Article

Hybrid 2D–CMOS microchips for memristive applications

https://doi.org/10.1038/s41586-023-05973-1

Received: 4 March 2022

Accepted: 17 March 2023

Published online: 27 March 2023

Open access

Kaichen Zhu[1,11], Sebastian Pazos[1,11], Fernando Aguirre[1], Yaqing Shen[1], Yue Yuan[1], Wenwen Zheng[1], Osamah Alharbi[1], Marco A. Villena[1], Bin Fang[1], Xinyi Li[2], Alessandro Milozzi[3], Matteo Farronato[3], Miguel Muñoz-Rojo[4,5], Tao Wang[6], Ren Li[7], Hossein Fariborzi[7], Juan B. Roldan[8], Guenther Benstetter[9], Xixiang Zhang[1], Husam N. Alshareef[1], Tibor Grasser[10], Huaqiang Wu[2], Daniele Ielmini[3] & Mario Lanza[1]✉

Exploiting the excellent electronic properties of two-dimensional (2D) materials to fabricate advanced electronic circuits is a major goal for the semiconductor industry[1,2]. However, most studies in this field have been limited to the fabrication and characterization of isolated large (more than 1 μm^2) devices on unfunctional SiO_2–Si substrates. Some studies have integrated monolayer graphene on silicon microchips as a large-area (more than 500 μm^2) interconnection[3] and as a channel of large transistors (roughly 16.5 μm^2) (refs. 4,5), but in all cases the integration density was low, no computation was demonstrated and manipulating monolayer 2D materials was challenging because native pinholes and cracks during transfer increase variability and reduce yield. Here, we present the fabrication of high-integration-density 2D–CMOS hybrid microchips for memristive applications—CMOS stands for complementary metal–oxide–semiconductor. We transfer a sheet of multilayer hexagonal boron nitride onto the back-end-of-line interconnections of silicon microchips containing CMOS transistors of the 180 nm node, and finalize the circuits by patterning the top electrodes and interconnections. The CMOS transistors provide outstanding control over the currents across the hexagonal boron nitride memristors, which allows us to achieve endurances of roughly 5 million cycles in memristors as small as 0.053 μm^2. We demonstrate in-memory computation by constructing logic gates, and measure spike-timing dependent plasticity signals that are suitable for the implementation of spiking neural networks. The high performance and the relatively-high technology readiness level achieved represent a notable advance towards the integration of 2D materials in microelectronic products and memristive applications.

Our 2 cm × 2 cm silicon microchips have been designed by means of Synopsys software and fabricated in a 200 mm silicon wafer in an industrial clean room using a 180 nm CMOS technology node (Fig. 1a and Extended Data Fig. 1). The circuits fabricated in this study consist of 5 × 5 crossbar arrays of one-transistor-one-memristor cells (1T1M, Fig. 1b,c and Supplementary Fig. 1), although some standalone memristors and CMOS transistors were fabricated for reference (Supplementary Fig. 2). The microchips have been designed to integrate the memristors into the back-end-of-line (BEOL) interconnections; that is, they have been terminated at the latest metallization layer (fourth in our wafer) and have been left without passivation. Hence, silicon oxide naturally grows on the wafers when they are extracted from the industrial clean room (Fig. 1d), which can be easily etched away to expose the tungsten vias (Fig. 1e and Supplementary Fig. 3). Then, a roughly 18-layer-thick sheet of hexagonal boron nitride (h-BN) (that is, roughly 6 nm), grown on a Cu substrate by means of chemical vapour deposition (CVD), was transferred on the microchips (Fig. 1f) using a low-temperature process (Methods). Finally, the h-BN on the contact pads was etched, and top electrodes made of different materials (that is, Au–Ti, Au or Ag) were patterned and deposited on the h-BN to finalize the circuits (Fig. 1g).

As the tungsten vias of the fourth metallization layer have a diameter of roughly 260 nm (Fig. 1h and Supplementary Fig. 3), the lateral size of the resulting h-BN memristors is, at most, 0.053 μm^2. Figure 1h shows a high-angle annular dark-field cross-sectional scanning transmission

[1]Materials Science and Engineering Program, Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [2]Institute of Microelectronics, Tsinghua University, Beijing, China. [3]Department of Electronics, Information and Bioengineering, Politecnico of Milan, Milan, Italy. [4]Department of Thermal and Fluid Engineering, Faculty of Engineering Technology, University of Twente, Enschede, the Netherlands. [5]Institute of Micro and Nanotechnology, IMN-CNM, CSIC (CEI UAM+CSIC), Madrid, Spain. [6]Institute of Functional Nano and Soft Materials, Collaborative Innovation Center of Suzhou Nanoscience and Technology, Soochow University, Suzhou, China. [7]Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [8]Department of Electronics and Computer Technology, Faculty of Sciences, University of Granada, Granada, Spain. [9]Department of Electrical Engineering and Media Technology, Deggendorf Institute of Technology, Deggendorf, Germany. [10]Institute for Microelectronics, TU Wien, Vienna, Austria. [11]These authors contributed equally: Kaichen Zhu, Sebastian Pazos. ✉e-mail: mario.lanza@kaust.edu.sa

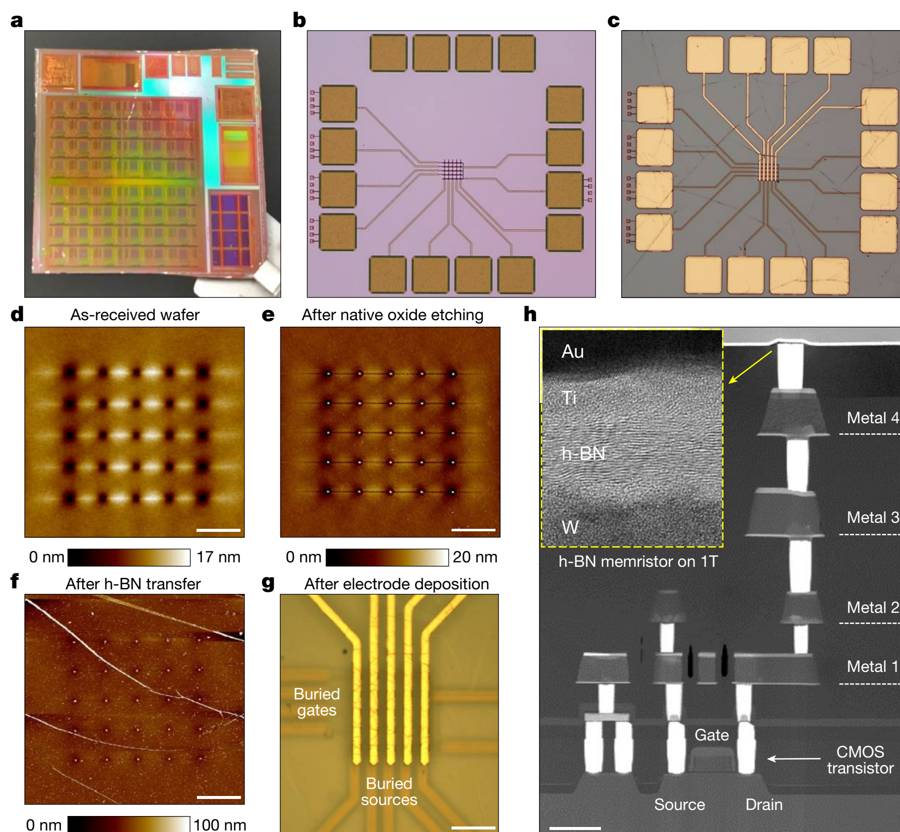


Fig. 1 | Fabrication of hybrid 2D–CMOS memristive microchips. **a**, Photograph of the 2 cm × 2 cm microchips containing the CMOS circuitry. **b**,**c**, Optical microscope images of a part of the microchip containing a 5 × 5 crossbar array of 1T1 M cells, as received (**b**) and after fabrication (**c**). The size of the squared pads is 50 µm × 50 µm. **d**–**f**, Topographic maps collected with atomic force microscopy of the vias in the 5 × 5 crossbar arrays on the wafers as received (**d**), after native-oxide etching (**e**) and after the transfer of the h-BN sheet (**f**). **g**, Optical microscope image of a finished 5 × 5 crossbar array of 1T1M, that is, after h-BN transfer and top electrodes deposition. **h**, High-angle annular dark-field cross-sectional scanning transmission electron microscope image of a 1T1M cell in the crossbar array. The inset, which is 20 nm × 16 nm, shows a cross-sectional TEM image of the Au–Ti–h-BN–W memristor on the via; the correct layered structure of h-BN can be seen. Scale bars, **d**–**f**, 10 µm; **g**, 25 µm; **h**, 600 nm.

electron microscope image of a 1T1M cell (with top Au–Ti electrode) in the crossbar array (Supplementary Figs. 4 and 5). The correct layered structure of the h-BN stack is confirmed before and after transfer by means of cross-sectional transmission electron microscopy (TEM) (inset in Fig. 1h and Extended Data Fig. 2). Nano-chemical analyses by means of electron energy loss spectroscopy demonstrate the correct composition of the h-BN sheet (Extended Data Fig. 3). The optical microscope images (Fig. 1c) reveal that the h-BN sheet does not crack during the transfer; this is an important advantage of using roughly 6-nm-thick 2D layered materials, and it increases the yield of the devices and circuits compared to counterparts using monolayer 2D materials[6].

Electronic memory

When we apply sequences of ramped voltage stresses (RVS) to several standalone 0.053 µm^2 Au–Ti–h-BN–W structures, most (roughly 90%) of them show erratic current fluctuations and no resistive switching (RS) is observed (Fig. 2a); the currents do not reach linear regime (that is, dielectric breakdown) even if we apply 11 V. This is striking because most (more than 75%) h-BN devices with larger areas (25 µm^2) show dielectric breakdown voltages (V_{DB}) between 3 and 11 V followed by filamentary non-volatile bipolar RS[6]. The reason should be the lower probability to find clusters of defects in small devices, which remarkably increases V_{DB} (ref. 7). Few (roughly 10%) 0.053 µm^2 Au–Ti–h-BN–W structures show V_{DB} between roughly 2.5 and 4 V followed by filamentary non-volatile bipolar RS (if a current limitation 1 mA or higher is applied, Supplementary Fig. 6). However, the endurance is only around 100 cycles, mainly due to the poor controllability of the current across the memristor and the overshoot during the dielectric breakdown[6].

On the contrary, the CMOS transistor in the 1T1M cell can precisely control the current across the h-BN memristor and avoid the current overshoot, which results in outstanding performance. First, we obtain the output characteristic of one standalone CMOS transistor by applying a constant voltage to the gate (V_G) and a RVS to the drain (V_{DS}), and measuring the drain-to-source current (I_{DS}); the CMOS transistor works correctly as expected (Fig. 2b). And second, we measure the 1T1M cell by applying RVS at the top Au–Ti electrode of the memristor while keeping the source terminal of the transistor grounded and simultaneously applying a constant V_G. When a sequence of RVS is applied to the top electrode of the Au–Ti–h-BN–W structure using V_G = 1.1 V, most 1T1M cells show non-volatile bipolar RS (Fig. 2c and Extended Data Fig. 4). The high state resistances (R_{HRS} of roughly 200 MΩ and R_{LRS} of roughly 200 KΩ)—beneficial to reduce power consumption—the non-linearity of the currents in both states and the progressive state transitions indicate that the RS is non-filamentary[8]. However, we do see an activation process, as the first RVS slightly increases the conductance of the devices (that is, softly degrades the h-BN stack, blue line in Fig. 2c and Extended Data Fig. 4). In the first microchip that we fabricated, this stable non-filamentary bipolar RS regime was observed in 32 out of 40 cells (yield 80%), and in the last one it was observed in 25 out of 25 devices (yield 100%). On the contrary, standalone Au–Ti–h-BN–W structures of 0.053 µm^2 and 1T1M cells without h-BN never showed this behaviour; this confirms that the RS is produced by the h-BN stack and that the CMOS transistor is key to control its soft degradation. Note

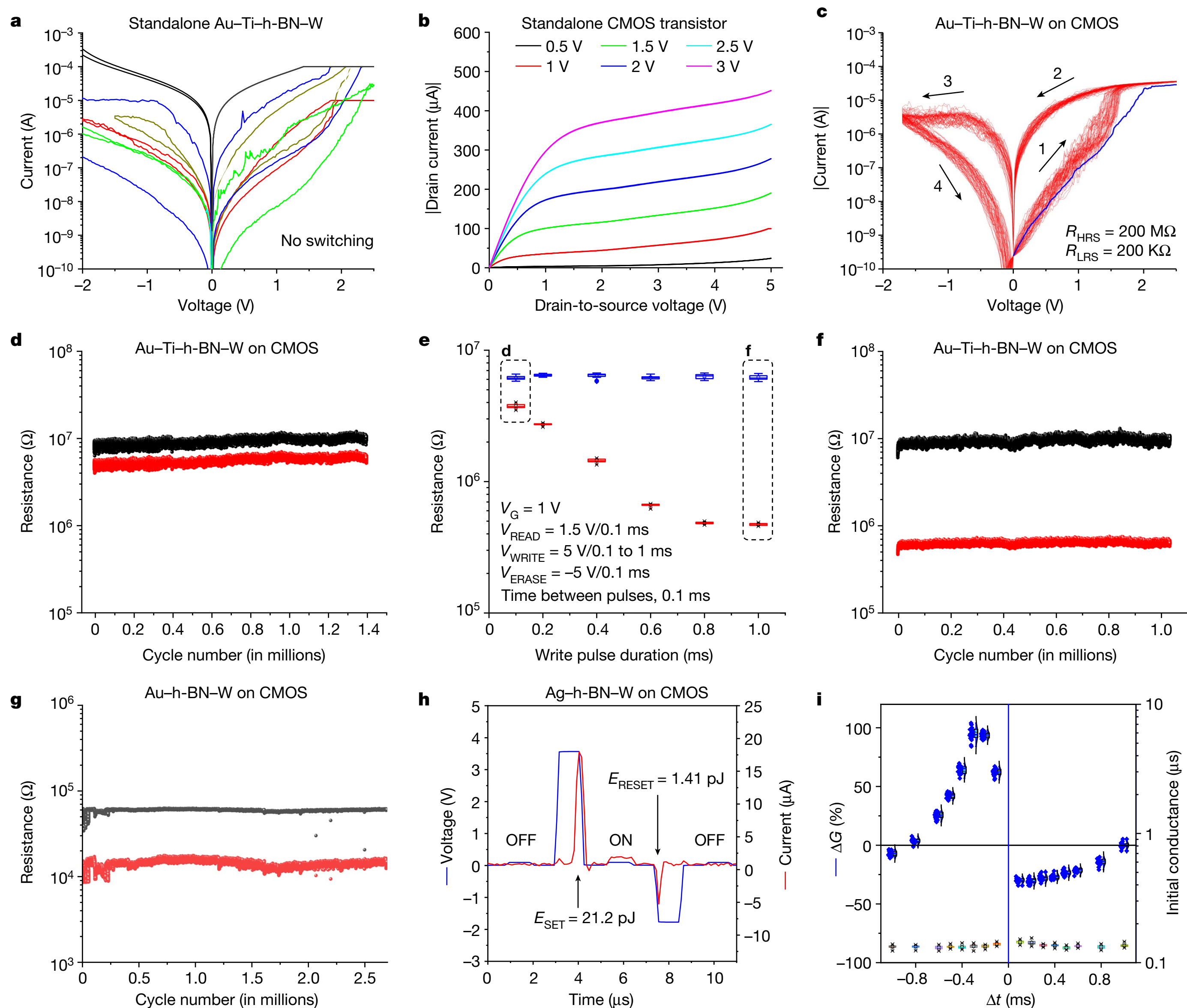


Fig. 2 | Electrical characterization of h-BN–CMOS based 1T1M cells. **a**, Electrical characterization of 0.053 μm² Au–Ti–h-BN–W structures, showing erratic current fluctuations and no stable RS (each colour line corresponds to one RVS with two polarities). **b**, Typical output characteristic of all standalone CMOS transistors. **c**, Typical non-volatile bipolar RS measured in most 1T1M cells with an Au–Ti–h-BN–W memristor and a CMOS transistor (when applying $V_G = 1.1$ V). **d**–**f**, Endurance plots of 1T1M cell showing around 1.4 million cycles and 1 million cycles for write pulse durations of 0.1 ms (**d**) and 1 ms (**f**). **e**, R_{LRS} and R_{HRS} in a 1T1M cell when applying pulsed voltage stresses of different durations. **g**, Endurance plot showing non-volatile bipolar RS at $V_G = 1$ V for 1T1M cells using Au–h-BN–W memristors. All the endurance tests have been conducted following the recommended characterization process described in ref. 42. **h**, Voltage and current versus time in a 1T1M cell with Ag–h-BN–W memristor, showing a low switching energy. **i**, STDP characteristic of the 1T1M cell with Au–Ti–h-BN–W memristor. Before STDP characterization, the devices are always tuned to the same initial conductance (lower box charts, which relate to the right *y* axis).

that in the standalone Au–Ti–h-BN–W structures the current is limited using the semiconductor parameter analyser, which activation time is long (roughly 70 μs) and parasitic capacitance is high (roughly 300 pF, related to the cables)[9]; on the contrary, in the 1T1M cell the series transistor acts as an instantaneous current limitation (it cannot drive more current than that allowed by the size of its channel) and the parasitic capacitance is much lower (roughly 50 pF, internal connections in the microchip), which reduces the duration of the switching transient and undesired currents across the Au–Ti–h-BN–W structure[10]. The values of R_{HRS} and R_{LRS} are stable over time, and multiple stable conductance levels can be programmed either by adjusting V_G during the set process (which fixes R_{LRS}) and/or by adjusting the end voltage of the negative RVS (which fixes R_{HRS}, Extended Data Fig. 5).

The most surprising observation, however, relates to the endurance, which readily reaches 2.5 million cycles (Fig. 2d–f) when applying sequences of pulsed voltage stresses. Under this type of stress, the values of R_{HRS}, R_{LRS} and R_{LRS}/R_{HRS} can be accurately controlled in three different ways: by tuning the duration of the write pulse, by tuning the amplitude of the write pulse, and by tuning the amplitude of the erase pulse (Fig. 2d–f and Extended Data Fig. 6). This endurance is very high considering the small size of the memristors (Supplementary Note 2), and similar to that of commercial metal-oxide-based resistive random access memories (0.5 million cycles)[11] and phase-change memories (10 million cycles)[12,13]. However, the switching time of the 1T1M cells using top Au–Ti electrodes is rather long (t_{SET} of 232 μs and t_{RESET} of 783 ns, Extended Data Fig. 7).

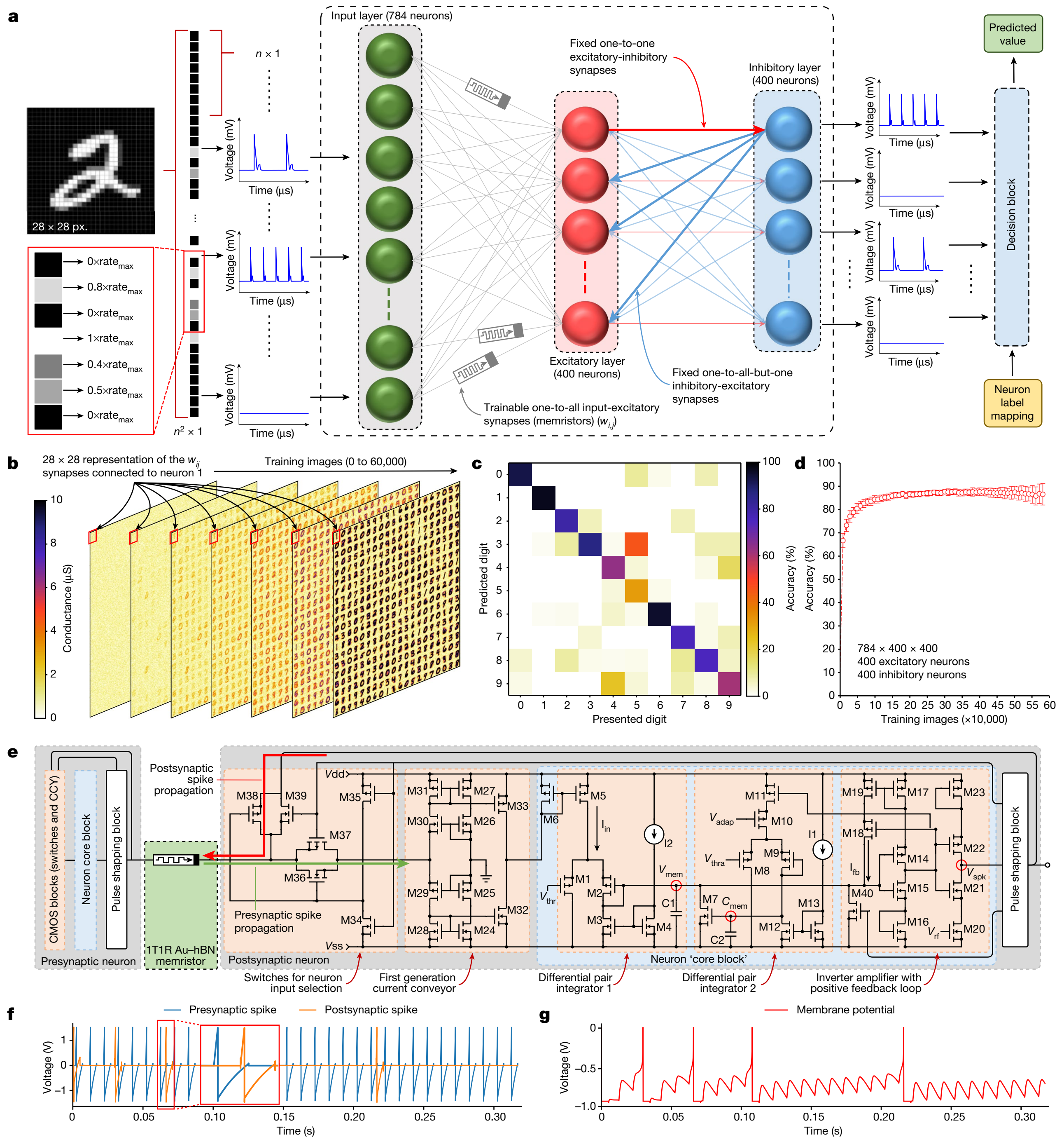


Fig. 3 | Implementation of a SNN using CMOS–h-BN based 1T1M cells. **a**, Structure of the considered SNN. Each MNIST image is reshaped as a 784 × 1 column vector, and the intensity of the pixels is encoded in terms of the firing frequency of the input neurons. The only trainable synapses are those connecting the input layer with the excitatory layer, and they are modelled with the STDP characteristic of the CMOS–h-BN based 1T1M cells. The learning is unsupervised, and the neurons are labelled only after the training. These label-neuron assignments are then feed to the decision block altogether with the firing patterns of the neurons, to infer the class of the image presented in the input. **b**, Evolution of the synaptic connections between the input and excitatory layers during training for the case of 400 excitatory and/or inhibitory neurons. The red square identifies 784 synapses arranged in a 28 × 28 representation. **c**, Confusion matrix indicating the classification accuracy for each class from the dataset. **d**, Classification accuracy as a function of the number of presented training images for the neural network comprising 400 excitatory and/or inhibitory neurons. The error bars show the standard deviation for 50 Monte Carlo simulation runs for every accuracy point. **e**, Circuit schematic of the proposed neuron–synapse–neuron block combining h-BN based 1T1M cells and CMOS circuitry. The colours indicate the complete neuron (grey surrounding box), the core block (light-blue box) and the individual building blocks (light-red boxes). CCY, current conveyor. **f**, SPICE simulation of the pre- and postsynaptic signals applied to the CMOS–h-BN based 1T1M. **g**, SPICE simulation of the neuron's membrane potential. The firing events progressively separate from each other due to the adaptative firing threshold.

The properties of the 1T1M cells can be adjusted using different top electrodes (Extended Data Fig. 8). When Au electrodes are used, the devices show reliable switching at lower state resistances (Fig. 2g), as well as shorter switching time (t) and lower switching energy (E), and when Ag electrodes are used, these values can be pushed down to t_{SET} = 680 ns, t_{RESET} = 60 ns, E_{SET} = 21.11 pJ and E_{RESET} = 1.41 pJ (Fig. 2h). The

reasons behind these observations are the lack of an interfacial Ti layer (which is prone to absorb oxygen, increasing the out-of-plane resistance) and the higher conductivity and diffusivity of Au^{X+} and Ag^{X+} ions[14] (Supplementary Note 1). The performance observed in h-BN–CMOS 1T1M cells using Au–Ti electrodes may allow them to cover niche applications between NAND Flash and DRAM within the memory hierarchy (for example, persistent memory), and when using Au or Ag electrodes their performance may be valid for low-power application-specific integrated circuits within the internet-of-things[15] (Supplementary Fig. 12).

Data computation

On the basis of the above measured performance metrics, the hybrid 2D–CMOS 1T1M cells show good potential for data computation. The high R_{HRS}/R_{LRS} ratio and the stability of the resistive states over time allows us to implement in-memory computing operations taking advantage of the internal connections of our 5 × 5 crossbar array of 2D–CMOS 1T1M cells. As a proof of concept, we realized 'or' and 'implication' operations (Extended Data Fig. 9), although more sophisticated operations could be easily realized by modifying the interconnections between the devices by means of custom design.

Furthermore, the 1T1M cells with Au–Ti–h-BN–W memristors show spike-timing dependent plasticity (STDP) when applying pairs of pulsed voltage stresses displaced in time at the input and output (Fig. 2i). This non-volatile RS performance is very attractive to construct electronic synapses for spiking neural networks (SNNs)[16], which consume less energy than traditional deep neural networks[17].

Although implementing by means of hardware a reliable 2D materials-based memristive SNN capable of competing with state-of-the-art developments[18,19] is not yet achievable due to the lower maturity of these materials, we can analyse the performance of a SNN made of memristors that show STDP characteristics such as those in Fig. 2i (Supplementary Note 3). First, we fit the measured STDP data from Fig. 2i, including the device-to-device variability, using an exponential decaying model to implement the learning rule (Supplementary Fig. 13). Second, we simulate a SNN to demonstrate the unsupervised learning capability (Fig. 3a), and benchmark it by classifying the images from the Modified National Institute of Standards and Technology (MNIST) database of handwritten digits[20,21] (Methods). The SNN has 784 input neurons, an excitatory layer of 400 neurons and an inhibitory layer of 400 neurons, plus a decision block that determines which is the most probable digit (0–9) represented by the input pattern. We trained the SNN with the complete MNIST dataset and evaluated the accuracy every 1,000 images. Figure 3b–d shows the main three figures-of-merit for this type of SNNs (that is, evolution of the synaptic weights with the number of training images, confusion matrix of the network and the training accuracy versus number of training images) and all of them indicate an excellent performance. To account for the device variability, we considered a Monte Carlo simulation with 50 iterations that randomizes the exponential fitting of the STDP plot and the initial value of the synapses, and the deviations observed in the accuracy are very low (less than 5%, Fig. 3d and Supplementary Fig. 14). The best average accuracy reaches roughly 90%, which is a very high value considering the simplicity of the SNN and the unsupervised training protocol (Supplementary Table 4).

We also propose a CMOS circuit for the hardware implementation of an electronic neuron based on our h-BN memristors (Fig. 3e), which is capable of accounting for the adaptative firing threshold and the refractory period after firing (pre- and postsynaptic traces and the evolution of the membrane potential, simulated by means of SPICE, in Fig. 3f,g).

Discussion

Very few commercial electronic products today already include 2D materials, and the ones that do (sensors[22], specialty cameras[23]) use very low integration density (more than 100 μm² per device)—because in larger devices the local defects in the 2D material are not so detrimental. Our hybrid 2D–CMOS microchips are still far from being ready for production, but we can safely claim that our work represents the highest performance and technology readiness level ever achieved in high-integration-density 2D materials-based electronic devices or circuits. The electrical characteristics of the h-BN memristors connected to a CMOS transistor are by orders of magnitude superior to those of standalone h-BN memristors[6,24–27] and h-BN memristors connected to 2D materials-based transistors[28,29].

The voltages needed to switch our devices (from ±1.4 to ±5 V) are low compared to other prototypes in the field of 2D materials (even more than 20 V)[30–32], but still higher than that used at the 180 nm CMOS node. Nevertheless, this is not an impediment for the development of this technology, as there are many commercial microchips that operate at much higher voltages; that is the case for all Flash memories[33] (state-of-the-art 3D-NAND Flash memories are programmed at around 20 V)[34] and all bipolar-CMOS microchips for automotive applications (which require up to 40 V)[35]. Strategies to fabricate wafers with devices that operate at different voltages are widespread[36], and many companies[37,38] offer versions of their 180 nm CMOS technology that operate at high voltages greater than 18 V. Note that prototype memristive devices developed by companies also operate at ±5 V (ref. 39).

We finally remark that, at a first glance, the use of Au and Ag electrodes may not appear ideal because they are categorized as contaminant in front-end-of-line (FEOL) processes. However, our h-BN memristors are integrated in the latest metallic layer of the BEOL interconnections (Fig. 1h), where Au pads, liners and wires are usually used[40] (Extended Data Fig. 10). The semiconductor industry has also developed ferroelectric memories with high content of Iridium[41] (a contaminant material forbidden in FEOL processes), and companies working in the field of 2D materials use Au electrodes in their studies and FEOL prototypes (Supplementary Table 5). Hence, the use of Au, Au–Ti or Ag electrodes in our hybrid 2D–CMOS microchips for memristive applications does not prevent their adoption by the industry.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-05973-1.

1. Akinwande, D. et al. Graphene and two-dimensional materials for silicon technology. *Nature* **573**, 507–518 (2019).
2. *The International Roadmap for Devices and Systems* (IEEE, 2021); https://irds.ieee.org/editions/2021.
3. Goossens, S. et al. Broadband image sensor array based on graphene–CMOS integration. *Nat. Photon.* **11**, 366–371 (2017).
4. Lin, Y. et al. Wafer-scale graphene integrated circuit. *Science* **332**, 1294–1297 (2011).
5. Han, S., Garcia, A. V., Oida, S., Jenkins, K. A. & Haensch, W. Graphene radio frequency receiver integrated circuit. *Nat. Commun.* **5**, 3086 (2014).
6. Shen, Y. et al. Variability and yield in h-BN-based memristive circuits: the role of each type of defect. *Adv. Mater.* **33**, 2103656 (2021).
7. Chen, A. Forming voltage scaling of resistive switching memories. In *Proc. 71st Device Research Conference* 181–182 (IEEE, 2013).
8. Sawa, A. Resistive switching in transition metal oxides. *Mater. Today* **11**, 28–36 (2008).
9. 2600B Source Measure Units (SMU) Instruments. *Tektronix* https://www.tek.com/en/datasheet/2600b-source-measure-units-smu-instruments (2023).
10. Chen, A. Current overshoot during set and reset operations of resistive switching memories. In *Proc. 2012 IEEE International Reliability Physics Symposium (IRPS)* MY.2.1–MY.2.4 (IEEE, 2012)
11. Non-volatile memory with very small operating current ReRAM. *Fujitsu* https://www.fujitsu.com/jp/group/fsm/en/products/reram/ (2023).
12. Analyzing Intel-Micron 3D XPoint: The Next Generation Non-Volatile Memory. *Anandtech* https://www.anandtech.com/show/9470/intel-and-micron-announce-3d-xpoint-nonvolatile-memory-technology-1000x-higher-performance-endurance-than-nand (2023).
13. Kau, D. C. et al. Stackable cross point phase change memory. In *Proc. 2009 IEEE International Electron Devices Meeting (IEDM)* 1–4 (IEEE, 2009).

14. Yang, Y. et al. Electrochemical dynamics of nanoscale metallic inclusions in dielectrics. *Nat. Commun.* **5**, 4232 (2014).
15. Lanza, M. et al. Memristive technologies for data storage, computation, encryption, and radio-frequency communication. *Science* **376**, eabj9979 (2022).
16. Yang, J., Strukov, D. & Stewart, D. Memristive devices for computing. *Nat. Nanotech.* **8**, 13–24 (2013).
17. Merolla, P. A. et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**, 668–673 (2014).
18. Li, C. et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nat. Commun.* **9**, 2385 (2018).
19. Ishii, M. et al. On-chip trainable 1.4M 6T2R PCM synaptic array with 1.6K stochastic LIF neurons for spiking RBM. In *Proc. 2019 IEEE International Electron Devices Meeting (IEDM)* 14.2.1–14.2.4 (IEEE, 2019).
20. Diehl, P. U. & Cook, M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Front. Comput. Neurosc.* **9**, 1662–5188 (2015).
21. LeCun, Y., Cortes, C. & Burges, C. MNIST handwritten digit database (ATT Labs, 2010); http://yann.lecun.com/exdb/mnist/
22. Graphenea card. *Graphenea* https://www.graphenea.com/ (2023).
23. Graphene Flagship industrial partner Emberion launches wide spectrum array sensor at Laser World of Photonics. *Graphene Flagship* https://graphene-flagship.eu/graphene/news/emberion-launches-hyperspectral-array-sensor-at-laser-world-of-photonics/ (2019).
24. Kim, M. et al. Analogue switches made from boron nitride monolayers for application in 5G and terahertz communication systems. *Nat. Electron.* **3**, 479–485 (2020).
25. Shi, Y. et al. Electronic synapses made of layered two-dimensional materials. *Nat. Electron.* **1**, 458–465 (2018).
26. Qian, K. et al. Hexagonal boron nitride thin film for flexible resistive memory applications. *Adv. Funct. Mater.* **26**, 2176–2184 (2016).
27. Zhuang, P. et al. Nonpolar resistive switching of multilayer-hBN-based memories. *Adv. Electron. Mater.* **6**, 1900979 (2020).
28. Yang, R. et al. Ternary content-addressable memory with MoS_2 transistors for massively parallel data search. *Nat. Electron.* **2**, 108–114 (2019).
29. Wang, C. et al. 3D monolithic stacked 1T1R cells using monolayer MoS_2 FET and hBN RRAM fabricated at low (150 °C) temperature. In *Proc. 2018 IEEE International Electron Devices Meeting (IEDM)* 22.5.1-22.5.4 (IEEE, 2018).
30. Wu, F. et al. Vertical MoS_2 transistors with sub-1-nm gate lengths. *Nature* **603**, 259–264 (2022).
31. Liu, L. et al. Uniform nucleation and epitaxy of bilayer molybdenum disulfide on sapphire. *Nature* **605**, 69–75 (2022).
32. Marega, G. et al. Logic-in-memory based on an atomically thin semiconductor. *Nature* **587**, 72–77 (2020).
33. Goda, A. 3-D NAND technology achievements and future scaling perspectives. *IEEE Trans. Electron Devices* **67**, 1373–1381 (2020).
34. Lin, W. L. et al. Grain boundary trap-induced current transient in a 3-D NAND flash cell string. *IEEE Trans. Electron Devices* **66**, 1734–1740 (2019).
35. Automotive Standard Analog Robustness and Performance. Brochure number BR2206AUTOICS. *ST Microelectronics* https://www.st.com/en/automotive-analog-and-power.html#documentation (2023).
36. Niimi, H., Khamankar, R. & Alshareef, H. N. Method for non-thermally nitride gate formation for high voltage devices. US patent 6,566 B2 (2004).
37. Mixed-Signal/CMOS. *Tower Semiconductor* https://towersemi.com/technology/mixed-signal-cmos/ (n.d.).
38. Minixhofer, R. et al. A 120 V 180 nm high voltage CMOS smart power technology for system-on-chip integration. In *Proc. 22nd International Symposium on Power Semiconductor Devices & IC's (ISPSD)* 75–78 (IEEE, 2010).
39. Saitoh, M. et al. HfO_2-based FeFET and FTJ for ferroelectric-memory centric 3D LSI towards low-power and high-density storage and AI applications. In *Proc. 2020 IEEE International Electron Devices Meeting (IEDM)* 375–378 (IEEE 2020).
40. Goodman, P. Current and future uses of gold in electronics. *Gold Bull.* **35**, 21–26 (2002).
41. Summerfelt S. R. et al. Ferroelectric capacitor stack etch cleaning methods. US patent 7,220,600 B2 (2007).
42. Lanza, M. et al. Standards for the characterization of endurance in resistive switching devices. *ACS Nano* **15**, 17214–17231 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

Methods

Microchip fabrication

The metal-oxide-semiconductor field-effect-transistor circuits were fabricated in a standard CMOS foundry. The size of the wafers is 200 mm, and the technology node was 180 nm. Each wafer contained 60 chips with a size of 2 cm × 2 cm, and each of them contained different circuits, including the unfinished memristors, finished transistors and unfinished 5 × 5 crossbar arrays of 1T1M cells. As we received the silicon wafer without a passivation layer, first we cut it to separate the microchips and then etched the native oxide by immersing them in a diluted hydrofluoric acid solution (10:1) for 1 min to etch the native oxide (SiO_2). This step was carried out to expose the conductive tungsten vias to make a good electrical contact with 2D materials. Second, we transferred a sheet of multilayer h-BN (previously grown by CVD on a Cu foil) using a wet transfer method. One layer of polymethyl methacrylate (PMMA) with thickness of around 300 nm was spin coated on the cut h-BN. The PMMA–h-BN–Cu sample was deposited on a $FeCl_3$ solution (0.1 g ml^{-1}) to etch the Cu substrate and, once the Cu disappeared, the resulting PMMA–h-BN sample was washed in diluted HCl solution (1 mol l^{-1} for 1 min) and deionized water (for 1 h). The PMMA–h-BN sample was picked up using the native oxide-free CMOS microchip and dried naturally in a dry box. The PMMA was then removed by immersing the sample in acetone for 24 h. Third, we used photolithography (mask aligner from SUSS MicroTec, model MJB4) to expose the h-BN on top of the metallic pad. Then we used a dry etching method with Ar–O_2 plasma (Plasma Cleaner from PVA TePla America Inc., Model IoN 40) to etch the h-BN (300 W for 10 min) and expose the pads. Finally, we used photolithography, electron beam evaporation (Kurt J. Lesker, model PVD75) and a lift-off process (rinse in acetone for 1 min) to pattern and deposit the top electrodes and/or drain electrodes (3 nm Ti with 40 nm Au on top without breaking the vacuum, or 50 nm Au or 50 nm Ag). The process was simple and reproducible, although we believe it could be considerably improved if optimized methods in an industrial clean room are used. Ideally, the h-BN should be grown in large CVD systems (for example, Aixtron[43]) and transferred on the wafers before cutting them in multiple microchips, using methods like laser debonding[44]. Note that even large companies are still using small (7.6 cm or 3-inch) tube furnaces to grow the h-BN for their prototypes[45]. We also confirm that getting the microchips finalized and etch the passivation film before transferring the h-BN works well.

Device characterization

The morphology of the devices was investigated by an optical microscope (DM 4000M, Leica), AFM (Dimension Icon, Bruker) and TEM (Titan Themis, FEI). The thin lamellae for TEM inspection were prepared using a scanning electron microscope provided with focused ion beam (Helios G4 UX, Thermo Fisher Scientific). The electrical characterization was performed by using two probe stations (both M150, Cascade) connected to different semiconductor parameter analysers: a Keithley 4200 and a Keysight B1500A. All the *IV* curves under DC voltages were collected using the Keithley 4200 in the ramped voltage sweep mode, for which three source-measure units are needed for drain, source and gate. Also, all the *IV* curves under pulse mode were collected by Keysight B1500A with two Waveform Generator/Fast Measurement Units connected to the drain and source. An Agilent E3631A DC Power Supply was used to apply constant voltage stress on the transistor gate as gate voltage for pulse measurement. All the endurance plots were collected using the recommended method described in ref. 42.

SNN simulation

The SNN architecture[20] has been developed using Brian2 (ref. 46), an SNN simulator written in Python. The learning process is based on the empirical measurement of STDP made in the 1T1M cells combining an Au–Ti–h-BN–W memristor on a CMOS transistor. We considered the variability of the network by running a Monte Carlo engine. We benchmarked the accuracy of the SNN during image classification[20] of the MNIST dataset of handwritten digits[21] under an unsupervised learning scheme. We propose a circuit-level model for the neuron–synapse–neuron system, as well as its implementation in SPICE. A detailed description of the SNN and its performance is given in Supplementary Note 3.

Data availability

The data needed to evaluate the conclusions in this work are publicly available online at https://doi.org/10.5281/zenodo.7607096. The datasets that we used for benchmarking are publicly available in ref. 21. The training methods are provided in ref. 20.

Code availability

The simulator Brian2 used here is publicly available in ref. 46. The codes used for the simulations described in Methods are publicly available online at https://doi.org/10.5281/zenodo.7607096.

43. Gomollon-Bel, F. Cost effective, large scale graphene with AIXTRON systems. *Graphene Flagship* https://graphene-flagship.eu/graphene/news/cost-effective-large-scale-graphene-with-aixtron-systems/ (2019).
44. Phommahaxay, A. et al. The growing application field of laser debonding: from advanced packaging to future nanoelectronics. In *Proc. 2019 International Wafer Level Packaging Conference (IWLPC)* 1–8 (IEEE, 2019).
45. Chen, T. A. et al. Wafer-scale single-crystal hexagonal boron nitride monolayers on Cu (111). *Nature* **579**, 219–223 (2020).
46. Stimberg, M., Brette, R. & Goodman, D. F. Brian 2, an intuitive and efficient neural simulator. *eLife* **8**, e47314 (2019).

Acknowledgements This work has been supported by the Ministry of Science and Technology of China (grant nos. 2019YFE0124200 and 2018YFE0100800), the National Natural Science Foundation of China (grant no. 61874075) and the Baseline funding scheme of the King Abdullah University of Science and Technology.

Author contributions K.Z., S.P. and M.L. designed the experiments. X.L. and H.W. provided the CMOS wafers. K.Z., Y.S., Y.Y., W.Z. and O.A. fabricated the hybrid 2D–CMOS microchips. K.Z. and S.P. characterized the microchips. K.Z. measured the logic gates. R.L., H.F. and T.W. assisted in electrical characterization. F.A. did the simulation of the neural network. M.A.V., B.F., A.M. and M.F. assisted in the fabrication, characterization and simulation of structures included in intermediate version of this manuscript that were finally removed. M.M.-R., J.B.R., G.B., X.Z., H.A., T.G. and D.I. gave technical recommendations on fabrication, characterization and/or simulations. M.L. wrote the manuscript, which was revised by all the authors.

Competing interests The authors declare no competing interests.

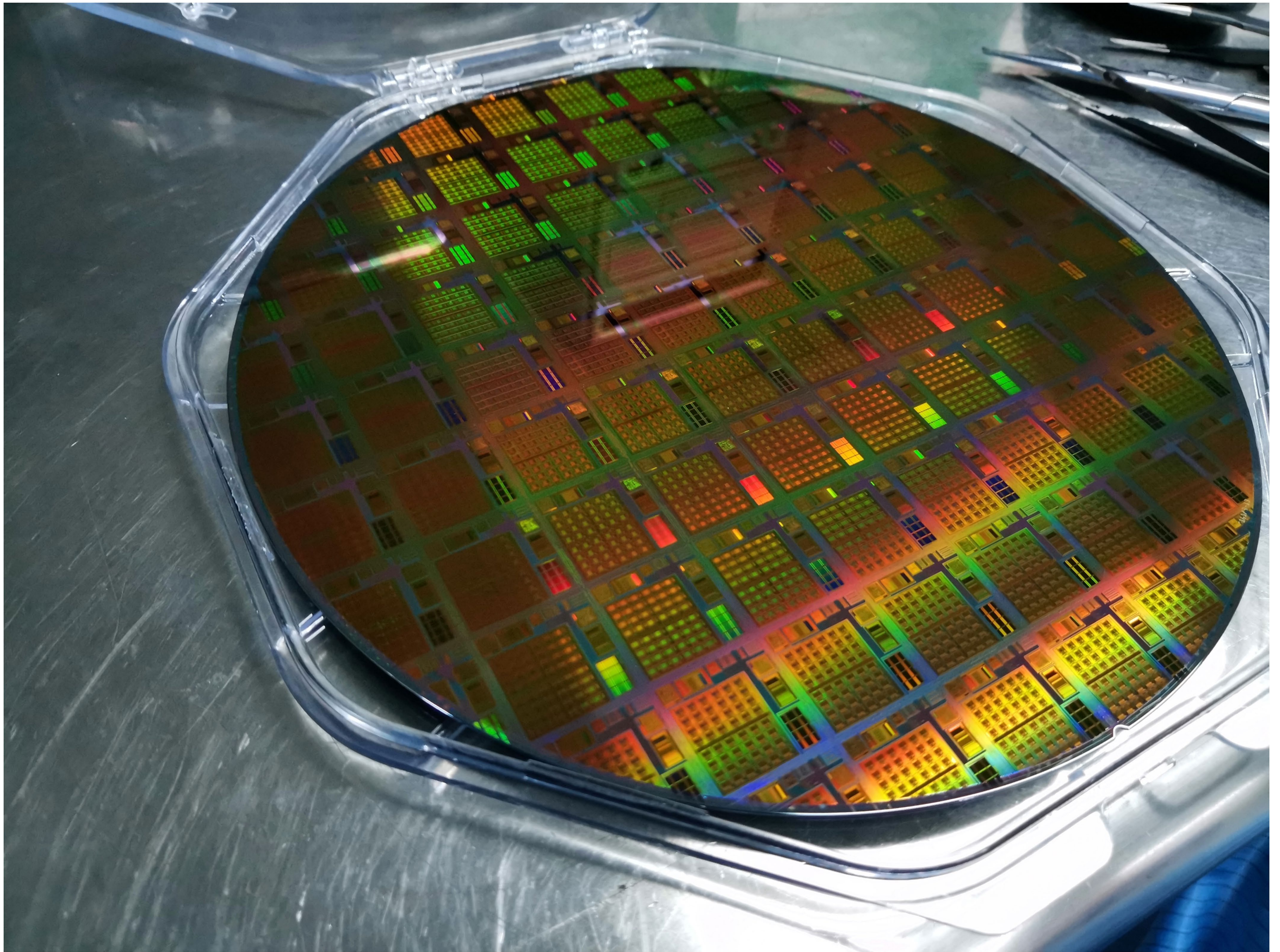
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-05973-1.

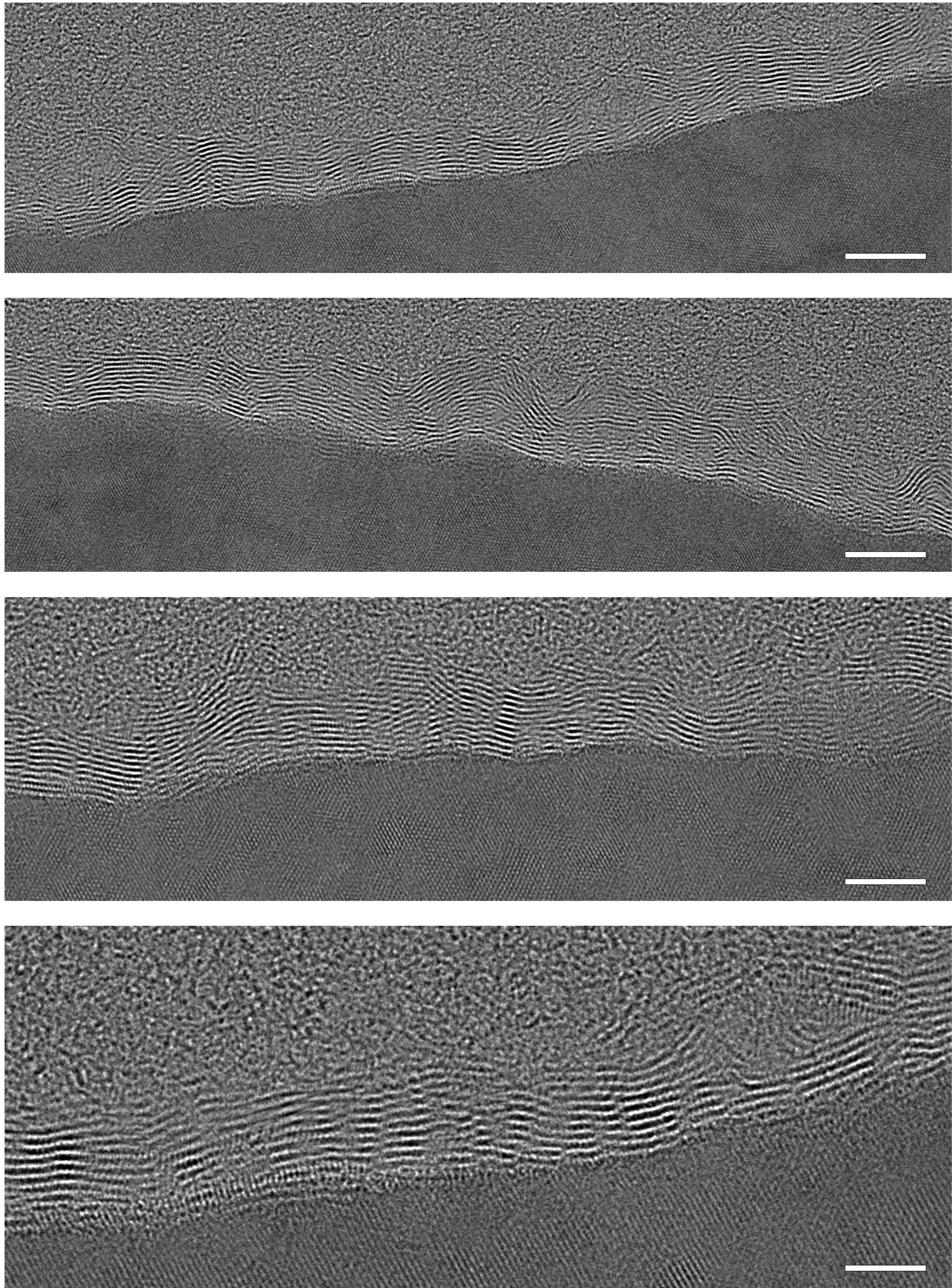
Correspondence and requests for materials should be addressed to Mario Lanza.

Peer review information *Nature* thanks Deji Akinwande, Lain-Jong Li and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

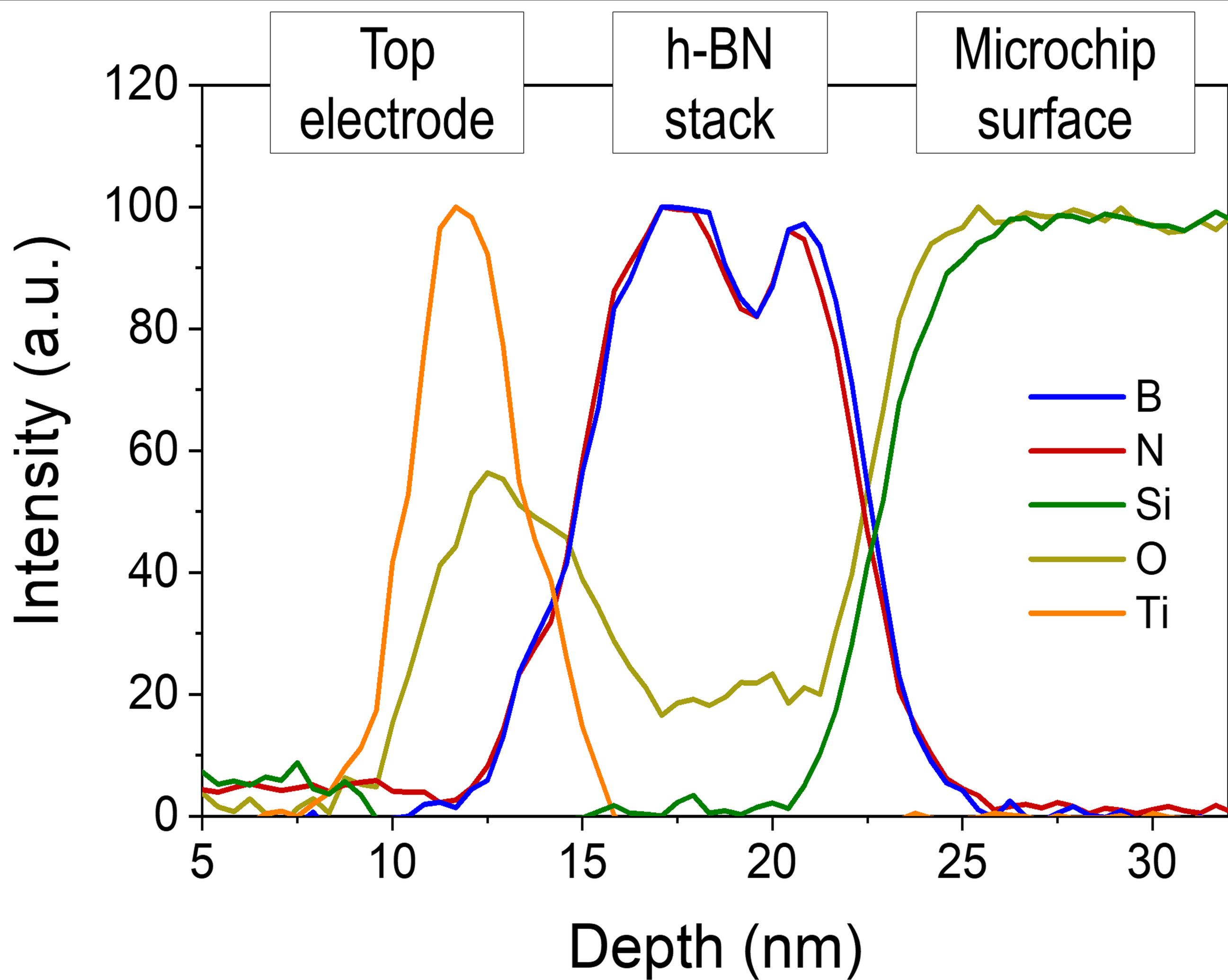
Reprints and permissions information is available at http://www.nature.com/reprints.



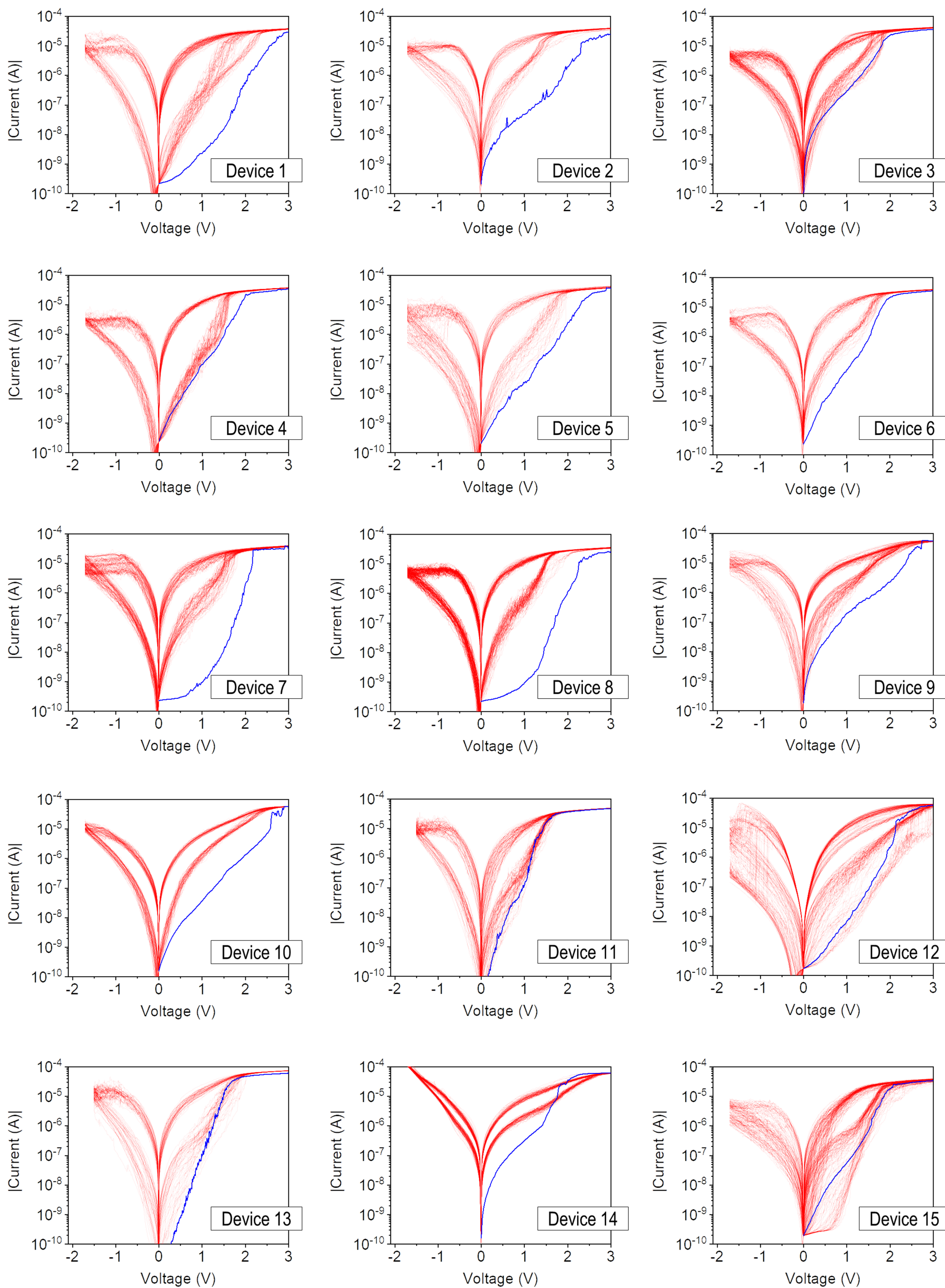
Extended Data Fig. 1 | As-received silicon wafers. Photo of the 200 mm wafer received from the foundry containing all the 2 cm × 2 cm microchips, made of CMOS circuits.



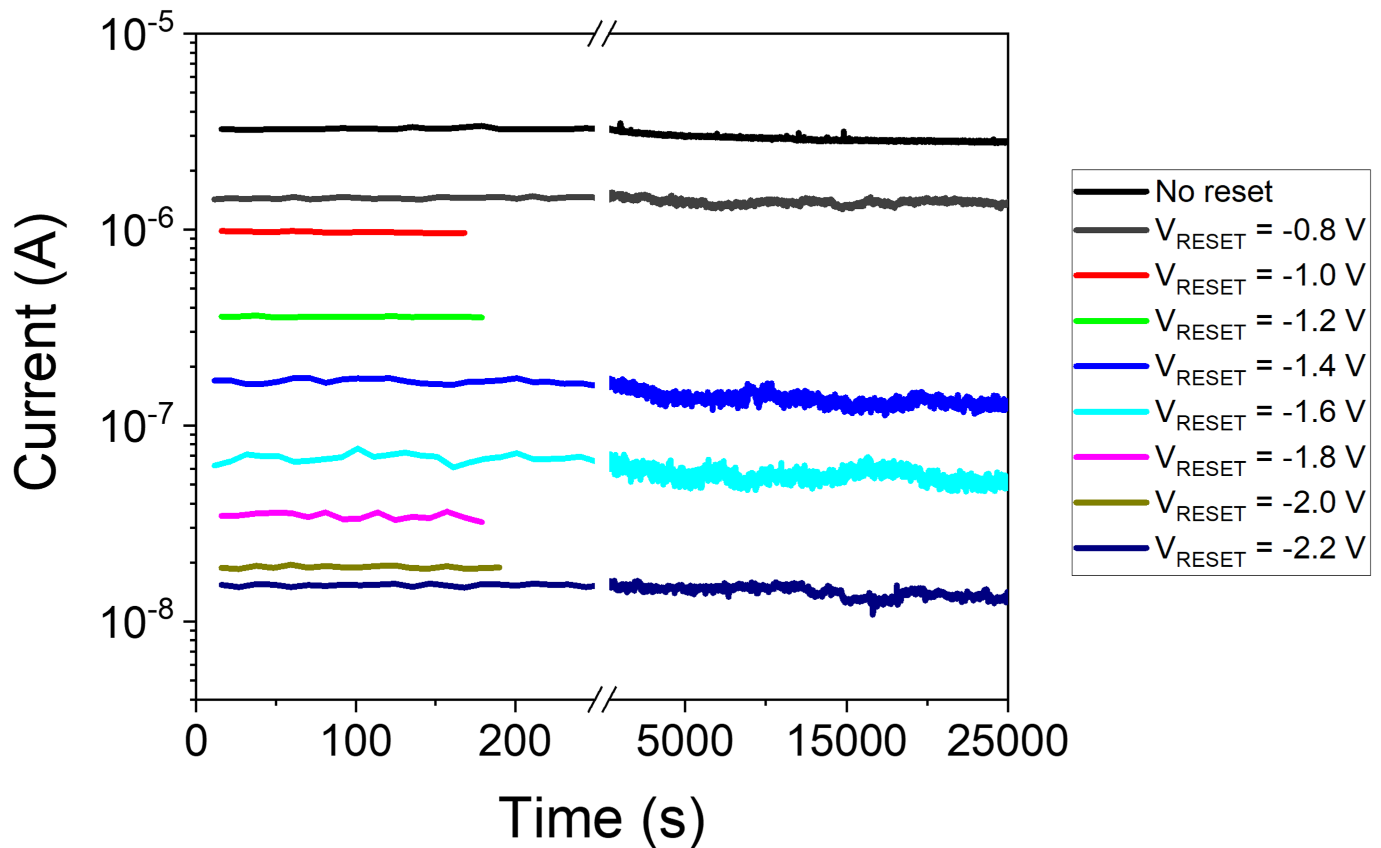
Extended Data Fig. 2 | Morphological analysis of the h-BN. Cross-sectional TEM images of the as-grown multilayer h-BN sheet on Cu. The layered structure is evident, although it contains local defects (i.e., lattice distortions) produced during the CVD process – they are not related to the FIB process because our exfoliated samples never show these features. Those native defects are necessary for the observation of memristive effect. The scale bars (from top to down) and 6 nm, 6 nm, 4 nm and 3 nm.



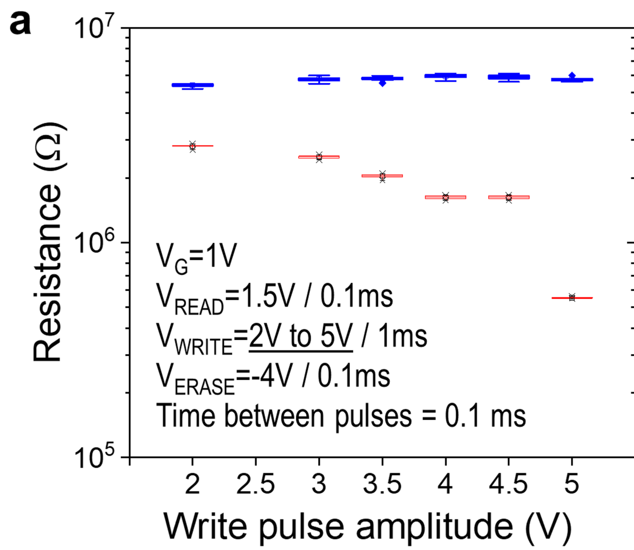
Extended Data Fig. 3 | Nano-chemical characterization of the h-BN stack. Electron energy loss spectroscopy signal showing the correct structure of the h-BN stack on the CMOS microchip. The image also shows that the titanium electrode can absorb a remarkable amount of oxygen. This oxygen peak is not observed when using gold or silver electrodes.

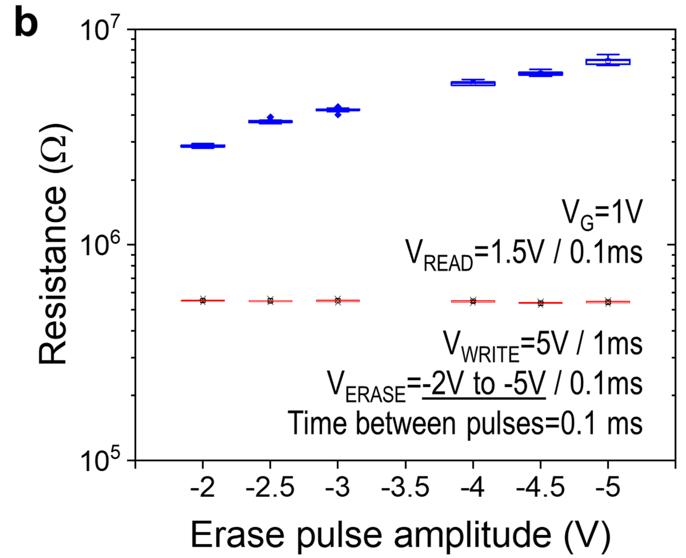


Extended Data Fig. 4 | Electrical characterization of multiple h-BN memristors. Current versus voltage plots collected for different 1T1M cells using Au–Ti–h-BN–W memristors with a lateral size of 0.053 μm². The plots have been drawn with 75% transparency to emphasize the general trend. Each plot contains 50 cycles or more. The first ramped voltage stress is shown in blue colour; it is normally more insulating than the rest, indicating the need of an activation step that softly degrades the h-BN and produces RS. However, no filament is completely formed (see Supplementary Note 1).

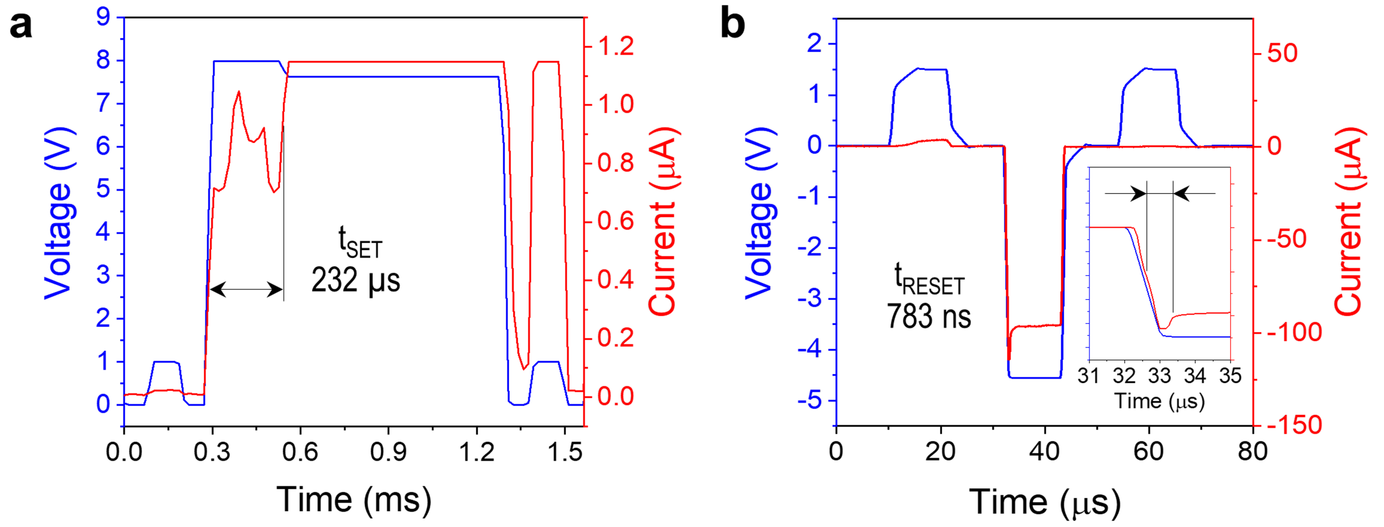


Extended Data Fig. 5 | Multilevel RS in 1T1M cells. Current versus time plots for a 1T1M cell when using different end voltage of the reset RVS. Multiple stable conductance levels can be programmed. The composition of the top electrode is Au–Ti. The read voltage applied is 0.1 V. We measured the retention of several states up to 7 h, and the rest up to 200 s due to laboratory availability. The devices show no sign of conductance degradation for any measured state after 7 h, although typical drift (less than 5%) is detected, which is acceptable. According to the Semiconductor Research Corporation [V. Zhirnov, Decadal Plan for Semiconductors: New Trajectories for Memory and Storage, presented in the 2022 Non-Volatile Memory Technology Symposium 2022, December 7th-11, Stanford, USA], the unsaved bytes turnover and lifetime for multiple memory applications is currently less than 15 h, meaning that no retention time above 10 years will be needed. Similarly, the requirements in terms of retention time for memristive electronic synapses is much more relaxed than in most electronic memories, and correct functioning with short retention times of few hours has been claimed [see Kuzum, D., Yu, S., & Wong, H. P. Synaptic electronics: materials, devices and applications. Nanotechnology 24, 382001 (2013)].

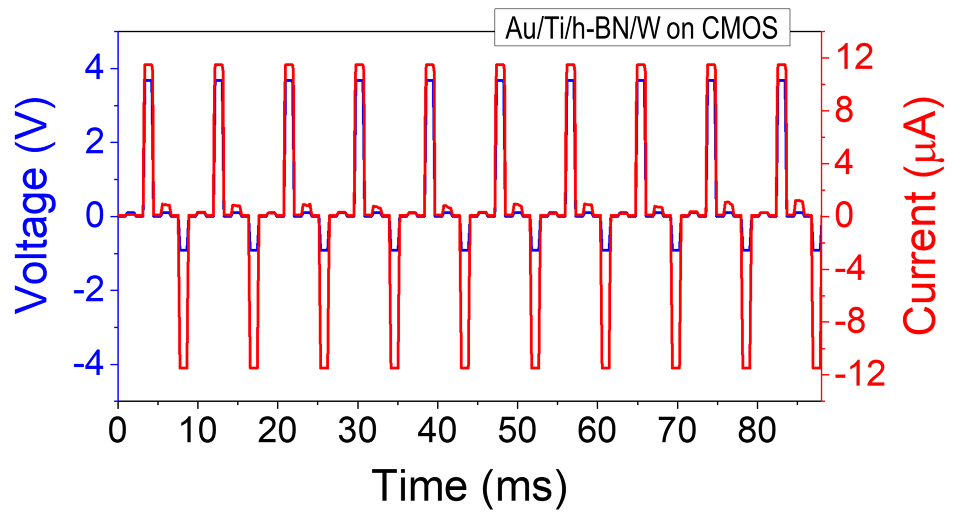


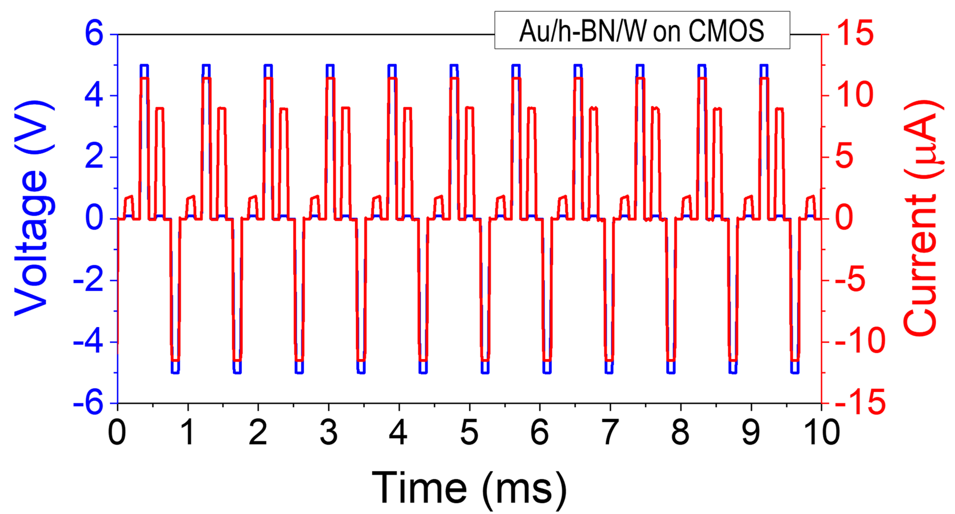


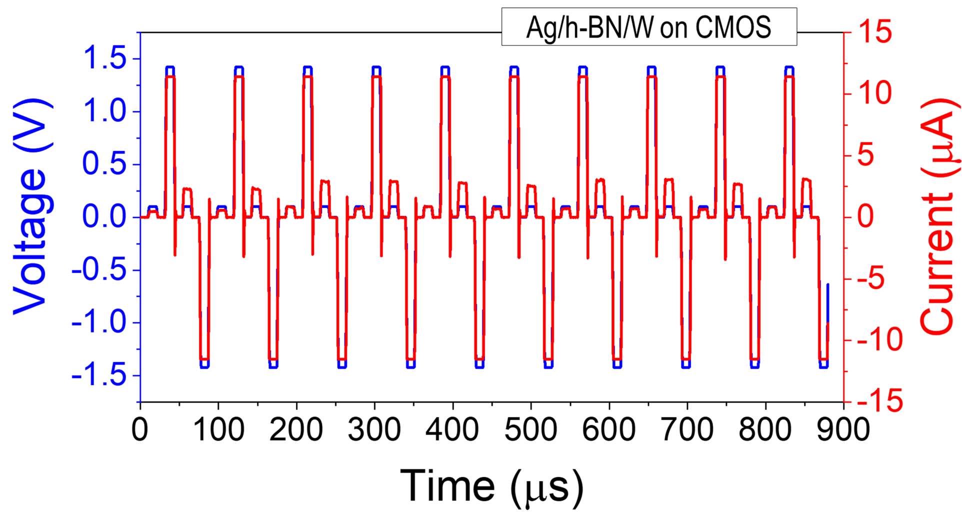
Extended Data Fig. 6 | Adjustment of R_{LRS}/R_{HRS} in 1T1M cells with Au–Ti top electrodes. Values of R_{HRS} and R_{LRS} depending on the pulse amplitude of write and erase operation. The value of the R_{LRS}/R_{HRS} window can also be adjusted in this way.



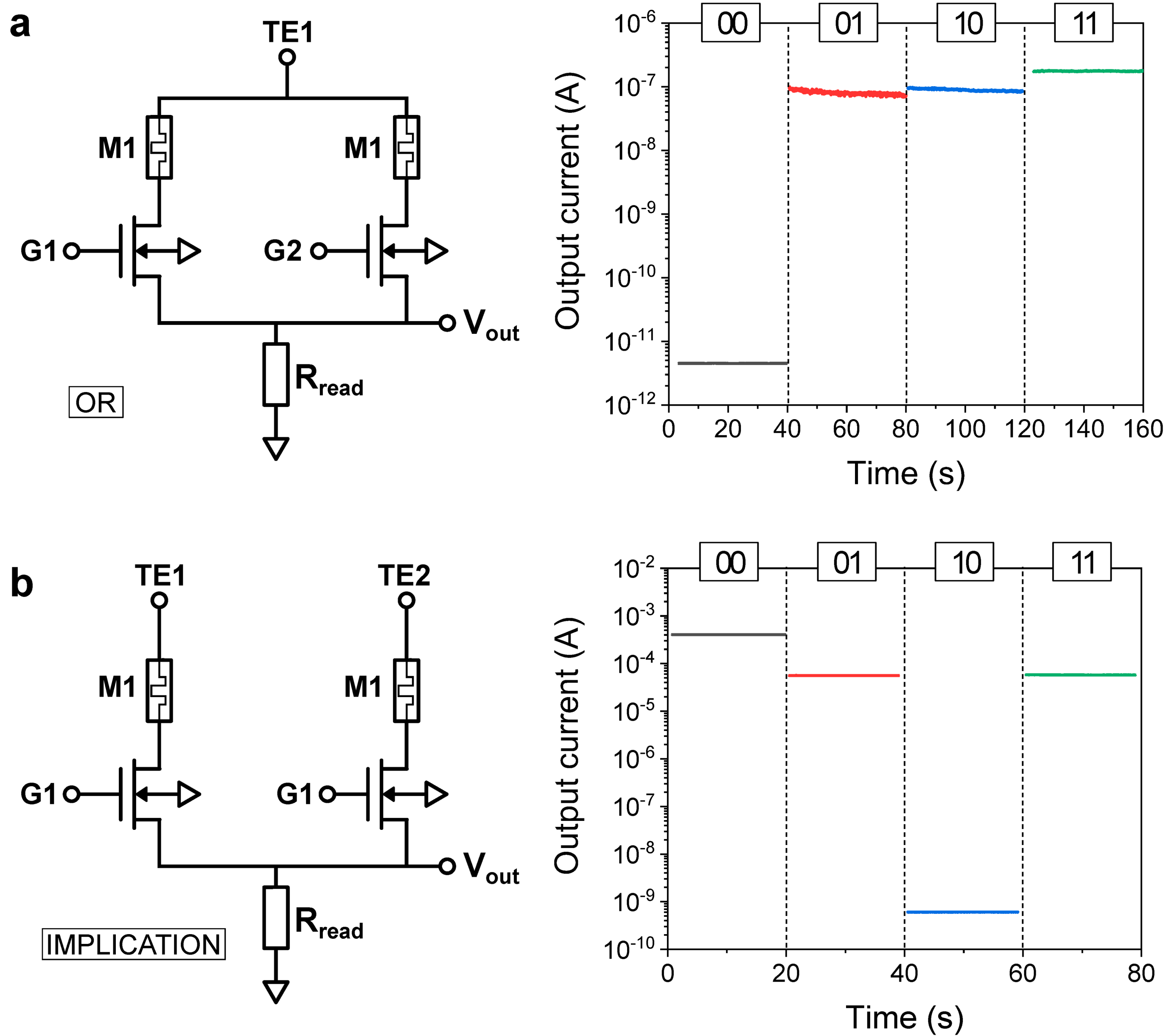
Extended Data Fig. 7 | Switching time of 1T1M cells using top Au–Ti electrodes. Voltage and current versus time plots showing the non-volatile bipolar RS with high temporal resolution. The switching time for the set transition is 232 μs, and the one for the reset transition is 783 ns.



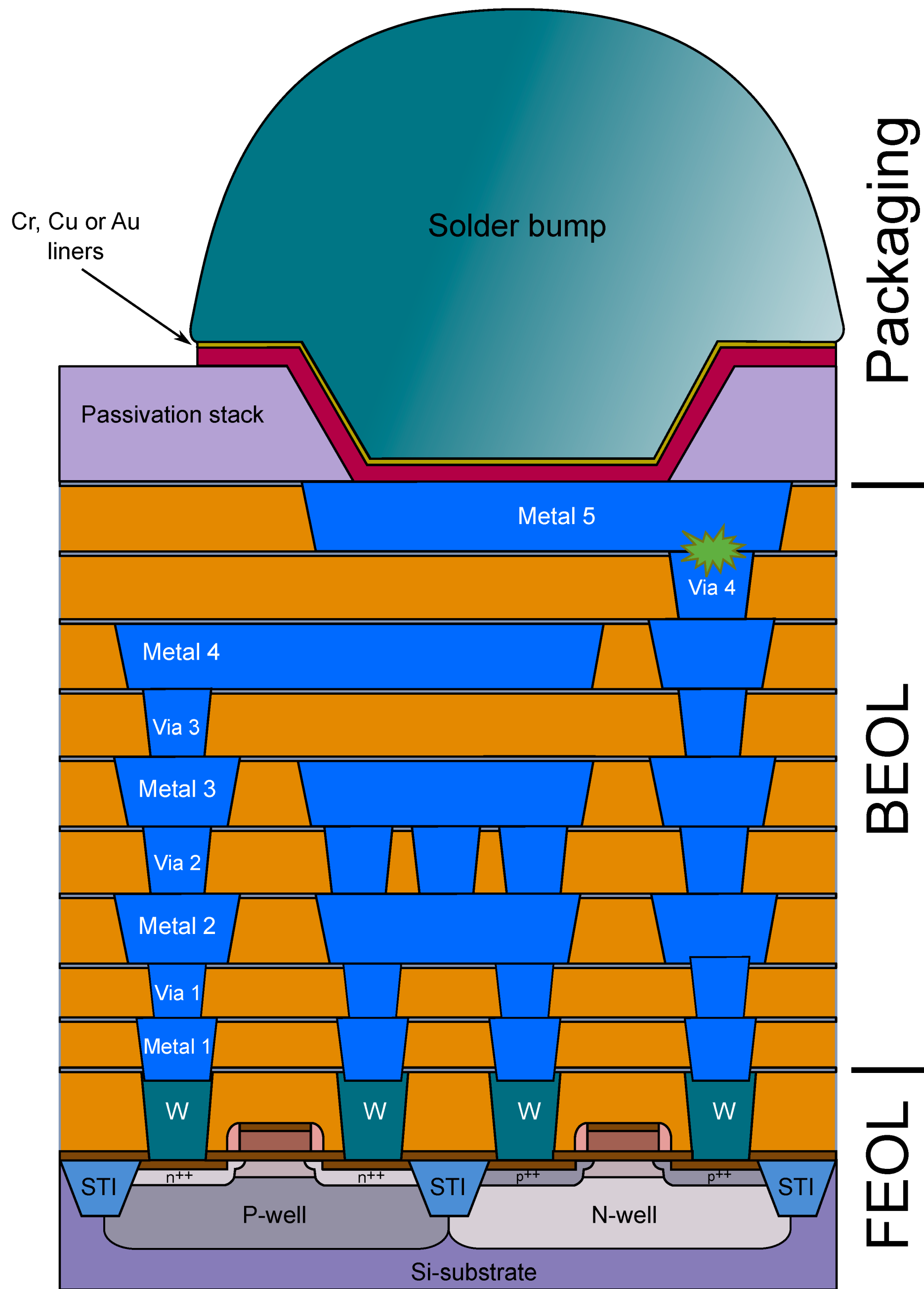




Extended Data Fig. 8 | Non-volatile bipolar RS in the 1T1M cells depending on the chemical composition of the top electrode. Voltage and current versus time plots showing reproducible non-volatile bipolar RS in devices with Au–Ti, Au and Ag top electrode. The operation voltages and switching times are different, providing enough flexibility to operate in different regimes.



Extended Data Fig. 9 | In-memory computation with crossbar arrays of h-BN memristors. a,**b**, Schematic and current versus time plot for "or" and "implication" logic operations (respectively) made with the crossbar array of 1T1M cells in the hybrid 2D/CMOS microchip. The value of the output current changes depending on the conductance of the memristors M1 and M2, which is indicated as "00", "01", "10" and "11" on each segment of the plot ("0" and "1" means low and high conductance, respectively).



Extended Data Fig. 10 | Non-exhaustive list of materials for microchip manufacturing. Schematic representation of a microchip; the position of our memristor is indicated with a green star. Our memristors are integrated in the last metallization layer of the BEOL, where Au liners are often employed, and very far from the transistors in the FEOL. Apart from this, Au wire bonding to the socket is also often employed. Hence, the use of Au electrodes of the memristors is not a concern for this type of devices/circuits. STI means shallow trench isolation.