Revisiting Charge Trapping/Detrapping in Flash Memories From a Discrete and Statistical Standpoint—Part I: V_T Instabilities

Giovanni M. Paolucci, Christian Monzio Compagnoni, *Senior Member, IEEE*, Carmine Miccoli, Alessandro S. Spinelli, *Senior Member, IEEE*, Andrea L. Lacaita, *Fellow, IEEE*, and Angelo Visconti, *Member, IEEE*

I. Introduction

TRAPPING and detrapping of charge in the gate oxide of MOS devices are among the most relevant reliability issues for nanoscale logic and memory technologies. Reduction of device dimensions, in fact, has dramatically increased the impact that single electrons and holes in the oxide have on both the threshold voltage (V_T) and the drain current (I_D) of the transistor or the memory cell [2]–[4]. This increase has not appeared just in terms of a higher average value of the threshold-voltage shift (ΔV_T^1) and drain current shift (ΔI_D^1) following a single charge trapping/detrapping event but also, and more important, in the possibility to statistically encounter cases where ΔV_T^1 and ΔI_D^1 display extremely high values [3], [5]–[12]. High ΔV_T^1 and ΔI_D^1 have been attributed not only to the reduced device area but also to 3-D device electrostatics and atomistic doping effects, resulting in percolative channel conduction [7], [13]–[15]. As a result, discreteness of trapping and detrapping events could have been clearly detected in many processes limiting device reliability, e.g., random telegraph noise (RTN) [16]–[20], bias temperature instability [10], [11], [21], [22], and cycling-induced V_T instabilities in Flash memories [12]. Despite the fact that the microscopic oxide defects and the physical mechanisms involved in the previous processes are different, the experimental observation of these processes in nanoscale technologies has highlighted some common features of the charge trapping/detrapping events from which they originate, such as the presence of an exponential tail in the ΔV_T^1 and ΔI_D^1 statistics [7], [10]–[12], [14], [15], [23], the statistical nature of the trapping/detrapping events [12], [20], [21], and the dispersion of the trapping (τ_t) and detrapping (τ_d) time constants over many decades of time [12], [22], [24]–[28].

Manuscript received March 11, 2014; revised April 29, 2014, May 6, 2014, and May 26, 2014; accepted May 27, 2014. Date of publication June 18, 2014; date of current version July 21, 2014. The review of this paper was arranged by Editor E. Rosenbaum.

G. M. Paolucci, C. Monzio Compagnoni, and A. S. Spinelli are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan 20133, Italy (e-mail: paolucci@elet.polimi.it; monzio@elet.polimi.it; spinelli@elet.polimi.it).

C. Miccoli is with Process Research and Development, Micron Technology Inc., Boise, ID 83716 USA (e-mail: cmiccoli@micron.com).

A. L. Lacaita is with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan 20133, Italy, and also with the Consiglio Nazionale delle Ricerche, Istituto di Fotonica e Nanotecnologie, Milan 20133, Italy (e-mail: lacaita@elet.polimi.it).

A. Visconti is with Process Research and Development, Micron Technology Inc., Agrate Brianza 20864, Italy (e-mail: aviscont@micron.com).

Extending the preliminary work presented in [29], in this paper, and in the corresponding Part II [1], we revisit charge trapping and detrapping in Flash memories considering charge discreteness, statistical charge capture and emission with τ_t and τ_d widely distributed over the logarithmic time axis, and statistical distribution of ΔV_T^1. In this Part I, we start addressing V_T instabilities coming from charge detrapping from the cell tunnel oxide, investigating the statistical properties of the detrapping process and the resulting statistical distribution of the displacement of device threshold voltage (ΔV_T) over time. The results have a quite general validity and may be applied straightforwardly to different reliability issues involving charge trapping/detrapping in the oxide of MOS devices, first of all to bias temperature instability in MOSFETs. Moreover, the detailed analyses shown here will be the starting point for the development in Part II of a comprehensive model accounting not only for charge detrapping but also for charge trapping during device operation, resulting in a powerful tool for the predictive investigation of cycling-induced V_T instabilities in Flash memories under whatever on-field usage condition.

II. Statistical Analysis of the Detrapping Process

A. Statistical Dispersion of τ_d

Fig. 1 shows an example of the average ΔV_T ($\langle \Delta V_T \rangle$) transient measured on a nanoscale NAND Flash array during a

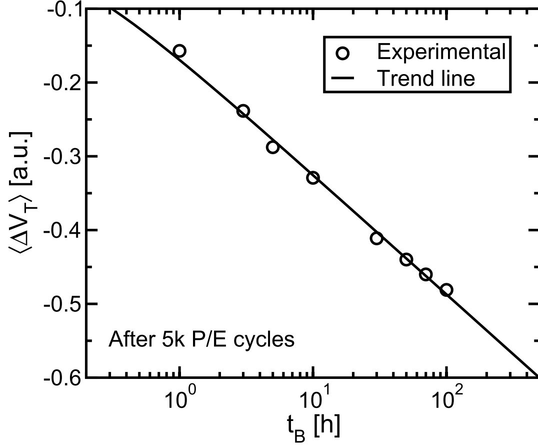

Fig. 1. Typical $\langle \Delta V_T \rangle$ transient measured on a nanoscale NAND Flash array during a postcycling data retention experiment from the programmed state.

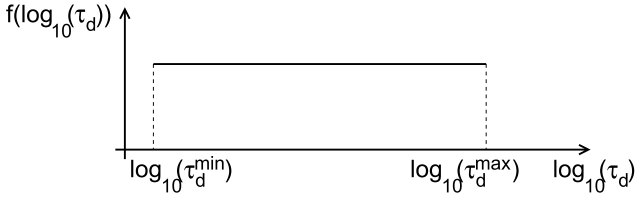

Fig. 2. Schematics for the pdf of the $\log_{10}(\tau_d)$ assumed in this Part I of our paper.

post-cycling data retention experiment from the programmed state, clearly highlighting a logarithmic decrease of V_T with elapsing time [24]–[26], [30]. The transient arises from the neutralization of the negative charge of single electrons trapped in the tunnel oxide, commonly referred to as electron detrapping from the tunnel oxide, and its logarithmic trend has always been recognized as a strong proof that the time constant τ_d of the discrete detrapping events contributing to the process has a very wide statistical spread, spanning many decades of time [12], [24]–[26]. The physical origin of the τ_d spread may be found in the statistical dispersion of the space or energy position of the trapped charge in the oxide [22], [24], [25], in 3-D electrostatics and atomistic doping [27], [28], or in the variability in the structural parameters of the oxide defects where charge is trapped [21]. However, irrespective of the details of the microscopic process leading to a detrapping event and of the physical origin of the τ_d spread, what is really necessary to assume to explain the logarithmic trend of ΔV_T in Fig. 1 is that τ_d is uniformly distributed over the logarithmic time axis. From this starting point and completely neglecting a lower level physical analysis, we assumed that each program/erase cycle on the memory cells contributes to the trapping of electrons that can then be detrapped during a subsequent data retention experiment with τ_d uniformly distributed over the logarithmic time axis, as shown in Fig. 2. Note that the lower ($\tau_d^{\min}$) and upper ($\tau_d^{\max}$) bounds of τ_d in Fig. 2 have been introduced only for mathematical convenience in our analysis and do not have any physical meaning, as will clearly appear in Part II of this paper [1], when discussing the choice of their value.

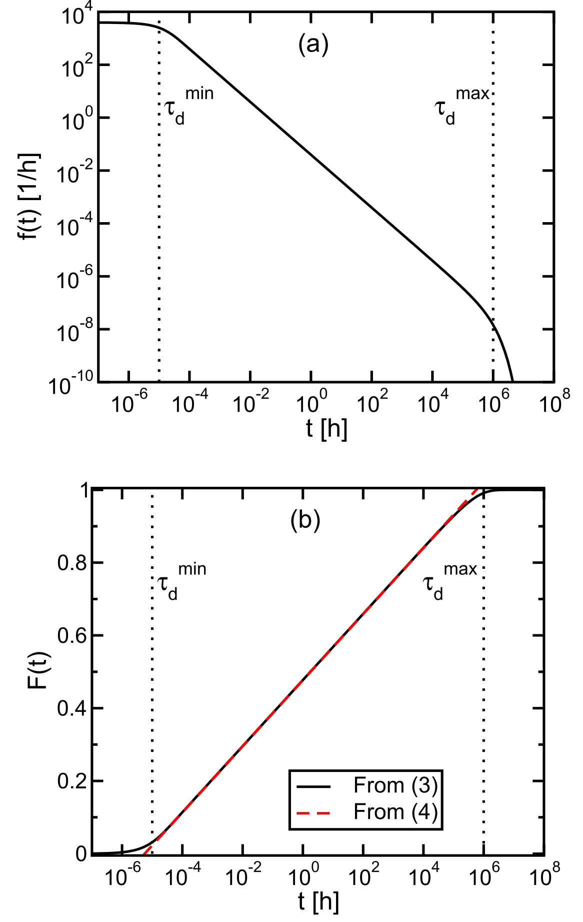

Fig. 3. (a) Pdf and (b) cdf of the time t at which a detrapping event occurs when $\log_{10}(\tau_d)$ is uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h. In (b), the results obtained from (3) and from its approximation for $\tau_d^{\min} < t < \tau_d^{\max}$ given by (4) are shown.

With the statistical distribution of $\log_{10}(\tau_d)$ shown in Fig. 2, the probability density function (pdf—f) of τ_d results in

$$f(\tau_d) = \frac{1}{\tau_d \cdot \ln\left(\tau_d^{\max}/\tau_d^{\min}\right)} \tag{1}$$

for $\tau_d^{\min} < \tau_d < \tau_d^{\max}$, while $f(\tau_d) = 0$ for $\tau_d < \tau_d^{\min}$ and $\tau_d > \tau_d^{\max}$. From $f(\tau_d)$, the pdf and the cumulative distribution function (cdf—F) of the time t at which a detrapping event takes place can be straightforwardly calculated as

$$\begin{aligned} f(t) &= \int_{\tau_d^{\min}}^{\tau_d^{\max}} f(t|\tau_d)\, f(\tau_d) d(\tau_d) \\ &= \frac{e^{-t/\tau_d^{\max}} - e^{-t/\tau_d^{\min}}}{t \cdot \ln\left(\tau_d^{\max}/\tau_d^{\min}\right)} \end{aligned} \tag{2}$$

$$F(t) = 1 - \frac{\int_{t/\tau_d^{\max}}^{\infty} x^{-1} e^{-x} dx - \int_{t/\tau_d^{\min}}^{\infty} x^{-1} e^{-x} dx}{\ln\left(\tau_d^{\max}/\tau_d^{\min}\right)} \tag{3}$$

where $f(t|\tau_d) = 1/\tau_d \cdot \exp(-t/\tau_d)$ represents the pdf of the time at which a detrapping event takes place given that the detrapping time constant is τ_d. Fig. 3(a) shows that $f(t)$ displays a $1/t$ behavior for $\tau_d^{\min} < t < \tau_d^{\max}$, meaning that in the presence of an average number of trapped electrons $\langle N_t \rangle$, a detrapping current $I_d \propto q\langle N_t \rangle / t$ is predicted,

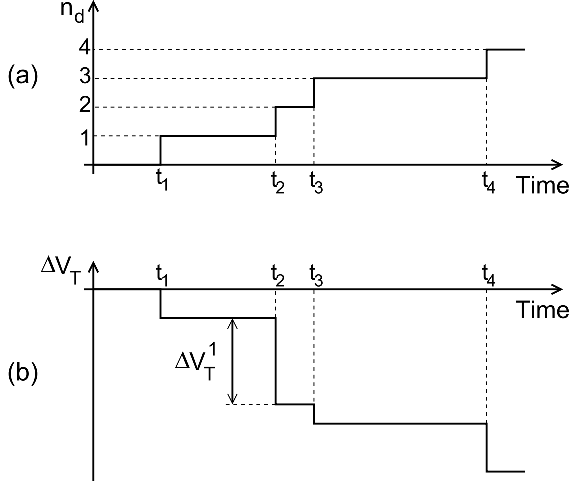

Fig. 4. Schematics for (a) the n_d transient obtained during a detrapping process and (b) corresponding ΔV_T transient.

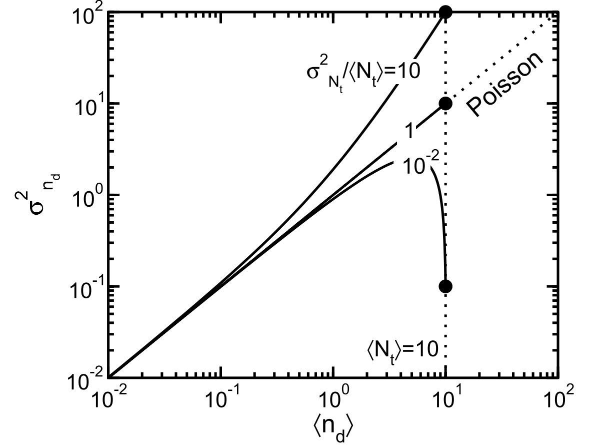

Fig. 5. Calculated $\sigma^2_{n_d}$ versus $\langle n_d\rangle$ relationship for different $\sigma^2_{N_t}$ and $\langle N_t\rangle = 10$. A uniform distribution between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h was assumed for the $\log_{10}(\tau_d)$.

i.e., proportional to $1/t$ and in agreement with experimental observations [24]. Fig. 3(b) shows, in turn, that, in the same range of t, $F(t)$ displays a logarithmic growth which can be well reproduced by the following first order approximation of (3) [31]:

$$F(t) \simeq \frac{\gamma + \ln\left(t/\tau_d^{\min}\right)}{\ln\left(\tau_d^{\max}/\tau_d^{\min}\right)} \tag{4}$$

where γ is the Euler–Mascheroni constant ($\simeq$0.577). This logarithmic trend will be linked to the logarithmic ΔV_T transient in Section III-A.

B. Statistics of Detrapping Events

Assuming now that a cell in the array has a number N_t of trapped electrons as a result of a previous cycling experiment, the number of detrapping events (n_d) taking place up to a time t in a post-cycling data retention experiment has a staircase increase with t, as shown in Fig. 4(a). The steps of the staircase are located at the random times t_i when the ith detrapping event takes place. Considering the detrapping events as independent, $n_d(t)$ follows a binomial distribution and the probability $P[n_d(t)|N_t]$ of having n_d detrapping events up to time t given N_t is

$$P[n_d(t)|\, N_t] = \binom{N_t}{n_d} F(t)^{n_d}\,[1 - F(t)]^{N_t - n_d}. \tag{5}$$

The resulting average number of detrapping events is, therefore, $\langle n_d(t)\rangle = N_t \cdot F(t)$, and the corresponding variance is $\sigma^2_{n_d}(t) = N_t \cdot F(t)[1 - F(t)]$.

When considering the statistical distribution of $n_d(t)$ among the cells of a memory array, the statistical dispersion of N_t must be accounted for. Irrespective of what statistics N_t obeys, in Appendix A we show that quite simple expressions can be derived for $\langle n_d(t)\rangle$ and $\sigma^2_{n_d}(t)$

$$\langle n_d(t)\rangle = \langle N_t\rangle \cdot F(t) \tag{6}$$
$$\sigma^2_{n_d}(t) = \langle N_t\rangle \cdot F(t)\,[1 - F(t)] + \sigma^2_{N_t} \cdot F(t)^2 \tag{7}$$

where $\sigma^2_{N_t}$ is the variance of N_t. From (7), $\sigma^2_{n_d}(t)$ will be equal, larger or lower than $\langle n_d(t)\rangle$ depending on that $\sigma^2_{N_t}$ is equal, larger or lower than $\langle N_t\rangle$, respectively, irrespective of $F(t)$. This is shown with some representative examples in Fig. 5, where $\langle N_t\rangle$ was assumed equal to 10: for $t \to \infty$, the curves must reach the condition $\langle n_d\rangle = \langle N_t\rangle$ and $\sigma^2_{n_d} = \sigma^2_{N_t}$, and this makes the $\sigma^2_{n_d}$ versus $\langle n_d\rangle$ relationship grow as, more and less than that of a Poissonian process for $\sigma^2_{N_t} = \langle N_t\rangle$, $\sigma^2_{N_t} = 10 \times \langle N_t\rangle$, and $\sigma^2_{N_t} = 10^{-2} \times \langle N_t\rangle$, respectively.

Despite the condition $\sigma^2_{N_t} = \langle N_t\rangle$ assures that $\sigma^2_{n_d}(t) = \langle n_d(t)\rangle$, this does not mean that the detrapping process is Poissonian. In Appendix B we show, however, that in the case N_t has a Poissonian distribution, then $n_d(t)$ is the result of a nonhomogeneous Poisson process [32] with intensity $\lambda(t)$ such that $\int_0^t \lambda(x)dx = \langle n_d(t)\rangle = \langle N_t\rangle F(t)$. In this case, the pdf and cdf of the statistical distribution of the time stretch elapsing between the $(i-1)$ th and ith detrapping event (dt_i) can be calculated according to the following formulas [33], valid in the case $i \ll \langle N_t\rangle$:

$$f(dt_i) = \int_0^\infty \frac{\lambda(x)\,\lambda(x + dt_i)\,\langle n_d(x)\rangle^{i-2} e^{-\langle n_d(x+dt_i)\rangle} dx}{(i-2)!}$$
$$F(dt_i) = \int_0^\infty \frac{\lambda(x + dt_i)\,\langle n_d(x)\rangle^{i-1} e^{-\langle n_d(x+dt_i)\rangle} dx}{(i-1)!}.$$

Fig. 6 shows the complementary value of $F(dt_i)$ in the case the $\log_{10}(\tau_d)$ is uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h and $\langle N_t\rangle = 10$. The most important result appearing from the figure is that $F(dt_i)$ shifts toward higher dt_i values for increasing i, i.e., the duration of the time stretch required for a detrapping event grows as the detrapping process proceeds. This effect, which can be clearly observed also monitoring $\langle dt_i\rangle$ as a function of i in Fig. 7, has a twofold origin. First, as the detrapping process proceeds, the number of electrons trapped in the oxide decreases, making the next dt_i the result of the detrapping competition among a lower number of possible detrapping events. Second, as the time at which a detrapping event occurs is on average equal to its τ_d, the τ_d spectrum of Fig. 2 is emptied from left to

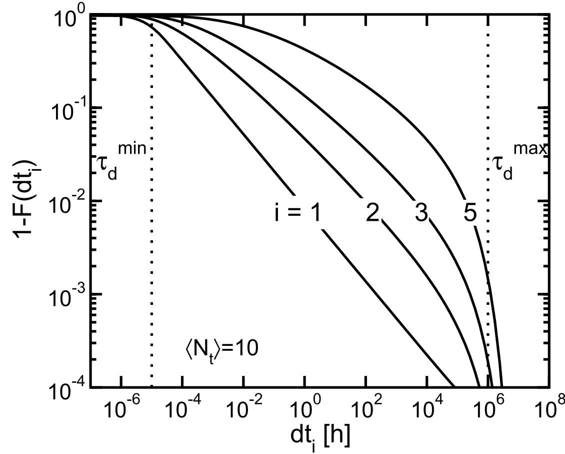

Fig. 6. Calculated complementary cumulative distribution function of the interarrival time dt_i in the case N_t has a Poissonian distribution with average value $\langle N_t \rangle = 10$ and $\log_{10}(\tau_d)$ has a uniform distribution between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h. The results are shown for increasing i.

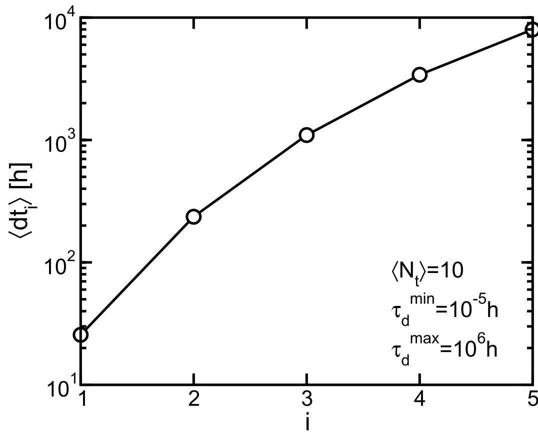

Fig. 7. Calculated $\langle dt_i \rangle$ as a function of i for the case considered in Fig. 6.

right as detrapping proceeds, i.e., the detrapping process is fed by events with longer and longer τ_d as time elapses and i grows. Both these effects contribute to the decrease with time of the rate of the process governing charge detrapping, making the detrapping events more far-between in time. Note, however, that this happens keeping a Poissonian behavior for the detrapping process, thanks to the Poissonian distribution of N_t among the cells and, moreover, to the implicit assumption of no correlation among the detrapping events in our analysis. If any correlation among the detrapping events existed, a sub-Poissonian behavior for the detrapping process could emerge, as clearly observed in the electron injection process toward the cell floating gate during Fowler–Nordheim programming, where the electrostatic feedback following each injection event reduces the probability for next electron injections [34]–[36].

As a final remark, note that if the number of trapped electrons is Poisson distributed among the cells at the beginning of the detrapping process, it remains Poisson distributed as time elapses, with average value equal to $\langle N_t \rangle - \langle n_d \rangle$ (see the results in Appendix B).

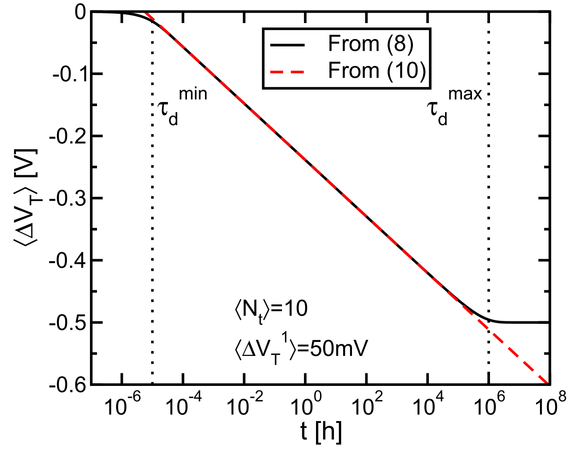

Fig. 8. Calculated $\langle \Delta V_T(t) \rangle$ transient in the case $\langle N_t \rangle = 10$, $\langle \Delta V_T^1 \rangle = 50$ mV, and $\log_{10}(\tau_d)$ uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h. The logarithmic approximation coming from (10) is also shown.

III. ΔV_T Statistics

A. $\langle \Delta V_T(t) \rangle$ and $\sigma^2_{\Delta V_T}(t)$

The neutralization of a single electron charge in the tunnel oxide gives rise to a step reduction of cell V_T, whose amplitude has a large statistical spread in nanoscale technologies [7], [10]–[12], [14], [15], [23]. As a consequence of this spread, the $n_d(t)$ waveform of Fig. 4(a) may transform in the $\Delta V_T(t)$ transient of Fig. 4(b), where the statistical dispersion of the achieved V_T loss after time t comes not only from randomness in the number $n_d(t)$ of detrapping events taking place up to t but also from randomness in the V_T shift coming from a single detrapping event ΔV_T^1. Assuming the ΔV_T^1 of a single detrapping event to be independent both of its τ_d [27], [28] and of the ΔV_T^1 of other detrapping events [7], the average value ($\langle \Delta V_T(t) \rangle$) and variance ($\sigma^2_{\Delta V_T}(t)$) of $\Delta V_T(t)$ are given by (see calculations in Appendix C)

$$\langle \Delta V_T(t) \rangle = -\langle \Delta V_T^1 \rangle \cdot \langle n_d(t) \rangle \quad (8)$$

$$\sigma^2_{\Delta V_T}(t) = \langle \Delta V_T^1 \rangle^2 \cdot \sigma^2_{n_d}(t) + \sigma^2_{\Delta V_T^1} \cdot \langle n_d(t) \rangle. \quad (9)$$

Fig. 8 shows the calculated $\langle \Delta V_T(t) \rangle$ transient in the case of $\langle N_t \rangle = 10$, $\langle \Delta V_T^1 \rangle = 50$ mV, and $\log_{10}(\tau_d)$ uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h. The transient displays a logarithmic trend with t in the range $\tau_d^{\min} < t < \tau_d^{\max}$, coming from the approximation of $F(t)$ given by (4), resulting in

$$\langle \Delta V_T(t) \rangle \simeq -\langle \Delta V_T^1 \rangle \cdot \langle N_t \rangle \cdot \frac{\gamma + \ln\left(t/\tau_d^{\min}\right)}{\ln\left(\tau_d^{\max}/\tau_d^{\min}\right)}. \quad (10)$$

The logarithmic trend well reproduces what typically observed experimentally for $\langle \Delta V_T(t) \rangle$ [24]–[26], [30] (Fig. 1), and can be synthetically characterized by its slope α given by

$$\alpha = \frac{\langle \Delta V_T^1 \rangle \cdot \langle N_t \rangle}{\ln\left(\tau_d^{\max}/\tau_d^{\min}\right)}. \quad (11)$$

From (11), our model directly highlights that the slope α of the $\langle \Delta V_T(t) \rangle$ transient is proportional to the average V_T

shift coming from a single detrapping event and the spectral density of trapped electrons along the logarithmic time axis, i.e., $\langle N_t \rangle / \ln(\tau_d^{\max}/\tau_d^{\min})$. This result establishes a strong connection between the spectral distribution of τ_d given in Fig. 2 and the $\langle \Delta V_T(t) \rangle$ transient, which will be used in Part II of this paper to explain the shape of the detrapping transient in more complex situations where the τ_d spectrum along the logarithmic time axis is not uniform. As a final remark on Fig. 8, note that the displacement of $\langle \Delta V_T(t) \rangle$ from the logarithmic trend coming from its saturation for $t > 10^6$ h is just the result of assuming $\tau_d^{\max} = 10^6$ h, i.e., of assuming no detrapping events with τ_d longer than 10^6 h and, therefore, the neutralization of all of the trapped electrons at shorter times. Note that this is an arbitrary assumption and that the experimental data shown in Part II of this paper [1] do not show saturation of the detrapping transient.

In the case N_t is Poisson distributed, (9) gives a relationship of direct proportionality between $\sigma^2_{\Delta V_T}(t)$ and $\langle \Delta V_T(t) \rangle$

$$\sigma^2_{\Delta V_T}(t) = -\langle \Delta V_T(t) \rangle \cdot \left(\langle \Delta V_T^1 \rangle + \frac{\sigma^2_{\Delta V_T^1}}{\langle \Delta V_T^1 \rangle} \right). \quad (12)$$

This linear relationship reflects the Poisson nature of the detrapping process in this case, with the proportionality constant being not only $\langle \Delta V_T^1 \rangle$ but also including a term depending on $\sigma^2_{\Delta V_T^1}$ to account for the statistical spread of ΔV_T^1. As a result, $\sigma^2_{\Delta V_T}(t)$ has a logarithmic increase with t in the range $\tau_d^{\min} < t < \tau_d^{\max}$, reproducing the $\langle \Delta V_T(t) \rangle$ behavior of Fig. 8 with just a scale factor.

B. Statistical Distribution of ΔV_T

The pdf of $\Delta V_T(t)$ can be calculated as

$$f(\Delta V_T(t)) = \sum_{n_d=0}^{\infty} f(\Delta V_T | n_d) P[n_d(t)] \quad (13)$$

where $f(\Delta V_T | n_d)$ is the probability to have a V_T loss equal to ΔV_T given that n_d detrapping events took place. Assuming that the ΔV_T^1 of the single detrapping events are independent and identically distributed, $f(\Delta V_T | n_d)$ can be obtained from the n_d-times convolution of the pdf of ΔV_T^1 for $n_d \geq 1$ ($f(\Delta V_T | n_d) = \delta(\Delta V_T)$ for $n_d = 0$)

$$f(\Delta V_T | n_d) = \bigotimes^{n_d} f(\Delta V_T^1). \quad (14)$$

Therefore, the Fourier transform of $f(\Delta V_T | n_d)$ with respect to ΔV_T is given by

$$\mathscr{F}\{f(\Delta V_T | n_d)\} = \left[\mathscr{F}\{f(\Delta V_T^1)\}\right]^{n_d} \quad (15)$$

and, therefore, the Fourier transform of $f(\Delta V_T(t))$ results in

$$\mathscr{F}\{f(\Delta V_T(t))\} = \sum_{n_d=0}^{\infty} \left[\mathscr{F}\{f(\Delta V_T^1)\}\right]^{n_d} P[n_d(t)]. \quad (16)$$

Considering now the most interesting case where N_t is Poisson distributed, making, in turn, $n_d(t)$ follow a Poisson statistics

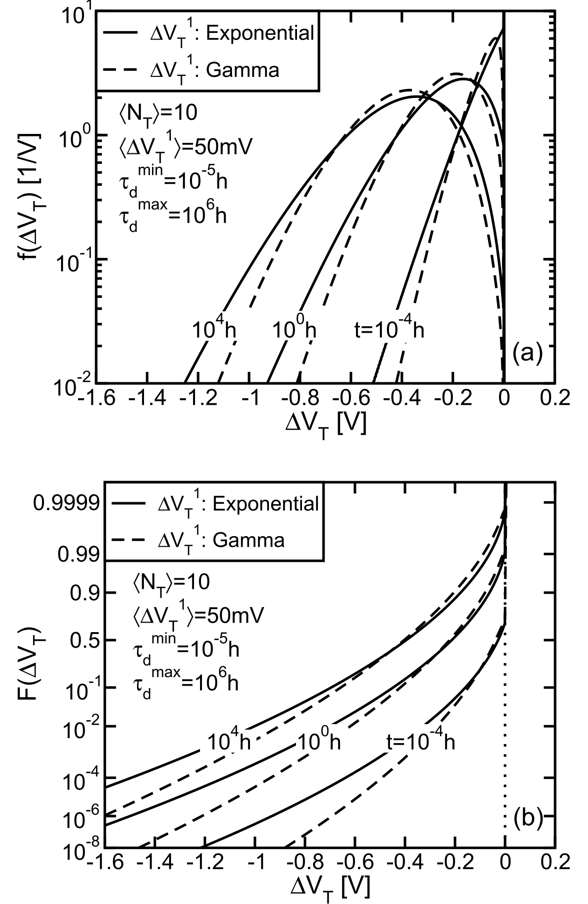

Fig. 9. Calculated (a) pdf and (b) cdf of the ΔV_T resulting from electron detrapping up to time t in the case $\langle N_t \rangle = 10$, $\langle \Delta V_T^1 \rangle = 50$ mV, and $\log_{10}(\tau_d)$ uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h. The results for the exponential and the gamma ΔV_T^1 distributions reported in Fig. 10 are shown.

with average value $\langle n_d(t) \rangle$, (16) becomes

$$\mathscr{F}\{f(\Delta V_T(t))\} = e^{-\langle n_d(t) \rangle} \sum_{n_d=0}^{\infty} \frac{\left[\mathscr{F}\{f(\Delta V_T^1)\} \langle n_d(t) \rangle\right]^{n_d}}{n_d!}$$
$$= e^{\langle n_d(t) \rangle [\mathscr{F}\{f(\Delta V_T^1)\} - 1]} \quad (17)$$

allowing the calculation of $f(\Delta V_T(t))$ by an inverse-transform operation

$$f(\Delta V_T(t)) = \mathscr{F}^{-1}\left\{e^{\langle n_d(t) \rangle [\mathscr{F}\{f(\Delta V_T^1)\} - 1]}\right\}. \quad (18)$$

Equation (18) represents an extremely powerful formula for the statistical analysis of V_T instabilities coming from electron detrapping, rigorously accounting for a Poissonian spread of N_t among the cells and for any kind of τ_d and of ΔV_T^1 statistical dispersion, with the only requirement that detrapping events are independent. From the standpoint of computational burdens, this equation greatly simplifies the calculation of the pdf of ΔV_T with respect to the use of (13) and (14), requiring basically just one direct and one inverse Fourier transform operation. In so doing, the whole pdf and, in turn, cdf of ΔV_T can be easily calculated down to very low probability levels, as shown in Fig. 9. Here, $f(\Delta V_T(t))$ (a) and $F(\Delta V_T(t))$ (b) are shown in the case of $\langle N_t \rangle = 10$,

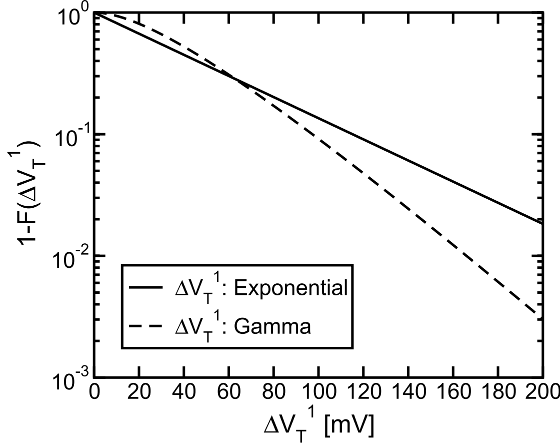

Fig. 10. Exponential and gamma ΔV_T^1 statistics assumed to investigate ΔV_T in Fig. 9.

$\langle \Delta V_T^1 \rangle = 50$ mV, and $\log_{10}(\tau_d)$ uniformly distributed between $\tau_d^{\min} = 10^{-5}$ h and $\tau_d^{\max} = 10^6$ h, for different times t since the beginning of the detrapping process. The results have been obtained assuming the two ΔV_T^1 statistics shown in Fig. 10, being a purely exponential distribution (continuous line) and a gamma distribution with shape and scale parameters equal to 2 and 25 mV, respectively (dashed line). While exponential distributions have been clearly reported for ΔV_T^1 in the case of electrons trapped close to the channel surface of decananometer devices [7], [10], [11], [14], [15], [23], the gamma distribution is more suitable when electrons are stored far from the silicon interface [3], [9] and in nanoscale devices [4]. In both cases, analytical expressions exist for $\mathscr{F}\{f(\Delta V_T^1)\}$, further simplifying the use of (18). Moreover, in the case of an exponential ΔV_T^1 distribution, $f(\Delta V_T(t))$ has the following analytical expression, derived in Appendix D:

$$f(\Delta V_T(t)) = e^{-\langle n_d(t)\rangle} \left[\delta(\Delta V_T) + e^{-\Delta V_T/\langle \Delta V_T^1 \rangle} \cdot \sqrt{\frac{\langle n_d(t)\rangle}{\langle \Delta V_T^1\rangle \Delta V_T}} \cdot I_1\left(2\sqrt{\frac{\langle n_d(t)\rangle \Delta V_T}{\langle \Delta V_T^1\rangle}}\right)\right]. \quad (19)$$

The results in Fig. 9 clearly reveal that the ΔV_T distribution enlarges and shifts toward negative values as time elapses. Moreover, as the exponential and the gamma statistics for ΔV_T^1 shown in Fig. 10 have the same $\langle \Delta V_T^1 \rangle = 50$ mV but different $\sigma^2_{\Delta V_T^1}$, the corresponding ΔV_T curves of Fig. 9 display the same average shift as time elapses but different enlargements. In particular, curves coming from the exponential ΔV_T^1 statistics (continuous lines) are wider than those coming from the gamma statistics (dashed lines), as $\sigma^2_{\Delta V_T^1}$ is equal to 2500 mV^2 in the first case and to 1250 mV^2 in the second. In addition to that, the low-ΔV_T tails of the distributions of Fig. 9 are higher in the exponential than in the gamma case, reflecting the higher tail of the exponential ΔV_T^1 statistics of Fig. 10. These results highlight the extreme importance of the ΔV_T^1 statistics on the ΔV_T distributions coming from electron detrapping in nanoscale technologies, where not only the statistical nature of the detrapping process but also the statistical spread of the V_T shift coming from each detrapping event must be accounted for a careful reliability analysis.

As a final remark on the distributions shown in Fig. 9, note that these are bounded to the negative ΔV_T half-plane, meaning that (18) does not allow for positive ΔV_T values. Although this is an obvious result considering that (18) describes the consequences of a process leading to the neutralization of the electron charge trapped in the cell tunnel oxide, it calls for the inclusion of other physical phenomena when comparing the modeling results with the experimental data. A lot of experimental evidence has clearly highlighted, in fact, that, despite cells display on average a negative ΔV_T during postcycling data retention experiments from the programmed state, some of them may increase their V_T [12], [25]. Positive ΔV_T have been attributed to hole detrapping, i.e., to the neutralization of a positive charge in the tunnel oxide, to electrons trapped close to the source or drain junction being detrapped toward the floating gate, and to RTN [12], [25]. Without the inclusion of these effects, or at least of some of them, the comparison of the results coming from (18) with the experimental data is meaningless, as meaningless is the comparison without a correct analysis of the experimental procedure used to characterize post-cycling V_T instabilities. These points are dealt with in detail in Part II of this paper.

IV. Conclusion

This paper presented a discrete and statistical reexamination of the electron detrapping process and of the consequent V_T instabilities in Flash memories. We showed that in the most interesting case where the number of trapped electrons feeding the detrapping process is Poisson distributed among the cells, detrapping events are the result of a nonhomogeneous Poisson process and a simple and powerful formula [namely, (18)] allows the calculation of the ΔV_T statistics. These results, which can be applied also to different reliability phenomena involving charge trapping/detrapping in MOS devices, pave the way to the development of a comprehensive statistical model able to deal with V_T instabilities under whatever on-field usage of the memory array in Part II of this paper, where detailed comparisons with the experimental data will be provided.

Appendix A

For an arbitrary N_t statistics, $\langle n_d(t)\rangle$ results

$$\begin{aligned}\langle n_d(t)\rangle &= \sum_{n_d=0}^{\infty} n_d \cdot P[n_d(t)] \\ &= \sum_{n_d=0}^{\infty} n_d \cdot \sum_{N_t=0}^{\infty} P[n_d(t)\,|N_t] \cdot P[N_t] \\ &= \sum_{N_t=0}^{\infty} P[N_t] \cdot \sum_{n_d=0}^{N_t} n_d P[n_d(t)\,|N_t] \\ &= \sum_{N_t=0}^{\infty} P[N_t] \cdot N_t F(t) \\ &= \langle N_t\rangle \cdot F(t)\end{aligned}$$

and $\sigma_{n_d}^2(t)$ is given by

$$\begin{aligned}
\sigma_{n_d}^2(t) &= \sum_{n_d=0}^{\infty} [n_d - \langle n_d(t)\rangle]^2 \cdot P[n_d(t)] \\
&= \sum_{n_d=0}^{\infty} [n_d - \langle n_d(t)\rangle]^2 \cdot \sum_{N_t=0}^{\infty} P[n_d(t)|N_t]\, P[N_t] \\
&= \sum_{N_t=0}^{\infty} P[N_t] \\
&\quad \cdot \sum_{n_d=0}^{N_t} \left[n_d^2 + \langle n_d(t)\rangle^2 - 2n_d\langle n_d(t)\rangle\right] P[n_d(t)|N_t] \\
&= \sum_{N_t=0}^{\infty} P[N_t] \cdot \Big[N_t F(t)(1-F(t)) + N_t^2 F(t)^2 \\
&\qquad + \langle N_t\rangle^2 F(t)^2 - 2\langle N_t\rangle N_t F(t)^2\Big] \\
&= \langle N_t\rangle \cdot F(t)[1-F(t)] + \sigma_{N_t}^2 \cdot F(t)^2 .
\end{aligned}$$

Appendix B

Assuming that N_t is Poisson distributed, we have

$$\begin{aligned}
P[n_d(t)] &= \sum_{N_t=n_d}^{\infty} P[N_t]\, P[n_d(t)|N_t] \\
&= \sum_{N_t=n_d}^{\infty} P[N_t] \binom{N_t}{n_d} F(t)^{n_d} [1-F(t)]^{N_t-n_d} \\
&= \frac{e^{-\langle N_t\rangle}}{n_d!} F(t)^{n_d} \sum_{N_t=n_d}^{\infty} \frac{\langle N_t\rangle^{N_t} [1-F(t)]^{N_t-n_d}}{(N_t-n_d)!} \\
&= \frac{[\langle N_t\rangle F(t)]^{n_d}}{n_d!} e^{-\langle N_t\rangle F(t)}
\end{aligned}$$

and therefore, $n_d(t)$ is Poisson distributed with average value $\langle n_d(t)\rangle = \langle N_t\rangle F(t)$. Moreover, the probability to have N_t' trapped electrons at time t when N_t electrons were trapped at time 0 is given by

$$\begin{aligned}
&P\left[N_t'(t)\right] \\
&= \sum_{N_t=N_t'}^{\infty} P\left[N_t'(t)|N_t\right] P[N_t] \\
&= \sum_{N_t=N_t'}^{\infty} \binom{N_t}{N_t-N_t'} F(t)^{N_t-N_t'} \cdot [1-F(t)]^{N_t'} \frac{\langle N_t\rangle^{N_t}}{N_t!} e^{-\langle N_t\rangle} \\
&= \frac{[1-F(t)]^{N_t'} \cdot F(t)^{-N_t'} \cdot e^{-\langle N_t\rangle}}{N_t'!} \cdot \sum_{N_t=N_t'}^{\infty} \frac{F(t)^{N_t}}{(N_t-N_t')!} \langle N_t\rangle^{N_t} \\
&= \frac{[1-F(t)]^{N_t'} \cdot e^{-\langle N_t\rangle}}{N_t'! \cdot \langle N_t\rangle^{-N_t'}} \cdot e^{\langle N_t\rangle F(t)} \\
&= \frac{[\langle N_t\rangle (1-F(t))]^{N_t'}}{N_t'!} \cdot e^{-\langle N_t\rangle[1-F(t)]}
\end{aligned}$$

and, therefore, $N_t'(t)$ is Poisson distributed with average value $\langle N_t\rangle[1-F(t)] = \langle N_t\rangle - \langle n_d(t)\rangle$.

Appendix C

For arbitrary N_t and ΔV_T^1 statistics and assuming the ΔV_T^1 of a single detrapping event to be independent of both its τ_d and the ΔV_T^1 of other detrapping events, $\langle \Delta V_T(t)\rangle$ can be calculated as

$$\begin{aligned}
\langle \Delta V_T(t)\rangle &= \int_0^{\infty} \Delta V_T \sum_{n_d=0}^{\infty} f(\Delta V_T|n_d)\, P[n_d(t)]\, d\Delta V_T \\
&= -\langle \Delta V_T^1\rangle \cdot \sum_{n_d=0}^{\infty} n_d\; P[n_d(t)] \\
&= -\langle \Delta V_T^1\rangle \cdot \langle n_d(t)\rangle
\end{aligned}$$

where $f(\Delta V_T|n_d)$ is the probability to have a V_T loss ΔV_T given that n_d detrapping events took place. Similarly, $\sigma_{\Delta V_T}^2(t)$ can be calculated as

$$\begin{aligned}
\sigma_{\Delta V_T}^2(t) &= \int_0^{\infty} [\Delta V_T - \langle \Delta V_T(t)\rangle]^2 \\
&\qquad \cdot \sum_{n_d=0}^{\infty} f(\Delta V_T|n_d)\, P[n_d(t)]\, d\Delta V_T \\
&= \sum_{n_d=0}^{\infty} P[n_d(t)] \cdot \Big[n_d \sigma_{\Delta V_T^1}^2 + \langle \Delta V_T^1\rangle^2 n_d^2 \\
&\qquad + \langle \Delta V_T(t)\rangle^2 + 2\langle \Delta V_T(t)\rangle n_d \langle \Delta V_T^1\rangle\Big] \\
&= \sigma_{\Delta V_T^1}^2 \langle n_d(t)\rangle + \langle \Delta V_T^1\rangle^2 \left[\sigma_{n_d}^2(t) + \langle n_d(t)\rangle^2\right] \\
&\quad + \langle \Delta V_T(t)\rangle^2 + 2\langle \Delta V_T(t)\rangle \langle \Delta V_T^1\rangle \langle n_d(t)\rangle \\
&= \langle \Delta V_T^1\rangle^2 \cdot \sigma_{n_d}^2(t) + \sigma_{\Delta V_T^1}^2 \cdot \langle n_d(t)\rangle.
\end{aligned}$$

Appendix D

The pdf of $\Delta V_T(t)$ can be calculated as

$$f(\Delta V_T(t)) = \sum_{n_d=0}^{\infty} f(\Delta V_T|n_d)\, P[n_d(t)] .$$

Assuming an exponential distribution for ΔV_T^1, $f(\Delta V_T|n_d)$ is given by

$$f(\Delta V_T|n_d) = \frac{e^{-\Delta V_T/\langle \Delta V_T^1\rangle} \left(\Delta V_T/\langle \Delta V_T^1\rangle\right)^{n_d}}{\Delta V_T\,(n_d-1)!}$$

for $n_d \geq 1$, while $f(\Delta V_T|n_d) = \delta(\Delta V_T)$ for $n_d = 0$. Considering now N_t as Poisson distributed, $f(\Delta V_T(t))$ becomes

$$\begin{aligned}
&f(\Delta V_T(t)) \\
&\quad = e^{-\langle n_d(t)\rangle} \cdot \Bigg[\delta(\Delta V_T) + e^{-\Delta V_T/\langle \Delta V_T^1\rangle} \frac{\langle n_d(t)\rangle}{\langle \Delta V_T^1\rangle} \\
&\qquad\qquad \cdot \sum_{n_d=1}^{\infty} \frac{\left(\langle n_d(t)\rangle \Delta V_T/\langle \Delta V_T^1\rangle\right)^{n_d-1}}{n_d!\,(n_d-1)!}\Bigg] \\
&\quad = e^{-\langle n_d(t)\rangle} \Bigg[\delta(\Delta V_T) + e^{-\Delta V_T/\langle \Delta V_T^1\rangle} \\
&\qquad\qquad \cdot \sqrt{\frac{\langle n_d(t)\rangle}{\langle \Delta V_T^1\rangle \Delta V_T}} \cdot I_1\left(2\sqrt{\frac{\langle n_d(t)\rangle \Delta V_T}{\langle \Delta V_T^1\rangle}}\right)\Bigg]
\end{aligned}$$

where I_1 is the modified Bessel function of the first kind.

References

[1] G. M. Paolucci, C. M. Compagnoni, C. Miccoli, A. S. Spinelli, A. L. Lacaita, and A. Visconti, "Revisiting charge trapping/detrapping in Flash memories from a discrete and statistical standpoint—Part II: On-field operation and distributed-cycling effects," *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2811–2819, Aug. 2014.

[2] K. Prall and K. Parat, "25 nm 64 Gb MLC NAND technology and scaling challenges," in *Proc. IEEE IEDM*, Dec. 2010, pp. 5.2.1–5.2.4.

[3] C. Monzio Compagnoni *et al.*, "Comprehensive investigation of statistical effects in nitride memories—Part II: Scaling analysis and impact on device performance," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2124–2131, Sep. 2010.

[4] S. M. Amoroso *et al.*, "Investigation of the RTN distribution of nanoscale MOS devices from subthreshold to on-state," *IEEE Electron Device Lett.*, vol. 34, no. 5, pp. 683–685, May 2013.

[5] P. Fantini, A. Ghetti, A. Marinoni, G. Ghidini, A. Visconti, and A. Marmiroli, "Giant random telegraph signals in nanoscale floating-gate devices," *IEEE Electron Device Lett.*, vol. 28, no. 12, pp. 1114–1116, Dec. 2007.

[6] A. S. Spinelli, C. Monzio Compagnoni, R. Gusmeroli, M. Ghidotti, and A. Visconti, "Investigation of the random telegraph noise instability in scaled Flash memory arrays," *Jpn. J. Appl. Phys.*, vol. 47, no. 4S, pp. 2598–2601, 2008.

[7] A. Ghetti, C. Monzio 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.

[8] T. Nagumo, K. Takeuchi, S. Yokogawa, K. Imai, and Y. Hayashi, "New analysis methods for comprehensive understanding of random telegraph noise," in *Proc. IEEE IEDM*, Dec. 2009, pp. 759–762.

[9] A. Mauri *et al.*, "Comprehensive investigation of statistical effects in nitride memories—Part I: Physics-based modeling," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2116–2123, Sep. 2010.

[10] J. Franco *et al.*, "Impact of single charged gate oxide defects on the performance and scaling of nanoscaled FETs," in *Proc. IEEE IRPS*, Apr. 2012, pp. 5A.4.1–5A.4.6.

[11] B. Kaczer *et al.*, "The relevance of deeply-scaled FET threshold voltage shifts for operation lifetimes," in *Proc. IEEE IRPS*, Apr. 2012, pp. 5A.2.1–5A.2.6.

[12] C. Miccoli *et al.*, "Resolving discrete emission events: A new perspective for detrapping investigation in NAND Flash memories," in *Proc. IEEE IRPS*, Apr. 2013, pp. 3B.1.1–3B.1.6.

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

[14] A. Ghetti, M. Bonanomi, C. Monzio Compagnoni, A. S. Spinelli, A. L. Lacaita, and A. Visconti, "Physical modeling of single-trap RTS statistical distribution in flash memories," in *Proc. IEEE IRPS*, May 2008, pp. 610–615.

[15] A. Ghetti *et al.*, "Scaling trends for random telegraph noise in deca-nanometer Flash memories," in *Proc. IEEE IEDM*, Dec. 2008, pp. 835–838.

[16] H. Kurata *et al.*, "Random telegraph signal in flash memory: Its impact on scaling of multilevel flash memory beyond the 90-nm node," *IEEE J. Solid-State Circuits*, vol. 42, no. 6, pp. 1362–1369, Jun. 2007.

[17] C. Monzio Compagnoni, R. Gusmeroli, A. S. Spinelli, A. L. Lacaita, M. Bonanomi, and A. Visconti, "Statistical model for random telegraph noise in Flash memories," *IEEE Trans. Electron Devices*, vol. 55, no. 1, pp. 388–395, Jan. 2008.

[18] J. P. Campbell *et al.*, "Random telegraph noise in highly scaled nMOSFETs," in *Proc. IEEE IRPS*, Apr. 2009, pp. 382–388.

[19] J. P. Campbell *et al.*, "Large random telegraph noise in sub-threshold operation of nano-scale nMOSFETs," in *Proc. IEEE ICICDT*, May 2009, pp. 17–20.

[20] T. Nagumo, K. Takeuchi, T. Hase, and Y. Hayashi, "Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps," in *Proc. IEEE IEDM*, Dec. 2010, pp. 628–631.

[21] T. Grasser *et al.*, "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.

[22] G. Pobegen and T. Grasser, "On the distribution of NBTI time constants on a long, temperature-accelerated time scale," *IEEE Trans. Electron Devices*, vol. 60, no. 7, pp. 2148–2155, Jul. 2013.

[23] K. Fukuda, Y. Shimizu, K. Amemiya, M. Kamoshida, and C. Hu, "Random telegraph noise in flash memories—Model and technology scaling," in *Proc. IEEE IEDM*, Dec. 2007, pp. 169–172.

[24] R. Yamada, Y. Mori, Y. Okuyama, J. Yugami, T. Nishimoto, and H. Kume, "Analysis of detrap current due to oxide traps to improve flash memory retention," in *Proc. 38th Annu. IEEE IRPS*, 2000, pp. 200–204.

[25] N. Mielke *et al.*, "Flash EEPROM threshold instabilities due to charge trapping during program/erase cycling," *IEEE Trans. Device Mater. Rel.*, vol. 4, no. 3, pp. 335–344, Sep. 2004.

[26] N. Mielke, H. Belgal, A. Fazio, Q. Meng, and N. Righos, "Recovery effects in the distributed cycling of Flash memories," in *Proc. 44th Annu. IEEE IRPS*, Mar. 2006, pp. 29–35.

[27] N. Castellani, C. Monzio Compagnoni, A. Mauri, A. S. Spinelli, and A. L. Lacaita, "Three-dimensional electrostatics- and atomistic doping-induced variability of RTN time constants in nanoscale MOS devices—Part I: Physical investigation," *IEEE Trans. Electron Devices*, vol. 59, no. 9, pp. 2488–2494, Sep. 2012.

[28] C. Monzio Compagnoni, N. Castellani, A. Mauri, A. S. Spinelli, and A. L. Lacaita, "Three-dimensional electrostatics- and atomistic doping-induced variability of RTN time constants in nanoscale MOS devices—Part II: Spectroscopic implications," *IEEE Trans. Electron Devices*, vol. 59, no. 9, pp. 2495–2500, Sep. 2012.

[29] G. M. Paolucci *et al.*, "A new spectral approach to modeling charge trapping/detrapping in NAND Flash memories," in *Proc. IRPS*, 2014.

[30] C. Miccoli, C. Monzio Compagnoni, S. Beltrami, A. S. Spinelli, and A. Visconti, "Threshold-voltage instability due to damage recovery in nanoscale NAND Flash memories," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2406–2414, Aug. 2011.

[31] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. New York, NY, USA: Dover, 1964.

[32] P. A. W. Lewis and G. S. Shedler, "Simulation of nonhomogeneous Poisson processes by thinning," *Naval Res. Logistics Quart.*, vol. 26, no. 3, pp. 403–413, 1979.

[33] L. A. Baxter, "Reliability applications of the relevation transform," *Naval Res. Logistics*, vol. 29, pp. 323–330, Jun. 1982.

[34] C. Monzio Compagnoni *et al.*, "First evidence for injection statistics accuracy limitations in NAND Flash constant-current Fowler–Nordheim programming," in *Proc. IEEE IEDM*, Dec. 2007, pp. 165–168.

[35] C. Monzio Compagnoni, A. S. Spinelli, R. Gusmeroli, S. Beltrami, A. Ghetti, and A. Visconti, "Ultimate accuracy for the NAND flash program algorithm due to the electron injection statistics," *IEEE Trans. Electron Devices*, vol. 55, no. 10, pp. 2695–2702, Oct. 2008.

[36] C. Monzio Compagnoni, R. Gusmeroli, A. S. Spinelli, and A. Visconti, "Analytical model for the electron-injection statistics during programming of nanoscale NAND flash memories," *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 3192–3199, Nov. 2008.