Extended Abstracts of the 2013 International Conference on Solid State Devices and Materials, Fukuoka, 2013, pp590-591

Thermal and Plasma Treatments for Improved (Sub-)1nm EOT Planar and FinFET-based RMG High-k Last Devices and Enabling a Simplified Scalable CMOS Integration Scheme

A. Veloso, G. Boccardi, L.-Å. Ragnarsson, Y. Higuchi[1], H. Arimura[2], J. W. Lee[2], E. Simoen, M. J. Cho, Ph. J. Roussel, V. Paraschiv, X. Shi, T. Schram, S. A. Chew, S. Brus, A. Dangol, E. Vecchio, F. Sebaai, K. Kellens, N. Heylen, K. Devriendt, H. Dekkers, A. Van Ammel, T. Witters, T. Conard, I. Vaesen, O. Richard, H. Bender, R. Athimulam, T. Chiarella, A. Thean, and N. Horiguchi

Imec, [1]assignee at Imec from Panasonic, Kapeldreef 75, 3001 Leuven, Belgium; [2]also at K. U. Leuven, Belgium

Tel.: +32-16-28 17 28, Fax: +32-16-28 17 06, Email: Anabela.Veloso@imec.be

Abstract

We report on aggressively scaled RMG-HKL planar and multi-gate FinFET-based devices, systematically investigating the impact of post high-k deposition thermal (PDA) and plasma (SF_6) treatments on device characteristics, and providing a deeper insight into underlying degradation mechanisms. We demonstrate that: 1) substantially reduced J_G and noise values can be obtained for both type of devices with PDA and F incorporation in the gate stack by SF_6, without EOT penalty; 2) SF_6 also enables improved mobility and reduced N_{it} down to narrower fin devices (W_{Fin}≥5nm), mitigating the impact of fin patterning, fin corners and fin sidewalls crystal orientations, while allowing a simplified dual-EWF metal CMOS scheme suitable for both device architectures and which maximizes the space for gate metallization; 3) PDA also yields lower PMOS $|V_T|$, and substantially improved NBTI lifetime and hot-carrier (HC) immunity thanks to its reduction of bulk defects which is shown to be key in the (sub-)1nm EOT regime.

Introduction

Following successful implementation into device manufacturing [1], high-k/metal gate faces key challenges on effective work function (EWF), gate leakage (J_G), variability and reliability control for ultra-thin EOT/T_{inv}. For (sub-)22nm nodes, introduction of alternative device architectures to planar bulk, such as FinFET-based multi-gate devices [2,3], also requires 3D compatible integration options. In this work, we provide a thorough evaluation of post HfO_2 deposition treatments (PDA and SF_6-plasma) for planar *vs.* FinFET devices with different Si crystal orientations. In agreement with the previously reported tendency of F to segregate to the IL-SiO_2/HfO_2 and Si/IL-SiO_2 interfaces, passivating oxygen vacancies and interface traps by forming stronger Hf–F and Si–F bonds [4], lower N_{it} values are obtained here with SF_6, which also enables a simplified dual-EWF metal RMG-HKL CMOS scheme for scaled, high aspect-ratio gate trenches. Furthermore, improved BTI and HC reliability for (sub-)1nm EOT [5], where bulk defects are demonstrated to play a key role, is achieved here with PDA.

Device fabrication

The process flow used for device fabrication is illustrated in Fig.1 [6,7]. Source/drain silicide is done after the RMG module, allowing introduction of a higher temperature PDA. Fig.2 shows schematics of evaluated RMG-HKL stacks, using O_3-oxidation for the interfacial layer (IL)-SiO_2 prior to HfO_2 growth. A SF_6-plasma optimized to increase the number of ions reaching the bottom of narrow, high aspect-ratio gate trenches is used to incorporate F in the gate stack and to selectively remove the p-EWF/barrier metals (TiN/TaN) from NMOS areas in a dual-EWF metal CMOS scheme.

Results and discussion

Fig.3a shows that while PMOS $|V_T|$ is reduced with PDA, the opposite occurs upon exposure of HfO_2 to SF_6. Similar trends are observed for planar FETs and FinFETs, but with considerably smaller shifts [and similar ITP characteristics (Fig.3b)] for the latter. Mobility wise, Fig.4 shows reduction of high-field mobility, typically caused by increased surface roughness, with PDA mostly in planar devices, with similar low-field, peak mobility values. A slight improvement occurs with SF_6. Fig.5 shows trap density (N_t) values by charge pumping, with results at high frequencies (e.g., 1MHz) corresponding mostly to the response of traps located at the Si/IL-SiO_2 interface (N_{it}). Similarly to ΔV_T, opposite ΔN_{it} trends occur for PDA *vs.* SF_6: increased N_{it} with PDA and for higher PDA temperatures, reduced N_{it} with F incorporation in the gate stack. N_{it} values extracted for FinFETs are initially higher than for planar due to defects introduced at fin patterning, but Figs.5,6 show that they can be substantially reduced by using the same SF_6 process as in planar, while also helping to mitigate the differences seen for fin sidewalls with different crystal orientations {higher dangling bonds in (110) *vs.* (100) fin side surfaces [8]} down to W_{Fin}~5nm. A less steep slope for $N_{it,total} = N_{it}$(fin top surface + fin sidewalls) *vs.* W_{Fin} for SF_6 and PDA devices indicates higher gate integrity at fin corners, likely due to occurrence of some Si reflow (corner rounding) at 800°C PDA, and defects passivation by Hf–F and Si–F bonds with SF_6. However, whereas HfO_2 exposure to SF_6 leads to a modest NBTI lifetime improvement, that is substantial with PDA. This is seen in Fig.7 for planar FETs, which also show considerably smaller estimated bulk trap densities [= total effective trap density (N_{teff}) from ΔV_T of NBTI - N_{it}] with PDA: ~2.1 and 1.4× lower N_{teff} using 800°C PDA and SF_6, respectively. Bulk defects seem thus to play a key and dominant role in NBTI, being mostly pre-existing defects contributing to charge trapping and de-trapping during stress and relaxation as inferred by the larger recoverable (R) *vs.* permanent (P) BTI degradation components in Fig.7. In addition, while reference and SF_6 devices exhibit similar R *vs.* t_{stress} slopes, these are less steep with PDA, indicating less generation of new defects during stress. As for FinFETs, Fig.8 shows similar NBTI trends, with substantially improved lifetime for decreasing W_{Fin}. For W_{Fin}≤20nm, (100) *vs.* (110) fin sidewalls are beneficial. Since N_{it} [lower for the (100) orientation] was shown not to be determinant for planar FETs NBTI at EOT≤1nm, this difference is thought to be mostly due to some impact of the Si crystal orientation on the IL-SiO_2 growth. Fig.9 also shows improved HC immunity for narrow-fin FinFETs built with PDA, with their smaller ΔI_D and ΔV_T at high V_G=V_D indicating less charge trapping into bulk defects, thus corroborating the BTI and N_{it} assessment of reduced bulk defects presence. LF-noise analyses show in Figs.10a,b improved normalized input-referred noise spectral density values (~3-4.5× lower S_{VG}f *vs.* reference) with PDA and SF_6. In agreement with BTI results, further gain is obtained for FinFETs with (100) fin sidewalls, with all devices largely following 1/f noise behavior (Fig.10c). Proportionality of the normalized current noise spectral density (S_I/I_D^2) with $(g_m/I_D)^2$ for reference and fluorinated devices (Fig.10d), at lower $|I_D|$, points towards carrier number fluctuations or oxide trapping in correlation with mobility fluctuations at the origin of LF-noise, whereas that does not seem to be always the case with PDA.

Lastly, a simplified dual-EWF metal RMG-HKL CMOS scheme based on the use of SF_6 is shown in Fig.11. After selectively removing with SF_6 the p-EWF/barrier metals covering HfO_2, XPS results in Fig.12 show clear F incorporation in the stack, with Hf-F bonds mostly present at the surface. An optimized clean/strip is then required to avoid washing away HfF_4 (Fig.13), preserving the high-k dielectric physical thickness and EOT. Indeed, Figs.14a,b show similar EOT-J_G, V_T nFETs are obtained by depositing the n-EWF metal after the HfO_2 growth [9] or following the SF_6 removal of the p-EWF/barrier metals from NMOS areas. No impact on device properties from a longer HfO_2 exposure to SF_6 (similar V_T, $\sigma(V_T)$, J_G, and RO performance in Fig.14) ensures a wide process window and full compatibility with FinFETs, while maximizing the gate trench space left to be filled with n-EWF/fill-metallization. Interesting also to note the potential attraction of using PDA in this scheme to obtain nFETs with ~10× lower J_G at given V_T (Fig.14b).

Conclusions

Reduced J_G and noise values achieved by using PDA and a SF_6-plasma after HfO_2 growth in planar and FinFET devices, without EOT penalty. SF_6 also improves mobility and reduces N_{it} for W_{Fin}≥5nm, allowing a simplified, highly scalable dual-EWF metal RMG-HKL CMOS scheme suitable for both device architectures. Substantially improved NBTI lifetime and hot-carrier immunity for (sub-)1nm EOT are obtained by bulk defects reduction with PDA.

References

[1] K. Mistry *et al.*, IEDM *Tech. Dig.* 2007, 247; [2] H. Kawasaki *et al.*, IEDM *Tech. Dig.* 2008, 237; [3] C. Auth *et al.*, VLSI *Tech. Dig.* 2012, 131; [4] W.-C. Wu *et al.*, IEDM *Tech. Dig.* 2008, 405; [5] M. J. Cho *et al.*, IEEE *Trans. Elect. Dev.* 59(8) 2012, 2042; [6] A. Veloso *et al.*, VLSI *Tech. Dig.* 2012, 33; [7] G. Boccardi *et al.*, SSDM *Tech. Dig.* 2012, 723; [8] S. Maeda *et al.*, IRPS *Tech. Dig.* 2004, 8; [9] A. Veloso *et al.*, VLSI *Tech. Dig.* 2013 (to be published).

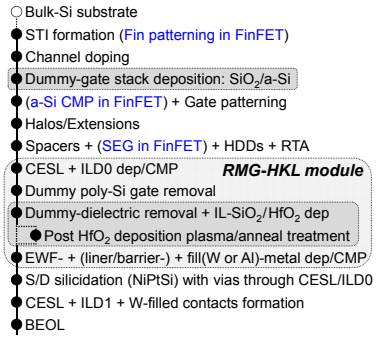

Fig.1 – Schematics of the process flow used for fabrication of replacement metal gate/high-k last (RMG-HKL) planar and multi-gate FinFET-based devices.

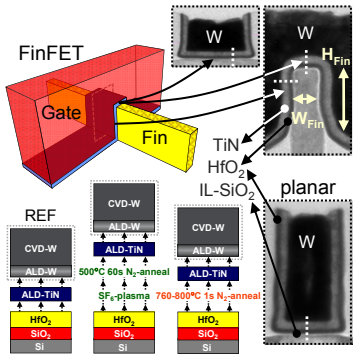

Fig.2 – Schematics of gate stacks (as-deposited) evaluated in this work for both planar (bottom right, TEM) and FinFET-based (on top) RMG-HKL devices.

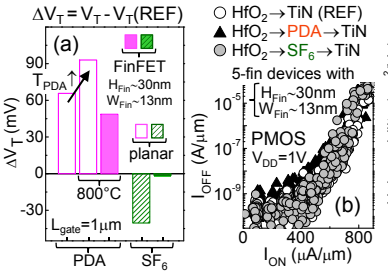

Fig.3 – PDA after HfO_2 growth results in lower PMOS $|V_T|$, while the opposite occurs with high-k exposure to SF_6. Similar trends are seen for planar (a) and FinFET devices [(a,b); ITP in (b)], but with smaller V_t shifts for the latter.

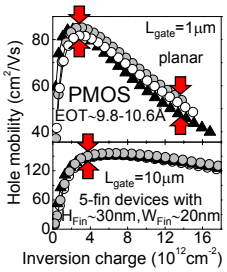

Fig.4 – Slight mobility improvement with SF_6 for FinFETs and planar FETs (no EOT penalty). PDA impacts high-field values mostly in planar.

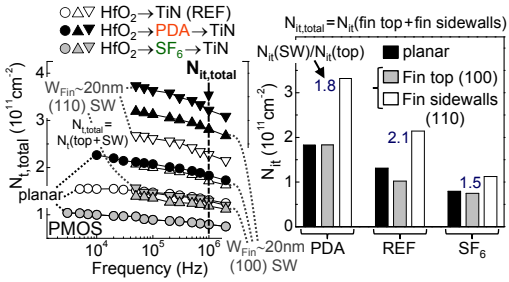

Fig.5 – Interface trap density (N_{it}) estimated from charge pumping is higher for FinFETs with (110) *vs.* (100) fin vertical sidewalls (SW) and in comparison to planar bulk devices. In all cases, F incorporation in the stack by a SF_6-plasma results in reduced N_{it}, whereas PDA increases it.

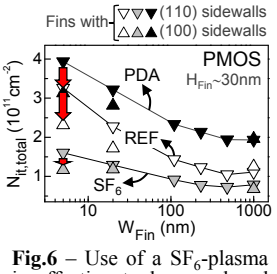

Fig.6 – Use of a SF_6-plasma is effective to keep reduced $N_{it,total}(W_{Fin})$ for devices with (110) or (100) fin sidewalls. A less steep $N_{it,total}$-W_{Fin} slope with SF_6 and also PDA indicates less defects at fin corners in both cases.

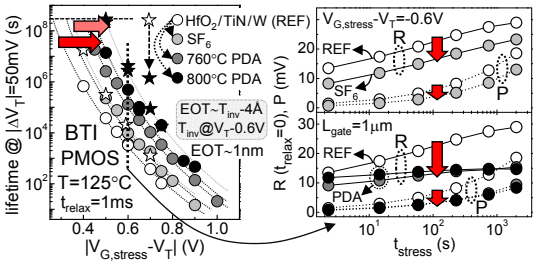

Fig.7 – NBTI lifetime is significantly improved with PDA for planar FETs, corresponding to lower total effective trapped charge density values and to recoverable (R) and permanent (P) degradation components with a less steep slope *vs.* t_{stress}.

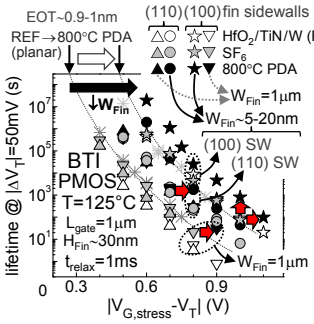

Fig.8 – Similarly to planar, also substantial NBTI improvement is obtained with PDA in FinFETs, a more modest gain with SF_6. BTI degradation is reduced for (100) *vs.* (110) fin sidewalls and for narrower fin devices.

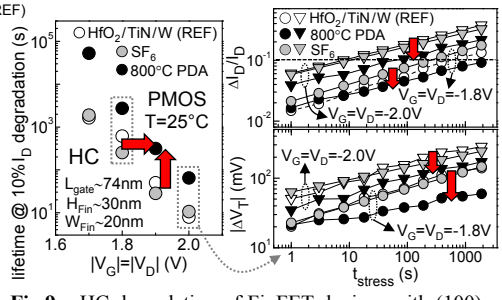

Fig.9 – HC degradation of FinFET devices with (100) fin sidewalls is reduced by including PDA in the flow. The smaller I_D degradation ($\Delta I_D/I_D$) or $|\Delta V_T|$ at high V_G=V_D corresponds thus mostly to less charge trapping into bulk defects and to no significant impact from N_{it}.

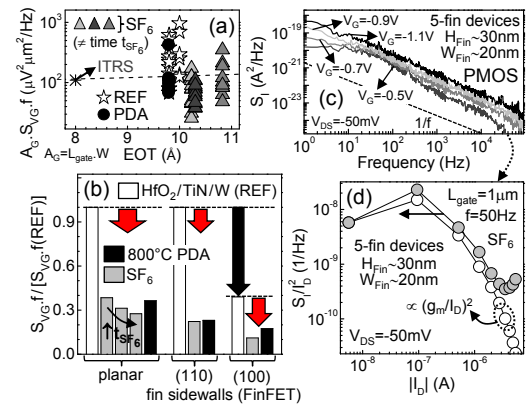

Fig.10 – Normalized S_{VG} values for RMG-HKL devices are in line with the ITRS 1/f noise roadmap [planar data in (a)]. Noise is considerably reduced with PDA and SF_6 for both planar and narrow-fin devices, and (100) fin sidewalls are beneficial (b). c,d) show examples of LF-noise spectra and the S_I/I_D^2 correlation with $(g_m/I_D)^2$, respectively.

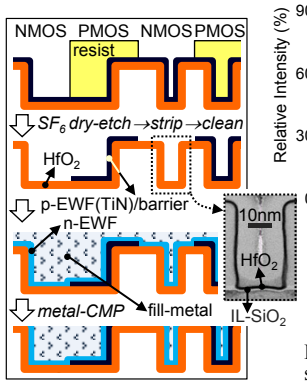

Fig.11 – Simplified dual-EWF metal RMG-HKL CMOS scheme, planar and FinFET compatible, with a SF_6-plasma removing the p-EWF/barrier metals from the NMOS areas.

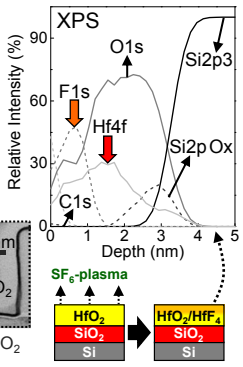

Fig.12 – XPS analyses show F incorporation into HfO_2 upon exposure to a SF_6-plasma. Hf-F bonds are mainly present at the surface, with only limited amount of F detected in the bulk of the HfO_2.

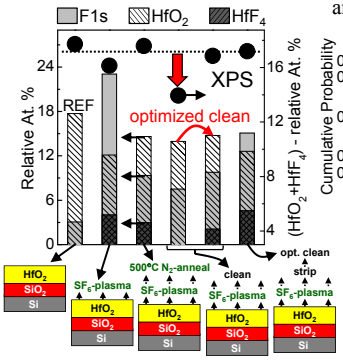

Fig.13 – An optimized clean after strip is required to be able to use a SF_6-plasma to remove metals from inside gate trenches without impacting the physical thickness of the high-k dielectric underneath, and with F incorporated in it ($\Rightarrow N_{it}$ reduction).

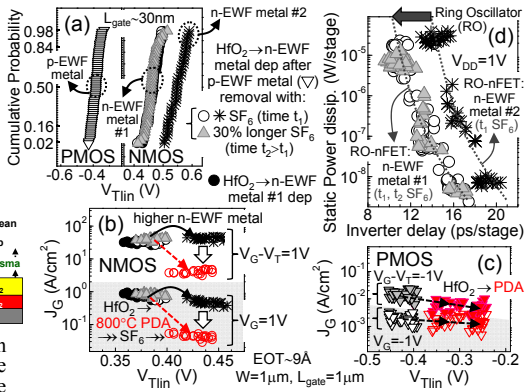

Fig.14 – a,b) Similar V_T, J_G, and EOT are obtained for nFETs built with the n-EWF metal deposited after HfO_2 growth or after using the SF_6-based CMOS process. SF_6 over-etch has no impact on devices (a,b) and circuits [RO in (d)]. b) shows that PDA significantly reduces J_G at a given V_T for nFETs, with smaller ΔJ_G for pFETs (c).