# Evidence of Widely Distributed Time Constants in the Vertical Charge Loss of 3-D Charge-Trap NAND Flash Memories

David G. Refaldi, Gerardo Malavena, *Member, IEEE*, Luca Chiavarone, Alessandro S. Spinelli, *Senior Member, IEEE*, and Christian Monzio Compagnoni, *Senior Member, IEEE*

***Abstract*— In this letter, through experimental evidence collected at room temperature and in the deep-cryogenic regime, we demonstrate that the so-called vertical charge loss from the gate stack of 3-D charge-trap NAND Flash memories is a process featuring widely distributed time constants. Results reveal that the turn-on of these time constants depends on the temperature at which the program operation is carried out. This, combined with the dependence of the time constants on the temperature at which vertical charge loss is monitored, gives rise to an apparent activation energy of the cell threshold-voltage transient during data retention that is close to zero when data retention and program occur at the same temperature. This phenomenology must be carefully taken into account when trying to extend the working temperature of NAND Flash memories down to the deep-cryogenic regime.**

***Index Terms*— NAND flash memory, cryogenic temperature regime, semiconductor device reliability.**

## I. Introduction

THANKS to its cost effectiveness, storing charge in a dielectric layer with a high density of microscopic defects has become the elective solution for 3-D NAND Flash memories [1], [2], [3]. This so-called charge-trap storage approach, however, brought in some relevant changes in the reliability physics of the technology with respect to floating-gate storage. The possibility for the charge stored in a memory cell to laterally migrate along the charge-trap layer during data retention [4], [5], [6] and nonuniform charge storage over cell channel [7] are just examples of physical phenomena which play a specific key role in the reliability of charge-trap cells.

In this letter, we experimentally address the charge loss from the gate stack to the channel, commonly referred to as *vertical charge loss* [8], [9], [10], [11], of 3-D charge-trap NAND Flash memory cells at temperature ($T$) equal to 300K and 15K. Data provide clear evidence that this process features a broad spectrum of time constants. The $T$ dependence of these time constants and of their turn-on gives rise to a nearly zero *apparent* activation energy for the threshold-voltage ($V_T$) transients during data retention when both Program and data retention occur at the same $T$. These results are of utmost importance to extend the working $T$ of 3-D NAND Flash memories down to the deep-cryogenic regime.

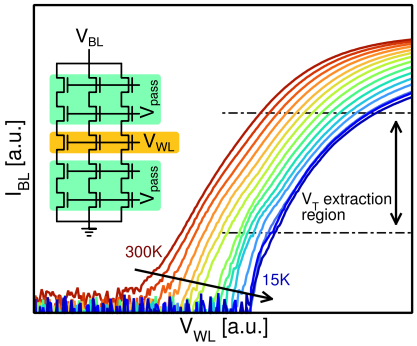


Fig. 1. $I_{BL} - V_{WL}$ trans-characteristic of the 3-D NAND Flash test element investigated in this work, at different $T$. The inset shows a schematic description of the test element and of the Read conditions. A log scale was used for $I_{BL}$.

## II. Experimental Samples, Tests and Results

The samples investigated in this work are test elements for 3-D charge-trap NAND Flash memory arrays featuring vertical polysilicon channels and 10 stacked wordlines (WLs). Cell dimensions correspond to the state-of-the-art for the 3-D NAND Flash technology. A few thousands of bitlines (BLs) in

Manuscript received 6 June 2024; revised 4 July 2024; accepted 25 July 2024. Date of publication 29 July 2024; date of current version 27 September 2024. The review of this letter was arranged by Editor H. Wu. *(Corresponding author: David G. Refaldi.)*

David G. Refaldi, Gerardo Malavena, Alessandro S. Spinelli, and Christian Monzio Compagnoni are with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: davidgianluigi.refaldi@polimi.it).

Luca Chiavarone is with the Process Research and Development, Micron Technology Inc., 20871 Vimercate, Italy.

Color versions of one or more figures in this letter are available at https://doi.org/10.1109/LED.2024.3435345.

Digital Object Identifier 10.1109/LED.2024.3435345

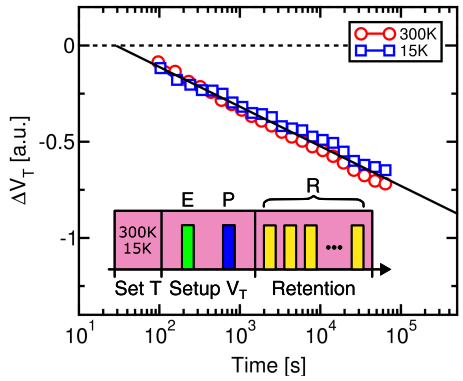


Fig. 2. Schematic description of the first experimental test used to address cell data retention (E=Erase, P=Program, R=Read) and $\Delta V_T$ transients obtained from it at $T$ = 300K and 15K. All the voltages reported in this letter are normalized to the same arbitrary constant.

the test element are short-circuited to monitor data retention on a large number of cells in parallel. To this aim, all the cells in the strings were, first, uniformly subjected to an Erase and Program operation. After increasing delay times since Program, then, the BL current ($I_{BL}$) was collected during a sweep of the voltage applied to a selected WL ($V_{WL}$), with all the other WLs under pass condition. Fig. 1 shows a schematic for the voltages applied to collect the $I_{BL}$ vs. $V_{WL}$ trans-characteristic along with a representative example of the curve obtained at different $T$ from 300K to 15K (see [12] for a description of the setup used to carry out cryogenic measurements). Cell data retention was addressed by extracting an equivalent $V_T$ as the $V_{WL}$ value leading to a constant $I_{BL}$ and monitoring its shift ($\Delta V_T$) from the first Read operation as a function of time elapsed since Program. Fig. 2 schematically shows the data retention test and the results obtained when all of its phases are carried out at the same $T$ equal to either 300K or 15K (note that nothing was changed in the Program operation at the two $T$, achieving approximately the same amount of stored charge in the cells thanks to the weak $T$ activation of Fowler-Nordheim tunneling [13]). Very similar $\Delta V_T$ transients appear from the test, suggesting a negligible $T$ activation of cell data retention, *i.e.*, an apparent activation energy close to zero. Similar results were obtained at intermediate $T$.

To delve further into the $T$ dependence of the cell data retention phenomenology, a second experimental test was devised. This is schematically shown in Fig. 3(a). After setting $T$ to 300K and performing the Erase and Program operations, data retention was monitored over three phases. During the first (Phase 1), $T$ remains at 300K and $\Delta V_T$ follows the same trend observed in Fig. 2. Then, $T$ is reduced down to 15K (Phase 2), leading to a freeze of $\Delta V_T$. Finally, $T$ is brought back to 300K (Phase 3), making $\Delta V_T$ display a rapid drop and approach the extrapolation of the trendline observed during Phase 1. From these results, the $\Delta V_T$ transient seems to display a significant $T$ activation, in striking contrast with what observed in Fig. 2. In particular, the $T$ activation of the $\Delta V_T$ transient is strong enough to make Phase 2 completely negligible. This is confirmed by the fact that, when the part of the transient corresponding to Phase 3 is replotted neglecting

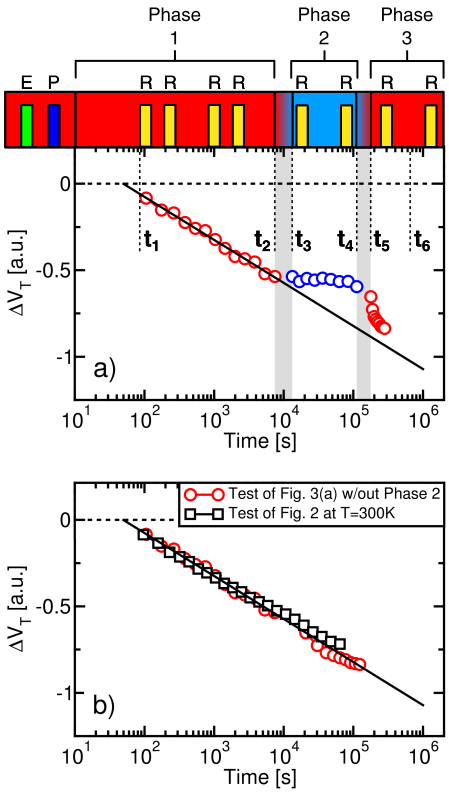


Fig. 3. (a) Schematic description of the second experimental test used to address cell data retention and $\Delta V_T$ transient obtained from it. Note that the value of $\Delta V_T$ at instant $t_3$ was set to match that at instant $t_2$ to compensate for the $T$ dependence of the $V_T$ Read operation, as highlighted in Fig. 1. The shaded bars highlight the stretches of time needed for the change of $T$ from 300K to 15K and vice versa. (b) Comparison between the transient corresponding to $T$ = 300K in Fig. 2 (black curve) and the transient obtained from the data reported in part (a) when neglecting all the blue points of Phase 2 and substrating the duration of Phase 2 from the time associated with the red points of Phase 3 (red curve).

the time spent at $T$ = 15K, the very same curve resulting from a single test at $T$ = 300K is obtained, as shown in Fig. 3(b).

The evidence provided by Fig. 2 and Fig. 3 reveals that $\Delta V_T$ follows a quite different behavior at $T$ = 15K in the two tests, steadily decreasing over time in the former and staying freeze over Phase 2 in the latter. This is clearly shown in Fig. 4 and may appear far more puzzling when considering that the $I_{BL}$ vs. $V_{WL}$ curves corresponding to the first Read operation during Phase 2 in the test of Fig. 3 and to the first Read operation in the test of Fig. 2 are almost matched (see the inset of Fig. 4). This excludes that the different $\Delta V_T$ behavior observed at $T$ = 15K in the two tests arises from a different cell placement over the $V_T$ axis.

## III. Discussion

To explain the experimental evidence provided in Sect. II it is worth noting, first of all, that the detected $\Delta V_T$ transients arise from vertical charge loss. The impact of lateral charge migration during data retention is, in fact, minimized by the uniform programming pattern of the cells along the NAND string [5], [8]. This is further confirmed by the fact that the same results were obtained over subsequent repetitions

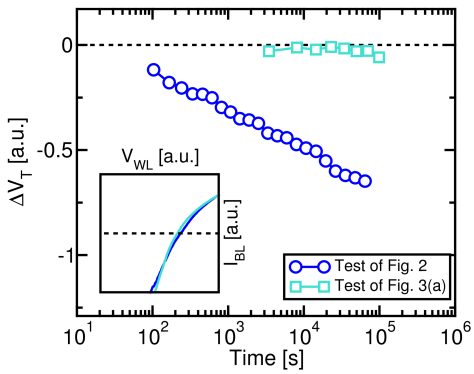


Fig. 4. Comparison of the $\Delta V_T$ transient obtained from the test in Fig. 2 at $T = 15$K with that obtained during Phase 2 of the test in Fig. 3(a) (in the latter case, $\Delta V_T$ and time are here evaluated with respect to the first Read operation carried out during Phase 2). The inset shows the $I_{BL} - V_{WL}$ curve at the first Read operation carried out at 15K in the tests (the two curves are almost matched because the $\Delta V_T$ occurring during Phase 1 of the test in Fig. 3(a) is quantitatively similar to the discrepancy in the outcome of the Program operation arising from the $T$ activation of Fowler-Nordheim tunneling).

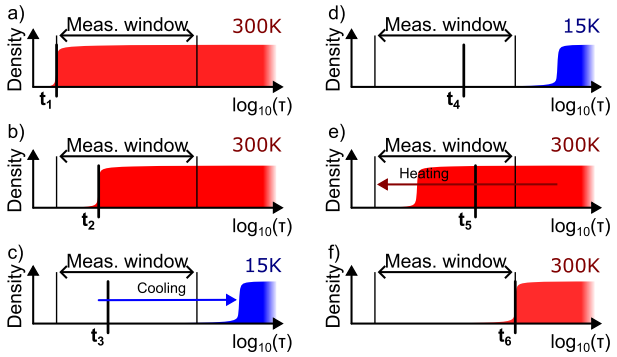


Fig. 5. Evolution of the spectrum of the emission time constants of electrons over the test in Fig. 3(a).

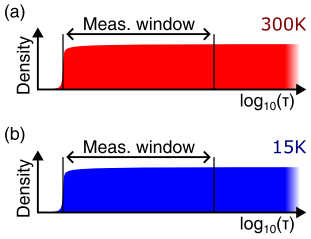


Fig. 6. Spectrum of the emission time constants of electrons at the first read operation of the test in Fig. 2 at (a) $T = 300$K and (b) $T = 15$K.

of the tests in Figs. 2-3(a) (not shown). This also excludes any possible impact on the test results of the annealing of charged microscopic defects in the tunnel dielectric stack, a phenomenology commonly referred to as charge detrapping by the NAND Flash community [14], [15], [16].

With vertical charge loss being the dominant process giving rise to $\Delta V_T$ in our experiments, the results obtained from the test in Fig. 3(a) can be explained through a physical picture based on the statistical distribution of the time constants for electron emission from the microscopic traps in the cell gate stack. This statistical distribution can be traced back to the wide spread in the spatial and energetic position of traps in charge-trap layers and, for the sake of simplicity, can be considered as uniform along the logarithmic time axis. This is schematically shown in Fig. 5(a), where the distribution of the emission time constants of electrons is reported at the time ($t_1$) corresponding to the first Read operation in the test of Fig. 3(a). At that instant, all the electrons with time constant shorter than $t_1$ have already been emitted. As time elapses, then, some other electrons are emitted from the gate stack of the memory cells, emptying the spectrum of time constants from left to right. This results in a certain $\Delta V_T$ and in the distribution of time constants shown in Fig. 5(b) when the end of Phase 1 is reached (time $t_2$). With the following reduction of $T$ down to 15K (time $t_3$), the spectrum of time constants shifts rightwards, as shown in Fig. 5(c). This is due to the $T$ activation of the time constants [17] and its outcome is that no electron has now a time constant within the observation window of data retention. As time elapses and the end of Phase 2 is reached (time $t_4$), therefore, $\Delta V_T$ remains flat and no change takes place in the spectrum of time constants, as shown in Fig. 5(d). However, when Phase 3 of the test begins and $T$ is moved back to 300K (time $t_5$), the spectrum of time constants shifts back to short times, as depicted in Fig. 5(e). As a consequence, some electrons have now an emission time constant so short to be emitted from the gate stack and make $\Delta V_T$ increase in magnitude. Finally, at the end of Phase 3 (time $t_6$), all the spectrum of time constants has been emptied up to the total observation time, as shown in Fig. 5(f).

Within the same physical picture, Fig. 6 shows that the similar $\Delta V_T$ transients obtained at $T = 300$K and 15K in the test of Fig. 2 can be attributed to a similar distribution of emission time constants at the two $T$ when the first read operation is carried out. This can be explained by assuming that the Program operation stores electrons in microscopic traps with different spatial and energetic position at $T = 300$K and 15K, with the final outcome that those traps capturing electrons during a Program operation at either $T$ have comparable emission time constants falling within the observation window of data retention. This, in turn, means that vertical charge loss displays widely distributed time constants, whose turn-on depends on the $T$ at which Program is carried out.

## IV. Conclusion

In this letter, we proved that vertical charge loss in 3-D charge-trap NAND Flash cells is a process featuring widely distributed time constants, whose turn-on depends on the $T$ at which the program operation is carried out. This has relevant consequences on the data retention transient of the cells and on technology operation in the deep cryogenic regime.

## References

[1] H. Tanaka, M. Kido, K. Yahashi, M. Oomura, R. Katsumata, M. Kito, Y. Fukuzumi, M. Sato, Y. Nagata, Y. Matsuoka, Y. Iwata, H. Aochi, and A. Nitayama, "Bit cost scalable technology with punch and plug process for ultra high density flash memory," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2007, pp. 14–15, doi: 10.1109/vlsit.2007.4339708.

[2] C. Monzio Compagnoni, A. Goda, A. S. Spinelli, P. Feeley, A. L. Lacaita, and A. Visconti, "Reviewing the evolution of the NAND flash technology," *Proc. IEEE*, vol. 105, no. 9, pp. 1609–1633, Sep. 2017, doi: 10.1109/JPROC.2017.2665781.

[3] A. Goda, "3-D NAND technology achievements and future scaling perspectives," *IEEE Trans. Electron Devices*, vol. 67, no. 4, pp. 1373–1381, Apr. 2020, doi: 10.1109/TED.2020.2968079.

[4] A. Maconi, A. Arreghini, C. Monzio Compagnoni, G. Van den Bosch, A. S. Spinelli, J. Van Houdt, and A. L. Lacaita, "Impact of lateral charge migration on the retention performance of planar and 3D SONOS devices," in *Proc. ESSDERC*, 2011, pp. 195–198, doi: 10.1109/ESSDERC.2011.6044201.

[5] H.-J. Kang, N. Choi, S.-M. Joe, J.-H. Seo, E. Choi, S.-K. Park, B.-G. Park, and J.-H. Lee, "Comprehensive analysis of retention characteristics in 3-D NAND Flash memory cells with tube-type poly-Si channel structure," in *VLSI Symp. Tech. Dig.*, 2015, pp. 1–2, doi: 10.1109/VLSIT.2015.7223670.

[6] A. Padovani, M. Pesic, M. A. Kumar, P. Blomme, A. Subirats, S. Vadakupudhupalayam, Z. Baten, and L. Larcher, "Understanding and variability of lateral charge migration in 3D CT-NAND flash with and without band-gap engineered barriers," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Mar. 2019, pp. 1–8, doi: 10.1109/IRPS.2019.8720566.

[7] B. Choi, S. H. Jang, J. Yoon, J. Lee, M. Jeon, Y. Lee, J. Han, J. Lee, D. M. Kim, D. H. Kim, C. Lim, S. Park, and S.-J. Choi, "Comprehensive evaluation of early retention (fast charge loss within a few seconds) characteristics in tube-type 3-D NAND Flash memory," in *VLSI Symp. Tech. Dig.*, 2016, pp. 1–2, doi: 10.1109/VLSIT.2016.7573385.

[8] G. Nicosia, N. Righetti, and Y. Dong, "Distributed cycling in charge trap-based 3D NAND arrays: Model and qualification tests implications," in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2023, pp. 1–4, doi: 10.1109/IMW56887.2023.10145969.

[9] J. Park, G. Yoon, D. Go, D. Kim, U. An, J. Kim, J. Kim, and J.-S. Lee, "Decomposition of vertical and lateral charge loss in long-term retention of 3-D NAND Flash memory," in *Proc. IRPS*, 2023, pp. 1–4, doi: 10.1109/IRPS48203.2023.10117868.

[10] Y. H. Liu, T. C. Zhan, Y. S. Yang, C. C. Hsu, A. C. Liu, and W. Lin, "Impact of trapped charge vertical loss and lateral migration on lifetime estimation of 3-D NAND flash memories," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Mar. 2023, pp. 1–6, doi: 10.1109/IRPS48203.2023.10118289.

[11] L. Chiavarone, G. Nicosia, N. Righetti, and Y. Dong, "Experimental segmentation of vertical charge loss mechanisms in charge trap-based 3D NAND arrays," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2024, p. 3, doi: 10.1109/irps48228.2024.10529408.

[12] D. G. Refaldi, G. Malavena, M. Giulianini, L. Chiavarone, A. S. Spinelli, and C. Monzio Compagnoni, "First evidence of SET-like behavior of 3-D NAND flash cells in the deep-cryogenic regime," *IEEE Trans. Electron Devices*, vol. 71, no. 2, pp. 1066–1071, Feb. 2024, doi: 10.1109/TED.2023.3344066.

[13] Y. Aiba, H. Tanaka, T. Maeda, K. Sawa, F. Kikushima, M. Miura, T. Fujisawa, M. Matsuo, and T. Sanuki, "Cryogenic operation of 3D flash memory for new applications and bit cost scaling with 6-bit per cell (HLC) and beyond," in *Proc. 5th IEEE Electron Devices Technol. Manuf. Conf. (EDTM)*, Apr. 2021, pp. 1–3, doi: 10.1109/EDTM50988.2021.9421051.

[14] G. M. Paolucci, C. Monzio Compagnoni, C. Miccoli, A. S. Spinelli, A. L. Lacaita, and A. Visconti, "Revisiting charge trapping/detrapping in flash memories from a discrete and statistical standpoint—Part I: $V_T$ instabilities," *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2802–2810, Aug. 2014, doi: 10.1109/TED.2014.2327661.

[15] D. Resnati, G. Nicosia, G. M. Paolucci, A. Visconti, and C. M. Compagnoni, "Cycling-induced charge trapping/detrapping in flash memories—Part I: Experimental evidence," *IEEE Trans. Electron Devices*, vol. 63, no. 12, pp. 4753–4760, Dec. 2016, doi: 10.1109/TED.2016.2617888.

[16] D. Resnati, G. Nicosia, G. M. Paolucci, A. Visconti, and C. Monzio Compagnoni, "Cycling-induced charge trapping/detrapping in flash memories—Part II: Modeling," *IEEE Trans. Electron Devices*, vol. 63, no. 12, pp. 4761–4768, Dec. 2016, doi: 10.1109/TED.2016.2617890.

[17] G. Nicosia, A. Mannara, D. Resnati, G. M. Paolucci, P. Tessariol, A. S. Spinelli, A. L. Lacaita, A. Goda, and C. Monzio Compagnoni, "Characterization and modeling of temperature effects in 3-D NAND flash arrays—Part II: Random telegraph noise," *IEEE Trans. Electron Devices*, vol. 65, no. 8, pp. 3207–3213, Aug. 2018, doi: 10.1109/TED.2018.2839904.

Open Access funding provided by 'Politecnico di Milano' within the CRUI CARE Agreement