A Physics-Based Program-Retention Joint Model of Charge-Trap Transistor

Yi Xiao, Haozhang Yang, Jiaqi Li, Chenglin Ye, Nan Tang, YongJia Wang,
Zheng Zhou[#], Xiaoyan Liu, Jinfeng Kang, Peng Huang[*]
School of Integrated Circuits, Peking University, Beijing 100871, China;
Beijing Advanced Innovation Center for Integrated Circuits, Beijing 100871, China
E-mail: [#]zhouzime@pku.edu.cn, [*]phwang@pku.edu.cn

***Abstract*—We present a physics-based program-retention joint model of charge-trap transistor (CTT). The model quantitatively evaluates electrons injection into and escape from bulk trap of high-k layer in gate dielectric. Excellent agreement between the model results and experiment data from 22 nm Fully-Depleted Silicon-on-Insulator (FDSOI) charge-trap transistors confirms the model's validity. Guided by the model, a compact retention model enabled threshold voltage shift (ΔV_{TH}) drift compensation scheme is proposed and validated by experiment. This scheme compensates 97% of current drift and corrects all of the tested programmed states into the target range.**

Keywords—Charge-Trap Transistor, Program-Retention Model, FDSOI

I. Introduction

There is an increasing demand for embedded memories in low-power IoT devices and in-memory computing. Conventional embedded memories such as embedded Flash and emerging memories such as resistive random access memory (RRAM) and phase change random access memory (PCRAM) require extra process steps compared with standard CMOS process flow [1,2], which may cause increased cost and lowered yield. While the charge-trap transistor (CTT) is fabricated by standard CMOS process flow and has decent retention [1-3], which makes it a promising candidate for low-power embedded memories and in-memory computing. However, due to the thin interface layer and the lack of blocking layer compared to Flash memory [4], CTT exhibits threshold voltage shift (ΔV_{TH}) drift during retention [5], which hinders accurate data storage and degrades the accuracy of in-memory computing. Compensating the drift is a potential method for addressing this issue. In this work, we propose a physics-based program-retention joint model of CTT validated by experiments on 22nm FDSOI charge-trap transistors (FDSOI CTTs) and a compact model enabled ΔV_{TH} drift compensation scheme for mitigating drift using our model.

II. FDSOI Charge-Trap Transistor

The schematic and TEM image of measured FDSOI CTT are shown in Fig. 1(a) and (b). The gate dielectric is composed of an interface layer (IL) and a high-k layer (HL) with traps capturing the electrons [6]. During program, electrons are injected into the traps in the high-k layer by tunneling when applying a positive voltage (V_{pgm}) on the gate electrode, and the source and drain electrodes remain at 0V. As shown in Fig.1(c), the transfer curve after program shifts right, showing a program-induced threshold voltage shift. Fig. 1(d) shows the retention behavior after program. We can observe obvious ΔV_{TH} drift during retention. Due to the thin interface layer and the lack of blocking layer of the FDSOI CTT, electrons are relatively easier to de-trap from traps in the high-k layer compared to those in Flash memory, which causes the ΔV_{TH} drift.

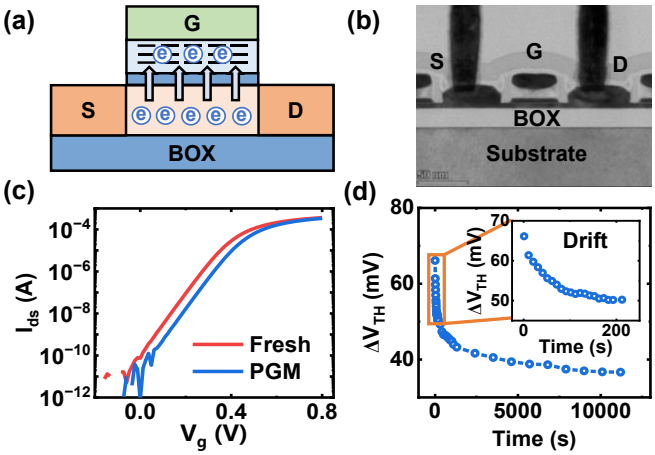

Fig. 1. (a) Schematic illustration and mechanism of FDSOI CTT. (b) TEM image of FDSOI CTT. (c) The transfer curve of FDSOI CTT before and after program (PGM) shows a positive shift of V_{TH} after PGM. (d) The retention behavior after program of FDSOI CTT, showing ΔV_{TH} drift.

III. Physics-Based Program-Retention Joint Model

Fig. 2 shows the overall modelling methodology. Given the input V_{pgm} and the initial spatial density of the filled trap $n_{t0}(x)$, the electric field (E-field) is calculated by one-dimensional Poisson equation, assuming an averaged constant E-field in the high-k layer. The trap/de-trap mechanism of the trap considers the electrons injection into and escape from the trap by tunneling. From the calculated E-field, injection current is calculated by tunneling using the Wentzel–

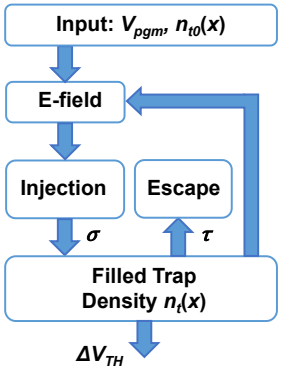

Fig. 2. The overall modelling methodology, where $n_t(x)$ is solved iteratively with the input V_{pgm} and the given initial condition $n_{t0}(x)$.

Kramers–Brillouin (WKB) approximation. With the constant capture cross section σ, the electron injection rate of the trap by tunneling is calculated. With the escape time constant τ, the electron escape rate from the filled trap is calculated. During program, both electrons injection and escape are taken into account. During retention, only electrons escaping from the trap is considered and the injection current is neglected. With the electron injection rate and escape rate, the spatial density of the filled trap $n_t(x)$ is calculated by iteratively solving trap/de-trap differential equation written as

$$\frac{dn_t(x)}{dt} = \frac{J_T(x)}{q}\sigma(N_t - n_t(x)) - \frac{n_t(x)}{\tau_{t-g}(x)} - \frac{n_t(x)}{\tau_{t-c}(x)} \quad (1)$$

where x is the distance from the HL/IL interface in the high-k layer, $J_T(x)$ the tunneling current, q the charge of an electron, σ the capture cross section of the considered trap, N_t the spatial density of the trap and $\tau_{t\text{-}g}(x)$ and $\tau_{t\text{-}c}(x)$ the escape time constant towards gate and channel respectively. The first term on the right hand side of the equation (1) corresponds to the electrons injection and the second and third term correspond to the electrons escape. The E-field is further updated according to the spatial density of the filled trap $n_t(x)$. With the calculated $n_t(x)$ by equation (1), the ΔV_{TH} is obtained by the filled trap density $n_t(x)$ written as

$$\Delta V_{TH} = \int_0^{t_{HL}} \frac{q n_t(x)}{\varepsilon_{HL}}(t_{HL} - x)dx \quad (2)$$

where t_{HL} and ε_{HL} are the thickness and dielectric constant of the high-k layer.

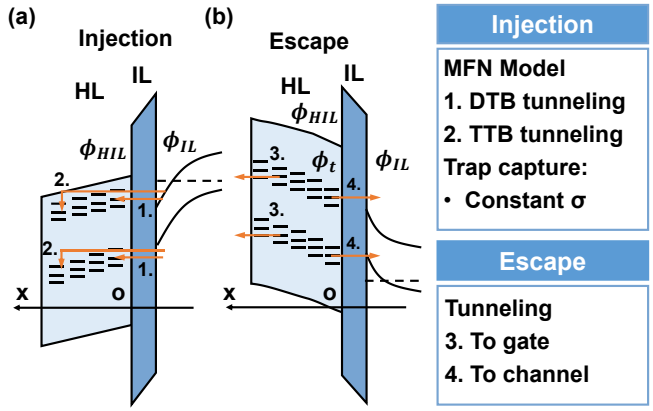

Fig. 3. (a)Illustration of electrons injection mechanisms. Electrons tunnel from channel to the traps in the HL. Tunneling current is modelled by MFN model with constant capture cross section (σ). (b). Illustration of electrons escape mechanisms. Electrons escape to gate and channel by direct tunneling.

Fig. 3(a) and (b) show the energy band diagram of the CTT during program and retention and the specific injection and escape components. We consider spatially uniform bulk traps in the high-k layer falling in two energy levels [7]. The tunneling current can be calculated using modified Fowler-Nordheim (MFN) model [8] with quantum barrier lowering [9]. Quantum barrier lowering is given by

$$\phi_{IL} = \phi_{IL0} - \theta F_{IL}^{2/3} \quad (3)$$

where ϕ_{IL0} is the potential barrier height of the interface layer and the channel without quantum barrier lowering, θ an adjustable parameter and F_{IL} the E-field in the interface layer. Using WKB approximation, the tunneling current can be expressed as either double trapezoidal barrier (DTB) or trapezoidal-triangular barrier (TTB) forms, depending on the position-dependent barrier shape. For the position where ϕ_{IL}-ϕ_{HIL}-$F_{IL}t_{IL}$-$F_{HL}x$>0, the potential barrier is double trapezoidal and the tunneling current can be written as

$$J_T(x) = \frac{A_n F_{IL}^2}{\phi_{IL}} \exp(-B_1(\phi_{IL}^{3/2} - (\phi_{IL} - F_{IL}t_{IL})^{3/2})/F_{IL}) \quad (4)$$
$$\exp(-B_2(\phi_1^{3/2} - (\phi_1 - F_{HL}x)^{3/2})/F_{HL})$$

where A_n, B_1 and B_2 are adjustable parameters, t_{IL} the thickness of the interface layer, F_{HL} the averaged electric field in the high-k layer, ϕ_{HIL} the conduction band offset of the HL and the IL and the intermediate ϕ_1 is written as

$$\phi_1 = \phi_{IL} - \phi_{HIL} - F_{IL}t_{IL} \quad (5)$$

For the position where ϕ_{IL}-ϕ_{HIL}-$F_{IL}t_{IL}$-$F_{HL}x$<0, the potential barrier is trapezoidal-triangular and the tunneling current can be written as

$$J_T(x) = \frac{A_n F_{IL}^2}{\phi_{IL}} \exp(-B_1(\phi_{IL}^{3/2} - (\phi_{IL} - F_{IL}t_{IL})^{3/2})/F_{IL}) \quad (6)$$
$$\exp(-B_2\phi_1^{3/2}/F_{HL})$$

The time constants of electrons escaping from traps to gate and channel can be given by tunneling [10] and a field dependent $\tau_1(F_{HL})$, which are written as

$$\tau_{t-g}(x) = \tau_1(F_{HL})\exp(K_{\tau,g}) \quad (7)$$

$$\tau_{t-c}(x) = \tau_1(F_{HL})\exp(K_{\tau,c}) \quad (8)$$

respectively, where $\tau_1(F_{HL})$ is written as

$$\tau_1(F_{HL}) = \tau_0 \exp(-b_t\sqrt{F_{HL}}) \quad (9)$$

$K_{\tau,g}$ and $K_{\tau,c}$ are derived from WKB approximation assuming a trapezoidal barrier shape and double trapezoidal barrier shape respectively considering the relatively small E-field in the high-k layer. $K_{\tau,g}$ and $K_{\tau,c}$ can be written as

$$K_{\tau,g} = -B_2(\phi_t^{3/2} - (\phi_t - F_{HL}(t_{HL} - x))^{3/2})/F_{HL} \quad (10)$$

$$K_{\tau,c} = -B_3((\phi_t - F_{HL}x)^{3/2} - \phi_t^{3/2})/F_{HL} - B_1(\phi_t + \phi_{HIL} - F_{IL}t_{IL} - F_{HL}x)^{3/2}/F_{IL} + B_1(\phi_t + \phi_{HIL} - F_{HL}x)^{3/2}/F_{IL} \quad (11)$$

where B_3 is an adjustable parameter and $q\phi_t$ the trap energy level from the conduction band of the high-k layer of the trap.

IV. Results and Discussion

A. Model Validation

The program-induced ΔV_{TH} and the corresponding retention behavior of different program time and gate voltages are measured and calculated in Fig. 4(a) and (b).

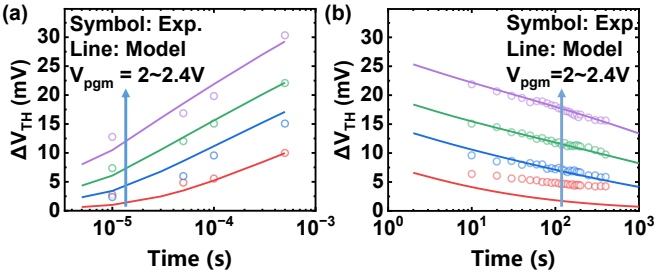

Fig. 4. (a) Comparison between experiment and model of program characteristics at different program time and gate voltages (V_{pgm}). (b) Comparison between experiment and model of retention behavior after program in (a).

The model shows agreement with experiment. Both program and retention show logarithmic behavior. The slight deviation of retention behavior of the low ΔV_{TH} case may be attributed to the simplified assumption of discrete energy levels rather than continuously distributed energy levels. The measured and calculated incremental step pulse programming (ISPP) curves and their corresponding retention characteristics of different program pulse width (T_{pgm}) are shown in Fig. 5(a) and (b). Our model accurately captures the behavior of the measured data. All the data is measured at subthreshold region of the device. The parameters used for the measured device are listed in Table. 1.

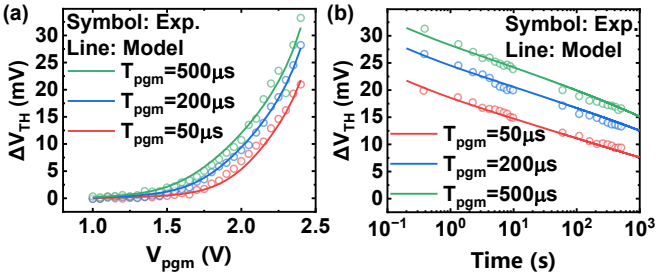

Fig. 5. (a) Comparison between experiment and model of ISPP curve. T_{pgm} denotes the program pulse width. (b) Comparison between measurement and model results of retention behavior after ISPP in (a).

B. *Drift compensation scheme*

Guided by our model and experiment, a compact retention model enabled ΔV_{TH} drift compensation scheme is proposed for drift compensation. Since both model calculation and experiment gives a logarithmic retention behavior, the compact retention model can be written as

$$\Delta V_{TH,dr} = -\alpha \ln(t / \tau_{eff} + 1) \tag{12}$$

where α and τ_{eff} are obtained by fitting of the model calculation of retention curve.

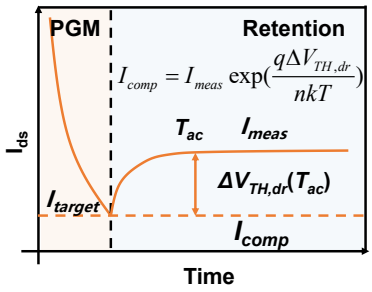

Fig. 6. Drift compensation scheme to alleviate the ΔV_{TH} drift ($\Delta V_{TH,dr}$) during retention. The measurement of current is done in subthreshold region to reduce read disturb. $\Delta V_{TH,dr}$ is obtained by the compact retention model.

As shown in Fig. 6, given the access time T_{ac}, the threshold voltage drift during retention $\Delta V_{TH,dr}$ due to de-trapping can be calculated by the compact model, thus the compensated current I_{comp} can be obtained from measured current I_{meas} by expression in subthreshold region written as

$$I_{comp} = I_{meas} \exp(\frac{q\Delta V_{TH,dr}}{nkT}) \tag{13}$$

where k is the Boltzmann constant, T the temperature and n is a coefficient related to subthreshold slope of the device given as

$$n = \frac{q \cdot SS}{kT \ln(10)} \tag{14}$$

where SS is the subthreshold slope of the device. Fig. 7 shows the measurement of retention behavior for multiple times and the compensated results by the compact model with α = 1mV, τ_{eff} = 95ms. ISPP and write verify scheme is adopted for program and the program stops when the verified current falls within the target range. The target current is 72.5nA and the target range is set to be ±4nA. Throughout the tested period, most of the compensated current states fall within the target range. Fig. 8 shows the current distribution with and without compensation 210s after program, showing 0.6nA and 22.8nA of averaged drift respectively. With the proposed compact model and the drift compensation scheme, 97% of the averaged current drift is compensated and all of the tested programmed states after compensation are corrected into the target range.

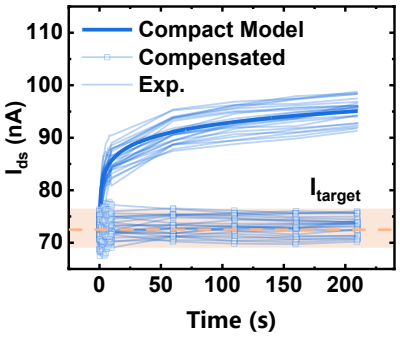

Fig. 7. The measurement of retention behavior for multiple times and the compensated results with compact model. The compact model fits the experiment results well with α = 1mV, τ_{eff} = 95ms.

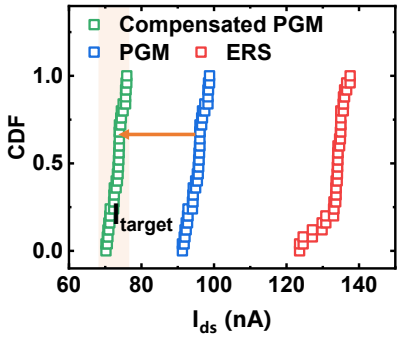

Fig. 8. The current distribution with and without compensation at 210s after program. 97% of current drift is compensated and all of the tested programmed states are corrected into the target range.

V. Conclusion

A physics-based program-retention joint model of charge-trap transistor has been proposed, incorporating the mechanisms of electrons injection and escape by tunneling. The model accurately captures the program and retention characteristics in measured 22 nm FDSOI charge-trap transistors. Guided by model predictions, a compact retention model enabled ΔV_{TH} drift compensation scheme is developed, effectively mitigating 97% of the current drift and correcting all of the tested programmed states into the target range. These results demonstrate that the proposed model offers a promising approach for addressing the ΔV_{TH} drift issue in charge-trap transistor.

Table. 1 The parameters used for the measured device.

Model Parameters	
A_n	3.33×10^{-5} A/(m^{2}V)
B_1	9.66×10^{9} m^{-1}
B_2	1.82×10^{10} m^{-1}
B_3	2.93×10^{10} m^{-1}
θ	1.6×10^{-7} V$^{1/3}$m$^{-2/3}$
ϕ_{IL0}/ϕ_{HIL}	3.0V/1.0V [11]
ϕ_{t1}/ϕ_{t2}	1.1V [12]/2.5V [7]
N_{t1}	5.65×10^{19}cm^{-3}
N_{t2}	2.00×10^{18}cm^{-3}
τ_0	3×10^{-12} s
b_t	5×10^{-4}V$^{-1/2}$/m$^{-1/2}$
σ	1×10^{-14} cm^2 [13]
t_{IL}/t_{HL}	0.8nm/2.0nm [14]

ACKNOWLEDGMENT

This work was supported in part by the National Key Research Plan of China under Grant 2023YFB4402400; in part by the National Natural Science Foundation of China Program under Grant 62034006 and Grant 62474005; in part by the STIC under Grant QYJS-2022-2200-B; in part by the 111 Project Program under Grant B18001.

REFERENCES

[1] S. Nouri, and S. S. Iyer, "An 8T eNVSRAM Macro in 22nm FDSOI Standard Logic with Simultaneous Full-Array Data Restore for Secure IoT Devices", International Solid-State Circuits Conference (ISSCC), 2023.

[2] S. Qiao, S. Moran, D. Srinivas, S. Pamarti, and S. S. Iyer, "Demonstration of analog compute-in-memory using the charge-trap transistor in 22 FDX technology", International Electron Devices Meeting (IEDM), 2022.

[3] F. Khan, D. Moy, D. Anand, E. H.-Schroeder, R. Katz, L. Jiang, E. Banghart, N. Robson, and T. Kirihata, "Turning logic transistors into secure, multi-time programmable, embedded non-volatile memory elements for 14 nm FINFET technologies and beyond", Symposium on VLSI Technology and Circuits (VLSI), 2019.

[4] S. Kim, S. Yong, W. Kim, S. Kang, H. Park, K. Yoon, D. Sheen, S. Lee, and C. Hwang, "Review of Semiconductor Flash Memory Devices for Material and Process Issues", Advanced Material., 2023, vol. 35, 43.

[5] J. Viraraghavan, D. Leu, B. Jayaraman, A. Cestero, R. Kilker, M. Ying, J. Golz, R. R. Tummuru, R. Raghavan, D. Moy, T. Kempanna, F. Khan, "80Kb 10ns read cycle logic Embedded High-K charge trap Multi-Time-Programmable Memory scalable to 14nm FIN with no added process complexity", Symposium on VLSI Circuits (VLSI-Circuits), 2016.

[6] F. Khan, M. S. Han, D. Moy, R. Katz, L. Jiang, E. Banghart, N. Robson, T. Kirihata, J. C. S. Woo, and S. S. Iyer, "Design Optimization and Modeling of Charge Trap Transistors (CTTs) in 14 nm FinFET Technologies", IEEE Electron Devices Letters (EDL), 2019, vol. 40, 7.

[7] G. Rzepa, J.Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P.Weckx, M. Jech, T. Knobloh, M. Waltl, P.J. Roussel, D. Linten, B. Kaczer, and T. Grasser, "Comphy — A compact-physics framework for unified modeling of BTI", Microelectronics Reliability, vol. 85, pp. 49-65.

[8] H. Bachhofer, H. Reisinger, E. Bertagnolli, and H. von Philipsborn, "Transient conduction in multidielectric silicon–oxide–nitride–oxide semiconductor structures", Journal of Applied Physics, 2001, vol. 89, 5.

[9] W. Quan, D. M. Kim, and M. K. Cho, "Unified compact theory of tunneling gate current in metal–oxide–semiconductor structures: Quantum and image force barrier lowering", Journal of Applied Physics, 2002, vol. 92, 7.

[10] M. Kim, S. Kim, and H. Shin, "A Compact Model for ISPP of 3-D Charge-Trap NAND Flash Memories", IEEE Transactions on Electron Devices (TED), 2020, vol. 67, 8.

[11] E. Bersch, M. Di, S. Consiglio, R. D. Clark, G. J. Leusink, and A. C. Diebold, "Complete band offset characterization of the $HfO_2/SiO_2/Si$ stack using charge corrected x-ray photoelectron spectroscopy", Journal of Applied Physics, 2010, vol. 107, 4.

[12] Y. Zhang, Y. Y. Shao, X. B. Lu, M. Zeng, Z. Zhang, X. S. Gao, X. J. Zhang, J.-M. Liu, and J. Y. Dai, "Defect states and charge trapping characteristics of HfO2 films for high performance nonvolatile memory applications", Applied Physics Letters, 2014, vol. 105, 17.

[13] C. Z. Zhao, J. F. Zhang, and M. B. ZahidB. Govoreanu, G. Groeseneken, and S. De Gendt, "Determination of capture cross sections for as-grown electron traps in stacks", Journal of Applied Physics, 2006, vol. 100, 9.

[14] J. Xu, M. Fan, Z. Sun, F. Liu, X. Liu, "Assessment of Traps-induced Noise in FDSOI MOSFETs", International Conference on Solid State Devices and Materials (SSDM), 2023.