Tuning resistive switching behavior by controlling internal ionic dynamics for biorealistic implementation of synaptic plasticity

*Sangmin Yoo, Yuting Wu, Yongmo Park and Wei D. Lu**

S. Yoo, Y. Wu, Y. Park, Prof. W. D. Lu

Department of Electrical Engineering and Computer Science

The University of Michigan

Ann Arbor, MI 48109, USA
*E-mail: wluee@umich.edu

Keywords: 2nd order memristors, STDP, neuromorphic computing, SNN, correlation detection

-Abstract

Memristive devices have demonstrated rich switching behaviors that closely resemble synaptic functions and provide a building block to construct efficient neuromorphic systems. It has been demonstrated that resistive switching (RS) effects are controlled not only by the external field, but also by the dynamics of various internal state variables that facilitate the ionic processes. The internal temperature, for example, works as a second-state variable to regulate the ion motion and provides the internal timing mechanism for the native implementation of timing- and rate-based learning rules such as spike timing dependent plasticity (STDP). In this work, we show that the 2nd state-variable in a Ta_2O_5-based memristor, its internal temperature, can be systematically engineered

by adjusting the material properties and device structure, leading to tunable STDP characteristics with different time constants. When combined with an artificial post-synaptic neuron, the 2nd-order memristor synapses can spontaneously capture the temporal correlation in the input streaming events.

1. Introduction

The exponential growth of integrated circuits' density and functionality due to transistor scaling has recently slowed down as the device size approaches sub-10nm[1]. To meet the demands of modern applications such as artificial intelligence (AI), autonomous vehicles, and internet-of-things (IoT), advanced devices and computing architectures are needed to achieve efficient large-scale data analysis and storage in real time. Memristive devices offer high scalability, non-volatile storage and rich switching dynamics that make them appealing candidates for in-memory computing and neuromorphic computing systems.[2–4] A standard memristor is a two-terminal device with a switching medium between the two electrodes (**Figure 1**a). It stores data in the form of different conductance values, as a result of internal ion (e.g. oxygen vacancy V_O) re-distribution in the medium driven by the externally applied electric field.[5,6] Meanwhile, Joule heating induced during the process exponentially increases the drift velocity and the diffusivity of oxygen vacancies, and thus plays a critical role in the RS process[7]. In our previous work[8], we have demonstrated that the internal device temperature T can be considered as a 2nd state-variable, which though not directly measurable can strongly affect the 1st state-variable (the size of the oxygen vacancy-based filament) evolution and the RS characteristics. Since the internal temperature rises during programming and spontaneously decays after removal of the applied pulses, heat can be accumulated from multiple pulses, depending on how closely the pulses are applied. Thus, the device conductance change is not only determined by the present pulse but also by the relative timing of the pulses in the near history, allowing these so-called 2nd-order memristive devices to natively implement important synaptic plasticity effects such as spike

timing dependent plasticity (STDP). Essentially, the spontaneous decay of the internal temperature can be regarded as playing a similar role as the Ca^{2+} concentration in a biological synapse[9] to decode spike timing information, making it possible to achieve bio-realistic implementation of neuromorphic systems[10].

However, previous studies on 2nd order memristors lack controllability of the 2nd state-variable. Instead, different shapes of signals[3,11,12] or additional heating pulse were used to achieve the desired STDP function[8,13]. In this work, through detailed physical modeling and experiments, we study methods to control the 2nd state-variable (internal temperature) and its effects on the switching dynamics of a 2nd-order memristor, including how the STDP time constant can be modified. When using the natively implemented STDP learning rule for unsupervised learning in a spiking neural network[14], we show the 2nd-order devices can naturally uncover the correlation pattern in input spiking events.

2. Results and discussions

As discussed earlier[8], a memristor's 2nd state variable can play a critical role to the evolution of the 1st state variable and the resulting RS characteristics. For example, the internal temperature T exponentially changes the ion diffusivity (D) and drift velocity (v):

$$D = \frac{1}{2} \cdot a^2 f exp(-\frac{E_a}{kT}) \quad (1)$$

$$v = af exp(-\frac{E_a}{kT}) \cdot \sinh(qaE/2kT) \quad (2)$$

where f is the escape-attempt frequency, a is the effective hopping distance, and E_a is the activation energy for V_O migration.[15] These parameters and temperature gradient drive the oxygen vacancy drift/diffusion, as described below:

$$\frac{dn_D}{dt} = \nabla \cdot \left(D\nabla n_D - vn_D + DS_{n_D}\nabla T\right) \quad (3)$$

Where n_D is the local oxygen vacancy density, $D\nabla n_D$ and vn_D are the Fick diffusion flux and the drift flux terms, and $DS_{n_D}\nabla T$ corresponds to the Soret diffusion effect, respectively.[15,16] Thus, a memristor's RS characteristics are strongly affected by the internal Joule heating and heat dissipation dynamics, which can be mathematically described as

$$\frac{dQ}{dt} = -\frac{kA(T_1(t)-T_2)}{d} = c\rho\frac{dT(t)}{dt} \quad (4)$$

where k is the thermal conductivity, A is the area of the device, d is the distance between two temperature points T_1 and T_2, c is the specific heat and ρ is the material density (Figure 1a). Here T_1 represents the switching region temperature, and can be calculated as a function of t:

$$T_1(t) = (T_1(0) - T_2)\,exp\left(-\frac{kA}{c\rho d}t\right) + T_2 \quad (5)$$

Where T_2 is the ambient temperature. From (5), the time t required for T_1 to reach a specific temperature can be derived:

$$t = \ln\left(\frac{T_1(0)-T_2}{T_1(t)-T_2}\right)\frac{c\rho d}{kA} \quad (6)$$

where the dimensions of the device structure and the intrinsic material properties (c, ρ, d, k, A) are known beforehand. Based on this understanding, we aim to control the internal temperature dynamics by adjusting these material and device parameters.

Since temperature rise and heat dissipation normally occur very fast (e.g. nanoseconds), it is generally desirable to slow down the temperature dynamics to more efficiently take advantage of the

2nd-order effects. For example, if the devices are used to directly process biological spike trains, a time constant ~ ms is desired. To achieve a longer time constant, Equation (6) suggests that heat conduction paths with smaller A and lower k are required. Experimentally, we inserted heat insulating (HI) layers (NiCr), whose thermal conductivity value is about one-fourth of the top and bottom electrode (Pd), between T_1 and T_2 on both sides (Figure 1b) to test k dependency. We also tested devices with different sizes, from 1μm^2, 4μm^2, 20μm^2, 40μm^2 to 100μm^2, to verify A dependency. Figures 1b,c show the schematic cross-section view and the top view of such a thermally enhanced device having heat insulating layers in the extended top and bottom electrodes, respectively.

A detailed device model[15] based on the COMSOL Multiphysics tool was used to calculate the temperature profile and the oxygen vacancy density at various stages during device switching. Figure 1d shows the V_O profile of the devices immediately after the forming process, with the upper panel showing the reference device without the HI layer (standard device), and the lower panel showing the device with the inserted HI layers (HI-device). The two cases show different V_O distribution profiles that lead to different conductance values. Figure 1e plots the internal temperature distribution immediately after filament formation for the two cases. Thanks to the HI layers, the proposed HI-device can maintain the induced heat better, as shown in the lower panel. The enhanced thermally assisted ion diffusion/drift in turn results in the wider V_O filament in the switching layer of the HI-device, compared with the standard device (Figure 1d).

Figure 1f shows the electrical measurement results for a standard device and a HI-device, respectively. Compared with the standard device, the HI-device can be switched at a lower voltage and results in a higher on-current, which can be explained by its better heat accumulation capability as supported by the simulated temperature profiles and V_O distribution profiles discussed above (Figures 1d-e). Similarly, better heat trapping also allows the HI-device to reset at a lower voltage, due to its

elevated mobility at the elevated internal temperature, as explained by Equation (2). These experimental RS characteristics are in turn reproduced and supported by the Multiphysics simulation results (Figure 1g), for the same processes and device structures.

The Ta_2O_5-based memristors also exhibit analogue conductance modulations (Figure 1h,i), when stimulated with consecutive potentiation or depression pulses. In the measurements, 100 depression pulses were applied first to induce long-term depression (LTD), followed by 100 potentiation pulses to induce long-term potentiation (LTP). Due to better heat accumulation, the HI-device (Figure 1i) shows a wider conductance modulation range, even though pulses with lower amplitude (0.6V/0.85V) were applied during the LTP/LTD processes.

Temporal properties of these 2nd order memristors are also verified both computationally and experimentally. Simple, non-overlapping square pulses were used in these tests, as depicted in **Figure 2**a. Pre- and Post- synaptic pulses were applied to the top and bottom electrode, respectively. Each pulse is designed to have a low amplitude that by itself is not enough to evoke conductance change. As discussed earlier, the time gap between pulses (Δt) plays a critical role in determining the internal temperature and allows the device to naturally decode input timing information. Figure 2b shows the simulated internal device temperature induced by two pulses with different Δt. As expected, smaller Δt leads to higher internal temperature because the residual heat from the 1st pulse is largely accumulated during the 2nd pulse. To characterize the temporal dynamics of the heat decay process, we define the time window as the time gap between the two pulses in a pulse pair within which a measurable conductance change (*i.e.* 1μS) can be induced. When Δt is within the heat decay's time window the sufficiently elevated temperature experienced at the 2nd pulse can induce a conductance change (Figure 2c). In particular, the polarity and amplitude of the conductance change depend on the sign and value of Δt , leading to the natively implemented spike-timing dependent plasticity rules, as

shown in Figure 2c. The experimental data are consistent with simulation results based on the Multiphysics device model, further supporting the roles of the internal dynamics in native STDP implementation.

More results can be obtained by repeatedly applying the STDP pulse pairs and observe the device response. The first half of Figure 2d shows the device response to 80 pulse pairs with a preceding post-synaptic pulse followed by a pre-synaptic pulse, while the 2nd half showed subsequent response to 80 pulse pairs with a preceding pre-synaptic pulse followed by a post-synaptic pulse. The measurements were also performed for two Δt conditions. While pulse pairs with a 100ns time gap causes clear conductance modulations, the 10μs time gap pulse pairs cannot evoke conductance modulation even after repeated applications. These measurements again illustrate the importance of the short-term dynamics of the internal temperature to achieve STDP characteristics.

The device model further suggests how to tune the internal short-term heat dynamics. For example, based on Equation (6), the heat dynamics depend on material properties such as the heat capacity c, mass density ρ, thermal conductivity k, and device geometry A and d. In the following, we discuss experimental and simulation results by tuning these parameters to control the 2nd order memristor's time constant and the STDP behavior.

First, a larger device area facilitates faster cooling, as evidenced in Equation (3). **Figure 3**a shows the simulated thermal distribution inside memristors having different electrode areas of 1μm^2 and 100μm^2, respectively. The results are captured at the beginning of the forming process. One can see that the peak temperature, which drives the ion migration, is significantly lower in the larger area device. This effect may be explained by the fact that the electrodes act as heat sinks to dissipate heat

produced by Joule heating, so devices with wider electrodes will have much faster heat dissipation than devices with narrower ones, thanks to its larger thermal conductance $(\frac{kA}{d})$ [17].

STDP characteristics measured from standard devices with different areas support these simulation results (Figure 3b). In the measurement, each conductance change value is calculated after the application of 100 pulse pairs, comprised of a pre-synaptic pulse of 0.9V and a post-synaptic pulse of 0.8V with 100ns pulse width, at a given Δt, and the results are averaged from 3 such measurements to reduce variations. The device is then reset to the same initial condition for the next measurement. t_1 and t_{100} represent time windows of the $1\mu m^2$ and $100\mu m^2$ device, respectively. As expected, the device with a smaller electrode area ($1\mu m^2$) can afford a longer Δt (time window) for STDP owing to the slower cooling effect, while the device with a larger electrode area ($100\mu m^2$) requires a much shorter time window. Figure 3c plots the dependency of the required time window on the device area: the experimentally measured time window is inversely proportional to the device area, agreeing with the analytical expression in Equation (6).

Besides the time window Δt required for minimum conductance change, the characteristic time constant τ in the STDP response can also be obtained by fitting the LTP and LTD portions following:

$$\Delta G = B_G * exp\left(-\frac{|\Delta t|}{\tau} \right) \quad (7)$$

The STDP time constants are extracted from devices with different areas presented in Figure3b and are plotted in Figure 3d, showing the STDP time constant also decreases with increasing device area. This observation can also be explained by the device model. Specifically, based on Equation (6) and (7), τ can be derived as:

$$\tau = \frac{c\rho d}{kA} \cdot \frac{\ln(\frac{T_1(t)-T_2}{T_1(0)-T_2})}{\ln(\frac{\Delta G}{B_G})} \quad (8)$$

, showing inverse dependence of A.

Figure 3e and Figure 3f show the dependency of STDP characteristics on pulse amplitudes for the two cases. The pulse amplitude has a strong effect on the size of the conductance change, but a weak effect on the STDP time constant. This observation can again be explained by the heat dynamics (*i.e.* Equation (6) and (8)), where the internal short-term dynamics is only determined by material and device parameters where the applied pulse amplitude only has an indirect effect. On the other hand, tuning the pulse amplitude may be useful if a larger or smaller STDP effect is desired without changing its temporal characteristics.

We next study the 2nd-order memristor's dependency on the thermal conductivity in the thermal conducting path. As introduced in Figure 1b and 1c, an additional low thermal conductivity layer (HI layer) is added between the anticipated high temperature source point (T_1 in Figure 1b) and the cooling point (T_2 in Figure 1b) to slow down the cooling process. The effect of the HI layer is first simulated, as shown in **Figure 4**a (with HI) and 4b (without HI). Each frame in the figures corresponds to a time instant after the application of a programming pulse, and the plots show the internal temperature distribution evolution for the two cases. As suggested by Equation (6) and (8), HI helps the device to reach a higher internal temperature, and traps heat longer than the standard device without the HI layer. Figure 4c plots the evolutions of the peak internal temperature for the two cases, highlighting the effects of inserting the HI layer in increasing the peak temperature and slowing down the heat dissipation process.

These effects were then tested in experimentally fabricated devices with the HI layer using post-synaptic pulse of 0.7V and pre-synaptic pulse of 0.9V with 100ns pulse width. Specifically, Figure 4d plots the STDP results obtained from a HI-device and a standard device, both having the same area of 1μm^2. While the standard device exhibits an STDP time window shorter than 1μs, the HI-device exhibits a much longer time window of ~1ms. The time constant of the 1μm^2 standard device and HI-device shown in Figure 4d are (τ_d: 808.4ns, τ_p: 728.9ns) and (τ_d: 496.0μs, τ_p: 353.2μs), respectively. The ability to slow down the heat dissipation and increase the STDP time window to ms level may open opportunities for the 2nd-order memristor devices and networks to directly interact with temporal signals at these time scales, *i.e.* neural recordings from biological networks[18]. Similar to the standard device, the time constant of the HI-device can also be adjusted by tuning the device area (A_1), as shown in Figure 4e. with time constants of (τ_d: 496.0μs, τ_p: 353.2μs) and (τ_d: 465.3μs, τ_p: 324.7μs), for the 1 μm^2 and 4 μm^2 devices, respectively. The fitting parameters for Equation (7) are summarized in Table 1 below.

Table 1. Fitting parameters for the measurement data of devices.

Device	Layer stack	Cross-sectional Area	Time constant of LTP (τ_p)	Time constant of LTD (τ_d)	Fitting parameter of LTP, B_{Gp}	Fitting parameter of LTD, B_{Gd}
Standard device	Pd (35nm) –	1μm^2	728.9ns	808.4ns	0.00282	-0.00286
	TaO_x (30nm)– Ta_2O_5 (4nm) – Pd (30nm)	100μm^2	58.6ns	68.8ns	0.04373	-0.01120

HI-device	NiCr (40nm) - Pd (35nm) - TaO_x (30nm) - Ta_2O_5 (4nm) - Pd (60nm) – NiCr (70nm)	$1\mu m^2$	353.2µs	496.0µs	0.00243	-0.00208
		$4\mu m^2$	324.7µs	465.3µs	0.00156	-0.00141

Finally, we verify the feasibility to perform efficient temporal computation with a 2^{nd}-order memristor network through the natively implemented of STDP learning rule with device-fitted simulation. Spiking neural networks have attracted increasing interest in recent years, as they offer an appealing opportunity for highly efficient computation with sparse, binary spikes. It has been shown that STDP can help learn the input correlations by exploiting the timing information embedded in the spiking sequences[19], and STDP memristor-based networks have been used to perform synaptic connection pattern reconstruction[18], coincidence detection[20], and correlation pattern detection[21,22]. However, in general these prior studies do not rely on internal dynamics to perform computing, and the conductance changes need to be first computed in software and then programmed into the devices. Below we show that by taking advantage of the internal dynamics, the 2nd order memristors can directly uncover the temporal features hidden in the input spiking events, without external software computation and manual weight updates.

In this study, the inputs are 100 spike trains each consisting of around 1000 spiking events, where the different input streams have different degrees of correlation (see Methods for more details). As shown in **Figure 5**b, each pre-synaptic input stream is fed into one memristor device, whose normalized conductance value represents the synaptic efficacy. The summation of the weighted post-synaptic potential then updates the membrane potential V_m of the post-synaptic neuron based on the leaky-integrate-and-fire (LIF) neuron model. The post-synaptic neuron fires an action potential if V_m

surpasses the defined threshold value V_{th} (Figure 5a), and V_m is reset to the resting state after firing. During this process, the 2nd-order memristive devices naturally adapt their conductance states based on the relative timing between the pre- and post-synaptic spikes. Compared with random inputs, correlated inputs have a higher probability to precede a post-synaptic spike, as the arrival times have higher chance to coincide with each other to cause a larger increase in V_m[19]. The pre-post spike pairs then lead to potentiation of the corresponding synapses, which further increases the tendency to induce post-synaptic events. After this process, the correlated inputs can be identified from the uncorrelated ones, reflected by the clear separation of device conductance into several groups based on the degree of input correlation, as shown in Figure 5d and 5e. Inputs with higher correlation induce faster learning and higher memristor conductance values, while uncorrelated inputs do not induce significant conductance changes, as the LTP and LTD effects essentially cancel out each other, resulting in conductance distributions that match the correlation pattern in the inputs.

Beyond the assumption that all devices are identical, we performed additional studies to evaluate how device-to-device (D2D) variations would impact the system performance. D2D variations drawn from a Gaussian distribution have been added to the amplitude (A_p) and relaxation time (τ_p) of the STDP behavior. Standard deviations from 5%-15% have been simulated, which is comparable to the experimental measurement results in actual devices. Figure 5f shows that with increasing variations in A_p, the separation between groups of different correlation coefficients start to shrink. On the other hand, even with 15% percentage of standard variation, no overlapping between groups is observed. We also studied the influence of D2D variations of relaxation time. Figure 5g shows that the final histograms with different levels of τ_p variations are almost identical, meaning that the system is insensitive to the variation of relaxation time. Therefore, the purposed system exhibits robustness to D2D variations.

3. Conclusion

We demonstrate the ability to control the STDP characteristics, including the characteristic time constants of a 2nd order memristor, by controlling the device's material parameters and structural dimensions. In contrast, different pulse amplitudes modify the size of the conductance changes without changing the characteristic time window. The ability to slow down the heat dissipation processes and extend the STDP time window allows the 2nd-order memristors to potentially be tailored to process temporal signals with specific temporal characteristics, such as neural recordings from biological networks. When integrated with a post-synaptic artificial neuron, the 2nd order memristor networks can capture the correlation in the pre-synaptic spikes in an unsupervised fashion, leading to inhomogeneous conductance distributions that match the correlation pattern of the input data.

Experimental Section/Methods

Device model: We developed the COMSOL Multiphysics model following the previous work[15], and extends it to 3D device structures. This allows us to tune the electrode material properties and the structural dimension to study their effects on the internal dynamics of the 2nd-order memristor. We assumed that the top and bottom electrodes are effective heat sinks so that 1μm away from the device the temperature of the electrode reaches room temperature. The memristive devices and interconnecting wires are assumed to be surrounded by silicon dioxides. The peak temperature value in the switching layer is obtained from the temperature profile simulation at given time instances. Material properties, including heat capacity, thermal conductivity, and mass density are chosen from reported thin film values from literature[23–25]. The thermal conductivity is obtained using the following reported model as a function T:[26]

$k = 0.2T+4$ (W/(m·K))

We used two separate physical structures as schematically shown in Figure 1a),b) to model devices with and without the NiCr-based HI layer, respectively. Other physical assumptions are identical to our previous work[15]. The parameters used are listed in a table below.

Table 2. Material properties used in the COMSOL Multiphysics simulation.

Material	Specific Heat [J/(kg·K)]	Thermal conductivity [W/(m·K)]	Density [kg/m^3]
NiCr	380	17	7750
Pd	240	71.2	1202

Device fabrication: We fabricated TaO_x based memristors with different areas through photolithography. The layout for the standard devices includes memristors with sizes from 1μm^2 to 100μm^2. Device fabrication starts with 35nm Pd bottom electrode deposition by photolithography, e-beam evaporation and lift-off. It is followed by sputtering of 30nm TaO_x using a Ta metal target in an Ar/O_2 gas mixture (3% O_2) at 400°C. The pressure of the gas is ~5mTorr. A 4-nm $Ta_2\ O_5$ switching layer is then deposited by RF sputtering using a $Ta_2\ O_5$ ceramic target in Ar with a pressure of ~5 mTorr. Top electrode is deposited in the same way as the bottom electrode[8,27]. For the HI-device, 40nm and 70nm NiCr HI layers are added before the deposition of the bottom electrode and after that

of the top electrode, respectively. The thickness of the top electrode is increased to 60nm to ensure continuous connection over the step edges of the added HI layer.

Correlation detection simulation: The inputs have the same average firing rate (r) of ~500Hz. Group 1 (neuron 0-9) has the pairwise correlation parameter (c) of around 0.1, and group 2 (neuron 10-19) has correlation of around 0.2. The spiking events of the other 80 neurons are uncorrelated. The generation of the correlated spiking inputs follows the methods described in [19]. The correlated spike trains $X_k(t)$ are produced by conditioning their firing probabilities on the activity of a common reference spike train $X_O(t)$. The equations of the conditional probabilities are shown below.

$$\vartheta = P(X_k(T) = 1|X_o(T) = 1) = r\Delta T + \sqrt{c}(1 - r\Delta T) \quad (9)$$

$$\varphi = P(X_k(T) = 1|X_o(T) = 0) = r\Delta T(1 - \sqrt{c}) \quad (10)$$

The pairwise correlation between two generated spike trains $X_i(t)$ and $X_j(t)$ can be calculated as following, which is equal to c.

$$E[X_iX_j] = \sum_{m=0,1} P(X_i(T) = 1|X_o(T) = m) * P(X_j(T) = 1|X_o(T) = m)$$

$$= \vartheta^2 r\Delta T + \varphi^2(1 - r\Delta T) = (r\Delta T)^2 + c(r\Delta T)(1 - r\Delta T) \quad (11)$$

$$Corr(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}} = \frac{E[X_iX_j] - E[X_i]E[X_j]}{\sqrt{\left(E[X_i^2] - E[X_i]^2\right)\left(E[X_j^2] - E[X_j]^2\right)}}$$

$$= \frac{(r\Delta T)^2 + c(r\Delta T)(1 - r\Delta T) - (r\Delta T)^2}{(r\Delta T)(1 - r\Delta T)} = c \quad (12)$$

The artificial neuron is modelled as a leaky-integrate-and-fire neuron. The membrane potential is updated every 2μs according to equation (13), where τ_m=100μs and the threshold voltage is set as 5.

$$\frac{dV_m(t)}{dt} = -\frac{V_m}{\tau_m} + \sum_{i=1}^{n} w_i X_i(t) \quad (13)$$

A multiplicative STDP learning rule is used for numerical simulation of conductance updates, where the parameters are fitted to the measurement data (area = 1μm^2 with HI layer) in Figure 4e. The parameters are set as following: A_p=0.23, A_d =0.23, τ_p=56.3 μs, τ_d=123.2 μs, η=0.01, stdp_window=200μs.

$$w = \begin{cases} w - \eta * \mathrm{w} * \mathrm{A_d} * \exp\left(\frac{\Delta t + 100\mathrm{ns}}{\tau_d}\right), & -200\mu\mathrm{s} \le \Delta t < 0 \\ 0, & \Delta t = 0 \\ w + \eta * (1.0 - w) * \mathrm{A_p} * \exp\left(\frac{\Delta t - 100\mathrm{ns}}{\tau_p}\right), & 0 < \Delta t \le 200\mu\mathrm{s} \end{cases} \quad (14)$$

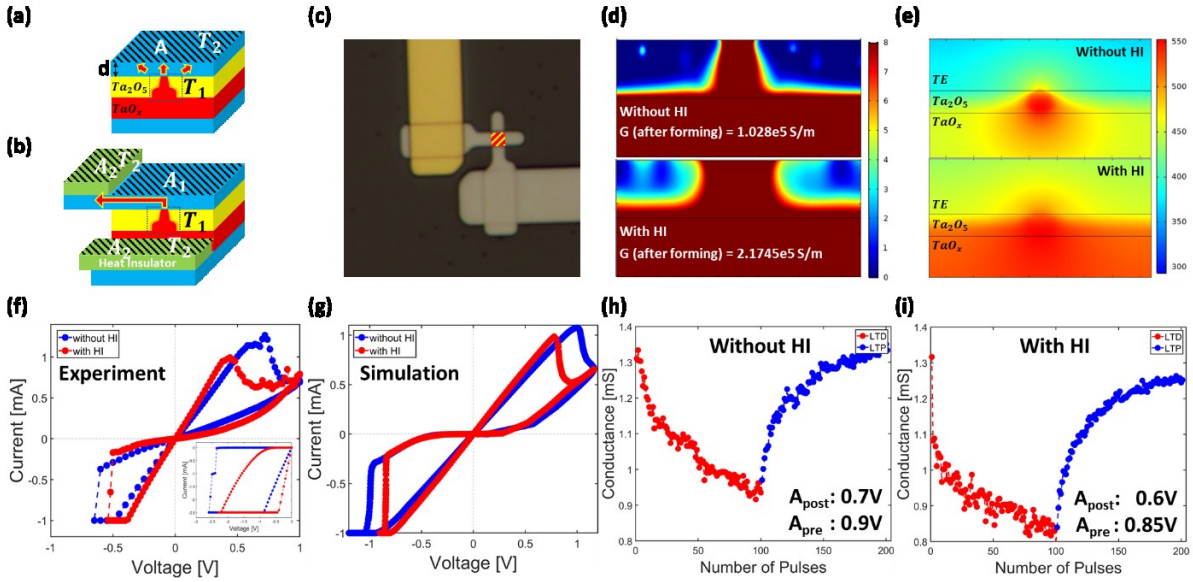


Figure 1. a-b) Schematic device structures for a device without (a) and with (b) the HI layers. c) Top view of a fabricated HI-device. d) Oxygen vacancy and e) Temperature profile of the two cases, showing the device states immediately after forming. f) Experimentally measured DC I-V curves of the two types of devices. g) Simulated DC I-V curves of the two types of devices from the COMSOL Multiphysics model. h-i) Analog conductance modulation results through consecutive pulses for the standard device (h) and HI-device (i).

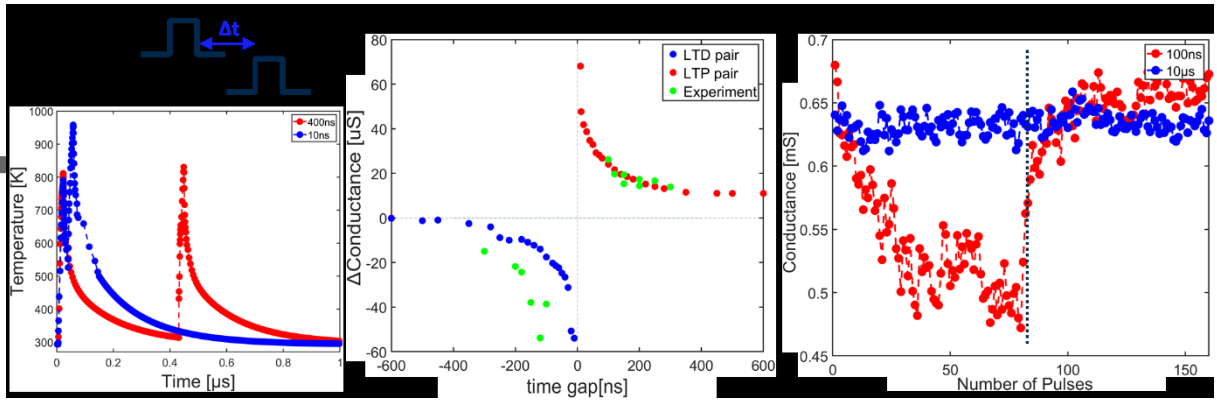


Figure 2. a) Schematic of the pre- and post-synaptic pulse pairs with a time gap (Δt) for STDP studies. b) Simulated internal temperature dynamics induced by pulse pairs, for two different Δt cases, where the maximum temperature inside the device was recorded as a function of time. c) Simulated (blue and red dots) and measured (green dots) STDP behaviors natively implemented in the standard device with device area 100μm^2. d) Long term conductance changes induced by sequentially applying pulse pairs, for two scenarios with Δt of 100ns and 10μs, respectively.

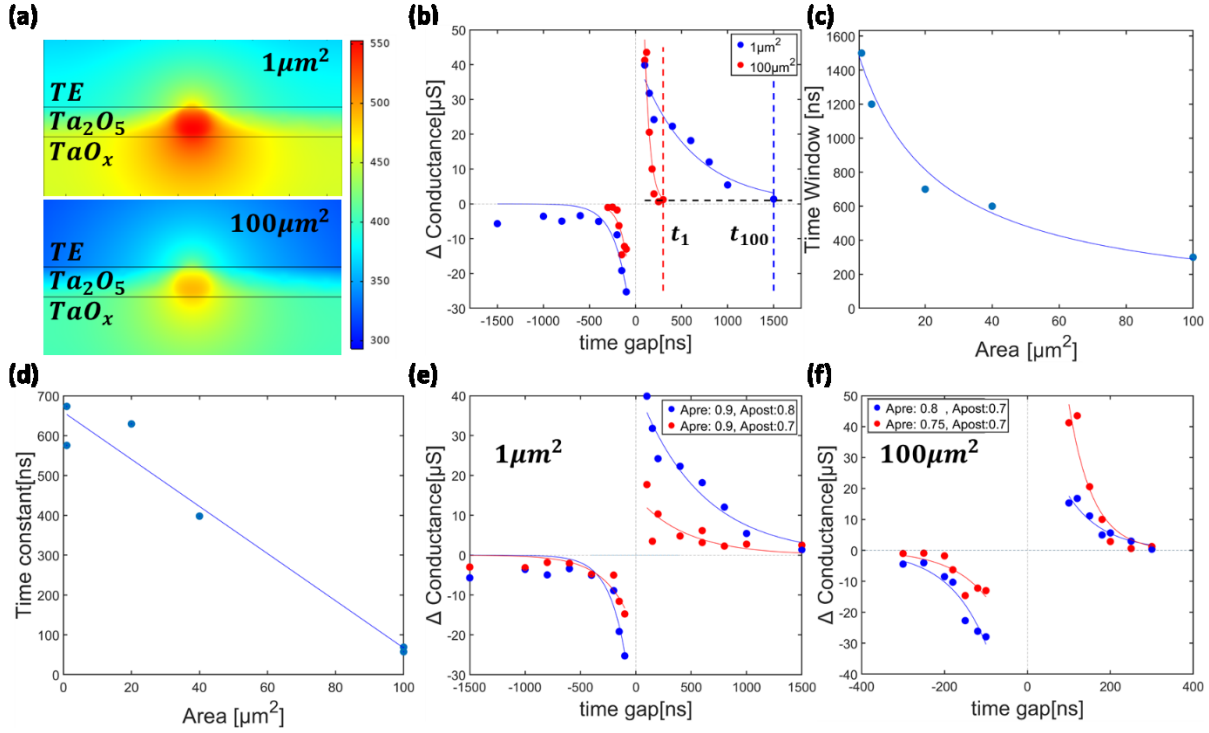


Figure 3. a) Internal temperature profiles of two standard devices with areas of 1μm^2 and 100μm^2, respectively. The temperature profiles were captured at the beginning of the filament formation. b) STDP behaviors experimentally measured from two fabricated devices with areas of 1μm^2 and 100μm^2, respectively. c) The extracted STDP time window as a function of device area. The solid line is a rational function fitting following Equation 6. d) The extracted STDP time constant as a function of the device area, along with a linear fit. e-f) STDP characteristics obtained with different pulse amplitudes for the two devices with areas of (e) 1μm^2 and (f) 100μm^2, respectively.

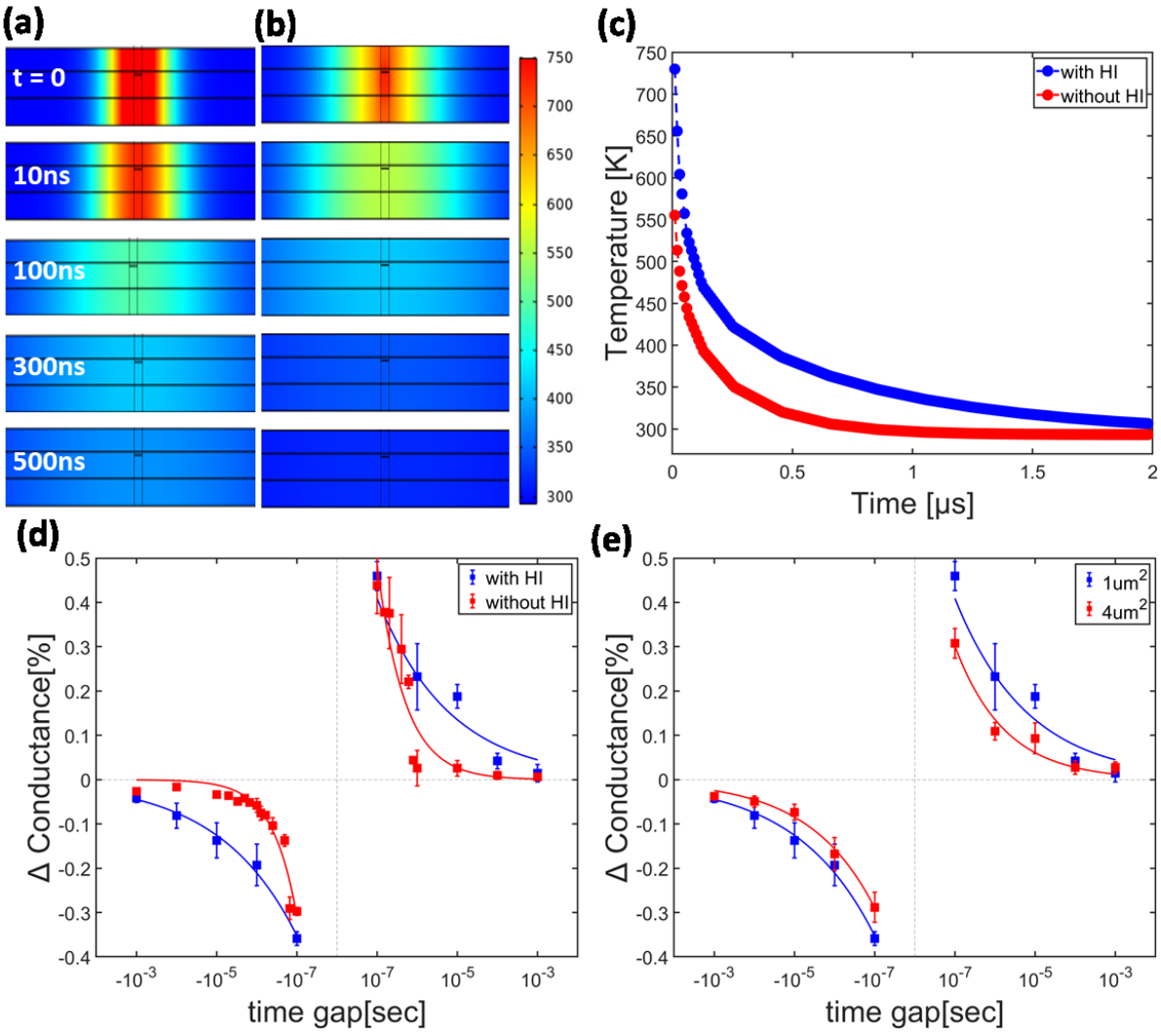


Figure 4. a-b) Visualization of the internal heat dissipation dynamics for an HI-device (a) and a standard device (b). The temperature profile was calculated at different time instances after the removal of a programming pulse. c) Internal peak temperature evolutions for the two devices obtained through COMSOL simulation. The peak temperature was recorded immediately after the removal of a programming pulse. d) STDP characteristics measured from the standard and the HI-device with 1μm^2 area, respectively. e) STDP characteristics measured from the HI-device having different areas.

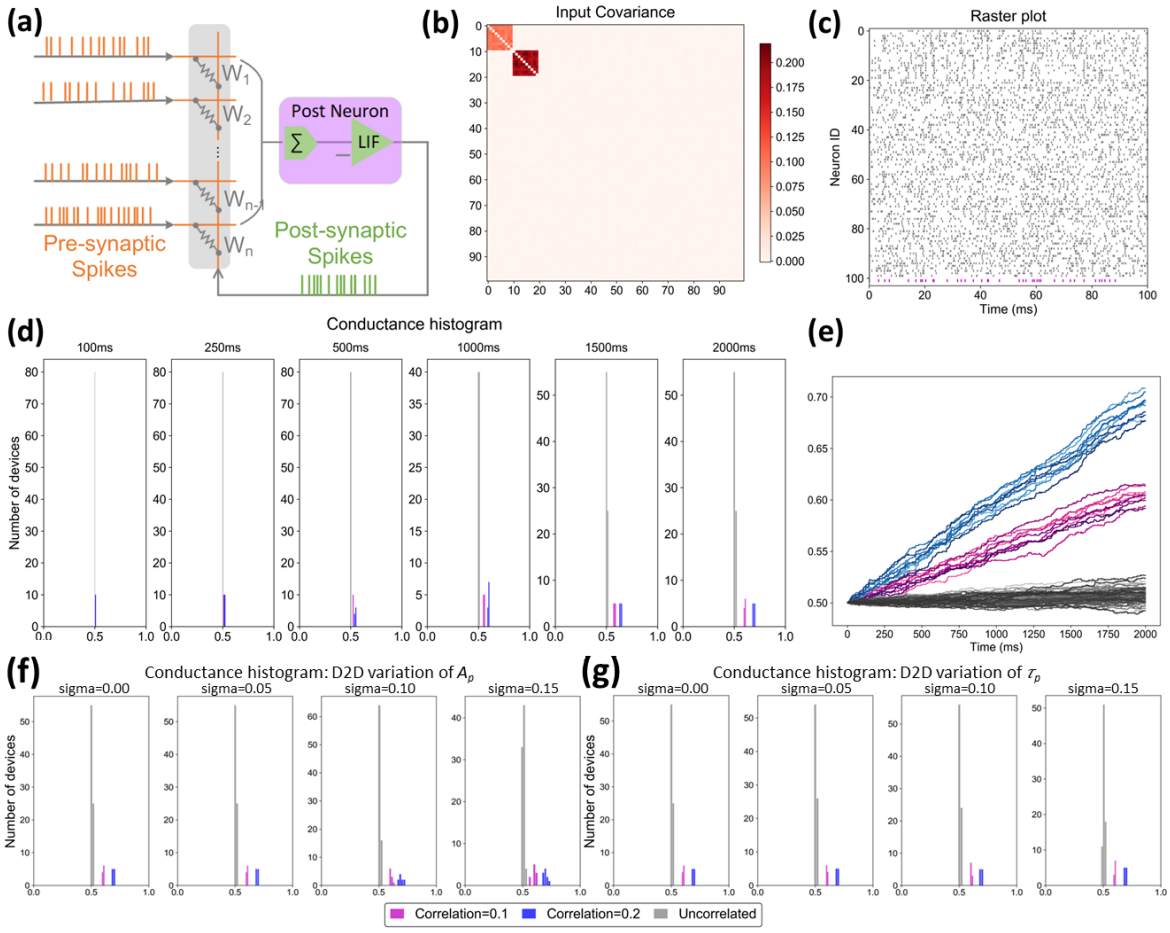


Figure 5. a) Scheme showing the correlation detection setup. The network consists of a single post-synaptic LIF neuron and an array of 2nd-order memristor synapses. b) Ground truth covariance matrix between the 100 input spike trains with fixed firing rate at 500Hz. The correlation within group 1 (Neuron 0-9) is around 0.1, and the correlation within group 2 (Neuron 10-19) is around 0.2. The other 80 spike trains are uncorrelated. c) Raster plot of the 100 pre-synaptic neurons (grey) and the post-synaptic neuron (magenta), showing the neurons' firing patterns. d) Histogram and e) Line plot of the normalized device conductance evolution upon application of the input spike trains, showing the clear separation of device conductance values that form groups corresponding to different input

correlation coefficients. The synaptic devices receiving inputs with the highest correlation (10-19) evolve to form a group having the highest conductance values, while the synaptic devices receiving uncorrelated inputs (20-99) experience essential no conductance changes. f) and g) Histogram of the normalized final device conductance with different levels of D2D variations of f) A_p g) τ_p. The D2D variations are assumed to be Gaussian in the studies, and the percentage values represent the standard deviation of the D2D variations normalized against the mean value.

Acknowledgements

The authors would like to thank Dr. S. H. Lee and Dr. M. Zidan for helpful and stimulating discussions. This work was supported in part by the National Science Foundation through awards ECCS-1915550 and DMR-1810119.

Received: ((will be filled in by the editorial staff))
Revised: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

References

[1] I. S. Devised, I. For, T. Assessment, I. S. Without, R. To, C. Considerations, P. To, I. Products, O. R. Equipment, **2013**.

[2] F. Cai, J. M. Correll, S. H. Lee, Y. Lim, V. Bothra, Z. Zhang, M. P. Flynn, W. D. Lu, *Nat. Electron.* **2019**, *2*, 290.

[3] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, W. Lu, *Nano Lett.* **2010**, *10*, 1297.

[4] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, E. Eleftheriou, *Nat. Nanotechnol.* **2020**, *15*, 529.

[5] C. Du, W. Ma, T. Chang, P. Sheridan, W. D. Lu, *Adv. Funct. Mater.* **2015**, *25*, 4290.

[6] Y. J. Jeong, S. Kim, W. D. Lu, *Appl. Phys. Lett.* **2015**, *107*, DOI 10.1063/1.4934818.

[7] S. Kim, S. J. Kim, K. M. Kim, S. R. Lee, M. Chang, E. Cho, Y. B. Kim, C. J. Kim, U. -In Chung, I. K. Yoo, *Sci. Rep.* **2013**, *3*, 1.

[8] S. Kim, C. Du, P. Sheridan, W. Ma, S. Choi, W. D. Lu, *Nano Lett.* **2015**, *15*, 2203.

[9] S. Yang, Y. Tang, R. S. Zucker, *J. Neurophysiol.* **1999**, *81*, 781.

[10] M. A. Zidan, Y. J. Jeong, W. D. Lu, *IEEE Trans. Nanotechnol.* **2017**, *16*, 721.

[11] S. Yu, S. Member, Y. Wu, R. Jeyasingh, D. Kuzum, H. P. Wong, *IEEE Trans. Electron Devices* **2011**, *58*, 2729.

[12] T. Serrano-Gotarredona, T. Masquelier, T. Prodromakis, G. Indiveri, B. Linares-Barranco,

Front. Neurosci. **2013**, *7*, 1.

[13] Y. Guo, H. Wu, B. Gao, H. Qian, *Front. Neurosci.* **2019**, *13*, 1.

[14] P. Diehl, M. Cook, *Front. Comput. Neurosci.* **2015**, *9*, 99.

[15] S. H. Lee, J. Moon, Y. J. Jeong, J. Lee, X. Li, H. Wu, W. D. Lu, *ACS Appl. Electron. Mater.* **2020**, *2*, 701.

[16] S. Kim, S. Choi, W. Lu, *ACS Nano* **2014**, *8*, 2369.

[17] L. Boteler, A. Lelis, M. Berman, M. Fish, *2019 IEEE 7th Work. Wide Bandgap Power Devices Appl. WiPDA 2019* **2019**, 265.

[18] Y. Wu, J. Moon, X. Zhu, W. D. Lu, *Adv. Intell. Syst.* **2021**, *3*, DOI 10.1002/aisy.202000276.

[19] R. Gütig, R. Aharonov, S. Rotter, H. Sompolinsky, *J. Neurosci.* **2003**, *23*, 3697.

[20] M. Prezioso, M. R. Mahmoodi, F. M. Bayat, H. Nili, H. Kim, A. Vincent, D. B. Strukov, *Nat. Commun.* **2018**, *9*, 1.

[21] T. Tuma, M. Le Gallo, A. Sebastian, S. Member, *IEEE Electron Device Lett* **2016**, *37*, 1238.

[22] A. Sebastian, T. Tuma, N. Papandreou, M. Le Gallo, L. Kull, T. Parnell, E. Eleftheriou, *Nat. Commun.* **2017**, *8*, 1.

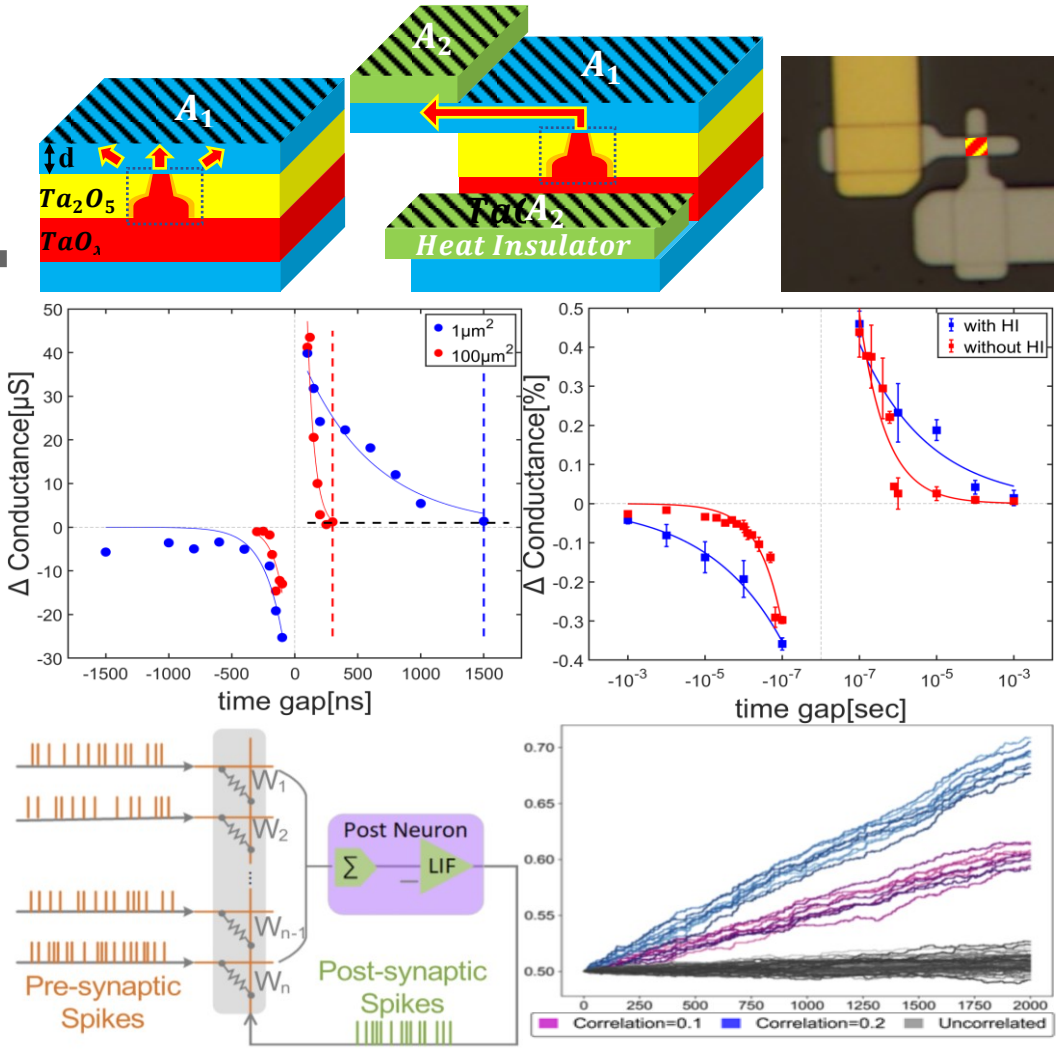
[23] "Nickel-Chromium Alloys (NiCr)," can be found under https://www.reade.com/products/nickel-chromium-alloys-nicr, **n.d.**

[24] AZO, "Palladium (Pd) - Properties, Applications," **2013**.

[25] "Technical data for Palladium," **n.d.**

[26] X. Zhang, H. Xie, M. Fuji, H. Ago, K. Takahashi, T. Ikuta, H. Abe, T. Shimizu, *Appl. Phys. Lett.* **2005**, *86*, DOI 10.1063/1.1921350.

[27] J. Lee, W. Schell, X. Zhu, E. Kioupakis, W. D. Lu, *ACS Appl. Mater. Interfaces* **2019**, *11*, 11579.



TOC Graphic:

By tuning material properties and device structures, the internal dynamics of a 2nd-order memristor can be tailored, allowing it to natively exhibit timing-based learning rules such as spike-timing dependent plasticity (STDP) with adjustable time constants. These memristors can be used to form bio-realistic networks and naturally process temporal data such as detecting correlation patterns in input spike trains.