Article

Asymptotic Behavior of Memristive Circuits

Francesco Caravelli

Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA; caravelli@lanl.gov

Abstract: The interest in memristors has risen due to their possible application both as memory units and as computational devices in combination with CMOS. This is in part due to their nonlinear dynamics, and a strong dependence on the circuit topology. We provide evidence that also purely memristive circuits can be employed for computational purposes. In the present paper we show that a polynomial Lyapunov function in the memory parameters exists for the case of DC controlled memristors. Such a Lyapunov function can be asymptotically approximated with binary variables, and mapped to quadratic combinatorial optimization problems. This also shows a direct parallel between memristive circuits and the Hopfield-Little model. In the case of Erdos-Renyi random circuits, we show numerically that the distribution of the matrix elements of the projectors can be roughly approximated with a Gaussian distribution, and that it scales with the inverse square root of the number of elements. This provides an approximated but direct connection with the physics of disordered system and, in particular, of mean field spin glasses. Using this and the fact that the interaction is controlled by a projector operator on the loop space of the circuit. We estimate the number of stationary points of the approximate Lyapunov function and provide a scaling formula as an upper bound in terms of the circuit topology only.

Keywords: memristive circuits; spin models; disordered systems

1. Introduction

The study of memristors has become recently an area of interest [1–3] for a variety of reasons. This is not only due to the fact that memristors are considered by many the fourth circuit element, but also due to their potential applications and their interesting dynamics. In particular, it has been understood from a purely computational theory perspective that the combination of memristive and CMOS components leads to universal computing machines, called memcomputers [4]. From a theoretical perspective, circuits made of memristors have been shown to exhibit non-trivial dynamics which can in principle serve for computational purposes [5–8]. Memristors are passive components which can be thought of as a time varying resistance sensitive to the either the current or the voltage, which in turn depends on a dynamical internal state variable. The main characteristic of a memristor is a pinched hysteresis loop in the Current-Voltage diagram when controlled in alternate voltage [9–11]. The aim of this paper is characterizing the asymptotic behavior of purely memristive circuits (i.e., only memristors) for general circuit topology. An open question which has not insofar been answered at a theoretical level is what is the role of the circuit topology in the relaxation process of memristive circuits. It is in fact thought that the topology plays a role in the optimization capabilities of these circuits with memory (also called information overhead by Di Ventra and Traversa [4]). We provide a precise answer for the simpler class of circuits in which we know the exact role of topology, and show that the dynamical interactions across the circuit map into a coupling term between the memristors for the Lyapunov function for the dynamics. Specifically, we show that memristive circuits, if controlled with constant voltage, can perform naturally a local unconstrained optimization, specifically a Quadratically

Unconstrained Binary Optimization (QUBO) [12]. In order to gain some theoretical understanding on the complexity of the asymptotic states we employ the Kac-Rice formula (average number of stationary points) to provide a rough upper bound on the complexity of the function. In the process of studying random circuits, we also show the connection between memristor dynamics on random circuits and the Sherrington-Kirkpatrick model.

Our results not only establish a direct and analytical connection between (local) optimization and memristors, but potentially introduces a new class of (heuristic) optimization algorithms for combinatorial problems based on the first order dynamical equation studied in the present paper. In fact, an interesting byproduct of the present paper is that the Lyapunov function we derived goes beyond the physically implementable circuits, for example, in principle it applies to cases in which the projection operator is replaced with (semi-)positive operators. As an application, we map a classical problem of stock returns optimization (Markowitz) into ours, and test the performance of the found optimization procedure. We provide an instance of optimization of the Nikkei 225 dataset in the Markowitz framework, and show that it is competitive at least compared to exponential annealing, which usually performs poorly on hard combinatorial problems.

2. Memristive Circuits

Before we introduce the dynamics for a generic circuit, we first briefly discuss the type of memristors under scrutiny [9]. Specifically, we consider memristors whose internal dynamics (the parameter w) depends on the current only, and in which the resistance depends linearly on an internal parameter w, and satisfy the linear relation $R(w) = R_{on}(1 - w) + R_{off}w$, with $0 \leq w(t) \leq 1$. We consider in particular the time evolution of a single memristor with diffusive dynamics [13,14] and for an applied voltage S:

$$\frac{d}{dt}w(t) = \alpha w(t) - \frac{R_{on}}{\beta}I = \alpha w(t) - \frac{R_{on}}{\beta}\frac{S}{R(w)} \tag{1}$$

which shows that a competition between drift and decay occurs. This type of memristors has been considered for machine learning applications. With our convention, R_{on} and R_{off} are the limiting resistances for $w = 0$ and $w = 1$ respectively ($R_{off} > R_{on} > 0$), and α and β are constants which set the timescales for the relaxation and excitation of the memristor respectively. It is worth mentioning that our model is different from the one introduced in Reference [9] as it has not only opposite polarity, but also a decay function which is slightly different. In a recent paper [15] it has been observed that mean field theory and techniques from statistical physics can be applied to study a specific circuit topology. In that paper, it was also noted that a Lyapunov function exists, and that zero temperature mean field theory provides a good estimate for the average asymptotic dynamics for a single mesh of memristors. In the present paper we extend some of these results to a more general class of memristive circuits, which sheds new light onto this type of nonlinear systems. Specifically, in Reference [5] the following differential equation was derived for a generic but purely memristive circuit:

$$\frac{d\vec{W}}{dt} = \alpha\vec{W} - \frac{1}{\beta}(I + \xi\,\Omega W)^{-1}\Omega\vec{S}(t), \tag{2}$$

where $\vec{S}$ is the vector of voltage sources in series to each memristor, while I is the identity matrix. It is immediate to observe that the nonlinearity is controlled by the parameter $\xi = \frac{R_{off} - R_{on}}{R_{on}}$. It is important to note that Ω and W are matrices, while I is the identity matrix. W is the diagonal matrix with the memristor memory values $w_i(t)$ on the diagonal, meanwhile $\Omega = A^t(AA^t)^{-1}A$ is a projector operator on the fundamental loops of the circuit, being A the cycle matrix of the directed graph representing the circuit flows. Given an orientation of each loop and edge, then $A_{i\beta}$ has dimensionality $L \times N$, where N is the number of memristors and L the number of fundamental loops. If a memristor β has the same orientation of a loop i, then $A_{\beta i} = 1$, and -1 if opposite. $A_{\beta i} = 0$ if the memristor β does not belong to

the loop i. Thus the matrix A has only $-1, +1, 0$ values (The matrix Ω is not diagonal because edges of a graph can belong to more than one loop.). The richness of the dynamicsl of this equation has been characterized in Reference [5], while the locality properties of Ω in Reference [16]. The important observation that we anticipate is that the dynamical Equation (2) has at least one Lyapunov function which we have derived analytically. While the derivation is provided in the Appendix A, here we provide only its functional form:

$$
\begin{aligned}
L(W) \;=\;& -\frac{\alpha}{2}\sum_i W_i^2 - \frac{\alpha\xi}{3}\sum_i \Omega_{ii} W_i^3 \\
& -\; \alpha\xi \sum_{i\neq j}\Omega_{ij}W_i W_j^2 + \frac{1}{\beta}\sum_{ij}W_i\Omega_{ij}S_j,
\end{aligned}
\tag{3}
$$

from which a few key facts can be immediately observed. For instance, the role of Ω is the one of a coupling matrix, as one would expect, and that it is not quadratic in the internal memory variables. Equation (3) satisfies $\frac{d}{dt}L(W) \leq 0$ whenever $M_i = -2\alpha\xi\sum_{j\neq i}W_j\Omega_{ji}W_i$ is small (thus for very weakly interacting memristors, as it depends only on the offdiagonal terms) and $\frac{d}{dt}L(W) = 0 \leftrightarrow \frac{d}{dt}\vec{W} = 0$. One can obtain a more precise bound on the derivative of the Lyapunov function. Let us define $||\Omega\vec{S}||^2 = N^2 s^2(N)$. We have proved in the Appendix A that if
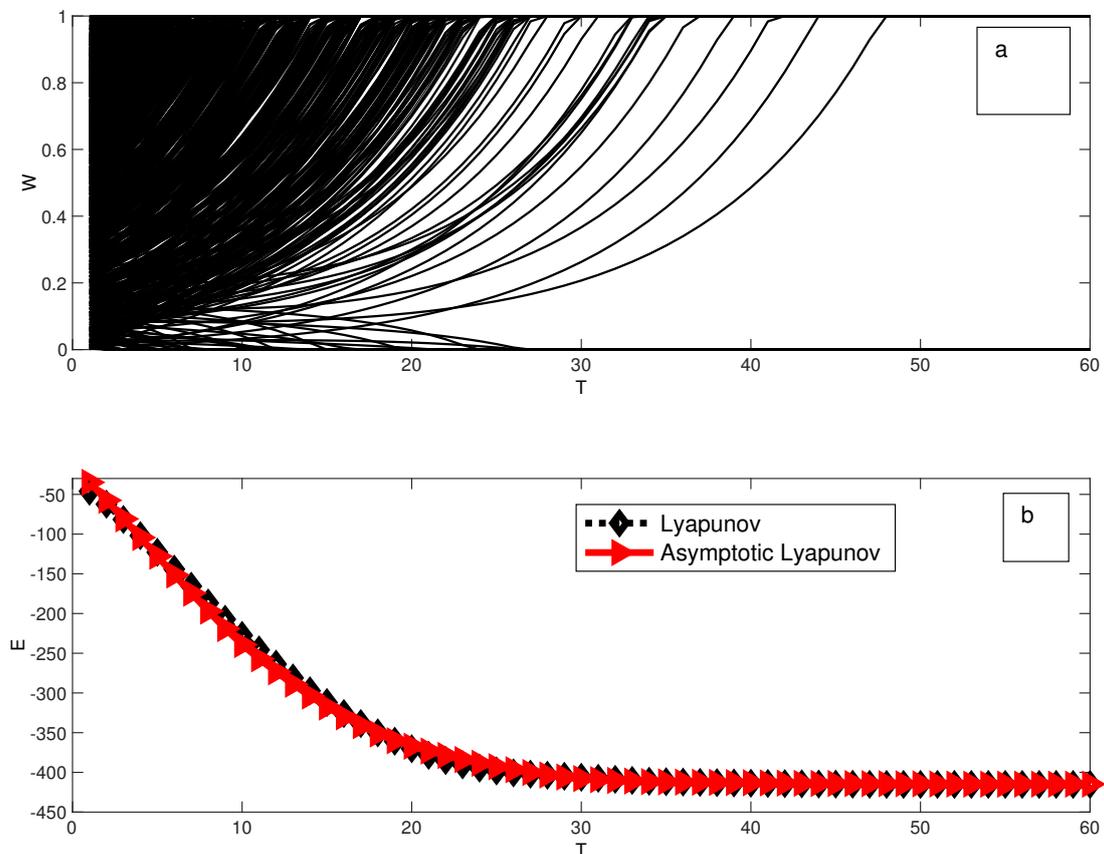
$$
4\xi^2(1+\xi)\bar{\Omega}^2 < \frac{s(N)^2}{\alpha^2\beta^2} - 2\frac{s(N)}{\alpha\beta},
\tag{4}
$$

then $\frac{d}{d}L(W) < 0$, where $\bar{\Omega} = \max_{ij}|\Omega_{ij}|$. It is interesting to note that the order parameter $\frac{s(N)}{\alpha\beta}$ is a generalization of the one found in Reference [15] and which controlled the asymptotic state of the effective mean field circuit. Here again it plays a role. Note that $\xi^2(1+\xi)\bar{\Omega}^2$ is what controls the height of the parabola, and intuitively the larger then nonlinearity the higher the voltage has to be for the system to relax according to the Lyapunov function. Thus, Equation (4) establishes a dynamical phase diagram.

We stress that the dynamics is constrained in the hypercube $[0,1]^N$, where N is the number of memristors, and thus the Lyapunov function above when a memristor reaches the boundaries of this domain. However, if we now use the fact that the dynamics is controlled with constant voltage each memristor will eventually reach the asymptotic value 1 or 0. In this asymptotic limit, we note that the function above can be interpreted as a spin-like model asymptotically. In Reference [15] this phenomenon has been partly explained by the emergence of unstable fixed points in the dynamics (although only in the mean field approximation). For the specific case of the Equation (2) however, it is necessary to go beyond this approximation, for which each memristors will have (intuitively) a different fixed point associated to it. This phenomenon, which is circuit dependent, is associated to the structure of the couplings Ω_{ij}, and will be studied elsewhere. Anyhow, for large times the Lyapunov function can be approximated by ($W_j^n \approx W_j$)

$$
\begin{aligned}
L(W(t\gg 1)) \;\approx\;& -[\frac{\alpha}{2}\sum_i W_i + \frac{\alpha\xi}{3}\sum_i \Omega_{ii} W_i \\
& +\; \alpha\xi\sum_{i\neq j}\Omega_{ij}W_i W_j - \frac{1}{\beta}\sum_{ij}W_i\Omega_{ij}S_j] \\
\equiv L_a(W) \;=\;& -[\sum_i W_i\left(\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij}S_j\right) \\
& +\; \alpha\xi\sum_{i\neq j}\Omega_{ij}W_i W_j],
\end{aligned}
\tag{5}
$$

where we have used the fact that $W_i^n = W_i$ if $W_i = 1, 0$. An effective external field $h_i = \frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\Omega_{ij}S_j$ emerges in this asymptotic approximation. We also test numerically both the Lyapunov function of Equations (A1) and (6) in Figure 1. Meanwhile in Figure 1a we plot the evolution of each memory element, in Figure 1b we plot the evolution of both the Lyapunov function and the asymptotic approximation of it, which remains close to the exact one. On the other hand, in Figure 2 we show the difference between the obtained Lyapunov function and the asymptotic one as a function of time. This mapping is reminiscent of the case of continuous neuronal networks introduced by Little [17] and then Hopfield in a series of important papers [18,19]. The Little-Hopfield model has sparked interest from the Statistical Physics community since the very beginning [20]. In the past years this particular line of study has been subject of scrutiny by many experimental groups [21–23]. The key difference is that in each case, these were studied in conjunction with ordinary and active (CMOS) electronic components to build Hopfield learning networks. We argue instead that memristive circuits *per se* form a special kind of Hopfield network defined by the Lyapunov function above, without the need of extra components.

Figure 1. Dynamics of the memory is shown in Figure (**a**) and the corresponding evolution of the Lyapunov function of Equation (A1) ((**b**), continuous line) for 8750 memristors and for the asymptotic function of Equation (6) Figure ((**b**)-dashed line). We considered $\alpha = 0.1$, $\beta = 1$ and $\xi = 10$, for Ω_{ij} from a random circuit of the Erdos-Renyi type with $p = 0.9$. The sources $\vec{S}$'s elements were chosen at random between $[-0.05, 0.05]$.
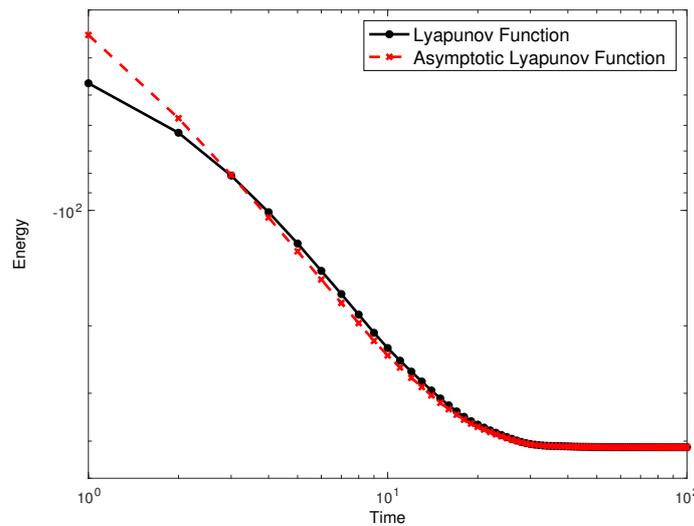
Figure 2. Dynamics of the Energy (**a**) and corresponding evolution of the Lyapunov function (**b**) in Figure 1.

Functional Complexity for Random Circuits

We now study random circuits, in particular for the function complexity of Equation (6). First, this will allow us to show an interesting connection to the field of disordered systems, and in this approximation provide estimates of the complexity of the Lyapunov function. In fact, minimizing quadratic function with discrete variables is in general a hard problem to solve. We are however in the situation in which the problem at hand is in nature both continuous in time and in the variables (being these between 0 and 1), but naturally provides an answer for the more complicated case usually associated to finding the ground state of a frustrated spin systems. Despite the fact that we are not able in the present paper to provide a complete answer to which optimization class a system of purely memristive circuits belongs to, we try to provide some answers to the questions above using some techniques introduced to study random polynomials [24,25]. This will be important in light of the fact that we have shown that memristors perform a local optimization, rather than a global one, and that the circuit constraints should enter somehow.

The number of stationary points of a generic multi-varied function $L(W)$ can be estimated by:

$$\# = \sum_{\{\vec{w}=1,0\}} \det(\partial^2_{\partial w_i \partial w_j} L(\vec{w})) \prod_i \delta \left(\partial_{w_i} L(\vec{w}) \right). \tag{6}$$

Thus, if we aim to consider the expected number of stationary points #, then it does make sense to consider the (quenched) average of the quantity in Equation (6) for the distribution for $\langle \cdot \rangle_{P(\Omega)}$ with respect to the coupling matrix Ω. One important observation is that because Ω is a projector, $\det(\partial_{W_i} \partial_{W_j} L(W))$ can be calculated exactly without knowledge of the elements distribution, but only using the fact that Ω is a projector (details in the Appendix B). We find that the average number of stationary points can be split in the form:

$$\langle \# \rangle = (\sqrt{3} \alpha \xi)^N \left(1 - \frac{1}{\sqrt{N}} \right)^L Z(\vec{S}, Q) \tag{7}$$

where L is the number of fundamental loops of the circuit, and where:

$$Z(\vec{S}, P(\Omega)) = \langle \sum_{\{\vec{w}=1,0\}} \prod_i \delta \left(\partial_{w_i} L(\vec{w}) \right) \rangle_{P(\Omega)}. \tag{8}$$

depends on the distribution of the elements of the matrix Ω.

We first need, then, to understand the distribution $P(\Omega)$ for random circuits. Since the analytical calculation of such distribution goes beyond the scope of this paper, we perform such analysis numerically. We first generate random circuits with Erdos-Renyi graph above the percolation threshold ($p = 0.7$). We are interested, in particular, in the scaling with respect to the number of memristors N. For fixed Ω, the distribution $P(\Omega)$ of off-diagonal elements is shown in Figure 3a. We observe that although this seems to be unimodal, a careful look shows that it is not. We observe in Figure 3b that there is a non-zero probability of having elements not distributed around zero, but that these are orders of magnitudes smaller than the central distribution. While it is not surprising that the distribution is not completely random (thus, there are correlations among elements) since these are constrained by $\Omega^2 = \Omega$, it is surprising to note that we can roughly approximate this distribution with a unimodal one. We are in particular interested in how the width of the distribution scales with the number of memristors. This can be calculated exactly if we know the scaling of the diagonal elements, as we will see below. In Figure 4 we show how the diagonal elements of Ω scale with N. We fit first the diagonal elements, precisely $1 - \Omega_{ii}$, showing that it scales as a power law to a good approximation. If we perform a fit using the functional form $\Omega_{ii} \approx 1 - \frac{c}{N^\alpha}$, we obtain the best fit values $c = 1.74557 \approx \sqrt{3}$ and find that $\alpha = 0.5043 \approx 1/2$. We will, since now on, consider these values in what follows. It is important to note that the scaling above is sufficient to obtain a scaling for the off-diagonal elements as well. We define the matrix of off-diagonal element G, as $\Omega \approx (1 - \sqrt{\frac{3}{N}})I + G$. Since we have that $\Omega^2 = \Omega$, it is easy to see that G must satisfy the equation $G^2 + (1 - 2\sqrt{\frac{3}{N}})G + \sqrt{\frac{3}{N}}(\sqrt{\frac{3}{N}} - 1)I = 0$. For $N \gg 1$ this implies that also G satisfies the equation $G^2 = G$ as well. We can at this point solve for an eigenvalue equation for G, and obtain $\lambda = \{\sqrt{\frac{3}{N}}, -1 + \sqrt{\frac{3}{N}}\}$. Since this must be a projector for large values of N, we must have $\Omega \approx I + \sqrt{\frac{3}{N}}Q$, where $G = \sqrt{\frac{3}{N}}Q$, with Q again a projector. Thus, the scaling approximation obtained for random Erdos-Renyi type circuits, we have $\Omega_{ii} \to 1$, meanwhile $\Omega_{ij} \approx \frac{\sqrt{3}}{\sqrt{N}}Q_{ij}$. This is the scaling we will use in the following.
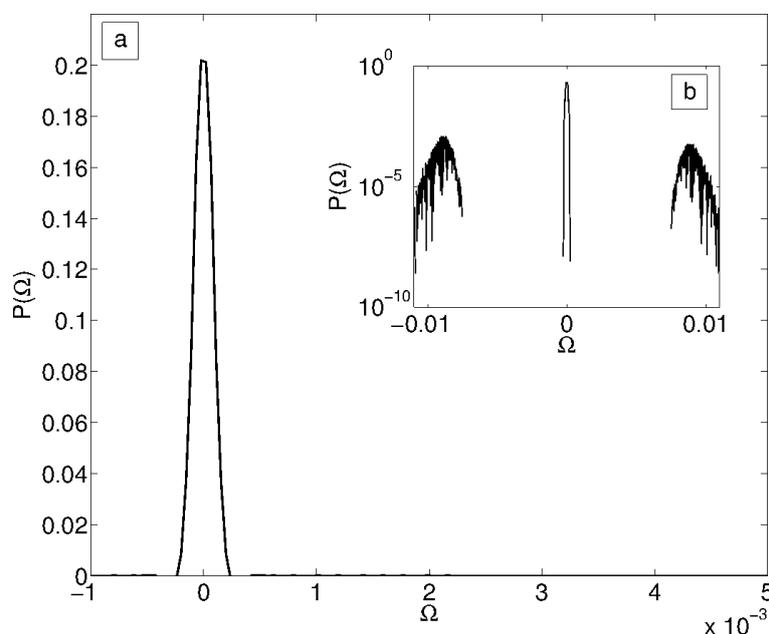
Figure 3. Distribution of the elements of the matrix Ω in the case of an Erdos-Renyi graph. In (**a**) we plot the distribution density, while in (**b**) the logarithm, which show that the distribution is multimodal. However, we see that there are roughly 5 orders of magnitudes between the bulk of the probability distribution and the two smaller values of the density at ± 0.01, which shows that the distribution can be well approximated with an unimodal one.
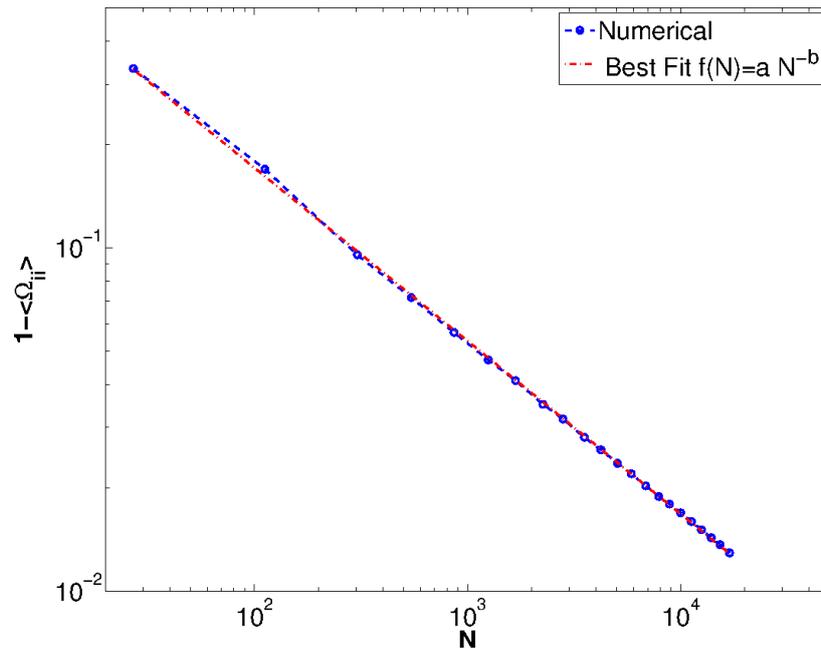
Figure 4. Scaling of the components Ω_{ii} as a function of the number of memristors for randomly generated circuits.

While in what follows we show only the intermediate results, all the exact calculations are provided in the Appendix B. We can in fact a this point calculate Z in Equation (6) for random circuits. We assume that $P(Q_{ij}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{Q_{ij}^2}{2\sigma^2}}$, where σ is an effective width for $N = 1$. In this case, calculating Z can be done by means of Gaussian integrals. In this case, we obtain, in the limit $\alpha\xi \gg 1$:

$$\langle \# \rangle \approx \left(1 - \frac{1}{\sqrt{N}}\right)^L \left(\frac{\sqrt{3}}{\sqrt{\pi}\sigma}\right)^N. \tag{9}$$

We now use Euler's relation $L = N - V + \chi$, where V is the number of vertices, χ the Euler characteristic and use the Erdos-Renyi relation $N = p(V)\frac{V(V-1)}{2}$ where $p(V)$ is the edge probability. Now we have that, since $\chi > 0$ and only topological, while L scales linearly for large N in the number of memristors. We now observe that given $\rho = \frac{\sqrt{3}}{\sqrt{\pi}\sigma}$, we have that in the limit $N \to \infty$, $\langle \# \rangle$ goes to infinity for $\rho > 1$, while it goes to zero for $\rho \leq 1$, as $\lim_{N\to\infty}(1 - \frac{1}{\sqrt{N}})^N = 0$. We thus obtain the result that for $\sigma > \frac{\sqrt{3}}{\sqrt{\pi}}$ the Lyapunov function seems to have a large number of stationary points and thus a hint of the hardness of the minimization problem.

It is easy at this point to provide an approximate mapping between the asymptotic Lyapunov function and the Hamiltonian of a mean-field spin glass of the Sherrington-Kirkpatrick type with an external field:

$$L(\sigma) = \sum_i \tilde{h}_i \sigma_i + \frac{\alpha\xi\sqrt{3}}{\sqrt{N}} \sum_{i \neq j} Q_{ij}\sigma_i\sigma_j \tag{10}$$

where we used the mapping $W_i = \frac{1}{2}\sigma_i - \frac{1}{2}$ with $\sigma_i = \pm 1$ and $\tilde{h}_i = \frac{h_i}{2} - \frac{1}{2}\sum_j Q_{ij}$ and disregarded an unimportant constant which only shifts the function. The result above is a generalization of what has been found in Reference [26] in the limit $\xi \ll 1$, for example, the fact that the dynamics of memristors follows a constrained gradient descent (Rosen projections). Moreover, it provides a physical system to test experimentally the physics of mean field spin glasses [27].

3. Asymptotic State Recollection and Combinatorial Optimization

One important question for memristive circuits is what is the asymptotic state they reach. As we have argued, this can be answered by looking at the minima of the Lyapunov function for the case of constant voltages. Although this is a hard problem to solve in its full generality, we can use some approximate analytical formulae to provide an answer in at least for some region of the parameters. First, we consider the case $\alpha = 0$. It is easy to see that we can integrate the equation [26] to obtain:

$$\vec{W} + \frac{\xi}{2}\Omega\vec{W}^2 = -\frac{1}{\beta}\Omega\vec{S}(t - t_0) + \vec{c} \tag{11}$$

where c is an integration constant. Equation (12) is a *unimodular quadratic vector equation*, which is the special case of vector equations of the form $\vec{x} + \vec{b}(\vec{x}, \vec{x}) = \vec{c}$. This equation does not have an exact solution in closed form but several numerical methods have been developed [28]. Here we use a heuristic method to identify what are the asymptotic values of $\vec{W}$. Again we use the fact that in the limit $t \to \infty$, we observe numerically the asymptotic values of W are either 0 or 1. For $\xi = 0$, we have $\vec{W}(t = \infty) = \frac{1 - \text{sign}(\Omega\vec{S})}{2}$. For $\xi > 0$ and in the asymptotic limit, $\lim_{t\to\infty} \vec{W}^2 \approx \vec{W}$. Using this approximation, we obtain:

$$(I + \frac{\xi}{2}\Omega)\vec{W} = -\frac{1}{\beta}\Omega\vec{S}(t - t_0) + \vec{c}. \tag{12}$$

We then can obtain the asymptotic formula:

$$\vec{W}(t = \infty) \quad \approx \quad \frac{1 - \text{sign}((I + \frac{\xi}{2}\Omega)^{-1}\Omega\vec{S})}{2} = \frac{1 - \text{sign}(\Omega\vec{S})}{2}, \tag{13}$$

where we note that in the last equation we have used the fact that $1 + \xi/2 > 0$ if $0 < R_{on} < R_{off}$. However, we note that the exactness of the asymptotic behavior depends only on ξ and not on α. The results on how well Equation (13) predicts the asymptotic behavior of the circuit as a function of ξ are shown in Figure 5. Despite our heuristic method and approximation, the obtained values are a good approximation for small values of ξ of the real system. Using the formula above, we now provide the connection with the state recollection of neural networks. We note that the projection operator Ω can be written as [16] $\Omega_{ij} = \sum_{l=1}^{L} \tilde{A}_i^l \tilde{A}_j^l$, where $\tilde{A}_i^l$ is the orthonormalized loop matrix of the circuit. From the theory of neural networks we are aware that the number of stored patterns equals the number of independent eigenvectors of the interaction matrix Ω, which is purely topological, as argued in Reference [26]. In the case of a purely memristive circuit the number of independent memory units number is constrained by the topology of the circuit and it depends on the number of fundamental loops [26]. As we have noted before, the asymptotic behavior of the circuit can be approximated by (where we reinstate α)

$$\vec{W}(t = \infty) \approx \frac{1 - \text{sign}(((1 - \alpha)I + \frac{\xi}{2}\Omega)^{-1}\Omega\vec{S})}{2},$$

where we realize that the source vector $\vec{S}$ can act as the recollection mechanism. If $\vec{S} = \sum_l \rho_l \tilde{A}^l$, for a certain proportionality constant ρ, then $(\Omega\vec{S})_i = \sum_l \rho_l \tilde{A}_i^l$. Thus, the asymptotic configuration of the memristors will be given by the approximate value:

$$\begin{aligned} \vec{W}(t = \infty) \quad &\approx \quad \frac{1 - \text{sign}(((1 - \alpha)I + \frac{\xi\rho}{2}\tilde{A}^l)^{-1})}{2} \\ &= \quad \frac{1 - \text{sign}\left((1 - \alpha)I + \frac{\xi\rho}{2}\tilde{A}^l\right)}{2}, \end{aligned} \tag{14}$$

which shows that stored patterns are indeed in the loop basis.
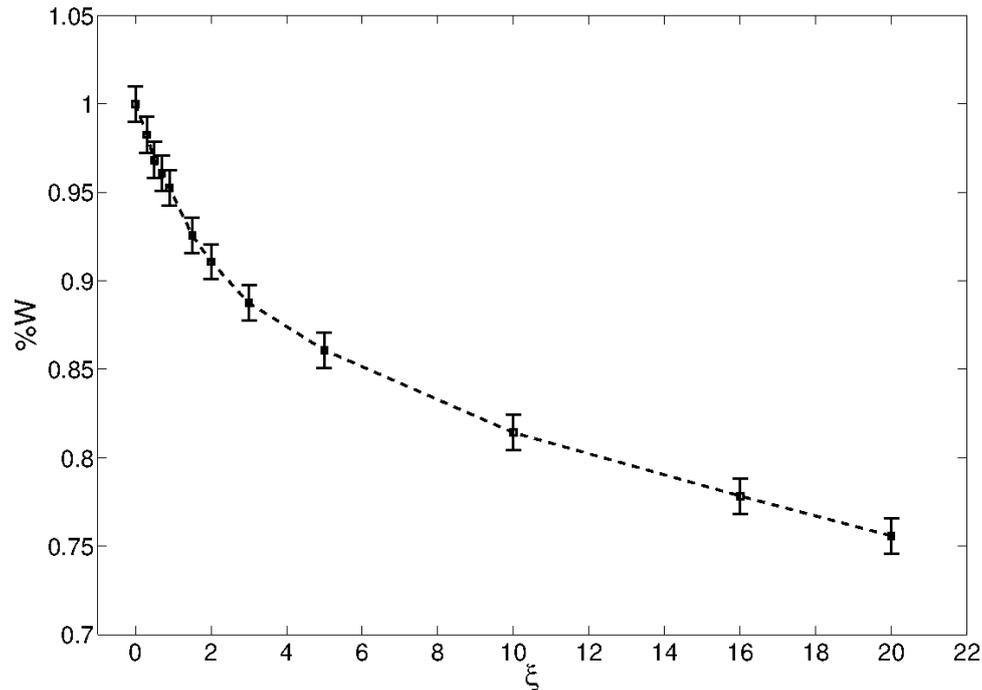
Figure 5. Percentage of exactly predicted values of memristors according to Equation (13) as a function of ξ. The error bar are calculate from 100 samples, with approximately 800 memristors (on average) and $\xi = 10$, $\alpha = 0.1$ and $\beta = 1$, with the vector S drawn at random between $[-0.5, 5]$.

The analysis above is reminiscent of the asymptotic state recollection in the Hopfield-Little model.

We now ask the converse question of how close are the asymptotic states to minima obtained from the Lyapunov functional. As we have seen in the previous sections memristors can be used to provide solutions (possibly sub-optimal) to hard combinatorial problems. This claim should be taken, if not with skepticism, with some care. First we note that $L(W)$ is not positive definite (although it is bounded) and thus we have an asymptotically local stable equilibrium rather than a global one. We are interested in providing some benchmarks on how well memristors find the minimum of the function $L_a(W)$ compared to other optimization techniques. Given 100 samples of random circuits based on Erdos-Renyi random graphs with $p = 0.7$, in Figure 6 we show the energy attained using memristors in 60 time steps (red dots), using a Metropolis algorithm with exponential annealing (blue) with over 8000 steps and using the minimum of 100 random configurations as a reference (blue dots). We see that despite the fact that memristors do not seem to reach the absolute minimum, the energy is closer to the Metropolis result than to the random configuration one. Although it is easy to see that the dynamical equations do not reach the Metropolis attained minimum, it did take less than 100 iterations for the memristive system to converge using a simple Euler-Newton first order integration. This already shows that, as anticipated, memristors perform a local optimization in this regime. In the next section, however, we apply the algorithm to a specific dataset, where this optimization has some advantage over the exponential annealing.
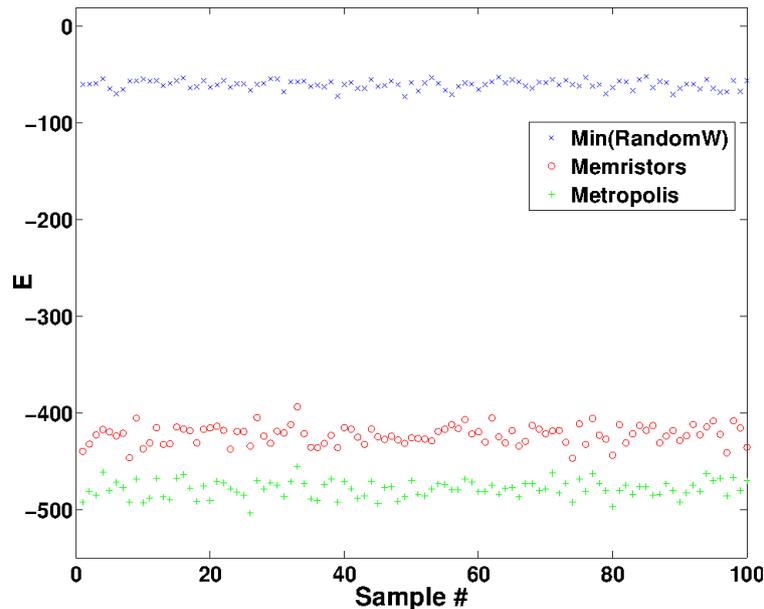
Figure 6. Obtained minimum energy using the minimum over 100 random binary values (blue crosses), roghly 8000 steps of exponential annealing (green pluses) and 60 time steps using the memristor dynamics (red circles). We generated each sample with approximately 800–900 memristors, fixing the percolation parameter to $p = 0.7$, with $\xi = 10$, $\alpha = 0.1$ and $\beta = 1$. The numerical integration was performed using a simple Euler-Newton integration with integration step $dt = 0.1$ and total number of steps $T = 1000$. We truncated the Metropolis annealing to $10 * N$ time steps, with N the number of memristors and with an annealing rate $\lambda = 0.995$ and an exponential annealing law for the temperature given by $Temp(k) = 100 * \lambda^k$.

4. An Application: Minimization of a Quadratic Function

We have argued that memristors' dynamics can be thought, when these are connected in a circuit, as reaching an asymptotic state which is among the minima of a certain Lyapunov function. Albeit real memristors can have a dynamics which is far from ideal, we would like to consider if these could in principle be used for any sort of purpose which goes beyond electronics. In fact, note that (2) can be used as a quick meta-heuristic method to search for the minimum of a function, or to be given as input to other more complicated and efficient optimization techniques. In realistic circuits, the quadratic form matrix Ω which occurs in the Lyapunov function is a projector on the cycle space of the graph. Thus, problems for which analog memristive circuits can be used are naturally those that can be embedded in the cycle basis of a graph [29]. However, in principle we can simulate the system also for matrices Ω that are not necessarily projectors, as we have shown and as we discuss further below. We can thus perform a preliminary test of the applicability of analog memristive circuits in a problem of optimization and see how the circuit performs compared (for instance) to simulated annealing.

We thus consider a simple application of practical importance, for example, studying the problem of investment in a set of assets. For this purpose, we use the 225 Nikkei dataset which is used for benchmarking heuristic optimization algorithms [30], available in Reference [31]. We can use in fact a memristor-inspired optimization scheme, taking advantage of the fact that it works beyond projector operators Ω. We mention that the proof of the Lyapunov function relies only on the matrix $I + \xi \frac{\Omega W + W \Omega}{2}$ being positive definite and thus applies also to any positive matrix and to some non-positive ones (see Appendix A for details) for ξ small enough. Although we cannot use these in a hardware circuit, we can solve the differential equation numerically to infer a minimum heuristically. Using this idea, we ask in which of the assets of portfolio we should invest in (a yes-no answer). The dataset is composed of 225 assets, including returns and the covariance matrix and can thus use the combinatorial Markowitz functional [32]. If we use the fact that maximizing a quadratic function is equivalent to

minimizing the function with a minus sign in front, we can write an equality between the Markowitz function for binary variables and the memristive one:

$$
\begin{aligned}
M(W) &= \sum_i \left(r_i - \frac{p}{2}\Sigma_{ii} \right) W_i - \frac{p}{2}\sum_{i\neq j} W_i \Sigma_{ij} W_j \\
-L(W) &= \sum_i W_i \left(\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij}S_j \right) \\
&\quad - \alpha\xi \sum_{i\neq j} \Omega_{ij} W_i W_j
\end{aligned}
\tag{15}
$$

where p is a trade-off parameter, $\vec{r}$ are the returns and Σ is the covariance matrix. From imposing the equality between the two functionals, we observe the necessity of imposing $\Omega = \Sigma$ and thus

$$
\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij}S_j = r_i - \frac{p}{2}\Sigma_{ii}, \quad \frac{p}{2} = \alpha\xi.
\tag{16}
$$

We can thus obtain the source vector to try to force the system towards the right minimum:

$$
\vec{S} = \beta\Sigma^{-1}\left(\frac{\alpha}{2} + (\frac{p}{2} + \frac{\alpha\xi}{3})\vec{\eta} - \vec{r} \right)
\tag{17}
$$

where $(\eta)_i = \Sigma_{ii}$, for example, the variance of each return. We have the freedom to choose arbitrarily β but either ξ or α would be fixed. We observe that we need to invert the covariance matrix, which is a slow but polynomial in the number of variables. Moreover, we see that the harder the problem to solve (the norm of the inverse of Σ), the smaller β has to be chosen, from which we infer the slowness of the process.

We are now in the position to test the memristive minimizationi against an exponential annealing process, which is known to perform poorly in the combinatorial setting. The results for the case of the Nikkei dataset are shown in Figure 7, comparing the Metropolis annealing to the case of the computational time and in the final value obtained. We performed three different tests. First, we randomized the initial states and performed 100 Simulated Annealing procedures [33], with initial temperature $T = 100$, an exponential annealing rate $\lambda = 0.995$ ($T_k = \lambda^k T_0$) and a number of time steps of $N_T = 500 * N$, where N was the number of assets 100. On the Nikkei 225 dataset, the maximum obtained with the simulated annealing was $r_{max} \approx 9.15$. With the identical initial states, the memristive optimization has obtained a maximum of $r_{max} \approx 14.23$. We have then used the final state of the memristive optimization output as an input to an annealing procedure but with a lower initial temperature $T_0 = 0.025$, thus effectively fast-forwarding the annealing procedure. This combined optimization obtained the absolute maximum, albeit of only less than a percentage better than the memristive optimization, of $r_{max} \approx 14.40$.
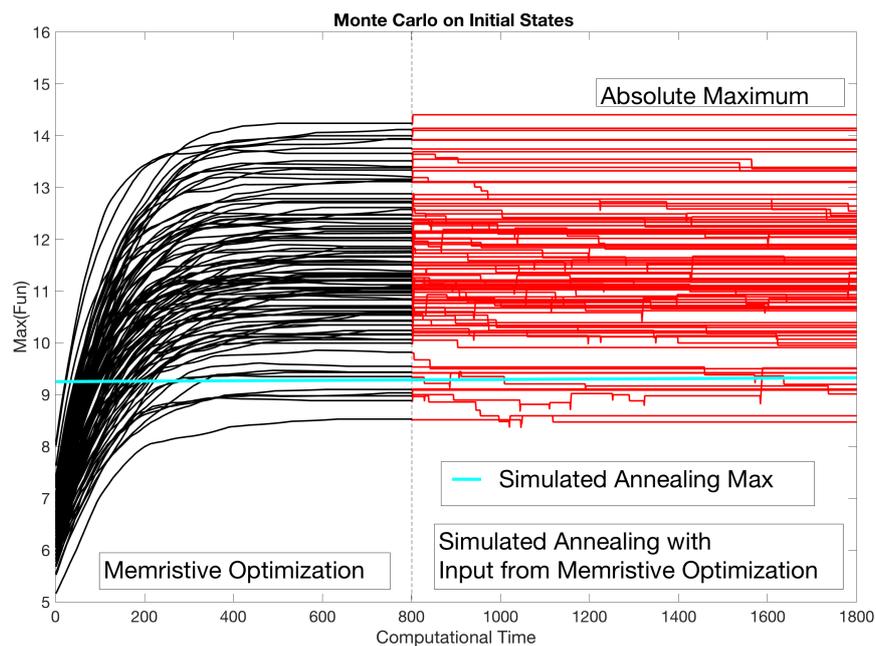
Figure 7. Maximization of the return for the Nikkei 225 dataset using a Monte Carlo procedure. We compare the results between the best of 100 Simulated Annealing procedures (the green line in the plot ≈ 9.15) versus with the memristive optimization proposed in this paper (black lines), with identical initial conditions. The simulated annealing was conducted with an initial temperature $T = 100$ and an exponential annealing rate $\lambda = 0.999$ (and effective time steps N = 500). We have then used the final states of the Memristive Optimization procedure as an input to a simulated annealing procedure with lower temperature $T_0 = 0.025$ and identical annealing rate, which obtained the absolute maximum of the return.

5. Concluding Remarks

In the present paper we have analyzed the asymptotic dynamics of memristive circuits, showing that an approximate Lyapunov function for the dynamics exists for memristors with slow decay. Because of the properties of memristors, we have argued that asymptotically the Lyapunov function of a memristive circuit can be written as a quadratic function on spin-like variables and have connected the coupling matrix to the projector operator on the loop space of the graph [34]. This result shows a direct connection between purely memristive circuits and Hopfield networks, which we argue minimize a similar type of Lyapunov function. The internally stored patterns are in this case the cycles of the graph. This is interesting for various reasons. Insofar it had been argued that a certain degree of external control was necessary in order to perform computation or use memristors as a memory device, either via the use of CMOS or the introduction of capacitors into the network. This paper provides sufficient evidence for the use of memristors in their completely analog regime for computational purposes. This comes at a cost, which is the embedding of the problem of interest in the cycle basis of the graph, which can be a rather nontrivial problem when it is solvable.

The Lyapunov function for the dynamics of memristive circuits can then be used for various purposes. We have first focused on the complexity of these functions in the case of random circuits. We have provided numerical evidence and argued that the couplings scale as in the case of mean-field spin glasses, that is, the Sherrington-Kirkpatrick model, although only if neglecting correlations between matrix elements of lowest order. Using this approximation we have provided approximate formulae for the number of stationary points, provided some topological considerations. This result explains the importance of the initial conditions for the asymptotic state of the memristors. In fact, it is known (for instance in the case of a glassy system) that for a system which has a large set of local minima there is a large sensitivity to the initial conditions.

As a test-bed of these ideas, we have used the memristor dynamics to see how well (or bad) the dynamics leads to the minimization of a certain quadratic functional We have compared the minimization of a combinatorial problem using both simulated annealing and the memristive dynamics introduced in this paper. Although the memristor dynamics could not reach a global minimum, the speed of convergence and the closeness to the Metropolis result suggest the use of these mixed dynamical-combinatorial algorithm to provide quick answers to combinatorial problems on spin-like variables (or alternative 0–1 variables, thus a quadratic unconstrained binary optimization). We do not claim these answers to be optimal but we find nonetheless interesting that a rather simple procedure like the one we did performed remarkably well compared to simulated annealing.

The main result of this paper however remains the Lyapunov functional for the study of the dynamics of memristors and to understand the relaxation properties of these circuits (at least when these are controlled with constant voltage). We have shown a direct connection between optimization and memristive circuits, as advocated by other authors, for instance [35–37]. It is inspiring to think that the minimization of the functional depends on the balance between reinforcement-decay properties of the memristive dynamics, along the lines of other similar algorithms [38] which use collective dynamics as heuristic optimization methods.

Thus, this paper provides needed background work to understand the dynamics of circuits with memory, their sensitivity to initial conditions and their use for computational purposes in the fully analog regime.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Properties of $L(W)$

Appendix A.1. Derivation of $L(W)$

The Lyapunov functional provided in the paper is given by:

$$
\begin{aligned}
L(\vec{W}) &= -\frac{\alpha}{2} \sum_i W_i^2 - \frac{\alpha\xi}{3} \sum_i \Omega_{ii} W_i^3 \\
&\quad - \alpha\xi \sum_{i\neq j} \Omega_{ij} W_i W_j^2 + \frac{1}{\beta} \sum_{ij} W_i \Omega_{ij} S_j.
\end{aligned} \tag{A1}
$$

In order to derive the functional above, we use the differential equation for the memristive dynamics:

$$
(I + \xi\Omega W)\frac{d\vec{W}}{dt} = (I + \xi\Omega W)\alpha\vec{W} - \frac{1}{\beta}\Omega\vec{S}, \tag{A2}
$$

for the case of constant voltages $\vec{S}$. At this point we can obtain the functional form of the derivative of the Lyapunov function:

$$
\begin{aligned}
\frac{d}{dt}L(\vec{W}) &= \partial_t L(W) + \sum_i \delta_{w_i} L(W) \frac{d}{dt} W_i \\
&= -\sum_i \frac{dW_i}{dt}\left[(I + \xi\Omega W)\alpha\vec{W} - \frac{1}{\beta}\Omega\vec{S}\right]_i + \vec{M}\cdot\frac{d\vec{W}}{dt} \\
&= -\sum_{ij} \frac{dW_i}{dt}\,(I + \xi\Omega W)_{ij}\,\frac{dW_j}{dt} + \vec{M}\cdot\frac{d\vec{W}}{dt} \\
&= -||\frac{d\vec{W}}{dt}||^2_{(I+\xi\Omega W)} + \vec{M}\cdot\frac{d\vec{W}}{dt}
\end{aligned}
\tag{A3}
$$

where we assumed that $\partial_t L(W) = 0$, as $\vec{S}$ is constant in time. We also introduced the quantity $M_i = -2\alpha\xi\sum_{j\neq i} W_j\Omega_{ji}W_i$. The variation of the functional is provided in the subsection below in detail. First we prove that $I + \xi\Omega W$ is a positive definite matrix. We use the matrix similarity for eigenvalues, $I + \xi\Omega W \sim I + \xi\sqrt{W}\Omega\sqrt{W}$ to show that given a vector $\vec{b}$, its norm is $\vec{b}^t(I + \xi\sqrt{W}\Omega\sqrt{W})\vec{b} = ||\vec{b}||^2 + \xi||\sqrt{W}\vec{b}||_\Omega$. Since Ω is a projector and is semipositive, then we have that for any non-zero vector $\vec{b}$, $\vec{b}^t(I + \xi\sqrt{W}\Omega\sqrt{W})\vec{b} > 0$. We now observe that the first term in Equation (A4) is always negative, meanwhile the second term is always positive, since the inverse of a positive operator is always positive. However, meanwhile the first term does not scale with any parameter, we have that the second term scales roughly as $\alpha^2\xi^2$. Also, the min norm of the operator $I + \xi\Omega W$ is of order one, which implies that the sup norm of its inverse is also of order one, for arbitrary values of ξ. As a comment, we see that up to this point, this is enough to prove that if Ω is diagonal the second term disappears, as $\vec{M} = 0$. This completes the proof that $\frac{d}{dt}L(\vec{W}) \leq 0$ for the case of Ω diagonal, which is trivial. Also, note that we have $\frac{d}{dt}L(\vec{W}) = 0$ if and only if $\frac{d\vec{W}}{dt} = 0$. This confirms the fact that $L(\vec{W})$ is a Lyapunov function in this regime and thus applies also for weakly interacting memristors.

We are interested in the more general case of arbitrary Ω. For this purpose, we consider the following bounds. We have

$$
-||\frac{d\vec{W}}{dt}||^2_{I+\xi\Omega W} + \frac{d\vec{W}}{dt}\cdot\vec{M} \leq -||\frac{d\vec{W}}{dt}||^2 + ||\frac{d\vec{W}}{dt}||\cdot||\vec{M}|| = ||\frac{d\vec{W}}{dt}||\left(||\vec{M}|| - ||\frac{d\vec{W}}{dt}||\right)
\tag{A4}
$$

where we used the Cauchy-Schwarz inequality and the conservative lower bound $||\frac{d\vec{W}}{dt}||^2_{I+\xi\Omega W} \geq ||\frac{d\vec{W}}{dt}||^2$. We know already that if $\frac{d}{dt}\vec{W} = 0$, then $\frac{d}{dt}L = 0$. Let us focus on $\frac{d}{dt}\vec{W} \neq 0$. Thus, a conservative bound suggests that we need to bound it via $\sup_{\vec{W}}||\vec{M}||$ and $\min_{\vec{W}}||\frac{d}{dt}\vec{W}||$. We have shown already that if Ω_{ij} is diagonal then the proposed Lyapunov function has negative derivative, so we can make a very loose upper bound. We now have the following bound on $||\vec{M}||$:

$$
||\vec{M}|| = 2\alpha\xi\sqrt{\sum_i\sum_{j\neq i}\sum_{j\neq k} W_i^2\Omega_{ij}W_j\Omega_{ik}W_k} \leq 2\alpha\xi\sqrt{\sum_{j\neq i, k\neq i}\Omega_{jk}^2} \leq 2\alpha\xi\bar{\Omega}N,
\tag{A5}
$$

where $\bar{\Omega} = \sup|\Omega_{ij}|$ is the maximum absolute value of the elements of Ω_{ij}. On the other hand, we have

$$
\begin{aligned}
\min_{\vec{W}}||\frac{d}{dt}\vec{W}|| &= \min_{\vec{W}}||\alpha\vec{W} - \frac{1}{\beta}(I + \xi\Omega W)^{-1}\Omega\vec{S}|| \\
&= \min_{\vec{W}}\sqrt{\alpha^2||\vec{W}||^2 + \frac{1}{\beta^2}||(I + \xi\Omega W)^{-1}\Omega\vec{S}||^2 - 2\frac{\alpha}{\beta}\vec{W}\cdot(I + \xi\Omega W)^{-1}\Omega\vec{S}} \\
&= \sqrt{\alpha^2\min_{\vec{W}}||\vec{W}||^2 + \frac{1}{\beta^2}\min_{\vec{W}}||(I + \xi\Omega W)^{-1}\Omega\vec{S}||^2 - 2\frac{\alpha}{\beta}\max_{\vec{W}}\vec{W}\cdot(I + \xi\Omega W)^{-1}\Omega\vec{S}} \\
&= \sqrt{\frac{1}{\beta^2}||(I + \xi\Omega)^{-1}\Omega\vec{S}||^2 - 2\frac{\alpha}{\beta}\max_{\vec{W}}\vec{W}\cdot(I + \xi\Omega W)^{-1}\Omega\vec{S}}.
\end{aligned}
\tag{A6}
$$

Now, since Ω is a projector, we have $(I + \xi\Omega)^{-1}\Omega = \sum_{j=0}^\infty(-1)^j\xi^j\Omega^{j+1} = \sum_{j=0}^\infty(-1)^j\xi^j\Omega = \frac{1}{1+\xi}\Omega$. Moreover,

$$
\max_{\vec{W}}\vec{W}\cdot(I + \xi\Omega W)^{-1}\Omega\vec{S} \leq \frac{N}{(I + \xi)^{-1}}||\Omega\vec{S}||
\tag{A7}
$$

and thus

$$min_{\vec{W}}||\frac{d}{dt}\vec{W}|| = \frac{1}{\sqrt{1+\xi}}\sqrt{\frac{1}{\beta^2}||\Omega\vec{S}||^2 - 2\frac{\alpha}{\beta}N||\Omega\vec{S}||}. \tag{A8}$$

We thus find that if

$$2\alpha\xi\bar{\Omega}N - \frac{1}{\sqrt{1+\xi}}\sqrt{\frac{1}{\beta^2}||\Omega\vec{S}||^2 - 2\frac{\alpha}{\beta}N||\Omega\vec{S}||} < 0 \tag{A9}$$

the Lyapunov function we introduced has negative derivative always. Let us now define $||\Omega\vec{S}||^2 = N^2 s^2(N)$. We can rewrite the inequality as

$$4\xi^2(1+\xi)\bar{\Omega}^2 < \frac{s(N)^2}{\alpha^2\beta^2} - 2\frac{s(N)}{\alpha\beta}. \tag{A10}$$

Because $s(N) \sim O(1)$ in the physical case (e.g., the single voltage element does not scale with the size of the system), the bound above gives a meaningful relationship between the nonlinearity parameter ξ and the mean field dynamical properties of the system, $\frac{s}{\alpha\beta}$, which had been already introduced in Reference [15] in a mean field toy model of memristive dynamics. It is interesting to note that if we define $\frac{s}{\alpha\beta} = q$, the equation above forms a parabola and an inequality of the form

$$-q^2 + 2q + c < 0, \tag{A11}$$

which characterizes the area below the curve. Thus, via the bound we see that below the parabola the system is going into a minimum of the Lyapunov function.

This shows what we had anticipated, for example, the fact that the Lyapunov function we defined has a negative derivative and because we have that $\frac{d}{dt}L(W) = 0 \leftrightarrow \frac{d}{dt}\vec{W} = 0$ we have the required property for ξ small enough. Also, being $L(W)$ defined on a compact and being smooth, we know it must be bounded from below, which concludes the proof. Clearly, how proof has a main fallacy: it applies to continuous dynamical systems with continuous derivatives. If certain memristors reach the boundary values, discontinuities appear in the dynamics. Simulations however show that the Lyapunov functional we obtained is an upper bound to the one obtained numerically and including the constraints.

Appendix A.2. Variation

Let us now calculate the variation of the functional of Equation (A4). Again, we consider the functional:

$$\begin{aligned}
L(\vec{W}) &= -\frac{\alpha}{2}\sum_i W_i^2 - \frac{\alpha\xi}{3}\sum_i \Omega_{ii}W_i^3 \\
&\quad - \alpha\xi\sum_{j\neq i}\Omega_{ji}W_jW_i^2 + \frac{1}{\beta}\sum_{ji}W_j\Omega_{ji}S_i
\end{aligned} \tag{A12}$$

We have that:

$$\begin{aligned}
\delta_{W_j}L(W) &= \delta_{W_j}\Big(-\frac{\alpha}{2}\sum_i W_i^2 - \frac{\alpha\xi}{3}\sum_i \Omega_{ii}W_i^3 \\
&\quad - \alpha\xi\sum_{i\neq j}\Omega_{ji}W_jW_i^2 + \frac{1}{\beta}\sum_{ji}W_i\Omega_{ji}S_i\Big) \\
&= -\alpha W_j - \alpha\xi\Omega_{jj}W_j^2 \\
&\quad - \alpha\xi\sum_{i(i\neq j)}\Omega_{ji}W_i^2 + \frac{1}{\beta}\Omega_{ji}S_i - 2\alpha\xi\sum_{i(i\neq j)}W_i\Omega_{ij}W_j
\end{aligned} \tag{A13}$$

which implies

$$
\begin{aligned}
\delta_{\vec{W}} L(\vec{W}) &= -\alpha \vec{W} - \alpha \xi \Omega \vec{W}^2 + \frac{1}{\beta} \Omega \vec{S} - (I + \xi \Omega W)\frac{d}{dt}\vec{W} + \vec{M} \\
&= -\left((I + \xi \Omega W)\alpha \vec{W} - \frac{1}{\beta}\Omega \vec{S} \right) - (I + \xi \Omega W)\frac{d}{dt}\vec{W} + \vec{M} \quad \text{(A14)} \\
&= -(I + \xi \Omega W)\frac{d}{dt}\vec{W} + \vec{M}
\end{aligned}
$$

where on the last line we used the equations of motion and defined $M_i = -2\alpha\xi \sum_{j(i \neq j)} W_j \Omega_{ji} W_i$. For random networks, This mapping is reminiscent of the case of continuous neuronal networks introduced by Little [17] and then Hopfield in a series of important papers [18,19]. The Little-Hopfield model has sparked interest from the Statistical Physics community since the very beginning [20]. In the past years this particular line of study has been subject of scrutiny by many experimental groups [21–23]. The key difference is that in each case, these were studied in conjunction with ordinary and active (CMOS) electronic components to build Hopfield learning networks. We argue instead that memristive circuits *per se* form a special kind of Hopfield network defined by the Lyapunov function above, without the need of extra components.

Appendix B. Complexity of the Lyapunov Functional via Kac-Rice Formula

This section provides the details for the average number of stationary points of the Lyapunov functional above in the asymptotic regime. Let us now consider the following average:

$$
\langle \# \rangle = \left\langle \sum_{\{\vec{w}=1,0\}} \det(\partial^2_{\partial w_i \partial w_j} L(\vec{w})) \prod_i \delta\left(\partial_{w_i} L(\vec{w}) \right) \right\rangle_{P(\Omega)} \quad \text{(A15)}
$$

where $L(\vec{w})$ is the asymptotic functional in Equation (A4). We can see that

$$
\begin{aligned}
\partial_{\partial w_i} L(\vec{w}) &= -\left(\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij} S_j \right) \\
&\quad + \alpha\xi \sum_j \left(\Omega_{ij} - \Omega_{ii}\delta_{ij} \right) w_j
\end{aligned} \quad \text{(A16)}
$$

and thus

$$
\partial^2_{\partial w_i \partial w_j} L(\vec{w}) = \alpha\xi \left(\Omega_{ij} - \Omega_{ii}\delta_{ij} \right) \quad \text{(A17)}
$$

We are interested in an asymptotic upper bound on the number of local minima of the function $L(w)$. We first proceed by calculating the determinant. If N is the number of memristors, then $\det(\alpha\xi (\Omega_{ij} - \Omega_{ii}\delta_{ij})) = (\alpha\xi)^N \det(\Omega_{ij} - \Omega_{ii}\delta_{ij})$. In the scaling observed above, we have noted that $\Omega_{ii} \approx 1 - \frac{\sqrt{3}}{\sqrt{N}}$, meanwhile $\Omega_{ij} \approx \frac{\sqrt{3}}{\sqrt{N}} Q_{ij}$. Using this scaling, we have: $\det\left(\Omega_{ij} - \Omega_{ii}\delta_{ij} \right) \approx \frac{3^{\frac{N}{2}}}{N^{\frac{N}{2}}} \det\left(Q_{ij} - (\sqrt{N} - \sqrt{3})\delta_{ij} \right) \approx_{N \gg 1} \det\left(Q_{ij} - \sqrt{N}\delta_{ij} \right)$, with Q a projector which can be written as

$$
Q_{ij} = \sum_{l=1}^{L} v_i^l v_j^{l^t}. \quad \text{(A18)}
$$

We now use the formula:

$$
\det(v_i^l v_j^{l^t} + A) = (1 + v^{l^t} A^{-1} v)\det(A) \quad \text{(A19)}
$$

if A is invertible and write:

$$
\det\left(Q_{ij} - \sqrt{N}\delta_{ij} \right) = N^{\frac{N}{2}} \prod_{l=1}^{L} (1 + v^{l^t} A_l^{-1} v) \quad \text{(A20)}
$$

with $A_l = \sum_{k=l+1}^{L} v_i^k v_j^{kt} - \sqrt{N}\delta_{ij}$. Since $\sum_{k=l+1}^{L} v_i^k v_j^{kt}$ is also a projector on the subspace, we have:

$$A_l^{-1} = \frac{1}{\sqrt{N}(1-\sqrt{N})} \sum_{k=l+1}^{L} v_i^k v_j^{kt} - \frac{1}{\sqrt{N}} I. \tag{A21}$$

It is now easy to see that since $v^l \cdot v^{l'} = 0$ if $l \neq l'$, we have that

$$v^{l^t} A_l^{-1} v = -\frac{1}{\sqrt{N}},$$

and thus we have the remarkable result that in the limit described above, we have

$$\begin{aligned}
\det\left(Q_{ij} - \sqrt{N}\delta_{ij}\right) &= \frac{3^{\frac{N}{2}}}{N^{\frac{N}{2}}} \left(1 - \frac{1}{\sqrt{N}}\right)^L N^{\frac{N}{2}} \\
&= 3^{\frac{N}{2}} \left(1 - \frac{1}{\sqrt{N}}\right)^L
\end{aligned} \tag{A22}$$

which is independent from the distribution of Ω's elements but only on the scaling we used. We thus obtain a rough scaling of the number of minima, given by:

$$\langle \# \rangle = (\sqrt{3}\alpha\xi)^N \left(1 - \frac{1}{\sqrt{N}}\right)^L Z(\vec{S},\Omega) \tag{A23}$$

Up to here no assumption has been made on the distribution of Ω_{ij}, if not the scaling observed numerically.

In order to evaluate $Z(\vec{S},\Omega)$, however, assumptions have to be made. We now focus on evaluating the term:

$$Z(\vec{S}, P(\Omega)) = \langle \sum_{\{\vec{w}=1,0\}} \prod_i \delta\left(\partial_{w_i} L(\vec{w})\right) \rangle_{P(\Omega)} \tag{A24}$$

which can be written as:

$$Z(S, P(\Omega)) = \frac{1}{(2\pi)^N} \int \prod_{ij} d\Omega_{ij} \sum_{\{w\}} \prod_i \int_{-\infty}^{\infty} d\eta_i e^{i\eta_i\left(-\left(\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij}S_j\right) + \alpha\xi \sum_j (\Omega_{ij} - \Omega_{ii}\delta_{ij})w_j\right)} P(\Omega_{ij}). \tag{A25}$$

We will now introduce again the scaling $\Omega_{ii} = 1 - \frac{1}{\sqrt{N}}$ and $\Omega_{ij} = \frac{\sqrt{3}}{\sqrt{N}}Q_{ij}$. We now consider the case in which $P(Q_{ij}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{Q_{ij}^2}{2\sigma^2}}$. We can use a this point the formula:

$$\int_{-\infty}^{\infty} dx\, e^{iqx - \frac{x^2}{2b^2}} = \sqrt{2\pi b^2} e^{-\frac{b^2 q^2}{2}} \tag{A26}$$

and write:

$$\begin{aligned}
Z(S, P(\Omega)) &= \frac{1}{(2\pi)^N} \int \prod_{ij} d\Omega_{ij} \sum_{\{w\}} \prod_i \int_{-\infty}^{\infty} d\eta_i e^{i\eta_i\left(-\left(\frac{\alpha}{2} + \frac{\alpha\xi}{3}\Omega_{ii} - \frac{1}{\beta}\sum_j \Omega_{ij}S_j\right) + \alpha\xi \sum_j (\Omega_{ij} - \Omega_{ii}\delta_{ij})w_j\right)} P(\Omega_{ij}) \\
&= \frac{1}{(2\pi)^N} \sum_{\{w\}} \prod_i \int_{-\infty}^{\infty} d\eta_i e^{i\eta_i\left(\alpha\xi(1-\frac{\sqrt{3}}{\sqrt{N}})w_j - (\frac{\alpha}{2} + \alpha\xi(1-\frac{\sqrt{3}}{\sqrt{N}}))\right)} \\
&\quad \cdot \int_{-\infty}^{\infty} \prod_{ij} dQ_{ij} e^{i\frac{\sqrt{3}}{\sqrt{N}}\eta_i Q_{ij}\left(\alpha\xi w_j + \frac{1}{\beta}S_j\right) - \frac{Q_{ij}^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2}} \\
&= \frac{1}{(2\pi)^N} \sum_{\{w\}} \prod_i \int_{-\infty}^{\infty} d\eta_i e^{i\eta_i\left(\alpha\xi(1-\frac{\sqrt{3}}{\sqrt{N}})w_j - (\frac{\alpha}{2} + \alpha\xi(1-\frac{\sqrt{3}}{\sqrt{N}}))\right)} e^{-N\frac{\eta_i^2\sigma^2\left(\alpha\xi w_j + \frac{1}{\beta}S_j\right)^2}{3}{2}}
\end{aligned}$$

and if we define $a(w) = \alpha\xi(1 - \frac{\sqrt{3}}{\sqrt{N}})w_j - \left(\frac{\alpha}{2} + \alpha\xi(1 - \frac{\sqrt{3}}{\sqrt{N}})\right)$ and $b_i(w) = \left(\alpha\xi w + \frac{1}{\beta}S_i\right)$, we can write the integral as:

$$
Z(S, P(\Omega)) = \left(\frac{1}{2\pi}\frac{\sqrt{6\pi}}{\sqrt{N\sigma^2}}\right)^N
$$

$$
\cdot \prod_i \left(\frac{e^{-\frac{3a(1)^2}{2N\sigma^2 b_i(1)^2}}}{b_i(1)} + \frac{e^{-\frac{3a(0)^2}{2N\sigma^2 b_i(0)^2}}}{b_i(0)}\right) \tag{A27}
$$

We now rescale $\sigma^2 \to \sigma^2/N$ in order for σ to be physical in the limit $N \to \infty$. We thus obtain, in the approximation in which $S_i = S$, $Z(S)$ becomes:

$$
Z(S) = \left(\frac{\sqrt{3}}{\sqrt{\pi\sigma^2}}\right)^N \left(\frac{e^{-\frac{3a(1)^2}{2\sigma^2 b(1)^2}}}{b(1)} + \frac{e^{-\frac{3a(0)^2}{2\sigma^2 b(0)^2}}}{b(0)}\right)^N . \tag{A28}
$$

In the limit $\alpha\xi \gg S \gg 1$, we have $\frac{e^{-\frac{3a(1)^2}{2\sigma^2 b(1)^2}}}{b(1)} + \frac{e^{-\frac{3a(0)^2}{2\sigma^2 b(0)^2}}}{b(0)} \approx \frac{1}{\alpha\xi}$. Thus, we have the final formula:

$$
\langle\#\rangle \approx \left(1 - \frac{1}{\sqrt{N}}\right)^L (\frac{\sqrt{3}}{\sqrt{\pi}\sigma})^N. \tag{A29}
$$

which shows two regimes. If $\sigma < \sqrt{3/\pi} \approx 0.977205$, the number of stationary points increases exponentially.

References

1. Indiveri, G.; Liu, S.-C. Memory and information processing in neuromorphic systems. *Proc. IEEE* **2015**, *103*, 1379–1397. [CrossRef]
2. Avizienis, A.V.; Sillin, H.O.; Martin-Olmos, C.; Shieh, H.H.; Aono, M.; Stieg, A.Z.; Gimzewski, J.K. Neuromorphic Atomic Switch Networks. *PLoS ONE* **2012**, *7*, e42772. [CrossRef] [PubMed]
3. Stieg, A.Z.; Avizienis, A.V.; Sillin, H.O.; Martin-Olmos, C.; Aono, M.; Gimzewski, J.K. Emergent Criticality in Complex Turing B-Type Atomic Switch Networks. *Adv. Mater.* **2012**, *24*, 286–293. [CrossRef] [PubMed]
4. Traversa, F.L.; di Ventra, M. Universal memcomputing machines. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 2702–2715. [CrossRef] [PubMed]
5. Caravelli, F.; Hamma, A.; di Ventra, M. Scale-free networks as an epiphenomenon of memory. *Europhys. Lett. EPL* **2015**, *109*, 28006. [CrossRef]
6. Caravelli, F. Trajectories entropy in dynamical graphs with memory. *Front. Robot. AI* **2016**, *3*, 18. [CrossRef]
7. Caravelli, F.; Carbajal, J.P. Memristors for the curious outsiders. *Technologies* **2018**, *6*, 118. [CrossRef]
8. Zegarac, A.; Caravelli, F. Memristive networks: From graph theory to statistical physics. *Europhys. Lett. EPL* **2019**, *125*, 10001. [CrossRef]
9. Strukov, D.B.; Snider, G.; Stewart, D.R.; Williams, R.S. The missing memristor found. *Nature* **2008**, *453*, 80–83. [CrossRef]
10. Chua, L.O.; Kang, S.M. Memristive devices and systems. *Proc. IEEE* **1976**, *64*, 209–223. [CrossRef]
11. Yang, J.J.; Strukov, D.B.; Stewart, D.R. Memristive devices for computing. *Nat. Nanotechnol.* **2013**, *8*, 13–24. [CrossRef] [PubMed]
12. Glover, F.; Kochenberger, G.; Du, Y. A Tutorial on Formulating and Using QUBO Models. *arXiv* **2019**, arXiv:1811.11538.
13. Ohno, T.; Hasegawa, T.; Nayak, A.; Tsuruoka, T.; Gimzewski, J.K.; Aono, M. Sensory and short-term memory formations observed in a Ag$_2$S gap-type atomic switch. *Appl. Phys. Lett.* **2011**, *99*, 203108. [CrossRef]

14. Wang, Z.; Joshi, S.; Savel'ev, S.E.; Jiang, H.; Midya, R.; Lin, P.; Hu, M.; Ge, N.; Strachan, J.P.; Li, Z.; et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nat. Mater.* **2017**, *16*, 101–108. [CrossRef] [PubMed]

15. Caravelli, F.; Barucca, P. A mean-field model of memristive circuit interaction. *Eur. Phys. Lett.* **2018**, *122*. [CrossRef]

16. Caravelli, F. Locality of interactions for planar memristive circuits. *Phys. Rev. E* **2017**, *96*, 052206. [CrossRef]

17. Little, W.A. The existence of persistent states in the brain. *Math. Biosci.* **1974**, *19*, 101–120. [CrossRef]

18. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [CrossRef]

19. Hopfield, J.J.; Tank, D.W. Computing with neural circuits: A model. *Science* **1986**, *233*, 625–633. [CrossRef]

20. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Spin-glass models of neural networks. *Phys. Rev. A* **1985**, *32*, 2. [CrossRef]

21. Hu, S.G.; Liu, Y.; Liu, Z.; Chen, T.P.; Wang, J.J.; Yu, Q.; Deng, L.J.; Yin, Y.; Hosaka, S. Associative memory realized by a reconfigurable memristive Hopfield neural network. *Nat. Commun.* **2015**, *6*, 7522. [CrossRef]

22. Kumar, S.; Strachan, J.P.; Williams, R.S. Chaotic dynamics in nanoscale NbO_2 Mott memristors for analogue computing. *Nature* **2017**, *548*, 318–321. [CrossRef]

23. Tarkov, M. Hopfield Network with Interneuronal Connections Based on Memristor Bridges. In Proceedings of the 13th International Symposium on Neural Networks, ISNN 2016, St. Petersburg, Russia, 6–8 July 2016; pp. 196–203.

24. Adler, R.J.; Taylor, J.E. *Random Fields and Geometry*; Springer: New York, NY, USA, 2007.

25. Fyodorov, Y.V. High-Dimensional Random Field and Random Matrix Theory. *Markov Process. Relat. Fields* **2015**, *21*, 483–518.

26. Caravelli, F. The mise en scene of memristive networks: effective memory, dynamics and learning. *Int. J. Par. Dist. Syst.* **2018**, *33*, 350–366. [CrossRef]

27. Parisi, G. A sequence of approximate solutions to the S-K model for spin glasses. *J. Phys. A* **1980**, *13*, L-115. [CrossRef]

28. Poloni, F. *Algorithms for Quadratic Matrix and Vector Equations*; Theses (Scuola Normale Superiore); Edizioni della Normale: Pisa, Italy, 2011.

29. Kavitha, T.; Liebchen, C.; Mehlhorn, K.; Michail, D.; Rizzi, R.; Ueckerdt, T.; Zweig, K.A. Cycle bases in graphs characterization, algorithms, complexity, and applications. *Comput. Sci. Rev.* **2009**, *3*, 199–243. [CrossRef]

30. Chang, T.J.; Meade, N.; Beasley, J.E.; Sharaiha, Y.M. Heuristics for cardinality constrained portfolio optimisation. *Comput. Oper. Res.* **2000**, *27*, 1271–1302. [CrossRef]

31. 225 Nikkei Asset Dataset, OR-Library. Available online: http://people.brunel.ac.uk/~mastjjb/jeb/orlib/files/port5.txt (accessed on 12 August 2019).

32. Markowitz, H. Portfolio Selection. *J. Financ.* **1952**, *7*, 77–91.

33. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by Simulated Annealing. *Science* **1983**, *220*, 671–680. [CrossRef]

34. Caravelli, F.; Traversa, F.L.; di Ventra, M. The complex dynamics of memristive circuits: Analytical results and universal slow relaxation. *Phys. Rev. E* **2017**, *95*, 2. [CrossRef]

35. Traversa, F.L.; Ramella, C.; Bonani, F.; di Ventra, M. Memcomputing NP-complete problems in polynomial time using polynomial resources and collective states. *Sci. Adv.* **2015**, *1*, e1500031. [CrossRef]

36. Pershin, Y.V.; di Ventra, M. Solving mazes with memristors: A massively-parallel approach. *Phys. Rev. E* **2011**, *84*, 046703. [CrossRef]

37. Pershin, Y.V.; di Ventra, M. Self-organization and solution of shortest-path optimization problems with memristive networks. *Phys. Rev. E* **2013**, *88*, 013305. [CrossRef]

38. Dorigo, M.; Gambardella, L.M. Ant colonies for the traveling salesman problem. *Biosystems* **1997**, *43*, 73–81. [CrossRef]