

# Novel Low Thermal Budget CMOS RMG: Performance and Reliability Benchmark Against Conventional High Thermal Budget Gate Stack Solutions

J. Franco, *Member, IEEE*, H. Arimura, J.-F. de Marneffe, S. Brus, R. Ritzenthaler, K. Croes, B. Kaczer, and Naoto Horiguchi

**Abstract**— Low thermal budget gate stack fabrication is a key enabler for upcoming CMOS technology innovations, such as Sequential-3D integration and CFETs. In this paper, we discuss the impact on gate stack properties of the novel low temperature atomic hydrogen and oxygen treatments that we have recently demonstrated on MOS capacitors. Next we evaluate on a planar CMOS transistor platform various low thermal budget RMG gate stacks based on (combinations of) these treatments, by benchmarking them against high temperature RMG and Gate-First counterparts. This first demonstration on both  $p$ - and  $n$ -channel transistors with properly tuned gate work functions allows to benchmark electron and hole mobilities, and to assess stochastic variability ( $V_{th}$  mismatch), NBTI and PBTI reliability, and the performance ( $I_{ON}$ - $I_{OFF}$ ) and Channel Hot Carrier reliability (Safe Operating Area) of short channel nMOS. We find that the atomic hydrogen treatment, originally developed to improve pMOS NBTI reliability, substantially boosts also the nMOS device performance and reliability due to a remarkably improved interfacial layer quality, matching high-temperature counterparts and thus providing a complete solution for low thermal budget CMOS RMG at a competitive EOT's.

**Index Terms**—Bias Temperature Instabilities, Stacked CMOS, Sequential 3D Integration, CFETs, Low Thermal Budget Gate Stack, Radical Treatments.

## I. INTRODUCTION

Low thermal budget gate stack fabrication is a key enabler for several upcoming CMOS technology innovations. In contemporary High-k Metal Gate (HKMG) Replacement Gate (RMG) flows a so-called ‘reliability anneal’ is performed after HK deposition to cure dielectric defects [1]. This high-temperature (T) process step ( $\sim 850^\circ\text{C}$ ) is not suitable for Sequential 3D integrations of stacked transistors (Fig. 1), as top device tiers need to be fabricated at reduced thermal budget [2] to preserve the functionality of the bottom tiers already in place. For monolithic CFET [3] fabrication, in a “RMG-last” flow the high-T reliability anneal can degrade the contact performance.

Submitted for review on July 14, 2023. This work was performed under the imec CMOS partner program; the authors acknowledge the European Commission and regional authorities for their support of

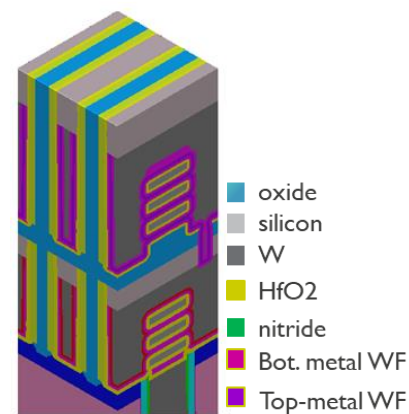


Fig. 1. Illustration (Coventor® image) of a CFET consisting of two groups of stacked nanosheets ( $p$ - and  $n$ MOS). A tight vertical pitch poses limits on the total gate thickness; either Sequential 3D or monolithic integration flows pose limits on the top tier fabrication thermal budget.

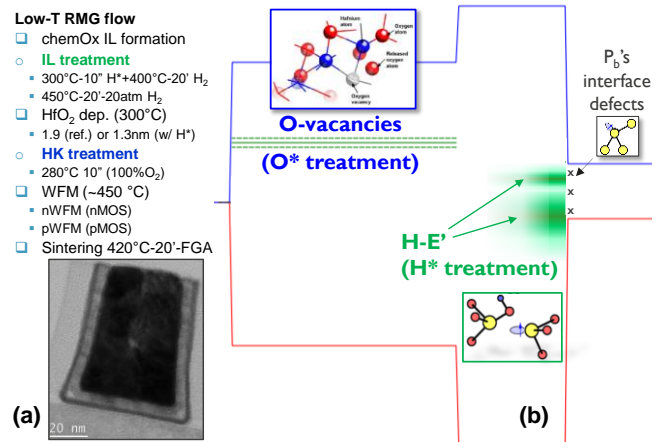


Fig. 2. (a) Schematic of the low-T RMG flow, with insertion of the novel  $H^+$  and  $O^+$  treatments before and after the HK deposition, respectively. The inset shows a TEM of one such fabricated gate. (b) Band diagram illustrating the  $\text{SiO}_2$  and  $\text{HfO}_2$  defect bands (related to the hydroxyl- $E'$  and oxygen vacancies defects, respectively) targeted by each treatment.

European collaborative projects. All authors are with imec, Kapeldreef 75, 3001 Leuven – Belgium (e-mail: Jacopo.Franco@imec.be).

Conversely, moving the RMG module to earlier in the flow (“RMG-first”) imposes stringent thermal stability requirements for the HKMG stacks to endure the epi and contact module thermal budgets (max  $T \sim 525^\circ\text{C}$  for several hours), which would be particularly challenging for the tight effective Work Function (eWF) control [4] required by multi- $V_{\text{th}}$  technologies. Furthermore, a standard RMG “reliability anneal” requires the deposition of a (relatively thick, typ.  $\sim 7\text{nm}$ ) sacrificial gate (TiN/a-Si), which is then replaced with the final WFM after a high-T spike anneal [1]; this integration flow, and in particular the removal of the sacrificial gate, become challenging for stacked nanosheets with tight vertical pitches (see Fig. 1). The development of novel low thermal budget RMG stacks discussed here (Fig. 2a), with simplified integration and with competitive performance and reliability as compared to state-of-the-art high thermal budget RMG, addresses these concerns.

In order to develop effective low thermal budget gate stack treatments, it is crucial to understand the microscopic nature of the electrically-active defects present in the as-deposited dielectric stack. In [5] we have shown that hydroxyl- $E'$  ( $H-E'$ ) defects [6] are abundant in  $\text{SiO}_2$  (interfacial) layers grown at reduced  $T$  ( $\leq 600^\circ\text{C}$ ) and not subsequently annealed at high  $T$  ( $\geq 700^\circ\text{C}$ ); this defect structure form at stretched Si-O-Si bonds, which exist due to unrelaxed interface strain: the incorporation of an OH-group (recall hydrogen is ubiquitously available in the gate stack fabrication flow, including during the formation of the  $\text{SiO}_2$  interfacial layer by wet chemical oxidation – chemOx – or by in-situ steam generation – ISSG) releases the strain locally, but leaves behind a Si dangling bond (see Fig. 2b). In [7] we have shown that the density of these defects peaks at the Si/ $\text{SiO}_2$  interface, and decays exponentially across the IL depth with a characteristic depth of 0.5nm, in agreement with the interface strain profile. Ab-initio calculations predict the  $H-E'$ s to behave as amphoteric defects [8], with charge transition levels located below the Si conduction band (and thus potentially detrimental for nMOS mobility) and above and below the Si valence band (thus behaving either as positive “fixed” charge or as a NBTI hole trap [5], and potentially detrimental for pMOS mobility). In  $\text{HfO}_2$  instead, oxygen vacancies ( $V_{\text{O}}$ ) with charge transition level above the Si conduction band (see Fig. 2b) are often quoted as the primary

responsible for electron trapping (and thus degraded nMOS PBTI reliability) [9]. To passivate efficiently the  $H-E'$ s in the  $\text{SiO}_2$  interfacial layer (IL) and the  $V_{\text{O}}$ s in  $\text{HfO}_2$  at low temperature, a remote plasma treatment was developed to provide hydrogen ( $\text{H}^*$ ) [5,7,10] and oxygen ( $\text{O}^*$ ) [11] radicals at suitable insertion points during the RMG flow, and efficiently passivate defects at low  $T$  ( $\leq 300^\circ\text{C}$ ). The  $\text{H}^*$  can deactivate the  $H-E'$  defect either by passivating its Si dangling bond, or by removing the OH- or H- present at the defect sites (releasing a  $\text{H}_2\text{O}$  or a  $\text{H}_2$  molecule respectively), and thus transforming the defect configuration into an (electrically-inactive) oxygen vacancy [6], or back into a (stretched) Si-O-Si bond. The  $\text{O}^*$  instead appears to passivate efficiently  $V_{\text{O}}$ s in  $\text{HfO}_2$  at low- $T$ .

In Section II we summarize the impact of the  $\text{H}^*$  IL treatment on the electrical properties of a pMOS gate stack with a  $p$ -WFM metal (WFM, TiN), as assessed on planar MOS capacitors. In Section III, a similar analysis is reported regarding the impact of the  $\text{O}^*$  HK treatment on a nMOS gate stack with  $n$ -WFM (TiAl-based). In Section IV, we evaluate various low thermal budget RMG gate stacks comprising  $\text{H}^*$ ,  $\text{H}_2$  and  $\text{O}^*$  treatments (see Table I) on a planar CMOS transistor platform and benchmark them against high- $T$  RMG and Gate-First HKMG counterparts. We find that, due to the induced remarkable improvement of the IL quality, the  $\text{H}^*$ - $\text{H}_2$  treatment boosts both the pMOS and nMOS carrier mobilities and both NBTI and PBTI reliability. Furthermore, a substantial improvement of the short channel nMOS Safe Operating Area is also observed, due to a reduced non-equilibrium charge trapping in the IL during Channel Hot Carrier stress. Thanks to the simultaneous improvement of pMOS and nMOS device metrics, and to a better EOT scalability as compared to the  $\text{O}^*$  HK treatment, the  $\text{H}^*$ - $\text{H}_2$  IL treatment comes out as a complete stand-alone solution for low thermal budget CMOS gate stacks.

## II. EFFECTS OF THE $\text{H}^*$ IL TREATMENT

The  $\text{H}^*$  treatment is performed right after formation of the IL (see Fig. 2a) to maximize the passivation of the pre-existing  $\text{SiO}_2$   $H-E'$ . Performing the  $\text{H}^*$  treatment later in the flow results in an insufficient passivation degree (not shown), likely due to

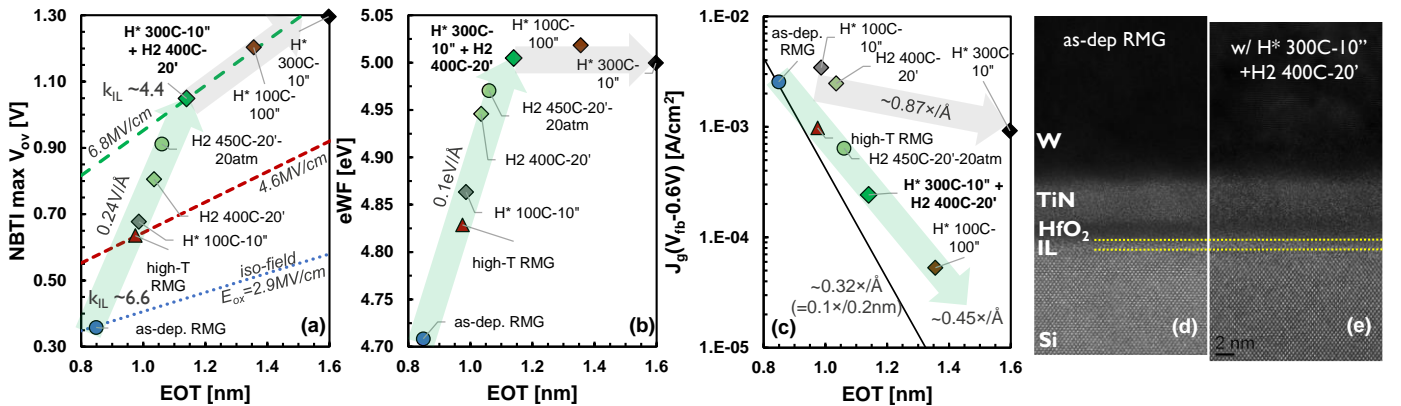


Fig. 3. (a) NBTI max. operating  $V_{\text{ov}}$  (defined as the overdrive inducing a 2.5mV shift after 1ks stress at  $25^\circ\text{C}$ , which projects to a  $\sim 50\text{mV}$  shift after 10-year at  $125^\circ\text{C}$  for a typ. kinetics) of low-T RMG gate stacks with different  $\text{H}^*/\text{H}_2$  IL treatments, plotted vs. EOT (note: a standard high-T RMG gate stack is also shown as ref). (b) eWF. (c) and (c) gate leakage density at  $V_{\text{G}} = V_{\text{fb}} - 0.6\text{V}$  plotted vs. EOT for each gate stack. (d) HR-TEM's of the as-deposited RMG gate stack, and (e) with insertion of the  $\text{H}^*$ - $\text{H}_2$  IL treatment.

the radicals reacting at other sites in the dielectric stack volume before reaching the H-E' preferentially located near the channel interface. A limited concentration of hydrogen radicals can be conveniently generated in-situ without plasma processing by annealing the IL at 400-450°C in H<sub>2</sub> before HKMG deposition, exploiting the native interface P<sub>b</sub> centers (see Fig. 2b) as hydrogen molecule cracking sites [12]. In this case, the generated density of H\* equals (at most) the native surface density of P<sub>b</sub> defects, i.e., 5-7×10<sup>12</sup> /cm<sup>2</sup>, irrespective of the dielectric thickness: this simplified approach is therefore not suitable for thicker SiO<sub>2</sub> layers [5], or if applied after HK deposition, as the generated radicals can “get stuck” at other sites in the dielectric volume before they can reach the closest H-E' near the interface. This explains why standard anneals in molecular hydrogen or forming gas performed at the end of the device fabrication flow are not sufficient to passivate the H-E' and improve adequately the quality of low-T IL's [5].

Fig. 3 depicts the impact of various H\* and H<sub>2</sub> IL treatments on the electrical properties of a pMOS gate stack, as assessed on MOS capacitors. Stronger H\*/H<sub>2</sub> treatments (i.e., longer duration or higher T) result in larger NBTI improvement, but also in increased EOT (Fig. 3a). Notice the reliability improvement is at first much stronger than the expectation related to the reducing oxide electric field (due to the increasing EOT), confirming a genuine reduction of the dielectric trap density; eventually, for the strongest treatments, the NBTI improvement rate saturates towards an iso-electric field line. The eWF (Fig. 3b) also increases upon the H\*/H<sub>2</sub> treatments, suggesting a reduction of the positive charge in the IL: this is readily explained by the H-E' defect band spanning above and below the Si E<sub>v</sub> [6,8] (see Fig. 2b), and thus behaving partially as positive “fixed” charge, and partially as NBTI hole traps. The eWF change eventually saturates for longer treatments, suggesting that the number of passivable H-E' defects fades out (consistent with the saturation of the NBTI improvement rate). The H\*/H<sub>2</sub> treatments are found to also reduce the gate leakage density (Fig. 3c); the J<sub>g</sub> reduction rate vs. EOT increase is however incompatible with an IL thickness increase (as the latter would induce a stronger J<sub>g</sub> reduction of ~10× per 2Å); indeed High Resolution-Transmission Electron Micrographs (HR-TEM's, Figs. 3d,e) do not reveal any significant difference

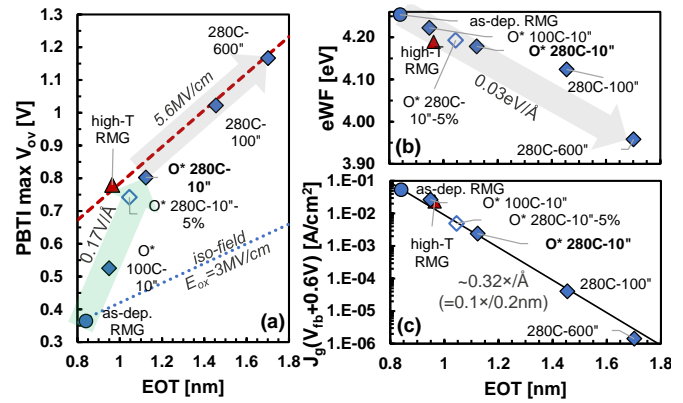


Fig. 4: (a) PBTI max. operating V<sub>ov</sub> of low-T RMG gate stacks with different O\* HK treatments, plotted vs. EOT (note a standard high-T RMG gate stack is also shown as ref). (b) eWF and (c) gate leakage density at V<sub>G</sub>=V<sub>fb</sub>+0.6V plotted vs. EOT for each gate stack.

in the IL thickness w/o or w/ the H\* treatment. We therefore ascribe the reduced leakage to a bandgap widening of the chemOx upon defect passivation, towards the nominal 9eV of thermal SiO<sub>2</sub> [10]. Similarly, the EOT increase up to ~3Å observed for moderate H\*/H<sub>2</sub> treatments is ascribed to the lowering of the chemOx relative permittivity from ~6.6 (as-deposited, i.e., containing a large density OH group) towards ~3.9 (nominal for high quality SiO<sub>2</sub>) upon removal of OH-groups at the defect sites, and thus should be considered as an unavoidable side effect of the H-E' defect deactivation. For stronger H\* treatments, an EOT increase up to ~8Å is observed, without a corresponding J<sub>g</sub> reduction (Fig. 3c): this is ascribed to suboxide formation, likely at the interface P<sub>b</sub> sites de-passivated by the excessive presence of H radicals [7]. To avoid this detrimental effect, the strongest H\* treatment condition (300°C-10") is coupled with an early H<sub>2</sub> sintering anneal (400°C-20') to restore an optimal passivation of the P<sub>b</sub>'s already before HK deposition [7]. The H\*-H<sub>2</sub> combination results in the largest NBTI and eWF improvement with a moderate EOT increase (~3Å) solely related to the IL k-value change. A H<sub>2</sub> IL treatment only instead minimizes the EOT penalty (~1.5Å) while still yielding a superior NBTI reliability as compared to a high-T RMG ref. These two IL treatment conditions have been selected for the transistor benchmark discussed later in Section IV.

	Gate Stack	Low-T IL treatment	HfO <sub>2</sub> [nm]	Low-T HK treatment	HK PDA	Reliability anneal	nMOS WFM	pMOS WFM
Low thermal budget	● RMG as-dep.	-	1.9	-	-	-	TiN/TiAl/TiN (1/4/3 nm)	TiN (5nm)
	⊠ HP H <sub>2</sub> on IL (thin HK)	H <sub>2</sub> 450°C-20'-20atm	1.3	-	-	-	-	TiN (5nm)
	▲ H*+H <sub>2</sub> on IL (thin HK)	H* 300°C-10" + H <sub>2</sub> 400°C-20'-1atm	1.3	-	-	-	TiN/TiAl/TiN (1/4/3 nm)	TiN (5nm)
	■ O* on HK	-	1.9	O* 280°C-10"	-	-	TiN/TiAl/TiN (1/4/3 nm)	-
	◆ H*+H <sub>2</sub> on IL +O* on HK (thin HK)	H* 300°C-10" + H <sub>2</sub> 400°C-20'-1atm	1.3	O* 280°C-10"	-	-	TiN/TiAl/TiN (1/4/3 nm)	TiN (5nm)
High thermal budget refs.	◆ PDA 850C ref.	-	1.9	-	He 850°C-1.5"	-	TiN/TiC/TiN (0.5/2/3nm)	TiN (8nm)
	● SSA 850C Rel. Ann. ref.	-	1.9	-	-	2/5nm TiN/a-Si +He 850°C-1.5"	TiN/TiAl/TiN (1/4/3 nm)	-
	✗ Gate-First HKMG refs.	-	-	-	-	-	-	-

Table I. Experimental RMG splits comprising H\*, H<sub>2</sub>, O\*, and H\*+H<sub>2</sub>+O\* low-T treatments. An “as-deposited” RMG is used as low-T ref., while RMG splits with 850°C post-HK or post-metal “reliability” anneals or a first-generation HKMG with Gate-First integration are used as high-T refs. For V<sub>fb</sub> tuning, a low WFM stack comprising TiN and TiAl(C) is used for nMOS, while a high WFM (TiN) is used for pMOS (see Fig. 5). All the H\* splits have a scaled HfO<sub>2</sub> to compensate for the expected EOT penalty due to the reduced IL permittivity.

### III. EFFECTS OF THE O\* HK TREATMENT

The O\* HK treatment [11] yields a substantial PBTI reliability improvement at processing T's  $\leq 280^\circ\text{C}$  (Fig. 4a). A 280C-10'' O\* exposure yields a PBTI reliability comparable to the high-T RMG ref. with an EOT penalty of just  $\sim 1.5\text{\AA}$  (or  $\sim 3\text{\AA}$  w.r.t. the the as-deposited RMG ref.) and is thus selected as the HK low-T treatment condition for the transistor benchmark discussed later in Section IV. Longer treatments result in larger reliability improvement but also larger EOT increases (up to  $\sim 9\text{\AA}$ ). The  $J_g$ -EOT relation (Fig. 4c) clearly reveals an increase of the IL physical thickness, due to the radicals oxidizing the Si surface. The EOT increase can be reduced by reducing the oxygen content in the remote plasma ambient (e.g., from 100% to 5%, topped up by He, open symbol in Fig. 4), but remains to some extent unavoidable. The O\* HK treatment also induces an eWF reduction (Fig. 4b), either by i) reduction of negative dielectric charge, or ii) increase of the total positive fixed charge in the thickened IL due to lower-T oxidation, or iii) enhanced electrostatic weight of the interface positive charge upon EOT increase: Notice, however, that the eWF modulation rate is much smaller compared to the H\* IL treatment (cf. Fig. 3b and 4b), and thus the eWF change can be almost neglected for the mild O\* treatments inducing acceptable EOT increases ( $\leq 3\text{\AA}$ ).

### IV. BENCHMARK OF LOW THERMAL BUDGET GATE STACKS WITH H\*, H<sub>2</sub> AND O\* TREATMENTS ON A CMOS PLATFORM

In this Section, various low thermal budget RMG gate stacks comprising H\*, H<sub>2</sub> and O\* treatments (Table I) are evaluated

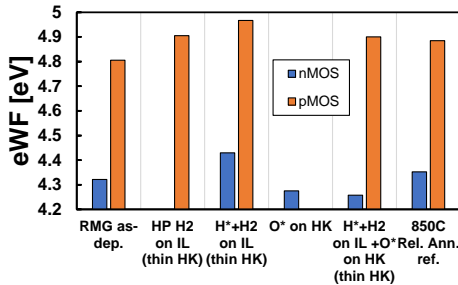


Fig. 5. eWF of each studied nMOS and pMOS split (see Table I), as estimated on the final devices.

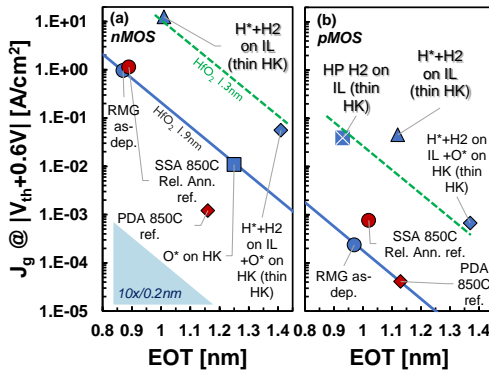


Fig. 6. Gate leakage density evaluated at  $V_{th} \pm 0.6\text{V}$  in (a) nMOS and (b) pMOS, plotted against the gate stack EOT (from C-V curves). A competitive EOT of 1nm is achieved with H\*/H<sub>2</sub> IL treatments combined with a scaled HK. The improved IL quality (increased bandgap) enables scaling of the HfO<sub>2</sub> from 1.9 nm (solid blue line) to 1.3nm (dashed green line), with a tolerable  $J_g$  penalty (nMOS  $J_g \leq 10\text{A}/\text{cm}^2$ ). Notice the O\* HK treatment always inducing a  $J_g$  reduction and an EOT penalty compatible with an increase of the IL thickness ( $10\times/0.2\text{nm}$ ).

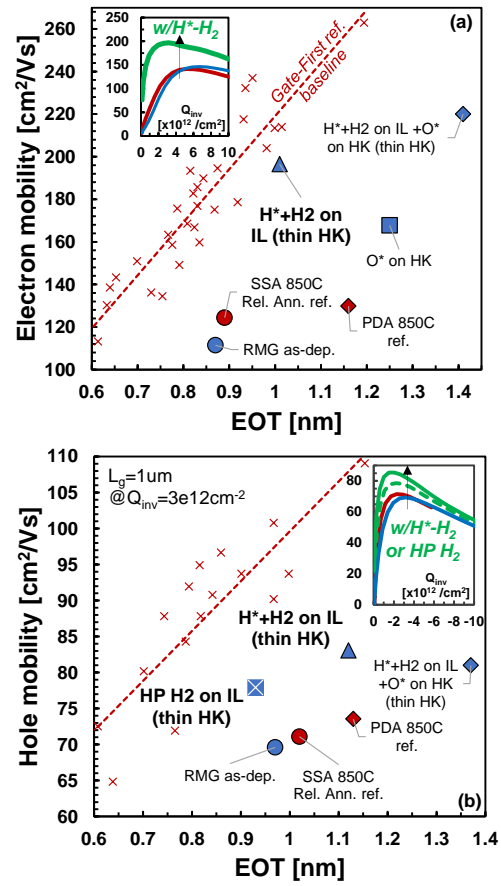


Fig. 7. Long channel split-CV (a) electron and (b) hole mobilities estimated at an inversion charge density of  $3 \times 10^{12} \text{cm}^{-2}$ . Insets: mobility vs.  $Q_{inv}$  curves for the as-dep. RMG (blue), w/ 850°C Rel. anneal (red), and w/ H\*/H<sub>2</sub> or H<sub>2</sub> IL treatment (green). The improved IL quality with the H\*/H<sub>2</sub> treatment enhances the mobilities at thin EOT almost up to the level of the Gate-First baseline, outmatching the high-T RMG counterparts. Note: no boosters (e.g., S/D stressors) were implemented in the planar transistor test vehicle.

on a planar CMOS transistor platform and benchmarked against high-T RMG and Gate-First HKMG counterparts. For proper  $V_{th}$  tuning (Fig. 5), consistently with the MOS capacitors results discussed in Sections II and III, a  $p$ -WFM (TiN) was deployed on the pMOS transistors, while the nMOS transistors employed a  $n$ -WFM stack (TiN/TiAl/TiN). To compensate for the inherent EOT penalty associated with the lower permittivity of the chemOx after H\*/H<sub>2</sub> treatments (see Fig. 3a), the HfO<sub>2</sub> thickness was scaled from 1.9nm to 1.3nm in all the gate stacks subject to IL treatment (see Table I), achieving competitive EOT's down to 1nm. The widened bandgap of the treated IL allows to maintain nMOS gate leakage at operating conditions within  $\sim 10\text{A}/\text{cm}^2$  (Fig. 6a). pMOS gate leakage does not represent a concern (Fig. 6b). The improvement of IL quality by H\*-H<sub>2</sub> treatment also results in the best mobility/EOT trade-off (Fig. 7). Both electron and hole mobilities are boosted substantially compared to the as-dep. RMG, outmatching the standard high-T RMG ref. and competing with the Gate-First HKMG baseline. Thanks to the mobility improvement, the  $I_{ON}$ - $I_{OFF}$  performance benchmarks of short channel nMOS (Fig. 8) are shown to be comparable to (or slightly better than) the conventional high-T RMG counterparts, despite the remaining  $\sim 1\text{\AA}$  EOT penalty of the H\*-H<sub>2</sub> treated gate stack (see Fig. 6a). The O\* HK treatment results in a less favorable performance

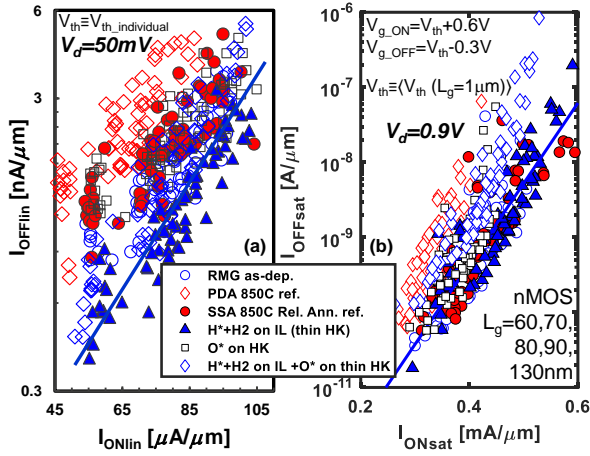


Fig. 8.  $I_{ON}/I_{OFF}$  benchmark of short channel nMOS at (a) low and (b) high  $V_d$ . Thanks to the improved IL quality and electron mobility, the performance of the  $H^*+H_2$  treated gate stack matches (or slightly outmatches) the as-dep. and high-T RMG refs. despite a  $\sim 1\text{\AA}$  thicker EOT (see. Fig. 7a).

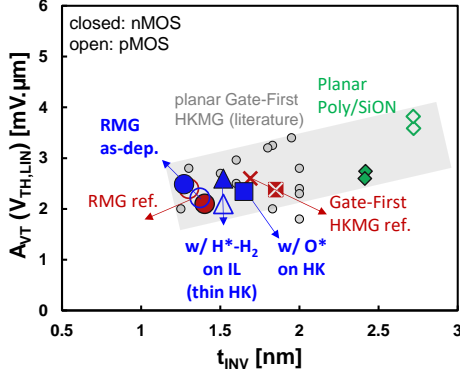


Fig. 9.  $V_{th}$  mismatch ( $A_{VT}$ ) as evaluated with matched pairs (Pelgram plot) of nMOS and pMOS device populations with gate areas in the range  $0.01\text{-}1\mu\text{m}^2$  and  $L_g$ 's in the range  $0.07\text{-}1\mu\text{m}$ . No increase of stochastic variability is observed in the low-T RMG gate stacks, irrespective of the applied treatments.

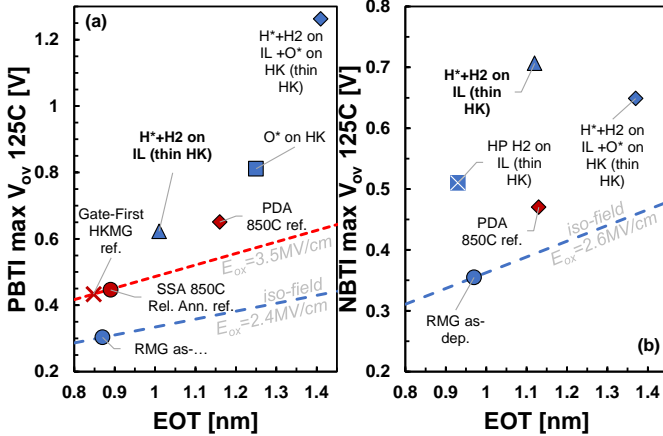


Fig. 10. BTI maximum operating overdrive of long channel ( $L_g=10\mu\text{m}$ ) (a) nMOS and (b) pMOS as estimated with fast-sensing ( $t_{\text{sense}}=1\text{ms}$ ) eMMS stress sequences at various stress  $V_g$ 's (failure criterion:  $\Delta V_{th}=6.65\text{mV}$  after 1ks stress at  $125^\circ\text{C}$ , corresponding to  $50\text{mV}$  at 10 years). The  $H^*+H_2$  IL treatment yields best NBTI and PBTI at 1nm EOT.

benchmark, due to the  $\sim 3\text{\AA}$  penalty associated with the IL regrowth, which cannot be compensated by a tolerable HK thickness reduction. A similar consideration applies to the gate stack subject to both  $H^*+H_2$  IL treatment and  $O^*$  HK treatment (total EOT penalty  $\sim 5.5\text{\AA}$ ).

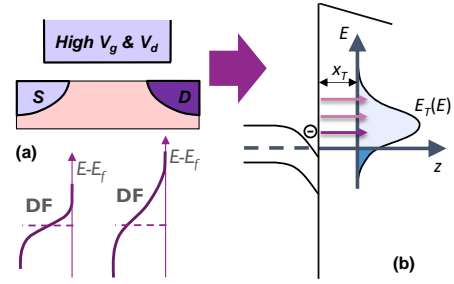


Fig. 11. (a) Sketch of the non-equilibrium carrier energy distribution function (DF) towards the drain side of a short  $L_g$  device stressed with high  $V_g$  and high  $V_d$  (Channel Hot Carrier stress). (b) Highly energetic carriers can get trapped also in dielectric defect levels at higher energies.

This first demonstration of low thermal budget RMG gate stacks in a scaled CMOS transistor platform allows us to investigate also the stochastic device variability, by utilizing statistical  $V_{th}$  datasets of matched pairs across wafer [13]. No degradation of the matching parameter ( $A_{VT}$ ) is observed in the devices with low thermal budget RMG as compared to high-T counterparts, irrespective of the deployed treatments (Fig. 9).

Fig. 10 reports the BTI reliability benchmark of long channel transistors. Consistently with the MOS capacitor results presented in Section II, excellent pMOS NBTI reliability is achieved with the  $H^*+H_2$  IL treatment (Fig. 10b). Interestingly, the same gate stack also yields the best nMOS PBTI/EOT tradeoff (Fig. 10a), ascribed to the improved IL quality beneficial also for suppressing charge trapping into HK electron traps. In scaled  $L_g$  devices stressed at high  $V_g$  and  $V_d$ , the non-equilibrium carrier energy distribution at the channel drain side is expected to cause charge trapping in a wider energy range of dielectric defect levels (Fig. 11) [14]. Degradation maps [15] measured at various  $V_g/V_d$  stress combinations allow to estimate the “safe operating area” (SOA) of each nMOS flavor by taking into account also Channel Hot Carrier degradation (Fig. 12). Due to the large defect density in both the IL and HK when omitting the high-T “reliability anneal”, the as-dep. RMG ref. shows a small SOA compared to the high-T counterpart (Fig. 12b vs. a). The  $H^*+H_2$  IL treatment improves the overall reliability (Fig. 12c) thanks to improved IL quality and suppressed tunneling towards HK defects, yielding a nMOS SOA comparable to the high-T RMG ref. The operating  $V_{DD}$  enhancements induced by the various treatments, as estimated from the SOA of each gate stacks, are reported in Fig. 13: once again the  $H^*+H_2$  IL treatment results in the best reliability/EOT tradeoff, matching the high-T refs. at comparable EOT.

## V. CONCLUSIONS

We discussed novel low thermal budget RMG gate stack solutions, based on hydrogen and oxygen radical treatments. A  $H^*$  IL treatment at  $T \leq 300^\circ\text{C}$  reduces the density of  $H-E'$  defects in  $\text{SiO}_2$ , resulting in improved pMOS NBTI reliability, increased eWF due to a reduction of the IL positive fixed charge, and reduced gate leakage, at the cost of  $1\text{-}3\text{\AA}$  inherent EOT penalty related to the removal of OH groups lowering the  $k$ -value of the chemOx IL. An early  $H_2$  sintering anneal at  $400\text{-}450^\circ\text{C}$  before HKMG deposition provides a limited amount of  $H^*$  radicals without plasma processing, by cracking of the hydrogen molecule at native interface  $P_b$  defects, resulting in a

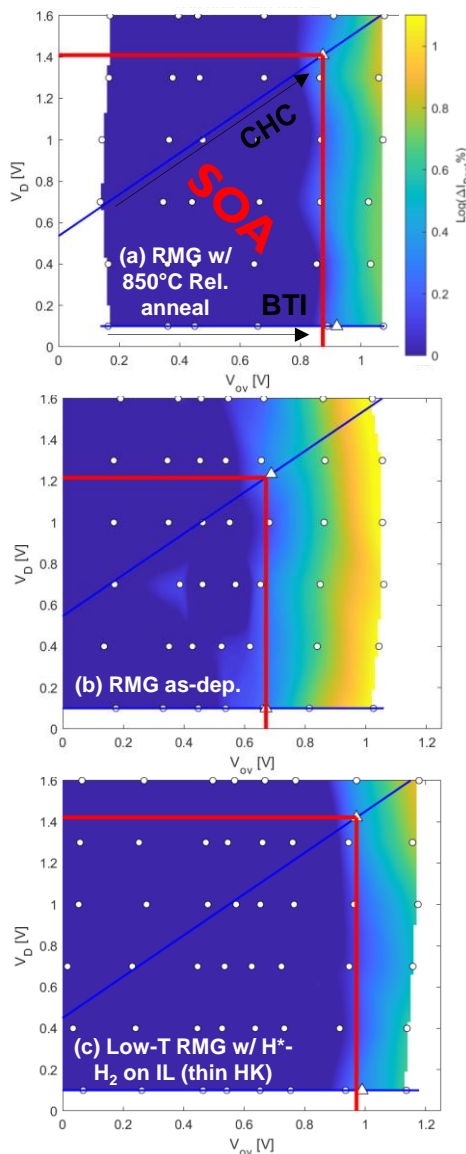


Fig. 12. Degradation map measured (a) in a ref. high-T RMG short  $L_g$  nMOS for different  $V_g/V_d$  bias combinations (each circle demarcates a tested condition). The device “safe operating area” (SOA, red rectangle) is determined within the stress  $V_{DD}$  (triangles) inducing a 2%  $I_{Dsat}$  drop after 0.5ks of stress at 25°C on the BTI (high  $V_g$ , low  $V_d$ ) or on the CHC ( $V_g=V_d$ ) cutlines. (b) The SOA is substantially reduced when omitting the high-T reliability anneal, but (c) is recovered with the H\*–H<sub>2</sub> low-T IL anneal.

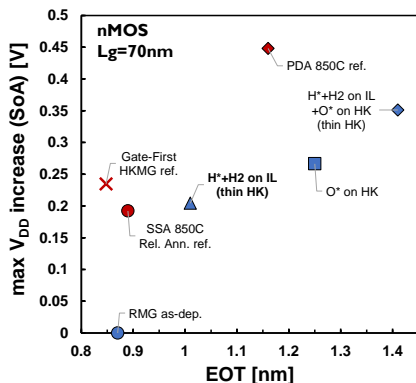


Fig. 13. Benchmark plot of the nMOS SOA  $V_{DD}$  increases (as estimated from the degradation maps, cf. Fig. 12) achieved with the various gate stack treatments, referred to the as-deposited RMG ref.

sizable improvement of the IL quality with a minimum EOT penalty of  $\sim 1\text{\AA}$ . An O\* HK treatment reduces the density of  $V_o$  defects in  $\text{HfO}_2$ , resulting in an improved nMOS PBTI reliability; however, this treatment induces an unavoidable EOT penalty of 2–3 $\text{\AA}$  due to re-oxidation of the Si surface. Low thermal budget RMG gate stacks employing (combinations of) these radical treatments were deployed on a planar CMOS transistor platform and benchmarked against conventional high-T gate stacks. The H\*/H<sub>2</sub> IL treatment was shown to provide a complete low-T RMG CMOS solution, due to the improved IL quality resulting in improved pMOS and nMOS performance and reliability. By deploying this IL treatment all the key electrical metrics of a ref. high-T RMG stack were matched (or outmatched), including hole and electron mobilities, NBTI and PBTI reliability,  $V_{th}$  matching, short channel nMOS performance and Safe Operating Area (Channel Hot Carrier reliability). The improved IL quality allowed also to scale the HK thickness to 1.3nm, partially compensating for the EOT penalty associated with the k-value decrease of the chemOx IL. The low thermal budget gate stack innovations presented here are expected to simplify the RMG integration in advanced device architectures (e.g., nanosheets), and to become enablers for upcoming CMOS technology innovations, such as Sequential-3D integration and CFETs.

#### ACKNOWLEDGMENT

Inspiring discussions with Prof. T. Grasser, A.-M. El-Sayed, M. Jech, D. Waldhoer (T.U. Wien, Austria), Profs. V. Afanas'ev and A. Stesmans (KU Leuven, Belgium), and E. Dentoni Litta (formerly at imec, Belgium) are gratefully acknowledged.

#### REFERENCES

- [1] B. Ye *et al.*, “NBTI Mitigation by Optimized HKMG Thermal Processing in a FinFET Technology”, in *IEEE Transactions on Electron Devices*, vol. 69, no. 3, pp. 905–909, March 2022, doi: <https://doi.org/10.1109/TED.2021.3139566> ;
- [2] A. Vandooren *et al.*, “Demonstration of 3D sequential FD-SOI on CMOS FinFET stacking featuring low temperature Si layer transfer and top tier device fabrication with tier interconnections”, *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, Honolulu, HI, USA, 2022, pp. 330–331, doi: <https://doi.org/10.1109/VLSITechnologyandCirc46769.2022.9830400> ;
- [3] S. Subramanian *et al.*, “First Monolithic Integration of 3D Complementary FET (CFET) on 300mm Wafers”, *2020 IEEE Symposium on VLSI Technology*, Honolulu, HI, USA, 2020, pp. 1–2, doi: <https://doi.org/10.1109/VLSITechnology18217.2020.9265073> ;
- [4] H. Arimura *et al.*, “Molybdenum Nitride as a Scalable and Thermally Stable pWFM for CFET”, *2023 IEEE Symposium on VLSI Technology*, Kyoto, Japan, 2023, pp. 1–2;
- [5] J. Franco *et al.*, “Atomic Hydrogen Exposure to Enable High-Quality Low-Temperature SiO<sub>2</sub> with Excellent pMOS NBTI Reliability Compatible with 3D Sequential Tier Stacking”, *2020 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, 2020, pp. 31.2.1–31.2.4, doi: <https://doi.org/10.1109/IEDM13553.2020.9372054> ;
- [6] T. Grasser *et al.*, “On the microscopic structure of hole traps in pMOSFETs”, *2014 IEEE International Electron Devices Meeting*, San Francisco, CA, USA, 2014, pp. 21.1.1–21.1.4, doi: <https://doi.org/10.1109/IEDM.2014.7047093> ;
- [7] J. Franco *et al.*, “Low Temperature Atomic Hydrogen Treatment for Superior NBTI Reliability—Demonstration and Modeling across SiO<sub>2</sub> IL Thicknesses from 1.8 to 0.6 nm for I/O and Core Logic”, *2021 Symposium on VLSI Technology*, Kyoto, Japan, 2021, pp. 1–2;
- [8] C. Wilhelmer *et al.*, “Ab initio investigations in amorphous silicon dioxide: Proposing a multi-state defect model for electron and hole capture”, in *Microelectronics Reliability* 139, pp. 114801, 2022, doi: <https://doi.org/10.1016/j.microrel.2022.114801>

- [9] A. Kerber and E. A. Cartier, "Reliability Challenges for CMOS Technology Qualifications With Hafnium Oxide/Titanium Nitride Gate Stacks", in *IEEE Transactions on Device and Materials Reliability*, vol. 9, no. 2, pp. 147-162, June 2009, doi: <https://doi.org/10.1109/TDMR.2009.2016954> ;
- [10] J. Franco *et al.*, "Low-temperature atomic and molecular hydrogen anneals for enhanced chemical SiO<sub>2</sub> IL quality in low thermal budget RMG stacks", *2021 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, 2021, pp. 31.4.1-31.4.4, doi: <https://doi.org/10.1109/IEDM19574.2021.9720657> ;
- [11] J. Franco *et al.*, "Low thermal budget PBTI and NBTI reliability solutions for multi-V<sub>th</sub> CMOS RMG stacks based on atomic oxygen and hydrogen treatments", *2022 International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, 2022, pp. 30.4.1-30.4.4, doi: <https://doi.org/10.1109/IEDM45625.2022.10019385> ;
- [12] A. Stesmans, "Interaction of Pb defects at the (111)Si/SiO<sub>2</sub> interface with molecular hydrogen: Simultaneous action of passivation and dissociation", in *AIP Journal of Applied Physics* 88, 489-497 (2000), doi: <https://doi.org/10.1063/1.373684> ;
- [13] M. J. M. Pelgrom, A. C. J. Duinmaijer and A. P. G. Welbers, "Matching properties of MOS transistors", in *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433-1439, Oct. 1989, doi: <https://doi.org/10.1109/JSSC.1989.572629> ;
- [14] M. Jech *et al.*, "Mixed Hot-Carrier/Bias Temperature Instability Degradation Regimes in Full {VG, VD} Bias Space: Implications and Peculiarities", in *IEEE Transactions on Electron Devices*, vol. 67, no. 8, pp. 3315-3322, Aug. 2020, doi: <https://doi.org/10.1109/TED.2020.3000749> ;
- [15] E. Bury *et al.*, "(Invited) Recent insights in CMOS reliability characterization by the use of degradation maps", *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Qingdao, China, 2018, pp. 1-4, doi: <https://doi.org/10.1109/ICSICT.2018.8565676> .