View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Indonesian Journal of Electrical Engineering and Informatics (IJEEI)

Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 7, No. 2, June 2019, pp. 331~337
ISSN: 2089-3272, DOI: 10.11591/ijeei.v7i2.1170

Reliability of graphene as charge storage layer in floating gate flash memory

M. Hilman Ahmad, N. Ezaila Alias, Afiq Hamzah, Zaharah Johari, M. S. Z. Abidin, Norlina Paraman, M. L. Peng Tan, Razali Ismail
School of Electrical Engineering, Universiti Teknologi Malaysia, Malaysia

Article Info

Article historys:

Received Mar 27, 2019
Revised Apr 27, 2019
Accepted May 18, 2019

Keywords:

Charge storage layer
Data retention
Flash memory
Graphene
Memory window

ABSTRACT

This study aims to investigate the memory performances of graphene as a charge storage layer in the floating gate with difference doping concentration of n-channel and p-channel substrates using Silvaco ATLAS TCAD Tools. The simulation work has been done to determine the performance of flash memory in terms of memory window, P/E characteristics and data retention and have been validated with the experimental work done by other researchers. From the simulation data, the trend of memory window at low P/E voltage is nearly overlapped between simulation and experimental data. The memory window at ±20V P/E voltage for n-channel and p-channel flash memory cell are 15.4V and 15.6V respectively. The data retention for the n-channel flash memory cell is retained by 75% (from 15.4V to 11.6V) whereas for the p-channel flash memory cell is retained by 80% (from 15.6V to 12.5V) after 10 years of extrapolation with -1/1V gate stress which shows that p-channel flash memory cell demonstrates better data retention compared to n-channel flash memory cell.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

N. Ezaila Alias,
Computational Nanoelectronics Research Group,
School of Electrical Engineering, Faculty of Engineering
Universiti Teknologi Malaysia, 81310 Johor, Malaysia.
Email: ezaila@utm.my

1. INTRODUCTION

Recently, the downsizing of silicon machinery is leading to a severe blockage of the coming era due to the complications of channel material such as short channel effects (SCE) in contemporary transistors [1, 2]. Since the vertical and lateral downsizing of flash memory cells lead to low program and erase (P/E) speed and large capacity of data storage for these cells but may results in expansion of capacitive coupling between the floating gate (FG) of neighboring cells which causes an extensive issue in threshold voltages (V_{th}) of the cells. It has reported that the thickness of conventional polycrystalline silicon (poly-Si) FG can be shortened to 7nm [3]. Despite that, an injected electron into thin poly-Si would be ballistically shifted across the FG resulting the P/E speed became slower and further create impact ionization in the blocking dielectric layer and consequently reduce the dielectric reliability. To corroborate with this issue with thin poly-Si FG, a thin metal layer is used as a FG in flash memory cell is capable of extinguishing the ballistic current in flash memory cells. The suitable metal would be worthy to use as a charge storage layer (CSL) in FG flash memory cell is graphene which having metallic properties and the thinnest neutrally stable material [4]. Being only one layer thick, they can further scale in the vertical direction and could offer reduced power dissipation because of smaller (SCE) [5]. Introduction of graphene as CSL in flash memory cell would ease significant trimming of the gate stack height, and further, avoided the contamination issues associated with the metal FG flash memory cell [6]. The properties of ultra-high mobility and ultra-thin structure may provide fast operation time and compact integration density respectively. Besides, its unique band structure could result in a new opportunity for nanoscale and low power devices. By these considerations, there has been interested in utilizing graphene for

denser, faster and less energy consuming in NVM [7]. Graphene transistors are predicted to be considerably faster than silicon transistors nowadays and create computer chips to become faster and more energy efficient [4].

From the previous review, multilayer graphene (MLG) and reduced multilayer graphene (rMLG) have been suggested as CSL in flash memory cells. In [4], it has been reported that MLG as CSL in FG flash memory cells with a memory window (MW) of 6.8V at 18V P/E voltage [4]. In [6], a large MW of 9.4V at ±20V P/E voltages has been investigated in rMLG as a CSL in flash memory structure. The data retentions in 10 years at room temperature is about 6.9V of MW with 74% charged retained while for 150°C is demonstrated by the MW of 2.8V with 30% charged retained [6].

The design of conventional flash memory products has been applied to the n-channel flash memory cells. However, through P/E cycles, the high voltage operation leads to high power consumption in flash memory cells. Thus, the p-channel flash memory cells which using band-to-band tunnelling induced hot electron injection (BBHE) was suggested for p-channel flash memory cells which may present several privileges such as excessive P/E speed, minimize power consumption and superior reliability [8]. The significant advantage of BBHE injection tunnelling method is its high efficiency of the injection which is described as the proportion of electrons composed at the FG to those generated in the channel, (I_G/I_D). The effectiveness of the injection of the p-channel flash memory cell can be 10^{-2} which is much considerable than the n-channel flash memory cell (approximate to 10^{-8}) which leads well suited for high-speed operation. The high injection efficiency of the BBHE is reached because band-to-band tunnelling (BBT) generated the electrons are efficiently accelerated by the lateral electric field in the drain depletion region [9, 10].

In this paper, we investigate the reliability of graphene-based FG non-volatile memory by observing the P/E characteristics, memory window and data retention in 10 years extrapolation under gate stress with a different type of channel. In this study, both programming and erasing of n-channel flash memory cell tunnelling are performed by Fowler-Nordheim (FN) tunnelling model while for p-channel flash memory cell, programming and erasing are performed by BBHE tunnelling and FN tunnelling respectively.

2. RESEARCH METHOD

The simulation works are divided into two main parts: (1) Device validation and (2) Performances of a p-channel flash memory cell with graphene as CSL as shown in Figure 1. In part 1, using Silvaco TCAD Tools, the virtual device structure of graphene-based FG non-volatile memory cell is simulated based on experimental work in [6], refer Table 1 and Figure 2 while the graphene properties parameters are referred to [11].

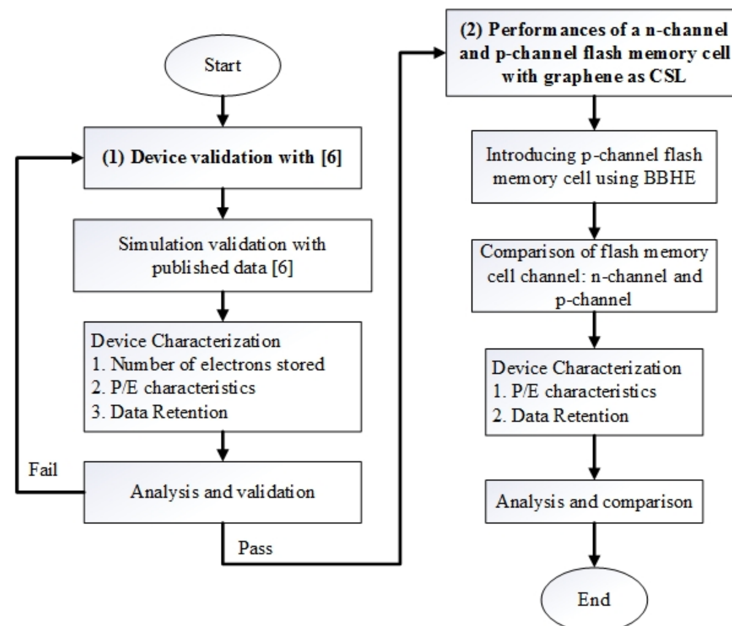


Figure 1. General simulation flowchart.

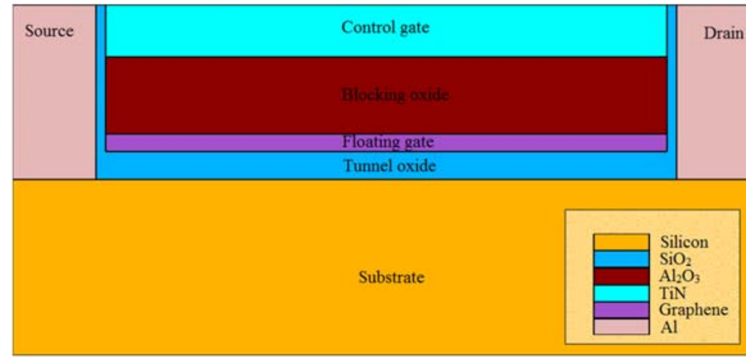


Figure 2. The device structure of graphene-based FG non-volatile memory cell.

The FN tunnelling model is used as the electron tunnelling through a barrier using the extracted FN parameters work in [12]. The number of electrons stored in FG and memory window for 1s from the simulation is validated with experimental data [6] refer Figure 4. The simulation work then continues until P/E characteristics and data retention using the same parameter. Data retention characteristics of graphene-based FG non-volatile cell under gate stress -1/1V is shown in Figure 6.

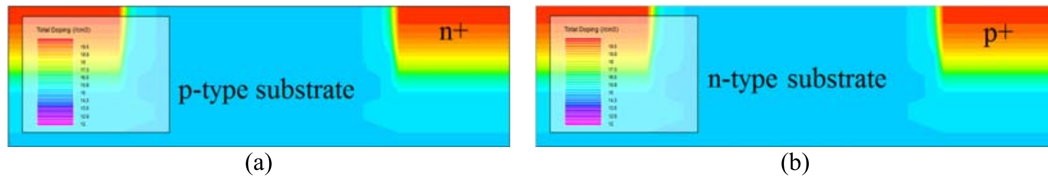
Table 1. Structure dimension and physical parameters of the graphene-based FG non-volatile memory cell used for simulation [6, 11].

Variable	Experimental	Simulation (n-channel)	Simulation (p-channel)
Gate dimension [6]			
Gate length, L_G (nm)	-	600	600
Tunnel oxide thickness, t_{ox} (nm)	8 (SiO_2)	8 (SiO_2)	8 (SiO_2)
Blocking oxide thickness, t_{IPD} (nm)	22 (Al_2O_3)	22 (Al_2O_3)	22 (Al_2O_3)
FG thickness, t_{FG} (nm)	~5	5	5
Control gate thickness, t_{CG} (nm)	15 (TiN)	15 (TiN)	15 (TiN)
Doping profile [11]			
Silicon substrate doping (cm^{-3})	1×10^{15} (p-type)	1×10^{15} (p-type)	1×10^{15} (n-type)
Souce and drain doping (cm^{-3})	1×10^{20} (n-type)	1×10^{20} (n-type)	1×10^{20} (p-type)

Table 2. Extracted FN and BBHE tunnelling coefficients for SiO_2 tunnelling oxide [12, 13].

Variable	Programming	Erasing
FN tunneling coefficients [12]		
A_{FN} (A/V^2)	1.23×10^{-6}	1.87×10^{-7}
B_{FN} (V/cm)	2.37×10^{8}	1.88×10^{8}
BBHE tunneling coefficients [13]		
BB_A (cm^{-1} V^{-2} s^{-1})	9.6615×10^{18}	-
BB_B (V/cm)	3.0×10^{7}	-
BB_γ	2.0	-

For part 2, the simulation works continue by changing the doping type of flash memory cells from n-channel to p-channel (refer Table 1) while all the cell dimensions and parameters remain unchanged (refer Figure 3). The device fabrication for flash MLG device in [6] did not mentioned the exact gate length of the device therefore 600 nm gate length is chosen for simulation due to the standard simulation value for MOSFET. For P/E operations, BBHE and FN tunnelling model was operated to capture the electrons tunnelling through tunnel oxide using the BBHE and FN coefficients that were initially extracted from the optimization work in [12, 13] (refer Table 2). BBHE tunnelling is applied for programming in p-channel flash memory cell due to very poor current-voltage (I-V) characteristics if FN tunnelling is applied. This is because BBHE injection can achieves until 40% reduction of electric field tunnel oxide, E_{ox} of electron flow as compared to FN tunneling barrier [10]. The simulation is done by applying fixed negative biased (-2.0V) to the drain and positive biased control gate (12, 14, 16, 18, 20V) for programming (BBHE tunnelling mechanism) and applying free biased to the drain and negative biased control gate (-12, -14, -16, -18, -20V) for erasing (FN tunnelling mechanism). The simulation continues to simulate the data retention until 10^4s and extrapolate until 10^8s (10 years).



(a) (b)

Figure 3. The doping concentration of (a) n-channel and (b) p-channel flash memory cell.

3. RESULTS AND ANALYSIS

3.1. Device Validation

The charge stored in the FG of graphene-based FG flash memory were extracted using Silvaco ATLAS Tool. The programming transients at 20V programming voltage is used for 1s programming time. Normally, the number of electrons stored in FG after 1s programming at 20V program voltage is about 10^{13} cm^{-2} [6]. Figure 4 shows the FG charge after 1s of programming time is $-2x10^{-14}$ C per unit micron. The electrons start to accumulate in the FG at 10^{-7} s. Therefore, the number of electrons stored in FG after 1s programming at 20V of the program voltage is about $1.25x10^{13}$ cm^{-2}.

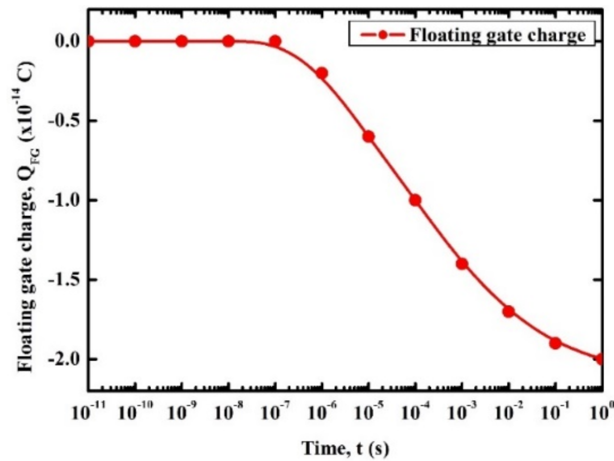
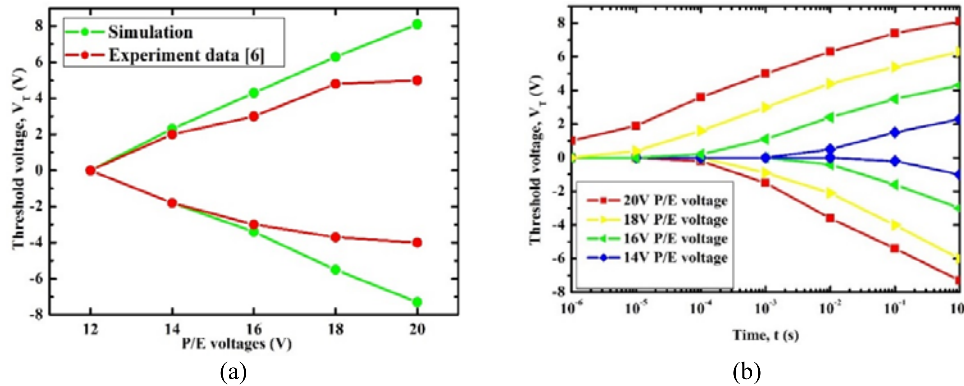


Figure 4. The number of electrons stored after in FG after 1s programming at 20V program voltage (n-channel flash memory cell).

The difference in P/E operations from 0V to ±20V P/E voltages for 1s P/E time is shown in Figure 5(a). The memory window in [6] is about 9.4V at ±20V P/E voltage whereas, for this proposed work, the memory window is 15.4V which is much larger than the experimental work [6]. Figure 5(b) shows the P/E characteristics for several P/E voltages for proposed work from ±14V to ±20V P/E voltages. The threshold voltage shift (ΔV_{th}) for both programming and erasing are increasing with increasing of P/E biases from ±14V to ±20V P/E voltages. Increasing the amplitude biases of the gate will lead to the easiest formation of inversion of the substrate surface. Since the flash memory cell are constructed from p-type silicon substrate (n-channel), application of positive bias would introduce an inversion layer of the silicon surface. Under high positive bias, the electrons were made inversion carriers would experience FN tunnelling via tunnel oxide, and these carriers will fall into the potential well which formed by the graphene sheets in the FG. During erase operation, a high negative bias is applied to the gate terminal. The holes will be accumulated in silicon surface, tunnelling across the tunnel oxide to the FG and recombine with the electrons stored in the FG [14]. The discrepancy between simulation and experimental is because experimental work represents exact behaviour of the object under test with specific measuring errors due to measurement, faulty equipment and also external force from an analyst during the experimental analysis, while, in simulation, only involves with the numerical model in the simulator.



(a) (b)

Figure 5. (a) Comparison of memory window between simulation results and experimental data of graphene-based FG flash memory (b) P/E characteristic of graphene-based FG flash memory for n-channel.

In [6], the charge stored in CSL flash memory cell was tested at room temperature also at high temperature. After 10 years of extrapolation, the memory window obtained is 6.9V (74% charges retained) at room temperature. The simulation was done at room temperature and showed that the wider memory window obtained by the proposed device about 11.6V (75% charges retain) as shown in Figure 6. The discrepancy between simulation and experimental data of memory window is due to non-homogeneity of rMLG in FG (~5nm) during the fabrication process while in a simulation work is precise of 5nm graphene FG. Besides, in experimental work, the fabrication of rMLG has used the method of thermal reduction of graphene oxide (GO). Figure 7(a) and 7(b) show the image of a schematic cross-sectional fabricated rMLG flash memory cell and the atomic force microscopy (AFM) image of the rMLG sheet respectively.

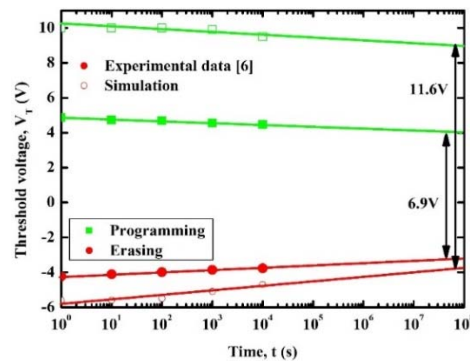
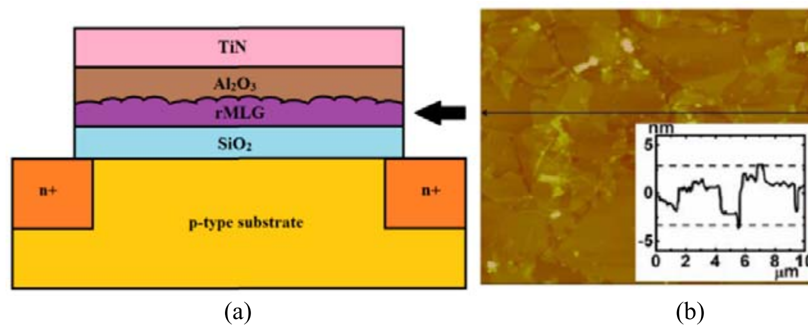


Figure 6. Data retention after 10 years extrapolation: comparison between the proposed device and experimental data [6].



(a) (b)

Figure 7. (a) The view of schematic cross-section rMLG flash memory cell and (b) image of rMLG sheet under AFM [6].

3.2. Performances of P-Channel Flash Memory Cell

P/E characteristics of both n-channel and p-channel flash memory cells are compared in Figure 8(a) and 8(b) respectively. Both P/E speed rise with the increment of P/E voltages from ±14V to ±20V gate biases. The p-channel flash memory cell manifests rapid programming speed compared to erasing at 1s P/E time. However, the erasing characteristics from Figure 8(b), the erasing speed for a p-channel flash memory cell is slower than that of in n-channel flash memory cell under the same voltages. This is because the creation of oxide charges under FN tunnelling mechanism occurs which leads to the increasing of oxide leakage current. When the negative bias applied to the gate, the electrons injection from the FG to n-substrate under FN tunnelling results in electron-hole pairs at the Si-SiO_2 interface and some of the generated holes are injected into the oxide and become trap in the oxide, and then it will increase the oxide field and causes the oxide leakage current. Figure 9 shows that the data retention of n-channel and p-channel flash memory cells after 10 years. The MW for n-channel and p-channel flash memory cells are 15.4V and 15.6V respectively. However, for data retention of p-channel flash memory cell is better compared to n-channel flash memory cell. For n-channel flash memory cell, the charge is retained by 75% (from 15.4V to 11.6V in 10 years) while in p-channel flash memory cell, the charge is retained by up to 80% (from 15.6V to 12.5V). This p-channel flash memory cell shows better reliability due to the BBHE injection mechanism for P/E operations. This superior injection mechanism is assumed due to the low oxide charge and having the longest lifetime compared with the other methods. This superiority is considered to be due to the fewer oxide charges which lead to reducing oxide damage. Besides, in p-channel flash memory cell, the programmed and the erased electron is the minority carrier, and a smaller number of electrons is injected and trapped into the oxide. Therefore, the gate disturbance of the p-channel flash memory cell is superior to that of the n-channel flash memory cell.

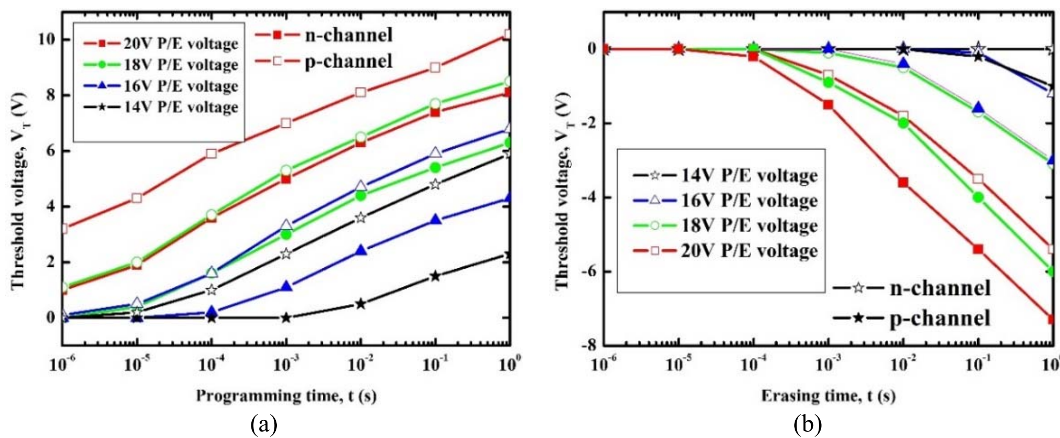


Figure 8. The threshold voltage shift of n-channel and p-channel memory cell for (a) program and (b) erase.

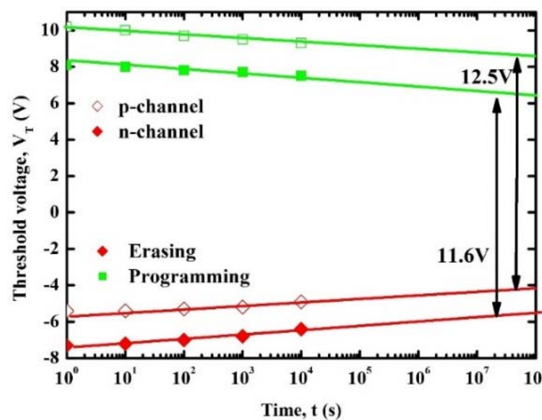


Figure 9. The data retention for both n-channel and p-channel flash memory cell.

4. CONCLUSION

The graphene as charge storage layer flash memory is successfully simulated and validated with experimental work. It shows that the number of electrons stored in graphene-based FG flash memory after 1s at 20V programming voltage is 1.25×10^{13} cm^{-2} while the experimental work is 10^{13} cm^{-2}. The memory window for n-channel and p-channel flash memory cell are 15.4V and 15.6V, respectively, at ± 20V P/E voltage. The data retention after 10 years for n-channel and p-channel flash memory cell are retained by 75% (from 15.4V to 11.6V in 10 years) and by 80% (from 15.6V to 12.5V), respectively. It is concluded that p-channel flash memory cell shows better data retention compared to n-channel flash memory.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the financial support of the Ministry of Higher Education (MOHE), Malaysia under the Research University Grant (GUP) Project No. Q.J130000.3551.07G45. Also thanks to the Research Management Center (RMC) of Universiti Teknologi Malaysia (UTM) for providing an excellent research environment in which to complete this work.

REFERENCES

[1] X. Wang, W. Xie, and J. B. Xu, "Graphene Based Non-Volatile Memory Devices," *Advanced Materials*, vol. 26, pp. 5496-5503, 2014.
[2] M. N. Hossain and M. H. Chowdhury, "Graphene Nanotechnology for the Next Generation Nonvolatile Memory."
[3] S. Raghunathan, T. Krishnamohan, K. Parat and K. Saraswat, "Investigation of Ballistic Current in Scaled Floating-gate NAND FLASH and a Solution," *2009 IEEE International Electron Devices Meeting*. pp. 819-822, 2009.
[4] A. Mishra, H. Kalita, M. Waikar, A. Gour, M. Bhaisare, M. Khare, M. Aslam and A. Kottantharayil, "Multilayer Graphene as Charge Storage Layer in Floating Gate Flash Memory," *4th IEEE International Memory Workshop. Milan*, Italy. 2012.
[5] S. Bertolazzi, D. Krasnozhon, and A. Kis, "Nonvolatile Memory Cells Based on MoS2/Graphene Heterostructures," *American Chemical Society Nano*, vol. 7, pp. 3246-3252, 2013.
[6] A. Mishra, A. Janardanan, M. Khare, H. Kalita, and A. Kottantharayil, "Reduced Multilayer Graphene Oxide Floating Gate Flash Memory with Large Memory Window and Robust Retention Characteristics," *IEEE Electron Device Letters*, pp. 1-3, 2013.
[7] P. Avouris, Y. M. Lin, F. Xia, D. B. Farmer, Y. Wu, T. Mueller, K. Jenkins, C. Dimitrakopoulos, A. Grill, "Graphene-based fast electronics and optoelectronics," *IEEE*, pp. 1-4, 2010.
[8] Y. H. Lin, H. C. You, J. Y. Hsu and J. H. Li, "P-Channel Gd_2O_3 Trapping Layer for SONOS-Type Flash Memory," *Applied Mathematics & Information Sciences Letters,* vol. 9, pp. 141-147, 2015.
[9] S. S. Chung, S. T. Liaw, C. M. Yih, Z. H. Ho, "N-Channel Versus P-Channel Flash EEPROM-Which one has better reliabilities," *IEEE Explore*, pp. 1-, 2001.
[10] T. Ohnakado, K. Mitsunaga, M. Nunoshita**,** H. Onoda, K. Sakakibara, N. Tsuji, N. Ajika, M. Hatanaka and H. Miyoshi, "Novel Electron Injection Method Using Band-to-Band Tunneling Induced Hot Electron (BBHE) for Flash Memory with a P-channel Cell," *IEEE Explore,*pp. 1-4, 1995.
[11] T. Thingujam, K. Jolson, M. Kumar, S. K. Sarkar. "TCAD based modeling and simulation of Graphene Nanostructured FET (GFET) for High Frequency Performance," *Journal of Engineering Technology*, vol. 6(1), pp. 1-5, 2017.
[12] A. Hamzah, N. E. Alias, and R, "Ismail. Low-voltage high-speed programming gate-all-around floating gate memory cell with tunnel barrier engineering," *Japanese Journal of Applied Physics*, vol. 57: 1-8, 2018.
[13] *Atlas User's Manual Device Simulation Software*, Santa Clara. 2013.
[14] N. M. Hossain, M. H. Chowdhury, "Multilayer Graphene Nanoribbon Floating Gate Transistor for Flash," *ACM Journal on Emerging Technologies,* vol. 12(1), pp. 1-17, 2015.