A global modeling approach of the leakage phenomena in dielectrics

Jérémy Postel-Pellerin[1,*], *Gilles* Micolau[2], *Philippe* Chiquet[1], *Maminirina* Joelson[2], and *Jean-Baptiste* Decitre[3]

[1]Aix-Marseille University, IM2NP, CNRS, UMR 7334, rue Enrico Fermi, 13451 Marseille, France
[2]Avignon University, EMMAH, UMR 1114, INRA-UAPV, 301 rue Baruch de Spinoza, 84000 Avignon, France
[3]LSBB, UMS 3538, UNS/UAPV/CNRS/AMU, La Grande Combe, 84400 Rustrel, France

Abstract. Thanks to its low noise level, the LSBB environment provides particular environment to carry out high quality electrical characterizations. In this paper, we propose a complete modeling approach of the experimental results from our experimental microelectronic setup. The tested device is a Metal Oxide Semiconductor (MOS) floating gate capacitor which can be found in electrostatic non volatile memories such as Flash. The main idea is to characterize and model the leakage current through the tunnel oxide. We proposed, in a previous work, a model for charge loss considering a fractional Poisson process, involving only two parameters, expressed as a Mittag-Leffler (ML) function. Here, we also propose a combo of Fowler-Nordheim (FN) and Poole-Frenkel (PF) models for leakage currents, based on tunnel effect transport through the oxide. It gives the leakage current on a medium-to-long scale of time while the ML model can possibly take into account a shorter time step. The perspective is to find a relationship between these different models, used in various fields, to propose a generic model of phenomena involving leakage in complex and porous materials at different scales of time and space.

1 Introduction

Flash memory cells are based on the floating gate technology principle [1]. The most widespread solution to enable semiconductor memories to be non-volatile, that is to say able to keep information without any power supply, is to use MOS transistors whose threshold voltage is shifted by a charge stored in an isolated gate above the channel. Floating gate technologies consist in adding a second gate between the gate and the channel of a classical MOS transistor. This second gate, called "Floating Gate" (FG), can isolate charges to make the transistor threshold voltage variable. Charges are injected through a dielectric, in general Silicon dioxide SiO2, placed between the floating gate and the transistor channel, as presented in Fig. 1. Lastly, the two gates of this "transistor" are separated by a tri-layer stack oxide "Oxide/Nitride/Oxide" (ONO). Thus, the Flash elementary cell can be seen as a classical MOS transistor whose gate would be in series with a capacitor C_{PP}. The floating gate can now store a charge while the second electrode of the capacitor becomes the "Control Gate" (CG) of the cell.

*e-mail: jeremy.postel-pellerin@im2np.fr

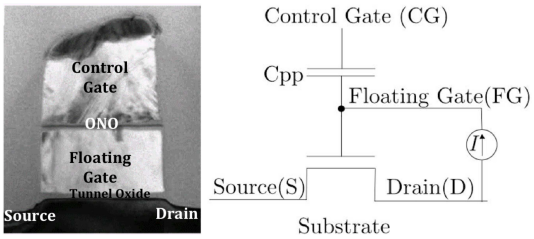

Figure 1. Picture of a floating gate structure (left) and its electrical scheme (right)

Barrier transparency in the tunnel oxide, which can be electrically modeled by a current source I, allows the injection of charges in the floating gate, shifting the MOS transistor threshold voltage V_T according to equation (1):

$$V_T = V_{T0} - \frac{Q_{FG}}{C_{PP}} \tag{1}$$

where V_{T0} is the natural threshold voltage of the cell, Q_{FG} the charge amount in the floating gate and C_{PP} the capacitance between the control gate and the floating gate.

Memory cell reliability, defined as the capability of said device to function over time, is a major issue for manufacturers and can be related to many parameters such as process and using conditions. Indeed, unceasing device scaling, decrease of dielectric thicknesses around the floating gate and high voltages applied on cells are many parameters altering memory reliability [2–5]. Thus, dielectrics' quality is a major issue due to their antagonistic roles: avoiding electric charge leakage currents during retention phase while being transparent enough during programming steps. A better understanding of these leakage currents is crucial to improve the whole quality of our memory cells, that's why we have to develop powerful methods to reach very low current levels.

2 Floating Gate Technique

Most of the electrical measurements developed to characterize semiconductor devices and especially Non-Volatile Memories are based on current measurements. Indeed, to study a major reliability aspect of these NVM we have to evaluate the very low-level leakage currents responsible for the charge loss during the retention phase. Nevertheless, these currents are not accessible through direct measurements, even with high-performance analyzers [6], so we have to use some indirect measurement techniques to reach lower level currents. One of the most widely used technique is the "Floating-Gate Technique" (FGT), based on the use of a MOS capacitor and a MOS transistor in parallel [7, 8]. Using this method requires i) protecting the wafer from mechanical perturbations (vibrations) and ii) (very) long time acquisition. In a classical laboratory environment, those two constraints are generally not easy to deal with. Indeed, they lead to use, for a very long time, huge, heavy and very expensive probers (mechanically insulated by air shocks), which is often economically not possible. The main idea of our experimental platform consists in developing a "cheap" platform, presented exhaustively with its methodology in [9–12]. The mechanical insulation is "naturally" done by the very peculiar environment of the underground LSBB Laboratory located in Apt (South of France). The direct result of the measurements is a Drain-Source current in the TMOS device. After treatment consisting in interpolation with the known characteristics of this later,

the electric charge in the floating gate is accessible versus time. Thus, the leakage current is the temporal derivative of the charge according to time. Because the measurements are generally over sampled [12], this operation cannot be simply done by direct numerical derivation as finite differences. This critical operation can be leaded by a modeling of the current, fitting correctly the charge curve. One instance of extracted leakage current is proposed in Fig. 2.

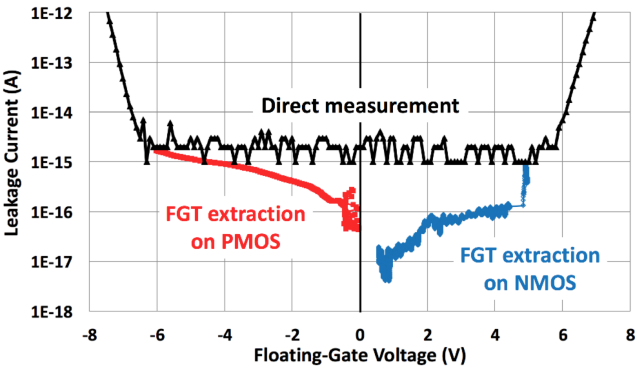

Figure 2. Results of the FGT extraction for both positive and negative floating-gate voltages

The curve of the leakage current from direct derivative of charge evolution is found to be highly noisy and is not suitable to define a clear trend for current leakage. It is directly linked to the over-sampling measurement process. Because we cannot anticipate the variation of the charge during time, we choose an arbitrarily short time between two measures. One way to estimate this leakage current is to have a model whom generated curves can be derivated without issue. Two ways are explored : "deterministic" loss of charge by known microelectronics currents and relaxation law generated by random processes.

3 Modeling of the leakage current

3.1 Conduction through dielectrics in microelectronics

In microelectronics, electrons can directly tunnel through dielectrics, but conduction can also be assisted by traps existing in a damaged dielectrics.

3.1.1 Tunneling currents

Quantum mechanics, through the resolution of the Schrödinger equation, makes it possible to predict the passage of particles through a barrier having finite width and height [13]. This non-zero probability of passage of carriers, called transmission coefficient, is the cause of various conduction phenomena through the oxide, such as direct tunnel or Fowler-Nordheim (FN) tunneling [14]. Due to the relatively thick tunnel oxide (around 10nm [5]) , direct tunneling is very low but the intern electric field is high enough to enable Fowler-Nordheim mechanism, driven by two parameters denoted α_{FN} and β_{FN} in equation (2):

$$J_{FN} = \alpha_{FN}.E_{tun}^2.exp\left(-\frac{\beta_{FN}}{E_{tun}}\right) \tag{2}$$

In the following of this study, we will only consider FN tunneling as a mechanism possibly responsible for observed leakage currents.

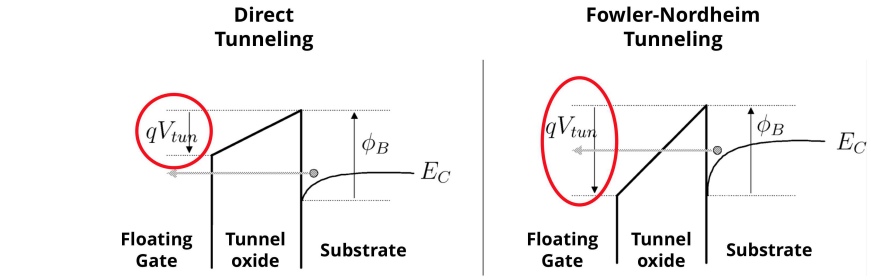

Figure 3. Tunneling mechanisms

3.1.2 Trap-assisted mechanisms

Due to the repeated programming/erasing operations on the memory cell (and thus electrons passing through dielectrics), tunnel oxide is progressively damaged and traps are created inside it, creating additional leakage paths. Several trap-assisted mechanisms have previously been proposed in literature [15, 16], described in Fig. 4.

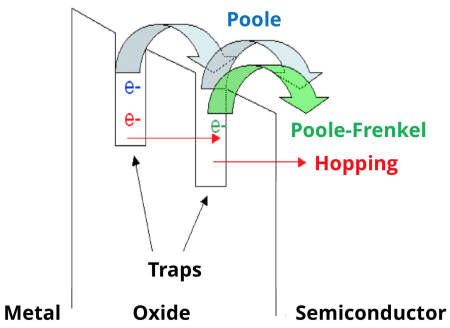

Figure 4. The main trap-assisted mechanisms in dielectrics

Tunnel oxide is thin enough so that the probability to have two (or more) traps aligned to create a conductive path is very low [17] and we will only consider in our modeling process the Poole-Frenkel mechanism, according to equation (3), driven by two parameters denoted α_{PF} and β_{PF}:

$$J_{PF} = \alpha_{PF}.E_{tun}.exp\left(q.\beta_{PF}.\frac{\sqrt{E_{tun}}}{kT}\right) \tag{3}$$

At the scale of an electrical characterization of a device, finally, only biases and flowing currents are recorded. FN and PF can be discriminated because they don't occur in the same zone of biases. Nevertheless, it is not simple to characterize a PF mechanism, because it depends on the number of traps, their geometrical characteristics, their electromagnetic behavior, etc. It can also be useful to consider that we are dealing with a transfert of material in a complex medium, described by a set of macroscopical parameters, in view to build relaxation models.

3.2 Relaxation models

In a previous paper, we have proposed the use of relaxation models, considering either a simple relaxation (Classical Poisson process) or a relaxation with a memory effect (Fractional Poisson process) [12].

3.2.1 Classical Poisson process

Indeed, we can consider electrons in the floating-gate as moving particles in a box with an opening to the outside. The number of particles inside the box is a random process N(t) following a binomial law and then the mean value n at time t follows ordinary differential equation (4):

$$\frac{d}{dt}n(t) + C_c.n(t) = 0 \tag{4}$$

where the parameter C_c is the probability that a given particle is instantaneously outside the box. The obvious solution is given by an exponential law (5):

$$n(t) = n_0.exp(-C_c.t) \tag{5}$$

where n_0 is the initial number of particles. It is a memoryless relaxation law, also known as the disintegration law.

The modeling of the charge evolution, using Classical Poisson process, is shown in Fig 6.

3.2.2 Fractional Poisson process

If a memory effect is taken into account i.e. the probability of particle to come out is now depending on the number of remaining particles, we have a Fractional Poisson Process [18], based on the Mittag-Leffler function and on the fractional Caputo derivative of order α [19–23]. The mean value n of number of particles at time t follows the fractional relaxation equation (6):

$$\frac{d^\alpha}{dt^\alpha}n(t) + C.n(t) = 0 \tag{6}$$

4 Comparison between models and experiments

We have first considered only the FN mechanism, occurring at "high" voltage, which has been fitted on the direct measurement. We can see in Fig. 5 that this mechanism does not explain the FGT extraction and a second mechanism is necessary for lower voltages. That's why we added the PF mechanism and the sum of the two contributions. The four parameters, $\alpha_{FN}, \beta_{FN}, \alpha_{PF}, \beta_{FN}$ have been optimized in order to model both direct measurement (over the noise level) and FGT extraction (Fig. 5).

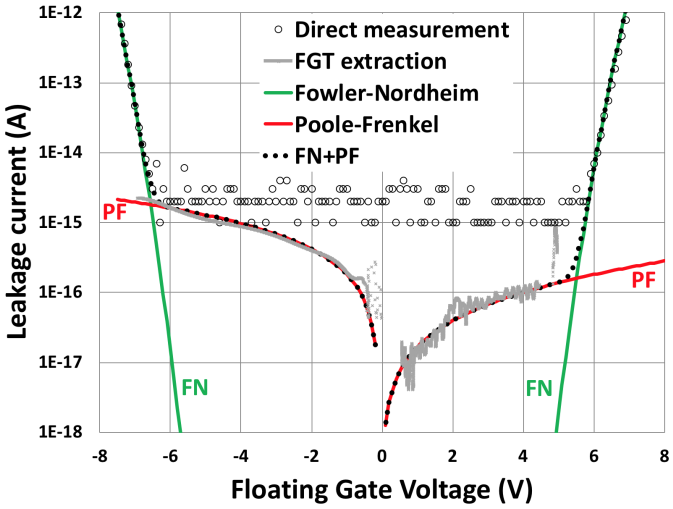

Figure 5. Modeling of both direct measurement and FGT extraction, using a combo of FN and PF mechanisms. $\alpha_{FN} = 7.1 \times 10^{-4} A/V^2$, $\beta_{FN} = 2.55 \times 10^{10} V/m$, $\alpha_{PF} = 1.26 \times 10^{-16} A/(V.m)$, $\beta_{FN} = 5 \times 10^{-7} (V.m)^{1/2}$

We present, then, the comparison between measurements and the relaxation processes models. The parameters of the relaxation models have been optimized in order to fit only the FGT measurements. Two models are explored : the classical Poisson model (one parameter C_c) and the fractional Poisson Processes. For this latests, two set of parameters are proposed. The first one (C_1, α_1) is optimized in order to fit the upper part of the experimental measurements of the charge (short times), while the second one (C_1, α_1) is optimized to fit the lower part of the curve (long times). Fig 6 shows the modeling of the charge evolution, using Classical Poisson (one parameter C_c) and Fractional Poisson processes (four parameters).

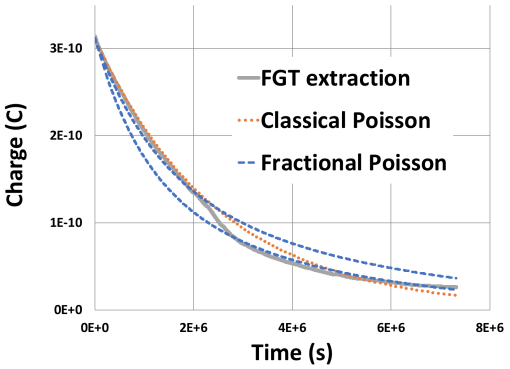

Figure 6. Three models of relaxation compared to the extracted values of the charge versus time. Two fractional models are proposed, for the upper part and the lower part of the curve. The parameters are : $C_c = 4 \times 10^{-7} s^{-1}$, $\alpha_1 = 0.8922$, $C_1 = 3.285 s^{-\alpha_1}$,$\alpha_2 = 0.964$, $C_2 = 2.85 s^{-\alpha_2}$

After a derivative step, based on the analytical formulation of the models, and since we know the floating-gate voltage for each time step, we obtain the leakage current evolution, presented in Fig. 7, for the three relaxation models.

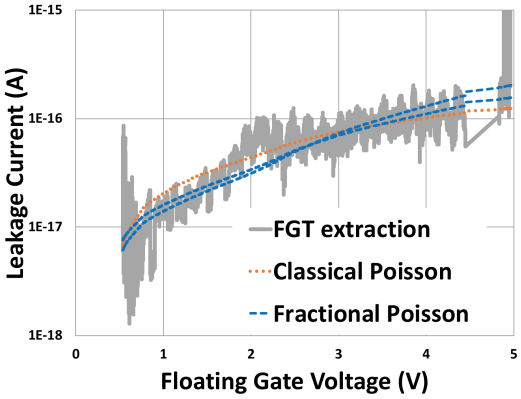

Figure 7. Three models of relaxation compared to the extracted values of the leakage current, from charge evolutions in Fig. 6.

It appears that the three relaxation models are close to the shape of the FGT measurements for biases between 0.5 and 4.5 V. They don't succeed in fitting upper biases of the floating gate. It can be interpreted by the fact that for biases greater than 4.5 V, the deterministic FN current becomes the major way of electrons transfert through the oxide. Looking at Fig.5, we denote that FN current is the greatest when the floating gate voltage rises around 5 V.
Let us insist on the fact that for the Poisson relaxation models, either classical or fractional, only the measurements of charge versus time are necessary, while for FN and PF models, a model is needed to link the floating gate voltage to the electric field in the oxide.
Finally, to summarize, we present in figure 8, in logarithmic scale, all the models and the measurements in the same figure. Fractional Poisson processes give a good agreement with measurements for biases lower than 1.5 V.

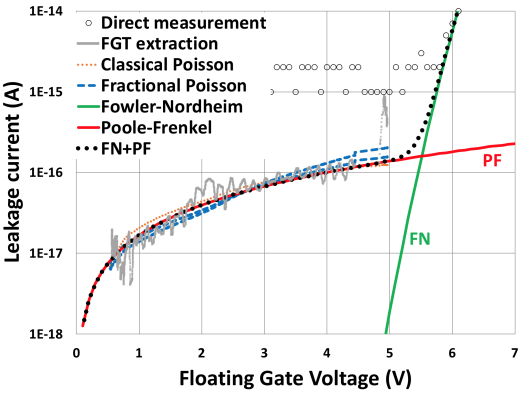

Figure 8. All models compared to the extracted values of the leakage current, from datas in Fig. 5 and 7.

5 Conclusion

We present here two different approaches for fitting ultimate measurements of current leakage in a floating gate device. The first one is based on deterministic physical formulations of charge transfert in insulators. The second approach is based on stochastic formulations involving processes with memory effect, according to the fact that the oxide could be seen as a complex material in which displacements are resulting from numerous interactions. These two approaches give a good agreement in the long time range behavior of the charge loss in the device, corresponding to low biases of the floating gate.
We can explore two main ways of perspective. The first one should be a theoretical approach linking fractional Poisson process with trap assisted currents as Poole-Frenkel. The second way of perspectives lyes in increase the capability of our experimental platform adding a capacitance meter (allowing to monitor the damages of the oxide) and a better mechanical and optical system (allowing larger wafers with smaller devices).

References

[1] W. Brown and J. Brewer, *Nonvolatile semiconductor memory technology: a comprehensive guide to understanding and to using NVSM devices* (IEEE Press, 1998)
[2] P. Canet et al., Proc. of IEEE ISCAS, pp. 846-849 (2001)
[3] P. Canet et al., Proc. of IEEE MWSCAS, pp. 1144-1147 (2000)
[4] P. Canet et al., Proc. of IEEE ESSCIRC, pp. 683-686 (2002)
[5] ITRS, *Process Integration, Devices, and Structures* (2013)
[6] Keysight, *B1500A Semiconductor Device Analyzer data sheet* (October 2016)
[7] B. Fischbein et al., IEEE Elec. Dev. Lett. **12**, pp. 713-715 (1991)
[8] F.H. Gaensslen et al., IEEE Elec. Dev. Lett. **1**, pp. 231-233 (1980)
[9] J. Postel-Pellerin et al., Proc. of IMEKO TC4, pp. 1041-1045 (2014)
[10] J. Postel-Pellerin et al., Acta Imeko **4**, pp. 36-41 (2015)
[11] J. Postel-Pellerin et al., Proc. of IEEE CAS, pp. 287-290 (2014)
[12] G. Micolau et al., Proc. of iDUST **2**, pp. 1-9 (2016)
[13] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Mécanique Quantique*, (Hermann, 1998)
[14] R.H. Fowler and L. Nordheim, Proc. of the Royal Soc. of London A **119**, pp. 173-181 (1928)
[15] R.E. Shiner et al., Proc. of IEEE IRPS, pp. 238 (1980)
[16] N.R. Mielke, Proc. of IEEE IRPS, pp. 106 (1983)
[17] J.H. Stathis, Journ. of Appl. Phys. **86**, pp. 5757-5766 (1999)
[18] N. Laskin et al., Comm. in Nonlin. Sci. and Sim. **8**, pp. 201-213 (2003)
[19] M.-C. Neel et al., Proc. of Dyn. Frac. Appl. (2010)
[20] J. Golder et al., Math. and Comp. in Sim. **81**, pp. 2181-2189 (2011)
[21] J. Golder et al., Water Sci. and Engineering **7**, pp. 32-40 (2014)
[22] D. Sierociuk et al., IEEE Trans. on Control Syst. Technology **21**, pp. 459-466 (2011)
[23] A.A. Kilbas et al., Integral Transform and Spec. Func. **15**, pp. 31-49 (2004)