Advantages of the FinFET architecture in SONOS and Nanocrystal memory devices

S. Lombardo,[a] C. Gerardi,[b] L. Breuil,[c] C. Jahan,[d] L. Perniola,[d] G. Cina,[b] D. Corso,[a]
E. Tripiciano,[b] V. Ancarani ,[b] G. Iannaccone,[e] G. Iacono,[b] C. Bongiorno,[a] J. Razafindramora,[d]
C. Garozzo,[a] P. Barbera,[a] E. Nowak,[d] R. Puglisi,[a] G. A. Costa,[b] C. Coccorese,[b]
M. Vecchio,[b] E. Rimini,[a] J. Van Houdt,[c] B. De Salvo,[d] and M. Melanotte [b]

[a] CNR-IMM, Stradale Primosole, 50, 95121 Catania, Italy, Tel.: +39 095 5968223, e-mail: salvatore.lombardo@imm.cnr.it;
[b] STMicroelectronics, FMG , Catania, Italy; [c] IMEC, Leuven, Belgium;
[d] CEA-LETI, Grenoble, France [e] University of Pisa, Italy

Abstract

Double-gate and tri-gate FinFET type memories with nitride (SONOS-like) or Si nanocrystals storage with minimum feature sizes of 10 nm were realized. Strong performance advantages in program / erase characteristics and reliability deeply linked to the FinFET architecture are demonstrated.

Introduction

Memories with discrete storage media like nanocrystals or nitride are a promising alternative to floating gate (FG) devices given the improved resistance to SILC defects and reduced parasitic capacitive couplings due to the FG [1]. Moreover, the combination to a FinFET architecture allow advantages in terms of suppression of short channel effects [2]. In this paper we show for the first time that the double or tri-gate structures allow a large improvement in terms of VT window, justified by the distribution of the electric field in the gate dielectric stack. This allows to increase tunnel oxide thickness with a resulting better retention.

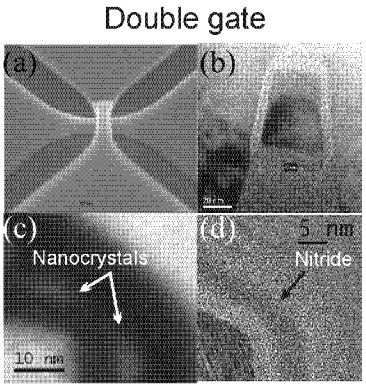

Fig. 1. TEM and SEM micrographs in plan (a) or in cross-sectional view of the Fin (b-d) of double-gate FINFLASH devices with nanocrystals (c) and nitride (d) trapping media.

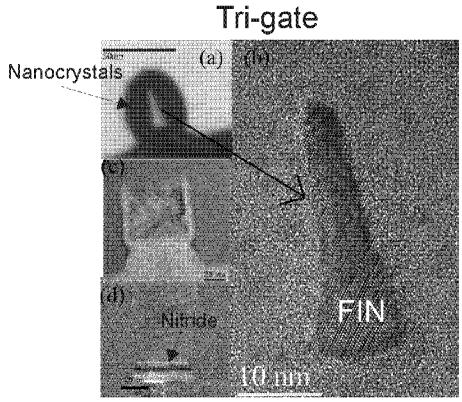

Fig. 2. Cross-section TEM micrographs of tri-gate FINFLASH devices with nanocrystals (a, b) and nitride (c, d) trapping media. (a-c) are sections orthogonal to the fin, (d) is along the channel length.

Device realization

FinFLASH devices were realized on SOI wafers by DUV lithography, except for the shortest devices in which both fin and gate mask were realized by e-beam lithography. Tri-gate and double gate structures were realized with both Si nanocrystals and silicon nitride layers as trapping media. Tunnel oxides were in the range from 3 to 6 nm. In nitride memories (SONOS-type) control dielectrics were HTO layers of thicknesses from 5 to 8 nm, while in nanocrystal devices ONO stacks of EOT from 10 to 15 nm were used.

Results and Discussion

Figs. 1 and 2 show TEM and SEM micrographs of the double gate and tri-gate devices structures, respectively. Channel lengths (L_{CH}) and widths (W) as narrow as 30 nm and 10 nm were realized (Fig. 2). These devices show regions with high curvature in the dielectrics, in correspondence with the fin corners.

A. Improvements on Short Channel Effects (SCE)

Though the overall thickness of the dielectric stack, including tunnel and control oxide and trapping medium is above 15

nm these devices show excellent performances in terms of SCEs. Subthreshold slopes recorded at V_{DS} = 1 V (Fig. 3) and DIBL resistance (Fig. 4) remain remarkable even for very short channel lengths and improve as the fin height or width are decreased. The ON current measured at V_G-V_T = 1.5 V and V_{DS} = 1 V (Fig. 5) are above 20 μA, largely exceeding the values of conventional Flash devices. Tri-gate devices can further improve I_{ON}, though this advantage is larger for higher widths.

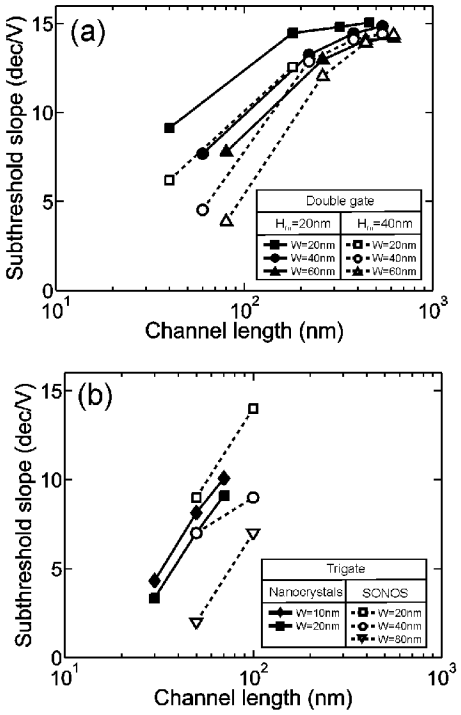

Fig. 3. Subthreshold slope measured at V_{DS} = 1 V as a function of channel length in double and tri-gate FINFLASH devices, in all cases with overall dielectric stack thickness above 15 nm. As the W or the fin height decrease, the slope improves well above the values achievable in standard planar Flash devices. In (b) fin heights range from 25 to 60 nm.

In double- and tri-gate FINFLASH structures the current in inversion is accumulated in correspondence to the fin corners, as evidenced by TCAD simulations, as in the example of Fig. 6. This explains why I_{ON} does not depend on fin height in double gate devices (Fig. 5).

B. Program / Erase Characteristics

Concerning the program / erase (P / E) characteristics these devices exhibit excellent performances. Fig. 7 shows the curves of CHE programming and FN tunneling erase, both very efficient at remarkably low gate and drain voltage. Note the ability of CHE programming by using "warm carriers" at V_D values lower than 3.2 V, consistent with [3]. Fig. 8 shows the P / E curves by FN tunneling of a double-gate FinFLASH cell with a 4.5 nm tunnel oxide. Note in particular that erase saturation takes place at remarkably higher voltages compared to standard planar SONOS devices [4]. This is an important feature since the erase saturation phenomenon in SONOS forces the use of reduced tunnel oxide thickness, with an associated retention limit at high temperature. This has been a strong motivation to the development of innovative concepts such as the NROM and the TANOS.

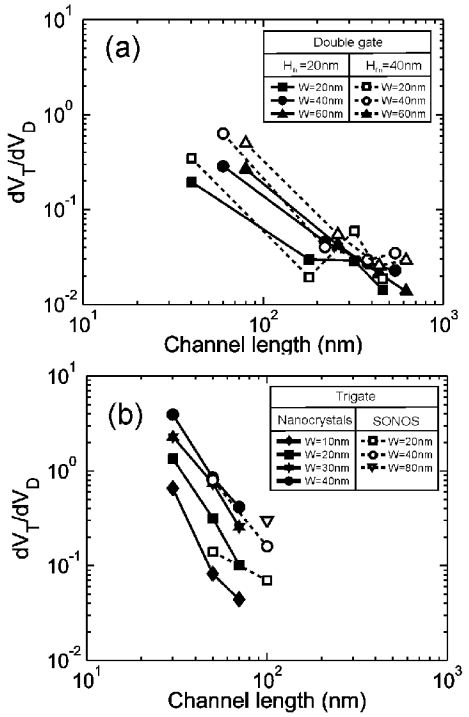

Fig. 4. DIBL measured with V_{DS} = from 0.1 V to 1 V as a function of channel length in double and tri-gate FINFLASH devices. As the W or the fin height decrease the DIBL improves dramatically, well above the values achievable in standard planar Flash devices.

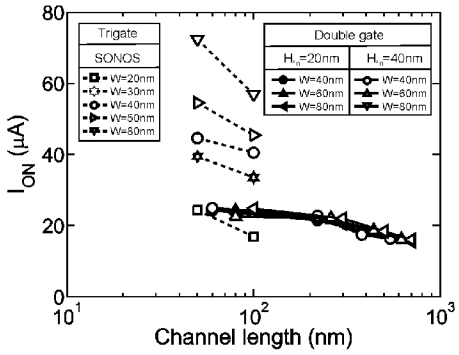

Fig. 5. ON current measured at V_{DS} = 1 V and V_G-V_T = 1.5 V as a function of channel length in double-gate and tri-gate FINFLASH devices. In double-gate devices I_{ON} is almost independent on fin height, indicating that the inversion charge is accumulated at the fin corners, as confirmed by TCAD simulations. In small W tri-gate devices the behavior is similar though I_{ON}, contrarily to the double-gate case, depends on W. However, as W tends to 0 the I_{ON} values tend to coincide with the double-gate case.

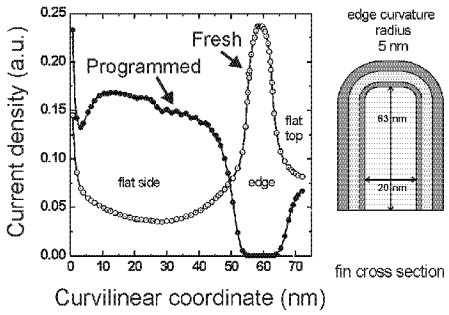

Fig. 6. Simulated current density in a tri-gate structure in the central cross section as a function of the curvilinear coordinate at the silicon-tunnel oxide interface (the origin corresponds to the bottom-left corner). For a gate overdrive of 1.5V, current in the fresh cell accumulates at the corners, whereas the current in the programmed cell only flows in the flat regions, because is the charge over the corners the most responsive to P/E operation.

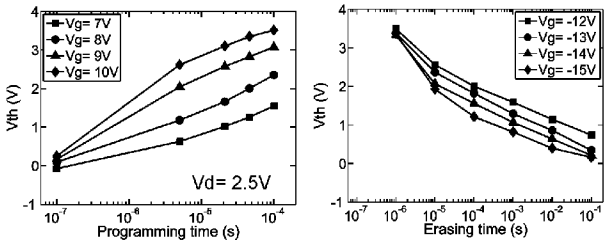

Fig. 7. Curves of programming by channel hot electron (CHE) injection and Fowler-Nordheim (FN) erase in a FINFLASH cell with Si nanocrystals, tunnel oxide thickness of 5 nm, HTO control of 8 nm, L_{CH} = 30 nm, W = 10 nm, fin height = 25 nm. Note the low VD value during CHE

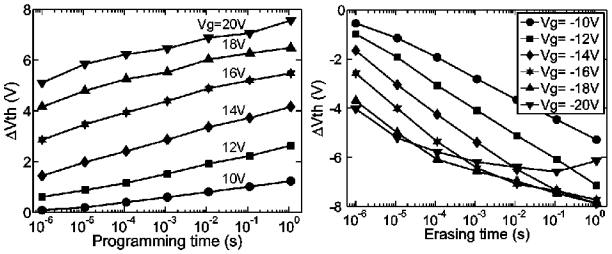

Fig. 8. Program / erase curves by FN tunneling in a SONOS-type FINFLASH cell with an ONO of 4.5 nm bottom oxide, 3.2 nm nitride, and 8 nm top oxide thickness. Erase saturation takes place only at very high (in absolute value) voltages, indicating a very effective suppression of electron back-injection from the gate.

A possible explanation for this large improvement is the following: Fig. 9 shows calculations of the electric field distribution in the gate stack under erase conditions. In particular the figure shows the electric field enhancement factor defined as the ratio of the electric field in the region of maximum curvature to the electric field in the flat regions as a function of the vertical coordinate along the gate stack, for different values of the curvature radius at the interface between the silicon fin and the tunnel oxide. Data are obtained from an analytical model, considering a charge-neutral nitride layer with 3 nm tunnel oxide, 5 nm nitride layer, 5 nm control oxide and confirmed by TCAD simulations. As can be seen, there is an effective field enhancement in the tunnel oxide, and a field suppression in the control oxide, whose strength increases with edge curvature. The high field in the part closer to the channel implies efficient programming and erase at the fin corners, while the much lower field at the top part implies suppression of the electron back-injection from the gate. These circumstances can explain the impressive P / E efficiency of the FinFLASH structures and the particularly good robustness to erase saturation.

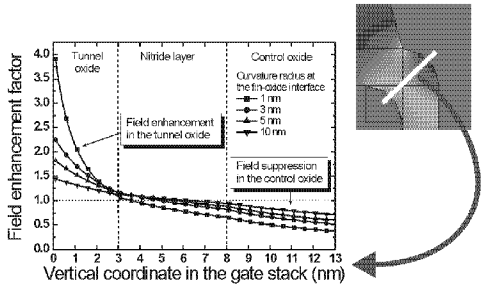

Fig. 9. Electric field enhancement factor defined as the ratio of the electric field in the region of maximum curvature to the electric field in the flat regions as a function of the vertical coordinate along the gate stack, for different values of the curvature radius at the interface between the silicon fin and the tunnel oxide. An effective field enhancement in the tunnel oxide and a field suppression in the control oxide are evident. These increase with edge curvature.

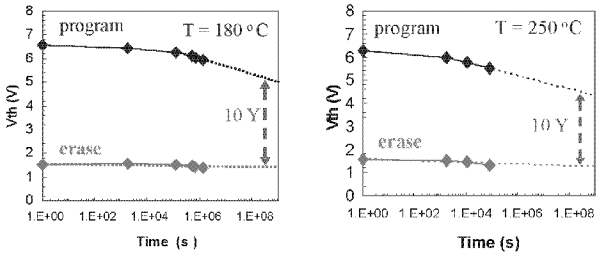

Fig. 10. Storage data at 180 and 250 °C in SONOS type FINFLASH cells as the one of Fig. 8. V_T window of about 3 V extrapolated at ten years at 250 °C are observed for the first time in this type of devices.

C. Retention and Cycling

This P / E characteristics are coupled with excellent performances for retention at high temperature, showing reliability well in excess of ten years at temperatures as high as 250 °C (Fig. 10), with extrapolated residual window of about 3 V. Concerning the robustness to extensive P / E cycling, Fig. 11 demonstrates excellent robustness up to 1E5 cycles, with a remarkably high VT window of about 8 V, in spite of the increase of field at the corners.

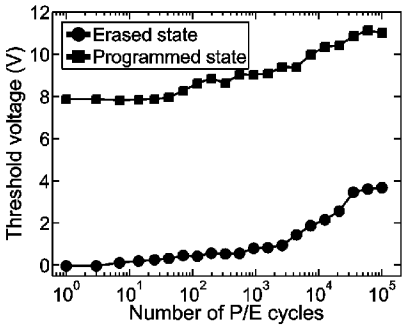

Fig. 11. FN P / E cycling of a SONOS type FINFLASH cell as those of Figs. 8 and 10. Programming and erase are obtained at V_G = +16 V t = 10 ms and V_G = -15 V t = 10 ms, respectively.

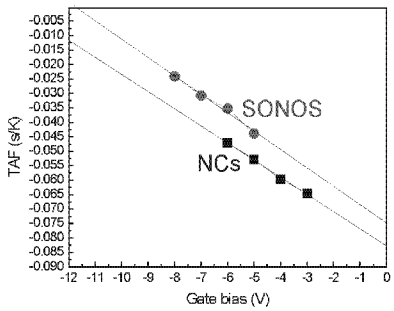

Fig. 12. Temperature acceleration factors in s / K as a function of accelerating de-trapping gate voltage in SONOS type and nanocrystal FINFLASH devices with an ONO stack 4.5 / 3 / 8 nm for the SONOS-type device, and a tunnel oxide of 5 nm and control ONO with EOT of 14 nm, for the nanocrystal device. Note that nanocrystals appear much more robust compared to SONOS-like devices with respect to gate stress.

Concerning the comparison between nitride and nanocrystal devices, it is known that nanocrystals imply a worsening of V_T window dispersion due to random fluctuations in the number of trapping dots, but this problem may be solved by increasing dot density or by dot self-organization [5]. Nanocrystals can, in fact, allow important advantages on cell performances. Fig. 12 demonstrates their better retention at low voltage, as shown by the comparison of the temperature acceleration factor, attributed to the suppression of Poole-Frenkel conduction, which on the contrary occurs for nitride traps [6]. Moreover since the better gate capacitive coupling of nanocrystals allow to decrease the P / E voltage, robustness to cycling improves (Fig. 13), in excess to 1E6 cycles. Note that though a ONO control dielectric is used, differently with respect to planar nanocrystal memories which require HTO [7], no irreversible charge trapping in the ONO layer takes place.

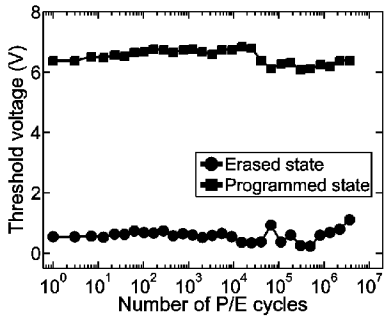

Fig. 13. FN P / E cycling of a nanocrystal FINFLASH cell with tunnel oxide of 4 nm and ONO of 4.5/5/9.5 nm. Programming and erase are obtained at V_G = +18 V t = 100 μs and V_G = -18 V t = 10 ms, respectively.

Summary

In summary, we show for the first time that in addition to the advantages of suppression of SCEs, the FINFLASH architecture allows to dramatically improve the V_T window (up to 8 V with 1E5 cycles) and the erase saturation problem, allowing thicker tunnel oxides with associated excellent retention, demonstrated up to 250 °C.

We gratefully acknowledge the support on TCAD of E. Guichard and A. Hoessinger of SILVACO. Work funded by the FP6 EU Project FINFLASH.

References

(1) B. De Salvo, C. Gerardi, S. Lombardo, T. Baron, L. Perniola, D. Mariolle, P. Mur, A. Toffoli, M. Gely, M .N. Semeria, S. Deleonibus, G. Ammendola, V. Ancarani, M. Melanotte, R. Bez, L. Baldi, D. Corso, I. Crupi, R. A. Puglisi, G. Nicotra, E. Rimini, F. Mazen, G. Ghibaudo, G. Pananakakis, C. Monzio Compagnoni, D. Ielmini, A. Lacaita, A. Spinelli, Y. M. Wan, K .van der Jeugd, IEEE-IEDM 2003.

(2) H. S. P. Wong, IBM J. Res. Dev. 2002.

(3) S.D. Suk, K.H. Yeo, K.H. Cho, M. Li, Y.Y. Yeoh, K.-H. Hong, S.-H. Kim, Y.-H. Koh, S. Jung, W.J. Jang, D.-W. Kim, D. Park and B.-I. Ryu, IEEE-VLSI Symp. (2007).

(4) R. van Schaijk, M. van Duuren, W. Y. Mei, K. van der Jeugd, A. Rothschild, M. Demand, Micr. Eng. 72, (2004).

(5) K.W. Guarini, C.T. Black, Y. Zhang, I.V. Babich, E. M. Sikorski, L. M. Gignac, IEEE-IEDM 2003.

(6) Jong Jin Lee, Se Hoon Lee, Heesoon Chae, Byung Yong Choi, Suk-Kang Sung, Seok Pil Kim, Eun Suk Cho, Choong Ho Lee, and Donggun Park, IEEE-IRPS, 530 (2006).

(7) R.F. Steimle, R. Muralidhar, R. Rao, M. Sadd 1, C.T. Swift, J. Yater, B. Hradsky, S. Straub, H. Gasquet, L. Vishnubhotla, E.J. Prinz, T. Merchant, B. Acred, K. Chang, B.E. White Jr., Micr. Rel. 47, 585 (2007).