UDC 004-076-4

A.B. Gnilenko

FLOATING GATE FLASH MEMORY CELL SIMULATION

Abstract. Flash memory is widely used in computers, information and communication electronic systems as a modern non-volatile memory device that is capable to store data without connection to power supply. For the efficient design of the flash memory devices, an accurate simulation of the memory cells becomes increasingly important. In the present paper, a flash memory cell with the floating gate enveloped by the control gate is simulated using TCAD tools. The memory cell is simulated during programming and erasing operations. I-V and transient characteristics are obtained for various tunnel oxide thicknesses and operating voltages to demonstrate the efficiency of the cell design under consideration.

Keywords: flash memory, non-volatile memory, floating gate, TCAD.

Introduction. Non-volatile memories (NVM), like EPROM, EEPROM and Flash have become indispensable elements of all modern electronic information and communication systems [1]. Their property to retain stored information not being connected to power supply appears vital especially for mobile systems. Nowadays, flash memory is widely used among other types of NVM to store information in portable devices.

Conventional flash memory cell is realized as a single-transistor device with an electrically isolated floating gate. Floating gate flash memory has been successfully used in compact portable memory storage devices for many years with continuous upgrading its characteristics [2]. The development of new design solutions as well as direct down-scaling cell dimensions made it possible in past years to succeed in obtaining higher programming and erasing speed, higher data-storage density and lower operating voltage [3]. A typical floating gate flash memory cell is a MOSFET-like transistor which uses an additional polysilicon gate layer completely isolated in oxide and located on top of p-doped silicon substrate. That floating gate layer is governed through a capacitive connection by a polysilicon control gate accessible for metal electrodes. The floating gate serves as a charge trapping layer to store data. The charge level on the floating gate determines the

logical state of the memory cell. The capacitive nature of the floating gate connection to silicon substrate layer and polysilicon control gate stipulates strong influence of the properties of surrounding insulators and shapes of the gates on memory operation characteristics.

The floating gate is insulated from n-channel of p-doped silicon substrate by thin tunnel oxide layer and from the control gate by inter-poly blocking layer, for which stacked layers of oxide-nitride-oxide (ONO-layer) are commonly used. The tunnel oxide is highly involved in all memory operations, like programming, reading, erasing and retaining. This layer must be thin enough to easily transfer charges to the floating gate during programming process and in the opposite direction for erasing. At the same time, the floating gate should be thick enough to prevent discharging the floating gate when the power is switched off and thereby to provide good charge retaining. The blocking layer between the gates is the second important insulating layer which should create an impenetrable for electrons potential barrier preventing electrons to leave the floating gate and reach the control gate properly biased for programming. But also, this gate must provide effective floating gate control with low programming voltage.

Commonly used floating gate flash memory design employs the control gate and the floating gate of the same length size and shape which are stacked on each other and separated by ONO blocking layer. To improve characteristics of the cell for effective memory operations one can change the shape and/or size of the gates as well as blocking and tunnel layers trading off between coupling and barrier properties [4].

In the paper a memory cell with the floating gate enveloped by the control gate is investigated by means of computer simulation. The design under consideration can improve floating gate control for low energy consumption and facilitate further down-scaling flash memory cells. The floating gate flash memory cell is simulated by Silvaco TCAD software package which is a modern simulation standard in computer aided design (CAD) of semiconductor devices allowing us to have a deep insight into the physics of processes in devices to reliably predict their behavior. The technological process of memory cell fabrication is simulated to construct the desired cell structure following methods used in semiconductor industry. Basic memory operations, like programming and erasing are simulated and electrical characteristics of the flash memory cell are obtained and discussed.

Simulation Tools. Silvaco TCAD is the sophisticated tool capable to simulate any type of semiconductor devices in both two and three dimensional domains. Silvaco

TCAD consists of two main components – process simulator ATHENA and device simulator ATLAS. ATHENA provides the physically-based simulation of technological processes used in semiconductor industry like ion implantation, diffusion, annealing, deposition, etching, etc. ATLAS makes it possible to predict the electrical behavior of semiconductor devices at specified bias conditions using different physical models – transportation, carrier statistic, mobility, recombination, etc.

In this paper ATHENA software is applied to simulate the fabrication of the floating gate flash memory cell and then, ATLAS software is used to calculate the memory cell electrical characteristics under the framework of the drift-diffusion transport model. This model consists of a set of fundamental equations including Poisson's equation, the continuity equations and the transport equations, which link together the electrostatic potential and the carrier concentrations inside a simulation domain.

The properties of charge carriers are described by Shockley-Read-Hall recombination model, transverse field, doping and temperature dependent mobility model and empirical impact ionization model under Fermi-Dirac carrier statistics. The hot electron injection model is used to simulate flash memory cell programming and Fowler-Nordheim tunneling for electrons and band-to-band tunneling model are applied to simulate erasure mechanism.

The set of the coupled non-linear partial differential equations are discretized on an irregular mesh of nodes within considered device domain using the finite-difference method. Then, an obtained discrete system of algebraic equations is solved by Gummel and Newton non-linear iteration techniques.

Flash Memory Cell Design. The technological process simulator ATHENA is applied to create a 2-D model of the floating gate flash memory cell, which can be used for obtaining its electrical characteristics with device simulator ATLAS. The structure of the memory cell is shown in Fig. 1 as a result of process modeling in ATHENA simulator.

The memory cell design is featured by a low profile obtained because the floating gate is being partly enveloped by the control gate. The relatively thin control gate "embraces" the floating gate from the top and sides creating good coupling for effective charge management.

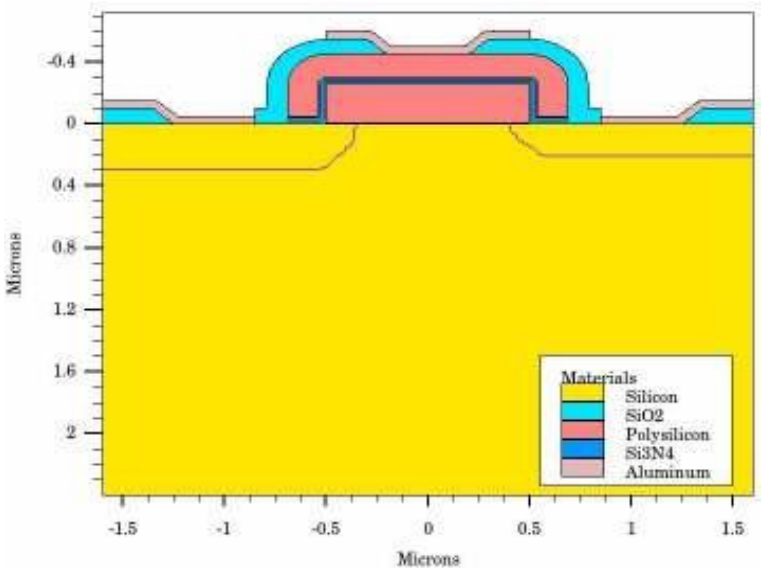

Figure 1 – Floating gate flash memory cell structure

The cell structure modeling is a sequence of technological stages to gradually form the desired design. The simulation is started from the choice of an initially p-doped bulk silicon region. To create n+-doped domains in p-doped silicon substrate for the source and the drain, ion implantation of arsenic is applied through the mask layer for obtaining the desired n-channel length. The floating gate is formed by the successive deposition of a thin layer of SiO2 for tunnel oxide and a polysilicon layer for the gate which are etched to obtain the required length of the gate. The successive deposition of the oxide-nitride-oxide (ONO) inter-poly layer, the control gate polysilicon layer and the insulating oxide layer forces the control gate to envelop the floating gate. For the design shown in Fig. 1, the floating gate and the control gate have the thickness of 0.25 μm and 0.15 μm, respectively. The tunnel oxide is 8 nm thick while ONO-layer has a 26 nm thickness being consisted of 6 nm SiO2 layer, 16 nm Si3N4 layer and 4 nm SiO2 layer stacked on top of each other. The cell design is finalized by the deposition and the etching of an aluminum layer which creates metallic contacts for the source, the drain and the control gate through preliminary made cavities in the insulating oxide layer.

Numerical Simulation Results. The electrical characteristics of the cell structure created in ATHENA process simulator is obtained using ATLAS device simulator. The key features of memory device simulation are the modeling basic memory operations, like programming and erasing. Both typical for MOSFET devices I-V characteristics, as well as transient characteristics are needed to analyze memory operation mechanisms. This investigation is mainly focused on the analysis of oxide thickness and operating voltage influence on the memory cell characteristics.

To estimate the threshold voltage value and the threshold voltage shift as a result of cell programming, I-V characteristics are obtained before and after programming for various tunnel oxide thicknesses. I-V characteristics calculated before programming with uncharged floating gate are demonstrated in Fig. 2 for a set of tunnel oxide thicknesses.

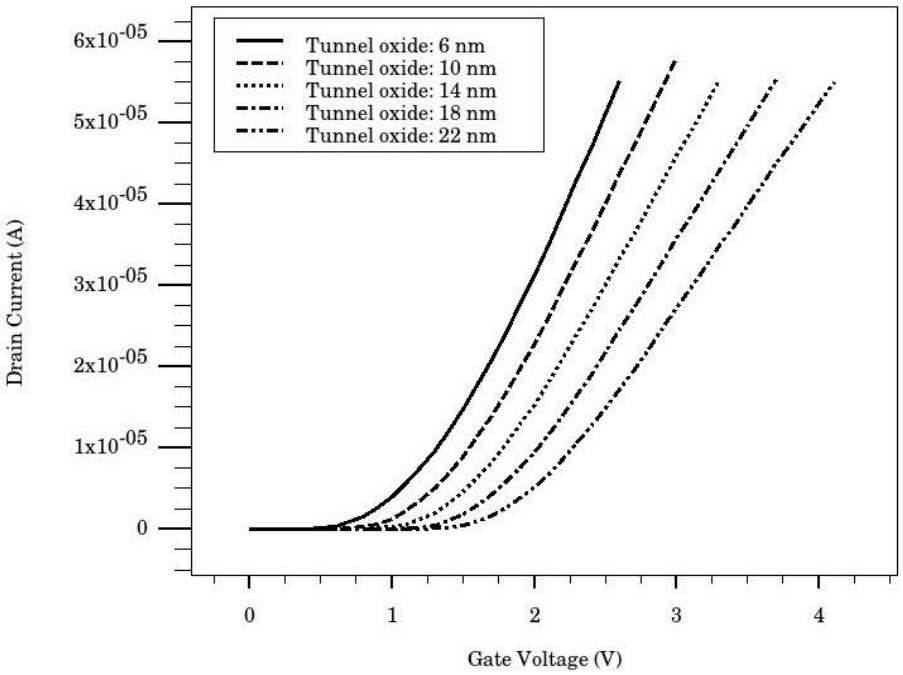

Figure 2 – Drain current versus gate voltage before programming

Programming is the memory operation that allows us to transfer charges to the floating gate through the tunnel oxide layer changing electrical properties of the device. For the simulation of programming, the drain is biased with 6 V and the gate – with 12 V leaving biasing conditions enough time for floating gate charging. I-V characteristics are obtained again after programming for the memory cell with charged floating gate as it is shown in Fig. 3. Comparing I-V characteristics before and after programming allows us to estimate threshold voltage shift which should be large enough to clearly recognize the state of the flash memory cell while reading operation.

For example, the threshold voltage shift is 10.14 V for 10 nm tunnel oxide thickness which is large enough to sense the memory cell state by memory device circuitry. The threshold shift remains significant in the wide interval of tunnel oxide thicknesses under applied to the electrodes voltages. This makes it possible to decrease the gate bias and lower energy consumption. The shape of the control gate can provide good coupling and draw acceptable charge level to the floating gate for thick enough inter-poly layer.

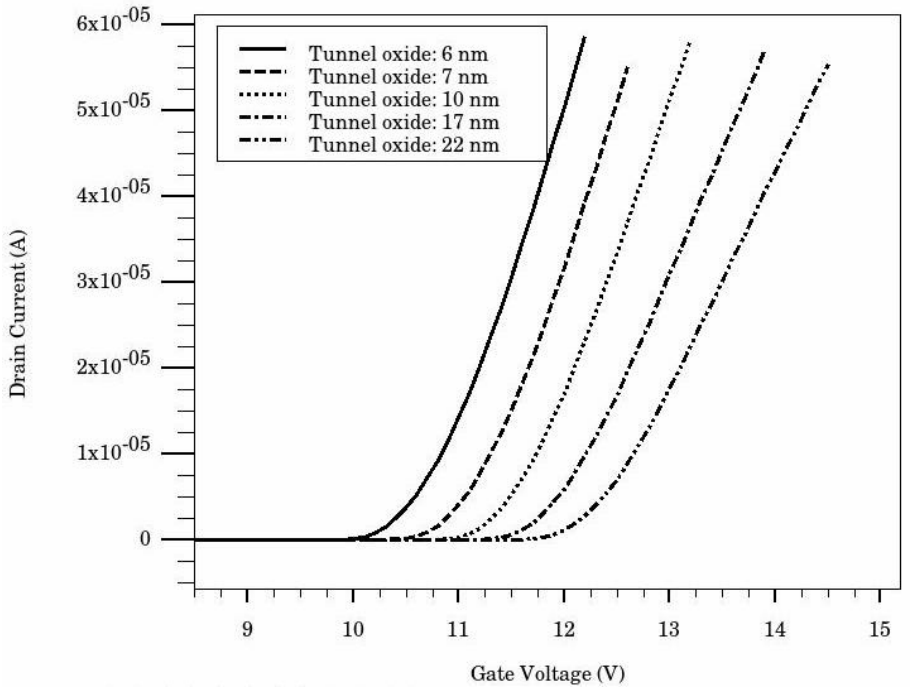

Figure 3 – Drain current versus gate voltage after programming

The transient characteristics of the floating gate integral charge versus logarithmically scaled time of the programming and erasing memory operation are demonstrated in Fig. 4 and Fig. 5, respectively.

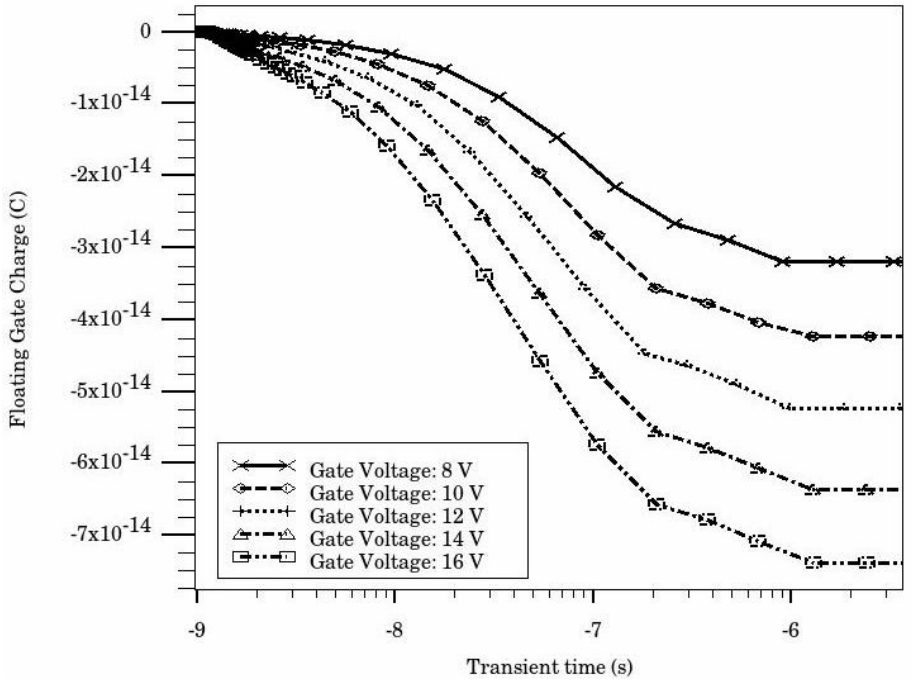

Figure 4 – Floating gate charge transient characteristics during programming

The programming operation is simulated for various voltages on the gate electrode. The transient characteristics allow us to estimate the duration of memory

operation and charge level obtained as a result of the operation. The programming time is appeared to be about 1 μs for the thickness of tunnel oxide 6 nm. The floating gate integral charge increases with the raise of the voltage applied to the gate electrode.

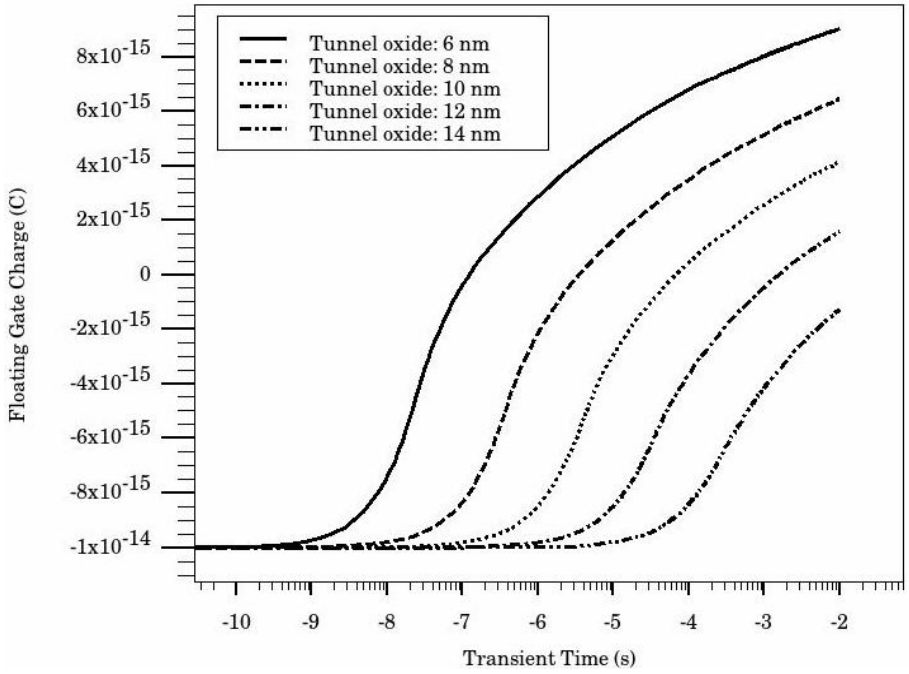

Figure 5 – Floating gate charge transient characteristics during erasing

The erasing operation is simulated by applying reduced bias 9 V to the drain of the memory cell with preliminary charged floating gate. The transient characteristics of the integral charge of the floating gate obtained for various tunnel oxide thicknesses are shown that erasing time is greatly dependent on the tunnel oxide thickness. The duration of the operation decreases when the oxide thickness is reduced. For tunnel oxide thickness of 6 nm a negative floating gate charge is reducing to zero level during 0.1 μs erasing process. The memory cell with tunnel oxide thickness of 10 nm can be erased to zero charge level during 0.1 ms. With the continuation of the erasing operation, the floating gate is charged positively.

Conclusions. The flash memory cell with the floating gate has been simulated using Silvaco TCAD software package. The control gate of the memory cell design investigated in the paper has reduced thickness and envelops the floating gate from top and sides providing good coupling for relatively thick inter-poly layer. I-V and transient characteristics of programming and erasing have been obtained for various tunnel oxide thicknesses and operating voltages. It has been demonstrated that the analyzed cell design is capable to provide a large shift of threshold voltages for programmed and un-programmed memory states, fast memory operations and

low operation voltages. Thus the considered floating gate transistor can be used as a low-profile and low-voltage memory cell to create effective flash memory devices.

REFERENCES

1. Advances in non-volatile memory and storage technology / Edited by Y.Nishi. – Elsevier Ltd., 2014. – 512 p.

2. Brewer J.E. Nonvolatile memory technologies with emphasis on flash / J.E. Brewer, M.G. Gill. – Wiley-IEEE Press, 2007. – 760 p.

3. Padovani A. Modeling strategies for flash memory devices / A. Padovani, L. Larcher, P. Pavan // Proc. of NSTI Nanotechnology Conf., 2011. – Vol. 2. – P. 762-767.

4. Zakaria M.R. An overview and simulation study of conventional flash memory floating gate device using concept F-N tunneling mechanism / M.R. Zakaria et al. // Proc. of V-th Int. Conf. on Intelligent Systems, Modeling and Simulation, 2014. – P. 775-780.

 ISSN 1562-9945