SiGe Channel Technology: Superior Reliability Toward Ultra-Thin EOT Devices—Part II: Time-Dependent Variability in Nanoscaled Devices and Other Reliability Issues

Jacopo Franco, *Student Member, IEEE*, Ben Kaczer, María Toledano-Luque, Philippe J. Roussel, Thomas Kauerauf, *Member, IEEE*, Jérôme Mitard, Liesbeth Witters, Tibor Grasser, *Senior Member, IEEE*, and Guido Groeseneken, *Fellow, IEEE*

***Abstract*—The time-dependent variability of nanoscaled $Si_{0.45}Ge_{0.55}$ pFETs with varying thicknesses of the Si passivation layer is studied. Single charge/discharge events of gate oxide defects are detected by measuring negative bias-temperature instability (NBTI)-like threshold voltage ($V_{\rm th}$) shift relaxation transients. The impact of such individually charged defect on device $V_{\rm th}$ is observed to be exponentially distributed. SiGe channel devices with a reduced thickness of their Si passivation layer show a reduced average number of active defects and a reduced average impact per charged defect on device $V_{\rm th}$. Our model for the superior reliability of the SiGe channel technology previously proposed in Part I, which is based on the energy decoupling between channel holes and dielectric defects, is shown to also explain these experimental observations. Other reliability mechanisms, such as $1/f$ noise, body biasing during NBTI, channel hot carriers, and time-dependent dielectric breakdown, are also investigated. None of these mechanisms are observed to constitute a showstopper for the reliability of this promising novel technology.**

***Index Terms*—Body bias, channel hot carriers (CHCs), Ge, negative-bias temperature instability (NBTI), pMOSFET, reliability, random telegraph noise (RTN), SiGe, time-dependent dielectric breakdown (TDDB), time-dependent variability, $1/f$ noise.**

I. Introduction

In PART I, we have shown that SiGe channel pMOSFETs can solve the negative-bias temperature instability (NBTI) issue for ultrathin equivalent oxide thickness (EOT) devices [1]. A crucial effect of the Si passivation layer (cap) thickness was found, with thinner Si caps yielding a significant reliability improvement. Contrary to Si channel pMOSFETs, ultrathin EOT (∼0.7 nm) SiGe channel devices optimized with a high Ge fraction (i.e., 55%), a thick quantum well (QW; i.e., ∼6–7 nm), and a reduced Si cap thickness (i.e., $<$ 1 nm) were demonstrated to qualify for ten years of continuous reliable operation at nominal V_{DD}. We ascribed this superior NBTI robustness mainly to the reduced interaction between channel holes and preexisting dielectric defects due to a favorable channel Fermi-level alignment shift. A simple mathematical implementation of this energy-based model was shown to excellently reproduce all the experimental observations there made [1].

However, those studies were performed on large-area ($10 \times 1\ \mu\text{m}^2$) test devices, as customary for standard NBTI testing. On such large-area devices, the random properties of many defects in the gate oxide (e.g., for a defect density of $\sim 10^{11}$ cm^{-2}, one device contains $\sim 10^4$ defects) average out, yielding the same well-defined BTI degradation curve on each device [2], [3].

Conversely, recent works have shown that, as the device geometries scale toward atomistic dimensions, the number of dopant atoms, as well as the number of dielectric defects in each transistor, is reduced to numerable levels (e.g., for a defect density of $\sim 10^{11}$ cm^{-2}, a 90×35 nm^2 device would include an average of *only three defects per device*). As an implication, *both* the fresh device parameters and the parameter *shifts* during operation become statistically distributed [2]–[5]. In other words, both a time-zero (i.e., as-fabricated) variability and a considerable time-dependent variability (i.e., reduced reliability) arise.

As a consequence, the deterministic "*average*" lifetime which is normally assessed on large area devices should be replaced by lifetime distributions [3], [6]. This *time-dependent variability* (i.e., nanoscaled NBTI reliability) can be studied in terms of *statistical measurements of the charging* and *discharging of individual defects* [7]–[9]. We and others have recently shown that the properties of individual charged gate oxide defects can be directly observed and measured by looking at the individual discharge events visible in NBTI-like $\Delta V_{\rm th}$ transients recorded on nanoscaled devices [3], [5], [10]. This approach recently led to the introduction of a novel defect characterization technique, which is the time-dependent defect spectroscopy [11].

Manuscript received July 16, 2012; revised October 2, 2012; accepted October 4, 2012. This work was supported by the European Commission under the Seventh Framework Programme (Collaborative project MORDRED, Contract 261868). The review of this paper was arranged by Editor J. S. Suehle.

J. Franco and G. Groeseneken are with imec, 3001 Leuven, Belgium, and also with the ESAT Department, Katholieke Universiteit Leuven, 3000 Leuven, Belgium (e-mail: Jacopo.Franco@imec.be).

B. Kaczer, M. Toledano-Luque, P. J. Roussel, T. Kauerauf, J. Mitard, and L. Witters are with imec, 3001 Leuven, Belgium.

T. Grasser is with the Technische Universität Wien, 1090 Vienna, Austria.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2012.2225624

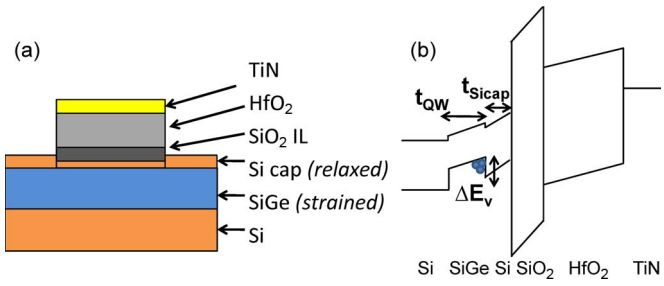

Fig. 1. (a) Sketch of the gate stack of SiGe devices used in this paper. (b) Band diagram in inversion: channel holes are confined into the SiGe QW due to the valence-band offset ΔE_v between the SiGe channel and the Si cap.

For these reasons, in this paper, we focus on the NBTI of nanoscaled $Si_{0.45}Ge_{0.55}$ pFETs and compare with the results obtained on large-area devices. In particular, we assess the impact of individually trapped charges on the characteristics of SiGe devices with varying Si cap thicknesses, since this gate-stack parameter was found to be crucial for the standard NBTI reliability on large-area devices [1]. The use of a SiGe channel is shown to offer a considerable reliability improvement also for deeply scaled devices, which is expected to significantly alleviate the time-dependent variability issue. This aspect can be understood in the framework of our previously proposed model of reduced interaction between channel holes and preexisting dielectric defects [1].

Finally, other reliability mechanisms including low-frequency $(1/f)$ noise, body biasing during NBTI, channel hot carriers (CHCs), and time-dependent dielectric breakdown (TDDB) are also investigated and are observed not to constitute showstoppers for the SiGe technology.

The broad set of experimental results reported here confirms once again the SiGe channel pMOS technology as a leading candidate for further scaled CMOS technology nodes [12].

II. Experiments

$Si_{0.45}Ge_{0.55}$ pFETs with metallurgic length $L_{\rm eff} \approx 35$ nm and width $W = 90$ nm were used to study the NBTI reliability of deeply scaled devices. The device gate stack consisted of a thin epitaxially grown SiGe channel, a Si cap with three different thicknesses in the range of 0.65–2 nm, a SiO_2 interfacial layer (∼0.8 nm), an HfO_2 dielectric (∼1.8 nm), and a TiN metal gate (see Fig. 1). More information about the process can be found in [12] and [13]. It is important to note that, since the device-to-device variability is known to depend on the doping level in the channel [4], [14], a beneficial effect is expected for SiGe devices owing to their undoped epitaxially grown channel. Therefore, for correct benchmarking, a ∼8-nm-thick undoped Si channel layer was epitaxially grown as a part of the Si reference gate stack [15].

The properties of individual defects, which ultimately control the time-dependent variability of a technology, were studied by means of NBTI-like measurement. After a defect charging phase (i.e., a NBTI-like stress), the channel current relaxation transients (ΔI_D) were recorded at a sense gate voltage $(V_{\rm Gsense})$ equal to the threshold voltage of the fresh device $(V_{\rm th0})$. The recorded ΔI_D transients were converted to $\Delta V_{\rm th}$ relaxation transients using the initial I_D–V_G measurement of the fresh device as reference [16]. The experimental temperature was 25 °C. A representative set of typical $\Delta V_{\rm th}$ relaxation transients recorded on nanoscaled SiGe devices is shown in Fig. 2(a).

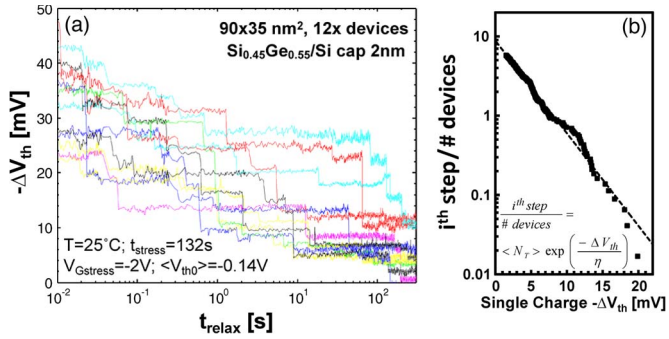

Fig. 2. (a) NBTI relaxation transients recorded on nanoscaled SiGe devices. For each device, multiple single-defect discharge events are visible. (b) Weighted CCDF plot of the individual $\Delta V_{\rm th}$ step heights observed on multiple (41) devices. Despite the intentionally undoped epitaxially grown SiGe channel, the $\Delta V_{\rm th}$ step heights appear to be exponentially distributed [4], with an average value $\eta \approx 3.9$ mV. The average number of defects per device $\langle N_T \rangle$ can be easily read in this plot as the intersection of the distribution with the y-axis [6].

III. Results

A. Individual Defect Discharge Events

As previously reported for Si devices [2], [5], the total $\Delta V_{\rm th}$ value observed after the same NBTI-like stress strongly varies from device to device [see Fig. 2(a)]. For each device, a different number of individual discharge events (i.e., a different number of previously charged defects) are visible. Each individual discharge causes a well-defined $\Delta V_{\rm th}$ step. The $\Delta V_{\rm th}$ step heights collected together into a cumulative distribution plot appear exponentially distributed [see Fig. 2(b); PDF: $f(\Delta V_{\rm th}, \eta) = \exp(-\Delta V_{\rm th}/\eta)/\eta$], with the average $\Delta V_{\rm th}$ step height η.

Individual discharge events causing the $\Delta V_{\rm th}$ step as large as ∼20 mV are easily observed in the $\Delta V_{\rm th}$ relaxation transients [probability of ∼1 in ∼240 observed defects; see Fig. 2(b)]. This value is about approximately nine times larger than expected from the charge-sheet approximation $\eta_0 = {\rm q}/{\rm C_{ox}} = q^* t_{\rm ox}/\varepsilon_0 \varepsilon_r W L_{\rm eff}$, with $t_{\rm ox}$ being the capacitance equivalent thickness of the gate stack (i.e., assuming the charged defect to be located at the channel interface) [17]. For comparison, it is worth to recall that the typical BTI failure criteria considered for process qualification range between 30 and 50 mV of the $V_{\rm th}$ shift. In other words, two such charged defects might already jeopardize the device reliability. Such anomalously large $V_{\rm th}$ shifts caused by single charged defects are ascribed to the intrinsic percolative nature of the current flow in nanoscaled devices, which is due to the nonuniform channel potential caused by random dopant fluctuations and other nanoscale issues as line edge roughness, metal gate granularity, etc. [4]. In the unlucky case of a gate oxide defect spatially located close to (i.e., on top) the critical point of a channel current percolation path, a single charge/discharge event can result in a significant change in the device current (i.e., observed in our experiment as a large $\Delta V_{\rm th}$ step).

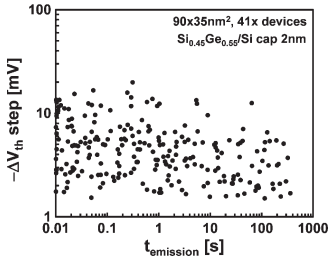

Fig. 3. Characteristic trap emission times and the corresponding $\Delta V_{\rm th}$ value caused by a single charged defect are clearly uncorrelated, just as recently observed for RTN defects [2], [17], [19]. (Note: The weak accumulation of several discharge events at 10 ms is an artifact related to the measuring speed of the used instrument).

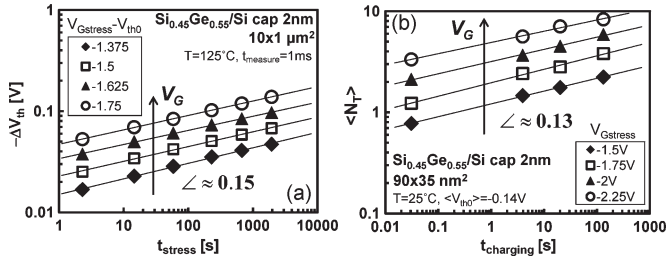

Fig. 4. (a) Large-area devices. The NBTI $\Delta V_{\rm th}$ follows a power law with the stress time and the stress gate voltage. (b) Nanoscaled devices. The average number of discharge events observed on each device (i.e., the average number of active traps per device $\langle N_T \rangle$) follows similar dependences.

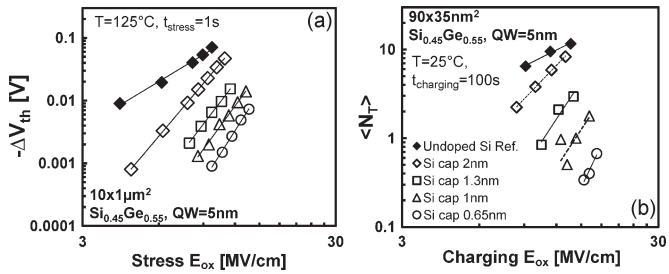

Fig. 5. (a) Large-area devices. SiGe channel pMOSFETs with a reduced Si cap thickness show reduced NBTI $\Delta V_{\rm th}$ and a stronger field acceleration leading to further improvement at lower operating fields. (b) Nanoscaled devices. The same experimental observations apply to the average number of charging/discharging defects per device ($\langle N_T \rangle$). Note: Very high equivalent oxide fields were needed for the preliminary charging phase in order to be able to observe active defects in SiGe devices with the thinnest Si cap ($\langle N_T \rangle$ as low as ∼0.33 at 15 MV/cm, i.e., one defect observed for every three measured devices).

The time-dependent device-to-device variability has been described by means of the average number of active (i.e., charging/discharging) oxide defects, $\langle N_T \rangle$, and the average $\Delta V_{\rm th}$ impact per defect, η, as previously discussed [6]. In order to estimate $\langle N_T \rangle$ and η for each gate stack studied here, the same sequence including a charging phase and a relaxation transient was repeated on a large set of nominally identical devices. The device set size was chosen to be large enough to capture the signatures of some hundred active defects for each gate stack (i.e., typically a few tens of devices but up to 160 devices for SiGe devices with a reduced Si cap thickness). The $\Delta V_{\rm th}$ steps observed in the relaxation traces were then collected into weighted complementary cumulative distribution function [CCDF; see Fig. 2(b)] plots, and a maximum-likelihood fit was performed in order to estimate the exponential distribution parameters η and $\langle N_T \rangle$.

This experimental technique is, to a certain extent, equivalent to standard random telegraph noise (RTN) studies. For RTN measurements normally performed at low gate bias ($V_G \approx V_{\rm th0}$), a large number of nanoscaled devices (up to thousands) are typically needed in order to capture the signature of a statistically significant number of oxide defects exchanging charges with the transistor channel [18]. Conversely, the experimental technique used here allows maximizing the number of observed defects per device, due to the preliminary charging ("perturbation") phase during which oxide defects with deeper energy levels become accessible due to a high oxide electric field. As a consequence, a reduced device sample size is sufficient to collect a statistically significant data set. In other words, the relaxation transients studied here can be interpreted as a nonsteady-state case of RTN, as discussed in [17], [19], and [20]. Furthermore, within this analogy, recent works have shown that the characteristic charging/discharging time of defects causing RTN are uncorrelated with the respective impact of each charged defect on the device characteristics [2], [17], [19]. This observation is confirmed by a typical set of $\Delta V_{\rm th}$ step heights observed with our experimental technique on SiGe devices with a 2-nm-thick Si cap plotted against their respective discharge time (see Fig. 3).

B. Average Number of Active Defects per Device ($\langle N_T \rangle$)

The $\Delta V_{\rm th}$ value induced by NBTI in large-area devices is known to approximately follow a power law with the stress time and with the stress gate overdrive voltage [see Fig. 4(a)]. Equivalently, for nanoscaled devices, the average number of discharge events observed on each device after the charging phase is observed to follow similar power-law dependences [see Fig. 4(b) [21]].

As discussed in Part I [1], for large-area devices, we found that a reduced Si cap thickness on SiGe pFETs results in a significantly reduced $\Delta V_{\rm th}$ at fixed stress conditions [see Fig. 5(a)]. In a similar way, a thinner Si cap yields a strong reduction of the average number of $\Delta V_{\rm th}$ steps (i.e., average number of active defects $\langle N_T \rangle$) in nanoscaled devices [see Fig. 5(b)]. It is worth noting that very low $\langle N_T \rangle$ values (as low as ∼0.33 for the short precharging time used here, i.e., only one defect observed per three measured devices) are observed for SiGe channel devices with the thinnest Si cap; such low $\langle N_T \rangle$ values complicate the experiment, requiring larger sample set (up to ∼160 devices were used for this particular gate-stack) to observe a sufficient number of charging/discharging events. However, as discussed later, such low $\langle N_T \rangle$ values might still jeopardize the reliability of a fraction of the device in a realistic device population (billions of devices).

C. Average $\Delta V_{\rm th}$ Impact per Charged Defect (η)

As aforementioned, the other parameter controlling the time-dependent variability of a given technology is η, i.e., the average $\Delta V_{\rm th}$ impact per individual charged oxide defect. When looking at the weighted CCDF plots of the $\Delta V_{\rm th}$ step heights observed on SiGe devices with two different Si cap thicknesses, a reduced η value is found for the thinnest Si cap (see Fig. 6, $\eta \approx 3.9$ mV for a 2-nm Si cap and $\eta \approx 1.8$ mV for a 0.65-nm Si cap). In order to benchmark more correctly the η values

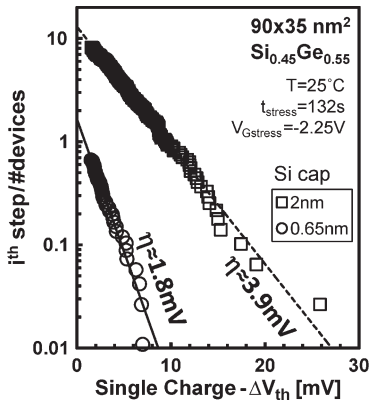

Fig. 6. Weighted CCDF plot of the $\Delta V_{\rm th}$ step heights observed on SiGe devices with two different Si cap thicknesses. The average $\Delta V_{\rm th}$ step height η is significantly reduced for the devices with the thinnest Si cap. Note also the reduced $\langle N_T \rangle$ (lower y-axis intercept).

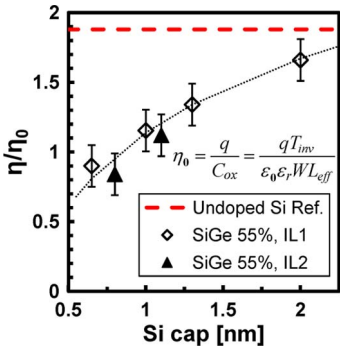

Fig. 7. Extracted average $\Delta V_{\rm th}$ step heights η for SiGe devices with different Si cap and for undoped Si channel devices, after a precharging phase at $E_{\rm ox} \approx 12$ MV/cm. SiGe devices with the thinnest Si cap show a significantly lower η ($\sim 2\times$). The observation is confirmed on SiGe devices with two different SiO_2 interfacial layer thicknesses. (Red dashed line) Benchmark value experimentally estimated on undoped Si channel ref. devices. The error bars on the estimated η values are related to the lower $\langle N_T \rangle$ observed for SiGe.

estimated on different gate stacks, a normalization of η for the expected $\Delta V_{\rm th}$ value per single charge calculated according to the charge sheet electrostatic approximation (see Section III-A) is proposed. The normalized η/η_0 values for SiGe devices with different Si cap thicknesses and with two different thicknesses of the SiO_2 interfacial layer are shown in Fig. 7 and benchmarked against a reference value measured on undoped Si channel devices. As discussed next, the observed reduction of η, combined with the reduced $\langle N_T \rangle$ value, promises *a significant improvement of the lifetime distribution of a realistic device population.*

Interestingly, also the emission time distributions (see Fig. 8) appear to be shifted toward lower values for thinner Si cap, suggesting a faster oxide defect discharge at low gate bias for these gate stacks, i.e., a faster NBTI relaxation. We already made a similar observation for large-area devices in [1]; as discussed below, this behavior is predicted by the energy-driven model that we propose.

The previously discussed experimental observations on nanoscaled SiGe devices with reduced Si cap thickness can be summarized as:

1) Reduced average number of active oxide defects $\langle N_T \rangle$;
2) Stronger stress electric-field dependence of $\langle N_T \rangle$;

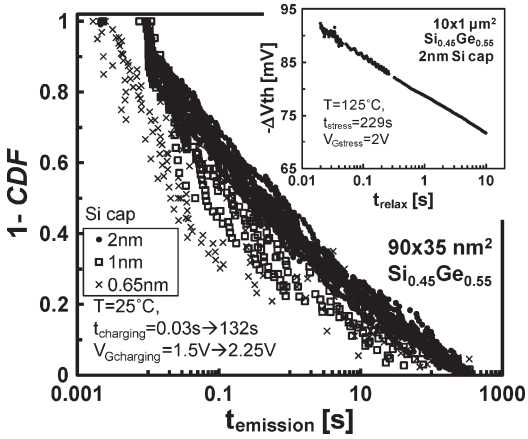

Fig. 8. One-CDF plot of all the charge emission times observed on multiple devices/traps after preliminary charging phases of different durations and at several voltages, on SiGe gate stacks varying Si cap thickness. The apparent distributions are reminiscent of the NBTI relaxation traces typically observed on large-area devices (inset: $\Delta V_{\rm th}$ relaxation trace for a 10×1 μm^2 SiGe device with a 2-nm Si cap). The emission time distributions appear shifted toward lower values for reducing Si cap thicknesses, suggesting a faster discharge of the trapped charges after stress removal, as previously observed for large-area devices in [1]. NB: A detailed study of a *single* trap emission times distribution can be found in [10].

3) Reduced average $\Delta V_{\rm th}$ impact per charged defect η;
4) Reduced average trapped charge emission time.

IV. Discussion: Consequences on the Time-Dependent Variability

As we have proposed in [6], the fraction of a realistic population (i.e., billions of devices) expected to be still functional after ten years of continuous operation can be estimated from the $\langle N_T \rangle$ and η values extracted on individual devices with the technique used in this paper. The calculation is based on the convolution of a Poisson distributed number of defects with the mean value $\langle N_T \rangle$, with an exponential distribution of impact per single charged defects on the device threshold voltage with a mean value η. The mathematical details can be found in [2] and [6].

Although $\langle N_T \rangle$ can be evaluated from relaxation transients recorded on individual nanoscaled devices as previously shown [see Fig. 2(b)], an easier and more robust estimation of $\langle N_T \rangle$ can be obtained from the NBTI-induced $\Delta V_{\rm th}$ data measured on large-area devices [see Fig. 5(a)], which include the contribution of typically thousands of defects per device [22]. Moreover, the dependences on the stress time and voltage, and the impact of a higher operating temperature (125 °C) are readily captured using the broad experimental data sets that we have previously presented in Part I [1]. For this purpose, the large-area device $V_{\rm th}$ shifts can be described with a simple power-law model, i.e.,

$$\Delta V_{\rm th} = A_{\Delta V_{\rm th}} (V_G - V_{\rm th0})^{\gamma} t_{\rm stress}^{n} \quad [\rm V] \tag{1}$$

where prefactor $A_{\Delta \rm Vth}$ is the $\Delta V_{\rm th}$ value that one would measure after 1-s stress at a gate overdrive of 1 V on large-area devices. This prefactor $A_{\Delta \rm Vth}$ can be easily converted into trapped charge density as

$$A_{\Delta N_{\rm ot}} = \frac{C_{\rm ox}}{q} A_{\Delta V_{\rm th}} = \frac{\varepsilon_0 \varepsilon_r}{t_{\rm ox}} \frac{A_{\Delta V_{\rm th}}}{q} \quad [\rm cm^{-2}]. \tag{2}$$

TABLE I
EXTRACTED PARAMETERS

Gate stack	N_0	γ	n	η/η_0
Si Ref.	8.64	3.03	0.13	1.88
SiGe – Si cap 2nm	1.45	4.78	0.15	1.66
SiGe – Si cap 1.3nm	0.60	5.28	0.17	1.34
SiGe – Si cap 1nm	0.33	5.52	0.18	1.15
SiGe – Si cap 0.65nm	0.11	6.36	0.19	0.9

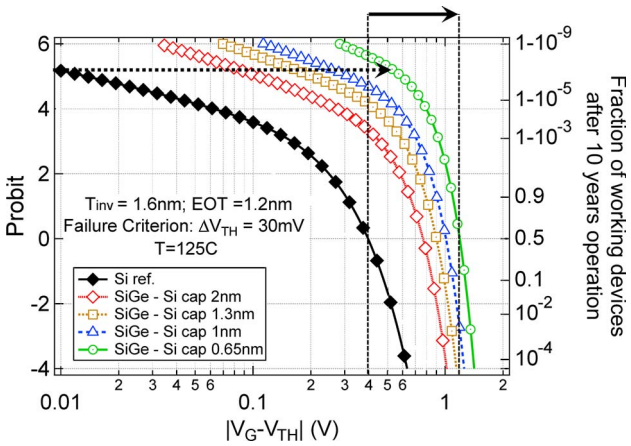

Fig. 9. Calculated fractions of working devices after ten year continuous operation at varying operating voltages for the different gate stacks studied here. A dramatic improvement of the distributions is apparent for SiGe devices with reduced Si cap thickness. Note: Large-area device lifetime would appear in this plot as a vertical dashed line (whole population fails above maximum allowed operating voltage, whereas it passes for lower voltages) with same median value (Probit $= 0$) of the respective nanoscaled device distribution. The reliability improvement previously observed in large area SiGe devices (demarcated by the solid arrow, distance at Probit $= 0$) is expected to be magnified at high percentiles (demarcated by the dotted arrow, at $\sim$1 ppb).

Next, assuming no additional oxide defectivity in deeply scaled devices, the power-law prefactor $A_{\Delta \mathrm{Not}}$ can be simply rescaled by the device area, i.e.,

$$N_{\mathbf{0}} = A_{\Delta N_{\mathrm{ot}}} W L_{\mathrm{eff}}. \tag{3}$$

Finally, $\langle N_T \rangle$ for the nanoscaled devices ($W = 90$ nm and $L_{\mathrm{eff}} = 35$ nm in our case) can be expressed as a function of the operating voltage and the stress time as

$$\langle N_T \rangle = N_{\mathbf{0}} (V_G - V_{\mathrm{th}\mathbf{0}})^{\gamma} t_{\mathrm{stress}}^{n}. \tag{4}$$

Parameters N_0, γ, and n extracted for the different gate stacks considered here are reported in Table I, together with the η values estimated on nanoscaled devices (see Fig. 7). A significantly reduced N_0, a significantly higher γ, and a slightly higher n are found for SiGe devices with reduced Si cap thickness w.r.t. Si ref. devices, in agreement with the experimental observations previously made in Part I [1].

Equation (4) was used to project to ten years of continuous operation and calculate the fraction of devices still working as a function of the operating gate overdrive voltage, as shown in Fig. 9. A dramatic improvement of the distribution for optimized SiGe devices is apparent, particularly at the high percentiles (e.g., $\sim$ one failure per billion devices).

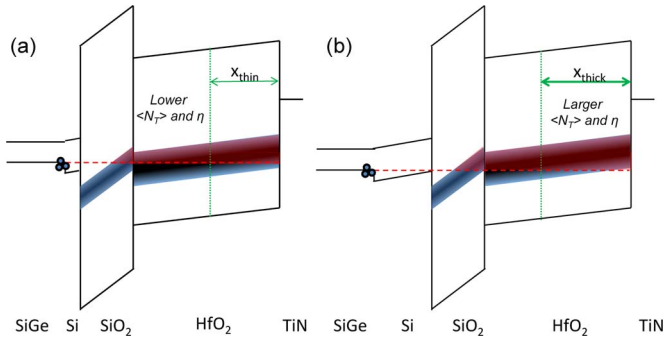

Fig. 10. Sketches of the SiGe gate stacks in the stress condition, showing the defect bands in SiO_2 and HfO_2 as from the model proposed in Part I for large-area devices. Due to the higher Fermi level in the SiGe channel, fewer defects are energetically favorable for channel holes. Moreover, the accessible defects are preferentially spatially located further from the channel, toward the gate, yielding a reduction of η. (a) Favorable alignment shift is maximized for the thinnest Si cap thickness, while (b) the additional voltage drop on a thicker Si cap (at fixed field) makes more defects on the channel side accessible, displacing the equivalent trapped charge centroid ($x_{\mathrm{thick}} > x_{\mathrm{thin}}$). This observation readily explains the larger $\langle N_T \rangle$ and η values observed for SiGe devices with thick Si caps.

Finally, we note that, due to the reduced $\langle N_T \rangle$ and η values, we expect both a reduced probability of RTN occurrence and a reduced RTN impact (amplitudes) in the optimized SiGe devices.

V. MODEL

In Part I, we proposed a model for the superior reliability of SiGe channel devices based on a favorable energy decoupling between channel holes and preexisting dielectric defects. We represented the preexisting dielectric defect energy levels as two defect bands located in the SiO_2 IL and in the HfO_2 band gaps. As depicted in Fig. 10, the Fermi level in the SiGe channel determines which part of the defect band is energetically favorable for trapping channel holes. For a thin Si cap, only a smaller part of the defect bands is accessible, while the larger voltage drop on a thicker Si cap (at fixed equivalent E_{ox}) lowers the Fermi level, thus making more defects accessible for channel holes. This interpretation, which we showed to be able to quantitatively match the NBTI data on large-area devices (cf. [1, Fig. 17]), also readily explains the experimental observation of a reduced $\langle N_T \rangle$ value in nanoscaled SiGe devices with thinner Si caps [see Fig. 5(b)].

Moreover, for thinner Si caps, the accessible defects are located on the gate side of the dielectrics, i.e., the equivalent trapped charge centroid is closer to the gate. Assuming that each ΔV_{th} step height results from the electrostatic effect of the charged defect convoluted with the nonuniform potential profile in the channel, the average electrostatic contribution might be reduced for the SiGe gate stack with a thin Si cap due to the preferential location close to the gate. This observation explains the observed reduced η for SiGe devices (see Fig. 7) [21]. Finally, the model can also explain the lower average charge emission time (i.e., faster NBTI relaxation) observed for SiGe devices with thin Si cap (see Fig. 8), since the higher energy difference between the charged defect levels and the SiGe channel is expected to favor the trapped charge emission once the stress condition is removed [17].

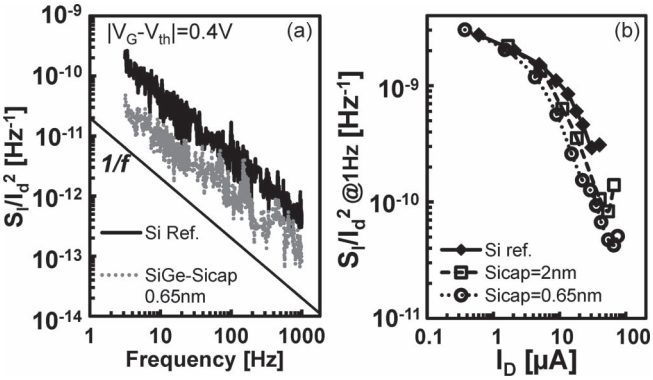

Fig. 11. (a) SiGe devices with a reduced Si cap thickness for optimized NBTI reliability also show reduced $1/f$ noise w.r.t. a Si ref. device with identical gate-stack (b) and also w.r.t. a SiGe device with thicker Si cap.

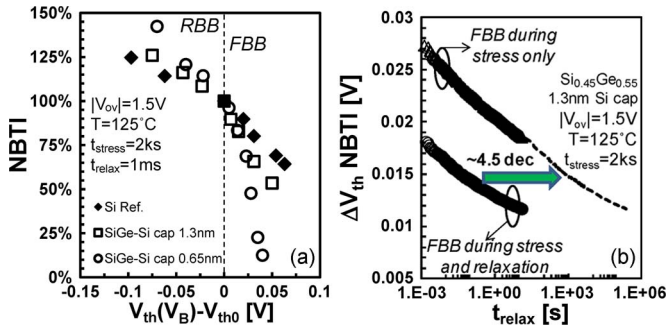

Fig. 12. (a) Body bias is found to modulate the oxide electric field at a given overdrive gate voltage and, therefore, the NBTI degradation. Due to the stronger electric-field dependence of NBTI observed in SiGe pMOSFETs with reduced Si cap thickness (see Fig. 5), more reduced NBTI with FBB is observed. (b) Significantly faster NBTI relaxation is observed when an FBB is applied during an entire stress/relaxation sequence.

VI. Other Reliability Mechanisms

In this section, the impact of other reliability mechanisms on the SiGe channel technology, namely, low-frequency $(1/f)$ noise, body biasing during NBTI, CHC, and TDDB, is briefly discussed.

A. Low-Frequency Noise

Similarly to BTI, $1/f$ noise has been recently ascribed to trapping and detrapping of channel carriers into oxide defects with widely distributed characteristic time constants [23]. The reduced interaction between carriers and oxide defects observed for SiGe devices due to the energy decoupling is therefore expected to also yield a reduced $1/f$ noise [24]. This observation is clearly confirmed when looking at the low-frequency noise spectra measured on a Si ref. device and on a SiGe device [see Fig. 11(a)]. Moreover, the noise reduction is observed to depend on the Si cap thickness [see Fig. 11(b)], in agreement with the BTI trends previously discussed.

B. Body Biasing During NBTI

The interplay of body bias and NBTI has received some research interest since it provides an alternative knob to independently control the inversion hole carrier concentration and the oxide electric field during device stress [25]. In particular, modulating the channel depletion width, the body bias can be used to modulate the gate overdrive and, therefore, the inversion charge at fixed gate voltage and fixed oxide field or to alternatively modulate the oxide field at fixed overdrive voltage and fixed inversion hole population [26].

The latter condition is of particular interest since it can be used to reduce NBTI without reducing the device performance. The stronger NBTI dependence on the oxide electric field observed in SiGe devices (see Fig. 5) yields further benefit when a NBTI stress condition is combined with a forward body bias (FBB). As shown in Fig. 12(a), an FBB reduces the NBTI degradation in large-area SiGe devices more significantly than on Si ref. devices. Furthermore, the FBB is observed to further speed up NBTI relaxation at low gate biases. This effect can be ascribed to enhanced hole detrapping at reduced relaxation oxide field due to the additional energy displacement of the defect levels w.r.t. the channel Fermi level. As one can observe in Fig. 12(b), this effect is substantial for SiGe devices (∼4.5-dec faster relaxation) [26].

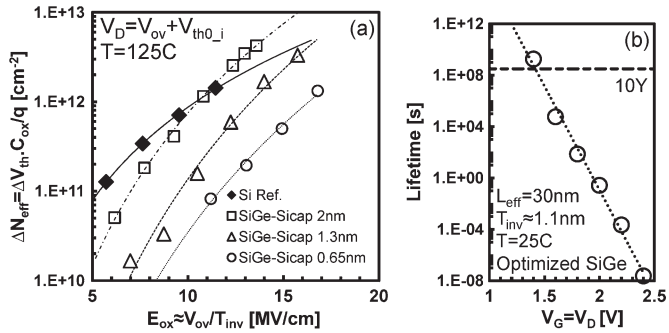

Fig. 13. (a) SiGe devices with a thick Si cap show lower CHC degradation w.r.t. Si ref. only at low stress conditions, while at high stress conditions, a higher degradation is apparent probably due to enhanced impact ionization in the small band-gap channel. However, the optimized SiGe gate stacks with a reduced Si cap thickness consistently show an overall reduced CHC degradation w.r.t. Si ref. (b) Extrapolated device lifetime under CHC stress for a ∼0.7-nm EOT ($T_{\rm inv} \approx 1.1$ nm) NBTI-optimized SiGe gate stack. CHCs do not constitute a showstopper for SiGe devices.

C. CHC

The use of a small band-gap semiconductor favors electron–hole pair generation in the channel by impact ionization. This effect is expected to enhance hot carrier degradation. As a consequence, poor hot carrier robustness has been reported for pure Ge-channel devices [27], [28]. However, during a typical CHC stress ($V_G = V_D = V_{\rm stress}$), a significant fraction of the total degradation in pMOSFETs is related to the residual NBTI effect at the source side of the channel [29]. The enhanced NBTI robustness of the optimized SiGe devices (i.e., the reduced charge trapping due to the favorable energy alignment shift of the channel Fermi level w.r.t. the dielectric defect levels) also significantly reduces the total degradation caused by the CHC stress [see Fig. 13(a)]. As we reported in [24], CHCs do not constitute a showstopper for the optimized ultrathin EOT SiGe devices [see Fig. 13(b)].

D. TDDB

TDDB is not considered a showstopper for Si channel pMOSFETs with EOT down to ∼0.7 nm, although a few soft breakdown events (i.e., nonfatal dielectric *wearout*) are

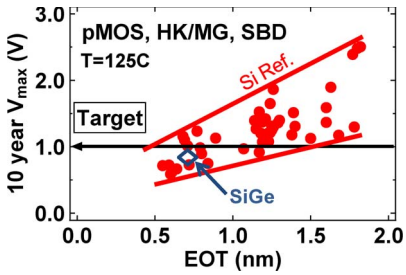

Fig. 14. Extracted maximum operating gate voltage using a soft breakdown (i.e., nondestructive) failure criterion for a ∼0.7-nm EOT NBTI-optimized SiGe gate stack shows no difference w.r.t. Si channel device benchmark data [30].

expected in ten years of continuous operation at nominal V_{DD} values [30]. As already noted for pure Ge channel devices in [31], no significant difference in the TDDB characteristic of SiGe devices w.r.t. their Si counterparts is observed (see Fig. 14). As we calculated in [24], for our ∼0.7-nm EOT optimized SiGe gate stack, less than ten soft breakdowns are expected on a 0.1-cm^2 chip operated for ten years at $V_{DD} =$ 1 V, in line with Si reference devices [30].

VII. Conclusion

The time-dependent variability of nanoscaled $Si_{0.45}Ge_{0.55}$ pFETs has been studied as a function of the Si cap thickness and compared with Si ref. devices. As in the Si counterparts, individual discharge events were visible in the NBTI-like $\Delta V_{\rm th}$ relaxation traces, with exponentially distributed step heights. The average number of discharge events followed the typical NBTI dependences on stress time and voltage observed for the total $\Delta V_{\rm th}$ on large-area devices. The use of a thinner Si cap on SiGe has been found to yield a significant reduction of the average number of active oxide defects $\langle N_T \rangle$ causing charge/discharge events and of the average $\Delta V_{\rm th}$ step height (η) per charged defect. These results confirm this technology to be extremely promising also for the reliability of nanoscaled devices.

Our previously proposed model based on energy decoupling between channel holes and preexisting dielectric defect energy levels [1] has readily explained the experimental observations in this paper as well, suggesting that fewer defects preferentially located further from the channel are accessible by channel holes for optimized SiGe devices.

Owing to this effect, a reduced low-frequency noise was also observed in SiGe devices. Moreover, an FBB during NBTI, reducing the oxide electric field at fixed inversion hole population, has been shown to yield a significant additional reliability improvement on SiGe devices due to the stronger NBTI field dependence observed for this technology. Other reliability mechanisms, such as CHCs and TDDB, have been shown not to be showstoppers.

The extensive experimental results reported here and in Part I strongly support SiGe technology as a leading candidate for future CMOS technology nodes, offering a complete solution to the reliability issue for ultrathin EOT nanoscaled pMOSFET devices.

References

[1] J. Franco, B. Kaczer, P. J. Roussel, J. Mitard, M. Cho, L. Witters, T. Grasser, and G. Groeseneken, "SiGe channel technology: Superior reliability toward ultra-thin EOT devices. Part I: NBTI," *IEEE Trans. Electron Devices*.

[2] B. Kaczer, T. Grasser, P. J. Roussel, J. Franco, R. Degraeve, L. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger, "Origin of NBTI variability in deeply scaled pFETs," in *Proc. IEEE IRPS*, 2010, pp. 26–32.

[3] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, P. Roussel, and M. Nelhiebel, "Recent advances in understanding the bias temperature instability," in *Proc. IEEE IEDM*, 2010, pp. 82–85.

[4] A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies, "RTS amplitude in decananometer MOSFETs: 3-D simulation study," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 839–845, Mar. 2003.

[5] V. Huard, C. Parthasarathy, C. Guerin, T. Valentin, E. Pion, M. Mammasse, N. Planes, and L. Camus, "NBTI degradation: From transistor to SRAM arrays," in *Proc. IEEE IRPS*, 2008, pp. 289–300.

[6] M. Toledano-Luque, B. Kaczer, J. Franco, P. J. Roussel, T. Grasser, T. Y. Hoffmann, and G. Groeseneken, "From mean values to distributions of BTI lifetime of deeply scaled FETs through atomistic understanding of the degradation," in *Proc. VLSI Symp. Technol.*, 2011, pp. 152–153.

[7] V. Huard, F. Cacho, Y. Mamy Randriamihaja, and A. Bravaix, "From defects creation to circuit reliability—A bottom–up approach," *Microelectron. Eng.*, vol. 88, no. 7, pp. 1396–1407, Jul. 2011.

[8] M. Nafria, R. Rodriguez, M. Porti, J. Martin-Martinez, M. Lanza, and X. Aymerich, "Time-dependent Variability of high-k based MOS devices: nanoscale characterization and inclusion in circuit simulators," in *Proc. IEEE IEDM*, 2011, pp. 127–130.

[9] B. Kaczer, S. Mahato, V. V. de Almeida Camargo, M. Toledano-Luque, P. J. Roussel, T. Grasser, F. Catthoor, P. Dobrovolny, P. Zuber, G. Wirth, and G. Groeseneken, "Atomistic approach to variability of bias-temperature instability in circuit simulations," in *Proc. IEEE IRPS*, 2011, pp. 915–919.

[10] M. Toledano-Luque, B. Kaczer, P. J. Roussel, T. Grasser, G. I. Wirth, J. Franco, C. Vrancken, N. Horiguchi, and G. Groeseneken, "Response of a single trap to ac negative bias temperature stress," in *Proc. IEEE IRPS*, 2011, pp. 364–371.

[11] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time dependent defect spectroscopy for characterization of border traps in metal–oxide–semiconductor transistors," *Phys. Rev. B, Condens. Matter*, vol. 82, no. 24, p. 245 318, Dec. 2010.

[12] J. Mitard, L. Witters, P. Garcia Bardon, P. Cristie, J. Franco, A. Mercha, P. Magnone, M. Alioto, F. Crupi, L. A. Ragnarsson, A. Hikavyy, B. Vincent, T. Chiarella, R. Loo, J. Tseng, S. Yamaguchi, S. Takeoka, W.-E. Wang, P. Absil, and T. Hoffmann, "Sub-nm EOT SiGe-55% pFETs for high-speed low-V_{DD} technology: A study from capacitor to circuit level," in *Proc. IEDM*, 2010, pp. 249–252.

[13] A. Hikavyy, R. Loo, L. Witters, S. Takeoka, J. Geypen, B. Brijs, C. Merckling, M. Caymax, and J. Dekoster, "SiGe SEG growth for buried channel p-MOS devices," *ECS Trans.*, vol. 25, no. 7, pp. 201–210, Sep. 2009.

[14] A. Ghetti, C. M. Compagnoni, A. S. Spinelli, and A. Visconti, "Comprehensive analysis of random telegraph noise instability and its scaling in deca-nanometer Flash memories," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1746–1752, Aug. 2009.

[15] G. Hellings, L. Witters, R. Krom, J. Mitard, A. Hikavyy, R. Loo, A. Schulze, G. Eneman, C. Kerner, J. Franco, T. Chiarella, S. Takeoka, J. Tseng, W.-E. Wang, W. Vandervorst, P. Absil, S. Biesemans, M. Heyns, K. De Meyer, M. Meuris, and T. Hoffmann, "Implant-free SiGe quantum well pFET: A novel, highly scalable and low thermal budget device, featuring raised source/drain and high-mobility channel," in *Proc. IEEE IEDM*, 2010, pp. 241–244.

[16] B. Kaczer, T. Grasser, P. J. Roussel, J. Martin-Martinez, R. O'Connor, B. J. O'Sullivan, and G. Groeseneken, "Ubiquitous relaxation in BTI stressing—New evaluation and insights," in *Proc. IEEE IRPS*, 2008, pp. 20–27.

[17] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, "The paradigm shift in understanding the bias temperature instability: From reaction-diffusion to switching oxide traps," *IEEE Trans. Electron Devices*, vol. 58, no. 11, pp. 3652–3666, Nov. 2011.

[18] S. Realov and K. L. Shepard, "Random telegraph noise in 45-nm CMOS: Analysis using an on-chip test and measurement system," in *Proc. IEEE IEDM*, 2010, pp. 624–247.

[19] T. Grasser, H. Reisinger, W. Goes, T. Aichinger, P. Hehenberger, P.-J. Wagner, M. Nelhiebel, J. Franco, and B. Kaczer, "Switching oxide traps as the missing link between negative bias temperature instability and random telegraph noise," in *Proc IEEE IEDM*, 2009, pp. 729–732.

[20] M. J. Kirton and M. J. Uren, "Noise in solid-state microstructures: A new perspective on individual defects, interface states, and low-frequency noise," *Adv. Phys.*, vol. 38, no. 4, pp. 367–468, Jan. 1989.
[21] J. Franco, B. Kaczer, M. Toledano-Luque, Ph. J. Roussel, P. Hehenberger, T. Grasser, J. Mitard, G. Eneman, L. Witters, T. Y. Hoffmann, and G. Groeseneken, "On the impact of the Si passivation layer thickness on the NBTI of nanoscaled $Si_{0.45}Ge_{0.55}$ pMOSFETs," *Microelectron. Eng.*, vol. 88, no. 7, pp. 1388–1391, Jul. 2011.
[22] M. Toledano-Luque, B. Kaczer, T. Grasser, P. J. Roussel, J. Franco, and G. Groeseneken, "Toward a streamlined projection of small device BTI lifetime distributions," in *Proc. WoDiM*, Dresden, Germany, 2012, to be published.
[23] B. Kaczer, T. Grasser, J. Martin-Martinez, E. Simoen, M. Aoulaiche, P. J. Roussel, and G. Groeseneken, "NBTI from the perspective of defect states with widely distributed time scales," in *Proc. IEEE IRPS*, 2009, pp. 55–60.
[24] J. Franco, B. Kaczer, G. Eneman, J. Mitard, A. Stesmans, V. Afanas'ev, T. Kauerauf, P. J. Roussel, M. Toledano-Luque, M. Cho, R. Degraeve, T. Grasser, L.-A. Ragnarsson, L. Witters, J. Tseng, S. Takeoka, W.-E. Wang, T. Y. Hoffmann, and G. Groeseneken, "6 Å EOT $Si_{0.45}Ge_{0.55}$ pMOSFET with optimized reliability ($V_{DD} = 1$ V): Meeting the NBTI lifetime target at ultra-thin EOT," in *Proc. IEEE IEDM*, 2010, pp. 70–73.
[25] Y. Mitani, H. Satake, and A. Toriumi, "Influence of nitrogen on negative bias temperature instability in ultrathin SiON," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 6–13, Mar. 2008.
[26] J. Franco, B. Kaczer, G. Eneman, P. J. Roussel, T. Grasser, J. Mitard, L.-A. Ragnarsson, M. Cho, L. Witters, T. Chiarella, M. Togo, W. Wang, A. Hikavyy, R. Loo, N. Horiguchi, and G. Groeseneken, "Superior NBTI reliability of SiGe channel pMOSFETs: Replacement gate, FinFETs, and impact of body bias," in *Proc. IEEE IEDM*, 2011, pp. 445–448.
[27] D. Maji, F. Crupi, E. Amat, E. Simoen, B. De Jaeger, D. P. Brunco, C. R. Manoj, V. R. Rao, P. Magnone, G. Giusi, C. Pace, L. Pantisano, J. Mitard, R. Rodriguez, and M. Nafria, "Understanding and optimization of hot-carrier reliability in germanium-on-silicon pMOSFETs," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1063–1069, May 2009.
[28] J. Franco, G. Eneman, B. Kaczer, J. Mitard, B. De Jaeger, and G. Groeseneken, "Impact of halo implant on the hot carrier reliability of germanium p-MOSFETs," *J. Vac. Sci. Technol. B, Microelectron. Process. Phenom.*, vol. 29, no. 1, pp. 01A804-1–01A804-4, Jan. 2011.
[29] J. Franco, B. Kaczer, G. Eneman, P. J. Roussel, M. Cho, J. Mitard, L. Witters, T. Y. Hoffmann, G. Groeseneken, F. Crupi, and T. Grasser, "On the recoverable and permanent components of hot carrier and NBTI in Si pMOSFETs and their implications in $Si_{0.45}Ge_{0.55}$-pMOSFETs," in *Proc. IEEE IRPS*, 2011, pp. 624–629.
[30] T. Kauerauf, R. Degraeve, L. Ragnarsson, P. Roussel, S. Sahhaf, G. Groeseneken, and R. O'Connor, "Methodologies for sub-1 nm EOT TDDB evaluation," in *Proc. IEEE IRPS*, 2011, pp. 7–16.
[31] B. Kaczer, B. De Jaeger, G. Nicholas, K. Martens, R. Degraeve, M. Houssa, G. Pourtois, F. Leys, M. Meuris, and G. Groeseneken, "Electrical and reliability characterization of metal-gate/HfO_2/Ge FETs with Si passivation," *Microelectron. Eng.*, vol. 84, no. 9/10, pp. 2067–2070, Sep./Oct. 2007.

Jacopo Franco (S'10) received the M.Sc. degree in electronic engineering from the Università della Calabria, Cosenza, Italy, in 2008. He is currently with the reliability group of Imec, Leuven, Belgium, working toward the Ph.D. degree at KU Leuven, Leuven.

Ben Kaczer received the Ph.D. degree in physics from The Ohio State University, Columbus, in 1998.

In 1998, he joined the reliability group of imec, Leuven, Belgium, where he is currently a Principal Scientist.

María Toledano-Luque received the Ph.D. degree in physics from the Universidad Complutense de Madrid, Madrid, Spain, in 2008.

Since 2011, she has been a Researcher at imec, Leuven, Belgium.

Philippe J. Roussel received the Diploma in electrical engineering from Industri Hogeschool of Ghent, Ghent, Belgium, in 1983.

In 1987, he joined imec, Leuven, Belgium, where he is currently working in a multidisciplinary team.

Thomas Kauerauf (M'07), received the Ph.D. degree from KU Leuven, Leuven, Belgium, in 2007.

Since 2006, he has been with imec, Leuven, where he currently works on the electrical characterization and reliability of high-k gate stacks.

Jérôme Mitard received the Ph.D. degree in microelectronic engineering from the Polytechnic University School of Marseille, Marseille, France, in 2003.

He is currently with imec, Leuven, Belgium, working on the integration of Ge and III–V channels.

Liesbeth Witters received the M.Sc. degree in chemical engineering from the Institut Français du Pétrole, Paris, France, in 1993.

Since 2001, she has been with imec, Leuven, Belgium, working on CMOS process development.

Tibor Grasser (SM'05) is currently a Professor at TU Wien, Vienna, Austria.

His research interests are centered around semiconductor device reliability issues, such as bias temperature instabilities and hot carrier degradadion.

Guido Groeseneken (F'05), received the Ph.D. degree in applied sciences from KU Leuven, Leuven, Belgium, in 1986.

Since 1987, he has been with imec, Leuven. Since 2001, he has been also a Professor at KU Leuven.