



## The mise en scène of memristive networks: effective memory, dynamics and learning

Francesco Caravelli

To cite this article: Francesco Caravelli (2017): The mise en scène of memristive networks: effective memory, dynamics and learning, International Journal of Parallel, Emergent and Distributed Systems, DOI: [10.1080/17445760.2017.1320796](https://doi.org/10.1080/17445760.2017.1320796)

To link to this article: <http://dx.doi.org/10.1080/17445760.2017.1320796>



Published online: 02 May 2017.



Submit your article to this journal [↗](#)



Article views: 22



View related articles [↗](#)



View Crossmark data [↗](#)



# The *mise en scène* of memristive networks: effective memory, dynamics and learning

Francesco Caravelli<sup>a,b</sup>

<sup>a</sup>Invenia Labs, Cambridge, UK; <sup>b</sup>London Institute of Mathematical Sciences, London, UK

## ABSTRACT

We discuss the properties of the dynamics of purely memristive circuits using a recently derived consistent equation for the internal memory variables of the involved memristors. In particular, we show that the number of independent memory states in a memristive circuit is constrained by the circuit conservation laws, and that the dynamics preserves these symmetry by means of a projection on the physical subspace. Moreover, we discuss other symmetries of the dynamics under various transformations of the involved variables, and study the weak and strong non-linear regimes of the dynamics. In the strong regime, we derive a conservation law for the internal memory variable. We also provide a condition on the reality of the eigenvalues of Lyapunov matrices. The Lyapunov matrix describes the dynamics close to a fixed point, for which show that the eigenvalues can be imaginary only for mixtures of passive and active components. Our last result concerns the weak non-linear regime, showing that the internal memory dynamics can be interpreted as a constrained gradient descent, and provide the functional being minimized. This latter result provides another direct connection between memristors and learning.

## ARTICLE HISTORY

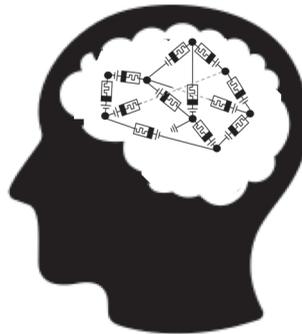
Received 13 November 2016  
Accepted 16 April 2017

## KEYWORDS

Exact results; memristive networks

## Complex Memristive Circuits for Learning:

- Circuit dynamics
- Effective memory
- Symmetries
- Oscillatory behavior
- Learning



We discuss the properties of the dynamics of purely memristive circuits using a recently derived consistent equation for the internal memory of the circuit. In particular, we show that the amount of memory in a memristive circuit is constrained by the conservation laws, and that the dynamics preserves these symmetry by means of a projection on this subspace. We obtain these results both for current and voltage controlled linear memristors. Moreover, we discuss other symmetries of the dynamics under various transformations, and study the weak and strong non-linear regimes.

In the strong regime, we derive a constrained conservation law for the internal memory. In particular, we are able to show that for the case of purely passive or active systems, the eigenvalues of the Jacobian are always real, implying that oscillations can emerge only for mixtures. Our last result concerns the weak non-linear regime, showing that the internal memory dynamics can be interpreted as a constrained gradient descent, and provide the functional being minimized. These results provide another direct connection between memristors and learning.

## 1. Introduction

There has been increasing interest in the properties of networks with memory. In the field of complex networks, memory is emerging as a new direction of study [1,2] in order to understand the properties of dynamical networks.

Memristors in particular have been attracting a renewed interest as these devices resemble swarms in solving certain optimization problems [3–6]. Memristors are 2-port devices which behave as resistances that change their values as a function of current or voltage. This type of memory is a common feature in many physical systems [7] and thus of general interest. Moreover, memristors have been proposed as building blocks for unconventional (or neuromorphic) computing [8,9], given that they are becoming easier to fabricate [10,11], although in specialised laboratories. It is thus interesting to study the behaviour of circuits of memristors, which we call memristive networks. These can serve also as simple models for further understanding the collective behaviour and learning abilities of many biological systems [12–16], including the brain [17,18] and its critical aspects. The behaviour of memristors is in spirit similar also to slime molds [19].

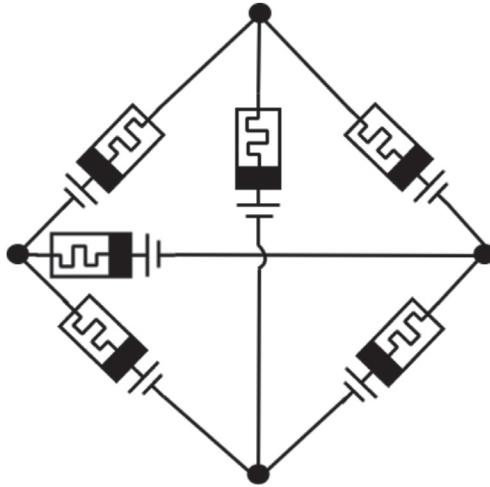
In a recent paper [20], a derivation of a ‘inner memory’ differential equation for purely memristive networks was obtained. It has been shown that several phenomena can be derived using this equation, such as a slow relaxation in the DC-controlled case, and an approximate analytical solution in the AC-controlled case. In order to derive such an equation, several graph-theoretic tools were used, which inherently showed the underlying freedom in describing the dynamics of the memory. In this paper, we further study such an equation and its underlying properties [21–26]. As an example, we provide an exact solution for the simple case of series of memristors in the *mean field* approximation, showing that it matches with the solution derived by simple circuit analysis. We then study the backbones of the dynamics: how the constraints structure typical of linear circuits is inherited by memristive networks dynamics. We describe what these constraints imply for the effective independent memory states. Also, we study the properties of the equation, its symmetries and variable transformations, and the weak and strong non-linear regimes.

We are moreover able to prove a constraint on the presence of oscillations around fixed points of the dynamics in the case in which there are no mixtures of passive and active elements. To conclude, we show in the limit of weak non-linearity that memristors perform ‘learning’: we cast the dynamics of the internal memory as constrained gradient descent, and provide the functional being minimized. Conclusions follow.

## 2. Memristive circuits

### 2.1. Memristive networks

We begin by briefly introducing the type of circuits we are interested in. First of all, we consider a particular class which, given a graph  $\mathcal{G}$  associated to the topology of the circuit, *each* edge of the graph can be replaced by a series of a voltage generator and a memristor. Thus, we do not consider the case in which inductors, resistors or capacitors are present in the circuit (although nothing obstruct a generalisation which includes these components). An example of such circuit is provided in Figure 1,



**Figure 1.** An example of memristive network considered in this paper. Given a graph  $\mathcal{G}$ , each edge of the graph can be associated with a series of a voltage generator and a memristor.

where the graph  $\mathcal{G}$  is a complete graph  $K_4$ . Also, we consider the case in which each memristor has a resistance which varies linearly as a function of the internal parameter  $w$ , e.g.

$$R(w) = R_{off}w + (1 - w)R_{on}, \quad (1)$$

where  $R_{off} \geq R_{on}$  are the two limiting resistances of the memristor, and  $0 \leq w \leq 1$ . We also consider a simple dynamics for the internal parameter  $w$ , which we identify as 'internal memory':

$$\frac{d}{dt}w = \alpha w - j \frac{R_{on}}{\beta} I. \quad (2)$$

In Equation (2),  $\alpha$  and  $\beta$  are the variables which set the timescales of the decay and the reinforcement due to the flow of the current respectively. The constant  $j = \pm 1$  is called polarity and is associated to the response (increasing or decreasing resistance) of the memristor to an external potential. Although here we describe only the theoretical properties of the dynamics, this type of memristors can be experimentally realised using atomic switches, which are based on silver ions filaments [27,28]. This type of memristor is called *ideal*, and it satisfies the current–voltage relationship typical of a resistor:  $V = R(w)I$ . Thus, it has the zero-crossing property, i.e.  $V = 0$  if  $I = 0$  and viceversa. Let us make a few comments. We consider the case of a voltage generator in series with the memristor simply because in the parallel case the dynamics of each specific memristor is trivial, and depends only on the voltage applied by the generator added in parallel. Moreover, the notation of Equation (1) differs from the one originally introduced in [21] by a transformation  $w \rightarrow 1 - w'$ . However, this change can be reabsorbed in the definition of polarity of each memristor, as  $\frac{d}{dt}w = -\frac{d}{dt}w'$  in Equation (2). We favour the notation of Equation (1) as the limiting internal variable value  $w = 0$  corresponds to the lower resistive state  $R_{on}$ . Physical memristors [27] satisfy a relaxation into the maximum resistance value  $R_{off}$ . We thus need to be careful, as in our notation  $\alpha > 0$  corresponds to a relaxation at zero potential to a  $R_{off}$  (insulating) state. Physically, this relaxation is related to an Ostwald ripening phenomenon [28].

Another point that we want here to make, is the distinction between active and passive components in our formalism. In the present paper, a passive component is an element which has the characteristics of a *positive resistance*. On the other hand, an *active* component is interpreted as a *negative resistance*, i.e. it satisfies  $V = -RI$ .

## 2.2. Graph theory, circuits and memristors

In the previous section we have introduced the simplest ideal memristor which will be considered in the bulk of this paper. In this section we recall the basic notions of graph theory which were used to derive the consistent internal memory differential equation in [20,29,30], which is the starting points for the analysis which follows. First of all, we consider a graph  $\mathcal{G}$  (a circuit) with  $N$  nodes and  $M$  edges (memristors) which describes the connectivity properties of the circuit. It is standard to start by choosing an orientation  $\mathcal{O}$  for the currents flowing in the circuit but, as we will see later, the analysis is independent from this choice from a physical point of view. In order for the graph to represent a circuit, the graph must be connected and the degree of each node  $i$  must satisfy  $d_i > 2$ , meaning each node is attached to at least two edges.

For the sake of clarity, we use latin indices for the edges, and greek indices for the nodes; greek indices with tildes will represent instead cycles on the graph. For instance, we will introduce a potential vector  $p_\alpha$ , and for each edge a current  $i_k$ .

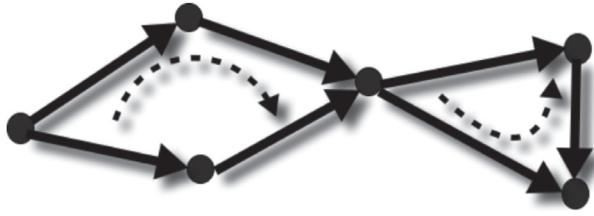
We now introduce a few mathematical definitions in order to clarify the discussion. Once an orientation  $\mathcal{O}$  has been assigned, and a set of oriented cycles is obtained, we can introduce two key matrices which will be used in the following: the directed incidence matrix  $B_{\alpha k}^{\mathcal{O}}$ , which is a matrix of size  $N \times M$ , and the cycle matrix  $A_{\xi \beta}^{\mathcal{O}}$ , which is of size  $C \times M$ , where  $C$  is the number of cycles of the graph,  $M$  the number of edges and  $N$  the number of nodes. The incidence matrix has an entry  $-1$  if an (oriented) edge is leaving a node,  $+1$  if it is incoming to a node, and  $0$  otherwise. The directed incidence matrix  $B_{\alpha k}$  labels edges on the rows and nodes on the columns:  $B_{\alpha k}$  takes values  $+1$  if an edge  $\alpha$  is incoming on a node  $k$ ,  $-1$  if it is outgoing, and  $0$  if the two are not incident. The cycle matrix has loop labels on its columns and edges on the rows:  $A_{\tilde{\alpha} \beta}$  has entry  $-1$  if the directed edge  $\beta$  is in the opposite direction of a chosen cycle  $\tilde{\alpha}$ ,  $+1$  if it shares the same orientation, and  $0$  if it does not belong to that cycle. In what follows, we will assume that an orientation for the cycles and the currents have been chosen, as in Figure 2.

One thing that should be stressed is that  $BB^t$  and  $B^tB$  are very different operators (where  $^t$  represents the matrix transpose): the former is usually called *laplacian* and is a matrix which acts on the set of *nodes*, meanwhile the latter is usually called *edge laplacian* [31,32] and acts on the set of *edges*. Both operators are however positively defined, as  $\tilde{n} \cdot BB^t\tilde{n} = (B\tilde{n}) \cdot (B\tilde{n}) \geq 0$ , and in the other case  $\tilde{n} \cdot BB^t\tilde{n} = (B^t\tilde{n}) \cdot (B^t\tilde{n}) \geq 0$ .

The conservation of the current at each node, the *first Kirchhoff law* or Current Law (KCL), can be written in terms of the incidence matrix  $B$  as  $\sum_{j=1}^M A_{\xi j} B_{\alpha j} = B^t \tilde{i} = 0$ . This set of equations contains a superfluous one. Thus, in order for  $B$  to have the linear independence of the rows, it is common practice to remove one of the rows and work with the *reduced incidence matrix*  $\tilde{B}$ . In the following, we will thus consider only results derived with this matrix rather than the full one and remove the  $\tilde{\cdot}$ .<sup>1</sup> The incidence matrix can also be used to calculate the voltage applied to each resistor from the potential at the nodes. Given a potential vector based on the nodes  $\tilde{p} = \{p_\xi\}$ , the vector of voltages applied to each resistor can be written as  $\{\tilde{v}\}_k = v_k = \sum_\xi B_{\xi k}^t p_\xi$ .

As for the case of the incidence matrix, also in the case of the cycle matrix one has to consider the reduced cycle matrix  $\tilde{A}$  when one row has been removed. The *second Kirchhoff law*, or Voltage Law (KVL) can be expressed mathematically as  $\sum_\beta A_{\tilde{\xi} \beta} v_\beta = 0$ . This implies that the voltage on each loop (or *mesh* in circuits) must be zero, which can also be written as  $\sum_\beta A_{\tilde{\xi} k} R_\beta i_\beta = 0$ . It is possible to write this alternatively as  $\sum_\beta A_{\tilde{\xi} \beta} \sum_\xi B_{\xi \beta}^t p_\xi = 0$ . Since this is true for arbitrary potential vectors  $p_\xi$ , this implies that in general one has  $B \cdot A^t = A \cdot B^t \equiv 0$ . Effectively, this equation represents the conservation of energy, which in the language of circuits is called *Tellegen's theorem*.

There are two ways to construct the basis of a circuit: either by using the tree and cycles decomposition, or alternatively by using the chords or the co-chords decomposition [33]. Here we consider the standard co-chords decomposition, which is based on *spanning trees*. Let us first introduce a maximal spanning tree  $\mathcal{T}$ , whose elements are called *co-chords*, and the set of edges of the graph not included in the tree, or *chords*, are given by  $\tilde{\mathcal{T}}$ . If there is more than one tree, each tree has the same



**Figure 2.** Example of fully oriented network: an orientation for each edge and for each cycle was made.

cardinality. Each chord element,  $\vec{T}$ , can be assigned to a cycle, called *fundamental loop*. The number of fundamental loops is constrained by the topology of the circuit, and is given by  $L = M - N + 1$ : this is the number of edges minus the cardinality of the tree,  $|\mathcal{T}| = N - 1$ . We resort to the appendix of [20] for all the details of the derivation of the equation for the internal memory dynamics. The important fact is however that using the Kirchhoff laws, it is possible to derive the following exact solution for the circuit currents, based only on the source vector  $\vec{S}(t)$ , the resistance matrix  $R$  and the cycle matrix [30]:

$$\vec{i} = A^t \vec{i}_c = -A^t (ARA^t)^{-1} A \vec{S}(t). \quad (3)$$

For the case of linear memristors, i.e.  $R(w) = R_{on}(1 - w) + wR_{off}$ , we have shown in [2] that it is possible to rearrange the right hand side in terms of only a projector operator  $\Omega = A^t(AA^t)^{-1}A$  on the space of cycles. This is done by carefully inverting only matrices which are invertible. The vector  $\vec{S}$  represents the voltage source in series to the memristances and is a vector with a number of entries equal to the number of memristors. This is due to the fact that in our definition each memristor is in series with a voltage generator. Since  $A$  is a reduced incidence matrix, then  $ARA^t$  (which is usually called *augmented cyclic matrix*) is always invertible for non-zero (thus also negative) resistances [34].

Specifically, in [20] the following exact equation for the internal memory of an ideal memristive circuit was derived:

$$\frac{d\vec{W}}{dt} = \alpha \vec{W} - \frac{\mathcal{J}}{\beta} (I + \xi \bar{\Omega} W)^{-1} \bar{\Omega} \vec{S}, \quad (4)$$

where  $\bar{S} = P\vec{S}$  and  $\bar{\Omega} = A^t(APA^t)^{-1}AP = A^t(\bar{A}A^t)^{-1}\bar{A}^2 P$  and  $J$  are diagonal matrices made of only  $\pm 1$ . In the case of the matrix  $P$ , elements associated with  $-1$  are active elements (negative resistance), meanwhile elements with  $+1$  are passive (positive resistance). The matrix  $\mathcal{J}$  is a generalisation of the polarity of each memristor. In general, a change of current orientation can in fact be reabsorbed in this matrix. In the following, we set  $\mathcal{J} = I$  for simplicity. The memristor variables  $w_i$  enter the dynamics by means of the diagonal matrix  $W = \text{diag}(\vec{W}) = \text{diag}(w_1, \dots, w_n)$ , where with  $\text{diag}(\cdot)$  we mean the diagonal matrix with the input vector as diagonal elements. By definition, barred quantities ( $\bar{\Omega}$ ,  $\bar{S}$ ) depend on  $P$ . As stressed in [20], these equations are true in the *bulk*, e.g. when all the memristors are not close to  $w = \{0, 1\}$ , and has been derived assuming  $W$  invertible. This is in general not true close to the lower boundary  $w = 0$ , but in simulations we have observed a smooth behaviour also in such case [2]. This implies that together with the equations (4), one should impose the constraint  $0 \leq w_i \leq 1$  to have a faithful dynamics of a physical memristive system.

Of course, Equation (4) describes a specific class of memristors (ideal memristors) and is by no means general. It can however be interpreted as first approximations for the real dynamics of memristors with a pinched hysteresis loop, in which the resistance is linear in the internal parameter, and the internal parameter varies linearly in the current.

As we have seen in Section 2.2,  $A$  and  $B^t$  are dual, i.e. they satisfy  $A^t B = B^t A = 0$ . In general, for any projector operator  $\Omega = A^t (AA^t)^{-1} A$ , if  $AB^t = A^t B = 0$ , then it is possible to show that

$$I - A^t (AA^t)^{-1} A = I - \Omega = B(B^t B)^{-1} B^t, \quad (5)$$

e.g. the fact that  $\Omega$  can be written both in terms of  $B$  or  $A$ . Given the fact that  $B$  is numerically much easier to calculate than  $A$ , this turns useful when performing simulations.

Equation (4) may seem quite obscure at first. Below we provide an example where calculations can be made without approximations. We show that known results can be re-derived using the equations above.

### 2.2.1. A specific example: exact mean-field solution for memristors series

In order to see how the equations above can be applied, let us consider a simple enough case in which calculations can be performed analytically, and which are nonetheless not trivial: this is the case of a series of  $N$  current-controlled memristors, for which in principle Equation (4) would not be needed. In this case the use of the graph-theoretical machinery is an overkill which however provides insights in the meaning of the operator  $\Omega$ .

For a series of memristors, the adjacency matrix is a Toeplitz matrix. Thus, the projector operator on the incidence matrix can be written as  $\Omega_B = 1 - \frac{1}{N} \bar{u} \bar{u}^t$ , where  $\bar{u} = [1 \cdots 1]^t$  is a column vector of size  $N$  with only 1's. Thus,  $\Omega = 1/N \bar{u} \bar{u}^t$ . We can thus calculate the inverse  $(1 + \xi \Omega W)$  if  $W$  has the same entries. In fact, we can use in this case the Sherman–Morrison identity [35]: one has that for any  $k \in \mathbb{R}$ ,  $(I + k \bar{u} \bar{u}^t)^{-1} = I - \frac{k}{1+kN} \bar{u} \bar{u}^t$ , thus if all memristors have the same initial value, one has  $(1 + \xi w \Omega)^{-1} = (1 + \frac{\xi w}{N} \bar{u} \bar{u}^t)^{-1} = I - \frac{\xi w}{N(1+\xi w)} \bar{u} \bar{u}^t$ . We can calculate the rhs of the dynamical equations exactly:

$$\begin{aligned} \frac{d}{dt} w(t) \bar{1} &= -\frac{1}{\beta} \left( I - \frac{\xi w}{N(1+\xi w)} \bar{u} \bar{u}^t \right) \frac{1}{N} (\bar{u} \bar{u}^t) \bar{s} \\ &= -\frac{1}{\beta} \left( I - \frac{\xi w}{(1+\xi w)} \right) \frac{1}{N} \bar{u} \bar{u}^t \bar{s} \\ &= -\frac{1}{\beta} \frac{1}{(1+\xi w)} \frac{1}{N} \bar{u} \bar{u}^t \bar{s}. \end{aligned} \quad (6)$$

We note that  $\frac{1}{N} \bar{u} \bar{u}^t$  is a projector, which implies:

$$\frac{d}{dt} w(t) = -\frac{1}{\beta} \frac{1}{(1+\xi w(t))} \frac{1}{N} s(t). \quad (7)$$

Equation (7) is the same that would be obtained for a series of  $N$  identical memristors if there was only a source voltage which is homogeneous across the circuit. Using the fact that we are approximating all the memristors with the same parameter, we used the rule for a series of resistors, to obtain:

$$\begin{aligned} I &= \frac{V}{R} \equiv \frac{s(t)}{\sum_i R_i(w)} \equiv \frac{s(t)}{\sum_i (R_{on} w + (1-w) R_{off})} \\ &= \frac{s(t)}{N (R_{on} w + (1-w) R_{off})} = \frac{1}{N R_{on}} \frac{s(t)}{(1+\xi w)} \end{aligned} \quad (8)$$

and using the fact that  $\frac{d}{dt} w = \alpha w - \frac{R_{on}}{\beta} I$  we obtain the final Equation (7). We note that if  $\bar{w}$  is not uniform (i.e. when we do not use the mean field approximation), then it is not possible to neglect the correlation arising from the denseness of  $\Omega$ .

### 2.3. Network constraints and effective memory

Network constraints are fundamental in order to make precise the notion of *effective memory* in memristive networks. In order to see this, let us look at the constraints and how these affect the internal memory capacity. The analysis which follows below applies to the case of a *linear* relationship between the memristor' memory and either voltage or current. We consider first two specific models for the evolution of the internal memory in the 'bulk' (far from the boundaries). These are

$$\text{Current-controlled memristors: } \frac{d}{dt}w = a\vec{i}, \quad (9)$$

and

$$\text{Voltage-controlled memristors: } \frac{d}{dt}w = b\vec{v}, \quad (10)$$

which are two different types of memristors considered in the literature [36,37], and  $a$  and  $b$  are simply constants. In both Equations (9) and (10) one can use the network constraints in order to obtain information on the exact number of independent memory states stored by the circuit. In order to see this, we note first that the Kirchhoff current constraint can be written as:

$$B\vec{i} = 0, \quad (11)$$

and if we now combine the internal memory equation for the current-controlled memristors of Equation (9), it is easy to see that:

$$B\vec{i} \propto B \frac{d}{dt}\vec{w} = 0. \quad (12)$$

At this point we can use the tree and co-tree splitting to write the following linear relationship:

$$B \frac{d}{dt}\vec{w} = B_{\mathcal{T}} \frac{d}{dt}\vec{w}_{\mathcal{T}} + B_{\bar{\mathcal{T}}} \frac{d}{dt}\vec{w}_{\bar{\mathcal{T}}} = 0. \quad (13)$$

We thus obtain the final formula which connects the derivatives of the memory on the tree and the co-tree elements:

$$\frac{d}{dt}\vec{w}_{\mathcal{T}} = -B_{\bar{\mathcal{T}}}^{-1}B_{\mathcal{T}} \frac{d}{dt}\vec{w}_{\bar{\mathcal{T}}}. \quad (14)$$

This equation can be then integrated in time to obtain, for current-controlled memristors, the following result:

$$\vec{w}_{\mathcal{T}}(t) = -B_{\bar{\mathcal{T}}}^{-1}B_{\mathcal{T}} \left( \vec{w}_{\bar{\mathcal{T}}}(t) + \vec{w}_{\bar{\mathcal{T}}}^0 \right), \quad (15)$$

where the constant  $\vec{w}_{\bar{\mathcal{T}}}^0$  arises from integrating the equation over time. Before we provide an interpretation of the above result, we want first to show that such analysis applies also in the case of voltage controlled memristors. In this case, rather than the conservation of the currents at the nodes, we use the conservation of voltage across a cycle. This can be represented by the following equation dependent this time on the cycle matrix  $A$ :

$$A\vec{v} \propto A \frac{d}{dt}\vec{w} = 0. \quad (16)$$

If we use the tree-chords splitting again, we have the same formalism as before, by replacing  $B_{\mathcal{T}}$  with  $A_{\mathcal{T}}$ . We thus obtain:

$$\vec{w}_{\mathcal{T}}(t) = -A_{\bar{\mathcal{T}}}^{-1}A_{\mathcal{T}} \left( \vec{w}_{\bar{\mathcal{T}}}(t) + \vec{w}_{\bar{\mathcal{T}}}^0 \right). \quad (17)$$

Both Equations (15) and (17) are representations of the constraints of the network. In both cases, we can write the equation for the internal memory as:

$$\vec{w}_{\mathcal{T}}(t) = Q_{\mathcal{T}} \left( \vec{w}_{\bar{\mathcal{T}}}(t) + \vec{w}_{\bar{\mathcal{T}}}^0 \right) \quad (18)$$

where  $Q_{\mathcal{T}}$  is a linear operator which depends on the chosen spanning tree, and thus for the whole memory as

$$\begin{aligned}\vec{w}(t) &= (Q_{\mathcal{T}}\vec{w}_{\bar{\mathcal{T}}}(t), \vec{w}_{\bar{\mathcal{T}}}(t)) + (Q_{\mathcal{T}}\vec{w}_{\bar{\mathcal{T}}}^0, \vec{0}_{\bar{\mathcal{T}}}) \\ &= (Q_{\mathcal{T}}, I)\vec{w}_{\bar{\mathcal{T}}}(t) + (Q_{\mathcal{T}}\vec{w}_{\bar{\mathcal{T}}}^0, \vec{0}_{\bar{\mathcal{T}}})\end{aligned}\quad (19)$$

This is general, and it is valid both for current-controlled and voltage-controlled memristors, as long as these are linear at the first order approximation. It is easy to see that Equation (19) establishes an affine relationship between the internal memory and a subspace of chord memory. We can thus introduce the concept of *effective memory* of a memristive circuit  $\mathcal{G}$ , given by:

$$\Gamma(\mathcal{G}) = \frac{|E(\mathcal{G})| - |\mathcal{T}(\mathcal{G})|}{|E(\mathcal{G})|}\quad (20)$$

where  $|E(\mathcal{G})|$  is the number of memristive elements and  $|\mathcal{T}(\mathcal{G})|$  represents the cardinality of a maximal spanning tree in the circuit  $G$ . Since  $M$  can grow as the number of nodes of the circuit square, meanwhile  $|\mathcal{T}(\mathcal{G})|$  grows linearly in the number of nodes, this implies that denser circuits can effectively overcome the limitation of a smaller internal capacity. We note that the effective capacity of Equation (20) is well defined: this number is independent from the choice of the spanning tree, and thus is a relevant physical quantity, meanwhile Equation (19) implicitly depends on the choice of the spanning tree. Specifically, the number of ways in which Equation (19) can be written depends on the number of spanning trees of the circuit.

As simple as such argument may look, it shows that the effective memory in a memristive circuit lives on a submanifold of the internal memory. Once a spanning tree has been chosen, and the dynamical equations derived, the projection operator ensures that such sub-manifold is protected and well defined in the dynamics.

#### 2.4. Strongly and weakly non-linear regimes: two different limits for the dynamics

In this section we study the behaviour of the dynamics in the weak and strong non-linear regimes. There are at least two regimes that we would like here to describe:  $\xi \approx 0$ , which we call *weakly non-linear regime*, and  $\xi \rightarrow \infty$ , we call *strongly non-linear regime*. We focus here on the case of current-controlled memristors, but a similar analysis applies also to voltage-controlled memristors. The differential equation for the internal dynamics, for  $\alpha = 0$ , is given by:

$$\frac{d}{dt}\vec{W} = -\frac{1}{\beta}(I + \xi\bar{\Omega}W)^{-1}\bar{\Omega}\vec{S} = -\frac{1}{\beta}T(\xi)\vec{S},\quad (21)$$

where we introduced the definition of the operator  $T(\xi)$ . The two regimes of  $\xi$  can be understood from the analysis of the behaviour of the operator

$$T(\xi) \equiv \left(I + \left(\frac{R_{on}}{R_{off}} - 1\right)\bar{\Omega}W\right)^{-1}\bar{\Omega} = (I + \xi\bar{\Omega}W)^{-1}\bar{\Omega},\quad (22)$$

which we will now try to make precise from an operatorial point of view in both limits.

In the weakly nonlinear regime, i.e.  $R_{off} \approx R_{on}$ , the following Taylor expansion of the operator applies:

$$\lim_{\xi \rightarrow 0} T(\xi) \approx (I - \xi\bar{\Omega}W)\bar{\Omega} + O(\xi^2).\quad (23)$$

This regime will be studied in detail in Section 4, showing that we can identify the 'learning' abilities of the memristive circuit. In the strong non-linear regime instead, it does make sense to write

$$\lim_{\xi \rightarrow \infty} T(\xi) = \lim_{\xi \rightarrow \infty} \frac{1}{\xi} \left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1}\bar{\Omega}.\quad (24)$$

We note that  $\left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1}$  for large  $\xi$ , can be seen as the Tychonov regularisation of the inverse of the operator  $\bar{\Omega}W$ . Equation (4) was derived with the assumption that  $W$  is an invertible (diagonal) matrix, i.e. that no memristor reached the  $R_{on}$  state. We can thus write

$$\lim_{\xi \rightarrow \infty} \left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1} (\bar{\Omega}W) (\xi W)^{-1}, \quad (25)$$

and study for the time being how does  $\lim_{\xi \rightarrow \infty} \left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1} (\bar{\Omega}W)$  behave. The Tychonov regularisation converges to the Moore–Penrose pseudo-inverse of  $\bar{\Omega}W$ , implying that  $\lim_{\xi \rightarrow \infty} (\bar{\Omega}W) \left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1} (\bar{\Omega}W) = \bar{\Omega}W$ . If we multiply Equation (21) on the left by  $\bar{\Omega}W$ , we can write:

$$\begin{aligned} \lim_{\xi \rightarrow \infty} \bar{\Omega}W \frac{d\bar{W}}{dt} &= - \lim_{\xi \rightarrow \infty} \frac{1}{\xi\beta} (\bar{\Omega}W) \left(\frac{1}{\xi}I + \bar{\Omega}W\right)^{-1} (\bar{\Omega}W) W^{-1} \bar{S} \\ &= - \frac{1}{\xi\beta} \bar{\Omega}W W^{-1} \bar{S} = - \frac{1}{\xi\beta} \bar{\Omega} \bar{S}. \end{aligned} \quad (26)$$

From the equation above we derive a conservation law by integration over time, in the limit  $\xi \gg 1$ :

$$\bar{\Omega} \left[ \frac{\xi\beta}{2} (\bar{W}^2(t) - \bar{W}^2(t_0)) + \int_{t_0}^t \bar{S}(\tilde{t}) d\tilde{t} \right] = 0, \quad (27)$$

where  $\bar{W}^2$  means the vector with all the elements squared. In general, it is easy to see that this equation is true up to an arbitrary vector  $\vec{k}(t)$ , obtaining

$$\frac{\xi\beta}{2} (\bar{W}^2(t) - \bar{W}^2(t_0)) + \int_{t_0}^t \bar{S}(\tilde{t}) d\tilde{t} + (I - \bar{\Omega})\vec{k}(t) = 0, \quad (28)$$

which is the final conservation law in this limit, similar to what observed in [38]. Equation (28) is true only in the approximation in which the dynamics lies in the bulk, i.e.  $0 < W_i(t) < 1$ . Also, we note that in order to derive this conservation law, we needed to introduce  $W^{-1}$ , which is not invertible if some memristors are in the state  $W_i = 0$ , but the final result is independent from the inverse.

### 3. General properties of the dynamics

#### 3.1. Symmetries and dualities

Equation (9) satisfies several symmetries which we would like here to describe in detail. Let us first start by saying that the dynamical equations obtained depend on the choice of a spanning tree to begin with: the operator  $\bar{\Omega}$  should in fact be more correctly written as  $\bar{\Omega}_{\mathcal{T}}$  to be precise. The results we obtain do not depend on the choice of the tree  $\mathcal{T}$ , but the equations do. This is an example of a gauge degree of freedom. In addition, the equations depend on the choice of a direction of the currents on the circuit. However, these can be reabsorbed in the matrix  $\mathcal{J}$  introduced before, as these are simply the signs associated to the current. A general  $\mathcal{Z}_2$  symmetry is however easier to see for a global change of current signs: under a change of current direction,  $A \rightarrow -A$ . We note that  $\bar{\Omega}$  is independent from this transformation, as it depends on an even number of matrices  $A$ . Another symmetry of the dynamics is given by a change of active components to passive components and viceversa. Formally, this implies  $P \rightarrow -P$ : again, since  $P$  appears twice in  $\bar{\Omega}$ , the dynamics is unchanged.

Another symmetry to be expected is the transformation  $\bar{S} \rightarrow -\bar{S}$  and  $t \rightarrow -t$ , which reverses voltages and time, this however for the specific case of  $\alpha = 0$ .

Let us now consider a linear transformation of the  $w_i(t)$  involved, i.e.  $\vec{W}'(t) = O^{-1} \vec{W}(t)$ , where  $O$  is an invertible matrix. In this case,  $W(t) = \text{diag}(\vec{W}(t)) \rightarrow W'(t) = O^{-1} \text{diag}(W(t)) O$ . In order to

see this, let us look at how the equation transforms under a change of basis for  $W$ . We first note that  $\frac{d\bar{W}'}{dt} = O^{-1} \frac{d\bar{W}}{dt}$ . Thus:

$$\begin{aligned} \frac{d\bar{W}'}{dt} &= O^{-1} \frac{d\bar{W}}{dt} \\ &= \alpha O^{-1} \bar{W} - \frac{1}{\beta} O^{-1} (I + \xi \bar{\Omega} W)^{-1} O O^{-1} \bar{\Omega} O O^{-1} \bar{S} \\ &= \alpha \bar{W}' - \frac{1}{\beta} O^{-1} (I + \xi \bar{\Omega} O O^{-1} W)^{-1} O O^{-1} \bar{\Omega} O O^{-1} \bar{S} \\ &= \alpha \bar{W}' - \frac{1}{\beta} (I + \xi \bar{\Omega}' (O^{-1} W O))^{-1} \bar{\Omega}' \bar{S}', \end{aligned} \quad (29)$$

where we defined  $\bar{\Omega}' = O^{-1} \bar{\Omega} O$  and  $\bar{S}' = O^{-1} \bar{S}$ . This shows for instance that if we choose a basis in which  $W = \text{diag}(\bar{W})$ , then  $\bar{\Omega}$  will not be diagonal. If on the other hand we choose a basis in which  $\bar{\Omega}$  is diagonal, as a result  $W$  will likely not be diagonal, unless  $\bar{\Omega}$  and  $W$  commute. One thing that needs to be stressed is that  $O^{-1} W O \neq \text{diag}(\bar{W}') = W'(t)$ . If however  $O$  is a matrix which represents a specific permutation of the memristive element labels, then  $W'(t)$  will still be diagonal with the elements on the diagonal permuted accordingly.

One feature which becomes clear in Equation (9), is the fact that not all components of the source vector affect the evolution of the internal memory. In fact, we could add an arbitrary vector  $\Delta S = (I - \bar{\Omega})\bar{k}$  to  $\bar{S}$ : since  $\bar{S}$  enters the equation as  $\bar{\Omega}\bar{S}$ , one automatically has that

$$\bar{\Omega}(\bar{S} + \Delta S) = \bar{\Omega}\bar{S} + \bar{\Omega}(I - \bar{\Omega})\bar{k} = \bar{\Omega}\bar{S}. \quad (30)$$

This is a result which is reminiscent of the network constraints, or alternatively interpreted as a gauge freedom. Thus, we can easily decompose  $\bar{S}$  using the identity  $I = \bar{\Omega} + (I - \bar{\Omega})$ , as

$$\bar{S} = I\bar{S} = \bar{\Omega}\bar{S} + (I - \bar{\Omega})\bar{S}, \quad (31)$$

where the second term on the right hand side is in the orthogonal subspace respect to  $\bar{\Omega}$ , as  $\bar{\Omega}(I - \bar{\Omega}) = 0$ . Since in Equation (4) we have that the vector  $\bar{S}$  has the matrix  $\bar{\Omega}$  applied on the left, the components of  $\bar{S}$  orthogonal to  $\bar{\Omega}$  do not contribute to the dynamics. This result is important in light of the fact that the applicability of memristive circuits to, for instance, machine learning, depends on the ability to control the dynamics by means of external voltages.

### 3.2. Passive/active components and oscillations

In this section we wish to show that oscillations around an asymptotic fixed point can be present only when  $\bar{\Omega}$  is not symmetric. By construction, this happens when  $P \neq \pm I$ , being  $\bar{\Omega} = A^t(\bar{A}A^t)^{-1}\bar{A}$ . Physically, this represents the case in which only a mixture of active and passive components are present in the circuit (which are in our formalism represented by positive and negative resistances).

First we work out a simple exercise which will turn out to be useful later. Specifically, this will be in the case  $P = \pm I$ , for which  $\bar{\Omega} \rightarrow \Omega$ . One key element of the proof which follows below is the analysis of matrix similarity, for which a matrix has similar eigenvalues to another matrix. For instance, although  $\Omega W$  is not a symmetric matrix, it has always real eigenvalues. In order to see this, we note that the eigenvalues of any matrix product  $\bar{\Omega}W$  has the same eigenvalues of the matrix  $W^{\frac{1}{2}}\bar{\Omega}W^{\frac{1}{2}}$ . In this case, since  $W$  is diagonal and positive, the square root of the matrix is simply the square root of the diagonal elements. This is due to the fact that any matrix  $QMQ^{-1}$ , for any invertible matrix  $Q$ , has the same eigenvalues as those of  $M$ . In the following, we use the symbol  $\sim$  for similarity, i.e. matrices with similar eigenvalues ( $M \sim QMQ^{-1}$ ). If  $\bar{\Omega}$  is symmetric and real, then  $W^{\frac{1}{2}}\bar{\Omega}W^{\frac{1}{2}}$  has real eigenvalues as it is a

symmetric matrix. This implies that also  $(I + \xi \bar{\Omega} W)^{-1}$  has real eigenvalues. In fact, we have that

$$(I + \xi \bar{\Omega} W)^{-1} \sim W^{\frac{1}{2}} (I + \xi \bar{\Omega} W)^{-1} W^{-\frac{1}{2}} = (I + \xi W^{\frac{1}{2}} \bar{\Omega} W^{\frac{1}{2}})^{-1}.$$

Since the inverse of a symmetric matrix is symmetric, again its eigenvalues must be real. Also, since  $\bar{\Omega}$  is invariant under the transformation  $P \rightarrow -P$ , such analysis applies also for the inverse system, in which the number of passive and active component has been exchanged.

On the other hand, this is *not* true if  $\bar{\Omega}$  is not symmetric, and thus  $\bar{\Omega} W$  can have pairs of complex eigenvalues. Let us now assume that the spectrum of  $\bar{\Omega} W$  is  $(\lambda_1, \dots, \lambda_n, \lambda_{n+1}, \lambda_{n+1}^*, \dots, \lambda_{n+k}, \lambda_{n+k}^*)$ . Then, the spectrum of  $R = (I + \xi \bar{\Omega} W)^{-1}$  will be of the form

$$\sigma(R) = \left( (1 + \xi \lambda_1)^{-1}, \dots, (1 + \xi \lambda_n)^{-1}, \frac{1 + \xi \lambda_{n+1}^*}{(1 + \xi \lambda_{n+1}^*)(1 + \xi \lambda_{n+1})}, \frac{1 + \xi \lambda_{n+1}}{(1 + \xi \lambda_{n+1}^*)(1 + \xi \lambda_{n+1})}, \dots \right) \quad (32)$$

and thus still possibly contains pairs of complex eigenvalues. This observation is key to show that in the case in which  $\bar{\Omega}$  is symmetric (i.e. no mixture of active/passive components) there cannot be oscillations around fixed points.

In order to see this, let us now consider the linearized dynamics close to a fixed point  $W^*$ ,

$$\frac{d}{dt} \bar{W} \approx L|_{\bar{W}^*} \bar{W}, \quad (33)$$

where  $L|_{\bar{W}^*}$  is the Lyapunov matrix, given by

$$\begin{aligned} L_{ji} &= \partial_{w_i} f_j(\bar{W})|_{\bar{W}^*} \\ &= \partial_{w_i} \sum_k \left( (I + \xi \bar{\Omega} W)_{jk}^{-1} (\bar{\Omega} \bar{S})_k \right) |_{\bar{W}^*} \\ &= \partial_{w_i} \sum_k \left( \partial_{w_i} (I + \xi \bar{\Omega} W)_{jk}^{-1} \right) (\bar{\Omega} \bar{S})_k |_{\bar{W}^*}. \end{aligned} \quad (34)$$

In the calculations which follow we will use the formula  $\partial_{w_i} W_{jk} = \partial_{w_i} (w_j \delta_{jk}) = \delta_{ij} \delta_{jk}$ , and the formula  $\partial_s A^{-1} = -A^{-1} (\partial_s A) A^{-1}$  for any scalar quantity  $s$ . We thus have:

$$\begin{aligned} L_{ji}(W^*) &= -\xi \sum_{k_1, k_2, k_3, k_4} (I + \xi \bar{\Omega} W)_{jk_1}^{-1} \bar{\Omega}_{k_1 k_2} (\partial_{w_i} W_{k_2 k_3}) (I + \xi \bar{\Omega} W)_{k_3 k_4}^{-1} (\bar{\Omega} \bar{S})_{k_4} |_{\bar{W}^*} \\ &= -\xi \sum_{k_1, k_2, k_3, k_4} (I + \xi \bar{\Omega} W)_{jk_1}^{-1} \bar{\Omega}_{k_1 k_2} \delta_{ik_2} \delta_{k_2 k_3} (I + \xi \bar{\Omega} W)_{k_3 k_4}^{-1} (\bar{\Omega} \bar{S})_{k_4} |_{\bar{W}^*} \\ &= -\xi \left( (I + \xi \bar{\Omega} W^*)^{-1} \bar{\Omega} \right)_{ji} \sum_k \left( (I + \xi \bar{\Omega} W^*)^{-1} \bar{\Omega} \right)_{ik} \bar{S}_k, \end{aligned} \quad (35)$$

which is a rather complicated expression. In the first line of Equation (35) we have used the derivative for the inverse, in the second the identity for the derivative of the diagonal matrix  $W$ , and in the third line we have simply summed over the indices and renamed the remaining indices. We wish to understand now what are the conditions for which the matrix  $L$  above has only real eigenvalues (in which case no oscillations occur). This task can be achieved by showing that the matrix of Equation (35) is similar to a real symmetric operator. First we note that also the matrix  $L$  is of the form:

$$L = MD \quad (36)$$

where  $M$  is a full matrix and  $D$  is non-zero only on the diagonal. We do not consider any restriction on the elements of  $D$ : these can either be positive or negative for the time being. We assume however

that the fixed points  $W^*$  are such that  $w_i^* \neq 0$ . The diagonal elements of  $D$  are the vector elements  $(I + \xi \bar{\Omega} W^*)^{-1} \bar{\Omega} \bar{S}$  and are real, meanwhile  $M = (I + \xi \bar{\Omega} W^*)^{-1} \bar{\Omega}$ . First, we write:

$$\begin{aligned} L &= W^{*-1/2} W^{*1/2} (I + \xi \bar{\Omega} W^*)^{-1} W^{*-1/2} W^{*1/2} \bar{\Omega} W^{1/2} W^{-1/2} D \\ &= W^{*-1/2} (I + \xi W^{*1/2} \bar{\Omega} W^{*1/2})^{-1} W^{*1/2} \bar{\Omega} W^{*1/2} W^{*-1/2} D \\ &= W^{*-1/2} (I + \xi X)^{-1} X W^{*-1/2} D = MD \sim \sqrt{DM} \sqrt{D} \end{aligned} \quad (37)$$

In the first line we have used the identity  $I = W^{*-1/2} W^{*1/2}$ . The square root matrix exists for positively defined matrices, such as the diagonal matrix  $W^*$  which, by construction, is invertible. In the second line we have used the identity  $W^{*1/2} (I + \xi \bar{\Omega} W^*)^{-1} W^{*-1/2} = (I + \xi W^{*1/2} \bar{\Omega} W^{*1/2})^{-1}$ . In the third line we have implicitly introduced the definition  $X = W^{*1/2} \bar{\Omega} W^{*1/2}$ . We have also implicitly defined the matrix  $M = W^{*-1/2} (I + \xi X)^{-1} X W^{*-1/2}$ .

We now observe that  $(I + \xi X)^{-1} X$  is symmetric if and only if also  $X$  is symmetric. In order to see this, we use the identity  $((I + \xi X)^{-1} X)^t = X^t (I + \xi X^t)^{-1}$ ; we note that the identity  $((I + \xi X)^{-1} X)^t = X (I + \xi X)^{-1}$  holds due to the fact that  $X$  commutes with itself. Thus, the matrix  $M$  is symmetric if and only if  $X$  is symmetric. In general, we have that  $MD$  has the same eigenvalues of  $\sqrt{DM} \sqrt{D}$ . This implies that, if  $M$  is symmetric, the eigenvalues of  $MD$  are real if and only if  $\sqrt{D}$  is real. We can thus state that, given the fixed points of the dynamics  $W^*(\bar{S})$ , there are no oscillations if the following two conditions are satisfied:

- (1)  $X$  (and thus  $\bar{\Omega}$ ) is symmetric and
- (2)  $D_{ii} = \sum_k \left( (I + \xi \bar{\Omega} W^*(\bar{S}))^{-1} \bar{\Omega} \right)_{ik} \bar{S}_k > 0 \forall i$ .

All these facts put together show that  $L$  is similar to a real symmetric operator (and thus with real eigenvalues) if and only if the conditions above are satisfied. Let us note that in the proof we have not chosen a specific fixed point  $W^*$ , although the condition (2) depends implicitly only the source vector  $\bar{S}$  and the topology of the circuit. Since  $\bar{\Omega}$  is symmetric only if  $P = \pm I$ , this proves a weak result, e.g. the fact that is a constraint on the presence of oscillations around fixed points. In order to make this result stronger, we need to specify whether  $D_{ij}$  are effectively non-negative at the fixed point. In order to do that, we need to use the equations of motion and the fixed points structure. This is given by:

$$0 = \frac{d}{dt} \bar{W}(t) = \alpha \bar{W}(t) - \frac{1}{\beta} (I + \xi \bar{\Omega} W)^{-1} \bar{\Omega} \bar{S} \quad (38)$$

from which we obtain the fact that at the fixed point:

$$\alpha \bar{W}^* = \frac{1}{\beta} (I + \xi \bar{\Omega} W^*)^{-1} \bar{\Omega} \bar{S}. \quad (39)$$

We now notice that on the right hand side we have a vector which is the definition of the diagonal elements of  $D_{ij}$ . We can thus now confidently say that  $D_{ij} = \alpha \beta W^*$ . Since  $\alpha > 0$  and  $\beta > 0$ , and  $1 > W^* > 0$  by construction, this completes the proof that we had anticipated. Analogously to the *Barkhausen criterion* for circuits with feedback loops [39], this condition is necessary but not sufficient, i.e. there could be memristive circuits with passive/active elements mixtures which do not have oscillations.

#### 4. Dynamics as a constrained gradient descent optimization

In the previous sections we have studied the properties of the dynamics of purely memristive circuits. In this section we aim to look at the dynamics from another angle, which is spirit close to self-organizing maps as described long ago by Kohonen [40]. A precise statement which connects memristors to an

optimization problem will be made in what follows. We first consider one specific case as a warm up: the mean field problem in which we use the ansatz  $\vec{W} = w(t)\vec{1}$ , and for the case of only passive (or active) components. In this case, the vector

$$-\frac{1}{\beta} (I + \xi w(t)\Omega)^{-1} \Omega \vec{S}, \quad (40)$$

which appears on the right hand side of the differential equation in (4), can be interpreted as a tentative linear regression. Let us in fact assume that we aim to solve the equation:

$$\Omega \vec{r}(t) = \vec{S}(t) \quad (41)$$

given a specific regularizer [41] for the variable  $\vec{r}(t)$ . Such (ill-posed) equation can be solved by means of a Tychonov regularisation:

$$\min_{\vec{r}(t)} \|\Omega \vec{r}(t) - \vec{S}(t)\|^2 + \left\| \frac{1}{\sqrt{\xi w(t)}} \vec{r}(t) \right\|^2 \quad (42)$$

where the norm  $\|\cdot\|^2$  is the standard  $L_2$  vector norm, which gives

$$\vec{r}^* = \left( \frac{1}{\xi w(t)} I + \Omega^t \Omega \right)^{-1} \Omega \vec{S}(t). \quad (43)$$

Since  $\Omega$  is symmetric and a projector, one has

$$\vec{r}^* = \left( \frac{1}{\xi w(t)} I + \Omega \right)^{-1} \Omega \vec{S}(t), \quad (44)$$

which is proportional to the vector of Equation (40). The factor  $\vec{r}^*$  is interpreted as the time derivative of the internal state  $w(t)$  in the homogeneous case. The internal state thus moves in the direction which minimizes the least square above. This result hints toward the fact that memristive systems are performing a specific type of optimization. However, in the general case things are slightly more complicated and at the moment we do not have a complete answer of what type of optimization these systems are performing. Notwithstanding these difficulties, there is something we can say in weak non-linear regime,  $\xi \ll 1$ . We want to interpret Equation (4) as a specific dynamics which is of interest to machine learning, and in general to optimization problems. For simplicity, we consider the case  $P = I$  and  $\alpha = 0$ . Specifically, let us consider the following time-discretized dynamics:

$$\vec{W}(t+1) = \vec{W}(t) - \frac{dt}{\beta} (I + \xi \Omega W(t))^{-1} \Omega \vec{S}(t) \quad (45)$$

which, in the approximation  $\xi \ll 1$  can be written as

$$\begin{aligned} \vec{W}(t+1) &\approx \vec{W}(t) - \frac{dt}{\beta} (I - \xi \Omega W(t)) \Omega \vec{S}(t) \\ &= \vec{W}(t) - \frac{dt}{\beta} (I - \xi \Omega) \text{diag}(\Omega \vec{S}(t)) \vec{W}(t). \end{aligned}$$

We now use the fact that if  $\Omega$  is projector, then one can use the identity

$$\Omega \text{diag}(\Omega \vec{S}(t)) = \text{diag}(\Omega \vec{S}(t)), \quad (46)$$

from which we can derive:

$$\vec{W}(t+1) \approx \vec{W}(t) + \frac{dt(\xi-1)}{\beta} \Omega \text{diag}(\Omega \vec{S}(t)) \vec{W}(t). \quad (47)$$

The equation above can be rewritten as:

$$\vec{W}(t+1) \approx \vec{W}(t) + \mu \Omega \vec{\nabla}_{\vec{W}} f(\vec{W}), \quad (48)$$

where we defined  $\frac{dt(\xi-1)}{\beta} \equiv \mu$  and  $\vec{\nabla}_{\vec{W}} f(\vec{W}) = \text{diag}(\Omega \vec{S}(t)) \vec{W}(t)$ . Now that we have written the dynamical equation in this fashion, it is easy to realise that the dynamics is effectively a gradient descent procedure for a constrained optimization problem. We claim that such dynamics performs a constrained optimization of the type:

$$\text{minimize } f(\vec{W}) \text{ s.t.} \quad (49)$$

$$B\vec{W} = 0 \quad (50)$$

where  $f(W) = W_i^2$  and  $B$  is the directed incidence matrix. In order to see this, let us consider Rosen's gradient projection method to solve this optimization problem [42]. The basic assumption of the gradient projection method is that  $\vec{W}$  lies in the tangent subspace of the constraints. In order to provide an exact mapping, we consider first a general update rule given by:

$$\vec{W}_{t+1} = \vec{W}_t + \alpha \vec{z} \quad (51)$$

where both  $\vec{W}_{t+1}$  and  $\vec{W}_t$  satisfy are assumed to satisfy the linear constraint, and which depends on an arbitrary vector  $\vec{z}$ . We restrict our attention to the case in which the vector  $\vec{z}$  is in the steepest descent direction, and also satisfies  $B\vec{z} = 0$ . This condition ensures that if  $\vec{W}_0$  satisfies the linear constraint, then  $\vec{W}_t$  will also  $\forall t > 0$ . To be clear, the goal is to show the equivalence between the discrete dynamics of Equation (47) and the following optimization procedure:

$$\text{minimize } \vec{z} \cdot \vec{\nabla}_{\vec{W}} f(\vec{W}) \quad (52)$$

$$\text{s.t. } B\vec{z} = 0 \text{ and} \quad (53)$$

$$\vec{z} \cdot \vec{z} = 1. \quad (54)$$

where  $\vec{\nabla}_{\vec{W}} f(\vec{W}) = (\partial_{W_1} f(\vec{W}), \dots, \partial_{W_M} f(\vec{W}))$ . We now follow the procedure introduced by Rosen [42].

We introduce the Lagrange multipliers  $\vec{\lambda}$  and  $\mu$ , and the Lagrangian:

$$\mathcal{L}(\vec{z}, \vec{\lambda}, \mu) = \vec{z} \cdot \vec{\nabla}_{\vec{W}} f(\vec{W}) - \vec{s} \cdot B\vec{\lambda} - \mu(\vec{z} \cdot \vec{z} - 1). \quad (55)$$

The Euler–Lagrange equations for  $\vec{s}$  are given by:

$$\partial_{\vec{z}} \mathcal{L} = \vec{\nabla}_{\vec{W}} f(\vec{W}) - B\vec{\lambda} - 2\mu\vec{z} = 0. \quad (56)$$

If we multiply this equation by  $B^t$  on the left hand side, we obtain the equation:

$$B^t \vec{\nabla}_{\vec{W}} f(\vec{W}) - B^t B\vec{\lambda} = 0, \quad (57)$$

from which we can invert for the Lagrange multiplier  $\vec{\lambda}$ :

$$\vec{\lambda} = (B^t B)^{-1} B^t \vec{\nabla}_{\vec{W}} f(\vec{W}), \quad (58)$$

and thus using Equation (56) we finally obtain:

$$\vec{z} = \frac{1}{2\mu} (I - B(B^t B)^{-1} B^t) \vec{\nabla}_{\vec{W}} f(\vec{W}). \quad (59)$$

Such vector can be now re-inserted into the dynamical equation, which is now interpreted as a constrained gradient descent:

$$\vec{W}_{t+1} = \vec{W}_t + \frac{1}{2\mu} (I - B(B^t B)^{-1} B^t) \vec{\nabla}_{\vec{W}} f(\vec{W}) \equiv \vec{W}_t + \frac{1}{2\mu} \Omega \vec{\nabla}_{\vec{W}} f(\vec{W}). \quad (60)$$

It is easy at this point to identify, a posteriori, every element in this equation. The projector operator is given by  $\Omega$ . In the case in which only active or only passive element are present, we can use the duality between the loop matrix  $A$  and the incidence matrix  $B$ , to write  $\Omega = I - B(B^t B)^{-1} B^t$ . Thus, the constraint  $B\vec{W} = 0$  can be interpreted exactly as the conservation of memory in the circuit, and  $B$  promptly identified as the incidence matrix. The constant  $\frac{1}{2\mu} \equiv \frac{dt(1-\xi)}{\beta}$  is also obtained, and all it is left to us to do is to identify  $(\vec{\nabla}_{\vec{W}} f(\vec{W}))_i \equiv \sum_j \Omega_{ij} S_j W_i$ , from which after a simple integration we obtain the functional  $f(\vec{W}) = \sum_{ij} \frac{\Omega_{ij}}{2} S_j W_i^2$ . This interpretation is key to identify memristive networks as 'learning': gradient descent is in fact one of main training algorithms in machine learning and optimization, and in particular in neural networks. Such connection establishes memristive circuits as the perfect neuromorphic devices. Of course, this is not the first time this was suggested [43], but here we have provided further evidence of the above. For instance, in [44] it was shown that in the case of a memristor series one can use the equations for learning. Using the fact that in the case of a series  $\Omega = \frac{1}{N} \vec{u} \vec{u}^t$ , and that  $\sum_j u_j = N$ , we can show that the functional being minimized was  $f(\vec{W}) = \frac{1}{2N} \sum_{ij} S_i W_j^2$ .

This is interesting also for other reasons. First of all, it makes precise the notion of information overhead for the specific case of purely memristive systems. In a recent paper, Di Ventura and Traversa [4] put forward the suggestion that the graph topology is part of the optimization process. In the above we have just observed that the function being optimised is  $f(\vec{W}) = \sum_{ij} \frac{\Omega_{ij}}{2} S_j W_i^2$  in which both the external voltage sources *and* the network topology (through  $\Omega$ ) appears.

For the technological application of such statement, this poses the problem of engineering  $\Omega$  and choose  $\vec{S}$  in order to minimize the function of which one desires to find a minimum.

## 5. Conclusions

In the present paper we have made several steps towards understanding the collective behaviour of circuits with memory. We used a recently derived equation for the internal memory of an ideal and purely memristive system. Memristors, and in particular memristive circuits, are interesting devices with a very rich dynamical behaviour. Even for the simpler memristor type (linear), non-linear phenomena emerge at the dynamical level. In fact, such an equation establishes that the non-linearity is controlled by a single parameter, which is the ratio between the resistance in the insulating phase and the resistance in the conducting phase of the memristor. Here we focused on the technical aspects and properties of the derived equation and tried to create a link between the dynamics of the internal memory to a more standard machine learning approach. Specifically, we have described in detail the symmetries of Equation (4), and analysed the difference between purely passive (or active) systems and their mixtures. We have proven a condition on the existence of complex eigenvalues of the Lyapunov matrix. This was done by showing that the matrix which describes the dynamics close to a fixed point is similar to a real symmetric operator. We have shown that only for mixture of passive and active components one can have complex Lyapunov exponents.

We have also discussed what type of optimization purely memristive systems are performing using the internal memory equation. As it turns out, in the limit of weak non-linearity, the dynamical equation

can be casted in the form of a (linear) constrained gradient descent equation [42]. The functional being minimized was found to be a combination of sources and quadratic in the memristor internal parameters, weighted by the projector operator on the space of cycles. More complex optimizations require the introduction of other circuital elements [4] which we did not consider in the present paper.

## Notes

1. For the interested reader, we note that in the language of *discrete cohomology* [45], the incidence matrix represents the *boundary* operator  $\partial$ . Such representation exists for any oriented graph.
2. Such equation can be generalised to the case in which the internal memory of each memristor changes linearly in the voltage, rather than the current. The result was obtained in [20] and reads:

$$\frac{d\bar{W}}{dt} = \alpha\bar{W} - \frac{1}{\beta} (I + \xi W) (I + \xi \bar{\Omega} W)^{-1} \bar{\Omega} \bar{S}. \quad (61)$$

## Acknowledgements

We would like to thank Fabio Lorenzo Traversa and Massimiliano Di Ventra for various discussions concerning memristors. I would also like to thank the anonymous editors of a previous paper for saying that such equations were not useful. I would like to thank the anonymous referees and J.P. Carbajal for several comments which improved both the readability of this paper and spotting few incorrect equations.

## Disclosure statement

No potential conflict of interest was reported by the author.

## References

- [1] Caravelli F, Hama A, Di Ventra M. Scale-free networks as an epiphenomenon of memory. *EPL*. 2015;109(2):28006.
- [2] Caravelli F. Trajectories entropy in dynamical graphs with memory. *Front Robot AI*. 2016;3(18). arXiv:1511.07135.
- [3] Dorigo M, Gambardella LM. Ant colonies for the traveling salesman problem. *Biosystems*. 1997;43(2):73–81.
- [4] Traversa FL, Di Ventra M. Universal memcomputing machines. *IEEE Trans Neural Netw Learn Syst*. 2015;26(11):2702–2715. DOI: 10.1109/TNNLS.2015.2391182. preprint arXiv:1405.0931.
- [5] Traversa FL, Ramella C, Bonani F, et al. Memcomputing NP-complete problems in polynomial time using polynomial resources and collective states. *Sci Adv*. 2015;1(6):1500031.
- [6] Traversa FL, Di Ventra M. Polynomial-time solution of prime factorization and NP-hard problems with digital memcomputing machines. 2017; arXiv:1512.05064. DOI:10.1063/1.4975761
- [7] Pershin YV, Di Ventra M. Memory effects in complex materials and nanoscale systems. *Adv Phys*. 2011;60:145–227.
- [8] Indiveri G, Liu S-C. Memory and information processing in neuromorphic systems. *Proc IEEE*. 2015;103(8):1379–1397.
- [9] Traversa FL, Pershin YV, Di Ventra M. Memory models of adaptive behaviour. *IEEE Trans Neural Netw Learn Syst*. 2013;24:1437–1448.
- [10] Avizienis AV, Sillin HO, Martin-Olmos C, et al. Neuromorphic atomic switch networks. *PLoS one*. 2012;7(8):e42772.
- [11] Stieg AZ, Avizienis AV, Sillin HO, et al. Emergent criticality in complex turing B-type atomic switch networks. *Adv Mater*. 2012;24:286–293.
- [12] Di Ventra M, Pershin YV. The parallel approach. *Nat Phys*. 2013;9:200–202.
- [13] Pershin YV, Di Ventra M. Solving mazes with memristors: a massively-parallel approach. *Phys. Rev. E*. 2011;84:046703.
- [14] Adamatzky A, Chen G. *Chaos, CNN, memristors and beyond*. Singapore: World Scientific; 2013.
- [15] Pershin YV, Di Ventra M. Self-organization and solution of shortest-path optimization problems with memristive networks. *Phys. Rev. E*. 2013;88:013305.
- [16] Pershin YV, La Fontaine S, Di Ventra M. Memristive model of amoeba learning. *Phys. Rev. E*. 2009;80:021926.
- [17] Chialvo DR. Emergent complex neural dynamics. *Nat Phys*. 2010;6:744–750.
- [18] Pershin YV, Di Ventra M. Experimental demonstration of associative memory with memristive neural networks. *Neural Netw*. 2010;23:881–886.
- [19] Adamatzky A. *Physarum machines*. London: World Scientific; 2010. (World Scientific Series on Nonlinear Science Series A; vol. 74).
- [20] Caravelli F, Traversa FL, Di Ventra M. The complex dynamics of memristive circuits: analytical results and universal slow relaxation. *Phys Rev E*. 2017;95. 022140. arXiv: 1608:08651.

- [21] Strukov DB, Snider G, Stewart DR, et al. The missing memristor found. *Nature*. 2008;453:80–83.
- [22] Strukov DB, Likharev KK. CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices. *Nanotechnology*. 2005;16:888–900.
- [23] Chua LO, Kang SM. Memristive devices and systems. *Proc IEEE*. 1976;64:209–223.
- [24] Di Ventra M, Pershin YV, Chua LO. Circuit elements with memory: Memristors, memcapacitors, and meminductors. *Proc IEEE*. 2009;97(10):1717–1724.
- [25] Yang JJ, Strukov DB, Stewart DR. Memristive devices for computing. *Nat Nanotechnol*. 2013;8:13–24.
- [26] Demina VA, Erokhin VV, Emelyanov AV, et al. Hardware elementary perceptron based on polyaniline memristive devices. *Org Electron*. 2015;25:16–20.
- [27] Ohno T, Hasegawa T, Nayak A, et al. Sensory and short-term memory formations observed in a Ag<sub>2</sub>S gap-type atomic switch. *Appl Phys Lett*. 2011;99:203108.
- [28] Wang Z, Joshi S, Savel'ev SE, et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nat Mater*. 2017;16:101–108.
- [29] Nilsson JW, Riedel SA. *Electric circuits*. 9th ed. Boston: Prentice Hall; 2011.
- [30] Bollobas B. *Modern graph theory*. New York (NY): Springer Science; 1998.
- [31] Wilson RC, Aziz F, Hancock ER. Eigenfunctions of the Edge-based Laplacian on a graph. 2015; arXiv:1302.3433.
- [32] Zelazo D, Mesbahi M. Edge agreement: graph-theoretic performance bounds and passivity analysis. *IEEE Trans Autom Control*. 2011;56(3):544–555.
- [33] Polettini M. Cycle/cocycle oblique projections on oriented graphs. *Lett Math Phys*. 2015;105:89–107.
- [34] Riaza R. Cyclic matrices of weighted digraphs. *Discrete Appl Math*. 2012;160:280–290.
- [35] Sherman J, Morrison JW. Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix. *Ann Math Stat*. 1949;20:621.
- [36] Pfeiffer P, Egusquiza IL, Di Ventra M, et al. Quantum memristors. *Sci Rep*. 2016;6. Article number: 29507.
- [37] Di Ventra M, Pershin YV. On the physical properties of memristive, memcapacitive and meminductive systems. *Nanotechnology*. 2013;24:255201.
- [38] Gale E. The memory-conservation theory of memristance. In: *Proceedings of the 16th UKSim-AMSS International Conference of Modelling and Simulation (UKSIM2014)*; 2014. p. 598–603.
- [39] MacFarlane AGJ. The development of frequency-response methods in automatic control. *IEEE Trans Autom Control*. 1979;24(2):250–265.
- [40] Kohonen T. *Self-organization and associative memory*. Berlin: Springer-Verlag; 1987.
- [41] Golub GH, Hansen PC, O'Leary DP. Tikhonov regularization and total least squares. *SIAM J Matrix Anal Appl*. 1999;21(1):185–194.
- [42] Rosen JB. The gradient projection method for nonlinear programming. Part I. linear constraints. *J Soc Ind Appl Math*. 1960 mar;8(1):181–217.
- [43] Adamatzky A, Chua L. *Memristor networks*. Springer ed. (Heidelberg); 2014.
- [44] Carbajal JP, Dambre J, Hermans M, et al. Memristors models for machine learning. *Neural Comput*. 2015;27(3):725–747.
- [45] Ziedler E. Electrical circuits as a paradigm in homology and cohomology. In: *Quantum field theory III: gauge theory*. Leipzig: Springer-Verlag; 2011. p. 1009–1026.