



A FeFET-Based Hybrid Memory Accessible by Content and by Address

Cedric Marchand, Ian O'Connor, Mayeul Cantan, Evelyn Breyer, Stefan Slesazeck, Thomas Mikolajick

► To cite this version:

Cedric Marchand, Ian O'Connor, Mayeul Cantan, Evelyn Breyer, Stefan Slesazeck, et al.. A FeFET-Based Hybrid Memory Accessible by Content and by Address. IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, 2022, 8 (1), pp.19-26. 10.1109/JXCDC.2022.3168057 . hal-03727669

HAL Id: hal-03727669

<https://hal.science/hal-03727669>

Submitted on 3 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Received 10 February 2022; revised 5 April 2022; accepted 13 April 2022.
Date of publication 18 April 2022; date of current version 6 May 2022.

Digital Object Identifier 10.1109/JXDC.2022.3168057

A FeFET-Based Hybrid Memory Accessible by Content and by Address

CÉDRIC MARCHAND¹ (Associate Member, IEEE),
IAN O'CONNOR¹ (Senior Member, IEEE), MAYEUL CANTAN¹, EVELYN T. BREYER²,
STEFAN SLESAZECK², and THOMAS MIKOLAJICK^{2,3} (Senior Member, IEEE)

¹Univ Lyon, École Centrale de Lyon, CNRS, INSA Lyon, Université Claude Bernard Lyon 1, CPE Lyon, INL, UMR5270, 69130 Écully, France

²NaMLab gGmbH, 01187 Dresden, Germany

³Chair of Nanoelectronics, IHM, Dresden University of Technology (TU Dresden), 01069 Dresden, Germany

CORRESPONDING AUTHOR: C. MARCHAND (cedric.marchand@ec-lyon.fr)

This work was supported in part by the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement 780302 and in part by the French "Agence Nationale de la Recherche" under Project ANR-20-CE39-0006 (SECRET project). This publication reflects only the authors' view, and the commission is not responsible for any use that may be made of the information it contains.

ABSTRACT Emerging nonvolatile memory technologies are attracting interest from the system design level to implement alternatives to conventional von-Neumann computing architectures. In particular, the hafnium oxide-based ferroelectric (FE) memory technology is fully CMOS-compatible and has already been used for logic-in-memory architectures or compact ternary content addressable memory (TCAM) cells. These enable the tight combination of different functionalities in the same circuit to reduce implementation area and energy consumption. In this article, we propose a new hybrid memory circuit that combines TCAM and normal memory capability: the Ternary Content addressable and MEMory (TC-MEM). A 1-bit TC-MEM circuit is proposed and discussed in detail, both as a concept and through its implementation in a 28-nm ferroelectric field-effect transistor (FeFET) technology. Measurement results demonstrate the circuit functionality. We also discuss how to scale it to multibit circuits, as well as its use both as a TCAM and as a normal memory allowing the implementation of reversible functions using one memory table instead of two memory tables, and in-memory-computing concepts.

INDEX TERMS Ferroelectric field-effect transistor (FeFET), in-memory-computing (IMC), logic-in-memory, memory, ternary content addressable memory (TCAM).

I. INTRODUCTION

IN-MEMORY-COMPUTING (IMC) [1] is a viable approach to overcome limitations in current computing architectures due to the von-Neumann bottleneck. It allows computation to be performed directly inside the memory and leads to drastic reductions in the number of data transfers from the memory to the computing core. This improves the energy efficiency of computing systems [2] since data transfer is the most energy-consuming part of current computing architectures. IMC targets a wide range of applications such as hardware security, artificial intelligence (AI), image filtering, and many others [3], [4].

Recently, research interest has increased in ternary content-addressable memories (TCAMs) [5] due to opportunities afforded by emerging nonvolatile technologies. Novel compact circuit topologies for TCAM have been demonstrated using resistive switching memory (RRAM) [6], [7],

magnetic memory (MRAM) [8], and CMOS-compatible $Hf_xZr_{1-x}O_2$ -based ferroelectric (FE) technologies [ferroelectric random access memory (FeRAM) and ferroelectric field-effect transistor (FeFET)] [9], [10]. Coupling emerging nonvolatile technologies to TCAM functionality enables such circuits to be used in new applications, such as IMC [11], AI [12], [13], and edge computing [14]. Based on these new circuits and more specifically on the TCAM design using FeFET, few-shot learning and AI applications have been proposed [15].

However, TCAM circuits based on emerging technologies can only be used as TCAM. For some applications, it can be useful to have a memory that is reversible. In [16], such a combination is implemented in a standard 28-nm CMOS technology. Using FeFETs, such combinations have been demonstrated in [17] and [18]. However, these papers present a modification of the memory array that enables

both memory and TCAM operations. The presented results are based on HSPICE simulations using a predictive 45-nm technology. In addition, they sacrifice half of the memory in order to create dual functionalities. However, the combination of memory and TCAM inside the same circuit architecture is very interesting from the application point of view. In cryptography, for instance, some functions are used in one direction for encryption and in the other for decryption. In standard computing systems, the design of such functions requires two address spaces in the memory. As a consequence, combining TCAM and normal memory functions in the same circuit saves both implementation area and energy consumption. Such a circuit will not only be beneficial for cryptographic implementations but for all reversible operations needed in computing systems, especially in the context of energy-constrained edge computing for the Internet of Things (IoT). For example, applications such as image filtering, data compression, and edge AI (with the possibility of rapidly estimating Hamming distances [15]) will also benefit from the combination of both of these memory operations.

In this article, we introduce a new nonvolatile memory element based on FeFETs that combine TCAM and standard memory functionalities: the Ternary Content addressable and MEMory (TC-MEM). First, general information on operation with FeFET technology is presented in Section II. The design of a 1-bit TC-MEM cell is proposed and discussed in detail in Section III-A, and we show how to scale it to multibit cells in Section III-B. The results are obtained from a test chip manufactured in GLOBALFOUNDRIES 28-nm SLP technology. The results presented along with the design aspects show that the TC-MEM is able to combine both functionalities: TCAM and normal memory operations. For the TCAM operation, in addition to the full word search, the TC-MEM circuit also provides a partial word search. The TC-MEM offers the possibility to implement reversible functions, which can be precomputed directly inside the memory but also IMC functionalities. These different applications are discussed in Section IV.

II. FeFET TECHNOLOGY OVERVIEW

A. FeFET BASICS

The FeFET is a device where an FE capacitor is connected to the gate of a transistor. This can be done directly in the gate-stack of the transistor to make a single integrated device [front end of line (FEoL)] or with a metal connection [19] from a separate [back end of line (BEoL)] ferroelectric capacitor. In this article, we used fully integrated FEoL FeFETs (although the concepts can also be used in a BEoL approach), i.e., an FE oxide layer is included in the gate-stack of the transistor, as presented in Fig. 1(a). Fig. 1(b) shows the symbol used for the FeFET component in circuit designs presented in this article. Example FeFET layouts used in the manufactured test chip are presented in Section III, where FeFETs are marked by a green square layer surrounding transistors.

Considering the circuit manufactured with 28-nm technology from GLOBALFOUNDRIES, we used the following

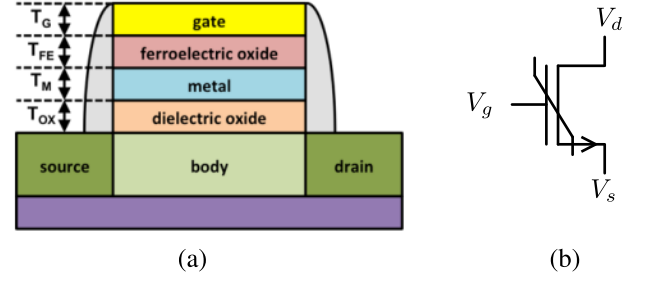


FIGURE 1. (a) Structure of the gate-stack for an FEoL FeFET. (b) Symbol used to represent the FeFET in circuit schematics.

parameters for the sizes of the different transistors.

- 1) *n*-type FeFET $\rightarrow W = 500$ nm and $L = 500$ nm.
- 2) nMOS $\rightarrow W = 80$ nm and $L = 36$ nm.
- 3) pMOS $\rightarrow W = 130$ nm and $L = 50$ nm.

We choose to use such large FeFETs to ensure proper functionality (in particular, reliable program/erase operations) [20] and provide a proof of concept of the TC-MEM circuit of 1 and 2 bit. The type of FeFET used in this article has already been extensively studied in the literature [20]–[22]. A summary of its operation is given in the following paragraph.

Depending on the polarization of the FE capacitor, the threshold voltage of the transistor varies. In the down polarization state (of the FE oxide layer), the FeFET will act almost as a normal transistor. We refer to this behavior as the low threshold voltage (LVT) state for the FeFET. In the up polarization state, the FE capacitor will prevent the transistor from switching from the OFF-state to the ON-state. We refer to this behavior as the high threshold voltage (HVT) state for the FeFET. To change the polarization of the FE capacitor, a negative or positive voltage pulse high and long enough to exceed the coercive electrical field of the FE layer has to be applied at the gate of the FeFET.

B. FeFET WRITING SCHEME

In this article, we propose to create a FeFET-based circuit that can be used as TCAM or as normal memory. To easily read both values (0 and 1), it is necessary to achieve a clear separation between the LVT and HVT states of the FeFET. This separation is achieved by carefully choosing the writing scheme of the FeFETs. To do so, it is crucial to characterize how the threshold voltage of the FeFET varies with writing pulses. We used single pulses with voltages from 2.25 to 5 V and a duration from 0.1 to 15 μ s.

The experimental results of this characterization are presented in Fig. 2, where the threshold voltage of a FeFET is recorded for varying amplitudes and time duration of the write pulse. It is worth noticing that an LVT (close to 0 V) can be achieved with a write pulse voltage above 4 V and a duration of 0.1 μ s. However, it is safer to use longer writing pulses to ensure that the polarization of the FeFET has been correctly switched, as demonstrated in [21]. Consequently, we choose to use a write amplitude of ± 4 V and a time

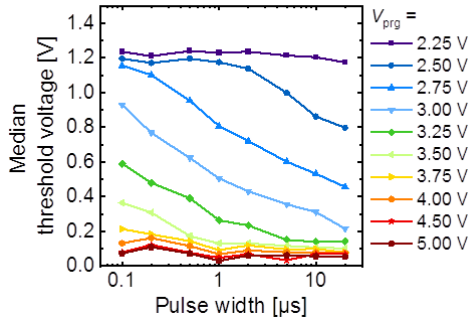


FIGURE 2. Variation of the threshold voltage of a FeFET transistor according to the amplitude and the time duration of the write pulse on the gate of the FeFET.

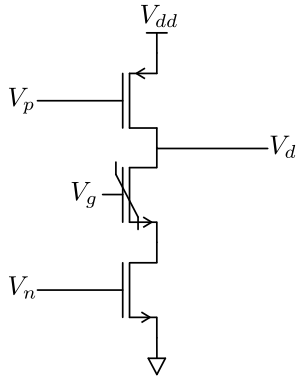


FIGURE 3. Single FeFET circuit used to plot the voltage characteristic shown in Fig. 4.

duration of 10 μ s. This implies that only the two strong polarization states (HVT and LVT) of the FeFETs are used, avoiding the analog behavior used in [23]. Consequently, even if a large FeFET is used to prove the concept of the TC-MEM, it will be possible to scale down the cell dimension without observing variability issues. Finally, we choose the symmetric program/erase voltages to simplify programming circuits even if Fig. 2 only shows the program case, and the erase operation might behave slightly different.

With the electrical characteristics of the write pulses chosen, it is now possible to define the writing scheme of the FeFET.

- 1) A positive 4-V pulse during 10 μ s will switch the FeFET into its HVT state. This is equivalent to storing a 0 inside the FeFET.
- 2) A negative 4-V pulse during 10 μ s will switch the FeFET into its LVT state. This is equivalent to storing a 1 inside the FeFET.

With this writing scheme, the circuit presented in Fig. 3 is used to plot the output (drain) voltage characteristic of the FeFET, as shown in Fig. 4. To generate this output voltage characteristic for each polarization of the FeFET, we use the following steps.

- 1) The FeFET is programmed or erased using our writing scheme.

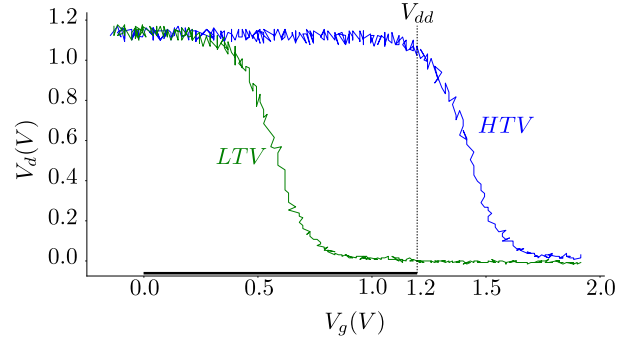


FIGURE 4. Voltage characteristic of the FeFET using ± 4 -V writing pulses during 10 μ s.

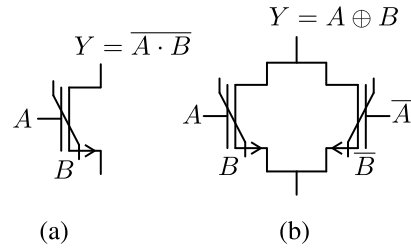


FIGURE 5. Basic 1-bit circuits implementing (a) memory cell and (b) TCAM.

- 2) The output V_d is precharged to V_{dd} (i.e., the pMOS transistor is passing, and the nMOS is not passing).
- 3) V_p and V_n are set to V_{dd} .
- 4) A voltage ramp from 0 to 2 V is applied; V_g and V_d are measured.

It is possible to see that, in the HVT state, the output voltage of the FeFET remains high (i.e., at logical level 1) for the entire gate voltage range (from 0.0 to 1.2 V). Conversely, when the FeFET is in its LVT state, the output shows that the FeFET acts as a conventional nMOS transistor, turning on when $V_g \approx 0.4$ V.

C. SINGLE-DEVICE PROGRAM/ERASE ENERGY CONSUMPTION

Considering the design of the test chip, one row of the chip contains different test structures, and it is difficult to extract the exact energy consumption of a single FeFET transistor. Furthermore, only N -type FeFETs are included in the circuit. This led to a particular type of circuit design to limit the energy consumption. Indeed, all test structures of the manufactured test chip use a dynamic logic design style, such as that presented in Fig. 3. Due to this design style, read energy depends mainly on the output load capacitance such that the measurement of the output current is not representative of a real design.

Concerning program/erase power consumption, four FeFETs are evaluated at the same time without the other test structures when characterizing the test chip. Since we focus on functionality and proof of concept rather than performance

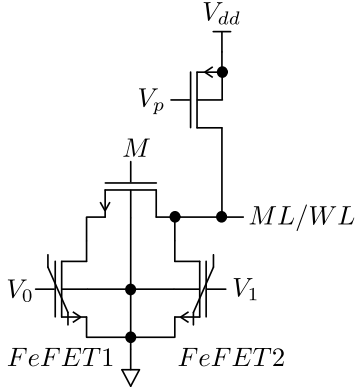


FIGURE 6. Schematic of the 1-bit TC-MEM cell.

in this article, we used large FeFETs such that the energy consumption is significantly higher than it would be for optimized circuits, such as presented in [17] and [18]. Indeed, large capacitors are necessary to guarantee reliable switching, but the capacitor size is directly linked to the FET size in the FEOL approach: this is one shortcoming of the FEOL FeFET. However, extracting an estimation of the power consumption on the test chip is interesting to make a projection and comparison with other technologies for future work.

To extract the energy consumption of one FeFET during write operations, the current is measured at the gate of the FeFET in order to characterize the energy used to change the polarization of the FeFET. The energy computed is then divided by 4 to recover the energy consumption of a single FeFET. According to these measurements, the write energy consumption of our FeFET transistor is given as follows:

- 1) $6.64e^{-10}$ J to write a 0 inside the FeFET;
- 2) $5.82e^{-10}$ J to write a 1 inside the FeFET.

III. DESIGN OF THE TC-MEM MEMORY

In this section, the design of the 1-bit TC-MEM cell is presented and scaled to 2 bit and more. The target function for the TC-MEM is to be accessible both for content search memory and normal (address search) memory operations.

A. 1-BIT TC-MEM CIRCUIT STRUCTURE

To design the 1-bit TC-MEM cell, two modes of operations are required. The first mode of operation for the TC-MEM cell corresponds to the normal memory mode. This is achieved with a single FeFET transistor, as presented in Fig. 5(a). The FeFET realizes a NAND operation between the value stored in its FE layer and the evaluating input to the gate, as presented in [21].

The second mode of operation corresponds to the TCAM mode, for which the structure is presented in Fig. 5(b). Two FeFETs connected in parallel can be programmed to realize an XOR operation [22]. The XOR operation is fundamental to TCAM operations, as presented in [10]. As shown in Fig. 5(b), the output “Y” is high if and only if both FeFETs are not conducting. Considering that one FeFET realizes a

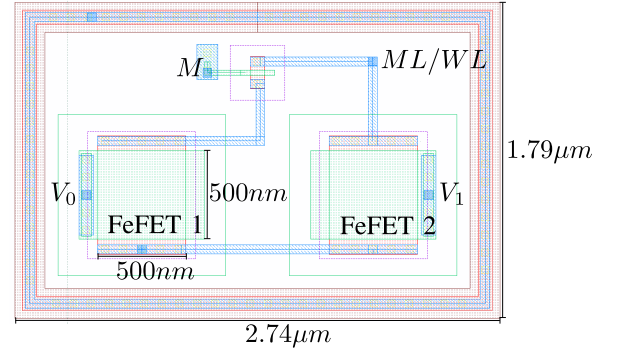


FIGURE 7. Layout of the 1-bit TC-MEM cell without the shared precharge transistor.

NAND operation, we can summarize the logical behavior of this circuit with the following equation:

$$\begin{aligned} Y &= (\overline{A \cdot B}) \cdot (\overline{\overline{A} \cdot \overline{B}}) \\ &= (\overline{A} + \overline{B}) \cdot (A + B) \\ &= A \oplus B. \end{aligned} \quad (1)$$

To create the 1-bit TC-MEM cell, both functionalities must be accessible within the same circuit. This is made possible by adding an nMOS transistor between the two FeFETs connected in parallel, as presented in Fig. 6. The layout corresponding to this circuit is shown in Fig. 7. In the circuit presented in Fig. 6, if the mode input (M) is low ($M = 0$), the nMOS transistor is in the OFF-state. This implies that only FeFET2 can change the output ML/WL (match/word line in TCAM/memory modes, respectively). This corresponds to the memory mode.

In the opposite state, if M is high ($M = 1$), the nMOS is in the ON-state. FeFET1 and FeFET2 are connected in parallel: this corresponds to the TCAM mode.

The V_p input is used to precharge the ML/WL line to V_{dd} before the evaluation phase (conditional discharge of the ML/WL line via the network of FeFETs). Due to pad limitations on the manufactured test chip, the precharge transistor is shared with multiple TC-MEM cells in the same pad row. Fig. 7 presents the layout of the 1-bit TC-MEM without the shared precharge transistor.

In Fig. 8, the result of the measurement of the 1-bit TC-MEM circuit is presented. The write operation is not presented, and only read operations in both TCAM and memory modes are shown. The read time used for these measurements is $10 \mu s$ for each evaluation. In the first part of Fig. 8, the TC-MEM circuit has a logical 1 stored inside, i.e., FeFET1 is in the LVT state and FeFET2 is in the HVT state. In this configuration, the output (ML/WL) in the TCAM mode is discharged only if V_0 is high (whatever the value of V_1) during evaluation. In the memory mode, the output remains high independently of the inputs, meaning that a 1 is read.

In the second part of Fig. 8, the value stored inside the TC-MEM cell is 0, i.e., FeFET1 is in the HVT state and

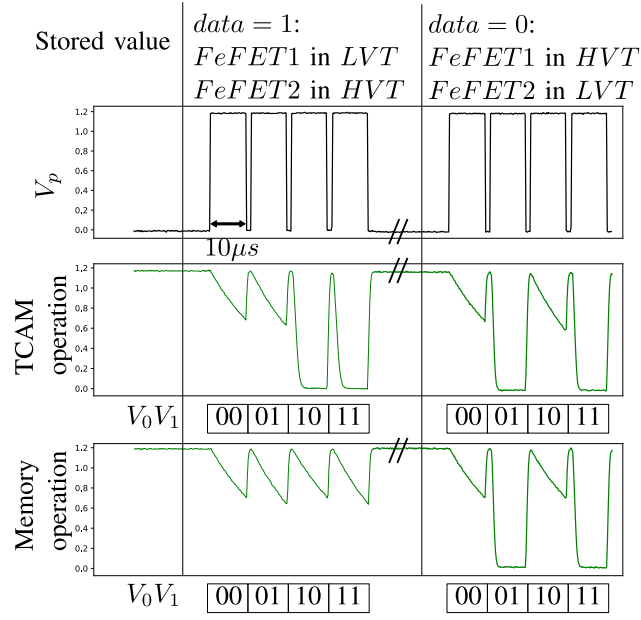


FIGURE 8. Output measurement of the 1-bit TC-MEM circuit in TCAM and memory models. For reading operation (TCAM and memory modes), V_0 and V_1 follow: $V_x = 0 \rightarrow V_x = 0$ V and $V_x = 1 \rightarrow V_x = 1.2$ V, with $x \in \{0; 1\}$.

FeFET2 is in LVT state. In the TCAM mode, ML/WL is discharged only if V_1 is high (whatever the value of V_0 during evaluation). In the memory mode, the output is discharged when $V_1 = 1$, meaning that a 0 is read.

From Fig. 8, it can be seen that, in the TCAM mode, if $V_0 = V_1$, the ML/WL output is the same independently of the value (0 or 1) stored inside the TC-MEM circuit. This means that the input combinations $V_0 = V_1 = 0$ and $V_0 = V_1 = 1$ correspond to the “don’t care” state of the TCAM. If $V_0 = V_1 = 0$, the output remains high for both values stored inside the cell—this “don’t care” state will be referred to as the “blocking” don’t care state. If $V_0 = V_1 = 1$, the output is discharged independently of the value stored inside the cell—this “don’t care” state will be referred to as the “passing” don’t care state.

In the memory mode, the value stored inside the TC-MEM is read if $V_1 = 1$. Table 1 summarizes this behavior and these features.

Section III-B explains how to scale the TC-MEM circuit to multiple bits cells.

B. SCALING THE TC-MEM TO 2 BIT AND MORE

To combine several 1-bit TC-MEM cells, it is important to keep in mind that we are targeting dual functionality. In [13], the multibit TCAM circuit is achieved by connecting different bits in parallel to the same match line. This is not possible for the TC-MEM circuit because it will cause issues in the memory mode. Indeed, even if FeFETs connected in parallel are disconnected in the memory mode, the multibit TC-MEM cell in parallel will create forbidden states for some inputs. This will imply a more complex control logic and a different

TABLE 1. Summary of the 1-bit TC-MEM cell behavior.

value	Mode	V_0	V_1	ML/WL	Result
0	TCAM	0	0	1	Blocking don't care
		0	1	0	Match
		1	0	1	No match
		1	1	0	Passing don't care
	RAM	0	0	1	-
		0	1	0	Read
1	TCAM	1	0	1	-
		1	1	0	Read
	RAM	0	0	1	-
		0	1	1	Read
		1	0	1	-
		1	1	1	Read

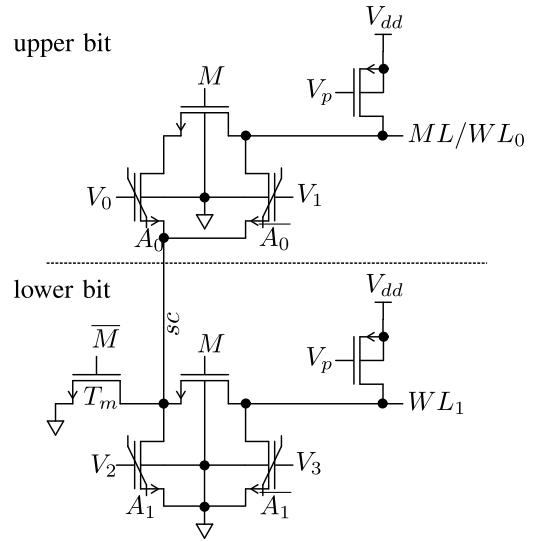


FIGURE 9. Schematic of a 2-bit TC-MEM circuit.

evaluation protocol for TCAM and memory modes. The other possibility is to connect multiple TC-MEM cells in series, as presented in Fig. 9.

The two 1-bit TC-MEM circuits are serially connected by the sc line. An additional nMOS transistor (T_m) is added to the circuit to avoid the creation of a floating node between lower and upper bitcells in the memory mode. In addition, T_m establishes the necessary ground connection to discharge the upper word line (WL_0) in the memory mode (i.e., it is possible to read the upper bit A_0). The T_m transistor is a regular nMOS transistor of the considered 28-nm SLP technology from GLOBALFOUNDRIES. Its dimensions are $W = 80$ nm and $L = 36$ nm. Fig. 10 presents the layout corresponding to the circuit presented in Fig. 9 without the precharge transistors, which are shared with other structures of the same pad row in the test chip.

The operation of the circuit is given as follows: in the TCAM mode ($M = 1$), the serial connection between the two bitcell of the TC-MEM does not lead to discharge of the

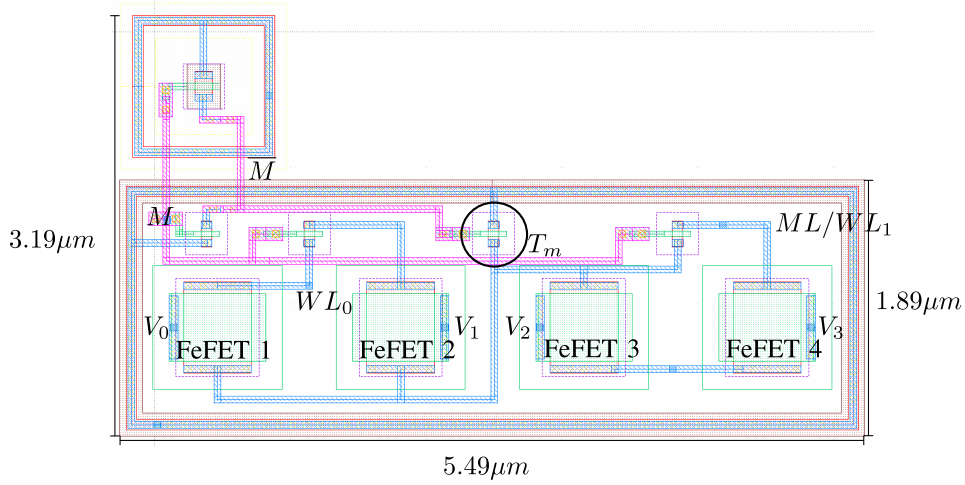


FIGURE 10. Layout of a 2-bit TC-MEM circuit with the M input inverter but without precharge transistors.

match line (ML) if one of the two bitcells does not match. Considering (1), this can be summarized by the following equation:

$$ML = (\overline{V_0 \cdot A_0} \cdot \overline{V_1 \cdot A_0}) + (\overline{V_2 \cdot A_1} \cdot \overline{V_3 \cdot A_1}). \quad (2)$$

Equation (2) implies that the match line will be low only if the correct value is presented for the entire 2-bit word.

In the memory mode ($M = 0$), the upper bitcell is disconnected from the lower bitcell due to the T_m transistors, which connects the source of both FeFET in the upper bitcell to the ground. In this way, it is possible to properly read both bits of the 2-bit TC-MEM using word lines WL_0 and WL_1 .

In addition, the 2-bit TC-MEM offers the possibility to perform a partial word search (in TCAM mode) by using the blocking and passing don't care states. For example, if only A_0 is considered during TCAM search, it is possible to force the lower bit (A_1) to its "passing" don't care state in order to search only for A_0 . In the same way, if only A_1 is considered, it is possible to force the upper bit (A_0) to any don't care state ("blocking" or "passing"). In the case of partial word search, all output lines serve as partial match lines. However, if the lower bit is in the "blocking" state, all the rest of the TC-MEM cells will be in a "blocking" don't care state.

Fig. 11 presents the output voltages ML/WL_0 and WL_1 measured experimentally on the test chip with stored values $A_0 = 0$ and $A_1 = 1$. Table 2 summarizes these results with the corresponding TCAM state (i.e., "blocking" or "passing" don't care state, partial search, and full search).

Table 3 presents the results in the memory mode. In Table 3, it is important to highlight that, because $A_0 = 1$, the fact that the lower bitcell output WL_0 is never discharged in the memory mode is a normal behavior.

C. COMPARISON OF THE TC-MEM WITH OTHER FeFET-BASED TCAM AND COMBINED CIRCUITS

Table 4 can be used to compare the TC-MEM with other FeFET-based circuits from [9], [17], and [18]. In this table, we extract the number directly from [9], [17], and [18]. This

TABLE 2. Summary of the 2-bit TC-MEM cell behavior in the TCAM mode with $A_0 = 0$ and $A_1 = 1$ from Fig. 11.

V_0	V_1	V_2	V_3	WL_1	ML/WL_0	Result
0	0	0	0	1	1	Blocking
0	0	0	1	1	1	Partial search A_1
0	0	1	0	0	1	Partial search A_1
0	0	1	1	0	1	Don't care
0	1	0	0	1	1	Blocking
0	1	0	1	1	1	Full search
0	1	1	0	0	0	Full search, Match
0	1	1	1	0	0	Partial search A_0
1	0	0	0	1	1	Blocking
1	0	0	1	1	1	Full search
1	0	1	0	0	1	Full search
1	0	1	1	0	1	Partial search A_0
1	1	0	0	1	1	Blocking
1	1	0	1	1	1	Partial search A_1
1	1	1	0	0	0	Partial search A_1
1	1	1	1	0	0	Passing

TABLE 3. Summary of the 2-bit TC-MEM cell behavior in the memory mode ($M = 0$) with $A_0 A_1 = 10$ stored inside.

V_1	V_3	WL_0	WL_1	Result
0	0	1	1	-
0	1	1	0	read A_1
1	0	1	1	read A_0
1	1	1	0	read A_0 and A_1

comparison is quite difficult because we are only proposing a proof of concept using large FeFET with test chip measurement, while the previous works propose simulation-based results on a 45-nm predictive technology. In consequence, the write energy is significantly higher using our measurements. However, assuming an ideal case where FeFETs are implemented with the minimum size allowed by the technology as done in previous works and the same write time, a linear projection on the write energy predicts write energy of 11.5 fJ per bit compared to 156 fJ per bit for [17] and 125 fJ per bit for [18]. In [9], the write energy is about 1.56 fJ per bit for TCAM only.

From the area point of view, the TC-MEM combined bitcell is a 1T-2FeFET circuit compared to 4T-2FeFET for [17] and 2T-2FeFET for [9].

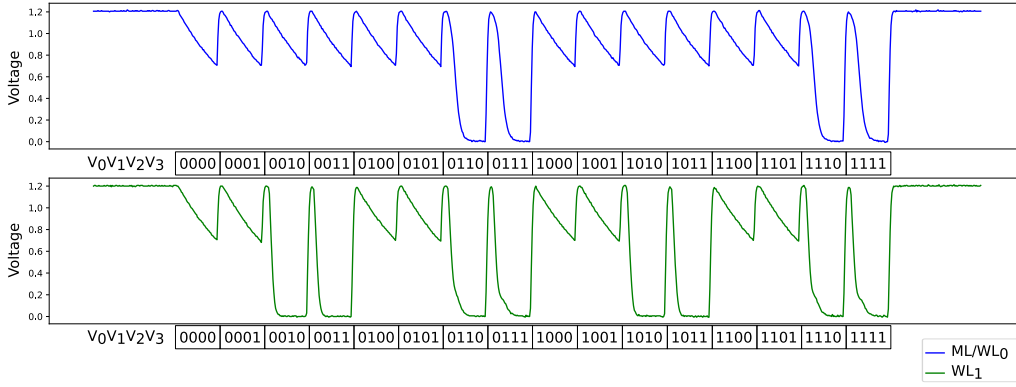


FIGURE 11. Measurement of WL_1 and ML/WL_0 of the 2-bit TC-MEM circuit with $A_0 = 0$ and $A_1 = 1$ in the TCAM mode.

TABLE 4. Comparison of different circuits from [9], [17], and [18] and this work from measurement and projection.

Work	[9]	[17]	[18]	This work	This work
Technology	45nm PTM	45nm PTM	45nm PTM	28nm	28nm
Results	simulation	simulation	simulation	measurement	projection
FeFET Area (W/L)	65nm/45nm	-	65nm/45nm	500nm/500nm	80nm/30nm
Bit cell Area	2FeFET (TCAM only)	4T-2FeFET	2T-2FeFET	1T-2FeFET	1T-2FeFET
Write time	10ns	-	10ns	10μs	10ns
Write energy (1-bit)	1.56fJ	156fJ	125fJ	1.2nJ	11.5fJ

The TC-MEM circuit can also be generalized to an n -bitcell using serial connections. The number of transistors ($\#T$) needed to realize an n -bit TC-MEM cell is

$$\begin{aligned} \#T &= n(2 \text{ FeFET} + 1 \text{ nMOS}) + (n-1)(1 \text{ nMOS}) \\ &= 2n \text{ FeFET} + (2n-1) \text{ nMOS}. \end{aligned} \quad (3)$$

Equation (3) does not take the precharge transistors into account because their number will depend on the implementation of the full TC-MEM array. If only one output line is set for each bit of a word, n precharge transistors will be needed. Considering a full TC-MEM array with n -bit word size, the read time in TCAM mode will strongly depend on the size n of the words because of the serial connections. This will limit the size of the possible word to match with speed constraints for complete circuit design.

IV. APPLICATIONS OPPORTUNITIES

In Section III, we detailed the design elementary bitcell of the TC-MEM memory and show how it is possible to scale it to a multibit circuit. In this section, we discuss the application opportunities brought by this circuit. The TC-MEM circuit opens new functionalities due to its dual behavior (TCAM and memory), for example, standard memory, TCAM, partial word search, reversible computing, and in-memory computing.

Standard memory and TCAM functionalities have already been presented in Section III. This section presents the additional functionalities.

A. REVERSIBLE COMPUTING

Reversible computing refers to applications where some functions and their inverse need to be executed. Let us consider a reversible function $f : x \rightarrow f(x)$. Storing the

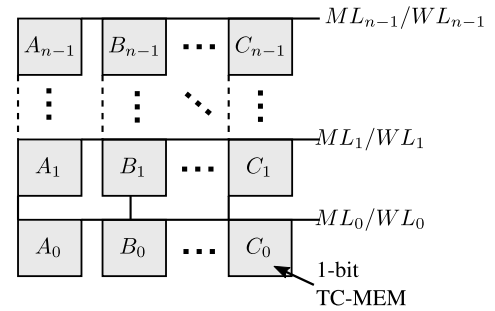


FIGURE 12. n -bit TC-MEM array simplified schematic.

results $f(x)$ inside the TC-MEM array allows us to compute f and f^{-1} .

- 1) In the memory mode, calling the address x gives $f(x)$.
- 2) In the TCAM mode, searching for the word $y = f(x)$ gives the inverse function: $f^{-1} : y \rightarrow f^{-1}(y) = x$.

This feature is particularly interesting in cryptography for example. Indeed, numerous cryptographic algorithms use substitution functions, and classic processing architectures require two address spaces in memory to implement the encryption and decryption functions. Using the TC-MEM, it is possible to implement encryption and decryption in a single address space. In addition, the symmetric structure of the TC-MEM (the value and its complement are stored) may lead to symmetric energy consumption to read “0” or “1” in both TCAM and memory modes, which is interesting against classical side-channel attacks, but further evaluation and such attacks must be tested.

B. PARTIAL WORD SEARCH

As explained in Section III, it is possible to search for only parts of words stored inside the TC-MEM. Let us consider the word $A = A_{n-1}, \dots, A_1 A_0$ of the TC-MEM array presented

in Fig. 12 and two integers x and y such that $0 < x < y < n - 1$. Three cases are possible.

- 1) The part $A_{n-1}A_{n-2}, \dots, A_y$ of A is searched: to do this, the remaining bits $A_{y-1}A_{y-2}, \dots, A_0$ of A have to be set to the “passing” don’t care state.
- 2) The part A_xA_{x-1}, \dots, A_0 of A is searched: to do this, the remaining bits $A_{n-1}A_{n-2}, \dots, A_{x+1}$ of A can be set to the “passing” or “blocking” don’t care state.
- 3) The part A_yA_{y-1}, \dots, A_x of A is searched: to do this, the least significant bits $A_{x-1}A_{x-2}, \dots, A_0$ of A have to be set to the “passing” don’t care state, while the most significant bits $A_{n-1}A_{n-2}, \dots, A_{y+1}$ of A can be set to the “passing” or “blocking” don’t care state.

In Fig. 11 and Table 2, this feature is proven to be functional with the partial research of A_0 or A_1 . However, the architecture proposed in this article cannot be used to search multiple parts of the same word simultaneously.

C. IN-MEMORY COMPUTING

It is possible to directly use the TC-MEM circuit in the memory mode to achieve and operations. Indeed, calling two words at a time (A and B) from Fig. 12, we get the following equation:

$$\forall i \in [0; n - 1], \quad WL_i = A_i \cdot B_i. \quad (4)$$

In the TCAM mode, if a value $X = X_{n-1} \dots X_1X_0$ is presented to the address containing the word A , the generalization of (2) gives the following equation:

$$\forall i \in [0; n - 1], \quad ML_i = \sum_{k=0}^i (A_k \oplus X_k). \quad (5)$$

Other logical functions, such as binary XOR, require the architecture of the TC-MEM array to be modified.

V. CONCLUSION AND FUTURE WORKS

In this article, we presented a new and hybrid memory circuit that combines TCAM and normal memory functionalities: the TC-MEM. The 1-bit TC-MEM circuit design was discussed in detail, and its scaling to multibit TC-MEM circuits was demonstrated. The results presented in this article have been measured on circuits fabricated using 28-nm FeFET technology and proved the dual functionalities of the TC-MEM. The utilization of other functionalities, such as partial search, reversible computing, and IMC, was also discussed.

The future investigation includes the use of the circuit to implement IMC operations to target AI or cryptographic computing. A full TC-MEM array with different word sizes will be also implemented to investigate read time operation with serial connection. Investigation on search delay and energy will be done on a full TC-MEM array for application-oriented evaluation. Finally, new generations of the TC-MEM will be explored in order to enhance the functionalities of this circuit.

REFERENCES

- [1] M. Mirzadeh, Y. O. Koçberber, B. Falsafi, and B. Grot. (2015). *Sort vs. Hash Join Revisited for Near-Memory Execution*. [Online]. Available: <http://infoscience.epfl.ch/record/209111>
- [2] Y. Wang *et al.*, “An energy-efficient nonvolatile in-memory computing architecture for extreme learning machine by domain-wall nanowire devices,” *IEEE Trans. Nanotechnol.*, vol. 14, no. 6, pp. 998–1012, Nov. 2015.
- [3] L. Ni, Y. Wang, H. Yu, W. Yang, C. Weng, and J. Zhao, “An energy-efficient matrix multiplication accelerator by distributed in-memory computing on binary RRAM crossbar,” in *Proc. 21st Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2016, pp. 280–285.
- [4] Y. Zhang *et al.*, “Recryptor: A reconfigurable in-memory cryptographic cortex-M0 processor for IoT,” in *Proc. Symp. VLSI Circuits*, Jun. 2017, pp. C264–C265.
- [5] Z. Ullah, M. K. Jaiswal, and R. C. C. Cheung, “Z-TCAM: An SRAM-based architecture for TCAM,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 2, pp. 402–406, Feb. 2015.
- [6] L. Zheng, S. Shin, S. Lloyd, M. Gokhale, K. Kim, and S.-M. Kang, “RRAM-based TCAMs for pattern search,” in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 1382–1385.
- [7] D. R. B. Ly *et al.*, “Novel 1T2R1T RRAM-based ternary content addressable memory for large scale pattern recognition,” in *IEDM Tech. Dig.*, Dec. 2019, p. 35.
- [8] R. Arakawa, N. Onizawa, J.-P. Diguët, and T. Hanyu, “Multi-context TCAM based selective computing architecture for a low-power NN,” in *Proc. 26th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, Nov. 2019, pp. 117–118.
- [9] X. Yin *et al.*, “An ultra-dense 2FeFET TCAM design based on a multi-domain FeFET model,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 9, pp. 1577–1581, Dec. 2018.
- [10] I. Bayram and Y. Chen, “NV-TCAM: Alternative interests and practices in NVM designs,” in *Proc. IEEE Non-Volatile Memory Syst. Appl. Symp. (NVMSA)*, Aug. 2014, pp. 1–6.
- [11] J.-F. Li, T.-L. Tsai, C.-L. Hsu, and C.-T. Sun, “Testing of configurable 8T SRAMs for in-memory computing,” in *Proc. IEEE 29th Asian Test Symp. (ATS)*, Nov. 2020, pp. 1–5.
- [12] P. Huang, R. Han, and J. Kang, “AI learns how to learn with TCAMs,” *Nature Electron.*, vol. 2, no. 11, pp. 493–494, Nov. 2019. [Online]. Available: <https://www.nature.com/articles/s41928-019-0328-9>
- [13] K. Ni *et al.*, “Ferroelectric ternary content-addressable memory for one-shot learning,” *Nat. Electron.*, vol. 2, no. 11, pp. 521–529, 2019.
- [14] S. Thomann *et al.*, “On the reliability of in-memory computing: Impact of temperature on ferroelectric TCAM,” in *Proc. IEEE 39th VLSI Test Symp. (VTS)*, Apr. 2021, pp. 1–6.
- [15] A. F. Laguna, X. Yin, D. Reis, M. Niemier, and X. S. Hu, “Ferroelectric FET based in-memory computing for few-shot learning,” in *Proc. Great Lakes Symp. VLSI*, New York, NY, USA, May 2019, pp. 373–378.
- [16] S. Jeloka, N. B. Akesh, D. Sylvester, and D. Blaauw, “A 28 nm configurable memory (TCAM/BCAM/SRAM) using push-rule 6T bit cell enabling logic-in-memory,” *IEEE J. Solid-State Circuits*, vol. 51, no. 4, pp. 1009–1021, Apr. 2016.
- [17] X. Zhang, X. Chen, and Y. Han, “FeMAT: Exploring in-memory processing in multifunctional FeFET-based memory array,” in *Proc. IEEE 37th Int. Conf. Comput. Design (ICCD)*, Nov. 2019, pp. 541–549.
- [18] D. Reis, A. F. Laguna, M. Niemier, and X. S. Hu, “Attention-in-memory for few-shot learning with configurable ferroelectric FET arrays,” in *Proc. 26th Asia South Pacific Design Autom. Conf.*, Jan. 2021, pp. 49–54.
- [19] S. K. Thirumala and S. K. Gupta, “Reconfigurable ferroelectric transistor—Part I: Device design and operation,” *IEEE Trans. Electron Devices*, vol. 66, no. 6, pp. 2771–2779, Jun. 2019.
- [20] S. Beyer *et al.*, “FeFET: A versatile CMOS compatible device with game-changing potential,” in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2020, pp. 1–4.
- [21] E. T. Breyer *et al.*, “Ultra-dense co-integration of FeFETs and CMOS logic enabling very-fine grained logic-in-memory,” in *Proc. 49th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2019, pp. 118–121.
- [22] I. O’Connor *et al.*, “Prospects for energy-efficient edge computing with integrated HfO₂-based ferroelectric devices,” in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Oct. 2018, pp. 180–183.
- [23] X. Yin *et al.*, “FeCAM: A universal compact digital and analog content addressable memory using ferroelectric,” *IEEE Trans. Electron Devices*, vol. 67, no. 7, pp. 2785–2792, Jul. 2020.