BTI Reliability Improvement Strategies in Low Thermal Budget Gate Stacks for 3D Sequential Integration

J. Franco*, Z. Wu[1], G. Rzepa[2], A. Vandooren, H. Arimura, L.-Å Ragnarsson, G. Hellings, S. Brus, D. Cott, V. De Heyn, G. Groeseneken[1], N. Horiguchi, J. Ryckaert, N. Collaert, D. Linten, T. Grasser[2], B. Kaczer
imec, Leuven – Belgium, [1]also at KU Leuven – Belgium , [2]TU Wien – Austria *Jacopo.Franco@imec.be

***Abstract*—Low thermal budget gate stacks will be required for novel integration schemes, such as 3D sequential stacking of CMOS tiers. We study the impact of a reduced thermal budget on BTI reliability, and we demonstrate two strategies to tolerate the inherently large high-k defect densities: i) replacing inversion mode devices with highly doped junction-less transistors, or ii) engineering dipoles at the interface between SiO_2 and HfO_2 to suppress the carrier-defect interaction. The latter approach is demonstrated for nMOS PBTI and, for the first time here, also for pMOS NBTI, as even this aging mechanism is controlled by high-k defects in ultra-thin EOT low thermal budget gate stacks.**

Introduction

3D sequential integration is a recently envisioned approach to increase CMOS functionality per die area by stacking transistors on top of each other (**Fig. 1**), or to co-integrate heterogeneous technologies on multiple tiers of the same wafer within a single fabrication flow [1,2]. Thermal budget management represents the most crucial challenge of this integration approach: thermal steps for the fabrication of the top tier should be limited to preserve the functionality of the bottom tier devices and interconnects (BEOL). Gate stack reliability is extremely sensitive to process temperature, particularly for high-k/metal gate (HKMG) technologies. In a gate-first integration flow, the high-k layer is exposed to the source/drain (S/D) activation anneal (~1000°C), which reduces the dielectric defect densities and minimizes Bias Temperature Instability (BTI) [3,4]. In more contemporary Replacement Gate (RMG) flows, a dedicated high-temperature 'reliability' anneal (~900°C) is typically performed after the deposition of the final gate stack to ensure sufficient stability [5,6]. Such high temperature steps are not suitable for top tier fabrication as they would degrade the BEOL of the bottom tier.

In this paper, we study first the impact of a limited thermal budget on the BTI reliability of the HKMG stack. By using our recently introduced physics-based BTI modeling framework ComPhy ("Compact Physical", [7]), we compare the oxide defect properties in an as-deposited low thermal budget SiO_2/HfO_2/TiN gate stack with the same gate stack exposed to a typical 'reliability' anneal, and with a commercial high-k-first 28nm HKMG technology, revealing larger oxide defect density at energy levels around the Si channel band edges. We then use the pathfinder predictive capabilities of ComPhy to define two possible strategies to guarantee sufficient BTI reliability despite the inherently large defect density in low thermal budget high-k dielectrics. The first strategy consists of replacing the standard inversion mode devices with junctionless devices, as the latter device type has been shown to offer improved reliability due to lower operating oxide fields [8]. We discuss how the reliability can be improved further by increasing the channel doping density in these devices, counter-intuitively to inversion mode transistors. The second strategy consists of engineering functional dipoles at the interface between SiO_2 and HfO_2 (by depositing a thin dipole-former interlayer with ALD) to 'shift up' or 'down' the energy levels of the high-k defects with respect to the Si conduction or valence band, for nMOS and pMOS reliability, respectively. We have demonstrated this latter strategy earlier to improve the PBTI reliability of Ge and InGaAs nMOSFETs [9,10] and of Si *n*-channel capacitors [11]. Here we demonstrate it on Si planar nMOSFETs, highlighting a positive correlation between channel electron mobility and PBTI reliability. Furthermore, we show for the first time that dipole engineering can improve also pMOS reliability, as for low thermal budget thin-EOT gate stacks with <1nm SiO_2 interfacial layer (IL) NBTI is also limited by high-k traps. *All these findings are confirmed on Si hardware.*

The results presented here open up the reliable use of low thermal budget high-k gate stacks for 3D sequential integration, but also for other novel integration concepts as, e.g., the embedding of thin film transistors in the BEOL, or to conveniently re-arrange the integration flow of standard Si CMOS which is currently dictated by the highest temperature steps, e.g., the S/D contacts are fabricated only after the gate stack 'reliability' anneal instead of directly after S/D epitaxy.

Experimental

To study the impact of different thermal budget on the top and bottom tiers of a 3D sequential integration separately, we fabricated CMOS wafers without applying any thermal budget after gate stack deposition ("Top": as-deposited, **Fig. 2a**) or by performing a 2h long Post-Metal Anneal (PMA) at 525°C to mimic the additional thermal budget that bottom devices see during top tier fabrication ("Bottom", **Fig. 2b**). Fully stacked wafers were also fabricated (Fig. 1), with *n*- or *p*-channel SOI junction-less devices on top of a standard planar Si CMOS (**Fig. 2c**). For all devices the gate stack comprised a chemical SiO_2 IL (~0.6nm as-dep., 1nm w/ PMA), ~1.8nm HfO_2, and 5nm TiN.

Impact of low thermal budget on HKMG stack reliability

Top tier devices which did not receive any high temperature anneal after HKMG deposition show poor PBTI and NBTI reliability (**Fig. 3a,b**): extremely large threshold voltage shifts (ΔV_{th}: ~20× and ~10× larger than the acceptable target for nMOS and pMOS respectively) and detrimentally weak BTI voltage accelerations are observed, in contrast to the same gate stack exposed to a ~900°C-1s 'reliability' anneal, and to a commercial 28 nm HKMG technology. Interestingly, the same gate stack yields almost sufficient BTI reliability when used in bottom tier devices (**Fig. 3c,d**), suggesting that the additional thermal budget exposure might make the implementation of a dedicated reliability anneal unnecessary for these devices.

To understand the origin of the poor top tier reliability, we used ComPhy [7] to model the ΔV_{th} measured in nMOS and pMOS devices subjected to a complex stress/recovery waveform and to more conventional measure-stress-measure patterns [12], with various voltages and at different temperatures (**Fig. 4**). An excellent match of the modeled BTI kinetics to the experimental data is achieved by properly calibrating shallow and deep defect band properties for both the SiO_2 and HfO_2 dielectric; these defect properties are compared to the ones extracted in [7] from BTI measurements of a commercial 28 nm technology (**Table I**). The reduced thermal budget results in a ~2× larger defect density in HfO_2, and especially in a significantly reduced mean trap energy level and mean thermal barrier for capture for both the SiO_2 and HfO_2 traps, possibly due to a less stiff, more disordered amorphous oxide. Therefore, a larger amount of oxide defects can trap channel carriers at operating oxide electric field (E_{ox}=3MV/cm), as compared to a standard high thermal budget technology (**Fig. 5**).

We demonstrate two alternative strategies for sufficient low thermal budget reliability: i) the use of highly doped junctionless transistors, which operate at lower E_{ox} compared to inversion mode devices, or ii) the insertion of dipole-former interlayers between SiO_2 and HfO_2 to shift the defect levels in the latter w.r.t. the Si channel bandgap.

Solution I: Junction-less devices optimized for minimum E_{ox}

Accumulation-mode junction-less devices offer a superior reliability as they operate close to the MOS flatband voltage [8]: on-state is achieved with a low E_{ox}, while an opposite E_{ox} is necessary to switch-off the current flow by depleting the thin channel of the device (**Fig. 6**) [note: the device V_{th} is therefore defined as the V_G required to fully deplete the channel [13], **Fig. 7**]. In a conventional inversion-mode MOSFET, the on-state E_{ox} is determined by the sum of the inversion charge (proportional to the gate overdrive voltage, V_{ov}=V_G-V_{th}) and the depletion charge: as such, a higher channel doping level results in a larger E_{ox} for a given V_{ov}, detrimental to reliability [14]. We argue that in a junction-less device instead, a higher doping can be beneficial for reliability: electrostatic simulations [15] show that for a given E_{ox} a larger V_{ov} (i.e., higher drive current) can be applied on a junction-less device for increasing doping densities (**Fig. 8**). Experimental data confirm that sufficient PBTI and NBTI reliability is achieved in junction-less devices with doping $>2\times10^{18}/cm^3$ (**Fig. 9**), despite the large defect density in the low thermal budget oxides. Note that the BTI trends are excellently reproduced in ComPhy (Fig. 9, lines), by using the same defect properties calibrated on low thermal budget inversion mode devices (Table I), and the specific $E_{ox}(V_{ov})$ relation implied by the junction-less electrostatics (cf. Fig. 8). We conclude that junction-less devices represent a convenient option for top tiers, as their low operating E_{ox} relieves the dielectric quality requirements. Moreover, these devices are inherently thermal-budget-friendly as they do not require S/D activation (note: channel doping is activated before bonding the top Si slab to the bottom tier passivation oxide [2]).

Solution II: dual interface dipole engineering for sufficient BTI reliability in inversion-mode CMOS

As discussed above, the main shortcoming of a low thermal budget HKMG stack is the low energy level of the oxide defects, which enhances charge trapping. Properly engineered interface dipoles can 'shift up' the shallow high-k defect band w.r.t. the Si conduction band (for improved PBTI), or 'shift down' the deep defect band w.r.t. the Si valence band (for improved NBTI), and therefore reduce the density of accessible defects to a level comparable to the high temperature HKMG stack of a commercial 28 nm technology (**Fig. 10**, cf. Fig. 5).

A. ComPhy pathfinding study: nMOS and pMOS

ComPhy simulations show that a dipole at the SiO_2/HfO_2 interface is very effective to improve PBTI: a ΔV_{th}<50mV is projected for 10-year operation at V_{ov}=0.7V at 125°C if the high-k shallow defect energy is increased by 0.4eV (**Fig. 11**). This improvement is virtually independent of the SiO_2 thickness, as PBTI is mostly controlled by high-k electron traps (SiO_2 electron traps are negligible, cf. Table I).

Interestingly, ComPhy simulations show that a similar approach can be effective also for improving NBTI, despite that in pMOS this mechanism is commonly ascribed to interfacial traps: in a low thermal budget HKMG stack, the deep defects in HfO_2 have a sufficiently high density to surpass the SiO_2 hole traps (**Fig. 12a**, dashed vs. dotted lines), and therefore a proper interface dipole shift can reduce charge trapping to a tolerable level. However, if a thick IL is used (**Fig. 12b**, 1nm vs. 0.6nm), the total density of hole traps in the larger SiO_2 volume becomes dominant, defeating the effectiveness of an interface dipole at the SiO_2/HfO_2 interface for improving the overall NBTI.

This pathfinding simulations suggest that by using a 0.6nm thin SiO_2 IL and by inducing a +0.4eV or a -0.4eV dipole shift at the SiO_2/HfO_2 interface, *PBTI- and NBTI-induced ΔV_{th} can be brought back to a level comparable to the (high thermal budget) commercial 28 nm ref. technology* (**Fig. 13**), guaranteeing >10-year reliable operation.

B. Experiment: LaSiOx dipole for nMOS PBTI

We have recently shown [11] that a ~0.3nm thin $LaSiO_x$ layer ALD-deposited between SiO_2 and HfO_2 can induce a ~0.4eV V_{fb} shift in *n*-Si MOS capacitors. We reproduce this here on planar MOSFETs, demonstrating low nMOS V_{th} with a TiN high work function metal. Moreover, the insertion of the dipole forming interlayer is observed to be beneficial also for channel electron mobility (**Fig. 14**) due to reduced carrier-defect interaction. The PBTI ΔV_{th} measured on $LaSiO_x$-inserted nMOSFETs are significantly reduced (~8×) w.r.t. the reference low thermal budget "top tier" gate stack, in excellent quantitative agreement with the ComPhy prediction (**Fig. 15**). Note that when combined with the higher thermal budget of the bottom tier, exceptional PBTI reliability can be achieved (Fig. 15, squares); hence, this gate stack represents an interesting option also for bottom devices.

C. Experiment: Al_2O_3 dipole for pMOS NBTI

To experimentally verify the effectiveness of interface dipoles for NBTI reliability, a negative dipole forming layer should be identified. It was previously reported [16], that Al_2O_3 is a strong negative dipole former on SiO_2. In order to achieve a sufficient interface dipole density for a ~0.4eV shift, a ~1nm thick Al_2O_3 layer is necessary (**Fig. 16**, quantitatively in-line with [16]). Additional trapping in such a thick layer might defeat the reliability improvement strategy. However, our previous BTI study of bulk Al_2O_3 oxide defects suggests a negligible trap density at energies close to the Si valence band (**Fig. 17**), making Al_2O_3 suitable for pMOS use (while its wide defect distribution around the Si conduction band makes it unsuitable for nMOSFETs [10]).

By depositing a 1nm Al_2O_3 layer on SiO_2 before HfO_2, a ~10× reduced NBTI ΔV_{th} is demonstrated (**Fig. 18**). In particular, a stronger NBTI voltage acceleration is observed when inserting Al_2O_3 (**Fig. 19**), quantitatively in line with the ComPhy prediction (solid lines) and comparable to best-in-class SiGe pMOSFETs [17], confirming an effective decoupling of the defect energy level w.r.t. the channel Fermi level. A ~0.2nm thin Al_2O_3, inducing ~0.24eV of dipole shift, is already sufficient to bring NBTI ΔV_{th} within specs, with a minimal EOT penalty (cf. Fig. 16a). We note the improvement achieved by inserting thicker Al_2O_3 layers is much larger than the E_{ox} reduction related to the thicker EOT (**Fig. 20**).

For proper pMOS V_{th} tuning, the effective work function shift related to the Al_2O_3 dipole might need to be compensated: this can be achieved by replacing the TiN metal with a more mid-gap work function metal such as TaN, maintaining the improved reliability (**Fig. 21**). Finally, we compare the reliability of the Al_2O_3-inserted gate stacks to the reference one at top- and bottom-tier thermal budgets (**Fig. 22**): in the as-deposited gate stack with a ~0.6nm native SiO_2 IL, a thicker Al_2O_3 layer results in larger eWF shift and maximum operating V_{ov}; in contrast, when applying a 525ºC-2h PMA to mimic the thermal budget seen by the bottom devices, the NBTI reliability is not impacted by the Al_2O_3 dipole: this is due to i) the higher thermal budget, yielding a sufficient baseline reliability, and ii) a thicker SiO_2 IL (1nm vs 0.6nm, cf. Fig. 22b) reducing the contribution of high-k deep traps to the overall NBTI, in line with the ComPhy prediction (cf. Fig. 12b).

Conclusions

We have demonstrated two strategies to achieve sufficient gate stack BTI reliability at low thermal budget, relevant for novel integration schemes such as 3D sequential. The first approach consists of using highly doped junction-less devices in the top tier to reduce the operating oxide electric field, and therefore the density of charging oxide traps. The second approach utilizes ALD dipole-forming interlayers ($LaSiO_x$ for nMOS and Al_2O_3 for pMOS) to engineer the energy alignment of the high-k defect levels (i.e., the main contributors to both PBTI and NBTI ΔV_{th} in low thermal budget gate stacks) w.r.t. the Si channel band edges. Both strategies were identified by using the pathfinding predictive capabilities of our BTI simulation framework ComPhy, and were demonstrated experimentally on Si hardware.

Acknowledgements: research funded by imec's core partner program.

References

[1] P. Batude *et al.*, in *Proc.* VLSI Tech. 2015;
[2] A. Vandooren *et al.*, in *Proc.* VLSI Tech. 2018;
[3] E. Bury *et al.*, in *Proc.* IRPS 2013;
[4] H. Arimura *et al.*, in *Proc.* IRPS 2014;
[5] G. Rzepa *et al.*, in *Proc.* IRPS 2017;
[6] B. Linder *et al.*, in *Proc.* IRPS 2016;
[7] G. Rzepa *et al.*, in *Micr. Rel.* 85, pp. 49-65 , 2018;
[8] M. Toledano-Luque *et al.*, in EDL 35(12), 2014;
[9] H. Arimura *et al.*, in *Proc.* IEDM 2016;
[10] J. Franco *et al.*, in *Proc.* IEDM 2017;
[11] J. Franco *et al.*, in *Proc.* IRPS 2017;
[12] B. Kaczer *et al.*, in *Proc.* IRPS 2008;
[13] J.-P. Colinge *et al.*, Nature Nanotechnology 5, pp. 225–229, 2010;
[14] J. Franco *et al.*, in *Proc.* IRPS 2016;
[15] J. R. Hauser and K. Ahmed, in *Proc.* AIP Conference 449, 235 ,1998 (CVC version 5.0);
[16] Y. Kamimuta *et al.*, in *Proc.* IEDM 2007;
[17] J. Franco *et al.*, in *Proc.* IEDM 2013.

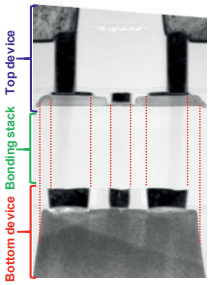


Fig. 1: TEM of a 3D structure showing stacked top and bottom tier devices with nanometric alignment [2].

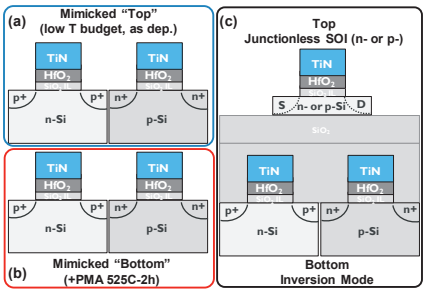


Fig. 2: Used test structures: planar CMOS with **(a)** as-deposited gate stack (compatible with top tier, "top"), or **(b)** with a 525ºC-2h PMA mimicking the thermal budget seen by the bottom tier during top fabrication ("bottom"); **(c)** stacked structures with SOI junction-less devices on bulk CMOS [2].

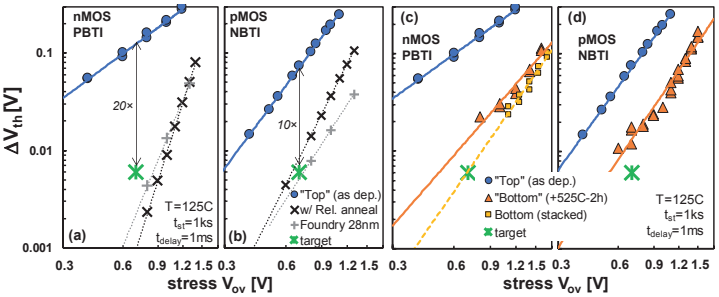


Fig. 3: Low thermal budget HKMG stack ("top") shows poor **(a)** PBTI and, **(b)** NBTI compared to the same stack with a high temperature 'reliability' anneal, or compared to a 28nm commercial technology [5,7]. In contrast, thanks to the additional thermal budget of the top tier fabrication, the "bottom" stack almost meets the **(c)** PBTI and **(d)** NBTI targets (ΔV_{th}<6mV at V_{ov}=0.7V, T=125C, t_{st}=1ks, rescaled from 50mV at 10 years) without a dedicated reliability anneal.

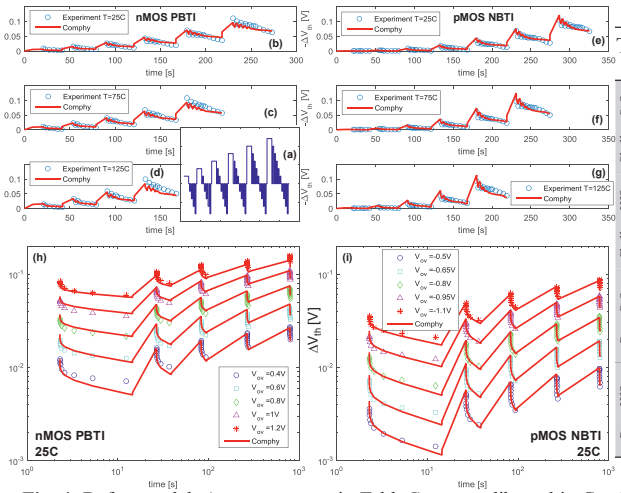


Table I

		Foundry 28 nm	Low Thermal Budget
Shallow SiO_2	N_t [/cm³]	1.5×10^{19}	1.17×10^{19}
	$\langle E_t \rangle \pm \sigma_{Et}$ [eV]	1.13 ± 0.15	$\mathbf{0.55} \pm 0.11$
	$\langle S \rangle \pm \sigma_S$ [eV]	3.82 ± 1.36	$\mathbf{2.66} \pm 1.07$
	R [1]	0.407	0.852
Shallow HfO_2	N_t [/cm³]	5.52×10^{20}	2.56×10^{20}
	$\langle E_t \rangle \pm \sigma_{Et}$ [eV]	1.2 ± 0.156	$\mathbf{0.5} \pm 0.105$
	$\langle S \rangle \pm \sigma_S$ [eV]	3.19 ± 0.77	$\mathbf{2.34} \pm 0.7$
	R [1]	0.587	0.556
Deep SiO_2	N_t [/cm³]	1.42×10^{20}	$\mathbf{2.05 \times 10^{20}}$
	$\langle E_t \rangle \pm \sigma_{Et}$ [eV]	-1.26 ± 0.25	$\mathbf{-0.92} \pm 0.23$
	$\langle S \rangle \pm \sigma_S$ [eV]	5.63 ± 2.67	5.49 ± 1.88
	R [1]	1.8	1.15
Deep HfO_2	N_t [/cm³]	2.95×10^{20}	$\mathbf{5.17 \times 10^{20}}$
	$\langle E_t \rangle \pm \sigma_{Et}$ [eV]	-0.17 ± 0.14	$\mathbf{-0.6} \pm 0.1$
	$\langle S \rangle \pm \sigma_S$ [eV]	6.5 ± 1.99	6.08 ± 1.53
	R [1]	0.514	1.58

Fig. 4: Defect models (see parameters in Table I) were calibrated in ComPhy [7] to excellently reproduce the nMOS PBTI (b,c,d,h) and (e,f,g,i) pMOS NBTI kinetics in a variety of stress/recovery patterns: **(a)** complex waveform comprising increasing stress voltages and decreasing discharge voltages, performed at **(b,e)** 25ºC, **(c,f)** 75ºC, **(d,g)** 125ºC; and **(h-i)** standard extended measure-stress-measure patterns [12] with various stress overdrive voltages (25ºC).

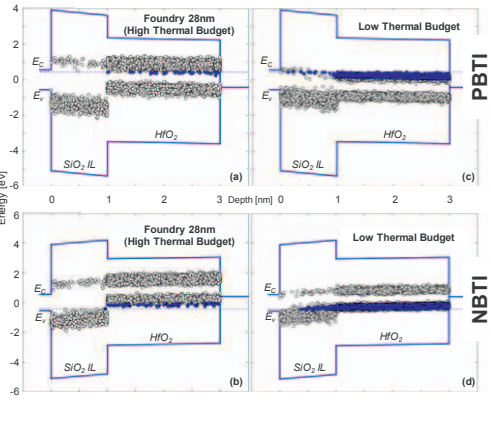


Fig. 5: Band diagrams showing the occupancy of shallow and deep defect bands (cf. Table I) in SiO_2 and HfO_2 at BTI stress condition ($E_{ox,IL}$=±3MV/cm), as calibrated in Comphy on **(a-b)** commercial 28nm nMOS and pMOS, **(c-d)** low thermal budget HKMG stack ("top", as deposited). All defect levels crossing the channel Fermi level due to the applied E_{ox} are depicted here as filled (as their charge state might change); note the larger density of charging defects in the low thermal budget stack, as compared to the commercial 28nm technology.

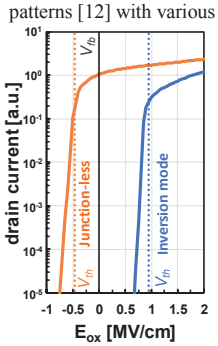


Fig. 6: Simulated I_D vs. E_{ox} for an inversion mode MOSFET and for a junction-less device (replotted from [8]). The latter operates close to V_{fb} with a low E_{ox}, while it requires a negative E_{ox} to switch off the current.

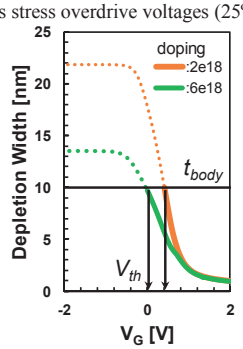


Fig. 7: Calculation [15] of the channel depletion width in a MOS stack for two different doping levels. In a junction-less device, the V_{th} is defined [13] as the V_G required to fully deplete the thin body (~10nm).

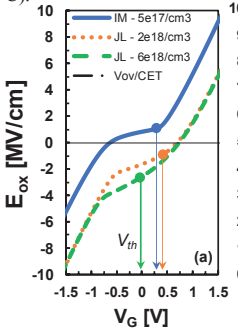
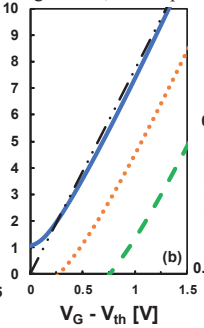


Fig. 8: Calculated E_{ox} for inversion- (N_d=5e17 /cm³) and accumulation-mode MOSFETs (N_d= 2-6e18/cm³) vs. **(a)** V_G, and vs. **(b)** $V_G - V_{th}$ (note: V_{th} defined at full depletion, cf. Fig. 7). At a given V_{ov}, junction-less show a reduced E_{ox} for increasing doping. Note the invalidity of the approx. $E_{ox} \sim V_{ov}$/CET for junction-less devices.



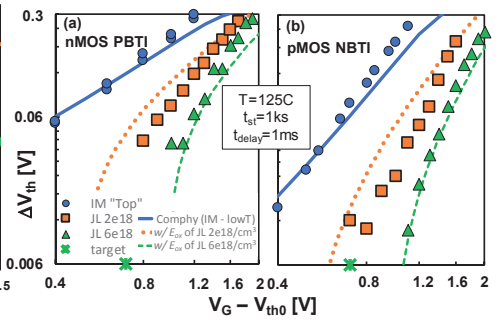


Fig. 9: **(a)** nMOS PBTI, and **(b)** pMOS NBTI ΔV_{th} measured in inversion mode and junction-less devices vs. V_{ov}. The same low thermal budget gate stack is used in all devices. The reduced ΔV_{th} in junction-less for increasing doping levels is quantitatively predicted by Comphy, accounting for the E_{ox} reduction in junctionless devices (cf. Fig.8b).

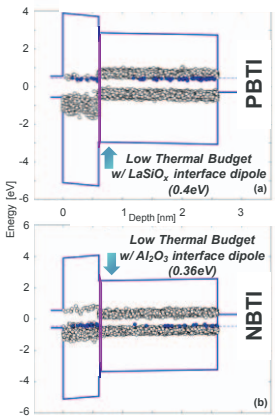


Fig. 10: Occupancy of shallow and deep defect bands in low thermal budget SiO_2 and HfO_2 at BTI stress bias ($E_{ox,IL}$=±3MV/cm), considering a ~0.4eV dipole **(a)** up-shift for improved PBTI, or **(b)** down-shift for improved NBTI.

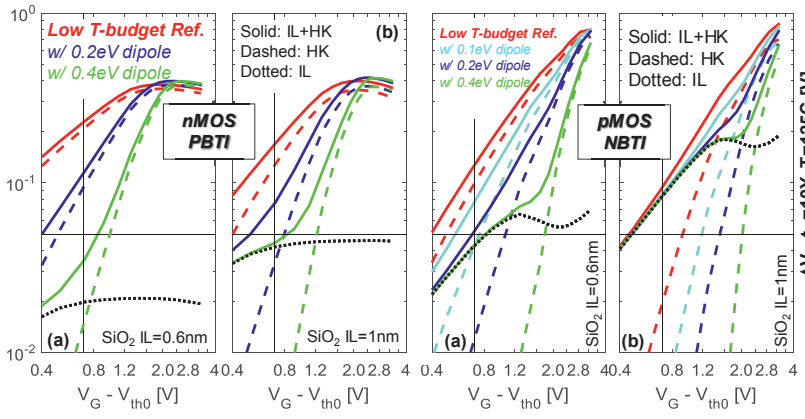


Fig. 11: Comphy simulations of PBTI ΔV_{th} (10Y,125ºC) for increasing V_{ov} (defect model for low thermal budget, cf. Table I). The impact of a dipole at the SiO_2/HfO_2 interface is emulated by shifting up the energy level of the high-k traps of 0.2 and 0.4eV. The dashed and dotted lines show the respective contributions of HfO_2 and SiO_2 traps. Two scenarios are considered **(a)** 0.6nm thin SiO_2 IL, **(b)** conventional 1nm SiO_2 IL. In both cases, a 0.4eV dipole is predicted to maintain PBTI ΔV_{th}(10Y,125ºC) within a 50mV target at V_{ov}=0.7V.

Fig. 12: Same as Fig. 11, for NBTI. The dipole impact is emulated by shifting the energy level of the HfO_2 deep traps down of 0.1, 0.2 and 0.4eV, for **(a)** a 0.6nm SiO_2 IL, and **(b)** a 1nm SiO_2 IL. A dipole shift is effective in improving NBTI only in the former case, while in the latter the reliability is limited by hole traps in the 1nm SiO_2 IL. (Note: the kink in the contribution of SiO_2 traps is due to gate interaction on trap occupancy at high V_G).

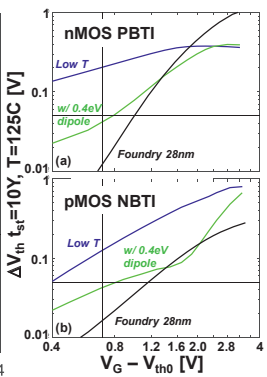


Fig. 13: Comparison of **(a)** PBTI and **(b)** NBTI ΔV_{th} (10Y, 125ºC) simulated for the low thermal budget ref. and the dipole-engineered stacks, compared to the projection for a commercial 28 nm technology.

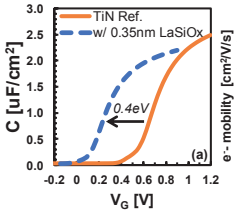


Fig. 14: **(a)** C-V curves of *n*-substrate MOS capacitors w/o and w/ $LaSiO_x$ interlayer. The insertion of $LaSiO_x$ between SiO_2 and HfO_2 improves the interface quality, resulting also in **(b)** enhanced electron mobility (both w/o and w/ 525ºC-2h PMA).

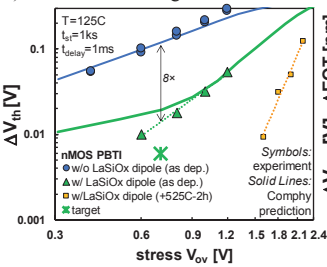


Fig. 15: Measured PBTI ΔV_{th} (t_{st}=1ks, 125ºC) for increasing V_{ov}, on the 'as-dep.' low thermal budget gate stack, w/o and w/ the $LaSiO_x$ layer. A ~8× ΔV_{th} reduction is observed, quantitatively in line with the prediction based on dipole shift (solid lines). Further improvement is observed with a PMA.

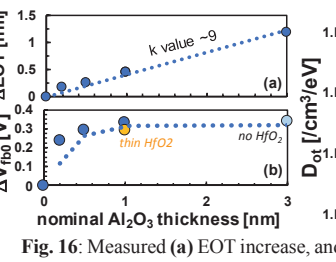


Fig. 16: Measured **(a)** EOT increase, and **(b)** V_{fb} tuning for increasing thicknesses of Al_2O_3 between SiO_2 and HfO_2. A ~0.24eV dipole shift is obtained with a minimum Al_2O_3 thickness with an EOT penalty of only ~1.8Å, while Al_2O_3 ≥1nm yields a max. shift of ~0.36eV.

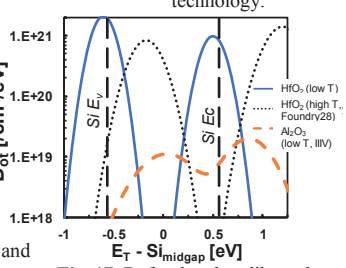


Fig. 17: Defect bands calibrated on HfO_2 (low and high thermal budget) compared to Al_2O_3 defect bands estimated in [10]; the latter shows negligible defect density around the Si valence band.

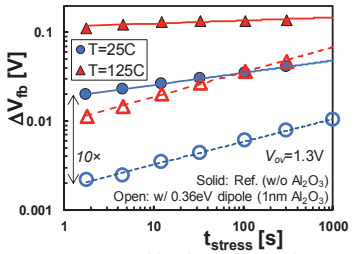
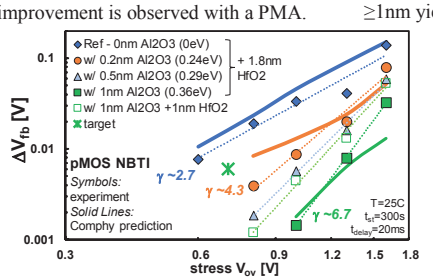


Fig. 18: NBTI kinetics measured on *p*-substrate capacitors at 25ºC and 125ºC (stress V_{ov}=1.3V), on the low thermal budget gate stack w/o and w/ Al_2O_3 interface dipole. Up to ~10× improvement is observed in the latter case.



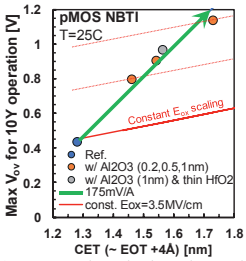


Fig. 19: NBTI shifts for increasing V_{ov} (t_{st}=300s, T=25ºC) measured on the low thermal budget gate stack w/o and w/ Al_2O_3 interlayers of increasing thicknesses. The insertion of a ~0.2nm Al_2O_3 layer brings the reliability within specs. Further improvement is obtained for larger dipole shifts, quantitatively in line with the Comphy prediction (solid lines). Note the extremely favorable voltage acceleration factor γ~6.7 obtained with a ~0.36eV dipole shift, in line with best-in-class SiGe pMOS reliability [17].

Fig. 20 (left): Estimated pMOS max V_{ov} for 10Y operation vs. the CET (~EOT+0.4Å) of each stack: the insertion of Al_2O_3 improves the reliability considerably more than expected due to the induced EOT increase, related to the beneficial dipole shift.

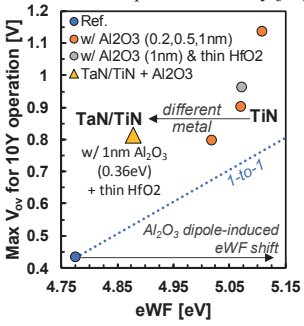


Fig. 21 (right): Estimated max V_{ov} vs. eWF. The insertion of Al_2O_3 increases eWF due to the dipole shift: the eWF can be tuned back by replacing TiN with a lower work function metal (e.g. TaN), maintaining the reliability improvement related to the high-k defect energy shift.

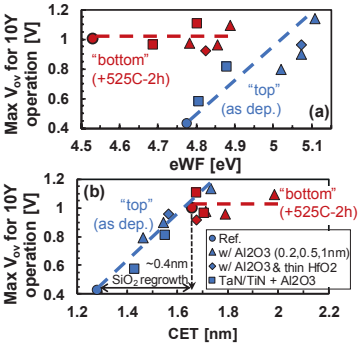


Fig. 22 (right): Max operating V_{ov} for the ref. gate stack and the various Al_2O_3-inserted stacks, both as-dep. ("top") or w/ 525ºC-2h PMA ("bottom"), plotted vs. **(a)** the eWF and **(b)** the gate stack CET (note: the PMA induces a ~0.4nm SiO_2 regrowth). At low thermal budget, the insertion of Al_2O_3 increases eWF, CET and max V_{ov}; on the "bottom" gate stacks instead, Al_2O_3 does not yield additional reliability improvement despite similar eWF and CET modulations: this is related to i) the thicker SiO_2 IL reducing the impact of the dipole shift on NBTI (cf. Fig. 12b), and to ii) the inherently improved reliability at high thermal budget (cf. Fig. 3d).