Supporting Information

3D-Structured Mesoporous Silica Memristor for Neuromorphic Switching and Reservoir Computing

Ayoub H. Jaafar,[a,b] Li Shao,[c] Peng Dai,[a] Tongjun Zhang,[a] Yisong Han,[d] Richard Beanland,[d] Neil T. Kemp,[b] Philip N. Bartlett,[c] Andrew L. Hector,[c] and Ruomeng Huang [a*]*

[a] School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK
[b] School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK
[c] School of Chemistry, University of Southampton, Southampton, SO17 1BJ, UK
[d] Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

Email: a.l.hector@soton.ac.uk; r.huang@soton.ac.uk

KEYWORDS: neuromorphic computing, mesoporous silica, reservoir computing, short term plasticity, temporal information processing

Supporting Information

Figure S1 demonstrates the vertical and horizontal cuts extracted from the GISAXS data shown in Figure 1b. In the horizonal-cut, the peak with the highest intensity, around 0.4 nm^{-1}, corresponds to the expected specular reflection (0.428 nm^{-1}) and Yoneda peak (0.384 nm^{-1}). The other peak observed between 0.58 nm^{-1} and 1.61 nm^{-1} are correspond to (020), (111) and (022) planes.

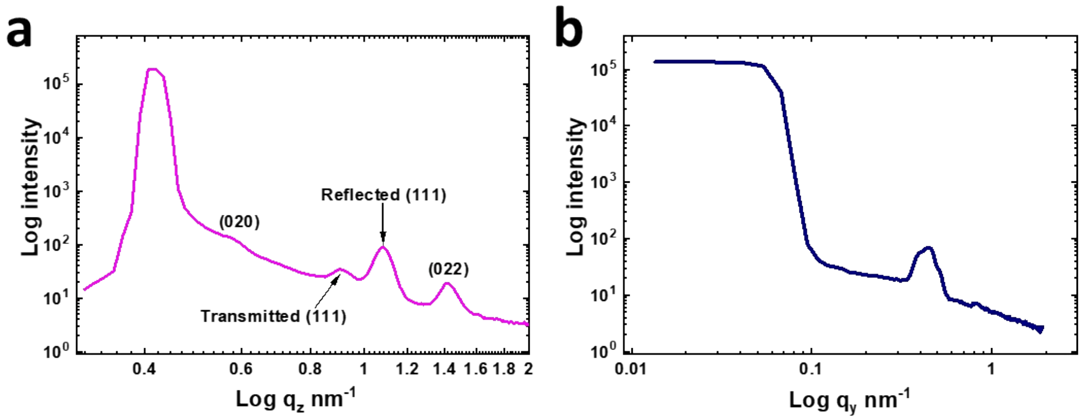

Figure S1 Vertical (a) and horizontal (b) integration of Figure 1b respectively. The incident angle was 0.3° and scanned for 20 minutes.

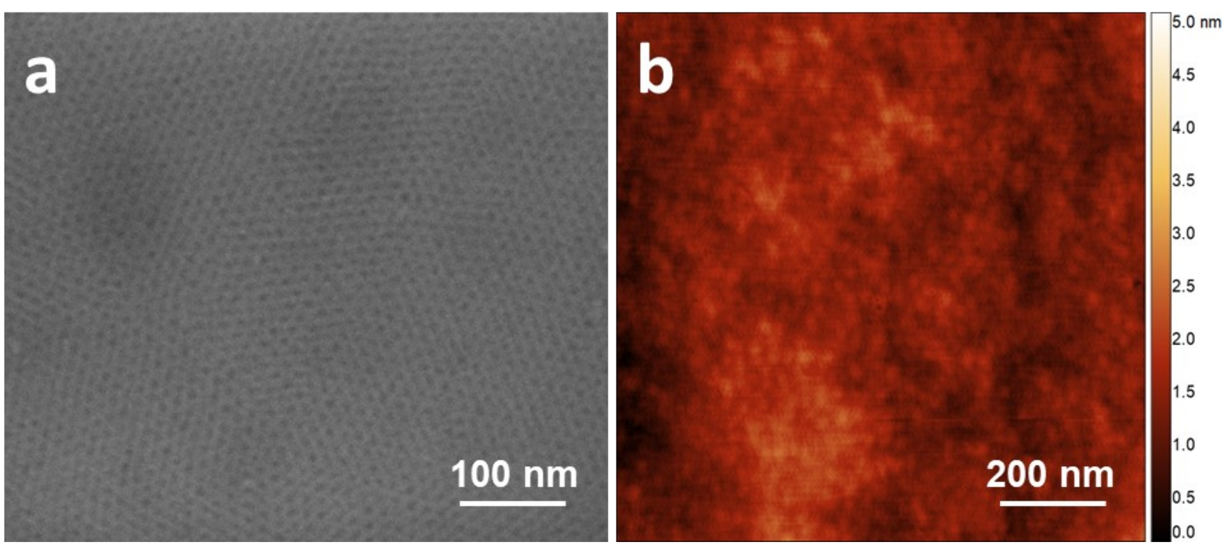

Figure S2 Top-view scanning electron microscopy image (a), and atomic force microscopy image (b) of the as-deposited mesoporous silica thin film.

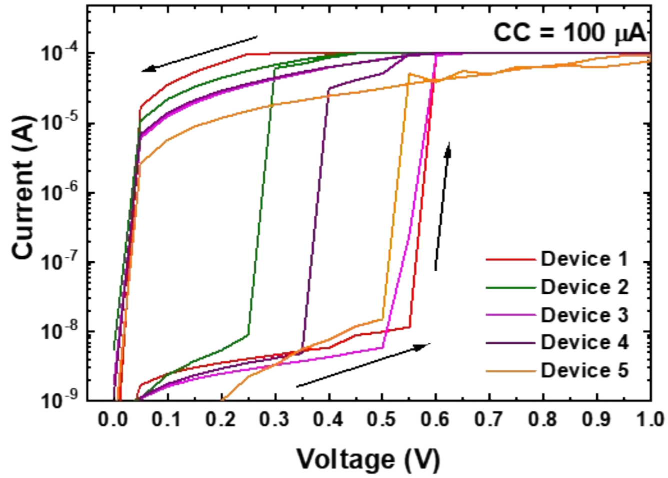

Figure S3 I-V characteristics of 5 different TiN/mSiO_2/Ag volatile memory devices under a CC of 100 μA.

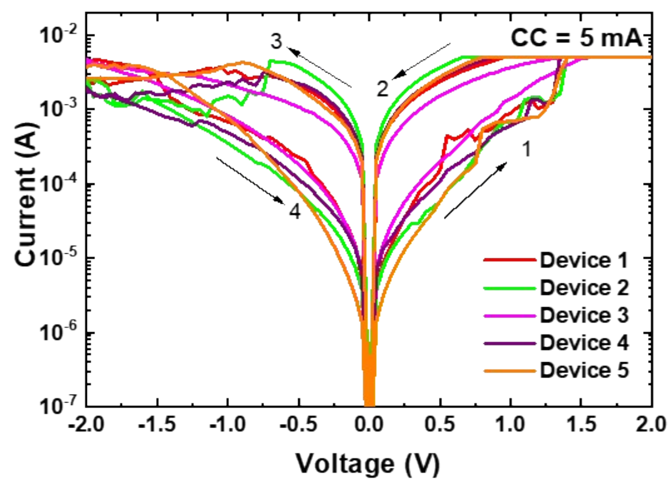

Figure S4 I−V characteristics for five different devices taken from different regions across the sample. Bipolar switching is observed in the devices when higher compliances currents are used.

The conduction mechanism of the nonvolatile memristor device is dominated by a space-charge-limited-current (SCLC), as presented in Figure S4. The log I-log V clearly shows four conduction regions for the transition from the HRS to the LRS. At low applied voltage region $V \leq 0.25$, the conduction mechanism is Ohmic, slope ~1, indicating charge transport via thermally generated free carriers. In the region of $0.3 \leq V \geq 0.8$, the current increases and shows the voltage square dependence, slope ~2. This is attributed to a trap-controlled-space-charge-limited-current (TC-

SCLC). At higher voltages in the region of 0.85 ≤ V ≥ 1.55, the current increases rapidly, slope ~3, switching the device to the LRS. We ascribe this to a trap-filled space-charge-limited-current (TF-SCLC). In the case of the LRS, the conduction mechanism is Ohmic, featuring by the linear dependence of current with voltage, slope ~1.

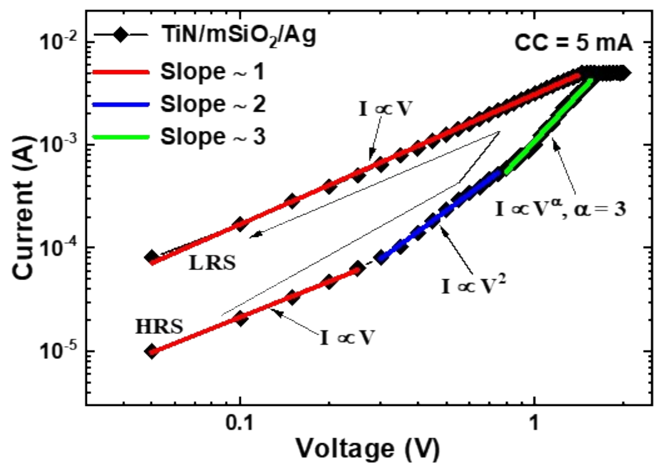

Figure S5 I-V sweep for the TiN/mSiO_2/Ag nonvolatile memristor device at CC of 5 mA demonstrating fits to the SCLC mechanism for the HRS to LRS transition.

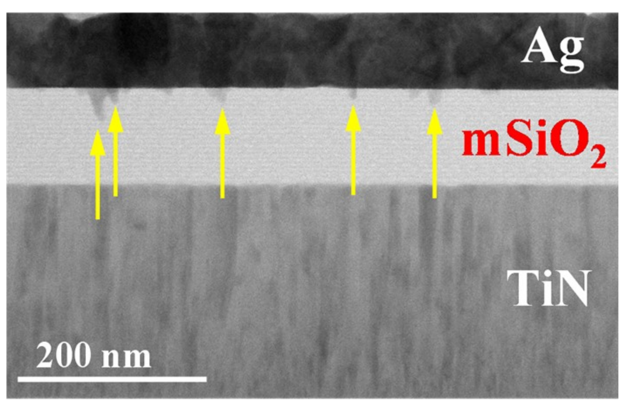

Figure S6 A cross section TEM image for the switched TiN/mSiO_2/Ag memristor device showing the conical nanoscale grown Ag filaments within the mSiO_2 insulator thin film.

A stretched-exponential based function (SEF) was used to evaluate the relaxation time of the volatile LRS state, as shown in Figure S7. The current level is modelled by an exponential equation $I(t) = I_0 e^{\left[-(t/\tau)^{\beta}\right]}$. Here $I(t)$ is the memory (resistance) level at a given time t, I_0 is the memory level at $t = 0$, τ is the characteristic relaxation time, which can be used to evaluate the forgetting rate. β is the stretch index and was fitted to be 0.5 in this work. The fitted relaxation time was found to be 600 μs.

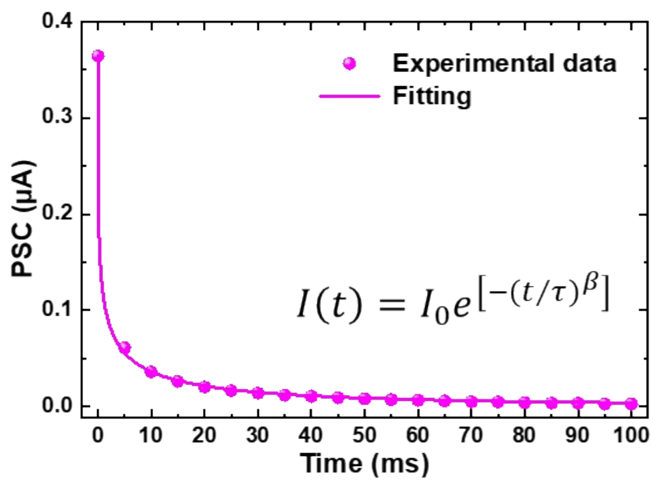

Figure S7 Retention data recorded (dots) after stoppage of SET pulses for the devices when in their volatile switching mode. The curve shown (solid line) is a stretched exponential fit (SEF) and indicates a relaxation time of 600 μs.

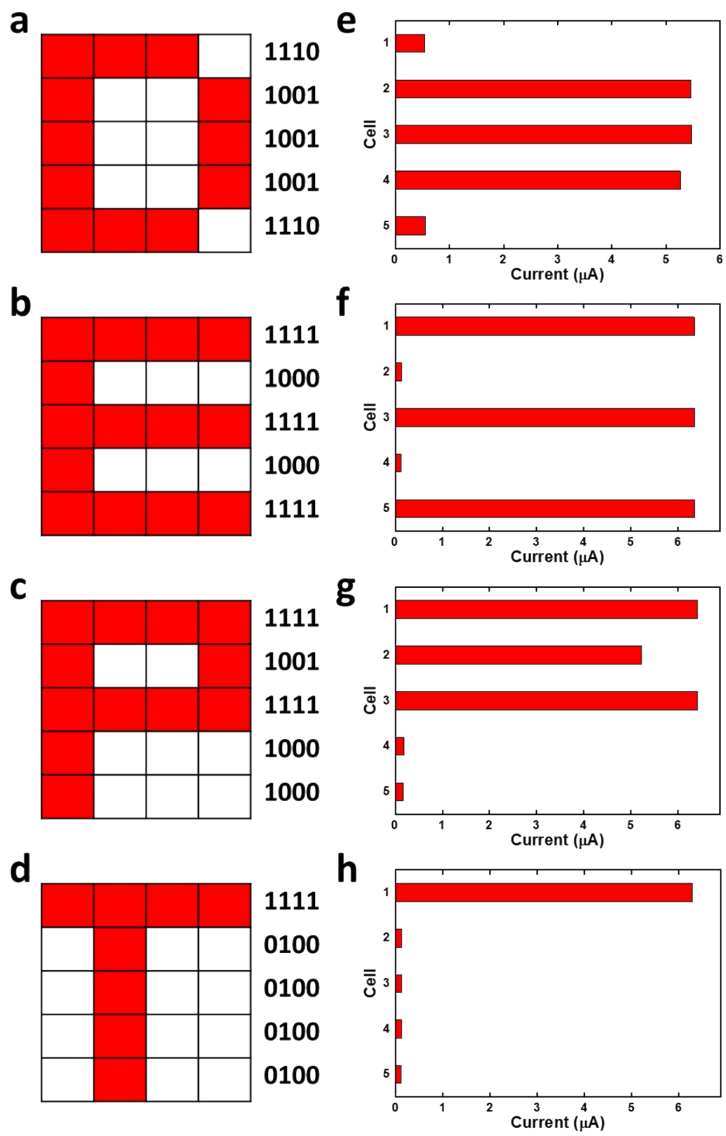

Figure S8 a) Letter "D", b) "E", c) "P", d) "T" divided into 5 x 5 pixels for letter recognition. e-f) the corresponding current responses after pulse sequence application for $mSiO_2$ memristor devices.

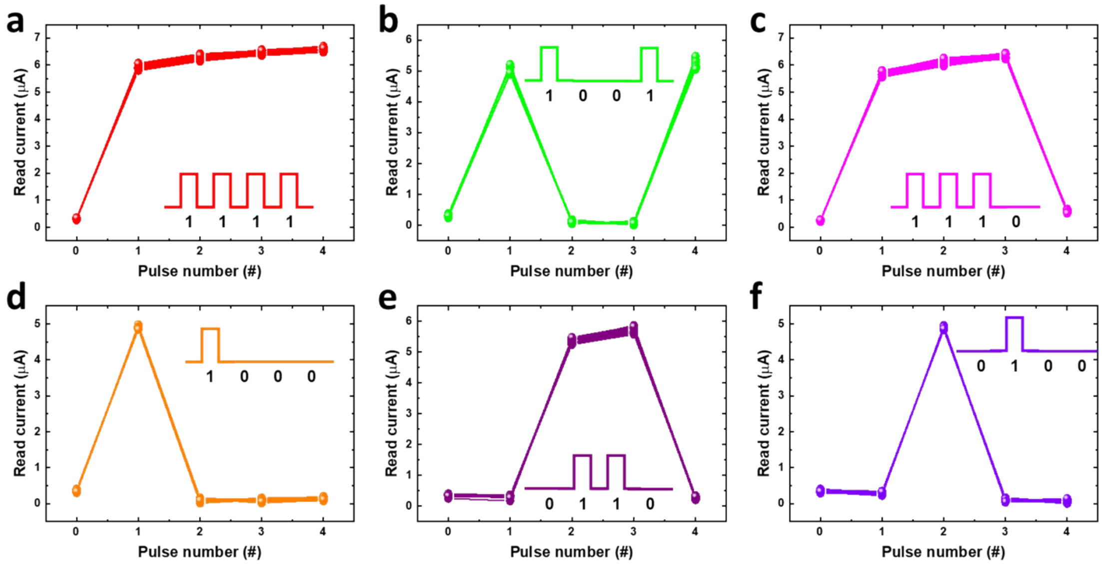

Figure S9 20 consecutive cycles for each of the six pulse streams, demonstrating the reproducible response of the $mSiO_2$ memristor devices to the input pulse streams.

The readout function is trained via the supervised learning algorithm to minimize the cross-entropy loss. Before the training process, the dataset, 570 groups currents-letter pairs, is experimentally measured from the memristors being fed with different pulse streams. Wherein, the current is scaled to the value with the unit of micron ampere (μA); the letters, "ADEPT", are represented by the digits of 0-4 in the dataset, respectively (shown in Table S1). The obtained dataset is then split into three parts, which are training set (80%), validation set (10%) and testing set (10%). The training set is utilized to train the neural network. The validation set works to monitor whether the neural network is overfitting or underfitting. The last part, testing set, that has not been seen by the network, is to measure the network performance. In the training and validation process, a softmax activation function (Equation S1) follows the weight matrix to scale the probability vector's value in the range of 0-1 with sum of 1. This non-linear operation is necessary and beneficial when applying the cross entropy (Equation S2) as loss function in a multinominal logistic regression. This loss function projects the probability distribution difference between prediction ($\hat{y}$) and ground truth (y), so softmax could prevent the value loss function less than 0 and make the training process converge fast and improve the network robustness. Note that the softmax activation function is not used in the testing process, since the non-linear behavior is hard to carry out in the circuit. However, this removal of softmax will not change the classification results. Although softmax normalizes the sum of the probability vector and enlarges the probability difference among the values in the vector, it does not change the value orders in the probability vector, so the largest value still stays in the same index as before removing softmax. Therefore, we can train the neural network with softmax to boost the learning speed and stability, but only use the weight matrix to do the classifications in practices. The previous discussed neural network training is processed on a Windows PC with the CPU of Intel Core i9-9900K and the GPU of NVIDIA RTX 2070 by the open-source machine learning package PyTorch. The detailed training hyperparameters are listed in Table S2.

$$x_i' = \frac{e^{x_i}}{\sum_i^n e^{x_i}} \#S1$$

$$CrossEntropy = -\sum_i^n y_i \log \hat{y}_i \#S2$$

Table S1 The mapping relationship between labels and letters.

Label	0	1	2	3	4
Letter	A	D	E	P	T

Table S2 The neural network training hyperparameters.

Hyperparameter	**Values**
Epochs	200
Learning rate	$1 \times 10^{-3} \times \left[1 - \mathrm{Max}\left(0, \frac{epoch - 100}{100}\right)\right]$
Batch size	8
Optimizer	Adam (default)
Loss function	Cross entropy