ARTICLES

Journal of Semiconductors
(2020) 41, 102105
doi: 10.1088/1674-4926/41/10/102105

Design, modelling, and simulation of a floating gate transistor with a novel security feature

H. Zandipour[1, †] and M. Madani[2]

[1]Department of Physics, Georgia Southern University, Savannah, GA 31419, USA

[2]Department of Electrical Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

Abstract: This study proposes a new generation of floating gate transistors (FGT) with a novel built-in security feature. The new device has applications in guarding the IC chips against the current reverse engineering techniques, including scanning capacitance microscopy (SCM). The SCM measures the change in the *C–V* characteristic of the device as a result of placing a minute amount of charge on the floating gate, even in nano-meter scales. The proposed design only adds a simple processing step to the conventional FGT by adding an oppositely doped implanted layer to the substrate. This new structure was first analyzed theoretically and then a two-dimensional model was extracted to represent its *C–V* characteristic. Furthermore, this model was verified with a simulation. In addition, the *C–V* characteristics relevant to the SCM measurement of both conventional and the new designed FGT were compared to discuss the effectiveness of the added layer in masking the state of the transistor. The effect of change in doping concentration of the implanted layer on the *C–V* characteristics was also investigated. Finally, the feasibility of the proposed design was examined by comparing its *I–V* characteristics with the traditional FGT.

Key words: floating gate transistor (FGT); scanning capacitance microscopy (SCM); metal–oxide–semiconductor (MOS) capacitance; non-volatile memory (NVM); reverse engineering

Citation: H Zandipour and M Madani, Design, modelling, and simulation of a floating gate transistor with a novel security feature[J]. *J. Semicond.*, 2020, 41(10), 102105. http://doi.org/10.1088/1674-4926/41/10/102105

1. Introduction

Non-volatile memories have many applications, including industrial, military, and general public[1–3]. A floating gate transistor (FGT) can be used to store a bit of information[4, 5]. For instance, in an FGT, the transistor is a metal–oxide semiconductor field-effect transistor (MOSFET) with a floating gate (FG) that can either be charged by positive or negative charges to represent an on or off switch[6, 7]. In other words, the charge on the FG will change the threshold voltage of the FGT transistor[8]. An array of FGTs can be used to construct an electrically erasable programmable read only memories (EEPROMs) or flash-EEPROM[9–11]. These types of memories are commonly used in microprocessor/microcontroller designs, which generally gives the user the ability to program the IC chips. More importantly, this would obfuscate the interconnections and functionality of an IC chip because some of the units are interconnected through the programmed memory. In this case, the traditional non-destructive reverse engineering methods such as X-ray imaging[12] are ineffective to extract the functionality of the chip because the units are not directly connected together but are connected through FGTs. However, there have been several attempts to recover the stored data on the FGT, which raises a concern in the security of the IC chip design that is a key issue for electronic chip manufacturers[13]. Almost all of the successful methods are destructive and need backside preparation to recover the data from each FGT (i.e. distinguishing an "on" transistor from and "off" one by detecting the charge on the FG). Several successful attempts have been made to recover the presence of charge on the FGs. Some notable techniques are scanning Kelvin probe microscopy (SKPM)[14], laser timing probe (LTP)[15] and scanning capacitance microscopy (SCM)[16, 17]. In SKPM, the entire silicon needs to be removed from the back of FGT. This method measures the surface potential (i.e. electric field) and is considered to be ineffective for the smaller feature size device technologies (holding less than 1000 electrons on the FG)[13]. Meanwhile, the LTP requires partial delayering of the substrate and remaining layer must be finely polished. It is worth mentioning that this delayering should not exceed the limit of 50–200 nm to keep the silicon undamaged[13]. This method measures the change in optical properties such as absorption coefficient, which is altered by a change in carriers density caused by a charged FG. This method also becomes inefficient for smaller feature size designs since the variation in optical properties would be infinitesimal and requires equipment with much higher accuracy. The scanning capacitance microscopy (SCM)[16, 17] is one these techniques and is believed to be among the most powerful ones[13]. It measures the capacitance between the control gate (the gate above the FG, which is accessible for programing) and the bottom of the substrate. The substrate needs to be delayered first but it does not require fine polishing. In the SCM setup, the control gate and substrate are connected to a high frequency small signal source with a variable DC offset voltage. By sweeping the DC offset voltage, a *C–V* curve is drawn. The shift in the *C–V* curve is then used to identify the state of each FGT. Unlike the two previous methods, the SCM is even capable of detecting the charge on the FG for smaller feature size technologies.

Correspondence to: H Zandipour, hzandipour@georgiasouthern.edu
Received 2 MARCH 2020; Revised 2 APRIL 2020.
©2020 Chinese Institute of Electronics

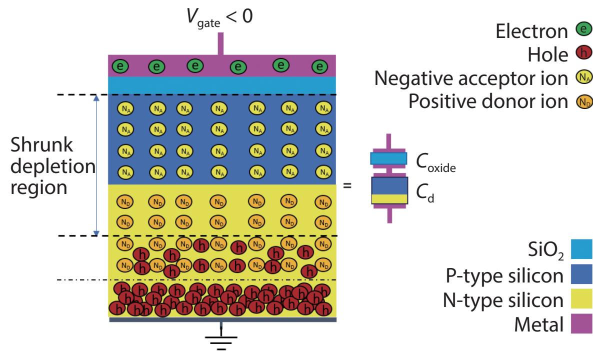

Fig. 1. (Color online) A MOPNS under the influence of small negative voltage applied to its gate.

In the following, a new FGT structure is presented and its C–V characteristics is modelled. Then, the trend of this model is compared to a 2D COMSOL simulation to validate the work objectives. In addition, the C–V curve obtained from the newly designed structure is compared to the conventional structure to investigate the effectiveness of this new design against SCM reverse engineering technique. Finally, an analysis of the I–V characteristics examines the feasibility of this design.

2. Proposed structures (MOPNS and MONPS)

Our proposal is to form an oppositely doped compensated silicon region by ion implanting of opposite dopant in the extrinsic substrate a distance away from the SiO_2–Si interface. This must be done in a way that the structure ends up with at least slightly thicker oppositely doped layer than the remaining substrate even after the back-side delayering (similar to the structure in Fig. 1). For our reference, we would like to call the structure with P-type substrate, the metal oxide P-type N-type semiconductor (MOPNS) and similarly the structure with N-type substrate, the metal oxide N-type P-type semiconductor (MONPS). From now on, the focus will be on the MOPNS, which could be easily altered to a MONPS for future applications. To start analyzing the structure, we consider the case that both regions have the same effective opposite dopant concentrations. One might say that if the substrate region depth is small enough (in tens of nm), then almost all its majority carriers will be diffused to the oppositely doped substrate (P-type region) and leave their negative ions behind; that is to say, the holes are being diffused to meet with the electrons for annihilation, and vice versa. This would happen until there are few majority carriers left in P-type region (i.e. it is depleted from its majority carries). Relatively, the N-type region will be depleted for the same depth when there is no voltage applied to the gate at thermal equilibrium because both regions have the same dopant concentration. The capacitance can also be found by treating the structure as two parallel plates capacitor by knowing the depth of this depleted region. The total capacitance, similar to an MOS, has a constant oxide capacitance (C_{oxide}) in series with the capacitance from the depleted silicon (C_d) below the oxide layer as it is represented in Fig. 1. Furthermore, we know that in the capacitance measurement, the bottom of the device is grounded that can be considered as an abundant source of holes or electrons but these electrons and holes cannot invade the depletion region since any excess charge

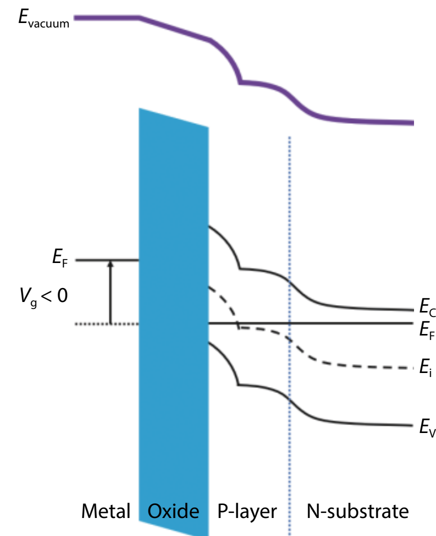

Fig. 2. (Color online) The upward band bending in energy band diagram when a small negative voltage is applied to the gate of a MOPNS.

would unbalance the created internal electric field. However, by adding small negative or positive potential to the gate (i.e. creating an external electric field) some of those holes or electrons are eventually going to push through the depletion region and make it tighter. To clarify this explanation, the graphical representation in Fig. 1 displays a MOPNS when a small negative potential is applied to its gate. In the high frequency case, increasing the applied potential of the gate causes this depletion region to shrink to a minimum value, which causes the capacitance to go up to a certain point. Since there is not enough time for the thermal generation of electrons or hole, the depletion region stands still. However, in the case of low frequency, the depleted region vanishes completely and only the oxide capacitance will be involved.

From the energy band diagram point of view, one might say that there will be an imposed potential over the oxide and the semiconductor when a voltage is applied to the gate. By considering the energy band diagram, this applied potential to the semiconductor would cause band bending (upward or downward depending on the polarities of the applied voltage) close to the oxide interface. For example, an upward band bending can be shown in the Fig. 2 when a negative potential is applied to the gate of a MOPNS.

By finding the depletion region's depth, we will be able to model the device capacitance. To reach this goal, one should start by finding the electric field. It is worth mentioning the boundary condition that assumes the electric field to be zero outside of the depleted N-type region ($E(x > x_d) = 0$) as it can be seen in Fig. 3.

The electric fields in both P-type (E_1) and N-type (E_2) regions can be found by integrating the charge density divided by the permittivity over each region (from Gauss's law). The electric field in each region can be found as:

$$E_1(x) = \frac{-q}{\varepsilon_{Si}}(x_d N_d + N_a x), \tag{1}$$

$$E_2(x) = \frac{-qN_d}{\varepsilon_{Si}}(-x + x_d), \tag{2}$$

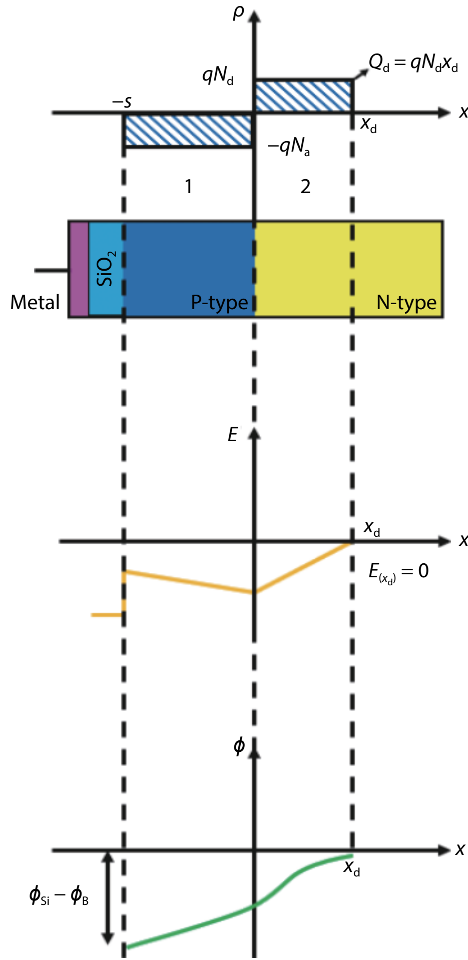

Fig. 3. (Color online) Graphical representation of charge density, electric field and electric potential of a MOPNS.

where q is the elementary charge, ε_{Si} is the permittivity of the silicon, N_a and N_d are the acceptor and donor concentration of the P-type and N-type, respectively, and x_d is the depth of the depleted N-type region.

By having the electric field, one can calculate the electric potential over the semiconductor by integrating the electric field (with a negative sign) over both regions. So, the potential drop over the semiconductor can be calculated as:

$$\phi_{Si} = \frac{q}{2\varepsilon_{Si}}\left[N_aS^2 - 2N_dx_dS - N_dx_d^2\right] + \phi_B, \tag{3}$$

where S is the depth of the P-type region that is considered to be depleted and ϕ_B is known as the built-in potential of the semiconductor.

We also know that the applied voltage to the gate (V_{gate}) would be the sum of the potential on the oxide (V_{oxide}), flat-band voltage (V_{fb}) and the potential drop over the semiconductor.

$$V_{gate} = V_{oxide} + V_{fb} + \phi_{Si}. \tag{4}$$

One can easily obtain the V_{oxide} by finding the electric field at the oxide interface ($E_1(-S)$). By considering the fact that electric displacement must be continuous at the silicon-oxide interface, Eq. (4) can be rewritten as:

$$V_{gate} = \frac{q}{C_{ox}}\left[N_aS - N_dx_d\right] + V_{fb} + \phi_{Si}. \tag{5}$$

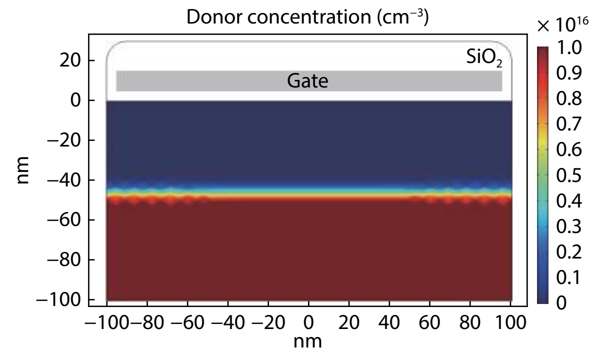

Fig. 4. (Color online) Donor concentration of a simulated MOPNS.

Now, one may calculate the x_d first and then the total capacitance of the device with respect to the applied gate voltage. After finding the x_d, one can obtain the depletion capacitance as:

$$C_d = \frac{\varepsilon_S}{x_d + S}. \tag{6}$$

3. Results and comparison

To validate the driven formula, a 2D COMSOL simulation has been done on a MOPNS with the same dimensions as Fig. 4 to find the low (0.1 Hz) and high (10^9 Hz) frequency capacitance of the device versus the offset voltage applied to the gate with the work function of 4 V. As can be seen in Fig. 4, the bottom N-type region has the donor concentration of $N_d = 10^{16}$ cm^{-3} that is considered to be equal to the acceptor concertation (not shown here) on the top P-type region. The amplitude of the small signal considered to be infinitesimal (V_{ac} = 100 μV) with respect to the change in the applied offset when the bottom of the structure is grounded (V_{Base} = 0 V). A comparison between the normalized C–V characteristics obtained from simulation and the simple obtained model can be found in Fig. 5. This demonstrates that our extracted model follows the similar pattern as the simulation, so the model could be used to describe the C–V characteristics of the proposed structure. However, the simulation provides greater details and also can take the effect of applied high frequency into account, which was not implemented in the simple model. For instance, it is worth noticing that the capacitance is slightly lower when a relatively strong positive voltage is applied to the gate than a relatively strong negative voltage. One may explain that the shrunk region is expected to be smaller for a relatively high negative than the positive potential, since the penetrating electrons feel a greater repelling force from the negative acceptor ions as they punch into the deletion region in contrast to penetrating holes, which feel a greater attractive force. The opposite is expected to be seen for a MONPS, of course. Since we are only interested in high frequency capacitance measurement in SCM, using the simulation results for the high frequency C–V curves is favourable.

To see how effective this design could be, a comparison between the high frequency C–V behaviours of an FG MOPNS and a traditional P-type FG MOS is depicted in the Fig. 6, which may also represent the outcome of the SCM technique. Both structures have the same structure size as Fig. 7 and also equal dopant concentrations as $N_a = N_d = 10^{16}$ cm^{-3}

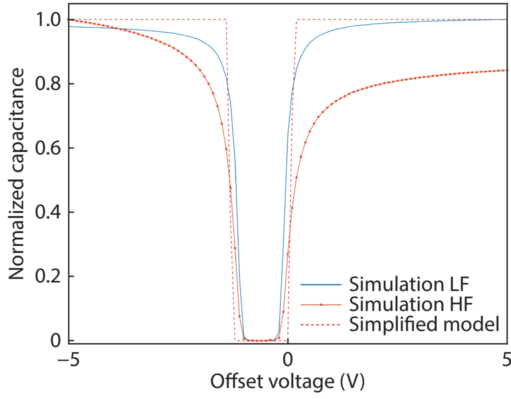

Fig. 5. (Color online) Comparison between the extracted model and simulation results (for HF and LF) of normalized $C-V$ curves for the MOPNS.

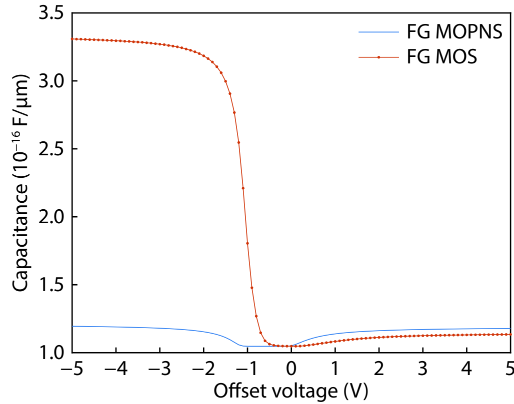

Fig. 6. (Color online) Comparison of HF C–V curves of the FG MOPNS and FG MOS with P-type substrate.

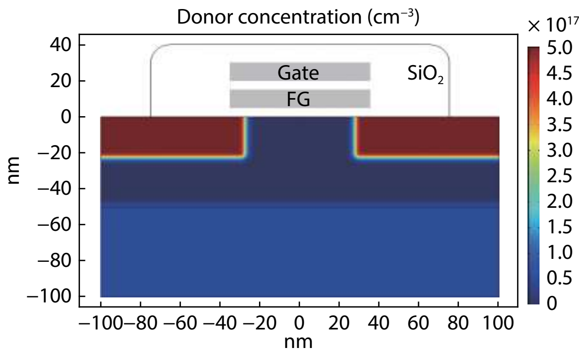

Fig. 7. (Color online) Donor concentration of a MOPNS transistor.

when their FGs are left un-charged. As can be seen in Fig. 6, in addition to a complete change in the C–V behaviour of the FG MOPNS compared to the P-type FG MOS, there is a significant drop (64%) in the maximum obtained capacitance as well as a serious drop (1600%) in the difference between the maximum and minimum.

To expand our investigation on the effect of the added layer, a comparison between an FG MOPNS with three different dopant concentrations of oppositely doped region (N-type) is done. Fig. 8 shows this comparison between three different donor concentrations of $N_{d1} = 10^{14}$ cm^{-3}, $N_{d2} = 10^{16}$ cm^{-3}, and $N_{d3} = 5 \times 10^{17}$ cm^{-3} for the oppositely doped region when the dopant of substrate is kept constant at $N_a = 10^{16}$ cm^{-3}. These dopant concentrations represent three

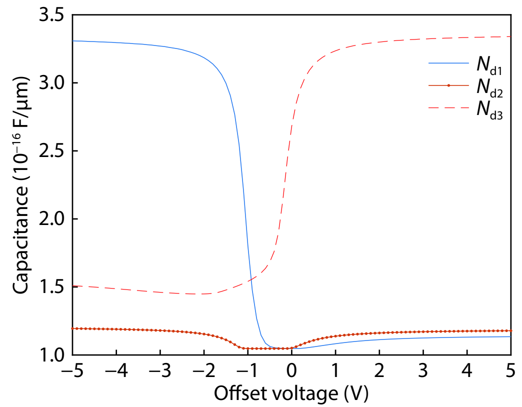

Fig. 8. (Color online) Comparison of HF C–V curve of the FG MOPNS with three different dopant concentrations for the oppositely doped region (N-type) when the dopant concentration of the P-type region is constant.

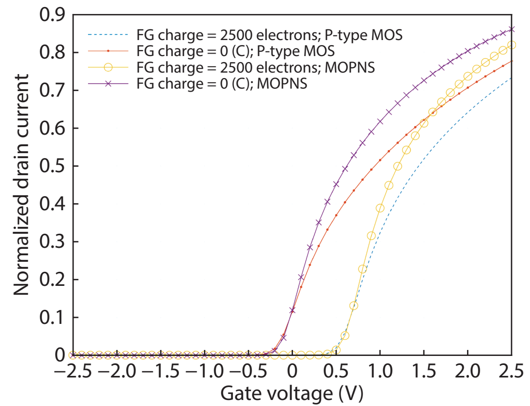

Fig. 9. (Color online) Comparison between I–V characteristics of FG MOPNS and FG MOS (P-type).

cases of lightly doped, the same dopant and highly doped respectively. As can be seen in this figure, the structure can simply be treated as a traditional P-type FG MOS when the dopant concentration of the implanted region is a lot smaller than the substrate, which notes the fact that the oppositely doped layer had no significant effect on the C–V curve of the device. Meanwhile, quite opposite is true when the dopant concentration of the implanted region is a lot higher than the substrate, and the structure changes its behaviour to an N-type FG MOS. Additionally, as was observed earlier, the C–V curve of the device alters to have a new characteristic (shown in Fig. 5) when the dopant concentration of the substrate is equal to the oppositely doped region. Results from equal and greater dopant concentration of the oppositely doped region are both desirable because they indicate a significant change in the C–V characteristic of the device in comparison to the traditional P-type FG MOS. Consequently, it can be said that the change in dopant concentration could add a security feature to this device. In other words, choosing the right device with the right C–V curve can become challenging.

Even though the obtained results from the C–V curve seems desirable, it is important to have a practical device otherwise the design would be useless. To have a functional device, the structure needs to have a similar functionality (i.e. I–V characteristic) as a simple P-type FG MOS. To take

this into account, normalized *I*–*V* characteristics of an FG MOPNS transistor and a P-type FG MOS transistor are depicted in Fig. 9. Both structures have the same size as Fig. 7. The FG MOPNS transistor have the dopant concentrations of $N_d = N_a = 5 \times 10^{16}$ cm^{-3} in the substrate and the donor concentration $N_{d(drain/source)} = 5 \times 10^{17}$ cm^{-3} for drain and source regions. These values are the same for the P-type FG MOS transistor. Their drains are held at 100 μV while the sources and the bases are grounded.

This comparison shows an insignificant change in the *I*–*V* characteristics. That is to say, the desirable shift in the *I*–*V* characteristic is clearly observed when the FG is charged, which supports the idea that our FG MOPNS transistor has similar functionality as a traditional P-type FG MOS transistor.

4. Conclusion

In this work, we have presented a new FG transistor design structure that can be used to safeguard the states of FG transistors against a vigorous reverse engineering skim, SCM. To develop a more efficient design, we represented the idea of implanting an oppositely doped region at the bottom of the silicon substrate, close to the oxide interface beyond delayering limit without changing the transistor normal operation characteristics. A simple model was extracted to describe the new design's *C*–*V* characteristic, which was then validated by the designated simulation. Later, a comparison between an FG MOPNS and a traditional P-type FG MOS showed a significant drop in the maximum measured capacitance, as well as a huge drop in the difference between maximum and minimum measured capacitance, when both the implanted region and the substrate have the same dopant concentration. Furthermore, it was observed that the dopant of this implanted region has a great influence on the *C*–*V* characteristic of the device. Our results have shown to have the same trend as a P-type FG MOS when the dopant concentration of the implanted layer is very low compared to the substrate. Meanwhile, the behavior observed to be similar to an N-type FG MOS when the dopant concentration of the implanted layer is very high compared to the substrate. In addition, our results show that the new device can be tuned to have different *C*–*V* characteristics. This adds to the complexity of choosing the right device with the right *C*–*V* characteristic, which may be considered as a security feature. Manufacturers may also take advantage of this result to obfuscate the stored data on their chip by implanting this layer at arbitrary spots with equal or greater dopant concentration than the substrate. Finally, the *I*–*V* characteristic of an FG MOPNS transistor were compared to an FG MOS transistor and very little difference in behavior was observed. This indicates that the proposed structure has the same functionality as a traditional FG MOS transistor. One might say that the attained results are encouraging and they point to a potentially practical design. It is worth noting that this design is not costly because the required ion implantation process is already implemented in the bipolar transistor fabrication to increase the collector's efficiency.

References

[1] Fujita O, Amemiya Y. A floating-gate analog memory device for neural networks. IEEE Trans Electron Devices, 1993, 40, 2029

[2] Wang G, Liu X, Wang W. Solution processed organic transistor nonvolatile memory with a floating-gate of carbon nanotubes. IEEE Electron Device Lett, 2018, 39, 111

[3] Jia X, Feng P, Zhang S, et al. An ultra-low-power area-efficient non-volatile memory in a 0.18 μm single-poly CMOS process for passive RFID tags. J Semicond, 2013, 34, 085004

[4] Lee J, Jeong Y, Jeong H, et al. Fabrication and characterization of a new EEPROM cell with spacer select transistor. IEEE Electron Device Lett, 2005, 26, 569

[5] Fang L, Kong W, Gu J, et al. A novel symmetrical split-gate structure for 2-bit per cell flash memory. J Semicond, 2014, 35, 074008

[6] Cacharelis P, Fong E, Torgerson E, et al. A single transistor electrically alterable cell. IEEE Electron Device Lett, 1985, 6, 519

[7] Dimaria D, Demeyer K, Dong D. Electrically-alterable memory using a dual electron injector structure. IEEE Electron Device Lett, 1980, 1, 179

[8] Wu J, Zhang L Q, Yao Y, et al. Investigation of dynamic threshold voltage behavior in semi-floating gate transistor for normally-off AlGaN/GaN HEMT. IEEE J Electron Devices Soc, 2017, 5, 117

[9] Kim Y M, Kim S J, Lee J S. Organic-transistor-based nano-floating-gate memory devices having multistack charge-trapping layers. IEEE Electron Device Lett, 2010, 31, 503

[10] Schauer H, Tran L V, Smith L. A high-density, high-performance EEPROM cell. IEEE Trans Electron Devices, 1982, 29, 1178

[11] Kolodny A, Nieh S, Eitan B, et al. Analysis and modeling of floating-gate EEPROM cells. IEEE Trans Electron Devices, 1986, 33, 835

[12] Holler M, Guizar-Sicairos M, Tsai E H R, et al. High-resolution non-destructive three-dimensional imaging of integrated circuits. Nature, 2017, 543, 402

[13] Quadir S E, Chen J, Forte D, et al. A survey on chip to system reverse engineering ACM. J Emerg Technol Comput Syst, 2016, 13, 1

[14] Henning A K, Hochwitz T, Slinkman J, et al. Two-dimensional surface dopant profiling in silicon using scanning Kelvin probe microscopy. J Appl Phys, 1995, 77, 1888

[15] Bidani L, Baharav O, Sinvani M, et al. Usage of laser timing probe for sensing of programmed charges in EEPROM devices. IEEE Trans Device Mater Reliab, 2014, 14, 304

[16] Matey J R. Scanning capacitance microscopy. Scan Microscopy Technol Appl, 1988, 897, 110

[17] Denardi C, Desplats R, Perdu P, et al. Descrambling and data reading techniques for flash-EEPROM memories. Application smart cards. Microelectron Reliab, 2006, 46, 1569