

## Perspective

## Memristive networks: From graph theory to statistical physics

A. ZEGARAC<sup>1,2,3</sup> and F. CARAVELLI<sup>4</sup><sup>1</sup> *ETH Zurich - 8092 Zurich, Switzerland*<sup>2</sup> *London Institute for Mathematical Sciences - 35a South Street, London W1K 2XF, UK*<sup>3</sup> *Invenia Labs - 27 Parkside Place, CB1 1JF Cambridge, UK*<sup>4</sup> *Theoretical Division (T4) and Center for Nonlinear Studies, Los Alamos National Laboratory Los Alamos, NM 87545, USA*received 30 November 2018; accepted 4 December 2018  
published online 15 January 2019PACS 05.45.-a – Nonlinear dynamics and chaos  
PACS 05.20.-y – Classical statistical mechanics  
PACS 07.50.Ek – Circuits and circuit components**Abstract** – We provide an introduction to a very specific toy model of memristive networks, for which an exact differential equation for the internal memory which contains the Kirchhoff laws is known. In particular, we highlight how the circuit topology enters the dynamics via an analysis of directed graph. We try to highlight in particular the connection between the asymptotic states of memristors and the Ising model, and the relation to the dynamics and statics of disordered systems.

perspective

Copyright © EPLA, 2019

**Introduction.** – This perspective paper is concerned with the open questions that we find interesting in the context of a seemingly simple toy model of memristive “endogeneous” dynamics. Some aspects, important for the understanding of the behavior of circuits of memristors, also apply to the equilibrium configuration of currents in a resistive network, as well as other systems in which Kirchhoff laws play an important role. By no means this paper is exhaustive, and for a broader overview of the topic of memristors we suggest the recent and less recent reviews [1–3]. Our aim is to emphasize two aspects of the dynamics of memristors which characterize the behavior of a circuit with memory: the rather non-trivial connection between the underlying circuit and the non-linear dynamics of these components, and the relationship between the physics of disordered systems and the dynamical asymptotic behavior of the circuit. In order to accomplish this task, we generalize the previously obtained equation to the case in which the disorder is present, and study the new equation.

In the late 2000s, researchers at Hewlett-Packard realized [4–6] that many transition metal dioxides had the properties, initially theorized by Chua [7,8] in the early 1970s, of possessing an internal memory and a hysteretic behavior; certain metal oxides such as those derived from tungsten or titanium have the interesting property that the resistance changes noticeably as a function of time.

The state of the resistance between two limiting values can be parametrized by a parameter  $w$ , which is constrained between 0 and 1. We will refer to this parameter as the *internal memory parameter*. For the case of titanium dioxide, the evolution of the resistance was described by the following two equations:

$$R(w) = R_{\text{on}}(1 - w) + wR_{\text{off}} \equiv R_{\text{on}}(1 + \xi w),$$

$$\frac{d}{dt}w(t) = \alpha w - \frac{R_{\text{on}}}{\beta}I(t),$$

initially studied for  $\alpha = 0$ , and where  $0 \leq w \leq 1$ ,  $\xi = \frac{R_{\text{off}} - R_{\text{on}}}{R_{\text{on}}}$ , and  $I(t)$  is the current flowing in the device at time  $t$ . If  $\alpha = 0$ , the second equation can be integrated and  $w$  shown to be directly related to the charge in the conductor. Albeit this model has been revisited several times, it still serves as a prototypical model of a memory-resistor: the memristor. Also, it became clear that memory is a very common feature of nanoscale components [9,10]. The interest in these components is due to the fact that memristors can serve for the purpose of neuromorphic computing [11]. For a more technology-oriented review of the subject, we suggest in particular the recent review [1].

Given the brevity of this article, we will focus on the stylized facts known for the vectorial differential equation, derived in [12]. The equation describes the time evolution

of the internal memory of a memristive circuit (circuit made of memristors only) for the case of homogeneous memristors ( $\alpha, \beta, R_{\text{on}}, R_{\text{off}}$  identical across the network):

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

where  $w_i(t)$  are the internal variables constrained between 0 and 1,  $W_{ij}(t) = w_i(t)\delta_{ij}$ ,  $I$  is the identity matrix,  $\Omega$  is a matrix which we will describe soon, and  $S_i(t)$  is a vector of the voltage sources. We can see that the constants  $\alpha, \beta$  and  $\xi$  control the decay and reinforcement time scales and the non-linearity in the equation, respectively. In principle, we could discuss the properties of this equation without referring to where it came from, but for a deeper understanding of its properties it is useful to understand its origins. We can think of the variables  $w_i$  as living on the edges of a graph, where  $\Omega_{ij}$  (which we anticipate to be a projector operator) contains the information about the topology of the graph. Each edge of the graph represents a resistive component.

There are several limits to the applicability of eq. (13). For instance, it has been derived only for ideal memristors, in which the time derivative of the internal parameter depends only linearly on the current, and is ideal: no parasitic capacitance or inductance are considered in the dynamics. Also, we consider only an endogenous dynamics, *e.g.*, there are no voltage or current generators in parallel to the system (circuit) under scrutiny. This implies that, for instance, the interesting dynamics of [13], in which it has been shown that memristors can be used to solve a maze, cannot be analyzed using the approach of this paper. Also, the possibility of solving the equation in full generality is an illusion, due to the constraints which make the differential equation discontinuous. What is the purpose then?

The advantages of using a toy model for analyzing a system which would be, otherwise, much more complicated are multiple. For instance, if the circuit is controlled with sinusoidal voltage and none of the memristor reaches the boundary, then the dynamics is continuous and a solution of the equation provides a solution for the evolution of each single memristor. A solution for small values of  $\xi$  and  $S$  when controlled with sinusoidal voltages has been provided in [12]. Also, when controlled with constant voltage, the dynamics of the circuit (as we will discuss below) is interesting enough to serve as a good toy model to the relaxation of more general circuits with memory. If, for instance, there are parasitic capacitance and inductance and the system is controlled with sinusoidal voltages, the dynamics of equation. (13) can still serve to analyze the dynamics for longer timescales (where now we replace  $R(t) \rightarrow Z(t)$ , the admittance). Also, how do the Kirchhoff constraints affect the dynamics? Are there any hidden symmetries? How do memristors interact in the short-term dynamics? Is there an emergent speed of light in the system? As we will see, these questions can be asked (and answered) with the toy model above.

For instance, the differential equation (13), written in this form, highlights some non-obvious symmetries of the dynamics of memristors. In fact, since we can always write  $\vec{S} = \Omega\vec{S} + (I - \Omega)\vec{S}$ , it is easy to see that we can add to  $\vec{S}$  any vector  $\vec{S}' = (I - \Omega)\vec{k}$ , which will not affect the dynamics. This form of freedom arises from the Kirchhoff constraints from which the differential equation has been derived. It is due to the Kirchhoff constraints that in principle the system could have long-range interactions; this is one of the formal arguments we cover in this paper.

Specifically, the first half of this paper focuses on the connection between  $\Omega$  and the graph (the circuit), while the second part on the properties of the differential equation which might be of interest to an audience of statistical physicists. Albeit most of the work discussed in this paper is not novel, there are some novel points of view that we wish to share along with recent numerical simulations.

**From graph theory...** – Let us first provide a simple explanation of the origin and applications of the matrix  $\Omega$ , as its use is not new and deeply connected to constrained flows, and thus resistive circuits. What we state below about the graph theoretical approach to memristors is true in fact for resistor networks as well. Let us consider a network of resistors connected *in series* to voltage generators; the graph  $G$  represents a circuit, and to each edge  $e_i$  of the graph we can associate a pair of variables  $(R_i, S_i)$ , where  $R_i$  is the resistance and  $S_i$  the voltage. If  $R_{ij} = R_j\delta_{ij}$  is the diagonal matrix of the resistances, then it is known that [12,14,15]

$$\vec{I} = -A^t(ARA^t)^{-1}A\vec{S}, \quad (2)$$

where the vector  $\vec{I}$  is the equilibrium configuration of the currents on each edge of the graph, and  $A$  is the cycle matrix of the graph, which we will define shortly. We note that in the case of unit resistances, we have

$$\vec{I} = -A^t(AA^t)^{-1}A\vec{S} = -\Omega\vec{S}, \quad (3)$$

where we can recognize  $A^t(AA^t)^{-1}A$  as a projector, which we denote by  $\Omega$  and which satisfies  $\Omega^2 = \Omega$ . It is not obvious to see this, but  $\Omega$  expresses the Kirchhoff laws [16] for the circuits. It is however interesting to note that  $\Omega$  is a generalization of the concept of effective resistance. Let us assume, for instance, that voltage is applied to only one edge in the network. We denote that edge by  $e_k$  and we label the vertices at its ends by  $v_1$  and  $v_2$ . Then we have that, since the current that flows into the resistor network must flow out, *i.e.*,  $\frac{S_k}{1+R_k} = I_k$ . It is easy to see, using eq. (3), that  $R_k$  is the definition of effective resistance between the nodes  $v_1$  and  $v_2$ , and thus we have

$$R_k = 1 + \frac{1}{\Omega_{kk}}.$$

In addition to this information,  $\Omega_{kl}$  will also contain the information about the current flowing through the resistance  $k$  in the network as we apply the voltage in series

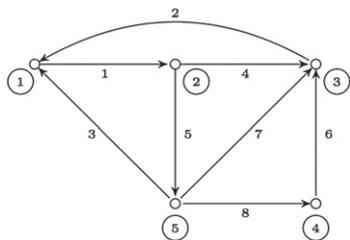


Fig. 1: A labelled directed planar graph.

on the resistance  $l$ . In order to be more precise about the properties of  $\Omega$ , we provide a quick graph theoretic introduction. A (directed) graph consists of two objects: vertices and edges. Vertices can be thought of as points and edges as lines that connect some of those points. We will label vertices as  $v_1, \dots, v_n$  and edges as  $e_1, \dots, e_m$ . Mathematically, we represent an edge starting at vertex  $v_i$  and ending at vertex  $v_j$  as an ordered pair  $(v_i, v_j)$ . We say that a graph is *planar* if it can be drawn in a plane without any of its edges intersecting.

*Incidence matrix:*  $\Omega_{BT}$ . Let  $G$  be a directed graph. One way of representing  $G$  is by specifying where each of its edges starts and where it ends. It is convenient to do this using a matrix. We call such a matrix an *incidence matrix*.

As an example, consider the graph labelled as in fig. 1.

Each column of its incidence matrix represents an edge: the first edge starts at vertex 1 and ends at vertex 2, so the first column of the matrix has entry 1 in the first row and entry  $-1$  in the second row. All the other entries in the first column are 0 because none of the other vertices are a part of that edge. By continuing this process for every edge, we get the incidence matrix  $B$  of the given graph:

$$B = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & 1 \end{pmatrix}. \quad (4)$$

More formally, if a graph  $G$  has  $v$  vertices and  $e$  edges, then the incidence matrix  $B$  of  $G$  is a  $v \times e$  matrix (*i.e.*, a matrix with  $v$  rows and  $e$  columns), whose entry  $(i, j)$  is defined as

$$B_{ij} := \begin{cases} 1, & \text{if } v_i \text{ is the initial vertex of the edge } e_j, \\ -1, & \text{if } v_i \text{ is the terminal vertex of the edge } e_j, \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

If  $B$  is an incidence matrix, we can define the projector operator  $\Omega_{BT}$ :

$$\Omega_{BT} = B^T (BB^T)^{-1} B, \quad \Omega_{BT}^2 = \Omega_{BT}. \quad (6)$$

If we try to compute  $\Omega_{BT}$  from the definition, we will find that the inverse  $(BB^T)^{-1}$  does not exist in general.

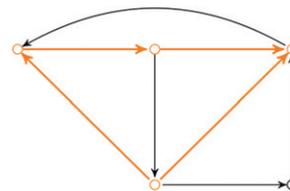


Fig. 2: An example of a cycle.

This can be solved by either considering the reduced incidence matrix (obtained by removing a row from the original incidence matrix) or by taking the pseudoinverse of the expression instead of the ‘‘regular’’ inverse.

*Cycle matrix:*  $\Omega$ . Before we can define the projector operator  $\Omega$ , we discuss a few more objects from graph theory. We define a *walk* on a directed graph  $G$  to be a sequence of vertices, say  $v_1, v_2, \dots, v_n$ , such that for every pair of consecutive vertices in the sequence there exists an edge that connects them, *i.e.*, for every  $i$  such that  $1 < i \leq n$  either  $(v_{i-1}, v_i) \in E(G)$  or  $(v_i, v_{i-1}) \in E(G)$ , where  $E(G)$  is the edge set of the graph  $G$ .

A *cycle* is a walk  $W = v_1 v_2 \dots v_n$  such that  $l \geq 3$ ,  $v_0 = v_n$  and the vertices  $v_i$ ,  $0 < i < n$  are distinct from each other and from  $v_0$ . An example of a cycle is shown in fig. 2.

The space spanned by the edges has a structure of a vector space. A *cycle space* is the subset of edge space that is spanned by all cycles of a graph. A graphical representation of this fact is shown in fig. 3. A cycle matrix  $A$  is a matrix whose columns form a basis of the cycle space. For example, if  $\{\vec{c}_1, \dots, \vec{c}_n\}$  is a set of column vectors that form a basis of the cycle space, then the cycle matrix is  $A = (\vec{c}_1, \dots, \vec{c}_n)$ . Finally, the *projector operator*  $\Omega$  on the cycle space of the graph is defined to be  $\Omega = A(A^T A)^{-1} A^T$ . The following useful identity connects the projector operator  $\Omega$  (based on the matrix  $A$ ) to the projector operator  $\Omega_{BT}$  (based on the incidence matrix  $B$ )  $\Omega = I - \Omega_{BT}$ .

*Locality.* One question which arises immediately is: given the fact that the Kirchhoff constraints introduce some sort of non-locality between resistors (at equilibrium) and memristors (out of equilibrium), it is worth mentioning a few results about the matrix  $\Omega$ . As we will see, this problem is connected to the graph embedding problem as well. Let us first focus on planar graphs. In [17] the following bound on locality of interactions for planar graphs was proved:

$$|\Omega_{i,j}| \leq e^{-zd(i,j)+\tilde{\rho}}. \quad (7)$$

For the purpose of this paper, we can think of  $z, \tilde{\rho}$  as constants and of  $d(i, j)$  as the distance between edges  $i$  and  $j$ . The full derivation can be found in [17]. Here we will focus on one of the key parts of the calculation and provide a no-go theorem for the generalization of the argument for arbitrary non-planar graphs. Finding an analytic expression for a quantity which involves an inverse of a potentially large matrix is a non-trivial problem. To overcome

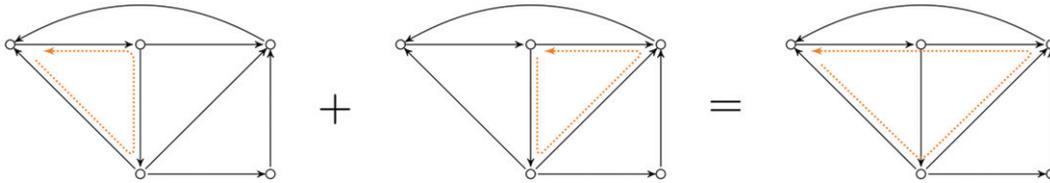


Fig. 3: Addition of cycles.

this in the case of  $\Omega = A(A^T A)^{-1} A^T$ , it was noticed that the expression for  $\Omega$  simplifies if the matrix  $A$  is orthonormalised first. If we denote by  $\tilde{A}$  the orthonormalised matrix, we get  $\Omega = \tilde{A} \tilde{A}^{-1}$ , which is a matrix product. The difficulty is now in calculating the orthonormalised matrix  $\tilde{A}$  from  $A$ . We proceed by using a non-algorithmic expression for the Gram-Schmidt process: the  $p$ -th column,  $A_p$ , of the orthonormalised matrix  $\tilde{A}$  is given by

$$\tilde{A}_p = \det \begin{pmatrix} \langle \vec{A}_1, \vec{A}_1 \rangle & \langle \vec{A}_1, \vec{A}_2 \rangle & \dots & \langle \vec{A}_1, \vec{A}_n \rangle & \vec{A}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \vec{A}_n, \vec{A}_1 \rangle & \langle \vec{A}_n, \vec{A}_2 \rangle & \dots & \langle \vec{A}_n, \vec{A}_n \rangle & \vec{A}_n \end{pmatrix}. \quad (8)$$

We will now show how the inner products  $\langle A_i, A_j \rangle$  can be expressed in terms of the adjacency matrix.

In the case of a grid graph  $G$ , we perform the following steps:

- First, we pick a basis of the cycle space of  $G$  as in fig. 4.
- Then we denote by  $G'$  the graph that has a vertex for each basis cycle of  $G$  and an edge between two vertices if the corresponding cycles in  $G$  are adjacent.
- Finally, we express the inner products as

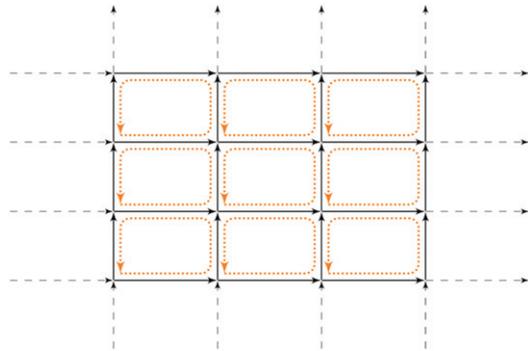
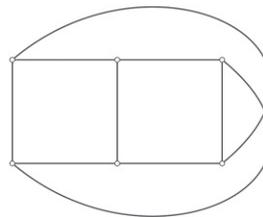
$$\langle A_i, A_j \rangle = \begin{cases} M_{i,j}, & \text{if } i = j, \\ |C_i|, & \text{if } i \neq j, \end{cases} \quad (9)$$

where  $M_{i,j}$  is the adjacency matrix of  $G'$  and  $|C_i|$  is the length of the cycle corresponding to the  $i$ -th vertex of  $G'$ .

The last step was key as it allowed manipulations which ultimately lead to the expression for the bound in [17].

Let us now focus on the case of non-planar graphs, where the problems arise. We found that the same method cannot be applied to the non-planar graphs, and here we provide an explanation of the reason. Physically, we would expect that for graphs that present a notion of distance (unlike, for instance, random graph, where the graph diameter scales as  $D \approx \log(N)$  in the number of nodes  $N$ ), a similar bound would apply. In the case of planar graphs, it was possible to choose a basis of the cycle space in such a way that the basis cycles bound the faces of the graph, greatly simplifying the proof.

A generalisation of this construction to non-planar graphs is not obvious due to the fact that in non-planar


Fig. 4: A choice of a basis on the grid graph  $G$ .

Fig. 5: An example of a non-planar graph,  $K_{3,3}$ .

graphs the notion of *faces* is not well defined. In an attempt to overcome this, we defined the faces of a non-planar graph  $G$  to be the faces of an embedding of  $G$  in some higher-dimensional closed orientable surface.

We illustrate this on the simplest non-planar graph,  $K_{3,3}$ , shown in fig. 5. Let  $T^2$  be a torus. An embedding of  $K_{3,3}$  in  $T^2$  is shown in fig. 6. We thus define the faces of  $K_{3,3}$  to be the connected components of  $T^2 \setminus \bar{K}_{3,3}$ , where  $\bar{K}_{3,3}$  denotes the embedding of  $K_{3,3}$  inside the torus  $T^2$ . To be able to use the method from [17], we have to find a basis of the cycle space consisting only of cycles that bound faces. However, the number of elements in the basis of a cycle space of  $K_{3,3}$  is 4. As can be seen in fig. 6 above, there are only 3 faces in the embedding of  $K_{3,3}$  in  $T^2$ . This is a problem which we elucidate further below. Setting aside the question of well-definedness of faces, it might seem that if we embed  $K_{3,3}$  in a different way or in some other closed orientable surface (for example a triple torus shown in fig. 7), we could get a sufficient number of faces. We prove below that there exists no embedding with sufficient number of faces.

Let us introduce the *Euler characteristic* of a graph  $G$  embedded in a surface  $S$  to be

$$\chi_{G,S} = |V| - |E| + |F|. \quad (10)$$

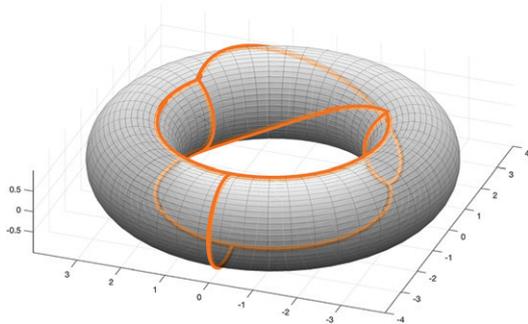
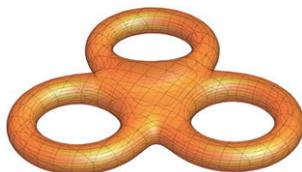

 Fig. 6: An embedding of  $K_{3,3}$  in a torus.


Fig. 7: A “triple torus”.

We want the number of faces to be greater than or equal to the dimension of the cycle space. Thus,

$$|F| \geq \dim \mathcal{C} = |E| - (|V| - 1), \quad (11)$$

where  $\mathcal{C}$  denotes the cycle space. Substituting eq. (11) into the expression for Euler characteristic given by eq. (10), we get

$$\begin{aligned} \chi_{G,S} &= |V| - |E| + |F| \\ &\geq |V| - |E| + (|E| - (|V| - 1)). \end{aligned}$$

That is, we need  $\chi_{G,S} \geq 1$ . A graph embedded in a surface has Euler characteristic equal to the Euler characteristic of that surface. We can also express the Euler characteristic of a surface in terms of its genus  $g$  as  $\chi_{G,S} = 2 - 2g$ . Hence, using the previously obtained inequality we get  $g \leq 1/2$ . That is, we can only find the sufficient number of cycles that bound faces in surfaces with less than  $1/2$  holes. The only such closed orientable surface is a sphere (with 0 holes), and a sphere is equivalent to a plane for all our purposes. Therefore, we cannot use the same approach as in the planar case. This leaves us with the open question of how to generalize the bound on the interaction strength between memristors to more general non-planar graphs.

Given that the circuit enters the equation only in  $\Omega$  and that it represents the Kirchhoff constraints, its study is useful. The locality bound is useful for bounding also the long-term dynamics of the system. As shown, for instance, in [17], these locality bounds imply an emergent speed of light in the system, similarly to the Lieb-Robinson bounds for quantum systems [18,19]:

$$|\langle w_i(t)w_j(0) \rangle| \leq K e^{-(d_{ij} - v_{\text{eff}}t)}, \quad (12)$$

where  $d_{ij}$  is the Hamming distance between the memristors and  $v_{\text{eff}}$  an effective speed of light. It is thus

interesting to study these bounds for more general circuits that are not necessarily planar.

**... to statistical physics.** Insofar we have focused on a network of resistors via the study of the matrix  $\Omega$ . An initial attempt at studying the statistical properties of dynamical graphs with memory was the one of [20], where the emergence of scale-free networks out of the endogeneous dynamics of excitable memristor-like components was observed. What about a network of memristors which satisfy Kirchhoff laws? In this case the non-linearity of the differential equation makes the analysis more complicated. The differential equation becomes vectorial as in eq. (13), where the variables  $\vec{w}$  are constrained on the hypercube  $[0, 1]^M$  with  $M$  being the number of memristors. It has been noted in [12] that for “complex enough” circuits, the average parameter  $\langle w \rangle = \frac{1}{N} \sum_i w_i$  relaxes slowly to the asymptotic values  $w = 1$  and  $w = 0$ .

The operator  $\Omega'$  plays an important role in the differential equation (1), as it is the only place where the graph topology enters. In the regime of large (random) graphs, slow relaxation can be observed in the average parameter  $\langle w \rangle$ . A power law type of relaxation had been observed in the relaxation of  $\langle w \rangle$  to the asymptotic values of  $w = 1$  and  $w = 0$  ( $t^{-\rho}$  with  $\rho \approx 0.92$ ). This feature is similar to what was observed experimentally in [21,22] for atomic switch networks. What seems to be a good parameter for the transition from fast to slow relaxation is the ratio between the number of (fundamental) circuit loops and memristors components, which is upper bounded by one.

In the case with disorder, *i.e.*, when not all memristors have similar properties, it can be shown that a generalized differential equation for the memories still exists and is given by

$$\frac{d}{dt} \vec{w}(t) = A \vec{w}(t) - B^{-1} (I + \Omega' W(t))^{-1} \Omega' T^{-1} \vec{S}(t), \quad (13)$$

where now  $A_{ij} = \alpha_i \delta_{ij}$ ,  $B_{ij} = \beta_i \delta_{ij}$  and where  $T = \delta_{ij}(1 + N_{ii})$  represents the *disorder* among the component. The matrix  $T$  characterizes the ratio  $\frac{R_{\text{off}}^i - R_{\text{on}}^i}{R_{\text{on}}^i}$  for each single memristor and  $N_{ii}$  represents the disorder at fixed network topology. The matrix  $T$  enters also in the modified projector operator  $\Omega'$ . The operator  $\Omega'$  is in fact a non-orthogonal projector, *i.e.*,  $\Omega' = A^t (A T A^t)^{-1} A T$ ; if  $N_{ij} = 0$ , we recover the previous equation with a symmetric projector operator. This type of non-orthogonal projectors was also found relevant for the mixture of active and passive memristive components, in which  $T$  can also take negative values [12,23].

Recent simulations, in which we have accounted for the disorder, have shown that for larger graphs and statistics, the relaxation of the average internal memory, once we factor in the disorder, is compatible with a logarithmic one. For instance, in fig. 8 we plot the average parameter  $\langle w \rangle$  that was obtained by averaging over different realizations of the disorder and for longer times. We see that

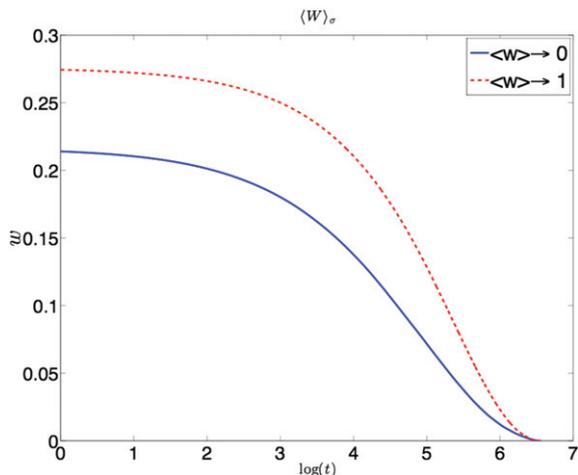


Fig. 8: Average relaxation for fixed  $\sigma = 0.05$ , averaged over 20 simulations, on a complete graph with 50 vertices (1225 edges). The simulations were performed using  $dt = 0.1$ ,  $\alpha = 0.003$ ,  $\beta = 100$  and were homogeneous across the system.

a  $\log(t)$  regime is established after an initial slow relaxation. This is one of the typical features of glassy systems which we find interesting in these rather simple systems. If we aim to use these systems for computation, glassiness in the dynamics will naturally imply a slow approach for the solution of the problem under scrutiny, for instance. It is interesting to note that a connection to disordered systems can be made a bit more precise by looking at the properties of the asymptotic states  $w_i$  as a function of the matrix  $\Omega$  even in the case with homogeneous memristors. It is interesting to note that the equation which describes the evolution of memristors, similarly to the case of resistors, also tries to solve another optimization problem, but this time more complicated: the QUBO, or Quadratically Unconstrained Binary Optimization and known in physics as the ground state of the Ising model. It has been noted first in [24] that for a model of mean-field memristor interaction a Lyapunov function, similar in spirit to the Blume-Capel model [25], exists. That model, in particular, is shown to be exactly solvable. In particular, it has been suggested in [24] that the the mean-field theory, similarly to the case of the Curie-Weiß model, provides a good estimate of the asymptotic value of the parameter  $\langle w \rangle$ , which can be considered an order parameter which parallels the mean magnetization of the Ising model. Given the fact that we are interested in the zero-temperature limit of the system, the mean-field theory is used at finite temperature and the temperature sent to zero at the end of the calculation. The averages are then intended as averages over the initial conditions for the dynamics. For instance, in fig. 9 we show the order parameter  $\langle w \rangle$  for the case  $\alpha > 0$  and  $\alpha < 0$  obtained via Monte Carlo and compared with the mean-field theory result as a function of the mean external voltage  $S$ . The case  $\alpha < 0$  is asymptotically stable, while the case  $\alpha > 0$  is asymptotically unstable. Yet, the information on the position of the

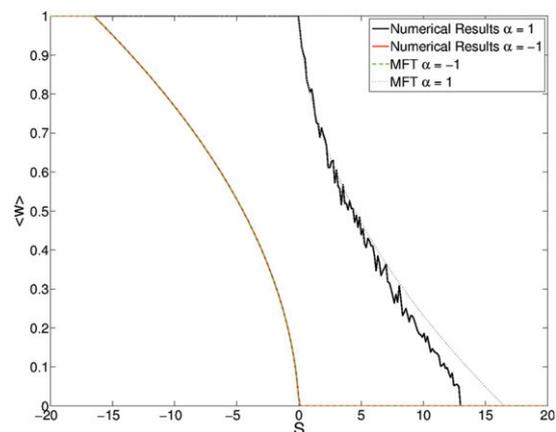


Fig. 9: Mean-field theory vs. numerical results for the asymptotic average parameter  $\langle w \rangle$ , taken from [24].

asymptotic fixed point of the dynamics can be used for obtaining information about the averages  $\langle w \rangle$ . In particular, this suggests that the mean-field theory of spin glasses could be an important source of inspiration to study more general systems [26].

For the more general case of eq. (13), it has been noted that the Lyapunov functional can be approximated by an Ising model (*i.e.*, binary, rather than continuous variables), in which the exchange interaction is proportional to  $\Omega_{ij}$  and with a non-zero effective external field. For the case of ideal memristive circuits,  $\Omega$  has to be a very specific matrix based on the circuit [27]. It has been observed numerically that for random graphs, the distribution of the elements  $\Omega_{ij}$  is a trimodal distribution; however, the bulk of the probability is well approximated by a zero-centered Gaussian distribution with  $\langle \Omega_{ij}^2 \rangle \approx \frac{1}{N}$ , where  $N$  is the number of memristors (or edges in the circuit). This implies memristive networks, reinforcing the connection to the theory of mean-field spin glasses [28].

From the point of view of optimization purposes, however, the system of differential equations can be simulated for arbitrary  $\Omega$  in principle, which suggests a heuristic optimization algorithm for the QUBO type of problems. Thus, the memristive differential equation can serve as a heuristic method for tackling a NP-complete problem such as QUBO [27]. These problems are NP-complete because there is no known algorithm that is better than exhaustive search: because of the binary nature of the variables, we necessarily have to explore all the  $2^N$  possible values of the variables  $\vec{w}$  to decide which extremum (or extrema) is better. In a certain sense, the memristive differential equation is a relaxation of the QUBO problem to continuous variables. This is the same class of the frustrated Ising model. The connection to NP-complete problems has also been observed in other memristor-based architectures [29,30]. The functional that the memristive networks are trying to locally maximize is the functional

$$M(W) = \sum_i \left( r_i - \frac{p}{2} \Sigma_{ii} \right) w_i - \frac{p}{2} \sum_{i \neq j} w_i \Sigma_{ij} w_j, \quad (14)$$

where  $r_i$  are the external fields and  $\Sigma$  the exchange interaction. The mapping between the optimization of the returns above and the equivalent memristive equation is

$$\Sigma = \Omega, \quad \frac{p}{2} = \alpha\xi, \\ \frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta} \sum_j \Omega_{ij} S_j = r_i - \frac{p}{2}\Sigma_{ii}, \quad (15)$$

We can obtain the vector  $S$  through inversion of the matrix  $\Sigma$ , if it is invertible. Algorithmically, there is still the freedom of choosing  $\xi$  and  $\alpha$  given  $p$ , but the two limits are different in nature:  $\xi \gg 1$  is the deep nonlinear regime, while  $\alpha \gg 1$  is the deep diffusive regime. Nonetheless, the mapping between the asymptotic dynamics and the Ising model allows the use of inference methods for the Ising model [31]. It is worth mentioning that other connections to statistical physics of disordered systems, and in particular highlighting the importance of memristor switching for the dynamics of memristive systems have been investigated in [32,33].

**Conclusions.** – In the present paper we have discussed the connection between the dynamics of memristors, graph theory and the properties of certain disordered systems. This paper is meant as an invitation to the more theoretically inclined researchers. We have highlighted the open problems in the analysis of a specific toy model of memristive endogenous dynamics and heuristic optimization algorithms for quadratically unconstrained binary optimization. We hope that the reader will use the toy model we have discussed in this paper as a playground for answering some precise mathematical questions regarding the interesting dynamics of memristors.

\*\*\*

FC acknowledges the support of NNSA for the U.S. DoE at LANL under Contract No. DE-AC52-06NA25396. AZ would like to thank Invenia Labs, the London Institute for Mathematical Sciences and the Los Alamos National Laboratory for funding and hospitality.

## REFERENCES

- [1] CARAVELLI F. and CARBAJAL J. P., *Technologies*, **6** (2018) 118 (engrXiv/c4qr9).
- [2] KUZUM D., YU S. and WONG H.-S. P., *Nanotechnology*, **24** (2013) 382001.
- [3] YANG J. J., STRUKOV D. B. and STEWART D. R., *Nat. Nanotechnol.*, **8** (2013) 13.
- [4] STRUKOV D. B., SNIDER G., STEWART D. R. and WILLIAMS R. S., *Nature*, **453** (2008) 80.
- [5] VALOV I. *et al.*, *Nat. Commun.*, **4** (2013) 1771.
- [6] YANG J. J., STRUKOV D. B. and STEWART D. R., *Nat. Nanotechnol.*, **8** (2013) 13.
- [7] CHUA L. O., *IEEE Trans. Circuit Theory*, **18** (1971) 13.
- [8] CHUA L. O. and KANG S. M., *Proc. IEEE*, **64** (1976) 209.
- [9] PERSHIN Y. V. and DI VENTRA M., *Adv. Phys.*, **60** (2011) 145.
- [10] DI VENTRA M. and PERSHIN Y. V., *Nat. Phys.*, **9** (2013) 200.
- [11] INDIVERI G. and LIU S.-C., *Proc. IEEE*, **103** (2015) 1379.
- [12] CARAVELLI F., TRAVERSA F. L. and DI VENTRA M., *Phys. Rev. E*, **95** (2017) 2.
- [13] PERSHIN Y. V. and DI VENTRA M., *Phys. Rev. E*, **84** (2011) 046703.
- [14] BOLOBAS B., *Modern Graph Theory, Springer Graduate Texts in Mathematics* (Springer Verlag, Berlin) 2012.
- [15] NILSSON J. W. and RIEDEL S., *Electric Circuits*, 9th edition (Pearson Education, Saddle River, NJ) 2011.
- [16] CARAVELLI F., *Int. J. Parallel Emergent Distrib. Syst.*, **33** (2017) 350.
- [17] CARAVELLI F., *Phys. Rev. E*, **96** (2017) 052206.
- [18] HASTING M. B., *Phys. Rev. Lett.*, **93** (2004) 140402.
- [19] SCHUCH N. *et al.*, *Phys. Rev. A*, **84** (2011) 032309.
- [20] CARAVELLI F., HAMMA A. and DI VENTRA M., *EPL*, **109** (2015) 28006.
- [21] AVIZIENIS A. V. *et al.*, *PLoS ONE*, **7** (2012) e42772.
- [22] STIEG A. Z., AVIZIENIS A. V. *et al.*, *Adv. Mater.*, **24** (2012) 286.
- [23] RIAZA R., *Discrete Appl. Math.*, **160** (2012) 280.
- [24] CARAVELLI F. and BARUCCA P., *EPL*, **122** (2018) 40008.
- [25] CAMPA A., DAUXOIS T. and RUFFO S., *Phys. Rep.*, **480** (2009) 57.
- [26] PARISI G., *Lett. Math. Phys.*, **88** (2009) 255.
- [27] CARAVELLI F., arXiv:1712.07046 (2018).
- [28] SHERRINGTON D. and KIRKPATRICK S., *Phys. Rev. Lett.*, **35** (1975) 1792.
- [29] TRAVERSA F. L. and DI VENTRA M., *Chaos*, **27** (2017) 023107.
- [30] TRAVERSA F. L., RAMELLA C., BONANI F. and DI VENTRA M., *Sci. Adv.*, **1** (2015) e1500031.
- [31] LOKHOV A. Y., VUFFRAY M., MISRA S. and CHERTKOV M., *Sci. Adv.*, **4** (2018) e1700791.
- [32] PERSHIN Y. V., SLIPKO V. A. and DI VENTRA M., *Phys. Rev. E*, **87** (2013) 022116.
- [33] SHELDON F. C. and DI VENTRA M., *Phys. Rev. E*, **95** (2017) 012305.