Effect Of Short Pulsed Program/Erase Cycling On Flash Memory Devices

Conference Paper · June 2016

5 authors, including:

Philippe Chiquet
Aix-Marseille Université
18 PUBLICATIONS **10** CITATIONS
SEE PROFILE

J. Postel-Pellerin
French National Centre for Scientific Research
41 PUBLICATIONS **63** CITATIONS
SEE PROFILE

P. Masson
University of Nice Sophia Antipolis
80 PUBLICATIONS **625** CITATIONS
SEE PROFILE

14th IMEKO TC10 Workshop on Technical Diagnostics
New Perspectives in Measurements, Tools and Techniques
for systems reliability, maintainability and safety
Milan, Italy, June 27-28, 2016

Effect Of Short Pulsed Program/Erase Cycling On Flash Memory Devices

P. Chiquet [1], J. Postel-Pellerin [1], C. Tuninetti [1], S. Souiki-Figuigui [1], P. Masson [2]

[1] Aix-Marseille Universite, IM2NP-CNRS UMR 7334, rue Frederic Joliot-Curie, 13451 Marseille, France, philippe.chiquet@im2np.fr
[2] Nice Sophia-Antipolis University EPOC, 06410 Biot, France, pascal.masson@unice.fr

***Abstract* - The present paper proposes to investigate the effect of pulsed Program/Erase cycling on Flash memory devices. Usually, electrical operations related to said devices involve the application of single long pulses to various terminals of the memory transistor to induce various tunneling effects allowing the variation of the floating gate charge. According to the literature, the oxide degradation occuring in such cases can be reduced by replacing DC stress by AC stress of the MOS-based devices. After a brief presentation of the functioning of the Flash memory transistors tested in this work, the experimental setup used to replace standard electric signals with short pulses will be described. Electrical results showing the benefits of programming and erasing non-volatile memories with short pulses will then be presented.**

1. INTRODUCTION

Over the course of the last few years, Flash memory cells, which can be found in nowadays common products like USB flash drives and smart cards, have become the dominant non-volatile memory device in the semiconductor industry. Flash memories are built around a floating gate transistor, which as shown in Fig.1, is a MOS transistor to which an extra tri-layer stack oxide "Oxide/Nitride/Oxide" (ONO) and a top electrode (control gate) have been added. Information is stored in the device as an electric charge located in the floating gate which, due to being surrounded by various oxide layers, retains this charge when the power supply is removed, thus assuring the non-volatile behavior of the Flash transistor [1].

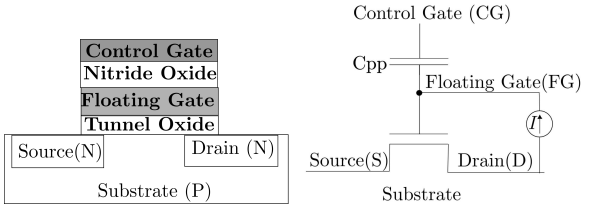

Fig. 1. Schematic (left) and electric (right) representation of a Flash floating-gate transistor.

According to the value Q_{FG} of this charge, two distinct logical states 0 and 1 can be differentiated, and the threshold voltage VT of the transistor can be calculated according to (1):

$$V_T = V_{T0} - \frac{Q_{FG}}{C_{PP}} \quad (1)$$

where V_{T0} is the natural threshold voltage of the cell and C_{PP} the control gate / ONO / floating gate capacitance [2].

The transition from one state to the other is obtained when electric signals are applied to some of the transistors terminals, by the injection or removal of electrons from the floating gate through the tunnel oxide. The logical states corresponding to the highest and lowest threshold voltage values of the floating gate transistor will from now on be referred to as the programmed and erased states respectively. Programming is usually achieved through Channel Hot Electron (CHE) injection in flash memories, while erasing is generally obtained thanks to Fowler-Nordheim (FN) injection [3].

Due to this functioning the tunnel oxide layer, which can be degraded due to repeated electrical operations, is critical to the reliability of the memory devices and various solutions can be considered in order to improve their overall reliability. The present work will focus on the use of short-pulsed signals, which consists in replacing long standard signals by a series of pulses of short plateau widths, as previous studies show that MOS-based devices are less impacted by AC stress than DC stress [4–10].

2. GENERATION OF SHORT-PULSED SIGNALS

A. Experimental setup

In order to investigate the endurance of the tested Flash memory cells under short pulses, a complete setup has been developped with the help of an Agilent Semiconductor Device Parameter Analyzer B1500A [11] equipped with two WGFMUs (Waveform Generator Fast Measurement Unit B1530A), two SPGUs (Semiconductor Pulse Generator Unit) and four SMUs (Source Monitor Unit), as described in Fig.2. The SMUs are used to measure the $I_{DS}(V_{GS})$ characteristics necessary to the reading (threshold voltage extraction at a given drain current value) of the cell state while the SPGU enable the definition of pulses for the successive programming/erasing operations. An Agilent 16440A selector switches between the SMU and SPGU during the endurance test. The additional WGFMU in the setup, connected via RSU (Remote-Sense and Switch Unit), enables the definition of arbitrary waveforms and the

dynamic measurement of the drain current, which can be useful when evaluating the consumption of the programming operation during the endurance test [12].

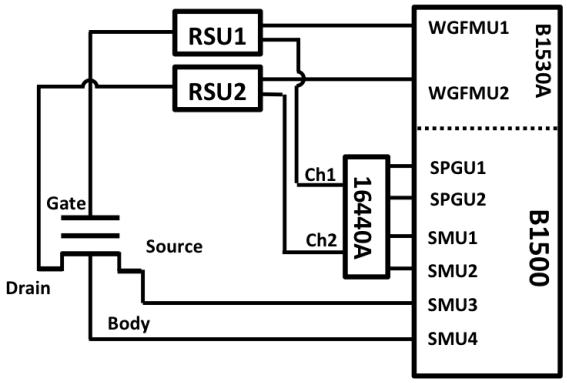

Fig. 2. Experimental setup used to perform endurance tests with current consumption measurements.

B. Definition of the applied signals

Short-pulsed signals have been chosen so as the programmed and erased threshold voltages measured for a fresh Flash device are close to the threshold voltage values obtained when standard signals are used. It has been experimentally observed that such a condition is met when the total plateau time spent on short pulses is equal to the duration of the standard signal.

Standard (STD) Flash CHE signals consist in simultaneous control gate and drain pulses of amplitude V_{PP} equal to +9V (during $5\mu s$) and +4.2V (during $1\mu s$) respectively, while standard FN erasing is obtained by applying a single $500\mu s$ pulse of amplitude V_{PP} equal to -18V on the control gate.

In order to obtain comparable threshold values using series of short pulses, the standard drain signal in CHE programming and the standard control gate signal in FN erasing have respectively been transformed into a series of 20 pulses of 50ns plateau time and a series of 10000 pulses of 50ns plateau time. In each case, the signals created with the experimental setup described above have been checked with the help of an oscilloscope as seen in Figs. 3 to 5.

3. ELECTRICAL RESULTS

A. Qualitative results and total window closure

Measurements carried out on numerous devices yielded electrical results highlighting the effect of pulsed program/erase operations on Flash memory endurance. Endurance of memory cells is mainly characterized by the evolution of their programming window, defined as the difference between the threshold voltages of the "programmed" and "erased" states, over numerous program/erase electrical operations.

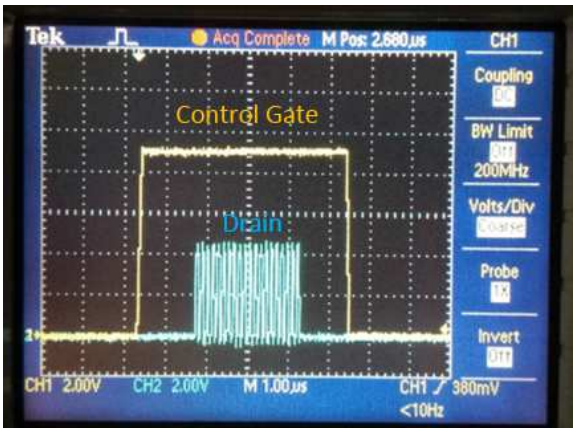

Fig. 3. Oscilloscope observation of the control gate (yellow) and drain (blue) signals for a short-pulsed CHE programming of the Flash memory cells.

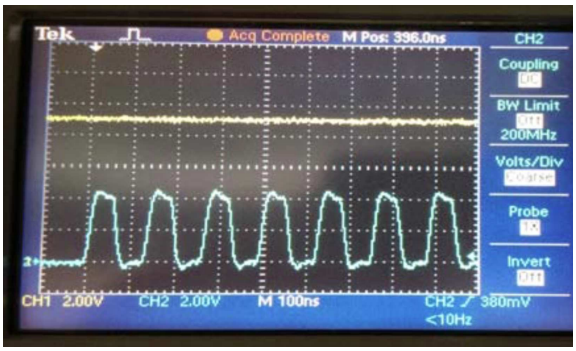

Fig. 4. Close-up observation of signals observed in Fig.3.

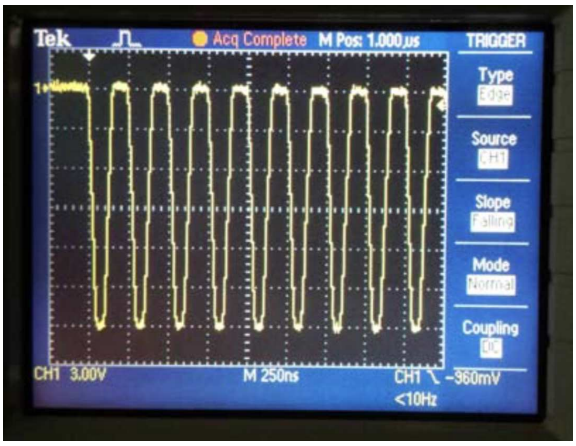

Fig. 5. Oscilloscope observation of the control gate signal of a Flash floating gate transistor during a FN erase operation.

In order to quantify the gradual programming window closure linked to device degradation, a criterion had to be chosen. In the present work, the total window closure ΔV_T

after k program/erase cycles, defined by (2) and (3), has been retained.

$$\Delta V_T = \Delta V_T^{prog} - \Delta V_T^{erase} \quad (2)$$

$$\Delta V_T^{prog,erase}(k) = V_T^{prog,erase}(k) - V_T^{prog,erase}(1) \quad (3)$$

According to it's very definiton, ΔV_T is calculated as the sum of two negative values and represents the combined shifts of the "programmed" and "erased" threshold voltages. This total window closure is plotted in Fig.6 as a function of the number of program/erase cycles up to 10^6 cycles. It can be seen that both the Short (20x50ns)/STD and STD/Short (10000x50ns) cyclings provide a decrease in programming window closure of bout 0.5V compared to the STD/STD cycling.

In order to understand the effect of the pulses' parameters on this observed endurance improvement, measurements involving different plateau amplitudes and duty cycles have been carried out.

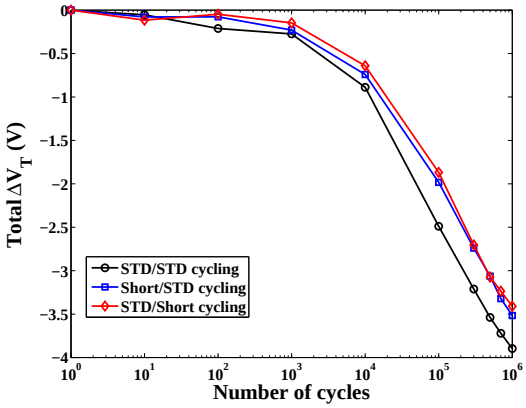

Fig. 6. Programming window closure up to 10^6 program/erase (STD/STD, Short/STD and STD/Short) cycles. Each measurement point has been averaged over ten Flash memory cells.

B. Effect of pulse amplitude

Electrical results presented in Fig.7 compare the programming window closure of the STD/Short (10000x50ns) and STD/STD cyclings for V_{PP} values of -17V, -18V and -19V. While a difference in programming window closure between the two cycling modes can be observed for the first two V_{PP} values after 10^4 cycles, erasing at V_{PP} = -19V with short pulses instead of the standard signal makes no difference. This could be explained by the fact that at a given duty cycle, the benefits of replacing DC stress with AC stress become noticeable only if the plateau time of the pulses is reduced as much as possible as the stressing field is increased [10]. Other studies have also shown that such results cannot be achieved at ambient temperature when plateau times of the applied pulses become long enough [4].

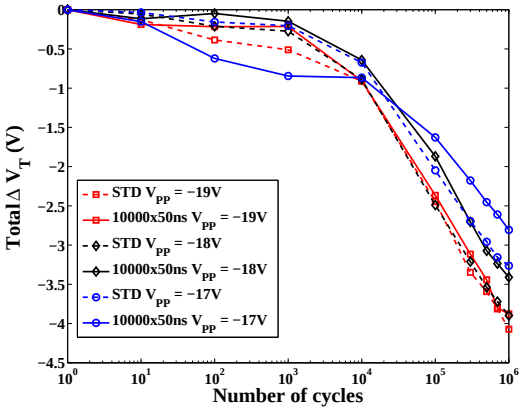

Fig. 7. Programming window closure up to 10^6 program/erase (STD/STD and STD/Short) cycles for various erasing V_{PP} values. Each measurement point has been averaged over ten Flash memory cells.

C. Effect of duty cycle

In order to achieve lower device degradation, allowing oxide relaxation during electrical stress is a common solution [4]. It has been achieved in this study by increasing the delay between the short pulses constituting the signals applied to the control gate and the drain of the transistor. The duty cycle, which is defined as the ratio of the plateau duration t_{pl} to the period value T as shown in Fig.8, is thus decreased.

Electrical results obtained after Short/Short cycling (V_{PP} = -18V) for duty cycles equal to 0.1 and 0.01 have been compared to experimental data obtained for STD/STD cycling in Fig.9. It can be observed that decreasing the duty cycle allows to reduce the programming window closure. Using experimental data from Fig.9 allows to plot ΔV_T as a function of the duty cycle as shown in Fig.10. The total window closure is shown to have a logarithmic behaviour with respect to duty cycle.

D. Measurement of device consumption

Energy consumption of non-volatile memory devices has been of recent interest as experimental possibilities offered with the coming of new parameter analysers in recent years [12]. It can be defined according the following relatioship (4) :

$$E_c = \int V_{DS}.I_{DS}dt \quad (4)$$

Results of time-resolved measurements of drain current performed during a programming operation before and after electrical cycling (STD/STD and Short/STD) can be observed in Fig.11 and Fig12. It is shown that device consumption typically increases with device degradation and the number of applied pulses in a single program/erase operation.

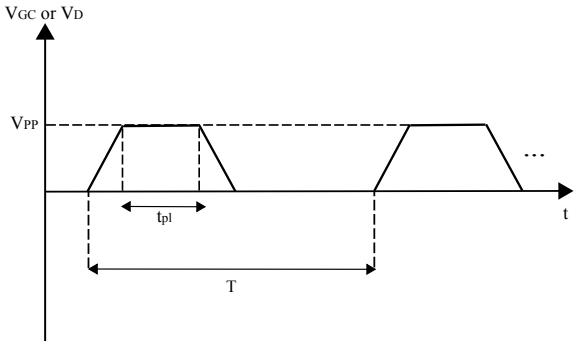

Fig. 8. Schematic representation of a period of the electrical signals applied to the control gate (V_{GC}) or drain (V_D) of the memory transistor. The duty cycle is calculated as the ratio of the plateau duration t_{pl} to the period value T.

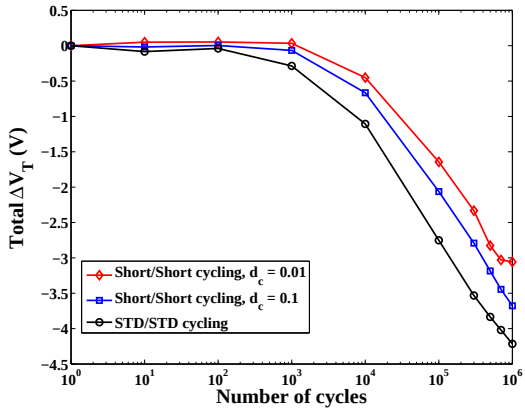

Fig. 9. Programming window closure up to 10^6 program/erase cycles obtained for various duty cycle values. Each measurement point has been averaged over ten Flash memory cells.

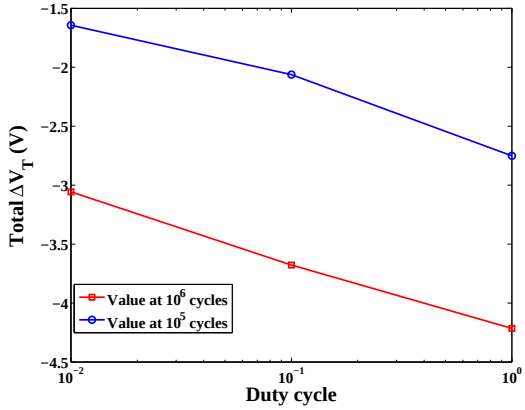

Fig. 10. Total window closure as a function of duty cycle obtained from data of Fig.9 at 10^5 and 10^6 cycles.

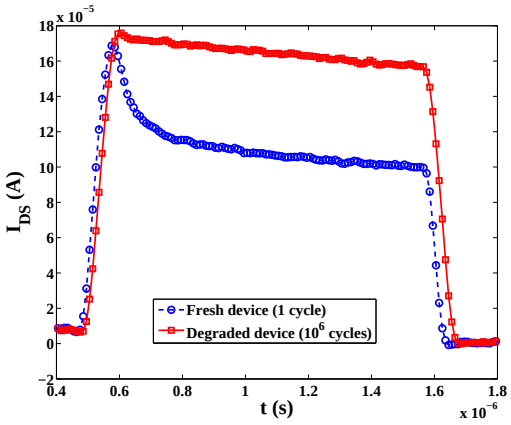

Fig. 11. Time-resolved measurement of the drain current I_{DS} during a standard programming signal before (circles) and after (squares) STD/STD program/erase cycling (10^6 cycles).

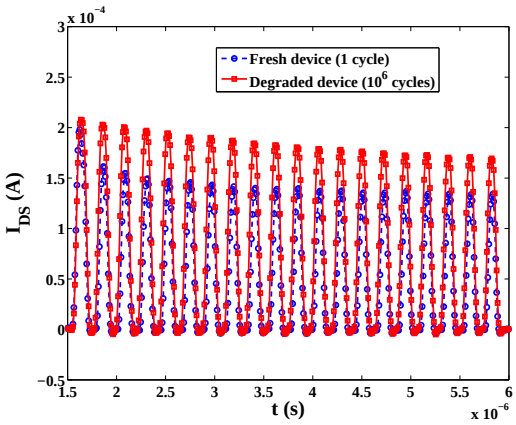

Fig. 12. Time-resolved measurement of the drain current I_{DS} during a short (20x50 ns) programming signal before (circles) and after (squares) short/STD program/erase cycling (10^6 cycles).

4. CONCLUSIONS

The impact of the use of short pulses on the endurance of Flash memories during the CHE programming and FN erasing operations have been investigated in this paper. The proposed experimental setup helped produce signals with plateau times as low as 50ns, allowing the observation of a decrease in programming window closure during program/erase cycling. The impact on this closure of parameters such as amplitude or duty cycle of the signals applied to the terminals of the memory transistors has been investigated.

As the study led in this paper shows, the endurance improvement obtained by replacing standard control gate and drain signals by series of short pulses can only occur at a cost as the overall signal durations and device consumption have been shown to increase. This trade-off could prove

useful in specific application fields (i.e. automotive and avionic industries) where reliability of the memory devices is the most important factor.

Morever, automotive and avionic applications typically require memory devices to operate within a much larger temperature range (typically -40°C to 150°C). As some other authors have shown, decreasing the duty cycle value yields better electrical results at higher temperatures [4], making the present results potentially even more interesting for such applications.

As a perspective, many other types of measurements could be considered to complete the study presented here. Firstly, measurements presented in this work could be reproduced using different plateau times but similar duty cycles to be able to understand the effect of the former on device degradation. Looking at the problem from a different angle, this experimental setup could also be used to find the electrical characteristics (mainly amplitude and plateau time) of short single pulses (as oppsosed to a series of pulses) that could replace the standard program/erase signals for an equivalent initial programming window. Finally, complementary measurements could be performed on test MOS capacitors to investigate the effect of short-pulsed electrical stress in terms of interface trap density [13] and stress induced leakage current (SILC), which are physical manifestations of oxide degradation that can lead to program/erase window closure in non-volatile memory devices.

ACKNOWLEDGMENTS

Authors wish to thank Jean-Luc Ogier (STMicroelectronics, Rousset) for his technical support.

REFERENCES

[1] W.Brown, J.Brewer, "Nonvolatile semiconductor memory technology: a comprehensive guide to understanding and to using NVSM devices", IEEE Press, 1998.

[2] P. Pavan, R. Bez, P. Olivo, E. Zanoni, "Flash memory cells - an overview", Proc. of the IEEE 85 (8) (1999), pp. 1248-1271.

[3] R.Laffont, R.Bouchakour, O.Pizzuto, J.-M.Mirabel, "A 0.18 um flash source side erasing improvement", Proceedings of NVMTS (2004), pp. 105-109

[4] B.Rebuffat, J-L.Ogier, P.Masson, M.Mantelli, R.Laffont, "Relaxation Effect on cycling on NOR Flash memories", Proc. of IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), 2015, pp. 613-616.

[5] A.Cester,A.Paccagnella, G.Ghidini, "Stress induced leakage current under pulsed voltage stress", Solid-State Electronics, vol.46, 2002, pp. 399-405.

[6] M.Nafria, J.Sune, D.Yelamos, X.Aymerich, "High field dynamic stress of thin SiO2 films", Microelectron. Reliab., vol.35, No.3, 1995, pp. 539-553.

[7] D.Caputo, R.Feruglio, F.Irrera, B.Ricco "Effect of pulsed stress on leakage current in MOS capacitors for non-volatile memory applications", Proc. of IEEE European Solid-State Device Research Conference (ESSDERC), 2002, pp. 567-570.

[8] F.Irrera, B. Ricco, "Pulsed tunnel programming of nonvolatile memories", IEEE Trans. On Electron Devices, vol.50, No.12, December 2003, pp. 2474-2480.

[9] A.Chimenton, F.Irrera, P.Olivo, "Improving performance and reliability of NOR-Flash arrays by using pulsed operation", Microelectronics Reliability, vol.46, 2006, pp. 1478-1481.

[10] F.Irrera, T.Fristachi, D.Caputo, B.Ricco, "Optimising Flash memory tunnel programming", Microelectronics Engineering, vol.72, 2004, pp. 405-410.

[11] Agilent Technologies, Agilent B15000 users guide, 2009.

[12] V.Della Marca, J.Postel-Pellerin, G.Just, P.Canet, J-L.Ogier, "Impact of endurance degradation on the programming efficiency and the energy consumtion of NOR Flash memories", Microelectron. Reliab., vol.54, No.9-10, 2014, pp. 2262-2265.

[13] B.Rebuffat, P.Masson, J-L.Ogier, M.Mantelli, R.Laffont, "Effect of AC stress on oxide TDDB and trapped charge in interface states" Proc. of International Symposium on Integrated Circuits (ISIC), 2014, pp. 416-419.