Extended Abstracts of the 2019 International Conference on Solid State Devices and Materials, Nagoya, 2019, pp397-398

Characterization of Crystalline Oxide Semiconductor FET/Si-FET Hybrid Structured Analog Multiplier for Artificial Neural Network

Yoshiyuki Kurokawa, Takeshi Aoki, Munehiro Kozuma, Hajime Kimura and Shunpei Yamazaki

Semiconductor Energy Laboratory Co., Ltd.
398, Hase, Atsugi, Kanagawa 243-0036, Japan
Phone: +81-46-248-1131, E-mail: kurokawa@sel.co.jp

Abstract

An analog multiplier with a hybrid structure where FETs whose channel layer is a crystalline oxide semiconductor are stacked on Si-FETs is fabricated and its characteristics are measured. The multiplier shows favorable multiplier characteristics. A demonstration in an artificial neural network (ANN) using our multiplier model with the acquired multiplier characteristics is performed. The inference accuracy of the ANN is comparable to that with ideal multiplication.

1. Introduction

Some studies to achieve multiply-accumulate (MAC) operation in an artificial neural network (ANN) by using analog operation have been proposed [1–3]. The use of analog operation is expected to result in fewer FETs in an arithmetic circuit and fewer data signal lines, for example.

An FET utilizing a crystalline oxide semiconductor [4], specifically a c-axis-aligned crystalline oxide semiconductor (CAAC-OS), for an active layer (OS-FET) exhibits an ultralow off-state current and can form memory with excellent charge retention. Thus, applications of OS-FETs to LSI [5–7] as well as an analog multiplier utilizing analog weight storage [8] have been proposed.

In this study, an analog multiplier employing an OS-FET/Si-FET structure [5–7] is fabricated so that its characteristics are measured. A handwritten digits recognition in an ANN using our multiplier model with the acquired multiplier characteristics is demonstrated.

2. OS-FET/Si-FET Analog Multiplier

Fig. 1 shows the circuit configuration and operating modes of the proposed analog multiplier consisting of an OS-FET, a Si-FET, and a capacitor. The OS-FET controls writing of charge to the capacitor. An ultralow off-state current of the OS-FET enables long-term charge retention. A read current is controlled by the signal VX.

Given that a read current is $I(w, x)$ when a voltage $W = W_0 + w$ (W_0: a bias voltage) is written (i.e., a weight w is written) to the retention node FN and a voltage $X = X_0 + x$ (X_0: a bias voltage) is applied (an input x is supplied) to VX. When the Si-FET operates in a saturation region, where $I(w,x) = (\beta/2)(W_0 + w + X_0 + x - V_{th})^2$ (V_{th}: threshold voltage), the following equation is satisfied [9]:

$$y = I(w,x) - I(0,x) - I(w,0) + I(0,0) = \beta w x. \quad (1)$$

That is, y calculated from the read current under four conditions where a weight 0 or w is written and an input 0 or x is supplied corresponds to the product of w and x.

Fig. 2 is an optical micrograph of the fabricated multiplier; 350-nm OS-FETs are fabricated to be stacked on 110-nm Si-FETs.

3. Measurement Results

Fig. 3(A) presents the measured characteristics of the fabricated multiplier. The read current $I(w, x)$ when the weight w is written and then the input x is swept is measured with various w. Here, both W_0 and X_0 are 1.5 V; the same applies to the following measurements. Fig. 3(B) shows the results of calculating the multiplier characteristics from the measured read current in Fig. 3(A) according to Eq. (1). The correlation coefficient between x and y with various w is 0.999 or higher, demonstrating good linearity.

Fig. 4 depicts the retention characteristics of the multiplier. The weight w has been written initially and then the read current $I(w, x)$ has been measured after predetermined times to calculate y (w, $x = -1.0$). The change of y at an elapsed time of 108,000 s is less than 4% with various w.

Fig. 5 shows the multiplier characteristics of our multipliers. The measurement is performed on twelve multipliers in the same manner as in Fig. 3 to calculate the multiplier characteristics. A variation in y between the multipliers is less than 5% with various w and x.

Table I compares our multiplier with other proposals [1–3]. In our multiplier, a variation between devices is small without special correction such as verification. This reveals stable data writing in the multiplier.

4. Artificial Neural Network Demonstration

Inference accuracy of an ANN using our multiplier is verified with a demonstration of handwritten digits recognition using MNIST datasets. First, the multiplier characteristics of our multiplier are approximated by the following model formula:

$$y_{Model}(w,x) = (w/w_M)^{\alpha_w}(x/x_M)^{\alpha_x}. \quad (2)$$

Here, w_M, x_M, α_w, and α_x are fitting parameters. Then, in order to incorporate the variation between devices into the model, $y_{Model}(w, x)$ is multiplied by $(1+\Delta)$, where Δ is a random number normally distributed with an average of 0 and a standard deviation of 0.05.

An ANN model is a three-layer fully connected network of an input layer, a hidden layer, and an output layer with 784, 100, and 10 neurons, respectively. The network is implemented using a Python program on a PC, and synaptic weights are obtained by learning. The inference accuracy of an ANN with an ideal multiplier in MAC operation is 97.89%. Meanwhile, that using our multiplier model instead of the ideal multiplier is 97.77%; i.e., the accuracy degradation is as small as 0.12%.

5. Conclusion

An OS-FET/Si-FET structured analog multiplier is fabricated, and the measurement results reveal favorable multiplier characteristics. ANN demonstration with our model with the acquired multiplier characteristics proves that our multiplier is a promising ANN multiplier.

References

[1] X. Guo et al., IEDM 2017, pp. 151–154.
[2] R. Mochida et al., VLSI Technology 2018, pp. 175–176.
[3] S. K. Gonugondla et al., ISSCC 2018, pp. 490–491.
[4] Y. Waseda et al., Mater. Trans. **59** (11), pp. 1691–1700 (2018).
[5] Y. Kurokawa et al., ASQED 2013, pp. 66–71.
[6] T. Aoki et al., ISSCC 2014, pp. 502–503.
[7] T. Ohmaru et al., ISSCC 2015, pp. 118–119.
[8] T. Aoki et al., SSDM 2017, pp. 191–192.
[9] A. Aslam-Siddiqi et al., JSSC **33** (10), pp. 1502–1509 (1998).

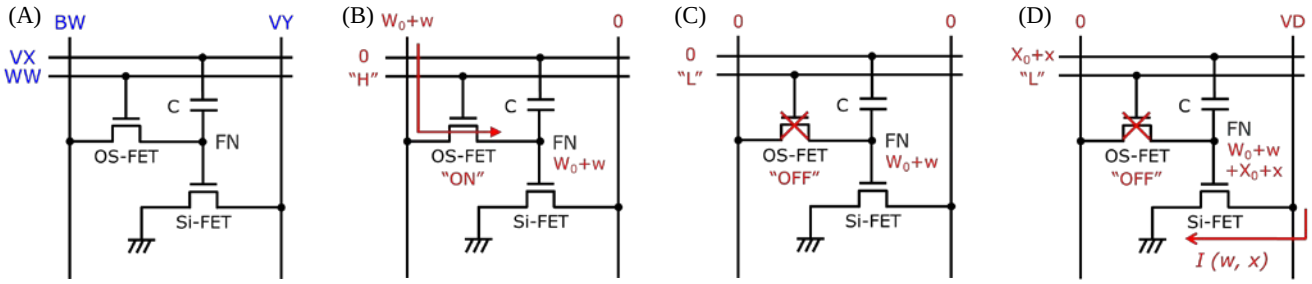

Fig. 1 (A) Circuit configuration of our multiplier. (B) In write mode, the OS-FET is turned on and the voltage $W = W_0 + w$ is written to the retention node FN. (C) In hold mode, the OS-FET is turned off. (D) In read mode, the voltage $X = X_0 + x$ is applied to VX and the output current flows.

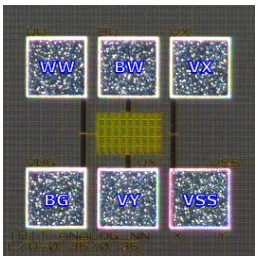

Fig. 2 Micrograph of the fabricated multiplier. One device can be selectively measured in an array of 9 × 16 multipliers.

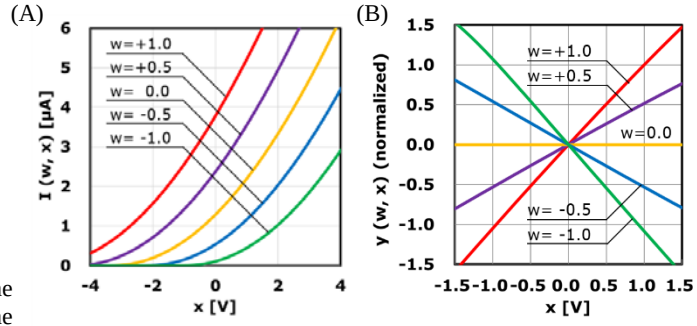

Fig. 3 (A) Measured x dependence of the read current $I(w, x)$ with various w in our multiplier. (B) Calculated multiplier characteristics, with y calculated from Eq. (1) using the measured read current.

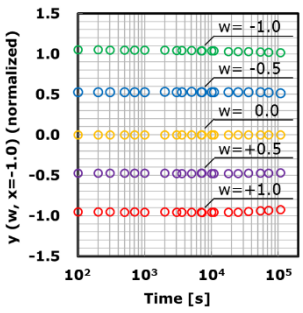

Fig. 4 Retention of our multiplier, showing a change in y over time in 108,000 s.

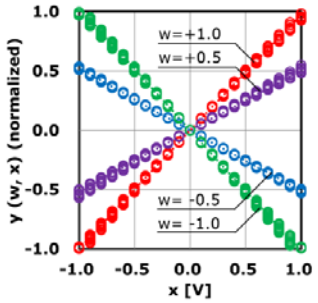

Fig. 5 Calculated multiplier characteristics of our multipliers. A variation in y between the multipliers is less than 5% with various w and x.

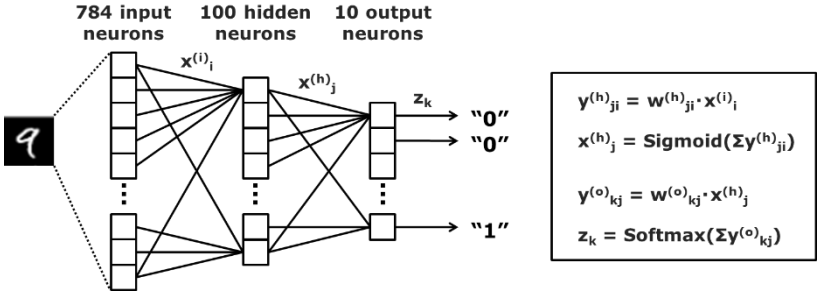

Fig. 6 ANN model (three-layer fully connected network) used for inference demonstration with MNIST datasets. The number of neurons is 784, 100, and 10 in the input layer, the hidden layer, and the output layer, respectively. Inference is performed with an ideal multiplier or our multiplier model replacing a multiplier in MAC operation.

Table I Comparison between this work and prior work

	This Work	IEDM 2017 [1]	VLSI 2018 [2]	ISSCC 2018 [3]
Technology	350-nm OS/110-nm Si	180 nm	180 nm	65 nm
Weight storage	CAAC-OS	eFlash	ReRAM	SRAM
Variation	< 5% (w/o verify)	~4% (w/ tuning)	> 10% (w/o verify) < 1.2% (w/ verify)	16% (w/o verify)
Nominal voltage	5 V (write) 3 V (read)	~11 V (write) 2.7 V (read)	1.8 V	1.0 V