Understanding charge traps for optimizing Si-passivated Ge nMOSFETs

P. Ren[1, 3], R. Gao[1], Z. Ji[1] (z.ji@ljmu.ac.uk), H. Arimura[2], J. F. Zhang[1], R. Wang[3], M. Duan[1], W. Zhang[1], J. Franco[2], S. Sioncke[2], D. Cott[2], J. Mitard[2], L. Witters[2], H. Mertens[2], B. Kaczer[2], A. Mocuta[2], N. Collaert[2], D. Linten[2], R. Huang[3], A. V.-Y. Thean[2], and G. Groeseneken[2]
[(1)]School of Engineering, Liverpool John Moores University, Liverpool L3 3AF, UK.
[(2)] IMEC, Leuven B3001, Belgium [(3)] Institute of Microelectronics, Peking University, Beijing 100871, China.

Background and key advance: Ge is an attractive channel material offering high hole and electron mobility, and therefore of interest for future p- and n-FET technologies. Ge nFETs can be made through two routes: GeO_2/high-k directly on Ge [1] or using a Si-passivated monolayer (ML) [2]. The former offers higher mobility but poor reliability [3], while the Si-passivated option has a better balance between mobility and reliability, making it promising for the debut of Ge CMOS [4]. However, significant trapping-induced PBTI compared with the Si counterpart is the key hurdle for its practical use. To optimize it, there is a pressing need for understanding the properties of these traps as well as their impact on time-dependent mobility and reliability. *In this work, for the first time, two types of electron traps are unambiguously identified in Ge nFETs, which are controlled respectively by a) the HK layer thickness and b) the growth conditions used for the Si-passivated layer. These different traps exhibit different impacts on mobility degradation. Based on this, process is improved, as experimentally verified with maximum operation overdrive enhanced by a factor of ~1.7 (4&6ML).*

Device fabrication: Si-passivated Ge nFETs were fabricated using a replacement metal gate high-k last process with the gate stack shown in **Fig.1**. After dummy gate removal and pre-cleaning, the thin Si layer was epitaxially grown on the Ge channel with either Si_3H_8/350ºC or SiH_4/500ºC. Laser annealing was performed at 750ºC. Other process differences are detailed in **Table 1**. Based on the process conditions, the devices are split into groups: D1&D2 each with 4&6 Si-ML were used to understand the traps, their origin and dependence on processing conditions. Based on this understanding, the process was designed to fabricate D3, verifying the expected improvement. All measurements are with a speed of 5µs.

Two types of traps: Oxide traps are investigated using the energy profiling technique [5] (**Fig.2a**). When charging bias, Vgch, is low, the profiles extracted from discharging after filling at different Vgch overlap well (**Fig.3a**). However, for higher Vgch, they deviate from each other (**Fig.3b**). This deviation is not due to the incomplete discharging (**Fig.2b**). Therefore, **Figs.3a&b** indicate that there exist two types of electron traps in the dielectric with different filling mechanisms: Type-A capture electrons without changing their energy levels (inset of **Fig.3a**); Type-B, after capturing electrons, shift energy levels down from ground level to charged level (inset of **Fig.3b**), resulting in the increased ΔVth under same surface potential after filling at higher Vgch (**Fig.3b**). Following procedure proposed in ref 5 (**Fig.4a**), Type-A are separated and given in **Fig.4b** and Type-B can be obtained by subtracting Type-A from total (**Fig.4b**) and given in **Fig.4c**. To compare the energy profile for different processes and MLs in the next section, ΔVth~(Vg-ΔVth) is converted to ΔNot~(Ef-Ec_SiO_2) (**Fig.5**) [5]. Type-A are both within Si bandgap and above Ec_Si, while Type-B are above Ec_Si before capturing electrons and shifted down below Ec_Si after the capture (**Fig.5**). Moreover, Type-A saturate within seconds during filling, while Type-B do not. This clear difference supports that they have different origins.

Physical origin of the traps: Two processes D1&D2 are used to investigate the physical origin of the traps. D2 shows much larger Type-A compared with D1, while within each group, Type-A do not change with Si-passivated layer thickness (**Fig.6a**). Si nFETs with SiO_2 dielectric is found to have insignificant Type-A (◆ in **Fig.6a**), excluding the SiO_2 layer as their origin. When gate stack is replaced with SiO_2/HfO_2, however, trapping increases substantially with HfO_2 thickness (**Fig.6b**), similar to the trend observed in Ge nFETs (**Fig.6a**). This strong correlation reveals that Type-A originate from HfO_2 layer and are dominated by HfO_2 thickness. For Type-B, they also show independence on Si-passivated layer thickness. D1, however, has much larger Type-B than D2, which is opposite to Type-A (**Fig.7a**), further supporting that they have different origins. Interestingly, Type-B are negligible in Si-nFETs with SiO_2/HfO_2, indicating they are specific to Ge process. Recent *ab-initio* atomic simulation suggests Ge segregation can introduce electron traps with ground levels above Ec_Si when neutral and shift down to their charged level into Si bandgap after capturing electrons [6]. Ge segregation can happen when growing Si-passivated layer: undesirable GeO_2 will be formed during the following oxidation step. The extremely low energy secondary ion mass spectroscopy (SIMS) shows intensity of segregation is controlled by Si precursor (**Fig.7b**) [7]. Therefore the larger Type-B in D1 process can be caused by the use of SiH_4, inducing more Ge segregated into the stack. *The experimental evidence (**Fig.4-7**) clearly reveals that there exist two different types of electron traps in Ge nFETs with different filling mechanisms and process dependence (**Fig.8**).*

Impact on mobility degradation: Mobility on fresh devices is one key parameter to present the technology improvement. However, its time-dependent degradation is rarely discussed. Mobility degradation did not always increase with total trapped charge (**Fig.9a**): when trapping reach $4x10^{11}$ cm^{-2} at the initial stage, mobility degrades little, then the follow-on $2x10^{11}$ cm^{-2} increase causes dramatic mobility degradation. By separating into Type-A and –B, interestingly, mobility degradation shows clear correlation against Type-B (**Fig.9c**), but no correlation with Type-A (**Fig.9b**). This also strongly supports that Type-B locate in SiO_2 layer close to the channel, exhibiting strong impact on mobility degradation. Type-A, on the other hand, are within HfO_2 layer (**Fig.8**) away from channel and therefore show little impact.

PBTI model for long term prediction: Two different traps in PBTI lead to a change of the time exponent of total ΔVth with overdrive voltage, Vg_ov=Vgch-Vth0, making long term prediction difficult through acceleration tests (**Fig.10** and inset). Since Type-A can be filled to saturation within seconds with their values known from direct measurement (**Fig.5a**), only Type-B need prediction which can be extracted by subtracting Type-A from the total degradation. For such extraction, constant surface potential during charging needs to be maintained. This is achieved by compensating the increasing ΔVth though dynamically adjusting Vgch to keep Vg_ov-ΔVth as constant (inset in **Fig.11a**). Type-A dominate at initial stage while Type-B gradually increase following power law against time (**Fig.11a**). Time exponent, n, becomes independent of both voltage and process (**Fig.11b**,**12a&b**). The observed decrease in m/n with thinner ML reflects a stronger defect-carrier coupling [3] (**Fig.12c&d**). Compared with trapping, interface states generation is negligible (**Fig.11c**). Therefore Ge nFETs PBTI can be modelled by Type-A and –B with **Eqn (1)** for long term prediction under any Vg_ov. The proposed model is validated using D2 4&6ML by comparing the predicted ΔVth with direct measurement under low Vg_ov. Good agreement has been achieved (**Fig.13**). It is worth noting that under device operation in which Vg_ov is fixed, due to trapping-induced surface potential lowering, Type-A is not a constant but decreases with time (**Fig.13**). With the proposed model, the maximum operation overdrive, Vg_ov(max) can also be evaluated (**Figs.14a-c**).

Process improvement: By identifying HfO_2 thickness and Ge segregation as two key parameters, PBTI is expected to be suppressed with 2nm HfO_2 and Si_3H_8 for Si growth (D3 process). This has been experimentally verified: Vg_ov(max)/EOT of D3 for both 4&6ML are enhanced by a factor of 1.7, comparing with D1 and D2 (**Fig.15**).

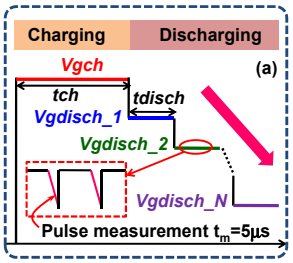
Conclusion: For the first time, two different types of electron traps are clearly identified in Ge nFETs with Type-A controlled by the HfO_2 layer thickness and Type-B by the Si growth induced Ge segregation. Only Type-B are responsible for mobility degradation and they do not saturate with stress time, while the opposite applies to Type A. A PBTI model is proposed and validated for the long term prediction.

[1] C. Lu, VLSI-T, 2015. [2] S. Sioncke, et al, ME, 2013. [3]. Franco, et al, IEDM, 2013. [4] H. Arimura, et al, IEDM, 2015. [5] Z. Ji, et al, TED, 2015, [6] S. C. Lu, et al, ISTEM, 2012. [7] B. Vincent, APL, 2009.

Si-cap Ge

Table 1. Samples used in this work

		HfO2 thickness	Si growth condition	SiO2 layer
W				
TiN				
HfO2	D1	2nm	SiH4/500C	~0.4nm (wet)
SiO2	D2	4nm	Si3H8/350C	~1.3nm (dry O3)
Si-ML				
Ge	D3	2nm	Si3H8/350C	~1.3nm (dry O3)



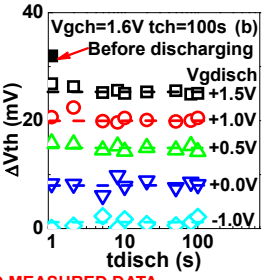
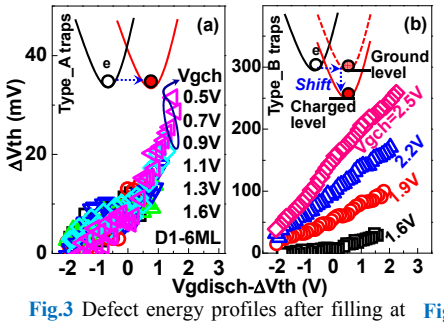
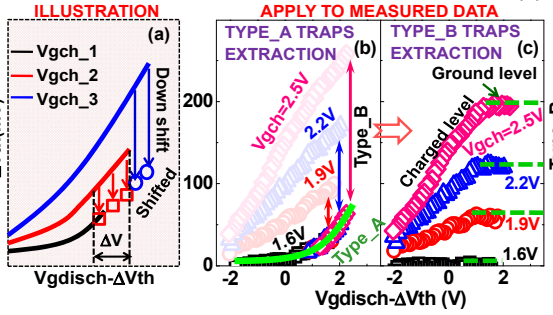


Fig.1 Gate stack of the RMG Si-passivated Ge nFETs.

Fig.2 **(a)** Waveform of Vg for defect energy profile extraction: After charging under Vg=Vgch, Vg is decreased by ΔVg to Vdisch_1. Discharging occurs and electron traps are periodically monitored by measuring Id-Vg until discharge completes. Vdisch_1 is then reduced to Vdisch_2 for the next discharging phase. **(b)** The flat regions against log(time) confirms the completion of discharging for each level.





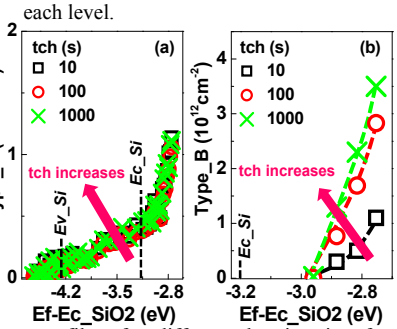
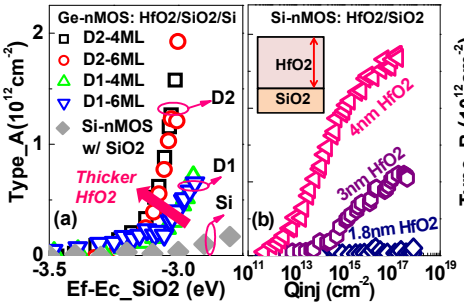
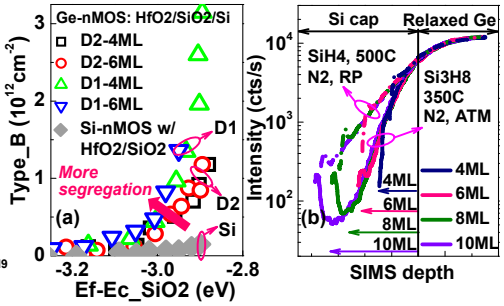


Fig.3 Defect energy profiles after filling at low **(a)** and high **(b)** Vgch. Plotting against Vgdisch-ΔVth ensures the comparison being under the same surface potential. Insets illustrate the charging mechanisms for two different types of electron traps.

Fig.4 **(a)** Illustration on separating Type-A: Starting with data for Vgch_1, data for next level of Vgch_2 shifted down to align them with data at the end of Vgch_1. This extends Type-A profile by ΔV. The process continues until reaching highest Vgch. The method is applied to real data for Type_A **(b)** and Type-B=Total -Type-A **(c)** extraction.

Fig.5 Energy profiles after different charging time for Type-A **(a)** and Type-B **(b)**. For Type-B, only ground levels are shown, which are from flat regions of Fig.4c.





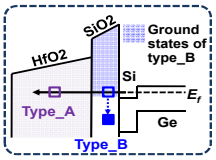


Fig.8 Spatial and energy distribution for Type-A and –B according to Fig.5-7. The empty symbols indicate its ground level and the solid symbol illustrate the charged level for Type-B.

Fig.6 **(a)** Process dependence for Type-A traps. Si-based nFET with SiO2 dielectric is shown for comparison. **(b)** Si-based nFET with HfO2/SiO2 dielectric shows significant trapping, decreasing for thinner HfO2.

Fig.7 **(a)** Process dependence for Type-B traps. Solid symbols represent Si-based nFET with SiO2/HfO2. **(b)** shows the Ge segregation intensity with different Si growth conditions [7].

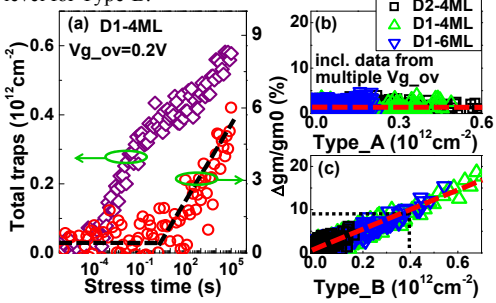
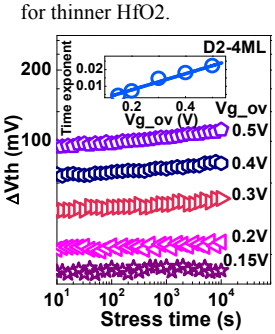
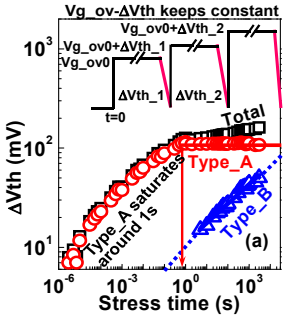


Fig.9 **(a)** Temporal kinetics for Δgm/gm0 and total traps (Type_A+Type_B). **(b)** shows Δgm/gm0 against Type-A and **(c)** Type-B respectively. Data obtained from various Vgch levels.





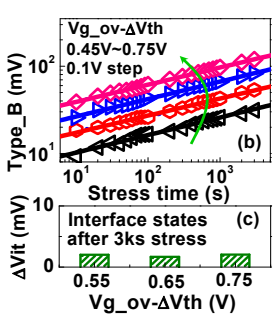


Fig.10 Conventional acceleration test on Ge nFETs. Inset shows its time exponent is not a constant for different overdrive, Vg_ov.

Fig.11 **(a)** Type-B kinetics from subtracting Type-A determined from Fig.4b from total degradation under constant Vg_ov-ΔVth condition. **(b)** The extracted Type-B traps under different Vg_ov-ΔVth. **(c)** Interface states generation is insignificant, w.r.t. trapping, evaluated by charge pumping.

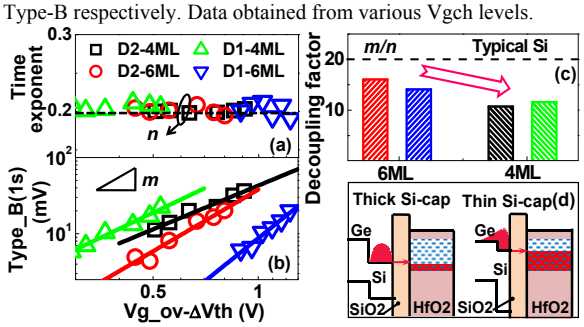
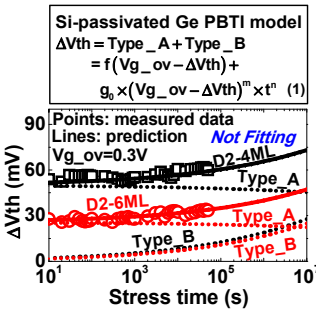
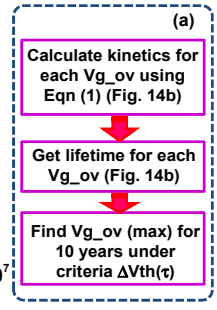
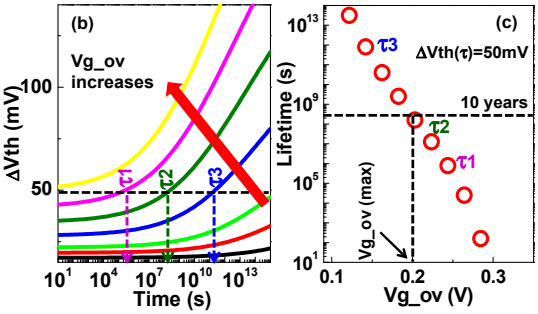
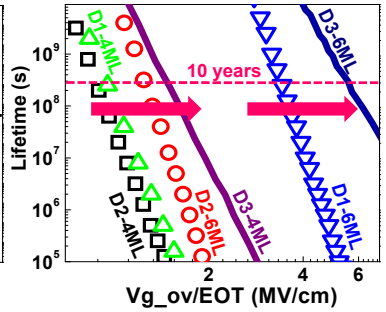


Fig.12 **(a)** Time exponent of Type-B shows unique value of 0.2 independent of stress voltage or processes. **(b)** Type-B follow power law against Vg_ov-ΔVth. **(c)** Decrease in m/n with thinner ML reflects a stronger defect-carrier coupling [3], as shown in **(d)**.









Eqn 1: PBTI Ge nFETs model for long term prediction. **Fig.13** Model validation by comparing prediction and measurement under low Vg_ov.

Fig.14 **(a)** Procedure for determining the maximum operation overdrive voltage, Vg_ov(max). **(b)** Lifetime, τ, for a given Vg_ov is calculated from the model. **(c)** By plotting lifetime against each Vg_ov(max) for 10 years lifetime can be determined for any allowed ΔVth(τ).

Fig.15 PBTI lifetime V.S. the oxide field (Vg_ov/EOT) of Ge nFETs. The improved process D3 has superior reliability.