Self-adaptation of networks of nonidentical pulse-coupled excitatory and inhibitory oscillators in the presence of distance-related delays to achieve frequency synchronization

L. Gil

Université Côte d'Azur, Institut de Physique de Nice, 17 rue Julien Lauprêtre, 06200 Nice, France

(Received 27 March 2023; accepted 1 September 2023; published 21 September 2023)

We show that a network of nonidentical nodes, with excitable dynamics, pulse-coupled, with coupling delays depending on the Euclidean distance between nodes, is able to adapt the topology of its connections to obtain spike frequency synchronization. The adapted network exhibits remarkable properties: sparse, anticluster, necessary presence of a minimum of inhibitory nodes, predominance of connections from inhibitory nodes over those from excitatory nodes, and finally spontaneous spatial structuring of the inhibitory projections: the furthest are the most intense. In a second step, we discuss the possible implications of our findings to neural systems.

DOI: 10.1103/PhysRevE.108.034211

I. INTRODUCTION

The properties and dynamics of a set of interacting elementary constituents is a very old subject of study in Physics. The elementary constituents are often identical, the connections at short distance, the geometry of low dimensionality and often homogeneous in such a way that the complete problem presents many symmetries and conserved quantities which greatly facilitate its study. None of this persists when one is interested in sets of biological cells, scientific collaborators, computing grids, water, electricity, travelers distribution networks, internet, or the spread of epidemics or opinions [1–3]. These networks are indeed made up of very inhomogeneous elements, with wide distribution of degrees of connectivity, often without any characteristic length of interaction. In addition to their ubiquity, these networks are also adaptive, characterized by their own dynamics of the weights of internode connections.

Influenced by Kuramoto's pioneering work in 1974, much of the literature has focused on the paradigmatic problem of synchronizing a network of coupled nonidentical oscillators. The existence of phase transition [4–6], of organization in modular structures, of clusters and chimera states has then been reported [7,8]. In particular, it has been shown that networks were able to self-organize to obtain global synchronization at a lower cost than that obtained by all-to-all homogeneously connected networks. Self-organization thus makes it possible to extract a better collective advantage from the specificities of each individuality [9–14].

However, the vast majority of this work concerns networks of phase oscillators (as opposed to excitable oscillators) coupled by a smooth, regular function of the phase difference (as opposed to pulse-coupled), and this despite the possible great importance of the latter type of network for understanding neural systems. This is because phase synchronization of excitable pulse-coupled oscillators in the presence of delays proportional to their distance is a geometric frustration problem not admitting a solution in general. The basic idea is as follows: for nodes A and B to train each other to spike in phase, the delay τ_{AB} between them must be an exact multiple of their interspike interval (ISI). Similarly, for B and C to spike in phase, τ_{BC} must be a multiple of ISI. But, unless you consider a very particular geometry (such as the one used in Refs. [15–18] where nodes are arranged on a circle and distances are measured along this circle), τ_{AC} is in general not a whole number of times ISI and therefore the spike of A participates in desynchronizing C. Along this argument, the pulse aspect of the coupling, i.e., the existence of an interaction only during a very short time interval compared to ISI, is fundamental. Indeed, the further the coupling is from a Dirac distribution, the less the proportionality relation between delay and ISI is constraining (see Appendix C).

Since phase synchronization is highly unlikely, what other forms of self-organization can be expected? Detailed numerical simulation of neural network models report the existence of frequency synchronization regimes: Ref. [19] simulates the activity of 10^5 neurons and 8.5×10^6 synaptic contacts randomly distributed on the surface of a sphere of radius 8 mm with submillisecond time resolution. The neurons interact via both local and long-distance connections. The ratio of excitatory to inhibitory neurons is 4/1. Neurons, both excitatory and inhibitory, are not identical and the parameters that describe their dynamics in the absence of coupling are randomly distributed around a mean value. Short-term depression and facilitation are taken into account through Markram's [20] phenomenological description of short-term synaptic plasticity. Long-term synaptic plasticity is taken into account through spike-timing dependent plasticity [21]. The main result of this numerical experiment is the observation of spontaneous self-organization of neurons into groups and repeatedly generated patterns of activity with millisecond precision of spike timing (in agreement with experimental observations [22]).

Networks of excitable pulse-coupled oscillators with delays proportional to distance can therefore be frequency synchronized. The tricky part is that these frequency-synchronized structures were obtained when we weren't really looking for them, since only biological mechanisms known to play a role in learning were involved in the network's adaptation dynamics. As a result, the network frequency synchronization could appear as a secondary effect of learning, possibly even fortuitous, but not as the main objective of the

network dynamics. Experimental observations do not resolve the ambiguity, as they focus mainly on proving the role of connections in network synchronization [23,24], with little evidence of the relationship between synchronization and learning.

Here we are not going to remove this ambiguity either, but simply bring a new element of reflection by approaching the problem from the opposite side: instead of starting from biological learning mechanisms and observing that they lead to a certain structural organization of networks and to synchronization, we are going to impose frequency synchronization (i.e., frequency synchronization is the stated goal and we treat it as an optimization problem) and report on the way the network self-organizes to achieve this. Our faraway hope is that subsequent comparisons between the theoretical organizations (for the purpose of learning and for the purpose of synchronization) and the effective organization of certain neural systems may help to resolve the ambiguity.

To our purpose, we use a greedy algorithm where each node modifies the weight of its incoming connections to best adjust its ISI to an external and common set point $\mathrm{ISI}_{\mathrm{sp}}$. Our mains results are:

(1) the frequency synchronization requires the mandatory presence of a minimum percentage of inhibitory nodes among excitatory ones.

(2) the nodes that spike at the same time and constitute repeatedly generated patterns of activity actually form anticlusters. This means that almost all of the connection weights are associated with interpattern links, while the mass of intrapattern connections is almost vanishing.

(3) During the adaptation process, the statistics of the connection weights converge to a log-normal distribution. The weight of outgoing connections from inhibitory nodes is significantly larger than would be expected if the weights were randomly distributed among the nodes. Those from the inhibitory nodes are on the contrary significantly less numerous. Moreover, we observe the spontaneous occurrence of a spatial structuring where the weight of the outgoing connections is greater and deviates all the more from the random distribution as the distance between the nodes is greater.

The study plan is as follows: First, the excitable dynamics model used will be described and the synchronization algorithm and its consequences on the network dynamics will be presented. The convergence of the algorithm will then be checked numerically. In a second step we present our results: (i) necessity of a minimum percentage of inhibitors, (ii) occurrence of death amplitude in the presence of a high percentage of inhibitors, (iii) the formation of anticlusters, and (iv) spatial distribution of the weights of the connections as a function of the distances and the excitatory-inhibitory nature of the connections. Finally, as the ingredients of our model have been chosen sparingly (excitability, distance-dependent delays, pulse-coupled, weighted summation of neighbor influence, excitatory or inhibitory projection, and the quest for frequency synchronization), the possible implications of our general results to genuine neural networks will be considered. In particular, we will discuss how our results are articulated with the Izhikevich's polychronization conjecture [25].

II. THE MODEL

A. Neuronal dynamics

To model a network of N pulse-coupled excitable oscillators, we use a point process framework [26]. The benefits of such a choice are multiple:

(1) The intrinsically probabilistic nature of the dynamics. We obtain a Poisson's distribution of interspikes interval for an isolated neuron without any effort.

(2) The perfect control of the dynamics of a neuron. The temporal evolution of an isolated neuron requires the integration of neither a dynamic system nor the computation of a nonlinear mapping but just corresponds to a shift in the state space.

(3) And above all a remarkable efficiency and speed of execution. The algorithm does not converge all the time, and even when it does, it can take several tens of millions of integration steps, hence the need to go fast.

The drawbacks are the consequence of the advantages: the dynamics of an isolated neuron is highly schematized, especially compared to the diversity of possible behaviors and to the precise modeling that could be done [27].

The state of neuron i at time t ($t \in \mathbb{N}$) is described by the variable $S_i(t)$ which takes discrete values in $[-T_i^r, T^s]$. T^s and T_i^r are integer values representing, respectively, the spike and the refractory durations. The neurons are not identical because they can differ by the duration of their refractory period T_i^r. The dynamics of S_i is composed by an alternation of a deterministic and a stochastic part. The deterministic part starts at time t^* whenever $S_i(t^*) = T^s$ and continues with

$t^* + \ldots$	0	1	...	$T^s - 1$	T^s	$T^s + 1$	...	$T^s + T_i^r - 1$	$T^s + T_i^r$
S	T^s	$T^s - 1$	...	1	-1	-2	...	$-T_i^r$	0

Note that during this deterministic sequence, S_i jumps from $+1$ to -1 without passing through 0. This is because we reserve $S_i = 0$ to describe the rest state, the one reached after the refractory period. The stochastic part starts at time t^{rest} whenever $S_i(t^{\mathrm{rest}}) = 0$ and is involved in the determination of the next state $S_i(t^{\mathrm{rest}} + 1)$

$$S_i(t^{\mathrm{rest}}) = 0 \Longrightarrow S_i(t^{\mathrm{rest}} + 1) = \begin{cases} T^s & \text{with prob} \quad p_i(t^{\mathrm{rest}}) \\ 0 & \quad '' \quad\quad 1 - p_i(t^{\mathrm{rest}}) \end{cases}, \tag{1}$$

with

$$p_i(t) = \mathcal{R}\left\{ p_0 + a \sum_{j=1}^{N} D_j W_{ij} H[S_j(t - \tau_{ij})] \right\}, \tag{2}$$

where

$$R(x) = \begin{cases} 0 & \text{if } x \leqslant 0 \\ x & \text{if } 0 \leqslant x \leqslant 1, \\ 1 & \text{if } x \geqslant 1 \end{cases} \quad H(n) = \begin{cases} 1 & \text{if } n > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where $p_0 \in [0, 1]$ and $a \geqslant 0$ are constant parameters, $D_j = \pm 1$ depending on whether j is excitatory or inhibitory, $W_{ij} \geqslant 0$ is the strength of the connection from j to i, and τ_{ij} is their propagation delay proportional to their Euclidean distance. The role of the function R is to guarantee that p_i is a probability, that is a positive number in [0,1]. The pulse-coupled character of the dynamics is modeled by the function H which takes nonzero values only when the neighbors spike at the right time.

When the neuron chains spikes without discontinuity, its dynamics is periodic and the interspike interval (ISI) reaches its minimum value $\Delta = T^s + T_i^r + 1$. In our simulations, we use $p_0 = 0.001$ such that the average ISI in absence of coupling ($a = 0$) is about 10^3 time steps.

B. Network geometry

In line with our objectives, the network is free to adapt as it wishes since it is all to all connected and that the weights of the connections W_{ij} can evolve without constraints between $[0, +\infty]$. However, the spatial positions of the nodes and consequently the propagation delays are determined once and for all at the beginning of the optimization process. In what follows, we discuss this initial distribution of positions that we want to be both random but with a well-defined smaller distance between neighbors [28]. Note that the distances between nodes correspond to the usual Euclidean distances in three-dimensional space (see Appendix B).

We proceed as follows: In a first step, N neurons are randomly distributed on the surface of a sphere of radius $R = 1$. The interneuron distances vary between 0 and $2R = 2.0$ and their initial distribution is shown in Appendix B, Fig. 13.

In a second step the spatial distribution of the nodes onto the surface is regularized to homogenize their surface density. This adjustment is achieved by subjecting the node i to repulsive $\sum_j 1/r_{ij}$ interactions. The repulsive forces are applied until the quality factor of the $\min_j(r_{ij})$ distribution is equal to 30 [28]. In the end, the nodes form an almost hexagonal network [with mesh $d_{\text{hex}} \simeq \langle \min_j(r_{ij}) \rangle$], with many penta-hepta topological defects (Appendix B, Fig. 14).

As the simulation is time discretized, all the delays τ_{ij} are expressed as integer unit of cdt the distance traveled by the information during a unit of time:

$$\tau_{ij} = \left[\frac{r_{ij}}{cdt} \right], \tag{4}$$

where [ ] stands for the integer value. An important parameter is then the number of time steps necessary to transmit the information from one node to its nearest neighbor. This number is equal to $\tau_{\min} = d_{\text{hex}}/cdt$. The maximum distance being $2R = 2$, the state of all the neurons must be stored in memory over a duration of $2\tau_{\min}/d_{\text{hex}}$ time steps. Therefore, for an economical management of the memory it is better to use a small value of $\tau_{\min}$ (in most of our simulations we used $\tau_{\min} = 3$, Appendix C for how to select parameters).

III. ALGORITHM

There are no strict and rigorous rules leading to the choice of the algorithm used. Rather, it is the result of a set of general considerations, analogies and heuristic arguments that we present below. Ultimately, the main rationale is that it effectively leads to synchronized solutions.

(1) We have deliberately chosen not to use a central control capable of accepting or rejecting a solution based on a global computation. The reason is that this kind of approach quickly becomes impractical with increasing N. On the contrary, we opted for a local, scalable and parallelizable approach.

(2) Following H. A. Simons' ideas in his paper "Architecture of the complexity" [29], the nodes of the network were imposed to be unable to perform complicated mathematical computations (such as, for example, gradient computations or predictions). We just expect each oscillator to be able to estimate its ISI and to compare it with the set point ISI_{sp}.

(3) We assume that the incoming weight adjustment is not done systematically at each time step but only when the node has just spiked.

(4) A node that has just tested a new local weight configuration but which ultimately does not adopt it, cannot force the rest of the network to return to its initial state configuration. This would require too much effort in terms of storage and transport of information. The node that did the test must continue on its way. Optimization must be done on the fly.

(5) The modification of the incoming connections of node A has a direct effect on its spike frequency. On the contrary, the effect of modifying its outgoing connections is obviously more indirect: When node A acts on the incoming connections of its neighbors, then their spike frequencies are modified and may act in return on the spike frequency of A. Both approaches are possible but we will limit ourselves in the algorithm to the most efficient one, i.e., the modification of incoming connections only

(6) We have chosen not to impose any a priori structure on the connection network. Each node is connected with all the others but the weights of the connections evolves without constraint, can vanish or, on the contrary, grow indefinitely. This is a very expensive choice in terms of computing time but which is absolutely necessary to let the network freely choose its own topology.

If, at time t, node i does not spike, then its incoming connections do not change. Now if it spikes at time t, then this node starts its remodelling activity by estimating the elapsed time interval ISI between its last two spikes. Let j be another node of the network ($j \neq i$) connected to i through W_{ij}. If the last spike of j took place at a time different from $t - \tau_{ij}$, then

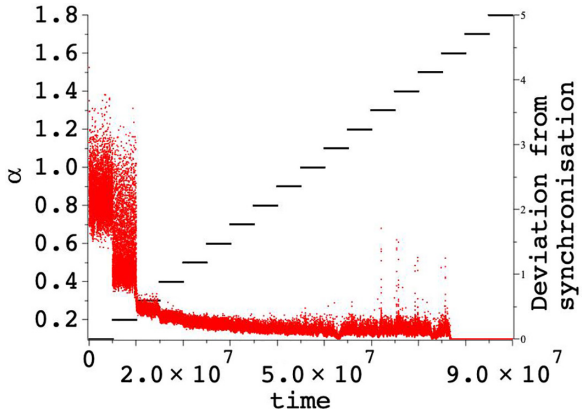

FIG. 1. Imposed time evolution of α (left axis, in black) and $G_s = \sum_{i=1}^{N}(\mathrm{ISI}_i - \mathrm{ISI}_{\mathrm{sp}})^2$ (right axis, in red) along the optimization process. The network consists in 300 nodes: for $T_r = 38$ there are 88 excitatory and 16 inhibitory nodes, for $T_r = 39$, 96, and 10 and for $T_r = 40$, 71, and 19. $T_s = 3$ and $a = 4$. The ISI set point is set at 45.

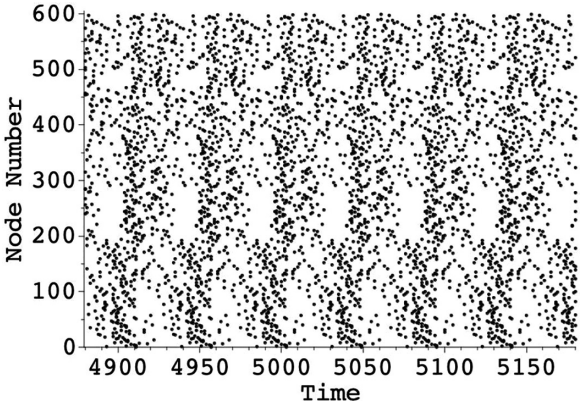

FIG. 2. Raster plot of the network activity. The configuration is [[38,[136,53],[39,[157,53]],[40,[170,31]]] and involved 600 nodes. $T_s = 3$, $p_0 = 0.001$, $b = 0.01$, $a = 4$, and $\mathrm{ISI}_{\mathrm{sp}} = 46$. The figure corresponds to the spatiotemporal dynamics after convergence of the optimization process.

i does not perceive any synaptic potential from j. It is then useless for itself to maintain the incoming connection W_{ij} and

$$W_{ij}(t+1) = W_{ij}(t)(1-b), \tag{5}$$

where b is a small positif real. On the contrary, if j spiked at $t - \tau_{ij}$, then its weight contribution is changed accordingly to

$$W_{ij}(t+1) = \max(0, W_{ij}(t) + \alpha\xi D_j(\mathrm{ISI} - \mathrm{ISI}_{\mathrm{sp}})), \tag{6}$$

where $\xi \in [0, 1]$ is a random uniformly distributed variable and $\alpha \geqslant 0$ stands for the modification amplitude. Interpretation of Eq. (6) is straightforward: if i detects that its ISI is higher than the set point (i.e., $(\mathrm{ISI} - \mathrm{ISI}_{\mathrm{sp}}) > 0$), then the incoming connections associated with inhibitory nodes ($D_j = -1$) will be decreased while those associated with excitatory nodes ($D_j = +1$) will be increased. As a result, W_{ij} and the probability for i to spike are increased. Conversely, when $\mathrm{ISI} < \mathrm{ISI}_{\mathrm{sp}}$, the same dynamics Eq. (6) leads to a decrease in the spike probability. Note that more sophisticated expressions can be considered for the weight change, but Eq. (6) can be understood as the unique linearization in the neighborhood of $\mathrm{ISI} \simeq \mathrm{ISI}_{\mathrm{sp}}$ of any mechanism imposing frequency synchronization.

It is important to realize that the algorithm is of the greedy type. Although the evolution of W_{ij} [Eq. (6)] imposes without any doubt that the ISI of node i will get closer to the set point, the simultaneous global convergence of all the nodes is absolutely not guaranteed: The convergence of a node can be done at the expense of another one.

Figure 1 represents a typical time evolution of the global deviation $G_s = \sum_{i=1}^{N}(\mathrm{ISI}_i - \mathrm{ISI}_{\mathrm{sp}})^2$ along the optimization process. While α is gradually increased by steps of 0.1, we observe a decrease of G_s to zero indicating that the system does evolve globally towards a frequency synchronization. However, this convergence is far from being uniform and takes rather the aspect of an avalanche dynamic where the local optimization of a node can provoke a cascade of events at the network level. Finally, when the global synchronization is reached, the network dynamics stops and the network does not evolve anymore.

Randomness is present in the dynamics through $p_i(t)$ [Eq. (1)] and ξ [Eq. (6)] and the initial geometrical distribution of the nodes. To investigate these effects, we perform two types of numerical experiments. All the simulations have in common the same parameter values (a, cdt, T_s, and ISI set point), the same initial W_{ij} values and they share the same distribution of refractory periods and [number of excitatory nodes, number of inhibitory nodes]: for $T_r = 38$, [80,20], for $T_r = 39$, [79,21], and for $T_r = 40$ [81, 19] (we introduce the notation [[38,[80,20]],[39,[79,21]],[40,[81,19]]] to designate such a configuration). However, the two groups differ by their initial distribution of the position of the nodes. The simulations of the first group (20 simulations) use a strictly identical geometrical distribution such that the origin of randomness is limited to $p_i(t)$ and ξ. We observe that the convergence toward a frequency synchronization regime is achieved for $\alpha > \alpha_c$ where α_c varies from one experiment to another with $\alpha_c \in [0.8, 1.2]$. Averaging over the 20 experiments, we found $\langle\alpha_c\rangle = 1.0 \pm 0.1$. Each of the simulations of the second group (10 simulations) uses its own, randomly generated, geometrical configuration. We found $\alpha_c \in [0.90, 1.60]$ with $\langle\alpha_c\rangle = 1.1 \pm 0.2$. Thus, we can see that (i) the two types of measures are consistent with each other, and (ii) the random distribution of node positions is an important source of fluctuations. Therefore, in what follows, each optimization process will be associated with a random draw of the position of the nodes.

IV. RESULTS

A. Spatiotemporal dynamics at convergence

At convergence, the spatiotemporal dynamics is characterized by the periodic succession of node patterns $\mathcal{D} = P_1, P_2 \ldots P_{\mathrm{ISI}_{\mathrm{sp}}}$ where $\mathrm{ISI}_{\mathrm{sp}}$ is the imposed interspike interval set point (Fig. 2). A pattern is constituted by the set of all nodes that spike at the same time. As the number of nodes varies from one pattern to another, the global firing rate oscillates periodically in time with the period $\mathrm{ISI}_{\mathrm{sp}}$ (Fig. 3). The patterns in the sequence $\mathcal{D}$ are 2 by 2 disjoint and their gathering constitutes the total set of nodes of the network. Therefore

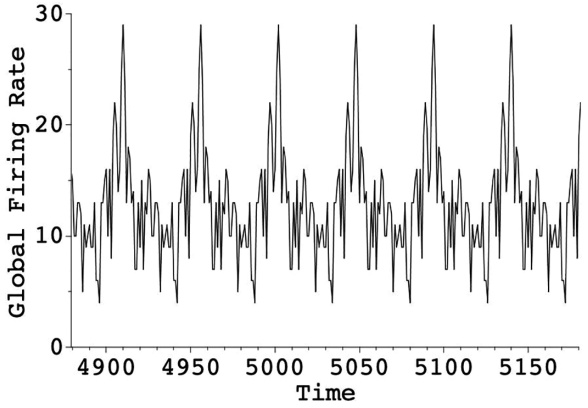

FIG. 3. Time evolution of the global firing rate associated with the raster plot in Fig. 2.

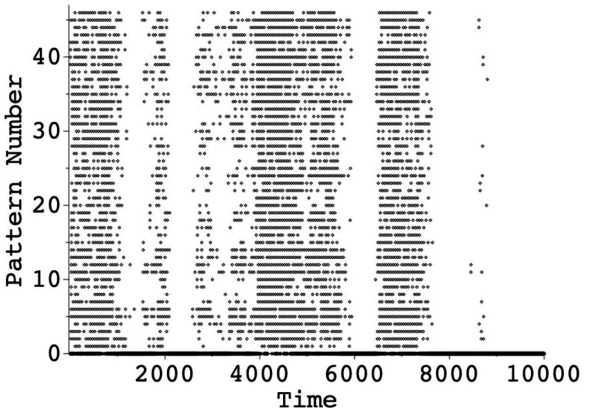

FIG. 5. Same as Fig. 4 but now $p_0 = 0.04$ such that the dynamics is strongly disrupted. As before, the numbers 1 to 46 on the vertical axis stand for the patterns $\mathcal{D} = P_1, P_2 ... P_{\mathrm{ISI_{sp}}}$, but now 0 is associated with any patterns that is not in the list $\mathcal{D}$. Pay attention to the difference in the horizontal scales: The one in Fig. 4 spans only a few $\mathrm{ISI_{sp}}$, while here it corresponds to more than 200.

they form a partition of the set of nodes. Figures 4 and 5 show typical temporal evolutions of the dynamics in the space of the patterns. On the vertical axis, the zero corresponds to any pattern that is not in the list $\mathcal{D} = P_1, P_2 ... P_{\mathrm{ISI_{sp}}}$. Figure 4 is the regular and periodic dynamics obtained after convergence of the optimization process. Figure 5 is obtained by freezing the dynamics of the network corresponding to Fig. 4 (i.e., W_{ij} are constant) and by increasing the background noise ($p_o = 0.04$). The global dynamics is found to be intermittent with phases of locking on the periodic solution at convergence, interspersed by episodes of more or less long stall with a complex dynamics.

B. Mandatory presence of inhibitory nodes and amplitude death

Numerous studies [23,24,30] have highlighted the fundamental role of inhibitory connections in the organization of neuronal systems either for the control of the excitatory network or for the fine tuning of the spike timings. This mechanism is so general that, accordingly to Ref. [23], "Synaptic inhibition itself can be synchronized by way of interactions within networks of inhibitory and excitatory neurons." It is therefore expected that our model also proves that frequency synchronization is only possible in the presence of a minimum number of inhibitory nodes.

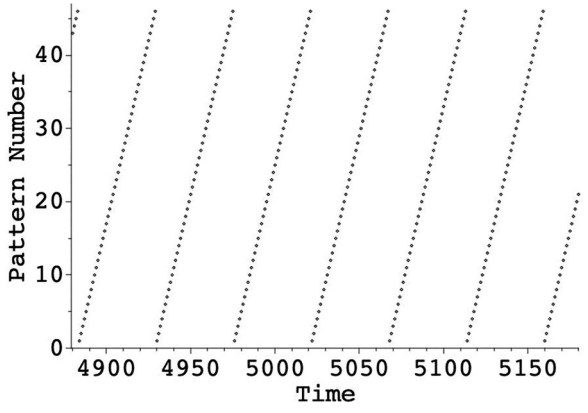

FIG. 4. Time evolution of the spatiotemporal dynamics in the space of the patterns. On the vertical axis, the numbers 1 to 46 stand for the patterns $P_1, P_2 ... P_{\mathrm{ISI_{sp}}}$ observed at the convergence of the optimization process in Figs. 2 and 3.

Each node being associated with a specific refractory period T_r, we should normally characterize a given network by its statistical distribution of T_r. Nevertheless, for the sake of simplicity, we have concretely limited ourselves to three distinct values (typically $T_r \in [38, 40]$). Tests with up to five values have been performed to check that this limitation was not relevant. The spike duration T_s being the same for all nodes, the set point for the interval between two spikes $\mathrm{ISI_{sp}}$ cannot be less than $\Delta_{\min} = T_{r_{\min}} + T_s + 1$ because our model [Eq. (1)] does not contain any mechanism capable of reducing the refractory period. However, it seems possible to impose an $\mathrm{ISI_{sp}}$ greater than $\Delta_{\max} = T_{r_{\max}} + T_s + 1$ because one expects the inhibitory neurons to cooperate to prohibit the spike over a duration longer than $T_{r_{\max}}$. Typically we impose either $\mathrm{ISI_{sp}} = \Delta_{\max} + 1$ or $\mathrm{ISI_{sp}} = \Delta_{\max} + 2$. Control simulations with $\mathrm{ISI_{sp}} = \Delta_{\max} + 5$ have been successfully performed. However, for even larger values, numerical convergence problems have been encountered.

We have conducted no less than 300 numerical experiments (Fig. 6). For each simulation, the initial position of the nodes is randomly generated. Then for each node, its value of T_r is chosen randomly and uniformly between the three values 38, 39, and 40. Finally the excitatory or inhibitory action of the node is randomly drawn: with a probability f_g the node is inhibitory, with a probability $1 - f_g$ it is excitatory. f_g changes with the experiments inside [0.05, 0.95]. For each simulation, α is increased in steps of 0.1 until a critical value α_c is reached for which a frequency synchronization regime is established. Value of α higher than 6.0 have not been investigated. Red points in Fig. 6 represents the set of (f_g, α_c) points. When several α_c are associated to the same value of f_g, it is the highest value of α_c that counts, the one that ensures the convergence towards the frequency synchronization whatever the initial geometry of the nodes and the optimization path taken. For $f_g \simeq 0$, the plot suggests a divergence of α_c associated with the impossibility of a global frequency synchronization

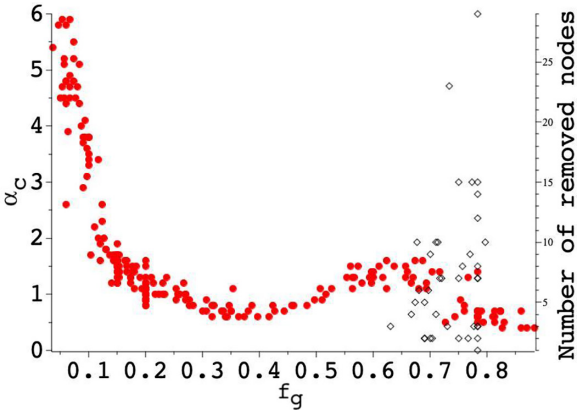

FIG. 6. The network consists of 300 nodes whose initial positions are randomly chosen on a sphere. The refractory period of each node is randomly chosen among the three values 38, 39, and 40 and its inhibitory/excitatory character is determined by drawing with a probability f_g (fraction of inhibitors). The ISI set point is 45. The red points (left axis, solid discs) stand for (f_g, α_c). The black ones (hollow diamonds) correspond to (f_g, n_{death}), where n_{death} is the number of nodes that have ceased to spike under the pressure of the inhibitory nodes along the optimization process.

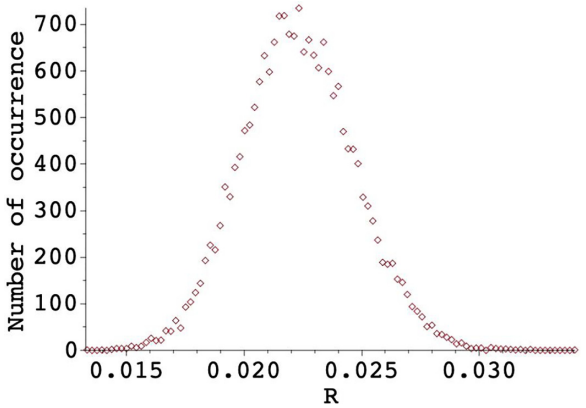

FIG. 7. Histogram of $R(\{W_{\text{rand}}\})$ defined in Eq. (7). We made 10 000 random draws and the histogram has 100 bins. The network consists of 300 nodes whose initial positions are randomly chosen on a sphere. The refractory period of each node is randomly chosen among the 3 values 38, 39, and 40 and the fraction of inhibitors $f_g = 0.15$. The ISI set point is 45 while the delay between two diametrically opposed nodes is $2/cdt = 33 < \text{ISI}_{\text{sp}}$. We find $\langle R(\{W_{\text{rand}}\})\rangle = 0.022 \pm 0.0025$ which implies that $R(\{W_{\text{cvg}}\})$ at convergence deviates from the mean value by more than 9.4 standard deviations.

in the absence of inhibitory nodes. For $f_g \simeq 1$, we observe the spontaneous death of a certain number n_{death} of nodes during the optimization process. At a given moment, under the action of their inhibitory connections, these nodes were unable to spike. And since a node that does not spike cannot change its incoming connections, the situation persists as long as the neighborhood action goes on.

C. Anticlusters structuring

For two nodes A and B to train each other to spike in phase, the delay τ_{AB} between them must be an exact multiple of the set point ISI_{sp}. Consequently, we expect and observe two very distinct operating regimes depending on whether the maximum delay between two nodes of the network $(2/cdt)$ is less or greater than ISI_{sp}.

We introduce

$$R(\{W\}) = \frac{\sum_{P_\mu \in \mathcal{D}} \sum_{i \in P_\mu, j \in P_\mu} W_{ij}}{\sum_{P_\mu \in \mathcal{D}} \sum_{P_{\mu' \neq \mu} \in \mathcal{D}} \sum_{i \in P_\mu, j \in P_{\mu'}} W_{ij}}, \qquad (7)$$

which, for a given configuration $\{W\}$, stands for the ratio between the total weight of the internal connections to each pattern P_μ and the total weight of the connections between two distinct patterns P_μ and $P_{\mu' \neq \mu}$. We compute $R(\{W_{\text{cvg}}\})$ where $\{W_{\text{cvg}}\}$ is the configuration network at the convergence of the optimization process. We compare the previous result with the distribution of $R(\{W_{\text{rand}}\})$ where $\{W_{\text{rand}}\}$ are derived from $\{W_{\text{cvg}}\}$ by randomly redistributing its weights among the nodes of the network.

We first consider the situation where $2/cdt < \text{ISI}_{\text{sp}}$ that corresponds to a "small" network where all nodes are within one interspike interval of each other. This regime correspond to the vast majority of our investigations because it is the most interesting and surprising situation. Figure 7 is a histogram of the values of $R(\{W_{\text{rand}}\})$ obtained after 10 000 draws of the random configuration $\{W_{\text{rand}}\}$. While $\langle R(\{W_{\text{rand}}\})\rangle \simeq 0.022$ with a standard deviation of 0.0025, the measured value at convergence is $R(\{W_{\text{cvg}}\}) = 1.9\,10^{-7}$, significantly smaller. It thus deviates from the random distribution by more than 9 standard deviations, which rules out any coincidence: therefore the patterns $P_\mu \in \mathcal{D}$ are characterized by a very strong anticluster structuring.

In the case of a network with $2/cdt = 100 > \text{ISI}_{\text{sp}}$, the situation is completely changed. In such a "large" network, each node can be linked to several distinct nodes shifted by exactly one ISI_{sp}. Then $R(\{W_{\text{cvg}}\})$ is no longer almost zero, but on the contrary is measured to be almost one standard deviation higher than $\langle R(\{W_{\text{rand}}\})\rangle$ (not shown). The anticluster structure is in competition with the connections between nodes belonging to the same pattern and is clearly less predominant. Figure 8 shows the connection weights repartition W_{ij} versus the delay τ_{ij} when both i and j belongs to the same given pattern P_μ (randomly chosen in the $\mathcal{D}$ sequence). We clearly observe that only internal connections with a delay equal to ISI_{sp} ot 2 ISI_{sp} are not vanishing.

D. Network sparseness

The Gini coefficient is a real number, between 0 and 1, that measures the rate of inequality of the distribution of a variable. It was originally developed in economics to measure the income inequality of a country's population. Applied to the case of connection weights, a null value of this coefficient would correspond to the homogeneous distribution of the mass, i.e., to the case where all W_{ij} are equal. On the contrary, a coefficient equal to 1 would mean that all the weights are zero, except for one and only one. For values of $f_g \simeq 0.2$ and the number N between 100 to 600 of nodes, we find a staggering value of 0.95, indicating that the optimized networks are particularly sparse with a very large majority of connections reduced to zero coexisting with a very few

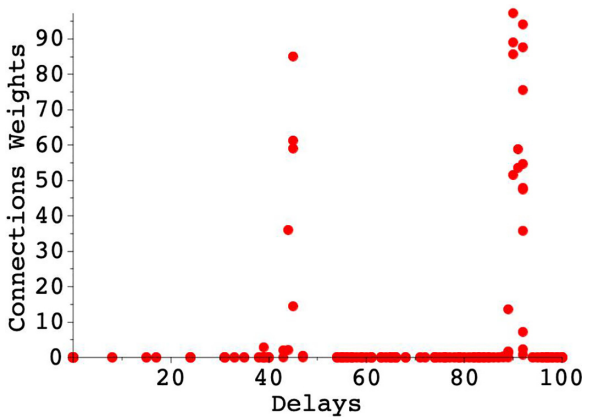

FIG. 8. The plot shows the set of points (τ_{ij}, W_{ij}) where i and j belong to the same pattern $P_\mu \in \mathcal{D}$. The networks has 600 nodes, $T_s = 3$, $a = 4$, and $p_0 = 0.001$. Their refractory periods are not identical and vary between 38 and 40. The ISI set point is fixed at 46 and $2/cdt = 100$. The fraction of inhibitory nodes is 20%. The first maximum is located at 46 (= $\mathrm{ISI}_{\mathrm{sp}}$) and the second at 92.

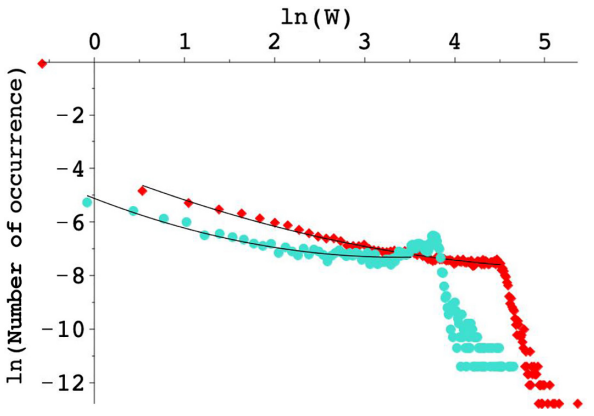

FIG. 9. Histogram with 200 bins of the weights of the connections W_{ij} in log-log scales. We integrated the results obtained for five distinct networks with $f_g = 0.2$ at convergence of the optimization process. The turquoise circles correspond to networks with 300 nodes while the red diamonds are associated with network with 600 nodes. Continuous lines are quadratic fits compatible with log-normal distributions.

number of very massive connections. Figure 9 shows a typical histogram of the connexion weights W_{ij} in log-log scales.

E. Predominance of projections from inhibitory nodes

Here we focus on the global masses of the network connections according to the excitatory or inhibitory nature of the nodes of departure and arrival. We introduce

$$P_{+\pm}(\{W\}) = \sum_{\substack{\{i \in [1,N] | D_i = +1\} \\ \{j \in [1,N] | D_j = \pm 1\}}} W_{ij} \tag{8}$$

where $P_{++}(\{W\})$ (respectively, $P_{+-}(\{W\})$) stands for the total mass of the excitatory ← excitatory connections (respectively, excitatory ← inhibitory) for the network configuration $\{W\}$. We define in the same way the other masses P_{-+} and P_{--} and we introduce the xy notation to designate any of the sign pairs $++$, $+-$, $-+$, and $--$.

The numerical values of the above quantities at the convergence of the optimization process do not have any meaning in themselves. Neither do their ratios since they depend on f_g. So we will proceed as for the demonstration of the anticluster structure in paragraph IV C, by comparing $P_{xy}(\{W_{\mathrm{cvg}}\})$ with the distribution of $P_{xy}(\{W_{\mathrm{rand}}\})$ where $\{W_{\mathrm{rand}}\}$ are derived from $\{W_{\mathrm{cvg}}\}$ by randomly redistributing its weights among the nodes of the network. The results are displayed in Table I. Since they differ from the mean values by several standard deviations, they are highly significant from a statistical point of view. They clearly shows a very net deficit in the mass of the connections from excitatory nodes, to the benefit of a substantial excess in the mass of the connections from inhibitory nodes.

F. Spatial distribution of the connexion weights

The question that interests us here is to know if there is a relationship between the propagation delay τ_{ij} between any 2 nodes i and j of the network and the weights W_{ij} (possibly W_{ji}) of their connections. For that purpose, we introduce the following definitions:

$$M_{+\pm}(\{W\}, \tau) = \sum_{\substack{\{i \in [1,N] \,|\, D_i = +1\} \\ \{j \in [1,N] \,|\, D_j = \pm 1\}}} W_{ij}\delta(\tau - \tau_{ij}), \tag{9}$$

where $\delta(n) = 1$ if $n = 0$ and cancels out for any other integer value. For the configuration $\{W\}$, $M_{++}(\{W\}, \tau)$ [respectively, $M_{+-}(\{W\}, \tau)$] is the sum of the masses of all the connections from an excitatory node to an excitatory one (respectively, inhibitory to excitatory) and separated by a propagation delay τ. By analogy, we define in the same way M_{-+} and M_{--}. We then proceed in the same way as for proving the anticluster feature of the optimized network or for proving the

TABLE I. Same configuration as in Fig. 2. P_{xy} are defined in Eq. (8). The rightmost column shows the difference between the measurements at convergence and the mean value in units of standard deviation. Undoubtedly, a large part of the mass has been allocated to the outgoing connections from the inhibitory nodes at the expense of the outgoing connections from the excitatory ones.

xy	$P_{xy}(\{W_{\mathrm{cvg}}\})$	$\langle P_{xy}(\{W_{\mathrm{rand}}\})\rangle$	standard deviation σ	$\frac{P_{xy}(\{W_{\mathrm{cvg}}\}) - \langle P_{xy}(\{W_{\mathrm{rand}}\})\rangle}{\sigma}$
$++$	557485	623876	3690	−18.0
$-+$	141147	159531	2792	−6.6
$+-$	224183	159530	2782	23.2
$--$	60525	40403	1516	13.3

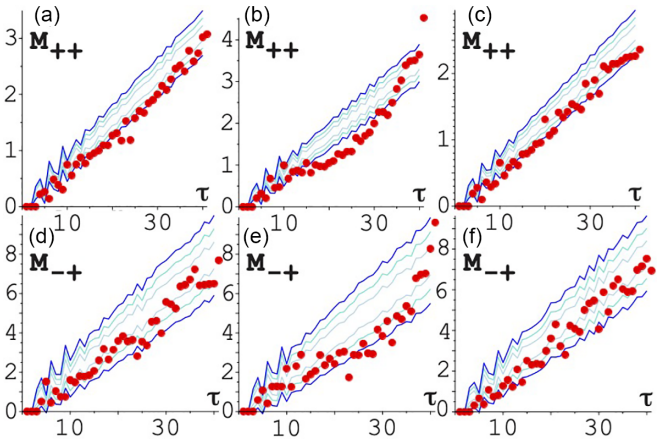

FIG. 10. The first row stands for the plot of $M_{++}(\{W\},\tau)$ versus τ while the second one with $M_{-+}(\{W\},\tau)$ versus τ. The columns correspond to the repetition of the measurement for 3 networks with 600 nodes, $T_s = 3$, $a = 4$, and $p_0 = 0.001$ but distinct random initial positions of the nodes. The refractory periods are not identical and vary between 38 and 40. The ISI set point is fixed at 46. The fraction of inhibitory nodes is 20%. Vertical axis are in arbitrary units, as the absolute value is irrelevant. Red points correspond to the converged optimized network while the numerous blue lines are associated with the random distribution of the weights among the network connections. The light, medium, and dark blue lines are, respectively, associated with mean value ±1, ±2, and ±3 standard deviations.

predominance of the inhibitory projections. We first compute $M_{xy}(\{W_{\rm cvg}\},\tau)$ where $\{W_{\rm cvg}\}$ is the configuration network at the convergence of the optimization process and then we compare the result with $M_{xy}(\{W_{\rm rand}\},\tau)$ where $\{W_{\rm rand}\}$ are derived from $\{W_{\rm cvg}\}$ by randomly redistributing its weights among the nodes of the network. The results are displayed in Figs. 10–12. For Figs. 10 and 11, the three columns correspond to the triple repetition of the numerical experiment by changing only the initial position of the nodes on the sphere. The red circles stand for the case of the optimized network $M_{xy}(\{W_{\rm cvg}\},\tau)$ while the numerous blue lines (light, medium, and dark) are

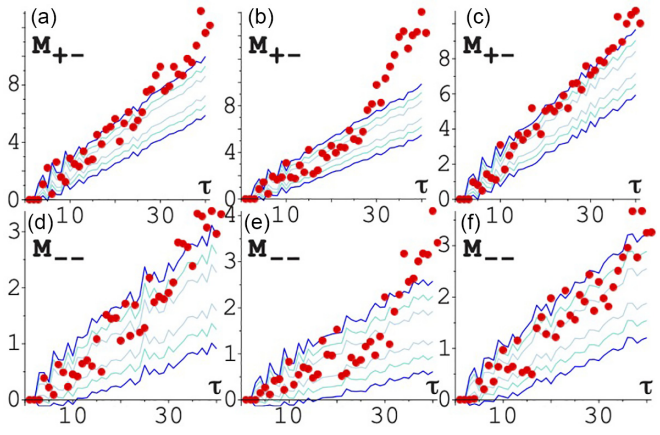

FIG. 11. Same regime of parameters as in Fig. 10, but now the first row deals with M_{+-} versus τ while the second one with M_{--} versus τ.

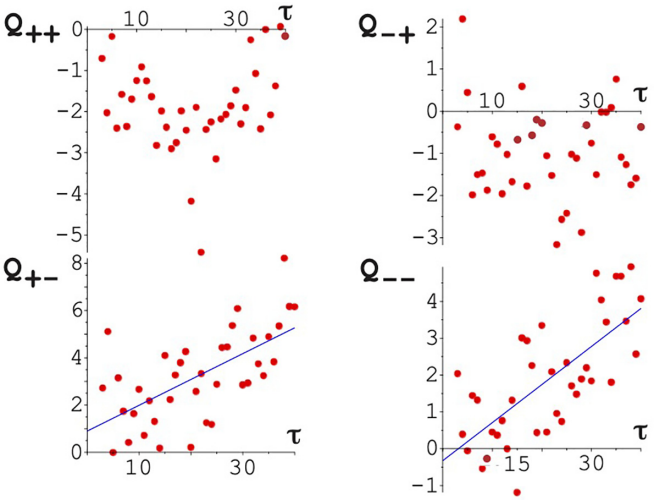

FIG. 12. Same regime of parameters as in the first column of Figs. 11 and 12. Black straight lines are linear regression fits. Q_{xy} as defined in Eq. (10), measures the deviation of $M_{xy}(\{W_{\rm cvg}\},\tau)$ from the mean value $\langle M_{xy}(\{W_{\rm rand}\},\tau)\rangle$ in units of standard deviation. Hence, on Fig. 10(a1) [respectively, Fig. 10(b1)], the fact that $M_{++}(\{W_{\rm cvg}\},\tau)$ [respectively, $M_{-+}(\{W_{\rm cvg}\},\tau)$] stays almost all the time in the blue crosses zone corresponds on the present top left plot (respectively, top right), to values of $|Q_{++}|$ (respectively, $|Q_{-+}|$) mostly smaller than 3. The situation is very different for Q_{--} and especially for Q_{+-}, where not only do the measurements deviate from the mean value by more than 3 standard deviations, but where there is also a clear upward trend materialized by the black solid line.

associated with 10 000 random draw repetitions and the respective position of the confidence intervals at ±1, ±2, and ±3 standard deviations. Some figures give the impression that the optimized values are compatible with a random configuration of the connection masses. Others, on the contrary, seem to indicate that they clearly deviate from it. To clarify the situation, we introduce

$$\mathcal{Q}_{-+} = \frac{M_{-+}(\{W_{\rm cvg}\},\tau) - \langle M_{-+}(\{W_{\rm rand}\},\tau)\rangle}{\sqrt{\langle M_{-+}(\{W_{\rm rand}\},\tau)^2\rangle - \langle M_{-+}(\{W_{\rm rand}\},\tau)\rangle^2}}, \quad (10)$$

which stands for the deviation from the mean value measured in units of standard deviation (also $\mathcal{Q}_{++}$, $\mathcal{Q}_{+-}$, and $\mathcal{Q}_{--}$) and plot it versus τ (Fig. 12). The analysis of the figures leads to the following remarks:

(1) For a given value of the delay, the values of M_{++} or M_{-+} associated with outgoing connections from excitatory nodes, do not deviate significantly (more than 3 standard deviations) from the mean value of the random distributions.

(2) Nevertheless, if for a given delay, the values of $M_{\pm+}$ were only due to chance, then from one delay to another we should observe an alternation of values larger and smaller than the average. The fact that a large majority of the values are below the mean value is statistically significant and is corroborated by the global $P_{\pm+}$ measurements.

(3) For outgoing connections from inhibitory nodes, we clearly observe that not only are $M_{\pm-}$ significantly above the random value, but also that this deviation increases with delay.

The further the inhibitory connection projects, the higher its weight.

V. DISCUSSION

A. Reminder of our results

We have just shown that a network of nonidentical nodes, with excitable dynamics, pulse-coupled, with coupling delays depending on the Euclidean distance between nodes, was able to adapt the topology of its connections to obtain spike frequency synchronization. The adapted network has the following remarkable properties:

(1) The spatiotemporal dynamics is organized in a periodic succession of patterns. A pattern is constituted by the set of nodes that spiked at the same time. The set of patterns forms a partition of the network. There are very few connections between nodes of the same pattern and the vast majority of connections concern nodes belonging to distinct patterns. This results in an anticluster structure.

(2) The network is very sparse.

(3) Inhibitory nodes play a fundamental role in frequency synchronization. Not only because frequency synchronization requires the presence of a minimum number of inhibitory nodes, but also because the total mass of outgoing connections from the inhibitory nodes is very significantly larger than if the connections were established randomly.

(4) We observe the spontaneous occurrence of a spatial organization of inhibitory nodes: The further the inhibitory connection projects, the higher its weight.

B. Comparison with neural networks

First, the need for a minimum percentage of inhibitory nodes is a result in total agreement with experimental observations [23,24]. Perhaps even more remarkable is that, as in our study, inhibitory neurons are mostly identified experimentally as playing a critical functional role not in the process of information storage, but in the temporal regulation of networks.

Second our observation of the organization of spatiotemporal dynamics into a periodic succession of patterns made up of nodes spiking together is strongly reminiscent of the spontaneous self-organization of neurons reported in numerical simulation [19] or in experimental observations [22], except that no mention is made of an anticluster structure in these publications. However, as the nodes of these structures spike at the same time and are not spatially homogenously distant, there is no reason for the brain to maintain such connections that are in fact useless. This prediction could be tested experimentally.

Unsurprisingly, the patterns differ in size and persistence from Refs. [19,22]. Indeed, whereas in our study nothing is intended to limit the quest for synchronization, there exist many biological mechanisms that could control the synchronization effort. For example, one can increase the duration of synaptic interaction to relax the geometric frustration character of the phase synchronization, introduce mechanisms to adapt the spike frequency of an isolated neuron, or simply modify the propagation times of potentials by taking into account the myelic sheaths.

C. About the Polychronization conjecture

Noting that the propagation delay between any individual pair of neurons is precise and reproducible with a submillisecond precision [31,32] and arguing that obtaining and maintaining such precision can only be understood if the spike-timing is of the highest importance for the brain, Izhikevich introduces the term Polychronization [25] to qualify such spiking activity and suggests that the periodic succession of patterns could play a crucial role in the information storage process. The idea is that the same neuron could belong to several distinct patterns, themselves belonging to different sequences of patterns. Depending on initial conditions and external stimuli, a specific sequence of patterns could then be selected. As there are many more possible patterns and sequences than neurons, the storage capacities would be gigantic.

We are not able to test this conjecture directly. Clearly, for an imposed ISI, our algorithm can converge on several possible network configurations. But once the configuration has been achieved, putting in action a new optimization process to force the frequency synchronization with an other $\mathrm{ISI}' \neq \mathrm{ISI}$ systematically erases the first configuration. Nevertheless, we can imagine several ways of getting around this problem. As the optimized network is extremely sparse, a first approach would be to freeze the small number of connections with a highest weights once and for all. Learning a new optimized configuration would then involve only those connections that have not been frozen. A second possible approach is the progressive construction of a network of optimized networks. We would proceed as follows: Consider a first network ($\mathcal{N}_1$) optimized to oscillate with ISI_1. Let P_1 be a pattern of the periodic succession. We then add to P_1 a set of new nodes (not in $\mathcal{N}_1$) to form a new network ($\mathcal{N}_2$). We then run our optimization algorithm to force $\mathcal{N}_2$ to oscillate with ISI_2, the optimization algorithm not affecting connections internal to $\mathcal{N}_1$ but acting only on connections internal to $\mathcal{N}_2$ and on those between P_1 and $\mathcal{N}_2$. As the nodes inside P_1 are not connected to each other (thanks to the anticluster structure), any ISI_2 value can be imposed on them since there are no internal propagation delays to satisfy. After convergence, the $\mathcal{N}_2$ network presents a dynamic consisting of the repetition with a ISI_2 period of a sequence of patterns forming a partition of $\mathcal{N}_2$. It is possible to force P_1 to be one of the elements of this partition so that its nodes form a single pattern of two distinct sequences of the $\mathcal{N}_1 + \mathcal{N}_2$ network, associated, respectively, with ISI_1 and ISI_2. The two sequences are simultaneously compatible when ISI_1 and ISI_2 are multiples of each other. Otherwise, it is conceivable that a stimulus acting on the nodes of P_1 will select one or the other sequence according to the frequency of the external forcing. The process can be repeated ad infinitum to create a network of networks. Works in these directions are in progress.

APPENDIX A: NONEXHAUSTIVE LIST OF RELATED WORKS

Pioneering work [33] deals with a network of identical integrate-and-fire pulse-coupled and excitatory units. The delay is not related to the distance between nodes but to

a maximum time beyond which the action of node j on node i is forgotten (i.e., reduces to zero). Two topologies are studied: a fully connected network and a two-dimensional regular mesh with local coupling. The dynamics converges to a frequency synchronized solution, where all nodes have the same ISI without spiking in unison. In Ref. [34], delays are now clearly associated with the time required for the action potential to propagate along the axon of each neuron. Identical excitatory neurons with an exponentially decreasing coupling with distance, give rise to waves (which implies a global synchronization in frequency but not in phase). Reference [35] investigated the effect of time delays on a set of two-dimensional identical excitatory oscillators. The oscillators are regularly distributed on a square grid and the interactions between oscillators A and B are delayed by an amount proportional to the distance r_{AB} between them. The weights of the connections first decrease as $1/r_{AB}$, then vanish for $r_{AB} > r_0$. The oscillators are not pulse-coupled. It is found that distance-dependent time delays induce various patterns including traveling rolls, square-like and rhombus-like patterns, spirals, and targets. Reference [36] considered the effects of distributed delays on amplitude death. Oscillators, whose amplitude must be described to eventually cancel it, are of Ginzburg-Landau type. They are identical and their dynamics are not excitable. Here the delays are not distance-dependant but are chosen randomly accordingly to a given probability distribution. It is showed that even a small spread in the delay distribution can greatly enlarge the set of parameters for which amplitude death occurs. The idea of the statistical distribution of delay was then taken up: to study a standard field model of neural excitatory and inhibitory populations [37,38], to investigate the coherent activity patterns in inhibitory, synaptically coupled, bursting Hindmarsh-Rose neurons [39], to demonstrate the widespread occurrence of dynamically maintained spike timing sequences in recurrent networks of pulse-coupled spiking neurons with large time delays [40].

Reference [41] studies the Rulkov mapping in the presence of a delay proportional to the interneuron distance and of a coupling strength proportional to the difference of the fast variables (coupling known as electrical as opposed to synaptic coupling known as pulse-coupling). The neurons are not identical, the dynamics of an isolated neuron is chaotic and the network organization allows a continuous modulation between a scale-free network with dominating long-range connections and a homogeneous network with mostly adjacent neurons connected. A time-averaged Kuramoto's order parameter (R) is measured. It is found that the most phase synchronized response ($R \simeq 0.4$) is obtained for the intermediate regime where long as well as short-range connections constitute the neural architecture.

APPENDIX B: NETWORK GEOMETRY

The position of each node i on the sphere of radius unit is given by its cartesian coordinates x_i, y_i and z_i with $x_i^2 + y_i^2 + z_i^2 = 1$. The distance r_{ij} between two nodes used to calculate the delay in information propagation is $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$. Figures 13–15 illustrate the properties of the random network of nodes after adjustment.

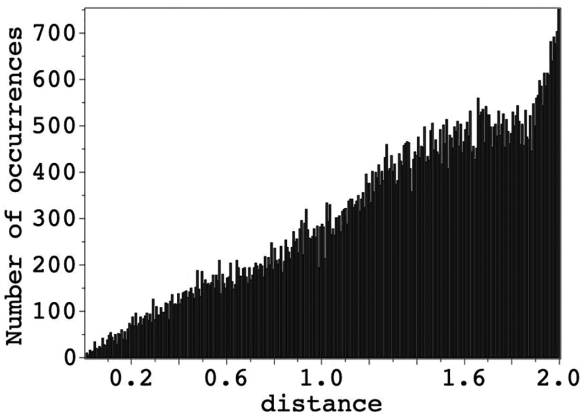

FIG. 13. Histogram of the interneurons distances before adjustment with 300 bins (vanishing distances are not taken into account). The network has 300 nodes randomly distributed on a sphere of unit radius.

APPENDIX C: GEOMETRIC FRUSTRATION

To be able to make comparisons, we first consider the classical Kuramoto's problem:

$$\partial_t \theta_i = \omega_i + \frac{1}{N}\sum_j W_{ij}\sin(\theta_i(t) - \theta_j(t - \tau_{ij}))$$
$$i \in [1, N] \tag{C1}$$

Looking for a phase synchronization regime $\theta_i(t) = \Omega t$ leads to

$$\Omega = \omega_i + \frac{1}{N}\sum_j W_{ij}\sin(\Omega\tau_{ij}). \tag{C2}$$

If we consider the N above equations [Eq. (C2)] as equations with unknown W_{ij}, then the ratio between the number of variables and the number of equations is $\zeta = N^2/N = N$.

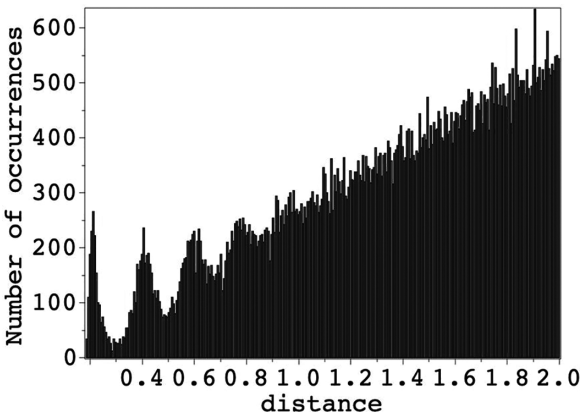

FIG. 14. Histogram of the interneurons distances after adjustment. The first peak (the farthest to the left) in the distribution is associated with $\min_j(r_{ij})$, i.e., the mesh of the hexagonal network. The ratio between the height of this peak and its width at half height defines the quality factor.

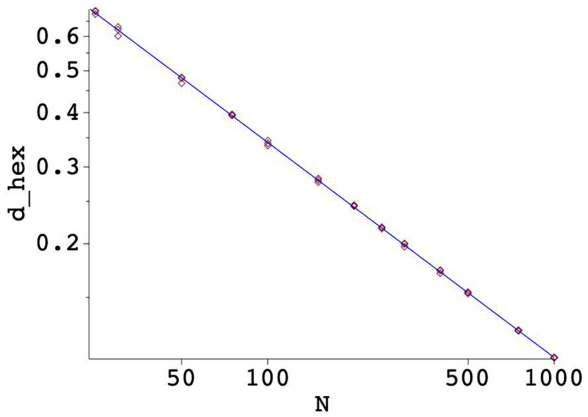

FIG. 15. Numerical investigation of the relationship between the average distance between a node and its nearest neighbor (d_{hex}) and the number N of nodes on the surface of the sphere. The blue straight line is the best power law fit $d_{\text{hex}} \simeq 3.41N^{-\frac{1}{2}}$.

This high value indicates that network configurations $\{W_{ij}\}$ with possible phase synchronization regimes are very common. Numerical simulations confirm this prediction (not shown).

The pulse-coupled regime between the nodes of a network is characterized by the ratio ξ between the spike duration and the interspike time interval (ISI) between two consecutive spikes. Our model which is discrete time is moreover characterized by the time interval dt between two time steps. This parameter does not appear explicitly, but implicitly through cdt which is the distance traveled by the information during dt [Eq. (4)]. cdt is then compared to d_{hex} to form a dimensionless number (see Appendix B, Fig. 15).

In our model, the probability for a neuron i at rest ($S_i = 0$) to spike is

$$P_i(t) = a\sum_j D_j W_{ij} H[S_j(t-\tau_{ij})], \tag{C3}$$

where $S_j(t)$ is the state of neuron j at time t, τ_{ij} is the distance induced delay between j and i, W_{ij} is the strength of the connection from j to i and $D_j = \pm 1$ depending on whether j is excitatory or inhibitory. $H(s) = 1$ if and only if $s > 0$ and vanishes otherwise. The minimal interspike interval (ISI) is then

$$\triangle = T^s + T^r + 1. \tag{C4}$$

We are looking for a phase synchronization regime with ISI $= T \geqslant \triangle$. We assume that all nodes spike at time $t = 0$ such that

$$S_i(0) = T^s \quad \forall i. \tag{C5}$$

Then the deterministic dynamics leads to $S_i(\triangle - 1) = 0$. The transition $S_i(\triangle - 1) = 0 \longrightarrow S_i(\triangle) = 0$ requires that $P_i(\triangle - 1) \leqslant 0$. The next transitions require $P_i(\triangle) \leqslant 0$, $P_i(\triangle + 1) \leqslant 0 ... P_i(T-2) \leqslant 0$, and $P_i(T-1) \geqslant 1$. This correspond to a total of $T - \triangle + 1$ inequalities associated with node i.

From Fig. 13 we roughly estimate that the density of nodes per unit of length located at a distance r from a given node is

$$\rho(r) \simeq \frac{N-1}{2} r. \tag{C6}$$

The first inequality $P_i(\triangle - 1) \leqslant 0$ involves only the nodes j that are at a delay $\tau_{ij} \in [\triangle - T^s, \triangle]$ from node i. These nodes are $\simeq \rho((\triangle - T^s)cdt)T^s cdt$ in number. For the next inequality the nodes involved are at delay $\tau_{ij} \in [(\triangle + 1 - T^s, \triangle + 1)]$ from node i and are $\simeq \rho((\triangle + 1 - T^s)cdt)T^s cdt$ in number. As $\rho(r)$ is increasing with r, the number of variables is larger for the second inequality than for the first. It will be also the case for the other following inequalities. Therefore the ratio number of variables/number of inequalities that we will retain is the smallest one

$$\zeta \simeq \frac{N-1}{2}[(\triangle - T^s)T^s]cdt^2. \tag{C7}$$

To finish our estimation, we have to express the dependence of cdt in N. Numerical investigations (Appendix B, Fig. 15) suggest that $d_{\text{hex}} \simeq 3.41/\sqrt{N}$ such that

$$cdt \simeq \frac{3.41}{\tau_{\min}\sqrt{N}} \tag{C8}$$

and

$$\zeta \simeq 5.8\frac{(\triangle - T^s)T^s}{\tau_{\min}^2} \tag{C9}$$

for large N. The above expression leads to important remarks:

(1) ζ does not depend on N. This is fundamentally different from the Kuramoto'case for which the probability of finding a network configuration that supports a synchronization regime increases linearly with N.

(2) the limit T^s negligible in front of T is obtained by making T^s tend towards 0 while maintaining constant T, $\triangle$ and $\tau_{\min}$. In this limit ζ is vanishing, which means that it is highly unlikely to accidentally run into a network configuration that can support a phase synchronization regime. This justifies our claim that the phase synchronization of a set of excitable, pulse-coupled oscillators with distance-dependent delay is a geometric frustration problem.

(3) In our numerical simulations we have indeed noticed that the network configuration search program struggled and sometimes failed to complete its task when T^s was too small in front of $\triangle$. The values of T^s that we have selected are therefore a compromise between values high enough for the algorithm to work and low enough for the network configurations to be influenced by geometric frustration.

(4) The temporal evolution of our model is discretized in time and is calculated only every dt. By scale change $dt \longrightarrow \frac{dt}{\lambda}$, T, T^s, $\triangle$, and $\tau_{\min}$ are, respectively, transformed into λT, λT^s, $\lambda\triangle$, and $\lambda\tau_{\min}$, but ζ is constant. Therefore our results should not be impacted by the discrete aspect of the temporal evolution.

Finally, it is interesting to note that, in the case where delays are not distance-dependent but chosen at random according to a distribution of characteristic width σ [37–40], then each of the $T - \triangle + 1$ inequalities $P_i(\triangle - 1) \leqslant 0$, $P_i(\triangle) \leqslant 0 .. P_i(T-2) \leqslant 0$ and $P_i(T-1) \geqslant 1$ involve $\simeq N\frac{T_s}{\sigma}$ unknown weights W_{ij}. The ratio between the number of

variables and the number of inequalities is then

$$\zeta \simeq N \frac{T_s/\sigma}{T - \triangle + 1}, \tag{C10}$$

i.e., proportional to N. The situation is then much less restrictive than in the previous case and explain the observations reported in the literature.

[1] T. Gross and B. Blasius, Adaptive coevolutionary networks: A review, J. R. Soc. Interface **5**, 259 (2008).

[2] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Synchronization in complex networks, Phys. Rep. **469**, 93 (2008).

[3] *Adaptive Networks Theory, Models and Applications*, edited by T. Gross and H. Sayama, NECSI Studies on Complexity Collection (Springer, Berlin, 2009).

[4] S. Bornholdt and T. Rohlf, Topological Evolution of Dynamical Networks: Global Criticality from Local Dynamics, Phys. Rev. Lett. **84**, 6114 (2000).

[5] S. Bornholdt and T. Röhl, Self-organized critical neural networks, Phys. Rev. E **67**, 066118 (2003).

[6] D. Garlaschelli, A. Capocci, and G. Caldarelli, Self-organized network evolution coupled to extremal dynamics, Nat. Phys. **3**, 813 (2007).

[7] P. M. Gleiser and D. H. Zanette, Synchronization and structure in an adaptive oscillator network, Eur. Phys. J. B **53**, 233 (2006).

[8] R. Gutiérrez, A. Amann, S. Assenza, J. Gómez-Gardeñes, V. Latora, and S. Boccaletti, Emerging Meso- and Macroscales from Synchronization of Adaptive Networks, Phys. Rev. Lett. **107**, 234103 (2011).

[9] M. Brede, Locals vs. global synchronization in networks of nonidentical Kuramoto oscillators, Eur. Phys. J. B **62**, 87 (2008).

[10] M. Brede, Synchrony-optimized networks of nonidentical Kuramoto oscillators, Phys. Lett. A **372**, 2618 (2008).

[11] L. Papadopoulos, J. Kim, J. Kurths, and D. S. Bassett, Development of structural correlations and synchronization from adaptive rewiring in networks of Kuramoto oscillators, Chaos **27**, 073115 (2017).

[12] T. Nishikawa and A. E. Motter, Symmetric States Requiring System Asymmetry, Phys. Rev. Lett. **117**, 114101 (2016).

[13] L. Gil, Optimally frequency-synchronized networks of nonidentical Kuramoto oscillators, Phys. Rev. E **104**, 044211 (2021).

[14] L. Timms and L. Q. English, Synchronization in phase-coupled Kuramoto oscillator networks with axonal delay and synaptic plasticity, Phys. Rev. E **89**, 032906 (2014).

[15] T.-W. Ko and G. B. Ermentrout, Effects of axonal time delay on synchronization and wave formation in sparsely coupled neuronal oscillators, Phys. Rev. E **76**, 056206 (2007).

[16] G. C. Sethia, A. Sen, and F. M. Atay, Clustered Chimera States in Delay-Coupled Oscillator Systems, Phys. Rev. Lett. **100**, 144102 (2008).

[17] J. Zhu and X. Liu, Locking induced by distance-dependent delay in neuronal networks, Phys. Rev. E **94**, 052405 (2016).

[18] T.-W. Ko, S.-O. Jeong, and H.-T. Moon, Wave formation by time delays in randomly coupled oscillators, Phys. Rev. E **69**, 056106 (2004).

[19] E. M. Izhikevich, J. A. Gally, and G. M. Edelman, Spike-timing dynamics of neuronal groups, Cereb. Cortex **14**, 933 (2004).

[20] H. Markram, Y. Wang, and M. Tsodyks, Differential signaling via the same axon of neocortical pyramidal neurons, Proc. Natl. Acad. Sci. USA **95**, 5323 (1998).

[21] G.-Q. Bi and M.-M. Poo, Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type, J. Neurosci. **18**, 10464 (1998).

[22] Y. Ikegaya, G. Aaron, R. Cossart, D. Aronov, I. Lampl, D. Ferster, and R. Yuste, Synfire chains and cortical songs: Temporal modules of cortical activity, Science **304**, 559 (2004).

[23] See for example the review by E. O. Mann and O. Paulsen, Role of GABAergic inhibition in hippocampal network oscillation, Trends Neurosci. **30**, 343 (2007); or that of M. Bartos, I. Vida, and P. Jonas, Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks, Nat. Rev. Neuroscience **8**, 45 (2007).

[24] J. A. Cardin, Inhibitory interneurons regulate temporal precision and correlations in cortical circuits, Trends Neurosci. **41**, 689 (2018).

[25] E. M. Izhikevich, Polychronization: Computation with spikes, Neural Comput. **18**, 245 (2006).

[26] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown, A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects, J. Neurophysiol. **93**, 1074 (2005).

[27] E. M. Izhikevich, Simple model of spiking neurons, IEEE Trans. Neural Netw. **14**, 1569 (2003).

[28] H. Wässle and H. J. Riemann, The mosaic of nerve cells in the mammalian retina, Proc. R. Soc. Lond. B. **200**, 441 (1978).

[29] H. A. Simon, The architecture of complexity, Proc. Amer. Phil. Soci. **106**, 467 (1962).

[30] A. Bibbig, R. D. Traub, and M. A. Whittington, Long-range synchronization of γ and β oscillations and the plasticity of excitatory and inhibitory synapses: A network model, J. Neurophysiol. **88**, 1634 (2002).

[31] H. A. Swadlow, Physiological properties of individual cerebral axons studied *in vivo* for as long as one year, J. Neurophysiol. **54**, 1346 (1985).

[32] H. A. Swadlow, Efferent neurons and suspected interneurons in motor cortex of the awake rabbit: Axonal properties, sensory receptive fields, and subthreshold synaptic inputs, J. Neurophysiol. **71**, 437 (1994).

[33] W. Gerstner, Rapid Phase Locking in Systems of Pulse-Coupled Oscillators with Delays, Phys. Rev. Lett. **76**, 1755 (1996).

[34] P. C. Bressloff and S. Coombes, Synchrony in an Array of Integrate-and-Fire Neurons with Dendritic Structure, Phys. Rev. Lett. **78**, 4665 (1997).

[35] S.-O. Jeong, T.-W. Ko, and H.-T. Moon, Time-Delayed Spatial Patterns in a Two-Dimensional Array of Coupled Oscillators, Phys. Rev. Lett. **89**, 154104 (2002).

[36] F. M. Atay, Distributed Delays Facilitate Amplitude Death of Coupled Oscillators, Phys. Rev. Lett. **91**, 094101 (2003).

[37] A. Hutt and F. M. Atay, Effects of distributed transmission speeds on propagating activity in neural populations, Phys. Rev. E **73**, 021906 (2006).

[38] A. Hutt and F. M. Atay, Analysis of nonlocal neural fields for both general and gamma-distributed connectivities, Physica D **203**, 30 (2005).

[39] X. Liang, M. Tang, M. Dhamala, and Z. Liu, Phase synchronization of inhibitory bursting neurons induced by distributed time delays in chemical coupling, Phys. Rev. E **80**, 066202 (2009).

[40] P. Gong and C. V. Leeuwen, Dynamically Maintained Spike Timing Sequences in Networks of Pulse-Coupled Oscillators with Delays, Phys. Rev. Lett. **98**, 048104 (2007).

[41] M. Gosak, R. Markovič, and M. Marhl, The role of neural architecture and the speed of signal propagation in the process of synchronization of bursting neurons, Physica A **391**, 2764 (2012).