Solid-State Electronics 113 (2015) 144–150

Contents lists available at ScienceDirect

Solid-State Electronics

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

Self-consistent simulation on multiple activation energy of retention characteristics in charge trapping flash memory

Sangyong Park [a], Seongwook Choi [b,*], Kwang Sun Jun [b], HuiJung Kim [b], SungMan Rhee [b], Young June Park [b]

[a] Flash TD Team, Semiconductor R&D Center, Samsung Electronics Co. Ltd., Seoul, Republic of Korea

[b] Department of Electrical and Computer Engineering, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Available online 3 June 2015

Keywords:
Charge trap memory
SONOS
Data retention
Lifetime estimation
Non-Arrhenius behavior
Multiple activation energy

ABSTRACT

Non-Arrhenius behavior has been reported in a various temperature range for the retention time of CT Flash memories. In order to understand the physical origin of the multiple activation energy due to the non-Arrhenius behavior, we conduct a simulation study using a 3D self-consistent numerical simulator developed in-house. As a result, it is found that both vertical and lateral charge transport in the conduction band of nitride layer are responsible for the non-Arrhenius retention characteristic. Also, the tunneling current through the bottom oxide and a lifetime criteria are turned out to be the key parameters which determine the multiple activation energy.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The Charge Trap (CT) Flash memory, based on the nitride trap layer, is one of the most promising candidates for next Flash memory generation due to its ability to be scaled down [1]. Especially, for the 3D vertical NAND (VNAND) flash structure, the CT type cell is an inevitable choice considering the ease of processing the 3D structures [2,3]. For actual production of flash cells, it is important to understand and properly model the data retention characteristics to guarantee the desired lifetime of the device. The widely used method predicting the data lifetime is an extrapolation using the activation energy, which is represented by the slope of the Arrhenius plot, as used in the conventional Floating Gate (FG) cell [4]. Usually, this method is used under the assumption that the activation energy is constant (satisfying Arrhenius equation) within the temperature of interest.

However, it has been continuously reported that the Arrhenius plot shows not a straight line but shows the multiple activation energies [5,6]. Several experimental data showing the non-Arrhenius relation is summarized in Fig. 1(a) for FG cells and Fig. 1(b) for CT memory cells [5–11]. This characteristics give rise to the problem that the expected lifetime varies according to the choice of the activation energy depending on the stress temperatures.

It is more interesting that there are different kind of non-Arrhenius behavior in Fig. 1: A decrease of activation energy ((ii)–(iii) in Fig. 1(a), (iii) in Fig. 1(b)), an increase of activation energy ((iv) in Fig. 1(b)) and recovery after the decrease of activation energy ((iv) in Fig. 1(a)). Therefore, for more accurate prediction of the device lifetime, a physical mechanism of the retention phase related to the non-Arrhenius behavior should be understood.

Some authors have been reported about the origin of the multiple activation energy in the CT memory but limited to the decreasing activation energy only [12]. Also, their theoretical model is based on the 1D model which cannot consider the lateral charge redistribution [12]. This may be problematic because, until recently, it has been continually reported that the lateral charge transport gives a significant effect to the charge retention [13,14]. The lateral migration was proved in many different ways, for example, the direct probing method (AFM) [15,16], electrical measurement [13,17] and theoretical simulation [13]. This lateral leakage cannot be eliminated by separating the nitride region for each cells because of the reliability degradation due to the cutting of the nitride layer [17]. Moreover, the nitride layer cannot be separated at all for 3D VNAND structure. Hence, it is crucial to consider the comprehensive models including the lateral transport for modeling the retention phase.

However, to the authors best knowledge, most of previous retention modeling and simulation works do not comprehensively considers the structural and physical effects; Some model does not consider the electron transport in the nitride region [12]. Others tried to include all the physical equations, yet they are limited in

* Corresponding author.
E-mail address: church7@snu.ac.kr (S. Choi).

http://dx.doi.org/10.1016/j.sse.2015.05.026
0038-1101/© 2015 Elsevier Ltd. All rights reserved.

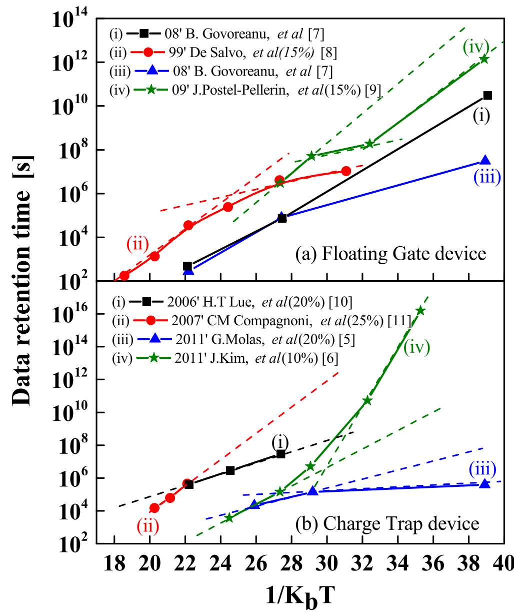

Fig. 1. Arrhenius plot for the retention time extracted from experimental data of (a) Floating Gate (FG) cells and (b) Charge Trapping memories. The percentage represents the criteria for the retention time.

the 1 dimensional simulation [17]. There is a report considering the 2D lateral transport [18]but it does not calculate the tunneling component and the channel region. Only recently, one can find some efforts to include 3D comprehensive model to the retention phase simulation [19,20] including author's previous work [21].

In this context, it is worthwhile to study the retention characteristics with 3D comprehensive physical models including the lateral/vertical transport, trapping/de-trapping and tunneling in a self-consistent manner. Hence, in this paper, we used the self-consistent 3D CT memory simulator [21–23] for investigating the non-Arrhenius behavior to explain all the trends in Fig. 1. In this way, the relationship between the multiple activation energies according to the temperature range and the stored charge transport in lateral as well as vertical direction could be understood.

The P/E cycling also play a central role for the retention as well as the lateral diffusion [14]. But, please note that it is beyond the scope of this paper to consider the new trap generation during the P/E cycling. Here, the pre-stress characteristic condition is only considered here because the pre-stress measurement data in Fig. 1 indicates the non-Arrhenius behavior. Based on the work on the pre-stress condition, we expect that the acceleration of retention due to the P/E cycling also could be accurately modeled.

2. Simulation structure and physics

2.1. Simulation structure

The simulation structures are based on SA-STI (Self-Aligned STI) SANOS ($Si/Al_2O_3/Si_3N_4/SiO_2/Si$) device with 4 nm tunneling oxide, 8 nm Si_3N_4, 14 nm Al_2O_3, and 30 nm channel length. In order to include the actual CT memory structure, the trap layer (Si_3N_4) is continuous along the active line as shown in Fig. 2. Because of the continuous trap layer, there is a leakage current in the lateral direction through this continuous layer. In this study, three adjacent cells are simulated and the device at the center (Gate 1 in Fig. 2) are programmed only while the adjacent cells (Gate 0 and Gate 2 in Fig. 2) are biased with 0 V. More study of the effect on

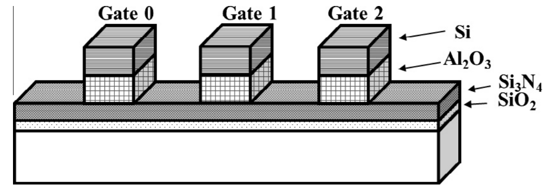

Fig. 2. Simulated structures of the SA-STI SANOS device with continuous nitride (Si_3N_4). The electron transport in the whole nitride region is calculated. We use Gate 1 for the program and retention while other cells are biased with 0 V.

the adjacent cells by the lateral charge spreading can be found in elsewhere [21].

2.2. Simulation models

For the comprehensive model for the CT memory cells, we included following key physics in our simulation code: (1) tunneling from the nitride trap to substrate conduction band, (2) tunneling from the nitride conduction band to substrate conduction band, (3) transport of trapped electron (via SRH and Poole Frankel), (4) transport in the nitride conduction band, (5) transport in the substrate channel region. We used a 3D in-house simulation code built for CT memory devices. Some key master equations and details for each mechanisms are summarized in the following [21–23].

2.2.1. The tunneling current through the tunneling oxide

The non-local tunneling current (J_n) through the tunneling (bottom) oxide is calculated with

$$J_n = vqnT_{ox} \tag{1}$$

where n is a charge concentration in the conduction band or traps, T_{ox} is a tunneling probability of the bottom oxide calculated by the WKB approximation and v is a hitting frequency. For the programming phase, n is the charge concentration in the substrate v is a hitting frequency in the quasi-bound state which is an electric field dependent at the channel surface.

2.2.2. The trapped charge loss mechanism

The transport of trapped charge is calculated using the Poole–Frankel model which can be written as

$$\frac{\partial n_c}{\partial t} = -\nabla \cdot J_n - c_n N_T n_c + v \exp\left(-\frac{E_t - \beta\sqrt{E}}{k_B T}\right) n_T \tag{2}$$

where n_c is a charge concentration in the conduction band, N_T is a concentration of traps with an energy level E_t, β is the Frenkel constant, E is a local electric field, c_n and v are constants. This equation is solved for the entire nitride region in 3D space.

2.2.3. The transport of carriers in the nitride conduction band

The electron transport in the nitride conduction band is solved with the drift–diffusion formalism than can be written as

$$J_n = q\mu_n n_c E + qD_n \nabla n_c \tag{3}$$

where μ_n is a mobility and D_n is a diffusion constant. This equation is also solved for the entire nitride region in 3D space.

2.2.4. Density gradient

The electron distribution in the channel region is calculated with the density-gradient formalism with quantum potential

$$\nabla \cdot (b_n \nabla \sqrt{n}) - \frac{\sqrt{n}}{2}\psi_{qn} = 0 \tag{4}$$

Download English Version:

https://daneshyari.com/en/article/747848

Download Persian Version:

https://daneshyari.com/article/747848

Daneshyari.com