Computational Capacity of Complex Memcapacitive Networks

DAT TRAN, Santa Clara University
CHRISTOF TEUSCHER, Portland State University

Emerging memcapacitive nanoscale devices have the potential to perform computations in new ways. In this article, we systematically study, to the best of our knowledge for the first time, the computational capacity of complex memcapacitive networks, which function as reservoirs in *reservoir computing,* one of the brain-inspired computing architectures. Memcapacitive networks are composed of memcapacitive devices randomly connected through nanowires. Previous studies have shown that both regular and random reservoirs provide sufficient dynamics to perform simple tasks. How do complex memcapacitive networks illustrate their computational capability, and what are the topological structures of memcapacitive networks that solve complex tasks with efficiency? Studies show that small-world power-law (SWPL) networks offer an ideal trade-off between the communication properties and the wiring cost of networks. In this study, we illustrate the computing nature of SWPL memcapacitive reservoirs by exploring the two essential properties: fading memory and linear separation through measurements of kernel quality. Compared to ideal reservoirs, nanowire memcapacitive reservoirs had a better dynamic response and improved their performance by 4.67% on three tasks: MNIST, Isolated Spoken Digits, and CIFAR-10. On the same three tasks, compared to memristive reservoirs, nanowire memcapacitive reservoirs achieved comparable performance with much less power, on average, about 99×, 17×, and 277×, respectively. Simulation results of the topological transformation of memcapacitive networks reveal that that topological structures of the memcapacitive SWPL reservoirs did not affect their performance but significantly contributed to the wiring cost and the power consumption of the systems. The minimum trade-off between the wiring cost and the power consumption occurred at different network settings of α and β: 4.5 and 0.61 for *Biolek* reservoirs, 2.7 and 1.0 for *Mohamed* reservoirs, and 3.0 and 1.0 for *Najem* reservoirs. The results of our research illustrate the computational capacity of complex memcapacitive networks as reservoirs in reservoir computing. Such memcapacitive networks with an SWPL topology are energy-efficient systems that are suitable for low-power applications such as mobile devices and the Internet of Things.

CCS Concepts: • **Hardware → Neural systems**; **Bio-embedded electronics**; **Emerging architectures**;

Additional Key Words and Phrases: Memcapacitor, memristor, reservoir computing, neural network, small-world power-law networks

Authors' addresses: D. Tran (corresponding author), Electrical and Computer Engineering, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053; email: dtran3@scu.edu; C. Teuscher, Electrical and Computer Engineering, Portland State University, 1900 SW Fourth Avenue, Suite 160, Portland, OR 97201; email: teuscher@pdx.edu.

ACM Journal on Emerging Technologies in Computing Systems, Vol. 17, No. 2, Article 17. Pub. date: February 2021.

17

1 INTRODUCTION

Emerging nanoscale devices, such as memcapacitors, offer possible solutions for implementing synapses, the essential blocks in building artificial neural networks that are brain-inspired computing architectures [Crutchfield et al. 2010]. Similar to the plasticity effect of synapses observed in biological systems, memcapacitors with a tunable value capability provide a way to emulate neurobiological circuits [Jeong and Lu 2018]. Studies have shown that memristive networks [Scharnhorst et al. 2018] and memcapacitive networks [Tran and Teuscher 2017a] are capable of performing neuromorphic tasks with high power efficiency. Although memristive networks have been successful in detecting brain signals of epilepsy seizure patients [Merkel et al. 2014], in generating higher harmonic waveforms [Demis et al. 2016], and in enhancing network dynamics by combing with passive components [Sheldon and Caravelli 2020], memcapacitive networks were able to perform complex tasks with much higher energy efficiency [Tran and Teuscher 2017b].

The intrinsic and nonlinear characteristics of memcapacitive networks provide an ideal computing substrate for reservoir computing (RC). As a brain-inspired computing paradigm, RC explains higher-order cognitive functions and the interaction of short-term memory with other cognitive processes [Rigotti et al. 2010]. In RC, computations take place by translating the tempo-spacial information of input streams into the intrinsic states of a dynamic substrate, the *reservoir*. An output layer then extracts the information from the reservoir for a particular target. With a simple training algorithm, which is the major advantage of RC, reservoirs can perform complex tasks with high efficiency [Bürger et al. 2015a; Tran and Teuscher 2017b]. Although the random structure is often considered as a probable platform for reservoirs [Scharnhorst et al. 2018], a regular structure such as a crossbar structure with memristors at the cross junctions of nanowires has enough dynamics to perform simple tasks [Du et al. 2017; Hassan et al. 2017].

Recent results from brain magnetic resonance imaging of patients reveal that neither random nor regular structures reflect the real structure of a biological brain [Yin et al. 2018]. Human brains are constructed in a small-world (SW) structure, and this structure helps the brains to react positively with radiation treatment. Furthermore, the SW network structure closely resembles the neural growth process [Bassett and Bullmore 2006], neural network functionality [Jian et al. 2018], and the structure of cerebral cortex [Eguiluz et al. 2005]. The SW structure explains how a biological brain is organized in distinctively functional units (or modular units) to minimize the cost of information processing while maximizing its capacity for growth and adaptation [Barbey 2018]. The SW network topology is an alternative structure for a dynamic substrate that brings RC one step closer to the real structure of a biological brain.

In this work, we study the computing capacity of memcapacitive network reservoirs and want to address the following: how do complex memcapacitive networks illustrate their computational capability, and what are the topological structures of memcapacitive networks that solve complex tasks with efficiency? We demonstrate the computing capacity of memcapacitive reservoirs by evaluating their fading memory property and linear separation through measurements of kernel quality. Both the fading memory and linear separation are essential in computing dynamics and are connected to the performance of the reservoirs [Abdel-Hamid et al. 2014; Büsing et al. 2010]. In quantifying the linear separation of the reservoirs, we also consider the effect of nanowire connections between memcapacitive devices and network nodes. The resistance of nanowires contributes to the overall resistance of reservoir networks [Demis et al. 2016; Lv et al. 2019]. Finally, we analyze how memcapacitive reservoirs change their behavior as their topology evolves from a regular organization to a completely random structure. Our results illustrate that memcapacitive networks have computing potentials to perform complex tasks with high energy efficiency.

2 BACKGROUND

2.1 Memcapacitive Devices

The memcapacitor is a memory and nonlinear device that utilizes the relationship between the charge q and the voltage v: $dq = Cdv$. Its capacitance depends on the external applied voltage v as follows:

$$q = C(x, v, t) v,$$
$$\frac{dx}{dt} = f(x, v, t), \tag{1}$$

where q is the charge on the device at time t; v is the applied voltage, x is the internal state of the device; C is the capacitance, which depends on the internal state x; and the function $f()$ describes the changes of the internal state x. The internal state x depends on characteristics of a device: device capacitance [Biolek et al. 2013], a filament's gap and cross-sectional area [Mohamed et al. 2014], or the radius of an elliptical interface and hydrophobic thickness [Najem et al. 2019]. The dynamic change of the internal state x connects to the nonlinear response of device capacitance. Setting device capacitance for synaptic applications is described in the supplementary materials.

The *Biolek* memcapacitor, a generic model, describes the response of a threshold memcapacitive device [Biolek et al. 2013]. The capacitance C of the device is a function of the internal state variable x and the applied voltage V_C:

$$q = CV_C,$$
$$\frac{dC}{dt} = f(V_C) * W(C, V_C),$$

where $f()$ describes the change of device state and the window function $W()$ sets limits to the internal state x. These functions are prescribed as

$$f(V_C) = \beta (V_C - 0.5 * [|V_C + V_t| + |V_C - V_t|]),$$
$$W(C, V_C) = \theta(V_C) * \theta(C_{max} - C) + \theta(-V_C) * \theta(C - C_{min}),$$
$$\theta(x) = \frac{1}{1 + e^{-x/b}},$$

where $\theta()$ is the step function, b is the constant coefficient for the model, β is a constant rate of change when $|V_C|$ is greater than the threshold voltage V_t, and C_{min} and C_{max} are the minimum and maximum values of the device capacitance. The primary function of the window function $W()$ is to prevent a distortion response of the device when its capacitance approaches the limit (either C_{min} or C_{max}) by reducing the rate of change in capacitance. This is a similar issue in a memristive model [Li et al. 2020].

The *Mohamed* model describes the memcapacitive behavior of a metal-oxide junction. Mohamed et al. [2014] observed that applied voltage $v(t)$ across the metal-oxide junction causes growth or shrinkage of the filament due to ion migrations in a metal-oxide device. The migration process affects simultaneously both the filament gap $x(t)$ and the filament cross-sectional area $m(t)$. The correlation of $x(t)$ and $m(t)$ determines the internal state of the device. The total capacitance C_{total} is composed of capacitance between filaments $C_{filament}$, gap capacitance C_{gap}, and oxide

capacitance C_{oxide} [Mohamed et al. 2014]:

$$C_{filament} = \frac{A}{d_1}\varepsilon_{gap}\frac{m(t)}{x(t)},$$

$$C_{gap} = \frac{A}{d_1}\varepsilon_{gap}\left[1 - m(t)\right],$$

$$C_{oxide} = \frac{A}{d_1}\varepsilon_{oxide}\left[1 - m(t)\right],$$

$$C_{tot} = C_{filament} + \frac{C_{gap}C_{oxide}}{C_{gap} + C_{oxide}},$$

where ε_{gap} is the gap permittivity, ε_{oxide} is the oxide permittivity, and d_1 and d_2 are the gap thickness and the oxide thickness. The state equations of $x(t)$ and $m(t)$ are explained in the work of Mohamed et al. [2014].

The recent study of Najem et al. [2019] illustrates the memory capacitive response in biomimetic membranes. The biomimetic membranes are composed of an elliptical and planar lipid bilayer between two lipid-coated aqueous droplets in oil [Najem et al. 2019]. The storage charge in the device is a function of its capacitance C and external applied $v(t)$. The capacitance $C(R, W)$ of the device is a correlation between the radius $R(t)$ of the elliptical interface and the hydrophobic thickness $W(t)$:

$$C\left[R(t), W(t)\right] = \frac{\varepsilon\varepsilon_0\left[a\pi R^2(t)\right]}{W(t)},$$

$$\frac{dR(t)}{dt} = \frac{1}{\xi_{ew}}\left\{\frac{a\varepsilon\varepsilon_0}{2W(t)}v(t) - k_{ew}\left[R(t) - R_0\right]\right\},$$

$$\frac{dW(t)}{dt} = \frac{1}{\xi_{ec}}\left\{-\frac{\varepsilon\varepsilon_0 a\pi R^2(t)}{W^2(t)}v(t) + k_{ec}\left[W_0 - W(t)\right]\right\},$$

where a is the planar ellipticity, ε is the effective dielectric constant, ε_0 is the vacuum permittivity, ξ_{ew} and k_{ew} are the effective damping and stiffness coefficients for the elliptical interface $R(t)$, ξ_{ec} and k_{ec} are the effective damping and stiffness coefficients for the hydrophobic thickness $W(t)$, R_0 is the initial radius, and W_0 is the initial hydrophobic thickness.

Current studies have revealed potential applications of memcapacitive devices. For example, the photoelectric memcapacitor ($Au/La_{1.875}Sr_{0.125}NiO_4/Au$) enables self-adaptive optical signal detection that is suitable for a human visual system [Zhao et al. 2020]. A chaotic system utilizes the nonlinear and memory characteristics of a memcapacitive device in image encryption [Sun et al. 2020]. The multistate memcapacitive device ($CH_3NH_3SnBr_3$) promises a high-density memory module [Qian et al. 2019]. The graphene nanoribbon memcapacitive device has three discrete states for multivalued logic design [Patel et al. 2018]. The memcapacitive network with neurotransistors can classify input signals [Wang et al. 2018]. Memcapacitive devices form logic gates to perform simple logic operations [Tran and Teuscher 2017a]. A recent study by Liu et al. [2020] demonstrates the possibility to fabricate a memcapacitive device from an organic donor-acceptor (D-A) copolymer that exhibits synaptic plasticity.

One major advantage of memcapacitors is low energy consumption compared to memristors. Memristors, by their nature, are resistors that consume power, whereas memcapacitors are storage devices that do not consume static power. The switching power of memcapacitors is also small due to their small values of capacitance, typically in a pF or nF range. With a similar form factor as memristors, memcapacitors are reliable, comparable to the performance of memristors, and sufficient for ultra-low-energy applications [Goswami et al. 2020; Tran and Teuscher 2017b]. Recent
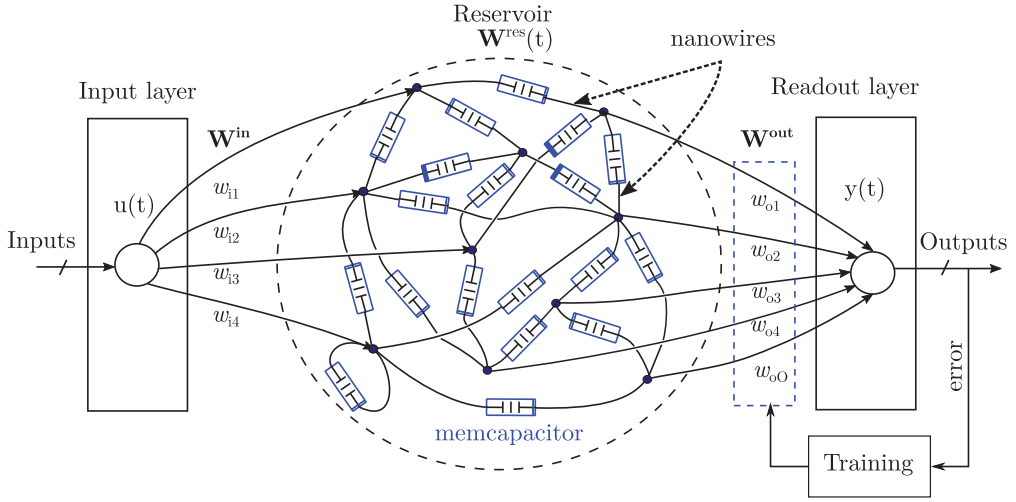
Fig. 1. A memcapacitive reservoir network. The reservoir is an electrical network of which memcapacitors are randomly connected between nodes through nanowires. The node voltages are the time-dependent state $x(t)$ of the reservoir that provide signals to the readout layer. $W^{res}(t)$ is the state matrix containing capacitance of memcapacitive devices at time t. Only the readout layer is trained with a simple algorithm.

studies have reported the discovery of memcapacitive dynamics in new nanocomposite materials [Cai et al. 2019; Khan and Lee 2016; Liu et al. 2020; Patel et al. 2018; Qian et al. 2019; Yang et al. 2017; You et al. 2016]. Such discovery opens a unique possibility for memcapacitive applications.

2.2 RC Architectures

The Multilayer Perceptron (MLP) is a common architecture for an artificial neural network. It is a feedforward network, composed of an input layer, one or more hidden layers, and an output layer. With the feedforward structure, MLP is not able to retain the information of input streams from previous timesteps. To overcome this limitation, recurrent connections are added to form a Recurrent Neural Network (RNN). The recurrent connections, however, significantly increase the training complexity of the RNN since a small change in the internal state of the RNN can potentially lead to chaotic behavior due to its recurrent feedback.

Reservoir Computing (RC) is an alternative to the traditional RNN for machine learning that avoids the training of large recurrent networks. In an RC architecture, inputs are mapped to the internal states of a reservoir and the outputs from the untrained reservoir provide input signals to a readout layer. The readout layer is trained with a simple algorithm to perform a particular task.

Figure 1 displays an example of a memcapacitive reservoir in which memcapacitors are randomly connected between nodes through nanowires. The input signal $u(t)$ is scaled by the input matrix W^{in}, whose values are randomly selected from the set of $\{-v, +v\}$. The time-dependent state $x(t)$ of the reservoir is determined by the internal capacitance of memcapacitive devices at time t. The internal capacitance of memcapacitive devices are presented as the state matrix $W^{res}(t)$, where $W^{res}(t) = [mc_1(t), mc_2(t), mc_3(t), \ldots, mc_n(t)])$. The reservoir state $x(t)$ is defined as

$$x(t + 1) = f\left[W^{res}(t)x(t) + W^{in}u(t)\right],$$

where f is the transforming function of a reservoir node. The output $y(t)$ at the output layer is the inner product of the reservoir state $x(t)$ and the output weight matrix W^{out}:
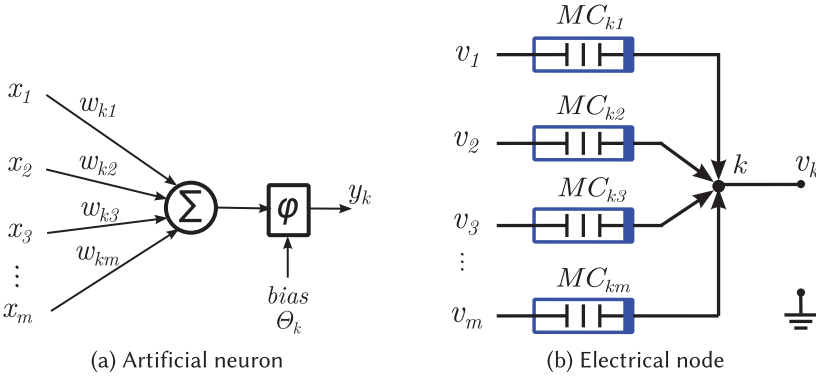
$$y(t) = x(t)W^{out}.$$

Fig. 2. Artificial neuron and electrical node. (a) An artificial neuron k performs a weighted sum on its inputs. (b) An electrical node k performs a summation on its input voltages.

In Figure 1, the reservoir is an electrical network of memcapacitors. The time-dependent state $x(t)$ of the reservoir is electrically determined by the combination of the device capacitance when the reservoir responds to input signal $u(t)$.

In memcapacitive reservoirs, a node is an electrical node that performs a similar operation to a node in traditional artificial neural networks. Its voltage is the sum of input voltages, similar to the output of an artificial neuron. In Figure 2(a), the artificial neuron k performs a weighted sum, which is similar to the operation of a biological neuron. The activation function $\varphi()$ transforms the weighted sum of its inputs $(x_1, x_2, x_3, \ldots, x_n)$ and its synaptic weights $(w_{k1}, w_{k2}, \ldots, w_{km})$ into an output signal y_k:

$$y_k = \varphi\left(\sum_{j=1}^{m} w_{kj}x_j + \theta_k\right), \tag{2}$$

where θ_k is a bias offset. The well-known activation function for the artificial neuron is sigmoidal, hyperbolic, logistic, or exponential. An electrical node k, shown in Figure 2(b), can perform a similar operation in which the capacitance of memcapacitive devices plays the role of a synapse. The sum of the total charge at node k is

$$q_{k1} + q_{k2} + q_{k3} + \cdots + q_{km} = 0$$
$$(v_1 - v_k)c_{k1} + (v_2 - v_k)c_{k2} + (v_3 - v_k)c_{k3} + \cdots + (v_m - v_k)c_{km} = 0$$
$$v_k(c_{k1} + c_{k2} + c_{k3} + \cdots + c_{km}) = v_1c_{k1} + v_2c_{k2} + v_3c_{k3} + \cdots + v_mc_{km}$$
$$v_k = \left(\frac{1}{c_{k1} + c_{k2} + c_{k3} + \cdots + c_{km}}\right)\sum_{j=1}^{m} c_{kj}v_j, \tag{3}$$

where q_k is the charge of the memcapacitor MC_k.

By rearranging Equation (3), we obtain

$$v_k = \varphi\left(\sum_{j=1}^{m} c_{kj}v_j\right), \tag{4}$$

$$\varphi(x) = \frac{x}{\sum_{j=1}^{m} c_{kj}}. \tag{5}$$

The total charge at node k is the weighted sum of input voltages and capacitive synapses (Equation (4)). The activation function $\varphi(x)$ converts the total charge into a voltage v_k as specified in

Equation (5). The voltage at the electrical node k is the weighted summation of input signals, similar to the output of an artificial neuron (Equation (2)).

In a traditional neural network, artificial neurons are connected through synapses to form a network. In reservoirs, memcapacitive devices take the role of synapses that connect nodes through nanowires to form electrical networks. Network topologies, as shown in studies, influence the dynamic response and contribute to the performance of reservoirs [Bürger et al. 2015b; Du et al. 2017; Hassan et al. 2017; Rodan and Tiňo 2011; Scharnhorst et al. 2018].

3 SMALL-WORLD POWER-LAW NETWORKS AS RESERVOIRS

In general, any system with enough dynamics is able to serve as a reservoir [Duport et al. 2012; Natschläger et al. 2002; Scharnhorst et al. 2018]. Random networks are probable platforms since they are easy to fabricate [Bürger et al. 2015b; Scharnhorst et al. 2018]. Regular networks also have sufficient dynamics to function as reservoirs to perform simple tasks [Du et al. 2017; Hassan et al. 2017].

Although both random and regular networks offer a potential computing substrate in the RC architecture, they do not reflect the real nature of physical [Watts and Strogatz 1998] and biological [Barbey 2018] neural networks. Real networks, such as online social networks, cultural networks, and the internet, often share a common feature, an SW property that is involved with a degree of interconnections at both local and global levels [Petermann and De Los Rios 2006].

Various studies have shown that the structure of SW networks describes the neural connections and the intercommunication lines between different parts of a biological brain [Achard et al. 2006; Eguiluz et al. 2005; Ma et al. 2018]. The data from magnetic resonance imaging scans of healthy and diseased brain states, for instance, reveal that the SW structure improves the resilience of human brains to overcome neural disease [Chockanathan et al. 2018]. SW networks, as neuroscience suggests, allow biological brains to balance resources to minimize the cost of information processing and to maximize the capacity for growth and adaptation [Barbey 2018]. SW networks closely resemble the topological structure of a brain-inspired computing network.

In this study, we investigate small-world power-law (SWPL) networks, which are the subset of the SW networks. Both the SW and SWPL networks are similar in the sense that they describe the fundamental characteristics of a small-world phenomenon. The main difference, however, is the formation of a network: an SW network arises from a regular ring network, whereas an SWPL network is formed from a grid network, shown in Figure 3. In the original SW model, nodes in a regular ring network have a uniform distribution to connect to any other nodes within the network [Watts and Strogatz 1998]. According to Petermann, the uniform distribution is not realistic, since each connection, local or global, is associated with a cost (wire connection) to the system, whose resources are often limited [Petermann and De Los Rios 2006]. In the SWPL model, the decay power law, $q(l) = l^{-\alpha}$, governs the formation of a connection. According to the decay power law, distant nodes have a smaller probability to form connections compared to nearby nodes for the reason that long connections require more resources (long connections) to form connections than short ones. The structure of an SWPL network reflects the nature of real networks, physical networks and biological neural networks, which have more local connections than global connections [Barbey 2018; Sporns et al. 2004].

Figure 3 illustrates a process of constructing an SWPL network from a grid network. The rewiring probability (or randomness factor), β, determines the probability of a link in the grid network being rewired. Depending on values of β, different network topology emerges, from a regular ($\beta = 0.0$) to a completely random ($\beta = 1.0$) network. The decay power-law function, $p(l)$, controls how a connection between nodes with a distance of l is formed. The decay exponent (the locality), α, describes whether a network is a highly local (α is large) or highly global ($\alpha = 0$)

(a) 8 x 4 regular network

(b) Small-world power-law network

(c) Uniform probability distribution between nodes
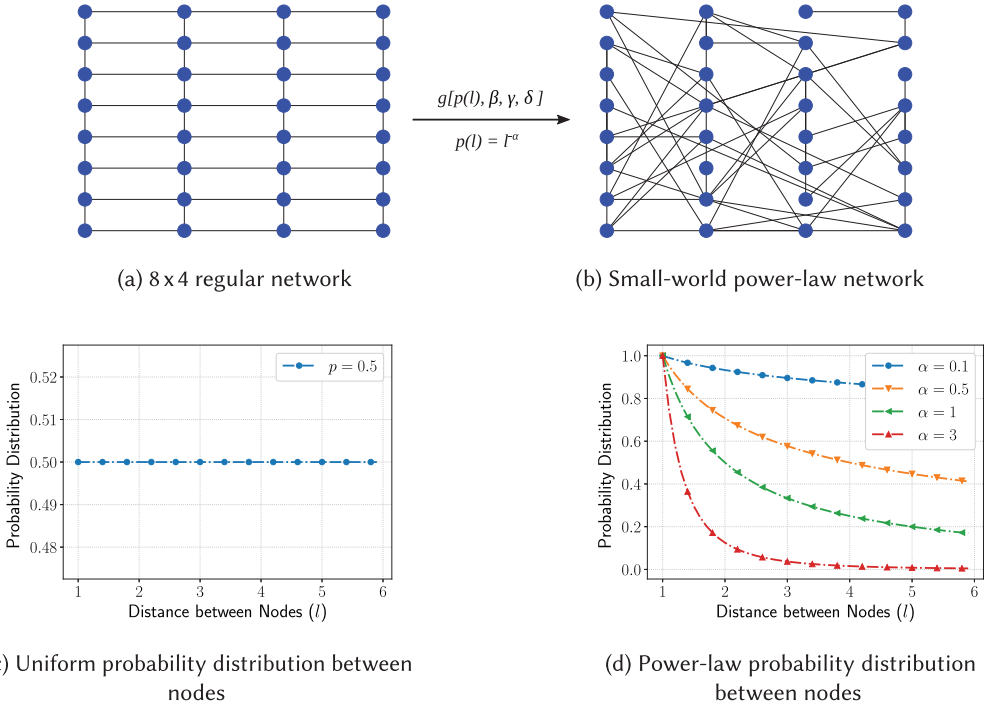
(d) Power-law probability distribution between nodes

Fig. 3. Process of creating a small-world power-law network from a regular network. The regular grid network (a), which has an uniform probability distribution between nodes (c), is transformed into a small-power network (b) by rewiring local connections with global links under the power-law probability distribution (d).

connection type. It has been shown that $\alpha < D + 1$ (D is the dimension of the regular network) is a necessary condition for the emergence of an SWPL network during the formation process [Petermann and De Los Rios 2006]. We also modify the formation model with two additional parameters, γ (removed links) and δ (added links), to allow flexibility in the formation of SWPL networks. Depending on different settings of the network parameters (D, α, β, δ, and γ), different SWPL structures will emerge (Figure 4).

4 METHODOLOGY

4.1 Device Models

In our study, we selected three models for memcapacitive networks: the *Biolek* behavior model, the *Mohamed* metal-oxide junction, and the *Najem* biomimetic membranes. Two memristive devices (*Chang* and *Oblea*) were used as references for comparison. The characteristics of the models are shown in Table 1.

4.2 Reservoir Properties

The computational capability of a reservoir is based on two essential properties: fading memory and separation property [Jaeger 2001; Maass et al. 2002].

4.2.1 Fading Memory. Fading memory allows the reservoir to map similar input signals of a finite length to similar internal states, independent of its initial states. Various studies have shown that fading memory significantly improves the dynamic response of reservoirs [Roy and Basu 2016; Snyder et al. 2013; Verstraeten et al. 2010]. For a memcapacitive reservoir, the fading effect depends
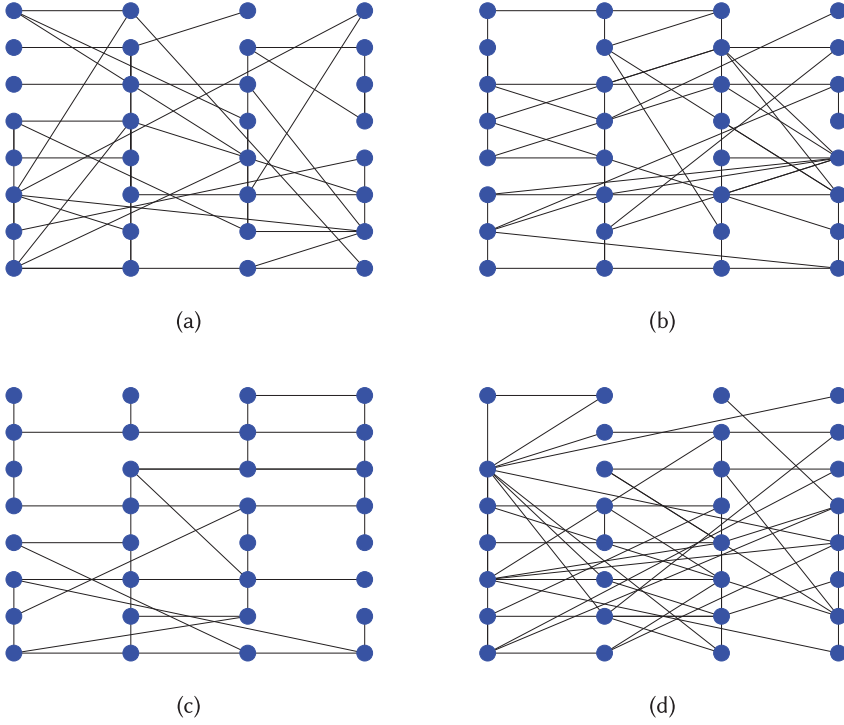
Fig. 4. SWPL networks deriving from the same 8 x 4 grid network with different network parameters. The parameters for each network are as follows: (a) $\alpha = 0.4, \beta = 0.8, \gamma = 20, \delta = 2$; (b) $\alpha = 1.2, \beta = 0.56, \gamma = 10, \delta = 4$; (c) $\alpha = 1.7, \beta = 0.22, \gamma = 17, \delta = 1$; and (d) $\alpha = 2, \beta = 0.9, \gamma = 8, \delta = 2$.

Table 1. Characteristics of Memcapacitive Devices

Model	Device	Type	Volatile	Values at w	
				w_{max}	w_{min}
Biolek [Biolek et al. 2013]†	*	MC	Yes	$100pF$	$1pF$
Mohamed [Mohamed et al. 2014]†	$Pt/Ag_2S/Pt$	MC	Yes	$6.5nF$	$1nF$
Najem [Najem et al. 2019]†	$KCl/C_3H_8O_3S$	MC	Yes	$0.3pF$	$272pF$
Chang [Jo et al. 2010]	Ag/Si	MR	Yes	$11k\Omega$	700Ω
Oblea [Oblea et al. 2010]	Ge_2Se_3/Ag	MR	No	$12k\Omega$	720Ω

†Thesemodels were modified to exhibit a state volatility, the essential property in RC.
*This model is not based on a physical device.
MC, memcapacitor; MR, memristor.

on the internal decay state of each device that contributes to the fading memory of the reservoir as a whole system. Since all memcapacitive models are for nonvolatile devices, the models were modified to express the state volatility by adding a decay term $f_d(x)$ (a negative quantity) to the state equation (Equation (1)):

$$\frac{dx}{dt} = f(x, v, t) + f_d(x), \qquad (6)$$

where v is the applied voltage, x is the internal state of the device, and the behavior function $f()$ determines the change of the internal state x. The experimental results showing the fading states of memcapacitive devices are presented in the supplementary materials.

4.2.2   Reservoir Kernel Quality. Kernel quality measures the linear separation property of a reservoir [Legenstein and Maass 2007]. The linear separation describes how the reservoir disassociates different input patterns, independent of a target output. Kernel quality is based on a set of n reservoir states generated by m different input stimuli. All reservoir states are collected into a matrix M of size $n \times m$. Each column is a state vector corresponding to a particular input. The rank r of this state matrix M is a measure of the computational power of the reservoir. A reservoir has the highest computational power when the matrix M has a maximum rank, or $r = m$. In this case, each column in the matrix M is linearly independent of the other columns.

4.2.3   Memory Capacity. The linear memory capacity is a standard measure of memory in RNNs. It quantifies the short-term memory (fading memory) of a recurrent network for reconstructing information from an input stream. Developed by Jaeger [2000], the memory capacity task involves how well a trained network is able to reproduce the input stream after various time lags [Jaeger 2002]. The memory capacity is calculated as the summation of the correlation coefficients between the observed output $y(t - k)$ of a trained network and the desired output $\hat{y}(t)$:

$$C_k = \frac{Cov^2\left[y(t - k), \hat{y}(t)\right]}{Var[y(t - k)]Var[\hat{y}(t)]},$$

$$MemCapacity = \sum_{k=1}^{K} C_k,$$

where C_k is the correlation coefficient at the delayed timestep k and K is the total number of delayed timesteps. The correlation coefficient, C_k, is greater than or equal to zero. When C_k is zero, the network has no short-term memory. When C_k is greater than 1.0, the network can reconstruct perfectly the input stream after k time lags. In our experiments, memcapacitive reservoirs of 60 nodes were simulated for 20 time lags ($K = 20$) of an input stream, whose random values were from uniform distributions on the interval $[-2.0, 2.0]$. A factor, v, scaled the input signal to prevent overstimulated conditions in reservoirs.

4.2.4   Synaptic Weights in Reservoirs. Jaeger [2001] observes that the spectral radius, λ, of the initial weight matrix must be less than 1 to maintain the stable response of an Echo State Network (ESN). When λ is greater than 1, signals circulating within the recurrent network are amplified indefinitely, which leads to chaotic behavior. As shown in Figures 1 and 2, the memcapacitive network is an electrical network that conforms to the circuit laws (Kirchhoff's circuit laws) for voltages and currents at network nodes. Consequently, the internal state (or the node voltages) of the memcapacitive network is always stable due to the circuit laws after a short settling time of the device capacitance. In this respect, the initial weight matrix of the memcapacitive network does not require the condition for its spectral radius. For memcapacitive reservoirs, all devices have C_{min} as their initial state. The response of memcapacitive reservoirs depends on how their internal states map out input signals. Here, the capacitance of memcapacitive devices plays a significant role in the task of mapping input signals. With scaling factor v, input signals are scaled to utilize the full capacitance range of memcapacitive devices. The optimal values of v are from experimental results for different memcapacitive networks.

4.3   Benchmark Tasks

4.3.1   MNIST Image Classification. MNIST is the acronym for *Mixed National Institute of Standards and Technology*. It is a simple benchmark to evaluate the performance of various brain-inspired networks [Du et al. 2017; Lee et al. 2015; Tran and Teuscher 2017a, 2017b]. The dataset is a collection of handwritten digits (28×28 pixels) that has 60,000 training and 10,000 testing images. Images were flattened into 784-pixel vectors as inputs to reservoirs for training and testing.

The amplitudes of the pixel vectors were scaled to a range of $[0, v]$ to avoid an overstimulated internal state of the reservoirs. Only subsets of training and testing sets, 1,000 training digits and 400 testing digits, were selected due to the simulation time of large reservoirs. Classifying MNIST digits is a simple and spatial task.

4.3.2 Isolated Spoken Digit Recognition. The spoken digit recognition is a complex benchmark for RC systems [Larger et al. 2017; Suarez et al. 2018]. The spoken digit dataset, created by Jackson [2016], contains the recordings of spoken digits in wave file formats at 8 kHz. The digit dataset has 1,500 recordings of digits 0 to 9 from various English speakers. The dataset was divided into two nonoverlapping sets: 1,000 digits for training and 500 digits for testing. The recordings of spoken digits were translated into input vectors of Mel-frequency cepstral coefficients, a common preprocessing technique for a speech recognition task [Abdel-Hamid et al. 2014; Bouvrie et al. 2008; Soures et al. 2017]. Each recording segment of a digit was divided into 3:4:3 ratio regions with 13 Mel-frequency cepstral features for each region [Clarkson and Moreno 1999]. In addition, each region was calculated for its delta and delta-delta coefficients to capture all features, and the final result is a 117-coefficient vector. The coefficient vectors were translated into input voltages and scaled to a range of $[-v, v]$ to avoid an overstimulating condition in reservoirs. Recognizing spoken digits is a complex task due to the temporal and spatial dimensions of input signals.

4.3.3 CIFAR-10. CIFAR-10 is another common benchmark [Ghayoumi et al. 2018]. It is a dataset collection of 60,000 color images with a size of 32×32 pixels. The images are categorized into 10 classes of objects: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. The dataset is divided into 50,000 images for training and 10,000 images for testing. CIFAR-10 images are high dimensional due to the image features represented by color intensity. Training and testing CIFAR-10 images require large-scale reservoirs with long simulation time. Preprocessing color images into grayscale images is an alternative for dimension reduction [Rigamonti et al. 2011]. Similar to the MNIST dataset, only subsets of training and testing data, 1,000 training digits and 400 testing images, were selected due to the simulation time of large reservoirs.

4.4 Nanowire Model and Power Calculation

As shown in Figure 1, each memcapacitive device is connected to nodes through nanowires. The nanowire resistance was modeled as a resistor R_{nw}, whose resistance is approximate $5\,\Omega/\mu m$ [Selzer et al. 2016]. To simplify the calculations of nanowire resistance, memcapacitive devices were assumed to have a uniform length of 10 nm [Dongale et al. 2015] and the nanowire length l_n was the Euclidean distance between nodes after subtracting the device length. The total nanowire resistance with a Euclidean distance l_{nw} is $R_{nw} = l_{nw} * 5\,\Omega/1\mu m$.

The nanowire resistor R_{nw} and the memcapacitor C_m form an RC circuit whose transient response in a charging phase is

$$v_{Cm}(t) = v_{in} + (V_o - v_{in})e^{\frac{-t}{R_{nw}C_m}}$$

$$v_{Rn}(t) = (v_{in} - V_o)e^{\frac{-t}{R_{nw}C_m}}$$

$$P_{nw}(n) = \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} \frac{v_{Rn}^2(k)}{R_{nw}} dk$$

$$P_{nw}(n) = \frac{1}{R_{nw}\Delta t} \int_{t_0}^{t_0+\Delta t} (v_{in} - V_o)^2 e^{\frac{-2k}{R_{nw}C_m}} dk, \tag{7}$$

where v_{in} is the node voltage at timestep n, V_o is the initial voltage in the capacitor C_m at timestep $n - 1$, C_m is the capacitance of the memcapacitor at timestep n, $P_{nw}(n)$ is the dissipated power consumed by the nanowire resistor R_{nw}, and Δt is the duration of a timestep.

Table 2. Reservoir Setting Parameters for Device-State Decay

Device Model	Type	Decay θ	Ts (μs)	Scaling Factors†	
				State S	t
Biolek [Biolek et al. 2013]	MC	3.00	1.0	S	t
Mohamed [Mohamed et al. 2014]	MC	3.00	1.0	$(S - 0.65451) * 1e4$	1
Najem [Najem et al. 2019]	MC	1.00	1.0	$(S + S_{min}) * 1e10$	1
Chang [Jo et al. 2010]	MR	1.13	1.0	$S/1e2 + 0.486$	1.0
Oblea [Oblea et al. 2010]	MR	0.15	0.1	$S/2e1 + 0.449$	$t/10.0$

† Since the internal states of memcapacitive devices are different, the measurements of device states were scaled to fit in a single display as shown in Figure 5(a).
MC, memcapacitor; MR, memristor.

At each timestep, the dissipated power in each nanowire was calculated with Equation (7) using the capacitance of the device at the timestep n. The calculation is a close approximation of actual power dissipated in each nanowire at each timestep.

4.5 Simulation Environment

Simulations were based on the software framework Python OGER, which is a comprehensive computing framework for a variety of datasets, reservoir node types, and training methods [Verstraeten et al. 2012]. An additional memcapacitive node, a memcapacitive network, is added to the existing set of reservoir nodes. The computation takes place inside the memcapacitive node by treating it as a temporarily capacitive network, which can be solved efficiently using the modified nodal analysis (MNA) algorithm [Wedepohl and Jackson 2002]. After calculating a timestep using MNA, the device capacitance is updated with the node voltages in the network to account for the state changes of memcapacitive devices. The process is repeated for consequent timesteps.

5 RESULTS

5.1 Fading Memory Property

The volatile effect of a memcapacitive device is guaranteed by the decay term $f_d()$ in Equation (6). It is not clear that a memcapacitive reservoir, as a collective and nonlinear network of memcapacitive devices, expresses a fading memory effect. Here, we investigated the fading phenomenon of memcapacitive reservoirs at the device level and the system level.

Four random spoken digits were selected as inputs to simulate the fading memory effect of three memcapacitive reservoirs (*Biolek, Mohamed,* and *Najem*) and two memristive reservoirs (*Chang* and *Oblea*). The reservoirs had 175 nodes with the parameters as follows: $\alpha = 1.02$, $\beta = 0.71$, $\delta = 0.0$, and $\gamma = 0.0$. The settings of the decay factors for the reservoirs depend on the physical property of each device, shown in Table 2. All reservoirs were simulated with a similar initial state. The spoken digits were translated into 117-coefficient vectors as input signals to the reservoirs. The input signals were scaled to an input range of $[-v, v]$ with $v = 8.0$ to provide enough stimuli to the reservoirs. Furthermore, each coefficient vector was translated into an active pulse of 400 timesteps. Using pulse signals is a common technique to evaluate the fading memory of a reservoir [Duport et al. 2012]. All reservoirs were simulated for 1,000 timesteps. The device states at each timestep were recorded and the simulation results were demonstrated in Figure 5(a).

Figure 5(a) shows the corresponding averages of device states when the reservoirs were simulated with four input coefficient vectors. Since the internal state and the response time are unique for each device, the measurements of the device states were scaled to fit into a single display for clarity. The scaling factors are in Table 2. As indicated by Vin_2, one of the 117 input pulses, the

(a) Fading memory at the device level   (b) Fading memory at the system level
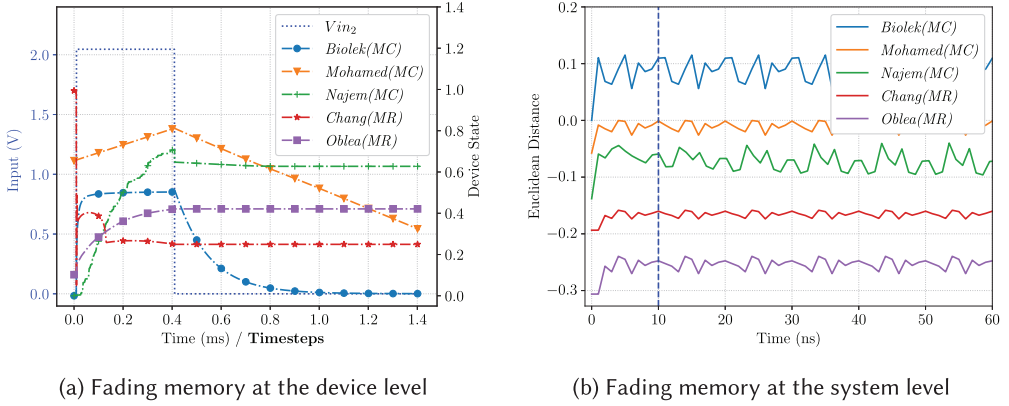
Fig. 5. Fading memory of memcapacitive reservoirs. The settings of reservoir parameters were $\alpha = 1.02$, $\beta = 0.71$, $\delta = 0.0$, and $\gamma = 0.0$. (a) Vin_2, one of 117 input pulses, shows the time intervals where the inputs were active $[0.0\ ms, 0.4\ ms]$ and inactive $[0.4\ ms, 1.4\ ms]$. Since the internal states are different and device dependent, the measurements of the device states were scaled to fit into a single display for comparison. The scaling factors are shown in Table 2. All device states were displayed with time measurements in milliseconds. (b) The plots of the Euclidean distance of other reservoirs were shifted downward for clarity. The vertical line is where the Euclidean distances of the reservoir states were approaching a cyclic steady state, indicating that the initial states of memcapacitive and memristive devices no longer had any effect on the reservoir states.

input signals were active in $[0.0\ ms, 0.4\ ms]$ and inactive in $[0.4\ ms, 1.4\ ms]$. During the active pulse $[0.0\ ms, 0.4\ ms]$, the device states of the memcapacitive reservoirs converged to a new state. Once the input signals were off ($t \geq 0.4\ ms$), the internal states of memcapacitive and memristive reservoirs gradually returned to a reset state, except the *Oblea* reservoir since the *Oblea* memristive device was a nonvolatile device. The gradual decay of the device states without active input signals illustrates the fading memory effect of the memcapacitive reservoirs at the device level. Without applied signals, the reservoir states, composed of the device states, gradually return to a reset state.

Next, we investigated how the memcapacitive reservoirs exhibit the fading memory effect at the system level by looking at the Euclidean distance of the reservoir states. Calculating the Euclidean distance between the trajectories of reservoir states for similar inputs but at different initial states is a common technique to understand the dynamic property of a reservoir [Verstraeten et al. 2007]. At a time t, the Euclidean distance $d_{u,v}(t)$ between the trajectories of two reservoir states $U(t)$ and $V(t)$ is defined as

$$d_{u,v}^2(t) = \sum_{i=1}^{N} [u_i(t) - v_i(t)]^2 ,$$

where $u_i(t)$ and $v_i(t)$ are the reservoir states (node voltages) at time t and N is the number of the reservoir nodes.

Each input signal was a series of 20 similar sequences. Each sequence was a combination of 10 random sounds of the same class digit. Each reservoir was simulated twice for 100 timesteps, and for each run the reservoir was initialized with a different state. The Euclidean distance was calculated using the reservoir states at each timestep. The timestep for reservoirs depended on the response of each device and is specified in Table 3. The network parameters α, β, δ, and γ were the same for all reservoirs.

Figure 5(b) presents the simulation results of the reservoir Euclidean distances. The plots of other reservoirs were shifted downward for clarity. The vertical line at $t = 10\ ns$ is when the Euclidean distances of the reservoir states were approaching a cyclic steady state. The cyclic steady-state

Table 3. Timestep Settings for Reservoir-State Decay

Device Model	Type	Timestep (ns)	Time Scale† ×1e7
Biolek [Biolek et al. 2013]	MC	1.0	100.0
Mohamed [Mohamed et al. 2014]	MC	10.0	10.0
Najem [Najem et al. 2019]	MC	10.0	10.0
Chang [Jo et al. 2010]	MR	100.0	1.0
Oblea [Oblea et al. 2010]	MR	100.0	1.0

†Time was scaled to fit the simulation results in a single display as shown in Figure 5(b). MC, memcapacitor; MR, memristor.

Table 4. Network Settings for Memory Capacity Task

Device Model	Type	N	Input†	O	α	β	θ	ts
Biolek [Biolek et al. 2013]	MC	60	1	14	1.2	0.4	0.181	1.05×10^{-10}
Mohamed [Mohamed et al. 2014]	MC	60	1	14	1.2	0.4	0.357	1.89×10^{-08}
Najem [Najem et al. 2019]	MC	60	1	14	1.2	0.4	0.165	1.51×10^{-11}
Chang [Jo et al. 2010]	MR	60	1	14	1.2	0.4	0.465	1.00×10^{-5}
Oblea [Oblea et al. 2010]	MR	60	1	14	1.2	0.4	0.624	2.04×10^{-7}

†The input signal is a one-dimensional vector of 100 elements whose values are from uniform distributions on the interval $[-2.0, 2.0]$ and scaled by a factor ν. The reservoirs were simulated for 20 time lags ($K = 20$) of the input signal. MC, memcapacitor; MR, memristor.

phenomenon suggests that the information contained in the initial conditions of the reservoir states is no longer relevant and that the fading memory effect, indeed, took place at the system level.

5.2 Memory Capacity Task

The input signal was a one-dimensional vector of 100 elements whose random values were from uniform distributions on the interval $[-2.0, 2.0]$. The input signal was scaled by a factor ν to prevent overstimulated conditions in memcapacitive reservoirs. Three memcapacitive reservoirs (*Biolek*, *Mohamed*, and *Najem*) of 60 nodes were simulated for 20 time lags ($K = 20$) of the input stream. The two memristive reservoirs, as references, were also simulated with the same input stream. The network settings of the memcapacitive and memristive reservoirs are shown in Table 4.

The measurements of memory capacity are shown in Figure 6. We observe that all reservoirs had a better memory capacity when the input scaling factor ν was more than 0.8. The memory capacity of the *Biolek* reservoir, however, was low compared to others. Since the changing internal state of the *Biolek* device is highly sensitive to the input amplitude and signal timestep, we suspect that the network settings of the *Biolek* reservoirs did not match with the characteristics of the device. In contrast, the *Chang* reservoirs had the highest measurements of the memory capacity. The results suggest that the network parameters were optimal for the device property. Finding optimal values of the network parameters for the *Biolek* reservoir is beyond the current scope of our study. The simulation results underline that both memcapacitive and memristive reservoirs were capable of short-term memory and that the reservoirs as whole systems underwent the effect of fading memory.

5.3 SWPL Network Property

As shown in the study of Jaeger [2001], the computational capacity of a reservoir depends on its dynamics. In the first experiment, we investigated how memcapacitive reservoirs behave when
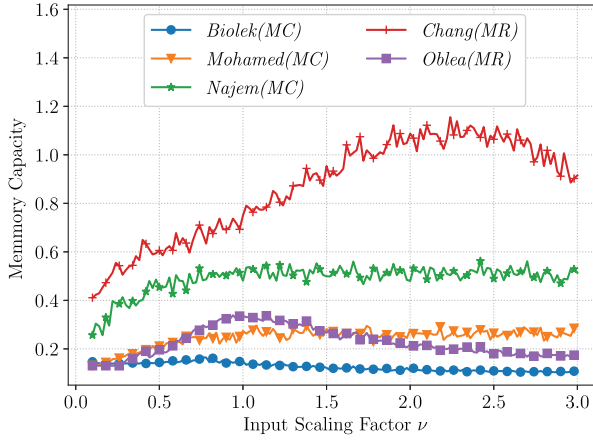
Fig. 6. Memory capacity task. All reservoirs had a better measurement of memory capacity when the scaling factor v was greater than 0.8. The results suggest that memcapacitive and memristive reservoirs exhibited short-term memory and the effect of fading memory.

Table 5. Timestep Settings for Ideal and Nanowire Reservoirs

| Device Model	Ideal†	Nanowire ($ts = n * \tau$)	
	ts	n†	$\tau = R_{min}C_{min}$
Biolek [Biolek et al. 2013]	7.23×10^{-9}	425.0	5.00×10^{-14}
Mohamed [Mohamed et al. 2014]	8.72×10^{-7}	10.0	7.72×10^{-11}
Najem [Najem et al. 2019]	5.93×10^{-6}	2.5	1.51×10^{-14}

†The values of ts for ideal reservoirs and the factor n for nanowire reservoirs were from experimental results. In Table 5, R_{min} was the minimum resistance of nanowire connection and C_{min} was the minimum capacitance of the memcapacitive device.
MC, memcapacitor; MR, memristor.

the number of devices changes by measuring their kernel quality. We also consider both cases of memcapacitive reservoirs: ideal reservoirs without the effect of nanowire connections and ones with the nanowire connections. The timestep settings for the memcapacitive reservoirs, based on our experimental results, are shown in Table 5. The settings of the timesteps are where the memcapacitive reservoirs achieved their highest performance.

In Figure 7(a), the results show that both the ideal and nanowire reservoirs had higher kernel quality measurements with a higher number of devices. As shown in Section 5.1, the states of the reservoirs depend on the internal states of the devices. Reservoirs with a higher number of devices have better dynamics to represent input signals compared to reservoirs with fewer devices. The measurements of the reservoir kernel quality reflect this phenomenon. Furthermore, the kernel quality of the reservoirs converging to a final value with a higher number of devices suggests that increasing the number of devices within the memcapacitive reservoirs would not significantly improve their dynamics.

A closer observation of the reservoir dynamics between the ideal and nanowire reservoirs is shown in Figure 7(b). Compared to the ideal reservoirs, the nanowire reservoirs had better dynamical responses with similar input signals. This is because the nanowire resistors and the capacitors of the memcapacitive devices formed RC circuits, whose responses are associated with the time constants of R and C. With similar numbers of the memcapacitive devices, the nanowire reservoirs display a higher kernel quality measurement. For instance, with a 2,000-device setting,
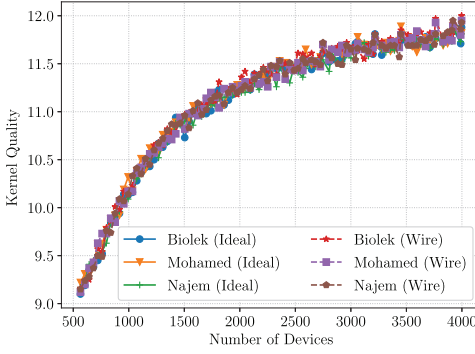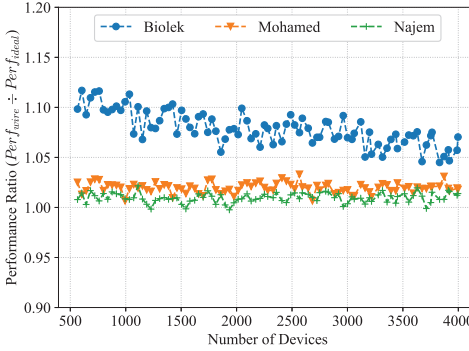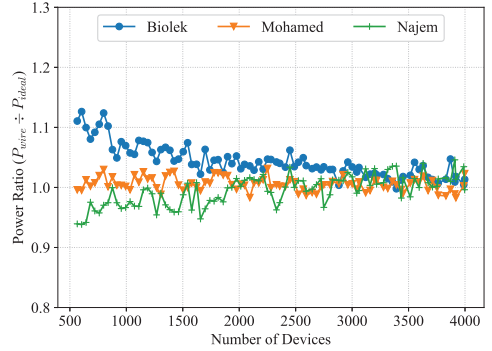
(a) Kernel quality measurements of reservoirs
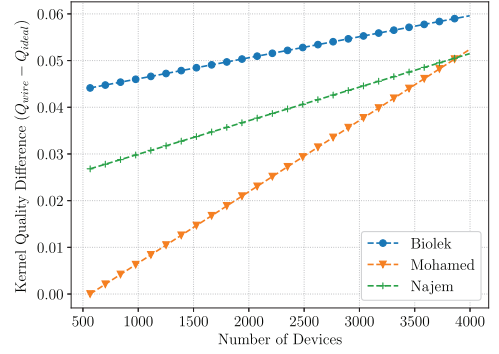
(b) Difference of kernel quality measurements of ideal and nanowire reservoirs.

Fig. 7. Kernel quality of memcapacitive reservoirs. The reservoir parameters were $\alpha = 1.2$, $\beta = 0.4$, $\delta = 0.0$, and $\gamma = 0.0$. (a) Kernel quality measurements of ideal and nanowire memcapacitive reservoirs. (b) Difference kernel quality measurements of memcapacitive reservoirs. The plotting data were fitted with a linear function.

(a) Performance ratio

(b) Power ratio

Fig. 8. Performance and power ratios between nanowire and ideal memcapacitive reservoirs on Isolated Spoken Digits. (a) The nanowire reservoirs, on average, performed better than the ideal reservoirs by 6.0%, 1.6%, and 0.8% for the *Biolek*, *Mohamed*, and *Najem* reservoirs, respectively. (b) However, the nanowire reservoirs consumed a similar amount of power compared to the ideal reservoirs on various reservoir sizes.

the kernel quality measurements of the nanowire reservoirs were better than the ideal reservoirs by 0.05, 0.035, and 0.022 for *Biolek, Mohamed,* and *Najem* reservoirs, respectively.

The results of performance and power ratios of the memcapacitive reservoirs on the Spoken Digit task are shown in Figure 8. We observe that the performances of the nanowire reservoirs were consistently higher than those of the ideal reservoirs over various reservoir sizes, shown in Figure 8(a). On average, compared to the ideal reservoirs, the performances of the nanowire reservoirs were better by 6.0%, 1.6%, and 0.8% for the *Biolek*, *Mohamed*, and *Najem* reservoirs, respectively. The power consumption, however, was similar for both the nanowire and ideal reservoirs, illustrated in Figure 8(b). On average, the power ratios between the nanowire and ideal reservoirs were 1.04, 1.01, and 0.99 for the *Biolek*, *Mohamed*, and *Najem* reservoirs, respectively. The results show that, compared to the ideal reservoirs, the nanowire resistance enhances the reservoir

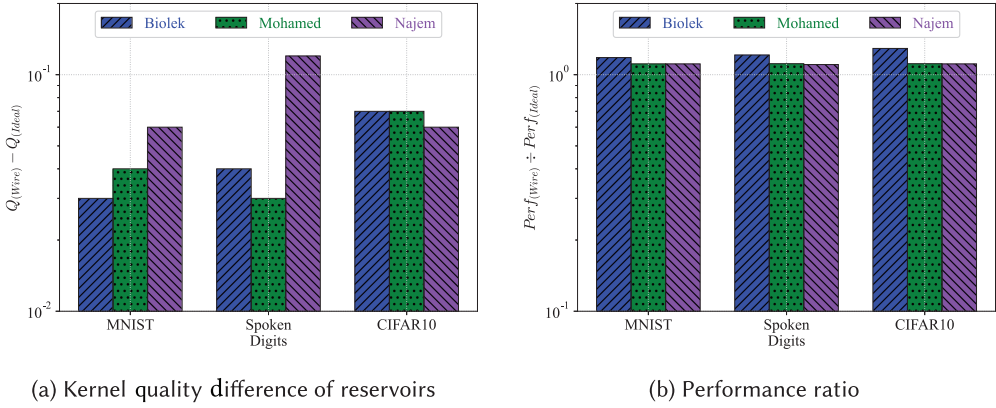(a) Kernel quality difference of reservoirs

(b) Performance ratio

Fig. 9. Kernel quality difference and performance ratio of ideal and nanowire reservoirs. Both ideal and nanowire reservoirs had similar network settings for three tasks: MNIST, Isolated Spoken Digits, and CIFAR10. (a) Q_{ideal} and Q_{wire} are the kernel quality measurements of ideal and nanowire reservoirs. (b) $Perf_{ideal}$ and $Perf_{wire}$ are the performances of ideal and nanowire reservoirs.

Table 6. Average Simulation Times for Ideal and Nanowire Reservoirs

	Simulation Time (seconds)†					
Device Model	Spoken Digits		MNIST		CIFAR10	
	Ideal	Wire	Ideal	Wire	Ideal	Wire
Biolek [Biolek et al. 2013]	0.96	43.29	49.60	650.17	46.71	332.12
Mohamed [Mohamed et al. 2014]	1.17	52.97	34.01	451.50	44.49	526.58
Najem [Najem et al. 2019]	0.89	54.80	28.12	363.99	66.65	734.50

†The hardware platform for all simulations is a Linux server of 64 cores, x86_64, 256GB RAM, and Ubuntu version of 16.04.6 (Xenial Xerus). The simulation times are the average times of simulating 100 instances of memcapacitive reservoirs with similar network settings.

dynamics and improves reservoir performance without increasing power consumption for Isolated Spoken Digits.

The better kernel quality measurements led to better classifications reflected in the simulation results on different tasks: MNIST, Isolated Spoken Digits, and CIFAR-10. In Figure 9, the results show the kernel quality measurements (Q_{ideal} and Q_{wire}) and the performances ($Perf_{ideal}$ and $Perf_{wire}$) of the ideal and nanowire reservoirs. Both ideal and nanowire reservoirs had similar network settings (α, β, δ, and γ) for all simulations. As shown in Figure 9(a), all memcapacitive reservoirs manifested a common trend in which the nanowire reservoirs had better kernel quality on the three tasks compared to the ideal ones. Having better kernel quality allowed the nanowire reservoirs to classify better inputs (Figure 9(b)). Compared to the ideal reservoirs, the combination of the nanowire resistance and device capacitance improved the dynamics of the nanowire reservoirs with the increase in the classification performances, on average, by 3%, 4%, and 7% for MNIST, Isolated Spoken Digits, and CIFAR-10. The nanowire reservoirs performed better than not only the ideal reservoirs but also their counterpart memristive reservoirs. In Figure 10, the nanowire memcapacitive reservoirs achieve similar performance (Figure 10(a)), compared to the memristive ones, but consumed much less power (Figure 10(b)), about 99×, 17×, and 277× less power on the three tasks, respectively.

The previous results (Figure 9) show that the interaction of the nanowire resistance and the device capacitance increased the dynamics of the nanowire reservoirs. It is unclear, however, how

(a) Performance of memcapacitive and memristive reservoirs

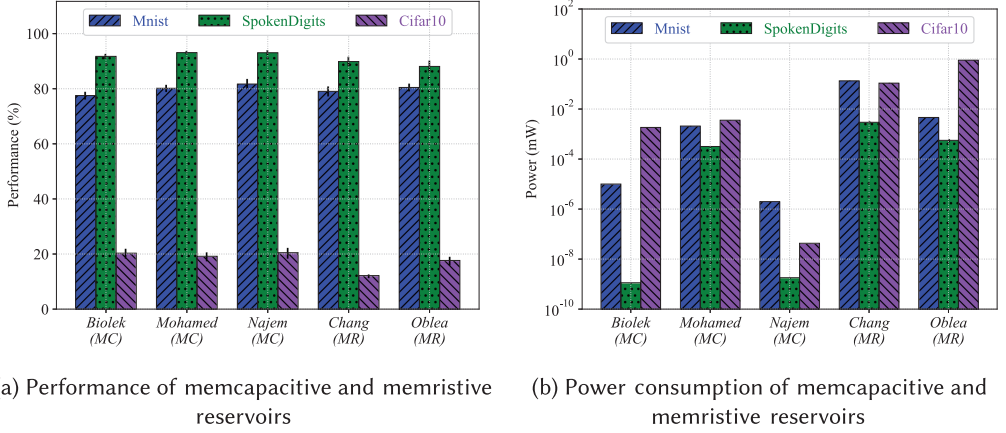(b) Power consumption of memcapacitive and memristive reservoirs

Fig. 10. Performance and power consumption of memcapacitive and memristive reservoirs. Compared to the memristive reservoirs, on average, the nanowire memcapacitive reservoirs achieved comparable performance on three tasks (MNIST, Isolated Spoken Digits, and CIFAR10) (a), but with much less power (b), about 99×, 17×, and 277× less, respectively.

the dissipated power in the nanowire resistance contributed to the overall power consumption of the reservoirs. In the next experiment, we studied the effect of the nanowire resistance on various topological structures of reservoir networks. As shown in Section 3, the characteristics of SWPL networks emerge with different settings for D, α, β, δ, and γ, particularly the randomness β and the locality α. Here, we explored the effect of the nanowire connects with different settings of the locality α ([0.0, 2.4]) and the randomness β ([0.0, 1.0]).

For each setting of α and β, 100 instances of a memcapacitive reservoir with similar network parameters were simulated to obtain average results. In our simulation results, the length of nanowire connections (L_{wire}) and the dissipated power in nanowire resistors (P_{nw}) represent the cost to systems, associated with different values of α and β. The dissipated power P_{nw} of nanowire resistors was included in the total power consumption of reservoirs. Since the Isolated Spoken Digit recognition task was more complex than classifying MNIST patterns and less dimensional than CIFAR-10 images, it was selected as a benchmark for the effect of nanowires on the topological change of memcapacitive networks. All memcapacitive reservoirs were trained with 1,000 digits and tested with 500 digits. The digit input signals were scaled to an input range of $[-\nu, \nu]$ to avoid overstimulating reservoirs. For each setting of α and β, the average results were obtained from the simulations of 100 instances of reservoirs that had the same network settings.

Since both α and β contribute to the dynamic change in reservoir structures, two simulation sweeps were performed: the locality and the randomness sweeps. For the locality sweep, several fixed values of β were chosen to represent different network structures, from a regular topology ($\beta = 0.0$) to a completely random structure ($\beta = 1.0$). Similarly, several fixed values of α, from a highly global connected type ($\alpha = 0.0$) to a highly local connected type ($\alpha = 2.4$), were selected for the randomness sweep. The classification performances of the memcapacitive reservoirs on the two sweeps are shown in Figure 11. From the results of both sweeps, the reservoirs achieved similar performance. The results suggest that the performance of the memcapacitive reservoirs is independent of their network structures.

According to the study of Verstraeten and Schrauwen [2009], the synaptic dynamics of node connections, represented by a reservoir weight matrix, determine the dynamic state of an ESN, a traditional RC. The dynamic state of an ESN is quantified by a minimal local Lyapunov
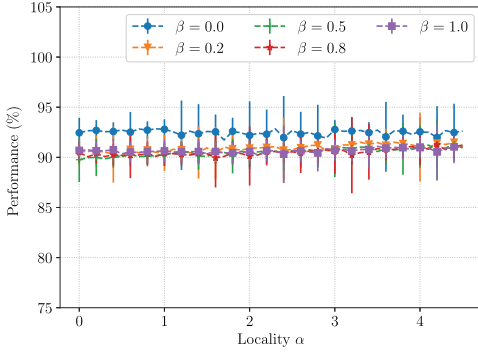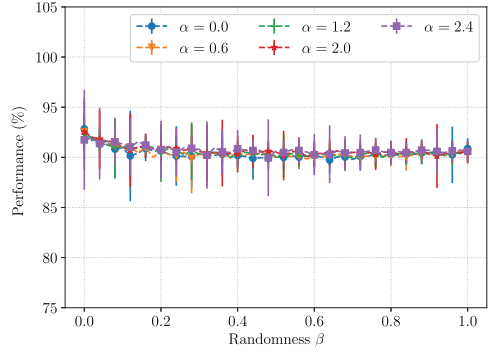
(a) Simulation sweep of locality α

(b) Simulation sweep of randomness β

Fig. 11. Performance of *Biolek* reservoirs from α and β sweeps. All reservoirs had a similar number of outputs (69 outputs) and nodes (420 nodes). (a) The values of β were from 0.0 (regular network) to 1.0 (completely random network). (b) The values of α were from 0.0 (highly global connected network) to 2.4 (highly local connected network).
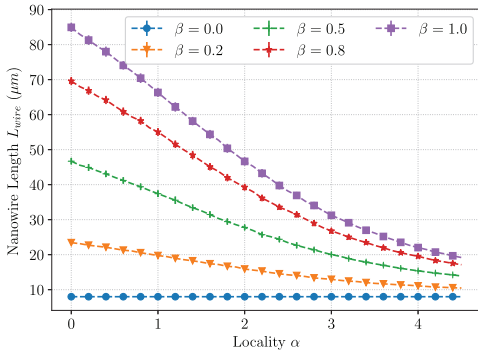
(a) Nanowire length on locality sweep
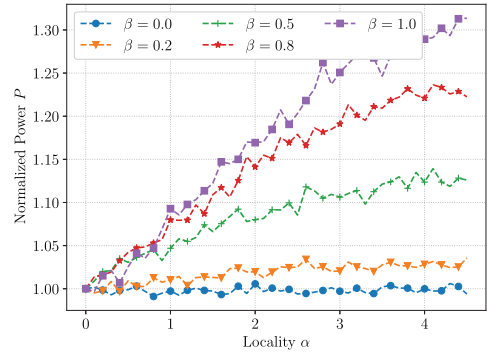
(b) Power consumption on locality sweep

Fig. 12. Total nanowire length and power consumption of *Biolek* reservoirs. All reservoirs had a similar number of outputs (69 outputs) and nodes (420 nodes); highly local connected reservoirs (large values of α) reduced wiring cost (a), represented as L_{wire}, but also consumed more power (b).

exponent based on the changes in the reservoir weight matrix. In this respect, the number of synaptic connections is the primary factor to influence the dynamics of a system, in which an ESN with a higher number of synaptic connections will have a better dynamic response. In our simulations, all memcapacitive reservoirs had the same number of reservoir nodes throughout the sweeps. With the same number of reservoir nodes, all reservoirs had the same number of memcapacitive devices to represent synaptic connections, which, in turn, contributed to the performance of the reservoirs. Since all reservoirs had a similar number of memcapacitive devices as synaptic connections, it is expected that reservoirs achieved similar performance as shown in Figure 11.

Although the topological changes in the network structure of the memcapacitive reservoirs did not affect their classification performances, they had a significant impact on wiring cost L_{wire} and power consumption P. Figure 12 illustrates that effect on the locality sweep. From Figure 12(a), we observe that the highly global connected reservoirs (small values of α) had a higher number of the total nanowire length (L_{wire}) compared to the highly local connected reservoirs (large values of

(a) Nanowire length on randomness sweep

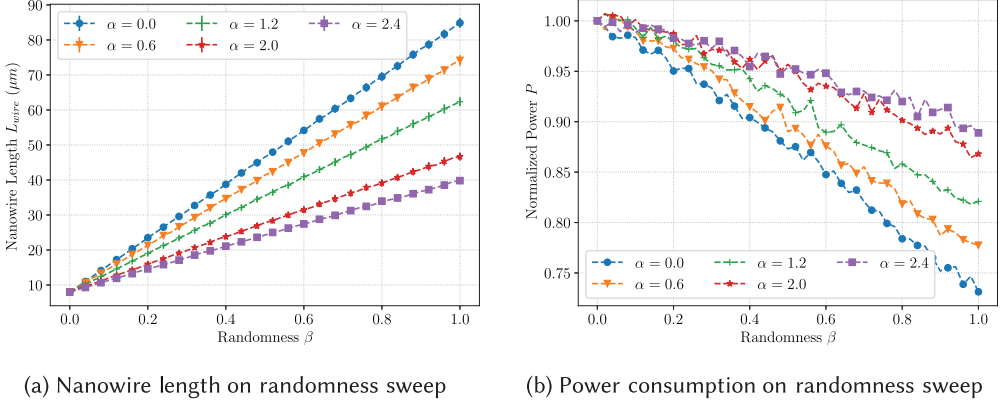(b) Power consumption on randomness sweep

Fig. 13. Total nanowire length and power consumption of *Biolek* reservoirs. All reservoirs had a similar number of outputs (69 outputs) and nodes (420 nodes); highly random reservoirs (large values of β) required more system resources, denoted by L_{wire}, (a) but consumed less power (b).

α). Recalling from Section 3, α is the exponent of the decay power-law function that governs the connections between nodes. A small value of α allows the formation of more global connections, whereas a large value of α elevates more the establishment of short connections. Consequently, the highly global connected reservoirs (small values of α) had more global connections, which resulted in a higher number of the total nanowire length, compared to the highly local connected ones.

In contrast to the results of the nanowire length, the highly local connected reservoirs consumed more power, compared to the highly global connected reservoirs. As shown in Section 4.4, the resistance of a nanowire connection plays an important role in connecting with the capacitor of a memcapacitive device to form an RC circuit. In an RC circuit, the time constant ($\tau = RC$) determines the charging/discharging rate of the capacitor. A small value of the resistor allows the capacitor to charge/discharge faster compared to a large one. Consequently, an RC circuit with a small resistor absorbs more energy (or power), indicated by the charge stored in the capacitor, than an RC circuit with a large resistor. In our case, the dissipated energy in the nanowire resistor is small, due to the small values of nanowire resistance ($m\Omega$), and the charge storage in memcapacitive devices is the dominant factor for energy consumption. The small nanowire resistance, however, does contribute to the time constants of the RC circuits within the memcapacitive reservoirs. For an RC with a global nanowire connection, the capacitor charges/discharges slower due to the high resistance of the long nanowire connection, compared to one with a local (or short) connection. As a result, highly global connected reservoirs, which have more long connections (or higher resistance), absorb more energy (or power) than highly local connected reservoirs. The simulation in Figure 12(b) reflects this phenomenon.

The simulation results of the randomness sweep are shown in Figure 13. Here, we observe that the highly random reservoirs (large values of β) required more system resources, expressed in L_{wire}, than the less random ones (Figure 13(a)). As explained in Section 3, β is the probability of a link in a regular network to be rewired during the formation of an SWPL network. With a higher value of β, more connections in a regular network are rewired with global and long connections to form an SWPL network. Subsequently, highly random reservoirs have a higher number of the total nanowire length, reflected in our simulation results.

In Figure 13(b), the power consumption of the memcapacitive reservoirs followed an opposite trend, in which the highly random reservoirs used less power compared to the less random ones.
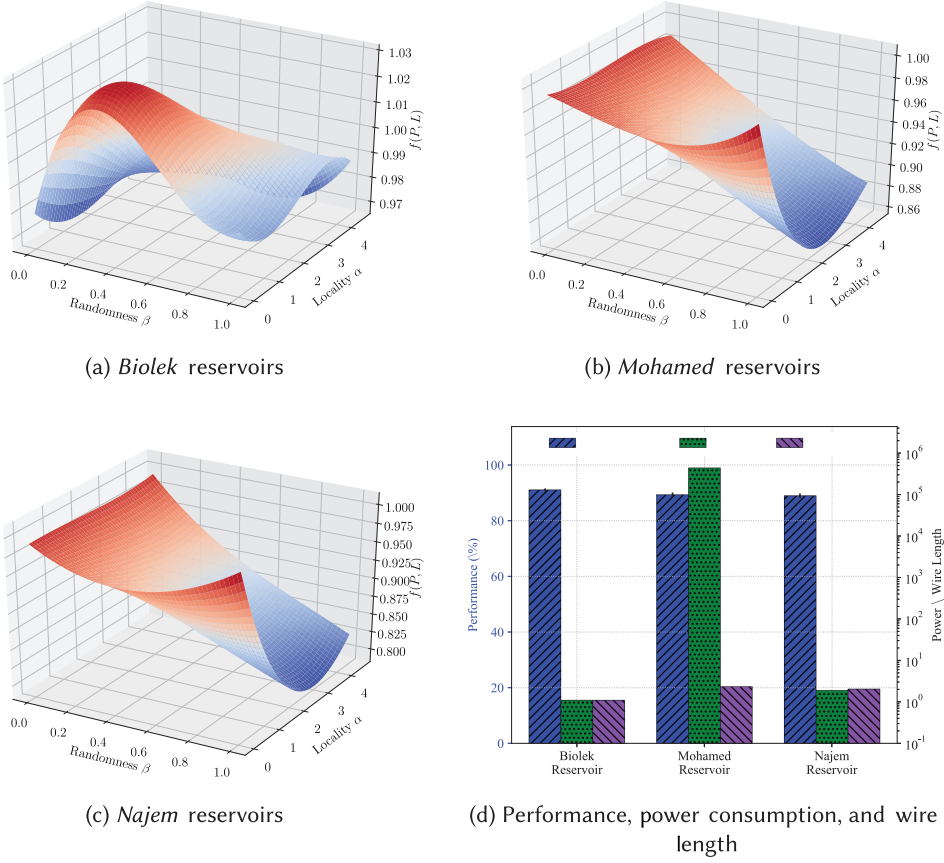
(a) *Biolek* reservoirs

(b) *Mohamed* reservoirs

(c) *Najem* reservoirs

(d) Performance, power consumption, and wire length

Fig. 14. The summation of power consumption P and wire length L_{wire} as a trade-off function $f(P, L)$ for different values of α and β. The interpolated points were based on the bivariate B-spline interpolation functions $bisplrep()$ and $bisplev()$, the two common functions for finding spreading points across a surface [Gentile et al. 2013]. The trade-off function $f(P, L)$ achieved its minimum value at different values of α and β: 4.5 and 0.61 for *Biolek* reservoirs (a), 2.7 and 1.0 for *Mohamed* reservoirs (b), and 3.0 and 1.0 for *Najem* reservoirs (c). The performances of the memcapacitive reservoirs at the minimum values of $f(P, L)$ are shown in (d).

Similar to what we saw in the locality sweep, the highly random reservoirs had more global connections or larger nanowire resistance. The large nanowire resistance, in turn, increased the time constant of RC circuits, which resulted in less energy absorption. Therefore, the highly random reservoirs consumed less power compared to the less random ones.

The simulation results from the locality and randomness sweeps illustrated the reciprocal relationship between the power consumption P and the wire length L_{wire} of memcapacitive reservoirs as they underwent topological changes for different values of α and β. Highly local connected reservoirs (high values of α) reduced the wiring cost, expressed in terms of the total wire length, but also consumed more power. In a similar respect, highly random reservoirs (high values of β) required more system resources but also depleted less power. Finding the values of α and β for a minimum trade-off between wiring cost and power consumption is an important task in designing a physical and practical system. Here, we defined a trade-off function as the sum of power consumption and nanowire length: $f(P, L) = P + L$. Both P and L were normalized to a range of

[0.0, 1.0]. The missing points (or the interpolated values) were based on the bivariate B-spline interpolation functions *bisplrep*() and *bisplev*(), the two common functions for finding spreading points across a surface [Gentile et al. 2013]. The plots of the trade-off function for the *Biolek, Mohamed,* and *Najem* reservoirs are shown in Figure 14.

The results show that the trade-off function achieved its minimum value at the different values of α and β: 4.5 and 0.61 for *Biolek* reservoirs (Figure 14(a)), 2.7 and 1.0 for *Mohamed* reservoirs (Figure 14(b)), and 3.0 and 1.0 for *Najem* reservoirs (Figure 14(c)). Since the dynamics and the power consumption of reservoirs depend on the physical characteristics of memcapacitive devices, it is expected that different memcapacitive reservoirs yield different trade-off values for α and β. The values of α and β, in general, suggest that highly local connected and highly random reservoirs result in a better trade-off between system resources (L_{wire}) and their power consumption.

Figure 14(d) displays the performance, the power consumption, and the total wire length of the memcapacitive reservoirs at the minimum value of the trade-off function $f(P, L)$ for comparison. We observe that all three memcapacitive reservoirs accomplished comparable performance, but the *Biolek* reservoir consumed the least power, compared to the *Mohamed* and *Najem* reservoirs. The low power consumption of the *Biolek* reservoirs was due to the small values of the memcapacitive device (in pF range). The result suggests that selecting a smaller-value memcapacitive device will further improve the power consumption of reservoirs without increasing wiring cost.

6 CONCLUSION

Emerging nanoscale memory devices (mem-devices), such as memcapacitors, offer a possible hardware platform for RC. Memcapacitive networks, composed of nonlinear memcapacitive devices, have the computational capability as reservoirs, indicated by their two essential properties: fading memory and linear separation. Our results show that memcapacitive reservoirs express the fading memory at both the device and system levels. The memcapacitive reservoirs illustrate the property of linear separation through their kernel quality measurements on three different tasks— MNIST, Isolated Spoken Digits, and CIFAR-10, in which nanowire reservoirs have a better dynamic response, compared to ideal reservoirs, and improve their performance, on average, by 4.67%.

Our study on the topological transformation, based on the change in locality α and randomness β, reveals that the topological structures of the memcapacitive SWPL reservoirs do not affect their performance but significantly contribute to the wiring cost and the power consumption of the systems. The minimum trade-off between the wiring cost and the power consumption occurs at different values of α and β: 4.5 and 0.61 for *Biolek* reservoirs, 2.7 and 1.0 for *Mohamed* reservoirs, and 3.0 and 1.0 for *Najem* reservoirs. The results of our research illustrate the computational capacity of complex memcapacitive networks, functioning as reservoirs in RC architecture. Such networks with an SWPL topology are energy-efficient systems that are applicable to low-power applications such as mobile devices and the Internet of Things.

SUPPLEMENTARY MATERIALS

The supplementary materials are in the file SupplementaryMaterials.pdf.

REFERENCES

Ossama Abdel-Hamid, Abdel-Rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu. 2014. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22, 10 (2014), 1533–1545.

Sophie Achard, Raymond Salvador, Brandon Whitcher, John Suckling, and Ed Bullmore. 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of Neuroscience* 26, 1 (2006), 63–72.

Aron K. Barbey. 2018. Network neuroscience theory of human intelligence. *Trends in Cognitive Sciences* 22, 1 (2018), 8–20.

Danielle Smith Bassett and Ed Bullmore. 2006. Small-world brain networks. *Neuroscientist* 12, 6 (2006), 512–523.

D. Biolek, M. Di Ventra, and Y. V. Pershin. 2013. Reliable SPICE simulations of memristors, memcapacitors and meminductors. *Radioengineering* 22, 4 (July 2013), 945–968.

Jake Bouvrie, Tony Ezzat, and Tomaso Poggio. 2008. Localized spectro-temporal cepstral analysis of speech. *Reconstruction* 7192, 6474 (2008), 5755.

Jens Bürger, Alireza Goudarzi, Darko Stefanovic, and Christof Teuscher. 2015a. Hierarchical composition of memristive networks for real-time computing. In *Proceedings of the 2015 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH'15)*. IEEE, Los Alamitos, CA, 33–38.

Jens Bürger, Alireza Goudarzi, Darko Stefanovic, and Christof Teuscher. 2015b. Computational capacity and energy consumption of complex resistive switch networks. *AIMS Materials Science* 2, 4 (2015), 530–545.

Lars Büsing, Benjamin Schrauwen, and Robert Legenstein. 2010. Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Computation* 22, 5 (2010), 1272–1311.

Jia-Wei Cai, Li-Xing Li, Chao Xu, Yang Feng, Ya-Nan Zhong, Jian-Long Xu, Xu Gao, and Sui-Dong Wang. 2019. Organic thin film memcapacitors. *Applied Physics Letters* 114, 4 (2019), 043302.

Udaysankar Chockanathan, Anas Z. Abidin, Adora M. DSouza, Giovanni Schifitto, and Axel Wismüller. 2018. Resilient modular small-world directed brain networks in healthy subjects with large-scale Granger causality analysis of resting-state functional MRI. In *Proceedings of SPIE 10578, Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*. 105780B.

Philip Clarkson and Pedro J. Moreno. 1999. On the use of support vector machines for phonetic classification. In *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 2. IEEE, Los Alamitos, CA, 585–588.

James P. Crutchfield, William L. Ditto, and Sudeshna Sinha. 2010. Introduction to focus issue: Intrinsic and designed computation: Information processing in dynamical systems-beyond the digital hegemony. *CHAOS* 20 (2010), 037101.

Eleanor C. Demis, Renato Aguilera, Kelsey Scharnhorst, Masakazu Aono, Adam Z. Stieg, and James K. Gimzewski. 2016. Nanoarchitectonic atomic switch networks for unconventional computing. *Japanese Journal of Applied Physics* 55, 11 (2016), 1102B2.

T. D. Dongale, K. P. Patil, S. B. Mullani, K. V. More, S. D. Delekar, P. S. Patil, P. K. Gaikwad, and R. K. Kamat. 2015. Investigation of process parameter variation in the memristor based resistive random access memory (RRAM): Effect of device size variations. *Materials Science in Semiconductor Processing* 35 (2015), 174–180.

Chao Du, Fuxi Cai, Mohammed A. Zidan, Wen Ma, Seung Hwan Lee, and Wei D. Lu. 2017. Reservoir computing using dynamic memristors for temporal information processing. *Nature Communications* 8, 1 (2017), 2204.

François Duport, Bendix Schneider, Anteo Smerieri, Marc Haelterman, and Serge Massar. 2012. All-optical reservoir computing. *Optics Express* 20, 20 (2012), 22783–22795.

Victor M. Eguiluz, Dante R. Chialvo, Guillermo A. Cecchi, Marwan Baliki, and A. Vania Apkarian. 2005. Scale-free brain functional networks. *Physical Review Letters* 94, 1 (2005), 018102.

Marc Gentile, Frederic Courbin, and Georges Meylan. 2013. Interpolating point spread function anisotropy. *Astronomy & Astrophysics* 549 (2013), A1.

Mehdi Ghayoumi, Miguel Gomez, Kate E. Baumstein, Narindra Persaud, and Andrew J. Perlowin. 2018. Local sensitive hashing (LSH) and convolutional neural networks (CNNs) for object recognition. In *Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA'18)*. IEEE, Los Alamitos, CA, 1197–1199.

Sreetosh Goswami, Santi P. Rath, Damien Thompson, Svante Hedström, Meenakshi Annamalai, Rajib Pramanick, B. Robert Ilic,2020. Charge disproportionate molecular redox for discrete memristive and memcapacitive switching. *Nature Nanotechnology* 15, 5 (2020), 380–389.

Amr M. Hassan, Hai Helen Li, and Yiran Chen. 2017. Hardware implementation of echo state networks using memristor double crossbar arrays. In *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN'17)*. IEEE, Los Alamitos, CA, 2171–2177.

Zohar Jackson. 2016. Free Spoken Digit Dataset (FSDD). Retrieved February 1, 2020 from https://github.com/Jakobovski/free-spoken-digit-dataset

Herbert Jaeger. 2001. The "echo state" approach to analyzing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report* 148, 34 (2001), 13.

Herbert Jaeger. 2002. *Short Term Memory in Echo State Networks*. GMD Report 152. German National Research Institute for Information Technology, Bonn, Germany.

YeonJoo Jeong and Wei Lu. 2018. Neuromorphic computing using memristor crossbar networks: A focus on bio-inspired approaches. *IEEE Nanotechnology Magazine* 12, 3 (2018), 6–18.

Xu Jian, Chen Fu Qin, Zhan Wang, Lei Du, Sun Xiao Meng, Peng Zu Lai, Suo Xue Ling, Wang Ting, Zhang Jun Ran, and Qiyong Gong. 2018. Disrupted functional network topology in children and adolescents with post-traumatic stress disorder. *Frontiers in Neuroscience* 12 (2018), 709.

Sung Hyun Jo, Ting Chang, Idongesit Ebong, Bhavitavya B. Bhadviya, Pinaki Mazumder, and Wei Lu. 2010. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Letters* 10, 4 (2010), 1297–1301.

A. K. Khan and B. H. Lee. 2016. Monolayer MoS_2 metal insulator transition based memcapacitor modeling with extension to a ternary device. *AIP Advances* 6, 9 (2016), 095022.

Laurent Larger, Antonio Baylón-Fuentes, Romain Martinenghi, Vladimir S. Udaltsov, Yanne K. Chembo, and Maxime Jacquot. 2017. High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification. *Physical Review X* 7, 1 (2017), 011015.

Chen-Yu Lee, Saining Xie, Patrick Gallagher, Zhengyou Zhang, and Zhuowen Tu. 2015. Deeply-supervised nets. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*. ACM, New York, NY, 562–570.

Robert Legenstein and Wolfgang Maass. 2007. Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks* 20, 3 (2007), 323–334.

Junrui Li, Zhekang Dong, Li Luo, Shukai Duan, and Lidan Wang. 2020. A novel versatile window function for memristor model with application in spiking neural network. *Neurocomputing* 405 (2020), 239–246.

Ruxin Liu, Ruixin Dong, Shuchao Qin, and Xunling Yan. 2020. A new type artificial synapse based on the organic copolymer memcapacitor. *Organic Electronics* 81 (2020), 105680.

Ziyu Lv, Qikun Hu, Zong-Xiang Xu, Junjie Wang, Zhonghui Chen, Yan Wang, Meng Chen, Kui Zhou, Ye Zhou, and Su-Ting Han. 2019. Organic memristor utilizing copper phthalocyanine nanowires with infrared response and cation regulating properties. *Advanced Electronic Materials* 5, 4 (2019), 1800793.

Xiaofen Ma, Guihua Jiang, Shishun Fu, Jin Fang, Yunfan Wu, Mengchen Liu, Guang Xu, and Tianyue Wang. 2018. Enhanced network efficiency of functional brain networks in primary insomnia patients. *Frontiers in Psychiatry* 9 (2018), 46.

Wolfgang Maass, Thomas Natschläger, and Henry Markram. 2002. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation* 14, 11 (2002), 2531–2560.

Cory Merkel, Qutaiba Saleh, Colin Donahue, and Dhireesha Kudithipudi. 2014. Memristive reservoir computing architecture for epileptic seizure detection. *Procedia Computer Science* 41 (2014), 249–254.

M. G. A. Mohamed, H. Kim, and T. Cho. 2014. Modeling of memristive and memcapacitive behaviors in metal-oxide junctions. *Scientific World Journal* 2015 (2014), 910126.

Joseph S. Najem, Md Sakib Hasan, R. Stanley Williams, Ryan J. Weiss, Garrett S. Rose, Graham J. Taylor, Stephen A. Sarles, and C. Patrick Collier. 2019. Dynamical nonlinear memory capacitance in biomimetic membranes. *Nature Communications* 10, 1 (2019), 1–11.

Thomas Natschläger, Wolfgang Maass, and Henry Markram. 2002. The "Liquid Computer": A novel strategy for real-time computing on time series. *Special Issue on Foundations of Information Processing of TELEMATIK* 8, 1 (2002), 39–43.

Antonio S. Oblea, Achyut Timilsina, David Moore, and Kristy A. Campbell. 2010. Silver chalcogenide based memristor devices. In *Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN'10)*. IEEE, Los Alamitos, CA, 1–3.

Jill Arvindbhai Patel, Zarin Tasnim Sandhie, and Masud H. Chowdhury. 2018. Ternary device using graphene memcapacitor for post binary era. In *Proceedings of the 2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS'18)*. IEEE, Los Alamitos, CA, 1130–1133.

Thomas Petermann and Paolo De Los Rios. 2006. Physical realizability of small-world networks. *Physical Review E* 73, 2 (2006), 026114.

W. H. Qian, X. F. Cheng, Y. Y. Zhao, J. Zhou, J. H. He, H. Li, Q. F. Xu, N. J. Li, D. Y. Chen, and J. M. Lu. 2019. Independent memcapacitive switching triggered by bromide ion migration for quaternary information storage. *Advanced Materials* 31, 37 (2019), e1806424.

Roberto Rigamonti, Matthew A. Brown, and Vincent Lepetit. 2011. Are sparse representations really relevant for image classification? In *Proceedings of the 2011 Conference on Computer Vision and Pattern Recognition (CVPR'11)*. IEEE, Los Alamitos, CA, 1545–1552.

Mattia Rigotti, Daniel D. Ben Dayan Rubin, Xiao-Jing Wang, and Stefano Fusi. 2010. Internal representation of task rules by recurrent dynamics: The importance of the diversity of neural responses. *Frontiers in Computational Neuroscience* 4 (2010), 24.

Ali Rodan and Peter Tiňo. 2011. Minimum complexity echo state network. *IEEE Transactions on Neural Networks* 22, 1 (2011), 131–144.

Subhrajit Roy and Arindam Basu. 2016. An online structural plasticity rule for generating better reservoirs. *Neural Computation* 28, 11 (2016), 2557–2584.

Kelsey S. Scharnhorst, Juan P. Carbajal, Renato C. Aguilera, Eric J. Sandouk, Masakazu Aono, Adam Z. Stieg, and James K. Gimzewski. 2018. Atomic switch networks as complex adaptive systems. *Japanese Journal of Applied Physics* 57, 3S2 (2018), 03ED02.

Franz Selzer, Carlo Floresca, David Kneppe, Ludwig Bormann, Christoph Sachse, Nelli Weiß, Alexander Eychmüller, Aram Amassian, Lars Müller-Meskamp, and Karl Leo. 2016. Electrical limit of silver nanowire electrodes: Direct measurement of the nanowire junction resistance. *Applied Physics Letters* 108, 16 (2016), 163302.

Forrest Sheldon and Francesco Caravelli. 2020. The computational capacity of Mem-LRC reservoirs. In *Proceedings of the Neuro-Inspired Computational Elements Workshop*. ACM, New York, NY, 1–4.

David Snyder, Alireza Goudarzi, and Christof Teuscher. 2013. Computational capabilities of random automata networks for reservoir computing. *Physical Review E* 87, 4 (2013), 042808.

Nicholas Soures, Lydia Hays, and Dhireesha Kudithipudi. 2017. Robustness of a memristor based liquid state machine. In *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN’17)*. IEEE, Los Alamitos, CA, 2414–2420.

Olaf Sporns, Dante R. Chialvo, Marcus Kaiser, and Claus C. Hilgetag. 2004. Organization, development and function of complex brain networks. *Trends in Cognitive Sciences* 8, 9 (2004), 418–425.

Laura E. Suarez, Jack D. Kendall, and Juan C. Nino. 2018. Evaluation of the computational capabilities of a memristive random network (MN3) under the context of reservoir computing. *Neural Networks* 106 (2018), 223–236.

Junwei Sun, Gaoyong Han, and Yanfeng Wang. 2020. Dynamical analysis of memcapacitor chaotic system and its image encryption application. *International Journal of Control, Automation and Systems* 18, 5 (2020), 1242–1249.

Dat Tran and Christof Teuscher. 2017a. Memcapacitive devices in logic and crossbar applications. *International Journal of Unconventional Computing* 13, 1 (2017), 35–57.

Dat Tran and Christof Teuscher. 2017b. Memcapacitive reservoir computing. In *Proceedings of the 2017 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH’17)*. IEEE, Los Alamitos, CA, 115–116.

David Verstraeten, Joni Dambre, Xavier Dutoit, and Benjamin Schrauwen. 2010. Memory versus non-linearity in reservoirs. In *Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN’10)*. IEEE, Los Alamitos, CA, 1–8.

David Verstraeten and Benjamin Schrauwen. 2009. On the quantification of dynamics in reservoir computing. In *Proceedings of the International Conference on Artificial Neural Networks*. 985–994.

David Verstraeten, Benjamin Schrauwen, Michiel d’Haene, and Dirk Stroobandt. 2007. An experimental unification of reservoir computing methods. *Neural Networks* 20, 3 (2007), 391–403.

David Verstraeten, Benjamin Schrauwen, Sander Dieleman, Philemon Brakel, Pieter Buteneers, and Dejan Pecevski. 2012. Oger: Modular learning architectures for large-scale sequential processing. *Journal of Machine Learning Research* 13, 1 (2012), 2995–2998.

Zhongrui Wang, Mingyi Rao, Jin-Woo Han, Jiaming Zhang, Peng Lin, Yunning Li, Can Li, et al. 2018. Capacitive neural network with neuro-transistors. *Nature Communications* 9, 1 (2018), 1–10.

Duncan . Watts and Steven H. Strogatz. 1998. Collective dynamics of “small-world” networks. *Nature* 393, 6684 (1998), 440–442.

L. M. Wedepohl and L. Jackson. 2002. Modified nodal analysis: An essential addition to electrical circuit theory and analysis. *Engineering Science & Education Journal* 11, 3 (2002), 84–92.

Paul Yang, Hyung Jun Kim, Hong Zheng, Geon Won Beom, Jong-Sung Park, Chi Jung Kang, and Tae-Sik Yoon. 2017. Synaptic transistor with a reversible and analog conductance modulation using a $Pt/HfO_x/n-IGZO$ memcapacitor. *Nanotechnology* 28, 22 (2017), 225201.

G. Yin, Y. Tian, P. Wang, L. Cao, and J. Lang. 2018. Radiation-induced changes in small world network of brain in patients with nasopharyngeal carcinoma: A structure MRI imaging study. *International Journal of Radiation Oncology · Biology · Physics* 102, 3 (2018), e564–e565.

T. You, L. P. Selvaraj, H. Zeng, W. Luo, N. Du, D. Bürger, I. Skorupa, S. Prucnal, A. Lawerenz, and T. Mikolajick. 2016. Capacitive switching: An energy-efficient, $BiFeO_3$-coated capacitive switch with integrated memory and demodulation functions. *Advanced Electronic Materials* 2, 3 (2016), 1500352.

Lei Zhao, Zhen Fan, Shengliang Cheng, Lanqing Hong, Yongqiang Li, Guo Tian, Deyang Chen, et al. 2020. An artificial optoelectronic synapse based on a photoelectric memcapacitor. *Advanced Electronic Materials* 6, 2 (2020), 1900858.