

Endurance and targeted programming behavior of HfO2-FeFETs

|  |  |  |
| --- | --- | --- |
| Haidi Zhou, Johannes Ocker | Martin Trentzsch, Stefan Dünkel | Thomas Mikolajick |
| Menno Mennenga, Marko Noack | Sven Beyer | *NaMLab gGmbH* |
| Stefan Müller | *GLOBALFOUNDRIES Dresden* | *Chair of Nanoelectronic Materials* |
| *Ferroelectric Memory GmbH* | *Module One LLC & Co. KG* | *TU Dresden* |
| Dresden, Germany | Dresden, Germany | Dresden, Germany |
| haidi.zhou@ferroelectric-memory.com | sven.beyer@globalfoundries.com | thomas.mikolajick@namlab.com |
| l | l |

***Abstract*— In this paper, recent advances on the development of Hafnium oxide (HfO2)-based ferroelectric field-effect transistors (FeFETs) are shown with respect to its**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **memory** | **window,** | **trapping** | **behavior** | **and** | **endurance** |

**characteristics. Although this novel ferroelectric memory cell shows superior characteristics such as device scalability, CMOS compatibility, fast access time and low power operation, the challenges for HfO2-based FeFET device lie with device variability and endurance. To investigate endurance failure in relation to charge trapping, different time delays were introduced to allow for detrapping of charges leading to**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **improved** | **endurance behavior.** | **Besides** | **the** | **continuous** |

**improvements in process technology which minimize trap densities, we here demonstrate a mitigation of device variability, using a targeted programming scheme, with which a significantly lower device variability is achieved.**

***Keywords— Device variability, Charge trapping, HfO2-based FeFET, Endurance failure, Targeted Programming***

I.INTRODUCTION

Since the first discovery of the ferroelectric properties of HfO2 thin films in 2007 originating from the non-centrosymmetric orthorhombic Pca21 crystal phase [1], [2] a significant interest has evolved for devices based on ferroelectric HfO2 (FE-HfO2) due to their numerous attractive features, including CMOS-compatibility, fast write time and low power operation. However, at the current stage, the HfO2-based FeFET devices are still not reaching their expected capability regarding endurance degradation and device variability. It is well known in the community that the degradation of the gate stack induced by charge trapping and trap generation is the main driving force for its endurance failure [3], [4]. Meanwhile the polycrystallinity of the deposited ferroelectric film can induce variability of film properties, which eventually leads to variation of device behavior as well [5], [6].

In order to analyze the deterioration of the gate stack caused by endurance cycling, a set of endurance tests with different time delays were implemented. In addition, to decrease the variability of the device performance across wafer, a targeted programming scheme is applied aiming at more uniform device performance at the wafer level.

II.EXPERIMENTAL DETAILS

*A.Device Fabrication*

The FeFET studied in this paper is embedded into GLOBALFOUNDRIES manufacturing proven 28 nm high-k metal gate (HKMG) super low power (SLP) CMOS platform on 300 mm wafers featuring a state of the art HKMG gate

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

stack comprising a 9 nm thick ferroelectric HfO2. A transmission electron microscopy (TEM) graph of the gate stack is shown in Fig. 1.

Gate

FE-HfO2

|  |  |
| --- | --- |
| Source | Drain |

50 nm

Fig. 1. TEM graph of the cross section of a HfO2-based FeFET fabricated a GLOBALFOUNDRIES 28 nm SLP technology platform.used for obtaining the experimental results in this paper

All measurements shown in this paper are performed on three different device geometries, that is 500 x 500 nm2, 200 x 200 nm2 and 80 x 80 nm2 correspondingly before shrinking by a factor of 0.9 according to the 28 nm technology PDK manuals. That means the final physical dimensions are 450 x 450 nm2, 180 x 180 nm2 and 72 x 72 nm2 correspondingly.

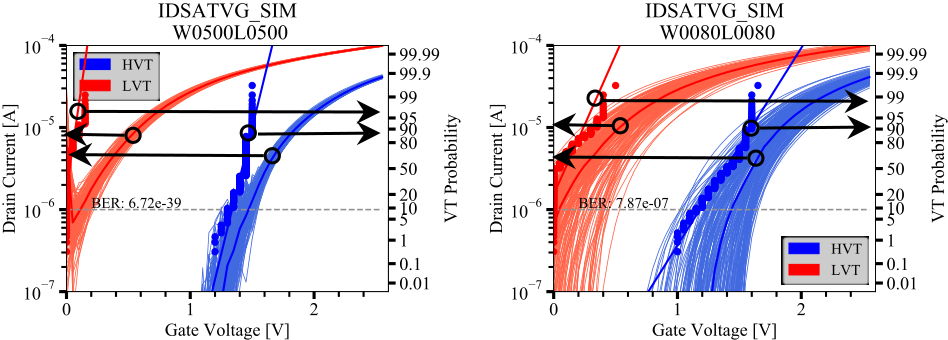
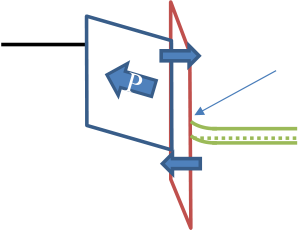
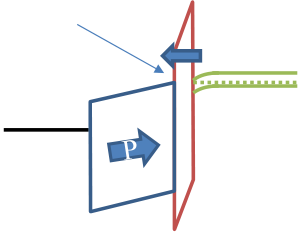
*B.Measurement Setup*

The test facility is equipped with a 300 mm fully automated wafer prober (Precio™ XL from Tokyo Electron) and a PXI-system from National Instruments for electrical measurement.

Within this measurement setup currents are measured from the source side of the FeFET and together with the applied gate voltage pulses the drain current gate voltage (*IdVg*)-curve can be obtained for each device, out of which the threshold voltage *Vth*was then extracted using a constant drain current criterion, in this case it is 1 µA \* *WG* (gate width) / *LG* (gate length). The whole wafer is split into two groups of 64 dies, one is cycled with 10 bipolar wake-up cycles and the other with 1000 bipolar wake-up cycles. Each group is then further split into two subgroups with each subgroup programmed into either low *Vth* (LVT) state or erased into the high *Vth*(HVT) state at the end of the test.

III.RESULTS AND DISCUSSION

The LVT and HVT extractions in this paper are executed after a 40 ms and 1 ms time delay after the program/erase pulse correspondingly. The criterion for choosing this specific time delay for the LVT and HVT extraction is, on the one hand, based on the short term retention test as shown in Fig. 5 which reveals electron and hole trapping behavior while on the other hand keeping the total test time at a reasonable range. As it



can be clearly seen in Fig. 5, the device shows a strong electron trapping and hardly any hole trapping. The LVT states become stable after about 40 ms while the HVT states are already stable after 40 µs due to negligible hole trapping. Thus, a time delay of 40 ms and 1 ms after program and erase pulse are considered appropriate to read the *Vth*by sweeping the gate voltage from 0 V to 2.5 V without being affected significantly by residual charge trapping.

In the following, recent advances and results on the HfO2-based FeFET development are reviewed, the memory window (MW) across wafer, trapping behavior and endurance characteristics are shown. Then, the endurance behavior with different delay times between program and erase pulse is compared and discussed. In the end, the effectiveness of a targeted programming scheme is verified.

after a write pulse (either program or erase) and after sufficient delay time, an additional pulse of same polarity compared to the write pulse but with variable amplitude is applied to the device. After this variable pulse, an immediate read-out is then performed. Since the ferroelectric switching has already occurred during the write pulse, the trapping behavior can then be measured independently. Fig. 7 shows that there is barely any hole trapping but a large amount of electron trapping independent of device geometries. This is attributed to the defect densities of the HfO2 material and the interface. The band diagram of the gate stack at both program and erase conditions (Fig. 3) shows that during program, electrons could potentially be injected into the gate stack and during erase process the electrons will tunnel back and some hole trapping can occur.

|  |  |  |
| --- | --- | --- |
| *A.Recent advances on FeFETs* | fter P reversa | fter P reversa |

The memory window wafer map of various device geometries after 10 bipolar wake-up cycles is shown in Fig. 4. It can be observed that with decreasing device size, the memory window decreases whereas the variability of LVT and HVT states increases. The reason for this could be that the deposited ferroelectric HfO2 film is only partially crystallized into the desired orthorhombic Pca21 phase due to the intrinsic polycrystalline characteristics of the film. Accordingly, statistical fluctuations in film composition might impact smaller devices more than larger devices. To get a deeper understanding of the influence from device sizes on its variability, a simulation of the FeFET with size 500 x 500 nm2 and 80 x 80 nm2 based on a grain model (described in detail in upcoming publications) is undertaken und the result is shown in Fig. 2 below.

|  |  |  |  |
| --- | --- | --- | --- |
| a) | 500 x 500 nm2 500 x 500 nm2 | b) | 80 x 80 nm2 |

Fig. 2. *IdVg* curves of simulated FeFETs based on 30 nm grain size. Dipole angle for each grain is fully randomly chosen between 0° and 90°, 0° means horizontal, i.e. no effect on Vth shift and 90° means vertical, i.e. maximum effect on Vth shift. In total 200 runs are executed and the read is swept from 0 V to 2.5 V.

As it can be seen from Fig. 2, the device size 500 x 500 nm2 has more grains and shows a smaller device variability compared to 80 x 80 nm2, which agrees well with the measurement data. The estimated bit error rate (BER) is 1.89e-19, 2.29e-5 and 1.12e-1 for device sizes of 500 x 500 nm2, 200 x 200 nm2 and 80 x 80 nm2 respectively as shown in Fig. 6. The calculation of BER assumes that the *Vth* distribution is normal and is calculated such that it represents the point at which extrapolated LVT normal distribution and extrapolated HVT normal distribution start to overlap. A closer look at the MW wafermap reveals that the cells with smaller MW can be found mainly around the wafer edge independent of the device geometries.

In order to study the influences of charge trapping on the endurance behavior of FeFET devices, we first must identify the trapping behavior of the film, thus, a set of trapping tests are implemented. The trapping test is implemented such that

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| order traps | - | e- trapping | - | e- bac tunne ing  - order traps  - generation  - |
| generation |
| - |
| - |

-

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ti | - | | Ti | htrapping | |
| FE-HfO2 | P-t pe ub | FE-HfO2 | P-t pe ub |
| nterfacia | | nterfacia | |
| a er | | a er | |

Fig.3. Band diagram of the FeFET gate stack at both program (+4 V for 10 µs) and erase (-4 V for 10 µs) condition.

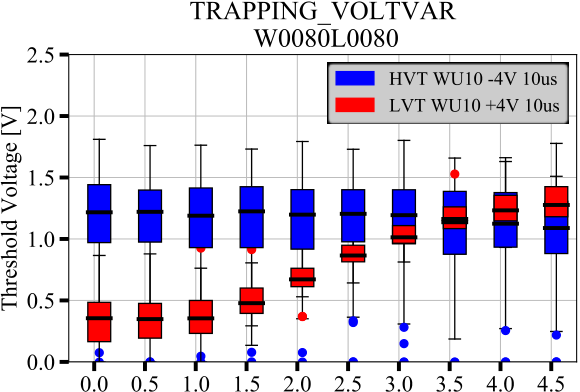
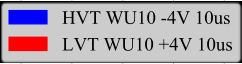
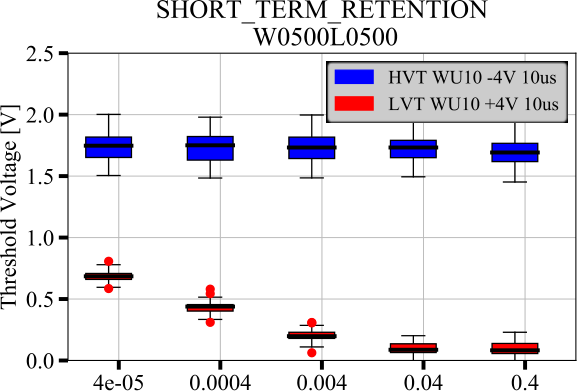
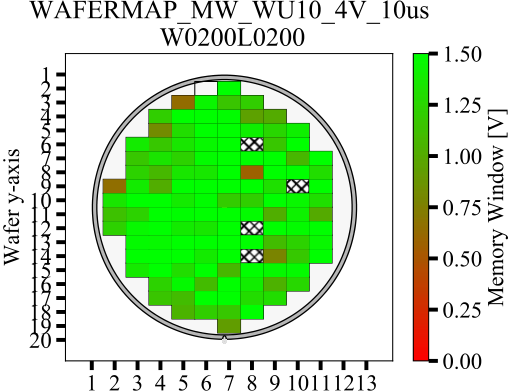
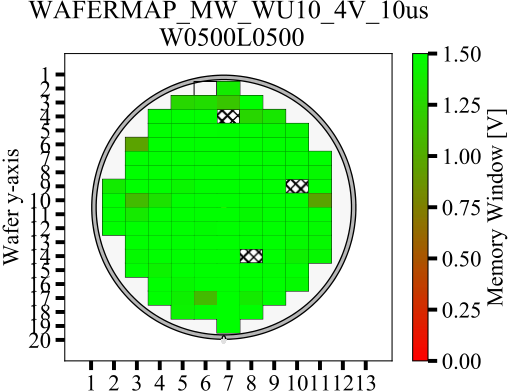
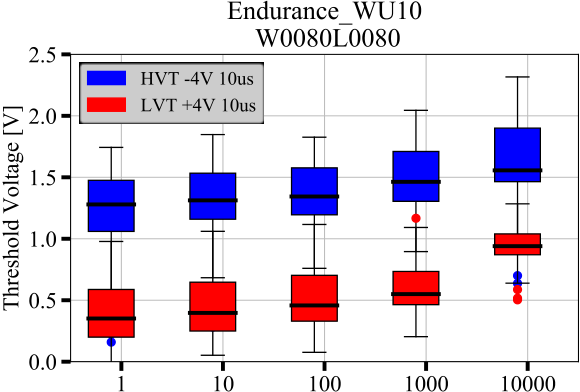
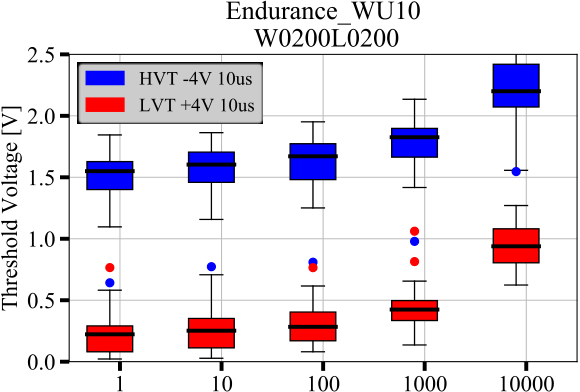
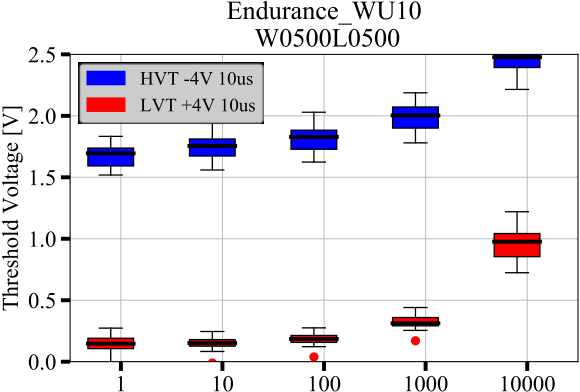
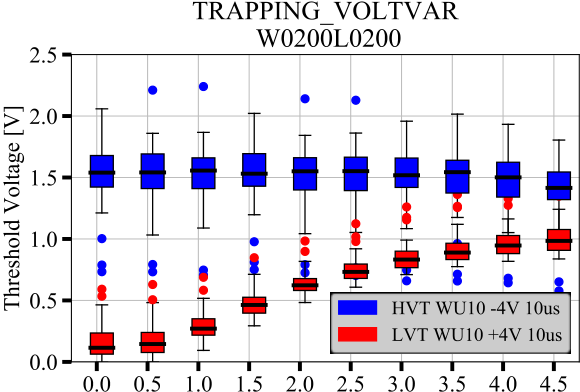
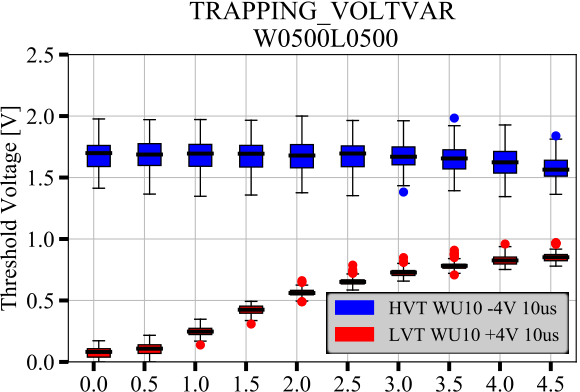
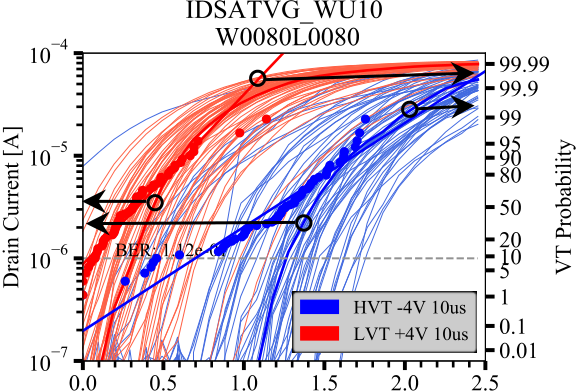
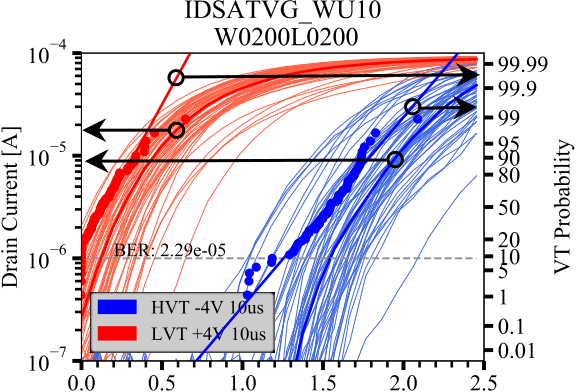
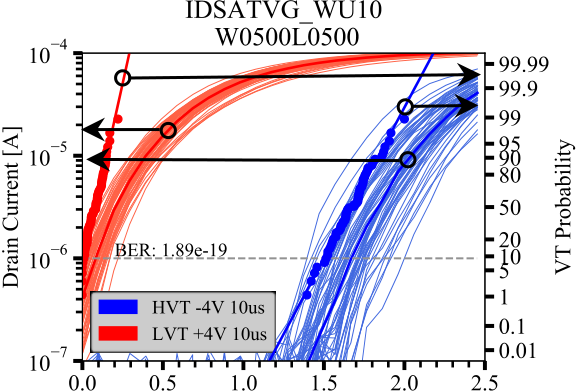
Hole trapping shifts the HVT to lower voltages whereas electron trapping shifts the LVT to higher voltages. Since this represents the opposite direction compared to the effect of polarization switching on *Vth*, charge injection leads to a MW decrease. As can be seen from Fig. 7, already starting from 1V, the electron trapping increases very rapidly and reaches its saturation at around pulse amplitudes of 3.5 V.

The endurance characteristics of the FeFET for 3 different device geometries is shown in Fig. 8. After 10 bipolar wake-up cycles the memory cell is stressed by an alternating program (+4 V for 10 µs) and erase (-4 V for 10 µs) pulse with a delay time of 10 µs after the program pulse. In between a specific number of bipolar cycles (in this case, 1, 10, 100, 1k and 10k) the cell is set into a programmed state, read out and then set into an erased state and read out to get the respective *Vth*. The boxplot from Fig. 8 shows the trend of both LVT and HVT movement as the device is cycled. Both move upwards due to the strong electron trapping behavior which cannot be compensated by hole trapping (see Fig. 7). During the cycling process the electrons which are trapped as a response to the positive program pulse can be released out of the gate stack or compensated by holes injected into the gate stack due to the subsequent negative erase pulse. The continuous injection and release of charges into the gate stack can lead to additional trap generation both within the HfO2 and at the interface between the interface oxide and the HfO2. [3]

*B.Passive detrapping for better endurance performance*

In order to get a better understanding of the correlation between charge trapping and endurance performance two sets of endurance tests were performed. Different delay times after the program and erase pulse were introduced and tests were performed separately. The tests were executed only on the 500 x 500 nm2 device geometry because this larger device geometry shows stronger *Vth* shifts during endurance cycling.

Results are shown in Fig. 9. As it can be clearly seen by



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Wafermap** | a) |  | b) |  |  |

Fig. 4. Wafermap of MW after applying 10 bipolar wake-up cycles for device geometries: a) 500 x 500 nm2 and b) 200 x 200 nm2. Greeen color indicates larger MW while red color shows smaller MW, squares with strips indicate contact problems. The program/erase conditions are +/- 4 V, 10 µs pulses

Fig. 5. Short term retention measurement of FeFETs with device geometry of 500 x 500 nm2. The program/erase conditions are +/-4 V for 10 µs.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IdVg and cumulative plots** | a) |  | b) |  | c) |  |

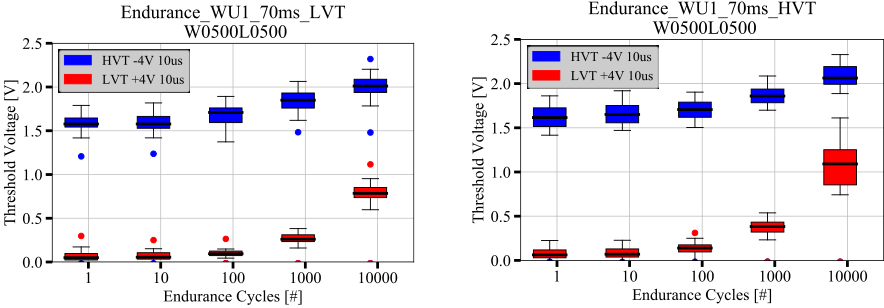
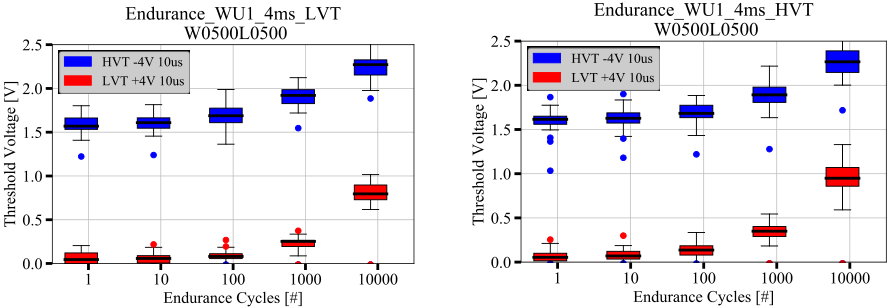
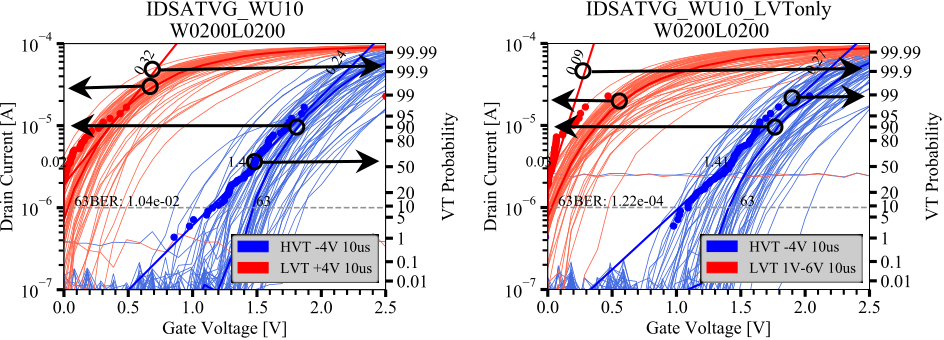
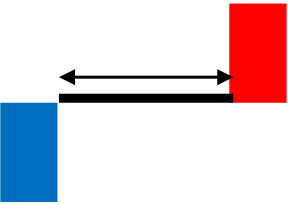
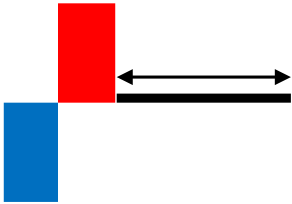
Fig. 6. *Id-Vg* curves and cumulative probability plot of Vth of the DUTs after applying a 10 bipolar wake-up cycles for different device geometries: a) 500 nm x 500 nm2, b) 200 nm x 200 nm2, and c) 80 nm x 80 nm2. Solid curves show the median *Id-Vg* curve from 62 devices aross wafer while solid lines show the fitting line of the probabililty plots. BER (Bit Error Rate) is calculated based on the assumption of a normal distribution of Vth across wafer. The program/erase conditions were chosen to be +/- 4 V, 10 µs.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Trapping** | a) |  | b) |  | c) |  |

Fig. 7. HVT and LVT trapping behavior of DUTs with 10 bipolar wake-up cycles for different device geometries: a) 500 x 500 nm2, b) 200 x 200 nm2, and c) 80 x 80 nm2. Blue boxes stand for the HVT devices while red boxes for the LVT devices, the median *Vth* value is shown in thick black line in the box. Trapping voltage is negative for HVT and positive for LVT. The program/erase conditions were chosen to be +/- 4 V, 10 µs pulses.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Endurance** | a) |  | b) |  | c) |  |

Fig. 8. Endurance test of DUTs with 10 bipolar wake-up cycles for different device geometries: a) 500 x 500 nm2, b) 200 x 200 nm2, and c) 80 x 80 nm2. Blue boxes stand for the HVT devices while red boxes for the LVT devices, the median *Vth* value is shown in thick black line in the box. The program/erase conditions for this bipolar cycling and endurance test are +/- 4 V, 10 µs pulses.



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 4 ms | a) | -4 V | +4 V | Tdelay | X N | c) | -4 V | Tdelay | +4 V | X N | Vth is below the targeted value, further program pulses will | |
| not be applied to the device anymore. For every test half of the | |
| 10 us | 10 us |
| wafer’s dies wi be chosen and 10 wake-up cycles at +/-4 | |
| V/10µs are performed. Thereafter a read-out is done und used | |
| 10 us | 10 us | as the reference state shown in Fig. 10a. Subsequently, the | |
| targeted programming scheme is applied to the device and the | |
| results are shown in Fig. 10b. The device geometry chosen is | |
| 200 x 200 nm2. As it is shown in Fig. 10b, a substantial | |
| improvement in standard deviation of LVT can be achieved | |
| by applying the targeted programming scheme to the device, | |
| thus the extrapolated BER can be improved by approximately | |
| 70 ms | b) | d) | two orders of magnitude. A more elaborate targeted | |
| programming involving time variation can be utilized in the | |
| future which should lead to even lower BER. | |
| a) | b) |

Fig. 9. Endurance test with two different time delays: 4 ms and 70 ms after

program pulse (a, b) and after erase pulse (c, d). The test sequence is shown above in the first row, the program/erase pulse for the endurance test is +/-4 V 10 µs and the time delay is added after program (left) and after erase (right) respectively.

comparing Fig. 9 to Fig. 8, longer delay time after program causes less upward shifting of the LVT and HVT states (Fig. 9a and 9b). A longer delay time after erase pulse mainly causes less upward shifting of the HVT state (see Fig. 9c and 9d). Especially the improvement in LVT shift with longer delay times after a program operation is attributed to the detrapping process. An active detrapping process by way of an immediate erase pulse after programming deteriorates the gate stack by hot holes injected from the silicon substrate. Hence less traps are generated in the gate stack if the detrapping process is not accelerated. This also explains why a longer delay time after an erase pulse is not preventing the LVT state from moving to higher voltages, since in this scenario an active detrapping process of trapped electrons can be assumed. However, for both extended delay times after program and after erase pulse, the HVT state showes less upward movement of about 200 mV during cycling. Referring back to the previous trapping test results, these already indicated that the film has barely any hole trapping. Hence, the general shift of HVT with cycling could be potentially due to deeply trapped electrons in the film which can be reduced by increasing delay after either program or erase pulse. In addition, the time delay after the program pulse leads to a significantly improved variability. To reduce the variability even further, a targeted programming scheme can be utilized (see next section).

Fig. 10. *IdVg* curves without (a) and with (b) targeted programming in comparison for 200 x 200 nm2. The readout for the graphs in a) is executed directly after 10 wake-up cycles, thereafter a targeted programming scheme is applied and the results are shown in b).

IV.CONCLUSION

In this paper some of the recent advancement of the embedded FeFET technology were shown and the more challenging aspects of this novel memory technology, i.e. endurance degradation and device variability with smaller device sizes were discussed. In order to get a deeper understanding about the correlation between endurance degradation and charge trapping, a set of endurance tests with different delay times (4 ms and 70 ms) were implemented. The results showed that if time delays after both program and erase pulses are introduced, both LVT and HVT shifts can be reduced. In addition, to decrease the device variability of smaller devices like 200 x 200 nm2 a targeted programming scheme with a sweeping range from +1 V to +6 V was applied to each device. The comparison to a single pulse write scheme showed that with targeted programming the device variability drops significantly and an improvement by orders of magnitudes in terms of BER can be achieved for scaled FeFETs.

REFERENCES

[1] T ösc e, “ r sta ine Hafnia and Zirconia ased Die ectrics for Memor pp ications,” 2010   
[2] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder, and U. Böttger, “Ferroe ectricit in hafnium oxide thin fi ms,” *Appl. Phys. Lett.*, vol. 99, no. 10, p. 102903, Sep. 2011, doi: 10.1063/1.3634052.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *C.Targeted programming for less device variability* | [3] | E. Yurchuk *et al.*, “Origin of the endurance degradation in the novel | | | | |
| HfO2-based 1T ferroelectric non-vo ati e memories,” *2014 IEEE Int.* | | | | |
| As it can be seen in Fig. 6, the device variability increases | [4] | *Reliab. Phys. Symp.*, pp. 2–2, 2014, doi: 10.1109/IRPS.2014.6860603. | | | | |
| as the device geometry gets smaller. In order to minimize this | E. Yurchuk *et al.*, “ harge-Trapping Phenomena in HfO₂-Based | | | | |
| [5] | FeFET-T pe onvo ati e Memories,” *IEEE Trans. Electron Devices*, | | | | |
| and get a lower BER, a targeted programming scheme can be |
| vol. PP, pp. 1–7, Jul. 2016, doi: 10.1109/TED.2016.2588439. | | | | |
| utilized. The targeted programming scheme is applied to the |
| J. Muller *et al*., "High endurance strategies for hafnium oxide based | | | | |
| LVT state only since in a AND structure of N-type FeFETs |
| ferroelectric field effect transistor," *2016 16th Non-Volatile Memory* | | | | |
| only a whole word line can be erased at one time. The test | [6] | *Technology Symposium (NVMTS)*, Pittsburgh, PA, 2016, pp. 1-7. | | | | |
| sequence is: The program pulse at the gate varies from 1 V all | H. Mulaosmanovic *et al*., "Evidence of single domain switching in | | | | |
| the way to 6 V in increments of 0.2V with respective pulse | hafnium oxide based FeFETs: Enabler for multi-level FeFET memory | | | | |
| cells," *2015 IEEE International Electron Devices Meeting (IEDM)*, | | | | |
| widths of 10 µs. 40 ms after each program pulse a read sweep |
| Washington, | DC, | 2015, | pp. | 26.8.1-26.8.3. |
| is performed, from which the Vth is extracted. If the extracted |
| doi: 10.1109/IEDM.2015.7409777 | | | | |