Analytic Modeling of the Bias Temperature Instability Using Capture/Emission Time Maps

T. Grasser[◇], P.-J. Wagner[◇], H. Reisinger[•], Th. Aichinger[*], G. Pobegen[†], M. Nelhiebel[‡], and B. Kaczer[○]

[◇]Institute for Microelectronics, TU Wien, Austria [•] Infineon, Munich, Germany
[*] Penn State University, USA [†] KAI, Villach, Austria [‡] Infineon, Villach, Austria [○] IMEC, Leuven, Belgium

Abstract

Despite a number of recent advances made in the understanding of the bias temperature instability (BTI), there is still no simple model available which can capture BTI degradation during DC or duty-factor (DF) dependent stress and the following recovery. By exploiting the intuitive features of the recently proposed capture/emission time (CET) maps [1, 2], we suggest an analytic model capable of handling a wide number of BTI stress and recovery patterns. As the model captures both the temperature- and bias- dependence of the degradation, it allows for realistic lifetime extrapolation. Compared to available models which do not consider the saturation of the degradation, our model predicts considerably more optimistic lifetimes.

Introduction

At the heart of the model stand the CET maps, which describe the wide distribution of capture and emission times. The CET maps have so far been extracted by numerically differentiating a set of ΔV_{th} recovery curves [1]. It has been shown that this approach can explain a wide class of both NBTI [2] as well as PBTI [3] stress and recovery patterns, including DC, AC, and DF stress. Although accurate, such a table-based model is valid for a single stress/recovery voltage/temperature combination only, becomes prone to numerical errors at lower stress conditions, and does not allow for extrapolation. In a first attempt to overcome these limitations, a log-normal distribution with higher-order polynomials for the mean and variance for the emission times was used [2], which is unfortunately at odds with physical models.

Analytic Capture/Emission Time Map Model

We base our CET map model on the capture and emission times as described by the non-radiative multiphonon model for charge exchange [4], with time-constants of the form $\tau = \tau_0 \exp(\beta E_{\text{A}})$. Rather than considering the various defect parameters impacting E_{A}, like the Huang-Rhys factor or the energy levels of the trap, we deal with the effective activation energy E_{A} directly. We consider the following: *(i)* Since both capture and emission are thermally activated processes [1, 4], we model the distribution of the activation energies rather than the time constants themselves. *(ii)* As the time constants are uncorrelated with the depth of the defect into the oxide [5], we use an effective prefactor $\langle \tau_0 \rangle$. *(iii)* Recent results have shown that BTI degradation consists of a recoverable component R which dominates the recovery over the whole experimental window, starting from a microsecond up to weeks [6]. Furthermore, a more permanent component P is observed, which is not fully permanent but merely recovers on timescales outside usual experimental windows [6]. By heating the sample, these time constants are dramatically reduced, leading to accelerated recovery also of P [7, 8]. Since in physical models the capture and emission times are correlated [9], we express the mean of the emission time as $\mu_{\text{e}} = \mu_{\text{c}} + \Delta\mu_{\text{e}}$. As such, both components can be described by regular bivariate normal distributions, see Fig. 1. *(iv)* While the temperature dependence is inherently considered by the distribution of the activation energies, the bias-dependence of the CET maps is modeled by assuming the amplitude of each component to follow $A = (V_{\text{s}}/V_{\text{s0}})^m$, with the stress voltage V_{s} and constants V_{s0} and m. Also, as previous studies on individual defects have shown [4], the mean values of the distribution are expected to approximately follow $\mu_{\text{c}} = \mu_{\text{c0}} + kV_{\text{s}}$ and $\Delta\mu_{\text{e}} = \Delta\mu_{\text{e0}} - kV_{\text{s}}$, with k a constant. The main effect of k is to shift the capture times to shorter values without affecting the emission times. However, the effect of the bias on the mean values was found to be small and completely negligible for the permanent component.

Analytic Degradation Model

Using the analytic model for the CET maps, the degradation in response to arbitrary digital switching between a given stress and recovery voltage can be calculated in a straightforward manner [1, 2]. We give explicit analytic solutions for the most important cases (see Fig. 2), while other cases follow analogously: *(i)* The simplest solution is obtained for a delay-free on-the-fly measurement (OTF), which is simply given by the integral of the capture-time distribution starting from the first measurement point t_0 up to the stress time t_{s}. *(ii)* In a measure-stress-measure (MSM) experiment, first all defects with capture time constants smaller than the stress time t_{s} are filled. Then, during the subsequent recovery, all defects with emission time constants smaller than the recovery time t_{r} are discharged. Thus, for any combination of t_{s} and t_{r}, the total ΔV_{th} is obtained by integrating the CET maps from 0 to t_{s} and t_{r} to ∞. As such, ΔV_{th} can be expressed via the cumulative distribution function of the bivariate normal distribution, see Fig. 3. *(iii)* Finally, for a DF stress with duty-factor α and period T, three regions of the CET map are filled in a different manner [2]: first, defects with emission times smaller than $(1-\alpha)T$ are completely discharged during each emission cycle and recharged in the following stress cycle. Second, defects with large emission times are not discharged during the emission cycle and charge up to $\tau_{\text{c}} < \alpha t_{\text{s}}$. Finally, in the transition region, defects with $\tau_{\text{e}} > \tau_{\text{c}}/\gamma$, with the on-off ratio $\gamma = \alpha/(1-\alpha)$, remain charged as well, resulting in a beveled box. Using basic integrals of the bivariate normal distribution, the total ΔV_{th} can be analytically determined at all stress and recovery times.

Model Evaluation

The model is first evaluated on a 2.2 nm SiON technology [10], where very detailed MSM data was acquired for the

construction of the CET maps. Recording of each dataset required about one to two weeks. Fig. 4 shows a comparison of the numerically determined CET map with the analytic model, which contains all essential features. In particular, the rightward slant of the distribution with increasing capture times, which previously necessitated the introduction of the higher-order polynomials for the mean and variance of the single normal distribution is well captured by the superposition of two bivariate normal distributions. By simultaneously extracting the analytic distribution for a number of datasets recorded at different V_s and T_L, a bias- and temperature-dependent analytic CET map model is obtained, see Fig. 5. A convincing comparison of the analytic model to experimental data for a number of V_s/T_L combinations is given in Fig. 6. Furthermore, Fig. 7 and Fig. 8 demonstrate how the analytic model captures the build-up of a more permanent degradation by low-frequency duty-factor cycling of the samples. Finally, the model is successfully evaluated against different samples in Fig. 9 (3 nm SiO_2) and Fig. 10 (2.5 nm HfSiO). As a comparison, extracted activation energy distributions are shown in Fig. 11, which, although qualitatively similar, show interesting differences. For instance, the HfSiON sample shows a large recoverable component with small activation energies, consistent with earlier reports [11].

Long-Term Stability of the CET Maps

In order for our model to be exact, the CET maps have to remain stable with repeated cycling. Interestingly, this is not always the case. It has already been previously observed that the amount of ΔV_{th} which can be cycled in a DF experiment can decrease [12]. In terms of the CET maps, which predict a constant amount of cycle-able charge, this implies that the density of the recoverable defects decreases with time. While nitrided samples have been reported to be relatively stable [12–14], as confirmed in Fig. 7 and Fig. 8, non-nitrided samples can display a significant reduction in R. This is shown in Fig. 12 for a 3 nm SiO_2 technology, with about 25% reduction after 30 cycles of 10 ks stress/10 ks recovery at 200 °C. Still, the degradation at large times is dominated by P, so that the overall accuracy of the model appears to be not dramatically affected by this effect. Another test [2, 3] for the stability of R is to compare the recovery traces after an AC stress of duration $2 \times t_s$ to a DC stress of duration t_s, which according to Fig. 2 have to merge at $t_r \approx t_s$. Two extreme cases are given in Fig. 13, for a non-nitrided and nitrided high-k oxide, with good agreement of the analytic model with the data.

Lifetime Prediction

Being physics-based rather than an empirical fit, the analytic model allows for lifetime extrapolation from higher V_s and T_L to operating conditions. Conventionally used power-law fits do not differentiate between the two components and do not consider saturation, leading to very pessimistic lifetimes. Also, the popularized reaction-diffusion model enforces a relatively large power-law slope of $1/6$ even at longer stress times [15], leading to even more pessimistic lifetime estimates. In contrast, our model predicts ΔV_{th} to basically follow the sum of two error-functions, which results in increasingly lower n with increasing stress time [16] until complete saturation (Fig. 14). This saturation level is predicted to be bias- but not temperature-dependent. As a consequence of this saturation, a dramatic increase in the predicted lifetimes is observed, see Fig. 15. Also, as already frequently suspected, extrapolation from higher V_s results in strong non-linearities and thus uncertainties due to the bias-dependence of the saturation level. On the other hand, as the saturation-level is temperature-independent, extrapolation from higher T_L is much better behaved and therefore to be recommended.

Conclusions

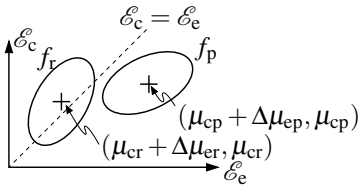
We have suggested a physics-based analytic model for BTI which covers DC, AC, and duty-factor dependent stress and the subsequent recovery as a function of stress voltage and temperature. Since the model is intuitively based on the occupancy of defects in the capture/emission time maps, it can be easily generalized to other more complicated stress/recovery patterns. Although the occasional reduction of the recoverable component in non-nitrided oxides is not well understood, the model provides an upper bound for the degradation which is considerably more optimistic than predictions obtained from previously suggested extrapolation methods.

Acknowledgments

The research leading to these results has received funding from the Austrian Science Fund (FWF) project n°23390-N24, the European Community's FP7 project n°261868 (MORDRED), and from the ENIAC project n°820379 (MODERN). Valuable comments by Ph. Roussel are appreciated.

References

[1] H. Reisinger, T. Grasser, W. Gustin, and C. Schlünder, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2010), pp. 7–15.
[2] H. Reisinger, T. Grasser, K. Ermisch, H. Nielen, W. Gustin, and C. Schlünder, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), pp. 597–604.
[3] K. Zhao, J. Stathis, B. Linder, E. Cartier, and A. Kerber, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), pp. 372–380.
[4] T. Grasser, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, and B. Kaczer, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2010), pp. 16–25.
[5] T. Nagumo, K. Takeuchi, T. Hase, and Y. Hayashi, in *Proc. Intl.Electron Devices Meeting (IEDM)* (2010), pp. 628–631.
[6] T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P.-J. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), pp. 605–613.
[7] A. Katsetos, Microelectronics Reliability **48**, 1655 (2008).
[8] C. Benard, G. Math, P. Fornara, J. Ogier, and D. Goguenheim, Microelectronics Reliability **49**, 1008 (2009).
[9] T. Grasser, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), (tutorial).
[10] T. Grasser, B. Kaczer, P. Hehenberger, W. Goes, R. O'Connor, H. Reisinger, W. Gustin, and C. Schlünder, in *Proc. Intl.Electron Devices Meeting (IEDM)* (2007), pp. 801–804.
[11] B. Kaczer, T. Grasser, J. Martin-Martinez, E. Simoen, M. Aoulaiche, P. Roussel, and G. Groeseneken, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2009), pp. 55–60.
[12] Y. Gao, A. Boo, Z. Teo, and D. Ang, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), pp. 935–940.
[13] Z. Teo, D. Ang, and K. See, in *Proc. Intl.Electron Devices Meeting (IEDM)* (2009), pp. 737–740.
[14] D. Ang, Z. Teo, T. Ho, and C. Ng, IEEE Trans.Dev.Mat.Rel. **11**, 19 (2011).
[15] S. Mahapatra, V. Maheta, A. Islam, and M. Alam, IEEE Trans.Electron Devices **56**, 236 (2009).
[16] H. Aono, E. Murakami, K. Okuyama, A. Nishida, M. Minami, Y. Ooji, and K. Kubota, Microelectronics Reliability **45**, 1109 (2005).
[17] W.-J. Tsay and P.-H. Ke, IEAS Working Paper No. 09-A011 (2009).
[18] B. Kaczer, A. Veloso, P. Roussel, T. Grasser, and G. Groeseneken, J.Vac.Sci.Technol.B **27**, 459 (2009).
[19] V. Huard, M. Denais, and C. Parthasarathy, Microelectronics Reliability **46**, 1 (2006).



$$f(x,y,\rho) = \phi(y)\,\phi\Big(\frac{x-\rho y}{\sqrt{1-\rho^2}}\Big)\Big/\sqrt{1-\rho^2}$$

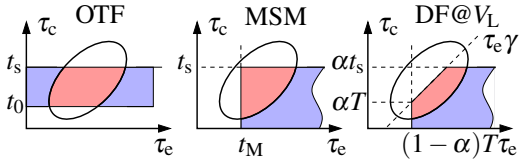
$$\tau_e = \tau_0 \exp(\beta(\mu_c + \Delta\mu_e + x\sqrt{\sigma_e^2+\sigma_c^2}))$$

$$\tau_c = \tau_0 \exp(\beta(\mu_c + y\sigma_c))$$

$$\rho = \sigma_c/\sqrt{\sigma_e^2+\sigma_c^2}, \qquad \beta = 1/k_B T_L$$

$$\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$$

Fig. 1: The analytic CET map model using two bivariate normal distributions f for the activation energies of the recoverable and permanent components R and P. The emission and capture times τ_c and τ_e are correlated with correlation coefficient ρ and obtained from the normalized variates x and y. T_L is the temperature and $\phi(x)$ the p.d.f. of the univariate normal distribution.



$$\textbf{OTF}: V = A\,(G(0,t_s) - G(0,t_0))$$

$$\textbf{MSM}: V = A\,G(t_r,t_s)$$

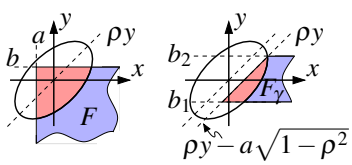
$$\textbf{DF@}V_L: V = A\,(G((1-\alpha)T, \alpha T) + G_\gamma((1-\alpha)T, \alpha t_s))$$

$$G(0,t_s) = \Phi(\tau_c^{-1}(t_s))$$

$$G(t_r,t_s) = F(\tau_e^{-1}(t_r), \tau_c^{-1}(t_s))$$

$$G_\gamma(t_r,t_s) = F_\gamma(\tau_e^{-1}(t_r), \tau_c^{-1}(t_s))$$

Fig. 2: Three examples on how the contribution V of each component to ΔV_{th} is calculated by simple analytical integration of the CET map. The recovery-free OTF scheme misses the lower part of the distribution ($\tau_c < t_0$), the measure-stress-measure scheme (MSM) the left part ($\tau_e < t_M$), and duty-factor stress with period T and duty-factor α has to consider a beveled box [2], determined by on/off ratio $\gamma = \alpha/(1-\alpha)$.



$$F(a,b,\rho) \approx \Phi(b) - \Phi(a/\rho) - \tfrac{1}{2}e^{c^2\rho^2/2} \times (e^{ca}(\Phi(b+c\rho) - \Phi(a/\rho + c\rho)) + e^{-ca}\Phi(a/\rho - c\rho)) \quad (\text{for } a < \rho b)$$

$$\approx \tfrac{1}{2}e^{c^2\rho^2/2 - ca}\Phi(b - c\rho) \quad (a \geq \rho b)$$

$$F(-\infty, b) = \Phi(b)$$

$$F_\gamma(a, b_1, b_2, \rho) = (\Phi(b_2) - \Phi(b_1))\,\Phi(a)$$

Fig. 3: The two integrals F and F_γ of the bivariate normal distribution f required to determine a wide range of stress/recovery patterns. While no exact solution for F exists, various approximations are available, with most of them being too crude for our purposes. The above approximation is derived using the method suggested in [17], with $c = \sqrt{3/(2(1-\rho^2))}$ and the normal integral $\Phi(x) = (1+\mathrm{erf}(x/\sqrt{2}))/2$.

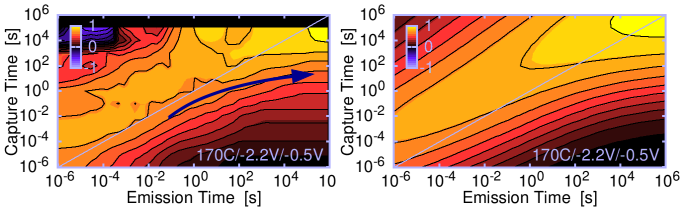


Fig. 4: Left: The CET map extracted from experimental MSM data by numerically differentiation [1, 9], which is limited to the experimental window. To evaluate the method, we used an extremely wide window of $t_s = [1\,\mu s \ldots 100\,ks]$ and $t_r = [1\,\mu s \ldots 9\,days]$. All maps are normalized and plotted using a signed log operator $\mathrm{sign}(g)\log_{10}(1+\kappa|g/g_{max}|)/\log_{10}(1+\kappa)$ with $\kappa = 100$ to bring out all important details. **Right**: The analytic CET map model of Fig. 1 fit to the same MSM data set, which contains all essential features.

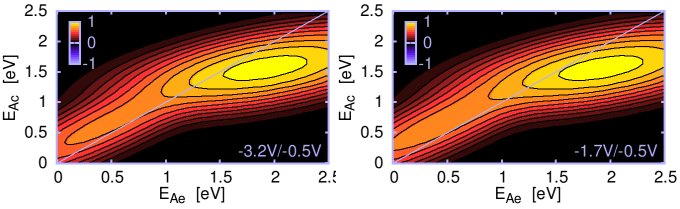


Fig. 5: The analytic activation energy map extracted from data recorded at the bias conditions $V_s = [-1.7\,V \ldots -3.2\,V]$ and $V_{relax} = -0.5\,V$ at two temperatures (170°C and 125°C). While the temperature dependence is inherently considered, the applied bias changes the map only slightly, left for $-3.2\,V$, right for $-1.7\,V$. The strongest changes is due to the change of the amplitudes A_r and A_p. Since both amplitudes change in a comparable way, this aspect is barely visible in the above plots due to the normalization to unity.

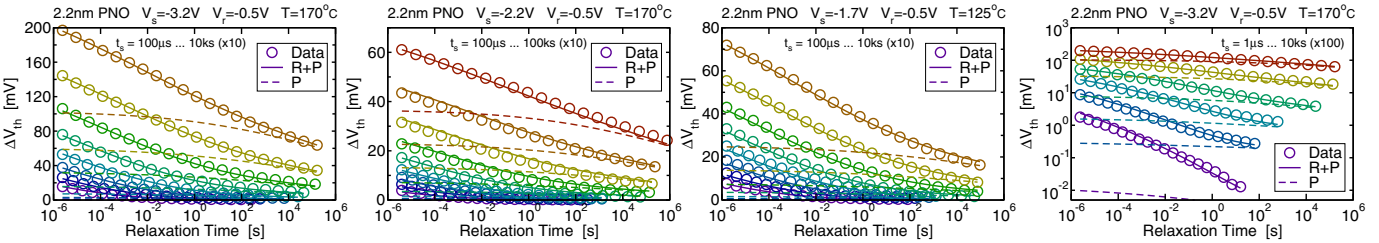


Fig. 6: Comparison of the analytic model using the activation energies of Fig. 5 to experimental data at different stress biases and temperatures. Excellent agreement is obtained for all stress and relaxation times in the extremely wide experimental window also on a logarithmic scale (right-most figure).

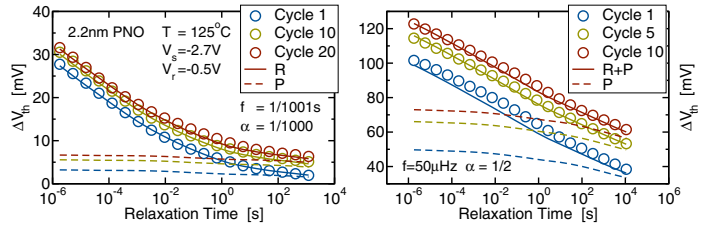


Fig. 7: Comparison of the analytic model against low-frequency duty-factor experiments, which favor the build-up of P. Excellent agreement is obtained in both cases. **Left**: Stress for 1 s and relax for 1 ks, repeated 20 times. **Right**: Stress for 10 ks and relaxed for 10 ks, repeated 10 times.

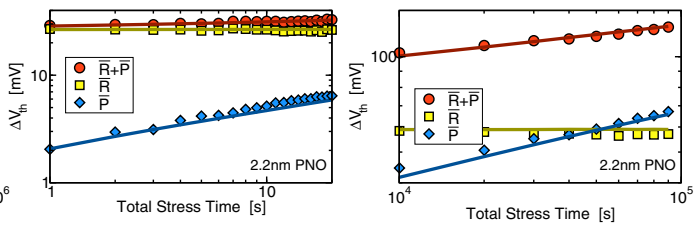


Fig. 8: Using the definitions $\overline{P} = \Delta V_{th}(t_{relax,last})$ and $\overline{R} = \Delta V_{th}(t_{relax,first}) - \overline{P}$, the data of Fig. 7 is replotted as a function of the total stress time. Consistent with the observations of [12], the amount of recoverable ΔV_{th} per cycle, $\overline{R}$, is stable. As a consequence, the CET map model (lines) is in excellent agreement with the data.

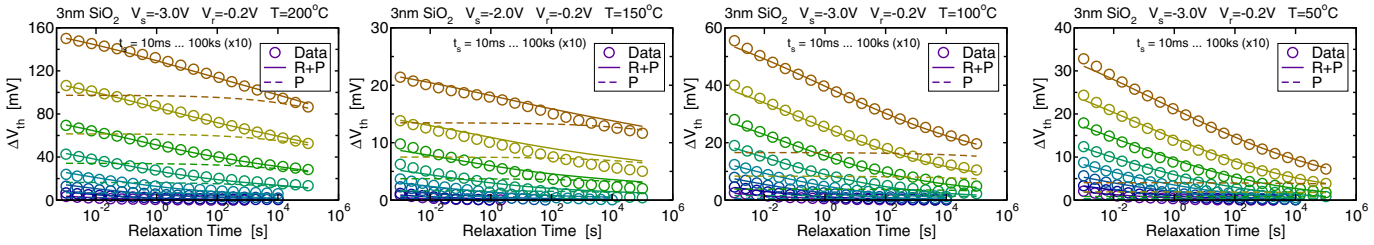


Fig. 9: Comparison of the analytic model to detailed MSM data for a 3 nm SiO_2 device.

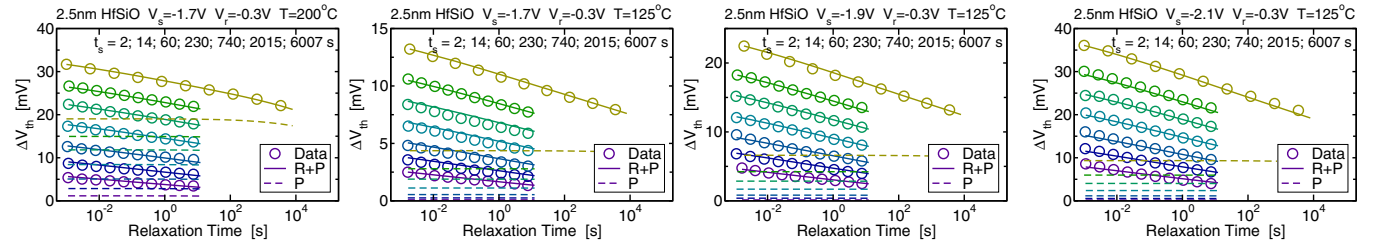


Fig. 10: Comparison of the analytic model to detailed MSM data for a EOT=2.5 nm HfSiO device [18].

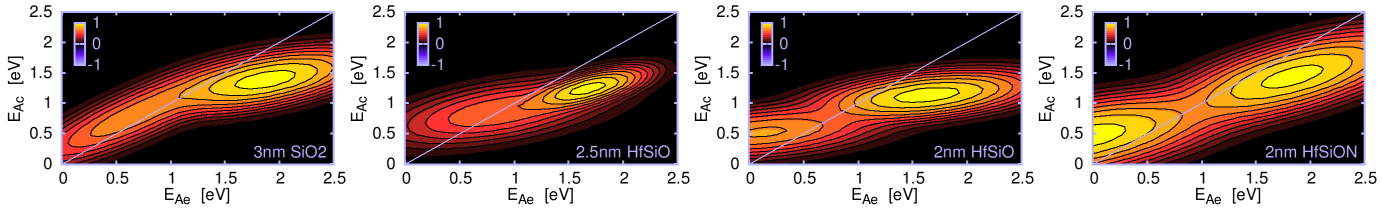


Fig. 11: The CET maps extracted for a number of different oxides. The first corresponds to Fig. 9 while the second belongs to Fig. 10. The trappy HfSiON from [11] (fourth map) shows a four-times larger NBTI degradation and an order of magnitude larger noise power than its non-nitrided HfSiO counterpart (third map).

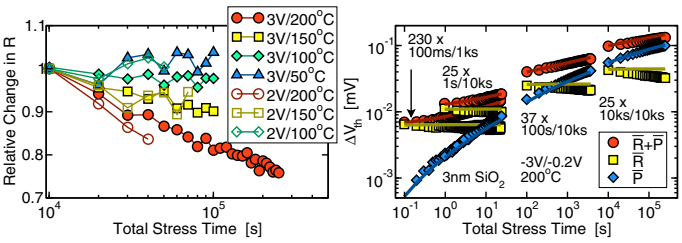


Fig. 12: Left: Similar low-frequency duty-factor experiment as in Figs. 7 and 8 but now performed on a non-nitrided oxide where the relative loss in $\overline{R}$ can be significant. Contrary to [12], in this device the effect is stronger at higher temperatures and appears to be only weakly field-dependent. **Right**: Repeated cycling between $-3\,\mathrm{V}$ and $-0.2\,\mathrm{V}$ dominantly builds up $\overline{P}$, maintaining the good accuracy of the model (lines).

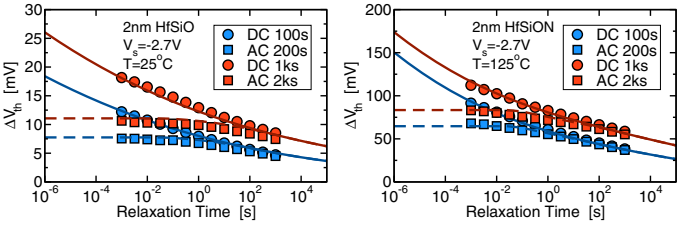


Fig. 13: Two extreme cases of recovery after DC and AC stress for the non-nitrided HfSiO device at 25 °C (left) and the nitrided HfSiON device at 125 °C. As already demonstrated in [2, 3] for PBTI, recovery after AC stress of duration $2 \times t_s$ merges with the DC recovery after stress of duration t_s at $t_r = t_s$. The analytic model (solid lines for DC and dashed lines for AC) reproduces the data very well.

Fig. 14: Extrapolated degradation at constant temperature (left) and constant bias (right). While in the shown measurement range the data with a delay of $1\,\mu\mathrm{s}$ (syms and dashed lines) can be well fitted with a power-law, the CET map model results in the sum of two error-functions, one for R and the other for P. When the data is extrapolated to zero delay (solid lines), the degradation level becomes larger but the slope decreases, which is beneficial for the lifetime. Furthermore, particularly for P a relatively strong curvature is observed (dotted lines), consistent with previous observations [10, 19]. Finally, the degradation saturates with the predicted final total degradation level depending only on the bias but not on the temperature.

Fig. 15: The calculated lifetimes ($\Delta V_{\mathrm{th}}(t_{\mathrm{lifetime}}) = 30\,\mathrm{mV}$) obtained from the CET map model extrapolated at constant temperature (left) and constant bias (right). Also shown are the results obtained from a power-law (PL) extrapolation fitted in the range $t_s = [1\,\mathrm{s} \ldots 10\,\mathrm{ks}]$ and to the prediction of the reaction-diffusion (RD) model [15]. For the last two extrapolation schemes data recorded with a delay of $1\,\mu\mathrm{s}$ was taken. For the conventional power-law extrapolation, the initial amplitude is lower but the slope higher, leading to a 'benefit' at higher V_s but an overly pessimistic lifetime at operating conditions ($V_G = 1.5\,\mathrm{V}$). The RD model postulates a saturated hole-trapping component at $t_s = 1\,\mathrm{s}$ and enforces a high longer-term slope of $1/6$, leading to very pessimistic lifetimes.