Reduction of the cell-to-cell variability in Hf$_{1-x}$Al$_x$O$_y$ based RRAM arrays by using program algorithms

E. Pérez, A. Grossi, C. Zambelli, P. Olivo, R. Roelofs, and Ch. Wenger

Abstract—In this report, we propose an effective route to reduce the cell-to-cell variability in 1T-1R based RRAM arrays by combining the excellent switching performance of Hf$_{1-x}$Al$_x$O$_y$ with an optimized Incremental Step Pulse with Verify Algorithm (ISPVA) for programming. The strongly reduced cell-to-cell variability improves the thermal and post-programming stability of the arrays, which is relevant for many applications of the RRAM technology. Finally, the retention study at 150 °C enables the prediction of the data storage capability.

Index Terms—RRAM, Hf$_{1-x}$Al$_x$O$_y$, variability, data retention.

I. INTRODUCTION

RESISTIVE Random Access Memories (RRAM) are based on the electrical modification of the conductance of Metal-Insulator-Metal stacks: the set operation switches the cell into a Low Resistive State (LRS), whereas reset operation brings the cell back to a High Resistive State (HRS) [1]. HfO$_2$ based RRAM is one of the most promising technologies for the next generation of non-volatile memory applications [2]. However, the large cell-to-cell variability in RRAM arrays is still a relevant issue. The random conductive filament formation and rupture process is generally regarded as a critical non-uniformity parameter of resistive switching. Ionic doping of HfO$_2$ is one effective way to improve the switching stability of the resistive switching [3]. Using a program algorithm is another option to reduce the resistance variability, although post-programming shifts were reported recently [4]. In this study, we demonstrate the drastic reduction of the cell-to-cell variability and post-programming instabilities using Hf$_{1-x}$Al$_x$O$_y$ as switching oxide and optimized program algorithms with verify procedures.

II. EXPERIMENTAL

The 1T-1R RRAM cells in the 4 kbit arrays are constituted by a NMOS transistor manufactured in IHP’s 0.25 μm CMOS technology, whose drain is connected in series to a variable resistor, as illustrated in Fig. 1. The resistor is a Metal-Insulator-Metal (MIM) device integrated on the metal line 2 of the CMOS process. This MIM resistor consists of a planar TiN/ Hf$_{1-x}$Al$_x$O$_y$ /Ti/TiN stack. One additional mask is required for patterning the MIM stack. The Hf$_{1-x}$Al$_x$O$_y$ layers with thickness of 6 nm are grown by Atomic Layer Deposition (ALD with an Al content of 10 %) using an ASM A412 Batch Furnace. After patterning the MIM cells with area of about 0.4 μm$^2$, an additional thin Si$_3$N$_4$ layer was deposited to protect the MIM cell.

Fig. 1: TEM cross-sectional image of 1T1R architecture with a NMOS access transistor (1T, marked by a blue square) and a 0.6×0.6 μm$^2$ MIM cell (1R, marked by a white square); BL, SL and WL denote the bit line, source line and word line, respectively.

III. RESULTS AND DISCUSSION

To activate the resistive switching behavior, the RRAM cells require a preliminary forming operation. This initial operation plays a fundamental role in determining the subsequent devices performance. Therefore, the Incremental Form and Verify (IFV) algorithm is used [5]. The functionality of the forming and programming algorithms has been verified over 1000 cycles on the arrays.

In order to suppress the reported post-program instabilities, the Incremental Step Pulse with Verify Algorithm (ISPVA) was applied [5], consisting of a sequence of increasing voltage pulses on the drain terminal during set operation and on the
source terminal during reset operation. After every pulse a read and verify operation is performed. Averaged, 6 Voltage pulses with a width of 12 μs are required for the set process, respectively 12 pulses with the same width for resetting the resistive cells.

The determination of the correct threshold values for the ISPVA is a trade-off between high On/Off ratio, high programming yield and low cell-to-cell variability. By simply increasing the On/Off ratio, the cell-to-cell variability will be raised too, while the programming yield will be reduced. In addition, the threshold values are also affected by the choice of the resistive oxide.

In order to determine the optimum threshold values for HfO$_2$ and Hf$_{1-x}$Al$_x$O$_y$, the ISPVA is applied by using the two different LRS thresholds of 18 and 30 μA. Similarly a current value of 5 μA was considered for HRS in order to target a HRS/LRS ratio of 3 respectively 6. The cumulative distributions of the read currents after reset and set programming cycles, using the ISPVA are illustrated in Fig. 2. The variability of the LRS can be optimized by the Al-doping of the HfO$_2$ layer and by choosing the correct threshold value of 30 μA, as shown in Fig. 2b).

Compared to recently reported distributions [6, 8], the window between HRS and LRS is well defined and the cell-to-cell as well as the cycle-to-cycle variability is strongly reduced.

The mean LRS current values of pure HfO$_2$ devices tend to increase with the number of cycles [7]. By doping with Al, this trend is eliminated. The current values of HRS and LRS of Hf$_{1-x}$Al$_x$O$_y$ remain constant during 1000 cycles. Moderate Al-doping of 10% stabilizes the HfO$_2$ matrix, improving the endurance properties by hindering oxygen-vacancies relaxation without severe change in switching parameters [8]. In addition, the Al-doping of HfO$_2$ reduces also the variability of the LRS current values, as illustrated in Fig. 3.

The evolution of the mean reset and set voltage values during the 1000 endurance cycles are illustrated in Fig. 4. The voltages values of the reset operation are a bit higher in pure HfO$_2$ than in Al-doped HfO$_2$, which is in line with reported trends [7]. The mean voltage values of Hf$_{1-x}$Al$_x$O$_y$ based cells are not impacted by cycling, while the values of HfO$_2$ cells are slightly influenced by cycling.

Fig. 2. Cumulative distributions of the read out currents after programming using the ISPVA after 1, 10, 100 and 1k cycles. The characteristics of HfO$_2$ RRAM cells are illustrated in a), while the impact of cycling on the Hf$_{1-x}$Al$_x$O$_y$ array is shown in b). The threshold values for the set and reset operations were marked as dotted lines.

Fig. 3. (a) Mean current values after reset (open circles) and set (closed circles) programming operations as function of cycling.
(b) Dispersion coefficient $\sigma^2/\mu(I_{LRS})$ (closed circles) and $\sigma^2/\mu(I_{HRS})$ (open circles) of the read out currents as function of cycling.

Fig. 4. Mean voltage values after reset (open circles) and set (closed circles) programming operations as function of cycling in Hf$_{1-x}$Al$_x$O$_y$ based RRAM arrays.

In order to evaluate the post-programming stability, the readout operation is continued for 100 sec after the HRS or LRS current thresholds are achieved. The programming as well as the reading operation is performed at -40 °C and 150 °C. The LRS and HRS resistances programmed by ISPVA algorithms are illustrated in Fig. 4. Short time instabilities (< 2 s) are
monitored; a few cross-bit cells are detected in the LRS/HRS window. But after 10 seconds, the HRS distribution reverts back above the threshold value. As shown in Fig 4, there isn’t any remarkable impact of temperature on the relaxation characteristics of the HRS.

Charging effects, filament rearrangement, redistribution of vacancies and dielectric relaxation processes in the resistive oxide layer could cause the short time instability effects. The LRS state remains stable after set programming, indicating the absence of filament rearrangements or charging effects. After finishing the reset ISPVA, dielectric relaxation processes could cause the shift of the HRS resistances to lower as well as larger values than the initial HRS resistance [9].

The data retention of HfO2 based RRAM devices can be improved by increasing the LRS currents [10]. An alternative approach to improve the retention is using ISPVA for programming Hf1-xAlxOy cells. The high temperature data retention is investigated at the temperature of 150 °C.

1000 Hf1-xAlxOy cells were programmed via ISPVA to LRS respectively to HRS then exposed to thermal stress up to 100 hours. The evolution of LRS and HRS distribution was monitored at log spaced sampling times from 0.1 to 100 hours, as illustrated in Fig 6.

As shown in Fig 7, the LRS currents continuously decrease with cycling numbers and retention time, which could be caused by oxygen diffusion related mechanism [11]. The degradation of the HRS is different, all currents increase similarly after a short time of baking followed by a more or less stable plateau. They are not additionally degraded by cycling. The number of so-called cross-bit cells; cells which current values shifted in between the threshold values for HRS (5 µA) and LRS (30µA) during the retention stress test are illustrated in Fig 5.

While the number of cross-bit cells in the HRS remains less than 15% of all Hf1-xAlxOy cells, the percentage of cross-bit cells in LRS is strongly raised by cycling and baking time. The low formation energy of oxygen vacancies in trivalent ion-doped HfO2, which reduces the variability of switching parameters, could also lead to worse retention results [12]. However, the shorter bond length between Al and O ions reduces also the unwanted diffusion of oxygen vacancies. The trade-off between good retention and low variability has to be balanced by choosing the correct amount of doping.

Fig. 6. Cumulative distribution of the read-out currents after 1 and 1000 cycles. The set and reset threshold values of the programming algorithms are illustrated by the dashed lines. The RRAM array is baked at 150 °C in log spaced sampling time rates from 0.1 to 100 hours. For clarity reasons, just 4 baking times are illustrated in these plots.

Fig. 7. Distribution of Hf1-xAlxOy based Cross-bit cells in between the LRS (open circles) and HRS (closed circles) threshold values as function of baking time and switching cycles.

Fig. 5. Post-programming distribution of resistances programmed by set and reset ISPVA. The programming operations of the arrays were performed at –40 °C (a) and 150 °C (b). The dotted lines represent the threshold values for set and reset programming.
REFERENCES


[9] X. Li, H. Wu, B. Gao, N. Deng and H. Qian “ Short time high resistance state instability of TaOx based RRAM devices”, DOI 10.1109/LED.2016.2630044, IEEE EDL.

