Extended Abstracts of the 2018 International Conference on Solid State Devices and Materials, Tokyo, 2018, pp147-148

Hot Carrier Programing Performance Study on 2T 40nm Flash Memory Cell Array For Variability Evaluation

T. Kempf[1,2,3], V. Della Marca[2], M. Mantelli[1], J. Postel Pellerin[2], J.-M. Portal[2], P. Masson[3], A. Regnier[1], S. Niel[1]
Phone: +33(0)4 4268 8428 E-mail: thibault.kempf@st.com
[1]STMicroelectronics, Rousset, France
[2]Aix-Marseille Université, IM2NP, CNRS, UMR 7334, 13397 Marseille, France
[3]Polytech'Lab / Nice Sophia-Antipolis University, Biot, France

Abstract

In this paper, we propose to use a memory array devices of 2T Flash cells, fully connected in parallel, to study their simultaneous programming by Source Side Injection (SSI). Single devices and test chip data are used to optimize the programming conditions against bitlines voltage drop, due to large current needed for parallel operation, and validate them through cycling experiments. We show that CAST I_D-V_G on this type of matrix allows identifying the memory technology variability and the extrinsic population of the written state.

1. Introduction

Due to the dropping of reliability margins, anticipating the variability of the memory cells population is mandatory for the development of new embedded Non Volatile Memory (eNVM) [1]. Consequently single devices characterizations have to be completed by test structures yielding a large statistics and a product-like approach. However, even if Micro-Controller Unit (MCU) for testing, or test chip (TC), has proved to be complementary to single device characterization [2] to evaluate the memory integration inside a specific design, its complexity increases the test time during technology development (Wafer Sort). On another hand, Cell Array Structure Test (CAST), initially patented by STMicroelectronics [3], have been widely used to study reliability concern of memory technologies such as extrinsic behavior of cycled memory cells during retention [4], or the BEOL and FEOL oxide breakdown. Indeed, by introducing an array of cells fully connected in parallel, these structures contain the representation of both the intrinsic and systematic memory variability as described **Fig. 1(a)**. However, with the bit size down scaling and its associated BEOL, voltage drop has appeared in these structures limiting their size [5]. Therefore, operation requiring high current, such as reading or programming by Hot Carrier (HC), are usually not affordable. Nevertheless, with the accession of Source Side Injection (SSI) mechanism, Flash consumption has decreased suggesting the possibility to program CAST in a written state.

In this study, we propose to enable SSI programming on our 40nm 2T Flash NOR technology. The comparison with single cell (SC) and TC are used to validate the programing operation conditions used on CAST. We demonstrate that CAST I_D-V_G of the programed state yields information about the memory threshold voltage (V_T) dispersion. Then cycling is proposed to validate the CAST programing operation.

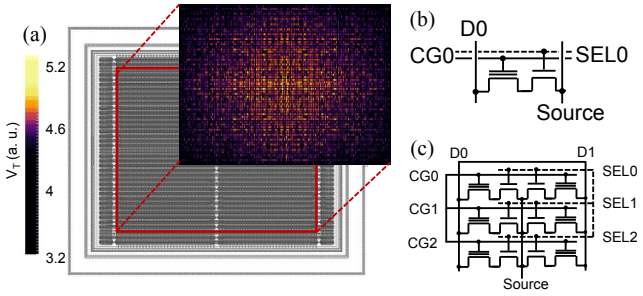

Fig. 1(a) 10k cells CAST layout and hypothetic V_T spatial dispersion extracted from test chip. (b) The CAST equivalent circuit. (c) The 2tr Flash memory equivalent circuit.

2. Technology details

Our Flash memory is a 2T 40nm NOR eNVM programmed by SSI and erased by Fowler-Nordheim (FN) composed of a select (SEL) used to access floating gate (FG) transistors as shown in **Fig. 1(b)**. In **Fig. 1(c)**, CAST device is similar to a memory array, except that all cells are connected in parallel (2048 cells per bitlines in the 10k CAST). A TC integrating 512kB of memory, with a similar array layout is used for the study and has been previously described in [2].

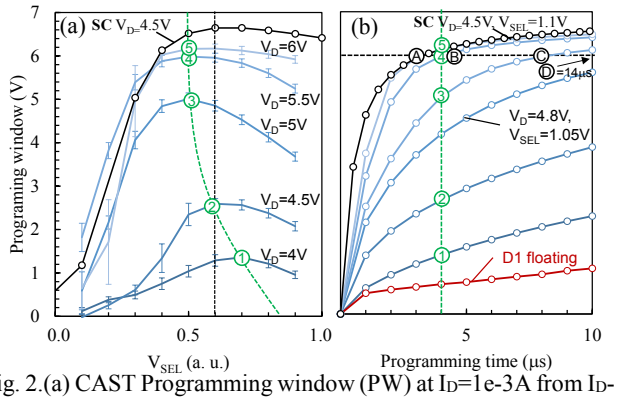

Fig. 2.(a) CAST Programming window (PW) at I_D=1e-3A from I_D-V_G versus V_{SEL} and V_D (for 8 devices) and (b) the corresponding CAST PW kinetic for the optimum V_{SEL} for each V_D. The black line is the SC behavior. The red line represents the CAST kinetic whose unselected cell drain (D1) is left floating. A to D represent the cycling conditions to obtain same initial PW used in **Fig. 4**.

3. SSI programing conditions and optimization

Fig. 2 describes the SC V_T dependence on V_{SEL} (**Fig. 2(a)**) under normal conditions (V_D=4.5V) and its kinetic (**Fig. 2(b)**). First, in order to obtain similar V_T and kinetic on CAST, we propose to ground the unselected cells (D1), when programing the cell (D0) that shares the same control gate (CG0) and

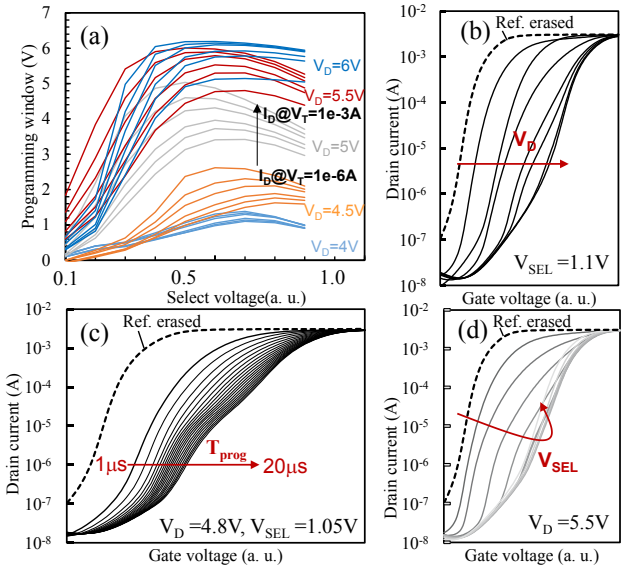

Fig. 3(a) Programming window extracted at I_D=1e-3, 3e-4, 1e-4, 3e-5, 1e-5 and 1e-6A for different V_D and V_{SEL}. Impact of V_D (b), V_{SEL}(c), and programming duration (T_{prog}) on the same die I_D-V_G.

select (SEL0) (see **Fig.1(c)**), in order to reduce voltage drop along metal lines. It allows a parallel current path for the programming operation only. As described in **Fig. 2(b)**, the programming kinetic using default conditions (red line) is very slow compared to the case of grounded D1 (blue lines). Moreover, as shown in **Fig. 2(a)**, higher V_D are needed compared to SC conditions because of the voltage drop in CAST. Using dynamic I_D of programming operation (not shown) and the method of [6], one can evaluate single cell resistance under programming (≈1MΩ), hence calculates a necessary V_D increase toward at least 5V (to obtain the required V_D on CAST nodes).

In **Fig. 2(a)**, the select voltage (V_{SEL}) impacts the V_T by decreasing, for high V_{SEL}, the cell resistance. Finally it shifts the CAST V_T optimum toward lower values of V_{SEL} compared to SC and decreases the maximum CAST V_T achievable. The **Fig 2(b)** shows that kinetics of optimum V_{SEL} for each V_D approach SC both in term of timing and V_T amplitude.

4. Population variability in CAST devices

While CAST V_T at high current represents the intrinsic population, V_T at lower I_D gives information about the CAST V_T dispersion. In **Fig. 3**, either extrinsic population or large V_T spread in CAST are spotted by looking at the I_D-V_G subthreshold slope or the presence of bump in the I_D-V_G curves. **Fig. 3(a)** clearly enlighten that the V_T dispersion is directly correlated to PW extracted at low I_D. Indeed, by decreasing V_D or V_{SEL}, slow cells are more impacted and are spreading the V_T distribution (**Fig. 3(a),(b)**). The programming kinetic of **Fig 3(c)** highlights the specific behaviour of these extrinsic cells compared to the main population: a slower kinetic as well as a lower saturated V_T (described also in **Fig. 3(a)** by V_D=6V curves). These cells may be directly linked to border array cells, hence to the spatial variability as described in [2].

5. Cycling CHE/FN versus cycling FN/FN

Cycling is done using the previously used FN/FN scheme on CAST and SC whereas FN/CHE scheme is applied on SC, CAST and TC, as described in **Fig. 4(a)**. While FN/FN shows a good agreement between CAST and SC in terms of PW closing, it is not representative of our Flash cycling degradation (shift of V_{Tp}). On the other hand, **Fig. 4(a) (b)**, CHE/FN cycling presents good agreement between SC, CAST and TC for PW closing as well as for the shifts of V_{Tp} and V_{Te} (without any impact of V_D). These results suggest that our CHE/FN CAST cycling generates similar degradation mechanisms compared to SC and TC cycling and that the methodology applied for HC operation for CAST is accurate.

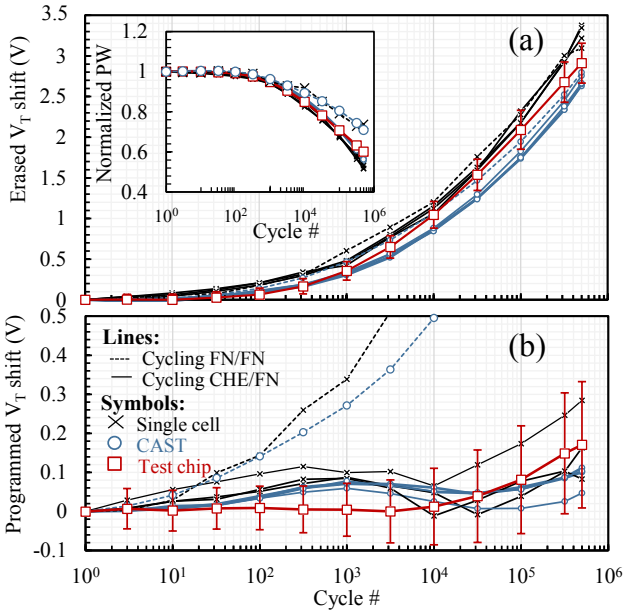

Fig 4. Cycling CHE/FN of SC (8 devices) for 3 different V_D (4.2V, 4.5V, 4.8V), CAST (2 devices) for 4 V_D (see **Fig. 2(b)**, A to D) and a TC sector (262144 cells) under default conditions. Cycling FN/FN: on 8 devices both SC and CAST.

6. Conclusions

In this paper, Hot Carrier programming of CAST memory array is successfully performed. Optimized programming conditions are needed in order to overcome the voltage drop while, thanks to the array layout design, the unselected cells reduce the access resistance. The correlation of cycling degradation between array, cells and test chip proves the accuracy of the programming scheme. Moreover, it enables to study the memory variability under various programming conditions. Through simple I_D-V_G characterizations, CAST gives useful information about V_T dispersion and the extrinsic population.

References

[1] R. Strenz, *Proc. of IEEE IEDM* (2011) 9.4.1.
[2] T. Kempf, *et al.*, *Accepted Proc. of IEEE IRPS* (2018) 6E.4-1.
[3] F. Pio, *et al.*, US Patent n°6,128,219, oct.3, 2000.
[4] R. Djenadi, *et al.*, Solid-State Electronics **78** (2012) 80.
[5] P. Canet, *et al.*, Microelectron. Reliab. **64** (2016) 36.
[6] T. Kempf, *et al.*, *Accepted Proc. of IEEE VLSI-TSA* (2018) 104.