Extended Abstracts of the 2002 International Conference on Solid State Devices and Materials, Nagoya, 2002, pp. 614-615

P10-4

Application of Cell Model to Adaptive Programming in SSI Flash

Chun-Mai Liu, Lawrence D. Engh, and Albert V. Kordesch

Winbond Electronics Corporation of America, 2727 North First St., San Jose, CA 95134, USA
(408-544-1773; fax 408-544-1786; cmliu0@winbond.com.tw)

Introduction

The source-side-injection (SSI) flash memory cell has been recently employed as the base of a multilevel storage technology [1, 2]. Method for measuring cell capacitance coupling ratios was reported [3]. Procedures for building a two-transistor cell modeling have been established [4] and a macro gate current model developed [5]. Consequently, accurate circuit model of the flash cell can be used for the development of program methods and for the optimization of memory circuit design.

Programming methods for multilevel storage were developed using either an iterative program scheme [6], or an open-loop scheme [7]. The iterative is usually more accurate, but lacks speed, while the open-loop has speed, but lacks precision. A parallel programming architecture can boost the overall memory system speed by as many as several thousand times, however its application is limited to page write and to higher density memories where overhead circuitry is efficiently shared. Consequently, improvement of single-cell programming speed is needed for general application.

In this work we have derived a concise cell model based on a two-transistor compact cell model with macro gate current module and integrated it into the algorithm equations for developing program algorithm with fast convergence.

Cell Model

Fig. 1 shows the SSI flash cell structure with two transistors merged in a split-gate configuration. This structure can be divided into a select gate transistor and a floating gate transistor. The gate voltage of FG (floating gate) transistor is found well represented by linear capacitive couplings of Vsg (select gate, the same as WL), Vcs (common source), Vc (common node) and Vb (substrate) under READ condition. TCAD simulations and experimental data prove that the representation of fixed coupling capacitance ratios can extendable into cell PROGRAM condition [3]. Accordingly, Fig. 2 shows the basic cell macro [5], which uses voltage controlled voltage sources and charging capacitor to generate Vfg, here the current source Ifg is the gate current by the source-side injected electrons.

Floating Gate Voltage

The floating gate voltage of the cell can be written as follows[1]:

$$V_{fg} = K_{sg}V_{sg} + K_{cs}V_{cs} + K_cV_c + \left(\frac{Q_{fg}}{C_T}\right) + W_{fg0} \quad (1),$$

where the K_{sg}, K_{cs} and K_c are the floating gate FG, common source CS and common node C voltage coefficients respectively. C_T is the total capacitance of the floating gate; V_{fg}, V_{sg}, V_{cs}, and V_c the voltages respectively; and Q_{fg} the FG stored charge. These coefficients, K's can be interpreted as "capacitance ratios", such that

$$K_{sg} = \frac{C_{sg}}{C_T},\ K_{cs} = \frac{C_{cs}}{C_T},\ K_c = \frac{C_c}{C_T} \quad (2),$$

where C_{sg}, C_{cs} and C_c as depicted in Fig. 1, representing the capacitances to select gate SG, common source CS and to common node C respectively, W_{fg0} is an integration constant.

Read Model

Fig. 3 shows the simulated source follower voltage Vsf of the flash READ operation vs. FG stored charge using the compact cell model. Here the read current Id is kept constant. This V_{sf} vs. Q_{fg} curve is found well approximated by a straight line:

$$V_{sf} \approx V_0 + \Lambda\left(\frac{Q_{fg} - Q_0}{C_T}\right) \quad (3),$$

where V_0, Q_0, and Λ are constants[2]. Here the source follower voltage Vsf is tracking common node voltage Vc, $V_{sf} \approx V_c(Read)$.

Program Model

Following the Lucky Electron Model [8], and under fixed select gate voltage Vsg and program current I_{bit} (= Id, bit line current), the gate current, Ig, can be described by a macroscopic equation[3] [5]:

$$I_g = C_1 I_{bit}\left(\frac{V_{fg} + k_1V_{cs} + B_{sg}}{C_2}\right)^2 \exp\left(\frac{-C_2}{V_{fg} + k_1V_{cs} + B_{sg}}\right) \quad (4)$$

Here C_1, C_2, k_1 and B_{sg} are constants[4]. The actual experimental $Log\ I_g$ curve varies only modestly as V_{fg}. Therefore within the range of cell operation, the gate current may be approximated as follows[5]:

$$I_g \approx I_{bit} \exp\left[\alpha_{00}\left(V_{fg} + k_1V_{cs}\right) - C_{00}\right] \quad (5)$$

Here α_{00}, C_{00}, and k_1 are constants.

Merged Model

As a program pulse of height Vpp (=Vcs) and of duration Δtpp is applied to the cell, Eq. (5) can be integrated to calculate the change from the initial state, denoted as "1", to final state, denoted as "2":

$$Q_{fg1} - Q_{fg2} = \int_1^2 I_g dt = I_{bit}\Delta tpp\ \exp\left[\alpha_{00}\left(V_{fg1\leftrightarrow 2} + k_1V_{cs}\right) - C_{00}\right] \quad (6)$$
$$= I_{bit}\Delta tpp\ \exp\left[\alpha_0 \frac{Q_{fg1\leftrightarrow 2}}{C_T} + K_{00}Vcs - C_{00}\right]$$

Here α_0, and K_{00}, are constants, and Q_{fg1} and Q_{fg2} denotes the FG stored charge of the initial and final states respectively. The Vfg dependency is further transformed to Qfg dependency in the second line[6]. We use $V_{fg1\leftrightarrow 2}$ .and $Q_{fg1\leftrightarrow 2}$ to denote the "average" V_{fg} and Q_{fg} values[7] between states "1" and "2", i.e. $V_{fg2} \le V_{fg1\leftrightarrow 2} \le V_{fg1}$ and $Q_{fg2} \le Q_{fg1\leftrightarrow 2} \le Q_{fg1}$. Employing Eq. (3), the FG stored charge Q_{fg} in Eq. (6) can be changed to the source-follower read voltage V_{sf}. Finally, with fixed program duration Δtpp and program current I_{bit}, we arrive at the master equation representing the concise cell model:

$$Log_{10}\Delta V_{sf} = Log_{10}\left(V_{sf1} - V_{sf2}\right) = mV_{sf1\leftrightarrow 2} + mKcs_{00}Vpp - G_{00} \quad (7),$$

where m, Kcs_{00} and G_{00} are model parameters. Here $V_{sf1\leftrightarrow 2}$ denotes an "average" V_{sf}, i.e. $V_{sf2} \le V_{sf1\leftrightarrow 2} \le V_{sf1}$. Fig. 4 shows a set of experimental $Log_{10}\Delta V_{sf}$ vs. V_{sf} data from the 0.5um SSI cell, where $V_{sf} = V_{sf2}$ and a set of fit lines according to the Log model, Eq. (7).

Adaptive Program

Adaptive program entails iterative program and read cycles that engage in real-time computing, where the program voltage Vpp (= Vcs) is determined from information of the current cell Vsf, the target Vsf and a cell model, such that the target Vsf can be fast approached. Fig. 5 shows the overall flowchart of an adaptive program scheme we have devised utilizing the concise model.

Specifically, by turning ΔVsf into a predictor ΔVsf_n, and $V_{sf1\leftrightarrow 2}$ to a target $Vsfnext_n$, and G_{00} to the cell speed parameter G_n, Eq. (7), can be readily rearranged for computing the pulse height Vpp of the program pulse with fixed program duration Δtpp and program current Ibit:

$$Vpp_n = K_p Log\Delta Vsfp_n + G_n - K_V Vsfnext_n \quad (8),$$

where n is the recursive index; Kp ~ 1/(m*Kcs_{00}; Kv ~ m; $\Delta Vsfp_n$ = the desired change of Vsf for the nth iteration; $Vsfnext_n$ = the target of the nth iteration. $Vsfnext_n$ and $\Delta Vsfp_n$, are related by:

$$Vsfnext_n = Vsf_n - \Delta Vsfp_n \quad (9),$$

where Vsf_n = the measured Vsf *before* the nth iteration.

Next, the error voltage is used for G-factor adaptation:

$$G_n = G_{n-1} + E\,(Vsf_n - Vsfnext_{n-1}) \quad (10),$$

where E is the strength of error correction; $Vsfnext_{n-1}$ = the target of the (n-1)th iteration.

The predictor $\Delta Vsfp$ can be expressed as a general function F:

$$\Delta Vsfp_n = F(T\arg et, Vsf_n, n) \quad (11),$$

where Target is the Vsf value of the final programming target. One choice of function F, which has increasingly aggressive targeting, is:

$$\Delta Vsfp_n = F(T\arg et, Vsf_n, n) = \frac{n}{Ntotal}(Vsf_n - T\arg et) \quad (12),$$

where Ntotal is the total number of iteration, and is set to 8 for the 0.5um cell based on empirical results.

Experimental and Results

These algorithm equations have been realized in both analog and digital forms, the former uses on-chip analog computing circuit and the latter digital computer control. Typically, the first several cycles produce very small changes to prevent over trim and adapt the model. The last two cycles drive the final accuracy. Fig. 6 shows the experimental data, where fast convergence to targets is demonstrated.

Conclusions

In this work, we have derived a concise cell model for real-time computing. This model has been successfully applied to design adaptive program method for multilevel flash storage devices.

Acknowledgments

We would like to thank Winbond management for their support and encouragement. We also would like to thank Dr. Geoff B. Jackson for stimulating discussions and for formulating the algorithm equations.

1 The substrate term $K_b V_b$ is left out, since $K_b \approx 0$.

2 Λ can be expressed as $1\Big/\left(1 - K_c + \frac{\gamma_{fg}}{2\sqrt{2\Phi_p + V_0}}\right)$ where γ_{fg}, Φ_p are FG transistor body factor and bulk potential respectively.

3 Although node C, the virtual source of FG is a poorly-defined geometrical point under program condition. Its voltage is tracking Vfg, i.e. Vc ~ Vfg.

4 The k_1Vcs term corresponds to the CS DIBL effect. Bsg depends on Vsg, Vtsg and Vtfg. The effects of Eox and of Em are approximated by –C2/(Vfg + k1Vcs + Bsg).

5 Expanding the term $\frac{C_2}{V_{fg} + k_1 V_{cs} + B_{sg}}$ at fixed point of V_{fg0}, V_{cs0}.

6 by assuming Vc ~ Vfg under program condition.

7 Where the $V_{sf1\leftrightarrow 2}$ and $Q_{sf1\leftrightarrow 2}$ situate will depend on the actual program condition such as pulse height Vpp and duration Δtpp, and the cell initial condition V_{sf1} or Q_{sf1}.

References

[1] C.-M. Liu, et al., *Proc. of VLSI-TSA*, pp. 187, 1999.
[2] C.-M. Liu, et al., *Extended Abstract of SSDM*, September, pp. 542, 1999.
[3] C.-M. Liu, et al., *Extended Abstract of SSDM*, August, pp 288, 2000.
[4] S.-P. Sim, et al., *Extended Abstract of SSDM*, September, pp. 544, 2001.
[5] C.-M. Liu, et al., *Extended Abstract of SSDM*, September, pp. 542, 2001.
[6] H. Tran, et al., *ISSCC Digest of Technical Papers*, p 270–271, Feb. 1996.
[7] M. Bauer, et al., *ISSCC Digest of Technical Papers*, p 132–133, Feb. 1995.
[8] Simon Tam, et al. *IEEE Trans. Electron Devices*, vol. ED-31, no. 9, pp. 1116, 1984

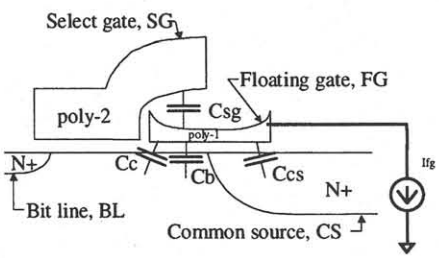

Fig. 1 The SSI flash cell--two transistors in a split-gate configuration with four components of FG capacitance and a current source for Ifg.

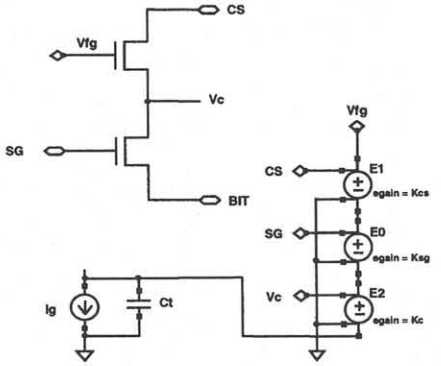

Fig. 2 Cell macro based on the two-transistor framework— voltage controlled voltage sources and a current source for Ifg.

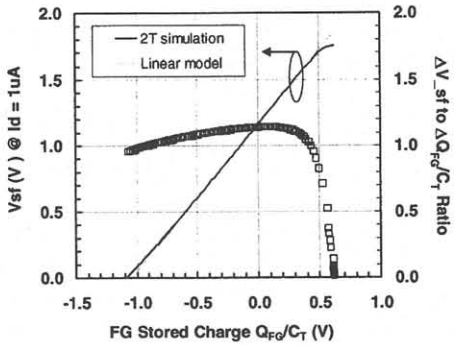

Fig. 3 Source follower Vsf vs. Floating gate stored charge Qfg— curve of 2-transistor compact model shows a near constant slope (Ibit = 1uA).

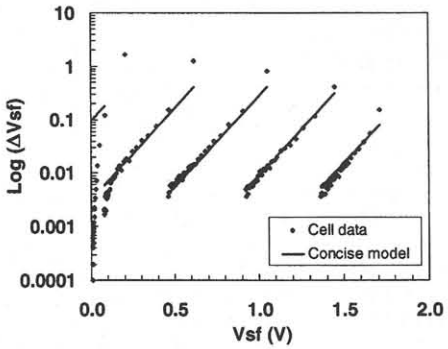

Fig. 4 The Log ΔVsf vs. Vsf— experimental data and calculated lines based on concise model (Ibit = 1uA).

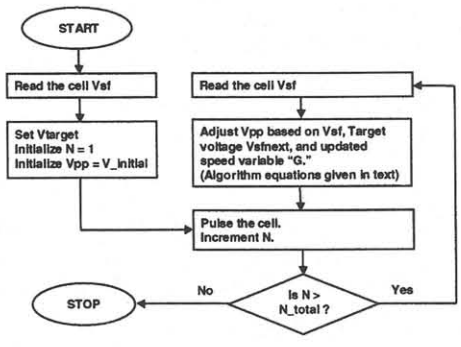

Fig. 5 Flowchart of the Adaptive algorithm shows the computation of Vpp, G and other variables.

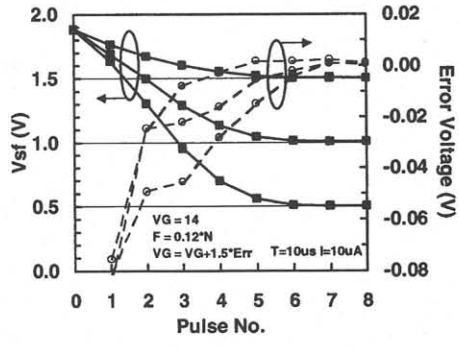

Fig. 6 Adaptive Programming— Vsf voltages and Error voltages measured from a digitally controlled test system.