Solid-State Electronics 56 (2011) 23–30

Contents lists available at ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

Comprehensive numerical simulation of threshold-voltage transients in nitride memories

Aurelio Mauri [a,*], Salvatore M. Amoroso [b,c], Christian Monzio Compagnoni [b,c], Alessandro Maconi [b,c], Alessandro S. Spinelli [b,c,d]

[a] Micron, R&D – Technology Development, 20041 Agrate Brianza (MI), Italy
[b] Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milano, Italy
[c] IU.NET, 20133 Milano, Italy
[d] Istituto di Fotonica e Nanotecnologie-Consiglio Nazionale delle Ricerche, 20133 Milano, Italy

ARTICLE INFO

Article history:
Received 16 March 2010
Received in revised form 5 August 2010
Accepted 1 November 2010
Available online 26 November 2010

The review of this paper was arranged by Prof. S. Cristoloveanu

Keywords:
Non volatile memories
Charge-trap memories
Tunneling
Semiconductor device modeling

ABSTRACT

This paper presents a comprehensive numerical modeling for the threshold-voltage transients of nitride-based memory devices during programming, erasing and data retention. The developed numerical tool self-consistently solves the Poisson, continuity and trapping equations in the nitride layer using a drift-diffusion formalism. The continuity equation has been discretized using the Scharfetter–Gummel scheme and a modified Gummel-map has been optimized to ensure fully convergence of the equations. The numerical model is able to describe the memory device operation for different gate bias regimes, therefore addressing both the program/erase and the retention conditions. Finally, numerical results are shown to carefully reproduce experimental data on template devices with different gate stack compositions, validating the physical assumptions and making the model a valuable tool for nitride memories investigation and design.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Charge-trap (CT) memory devices are the most promising alternative to the floating-gate technology in the NAND architecture.

The (poly)Silicon/Oxide/Nitride/Oxide/Silicon (SONOS) gate stack was one of the most successful implemented structures, but showed poor performances due to the compromise between data retention and program/erase (P/E) window (requiring a thin tunnel oxide to avoid program and erase saturation) [1].

The TaN/Al_2O_3/Nitride/Oxide/Silicon (TANOS) structure has been proposed to solve this compromise, featuring a thicker bottom oxide in the range of 4 nm for long data retention and disturb immunity and high work-function metal gate with high-k top dielectric for reduced erase saturation [2]. Barrier-engineered tunnel applied to TANOS (BET-TANOS) has been proposed in order to speed up erase transients with the help of the increased hole current through the ONO tunneling barrier [3].

In the recent years several modeling approach has been developed to describe charge traps devices. Most of them treat electron transport in nitride in approximated way [4,5] or cover only some aspects of the CT functionality [6] or are not completely fully consistent [7]. Only [8] introduces a self-consistent drift-diffusion solution for electron transport and trapping in the nitride, showing the impact of the electron transport on program/retention transients but neglecting the hole contribution. On the other hand hole transport, trapping and recombination is important when dealing with modeling of BET-TANOS where the behavior (erase) is completely ruled by hole injection from the substrate.

In this paper we present a comprehensive numerical model able to predict the threshold voltage (V_T) transients of SONOS, TANOS and BET-TANOS memory devices for different operating regimes, including the detailed description of electron and hole transport and trapping/detrapping in the nitride layer. To this aim, we have applied the semiconductor formalism to describe the carrier dynamics in the nitride, including the transport equations and two (one for each carrier type) new sets of equations taking into consideration the trapping/detrapping mechanisms that influence the carrier flow in the conduction and valence bands. The continuity equation has been discretized using the Scharfetter–Gummel scheme [9] and a modified Gummel-map [10] has been optimized to ensure fully convergence of the equations. The field dependence of the carrier mobilities is also taken into account.

The paper is organised as follows: in Section 2 the mathematical model is discussed, with subsections dedicated to the boundary conditions, the adopted mobility model and the emissivity models

* Corresponding author.
E-mail address: aureliogiance@micron.com (A. Mauri).

0038-1101/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sse.2010.11.004

for charge detrapping. The numerical implementation and the convergence scheme are presented in Section 3. Section 4 is dedicated to the discussion of the simulation results and their comparison with experimental data on different SONOS, TANOS and BET-TANOS devices. Conclusions are summarised in Section 5.

2. Numerical model

2.1. Basic equations

The equations implemented for modeling the V_T transients in charge trapping devices are a modification of the Arnett's system presented in [11], where we have added the contribution of holes in presence of amphoteric traps [13]. This results in highly coupled systems where appropriate decoupling scheme must considered as explained in Section 3. The system is:

$$\frac{\partial^2\psi(x,t)}{\partial x^2} = \frac{q}{\varepsilon_N}(n_c(x,t) - p_v(x,t) + n_t(x,t) - p_t(x,t)), \tag{1a}$$

$$\frac{\partial n_c(x,t)}{\partial t} = \frac{1}{q}\frac{\partial J_n(x,t)}{\partial x} - R_n, \tag{1b}$$

$$\frac{\partial p_v(x,t)}{\partial t} = -\frac{1}{q}\frac{\partial J_p(x,t)}{\partial x} - R_p, \tag{1c}$$

$$\frac{\partial n_t(x,t)}{\partial t} = \frac{J_n(x,t)}{q}\sigma_n[N_t - n_t(x,t) - p_t(x,t)] - e_n(x,t)n_t(x,t) - \frac{J_p(x,t)}{q}\sigma_{rn}n_t(x,t), \tag{1d}$$

$$\frac{\partial p_t(x,t)}{\partial t} = \frac{J_p(x,t)}{q}\sigma_p[N_t - p_t(x,t) - n_t(x,t)] - e_p(x,t)p_t(x,t) - \frac{J_n(x,t)}{q}\sigma_{rp}p_t(x,t). \tag{1e}$$

In the previous system, q is the electronic charge, ϵ_N is the nitride dielectric constant, N_t, n_t and p_t are the volume densities of traps and of trapped electrons and holes in the nitride, n_c and p_v are the free electron and hole volume densities in the nitride conduction and valence bands, e_n and e_p are the total electron and hole emissivities from the traps (units: s^{-1}) as resulting from the different physical mechanisms described in Section 2.4. In order to account for the different trapping rates of amphoteric traps in their neutral, positive-and negatively-charged state, four different trapping cross-sections (units: cm^2) have been used: σ_n and σ_p are the electron and hole capture cross-sections for neutral traps, while σ_{rn} and σ_{rp} are their counterparts in the case of negative-and positive-charged traps, respectively. Finally, R_n and R_p (units: $cm^{-3}\,s^{-1}$) describe the net electron and hole recombination rates due to carrier trapping/detrapping in presence of the electron and hole currents J_n and J_p that replace the classical SRH terms in a general semiconductor:

$$R_n = \frac{\partial n_t(x,t)}{\partial t} + \frac{J_n(x,t)}{q}\sigma_{rp}p_t(x,t) + \frac{J_p(x,t)}{q}\sigma_{rn}n_t(x,t), \tag{2a}$$

$$R_p = \frac{\partial p_t(x,t)}{\partial t} + \frac{J_p(x,t)}{q}\sigma_{rn}n_t(x,t) + \frac{J_n(x,t)}{q}\sigma_{rp}p_t(x,t). \tag{2b}$$

The system (1) includes the Poisson equation in the nitride, with electrostatic contributions of both free and trapped charges, and the continuity equations for electrons and holes in the conduction and valence bands. The last two equations of (1) describe electron and hole trapping and detrapping in presence of amphoteric traps. Note that, due to the amphoteric assumption, the trap density N_t is unique for both carriers (i.e., both the carriers interacts with the same physical centre). In addition to these equations, the hole and electron currents are calculated with the standard drift-diffusion formalism:

$$J_n(x,t) = qn_c(x,t)\mu_nF(x,t) + qD_n\frac{\partial n_c(x,t)}{\partial x}, \tag{3a}$$

$$J_p(x,t) = qp_v(x,t)\mu_pF(x,t) - qD_p\frac{\partial p_v(x,t)}{\partial x}, \tag{3b}$$

where F is the electric field, μ_n and μ_p are the electron and hole mobilities and D_n and D_p are the diffusion coefficients calculated by means of the Einstein relation ($D/\mu = k_BT/q$), being k_B the Boltzmann constant and T the absolute temperature. At high electric fields a modified Einstein relation should be used for the mobility–diffusion relation (e.g. Arora relation [14] and references therein). However, due to the negligible impact of diffusion process at high fields, this correction can be neglected.

2.2. Boundary conditions

To ensure the correct solution of system (1), appropriate boundary and initial conditions must be used for the program, erase and retention regimes, to account for the tunneling conduction of the dielectrics (open boundary condition). To this aim, we consider a one-dimensional nitride cell structure with the gate placed on the left side and a nitride layer extending from $x = 0$ (interface with top dielectric) to $x = x_N$ (interface with bottom dielectric), schematically shown in Fig. 1 under program conditions and in Fig. 2 for erase. In both figures we have drawn all the current fluxes involved in the dynamics: from bottom and top of the gate stack, from the traps towards gate and from/to conduction/valence band.

2.2.1. Program

Appropriate boundary conditions are given by the electron tunneling currents J_e^{BOT} flowing from the substrate to the nitride via the bottom dielectric and the tunneling conduction of the top dielectric. For electron injection, we considered the electron energy relaxation (see below) which forces a "non-classical" boundary conditions not applied at $x = x_N$ but rather at the relaxation coordinate x_r. The boundary conditions then become:

$$\begin{aligned} &J_n(x_r,t) = J_e^{BOT} + qn_c(x_r,t)\mu_nF(x_r,t) + qD_n\frac{\partial n_c(x_r,t)}{\partial x},\\ &J_p(x_N,t) = 0,\\ &J_n(0,t) = qn_c(0,t)v_{th}TC_e^{TOP},\\ &J_p(0,t) = 0,\\ &n_c(x,0) = p_v(x,0) = 0,\\ &n_t(x,0) = p_t(x,0) = 0, \end{aligned} \tag{4}$$

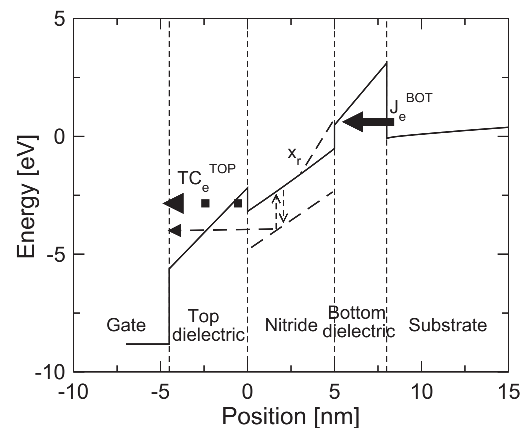

Fig. 1. Calculated nitride memory band profile under program conditions for a generic structure considered in this work. The electron injection and emission currents at the nitride edges are also shown.

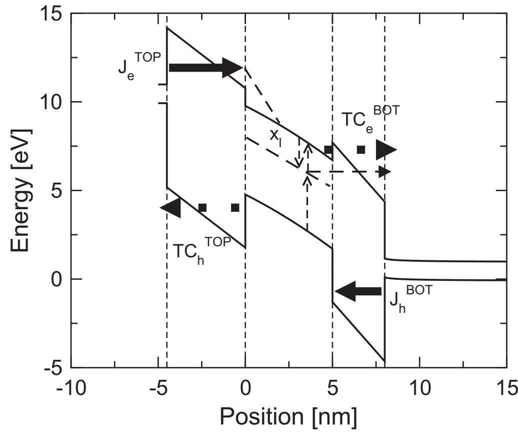

Fig. 2. Same as Fig. 1 but referred to the erase condition. The hole injection and emission currents are also shown.

where v_{th} is the thermal velocity and TC_e^{TOP} is the tunneling transmission coefficient of conduction-band electrons through the top dielectric, calculated with the Wentzel–Kramers–Brillouin (WKB) approximation. The first of the previous boundary conditions accounts for the tunneling current contribution to the drift-diffusion term. This term is evaluated at position $x = x_r$, representing the energy relaxation point of the injected electrons [15]. The relaxation mechanism has been proposed in [15] following the works by [16,17] on SiO_2. It is worth pointing out that, due to the presence of a relaxation length, thinner (than x_r) nitride layers should not exhibit trapping. This is confirmed by many experimental results of different research groups [18–20]. Finally, the boundary condition applied to the hole current represents also a good approximation for SONOS-type devices, having a rather large top oxide thickness (e.g., 6–8 nm) and therefore negligible hole injection from the gate to the nitride (for a typical SONOS devices the hole current coming from gate is 4–5 order of magnitude lower than the electron one).

2.2.2. Erase

The band diagram is shown in Fig. 2: the bottom interface is characterized by both hole injection from the substrate (J_h^{BOT}) and electron emission from traps directly in the substrate or through conduction band; electron injection (J_e^{TOP}) and possible hole emission must be accounted for the gate interface. The thermalization position for electrons injected from the gate is now called x_l, while zero-length thermalization was assumed for holes.

As a consequence, the boundary and initial conditions are:

$$\begin{aligned}
&J_n(x_l,t) = J_e^{TOP} + qn_c(x_l,t)\mu_n F(x_l,t) + qD_n\frac{\partial n_c(x_l,t)}{\partial x},\\
&J_p(0,t) = qp_v(0,t)v_{th}TC_h^{TOP},\\
&J_n(x_N,t) = qn_c(x_N,t)v_{th}TC_e^{BOT},\\
&J_p(x_N,t) = J_h^{BOT} + qp_v(x_N,t)\mu_p F(x_N,t) + qD_p\frac{\partial p_v(x_N,t)}{\partial x},\\
&n_c(x,0) = p_v(x,0) = 0,\\
&n_t(x,0) = n_0(x,0),\\
&p_t(x,0) = p_0(x,0).
\end{aligned} \tag{5}$$

Here, TC_h^{TOP} and TC_e^{BOT} are the tunneling transmission coefficients through the top and bottom dielectrics for valence band holes and conduction-band electrons, respectively. Moreover, $n_0(x,0)$ and $p_0(x,0)$ represent the trapped electron and hole concentrations in the nitride at the beginning of the erase as obtained at the end of programming.

2.2.3. Retention

Assuming negative/positive charge trapped in the nitride programmed/erased cell, the boundary conditions set an electron flux from the nitride conduction band (or from the gate) to the substrate and gate (or to nitride traps) and hole flux in the opposite direction.

$$\begin{aligned}
&J_n(0,t) = qn_c(0,t)v_{th}TC_e^{TOP},\\
&J_p(0,t) = J_h^{TOP} + qp_v(0,t)\mu_p F(0,t),\\
&J_n(x_N,t) = qn_c(x_N,t)v_{th}TC_e^{BOT},\\
&J_p(x_N,t) = J_h^{BOT} + qp_v(x_N,t)\mu_p F(x_N,t).
\end{aligned} \tag{6}$$

As initial condition we assumed zero free charge in the nitride bands and a trapped charge that reflects the distribution at the end of a programming/erasing step.

In addition to the previous conditions, the tunnel out flux from the traps towards the gate and/or the substrate is self-consistently taken in account for program/erase/retention operating regimes (Figs. 1 and 2).

2.3. Mobility model

Electron and hole mobilities in the nitride conduction and valence bands have been described by the following formula:

$$\mu(F,T) = \frac{\mu_0}{\left[1+\left(\frac{\mu_0 F}{v_{sat}}\right)^{\alpha}\right]^{1/\alpha}}, \tag{7}$$

where μ_0 is the low-field mobility of the carriers, $v_{sat} = 1 \times 10^7 \cdot (300K/T)$ cm/s is the velocity saturation value and $\alpha = 1.109 \cdot (T/300K)^{\beta}$, with $\beta = 0.66$. The previous equation represents a simplified version of the model reported in [21], able to include temperature and electric field effects on carrier transport. It is worth pointing out that usual mobility values quoted for nitride (see for example [22]) are related to Poole–Frenkel like conduction mechanisms. In our model, this process is already taken into account by the trapping–detrapping kinetics, and our mobility strictly refers to the transport properties of carriers in the conduction and valence bands where diffusion coefficient is constant and related only to the local-lattice parameter [23].

2.4. Emissivity models

The carrier emissivity phenomena must be carefully modeled due to the different ways a carrier can be emitted from a trap and the unknown relative weight of the different phenomena. We have considered the following contributions to the total emissivity coefficient e_n:

$$e_n = v_{th}P_{th} + v_{tt}P_{tt} + v_{ph}P_{ph}, \tag{8}$$

where P_{th} and P_{ph} are the thermal and multi-phonon assisted tunneling emission probabilities from a trap to the nitride conduction band, P_{tt} represents the direct trap-to-trap emission probability and v_{th}, v_{tt} and v_{ph} are the corresponding attempt-to-escape frequencies. The thermal emission coefficient has been calculated according to [24]:

$$P_{th} = e^{-(E_T - KF^{\gamma})/k_B T}, \tag{9}$$

where E_T is the (positive) trap energy position with respect to the nitride band and the term KF^{γ} accounts for the barrier reduction in presence of an electric field F. Trap-to-trap emission has been included by means of the transition state theory [25], resulting in:

$$P_{tt} = e^{-E_b/k_B T}, \tag{10}$$

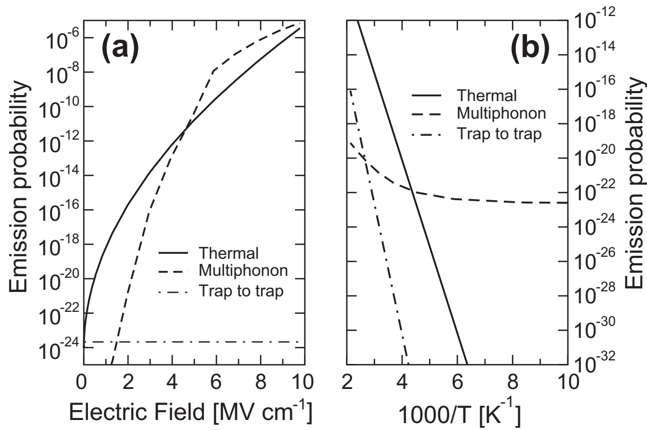

Fig. 3. Emission probability calculated for the different mechanisms as a function of the electric field for a 1.5 eV traps depth at 320 K temperature (a) and (b) as a function of temperature at a field of 2 MV cm^{-1}.

where E_b is the energy barrier involved in the transition between two traps ($0 < E_b < E_T$, related to the average distance between two traps in the nitride layer). This contribution results negligible from the fitting of the data presented in Section 4. This is mainly due to the fact that fits of Section 4 give a nitride traps density of the order of 10^{19}–10^{20} cm^{-3} that corresponds to an average trap-to-trap distance of the order of 2.2–4.6 nm, enough to consider non-interacting traps ($E_b \approx E_T$). Finally, the multi-phonon assisted tunneling probability to the nitride conduction band was calculated modifying [26] approach:

$$P_{ph} = \sum_{n=-\infty}^{\infty} I_n(z) e^{\left(nW_{ph}/(2k_BT) - S\coth(W_{ph}/2k_BT)\right)} TC(E_T + nW_{ph}), \tag{11}$$

where W_{ph} is the (single mode) phonon energy that is considered in the range of 40 meV, S is the Huang–Rhys factor [26], I_n is the modified Bessel function of order n evaluated at $z = S/\sinh(W_{ph}/(2k_BT))$ and TC is the transmission coefficient from the trap spatial position and energy $E_T + nW_{ph}$ to the nitride conduction band.

To estimate their relative weight, Fig. 3a shows the calculated emission probabilities of (9) and (11) as a function of the electric field at $T = 320$ K, assuming a trap energy depth $E_T = 1.5$ eV and $\gamma = 0.5$ (i.e. Poole–Frenkel emission): it can be seen that multi-phonon emission dominates over thermal emission at high electric fields, i.e. for P/E conditions, contrary to low electric field regime, i.e. for retention. The temperature dependence of the emission probabilities is displayed in Fig. 3b, referring to a fixed electric field $F = 2$ MV cm^{-1}. Multiphonon emission dominates at low temperatures, due to its low activation energy, while thermal emissions dominate at high temperatures.

3. Numerical treatment

In order to solve numerically system (1) and determine the V_T transient of the memory device under program, erase or retention conditions, the free and trapped electron and hole concentrations must be calculated as a function of x and t. To this aim, equations have been discretized in the spatial (index i, space step h_i) and in the time (index k, time step τ) domains using the standard central-difference scheme for the Poisson equation and the Scharfetter–Gummel approximation for the current density equations [9]. Referring to electrons (holes are treated in a similar way), the following system results where $B(y) = y/(e^y - 1)$ is the Bernoulli function and $B_{pos} = B(\mu/D(\psi_{i+1} - \psi_i))$ and $B_{neg} = B(\mu/D(\psi_i - \psi_{i+1}))$:

$$\begin{cases} \frac{\psi_{i-1}^k - 2\psi_i^k + \psi_{i+1}^k}{h_i^2} = q\frac{n_{c,i}^k + n_{t,i}^k - p_{v,i}^k - p_{t,i}^k}{\varepsilon_N} \\ \frac{n_{c,i}^k - n_{c,i}^{k-1}}{\tau} = \frac{D_{n,i}^k}{h_i^2}\left[n_{c,i-1}^k B_{neg,i-1} + n_{c,i+1}^k B_{pos,i+1} - n_{c,i}^k(B_{pos,i} + B_{neg,i})\right] - R_i^k \\ \frac{n_{t,i}^k - n_{t,i}^{k-1}}{\tau} = \frac{J_{n,i}^k}{q}\sigma_n\left[N_{t,i} - n_{t,i}^k - p_{t,i}^k\right] - e_{n,i}^k n_{t,i}^k - \frac{J_{p,i}^k}{q}\sigma_{rn} n_{t,i}^k. \end{cases}$$

The trapping dynamics is hidden in the term R_i^k allowing the use of tridiagonal solver for both the Poisson and continuity equations. The time solutions stability is obtained using implicit numerical methods: backward-Euler scheme for the continuity equations and two-step backward-differential-formula for the trapping equations.

Fig. 4 shows the adopted Gummel-map. The initial guess is used to solve the Poisson and continuity equations, then solving the coupled electron/hole trapping equations until convergence. The residuals of the main cycle for each equation are shown in Fig. 5: Poisson equation rapidly converges, while trapping equations require a higher number of iterations, though eventually converging even for strongly-coupled systems. Fig. 6 shows the calculated V_T transients (test case: ONA stack composition: 4.5/7.5/13 nm during P/E at $V_G = \pm 18$ V) for different mobility values ranging from 10^{-5} to 10^2 cm^2 V^{-1} s^{-1} and considering or not the self-consistency calculation. Self-consistent solution for program/erase transient is different from the not-self-consistent one (that means transport not correlated with electric field) only at the initial time (below $1e^{-7}$ μs): this confirms what was previously inferred in [11] about trapping efficiency in the nitride being only related to the trap properties and not to the free charge transport in bands. This is also the reason why [4,5] obtain good data fitting: electrical signature of the self-consistent calculation is not visible during program and erase measures. Moreover the V_T evolution does not depend

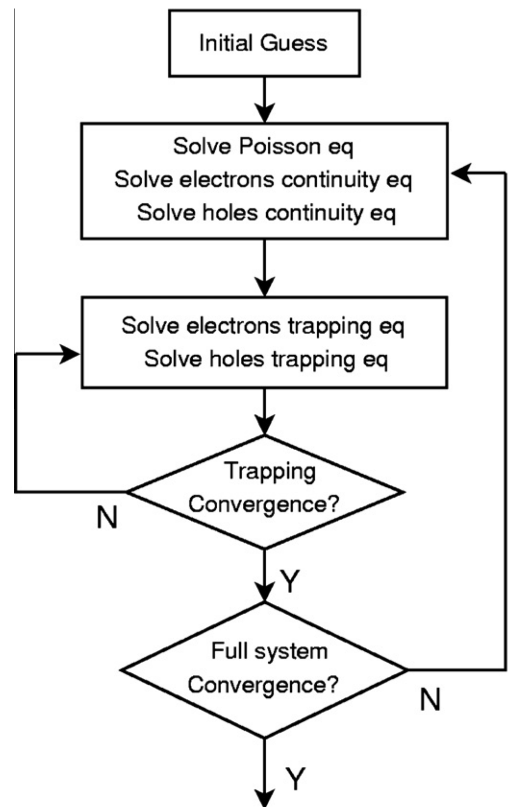

Fig. 4. Modified Gummel-map used: the initial guess is used to solve Poisson, electron and hole continuity equation; after that, the coupled trapping equations are brought to convergence and the full systems convergence is checked.

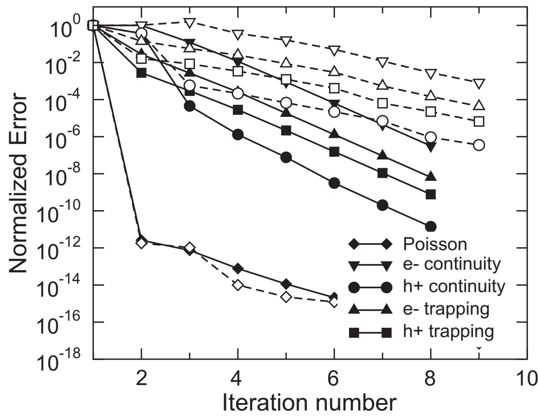

Fig. 5. Residual error of the different equations as a function of the iteration number for a strongly-coupled system (empty symbols) and for weakly coupled one (full symbols).

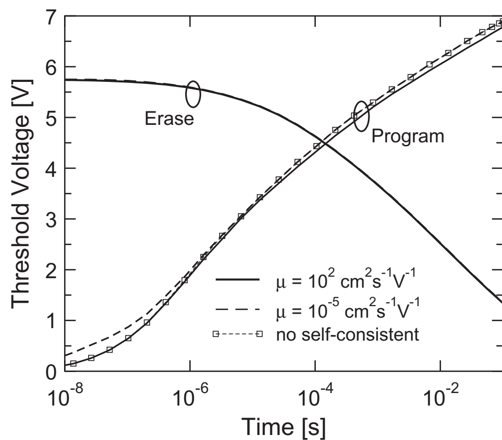

Fig. 6. Program/erase curves ($V_G = \pm 18$ V) for a TANOS structure showing the difference between self-consistency calculation and not. Moreover sensitivity to mobility variation from 10^{-5} to 10^{2} cm^2 V^{-1} s^{-1} has been shown.

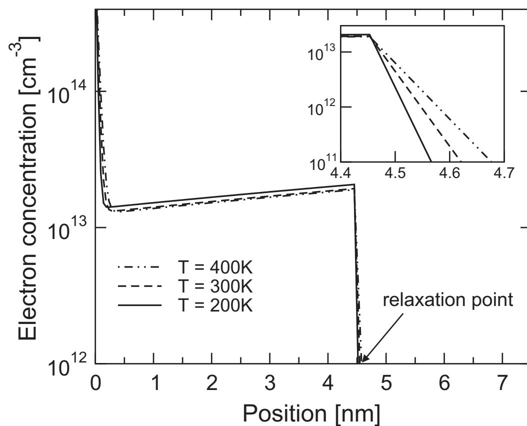

Fig. 7. Electrons back-diffusion in the nitride conduction band of a TANOS device during program, due to the relaxation mechanism at different temperatures. Inset is a zoom around the relaxation point.

on electron and hole mobilities, except for small differences in the early stage of the program operation, when very low mobility values are used. This is no strictly valid when the applied electric field is low, as in the case of retention transients where new phenomena appear. In fact in [12] a charge redistribution inside nitride is experimentally proven and to address correctly this, the redistribution must be the result of the charge transport in conduction band consistently with electric field and emission.

The effect of the boundary condition applied in the relaxation point x_r is shown in Fig. 7 where electrons back-diffusion from the energy relaxation point toward the injecting oxide interface is clearly highlighted in the inset as function of temperature.

4. Simulation results

To test our model we have applied the numerical code to a different tests case coming from published data. All the TANOS, SONOS and BET-SONOS/BET-TANOS devices have been considered where, apart from the different stack composition, different nitride layer has been used resulting in different trapping parameters. Due to the complexities and numbers of the physical phenomena considered in our model we list the fixed parameter applied to any curves: relaxation rate $\lambda = 3.1$ eV/nm, exponential for thermal detrapping $\gamma = 0.5$, phonon energy for multi-phonon emission $W_{ph} = 40$ meV, mobility parameters β and v_{sat} as declared in Section 2. Considering the wide range of data presented in literature the N_t, E_T, σ and e_n should be used as fitting parameters depending of nitride properties. Nevertheless the resulting fits showed below have given almost a constant value for escape frequencies (thermal $v_{th} \simeq 10^8$ s^{-1}, multi-phonon $v_{ph} \simeq 10^6$ s^{-1} and trap-to-trap tunneling $v_{tt} \simeq 0$ s^{-1}) and for the traps energy depth ($E_T \simeq 1.5$ eV for stoichiometric SiN and $E_T \simeq 1.3$ eV for Si-rich SiN with a gaussian distribution width $\sigma_E = 0.2$ eV). Also we have used for the dielectric relative constant the following: $\epsilon_{oxide} = 3.9$, $\epsilon_{nitride} = 6.5$,$\epsilon_{alumina} = 10.3$.

4.1. SONOS devices

Fig. 8 shows the comparison between our numerical results for program/erase and the experimental data extracted from [27], referring to a SONOS device with a 3/5/4.5 nm ONO stack. The trap spatial profile is assumed U-shape like, as suggested in [27] and in [28] for the silicon-rich nitride (but this is not relevant in fitting the results: a uniform distribution can be also assumed with the same average traps level). Calculated results were obtained with a bulk trap density $N_t = 4 \times 10^{19}$ cm^{-3} with a peak trap concentration at interfaces of $N_{int} = 2 \times 10^{20}$ cm^{-3}. Other parameters are $\sigma_n = 10^{-15}$ cm^2,$\sigma_{rn} = \sigma_{rp} = 5 \times 10^{-15}$ cm^2 and $\sigma_p = 10^{-19}$ cm^2. Note that

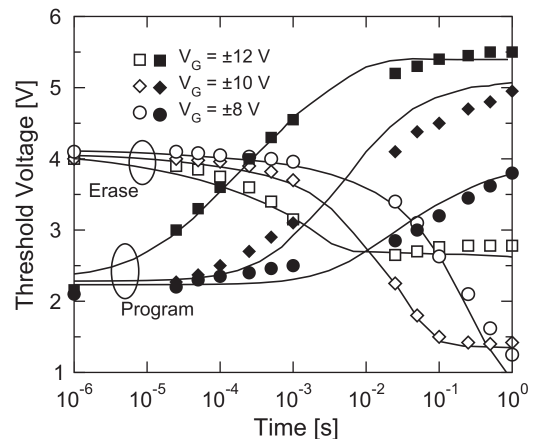

Fig. 8. Fitting of SONOS (3/5/4.5 nm of ONO stack transient program/erase curve obtained with the simulator code [27]). Line is the simulation while symbols are the experimental data.

our simulator is able to reproduce the saturating behavior during program (due to the balance between trapping and emission of electrons in the traps) and in erase, due to the leakage current through the top oxide: this is a typical behavior of the SONOS devices and it is not observed in the TANOS ones. The dis-agreement between simulation and data for the program case at V_G = +10 is related to the fact that we have used the same parameter for all the nitride fabrication processes (making a differenziation only for Si-rich and stoichiometric nitride as explained before) without any optimization for any single case; moreover we think that this difference can be also appointed to the intrinsic measure variability. Fig. 9 shows the free and trapped electron concentration in the nitride of program at a fixed time t = 1 ms and V_G = 10 V. The free electron concentration slightly decreases for deeper positions in the nitride due to the capture mechanism, while the trapped electron profile is correlated with the U-shape spatial distribution of the trap density. A similar result for holes is shown in Fig. 10 for the erase condition at a fixed time t = 1 ms and $V_G = -10$ V. The slight decrease of the hole concentration in the valence band is due to the low values of the trapping and recombination cross-section σ_p and σ_{rp}. The effect of the barrier for hole tunneling toward the gate, creating an accumulation layer, is also evident. We note that for this particular case the contribution to the threshold voltage by the trapped holes is negligible while in the case of BET-TANOS (see below) is the responsible of the negative threshold.

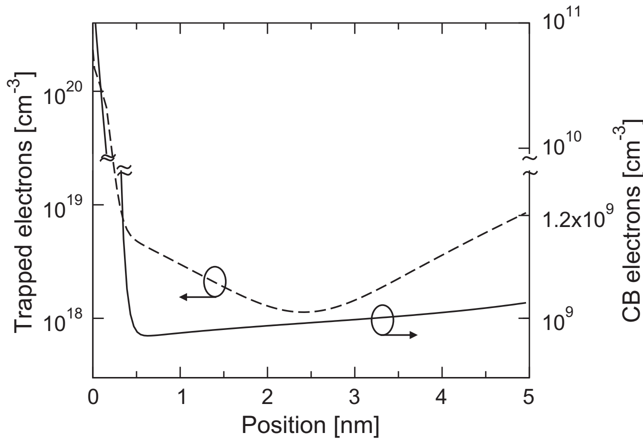

Fig. 9. Simulated electron concentration in the nitride traps (left) and in the nitride conduction band (right) of a SONOS device (3/5/4.5 nm of ONO stack) during program at t = 1 ms and V_G = +10 V.

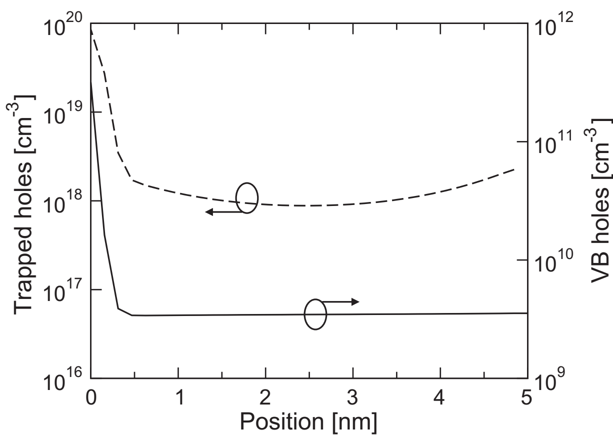

Fig. 10. Analogous to Fig. 9 but referred to holes during erase t = 1 ms and $V_G = -10$ V. Note the effect of the barrier between nitride and top oxide, resulting in accumulation of holes at the nitride boundary.

Fig. 11 shows the free electron concentration in the nitride conduction band during an erase operation. Electrons that are emitted from the nitride traps drift under the influence of the electric field toward the tunnel oxide interface, where they pile-up due to the oxide barrier before tunneling into the substrate. The V_T variation during erase is caused by electron loss from the nitride conduction band and traps to the substrate.

4.2. TANOS devices

TANOS devices has been considered [2] to overcome the limitation of the SONOS stack erase saturation by the adoption of a thicker high-K blocking oxide (Alumina) and a metal gate with high work-function (TaN). Comparison between simulation results and experimental data are shown in Fig. 12 where ONA stack is 4.5/6.5/15 nm [2]. A uniform trap spatial profile was assumed, with a trap density $N_t = 6 \times 10^{19}$ cm^{-3} and $\sigma_n = 10^{-14}$ cm^2. The model well reproduces the non-saturating behavior during erase due to the reduced back-tunneling electron current from gate.

4.3. Barrier-engineered tunnel devices

To explore the importance of the hole contribution in the modeling of the advanced charge traps devices we have analyzed data coming from [3] where the tunnel oxide of SONOS and TANOS

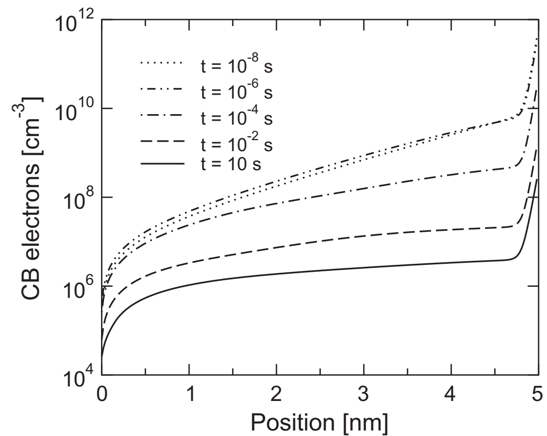

Fig. 11. Electron concentration in the nitride conduction band during an erase operation for a SONOS devices $V_G = -10$ V.

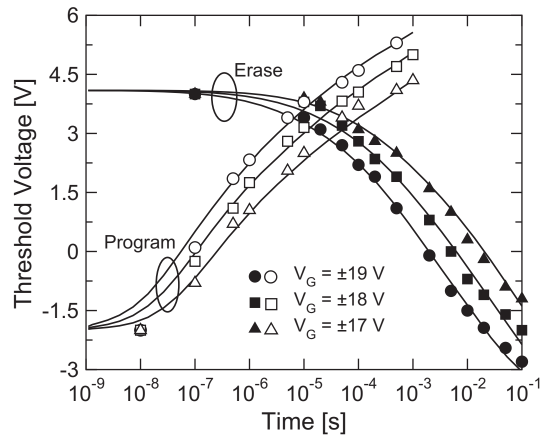

Fig. 12. Comparison against experimental data for program/erase characteristics of a TANOS device (ONA stack composition: 4.5/6.5/15 nm). Lines are simulation results; symbols are experimental data at $V_G = -18$ V.

has been replaced with an engineered tunnel made with an ONO stack to favour the hole current injection from the substrate during erase. Different compositions of the engineered tunnel (ONO) have been considered: in the case of BET-TANOS 3/2/1.5 nm, 1.5/2/2.5 nm, 1.5/2/3.0 nm while in the case of BET-SONOS 1.3/2/2.5 nm. We point out that due to the experimental approach the top oxide of the engineered barrier has been realized with the oxidation of the nitride giving a band structure similar to the SION layer. In Fig. 13 the comparison between simulation and data has been reported for erase transient curve at $V_G = -18$ V; the following parameter are used in the simulation: a uniform trap spatial distribution with a trap density $N_t = 6 \times 10^{19}$ cm^{-3}, $\sigma_n = 10^{-15}$ cm^2, $\sigma_{rn} = \sigma_{rp} = 2 \times 10^{-16}$ cm^2 and $\sigma_p = 5 \times 10^{-17}$ cm^2. The decreasing of the erase performances found in the case of BET-TANOS with ONO layer of 3/2/1.5 nm with the respect of the standard TANOS is due to the different EOT between the two architectures (higher for the BET-TANOS case) and to the fact the used ONO is too thick to have a significant hole injection. In the case of the other engineered tunnel stack, band structure increases a lot the hole current from the substrate resulting in a very low negative threshold at the end of the erase. Considering the data with BET-SONOS the difference with the respect to the BET-TANOS cases is due to the back-tunneling current coming from gate.

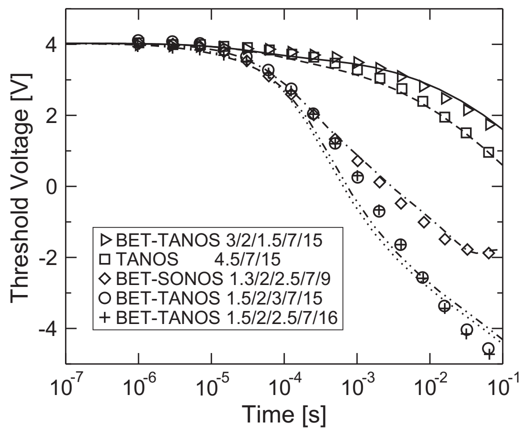

Fig. 13. Comparison between the model results and the experimental data in the case of BET-SONOS/TANOS devices. Line are the simulations and symbols the experimental data [3].

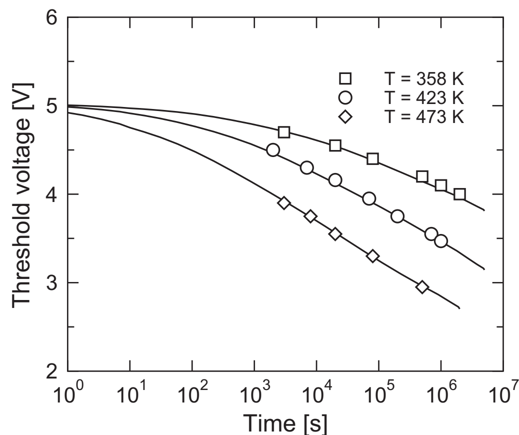

Fig. 14. Comparison of the retention simulation results with the experimental data [29] for a TANOS (4/5/10 nm of ONA stack). Line are the simulations and symbols the experimental data.

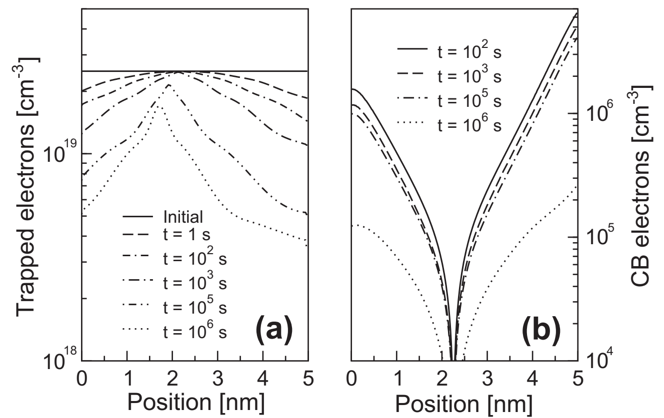

Fig. 15. (a) Trapped electrons and (b) conduction electrons during retention at 473 K for a TANOS device (4.5/6/10 nm of the ONA stack).

4.4. Retention

Fig. 14 shows a comparison between our retention simulations and experimental data at various temperatures for the TANOS structure considered in [29]. Fig. 15a shows the trapped electron profile for different retention times at $T = 473$ K: trap emission takes place towards the bottom oxide, consisting mainly in electron tunneling directly from traps to the substrate. Fig. 15b shows the corresponding conduction-band electrons: the large electron accumulation close to the bottom oxide with respect to the alumina side is related both to the higher trap emission (field-driven) and to the higher tunneling barrier.

5. Conclusions

We presented the mathematical and numerical implementation of a physics-based simulator for charge-trap memory devices, able to describe the program/erase and retention V_T transients. The simulator is based upon the semiconductor formalism for the description of carrier transport and trapping in the nitride layer, solving the equations using the Scharfetter–Gummel approach, and suitably extending it to consider the trapping phenomena with the help of the appropriate equations. Transient-time simulations convergence is obtained with a modified Gummel-map able to work also in the case of strong coupling among the equations. Simulations show very good agreement with respect to the experimental data in a large number of different operating and physical conditions, making the simulator a valuable tool for research and development of these devices.

Acknowledgements

Authors would like to thank A.L. Lacaita from Politecnico di Milano, P. Cappelletti, R. Bez and A. Benvenuti from Numonyx for discussions and support. This work has been partially supported by the European Commission under the FP7 research contract 214431 “GOSSAMER” and by MIUR under the FIRB Project No. RBIP06YSJJ.

References

[1] Wang SY, Lue HT, Lai EK, Yang LW, Yang T, Chen KC, et al. Reliability and processing effects of bangap engineered SONOS (BE-SONOS) flash memory. In: Proc IEEE int reliab physics symp; 2007. p. 171–5.
[2] Park Y, Choi J, Kang C, Lee C, Shin Y, Choi B, et al. Highly manufacturable 32Gb multi-level NAND flash memory with 0.0098 μm^2 cell size using TANOS (Si-Oxide-Al_2O_3–TaN) cell technology. In: Proc of IEEE IEDM tech dig; 2006. p. 35–40.

[3] Lai SC, Lue HT, Yang MJ, Hsieh JY, Wang SY, Wu TB, et al. MA BE-SONOS: a bandgap engineered SONOS using metal gate and Al_2O_3 blocking layer to overcome erase saturation. In: Proc of IEEE NVSMW; 2007. p. 88–9.
[4] Song YC, Liu XY, Wang ZY, Zhao K, Du G, Xia ZL, et al. Evaluating the effects of physical mechanism on program, erase and retention in charge trapping memory. In: Proc of IEEE SISPAD; 2008. p. 41–4.
[5] Paul A, Sridhar C, Gedam S, Mahapatra S. Comprehensive simulation of program, erase and retention in charge trapping flash memories. In: Proc of IEEE IEDM tech dig; 2006. p. 393–6.
[6] Furnemeont A, Cacciato A, Breuil L, Rosemeulen M, Maes H, De Meyer K, et al. Physical understanding and modeling of SANOS retention in programmed state. Solid State Electron 2008;52(4):577–83.
[7] Nasyrov KA, Gritsenko VA, Kim MK, Chae HS, Chae SD, Ryu WI, et al. Charge transport mechanism in metal–nitride–oxide–silicon structures. IEEE ED Lett 2002;23:336–8.
[8] Vianello E, Driussi F, Palestri P, Arreghini A, Esseni D, Selmi L, et al. Impact of the charge transport in the conduction band on the retention of SiNitride based memories. In: Proc of the ESSDERC; 2008. p. 107–10.
[9] Scharfetter D, Gummel H. Large-signal analysis of a silicon read diode oscillator. IEEE Trans of ED 1969;ED-16:64–77.
[10] Gummel HK. A self-consistent iterative scheme for one dimensional steady state transistor calculation. IEEE Trans of ED 1964;ED-11:455–65.
[11] Arnett PC. Transient conduction in insulators at high fields. J Appl Phys 1975;46:5236–43.
[12] Ghidini G, Galbiati N, Mascellino E, Scozzari C, Sebastiani A, Amoroso S, et al. Charge retention phenomena in CT silicon nitride: impact of technology and operating conditions. J Vac Sci Technol B, submitted for publication.
[13] Peterson M, Roizin Y. Density functional theory study of deep traps in silicon nitride memories. Appl Phys Lett 2006;89:053511.
[14] Arora VK. Drift diffusion and Einstein relation for electrons in silicon subjected to a high electric field. Appl Phys Lett 2002;80:3763–5.
[15] Mauri A, Monzio Compagnoni C, Amoroso S, Maconi A, Cattaneo F, Benvenuti A, et al. A new physics-based model for TANOS memories program/erase. In: Proc of IEEE IEDM tech dig; 2008. p. 555–8.
[16] Suzuki E, Hayashi Y, Yanai H. Transport processes of electrons in MNOS structures. J Appl Phys 1979;50:7001–6.
[17] Tomita T, Kamakura Y, Taniguchi K. Energy relaxation length for ballistic electron transport in SiO_2. Phys Stat Sol 1997;204:129–32.
[18] Lue H-T, Du P-Y, Wang S-Y, Lai E-K, Hsieh K-Y, Liu R, et al. A novel gate-sensing and channel-sensing transient analysis method for real-time monitoring of charge vertical location in SONOS type devices and its applications in reliability studies. In: Proc IEEE int reliab physics symp; 2007. p. 177–83.
[19] Ishida T, Mine T, Hisamoto D, Shimamoto Y, Yamada R. Anomalous electron storage decrease in MONOS nitride layers thinner than 4 nm. IEEE ED Lett 2008;29:920–2.
[20] An C-H, Soo Lee M, Kim H. Effects of Si_3N_4 thickness on the electrical properties of oxidenitrideoxide tunneling dielectrics. J Electrochem Soc 2008;155: G247–52.
[21] Canali C. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. IEEE Trans ED 1975;ED-22:1045–7.
[22] Katayama K. Measurement technique of carrier mobility in silicon nitride and its application to data retention in Monos memories. In: Proc IEEE int reliab physics symp; 2007. p. 656–7.
[23] Mott NF, Davis EA. Electronic processes in non-crystalline materials. Claredon Press; 1979.
[24] Manzini S. Electronic processes in silicon nitride. J Appl Phys 1987;62:3278–84.
[25] Voth GA, Chandler D, Miller WH. Rigorous formulation of quantum transition state theory and its dynamical corrections. J Chem Phys 1989;91:7749–60.
[26] Makram-Ebeid SS, Lanoo M. Quantum model for phonon assisted tunnel ionization for deep levels in a semiconductor. Phys Rev B 1982;25:6406–24.
[27] Wu KH, Chien HC, Tsai TK, Chang JW, Chan CC, Chen TS, et al. Phenomenal SONOS performance for next-generation flash memories. In: Proc of symposium on nano device technology; 2004. p. 35–40.
[28] De Salvo B, Ghibaudo G, Pananakakis G, Masson P, Baron TN, Fernandes A, et al. Experimental and theoretical investigation of nano-crystal and nitride-trap memory devices. IEEE Trans ED 2001;ED-48:1789–99.
[29] van den Bosch G, Furnemont A, Zahid MB, Degraeve R, Breuil L, Cacciato A, et al. Nitride engineering for improved erase performance and retention of TANOS NAND flash memory. In: Proc of NVSMW and ICMTD; 2008. p. 128–9.