Effects of Thermal Treatments on the Trapping Properties of HfO$_2$ Films for Charge Trap Memories

Sabina Spiga, Francesco Driussi, Alessio Lamperti, Gabriele Congedo, and Olivier Salicio

Effects of Thermal Treatments on the Trapping Properties of HfO₂ Films for Charge Trap Memories

Sabina Spiga*, Francesco Driussi¹, Alessio Lamperti, Gabriele Congedo, and Olivier Salicio

Laboratorio MDM, IIM-CNR, Via C. Olivetti 2, 20864 Agrate Brianza (MB), Italy
¹DIEGM—University of Udine, Via delle Scienze 206, 33100 Udine, Italy

Received November 23, 2011; accepted December 30, 2011; published online January 25, 2012

The charge trapping properties of HfO₂ thin films for application in charge trap memories are investigated as a function of high-temperature postdeposition annealing (PDA) and oxide thickness in the TaN/Al₂O₃/HfO₂/SiO₂/Si stack. The trap density (Nₜ) in HfO₂, extracted by simulating the programming transient, is in the 10¹³–10¹⁴cm⁻² range, and it is related to film thickness and PDA temperature. Diffusion phenomena in the stack play a significant role in modifying Nₜ in HfO₂ and the insulating properties of the Al₂O₃ layer. The memory performances for 1030°C PDA are promising with respect to standard stacks featuring Si₃N₄. ©2012 The Japan Society of Applied Physics

**Fig. 1.** XRR curves and fit (a) and GIXRD (b) patterns of Si/SiO₂/HfO₂/Al₂O₃ stack as a function of PDA temperature. In (b), reference patterns of monoclinic and orthorhombic HfO₂ are also shown [57385 and 71354 of ICSD 2010].

Charge trapping (CT) memory devices, where the information is stored in localized traps in thin dielectrics, such as Si₃N₄, are currently considered promising candidates to substitute the standard floating gate concept beyond the 16 nm technology node and to enable three-dimensional (3D) integration.⁰¹ Promising results have already been reported, especially for the TaN/Al₂O₃/Si₃N₄/SiO₂/Si (TANOS) stack,⁰² but the introduction of high-dielectric-constant (high-k) oxides as trapping layer alternative to Si₃N₄ could further improve the device performance in terms of trapping efficiency and equivalent oxide thickness (EOT) scaling.⁰³⁻⁰⁵ Indeed, the trapping efficiency of Si₃N₄ decreases with film thickness,⁰⁶ posing a limit to the further scaling of the nitride layer below 4 nm. It is also worth noting that in TANOS devices, the introduction of a high-work-function metal gate (such as TaN) and high-k oxides (such as Al₂O₃) as blocking layers represent consolidated solutions to improve program/erase efficiency and limit erase saturation with respect to Si/SiO₂/Si₃N₄/SiO₂/Si devices.² In particular, crystallized Al₂O₃ is the material providing the best compromise in terms of k-value (~9, higher than that of SiO₂), low trap-assisted tunnelling (TAT) leakage current through the blocking oxide, and reasonably large band gap and high conduction band offset (CBO ~ 2.5 eV) with respect to silicon.³⁹,₁⁰ However, given the high temperature (>900 °C) required to achieve crystalline Al₂O₃, the investigation of Al₂O₃/high-k/SiO₂ oxide multilayer properties and thermal stability as a function of high temperature treatments is of high interest for TANOS-like applications.

In this work, we investigate the charge trapping properties of HfO₂ single layers (k ~ 16) deposited by atomic layer deposition (ALD) and integrated in a TaN/Al₂O₃/HfO₂/SiO₂/Si (TAHOS) structure by combining physical analyses, electrical characterization of memory capacitors, and modelling of programming transients, as a function of HfO₂ thickness (ₜHfO₂) and postdeposition annealing (PDA) temperature in the 900–1100 °C range. By controlling and understanding the effects of ₜHfO₂ on charge trapping phenomena in the stack, we demonstrate the superior characteristics of HfO₂, which exhibits better charge trapping properties than Si₃N₄, without using engineered trapping layers.⁵

TAHOS stacks are composed of a 4.5-nm-thick thermally grown SiO₂ tunnel layer, plus HfO₂ ~6–16 nm nominal thickness), and Al₂O₃ (16.5 ± 0.1 nm thick, roughness σ = 0.8 ± 0.1 nm) layers deposited by ALD at 300 °C in a Savannah reactor (Cambridge Nanotech) using (MeCp)₂Hf(Me)(OMe), TMA and O₂ as Hf, Al and oxygen sources, respectively.¹¹ Stack PDA was performed in N₂ atmosphere at 900, 1030, or 1100 °C, followed by 15 nm TaN/30 nm W metal gate deposition by sputtering (capacitor area of 8 × 10⁻⁴ cm²). The metal gate underwent PDA at 900 °C in N₂ plus forming gas annealing at 400 °C. A reference TANOS capacitor (6 nm Si₃N₄) was also fabricated.¹¹

The structural and chemical properties of oxide multilayer stacks were characterized by X-ray reflectivity (XRR), grazing-incidence X-ray diffraction (GIXRD), and time of flight secondary ion mass spectrometry (ToF-SIMS) analyses. The XRR data and corresponding fit [Fig. 1(a) and Table I] are consistent with a trilayer structure (SiO₂/HfO₂/Al₂O₃) for all PDAs, with different ₜHfO₂ (HfO₂ electron density ρ₀ = 2.35 ± 0.05 e⁻/Å³) and same SiO₂ (4.5 nm thick, σ = 0.3 ± 0.1) and Al₂O₃ layers (14 ± 0.1 nm thick, σ = 0.8 ± 0.1 nm). Al₂O₃ shrinkage (i.e., ~15%) and densification [electron density increases from 0.95 to 1.17 (±0.05) e⁻/Å³] after high-temperature PDA are consistent with its crystallization in the γ phase.⁹,₁⁰ It is worth noting that while 900 °C is the onset temperature of Al₂O₃ crystallization, only PDA ≥ 1000 °C leads to full Al₂O₃ crystallization and better electrical characteristics.⁶,⁶ Concerning the HfO₂ trapping layer, the results of GIXRD
analyses in Fig. 1(b) show that the as-deposited 6 nm film is amorphous, whereas layers with thickness \( \geq 10 \) nm are already crystallized in a mixture of monoclinic and orthorhombic phases. PDA at 900°C leads to the crystallization of the 6 nm HfO\(_2\); for all stacks, the increase in PDA temperature leads to an increase in the monoclinic component. ToF-SIMS analyses were further performed in order to verify the chemical composition of the stack (Fig. 2). Each oxide layer is uniform throughout the film thickness, and the trilayer structure, evident in the as-deposited stack, is preserved for all PDAs without significant intermixing, in agreement with XRR data and fitting. On the other hand, changes in the tail observed in the ToF-SIMS profiles in Fig. 2(a) as a function of PDA indicate that diffusion phenomena took place during the thermal treatments. In particular, for PDA above 1030°C, Al diffusion into HfO\(_2\), and Hf diffusion into Al\(_2\)O\(_3\) is revealed, being stronger for 1100°C [Fig. 2(a)]. At the latter temperature, Si diffusion into HfO\(_2\) and Al\(_2\)O\(_3\) also takes place [Fig. 2(c)]. The consequent incorporation of dopant atoms in the charge trapping or blocking layers could affect the electrical properties and deserves careful consideration. Although quantitative information cannot be addressed from our ToF-SIMS data, we analysed the variation of the Al\(_{18}\)/\(^{18}\)O signal intensity ratio close to the HfO\(_2\)/SiO\(_2\) interface [Fig. 2(b)], using the as-deposited stack as a reference, to have more insight in the Al diffusion into the charge trapping layer. For the 16 nm films almost no variation is detected after 900°C, while the increase of Al\(_{18}\)/\(^{18}\)O signal ratio at 1030 and 1100°C demonstrates the increased Al incorporation into the HfO\(_2\) charge trapping layer along the full film thickness. The diffusion phenomena are enhanced in thin films [Fig. 2(b), squares], also likely due to a reduced grain size in the 6 nm HfO\(_2\) layer as shown by differences in XRD peak broadening [Fig. 1(b)].

Program and erase (P/E) transients are acquired for all TAHOS and reference TANOS capacitors [Figs. 3(a) and 3(b)] at gate voltages \( V_G \) between 12 and 20 V. Large memory windows \( > 10 \) V (interesting for multilevel programming) could be achieved using large \( |V_G| \) and long pulses (especially for the 16-nm-thick HfO\(_2\)), as well as very fast programming up to \( \Delta V_{FB} \approx 5 \) V for a 100 \( \mu \)s pulse and \( V_G = 20 \) V [inset of Fig. 3(a)]. For a given electric field applied to the tunnel oxide (\( F_{TO} \)) and PDA [Fig. 3(a)], with increasing \( t_{HfO2} \), the device speed is improved and larger flat band shifts (\( \Delta V_{FB} \)) are achieved, demonstrating the improved trapping capability of thick films, also with respect to the reference TANOS devices. For a given \( F_{TO} \) and \( t_{HfO2} \) [Fig. 3(b)], the programming efficiency of TAHOS stacks is slightly influenced by the PDA [except for a reduced speed at long pulse times and 1100°C PDA; but it is more evident for the 6 nm HfO\(_2\), inset of Fig. 3(b)]. On the other hand, erase efficiency increases with PDAs, likely due to better Al\(_2\)O\(_3\) insulating properties, which limits the electron injection from the gate and/or to large current through the tunnel oxide.\(^3\)\(^8\)
To better understand the trapping behavior of the fabricated HfO₂, experimental curves are compared with the results of the electrical model for program transients developed in refs. 11 and 12 [Fig. 3(a), dashed lines]. We note that at very large Vₑ and long programming time (t), the program curves slightly flatten (especially in the case of thin films). This effect is due to the saturation of the available traps, as demonstrated in Fig. 3(c), which shows the simulated trapped charge distribution along the HfO₂ thickness with increasing t of the 6 nm TAHOS device (PDA of 900 °C). In this respect, we modelled the trapping in the HfO₂ as a bulk effect, in agreement with previous works on CT memories. For large t, the trapped charge clearly saturates at the value of the trap density (Nₑ) and this causes a slowing down of the program curves seen in Fig. 3(a). Hence, the number of available traps is the limiting factor for the program speed at large t, and, by using Nₑ as a fitting parameter to reproduce the programming transients of the different devices, it is possible to study the influence of PDA on the trapping properties of HfO₂. As shown in Fig. 3(d), the 900 °C PDA curves are reproduced with constant Nₑ = 4.2 × 10^{19} cm⁻³ for all t_{HfO₂} values, suggesting that HfO₂ trap concentration is not dependent on t_{HfO₂}. The extracted Nₑ values are of the same order of magnitude as those reported in the literature for Si₃N₄. On the other hand, an overall Nₑ increase in HfO₂ is extracted for PDA at 1030 °C [Nₑ in the (4.5–8) × 10^{19} cm⁻³ range], likely due to the incorporation of Al atoms into the HfO₂ layer (Fig. 2). Indeed, the incorporation of trivalent dopant elements in HfO₂ could affect the formation of oxygen vacancies, which are considered to be the defects responsible for charge trapping in HfO₂. Moreover, the observed additional dependence of Nₑ on t_{HfO₂} (for 1030 °C PDA) could be related to the enhanced diffusion phenomena in thin films [Fig. 2(b)], where the Al incorporation extends to all film thicknesses. The good agreement between the model and experiments validates the assumption of a trapping in HfO₂ mainly due to a bulk effect, although it is not possible to exclude the finding that the increased overall trap density in thin films could also be partly related to a larger contribution of the interfaces, where additional defect centers could be created during the diffusion phenomena.

The extraction of Nₑ by PDA at 1100 °C is not illustrated in Fig. 3(d) since, for this PDA, a reliable quantitative extraction of the trap density is not possible. In fact, the noteworthy flattening of the experimental curves at large t [Fig. 3(b), inset] is not due to the saturation of the available HfO₂ traps only, but is also likely contributed by a large TAT through the Al₂O₃ layer. This is confirmed by the significant retention degradation with respect to devices annealed at 900 and 1030 °C [Fig. 4(a)], and it is likely related to the observed Hf and Si diffusion into Al₂O₃. All the above findings indicate that PDA at 1030 °C is the best condition to achieve large P/E windows, reasonable retention, and high trap density. Moreover, retention characteristics of TAHOS stacks are promising compared with reported results on stacks incorporating HfO₂ or engineered charge trapping layers, and they improve with increasing t_{HfO₂}, coming comparable to those of the reference TANOS for the 16 nm-thick HfO₂ [Fig. 4(b) and inset].

In summary, we demonstrated that the use of HfO₂ (with k value larger than Si₃N₄) in a TAHOS stack allows EOT scaling and improves the program efficiency and retention of memory devices. The trapping properties of HfO₂ are investigated by means of experimental analyses and using an electrical model. The Nₑ value in HfO₂ is in the 10^{19}–10^{20} cm⁻³ range, and it is related to film thickness and PDA temperature. Diffusion phenomena, especially for PDA > 1000 °C, play a significant role in modifying Nₑ in HfO₂ and the insulating properties of the Al₂O₃ blocking layer. Al incorporation into HfO₂ at 1030 °C increases the trap density, whereas Si and Hf diffusion into Al₂O₃ at 1100 °C lead to enhanced TAT conduction in the blocking layer, degrading retention and program efficiency. We therefore demonstrated that SiO₂/10–16 nm HfO₂/Al₂O₃ stacks annealed at 1030 °C represent a viable solution for improving TANOS performance with a good trade-off between the P/E window and retention. The possibility of using a single high-k layer, instead of an engineered multilayer trapping layer, is an additional advantage for ultra-scaled devices.

Acknowledgments The work was partially funded by the EU FP7-GOSSAMER project No. 214431 and by the IUNET consortium. The authors thank Numonyx for providing Si/SiO₂/Si and HfN stacks and Professor L. Selmi for manuscript revision and helpful discussion.

1) ITRS Roadmap 2010 [www.itrs.net].