Random Telegraph Noise-Induced Sensitivity of Data Retention to Cell Position in the Programmed Distribution of NAND Flash Memory Arrays

Davide Resnati, Christian Monzio Compagnoni, Senior Member, IEEE, Giovanni M. Paolucci, Carmine Miccoli, Alessandro S. Spinelli, Senior Member, IEEE, Andrea L. Lacaita, Fellow, IEEE, Angelo Visconti, Member, IEEE, and Akira Goda

I. INTRODUCTION

COMING from the uncontrollable capture and emission of single electrons in tunnel-oxide traps, random telegraph noise (RTN) is today one of the most relevant reliability issues for Flash memories [1]–[5]. RTN limits, in fact, the possibility of achieving and keeping narrow threshold-voltage ($V_T$) distributions in Flash arrays [6], [7], constraining the number of bits per cell that can be safely stored in multi-level devices. In particular, RTN broadens the $V_T$ distribution resulting from a program-and-verify operation [8] both in its upper and in its lower part [7], [9], with the broadening growing as time elapses during data retention.

In this letter, we show that RTN introduces a non-negligible contribution to the sensitivity of the $V_T$ transients during post-cycling data retention of nanoscale NAND Flash cells to cell position in the programmed array distribution. This RTN-induced sensitivity contributes to reducing the $V_T$ loss of cells in the lower part of the programmed distribution and to increasing the $V_T$ loss of cells in the upper part of the distribution. Results, which are not related to any fundamental data retention issue of cells in different parts of the array distribution, are explained as an artifact of the program-and-verify operation trying to compact the array $V_T$ distribution in the presence of RTN.

II. EXPERIMENTAL RESULTS

In order to investigate the sensitivity of the data retention $V_T$ transient of Flash cells to their position in the programmed array distribution, we moved two pages of a heavily cycled NAND array made with our 16 nm planar technology [10] to slightly different programming levels, as shown in Fig. 1. A program-and-verify algorithm [8] was used to fit all the cells of the two pages into two narrow and partially overlapped $V_T$ intervals, allowing to investigate the behavior of cells with the same initial $V_T$ ($V_{T0}$ from here on) but with different positions in the programmed distribution. For instance, cells in group A of page 1 and in group B of page 2 in Fig. 1 have the same $V_{T0}$ (shadowed region in the figure), but are, respectively, in the top and in the bottom part of their page distribution.
After page programming, the memory array underwent a bake experiment at $T_B = 110 \, ^\circ C$, which was periodically stopped to gather the $V_T$ of all the cells to monitor their high-temperature data retention (all the read operations were done at room temperature). Fig. 2 shows the average $V_T$ shift ($\langle \Delta V_T \rangle$) after a bake time $t_B = 1 \, h$ of cells with the same $V_{T0}$ in page 1 and in page 2. Results reveal, first of all, that cells display an average $V_T$ loss during bake (negative $\langle \Delta V_T \rangle$), coming from electron detrapping from their tunnel oxide [11]–[14]. Owing to the sensitivity of this process to the electric field in the tunnel oxide during the bake period [13], [15], an increase of the average $V_T$ loss of the entire page (see the open circles in the figure) appears when moving from page 1 to page 2. This average $V_T$ loss gives, however, a $\langle \Delta V_T \rangle$ dependence on $V_{T0}$ which is weaker than that followed by the cells inside each of the two pages, as clearly appearing from the comparison of the blue and of the red curve in Fig. 2 with the black dashed line (this represents a perfectly repeatable phenomenon observed in all of our data retention results, for whatever page programming level). As a consequence of this, cells in group $B$ of page 2 display a lower $\langle \Delta V_T \rangle$ than cells in group $A$ of page 1, despite the fact that they were initially at the same $V_{T0}$. Fig. 3 shows that this is true even for increasing $t_B$ and, therefore, in the presence of higher $V_T$ losses during the data retention transient.

In order to investigate the origin of the different data retention transient of cells in group $A$ and in group $B$, Fig. 4 shows the cumulative distribution of $\Delta V_T$ for these cells at $t_B = 1 \, h$ and $300 \, h$. Results reveal that, besides the displacement of their $\langle \Delta V_T \rangle$, the two groups of cells display far more relevant differences in the tail regions of their $\Delta V_T$ distribution. In particular, the upper tails of the two distributions reveal, first of all, the possibility that some cells display positive $\Delta V_T$ and, moreover, that this is more likely to occur for cells in group $B$. The comparison of the lower tails of the distributions reveals, instead, that cells in group $A$ are more prone to reducing their $V_T$ in the experiment than those in group $B$.

Since the tails of the data retention $\Delta V_T$ distribution may be significantly affected by RTN [14], [15], we explored the possibility that this noise source represents also the origin of the increased sensitivity of data retention to cell position in the page $V_{T0}$ distribution. To this aim, we repeated the data retention experiment with a lower $T_B$. Note, in fact, that the reduction of $T_B$ largely decreases charge detrapping from the tunnel oxide [12], [13], [15], while having only a minor impact on the statistics of RTN instabilities in the array [16]. Fig. 5 shows the statistical results for $\Delta V_T$ gathered on cells in the bottom part (group $B$) and in the top part (group $A'$) of the $V_{T0}$ distribution of a programmed page when $T_B = 55 \, ^\circ C$ and $t_B = 1 \, h$. The inset shows that the $\langle \Delta V_T \rangle$ of the entire page is $\simeq 0$, confirming that charge detrapping is almost negligible in this experiment and that RTN is the only mechanism impacting cell $V_T$. Despite this, however, cells display the same $\langle \Delta V_T \rangle$ trend inside the $V_{T0}$ distribution shown in Fig. 2, with those in the bottom part of the distribution having now positive $\langle \Delta V_T \rangle$ and those in the top part
increased their
VT
the first read operation on the page (Fig. 6b), with the former sensitivity of
/Δ1
These results reveal that RTN introduces a contribution to the
of cells previously observed in Fig. 4 also appear in Fig. 5. These results reveal that RTN introduces a contribution to the
VT
to cell position in the VT0 distribution, explaining the results of Figs. 2-3-4.

III. DISCUSSION
Data in Fig. 5 revealed that the RTN behaviors of cells in the bottom part and in the top part of a programmed page distribution display non negligible differences. This result can be explained with the help of Fig. 6, showing a schematic view of a page distribution considering the VT determining the end of the program-and-verify operation (a) and those resulting from the first read operation on the array (b). The program-and-verify algorithm stops programming each cell when it overcomes a selected level VPV and this should result, neglecting program noise contributions [17], in a VT distribution included between VPV and VPV + VS, where VS is the programming step amplitude [8]. However, immediately after the end of the program operation, RTN starts affecting cell VT, with the possibility for some cells to move below VPV (see red arrows in Fig. 6a) and for others to move above VPV + VS (blue arrows). This introduces a tail in the lower and in the upper part of the VT0 distribution resulting from the first read operation on the page (Fig. 6b), with the former and the latter made of cells that, respectively, reduced and increased their VT due to RTN. As a consequence, when cells are monitored during data retention using this distribution as a reference, those in its bottom and in its upper part will, respectively, most likely increase and decrease their VT due to RTN, owing to fluctuations towards their initial VT level. Although the average VT of the entire page is ≃ 0, this introduces a non zero ⟨ΔVT⟩ for cells far from the average VT0, as shown in Fig. 5 and explaining the results of Fig. 2. Note, however, that this does not give rise to a narrowing of the VT distribution as data retention proceeds, since cells moving from the sides towards the center of the distribution are replaced by more cells moving in the opposite direction.

IV. CONCLUSION
In this letter, we have highlighted a contribution given by RTN on the data retention ⟨ΔVT⟩ sensitivity to cell position in the programmed array distribution. Results have been explained with a simple physical picture for cell movements in the VT distribution in the time stretch in-between the program-and-verify operation and the first read operation used as reference for data retention assessment.

REFERENCES