Brain-Inspired Structural Plasticity through Reweighting and Rewiring in Multi-Terminal Self-Organizing Memristive Nanowire Networks

Original
Brain-Inspired Structural Plasticity through Reweighting and Rewiring in Multi-Terminal Self-Organizing Memristive Nanowire Networks / Milano, Gianluca; Pedretti, Giacomo; Fretto, Matteo; Boarino, Luca; Benfenati, Fabio; Ielmini, Daniele; Valov, Ilia; Ricciardi, Carlo. - In: ADVANCED INTELLIGENT SYSTEMS. - ISSN 2640-4567. - 2:8(2020), p. 2000096. [10.1002/aisy.202000096]

Availability:
This version is available at: 11696/64798 since: 2020-12-29T09:20:40Z

Publisher:
WILEY

Published
DOI:10.1002/aisy.202000096

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

16 September 2023
Brain-Inspired Structural Plasticity through Reweighting and Rewiring in Multi-Terminal Self-Organizing Memristive Nanowire Networks

Gianluca Milano, Giacomo Pedretti, Matteo Fretto, Luca Boarino, Fabio Benfenati, Daniele Ielmini,* Ilia Valov,* and Carlo Ricciardi*

Acting as artificial synapses, two-terminal memristive devices are considered fundamental building blocks for the realization of artificial neural networks. Current memristive crossbar architectures demonstrate the implementation of neuromorphic computing paradigms, although they are unable to emulate typical features of biological neural networks such as high connectivity, adaptability through reconnection and rewiring, and long-range spatio-temporal correlation. Herein, self-organizing memristive random nanowire (NW) networks with functional connectivity able to display homo- and heterosynaptic plasticity is reported thanks to the mutual electrochemical interaction among memristive NWs and NW junctions. In particular, it is shown that rewiring and reweighting effects observed in single NWs and single NW junctions, respectively, are responsible for structural plasticity of the network under electrical stimulation. Such biologically inspired systems allow a low-cost realization of neural networks that can learn and adapt when subjected to multiple external stimuli, emulating the experience-dependent synaptic plasticity that shape the connectivity and functionalities of the nervous system that can be exploited for hardware implementation of unconventional computing paradigms.

Cognitive functions of humans stem from the emergent behavior of biological neural networks composed of $\approx 10^{14} - 10^{15}$ synaptic connections in between neurons, whereby the high connectivity of the system provides robustness, adaptability, and fault tolerance.[1] Synaptic connections undergo rapid changes in synaptic strength in response to activity and recent history of the neuron (short-term plasticity) that shape information processing within the network.[2] These short-lived changes in synaptic weights can evolve in long-term changes that rely on alterations of synaptic connections, impacting the architecture and topology of the neural hardware. In addition to input-specific Hebbian changes in active synapses (homosynaptic plasticity), plasticity can be induced also at a larger population of synapses that were not active during the induction of Hebbian plasticity (heterosynaptic plasticity) and contributes to the stability and homeostasis of neural networks.[3] With the aim of emulating brain-inspired computing paradigms, neuromorphic functionalities have been implemented in artificial neural networks based on memristive device which functionalities are enclosed in nanoionic processes.[4-6]
Organized into large arrays, memristive devices acting as artificial synapses demonstrated supervised and unsupervised learning in conventional crossbar architecture.[7] Despite a bright future prospective for the development of next-generation artificial intelligence (AI) systems, it is hard for such rigid top-down architectures to emulate most typical features of biological neural networks such as high connectivity, adaptability through reconnection and rewiring, and long-range spatio-temporal correlation. Alternative ways using unconventional systems consisting of many interacting nano-parts have been proposed for the realization of biologically plausible architectures where the emergent behavior arises from a complexity similar to that of biological neural circuits.[8-19] However, these systems were unable to demonstrate bio-realistic implementation of structural plasticity including reweighting and rewiring and spatio-temporal processing of input signals similarly to our brain. Here, we report on emergent synaptic behavior of biologically inspired nanorarchitecture based on self-assembled and highly interconnected nanowire (NW) networks realized with a bottom-up approach, focusing on short-term changes. The operation principle of this is based on the mutual electrochemical interaction among memristive NWs and NW junctions composing the network and regulating its connectivity depending on the input stimuli. The functional connectivity of the system was shown to be responsible for heterosynaptic plasticity that was experimentally demonstrated and modeled in a multi-terminal configuration, where the formation of a synaptic pathway between two neuron terminals is responsible for a variation in synaptic strength also at nonstimulated terminals. These results highlight the ability of NW memristive architectures for building brain-inspired intelligent systems based on complex networks able to physically compute the information arising from multi-terminal inputs.

Taking inspiration from the recurrent connectivity of biological systems (Figure 1a), memristive devices based on highly interconnected NWs were realized by drop-casting Ag-NWs in suspension on a SiO$_2$ insulating substrate. Subsequent patterning of Au electrodes proceeds (Figure 1b and Figure S1, Supporting Information) without need of cleanroom facilities or nanolithographic steps. The high density of NW cross-point junctions ($\approx$10$^6$ NW junctions mm$^{-2}$) regulates the current/voltage distribution across the random network and ensures high connectivity of the system. Although single-crystalline Ag-NWs are highly conductive, the conductance of each NW junction is influenced by the mechanical stochasticity of the contact in between the crossed NWs and by the presence of an insulating polynvinylpyrrolidone (PVP) shell layer of $\approx$1–2 nm surrounding the Ag-NW core (Figure S2, Supporting Information). The PVP shell layer increases the NW junction resistance, thus representing one of the main issues for the realization of highly conductive electrodes based on Ag NWs.[20,21] However, PVP can be exploited as a solid electrolyte for exploiting cross-point junctions as electrochemical metallization memristive cells.[22] In this framework, the NW random network represents a nonlinear complex system that can be mapped onto a graph representation.[23] The NWs (nodes) are connected through memristive edges as shown in Figure 1c. When a voltage difference is applied between any couple of nodes in the system, the current flowing in the network is distributed according to the Kirchhoff’s current law and is regulated by the conductance (weight) of each memristive edge. The change of conductance in a single memristive edge is responsible for a redistribution of voltage/current across other nodes/edges of the network, inducing a cascade of conductance changes in other memristive edges through an avalanche effect that facilitates the emergence of spatially correlated structures of network activity. In particular, the NW network connectivity can be controlled by means of two different forms of synaptic plasticity involving different physical phenomena, different in nature:

1) “Reweighting” by manipulating the weight of connections in between NWs at cross-point junctions; As investigated by considering single NW cross-point junction devices (Figure 1d), an electrochemical potential difference applied between two intersection NWs induces anodic dissolution of Ag to form Ag$^+$ ions that migrate in the insulating shell layer under the action of the electric field and recrystallize to form a conductive bridge connecting the two NW cores as shown in Figure 1e. Single NW cross-point junction devices were observed to exhibit a wide range of pristine state resistances due to the mechanical stochasticity of the contact between intersecting NWs with the presence of the PVP coating layer (Figure 1f). However, under voltage sweep stimulation, cross-point junctions exhibited volatile resistive switching behavior characterized by an abrupt change of resistance in correspondence of the SET voltage that turns the device to a lower resistance state (Figure 1g). Additional electrical characterizations of single NW junctions are reported in Figure S3, Supporting Information, showing also that the SET event can be induced in both polarities due to the symmetric structure of the junction. Thus, the formation/rupture of an Ag conductive path at the NW intersection is responsible for the observed memristive behavior of NW cross-point junctions.

2) “Rewiring” by rupture/rewiring of NWs; As investigated in single NW devices (Figure 1h), Joule heating and electromigration-driven electrical breakdown events occurring in single NWs at high current densities[24,25] are responsible for the creation of a needle-like nanogap along the Ag-NWs. Interestingly, the electrical connection can be regenerated by forming a conductive filament within the nanogap assisted by field-driven electromigration and bipolar electrode effects[26] as shown in Figure 1d. Indeed, single NW devices characterized by a low-pristine state resistance exhibited breakdown events under proper stimulation (Figure 1i). After breakdown, the induced nanogaps (see inset of Figure 1j) behave as bipolar resistive switching elements exhibiting the typical pinched hysteretic loop in the I–V plane (Figure 1k). Additional electrical characterization of single NWs is reported in Figure S4, Supporting Information. Thus, the nanogap induced by the breakdown of a single NW starts to behave as a memristive element. As a consequence, it is worth noticing that the structural topology of the NW network can evolve depending on the network history, as breakdown events can divide single NW nodes into subnodes connected by a newly generated memristive edge.

The structural plasticity of the NW network related to the evolution of connectivity due to reweighting and rewiring effects of single NW cross-point junctions and single NWs leads to an emerging memristive response of the network. After initialization (Figure S5a,b, Supporting Information), the network measured in two-terminal configuration exhibits typical
memristive behavior (pinched hysteresis loop) in the $I-V$ plot (Figure 2a). By applying a positive voltage sweep from 0 to 1 V (sweep 1), the network changed from an initial high-resistance state (HRS) to a low-resistance state (LRS) during the SET process. Although the LRS is maintained during the voltage sweep from 1 to 0 V (sweep 2), voltage sweeps in the opposite polarity from 0 to $-0.8$ V (sweep 3) results in a RESET process turning the device to the initial HRS (sweep 4). Note that the switching polarities are dependent on the first stimulation during initialization. The endurance characteristics, tested by switching the device 300 times between HRS and LRS by means of full-sweep cycles, reveal that the memristive behavior was maintained over cycling (Figure S5c, Supporting Information). The time-evolution of the network stimulated by a constant subthreshold bias voltage over large time scales revealed nonequilibrium dynamics with persistent conductance fluctuations and metastability. The power-law dependence of the intrinsic noise fluctuations in these nonequilibrium dynamic systems of highly interconnected nonlinear elements suggests distributed connectivity and self-organized criticality characterized by the spontaneous emergence of complexity from simple local interactions$^{27,28}$ (Figure S6, Supporting Information). Similar power-law scaling behavior was observed in spontaneous neural oscillations generated in the human brain.$^{29}$ More importantly, network dynamics can be exploited for the emulation of short-term synaptic plasticity (STP) that regulates the information exchange and processing within biological neural networks.$^{30-32}$ By applying an over-threshold constant voltage bias, the network conductance (synaptic weight) between two pads (neuron terminals) can be gradually increased (facilitation), as shown in Figure 2b (details in S7,
Supporting Information). The gradual enhancement of network connectivity is related to cascade switching events of memristive elements constituting the network that self-selects the lowest-energy path for electronic conduction.\[9\] After facilitation, the synaptic weight gradually relaxes back toward the initial state, exhibiting a volatile behavior due to the spontaneous dissolution of Ag conductive filaments previously formed in memristive elements of the network.\[33\] Note that an increase in the voltage bias stimulation results in a more intense potentiation of network conductivity, as well as in a longer relaxation time. An important aspect is that the network can be programmed by pulses down to the μs timescale (S8, Supporting Information). The above-described network response to electrical stimuli allows the implementation of synaptic functionalities such as paired-pulse facilitation (PPF).\[34,35\] Indeed, the network stimulation by means of short voltage pulses (mimicking action potentials) repetitively applied to the presynaptic pad with short interpulse time intervals results in a gradual increase in the network conductivity as a function of the number of applied pulses. The emergence of PPF during voltage pulse stimulation and the subsequent spontaneous relaxation of the network conductance are reported in Figure 2c. The change in synaptic weight (Δω) and relaxation time can be modulated by changing the voltage pulse amplitude, with higher voltage pulses resulting in larger changes of Δω and longer relaxation times. Notably, PPF can be cyclically induced after device relaxation (S9, Supporting Information).

Due to the network connectivity, synaptic plasticity can be induced in each pair of nodes connected at least by one pathway of memristive edges. This means that a complex neural network composed of multiple synaptic pathways can be reproduced by a multi-terminal memristive network, in which each pad represents a “neuron terminal” (Figure 3a,b). In this configuration, the synaptic weight between any given pair of “neuron terminals” is regulated by the collective response of the network to external

Figure 2. Memristive and neuromorphic functionalities of NW networks. a) Resistive switching behavior of the NW network measured in a two-terminal configuration as schematized in the inset. b) Gradual increase in conductance (synaptic facilitation) observed by applying a 10 s voltage pulse and subsequent conductance relaxation due to volatile resistive switching behavior read with a voltage of 50 mV. c) Experimental demonstration of PPF in NW networks with gradual increase in conductance by stimulating the network with a train of 500 μs voltage pulses separated by 500 μs interpulse intervals (left panel) and relaxation process after stimulation (right panel) recorded by applying a constant voltage of 50 mV. The change in the synaptic weight (Δω) after 100 pulses is reported in the inset as a function of the applied pulse amplitude.
stimuli, endowing the system with intrinsic heterosynaptic plasticity. Note also that the behavior of each synaptic pathway is the result of the functional high connectivity of the system that averages the behavior of the multitude of all NW junctions. In biological systems, heterosynaptic plasticity is related to synaptic interactions that are responsible for a change in the strength of synapses that are not directly stimulated in addition to the specifically stimulated ones, providing distinct computational and learning properties to the network.\cite{3,36} Previous attempts showed that such synaptic interactions can be emulated in top-down fabricated devices based on Ag nanoclusters\cite{37} or 2D materials.\cite{38,39} Also, modulation of synaptic activity was reported by pulse shape engineering\cite{40,41} or light stimulation.\cite{42} Here, Ag-NW network shows that the stimulation of a synaptic pathway (S10, Supporting Information) results in synaptic weight changes not only in the directly stimulated synapse but also in other nonstimulated conductive pathways.\cite{43} Modeling of the system through a grid graph of resistances, with the blue path representing the stimulated synaptic pathway, was modeled through the formation of a lower resistance path between the stimulated nodes (Figure 3d). Modeling shows that the direct stimulation of a synapse results in a redistribution of voltage/current across the whole network and brings about a change in the effective conductance of other synaptic pathways (S12, Supporting Information). In this framework, the synaptic network activity can be mapped onto a correlation map where each pixel represents the resistance (synaptic strength) between a specific couple of pads (neurons). Figure 4a reports the correlation map of resistance variations (ΔR) across the network after direct stimulation of the synapse connecting neuron terminals I and II. In addition to a potentiation of the directly stimulated synaptic connection that exhibited the highest variation of resistance, the stimulation resulted also in changes in the strength of synaptic connections between nonstimulated neuron terminals. Note that experimental results are in qualitative accordance with model predictions (Figure 4b). Interestingly, the change in synaptic strength in nonstimulated synapses depends on the spatial location of the corresponding terminals. Larger changes in the synaptic weight can be observed in synaptic pathways directly connected to previously stimulated terminals I or II (synapse I–III, I–IV, I–V, II–III, II–IV, II–V), whereas almost no changes are observed in other spatially distant synaptic pathways (synapse III–IV, III–V, IV–V). This is because the synaptic

---

**Figure 3.** Implementation of heterosynaptic plasticity in multiterminal memristive networks. a) Schematization of the multi-terminal memristive NW network device with highlighted pads used during heterosynaptic experimental demonstration and b) corresponding biological representation of the device, where synaptic interactions lead to heterosynaptic plasticity characterized by a change in the synaptic weight of synapses that are not directly stimulated. c) Direct stimulation of synapse I–III in the NW network results in synaptic weight changes not only in the directly stimulated synapse but also in other nonstimulated conductive pathways. d) Modeling of the system through a grid graph of resistances, with the blue path representing the stimulated synaptic pathway.
pathway resistance is not strongly influenced by a peripheral change in network connectivity. Instead, the stimulation of a more central synaptic pathway, such as synapse I–III (Figure 4c,d), results in more generalized changes in synaptic weights in the system. These results, corroborated by considering also other pad configurations (Figure S10a–d, Supporting Information), reveal a strong dependence of heterosynaptic changes on the spatial location of the primary plasticity effect. In addition, the short-term behavior of heterosynaptic facilitation was investigated by monitoring the evolution over time of synaptic weights after stimulation. Experimental data and modeling of the evolution of ΔR over time in synaptic pathways connecting terminal I to other terminals after direct stimulation of either synapse I–II or synapse I–III are reported in Figure 4e–h, respectively. Time course of synaptic weights after different stimuli are reported in Figure S10e–h, Supporting Information. Results demonstrated that the device exhibits short-term heterosynaptic plasticity, with the strength of all synaptic pathways that tends to restore to the initial conditions over time due to the previously discussed memristive network relaxation (details of heterosynaptic relaxation modeling are reported in S13, Supporting Information). Apart from experimental fluctuations, experiment and model clearly show a similar spatio-temporal correlation. In this framework, the collective state of the network that depends on the physical location and time correlation of electrical stimuli represents a physical “reservoir” characterized by high nonlinear dynamics that can be exploited for reservoir computing (a feasible route for implementation of reservoir computing is proposed in S15, Supporting Information).

In conclusion, the electrochemically controlled connectivity of the memristive random NW network exhibits functional and structural plasticity that mimics the behavior of biological neural circuits by displaying homo- and heterosynaptic plasticity and activity-dependent changes in the connectivity map due to reweighting and rewiring effects. In contrast to conventional neural networks realized with a top-down approach based on memristive devices or transistors, these biologically inspired systems allow a low-cost realization of neural networks fabricated through a bottom-up approach that can learn and adapt when subjected to external stimuli, strictly mimicking the processes of experience-dependent synaptic plasticity that shape the connectivity and functionalities of the nervous system. These results represent a radically new approach toward the development of biologically inspired intelligent systems able to physically compute data from multiple inputs through the Kirchhoff’s laws. Furthermore, they highlight the way for the hardware implementation of unconventional computing paradigms where an input signal has to be mapped into a higher dimensional output. In this framework, there are several potential applications where the properties of the NW networks could...
provide a denser, more biorealistic hardware implementation. For instance, the randomness of the pristine NW connections may enable the realization of extreme learning machines and physical unclonable functions. The dynamic behavior of the memristive NW paves the way for efficient reservoir computing and spatio-temporal recognition. These properties, combined with the extremely low cost and scalability of the bottom-up approach, makes our memristive NW networks a breakthrough enabling technology for neuromorphic computing.

Experimental Section

NW Characterization: Ag-NWs with diameter of 115 nm and length of 20–50 μm in isopropyl alcohol suspension were purchased from Sigma-Aldrich. Chemical and structural characterization of NWs are reported in Figure S2, Supporting Information. Transmission electron microscopy (TEM) was performed by means of a FEI Tecnai F20ST equipped with a field emission gun (FEG) operating at 200 kV, X-ray photoelectron spectroscopy (XPS) was performed using a Kr source with energy of 1486.6 eV and using the C 1s peak position (284.8 eV) as calibration, dispersing Ag NWs on a SiO2 substrate.

Fabrication of NW Memristive Networks: NW random networks were fabricated by drop-casting and spontaneous solvent evaporation of Ag-NWs in isopropyl alcohol suspension (≈0.13%) on a SiO2 (1 μm)/Si commercial substrate. The network structural topology was characterized using field emission scanning electron microscopy (FE-SEM; Zeiss Merlin). The density of NW junctions was quantified through a rough estimation by counting the number of NW interconnections in random portions of SEM images. Metallic Au pads were realized on the NW network by sputtering and shadow mask (details in S1, Supporting Information). A sputter etching process was performed to remove the PVP layer surrounding the Ag NWs in correspondence of the electrode before Au deposition, thus allowing direct contact of the Au electrode with the Ag NW core.

Fabrication of Single Cross-Point NW Junctions and Single NW Devices: To fabricate devices based on single cross-point NW junctions and on single NWs (Figure 1d,h, respectively), Ag-NWs were distributed by drop-casting on an insulating SiO2 substrate prepatterned with a submillimeter probe circuit realized by direct laser writing lithography and Ti/Au deposition. Then, selected single cross-point NW junctions or single NWs were connected to the probe circuit by Pt deposition through an ion beam-induced deposition (IBID) with a gas injection system (GIS) in a FEI Quanta 3D Microscope. Note that the realization of contacts by this technique 1) ensures direct contact between the deposited Pt and the Ag-NW core as the few-nanometer thick PVP coating shell-layer is easily removed in the Pt/Ag contact area during IBID of Pt and 2) avoid any interaction of NWs with solvents and polymers necessarily used during conventional electron beam lithography (EBL) that can interact with the PVP coating layer and alter the original structure of the NW.

Two-Terminal Electrical Measurements: Electrical characterizations of single cross-point NW junctions, single NW devices, and NW networks in two-terminal configuration were performed using a Keithley 4200 semiconductor device analyzer equipped with pulse measuring units (PMUs) and coupled with a SemiProbe probe station. Electrical characterization of NW networks in two-terminal configuration (Figure 2) was performed by considering Au electrodes separated by ≈7 mm. I–V cycles shown in Figure 2a were performed by applying a voltage sweep of 0.27 V s⁻¹ and by externally imposing a compliance current (CC) of 20 mA. The change in the synaptic weight (Δw) reported in Figure 2c was calculated as Δw = [G(n) - G(1)]/G(1), where G(n) is the conductance of the synaptic pathway during the last pulse of the applied pulse train (n = 100). During measurements of the network relaxation process, a stress voltage of 50 mV was used for monitoring the network resistance to minimize the influence of the stress voltage on the relaxation process while ensuring at the same time a high signal-to-noise ratio to monitor the current time trace. All measurements were performed in air at room temperature.

Multi-Terminal Characterization: For multi-terminal characterization, the device was arranged in a conventional probe station equipped with multiple electrical probe tips controlled by micromanipulators. A Keithley 707 switch matrix was connected to the probes and the instruments to correct routing the measurements, allowing a sequential selection of each combination of pad pairs, while keeping other pads floating. The electrical characterization was conducted by using a TTI-TGA 1202 arbitrary waveform generator with a 100 MHz bandwidth able to deliver rectangular voltage pulses of various duration and amplitude, whereas the current flowing into the sample was collected by a Lecroy WaveSurfer 3024 oscilloscope with 200 MHz bandwidth and 4 GSamples s⁻¹ sampling rate. All measurements were controlled with general purpose interface bus (GPIB) connection by a remote desktop PC. Heteroelasticity was investigated by monitoring the resistance evolution in between all the considered synaptic pathways after direct stimulation of a selected synapse. Direct stimulation of a selected synapse was performed by applying a 1 s voltage pulse of 8 V at its terminals (Figure S8, Supporting Information), whereas electrical resistance between each couple of pads was sequentially read by applying a 100 μs voltage pulse with amplitude of 0.1 V. Correlation maps of resistance variation (ΔR) reported in Figure 4a–d and Figure S10a–d, Supporting Information were obtained by comparing resistance maps before and after the direct stimulation of a selected synapse. The resistance variation was calculated as ΔR = (Rpost_stimulus - Rpre_stimulus) and the variation of ΔR has to be intended as a decrease in resistance of the synaptic pathway after stimulation. The heterosynaptic relaxation process shown in Figure 4e–h and Figure S10e–h, Supporting Information was recorded by monitoring the evolution of ΔR over time of the selected synaptic pathways. Note that color maps limits were restricted to positive ΔR values for better data visualization, experimental data with ΔR < 0 arising from experimental data fluctuations were considered as tail values.

Modeling the Multi-Terminal NW Network Device: Simulation of the network response was performed in Python using the NetworkX package. Details of model implementation and the schematic flow used for modeling are shown in S11 and S12, Supporting Information, respectively, while modeling of heterosynaptic relaxation is reported in S13, Supporting Information.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The support by Mauro Raimondo in helping with SEM measurements, by Katarzyna Bejtka for performing TEM measurements, by Salvatore Guastella for performing XPS measurements, and by Thomas Poessinger for helping with graphics is gratefully acknowledged. Device fabrication was performed at “Nanofacility Piemonte”, a facility supported by the “Compagnia di San Paolo” foundation.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

heterosynaptic plasticity, memristive systems, nanoarchitectures, nanowire networks, neuromorphic systems


