NBTI from the perspective of defect states with widely distributed time scales

B. Kaczer, T. Grasser1, J. Martin-Martinez2, E. Simoen, M. Aoulaiche, Ph. J. Roussel, G. Groeseneken3
IMEC, Kapeldreef 75, B-3001 Leuven, Belgium
1 TU Wien, Austria
2 UAB Barcelona, Spain
3 also ESAT, KU Leuven, Belgium
phone: +32 (16) 281-557, fax: +32 (16) 281-706, e-mail: Ben.Kaczer@imec.be

Abstract—Broad similarity between negative bias temperature instability (NBTI) relaxation and 1/f noise is observed. Individual transitions in NBTI relaxation in small pFETs are observed and Poisson defect number statistics is inferred. Finally, it is argued that the wide distribution of defect times should be considered in addition to defect number variation in small devices.

Keywords- pFET, reliability, negative bias temperature instability (NBTI), 1/f noise, single charge events

I. INTRODUCTION

Bias-temperature instabilities (BTI) in CMOS technologies with both conventional and advanced gate stacks remain at the forefront of reliability concerns. Central to the negative BTI (NBTI) mechanism is the existence of defect states with a wide distribution of time scales. This concept, constituting the chief complication for the proper evaluation and hence understanding of NBTI, is further developed here.

We first note that the very wide time scale distribution of defect states, observable e.g. in NBTI relaxation experiments, is also standardly invoked to explain the so-called 1/f noise spectra [1], observed in our devices. We further observe many other similarities in the properties of the NBTI relaxation component [2] and the FET noise spectrum, be it nitrogen concentration, temperature, voltage, or NBTI-stress dependences. This leads us to speculate that both are in fact due to the same (gate oxide) defects.

We further visualize the picture of discrete states by monitoring individual transitions in NBTI relaxation in small pFETs, allowing us to directly link the threshold voltage shift \( \Delta V_{th} \) with the number of charged defects. Poisson defect number statistics is demonstrated.

Finally, we discuss the effect of the widely distributed time scales on the NBTI measurements and the switching of a digital inverter. We illustrate this by showing that defects with time scales visible in standard NBTI measurements could differ substantially from those active during inverter switching. We conclude that the distribution of defect times should be considered in addition to defect number variation in small devices and show how certain applications, such as SRAM, could be adversely affected.
II. EXPERIMENTAL

pFET devices with 1.4 nm EOT SiON / poly-Si gates [2] were used in most measurements. Where noted, pFETs with 2.0 – 2.4 nm CET HfSiO gate oxides and TaN gates were used, without or with NH₃ or DPN nitridation [3]. Extended measure-stress-measure (eMSM) technique was employed, alternating gate voltage \( V_G \) between \( V_{\text{stress}} \) and \( V_{\text{meas}} \) (chosen \( \approx V_{\text{th}} \)) [2] at temperature \( T = 125 ^\circ \text{C} \) on 10 \( \mu \text{m} \) wide and 0.25 – 0.5 \( \mu \text{m} \) long devices and analyzed following Ref. [4]. NBTI fluctuations were monitored on 0.25 \( \times \) 0.25 \( \mu \text{m}^2 \) SiON pFETs. Noise measurements were performed on a dedicated setup at low \( |V_G| \) to avoid stressing the tested device. NBTI stressing between noise measurements (Figure 5) was achieved by performing noise measurements at \( V_G = V_{\text{stress}} \).

III. RESULTS AND DISCUSSION

A. Similarities between NBTI relaxation and 1/f noise

Long, log(\( t \))-like behavior of \( \Delta V_{\text{th}} \) is typically observed in both the initial portion of NBTI degradation [5, 6] and the recovery phase. Figure 1a illustrates that the rate of degradation \( d\Delta V_{\text{th}}/d\tau_{\text{relax}} \) extracted from the log(\( \tau_{\text{relax}} \))-like \( \Delta V_{\text{th}} \) relaxation transient after even a very short, 0.1 s stress, follows \( 1/\tau_{\text{relax}} \) for over 7 decades. Such behavior is a signature of states with discharging time constants covering as many decades [2].

Incidentally, superposition of relaxation of states with widely distributed time scales is the standard explanation of the 1/f noise spectra [1], which are clearly observed in our pFETs (Figure 1b). This obvious similarity leads us to argue the possibility that the same states with widely distributed time scales in fact play a fundamental role in both NBTI and noise measurements.

This direct correspondence of both phenomena is further supported by several additional observations:

- Nitridation of a (high-k) gate stack results in a substantial increase in both the log(\( t \))-like relaxation component and the 1/f noise spectra, as documented in Figure 2 [8, 9].
- At low \( V_G \), the temperature dependences of both the relaxation rate and the 1/f noise are weak (Figure 3) [10].
- The relaxation component is observed to increase as \( (V_{\text{stress}} - V_{\text{meas}})^2 \) (Figure 4) [2,5]. Similarly, \( S_{\text{VG}} \) dependence goes like \( (V_G - V_{\text{th}})^2 \) as seen in Figure 2b.
- The impact of NBTI stress on both the noise (Figure 5) and the relaxation component of \( \Delta V_{\text{th}} \) (shown e.g. in Figure 6a of Ref. [2]) is relatively weak [11].

![Figure 2](image_url)

**Figure 2.** (a) Log-like relaxation component following stress at a fixed oxide electric field of \( \sim 9 \) MVcm\(^{-1} \) is substantially larger in nitrided gate stacks [3]. (b) Similarly, nitrided stacks show much stronger noise \( S_{\text{VG}} \) than their non-nitrided counterpart. \( S_{\text{VG}} \) is observed approximately increasing with \( (V_G - V_{\text{th}})^2 \).

![Figure 3](image_url)

**Figure 3.** At low \( V_G \), both the recoverable component rate [5] and noise spectrum \( S_{\text{VG}} \) have at most very weak temperature dependence.
The recoverable component is observed to increase as \((V_{\text{stress}} - V_{\text{meas}})^2\). Note that \(V_{\text{meas}} \approx V_{\text{th}}\) in our eMSM measurements.

Our NBTI model reported in Ref. [5], designed with the emphasis on capturing the wide temporal distribution of relaxing defects readily reproduces the 1/f noise spectra (Figure 6). The same figure also documents the weak \(T\) dependence at low \(V_G\) observed in Figure 3.

The correlation between \(S_{IV}\) and the relaxation component is also strong when comparing nitrided and non-nitrided stacks. As shown in Figure 2, the \(S_{IV}\) noise increases approximately tenfold in the nitrided stacks. Coincidentally, the recoverable component increases by the same factor in the nitrided stacks, as shown in Ref. 3. Conversely, we typically see that during stress the recoverable component is increasing only weakly [2], as is also documented by the very low power-law exponents reported by others [12]. This seems consistent with the relatively small increase in \(S_{IV}\) [9] after stress that by \(\Delta V_{th}\) consideration results in more than one decade increase in the total defect density (see Figure 5).

The most intriguing correlation is the apparent quadratic voltage dependence of the recoverable component (Figure 4) [2] and the noise spectra (Figure 2b). The \((V_G - V_{th})^2\) dependence of \(S_{IV}\) has been previously ascribed to channel carrier scattering [13]. In our NBTI model [5] the \((V_{\text{stress}} - V_{\text{th}})^2\) dependence of the recoverable component is explained by multiphonon-field-assisted tunneling of holes into defect precursors in the oxide [14]. This allows the model to reproduce the superlinear voltage dependence of \(S_{IV}\).

All the above arguments are a strong indication of a possible link between NBTI and 1/f noise. Irrespective of the existence of a common physical mechanism, we invoke the phenomenological similarities in the last subsection.

**B. NBTI relaxation in very small pFETs**

In small devices, noise typically manifests itself as random telegraph signal. Similarly, the states with widely distributed relaxation times behind NBTI can be readily visualized in small-area pFETs. Discrete NBTI relaxation \(\Delta V_{th}\) steps corresponding to individual discharging events are clearly apparent in the relaxation transient in Figure 7. The height of the steps varies significantly due to random spatial distribution in the channel plane [15, 16] and cannot be simply ascribed to defect depth in the oxide. By counting individual steps, the average step height is determined to be about 0.6 mV per \(q\), about 3× the value inferred from the simple \(q/C_{ox}\) assumption. The factor of 3× is consistent with Ref. [17].

It is moreover apparent from Figure 7 that the down-steps, responsible for the overall relaxation, are distributed over all measured decades. This directly confirms our assumption made previously on large pFETs. The same trend is apparent in Figure 8 showing the relaxation transients of 11 small devices. Again, counting individual steps allows us to construct the secondary y-axis in Figure 8.

![Figure 7](image_url) Individual down-steps due to deactivation of discrete defects are clearly visible in every decade of the relaxation transient taken on a small pFET, resulting in \(-\log(t_{relax})\) dependence. The average height of a down-step is \(-0.6\) mV.
typical time constants are log-uniformly

With the mean of the 11 transients behaves as \( \Delta V_{th} \) of a large device (e.g. Figure 1a), the device-to-device spread is not decreasing with increasing \( t_{relax} \). Normalized device-to-device variance \( \phi = \text{var}[n(t_{relax})]/<n(t_{relax})> \) with \( n(t_{relax}) \) from Figure 8 is close to unity (within the limited sample of 11 devices) for 0.6 mV/\( \phi \). The device-to-device spread in \( \Delta V_{th} \) remains approx. constant with increasing \( t_{relax} \).

C. NBTI from the perspective of defect states with widely distributed time scales

As already implied above, the defects created during NBTI stress do not have a single, but rather widely varying time constants. Below we show how to visualize this property and within this framework we briefly discuss how the fraction of active defects is affected by the different applications of NBTI stress. We explain on the example of a very small device how the distribution of the defect times should be considered.

We start by noting that in either NBTI relaxation or 1/f noise measurements, no maximum or minimum times are typically observed. For the sake of simplicity we therefore assume here that the time constants are log-uniformly distributed from times much shorter than the switching time of a pFET to very long, corresponding to lifetime of a CMOS application. To facilitate semi-quantitative simulations, defect states are represented by “RC” elements (Figure 10 [2]) with the total pFET \( \Delta V_{th} \) being proportional to the sum of voltages (“occupancies”) on all capacitors. To facilitate our explanation, we first consider that all RC elements have the same weight and can be partially occupied, which would emulate the behavior of a large-area device.

Figure 11a visualizes the log-uniformly distributed “states” with different time constants sequentially charging as the consequence of constant DC stress, a.k.a. the “on-the-fly” (OTF) NBTI measurement. As expected, the fastest states charge first, with slower states becoming occupied as the stress time increases. Upon the removal of the stress (Figure 11b), the fastest states are quickly discharged, illustrating why we observe only a fraction of all states during eMSM recovery sampling (e.g. Figures 7 and 8).

In the case of a digital alternating (“AC”) waveform, some states will be only partially represented in \( \Delta V_{th} \) [19]. This is illustrated in Figure 12 for the varying Duty Factor of the AC stress waveform. States with very short time constants will be again instantly occupied upon the application of stress (cf. Figure 11a). Conversely, states with time constants comparable to the reciprocal frequency of the applied signal will respond and will be (on average) only partially occupied.
In a digital inverter, the pFET $V_{th}$ shift influences the delay only when the pFET is turning on [20]. This case is visualized in Figure 13 for a 50% Duty Factor AC input signal. As can be seen in this example, only a subset of all states will be affecting this inverter transition. Comparison with Figure 11 illustrates the discrepancy with the fraction of charged states observed in both most common NBTI tests. In particular, the OTF test substantially overestimates the number of charged defects, i.e., the $\Delta V_{th}$ over the real-world case.

On the other hand, in some other applications, such as the SRAM, some pFETs can be continuously on, and hence stressed, for extensive periods of time. As we have shown above, the number of defects in such minimum-size pFETs will be governed by the Poisson distribution. To calculate the number of charged defects for each pFET instance at a given stress time, the widely distributed time scales of the defects need to be considered as well. The result of such calculation in Figure 14 illustrates that if a particular pFET exists in the population with a large number of defects and their basic temporal and device-to-device properties were quantitatively visualize active defects during the different transitions in NBTI relaxation in small pFETs were observed. Primarily, this correspondence is due to both phenomena being caused by states with widely distributed time scales. Similarly to random telegraph signal, individual transitions in NBTI relaxation in small pFETs were observed and their basic temporal and device-to-device properties were discussed. A distributed “RC” circuit was used to semi-

IV. CONCLUSIONS

Broad similarity between NBTI relaxation and 1/f noise was observed. Primarily, this correspondence is due to both phenomena being caused by states with widely distributed time scales. Similarly to random telegraph signal, individual transitions in NBTI relaxation in small pFETs were observed and their basic temporal and device-to-device properties were discussed. A distributed “RC” circuit was used to semi-

ACKNOWLEDGMENTS

This work was carried out as part of IMEC’s Industrial Affiliation Program funded by IMEC’s core partners and has been partially supported by the Spanish MICINN (BES-2005-10870).

REFERENCES


