", IEEE J. of Solid-State Circuits 46, pp. 97-106, 2011.

[4] Y.H. Kang, S. Hong, "A Simple Flash Memory Cell model for Transient Circuit Simulation", Electron Device Lett. 26, pp. 563-565, 2005.

[5] P. Pavan, L. Larcher, A. Marmiroli, «Floating Gate Devices: Operation and Compact Modeling », Kluwer Academic Publishers, Boston, ISBN 1-4020-7731-9, Feb. 2004.

[6] P. Pavan, L. Larcher, A. Marmiroli, «Floating Gate Devices: Operation and Compact Modeling», Technical Proceedings Nanotech 2004, Workshop on Compact Modeling, Boston MA, March 7-11, 2004, ISBN 0-9728422-6-8.

[7] L. Larcher, P. Pavan, S. Pietri, L. Albani, A. Marmiroli, "A New Compact DC Model of Floating Gate Memory Cells without Capactitive Coupling Coefficients," IEEE Trans. Electron Devices 49, pp. 301-307, 2002.

[8] W. Liu, X. Jin, K. M. Cao, an C. Hu, "BSIM4.0.0 MOSFET Model User's Manual," 2000.

[9] www.epfl.ch/ekv/

[10] G. Gildenblat, et al, Proc. 1997 International Semiconductor Device Research Symposium, p. 33, 1997.

[11] http://home.hiroshima-u.ac.jp/usdl/HiSIM.html

[12] L. Larcher, P. Pavan, L. Albani, T. Ghilardi, D. Cantarelli, A. Marmiroli, "Bias and W/L dependence of capacitive coupling coefficients in Floating Gate Memory cells" IEEE Trans. Electron Devices 48, pp. 2081-2089, 2001.

[13] P. Pavan, L. Larcher, M. Cuozzo, P. Zuliani, A. Conte, "A Complete Model for E^2PROM Memory Cells for Circuit Simulations," IEEE Trans. on CAD, Vol. 22, N.8, pp. 1072-1079, 2003.

[14] P. Pavan, L. Larcher, A. Marmiroli, "Floating Gate Devices: Operation and Compact Modeling", Technical Proceedings Nanotech 2004, Workshop on Compact Modeling, Boston MA, March 7-11, 2004, ISBN 0-9728422-6-8.

[15] L. Larcher, S. Bertulu, P. Pavan, "SILC effects on E^2PROM memory cell reliability," IEEE Transactions on Device and Materials Reliability, 2, pp. 13-18, 2002.

[16] L. Larcher, A. Paccagnella, and G. Ghidini, "A model of the stress-in- duced leakage current in gate oxides," IEEE Trans. Electron Devices 48, pp. 285-288, 2001.

[17] L. Larcher, A. Padovani, P. Pavan, P. Fantini, A. Calderoni, A. Mauri, A Benvenuti, "Modeling nand Flash Memories for IC Design", Electron Device Lett. 29, pp. 1152-1154, 2008.

[18] C. H. Lee, K. In Choi, M. K. Cho, Y. H. Song, K. C. Park, and K. Kim, "A novel SONOS structure of SiO_2-SiN-Al_2O_3 with TaN metal gate for multi-giga bit Flash memories," in IEDM Tech. Dig., pp. 613-616, 2003.

[19] A. Padovani, L. Larcher, P.Pavan, "Compact Modeling of TANOS Program/Erase Operations for SPICE-like Circuit Simulations", submitted to Microelectronics Journal, 2011.

[20] A. Padovani, L. Larcher, D. Heh, and G. Bersuker, "Modeling TANOS memory program transients to investigate charge trapping dynamics," IEEE Electron Device Lett. 30, pp. 882-884, 2009.

[21] A. Padovani, L. Larcher, D. Heh, G. Bersuker, V. Della Marca and P. Pavan, "Temperature effects on metal-alumina-nitride-oxide-silicon memory operations," Appl. Phys. Lett. 96, p. 223505, 2010.