A Physics-based Compact Model for ULTRARAM Memory Device

Abhishek Kumar[1], M. Ehteshamuddin[1], Anand Bulusu[1], Shruti Mehrotra[2], and Avirup Dasgupta[1]

[1]Indian Institute of Technology Roorkee, India, [2]GlobalFoundries, India
Email: abhishek_k1@ece.iitr.ac.in, avirup@ece.iitr.ac.in

2024 8th IEEE Electron Devices Technology & Manufacturing Conference (EDTM) | 979-8-3503-7152-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EDTM58488.2024.10512293

Abstract

ULTRARAM is a non-volatile memory (NVM) device utilizing triple barrier resonant tunnelling (TBRT) that exhibits high endurance (>10^7), retention (>1000 years), and ultra-low switching energy per unit area characteristics. In this paper, we present a compact model of the transfer characteristics of the ULTRARAM device. This model captures the trapping and de-trapping of charge in the floating gate (FG) using TBRT physics and is used to calculate the device characteristics. The model has been implemented in Verilog-A and validated with TCAD as well as experimental data.

Keywords: ULTRARAM, memory, resonant tunnelling.

Introduction

The traditional memory industry is dominated by charge-based memories such as static random-access memory (SRAM), dynamic random-access memory (DRAM), and flash [1]. However, each of these memory technologies has its own drawbacks. Flash memory is a suitable candidate for data storage to date. However, it needs a high voltage for non-volatility, which limits endurance [2]. DRAM is faster than flash with a smaller operating voltage but suffers from destructive reading and requires persistent data refreshing, increasing energy consumption, degrading performance, and limiting scaling capacity. SRAM is the fastest among charge-based memories but requires a large footprint even for 1-bit storage.

The ideal memory device would possess several key characteristics, including fast write/read speeds (<ns), low operating voltage (<1V), low energy consumption (∼ fJ/bit for write/read), long data retention time (>10 years), high write/read cycling endurance (> 10^{17} cycles), and excellent scalability (<10 nm) [3]. Achieving all of these ideal characteristics in a single universal memory device is a difficult task. To address this, various emerging non-volatile memory (NVM) technologies have been pursued to fulfil some of these ideal characteristics, like PCRAM, RRAM, STT-RAM, Fe-RAM, and FeFET [4].

Recently, a new memory technology called ULTRARAM has been demonstrated [5], [6]. It is a non-classical charge-trapping-based NVM that exhibits fast, non-volatile, high endurance, and ultra-low switching energy per unit area characteristics. It breaks the paradigm of the unachievable universal memory idea. The state is determined by the presence or absence of electrons in a FG. Unlike a single SiO_2 barrier in a flash, the novelty comes from the InAs/AlSb triple barrier resonant tunnelling (TBRT) mechanism [7], as shown in Fig. 1. The TBRT structure provides a high-potential electron barrier with no bias and allows fast resonant tunneling at program/erase pulse (±2.5V) with switching energy per unit area 1000 times lower than flash [8]. The device exhibits retention of over 1000 years and degradation-free endurance of over 10^7 program/erase cycles [6].

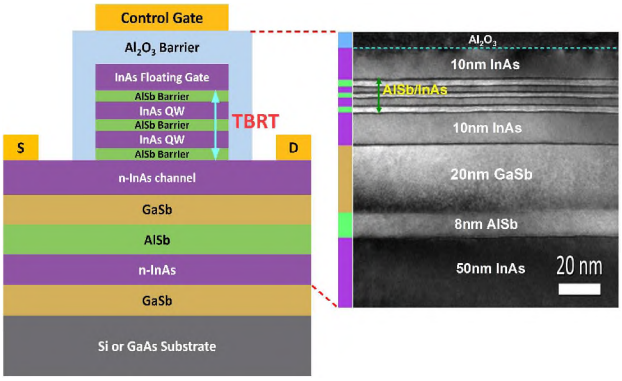


Fig. 1: Schematic of ULTRARAM memory device and the corresponding TEM image [5].

Since ULTRARAM is in the early stage of development, a physics-based compact model is indispensable for detailed performance analysis as well as device and circuit-level optimization.

Model Development

We have developed a physics-based model for TBRT which is coupled with a surface potential-based FET model to capture real-time device characteristics with high accuracy.

A. TBRT Model

The TBRT model is based on the approach given in [9]. This model is generic in nature and can work with any compound semiconductor TBRT framework. In this formulation, the following assumptions have been made: (a) the tunnelling process preserves the transverse momentum; (b) the distribution of the electric field is uniform in the structure; and (c) the effective mass approximation [10]. Under the above mentioned

conditions, the current density can be described as [9]

$$J = \frac{q_e m^* kT}{2\pi^2\hbar^3} \int_0^\infty T(E_x, V) D(E_x, V) dE_x, \quad (1)$$

where q_e is the charge of an electron, m^* is the effective mass of the electron, k is the Boltzmann's constant, T is the absolute temperature, $\hbar$ is the reduced Planck's constant, E_x is the longitudinal energy, V is the potential applied to the structure and $D(E_x, V)$ is the supply function and can be given as

$$D(E_x, V) = ln\left[\frac{1 + exp(\frac{E_f - E_x}{kT})}{1 + exp(\frac{E_f - E_x - q_e V}{kT})}\right]. \quad (2)$$

$T(E_x, V)$ is the tunnelling transmission coefficient approximated as [10]

$$T(E_x, V) = \frac{\Gamma^2}{(E_x - (E_R - q_e V n))^2 + \Gamma^2}, \quad (3)$$

where Γ is the half-width of resonance level, E_R is the energy of resonant level relative to the bottom of the well at its center and n is the voltage drop factor at the center of the well. For equal width barrier, n is 0.5, otherwise it can be determined from analysis or used as fitting parameter. For small Γ compared to the thermal energy, $T(E_x, V)$ is very small, i.e., $E_x \approx E_R - eVn$. Now, integrating (1) using (2) and (3) gives

$$J = J_0 D(E_x, V)\left[\frac{\pi}{2} + tan^{-1}\left(\frac{E_R - q_e V n}{\Gamma}\right)\right], \quad (4)$$

where $J_0 = q_e m^* kT/(2\pi^2\hbar^3)$.

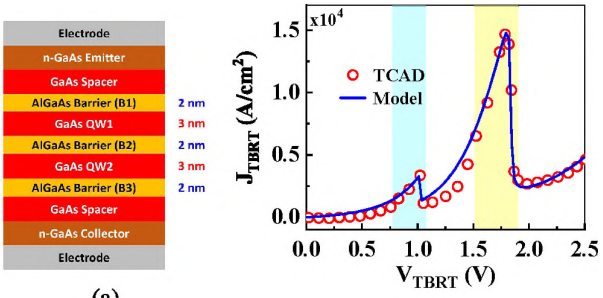
In TBRT, there are two resonance levels. The second resonance can be explicitly included using additional terms similar to (4). This form produces the double peaks due to resonant tunneling. To capture the thermal excitation over a barrier, an additional thermionic component can be added as $J_{th} = H(e^{eVn_3/kT} - 1)$, where H and n_3 can be used as fitting parameters. To generalize this model for any III-V material based TBRT, we can rewrite (4) as (5).

B. Drain Current Model

The current density, calculated from the TBRT model, is further used to calculate the charge trapping and de-trapping in the FG of the ULTRARAM. An R-C network has been used to integrate the TBRT current over the applied input duration to obtain the real time FG charge. Further, this charge is used to calculate the threshold voltage shift in program/erase state due to charge stored/erased in the FG. The drain current of the device can be expressed as [11]

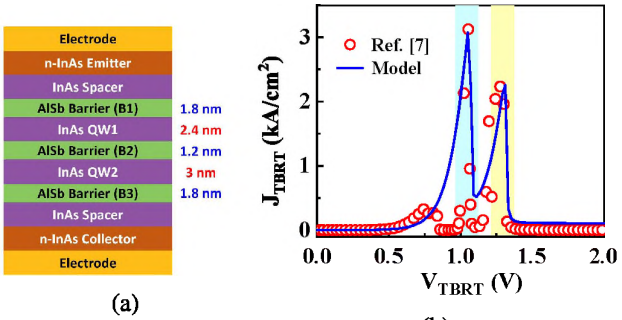
$$I_{ds} = \mu_{eff} C_g \frac{W}{L}\left(V_{gs,eff} - V_{off} - \frac{Q_{FG}}{C} - \psi_m\right)\psi_{ds}, \quad (6)$$

where μ_{eff} is the effective mobility, C_g is the gate capacitance, W is the device width, L is the channel length, $V_{gs,eff}$ is the effective applied gate voltage, V_{off} is the cut-off voltage, Q_{FG} is the charge in the FG, and ψ_{ds} is the drain to source surface potential. $\psi_m = (\psi_s + \psi_d)/2$, where ψ_s and ψ_d are the surface potentials at the source and drain side, respectively. ψ_s and ψ_d were calculated as given in [11].



(a) (b)

Fig. 2: (a) Schematic of GaAs/AlGaAs TBRT structure used for TCAD simulations. (b) Validation of model with TCAD data. The following TBRT parameter values have been used to fit the data: $A_1 = 4.6 \times 10^6$, $A_2 = 44 \times 10^4$, $B_1 = 3.4 \times 10^{-2}$, $B_2 = 2 \times 10^{-2}$, $C_{1,2} = 2.4 \times 10^{-1}$, $D_1 = 10^{-10}$, $D_2 = 10^{-3}$, $n_1 = 9.5 \times 10^{-2}$, $n_2 = 10.9 \times 10^{-2}$, $H = 10$, and $n_3 = 4.1 \times 10^{-2}$.



(a) (b)

Fig. 3: (a) Schematic of AlSb/InAs TBRT structure with corresponding thickness of multiple barriers and QWs. (b) Validation of model with the data given in [7]. The following TBRT parameter values have been used to fit the data: $A_1 = 16 \times 10^4$, $A_2 = 14 \times 10^4$, $B_{1,2} = 10^{-3}$, $C_{1,2} = 0.3$, $D_{1,2} = 10^{-3}$, $n_{1,2} = 28 \times 10^{-2}$, and $H = 0$.

Results and Discussion

Fig. (2) shows a GaAs/AlGaAs TBRT structure simulated in TCAD. The model captures the TBRT current accurately. In Fig. (3), we have validated the model with the AlSb/InAs based TBRT [7], where AlSb is the barrier material and InAs is the QW. Note that there is negligible contribution of thermionic current at room temperature. The same TBRT has been used in the ULTRARAM gate stack.

Further, the complete drain current model has been used to simulate the ULTRARAM characteristics. Fig.

$$J_{tbrt} = A_1 ln\left[\frac{1+exp(\frac{B_1-C_1+q_eVn_1}{kT})}{1+exp(\frac{B_1-C_1+q_eV(n_1-1)}{kT})}\right]\left[\frac{\pi}{2}+tan^{-1}\left(\frac{C_1-q_eVn_1}{D_1}\right)\right]$$
$$+ A_2 ln\left[\frac{1+exp(\frac{B_2-C_2+q_eVn_2}{kT})}{1+exp(\frac{B_2-C_2+q_eV(n_2-1)}{kT})}\right]\left[\frac{\pi}{2}+tan^{-1}\left(\frac{C_2-q_eVn_2}{D_2}\right)\right]+H(e^{q_eVn_3/kT}-1) \quad (5)$$

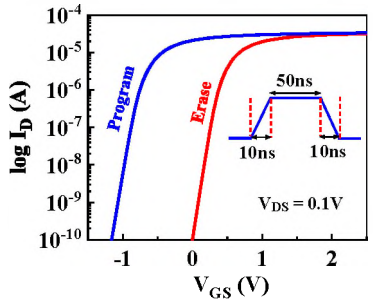


Fig. 4: $I_D - V_{GS}$ characteristics simulated from the model with the applied gate bias of $\pm$2.5V.

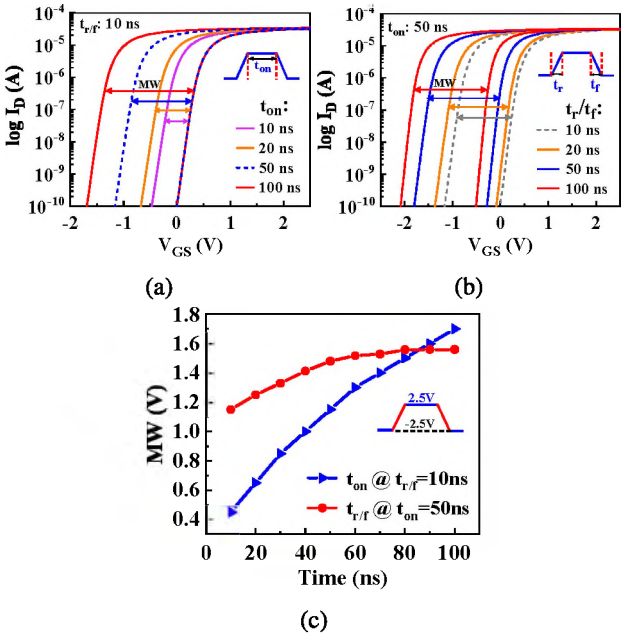


Fig. 5: Variations in MW of the device for (a) pulse width, (b) rise/fall time of the pulse, and (c) combinations of both. All the TBRT model parameters are same as used in Fig. (3).

(4) shows the I_D-V_{GS} results simulated from the model for the programmed and erased states. The obtained memory window (MW) depends on the input waveform, which is accurately captured in real time by the proposed model. Variation with pulse width (t_{on}) and rise/fall time ($t_{r/f}$) is shown in Fig. (5) at $V_{DS} = 0.1V$.

We have also validated the model with experimental data of ULTRARAM as shown in Fig. (6) [6]. The model shows excellent predictive accuracy.

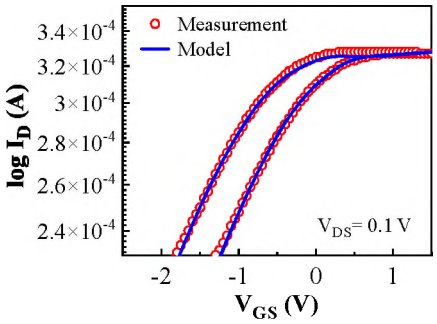


Fig. 6: Validation of model with experimental $I_D - V_{GS}$ characteristics [6] of ULTRARAM at $V_{DS} = 0.1V$.

Conclusion

A novel physics based compact model of ULTRARAM has been proposed. The model is able to capture real-time device characteristics for any input waveform and shows excellent accuracy with both TCAD and measured data for TBRT and ULTRARAM. The proposed model can be used for any III-V based framework. This SPICE-compatible model has been implemented in Verilog-A and is directly usable for circuit analysis.

Acknowledgement

We are thankful to Prof. M. Hayne and Dr. P. Hodgson from Lancaster University, UK, for their help with this work.

References

[1] A. Sebastian et al., Nature nano-tech. 15, no. 7, 529-544, 2020.
[2] T. Schenk et al., Reports on Progress in Physics 83, no. 8 (2020): 086501.
[3] S. Yu et al., IEEE SSC Magazine 8, no.2, 43-56, 2016.
[4] A. Chen, SSE 125 (2016): 25-38.
[5] D. Lane et al., IEEE TED, vol. 68, no. 5, pp. 2271-2274, May 2021.
[6] P. D. Hodgson et al., Advanced Electronic Materials 8, no. 4 (2022): 2101103.
[7] D. Lane et al., J. Physics D, 54, no. 35 (2021).
[8] D. Lane et al., IEEE EDTM, pp. 1-3, 2021.
[9] R. Tsu and L. Esaki, APL 22, no. 11 (1973): 562-564.
[10] J. N. Schulman et al., IEEE EDL, vol. 17, no. 5, pp. 220-222, May 1996.
[11] S. Khandelwal et al., "ASM-HEMT 101.0. 0 Advanced SPICE Model for HEMTs."