Supplementary Information

Memristive Monte Carlo DropConnect crossbar array enabled by device and algorithm co-design

Do Hoon Kim,‡ Woon Hyung Cheong,‡ Hanchan Song, Jae Bum Jeon, Geunyoung Kim and Kyung Min Kim*

‡These authors equally contributed to this work.

D. H. Kim, H. Song, J. Jeon, G. Kim, Prof. K. M. Kim
Department of Materials Science and Engineering
Korea Advanced Institute of Science and Technology (KAIST),
Daejeon 34141, Republic of Korea
*E-mail: km.kim@kaist.ac.kr

W. H. Cheong
Applied Science Research Institute
Korea Advanced Institute of Science and Technology (KAIST),
Daejeon 34141, Republic of Korea.

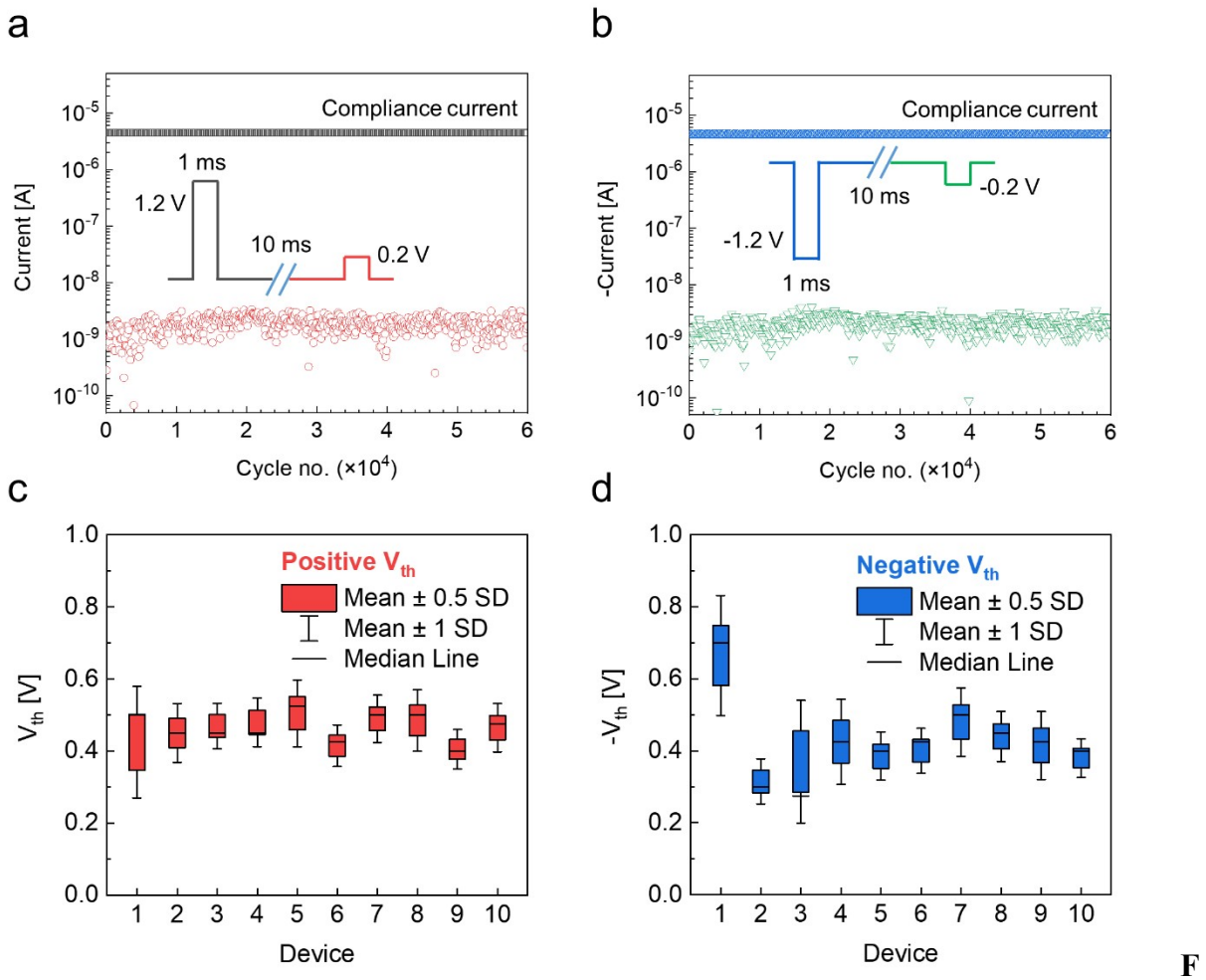
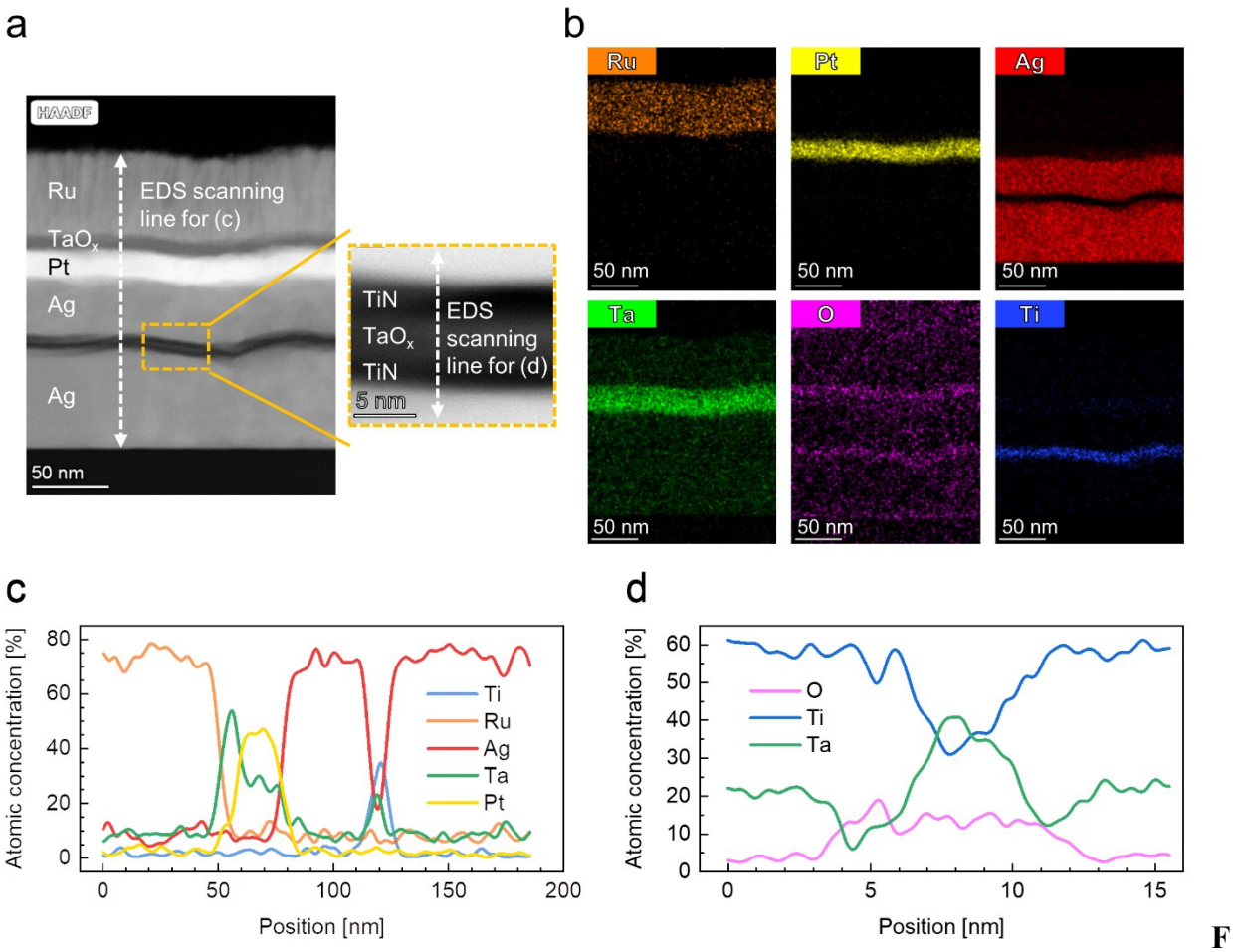


Fig. S1 (a, b) Endurance test results of the Ag/TiN/TaO_x/TiN/Ag bipolar selector under positive bias (a) and negative bias (b). One cycle consists of an ON pulse with 1.2 V amplitude and an OFF pulse with 0.2 V amplitude. The duration of the ON and OFF pulses is 1 ms. The interval between the ON and OFF pulses is 10 ms. In order to prevent non-volatile resistive switching, a series resistor of 200 kΩ was connected, limiting the current shooting similar as applying the compliance current. (c, d) Distribution of threshold voltages (V_{th}) under positive bias (c) and negative bias (d) obtained from 10 bipolar selector devices.



ig. S2 Transmission electron microscopy (TEM) cross-section image and energy dispersive spectroscopy (EDS) analysis of a bipolar 1S1M device. (a) Left panel shows cross-section image of the bipolar 1S1M device with material stack of Ru/TaO_x/Pt/Ag/TiN/TaO_x/TiN/Ag. Right panel shows magnified cross-section of TiN/TaO_x/TiN. (b) EDS mapping images of the bipolar 1S1M device. (c) EDS line profile of Ru/TaO_x/Pt/Ag/TiN/TaO_x/TiN/Ag. The corresponding line is shown in the left panel of (a). (d) EDS line profile of TiN/TaO_x/TiN. The corresponding line is shown in the right panel of (a).

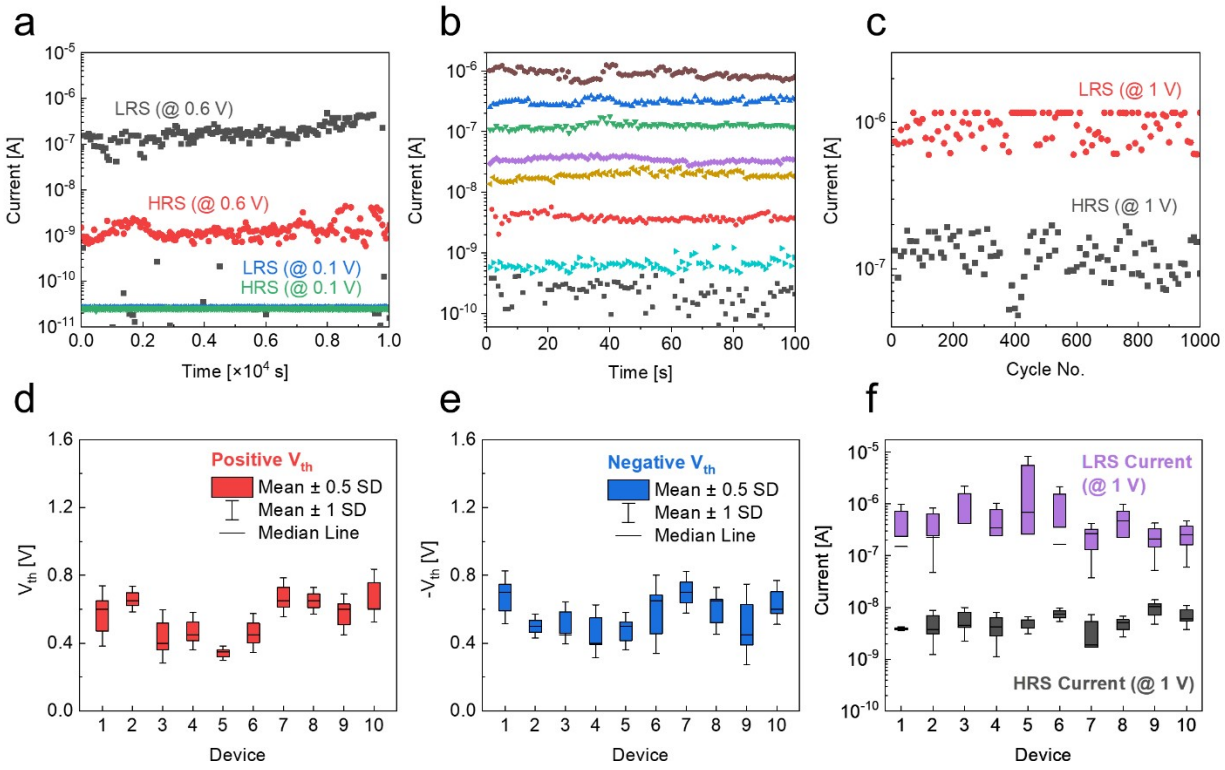


Fig. S3 (a) Retention characteristics of the Ru/TaO_x/Pt/Ag/TiN/TaO_x/TiN/Ag 1S1M device for 10,000 s. The read voltages were 0.1 V for the selector's OFF state and 0.6 V for the selector's ON state. At the selector OFF condition (0.1 V reading), both the LRS (blue) and HRS (green) was read as the OFF state. At the selector ON condition (0.6 V reading), the LRS (black) and HRS (red) were clearly distinguished. Among the LRS at 0.6 V, some was read as the OFF state because the switch was not turned on due to its stochastic switching behavior. (b) The retention characteristics for the 8 resistance states (3-bits) of the 1S1M device for 100 s under the 0.6 V read voltage. (c) The cycling endurance results of the 1S1M device for 1,000 cycles. 3 V and -4 V voltage pulses with 100 μs width were applied for set and reset switching, respectively. 1 V voltage pulse with 1 ms width was applied for reading. (d, e) Threshold voltage (V_{th}) distribution at positive bias (d) and negative bias (e) collected from 10 bipolar 1S1M devices. (f) HRS and LRS current values under the 1 V read voltage from 10 bipolar 1S1M devices.

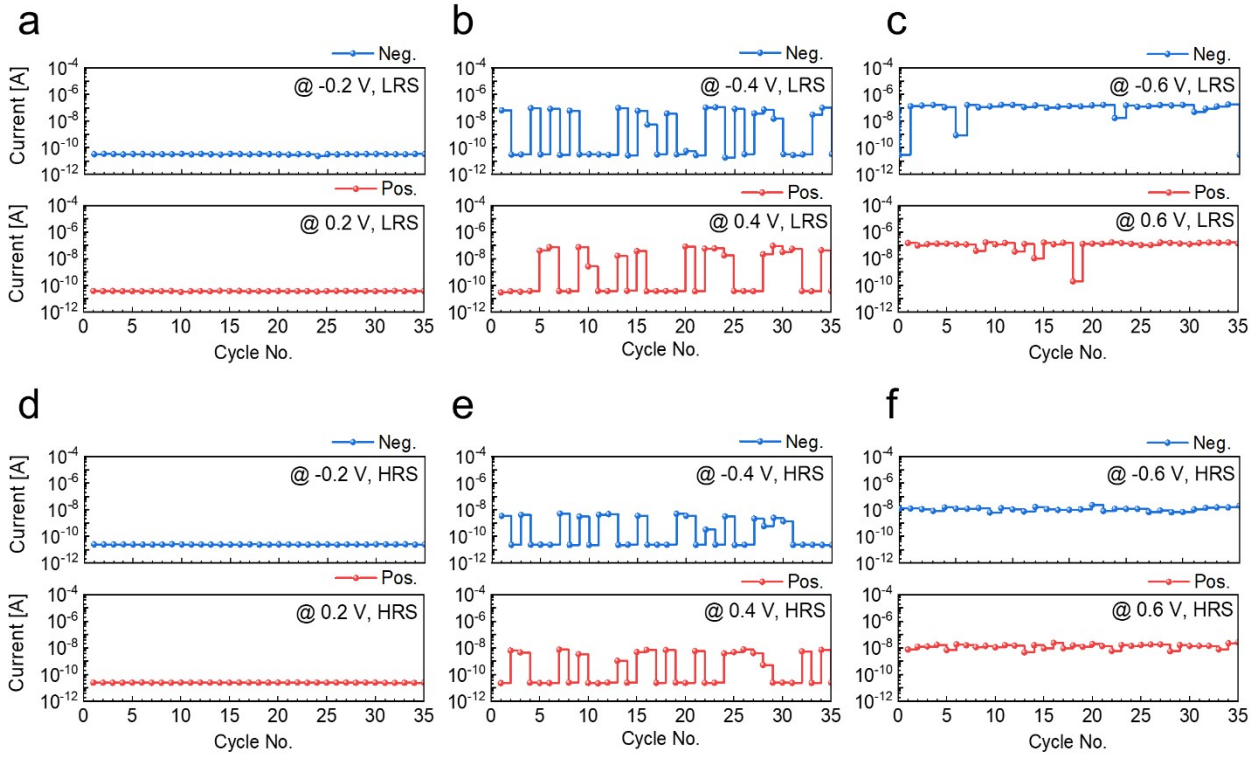
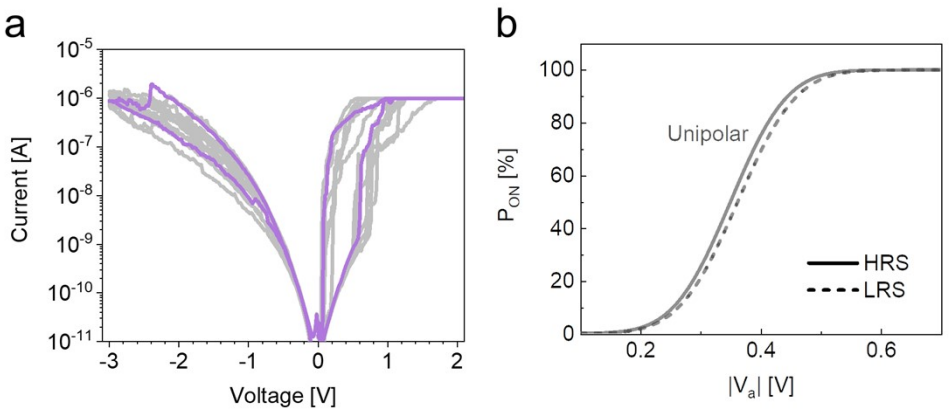


Fig. S4 Voltage-dependent probabilistic readout behavior of a bipolar 1S1M device. (a)-(c) Readout results when the read voltages were ±0.2, ±0.4, and ±0.6 V, respectively, after programming the device to the LRS (~5.0 MΩ). The upper panels correspond to readout with a negative bias, while the lower panels correspond to a positive one. (d)-(f) Readout results after programming the device to the HRS (~47 MΩ) at the identical read voltages. At ±0.2 V, switching of the selector hardly occurred, resulting in the reading of only the OFF state of the selector. At ±0.4 V, close to the average V_{th}, the switching probability was approximately 50 %, allowing for stochastic LRS or HRS readout. At ±0.6 V, the selector switched ON deterministically, enabling the reading of the LRS or HRS.



c

Selector		R_{off} @ 0.2 V (Ω)	VDF (HRS)	VDF (LRS)
Bipolar	Pos.	20.2 G	0.59%	0.043%
	Neg.	33.2 G	1.13%	0.027%
Unipolar		3.19 G	3.60%	0.27%

Fig. S5 A unipolar 1S1M device and switching probability of integrated devices. (a) DC I-V characteristics of the unipolar Ag/TiN/TaO_x/Pt selector-based 1S1M device. Stochastic reading occurs in the positive bias region, while threshold switching does not appear in the negative bias region. (b) P_{ON} (CDF of the fitted Gaussian distribution of the selector V_{th}) of the unipolar 1S1M device for the HRS and LRS cases. It was calculated via Equation (3), the same as Fig. 3f in the main text, where the voltage dividing effect of the selector and the memristor was also considered. (c) Table of parameters used when calculating the voltage dividing effect. The voltage dividing factor (VDF), a variable that indicates how much voltage drop occurs compared to the applied voltage to the integrated device using the selector's R_{OFF} and the memristor's HRS and LRS resistance values, is shown. When the memristor is HRS, the unipolar selector has a larger VDF since its R_{OFF} is much smaller than the bipolar selector's R_{OFF} at positive and negative bias, resulting in the deviation of P_{ON} in (b).

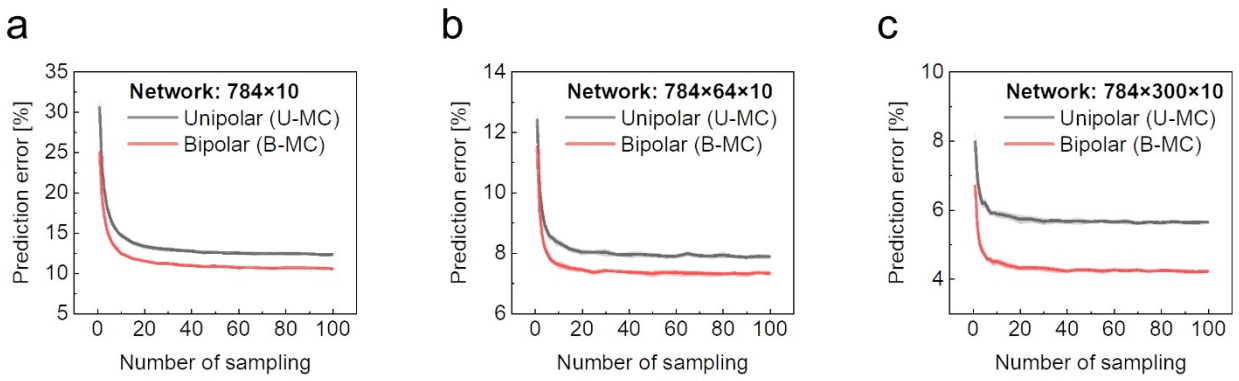


Fig. S6 Recognition error as a function of the MC sampling number for MC-DC networks of different sizes. The sizes of each network are (a) 784×10, (b) 784×64×10, and (c) 784×300×10. Smaller-sized networks exhibit significantly higher prediction errors compared to larger-sized networks. However, all networks consistently demonstrate that the prediction error saturates to a specific value as the sampling rate increases.

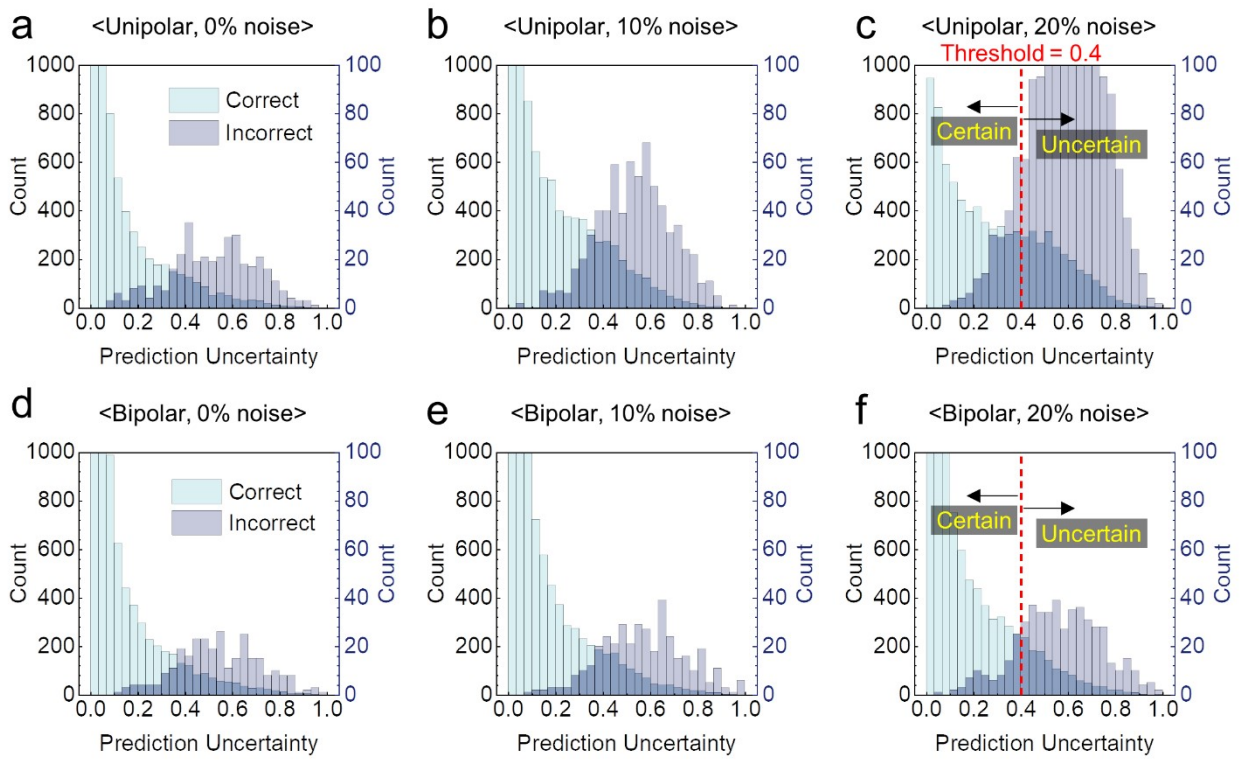


Fig. S7 Histograms of estimated prediction uncertainty for the noisy MNIST test dataset. (a)-(c) Distribution of prediction uncertainty for correct (light blue) and incorrect (deep blue) cases for 0, 10, and 20% noisy MNIST classification with the unipolar MC-DropConnect (MC-DC) 1S1M array (U-MC in the main text), respectively. (e)-(g) Distribution of prediction uncertainty for correct (light blue) and incorrect (deep blue) cases for 0, 10, and 20% noisy MNIST classification with the bipolar MC-DC 1S1M array (B-MC in the main text), respectively. The more the correct and incorrect distributions overlap, the more ambiguous it is to reject an incorrect answer with the estimated uncertainty value. In the case of a bipolar MC- DC 1S1M array, the overlap area remains small as the noise ratio increases, while the overlap area increases with the noise ratio in the unipolar MC- DC 1S1M array.

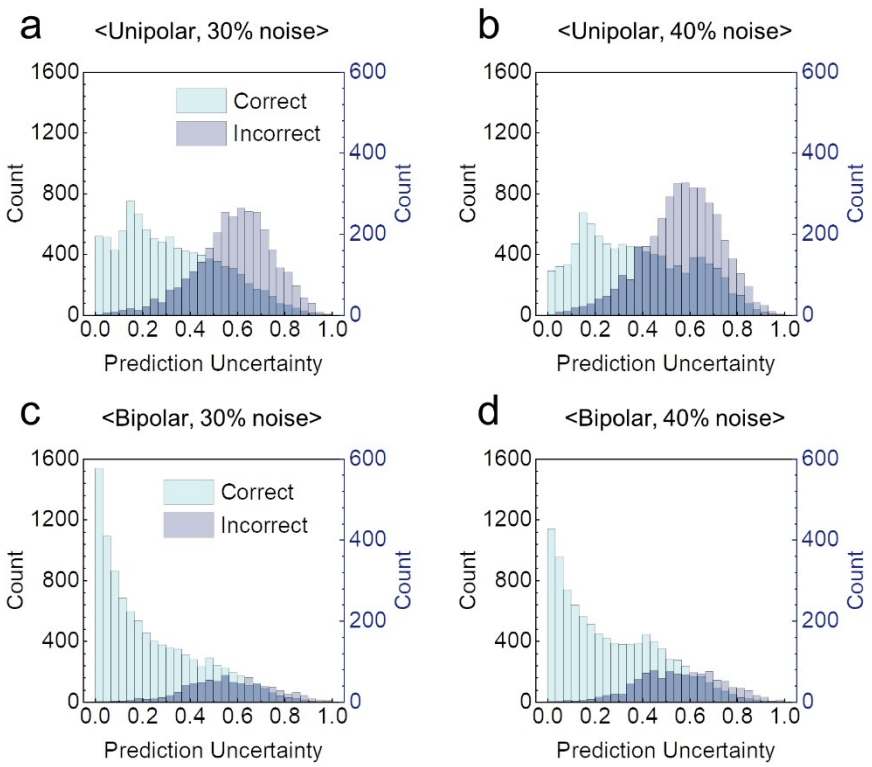


Fig. S8 Histograms of estimated prediction uncertainty for the noisy MNIST test dataset. (a), (b) Distribution of prediction uncertainty for correct (light blue) and incorrect (deep blue) cases for 30 and 40% noisy MNIST classification with the unipolar MC- DC 1S1M array (U-MC in the main text), respectively. (c), (d) Distribution of prediction uncertainty for correct (light blue) and incorrect (deep blue) cases for 30 and 40% noisy MNIST classification with the bipolar MC- DC 1S1M array (B-MC in the main text), respectively.

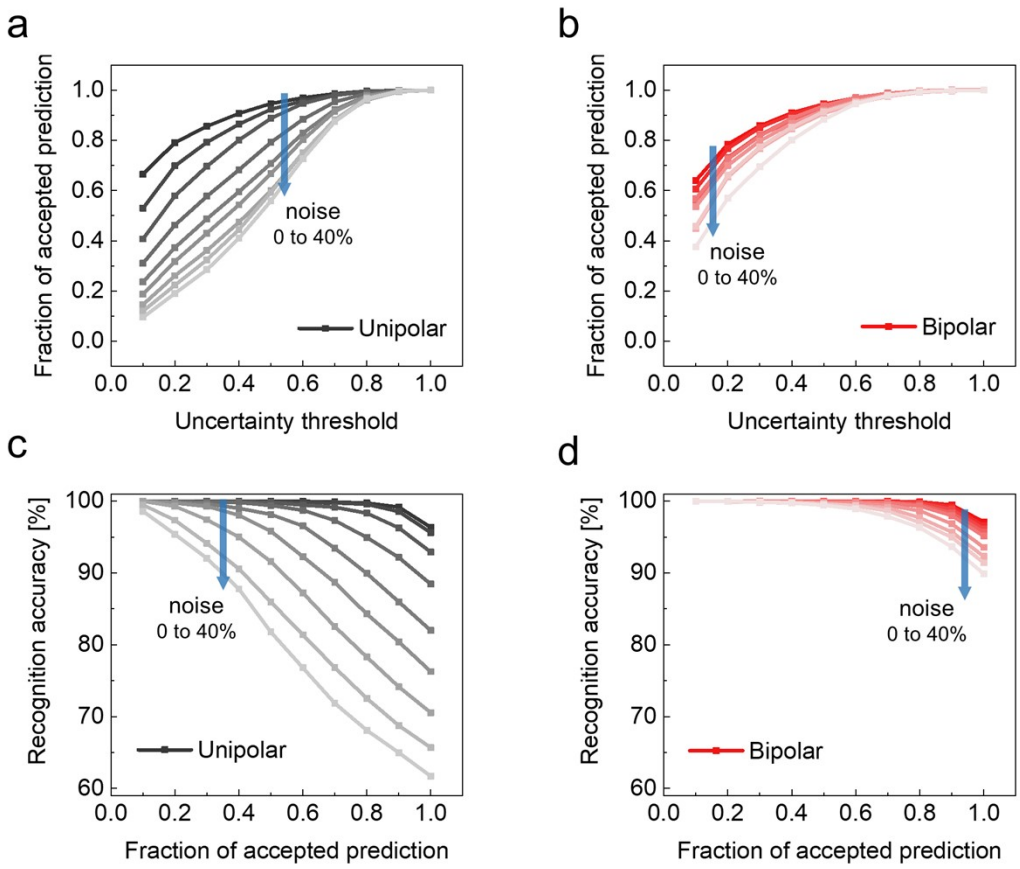


Fig. S9 Uncertainty-based prediction results. (a), (b) Fraction of accepted prediction as a function of uncertainty threshold, based on uncertainty estimation results of 10,000 MNIST test datasets by unipolar and bipolar MC-DC 1S1M arrays, respectively. Data with uncertainty higher than the corresponding uncertainty threshold value are rejected. (c), (d) Prediction accuracy on the noisy MNIST test dataset as a function of fraction of accepted prediction on unipolar and bipolar MC-DC 1S1M array, respectively. Only a subset of the corresponding fraction of predictions for the entire test dataset with the lowest uncertainty is retained, and the rest of the prediction is rejected.

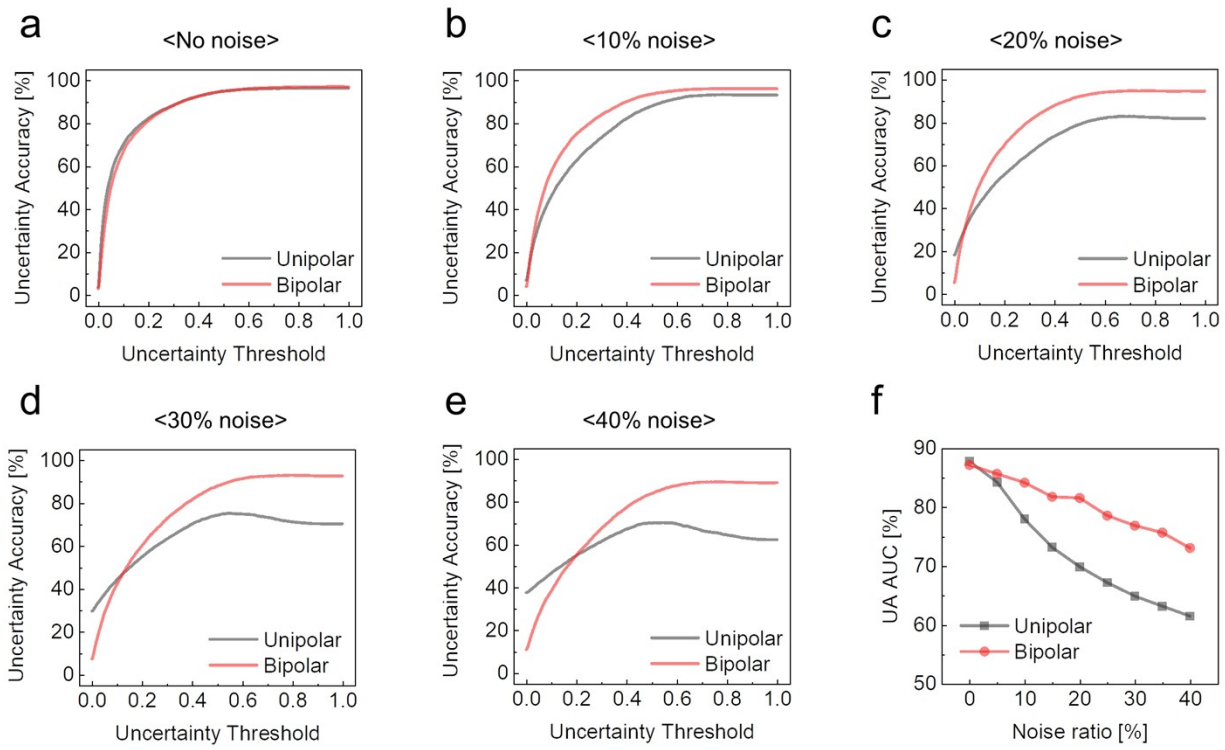


Fig. S10 The uncertainty accuracy (UA) as a function of the uncertainty threshold of the unipolar and bipolar MC-DC 1S1M array (also referred to as U-MC and B-MC in the main text, respectively). (a)-(e) show uncertainty accuracy values for the MNIST test dataset with 0, 10, 20, 30, and 40% noise, respectively. For all noise ratios, the bipolar case has greater uncertainty accuracy than the unipolar case. (f) Area under curve (AUC) metric of the uncertainty accuracy as a function of noise ratio of the unipolar and bipolar MC-DC hardware. Each data point represents the area under curve metric of the uncertainty accuracy curve of (a)-(e).