

Supporting Information

Hyperplane Tree-based Data Mining with Multi-functional Memristive Crossbar Array

Sunwoo Cheong^{1,†}, Dong Hoon Shin^{1,†}, Soo Hyung Lee¹, Yoon Ho Jang¹, Janguk Han¹, Sung Keun Shim¹, Joon-Kyu Han², Néstor Ghenzi^{1,3,}, and Cheol Seong Hwang^{1,*}*

¹Department of Materials Science and Engineering and Inter-university Semiconductor Research Center, College of Engineering, Seoul National University, Seoul, 08826, Republic of Korea

²System Semiconductor Engineering and Department of Electronic Engineering, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 04107, Republic of Korea

³Universidad de Avellaneda UNDAV and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Mario Bravo 1460, Avellaneda, Buenos Aires 1872, Argentina

* Correspondence: N.G. (e-mail: nghenzi@snu.ac.kr), C.S.H. (e-mail: cheolsh@snu.ac.kr)

Table of contents

- Supporting Figures S1-S30

- Supporting Notes 1-4

- Supporting Tables S1-3

- Supporting References

Supporting Table S1. Table of abbreviations and their meanings of the main text

| Abbreviation | Definition |
|---------------------|---|
| OD | Outlier detection |
| DC | Data clustering |
| mf-CBA | Multi-functional crossbar array |
| BC | Binary code |
| <i>M</i> | Hyperparameter for minority rate |
| MOD | Minority-based outlier detection |
| MBC | Minority binary code |
| HD | Hamming distance |
| X & Y | Do not care bit |
| T | Number of trees |
| H | Number of hyperplanes in each tree |
| <i>R</i> | Hyperparameter for outlier determination |
| OCV | Outlier count vector |
| K | Number of centroids in K-means clustering |
| CBC | Centroid binary code |

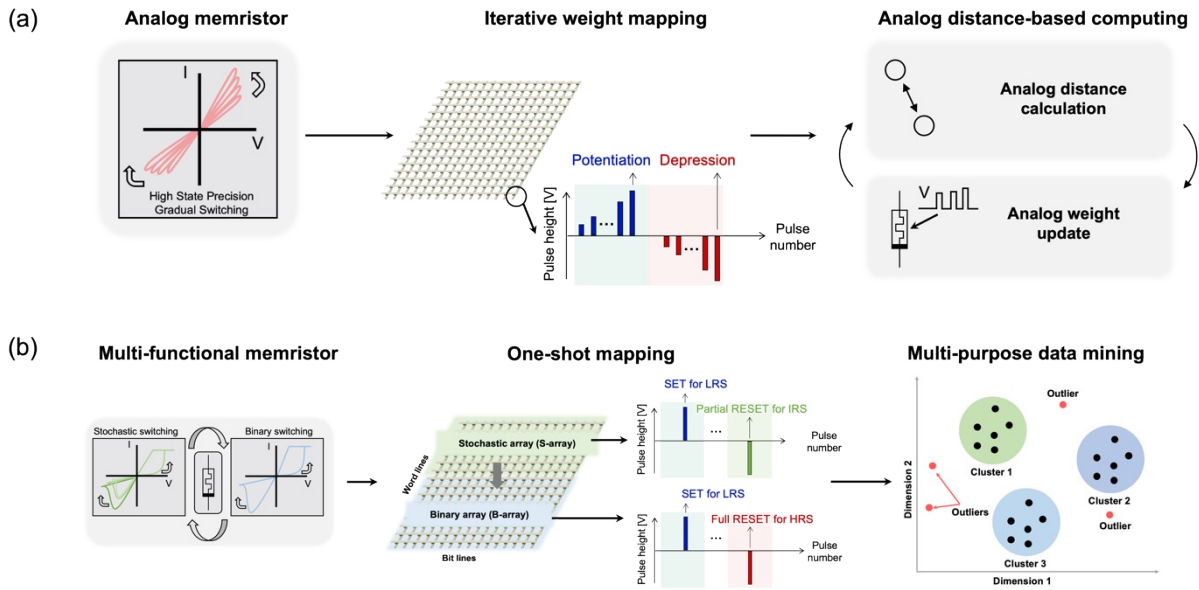


Figure S1. Analog domain VMM and proposed multi-functional CBA. (a) Conventional analog memristor-based VMM operation. For parallel computing, the state-tunable analog characteristics of the memristor are utilized to represent the weight of neural network or analog distances between data points. In this domain, the iterative pulse of weight potentiation/depression or incremental step pulse programming is applied to each cell in CBA, increasing programming time and energy. (b) Schematic design of the proposed mf-CBA for implementing outlier detection (OD) and data clustering (DC) algorithms. By changing the switching mode of the multi-functional memristor in arrays, the proposed mf-CBA could be reconfigured for stochastic or binary switching operations without additional computational overhead. The mf-CBA can be divided into two sub-arrays: (i) the S-array and (ii) the B-array. The mf-CBA configuration can be efficiently programmed using one-shot pulse mapping. In stochastic mode, memristors in the LRS state are subjected to pulses of intermediate voltage to result in stochastic IRS. In contrast, binary mapping is achievable in binary mode by applying set and reset pulses to both LRS and HRS states, respectively. These reconfigurable array mapping can implement the conceptual hyperplane and tree structure¹⁻³ toward the multi-purpose data mining hardware capable of simultaneously executing OD and DC.

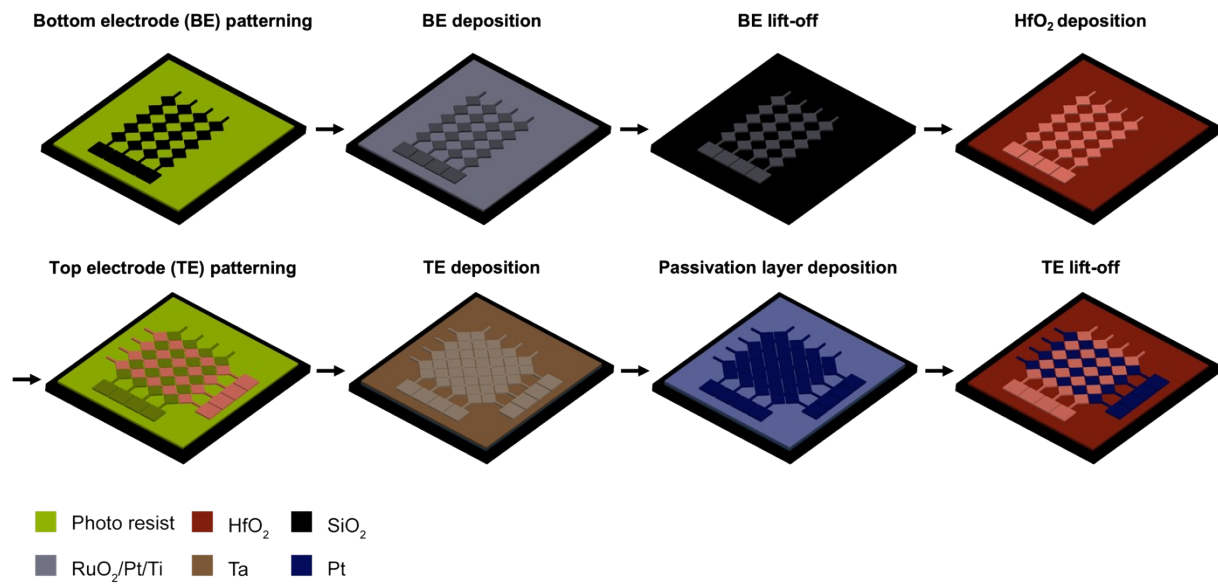


Figure S2. The fabrication process of the THR memristor-based multi-functional crossbar array.

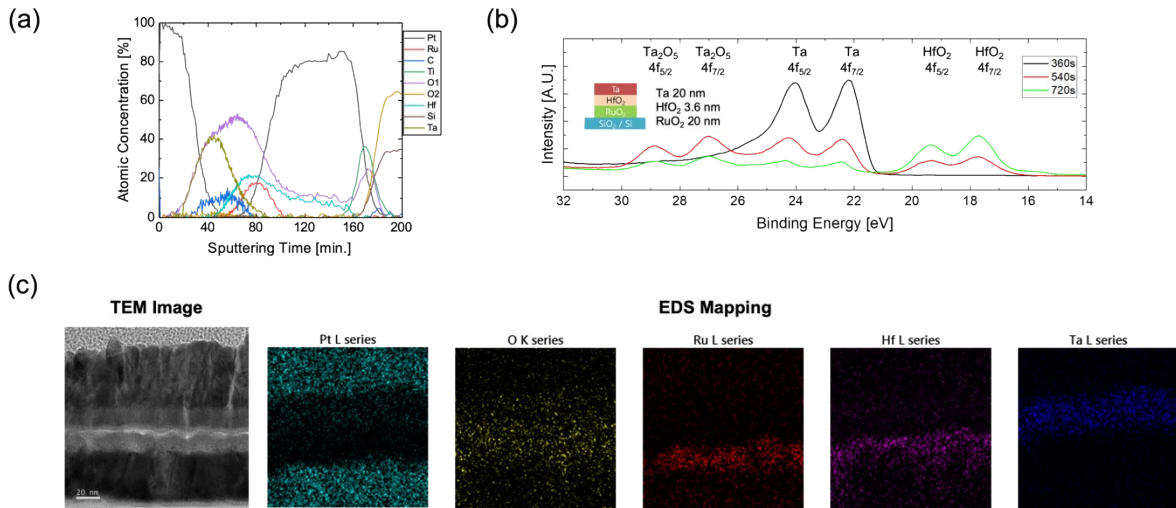


Figure S3. Structural and elemental analysis of the THR memristor. (a) Depth profile of the THR device through Auger electron spectroscopy. Elements of Ta, Hf, Ru, and oxygen concentrations are demonstrated. (b) X-ray photoelectron spectroscopy of the stacked film with different etching times. The emergence of the primary peak shifts gradually from metal Ta (22 to 24 eV) to Ta₂O₅ (27 to 29 eV) and subsequently to HfO₂ (16 to 18 eV), with the sputtering time increment.⁴⁻⁷ (c) EDS element mapping (Ta, Hf, O, Pt, and Ru) results obtained by cross-sectional TEM.

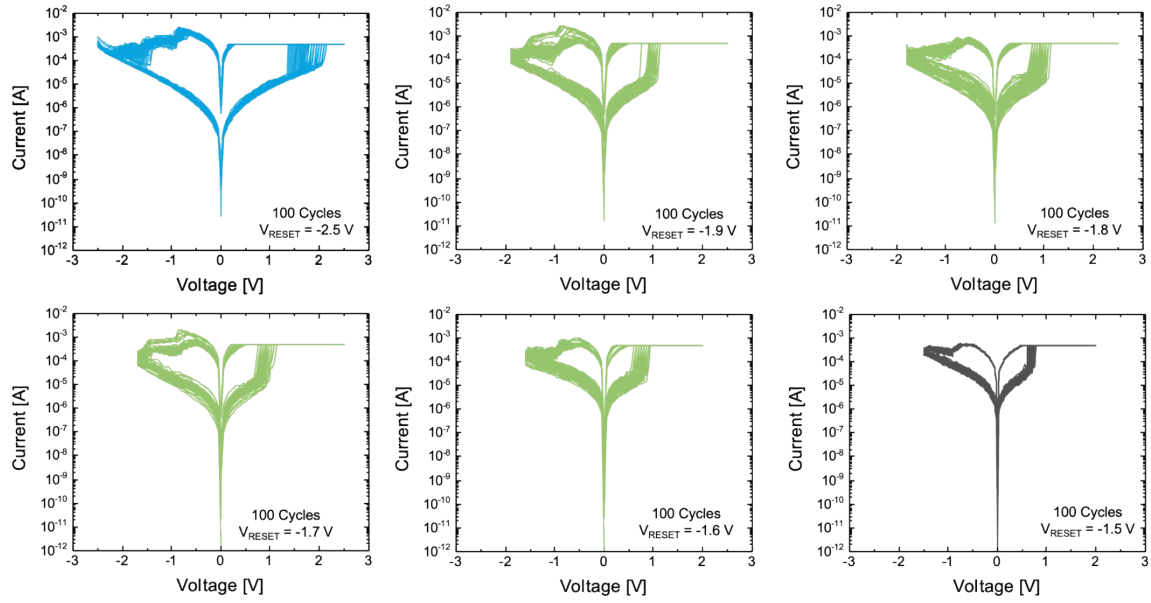


Figure S4. V_{RESET} -dependent I-V switching characteristics. When $V_{\text{RESET}} = -2.5$ V, stable binary switching of low state dispersion is observed, whereas significant resistance distribution in the IRS can be observed at V_{RESET} ranging from -1.6 V to -1.9 V. These V_{RESET} -dependent resistive switching stems from the two-step reset behavior of the THR memristor.⁸

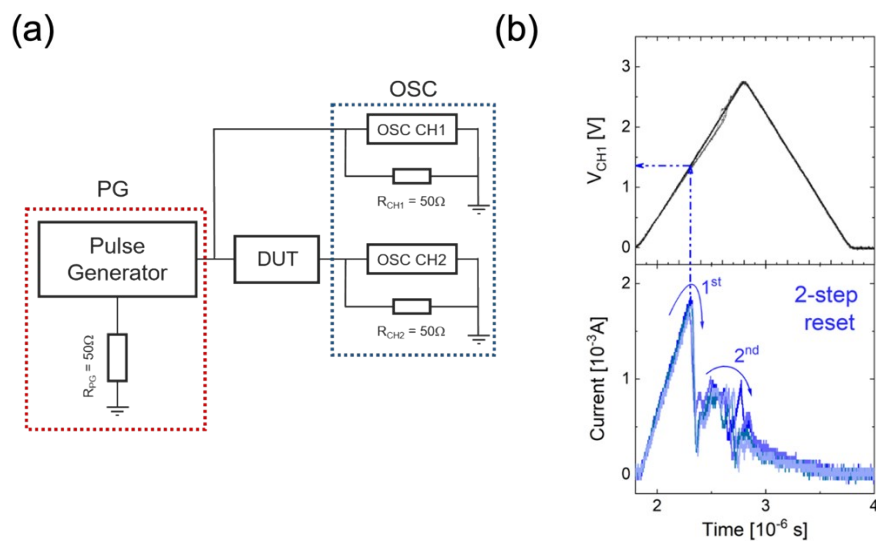


Figure S5. Reset behavior with a triangular reset pulse. (a) AC triangular pulse configuration to measure input voltage and current through the THR device. (b) μ s-scale triangular voltage pulse to show the reproducibility of the 2-step reset process. The stop voltage can control the mode of operation between stochastic and binary modes.

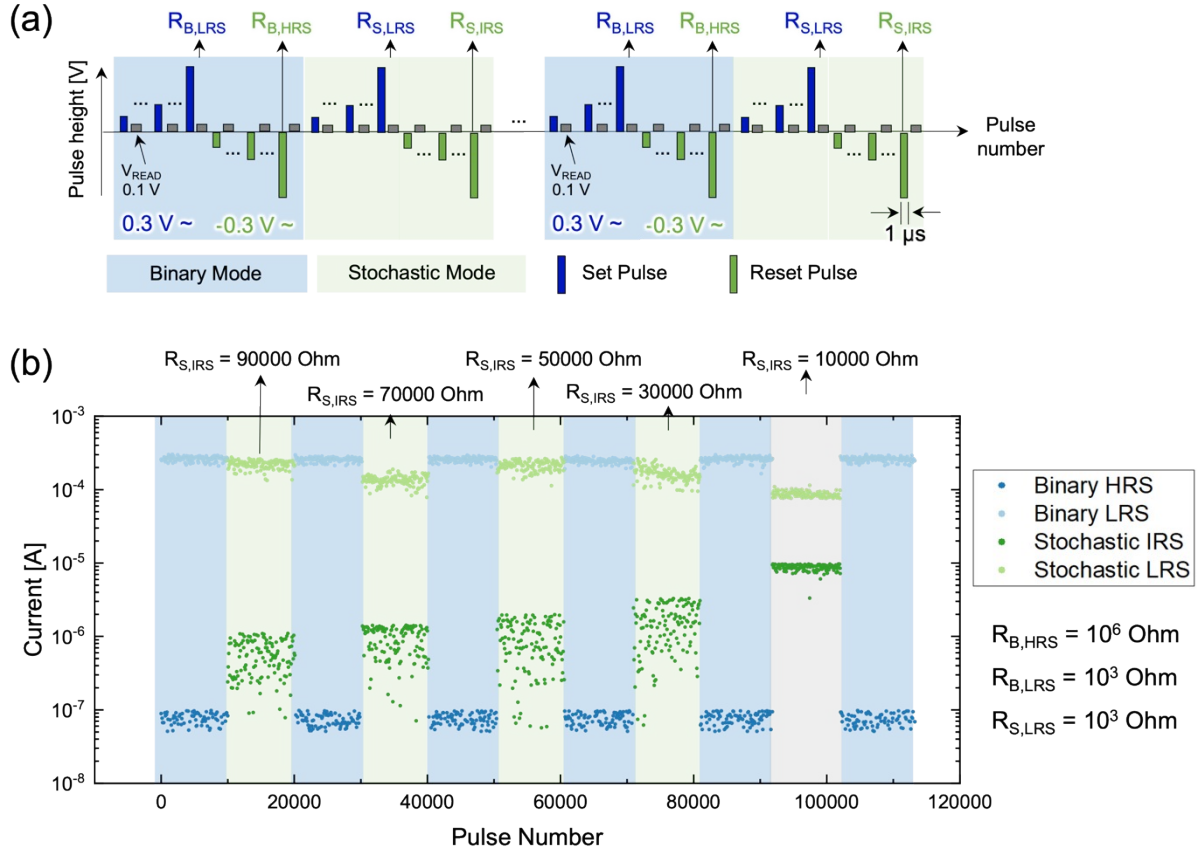


Figure S6. Pulse-driven mode reconfiguration. Pulsing application method and results of mode reconfigurations. (a) Pulsing scheme for multi-functional mode reconfiguration, where each binary mode and stochastic mode is repeated for 10^4 cycles. In binary mode, incremental pulses are applied to cycle between $R_{B,LRS}$ (10^3 Ohm) and $R_{B,HRS}$ (10^6 Ohm), while in stochastic mode, pulses cycle between $R_{S,LRS}$ (10^3 Ohm) and $R_{S,IRS}$. To demonstrate the stochastic resistive dispersion, the value of $R_{S,IRS}$ is varied from 90,000 Ohm to 10,000 Ohm with 20,000 Ohm interval after every 10^4 cycles. (b) The read current values ($V_{READ} = 0.1 V$) for HRS, IRS, and LRS. In binary mode, due to the abrupt set behavior, reached $R_{B,LRS}$ is measured as a value smaller than the predetermined 10^3 Ohms, exhibiting a memory window greater than 10^3 . In the case of $R_{S,IRS} = 10,000$ Ohm, the THR memristor resets only in the gradual reset region, which corresponds to $V_{RESET} = -1.5 V$ case in Figure S4, showing the small distribution of IRS states.

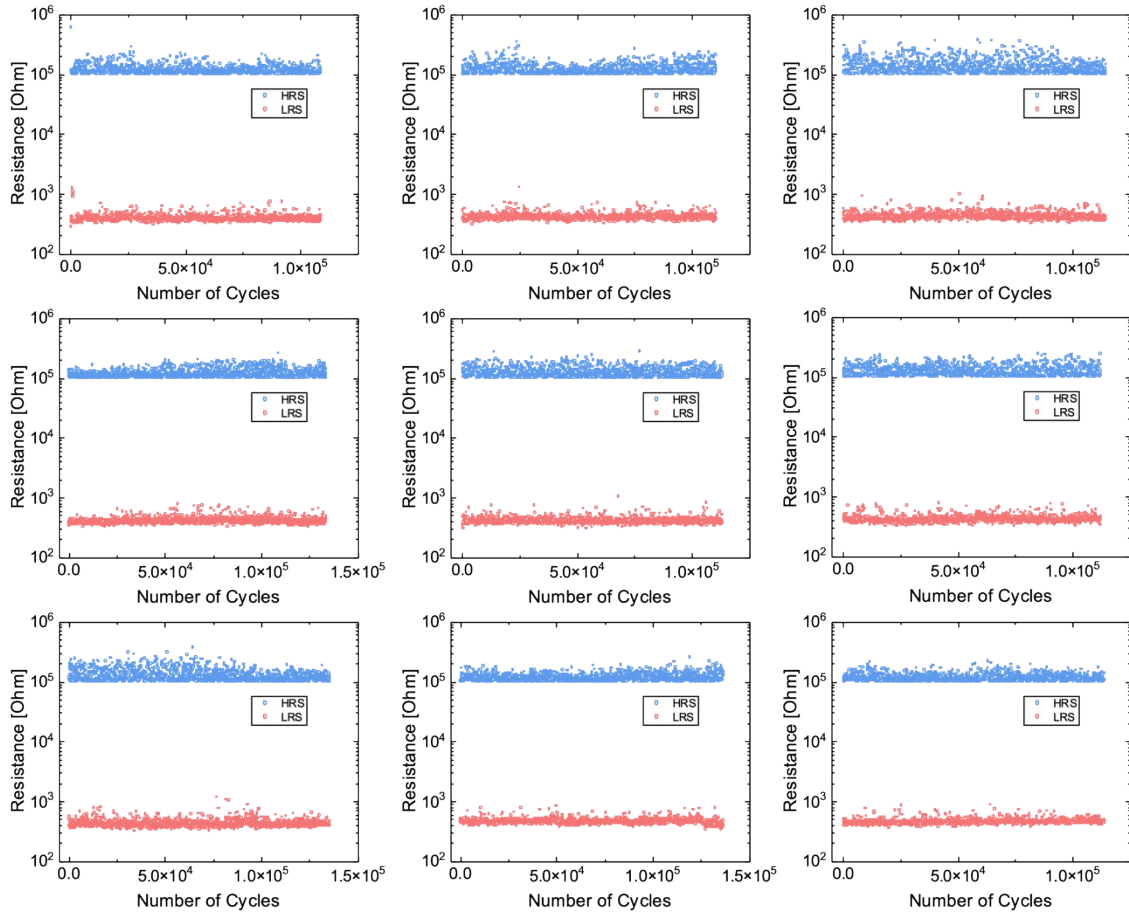


Figure S7. Device switching endurance of the THR memristor. Endurance tests were performed by a custom-built FPGA board (Experimental Section), assigning the target HRS resistance of 10^5 Ohm and the LRS resistance of 10^3 Ohm. All resistance values are calculated with output current at read voltage of 0.1 V. All devices show $> 10^5$ switching cycles without degradation or dielectric breakdown.

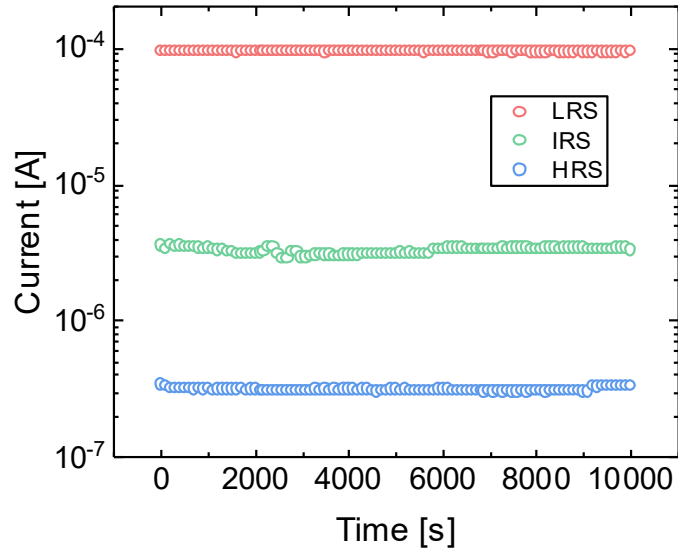


Figure S8. Retention characteristics of THR memristor at the temperature of 85 °C.

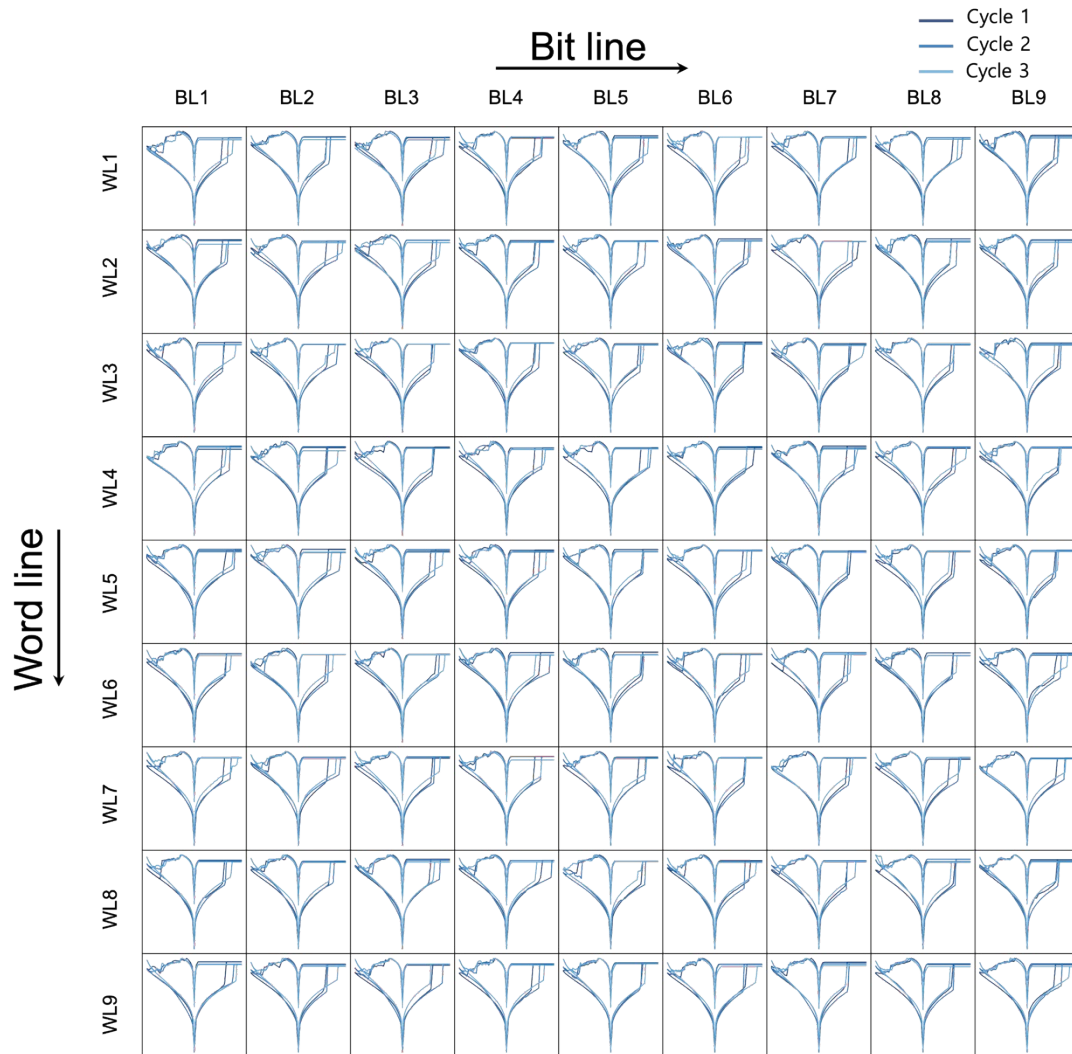


Figure S9. Binary mode I-V characteristics of 81 cells in 9 by 9 mf-CBA. Three separate I-V curves demonstrate the intra-array reproducibility of binary mode operation.

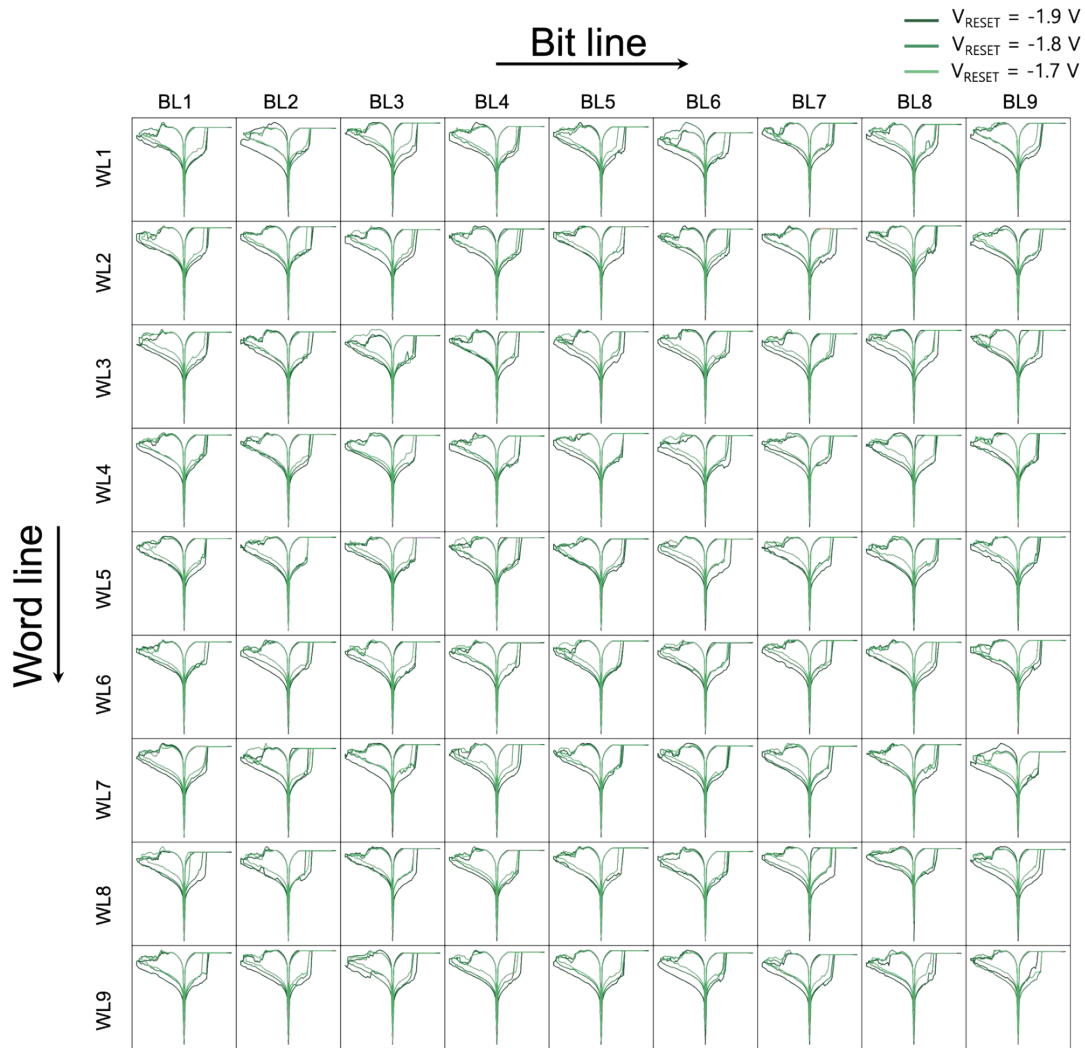


Figure S10. Stochastic mode I-V characteristics of 81 cells in 9 by 9 mf-CBA. Three different V_{RESET} were applied to demonstrate the intra-array reproducibility of stochastic mode operation.

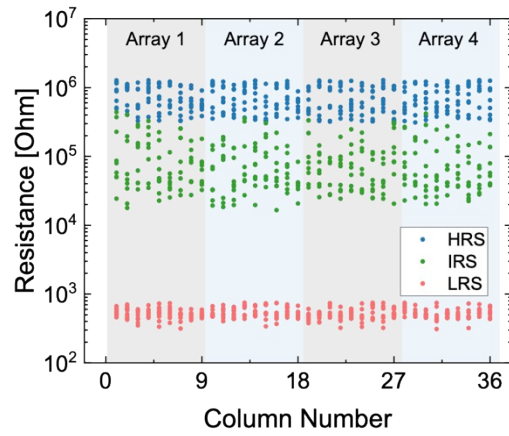


Figure S11. Inter-array variation of THR mf-CBA. HRS (blue), IRS (green), and LRS (red) resistance values were obtained from the four arrays, with 9 devices per column, showing consistent average resistance values and distributions across the inter-array scale.

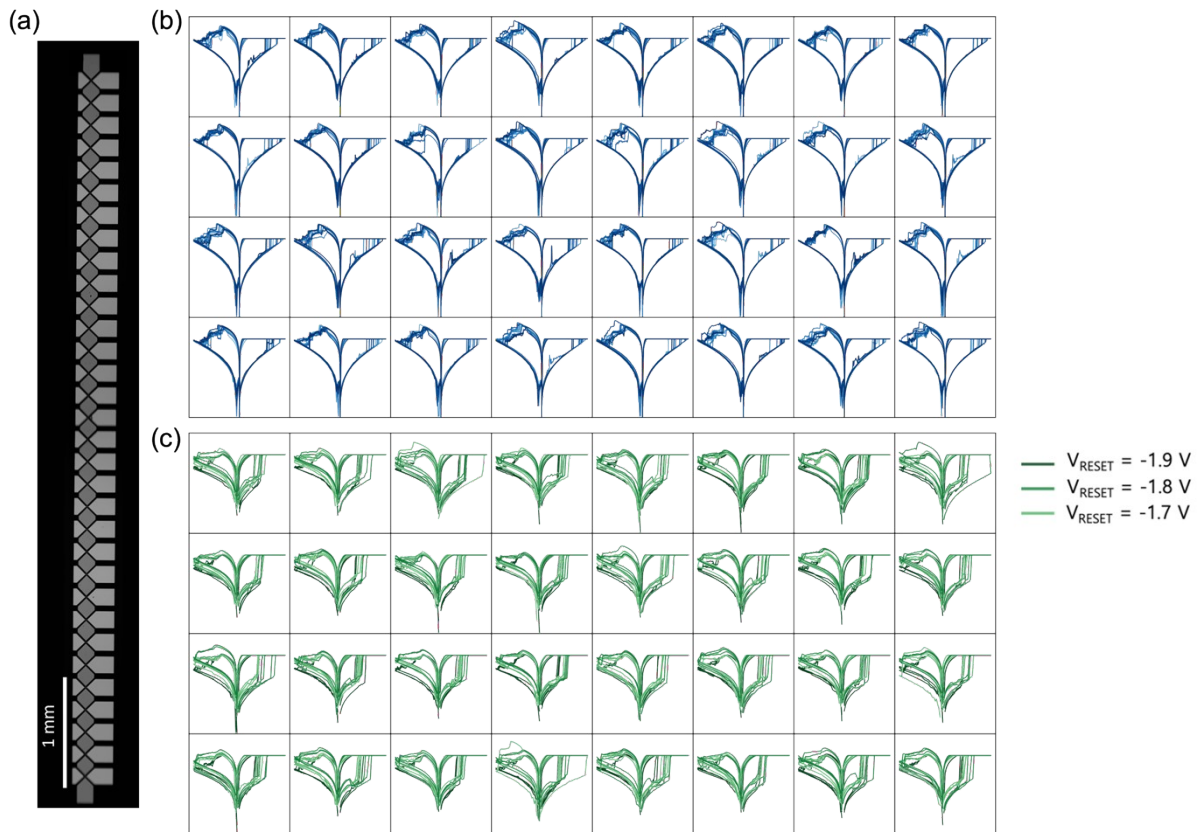


Figure S12. Binary and stochastic mode I-V characteristics of 32 cells in 32 by 1 THR array. (a) Optical microscope image of THR array. (b) Binary mode operation of ten I-V cycles. (c) Stochastic mode operation of three I-V cycles with different V_{RESET} .

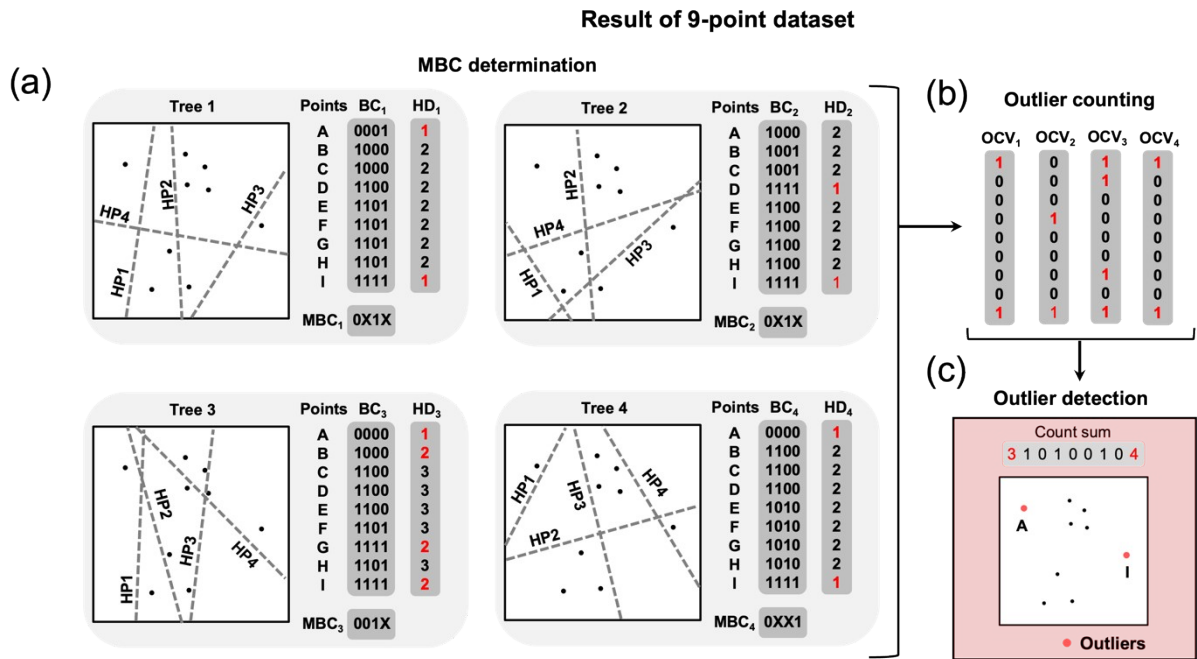


Figure S13. Schematic diagram of minority-based MOD process of 9-point dataset. (a) Randomly generated hyperplane tree structure and corresponding binary codes. Figure shows the calculated binary states of every data point using the MOD method of the main text. MBCs of each tree are determined with hyperparameter M of 25 %, and HDs of each tree are calculated between BCs and the MBC. For the outlier counting, hyperparameter R was settled to 25 %, which means that 2 with the lowest HD value among the 9 ($2/9 = 22 \% \sim 25\%$) components of the HD vector are supposedly assigned to be the outlier candidates. For trees #1, #2, and #4, there are two 1 and seven 2 HD components, so the data points with HD of 1 are assigned to be the outlier candidates, indicated by the red color in the corresponding HD vector. However, in the case of tree #3, the HD values are 1, 2, and 3, so the components with HD of 1 and 2 must be assigned to the outlier candidates. Here, four components have HD of 1 and 2 values, so all these four components are assigned to the outlier candidates. (b) Conversion of HD vectors into OCVs with the given hyperplane tree structure by changing the red and black colored values to 1 and 0, following the outlier count rule (the lower panel of Figure 3d of main text). (c) Final outlier detection results of the 9-point dataset. The two data points with the two highest count sum values (A with 3 and I with 4) are determined to be outliers, as highlighted in red. These steps complete the MOD process within the 9-point dataset.

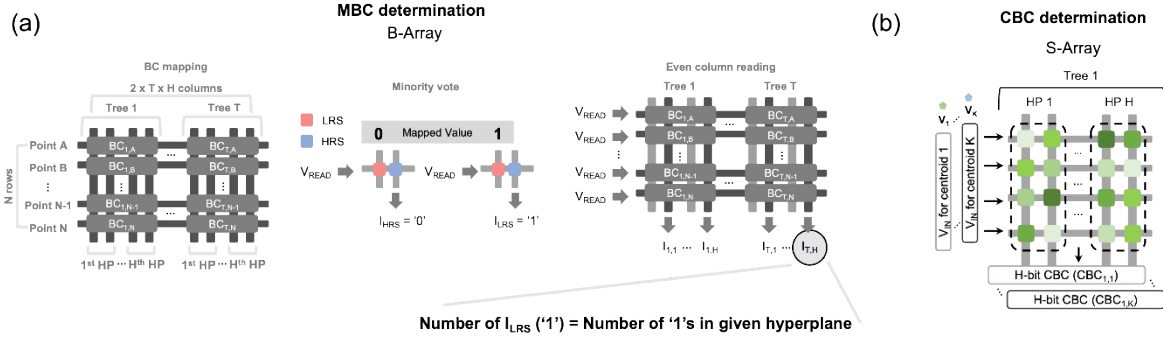


Figure S14. MBC and CBC determination process. (a) MBC requires the minority vote rule process based on the binary states of the dataset. Therefore, B-Array calculates the number of '1's of binary states based on a single hyperplane. In the BC-mapped B-array (left panel), applying V_{READ} to a row (word line) and current readout through the right column (bit line) corresponds to the direct binary value reading of the mapped value (middle panel). Applying V_{READ} to all word lines and grounding the even-numbered bit lines results in the current output, including the number of "1" for each hyperplane (column). Based on these current values with the ratio of I_{HRS} and I_{LRS} values, the MBC vectors are acquired in the software following the methods described in Figure 3 of the main text. (b) CBC introduces centroid into the given dataset, and the CBC can be determined with S-Array. To determine CBC, voltage inputs of randomly generated centroids are applied to the S-array. The S-array holds an identical hyperplane tree structure, used to calculate BCs.

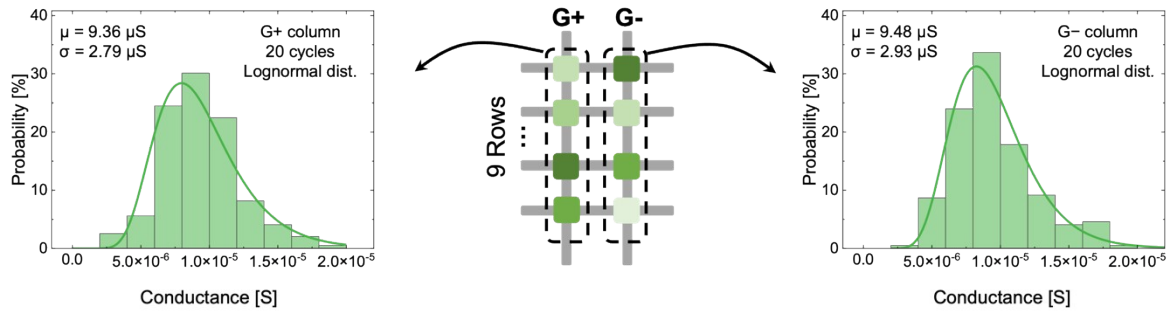


Figure S15. Single-column conductance distribution of S-array. When examining the conductance distribution for each of the nine cells comprising the columns representing G+ and G- values over 20 cycles, it is observed that both G+ and G- columns exhibit lognormal distributions, indicating similar distributions in terms of mean and standard deviation. Therefore, their difference comprises the Gaussian distribution of differential conductance shown in Figure 4d of the main text.

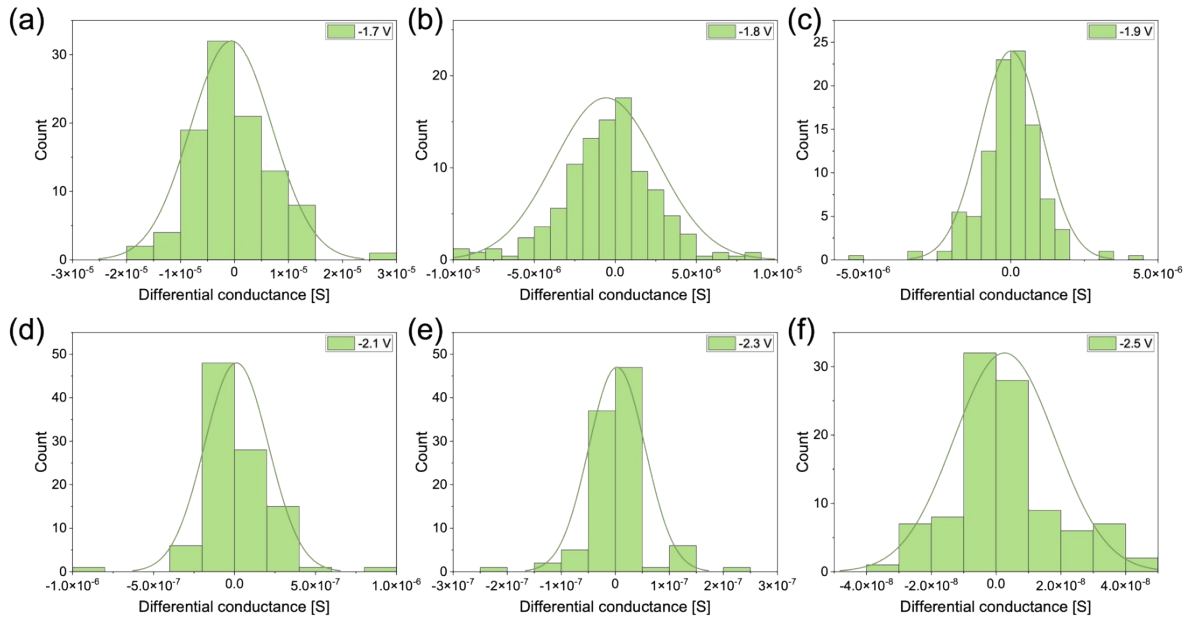


Figure S16. Histogram of 100 differential conductance trials at different V_{RESET} values. (a) $V_{\text{RESET}} = -1.7$ V. (b) $V_{\text{RESET}} = -1.8$ V. (c) $V_{\text{RESET}} = -1.9$ V. (d) $V_{\text{RESET}} = -2.1$ V. (e) $V_{\text{RESET}} = -2.3$ V. (f) $V_{\text{RESET}} = -2.5$ V.

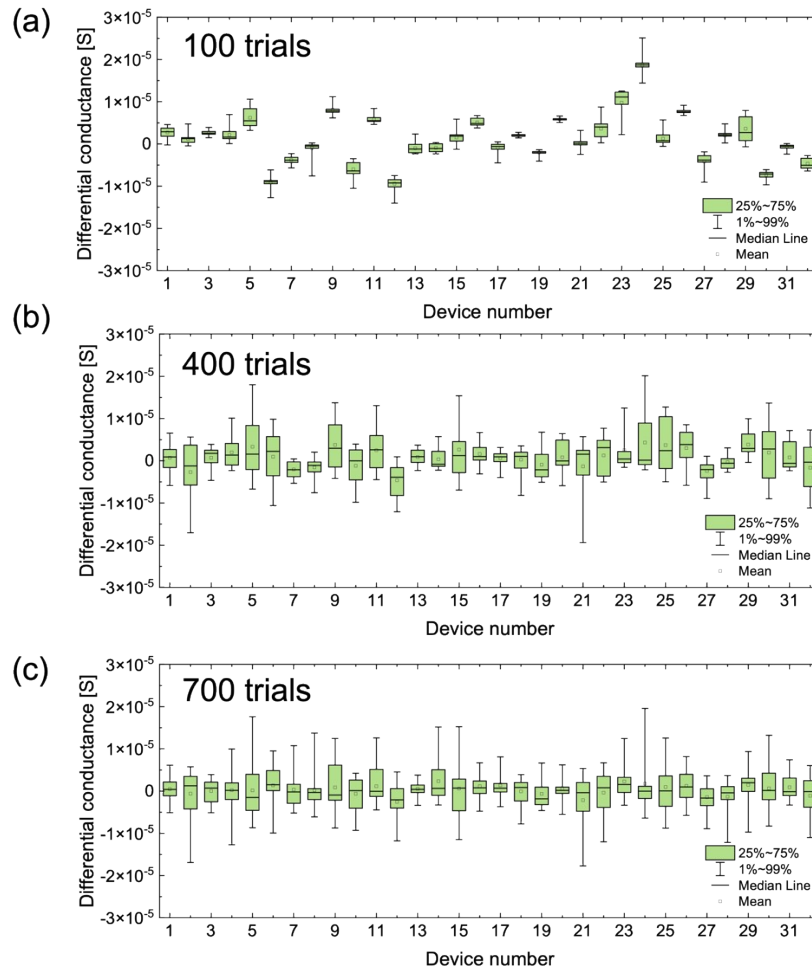


Figure S17. Differential conductance mapping distribution within a 32 by 1 THR array. Each device is reset to $V_{\text{RESET}} = -1.8 \text{ V}$ for every trial, and the differential conductance is calculated based on the difference in conductance mapped for each trial (state read at 0.1 V). (a) Results for 100 trials. (b) Results for 400 trials. (c) Results for 700 trials. As the number of trials increases, the distribution of differential conductance among devices becomes more uniform.

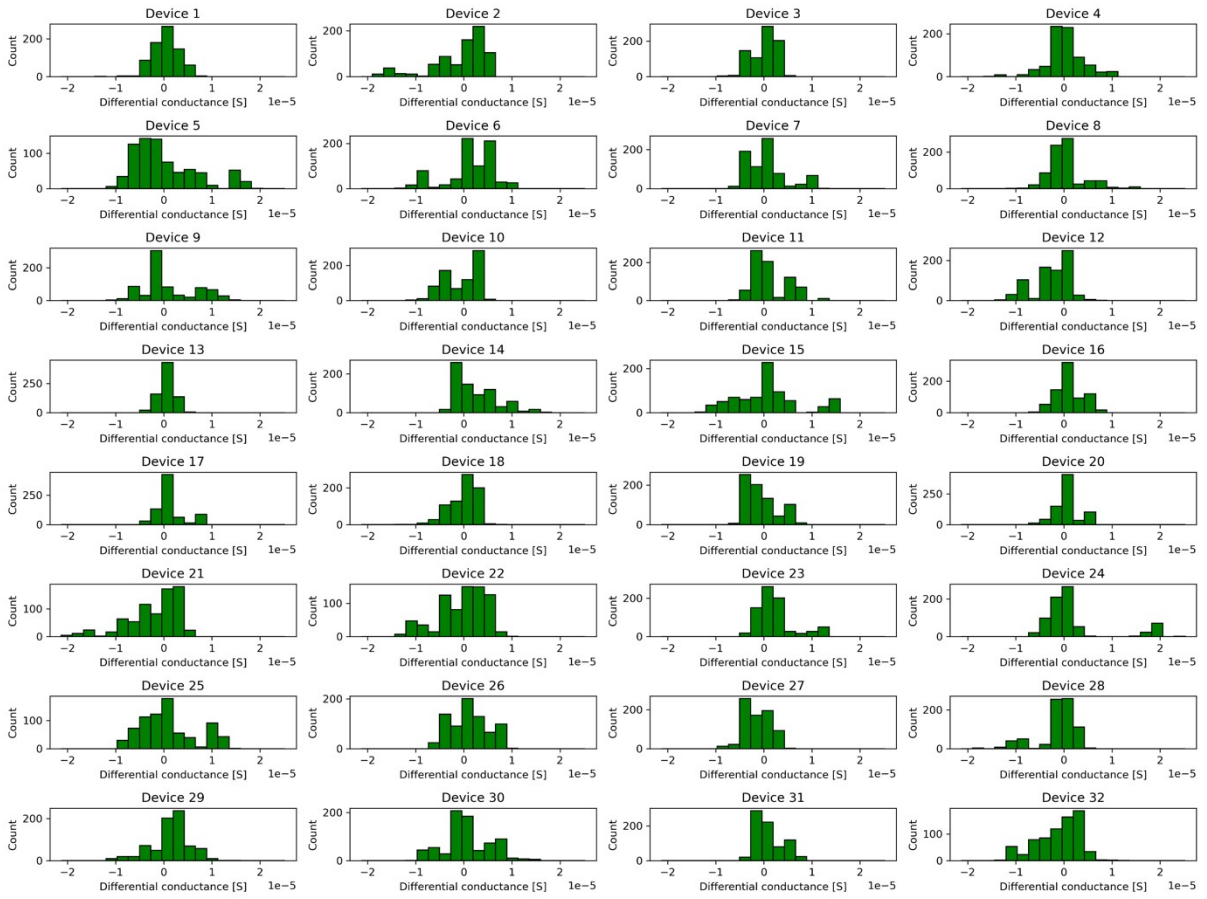


Figure S18. Histogram of differential conductances within a 32 by 1 THR array.

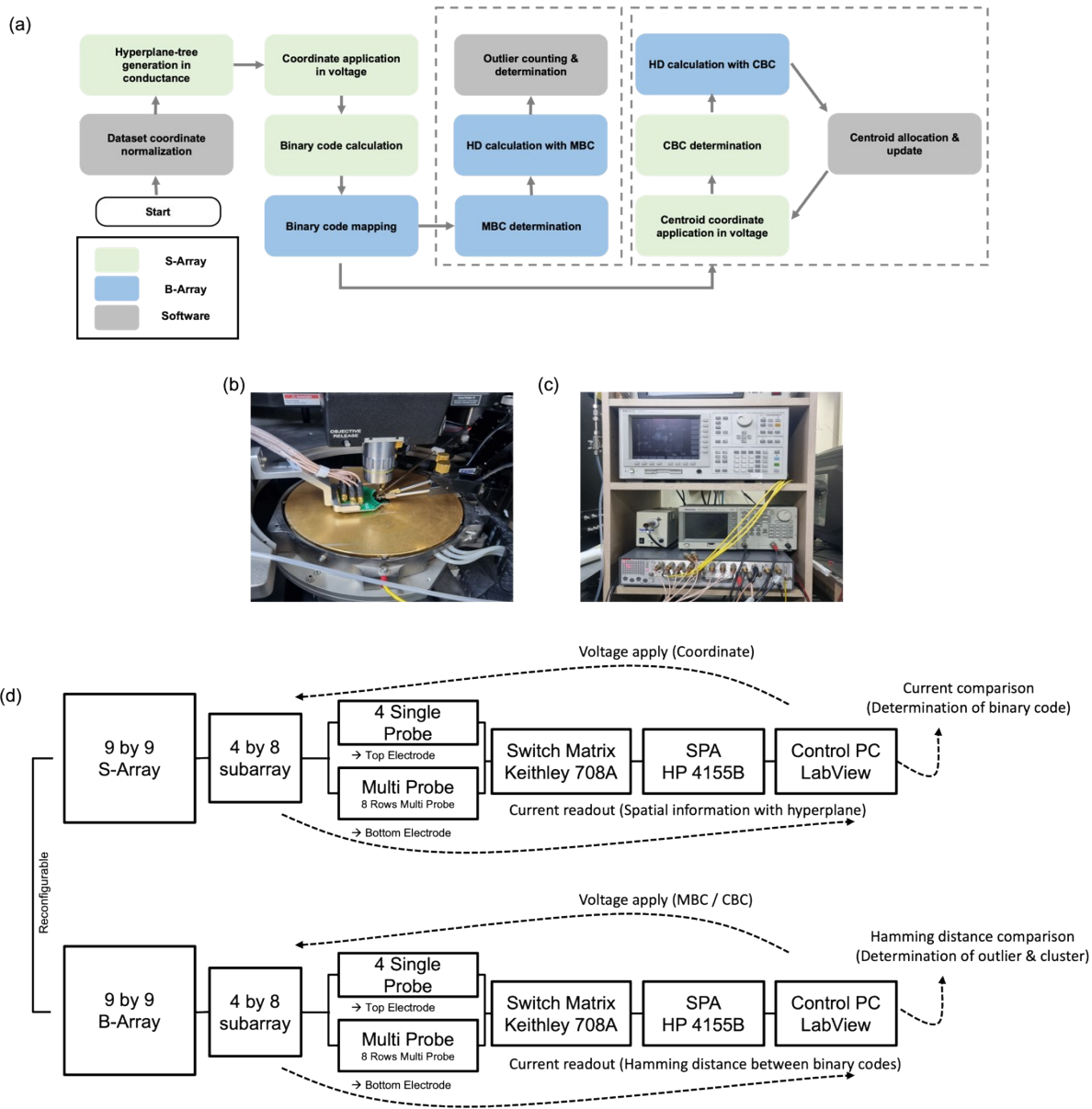


Figure S19. Data mining process flowchart and circuit block diagram of array measurement system. (a) Flow chart of proposed data mining method. (b) Photograph of the multi-probe. (c) Photograph of SPA and switch matrix. (d) 9 by 9 array measurement system.

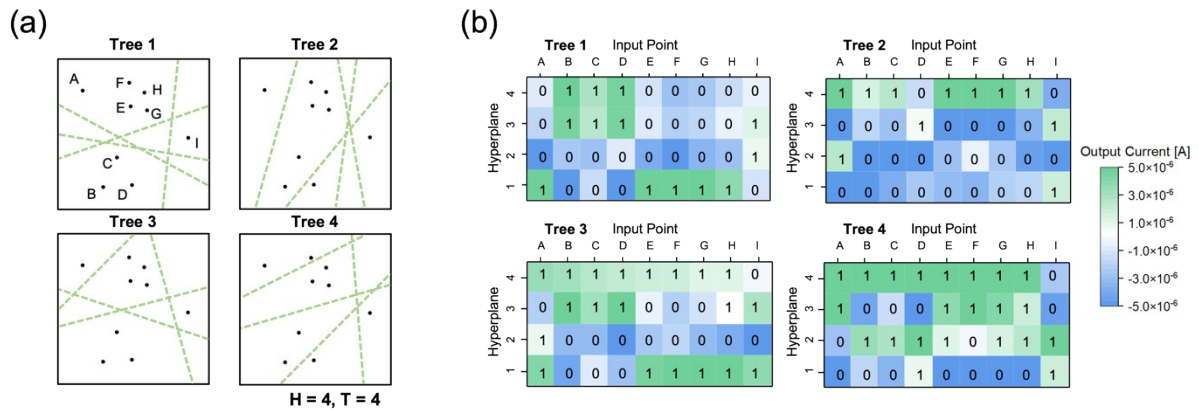


Figure S20. Generated hyperplanes and S-array current output. (a) Generated random 16 hyperplanes in four trees. (b) Measured current output by applying a voltage of point coordinates in the S-array corresponding to each tree. The sign of the output current, represented in green and blue, can identify the relative positional relationships of 9 points regarding a single hyperplane.

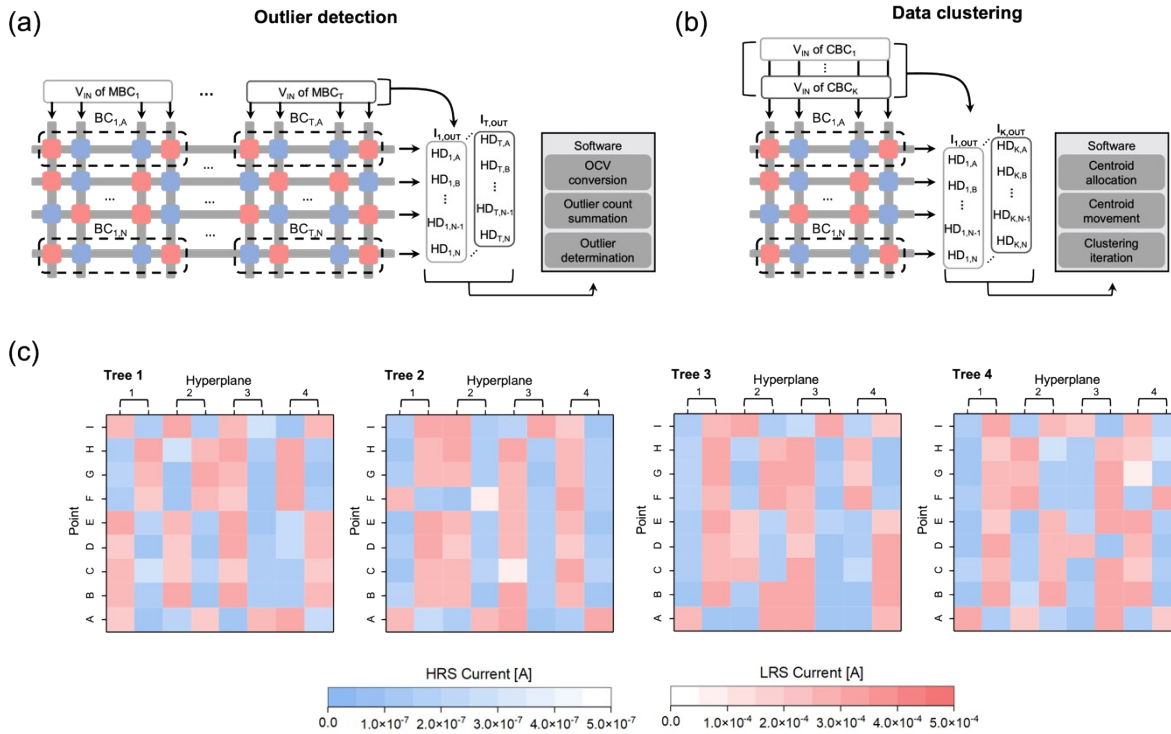


Figure S21. B-array operation schematic and output current of binary code-mapped B-array with 9-point dataset. (a) For OD, MBCs are applied as a voltage to bit lines to calculate HD between mapped BCs and applied MBCs as a current output in word lines. These calculated HDs are handled in software to detect outliers. (b) For DC, CBC calculated from S-array are applied as a voltage to bit lines, and corresponding HD current between BCs and CBCs are handled into software for K-means clustering. (c) Mapped BCs of four tree structures based on Figure S20 with four mf-CBAs. The current outputs read at 0.1 V are depicted, with LRS currents represented in red and HRS currents in blue.

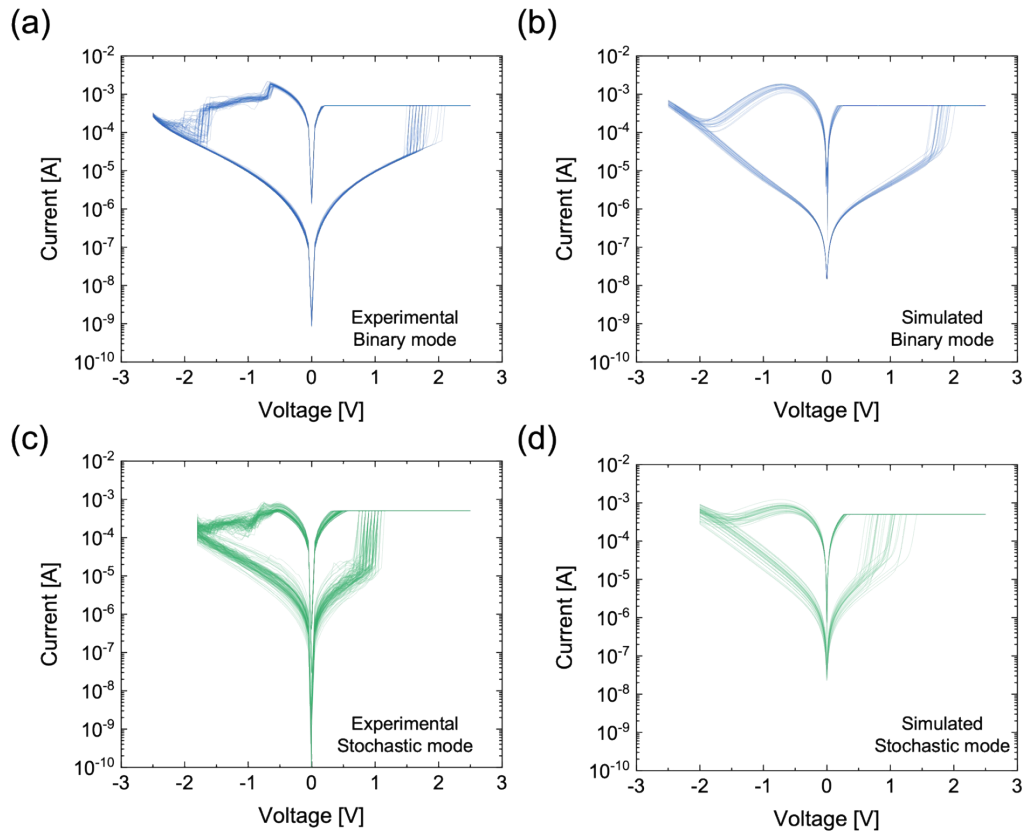


Figure S22. Simulated hardware characteristics of binary and stochastic mode switching behavior. (a) Experimental I-V curves in the binary mode THR device. (b) Simulated I-V binary response with the memory diode model explained in Supporting Note 3. (c) Experimental I-V response of the stochastic mode for the THR response. (d) Simulated I-V stochastic response with the memory diode model explained in Supporting Note 3.

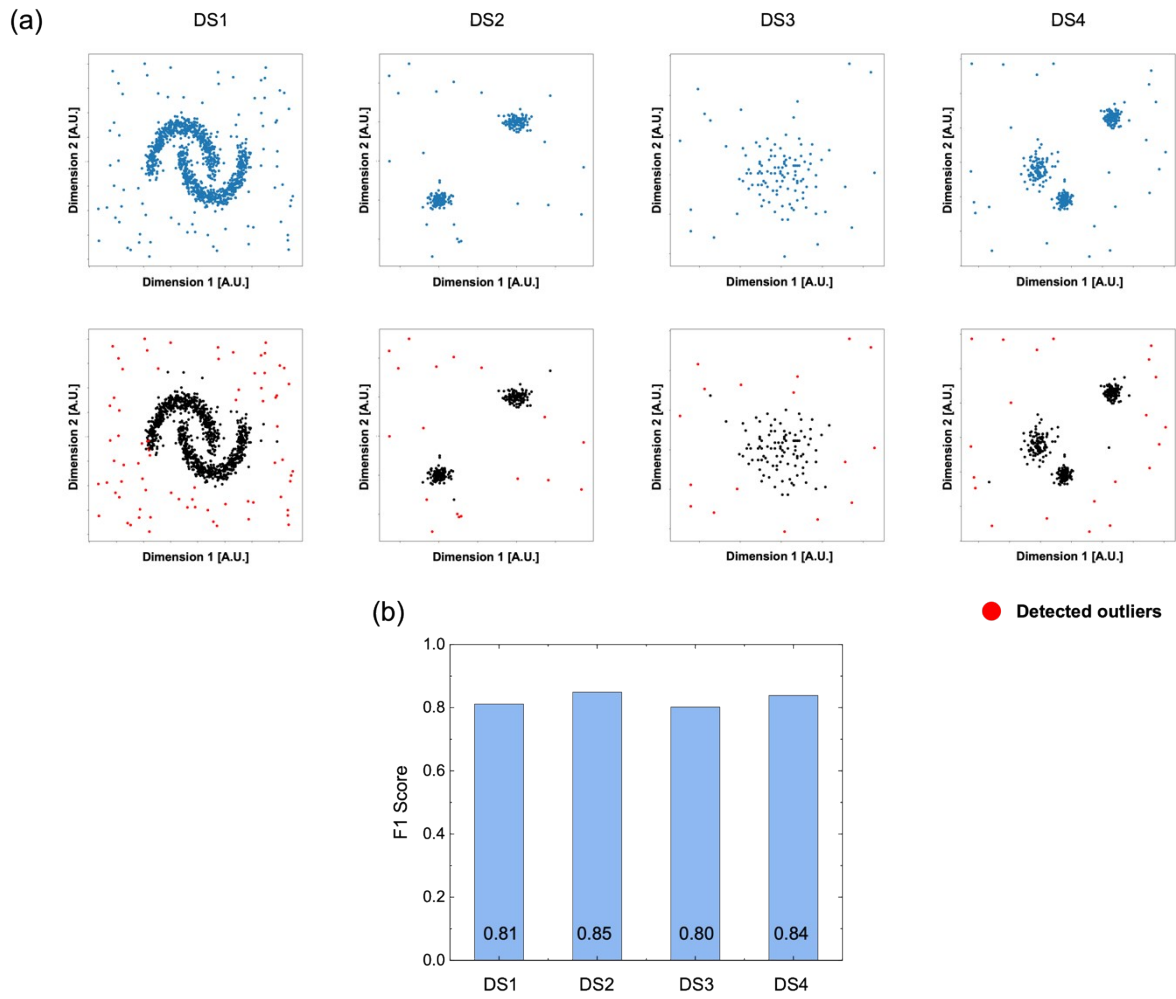


Figure S23. The MOD method results with additional datasets (DS1-DS4). (a) Detected outliers of a given dataset. (b) F1 score of MOD method of each dataset.

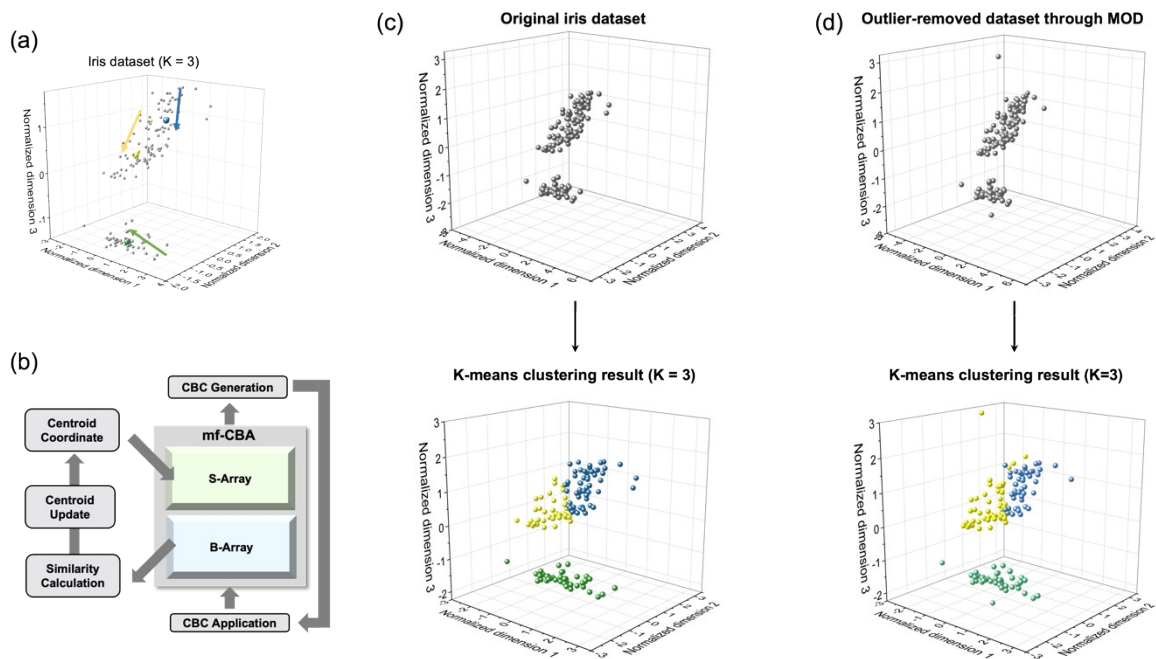


Figure S24. Comparison of K-means clustering results. (a) The centroid movement in the iris dataset ($K=3$). Grey points represent the original data points, and the green, blue, and yellow points show the generated centroids' update. (b) Process flow of K-means clustering with mf-CBA. (c-d) Comparison of the data clustering process with and without outliers of the iris dataset. Due to the robustness of the proposed MOD method, the proposed method clusters the original outlier-free iris dataset appropriately, as shown in (c). Meanwhile, the outlier-removed dataset with MOD also results in the appropriate clustering results, as shown in (d).

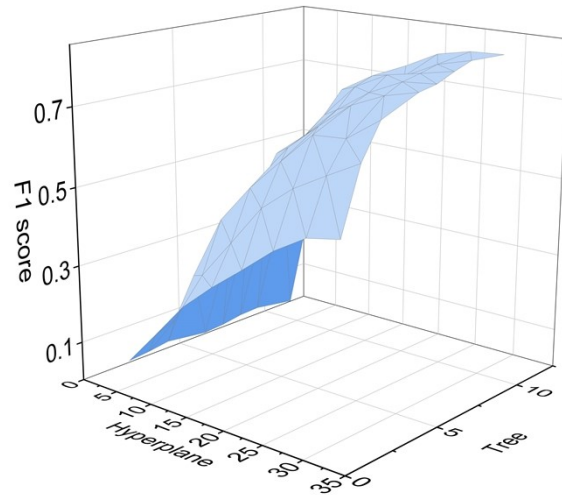


Figure S25. Combined effects of varying the number of hyperplanes and trees. The number of hyperplanes directly influences the number of valid minority bits involved in the outlier score, and the number of trees impacts the degree of the ensemble effect. Therefore, as both factors increase, the performance improves.

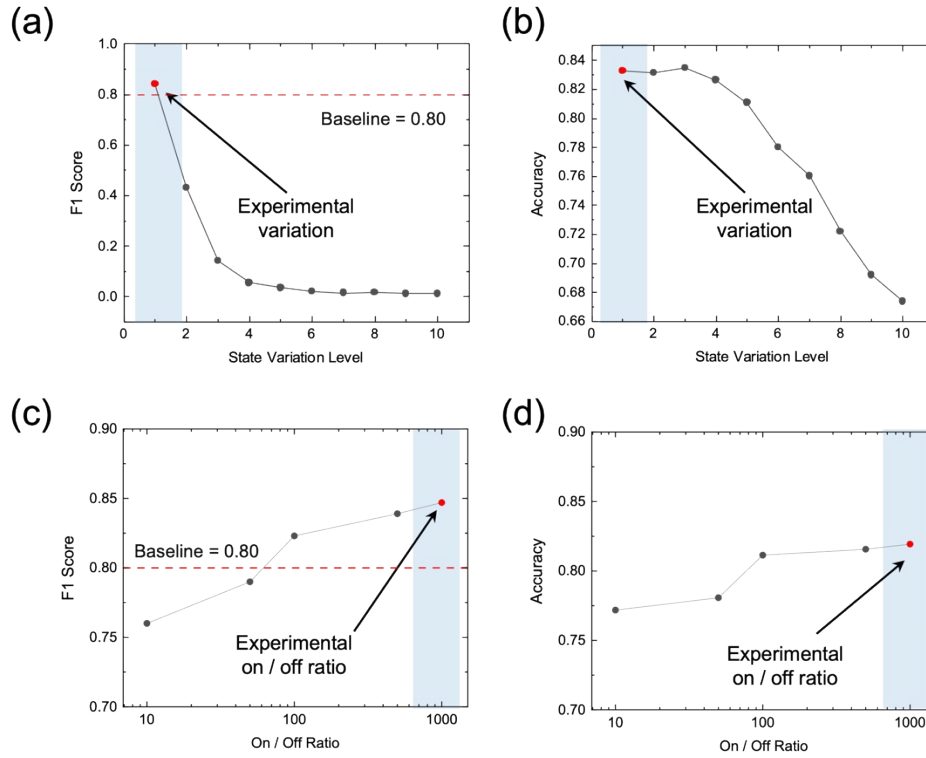


Figure S26. Performances of data mining tasks with iris dataset according to state variation level and memory window of the THR device. (a) The calculated F1 score of OD as a function of the assumed variation level. (b) The calculated accuracy of K-means clustering as a function of the assumed variation level. (c) The calculated F1 score of OD as a function of the assumed on/off ratio (G_{LRS}/G_{HRS}). (d) The calculated accuracy of K-means clustering as a function of the assumed on/off ratio. (a)-(d) validate the feasibility of the proposed mf-CBA for hardware implementation of the proposed data mining methods.

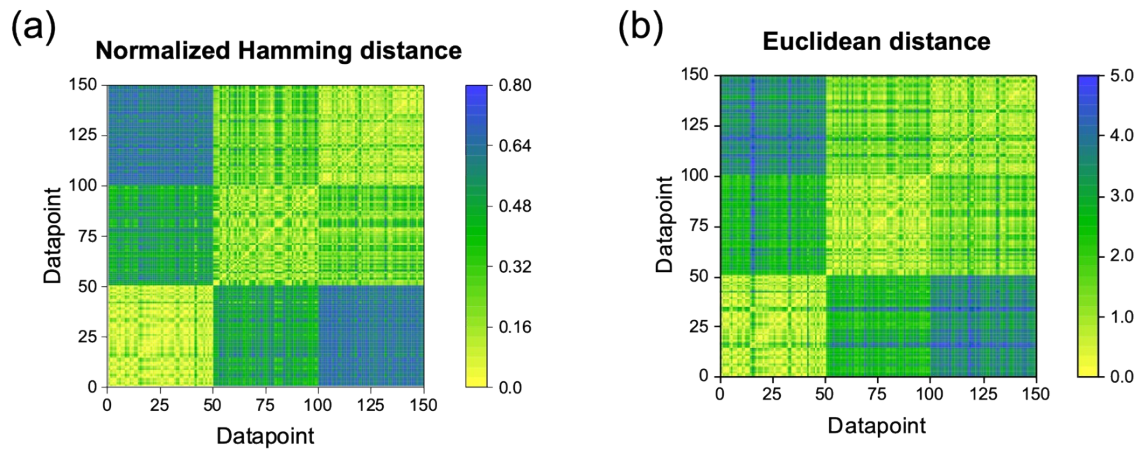


Figure S27. Comparison of distance calculation in the iris dataset. (a) and (b) shows the results of calculated HDs and Euclidean distances between all data points in the iris dataset. HD values were calculated with simulated hardware, and the software calculated the Euclidean distances. Both heatmaps exhibit comparable patterns and are divided into 9 areas, reflecting that the iris dataset is divided into three clusters, with the different areas representing the intra-cluster and inter-cluster distances. The diagonal areas correspond to intra-cluster distances, representing predominantly short-distance values. Conversely, the non-diagonal areas correspond to inter-cluster distances, demonstrating larger distance values than the diagonal case. This result shows that the binary projection of the S-array reflects spatial information accurately.

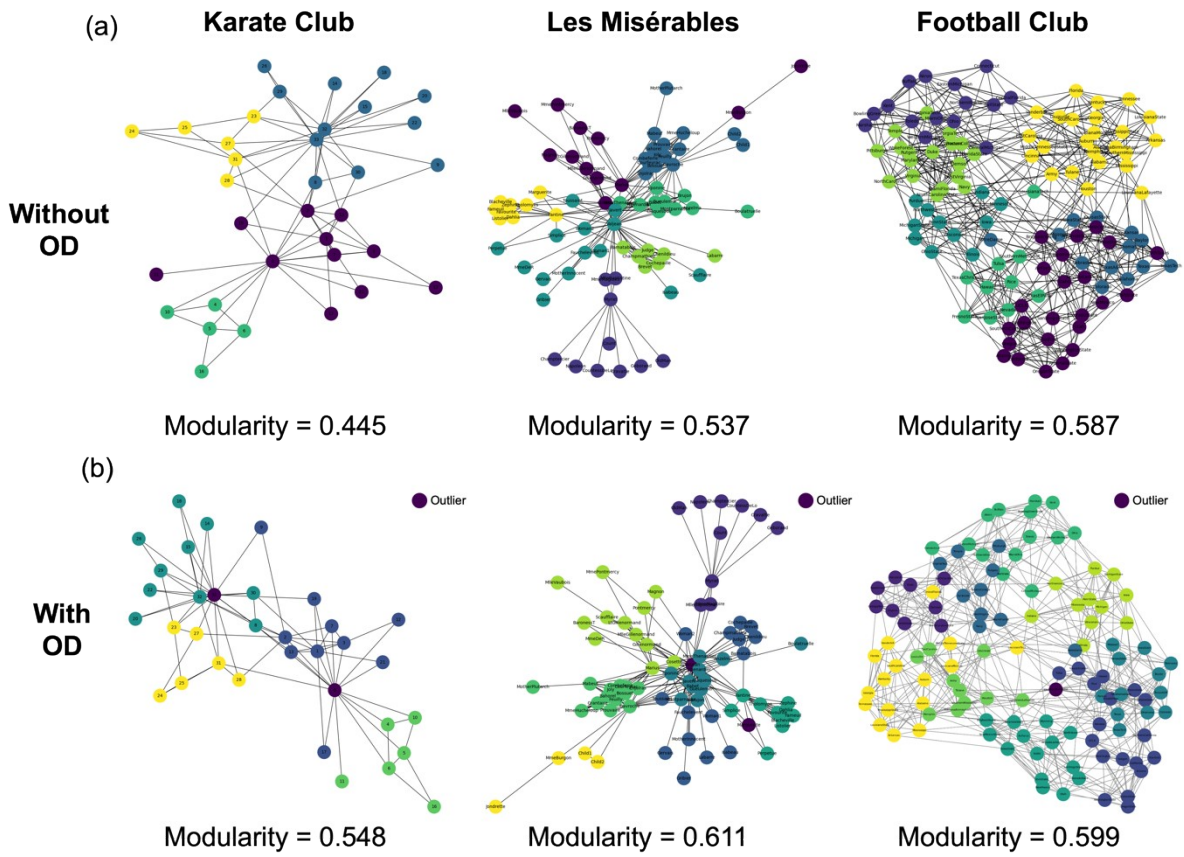


Figure S28. Community detection results in situations where OD was performed and situations where OD was not performed. (a) Cases without OD. (b) Cases with OD

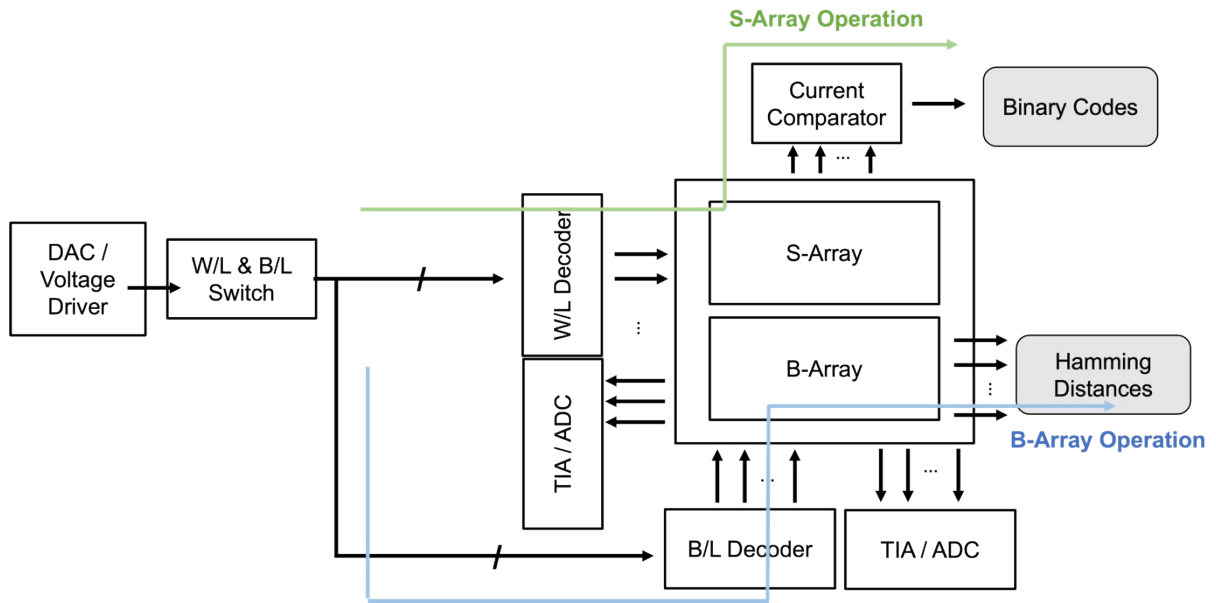


Figure S29. Schematic circuit diagram for mf-CBA operation. The selection of the mf-CBA is implemented with decoders controlled by word line/bit line switches to apply input voltages into the S- and B-arrays selectively. In the operation of the S-array, spatial coordinates are applied to word lines as voltage, and the output current from bit lines is compared with zero to determine binary codes from the hyperplane structure. In the operation of the B-array, MBC/CBC is applied as input voltages to bit lines, and the corresponding Hamming distances are calculated based on the output current.

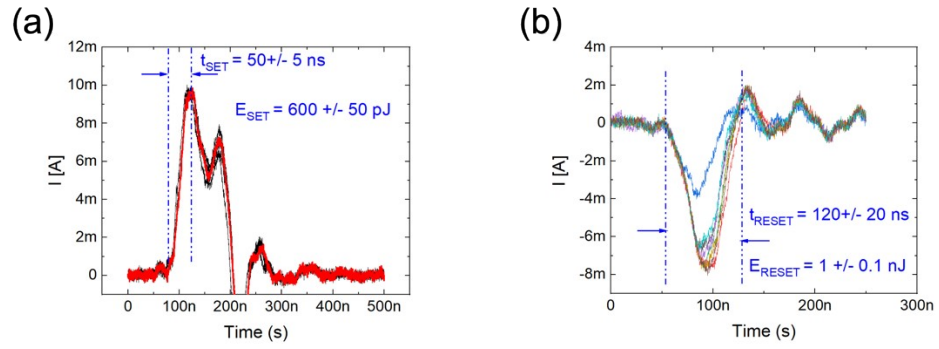


Figure S30. Output current curves of pulse switching for the THR device. (a) SET switching response with 3 V, 100 ns pulse. (b) RESET switching response with -1.9 V, 100 ns pulse. Due to the stochastic mode reset operation, reset response curves show more significant curve fluctuation than set response.

Supporting Note 1. Detailed description of the proposed OD and DC methods

1. The minority-based OD method

The following explains the three phases of the MOD method in detail. The first MBC determination begins by calculating the ratio of "1" s for each hyperplane. The upper middle panel of Figure 3d shows the "minority vote rule", which uses this ratio to determine the MBC. For the given hyperplane in the upper left panel, the only point I has a BC of "1", so the ratio of "1" to the total cases is 0.11 ($= 1/9$). When the hyperparameter M is arbitrarily taken as 0.25, the rule states that the "minority bit" for this specific hyperplane is "1" because the ratio of 0.11 < 0.25 . There could be another hyperplane, which produces the ratio $> (1-M)$, i.e., 0.75, of which the minority bit becomes "0" following the rule. When the ratio is between these two values, the minority bit for that hyperplane is assigned to be X (do not care bit). Therefore, the relationship between the tree and all the points in the given dataset can be represented by the MBC vector with H components when H hyperplanes are assigned for a given tree. It should be noted that each component in a given MBC vector represents the property of the hyperplane, not a data point. For the T numbers of trees, there are T H-long-MBC vectors, as shown in the upper right panel of Figure 3d. This approach prioritizes regions with fewer data points ("minority") as potential outlier locations.

Assigning several specific hyperplanes with the minority bit of X implies a pruning process for those hyperplanes lacking spatial minority information. When a specific hyperplane yields the "X" minority bit, information related to that hyperplane is excluded from subsequent phases. This exclusion is necessary because such a hyperplane tends to indicate a near-equal distribution of "0" s and "1" s, and, thus, to judge almost half of the data points as outliers, degrading the OD performance eventually. However, these hyperplanes are useful for similarity identification, as discussed later.

The second phase calculates the HD between the MBC and all data points' BCs, serving as an outlier score for each point because the MBC was designed to be close to outlier candidates. It should be noted that the HD counts the number of different components between two binary vectors, so the larger the HD the more different the two vectors.³³ The middle panel of Figure 3d shows this HD calculation process for each tree. In the case of a tree with an "X" state, the calculation discards it, and the HD is calculated with a code shorter than the original length. For example, if a tree with four hyperplanes has an MBC of (1, X, 0, 1), the second bits of BC and MBC vectors were excluded from the HD calculation. After the calculations, T HD vectors are acquired, each with a length of data point number.

Outliers are detected based on the calculated HD results in the last outlier counting phase. The lower panel of Figure 3d shows how each tree's outlier candidates are determined, using a hyperparameter R . The calculated HD vectors of each tree are converted to the outlier count vector (OCV), following the outlier counting rule, which assigns "1" or "0" to the components of OCV whether the HD of each data point is within $0 - R$ or $R - 100$, respectively, in ascending order. This assignment implies that the data points with high HD values, i.e., less similar to the MBC vector, tend to have the OCV component of "0". After performing this process for all trees, achieving T OCVs, each component in the OCVs is component-wise added to calculate the final outlier count. Finally, several data points with the highest final outlier count are designated as outliers. This ensemble approach yields robust OD results by utilizing the summed outlier counts across all trees. For more explanations, Figure S13 of the Supporting Information shows the results of applying these phases to the example 9-point dataset.

2. Modified K-means method

First, a predefined number of centroids (K) are created in the given Euclidean data space. Then, hyperplanes compress the spatial information of the centroids, producing a BC vector of the added centroids called the CBC vector, where each component is generated following equation (2). The tree concept is not used in this process, so one CBC vector is produced per centroid, and a single BC vector is generated for each data point. In this case, in contrast to the MOD case, hyperplanes with spatial minority are invalid in CBC (marked as Y, do not care bit) and also excluded from the subsequent process. Therefore, all generated hyperplanes are used for OD or DC without waste. Afterward, CBCs are utilized to calculate the pairwise HD with the remaining normal data points. Then, K HD vectors are created, and the data points with the smallest HD values in the component-wise comparison between the K HD vectors are assigned to that specific centroid, grouping the data points.^{3,4,34} For the given hyperplanes in the upper panel of Figure 3e, hyperplanes 1 and 3 coincide with the minority cases and points A and I are classified as outliers. In the middle panel, points B-C and E-H are classified as closer to centroid 1 (green dot) and 2 (blue dot), respectively.

Once such a grouping is over, each centroid's coordinates are updated to the average position of its assigned data points. This process repeats iteratively until the centroid's coordinates stabilize. The lower panel of Figure 3e shows the positions of the two saturated centroids after several iterations. Based on these positions, the seven normal data points are divided into two clusters.

Supporting Note 2. VMM operation and OD/PC method validation with mf-CBA

To validate the proposed OD and PC method with mf-CBA, the 9-point dataset in the main text was taken as an example. Here, the dataset is two-dimensional, so the hyperplane consists of three variables, two for w and one for b , respectively. Four mf-CBAs with 3 rows and 8 columns are utilized because one hyperplane requires two columns of S-array, and the given structure consists of four hyperplanes in four trees. The randomly distributed hyperplanes demonstrate the feasible operation of the suggested method using the THR memristor's stochastic switching. Afterward, the coordinates of each data point (x , y coordinates, and b value) are applied to three rows (word lines) in the form of voltages, ranging from -0.4 V to +0.4 V. Figure S20a depicts the generated four trees with four hyperplanes in each tree. Due to the stochastic switching of S-array, hyperplanes that divide the dataset are reconstructed from experimentally mapped conductances of mf-CBA. Figure S20b shows the calculated output current values of each point in the corresponding hyperplane tree structure. Here, the sign of the output current can identify the relative positional relationships of 9 points regarding a single hyperplane. Finally, these current output data are converted to BC as shown in the binary numbers (0 and 1) of Figure S20b, which are used in B-array for calculating the HD.

For the B-array operation to calculate HD, each tree's MBC and CBC vectors have H components, so their voltage inputs can be more conveniently made when input through B-array's column lines (bit lines), where the BC values of each data point are already mapped. For H-bit HD calculation, 2H bit lines are required for this method. Then, the HD for each data point is read through the current output at the corresponding row lines (word lines). Figure S21a and b shows the voltage application for MBC and CBC to the column lines, which produce the HD currents for OD and DC, respectively, in the respective row lines. Based on the calculated HD currents, the following processes in both tasks are conducted in software as in the block diagram. The outlier count rule converts the calculated HDs into OCVs in OD. Then, outliers are determined based on the final outlier count, the componentwise sum of OCVs. In DC, the calculated HDs are used as indicators to allocate data points to the nearest centroid for every iteration. Each centroid is moved to the average position of the assigned data points, which is repeated until all centroids' positions are stabilized. The moved centroid's CBC can be calculated with S-array as in Figure S14b to recalculate HD between the moved centroid's CBC and mapped BCs. After completing these processes, two outliers are detected in the 9-point dataset, and the remaining data points are divided into two clusters.

Figure S21c shows the current values (read at 0.1 V) when the BCs of the aforementioned 9-point dataset in the main text are mapped onto four B-arrays (trees), where the I_{HRS} and I_{LRS} are marked in blue and red backgrounds, respectively. For a data point in a single tree, 4-bit binary code can be mapped in 8 cells in a single row, so four mf-CBA with 9 rows (corresponds to the number of data points) and 8 columns (corresponds to the binary code mapping method, as in Figure 4f of main text) In other words, each tree consists of four double columns, and each double-column corresponds to one hyperplane. These cells represent the double-column binary mapping in Figure 4f of the main text, which is used to calculate HDs with MBC and CBC when input to the column lines in voltage form. Each mapped B-array contains 4-bit BCs of 9 points, resulting in the calculated HDs spanning from 0 to 4, as shown in Figure 4i of the main text.

Supporting Note 3. Device compact modeling of multi-functional THR memristor

In the context of the multi-functionality of the THR memristor and its filamentary-based resistive switching, a diode-like model was employed for the observed electrical response.⁹ This approach is supported by the quantum theory, which predicts that the current that flows through a filamentary-type structure can be approximated by a potential barrier with height E_B as given by the expression⁹:

$$I = eA \{ \exp[-\gamma(E_b - \mu_L)] - \exp[-\gamma(E_b - \mu_R)] \} \quad (S1)$$

, where μ_L and μ_R represent the electrochemical potentials at the left and right reservoirs of the structure, respectively, A and γ are constants related to the particular characteristics of the potential barrier, and e denotes the elementary charge of an electron. Under the application of a positive bias voltage $V > 0$, $\mu_L - \mu_R = eV$ so that Eq. (S1) becomes:

$$I(V) = I_0[\exp(\beta V) - 1] \quad (S2)$$

which is formally analogous to the ideal diode equation, I_0 represents the diode amplitude (which may exhibit voltage dependence through E_b), and $\beta \approx d \ln(I)/dV$ is the logarithmic conductance of the diode. For negative bias voltage $V < 0$, a similar equation, with the appropriate signs, holds. For a comprehensive model, series (R_S) and parallel resistances (R_{P1} , R_{P2}) are incorporated. Physically, R_S may represent a residual local potential barrier. In contrast, R_{P1} and R_{P2} may represent localized and area-distributed parallel leakage current pathways. The overall current-voltage relationship can thus be approximated by the expression⁹:

$$I(V) = \text{sgn}(V) \{ (\beta R_S)^{-1} W[\beta I_0 R_S d_1 \exp(\beta d_1 \exp(\beta d_1 (|V| + I_0 R_S)))] + d_1 (|V| G_{P1} - I_0) - |V| G_{P2} \} \quad (S3)$$

where $G_{P1} = 1/R_{P1}$, $G_{P2} = 1/R_{P2}$, $d_1 = (1 + G_{P1} R_S)^{-1}$, sgn is the sign function, and W is the Lambert function. For the Lambert function, we considered the following approximation, which is based on a Hermite-Padé expansion¹⁰:

$$W(x) = \ln(1+x) \left(1 - \frac{\ln(1 + \ln(1+x))}{2 + \ln(1+x)} \right) \quad (S4)$$

for real $x > 0$ with a relative error less than 10^{-2} .

Within the context of this work, the term I_0 in Eq. (S2) is modeled by employing sigmoidal-type threshold functions. This selection is motivated by the ability of these functions

to generate smooth transitions between the two resistance states during the application of a voltage signal:

$$I_0 = I_{0Min} + (I_{0Max} - I_{0Min}) \lambda(V) \quad (S5)$$

with

$$\lambda^{on/off}(V) = \lambda_{Max} \tanh \left[-\eta \left(V - \frac{V_{th}^{on/off}}{\beta_r} \right) \right] \quad (S6)$$

β_r regulates the threshold at which switching occurs between the two resistance states, and it is directly linked to the current compliance to mimic the behavior of Figure S22. $V_{th}^{on/off}$ is the threshold voltage to switch from the HRS to the LRS, and back to the HRS. The λ function plays the role of the state equation of the memristor.

Eqs. (S1)-(S6) show that the proposed model possesses essential characteristics for the compact modeling of memristors as it is analytic, continuous, and differentiable. The tractable mathematical functions accurately reproduce diverse electrical responses, including linear/non-linear and pinched/non-pinched I - V responses. η determines the transition rates for both switching processes. Supporting Table S2 shows the fitting parameters representing THR devices' switching characteristics, including the switching variations. Figure S22 shows the fitting results for the I - V curves in the binary and stochastic modes.

Supporting Table S2. Fitting Parameters for the binary and stochastic modes of the I-V responses of the THR-based memristor as shown in Figure S22.

| Parameter | Binary mode | Stochastic mode |
|------------------|--------------------|------------------------|
| η | 30 | 35 |
| α | 3.5 | 3 |
| $I_{O,min}$ | 150 ± 20 nA | 3 ± 2 μ A |
| $I_{O,max}$ | 3 ± 0.5 mA | 800 ± 300 μ A |
| R_s | 600 Ω | 1 k Ω |
| R_p | 10 k Ω | 10 k Ω |

Supporting Note 4. Energy consumption and latency calculation method

1. CPU-based digital computing system

The energy consumption for a CPU-based digital computing system was estimated as follows: thermal design power of CPU \times CPU utilization during data mining \times computational time.^{12,13} Through this method, energy consumption was calculated when performing the proposed method and a combination of the conventional algorithms (isolation forest + K-means).

The energy consumption for the proposed method was estimated as 45 W (thermal design power of CPU, Ryzen 5 4600H) \times 2.5 % (CPU utilization during data mining) \times 63.8 ms (computational time for data mining) = 71.8 mJ. And, the energy consumption for a combination of the conventional algorithms was estimated as 45 W (thermal design power of CPU, Ryzen 5 4600H) \times 2.4 % (CPU utilization during data mining) \times 43.2 ms (computational time for data mining) = 47.4 mJ.

2. Memristive computing system

The energy consumption of the S-array was the sum of switching energy ($E_{\text{switching}}$) for the mapping and static energy (E_{static}) for the VMM operations. Thus, the total energy was estimated using the following equations:

$$E_{\text{switching}} = \int V(t) \cdot I(t) \cdot dt \quad (\text{S7})$$

$$E_{\text{static}} = I_{\text{stochastic}} \cdot t_{\text{read}} \cdot (V_{\text{read},1} \cdot N_{\text{cell},1} \cdot N_{\text{dataset}} + V_{\text{in},1} \cdot N_{\text{centroid}}) \quad (\text{S8})$$

, where $E_{\text{switching}}$ is the energy for the set and reset process, $V(t)$ is the voltage till the switching process occurs, and $I(t)$ is the current that flows through the device during the switching event, as shown in Figure S24. The read voltage ($V_{\text{read},1}$) was 0.2 V on average, and its output current ($I_{\text{stochastic}}$) was 2 μA on average. The width of the read pulse, t_{read} , was 100 ns. The sum of S-array cells ($N_{\text{cell},1}$) depends on the S-array configuration.

Similarly, energy consumption calculations for the B-array were determined with equation S7 for the $E_{\text{switching}}$ during the mapping step, and the static energy was calculated for each iteration in the VMM operations as:

$$E_{\text{static}} = I_{\text{binary}} \cdot t_{\text{read}} \cdot N_{\text{dataset}} \cdot (V_{\text{read},2} \cdot N_{\text{cell},2} \cdot (1 - R_X) + V_{\text{read},2} \cdot N_{\text{cell},2} \cdot R_X \cdot N_{\text{centroid}}) \quad (\text{S9})$$

The read voltage ($V_{\text{read},2}$) was 0.1 V, and its output current (I_{binary}) was 0.05 mA, corresponding to the HRS and LRS average. The width of the read pulse, t_{read} , was 100 ns. The

sum of B-array cells ($N_{cell,2}$) depends on the B-array configuration. The ratio of X (R_X) is determined by the minority rate.

The energy consumption of the peripheral circuit on the mf-CBA was estimated as $P_{peri} \cdot t \cdot N_{element}$ based on 130-nm Si CMOS technology.¹⁴⁻¹⁷ The DAC, ADC, and comparator power were about 7.2 mW, 2.3 mW, and 0.2 mW. In addition, peripheral circuits, such as DAC, ADC, and voltage drivers, were estimated based on an 8-bit I/O system. Finally, this work's step-by-step results of energy calculations regarding latency and energy consumption are described in Supporting Table S3.

Supporting Table S3. Summarized step-by-step energy consumption and latency

| | This work | |
|---|------------------------|-----------------------|
| | Energy (mJ) | Latency (ms) |
| Hyperplane Tree Generation (S-array Mapping) | 3.3×10^{-3} | 2×10^{-4} |
| Spatial Data Compression (S-array VMM) | 0.9×10^{-3} | 1.9×10^{-2} |
| Binary Code Representation (B-array Mapping) | 24.6×10^{-3} | 1.6×10^{-2} |
| Hamming Distance Calculation (B-array VMM) | 1.3×10^{-3} | 3.8×10^{-3} |
| Peripheral Circuits | 135.4×10^{-3} | 2.1×10^{-4} |
| Total | 165.5×10^{-3} | 3.92×10^{-2} |

Supporting References

- 1 R. E. Banfield, L. O. Hall, K. W. Bowyer and W. P. Kegelmeyer, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007, **29**, DOI:10.1109/TPAMI.2007.250609.
- 2 D. Che, Q. Liu, K. Rasheed and X. Tao, in *Advances in Experimental Medicine and Biology*, 2011, vol. 696.
- 3 A. F. Veinott, *Oper. Res.*, 1967, **15**, DOI:10.1287/opre.15.1.147.
- 4 M. Khanuja, H. Sharma, B. R. Mehta and S. M. Shivaprasad, *J. Electron Spectros. Relat. Phenomena*, 2009, **169**, DOI:10.1016/j.elspec.2008.10.004.
- 5 O. Kerrec, D. Devilliers, H. Groult and P. Marcus, *Materials Science and Engineering B*, 1998, **55**, DOI:10.1016/s0921-5107(98)00177-9.
- 6 S. Rudenja, A. Minko and D. A. Buchanan, *Appl. Surf. Sci.*, 2010, **257**, DOI:10.1016/j.apsusc.2010.06.012.
- 7 L. Mai, D. Zanders, E. Subaşl, E. Ciftiyurek, C. Hoppe, D. Rogalla, W. Gilbert, T. D. L. Arcos, K. Schierbaum, G. Grundmeier, C. Bock and A. Devi, *ACS Appl. Mater. Interfaces*, 2019, **11**, DOI:10.1021/acsami.8b16443.
- 8 D. H. Shin, H. Park, N. Ghenzi, Y. R. Kim, S. Cheong, S. K. Shim, S. Yim, T. W. Park, H. Song, J. K. Lee, B. S. Kim, T. Park and C. S. Hwang, *ACS Appl. Mater. Interfaces*, 2023, DOI:10.1021/acsami.3c19523.
- 9 J. Matoušek, *Random Structures & Algorithms*, 2008, **33**, DOI:10.1002/rsa.20218.
- 10 J. Blasco, N. Ghenzi, J. Suae, P. Levy and E. Miranda, *IEEE Electron. Device Letters*, 2014, **35**, DOI:10.1109/LED.2014.2297992.
- 11 S. Winitzki, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2003, **2667**, DOI:10.1007/3-540-44839-x_82.
- 12 Y. Feng, Y. Zhang, Z. Zhou, P. Huang, L. Liu, X. Liu and J. Kang, *Nat. Commun.*, 2024, **15**, DOI:10.1038/s41467-024-45312-0.
- 13 H. Song, W. Park, G. Kim, M. G. Choi, J. H. In, H. Rhee and K. M. Kim, *Advanced Materials*, 2024, **36**, DOI:10.1002/adma.202400977.
- 14 F. Solis, Á. Fernández Bocco, A. C. Galetto, L. Passetti, M. R. Hueda and B. T. Reyes, *International Journal of Circuit Theory and Applications*, 2021, **49**, DOI:10.1002/cta.3029.
- 15 A. K. Adepur and K. K. Kolupuri, *International Journal on Cybernetics & Informatics*, 2016, **5**, DOI:10.5121/ijci.2016.5434.
- 16 F. Moradi, G. Panagopoulos, G. Karakonstantis, H. Farkhani, D. T. Wisland, J. K. Madsen, H. Mahmoodi and K. Roy, *Microelectronics J.*, 2014, **45**, DOI:10.1016/j.mejo.2013.09.009.
- 17 B. B. A. Fouzy, M. B. I. Reaz, M. A. S. Bhuiyan, M. T. I. Badal and F. H. Hashim, in *2016 International Conference on Advances in Electrical, Electronic and Systems Engineering, ICAEES 2016*, 2016.