The Impact of Self-Heating on Charge Trapping in High-*k*-Metal-Gate nFETs

Faraz Khan, Eduard Cartier, Chandrasekara Kothandaraman, J. Campbell Scott, Jason C. S. Woo, *Fellow, IEEE*, and Subramanian S. Iyer, *Fellow, IEEE*

***Abstract*— In this letter, charge trapping behavior in 22-nm technology high-*k*-metal-gate SOI CMOS logic devices is analyzed under various bias stress and self-heating conditions. It is observed that the charge trapping is not only dependent on the channel power density during stress, which is controlled by drain bias and device channel length, but is also strongly modulated by the device channel width. Thus, identical power densities in devices with different channel widths result in significantly different charge trapping behaviors. It is shown that device self-heating is strongly influenced by the device channel width and that the channel temperature during the charge injection process significantly impacts the magnitude and stability of the trapped charge. We discuss the implications of the findings for the application of high-*k*-metal-gate logic devices as embedded memory elements for non-volatile data storage in high-*k*-metal-gate CMOS technologies without added process complexity.**

***Index Terms*— Charge trapping, CMOS, high-k-metal-gate, self-heating.**

I. Introduction

HfO_2 USED as gate dielectric in high-k-metal-gate (HKMG) CMOS technologies is known to have oxygen vacancy related traps [1]–[3]. Bias stress induced charge trapping and defect generation in HfO_2 is strongly accelerated by temperature [4], [5]. Recently, we have proposed the use of HKMG logic nFETs under modified operation conditions, resulting in enhanced change trapping, for non-volatile data storage [2]; functional product prototype 64kb memory arrays were constructed and reliability was demonstrated. In this letter, we present the fundamentals of enhanced charge trapping and enhanced charge stability under the proposed operation conditions. We demonstrate the impact of device self-heating on charge trapping behavior in HKMG logic transistors and show that the magnitude as well as stability of the trapped charge is modulated by device self-heating, effects that can be used to optimize device design and operation conditions for the memory application.

Manuscript received November 17, 2015; accepted November 30, 2015. Date of publication December 2, 2015; date of current version December 24, 2015. The review of this letter was arranged by Editor T. Wang.

F. Khan was with IBM, Semiconductor R&D Center, Hopewell Junction, NY 12533 USA. He is now with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA 90095 USA, and also with GlobalFoundries, Advanced Technology Development, Malta, NY 12020 USA (e-mail: farazkhan@ucla.edu).

E. Cartier and C. Kothandaraman are with the Thomas J. Watson Research Center, IBM Research, Yorktown Heights, NY 10598 USA.

J. C. S. Woo and S. S. Iyer are with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA 90095 USA.

J. C. Scott is with the Almaden Research Center, IBM Research, San Jose, CA 95120 USA.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2015.2504952

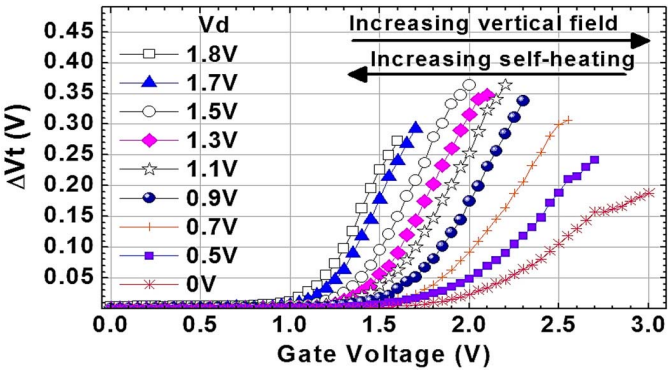

Fig. 1. ΔV_T as a function of PVRS stress with 10ms pulses at various fixed Vd's (W_{ch} = 1.04um, L = 20nm).

The self-heating enhanced charge trapping is found to be stable enough to consider the resulting device threshold voltage shifts (ΔV_T) as a mechanism for non-volatile data storage.

II. Experimental Details

Experiments were performed on devices fabricated on state-of-the-art 22nm high-performance SOI technology platform [6]. First, ΔV_T's are measured during pulsed gate voltage ramp sweeps (PVRS) for various fixed drain bias (V_d) conditions. Gate bias (V_g) was applied using 10ms pulses of increasing magnitudes in 50mV increments. After each pulse, the device V_T was measured within 10ms. Each device was ramped until breakdown and Fig. 1 shows the measured ΔV_T's until before breakdown. Details on the PVRS technique can be found in [7]. The pre-stress V_T of each device is ~280mV. We make two observations; Firstly, at higher V_d's (higher lateral field and self-heating), equivalent ΔV_T's are achievable at substantially lower V_g's. This is attributed to the impact of an enhanced level of hot carrier injection and charge trapping with increasing V_d as well as to enhanced charge trapping due to device self-heating [8] with increasing V_d. Secondly, the maximum achievable ΔV_T before device breakdown initially increases and then starts to decrease with increasing V_d. The breakdown of devices under low V_d conditions is electric field driven (high gate-to-drain bias, V_{gd}) whereas the breakdown of devices under high V_d conditions (which happens at much lower V_{gd}) is self-heating driven, which is a well-known phenomenon [9]. Shifts in ΔV_T vs. V_g trends before hard breakdown may be indicative of the beginning of soft breakdown [10].

Identical stress pulses (35ms at V_g = 2V and V_d = 1.3V) were then applied to devices with the same channel width (W_{ch}) but various channel lengths (L) and the V_T's

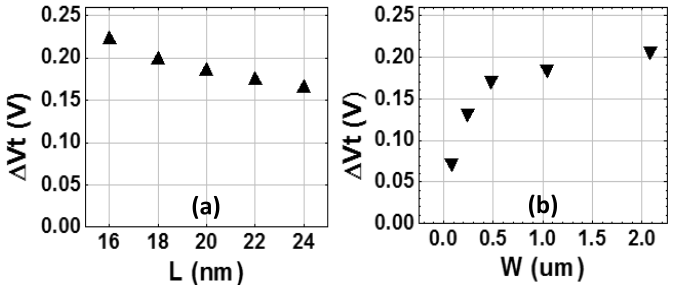

Fig. 2. Measured ΔV_T vs. (a) device channel length ($W_{ch} = 1.04$um) and (b) device channel width ($L = 20$nm).

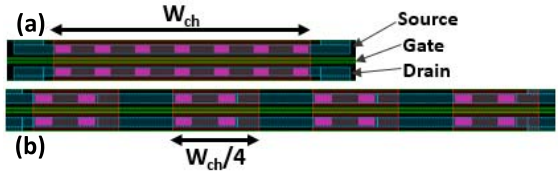

Fig. 3. (a) Single-finger and (b) Multi-finger device layout.

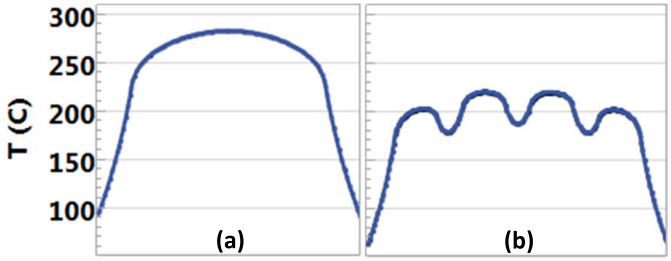

Fig. 4. Steady-state thermal profiles for a (a) single-channel and (b) multi-channel device in the W_{ch} direction, for an applied power of 4mW/um.

were measured within 10ms. It is seen that ΔV_T increases as L decreases (Fig. 2(a)), which is expected and consistent with increasing levels of hot carriers and self-heating (due to increase in lateral field) and decreasing V_T (due to short-channel effects) with decreasing L. However, our results also show that when identical stress pulses are applied to devices with the same L and various W_{ch}, ΔV_T increases with W_{ch} (Fig. 2(b)). This phenomenon of ΔV_T varying with W_{ch} (even when vertical and lateral fields and L are the same) is not readily explained by a uniform injection mechanism along W_{ch} and is attributed to the impact of self-heating, which is strongly modulated by W_{ch}. In this letter, the chuck temperature was always at 25C unless otherwise stated.

III. Thermal Simulations for Self-Heating

To better understand and quantify this phenomenon and to separate the impact of electric field from the thermal effects, single-finger devices vs. multi-finger (split-channel) devices were studied. Both devices have a total W_{ch} of 1.04um where each channel in the multi-finger devices, separated by a trench isolation, has a width of $W_{ch}/4$. Both devices are identical to each other except for the channel width and have a channel length of 20nm. Layouts of the two devices are shown in Figs. 3(a) and 3(b), respectively. First, channel thermal profiles of the two devices were analyzed. Thermal simulations were carried out using finite element analysis (Comsol™). Full 3D structural simulations of the devices are analyzed and solved

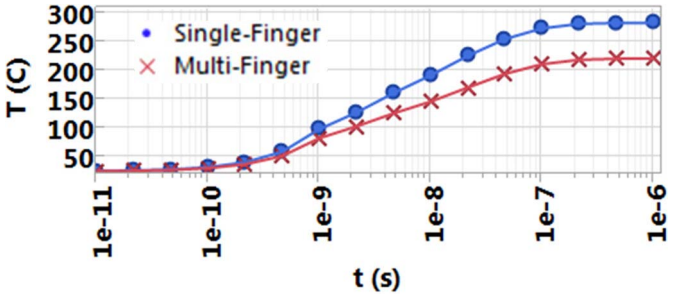

Fig. 5. Rise in channel temperature vs. stress time (4mW/um applied power).

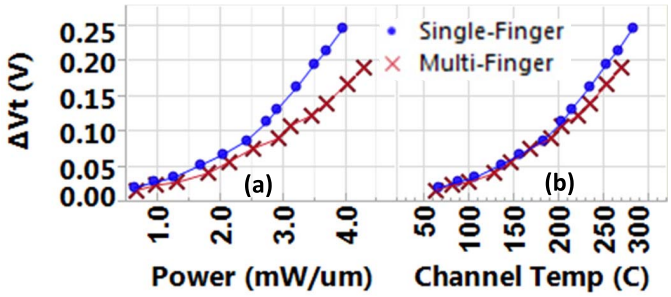

Fig. 6. Measured ΔV_T vs. (a) applied power density and (b) channel T during stress. It is observed that, at higher T's, the rate of increase in ΔV_T is higher.

for temperature distribution and heat flux. Figs. 4(a) and 4(b) show the channel temperature (T) profiles of the two devices for an applied power density of 4mW/um. It is clear that W_{ch}/active area significantly modulates device self-heating. In multi-finger devices, the area for vertical heat flow is effectively larger than the area for power dissipation. Additionally, the larger area for lateral heat dissipation and the higher number of contacts per unit width in multi-finger devices are also responsible for higher heat dissipation and thus a lower thermal resistance (R_{th}) as compared to single-finger devices. The extracted R_{th} for the single-finger device is $\sim 1.3\times$ compared to the multi-finger device (65.9 vs. 50.8 K/mW, respectively). Additionally, simulation results show that the devices reach thermal equilibrium within several hundred ns (Fig. 5).

Fig. 6(a) shows the measured ΔV_T vs. applied power density for devices that were used for the thermal simulations. The power was varied by varying V_d while $V_g = 2$V. It is seen that, for the same power density, ΔV_T for the single-channel device is considerably higher as compared to the split-channel device and the difference is greater at higher power densities. However, when plotted as a function of the calculated channel temperature (Fig. 6(b)), the ΔV_T characteristics of the two devices are almost identical except at very high temperatures where the single-channel device seems to have slightly higher ΔV_T. In other words, ΔV_T behaviors of the devices show a much stronger correlation to the self-heating temperature as compared to the applied power density. It is concluded from these results that the device self-heating temperature is a significant factor in modulating the charge trapping behavior.

IV. Charge Stability and Detrapping Characteristics

To evaluate the charge detrapping behavior, a set of identical devices was stressed at various fixed V_d's to achieve a

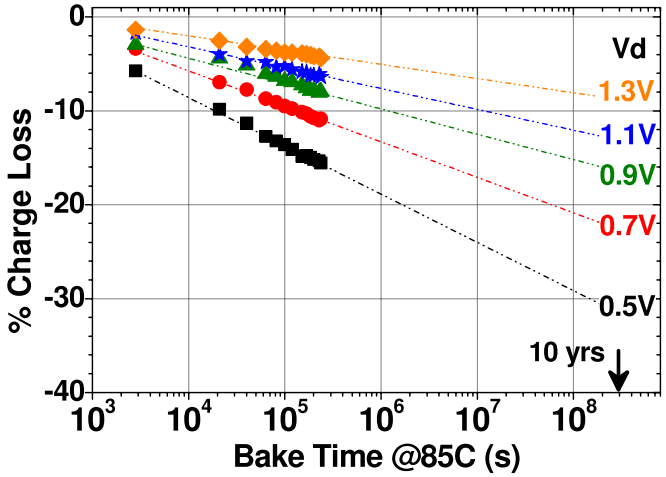

Fig. 7. Percentage charge loss vs. bake time @ 85C, for identical devices stressed at various fixed drain biases (W_{ch} = 1.2um, L = 20nm).

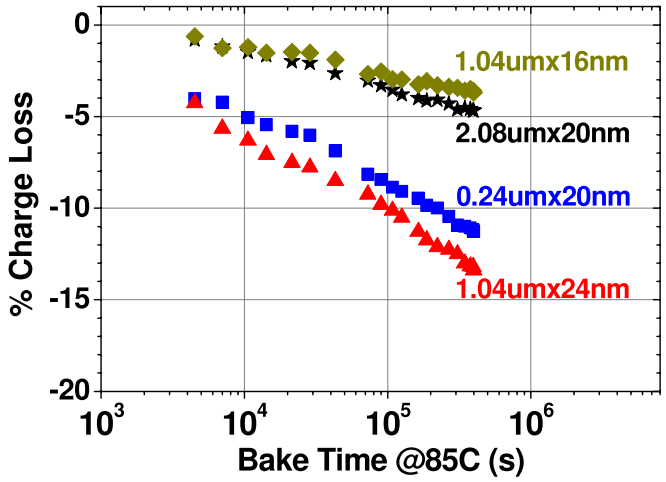

Fig. 8. Percentage charge loss vs. bake time @ 85C, for devices with various dimensions (W_{ch} × L, as labelled) stressed at V_d = 1.5V.

cumulative ΔV_T of ~250mV in each device, using PVRS, and then stored at an elevated temperature of 85C. Retention of the trapped charge in each of the devices was measured by monitoring the device V_T's as a function of time. The reduction in V_T's (loss of trapped charge) is plotted as a percentage of the initial values as shown in Fig. 7. It is observed that retention of the trapped charge shows a positive correlation to the stress drain bias, V_d.

Another set of devices with different channel widths (same length) and different channel lengths (same width) was stressed using PVRS at V_d = 1.5V to achieve a cumulative ΔV_T of ~265mV in each device and then stored at 85C. The retention of the trapped charge was measured as described above and is shown in Fig. 8. As can be seen, the trapped charge in wider and shorter devices has higher retention. The enhanced charge retention in wider devices is attributed to higher self-heating. The enhanced charge retention in shorter devices is attributed to a cumulative effect of higher self-heating due to higher power densities as discussed below and elevated levels of hot carriers due to higher lateral fields.

The higher stability of charge trapped at high temperatures (device self-heating induced), as compared to charge trapping at room temperature [1], [11], can be attributed to

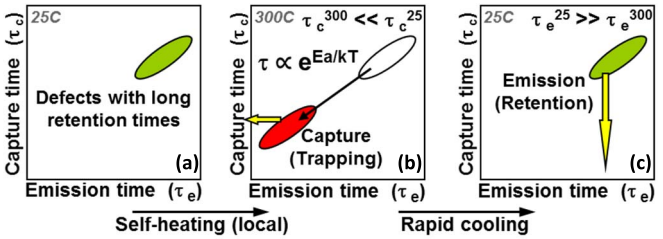

Fig. 9. Schematic of 'Capture-Emission Time Maps' for self-heating assisted charge trapping (adapted from [12]). (a) Defects with long emission times/good retention also have long capture times, (b) Capture times are reduced at elevated temperatures, and (c) Rapid quenching retains charge in defects with long emission times at low temperatures.

the fundamental nature of charge trapping and detrapping, which are thermally activated processes, wherein the capture and emission times of the trapped charge are directly correlated to their activation energies [12]. This is illustrated in Fig. 9. At low temperatures, stable traps with high activation energies (long capture times) require longer times to be filled (Fig. 9(a)). Self-heating induced high temperature enables access to these stable traps in shorter times, and they can be rapidly filled during the charge injection (Fig. 9(b)). Localization of self-heating leads to rapid cooling (in the ns range) after the stress conditions are removed, preventing charge detrapping as activation energies for the same can no longer be achieved, resulting in long emission times and enhanced retention (Fig. 9(c)). This understanding is consistent with the known properties of distributed oxide traps such as oxygen vacancies [12].

V. Summary and Conclusions

Charge trapping behavior in HKMG SOI CMOS devices has been analyzed and it is found that not only the magnitude but also the stability of the trapped charge increases with device self-heating during the charge injection process. The same magnitude of charge trapping can be achieved in much shorter times and has higher stability when the devices are stressed at higher drain biases or higher self-heating conditions. Device geometry and layout significantly modulate R_{th} and in turn the self-heating and charge trapping behavior.

The promising charge retention characteristics that have been demonstrated (extrapolated accelerated charge loss of <10% after 10 years at 85C) make it feasible to further explore the implementation of HKMG CMOS based devices as memory elements for scalable fully integrated system-on-chip applications such as embedded non-volatile memory (NVM) and potential alternative to existing one-time programmable technologies like eFUSE [13] for yield improvement, performance tailoring, field configurability, and data security enhancement. Further research work on HKMG CMOS device based NVM and its optimization is ongoing.

References

[1] E. Cartier, B. P. Linder, V. Narayanan, and V. K. Paruchuri, "Fundamental understanding and optimization of PBTI in nFETs with SiO_2/HfO_2 gate stack," in *Proc. IEDM*, Dec. 2006, pp. 1–4.

[2] C. Kothandaraman, X. Chen, D. Moy, D. Lea, S. Rosenblatt, F. Khan, D. Leu, T. Kirihata, D. Ioannou, G. LaRosa, J. B. Johnson, N. Robson, and S. S. Iyer, "Oxygen vacancy traps in Hi-K/metal gate technologies and their potential for embedded memory applications," in *Proc. IEEE IRPS*, Apr. 2015, pp. MY.2.1–MY.2.4.

[3] H. Hamamura, T. Ishida, T. Mine, Y. Okuyama, D. Hisamoto, Y. Shimamoto, S. Kimura, and K. Torii, "Electron trapping characteristics and scalability of HfO_2 as a trapping layer in SONOS-type flash memories," in *Proc. IEEE IRPS*, Apr./May 2008, pp. 412–416.

[4] E. Cartier and A. Kerber, "Stress-induced leakage current and defect generation in nFETs with HfO_2/TiN gate stacks during positive-bias temperature stress," in *Proc. IEEE IRPS*, Apr. 2009, pp. 486–492.

[5] F. Crupi, R. Degraeve, A. Kerber, D. H. Kwak, and G. Groeseneken, "Correlation between stress-induced leakage current (SILC) and the HfO_2 bulk trap density in a SiO_2/HfO_2 stack," in *Proc. 42nd Annu. IEEE IRPS*, Apr. 2004, pp. 181–187.

[6] S. Narasimha, P. Chang, C. Ortolland, D. Fried, E. Engbrecht, K. Nummy, P. Parries, T. Ando, M. Aquilino, N. Arnold, R. Bolam, J. Cai, M. Chudzik, B. Cipriany, G. Costrini, M. Dai, J. Dechene, C. DeWan, B. Engel, M. Gribelyuk, D. Guo, G. Han, N. Habib, J. Holt, D. Ioannou, B. Jagannathan, D. Jaeger, J. Johnson, W. Kong, J. Koshy, R. Krishnan, A. Kumar, M. Kumar, J. Lee, X. Li, C. Lin, B. Linder, S. Lucarini, N. Lustig, P. McLaughlin, K. Onishi, V. Ontalus, R. Robison, C. Sheraw, M. Stoker, A. Thomas, G. Wang, R. Wise, L. Zhuang, G. Freeman, J. Gill, E. Maciejewski, R. Malik, J. Norum, and P. Agnello, "22 nm high-performance SOI technology featuring dual-embedded stressors, epi-plate high-*k* deep-trench embedded DRAM and self-aligned via 15LM BEOL," *Proc. IEEE IEDM*, Dec. 2012, pp. 3.3.1–3.3.4.

[7] A. Kerber, S. A. Krishnan, and E. A. Cartier, "Voltage ramp stress for bias temperature instability testing of metal-gate/high-*k* stacks," *IEEE Electron Device Lett.*, vol. 30, no. 12, pp. 1347–1349, Dec. 2009.

[8] P. Su, K.-I. Goto, T. Sugii, and C. Hu, "Excess hot-carrier currents in SOI MOSFETs and its implications," in *Proc. 40th Annu. IRPS*, 2002, pp. 93–97.

[9] D. A. Dallmann and K. Shenai, "Scaling constraints imposed by self-heating in submicron SOI MOSFET's," *IEEE Trans. Electron Devices*, vol. 42, no. 3, pp. 489–496, Mar. 1995.

[10] N. Raghavan, K. L. Pey, and K. Shubhakar, "High-κ dielectric breakdown in nanoscale logic devices—Scientific insight and technology impact," *Microelectron. Rel.*, vol. 54, no. 5, pp. 847–860, 2014.

[11] G. Bersuker, J. H. Sim, C. S. Park, C. D. Young, S. Nadkarni, R. Choi, and B. H. Lee, "Mechanism of electron trapping and characteristics of traps in HfO_2 gate stacks," *IEEE Trans. Device Mater. Rel.*, vol. 7, no. 1, pp. 138–145, Mar. 2007.

[12] T. Grasser, P.-J. Wagner, H. Reisinger, T. Aichinger, G. Pobegen, M. Nelhiebel, and B. Kaczer, "Analytic modeling of the bias temperature instability using capture/emission time maps," in *Proc. IEEE IEDM*, Dec. 2011, pp. 27.4.1–27.4.4.

[13] C. Kothandaraman, S. K. Iyer, and S. S. Iyer, "Electrically programmable fuse (eFUSE) using electromigration in silicides," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 523–525, Sep. 2002.