

Variation in Natural Threshold Voltage of NVM Circuits due to Dopant Fluctuations and its Impact on Reliability

David Burnett, Jack Higman, Alex Hoefler, Chi-Nan Brian Li, and Peter Kuhn

Embedded Memory Center, Semiconductor Products Sector, Motorola,
6501 William Cannon Dr., MD: OE341, Austin, TX 78735
Phone: (512) 895-4631, Fax: (512) 895-8605, E-mail: david.burnett@motorola.com

Abstract

The statistical distribution of the natural threshold voltage (V_{T0}) of 512k-bit NVM circuit arrays has been studied for two different technologies. The major source of the V_{T0} variation is dopant fluctuations of the NVM well. An analytical model for the dopant fluctuations provides excellent agreement with the measured circuit V_{T0} variation and NVM cell mismatch for both technologies. The reliability implications of the V_{T0} variation are considered using charge leakage models for data retention.

Introduction

The intra-die V_T variations in logic, SRAM, and DRAM circuits due to intrinsic dopant fluctuations is a major concern as devices continue to scale [1-4]. For NVM circuits, the intra-die natural or charge-neutral V_T (the bitcell V_T with no charges on the floating gate, V_{T0}) variation can result in a large variation in the oxide electric-field under self-bias (proportional to $(V_T - V_{T0})$). This creates a reliability concern since recent data retention studies of cycled NVM circuits have shown the charge leakage to have an exponential dependence on the oxide field [8-11]. The V_{T0} variation can also have a large impact on the read disturb field, and it sets the minimum width of the erase and program distributions prior to compaction. Thus, it is important to understand the fundamental cause of the V_T variations and how to model it.

NVM Device V_T Mismatch and Model

Figure 1 shows the V_{T0} distributions of a typical 512k-bit array from two different NVM technologies. The NOR technology utilizes hot-carrier programming and source-side Fowler-Nordheim (FN) tunneling erase and was processed in a 0.25 μ m embedded technology [5]. The Uniform-Channel-Program-Erase (UCPE) technology uses FN tunneling for both programming and erase [6]. A UV-erase was carried out at the end of processing to remove any floating gate charges prior to the measurements. As observed, the total V_{T0} variation is over 350 mV for the UCPE array and over 700 mV for the NOR array.

Histograms of the threshold voltage mismatch of adjacent NVM cells from mini-arrays with floating-gate access (FG) and with standard bitcells (control gate, CG) are shown in Figs. 2 and 3. Each pair of cells shares the same drain

contact and has separate gate and source lines. The V_{T0} mismatch distributions follow a normal distribution for both types of mini-arrays. The standard deviation of the FG devices is 60% that of the CG bitcells, which is the same as the gate coupling ratio, α_G .

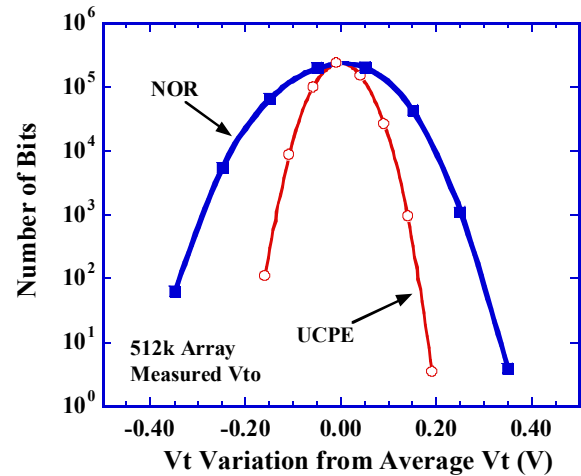


Fig. 1. Natural V_T (V_{T0}) distribution of a typical 512k-bit array from two different NVM technologies.

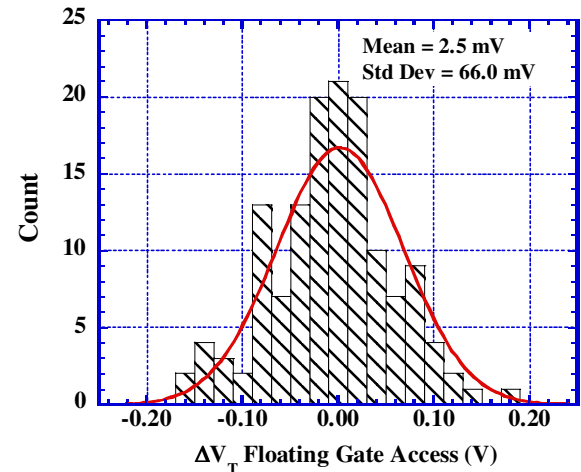


Fig. 2. Histograms of V_{T0} mismatch distribution for the floating-gate access device with nominal gate length and width from test structures with the NOR technology.

Figure 4 shows the standard deviation of the V_{T0} mismatch of the FG and CG devices with different L_{eff} for the NOR technology. An excellent fit to the data is generated using an

analytical model for the standard deviation of V_{T0} due to dopant fluctuations [7]:

$$\sigma V_{T0} = \frac{\sqrt[4]{4q^3 \epsilon_{Si} \phi_B N_A}}{\sqrt{3W_{eff} L_{eff}}} \left[\frac{k_B T}{q \sqrt{4q \epsilon_{Si} \phi_B N_A}} + \frac{T_{OX}}{\epsilon_{OX}} \right] \quad (1)$$

where N_A is the well doping density, T_{OX} is the tunnel oxide thickness, and W_{eff} and L_{eff} are the effective channel width and length. For a matched pair of devices, the standard deviation of the V_{T0} mismatch due to dopant fluctuations, $\sigma(\Delta V_{T0})$, should be $\sqrt{2} \sigma V_{T0}$ since V_{T0} is normally distributed. Then, the floating gate and control gate mismatches are modeled as $\sigma(\Delta V_{T0,FG}) = \sqrt{2} \sigma V_{T0}$ and $\sigma(\Delta V_{T0,CG}) = \sigma(\Delta V_{T0,FG})/\alpha_G$.

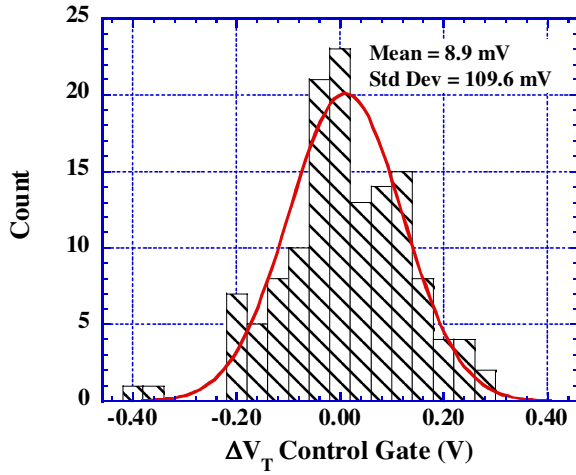


Fig. 3. Histograms of V_{T0} mismatch distribution for the control gate device with nominal gate length and width from test structures with the NOR technology.

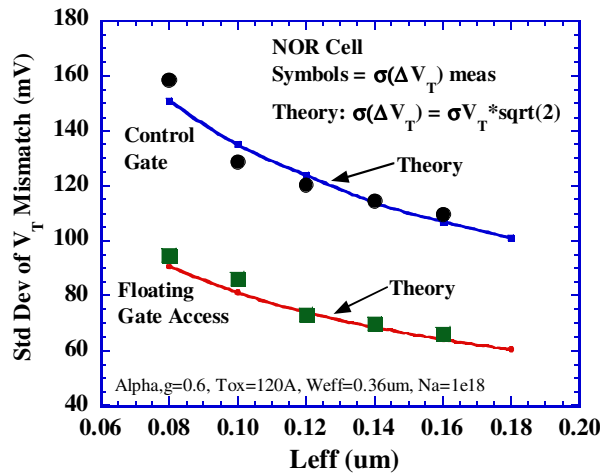


Fig. 4. Standard deviation of V_{T0} mismatch for the NOR cell with $Leff$ for both the control gate and floating gate access cells. The theory shows excellent agreement with the measured data.

Modeling of the Circuit V_t Distribution

In order to model the circuit V_{T0} distribution, the V_{T0} variation from intra-die process variation must be considered as well. The expected bitcell V_{T0} variation with active width, gate length, and floating gate width variations is much less than the expected V_{T0} variation from the dopant fluctuations for both technologies (Figs. 5 and 6). Other intra-die process variations, such as tunnel oxide and ONO thickness, should have even less effect. The $\sigma V_{T0,CG} = \sigma V_{T0}/\alpha_G$ of the UCPE technology is less than half that of the NOR technology primarily due to the difference in well doping. The UCPE technology uses the HV PWell while the NOR technology has about 6X higher doping density to provide adequate short-channel effects for the programming operation.

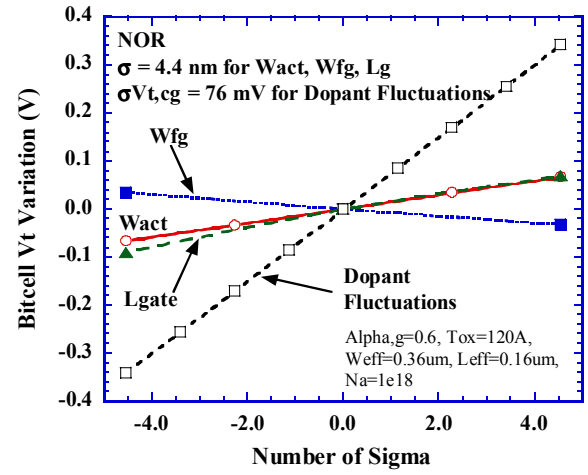


Fig. 5. Bitcell V_T variation for the NOR technology due to dopant fluctuations and process variations. The active, floating gate, and gate length dimensions are assumed to have a ± 20 nm variation across the 512k-bit block.

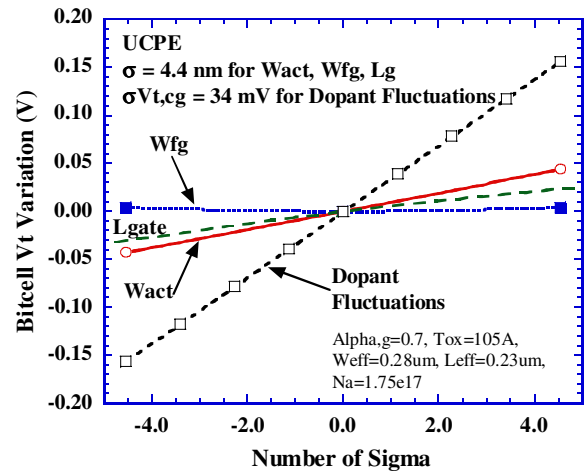


Fig. 6. Bitcell V_T variation for the UCPE technology due to dopant fluctuations and process variations. The active, floating gate, and gate length dimensions are assumed to have a ± 20 nm variation across the 512k-bit block.

Statistical simulations of the V_{T0} distribution for 512k-bit arrays show excellent agreement with the measured V_{T0} distributions for both technologies, as seen in Figs. 7 and 8. The four components of variation in Figs. 5 and 6 were used in the simulations. For the NOR technology, the measured $\sigma V_{T0,CG}$ of the array is about 80 mV, while, for the UCPE technology, the measured $\sigma V_{T0,CG}$ is about 39 mV. The array V_{T0} variation is dominated by the dopant fluctuation component as the calculated $\sigma V_{T0,CG}$ due to dopant fluctuations alone is 76 mV for the NOR technology and 34 mV for the UCPE technology. For a 512k-bit array, there are $\pm 4.6 \sigma$'s, so the total V_{T0} variation for the NOR technology due to dopant fluctuations is $9.2 \cdot (\sigma V_{T0,CG}) = 700$ mV.

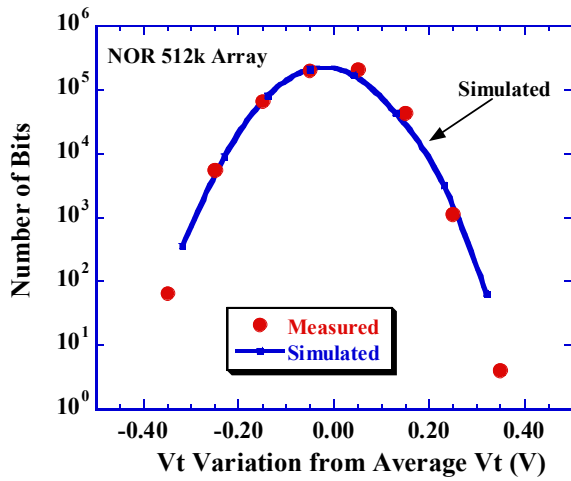


Fig. 7. Comparison of the measured and statistically simulated natural V_T distribution for a 512k-bit array with the NOR technology. The measured $\sigma V_{T0,CG}$ of the distribution is 80 mV and is dominated by the dopant fluctuation component.

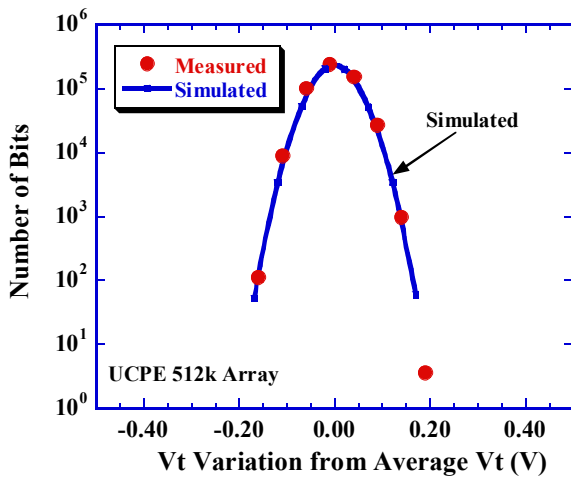


Fig. 8. Comparison of the measured and statistically simulated natural V_T distribution for a 512k-bit array with the UCPE technology. The measured $\sigma V_{T0,CG}$ of the distribution is 39 mV and is dominated by the dopant fluctuation component.

Implications for NVM Data Retention

After cycling, the charge leakage and TTF (Time-to-Fail) of Flash cells at low temperature can be described as

$$I = A \exp(E_{OX} / B) \quad (2)$$

$$TTF = C_T T_{OX} (B / A) \exp(-E_{OX}(V_{TF}) / B) \quad (3)$$

where A (a statistical distribution) and B are model parameters, C_T is the total cell capacitance, T_{OX} is the tunnel oxide thickness, V_{TF} is the fail threshold voltage, and the oxide electric field is calculated as $E_{OX} = \alpha_G (V_T - V_{T0}) / T_{OX}$ [8-11, $X=-1$ in 9,11]. Equation (3) is shown in reduced form, excluding the initial V_T term that has little contribution [9]. As seen in (3), the TTF has an exponential dependence on the oxide field, which in turn is normally distributed due to V_{T0} .

Fig. 9 shows the modeled drop in V_T due to charge loss from a self-bias field for the same A and B parameters but with a 1V V_{T0} variation. For a given sense amp trip or fail point, the time-to-fail (TTF) varies by 20X for a 1V V_{T0} variation as seen in Fig. 10. Although the TTF improves significantly as the fail V_T decreases, the TTF dependence on V_{T0} remains the same since $V_T \propto \ln(t)$ once significant charge loss sets in.

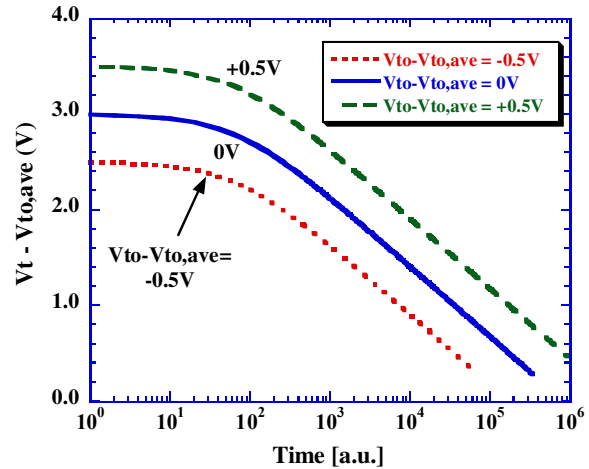


Fig. 9. V_T drop with time due to self-bias charge-loss for 3 bits with different V_{T0} . The leakage parameters are fixed to that of a worst-case bit from a 512k-bit block. Lower V_{T0} reduces the time to reach a fail V_T due to increased E_{OX} ($E_{OX} \propto (V_T - V_{T0})$), while higher V_{T0} increases the time to reach a fail V_T due to decreased E_{OX} . Note that the programmed V_T is assumed to vary with V_{T0} , as seen at $T=1$.

The effect of the V_{T0} variation on the circuit-level TTF is studied in Fig. 11. With fixed charge-loss parameters corresponding to the typical worst-case bit of a 512k block after 1000 cycles, the TTF is calculated using only the V_{T0} variation that is expected with an increasing number of blocks. The expected first fail time from 1000 blocks is about half of the expected first fail time from a single block for σV_{T0} of 80mV. This is much better than the predicted first fail time of less than 0.05X, as seen in Fig. 12, for 1000 blocks using an extracted A distribution [11, with $X=-1$], indicating that the charge-loss statistics are dominated by the

variation in the trap-assisted-tunneling conduction of the leaky paths [10,12] rather than the V_{T0} variation which plays a much smaller role.

Conclusion

The major component of the natural V_T variation of NVM circuits is the intrinsic dopant fluctuations of the well doping. Existing analytical models for the V_{T0} variation with dopant fluctuations provide excellent agreement with both circuit and device data. The implications of the V_{T0} variation on data retention are studied and quantified using charge-loss models.

Acknowledgment

The authors gratefully acknowledge the support and encouragement of Ko-Min Chang and John Bridwell.

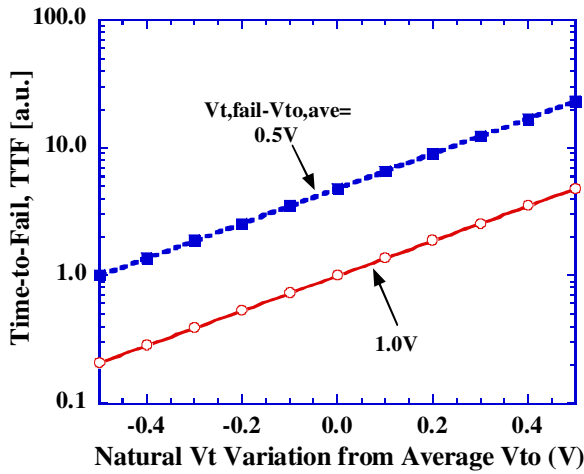


Fig. 10. Comparison of the time-to-fail (TTF) for two different fail V_T levels. The leakage parameters are representative of a worst-case bit from a 512k-bit block. A 1V range in V_{T0} leads to a 20X variation in TTF for the same leakage parameters.

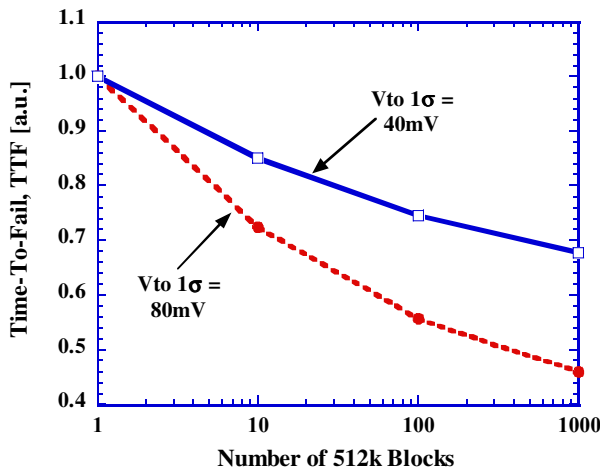


Fig. 11. The first-bit TTF for increasing number of 512k-bit blocks. The leakage parameter values (A , B) are fixed at those of the worst-case bit from one block. Then, the worst-case V_{T0} (lowest) out of the blocks is decreased statistically from its average value as the number of blocks increase.

References

- [1] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFET's, *IEEE Trans. Electron Devices*, vol. 41, pp. 2216-2221, 1994.
- [2] D. Burnett, K. Erington, C. Subramanian, and K. Baker, "Implications of fundamental threshold voltage variations for high-density SRAM and logic circuits," in *Tech. Digest VLSI Symp.*, 1994, pp. 15-16.
- [3] H.-S. Wong, D. Frank, P. Solomon, C. Wann, and J. Welser, "Nanoscale CMOS," *IEEE Proceedings*, vol. 87, 1999, pp. 537-570.
- [4] A. Bhavnagarwala, X. Tang, and J. Meindl, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE J. Solid-State Circuits*, vol. 36, pp. 658-665, 2001.
- [5] D. Burnett, *et al*, "Integration challenges for high-density, high-performance 0.25um embedded Flash memories," in *NVSMW Tech Dig.*, 2000, pp. 59-61.
- [6] C.-N. Li, *et al*, "A novel uniform-channel-program-erase (UCPE) Flash EEPROM using an isolated P-well structure," in *IEDM Tech Dig.*, 2000, pp. 779-782.
- [7] P. Stolk, F. Widdershoven, and D. Klaassen, "Modeling statistical dopant fluctuations in MOS transistors," *IEEE Trans. Electron Devices*, vol. 45, pp. 1960-1971, 1998.
- [8] Y. Okuyama, *et al*, "Anomalous leakage current model for retention failure in Flash memories," in *SSDM Tech Dig.*, 1999, pp. 530-531.
- [9] P. Kuhn, A. Hoefler, T. Harp, B. Hornung, R. Paulsen, D. Burnett, and J. Higman, "A reliability methodology for low temperature data retention in floating gate non-volatile memories," in *IEEE IRPS Tech. Dig.*, 2001, pp. 266-270.
- [10] H. Belgal, N. Righos, I. Kalastirsky, J. Peterson, R. Shiner, and N. Mielke, "A new reliability model for post-cycling charge retention of Flash memories," in *IEEE IRPS Tech. Dig.*, 2002, pp. 7-20.
- [11] A. Hoefler, J. Higman, T. Harp, and P. Kuhn, "Statistical modeling of the program/erase cycling acceleration of low temperature data retention in floating gate nonvolatile memories," in *IRPS Tech. Dig.*, 2002, pp. 21-25.
- [12] F. Schuler, R. Degraeve, P. Hendrickx, and D. Wellekens, "Physical description of anomalous charge loss in floating gate based NMV's and identification of its dominant parameter," in *IEEE IRPS Tech. Dig.*, 2002, pp. 26-33.

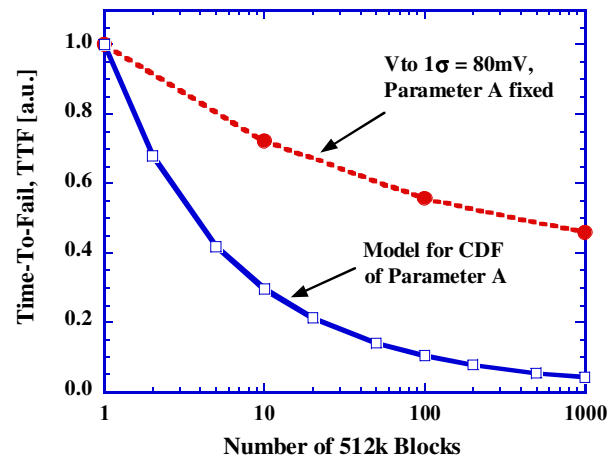


Fig. 12. The first-bit TTF for increasing number of 512k-bit blocks. The upper curve corresponds to only the V_{T0} variation with the leakage parameters (A , B) fixed while the lower curve uses the extracted A distribution (V_{T0} is fixed at the average V_{T0} value). The V_{T0} variation is observed to play a small role in the overall TTF degradation.