Microelectronic Engineering 96 (2012) 40–44

Contents lists available at SciVerse ScienceDirect

Microelectronic Engineering

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

Flash program modeling using nonquasi-static and tunneling techniques

Scott C. Wolfson *, Fat D. Ho

Univ. of Alabama in Huntsville, Department of Electrical and Computer Eng., Huntsville, AL 35899, USA

ARTICLE INFO

Article history:
Received 8 July 2011
Received in revised form 22 November 2011
Accepted 27 December 2011
Available online 28 March 2012

Keywords:
Flash memory
MOS memory circuits
Semiconductor device models
Simulation

ABSTRACT

We present a detailed and accurate physics based transient simulation for modeling flash memory programming characteristics using nonquasi-static and tunneling equations versus the typical Lucky-Electron Model. The result is a set of simple expressions that were originally developed for a MOSFET and adapted for use in floating gate memory. Of greater importance is the extensive use of physical parameters as opposed to the scale factors and probabilities used in other models. This technique allows floating gate memory designers to determine the nominal programming characteristics of single-level and multi-level memory cells prior to the fabrication process. This technique also allows designers to determine the effects of fabrication tolerances on the performance of the memory cell. The accuracy of this model was validated through comparison with experimental data and simulation results presented in several publications.

Published by Elsevier B.V.

1. Introduction

In today's cell phone and mobile computing industry there is an increasing need for flash memory with faster performance while requiring less power to operate and increased memory density. Through accurate simulation the effects of substrate material properties, temperature, device scaling, terminal voltage levels, doping concentration, terminal voltage slew rates, programming pulse durations and magnitudes can be analyzed prior to fabrication. This can aid in determining how design changes and fabrication tolerance variations effect performance and can greatly reduce the required design time, design cost and time to market for products using this technology.

The structure of a flash memory cell as shown in Fig. 1 is similar to that of an EEPROM where an intermediate or "floating" gate exists between the control gate terminal and the substrate surface. The floating gate acts as a capacitor capable of storing charge and the amount of stored charge is directly related to a shift in the overall threshold voltage of the memory cell. Likewise, the movement of electrons to the floating gate can be viewed as a gate current. Therefore, the keys to modeling the programming characteristics are accurately determining the gate current and determining the effects of stored charge on the overall threshold voltage.

While researching this topic we discovered that many, if not all, of the existing models were based on the probability intensive equations presented in [1] to calculate the programming gate current. The intent of this paper is to demonstrate an alternate and more physics base method to describe this process. The weak points of the Lucky Electron Model in [1] are its use of a series of probabilities to predict the channel to floating gate current and it only considers the electron injection at a point near the drain where the oxide electric field is at a maximum. In reality, electrons can traverse the oxide barrier at any point in the channel with the maximum contribution occurring at the drain end of the channel region. To produce a more physics based model we calculated the gate current using a nonquasi-static method presented in [2] combined with the tunneling equations presented in [3–5]. This new technique determines the gate current by integrating across the channel region to determine the contributions attributed by the charge coupling between the channel surface and the floating gate and tunneling between the channel surface and the floating gate. Also included is the minor contribution due to tunneling between the drain/source and the floating gate overlap region.

The following sections detail the equations used in the simulation development, simulation sequence and results achieved. The inclusion of the simulation sequence is meant to provide the reader with insight into the methods and numerical techniques used to perform the integration in the gate current equation and determining the surface potential at any point within the channel.

2. Model initial conditions

The capacitive equivalent circuit (Fig. 2) and the cross-sectional view of a floating gate device (Fig. 1) is the basis for a majority of the floating gate models [2,3]. From Figs. 1 and 2 the derivations of the initial floating gate charge (Q_{fg}) and initial floating gate voltage (V_{fg}) equations are possible. To increase the accuracy of the model we incorporated the contact potentials where $V'_g, V'_d, V'_s,$ and V'_b

* Corresponding author. Tel.: +1 256 876 8713; fax: +1 256 876 8688.
E-mail address: wolfsos@uah.edu (S.C. Wolfson).

0167-9317/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.mee.2011.12.019

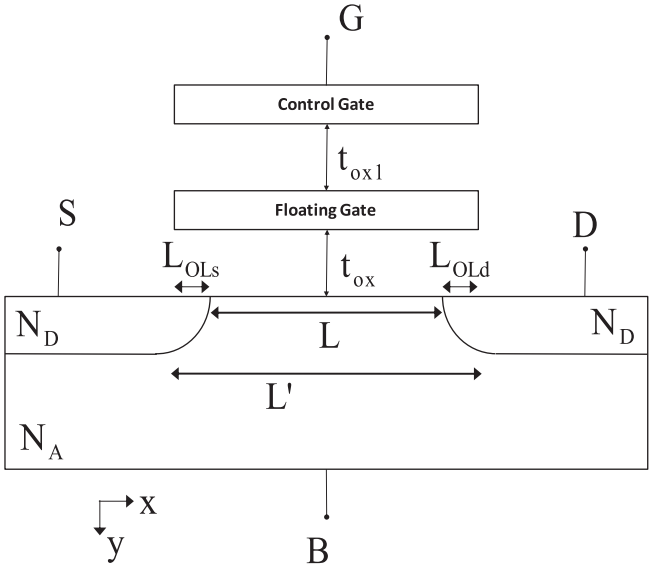

Fig. 1. Cross-sectional representation of an n-channel floating-gate memory cell, where t_{ox1} – control gate (CG) to floating gate (FG) oxide thickness, t_{ox} – FG to substrate oxide thickness, L_{OLd}, L_{OLs} – drain and source overlap lengths, W, L' – cell width and length, L – channel length.

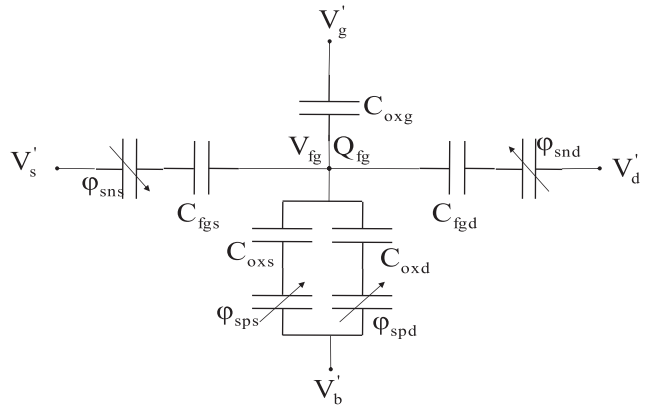

Fig. 2. Schematic representation of the floating gate capacitor equivalent circuit, where $C_{oxg} = (\varepsilon_{ox}/t_{ox1})WL'$ – CG to FG capacitance (F), $C_{oxs} = 0.5(\varepsilon_{ox}/t_{ox})\ WL$ – initial FG to source side substrate capacitance (F), $C_{oxd} = 0.5(\varepsilon_{ox}/t_{ox})\ WL$ – initial FG to source side substrate capacitance (F), $C_{fgd} = (\varepsilon_{ox}/t_{ox})WL_{OLd}$ – FG to drain terminal capacitance (F), $C_{fgs} = (\varepsilon_{ox}/t_{ox})WL_{OLs}$ – FG to source terminal capacitance (F), $C_t = C_{oxg} + C_{oxd} + C_{oxs} + C_{fgd} + C_{fgs}$ – total capacitance (F), φ_{snd}, φ_{sns} – substrate referenced surface potential between the FG and the drain/source terminals (V), φ_{spd}, φ_{sps} – substrate referenced surface potential between the FG and the drain/source sides of channel surface (V).

represent the terminal voltages plus the corresponding contact potentials. Likewise, V_G, V_D, V_S, and V_B represent the terminal voltages only. It is also important to note that the surface potential and capacitance of the channel region is not constant due to the charge sharing contributions of the drain and source. For this reason the drain and source sides of the channel region will be considered separately.

Under steady state conditions the equation for the initial floating gate charge is:

$$Q_{fg} = C_t V_{fg} - C_{oxg}V'_g - C_{fgd}(\varphi_{snd} + V'_d) - C_{fgs}(\varphi_{sns} + V'_s) - C_{oxd}\left(\varphi_{spd} + V'_b\right) - C_{oxs}\left(\varphi_{sps} + V'_b\right) \quad (1)$$

The equation for the initial floating gate voltage is easily produced by rearranging the terms in (1) and solving for V_{fg}.

$$V_{fg} = \left(\frac{1}{C_t}\right)\left[Q_{fg} + C_{oxg}\left(V'_g\right) + C_{fgd}(\varphi_{snd} + V'_d) + C_{fgs}(\varphi_{sns} + V'_s) + C_{oxd}\left(\varphi_{spd} + V'_b\right) + C_{oxs}\left(\varphi_{sps} + V'_b\right)\right] \quad (2)$$

A good initial value for φ_{sns} and φ_{snd} is 0.0 V while a good initial value for φ_{spd} and φ_{sps} is $2\varphi_{fp}$ where the Fermi potential φ_{fp} was determined using the equations presented in [1].

A value for the initial floating gate charge can now be calculated using the following equations and the parallel plate capacitance values:

$$V_{tcg\_initial} = \frac{C_t}{C_{fg}}V_{tfg} \quad (3)$$

$$Q_{fg} = C_{fg}(V_{tcg\_initial} - V_{tcg}) \quad (4)$$

where $V_{tcg\_initial}$ – initial threshold voltage w.r.t. the CG ($Q_{fg} = 0$), V_{tcg} – initial threshold voltage w.r.t. the CG ($Q_{fg} \neq 0$), V_{tfg} – initial threshold voltage w.r.t. the FG.

One can easily surmise that the equation for V_{tfg} is the standard MOSFET threshold voltage equation by realizing the cell structure of Fig. 1 is identical to a MOSFET when the control gate is removed. Also from this point forward the device capacitances associated with the channel surface will be determined using the equations presented in [6].

3. Gate current component due to charge coupling

The first component of the gate current $i_G(x,t)$ can be attributed to the charge coupling between the substrate surface and the floating gate. The method for deriving this current equation originated from the nonquasi-static continuity equation presented in [2]. The equation is:

$$\frac{\partial i_G(x,t)}{\partial x} = W\frac{\partial q'_g(x,t)}{\partial t} \quad (5)$$

where $q'_g(x,t)$ – gate charge per unit area (C/cm^2), x – channel position (cm), t – time (s).

Rearranging (5) produces the following:

$$i_G(t) = WL\frac{d}{dt}\int_0^1 q'_G(x,t)dx \quad (6)$$

where

$$q'_G(x,t) = C'_{ox}[V_{FG}(t) - V_{FB} - \psi_s(x,t)] - q'_o \quad (7)$$

where $C'_{ox} = (\varepsilon_{ox}/t_{ox})$ – FG to substrate capacitance per unit area (F/cm^2), $V_{FG}(t)$ – FG to source voltage at time t (V), V_{FB} – flat-band voltage (F), $\psi_s(x,t)$ – source reference surface potential relative to channel position x at time t (V), q'_o – charges in the oxide material (C/cm^2).

It is important to note that the gate current Eq. (6) takes on the form $i = CdV/dt$ when the equation for $q'_g(x,t)$ of (7) is inserted; therefore, this current component can be attributed to charge coupling. The complete solution for (6) is complicated by the fact that the surface potential is unknown for every point in the channel. The determination of this surface potential is accomplished by first realizing that under steady state conditions the current at the source, drain and every point in the channel region is assumed to be equal. This assumption produces the following equality [2]:

$$I_0(t) = I_{DS}(t) = \frac{W}{L}[f(\psi_s(L,t) - f(\psi_s(0,t))] = \frac{W}{x}[f(\psi_s(x,t)) - f(\psi_s(0,t))] \quad (8)$$

where $f(\psi_s(x,t))$ – function of the surface potential at channel position x where $f(\psi_s(0,t))$ is at the source and $f(\psi_s(L,t))$ is at the drain.