Chinese Physics Letters **42**, 068101 (2025) **Express Letter**

Ultrafast Ternary Content-Addressable Nonvolatile Floating-Gate Memory Based on van der Waals Heterostructures

Peng Song[1,2], Xuanye Liu[1,2], Jiequn Sun[1,2], Nuertai Jiazila[1,2], Chijun Wei[2], Hui Gao[1,2], Chengze Du[1,2], Hui Guo[1,2,3], Haitao Yang[1,2,3], Lihong Bao[1,2,3*], and Hong-Jun Gao[1,2,3]

[1]*Institute of Physics & Beijing National Center for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100190, China*
[2]*School of Physical Sciences and CAS Key Laboratory of Vacuum Physics, University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*Hefei Nationals Laboratory, Hefei 230088, China*

(Received 1 April 2025; accepted manuscript online 9 May 2025)

As a typical in-memory computing hardware design, nonvolatile ternary content-addressable memories (TCAMs) enable the logic operation and data storage for high throughout in parallel big data processing. However, TCAM cells based on conventional silicon-based devices suffer from structural complexity and large footprint limitations. Here, we demonstrate an ultrafast nonvolatile TCAM cell based on the $MoTe_2$/hBN/multilayer graphene (MLG) van der Waals heterostructure using a top-gated partial floating-gate field-effect transistor (PFGFET) architecture. Based on its ambipolar transport properties, the carrier type in the source/drain and central channel regions of the $MoTe_2$ channel can be efficiently tuned by the control gate and top gate, respectively, enabling the reconfigurable operation of the device in either memory or FET mode. When working in the memory mode, it achieves an ultrafast 60 ns programming/erase speed with a current on-off ratio of $\sim 10^5$, excellent retention capability, and robust endurance. When serving as a reconfigurable transistor, unipolar p-type and n-type FETs are obtained by adopting ultrafast 60 ns control-gate voltage pulses with different polarities. The monolithic integration of memory and logic within a single device enables the content-addressable memory (CAM) functionality. Finally, by integrating two PFGFETs in parallel, a TCAM cell with a high current ratio of $\sim 10^5$ between the match and mismatch states is achieved without requiring additional peripheral circuitry. These results provide a promising route for the design of high-performance TCAM devices for future in-memory computing applications.

DOI: 10.1088/0256-307X/42/6/068101 CSTR: 32039.14.0256-307X.42.6.068101

Introduction. In-memory computing represents an emerging computational architecture that is designed to overcome the von Neumann bottleneck inherent in conventional computer architectures—the physical separation between memory and processing units.[1–3] By performing computational tasks directly within memory cells, this paradigm significantly reduces data movement overhead, thereby enhancing both energy efficiency and computational performance.[4–8] Conventional silicon complementary metal-oxide-semiconductor (CMOS) technology exhibits inherent limitations in power consumption and physical footprint for the versatile requirements of in-memory computing.[9] This inherent limitation drives the critical need for novel device architectures featuring dual functionality in both computing and memory storage. As a typical in-memory computing hardware architecture, ternary content-addressable memories (TCAMs) can perform logic operations in the memory cell by comparing the input data and stored data, which is efficient for high throughout in parallel big data processing. Numerous memory cells have been implemented for construction of TCAMs, including static random-access memory (SRAM),[10] ferroelectric field-effect transistors (FeFETs),[11] spin-transfer torque random-access memory (STT-RAM),[12] resistive random-access memory (RRAM),[13] and non-volatile memories (NVMs).[14] However, achieving high-performance TCAM that simultaneously demonstrates high resistance ratios, compact footprint, and flexible mask configuration remains a significant challenge. Furthermore, prior approaches to non-volatile logic operations in the TCAM cell typically require either excessive numbers of transistors or complicated device architectures. These architectures involve heterogeneous integration with non-volatile memory devices, which exhibits constrained operation speed in memory.[10–14]

Two-dimensional (2D) materials and van der Waals heterostructures based on them demonstrate distinctive advantages for constructing multifunctional devices integrating both logic and memory functionalities, owing to their atomic-scale thickness, tunable band structures, and superior charge transport properties.[15–18] These inherent characteristics endow these materials with exceptional capabilities in device miniaturization and performance optimization, making them uniquely positioned for next-

*Corresponding author. Email: lhbao@iphy.ac.cn
© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

generation integrated circuit applications.[19,20] Recent advancements in 2D material-based floating-gate memories have demonstrated remarkable programming/erasing speeds reaching 20 ns.[21–23] Compared to conventional silicon-based field-effect transistors (FETs), existing 2D-material-based devices utilizing separate p-type and n-type FET architectures demonstrate limited functional superiority and compromised area efficiency. This technological gap highlights the significance of ambipolar semiconductors like $MoTe_2$, which uniquely integrates dual conduction mechanisms (electron and hole transport) within a monolithic material platform.[24] Such inherent bipolarity fundamentally enhances device configuration flexibility, particularly when synergized with innovative structural designs. Moreover, the emerging partial floating-gate (PFG) architecture represents a breakthrough in ambipolar device engineering, enabling non-volatile and dynamically reconfigurable switching between p-type and n-type operational modes through electrostatic doping modulation.[25,26] Therefore, the integration of memory elements with FETs based on ambipolar 2D semiconductor $MoTe_2$ can be synergistically operated to enable novel logic-in-memory architectures, especially for TCAMs.

In this work, we demonstrate an ultrafast nonvolatile TCAM cell by integration of two top-gated partial floating-gate field-effect transistors (PFGFETs) based on the $MoTe_2$/hBN/multilayer graphene (MLG) van der Waals heterostructure. The reconfigurable operation of the device in either memory or FET mode is achieved by tuning the carrier type in the source/drain and central channel regions of the $MoTe_2$ channel through applying the control gate voltage pulses and top gate voltage. When the PFGFET functions as a nonvolatile memory, it achieves an ultrafast 60 ns programming/erase speed with a current on-off ratio of $\sim 10^5$, excellent retention capability, and robust endurance. When the PFGFET serves as a reconfigurable transistor, unipolar p-type and n-type FETs are obtained by adopting ultrafast 60 ns control-gate voltage pulses with different polarities. The monolithic integration of memory and logic within a single PFGFET device enables the content-addressable memory (CAM) functionality. Finally, by integrating two PFGFETs in parallel, a TCAM cell with a high current ratio of $\sim 10^5$ between the match and mismatch states is achieved without requiring additional peripheral circuitry.

Results and Discussion. Figure 1(a) shows a schematic of our floating-gate field effect transistor device, consisting of a vertically stacked molybdenum ditelluride ($MoTe_2$)/hexagonal boron nitride (hBN)/multilayer graphene (MLG)/van der Waals heterostructure on a SiO_2/p^{++} Si substrate, where $MoTe_2$, hBN, MLG, SiO_2 and p^{++} Si serve as the channel, tunneling layer, floating gate, blocking layer and control gate, respectively. Figure 1(b) shows a false-color optical image of a typical device in which MLG (marked by the black dashed line) is directly placed on a 300 nm SiO_2/Si substrate, followed by sequential stacking of 14 nm hBN (marked by green dashed line) and 4.8 nm $MoTe_2$ (marked by red dashed line) on top of the MLG. The thickness of different functional layers was measured by atomic force microscopy (AFM), as shown in Fig. S1, Supporting Information (SI). The Raman spectrum of the 2H-$MoTe_2$ channel is shown in Fig. S2 in SI.[27] Figure 1(c) shows the transfer curves obtained by applying a bidirectional gate voltage sweep (from negative to positive and back to negative) on the control gate (V_{cg}) under a fixed source-drain bias voltage (V_{ds}) of 0.1 V. As the V_{cg} is swept from −40 to 40 V, the transfer curve exhibits a memory window with two similar stable states corresponding to electron-doped and hole-doped configurations, which is consistent with the ambipolar transfer behavior of $MoTe_2$-based filed-effect transistors [Fig. S3(a) in SI]. Retention and endurance are two key metrics for storage devices, which respectively indicate the ability to retain data and the ability to withstand rewrite operations. Figure 1(d) shows the retention of the device. When a positive control-gate voltage pulse ($V_{cg,\,pulse}$) (+35.3 V) with 60 ns width [Fig. S4(a) in SI] is applied, electrons tunnel from the source into the floating gate MLG sequentially through the $MoTe_2$ channel and the tunneling hBN, and ultimately blocked by the SiO_2 blocking layer.[23,28] After the positive $V_{cg,\,pulse}$ removed, electrons are stored in the floating gate and electrostatically modulate the $MoTe_2$ channel, exhibiting a p-type state [blue curve in Fig. 1(d)]. On the contrary, when a negative $V_{cg,\,pulse}$ (−36.7 V) with 60 ns width [Fig. S4(b)] is applied, holes tunnel from the source into the floating gate MLG and electrostatically modulate the $MoTe_2$ channel, resulting in a n-type state [red curve in Fig. 1(d)]. Owing to the high barrier height between MLG and the SiO_2 blocking layer as well as the hBN tunneling layer, the accumulated electrons and holes in the floating gate can be effectively retained after the removal of the $V_{cg,\,pulse}$ (Fig. S5 in SI). The two current states can be stably maintained for more than 1000 s with an on-off ratio of about 10^2 under a fixed V_{ds} of 0.1 V at $V_{cg} = 0$ V, as shown in Fig. 1(d). The endurance performance is demonstrated by repeatedly programming (P) and erasing (E) the memory cell [Fig. 1(e)], which shows that the on-off ratio still retains a value of about 10^2 without obvious degradation after operating for more than 100 cycles. To precisely evaluate the dynamic characteristics of the device, it is essential to monitor the current level of the channel after program/erase operations by applying voltage pulses with varying amplitudes. Figure 1(f) illustrates the impact of $V_{cg,\,pulse}$ with varying amplitudes (under a fixed ultrashort pulse width of ∼60 ns) on the current levels of program/erase states. To ensure the testing consistency, the floating gate was grounded before each program/erase operation to discharge accumulated charges. A distinct current transition is clearly observed starting at −22 V and 24 V, respectively, indicating the onset of Fowler-Nordheim (FN) tunneling process at these operation voltages. Fitting analysis of the tunneling current reveals a pronounced linear relationship between parameters $\ln(I/V_{gs}^2)$ and $1/V_{gs}$, which aligns with the FN tunneling current formula $\ln(I/V^2) = \ln(A) - B/V$, in which A and B are constants, I is tunneling current, and V is bias voltage (Fig. S6 in SI).[21,22] Although the ambipolar nature of $MoTe_2$ results in a limited on/off ratio for $MoTe_2$-based floating-gate memory devices, they still achieve 60 ns ultrafast programming/erasing speeds via FN tunneling,

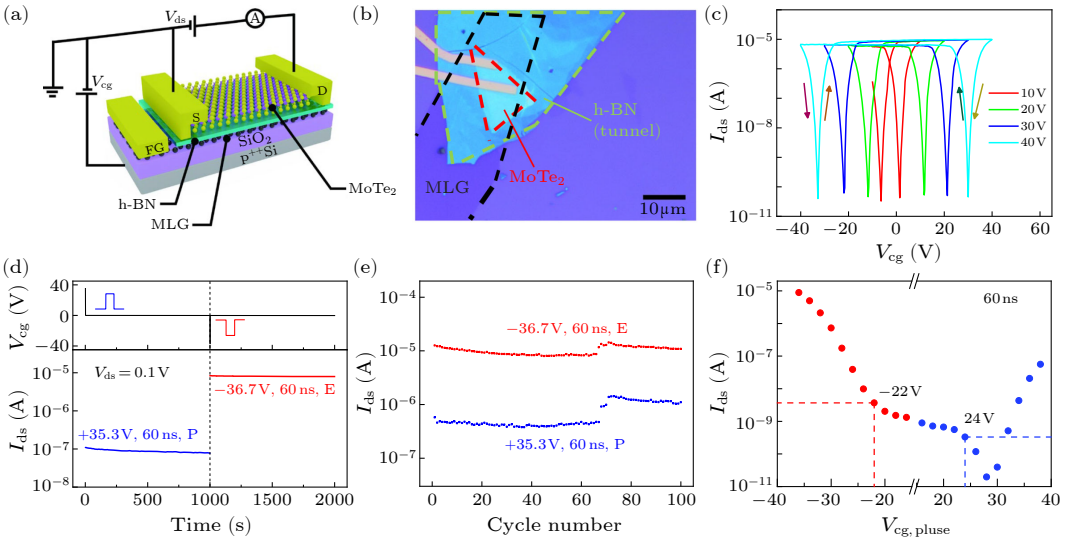


Fig. 1. Structure and characterization of the ultrafast $MoTe_2$ flash-memory. (a) Schematic of a ultrafast non-volatile floating-gate memory device based on a $MoTe_2$/hBN/MLG heterostructure, in which $MoTe_2$, hBN, MLG, SiO_2 and p^{++} Si serve as the channel, tunneling layer, floating gate, blocking layer and control gate, respectively.S and D stand for the source and drain electrodes, respectively. (b) Optical image of a typical memory device on an SiO_2/Si substrate. (c) Dual-sweeping transfer curves of the device by sweeping the control gate voltage within different ranges under a fixed V_{ds} of 0.1 V. (d) Current of the $MoTe_2$ channel can be toggled between the low-conductance state (program state) and high-conductance state (erase state) by applying a single +35.3 V or −36.7 V control gate voltage pulse (V_{cg}) with FWHM of 60 ns, respectively. (e) Endurance performance of the memory device executed with alternate voltage pulses (+35.3/−36.7 V programming/erasing pulses with FWHM of 60 ns). (f) A summary of the I_{ds} changes under different pulse amplitudes with a 60 ns pulse width at room temperature.

comparable to MoS_2- and InSe-based counterparts, while maintaining excellent retention and endurance characteristics.[21–23] Moreover, compared to unipolar materials like MoS_2 and InSe, the ambipolarity of $MoTe_2$ provides enhanced flexibility in device design.

To improve the on/off ratio and reduce the operational voltages of ultrafast floating-gate memory devices based on $MoTe_2$, we reconfigured the floating-gate memory device into the partial floating-gate (PFG) memory device.[25] Figure 2(a) shows a schematic of the PFG memory device architecture, where the source-drain regions and channel region are configured as a control-gate controlled floating-gate memory and a top-gate controlled FET, respectively, enabling the monolithic integration of memory and logic functions. The boundary of the bottom graphene layer (marked by black dashed line), bottom hBN (marked by green dashed line), $MoTe_2$ channel (marked by red dashed line) and top hBN (marked by purple dashed line) is shown in the optical image of a PFG memory device [Fig. 2(b)]. The thicknesses of the bottom hBN, $MoTe_2$, and top hBN layers are 11 nm, 2.7 nm, and 19 nm, respectively [Fig. 2(c)]. Figure 2(d) shows the transfer curves of the PFG memory device obtained by dually sweeping the control gate voltage (V_{cg}) between −40 and 40 V, with a fixed source-drain voltage (V_{ds} = 0.1 V) and varying top-gate voltages (V_{tg} = 0, 1, 2 V). The transfer characteristics of the PFG memory device under a positive V_{tg} exhibit clockwise hysteresis, two distinct memory states, and a large memory window, which is consistent with the behaviors in n-type floating-gate devices.[23] By introducing the top gate to tune the carrier density in the $MoTe_2$ channel, the program/erase voltages of the ultrafast floating gate memory device can be greatly reduced. As shown in Fig. 2(e),

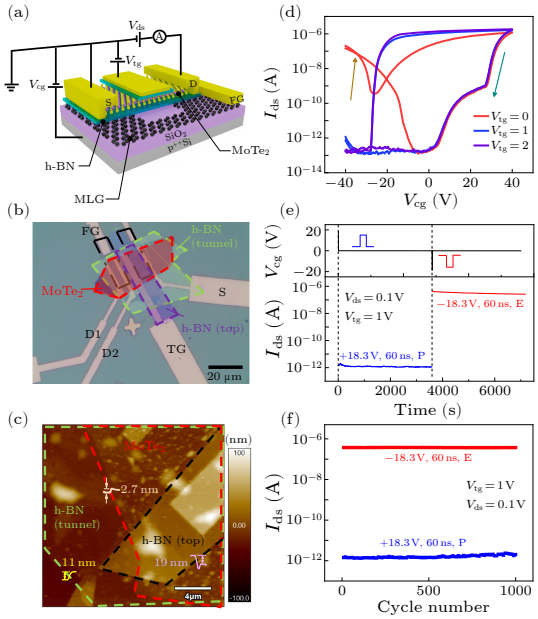


Fig. 2. Device configuration and memory performance of the $MoTe_2$-based PFGFET. (a) Schematic of the PFGFET structure equipped with both a floating gate and a top unit. (b) Optical image of a PFGFET based on hBN/$MoTe_2$/hBN/MLG heterostructure. (c) AFM topographic image of the PFGFET. (d) Transfer characteristics of the PFGFET device operating in the memory mode (V_{ds} = 0.1 V; V_{tg} = 0, 1, 2 V). (e) Time-evolution of current states of the PFGFET device after the programing (V_{cg} = +18.3 V, 60 ns) and erasing (V_{cg} = −18.3 V, 60 ns) operations, reading at V_{ds} = 0.1 V, V_{tg} = 1 V and V_{cg} = 0 V for over 3500 s. A large on-off ratio over 10^5 is achieved. (f) Robust endurance performance of the PFGFET memory device by alternating 60 ns width voltage pulse of V_{cg} = ±20 V, reading at V_{ds} = 100 mV, V_{tg} = 1 V and V_{cg} = 0 V.

the PFG memory device can be successfully programmed/erased by +18.3/−18.3 V pulses with 60 ns width. (The output voltage pulse waveforms are shown in Fig. S7 in SI). Furthermore, the enhanced tuning capability of the carrier density in the $MoTe_2$ channel by adoption of the top gate enables the great increment of the on/off ratio of the program/erase states to exceeding 10^5, as shown in Fig. 2(e). Both program and erase states can be stably maintained for more than 3500 s at a fixed V_{ds} of 0.1 V at $V_{tg} = 1$ V, demonstrating the robust retention characteristics of the PFG memory device. The endurance performance of the PFG device is measured by repeatedly programming and erasing the memory cell, which shows that the on-off ratio still retains $\sim 10^5$ without obvious degradation after operating for more than 1000 cycles [Fig. 2(f)].

Owing to the capability of the monolithic integration of memory cells and FET units in a single device, the partial floating gate (PFG) device can not only function as a nonvolatile memory but also a reconfigurable FET.[25] When the top gate is employed as the input terminal and the control gate is used as the trigger terminal, the PFG device is set into the FET mode. Compared with the majority of reported n-type[29–31] or ambipolar FET,[32–34] our PFG device can operate as an unipolar p-type or n-type FET. Figures 3(a) and 3(b) show the transfer characteristics of the PFG device obtained by sweeping the V_{tg} between −2 and 2 V after applying different V_{cg} pulses. Low off-state current (∼pA) and high on/off current ratio ($>10^5$) are simultaneously obtained for both the p- [Fig. 3(a)] and n-type [Fig. 3(b)] configurations. Figure 3(c) shows the lowest subthreshold swing (SS) values of ∼108 and ∼161 mV/dec for the p-type and n-type FETs, under $V_{cg,\,pulse}$ of +18.3 V and −18.3 V (60 ns width) respectively, further indicating the tuning capability and the excellent interface properties of our PFG device. When a positive $V_{cg,\,pulse}$ is applied, electrons tunnel into the MLG control-gate through the FN tunneling process. After removing the $V_{cg,\,pulse}$, the electrons trapped in the MLG control-gate electrostatically dope the $MoTe_2$ source-drain regions into p-type [Fig. 3(d)]. A negative V_{tg} modulates the central channel region of $MoTe_2$ into p-type as well. As a result, the entire $MoTe_2$ channel becomes homogenously p-type doped, enabling efficient hole-dominated conduction while suppressing electron transport [Fig. 3(e)]. Conversely, a negative V_{tg} modulates the central region of $MoTe_2$ into n-type. The formation of a PNP junction within the $MoTe_2$ channel confines both electrons and holes, suppressing carrier transport and resulting in a high-resistance off-state [Fig. 3(a)]. When a negative $V_{cg,\,pulse}$ is applied, holes tunnel into the MLG control-gate through the FN tunneling process. After removing the $V_{cg,\,pluse}$, the holes trapped in the MLG electrostatically dope the $MoTe_2$ source-drain regions into n-type [Fig. 3(f)]. With a positive V_{tg} applied, the $MoTe_2$ channel is electrostatically modulated into homogeneous n- type, enabling unipolar electron conduction with suppressed hole transport [Fig. 3(g)]. On the contrary, with a negative V_{tg} applied, the central region of the $MoTe_2$ channel is doped into p-type and an NPN junction is formed, suppressing both electron and hole transport and resulting in a high-resistance state [Fig. 3(b)].

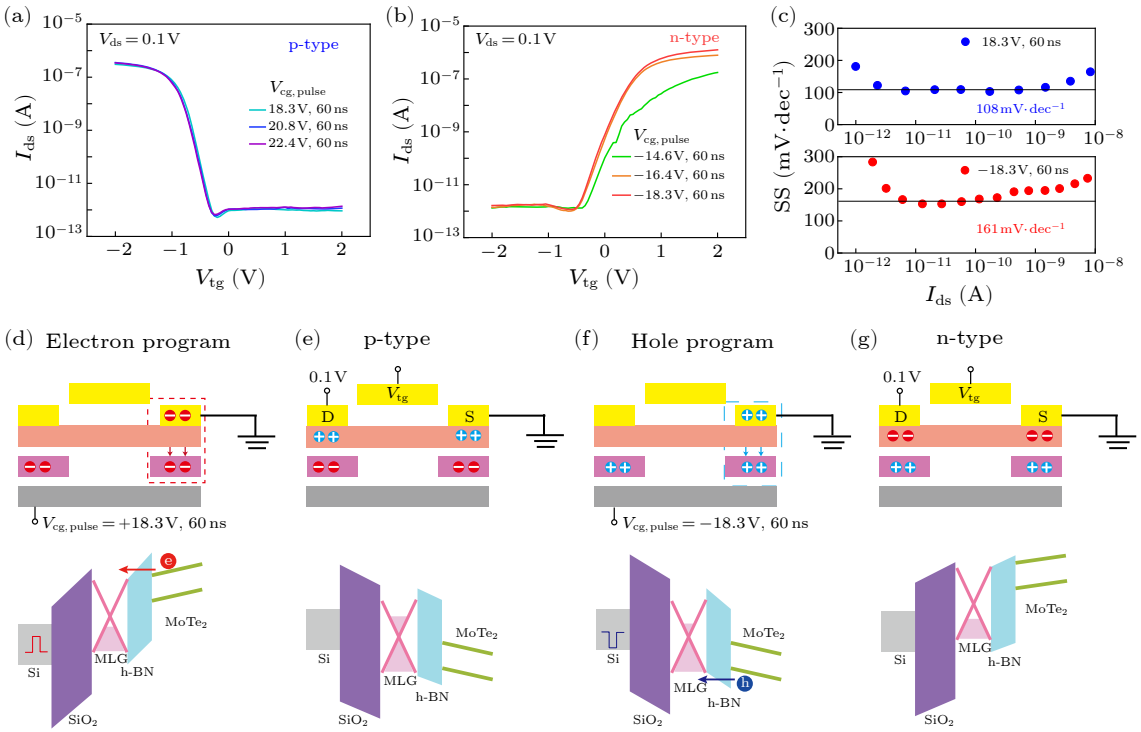
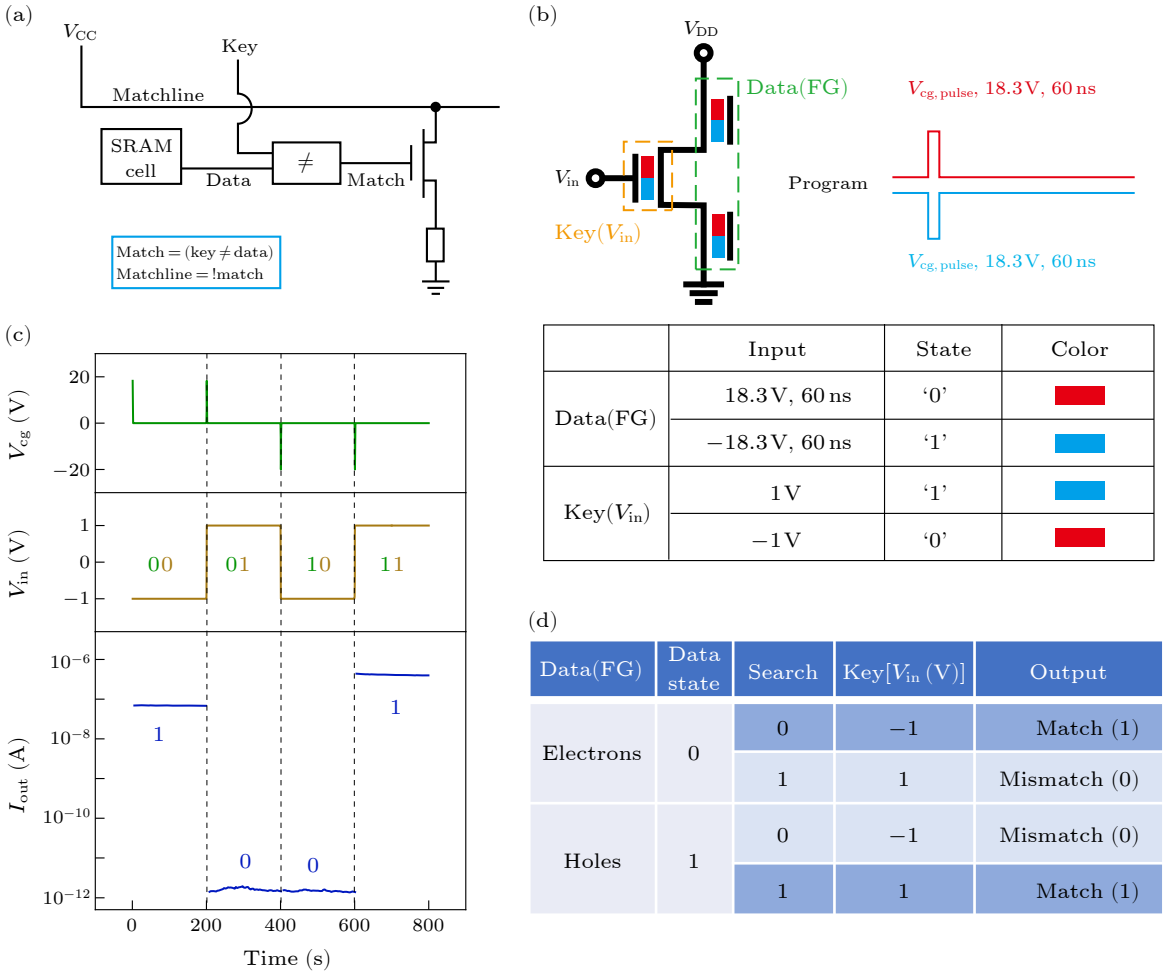


Fig. 3. Transport properties of the ultrafast $MoTe_2$-based PFGFETs in the FET mode. (a) and (b) Transfer curves of the device under p-type (a) and n-type (b) FET states. (c) SS of p-type (upper panel) and n-type (lower panel) FET. (d)–(g) Schematic of the energy band diagram of the PFGFET in different operation states (MLG, multi-layer graphene).



	Input	State	Color
Data(FG)	18.3 V, 60 ns	'0'	
	−18.3 V, 60 ns	'1'	
Key(V_{in})	1 V	'1'	
	−1 V	'0'	

(d)

Data(FG)	Data state	Search	Key[V_{in} (V)]	Output
Electrons	0	0	−1	Match (1)
		1	1	Mismatch (0)
Holes	1	0	−1	Mismatch (0)
		1	1	Match (1)

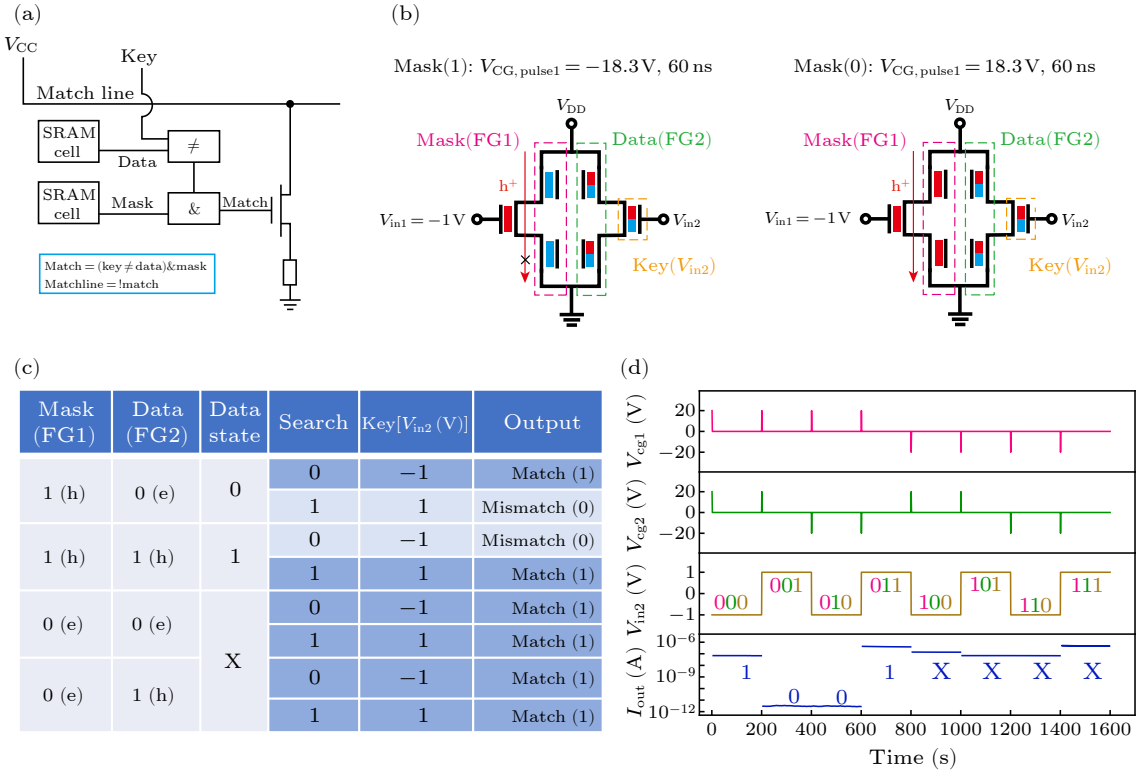
Fig. 4. Content-Addressable Memory (CAM) architecture based on PFGFETs. (a) Schematic diagram of a conventional CAM architecture comprising an SRAM cell and additional match-line logic circuitry. (b) Schematic diagram of a novel CAM architecture based on PFGFETs. The top-gate serves as the key input terminal. The floating-gate layer enables non-volatile data storage, and matching operations are implemented through fundamental FET-based logic mechanisms. (c) Reading of the current of the novel CAM cell using the setup in Fig. 4(b) for both match and mismatch states between the key and the stored data. (d) The state table and search operation in the novel CAM cell. The floating gate (FG) is used for storage and the top gate is used for search data.

Based on the above memory and FET operations, the PFGFET inherently embodies an XNOR (exclusive NOR) logic relationship between its memory cell and FET unit, thereby forming an intrinsic CAM architecture. This integration leverages the dual functionality of charge trapping (memory) and ambipolar transport (logic) within a single $MoTe_2$ heterostructure, enabling reconfigurable non-volatile logic-in-memory operations without additional peripheral circuitry. As schematically shown in Fig. 4(a), the conventional CAM is a memory architecture optimized for parallel content-based retrieval, comprising SRAM-based memory cells, comparator circuits, and a priority encoder. Although this architecture enables high-speed search operations by comparing the binary data stored in each memory cell with input data in constant time, the performance comes at the cost of high-power consumption and low storage density, as each cell requires dedicated comparator circuitry to support parallel search operations. To overcome these limitations, we proposed a PFGFET-based CAM architecture, as schematically depicted in Fig. 4(b). In this architecture, the top gate serves as the input search terminal, while the floating gate functions as the data storage node. The matching operation is executed via the intrinsic XNOR logic between the stored and input states. The logic state is defined as follows. For storage states, logic '0' (red) and '1' (blue) are defined by electrons ('0') and holes ('1') trapped in the floating gate, respectively. For input states, logic '0' (red) and '1' (blue) are defined by applied V_{tg} of −1 V ('0') and +1 V ('1'), respectively. Figure 4(c) shows the output current (I_{out}) in response of the different input states (with a fixed $V_{ds} = 0.1$ V) by defining a high output current as the 'match' (logic '1') state and a low output current as the 'mismatch' (logic '0') state. The corresponding logic truth table is shown in Fig. 4(d), which explicitly demonstrates the XNOR logic operation of the CAM. When searching for a specific stored state, the input logic value is compared with the stored state. If they are identical, the output is a match state (logic '1'); if they differ, the output is a mismatch state (logic '0'). Compared to conventional CAM architectures, the novel PFGFET-based CAM monolithically integrates memory cells, input terminals, and comparison logic within a single-device

structure. This integration enables ultrafast 60 ns non-volatile programming/erasing operations while eliminating peripheral circuits, thereby achieving higher storage density and lower dynamic power consumption.

TCAM is a memory architecture that enables fuzzy matching by introducing a mask bit to represent a 'don't care' (X) state. Figure 5(a) shows the schematic of a conventional TCAM architecture, which comprises two SRAM cells for storing mask and data bits, respectively, along with comparator circuits and a priority encoder. Each TCAM cell encodes three logic states (0, 1, or X) through the combinatorial control of a data bit and a mask bit. During search operations, all cells perform parallel comparisons with the input key. Cells in the X state automatically disregard matching criteria, requiring only valid bits (non-X states) to match for a hit signal to be triggered. The highest-priority match is then output via a priority encoder. This architecture achieves high-speed, flexible network rule lookup and is critical for real-time traffic processing in high-end network switches. However, conventional TCAM implementations suffer from circuit complexity and prohibitive power consumption, limiting their deployment to specialized hardware where throughput outweighs efficiency constraints.[10]

A TCAM architecture based on the van der Waals heterostructure can be constructed by integrating two parallel-connected PFGFETs without requiring additional circuitry, as shown in Fig. 5(b), in which the logical dependencies among the mask, data, and search key in TCAM can be realized based on the OR gate logic from parallel-connected circuits and the XNOR gate logic from PFGFETs. In the TCAM structure, Floating Gate 1 (FG1), Floating Gate 2 (FG2), and Top Gate 2 (TG2) are functionally assigned as the mask bit, data bit, and search key, respectively. The definition of logic states '0' and '1' are in accordance with those shown in Fig. 4(b). A constant voltage of −1 V is applied to Top Gate 1 (TG1). High output current is defined as a 'match' state, whereas low output current is defined as a 'mismatch' state. The operation logic is defined as follows. For Mask = '1' [Left panel in Fig. 5(b)], in PFGFET1, the $MoTe_2$ channel exhibits a NPN junction configuration, resulting in a high-resistance state. Owing to the parallel-connected configuration, TCAM operates in alignment with PFGFET2, functioning as a CAM. For Mask = '0' [Right panel in Fig. 5(b)], In PFGFET1, the $MoTe_2$ channel operates in a p-doped configuration, exhibiting a low-resistance state. Due to the parallel-connected configuration, the TCAM exhibits a low-resistance state [don't care (X)] regardless of the data and key states. The output state of TCAM is formulated as Output = $\overline{\text{Mask}} + (\overline{\text{Data} \oplus \text{Key}})$. The storage states and the data search scheme are shown in Fig. 5(c). Figure 5(d) shows a simulation result of the current characteristics for the TCAM based on the parallel



Mask (FG1)	Data (FG2)	Data state	Search	Key[V_{in2} (V)]	Output
1 (h)	0 (e)	0	0	−1	Match (1)
			1	1	Mismatch (0)
1 (h)	1 (h)	1	0	−1	Mismatch (0)
			1	1	Match (1)
0 (e)	0 (e)	X	0	−1	Match (1)
			1	1	Match (1)
0 (e)	1 (h)		0	−1	Match (1)
			1	1	Match (1)

Fig. 5. Ternary Content-Addressable Memory (TCAM) architecture based on PFGFETs. (a) Schematic diagram of a conventional TCAM architecture comprising two SRAM cells and priority encoder units. (b) Schematic diagram of a novel TCAM architecture based on two parallel-configured PFGFETs under two mask states. Top gate 1 (TG1) is continuously biased at −1 V and Top gate 2 (TG2) serves as the key input terminal, with floating gate 1 (FG1) holding mask states and floating gate 2 (FG2) storing data bits. (c) The state table and search operation in the novel TCAM cell. (d) Simulation of output current characteristics in the novel TCAM under varying mask states, data states, and key states through parallel circuit configurations.

current summation mechanism and the current characteristics of CAM, which exhibits consistency with those shown in Fig. 5(c). The ratio between I_{match} and I_{mismatch} reaches $\sim 10^5$. This high on/off current ratio enables the TCAM to reliably distinguish between 'match' and 'mismatch' states. Figure S8 summarizes the performance parameters of TCAM cell designs based on different technologies. A TCAM based on two PFGFETs features ultra-fast (60 ns) non-volatile memory, integrating mask storage, information storage, and logic circuits without the need for additional circuitry, which holds the great promise for future in-memory computing hardwares.

Conclusion. In summary, we demonstrated an ultrafast ternary content-addressable nonvolatile floating gate memory based on $MoTe_2$/hBN/MLG van der Waals heterostructures using a PFGFET architecture, which was enabled by the reconfigurable operation of the device into the ultrafast non-volatile floating-gate memory and unipolar (p-type or n-type) FET under different bias conditions of the control gate and top gate. When it functions as a memory element, it achieves ultrafast programming/erasing speeds of ∼60 ns, retention times exceeding 3,500 s, and endurance over 1,000 cycles. When it acts as a reconfigurable transistor, it can be switched between p-type and n-type FETs by applying 60 ns control-gate voltage pulses with different polarities, respectively, achieving low subthreshold swings of ∼108 and ∼161 mV/dec for electrons and holes conduction, respectively. The implementation of CAM employs an in-memory computing device architecture that simultaneously achieves both memory storage and logic operations. Furthermore, by integrating critical functionalities of non-volatile mask storage, non-volatile data storage, and logic operations within a single cell, a TCAM device was implemented with a stable current ratio of $\sim 10^5$ between the match and mismatch states. These results highlight the remarkable potential of the PFGFET architecture based on van der Waals heterostructures for future computing-in-memory applications.

Device Fabrication and Electrical Characterization. $MoTe_2$, hBN, and graphite bulk crystals were purchased from 2D Semiconductors, HQ Graphene and NGS Naturgraphit, respectively. First, the silicon chip with 285 nm thick SiO_2 was treated by O_2 plasma to change the hydrophilicity. Second, multilayer graphene (MLG) was mechanically exfoliated on the silicon chip and then annealed in vacuum furnace to remove the residue. After that, hBN and $MoTe_2$ were stacked sequentially onto the MLG layer by the dry transfer method.[35] It is worth noting that $MoTe_2$ was stacked onto the hBN layer in a glove box in an argon atmosphere. The contact electrodes of 6/60 nm Cr/Au were defined by standard electron-beam lithography followed by by thermal evaporation and lift-off process. Electrical measurements were performed with a Keithley 4200 semiconductor characterization system (4200-SCS) and a home-made electric circuit in a vacuum probe station at room temperature. Electrical characterization of the voltage pulses with FWHM of 60 ns was performed in the 4200-SCS system equipped with 4225-PMU and 4225-RPM units. The thicknesses of hBN and $MoTe_2$ layers were determined by a Bruker Dimension Edge atomic force microscope after all electrical tests.

Acknowledgements. This work was supported by the National Key Research & Development Projects of China (Grant No. 2022YFA1204100), National Natural Science Foundation of China (Grant No. 62488201), CAS Project for Young Scientists in Basic Research (YSBR-003), the Innovation Program of Quantum Science and Technology (2021ZD0302700).

References

[1] Wong H S P and Salahuddin S 2015 *Nat. Nanotechnol.* **10** 191
[2] Zidan M A, Strachan J P, and Lu W D 2018 *Nat. Electron.* **1** 22
[3] Boybat I, Le Gallo M, Nandakumar S R, Moraitis T, Parnell T, Tuma T, Rajendran B, Leblebici Y, Sebastian A, and Eleftheriou E 2018 *Nat. Commun.* **9** 2514
[4] Tong L, Peng Z R, Lin R F, Li Z, Wang Y L, Huang X Y, Xue K H, Xu H Y, Liu F, Xia H, Wang P, Xu M S, Xiong W, Hu W D, Xu J B, Zhang X L, Ye L, and Miao X S 2021 *Science* **373** 1353
[5] Sebastian A, Le Gallo M, Khaddam-Aljameh R, and Eleftheriou E 2020 *Nat. Nanotechnol.* **15** 529
[6] Shulaker M M, Hills G, Park R S, Howe R T, Saraswat K, Wong H S P, and Mitra S 2017 *Nature* **547** 74
[7] Li Y, Zhao M, Ma X D, Zhang L, Zhao S Z, Strupinski W, Zeng X Y, Zhang M J, and Hao Y F 2025 *Adv. Funct. Mater.* 2423333
[8] Chen H W, Xue X Y, Liu C S, Fang J B, Wang Z, Wang J L, Zhang D W, Hu W D, and Zhou P 2021 *Nat. Electron.* **4** 399
[9] Yang R, Li H T, Smithe K K H, Kim T R, Okabe K, Pop E, Fan J A, and Wong H S P 2019 *Nat. Electron.* **2** 108
[10] Pagiamtzis K and Sheikholeslami A 2006 *IEEE J. Solid-State Circuits* **41** 712
[11] Yin X Z, Chen X M, Niemier M, and Hu X S 2019 *IEEE Trans. Very Large Scale Integr. VLSI Syst.* **27** 159
[12] Matsunaga S, Miura S, Honjou H, Kinoshita K, Ikeda S, Endoh T, Ohno H, and Hanyu T 2012 *2012 Symposium on VLSI Circuits* (VLSIC) (Honolulu, HI, USA: IEEE) p. 44
[13] Chang M F, Lin C C, Lee A, Chiang Y N, Kuo C C, Yang G H, Tsai H J, Chen T F, and Sheu S S 2017 *IEEE J. Solid-State Circuits* **52** 1664
[14] Xiong X, Kang J Y, Liu S Y, Tong A Y, Fu T Y, Li X F, Huang R, and Wu Y Q 2022 *Adv. Mater.* **34** 2106321
[15] Xiong X, Huang M Q, Hu B, Li X F, Liu F, Li S C, Tian M C, Li T Y, Song J, and Wu Y Q 2020 *Nat. Electron.* **3** 106
[16] Huang M Q, Li S M, Zhang Z F, Xiong X, Li X F, and Wu Y Q 2017 *Nat. Nanotechnol.* **12** 1148
[17] Liu X Y, Li L X, Wei C J, Song P, Gao H, Wu K, Jiazila N, Sun J Q, Guo H, Yang H T, Zhou W, Bao L H, and Gao H J 2024 *Chin. Phys. B* **33** 127302
[18] Wei C J, Jiazila N, Liu X Y, Song P, Gao H, Sun J Q, Bao L H, Lin X, and Gao H J 2024 *Chin. Phys. Lett.* **41** 128501
[19] Gao A Y, Lai J W, Wang Y J, Zhu Z, Zeng J W, Yu G L, Wang N Z, Chen W C, Cao T J, Hu W D, Sun D, Chen X H, Miao F, Shi Y, and Wang X M 2019 *Nat. Nanotechnol.* **14** 217
[20] Migliato Marega G, Zhao Y F, Avsar A, Wang Z Y, Tripathi M, Radenovic A, and Kis A 2020 *Nature* **587** 72
[21] Wu L M, Wang A W, Shi J N, Yan J H, Zhou Z, Bian C, Ma J J, Ma R S, Liu H T, Chen J C, Huang Y, Zhou W, Bao L H, Ouyang M, Pennycook S J, Pantelides S T, and Gao H J 2021 *Nat. Nanotechnol.* **16** 882

[22] Liu L, Liu C S, Jiang L L, Li J Y, Ding Y, Wang S Y, Jiang Y G, Sun Y B, Wang J L, Chen S Y, Zhang D W, and Zhou P 2021 *Nat. Nanotechnol.* **16** 874

[23] Wang H, Guo H, Guzman R, Jiazila N, Wu K, Wang A W, Liu X Y, Liu L, Wu L M, Chen J C, Huan Q, Zhou W, Yang H T, Pantelides S T, Bao L H, and Gao H J 2024 *Adv. Mater.* **36** 2311652

[24] Pan J, Wu F, Wang Z D, Liu S J, Guo P W, Yin J J, Zhao B C, Tian H, Yang Y, and Ren T L 2024 *ACS Nano* **18** 28189

[25] Sun X X, Zhu C G, Yi J L, Xiang L, Ma C, Liu H W, Zheng B Y, Liu Y, You W X, Zhang W J, Liang D L, Shuai Q, Zhu X L, Duan H G, Liao L, Liu Y, Li D, and Pan A L 2022 *Nat. Electron.* **5** 752

[26] Gao H, Liu X Y, Song P, Wei C J, Jiazila N, Sun J Q, Wu K, Guo H, Yang H T, Bao L H, and Gao H J 2024 *Chin. Phys. B* **33** 127201

[27] Wu D, Guo C G, Zeng L H, Ren X Y, Shi Z F, Wen L, Chen Q, Zhang M, Li X J, Shan C X, and Jie J S 2023 *Light Sci. Appl.* **12** 5

[28] Yu J, Wang H, Zhuge F, Chen Z R, Hu M, Xu X, He Y H, Ma Y, Miao X S, and Zhai T Y 2023 *Nat. Commun.* **14** 5662

[29] Lin Z Y, Liu Y, Halim U, Ding M N, Liu Y Y, Wang Y L, Jia C C, Chen P, Duan X D, Wang C, Song F, Li M F, Wan C Z, Huang Y, and Duan X F 2018 *Nature* **562** 254

[30] Li N, Wang Q Q, Shen C, Wei Z, Yu H, Zhao J, Lu X B, Wang G L, He C L, Xie L, Zhu J Q, Du L J, Yang R, Shi D X, and Zhang G Y 2020 *Nat. Electron.* **3** 711

[31] Chen X Y, Xie Y F, Sheng Y C, Tang H W, Wang Z M, Wang Y, Wang Y, Liao F Y, Ma J Y, Guo X J, Tong L, Liu H Q, Liu H, Wu T X, Cao J X, Bu S T, Shen H, Bai F Y, Huang D M, Deng J N, Riaud A, Xu Z H, Wu C J, Xing S W, Lu Y, Ma S L, Sun Z Z, Xue Z Y, Di Z F, Gong X, Zhang D W, Zhou P, Wan J, and Bao W Z 2021 *Nat. Commun.* **12** 5953

[32] Pezeshki A, Hosseini Shokouh S H, Jeon P J, Shackery I, Kim J S, Oh I K, Jun S C, Kim H, and Im S 2016 *ACS Nano* **10** 1118

[33] Zhang H, Li C, Wang J L, Hu W D, Zhang D W, and Zhou P 2018 *Adv. Funct. Mater.* **28** 1805171

[34] Lee S J, Lin Z Y, Huang J, Choi C S, Chen P, Liu Y, Guo J, Jia C C, Wang Y L, Wang L Y, Liao Q L, Shakir I, Duan X D, Dunn B, Zhang Y, Huang Y, and Duan X F 2020 *Nat. Electron.* **3** 630

[35] Wu L M, Shi J N, Zhou Z, Yan J H, Wang A W, Bian C, Ma J J, Ma R S, Liu H T, Chen J C, Huang Y, Zhou W, Bao L H, Ouyang M, Pantelides S T, and Gao H J 2020 *Nano Res.* **13** 1127