Modeling the Thermal Characteristics of Stacked 2T0C Memory Array Based on $InGaZnO_4$ Thin-film Transistors

Song He[1], Haoxin Li[1], Guangwei Xu[1,*], Xinyi Tang[1], Yuanbiao Li[1], Jaewoo Kim[2], Tingting Gu[2], Xingkun Xue[2], Zelun Li[2], Handong Xu[2], Haiyang Dong[2], Kai Zhou[1], Xianqin Hu[1,2], Shibing Long[1]

[1]*School of Microelectronics, University of Science and Technology of China, Hefei 230026, China;*

[2]*Changxin Memory Technologies, Inc. Hefei 230601, China. Emails: xugw@ustc.edu.cn..*

***Abstract*—For the first time, the electrothermal characteristics of the stacked 2T0C memory array were modeled based on the material properties, electron transport mechanism of $InGaZnO_4$ (IGZO) and basic Fourier heat flow equation utilizing the technology computer-aided design (TCAD) tool. According to the proposed electrothermal model, the multi-layer stacked 2T0C memory array would bring severe self-heating issue from the thermal crosstalk among layers. The temperature rising of stacked 2T0C cells decrease the retention time severely due to the temperature instability of IGZO thin-film transistors TFTs. Furthermore, several advanced device-level strategies were proposed to ameliorate the heat accumulation problem.**

***Keywords*—IGZO, 2T0C DRAM, stacked memory array self-heating, TCAD simulation.**

I. Introduction

Nowadays, the dynamic random access memory (DRAM) is not only applied to the personal computer or server, but also plays a crucial role in various emerging network technology such as Artificial Intelligence [1], Cloud/Edge Computing, Internet of Things, etc. The increasing application scenes intensifies the demand for higher storage density of DRAM. Since the area consumption of traditional 1T1C DRAM is strictly limited by the need of sufficient capacitance, the new-type 2T0C configuration DRAM is expected to increase storage density further because of its potential for monolithic 3D stack [2]. $InGaZnO_4$ thin-film transistors (IGZO-TFTs) have been widely used in the drive backplane of high resolution and low power display with mature low temperature process and good large area uniformity since it was firstly invented in 2004 [3, 4]. The back end of the line (BEOL) compatibility makes the 2T0C DRAM configuration more promising. The IGZO-TFTs are introduced to 2T0C construction due to its extremely low off-state leakage [5]. It is generally considered that the low off-current of IGZO-TFTs is induced by the intrinsic low carrier concentration of IGZO deriving from its relatively wide bandgap (3.2 eV) and conduction restriction of electrons of the potential barrier above the conduction band [4]. Therefore, the specific storage capacitor is unnecessary for 2T0C DRAM configuration, which brings its stacking-compatibility. However, multi-layer stacked 2T0C memory array may bring severe self-heating issue from the thermal crosstalk among layers [6]. In addition, the thermal conductivity of IGZO (0.014 $W \cdot K^{-1}cm^{-1}$) is far lower in contrast to silicon (1.7 $W \cdot K^{-1}cm^{-1}$), which exacerbates the heat accumulation in the channel.

In this work, the stacked 2T0C memory model was built to predict the potential thermal issue utilizing the TCAD Sentaurus. According to this model, we predicted the potential self-heating issue in the scene of current-sensing reading approach [7]. The storage performance of the 2T0C memory cell degrades severely with the temperature rising. We also bulit several different stacked 2T0C cell structures and studied their thermal characteristics, which will give advanced device-level strategies to ameliorate the self-heating issue for 2T0C DRAM configuration.

II. Electrothermal Modeling

IGZO is normally a n-type semiconductor because of its extremely low hole mobility and native oxygen vacancy (V_O) as donor-like states satisfying Gaussian distribution. The tail states obey exponential distribution acting as electron traps below the conduction band and hole traps above the valence band [8, 9]. The density of tail states and V_O states can be expressed by

$$g_{CBa} = g_{ta} exp\left[\frac{(E - E_C)}{E_a}\right] \quad (1)$$

$$g_{VBd} = g_{td} exp\left[\frac{(E_V - E)}{E_d}\right] \quad (2)$$

$$g_{Gd} = g_d exp\left[\frac{(E - \lambda)^2}{\sigma^2}\right] \quad (3)$$

where E_C and E_V are conduction and valence band edge energies, g_{ta} and g_{td} are densities of tail states at $E = E_C$ and E_V acting as electron and hole traps, E_a and E_d are slopes of conduction and valence band tail states, g_d is the peak value, λ is the mean energy, and σ is the standard deviation of V_O states. The tail states below the conduction band determines the field-effect mobility of IGZO TFTs. In addition to the V_O states, the unintentional hydrogen diffusion from processing also acts as an important role of donor-like states in IGZO TFTs. In the effect of the ion scattering and phonon scattering mechanism, the temperature-dependent mobility of IGZO can be expressed by

$$\mu = \mu_0 \cdot \frac{2}{\left(\frac{T}{T_0}\right)^{\frac{3}{2}} + \left(\frac{T}{T_0}\right)^{-\frac{3}{2}} \cdot \frac{N_D}{N_{ref}}} \quad (4)$$

where T_0 is the room temperature which is 300 K here, N_{ref} is the reference concentration for doping-dependent mobility model, and μ_0 is the constant electron mobility. It's obvious that the mobility degradation is mainly induced by phonon scattering at higher temperature. Besides, the high-field velocity saturation model is implemented in our simulation as well. The basic material parameters of IGZO are shown in TABLE I, which are adjusted by the agreement with experimental transfer characteristics of IGZO-TFTs proposed by IMEC [10], as shown in Fig. 1. The details of the device model were illustrated in Fig. 2a.

TABLE I
KEY PARAMETERS USED IN SIMULATION

Symbol	Description	Value	Units
N_c/N_v	Effective conduction/valence band DOS	$5\times10^{18}/5\times10^{18}$	cm^{-3}
g_{ta}/g_{td}	Density of tails states at $E=E_c/E_v$	$4\times10^{19}/4\times10^{19}$	$cm^{-3}\cdot eV^{-1}$
E_a/E_d	Conduction/valence-band-tail slope	35/35	meV
E_g	Bandgap	3.20	eV
χ	Affinity	4.16	eV
ε	Permittivity	10	
μ_n/μ_p	Electron/hole mobility	20/0.1	$cm^2V^{-1}s^{-1}$
N_{ref}	Reference concentration for doping-dependent mobility model	1.25×10^{18}	cm^{-3}
V_{sat}	Saturation velocity	3.8×10^5	$cm\cdot s^{-1}$
τ_n/τ_p	Electron/hole lifetime	10/2	ns
g_d	Peak of V_O states	8×10^{19}	$cm^{-3}\cdot eV^{-1}$
λ	Mean energy of V_O states	0.85	eV
σ	Standard deviation of V_O states	180	meV
κ	Thermal conductivity	0.014	$W\cdot K^{-1}cm^{-1}$
C_p	Specific heat capacity	2.50	$J\cdot K^{-1}cm^{-3}$
R_{th}	Thermal resistance of top/bottom boundary	2.0/0.21	$K\cdot W^{-1}cm^{-2}$

According to the Fourier heat flow equation, the thermal behavior of stacked 2T0C memory array can be simply described by,

$$\rho C_p \frac{\partial T}{\partial t} - \nabla \cdot (\kappa \nabla T) = Q \quad (5)$$

where κ is the thermal conductivity, T is the temperature, t is time, C_p is the thermal capacity, ρ is the mass density of material, and Q is the thermal power. The joule heat of electron current occupies almost total heat of device which is shown as

$$Q = \frac{|\overrightarrow{J_n}|^2}{qn\mu_n} \quad (6)$$

where $\overrightarrow{J_n}$ is the current density in channel, q is the electron charge, n is the electron density and μ_n is the electron mobility of IGZO. It is assumed that the double-side cooling scheme with the heat dissipation paths of chip top and bottom were adopted in the stacked 2T0C Memory array. In addition, we only considered the thermal resistance of packaging on top and bottom surface in our simulation [6, 11]. The thermal resistance per square centimeter is calculated by

$$R_{th} = t \cdot \frac{1}{\kappa_p} \quad (7)$$

where t and κ_p is the thickness and thermal conductivity of packaging material. The packaging is described as 100 μm epoxy molding compound (EMC 5×10^{-3} $W\cdot K^{-1}cm^{-1}$) and 20 μm die attach film (DAF 9.5×10^{-3} $W\cdot K^{-1}cm^{-1}$). Then the thermal boundary condition is expressed by

$$\kappa\vec{n} \cdot \nabla T = \frac{T_{surface} - T_0}{R_{th}} \quad (8)$$

where $\vec{n}$ is the unit vector in the direction of the outer normal, $T_{surface}$ is the temperature of device boundary and T_0 is the constant temperature which is 300 K. Note that the periodic thermal boundary condition on the left and right side of device is adopted to simulate the 2T0C cells distribution horizontally of large-scale memory array, which is expressed by

$$T|_{x=x_{min}} = T|_{x=x_{max}} \quad (9)$$

$$\kappa\overrightarrow{n_1} \cdot \nabla T|_{x=x_{min}} + \kappa\overrightarrow{n_2} \cdot \nabla T|_{x=x_{max}} = 0 \quad (10)$$

where $\overrightarrow{n_1}$ and $\overrightarrow{n_2}$ are the unit vectors in the direction of the outer normal on both sides.

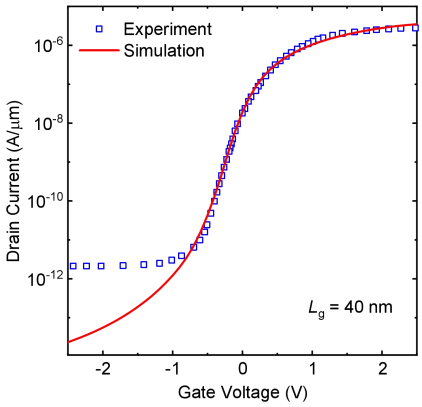

Fig. 1. Comparison between the experimental and simulated transfer characteristic curves. The simulated on-state I-V curve shows good agreement with experimental data [10]. The off-state region of the experimental transfer curve does not fit well with the simulation one due to the detection limitation of the test equipment.

III. RESULTS AND DISCUSSION

The 2T0C cell is implemented by connecting the source electrode of the write transistor (Wtr) to the gate electrode of the read transistor (Rtr) so that the oxide capacitance C_{ox} of the Rtr is used as the storage capacitor, as shown in Fig. 2a. Firstly, we simulated self-heating characteristics of 2T0C memory cell in single working cycle including a write operation followed by a read operation. For the write operation, the storage node (SN) is charged through the write transistor by applying high voltage to the write word-line (WWL) and write bit-line (WBL) during the write

operation. Then the write transistor is switched off by applying negative voltage on the WWL. During the read operation, the SN voltage can be sensed by electric current of read transistor by means of applying bias voltage on the read word-line (RWL) and read bit-line (RBL), as shown in Fig. 2b. The temperature distribution of device varying with time is shown in Fig. 2c. The heat accumulation in write operation is far lower than read operation and has almost dissipated completely before the read operation begins. The electron Joule heat in read operation is much more than write operation because of the constant current of read transistor in the scene of current-sensing reading approach. Therefore, the maximum temperature rising in read operation is much higher than write operation, as shown in Fig. 2d.

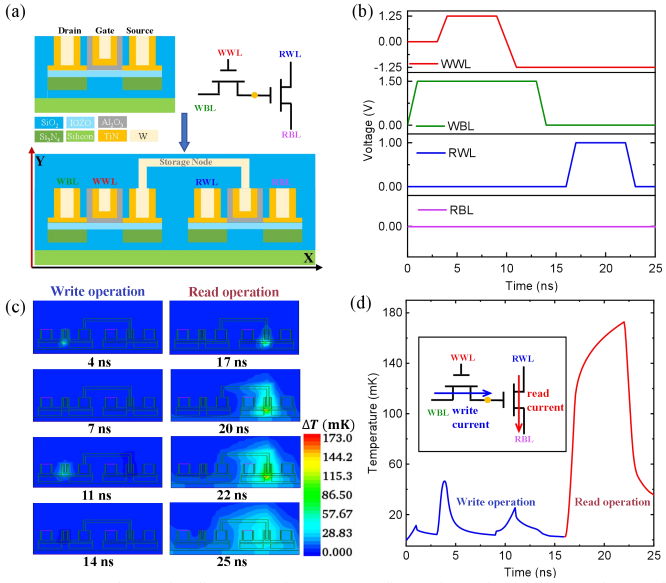

Fig. 2 (a) Schematic diagram of 2T0C configuration. (b) The waveform of write operation followed by read operation. (c) The temperature distribution and (d) maximum temperature rising of our 2T0C cell through write and read operation.

Several stacked 2T0C memory array structures were modeled to simulate the different thermal characteristics in device level. We keep SN voltage constant by applying constant positive voltage on the WWL and WBL and apply periodic pulse voltage on the RWL to imitate the DRAM cycle operation. Then the dynamic equilibrium working temperature was simulated. Note that we simply considered the heat generation of reading operation here according to the simulation results above and adopt the "worst case" that all the DRAM cells are read out at the same time. Therefore, every cell in the same layers is equivalent and the periodic thermal boundary condition is activated in simulation.

The equilibrium temperature is approximately linear rising with the increasing of stack layers since the total power of all layers linearly increase and the main thermal resistance originates from packaging of top and bottom surface, as shown in Fig. 3a. The decreasing of duty cycle of the pulse voltage applied on the RWL reduces the equilibrium temperature as the average power consumption lowers, as shown in Fig. 3b. With the oxide interlayer between 2T0C memory array layers getting thicker, the equilibrium temperature is almost unchanged but the critical time to reach thermal equilibrium extends since the total power and thermal resistance are nearly unchanged but the thicker oxide slows down the thermal crosstalk among layers, as shown in Fig. 3c. At last, the extra width of oxide sidewall between neighboring 2T0C cells in the same layer decreases the equilibrium temperature on account of the reduction of the thermal crosstalk between horizontally neighboring 2T0C memory cells, as shown in Fig. 3d. The thermal characteristics of different stacked 2T0C cell structures we proposed may give advanced guidance for the thermal management scheme of 2T0C DRAM configuration.

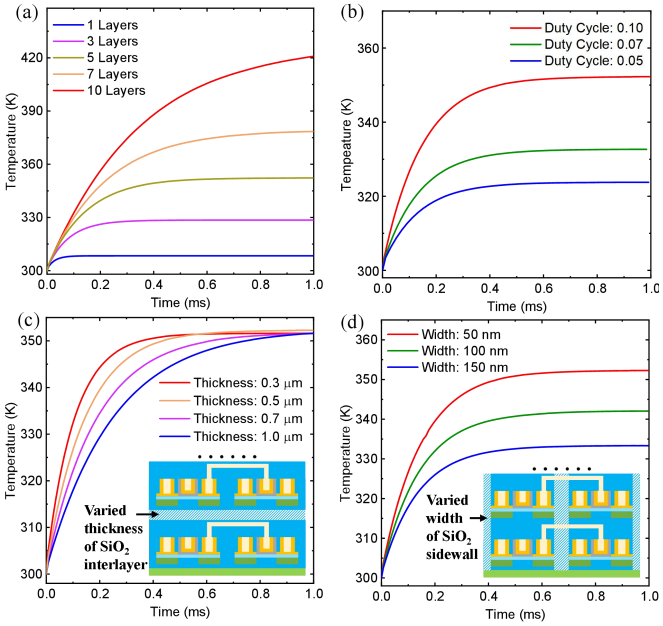

Fig. 3 The average temperature of stacked 2T0C memory array varying (a) number of layers, (b) duty cycle of periodic pulse voltage, (c) thickness of SiO_2 insulator interlayer, (d) width of SiO_2 sidewall. Default simulation parameters are 5 layers, 0.1 duty cycle, 0.5 μm thickness of SiO_2 interlayer and 50 nm width of SiO_2 sidewall, unless otherwise specified.

The temperature rising is deadly to the storage performance of the 2T0C memory cell. The temperature rising leads to the negative shift of threshold voltage and increasing off-state leakage of IGZO-TFTs, as shown in Fig. 4a. As temperature rises, the hydrogen emitted from the oxide layer introduces extra donor states and natively oxygen vacancy generates more electrons [12, 13], which are the main inducements of the temperature instability of IGZO-TFTs. Since the retention time of 2T0C memory cell is mainly determined by the off-current of the IGZO-TFTs, the temperature rising results in retention time decreasing greatly.

The storage performance in different temperature were simulated by applying voltage waveform on the 2T0C memory cell for retention time test as shown in Fig. 4b. The parasitic coupling disturbance between the WWL and SN results in the SN voltage decreasing with the WWL voltage lowering. The retention time is extracted at 0.1 V dropping of the written SN voltage, which is the SN voltage after the write operation finishes. As shown in Fig. 4c, the retention time of 2T0C cell degrades severely with temperature rising. As shown in Fig. 4d, with the temperature rising the written SN voltage rises a little due to on-state current increasing but fall greatly due to leakage current boosting of write transistor. More importantly, the increasing leakage leads to severe degradation of retention time which drops below 64 ms (typical fresh time) at 360 K and this critical temperature can be easily reached in seven-layer stacked 2T0C memory

array.

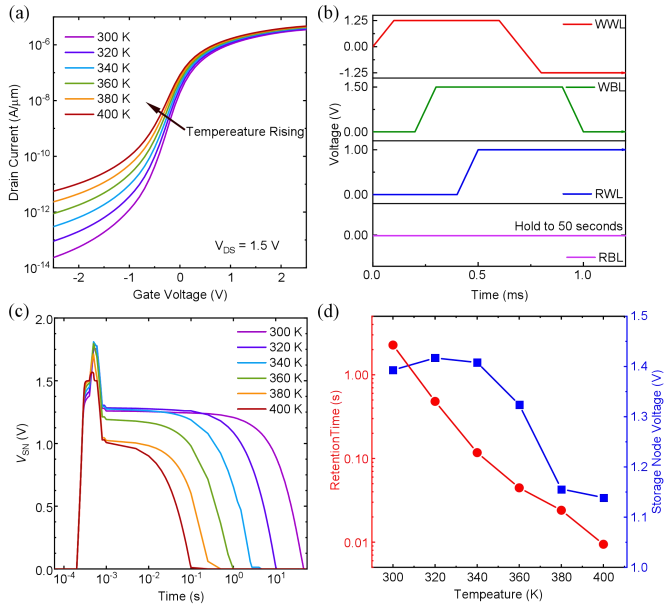

Fig. 4 (a) The transfer characteristics degradation of IGZO TFTs with temperature rising. (b) The waveform of retention time test. (c) The storage node voltage decreasing over time. (d) The temperature-dependent characteristics of retention time and storage node voltage.

IV. Conclusion

In conclusion, a coupling electro-thermal model of multi-layer stacked 2T0C memory array was built based on the material properties and electron transport mechanism of IGZO, as well as Fourier heat flow equation. The model indicates the temperature of the stacked 2T0C memory array boosts rapidly with the number of layers increasing. Since the temperature rising may results in severe storage performance degradation of 2T0C cell, the current-sensing reading approach needs further optimization to overcome the potential self-heating issue. The self-heating characteristics of different stacked 2T0C memory array structures we proposed may give advanced device-level guidance for the future design and fabrication by suggesting the potential thermal management scheme of 2T0C DRAM configuration. And there is on-going further research for thermal management and operation scheme of 3D stack structure which fits the actual application scene better.

V. Acknowledgements

This work was supported by NSFC under Grant U20A20207, Grant 61925110, Grant 61821091, the Opening Project of Key Laboratory of Microelectronics Devices and Integration Technology in IMECAS. This work was partially carried out at the Information Science Laboratory Center of University of Science and Technology of China (USTC).

References

[1] U. Schroeder, K. Cho and S. Slesazeck, “Materials for DRAM memory cell applications,” *Thin Films Silicon, Electron. Photonic Appl.*, pp. 369–401, 2017.

[2] A. Belmonte, H. Oh, N. Rassoul, G. L. Donadio, J. Mitard, H. Dekkers, R. Delhougne, S. Subhechha, A. Chasin, M. J. van Setten, L. Kljucar, M. Mao, H. Puliyalil, M. Pak, L. Teugels, D. Tsvetanova, K. Banerjee, L. Souriau, Z. Tokei, L. Goux and G. S. Kar, "Capacitor-less, long-retention (> 400 s) DRAM cell paving the way towards low-power and high-density monolithic 3D DRAM," *IEDM Tech. Dig.* , pp. 28.2.1-28.2.4, Dec 2020.

[3] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano and H. Hosono, "Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors," *Nature,* vol. 432, no. 7016, pp. 488-492, May 2004.

[4] T. Kamiya, K. Nomura and H. Hosono, "Present status of amorphous In-Ga-Zn-O thin-film transistors," *Sci. Technol. Adv. Mater.,* vol. 11, no. 4, Aug 2010.

[5] W. Lu, Z. Zhu, K. Chen, M. Liu, B.-M. Kang, X. Duan, J. Niu, F. Liao, W. Dan, X.-S. Wu, J. Son, D.-Y. Xiao, G.-L. Wang, A. Yoo, K.-Y. Cao, D. Geng, N. Lu, G. Yang, C. Zhao, L. Li and M. Liu, "First demonstration of dual-gate IGZO 2T0C DRAM with novel read operation, one bit line in single cell, I_{ON} = 1500 μA/μm@V_{DS} = 1V and retention time > 300 s," *IEDM Tech. Dig.* , pp. 26.4.1-26.4.4, Dec 2022.

[6] M. Li, T. Li and Y. Tang, "Improved thermal network modeling of die stacking DRAM and optimization," *Integration,* vol. 85, pp. 35-41, 2022.

[7] H. Oh, A. Belmonte, M. Perumkunnil, J. Mitard, N. Rassoul, G. L. Donadio, R. Delhougne, A. Furnemont, G. S. Kar and W. Dehaene, "Enhanced data integrity of In-Ga-Zn-Oxide based capacitor-less 2T memory for DRAM applications," *IEEE 51st European Solid-State Device Research Conference (ESSDERC)*, Sep 2021.

[8] S. Lee, A. Nathan, J. Robertson, K. Ghaffarzadeh, M. Pepper, S. Jeon, C.Kim, I. H. Song, U. I. Chung and K. Kinam, "Temperature dependent electron transport in amorphous oxide semiconductor thin film transistors," *IEDM Tech. Dig.* , pp. 14.6.1-14.6.4, Dec 2011.

[9] H.-H. Hsieh, T. Kamiya, K. Nomura and H. Hosono, "Modeling of amorphous $InGaZnO_4$ thin film transistors and their subgap density of states," *Appl. Phys. Lett.,* vol. 92, no. 13, Apr 2008.

[10] A. Belmonte, H. Oh, S. Subhechha, N. Rassoul, H. Hody, H. Dekkers, R. Delhougne, L. Ricotti, K. Banerjee, A. Chasin, M. J. van Setten, H. Puliyalil, M. Pak, L. Teugels, D. Tsvetanova, K. Vandersmissen, S. Kundu, J. Heijlen, D. Batuk, J. Geypen, L. Goux and G. S. Kar, "Tailoring IGZO-TFT architecture for capacitorless DRAM, demonstrating > 10^3 s retention, > 10^{11} cycles endurance and L_g scalability down to 14 nm," *IEDM Tech. Dig.*, Dec 2021.

[11] M. Rencz, "Thermal issues in stacked die packages," *Semiconductor thermal measurement and management IEEE twenty first annual IEEE symposium*, pp. 307-312, Mar 2005.

[12] S. Lee, A. Nathan, J. Robertson, K. Ghaffarzadeh, M. Pepper, S. Jeon, C.Kim, I. H. Song, U. I. Chung and K. Kinam, "Temperature dependent electron transport in amorphous oxide semiconductor thin film transistors," *IEDM Tech. Dig.* , pp. 14.6.1-14.6.4, Dec 2011.

[13] A. Chasin, J. Franco, K. Triantopoulos, H. Dekkers, N. Rassoul, A. Belmonte, Q. Smets, S. Subhechha, D. Claes, M. J. van Setten, J. Mitard, R. Delhougne, V. Afanas'ev, B. Kaczer and G. S. Kar, "Understanding and modelling the PBTI reliability of thin-film IGZO transistors," *IEDM Tech. Dig.* , pp. 31.1.1-31.1.4, Dec. 2021.