Degradation of time dependent variability due to interface state generation

M. Toledano-Luque[1], B. Kaczer[1], J. Franco[1], Ph.J. Roussel[1], M. Bina[2], T. Grasser[2], M. Cho[1], P. Weckx[1,3], and G. Groeseneken[1,3]

[1]imec, Kapeldreef 75, B3001 Leuven, Belgium. [2]TUWien, Austria. [3]KULeuven, Belgium. (toleda@imec.be)

Although only a few defects will be present in the gate dielectric of deeply-downscaled CMOS devices, their relative impact may become intolerable [1]. An individual charged defect can significantly alter the channel current I_D of a nm-sized FET, causing V_{TH} fluctuations (RTN) and time dependent ΔV_{TH} variability. Based on detailed understanding of *atomistic* devices, random dopant fluctuation (RDF) was identified as the main source of both time-zero variability and *giant* RTN fluctuations in ultra-scaled devices [2-3]. In order to reduce variability, the industry is moving to low doped channel devices [4]. Here we show that the adverse role of RDF is taken over by charged interface states. In addition to that, stress during operation induces interface state generation, which in turn causes V_{TH} drift and further increases RTN amplitudes and time-dependent ΔV_{TH} variability.

NBTI-like stress-and-relaxation measurements (inset of Fig.1) were used to assess the impact of individual charged gate oxide defects on the characteristics of a wide range of nanoscaled planar pFETs and pFinFETs. Representative sets of relaxation transients recorded on nanoscaled pFinFETs with two gate lengths L are shown in Fig. 1. Individual defect discharge events causing abrupt, discrete ΔV_{TH} steps are visible in the relaxation traces, described as the non-steady state case of drain current RTN [5]. Since several *single-charge* ΔV_{TH}'s are extracted from each device, this NBTI-like methodology represents a fast alternative to assess the RTN V_{TH} amplitude fluctuations [1]. According to Fig. 2a, the ΔV_{TH} step heights are exponentially distributed with expectation value η [6, 7]. Fig. 2b shows opposite trends of η and the number of charged traps N_T with the FinFET L and width W, resulting in $\eta \propto 1/A$ and $N_T \propto A$, with $A=(2H+W)\times L$.

Fig 1 also points out that after identical electrical stress the *total* ΔV_{TH} strongly varies among devices and becomes more widely distributed for the shorter pFinFETs, as emphasized in Fig. 3a for *total* ΔV_{TH} at t_{RELAX}=1ms. However, the resulting average ΔV_{TH} relaxation curve is L-independent (Fig. 3b). Fig. 4 further illustrates the *total* ΔV_{TH} distributions for different W's after identical stress conditions. Since the *total* ΔV_{TH} is always unipolar, the distribution, perfectly described by model in [5] (Table I), develop a long tail towards higher values (skewness) with decreasing pFET size. The constant $\langle \Delta V_{TH} \rangle$ as a function of L and W obtained in Figs. 3 and 4 is understood when considering the equations in Table I, since $\langle \Delta V_{TH} \rangle = \eta \times N_T$ is independent of L and W. Conversely, *the total* ΔV_{TH} *dispersion* $\sigma_{\Delta VTH}$, *increases with downscaling FinFET dimensions according to* $2\eta \times \langle \Delta V_{TH} \rangle$, thus, inversely with area A at constant $\langle \Delta V_{TH} \rangle$ degradation [BK]. The understanding of η thus requires special attention.

Fig. 5a shows the distribution of *single-carrier discharge* ΔV_{TH} normalized to the expected value according to the charge-sheet approximation $\eta_0 = q \times t_{inv}/A$ for different FinFET process flows. Note in Fig. 5b that although doped FinFETs should manifest reduced η values ($\eta/\eta_0 \sim 1$) due to the low doping level [2-4], *the ratio* η/η_0 *can reach high values, which will be in turn reflected in large* $\sigma_{\Delta VTH}$ (see Table I). Fig. 6 then displays the time-dependent BTI-induced ΔV_{TH} variability, $\sigma_{\Delta VTH}$, at a constant $\langle \Delta V_{TH} \rangle$=50mV as a function of the initial V_{TH0} deviation, σ_{VTH0}. The strong correlation between the two quantities demonstrates that identical sources are responsible for time-zero and time dependent variability—as also reflected in the algebraic similarity between σ_{VTH0} and $\sigma_{\Delta VTH}$ (see Table I). The ratio between the two deviations is about 0.7, however, $\sigma_{\Delta VTH0}$ will further increase with degradation (Table 1), eventually taking over the initial σ_{VTH0} for larger $\langle \Delta V_{TH} \rangle$ values.

It has been recently claimed by means of atomistic simulations that interface trapped charges are a potential source of variability [8]. To gain further insight, the interface trap density D_{IT} of a significant number of FinFET processes has been experimentally confronted against the respective η/η_0 value. Charge pumping as a function of FinFET W was used to accurately extract D_{IT} values at the top and the side walls [9], Fig. 7. Larger sidewall D_{IT}'s were obtained with respect to the top wall D_{IT}'s for all the tested FinFETs, see Fig. 8, due to Si surface orientation. Fig. 8 also illustrates that non-optimized scavenging processes deteriorate the Si/SiO_2 interface. Comparing Figs. 5a and 8, one can realize that an increase of the interface trap density is consequently reflected in an amplification of *single-carrier discharge* ΔV_{TH}'s. Fig. 9 further validates this statement—a clear correlation is observed between η/η_0 and D_{IT} for different FinFET process flows. *Charged interface states induce additional potential fluctuations in the channel, increasing preferential conduction paths between the drain and the source* (inset of Fig. 9). Charging and discharging of oxide traps over critical points of these conduction paths will produce significant fluctuations of the drain current, and consequently, time dependent V_{TH} variability.

This scenario is further strengthened by comparing the interface trap densities and the *single-carrier discharge* ΔV_{TH} distributions of fresh devices and after appreciable constant gate voltage stress. In Fig. 10a, a significant generation of the top and sidewall interface traps is observed. In parallel, an increase of the percentage of *single-carrier discharge* ΔV_{TH} values is detected in Fig. 10b, *implying an enhancement of the channel conduction percolation.* Consequently, *an increase of RTN amplitude and a degradation of the time-dependent variability during circuit operation are expected due to interface degradation*, as predicted in Fig. 11.

In conclusion, we have experimentally demonstrated that interface charged traps represent a new source of variability to be considered in low doped channel devices. RTN amplitude and time-dependent BTI-induced variability in low doped pFETs is further aggravated by interface trap generation during operation.

Refs: [1] T. Grasser, IEDM2010, [2] A. Asenov, TED2003, [3] A. Ghetti, IEDM2008, [4] K. J. Kuhn, TED2001, [5] B. Kaczer, EDL2010, [6] J. Franco, IRPS2012, [7] S. O'uchi, IEDM08, [8] X. Wang, IEDM12, [9] G. Kapila, EDL2007.

Acknowledgments: This work was performed as part of imec's Core Partner Program and partially supported by the EU under the 7th Framework Programme (Collaborative project MORDRED, contract No. 261868).

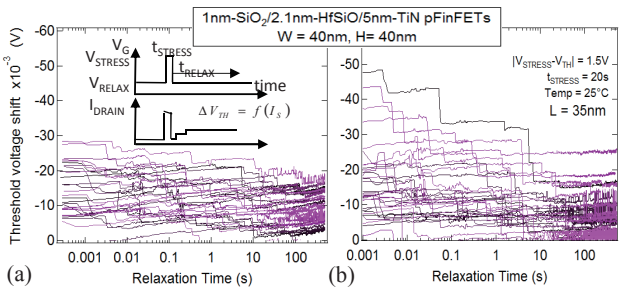


Fig. 1: NBTI-like relaxation transients obtained in (a) large area and (b) deeply-downscaled pFinFETs show single carrier discharge events. A drastic increase of ΔV_{TH} variability is visible with downscaling FinFET area, although similar degradation is observed *on average* (see Fig. 3).

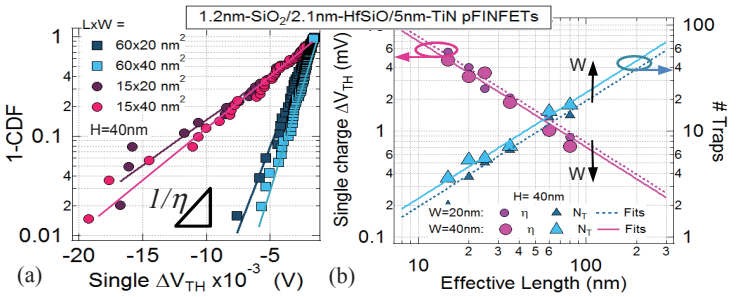


Fig. 2: (a) Complementary CDF of *single-carrier discharge* ΔV_{TH} follows to the first order an exponential distribution with average value η. (b) Opposite trends with L *and* W are observed for η and N_T, resulting in $\eta \propto 1/A$ and $N_T \propto A$ with $A=L\times(2H+W)$, H being the fin height. Thus, $\langle Total\ \Delta V_{TH} \rangle = \eta \times N_T$ is independent of L and W as shown in Figs. 3 and 4, respectively.

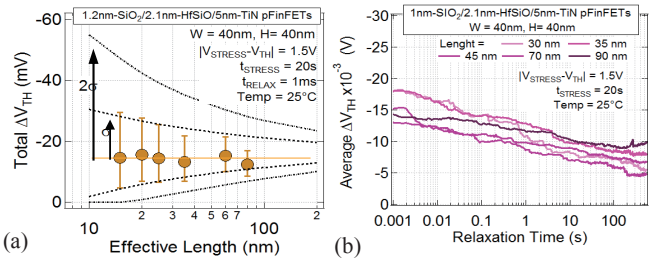


Fig. 3: (a) Mean ΔV_{TH} at t_{RELAX}=1ms, $\langle total\ \Delta V_{TH} \rangle$ ($=\eta \times N_T$), is independent of FinFET length, however, a large deviation ($\pm\sigma$) is measured for the shortest pFinFETs. (b) When the relaxation transients (Fig. 1) are averaged, similar ΔV_{TH} curves vs. t_{RELAX} are obtained *independently of FinFET area*, confirming η and N_T scaling with area.

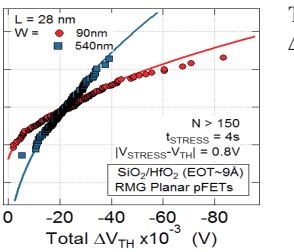


Fig. 4: *Total* ΔV_{TH} CDF (markers) for low EOT pFETs with different W are perfectly described by model (lines) in [5]. Note the distributions here are for negative values only, pivot at the median value and become wider with decreasing A.

Table I: Statistical description of *total* ΔV_{TH} and time-zero V_{TH}-variability σ_{VTH0}^2.

CDF, mean and deviation for *total* ΔV_{TH}:

$$H_{\eta,N_T}(\Delta V_{TH}) = \sum_{n=0}^{\infty} \frac{e^{-N} N_T^{\,n}}{n!} \left[1 - \frac{n}{n!}\Gamma\left(n, \left(\frac{\Delta V_{TH}}{\eta}\right)\right)\right]$$

$$\langle \Delta V_{TH} \rangle = \eta \times N_T \qquad \sigma^2_{\Delta V_{TH}} = 2\eta \times \langle \Delta V_{TH} \rangle$$

Skewness: $6\eta^2 \times \langle \Delta V_{TH} \rangle > 0$

Mean of *single-carrier discharge* ΔV_{TH}: $\eta \propto \frac{t_{ox}\sqrt{N_A}}{WL}$

BTI variability: $\sigma^2_{\Delta V_{TH}} = 2\frac{t_{ox}\sqrt{N_A}}{WL} \times \langle \Delta V_{TH} \rangle$

Charge sheet approx.: $\eta_0 = \frac{qt_{ox}}{WL}$ $\qquad$ V_{TH0} variability [4]: $\sigma^2_{V_{TH0}} \propto \frac{t_{ox}\sqrt{N_A}}{WL}$

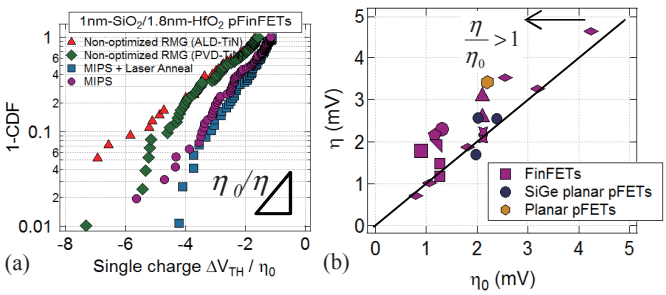


Fig. 5: (a) Larger impact per *single-carrier discharge* ΔV_{TH} is observed for non-optimized devices. Note the decrease of slope, i.e., increase of η. (b) Although low doped channel FETs are expected to be less sensitive to channel current percolations, large η values are obtained with respect to the expected η_0.

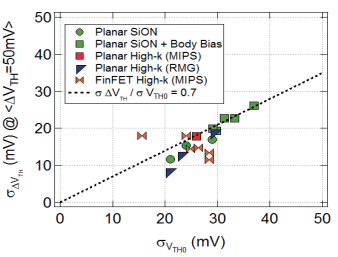
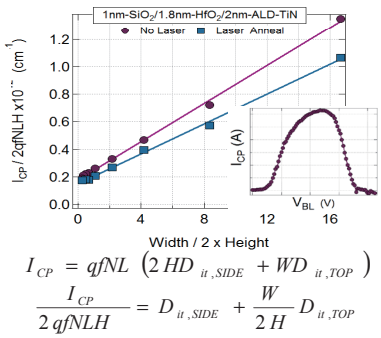


Fig. 6: The spread of the total ΔV_{TH} after NBTI ($\sigma_{\Delta VTH}$) is correlated to V_{TH0} variability (σ_{VTH0}) of fresh devices (Table I). For an average total $\langle \Delta V_{TH} \rangle$= 50 mV, the ratio is equal to 0.7.



$$I_{CP} = qfNL\left(2HD_{it,SIDE} + WD_{it,TOP}\right)$$

$$\frac{I_{CP}}{2qfNLH} = D_{it,SIDE} + \frac{W}{2H}D_{it,TOP}$$

Fig.7: Charge pumping as a function of fin width allows to independently estimate the side wall and top wall D_{IT} (see equations) [9].

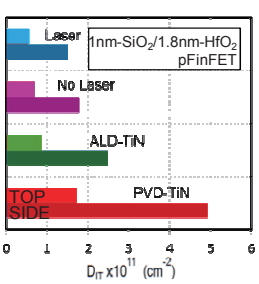


Fig. 8: A higher sidewall D_{IT} is observed for all the FinFETs. Non optimized devices can show 3× larger D_{IT} than optimized ones.

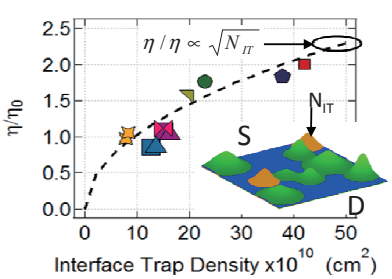


Fig. 9: A clear correlation is found between *single-carrier discharge* ΔV_{TH} and D_{IT}. ***Interface traps enhance channel conduction percolation between source and drain in the same way dopants do.***

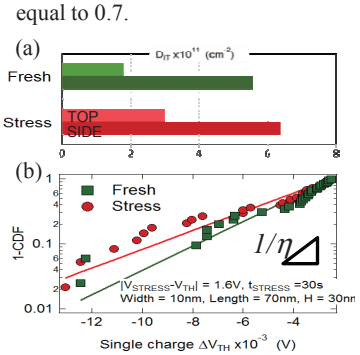


Fig. 10: (a) Generation of D_{IT} after electrical stress is reflected in (b) increase of *single-carrier discharge* ΔV_{TH}.

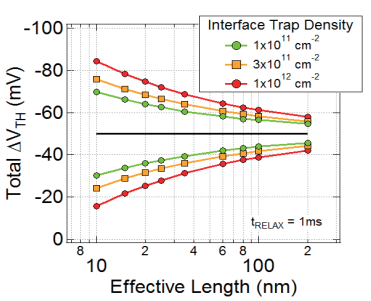


Fig. 11: Projected dispersion ($\pm\sigma_{\Delta VTH}$) of BTI induced V_{TH}-variability for different levels of D_{IT}. Increase of D_{IT} is reflected in η (Fig. 9), and in turn, in $\pm\sigma_{\Delta VTH}$ (Table I).