Network algorithmics and the emergence of the cortical synaptic-weight distribution

Andre Nathan and Valmir C. Barbosa
Programa de Engenharia de Sistemas e Computação, COPPE, Universidade Federal do Rio de Janeiro, Caixa Postal 68511, 21941-972 Rio de Janeiro, RJ, Brazil

(Received 6 November 2009; revised manuscript received 11 January 2010; published 19 February 2010)

When a neuron fires and the resulting action potential travels down its axon toward other neurons' dendrites, the effect on each of those neurons is mediated by the strength of the synapse that separates it from the firing neuron. This strength, in turn, is affected by the postsynaptic neuron's response through a mechanism that is thought to underlie important processes such as learning and memory. Although of difficult quantification, cortical synaptic strengths have been found to obey a long-tailed unimodal distribution peaking near the lowest values (approximately lognormal), thus confirming some of the predictive models built previously. Most of these models are causally local, in the sense that they refer to the situation in which a number of neurons all fire directly at the same postsynaptic neuron. Consequently, they necessarily embody assumptions regarding the generation of action potentials by the presynaptic neurons that have little biological interpretability. We introduce a network model of large groups of interconnected neurons and demonstrate, making none of the assumptions that characterize the causally local models, that its long-term behavior gives rise to a distribution of synaptic weights (the mathematical surrogates of synaptic strengths) with the same properties that were experimentally observed. In our model, the action potentials that create a neuron's input are, ultimately, the product of network-wide causal chains relating what happens at a neuron to the firings of others. Our model is then of a causally global nature and predicates the emergence of the synaptic-weight distribution on network structure and function. As such, it has the potential to become instrumental also in the study of other emergent cortical phenomena.

DOI: 10.1103/PhysRevE.81.021916 PACS number(s): 87.18.Sn, 87.19.lj, 89.75.Fb

I. INTRODUCTION

The strength of a synapse between a neuron's axon and another's dendrite is generally understood to be some measure of how influential an action potential fired by the presynaptic neuron can be on the buildup of such a potential in the postsynaptic neuron. While the physical entities whose measurement can be said to relate to synaptic strengths are various [1–4], recent experimental work involving measurements of the excitatory postsynaptic potential amplitude has revealed that synaptic strengths follow an approximately lognormal distribution, i.e., a long-tailed distribution that is unimodal and peaks near the lowest voltage values [4].

Understanding the processes that give rise to a distribution with these properties can be greatly enhanced by the construction of mathematical models that take into account the nature of each neuron involved (excitatory or inhibitory), the nature of a synapse's plasticity in terms of how its strength changes in response to inter-neuron signaling, and also the distribution of firings in time. Predictive models have been built with varying degrees of success [1,2,5,6], using the so-called synaptic weights as mathematical representations of synaptic strengths. The most successful of these models draw on relatively well established knowledge regarding the proportion of inhibitory neurons to be used and the rule to change synaptic weights [2,6].

Almost invariably, though, these models have relied on examining one single postsynaptic neuron, toward which firing patterns are directed that in essence seek to summarize the entire input history of the postsynaptic neuron by a simple stochastic process. Arguably this history is one of the most important elements in giving rise to the synaptic-weight distribution in a way that can be understood biologically [7], but in most current models there is no choice but to summarize it beyond retrieval. This happens because the models in question are all strictly local, allowing for no causal dependency between what happens at two neurons unless they are no farther apart from each other than one single synapse.

The one study that to our knowledge has avoided this pitfall is to be found in [6], where the authors consider a network of interconnected neurons along with a Hebbian mechanism for synaptic weights to change during the time-stepped simulation of a set of differential equations relating weights to firing rates. Nontrivial causal dependencies are then taken into account at all times, but the model's setup as tightly coupled differential equations in time blurs every detail of a firing's causal predecessors, which for all purposes remains irretrievable. The model we now introduce addresses the severe shortcoming of the history-oblivious models mentioned above, combining a network structure and algorithm with the proven mathematical elements of those models. In doing so, it also expands on the modeling capabilities of [6] by allowing the spatially remote causes of a neuron's firing to be fully exposed for detailed analysis.

Before proceeding, we note that the present work is part of the now decade-long effort to study the emerging properties of the various networks appearing in several biological, social, and technological domains. The reader is referred to the chapters collected in [8–10] for a general view of the main achievements. We find it surprising, though, that very few of such works have addressed the important question of how node functionality, together with network structure, affects the flow of information among nodes and perhaps even alters the global properties of interest. One example is the simple algorithm described in [11], whose independent ap-

 ©2010 The American Physical Society

plication at all nodes based only on local characteristics of the network gives rise to structures, simpler than the network taken as a whole, for efficient information dissemination. This paper's model is another example, but in a totally different context.

II. MODEL

The model has a structural component and an algorithmic one. The structural component is a directed graph D whose nodes correspond to neurons that can be either excitatory or inhibitory. For i and j two distinct nodes such that at least one of them is excitatory, an edge directed from i to j represents a synapse with associated nonnegative weight w_{ij}. No edge exists between two inhibitory nodes [12].

The algorithmic component turns each node in D into a simple simulator of the corresponding neuron, employing message passing on the edges along their directions to simulate the signaling through the corresponding synapses when nodes fire. Collectively, the nodes behave as an asynchronous distributed algorithm [13], here referred to as A, each executing a simple procedure P whenever receiving a message, possibly sending messages itself while executing P but remaining idle at all other times. Because nodes only do any processing in this reactive manner, at least one node is needed that initially executes P once without any incoming message to respond to and then starts behaving reactively like the others. We call such a node an initiator.

At node j, let v_j stand for the node's potential. Let also v^0 and v^t, with $v^0 < v^t$, be a node's rest potential and threshold potential, respectively, the same for all nodes. The effect of running P is for j to probabilistically decide whether to fire and, if it does fire, to send messages on all outgoing edges while setting v_j to v^0. If P is run as the initial processing by an initiator, then the firing occurs with probability 1 and P involves no actions other than the ones just described. If not, then let i be the sender of the triggering message. The firing occurs with probability $(v_j - v^0)/(v^t - v^0)$ after v_j has been updated to either $\min\{v^t, v_j + w_{ij}\}$, if i is excitatory, or $\max\{v^0, v_j - w_{ij}\}$, if i is inhibitory (so initializing v_j to some value in the $[v^0, v^t]$ interval ensures that it remains thus bounded and that the decision probability is legitimate). Then the weight w_{ij} is considered for an update.

The updating of w_{ij} seeks to mimic the commonly accepted generalization of the Hebbian rule embodied in the spike-timing-dependent plasticity principles [1,14], according to which the change incurred by a synapse's weight depends on the extent to which there is a causal dependency of what happens at a neuron upon the other's firing. As a general rule, the synaptic weight is increased (potentiated) if the postsynaptic neuron fires in response to the firing by the presynaptic neuron, decreased (depressed) otherwise. In either case the amount of change to the synaptic weight depends on how close in time the relevant firings are, becoming negligible with increasing separation.

Procedure P follows these principles by keeping track of the latest firing by j so that a decision can be made on whether to increase or decrease w_{ij}. If j does fire in response to the message received from i, then w_{ij} is increased. If it does not but the previous message received from any source did cause j to fire, then w_{ij} is decreased. The weight w_{ij} remains unchanged in all other cases. The actual amount of change to w_{ij} depends on whether it is to be increased or decreased, and so does the nature of the change (by a fixed amount or by proportion) [3,15,16]. An increase in w_{ij} is implemented by setting w_{ij} to $\min\{1, w_{ij} + \delta\}$ with $\delta > 0$, a decrease by setting w_{ij} to $(1-\alpha)w_{ij}$ with $0 < \alpha < 1$, thus ensuring that synaptic weights remain in the $[0,1]$ interval if so started.

Now assume that $v^t - v^0$ is significantly greater than 1 and that weights are indeed kept to the $[0,1]$ interval. Given the firing and weight-updating rules just described, at node j the most likely scenario concerning the average weight of the incoming edges is the following. Because any firing by j causes v_j to be reset to v^0, the reception of the next message is unlikely to cause another firing, as v_j gets set to at most $v^0 + 1$ and therefore the probability of the new firing is at most $1/(v^t - v^0)$. If d_j is the number of edges incoming to j, then it follows that the corresponding average weight undergoes an increase by δ/d_j due to the first message, and the result a decrease at the rate of α/d_j due to the second message (assuming the mean-field approximation in which the weight of every edge incoming to j is now the same). Then some number of weight-preserving messages may arrive, after which the same pattern of behavior gets repeated. For $k \geq 1$, using w_j^k to denote the average weight after the kth round yields

$$w_j^k = \left(1 - \frac{\alpha}{d_j}\right)\left(w_j^{k-1} + \frac{\delta}{d_j}\right), \tag{1}$$

with w_j^0 representing the initial average weight. Letting $w_j^* = \lim_{k\to\infty} w_j^k$, we obtain

$$w_j^* = \left(1 - \frac{\alpha}{d_j}\right)\frac{\delta}{\alpha}, \tag{2}$$

so

$$(1-\alpha)\frac{\delta}{\alpha} \leq w_j^* < \frac{\delta}{\alpha} \tag{3}$$

for any value of d_j. Thus, despite the aforementioned mean-field approximation, and even though the distribution of synaptic weights need not be the same as the distribution of the average weights incoming to the nodes, a first approximation is to expect the synaptic-weight distribution to peak somewhere near this interval. It also follows from Eq. (3), and from the fact that synaptic weights are bounded from above by 1, that we must have $\delta \leq \alpha$.

III. COMPUTATIONAL RESULTS

Running algorithm A starts with choosing one or more initiators, each of which executes P and then starts behaving like all other nodes. At any time it may happen that a node has more than one input message to process, in which case the order in which they are taken is the order of message reception. Because this order is in principle arbitrary, A is seen to acquire another degree of indeterminacy, in addition

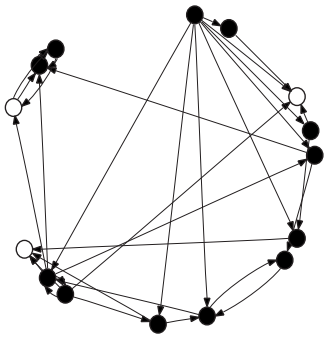

FIG. 1. Network topology. Restricted to two dimensions for visual clarity, a D instance comprises nodes positioned randomly on a radius-1 circle and edges, drawn as chords of the circle, that tend to be more abundant over lower Euclidean distances. Excitatory nodes are represented by filled circles, inhibitory nodes by empty circles.

to that which is already present owing to the probabilistic decisions.

We have conducted extensive computational experimentation with A on a graph D intended to model a simple cortex, in line with significant recent work that draws on the theory of graphs to help solve problems in neuroscience [17–25]. We regard D as a random graph but, unlike some of the early work on cortical modeling by such graphs [12], where fully random graphs [26] were used, we let D have a scale-free structure [27], with parameter as suggested by some of the more recent finds [28,29], though only loosely (since in these works it is not neurons, but rather larger functional clusters, that are investigated for their connectivity patterns). Thus, a randomly chosen node i in D has k outgoing edges with probability proportional to $k^{-1.8}$. Moreover, inspired by recent work on the modeling of cortical systems [30,31], we let each outgoing edge of i lead to another randomly chosen node j with probability proportional to $e^{-2d_{ij}}$, where d_{ij} is the Euclidean distance between i and j when the nodes of D are placed uniformly at random on a radius-1 sphere (Fig. 1), provided i and j are not both inhibitory.

All computational experiments have adhered to the methods described next, which refer to sequences of runs of algorithm A, each sequence comprising 10 000 runs. The first run in a sequence operates on initial node potentials and synaptic weights chosen randomly from the intervals $[v^0,v^t]$ and $[0,1]$, respectively, with $v^0=-15$ and $v^t=0$. Each subsequent run operates on the potentials and weights left by the previous run. For n the number of nodes in D, a new set of $0.05n$ initiators is chosen randomly at the beginning of each run. A run of A is implemented as a sequential program that selects the next node to be processed randomly (first out of the group of initiators for their first executions of P, then out of those nodes that have at least one message to be received). A new run in a sequence is only started after the previous one has died out (no more messages to be processed remain), which is guaranteed to happen eventually with probability 1. Unless otherwise noted, the remaining parameters used by procedure P are $\delta=0.01$ and $\alpha=0.05$.

It is important to realize that our sequential implementation of distributed algorithm A does in no way interfere with the massively parallel character that is inherent to A. In order to see this, notice that, even if a massively parallel implementation were possible (one processor per node), the concurrent handling of messages arriving at distinct nodes would still be restricted to causally disconnected messages, i.e., messages having nothing to do with one another (not even remotely). The relative order in which any such messages were handled would then be immaterial and could, therefore, be any order. All our sequential implementation does is to adopt a random order. In a similar vein, an alternative to requiring a run to die out completely before a new one is started is to do one single, longer run in which nodes with no message to be received are probabilistically given a chance to act as initiators every so often. While this alternative may seem biologically more plausible, we have found it to be essentially equivalent to the many-run approach outlined above. The latter, in turn, has had our preference because it allows for better termination control through how many runs constitute a sequence and how many initiators a run has.

All our results are averages over 50 000 independent sequences, of which each 500 sequences correspond to a new D instance (so there are 100 D instances overall). A D instance is constructed by first placing all nodes uniformly at random on a radius-1 sphere, then selecting the number of outgoing edges for each node. For $k>0$, a node receives k outgoing edges with probability $k^{-1.8}/\Sigma_{k'=1}^{n-1}(k')^{-1.8}$. Nodes are then chosen to be excitatory or inhibitory randomly, provided a certain proportion is respected, and the destination of each edge is decided. An outgoing edge of node i is set to lead to node $j\neq i$ with probability $e^{-2d_{ij}}/\Sigma_{j'\neq i}e^{-2d_{ij'}}$. In the latter, the i,j pair must comprise at least one excitatory node, and so must each of the i,j' pairs. The graph that is actually used in the run sequences is the giant strongly connected component (GSCC) of D, that is, the largest subgraph of D in which a directed path exists from any node to any other (and therefore so do directed cycles going through any two nodes) [32]. For the connectivity distribution and construction method in use this component comprises about $0.95n$ nodes on average.

Our results, here given for $n=1000$ and the well accepted proportion of $0.2n$ inhibitory nodes [12,33], show that the synaptic-weight distribution becomes analogous to the distribution unveiled by experimentation along the sequences of runs of algorithm A described above (Fig. 2). The process is gradual, leading the weights to become relatively concentrated around a single low-value mode while still allowing some residual probability to remain at the higher values. We note that, as predicted by the approximation suggested by Eq. (3) by virtue of the v^0 and v^t values in use, the synaptic-weight distribution does reach its maximum in the vicinity of the $[0.19,0.2]$ interval.

The long-term distribution is seen to stabilize even as the weights continue to evolve, thus suggesting the existence of an underlying weight dynamics whose effect on the overall distribution is nevertheless practically imperceptible. The existence of this persistent dynamics is revealed by the causal history of each terminal message reception (one that does not lead to the firing of the receiver), which can be significantly

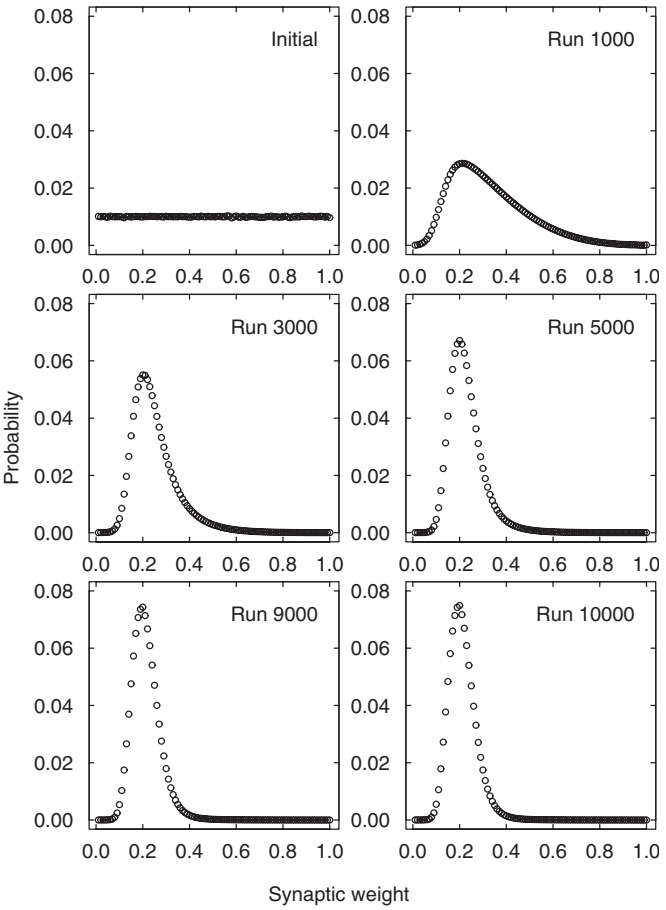

FIG. 2. The synaptic-weight distribution, shown after selected runs of algorithm A. Probabilities are binned to a fixed width of 0.01.

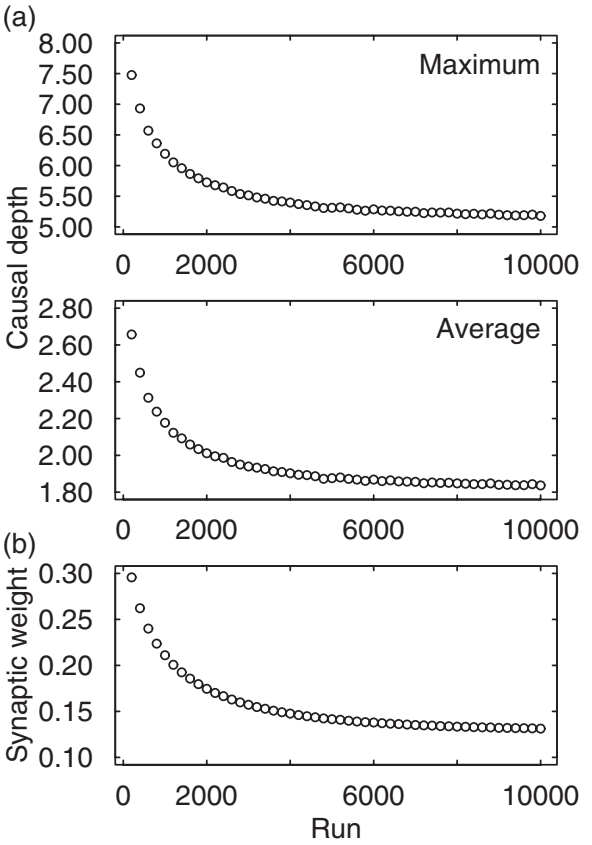

FIG. 3. Causal depth of a message reception and associated synaptic weights. The causal depth of a message reception is the size of its causal history, i.e., the number of firings that precede it along the chain of firings that begins at some initiator when it fires for the first time, each preceding the next by direct causation: given any two subsequent firings in this chain, the first entails the sending of a message whose reception triggers the second. (a) Maximum and average causal depth of terminal message receptions during the course of each run. (b) Average weight (before updates) of the synapses involved in the causal histories of terminal message receptions.

deep with respect to the relatively short average path of a scale-free network [34] [Fig. 3(a)]. Except for the initial sending of messages by initiators, the sending of every message causes a synaptic weight to be increased, unless it already equals 1. But weight-1 synapses are very rare, especially when arranged as a path in D, so the causal histories we have discovered do indeed hint at the existence of a dynamics of weight evolution in which weights both increase and decrease in complex patterns. Additional confirmation is provided by the average weight of the synapses involved in the causal histories of terminal message receptions, which is consistently less than 1 and also decreases throughout the runs as the synaptic-weight distribution settles [Fig. 3(b)].

For the same values of v^0 and v^t, we have observed convergence to an approximate lognormal distribution to occur also for other values of δ and α. Two of these final distributions are shown in Fig. 4 alongside the final distribution of Fig. 2. The two new distributions are for $\delta=0.015$ with $\alpha=0.05$ and $\delta=0.01$ with $\alpha=0.025$. Equation (3) predicts their peaks to be near the intervals [0.285,0.3] and [0.39,0.4], respectively, and this is seen to be true. It is also apparent from Fig. 4 that increasing the δ/α ratio seems to disrupt the approximate lognormal character of the final distribution. We had expected this to happen as the distribution's mode moved ever closer to 1 in the process of increasing δ/α, due

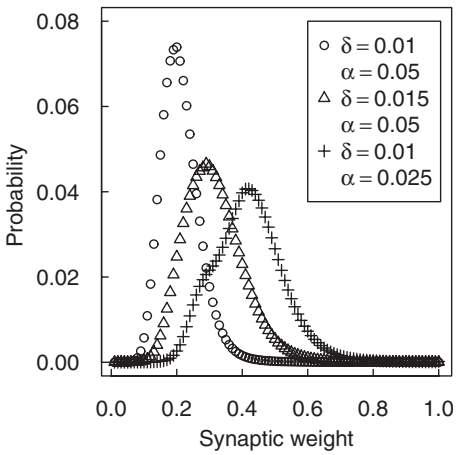

FIG. 4. Final synaptic-weight distributions for three combinations of δ and α values.

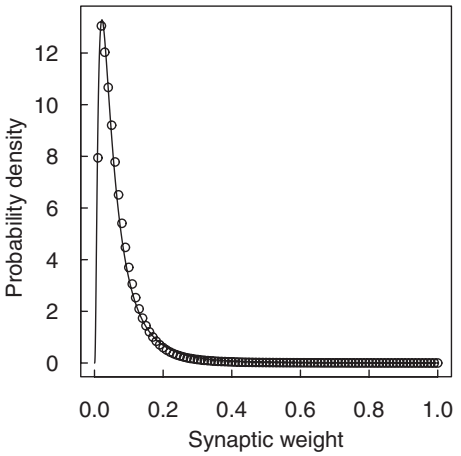

FIG. 5. Final synaptic-weight densities for $\delta=0.0002$ and $\alpha=0.04$. Densities are obtained by first binning probabilities to a fixed width of 0.01 and then dividing the resulting probabilities by the bin's width. For $w>0$ representing a synaptic weight, the solid line gives the lognormal density function $f(w)=0.438\exp[-(\ln w+2.997)^2/(2\times0.9111^2)]/w$, which corresponds to the one that in [4] fits experimental data, given by $g(w)=0.426\exp[-(\ln w+0.702)^2/(2\times0.9355^2)]/w$, after weights are rescaled by a constant factor from the [0,10] interval used in [4] down to the [0,1] interval we use. Thus, the two functions are such that $f(w)=10g(10w)$.

mainly to probability accumulation at the far end since weights are truncated at 1. What we see instead is that already for $\delta=0.01$ with $\alpha=0.025$ the distribution starts to deviate from having a lognormal appearance.

We have also found that varying the value of δ more widely so that the δ/α ratio is made to move substantially closer to 0 can be quite revealing. Specifically, for $\delta=0.0002$ and $\alpha=0.04$, the synaptic-weight distribution stabilizes in such a way that the corresponding density function is in excellent agreement, after synaptic weights have been appropriately rescaled, with the lognormal density function that in [4] is shown to provide the best fit to experimental data. This agreement is shown in Fig. 5, where the peak occurs at about 0.02, thus also indicating that for this combination of δ and α values, the prediction of Eq. (3), according to which the peak is to occur very near 0.005, breaks down.

IV. DISCUSSION

In D, the distributions of a node's out-degree (how many outgoing edges it has) and in-degree (how many incoming edges it has) inside the GSCC are as shown in Fig. 6. The expected out- or in-degree is 6.7. Except for the finite-size effects in the out-degree distribution near the highest degrees, the agreement with the power-law distribution of exponent −1.8 used to create the D instances is very good. Although this is known not to happen in general due to the conditioning on a node's being part of the GSCC [35], in the present case it can be attributed to the fact that, as remarked earlier, the GSCC is expected to encompass a sizable portion of D (roughly 95% of its nodes). As for the in-degree distribution, the one shown in Fig. 6 corresponds to the use of

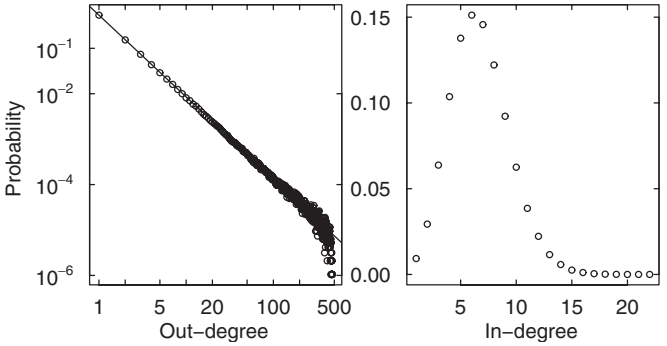

FIG. 6. Out- and in-degree distributions within the GSCC. Data are averages over 1000 graphs with $n=1000$. The solid line gives the power-law distribution of exponent −1.8.

$e^{-2d_{ij}}$ in the probabilistic decisions to deploy edges once the nodes' out-degrees have been selected.

In a more general formulation, these decisions occur with probabilities proportional to $e^{\lambda d_{ij}}$ for some $\lambda<0$ and the choice of which value of λ to use depends on whether one wishes D, on average, to fall below or above the phase transition that gives rise to the graph's GSCC. For the same power-law distribution of out-degrees we have been using, this phase transition is shown in Fig. 7, where the expected size of the GSCC is given for varying λ. The transition can be said to start to occur at about $\lambda=-25$ and to be consolidated roughly past $\lambda=-5$. The choice of $\lambda=-2$ we have used throughout is then to be understood as reflecting the need for most of the n nodes to be expected to lie within the GSCC.

Aside from the graph-related parameter λ and the δ/α ratio, whose variation we investigated in the previous section, our results have also relied on the 5% fraction of initiators and on the $v^t-v^0=15$ difference between the threshold and rest potentials. Increasing the former and decreasing the latter have the common effect of increasing the number of causally disconnected messages, in the sense discussed earlier, to be processed during each run. As a result, more weight alterations occur in a run and convergence to the final distribution occurs in fewer runs. The distribution itself,

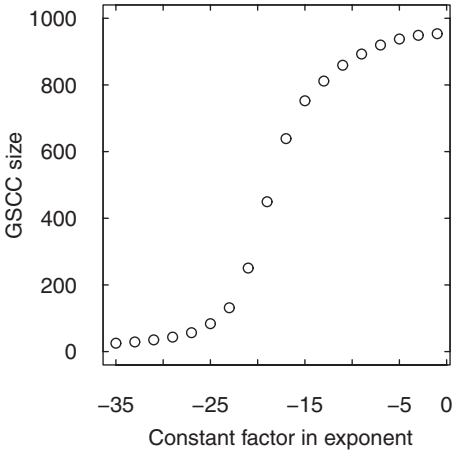

FIG. 7. Expected number of nodes in the GSCC for different values of the constant factor λ. Data are averages over 1000 graphs with $n=1000$.

however, remains unaltered, so both the initiator fraction and the $v^t - v^0$ difference are to be regarded merely as influencing how the run sequences progress in time. Of course, this has nothing to do with how long it takes to execute a run sequence to completion, since despite requiring fewer runs for stabilization each run requires more messages to be processed.

V. CONCLUDING REMARKS

Every run in every sequence involves a new group of initiators and as such provides new possibilities regarding the branching of causal histories and how they affect firings and weight changes throughout the network. Monitoring the traffic of messages as they traverse edges and reach nodes is then a means to do some quantification of how the cascading runs, with their intermingling causal trees rooted at many different initiators, cooperate in promoting the emergence of the synaptic-weight distribution.

We have found that the long-term distributions of how many runs traverse an edge or reach a node (see Fig. 8, which corresponds to the same run sequences as Figs. 2 and 3), allowing as they do for relatively high numbers with significant probabilities, suggest that some sort of information integration is taking place among portions of the network as the runs unfold. Perhaps such integration occurs in a sense similar to that which has been theorized recently regarding the emergence of higher functions such as consciousness [36]. If so, then network algorithmics such as we have discussed may come to provide a powerful framework to test the assumptions and eventual predictions of such theories.

We also remark that, although cerebral cortices do exhibit some structural properties that are typical of scale-free networks [37,38], such as the so-called small-world properties [39] and also the presence of hubs (nodes with very many outgoing edges), it is still uncertain whether a network description based on a scale-free distribution of a node's out-degree is appropriate. This is so despite the many other scale-free aspects of the cortex [40], especially if we consider that there do exist growth models that give rise to other types of out-degree distribution [37]. These models, however, are themselves not fully justified biologically, so it seems to us that the question has so far remained open. The difficulty in resolving it seems to be, essentially, related to the technologies available for measurement, which as far as we know have revealed the absence of a scale-free distribution of out-degrees only for very small groups of neurons [38]. So our decision to perform computations on networks having just this type of out-degree distribution is to be regarded as settling for the best granularity at which meaningful measurements can currently be performed, such as those of [28,29].

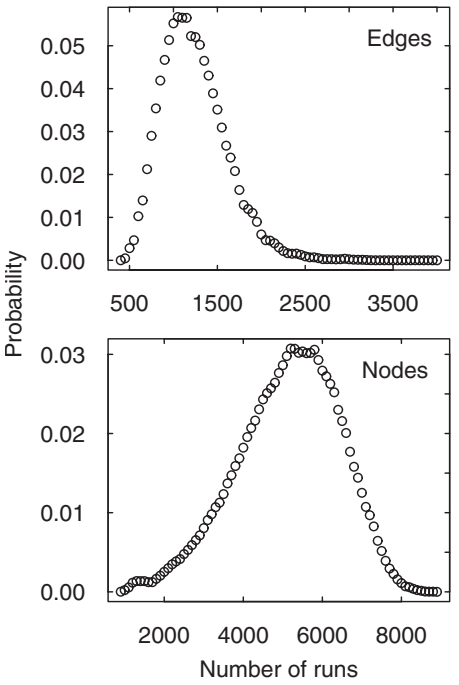

FIG. 8. Final distributions of the number of runs in which an edge is traversed or a node is reached. An edge is said to be traversed in a run when at least one message is sent along it during the course of that run. A node is said to be reached in a run when it receives at least one message during the course of that run. Probabilities are binned to a fixed width of 50 for edges, 100 for nodes.

One distinctive characteristic of our model has been the explicit incorporation of an algorithmic component to account for both the local processing that takes place at the nodes and the passing of messages among them. This immediately sets our approach apart from any form of "analog" modeling of the physical quantities involved in inter-neuron signaling, thus perhaps raising questions regarding biological plausibility. Our stance on this issue is in many senses akin to that of researchers in the field of artificial life, whose "life as it could be" motto refers precisely to the substitutability of physical substrates by pieces of code, provided function remains preserved [41]. This is what we have demonstrated our approach to be able to achieve.

ACKNOWLEDGMENTS

We acknowledge partial support from CNPq, CAPES, and a FAPERJ BBP grant.

[1] S. Song, K. D. Miller, and L. F. Abbot, Nat. Neurosci. **3**, 919 (2000).

[2] M. C. W. van Rossum, G. Q. Bi, and G. G. Turrigiano, J. Neurosci. **20**, 8812 (2000).

[3] A. Kepecs and M. C. W. van Rossum, Biol. Cybern. **87**, 446 (2002).

[4] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, PLoS Biol. **3**, 507 (2005).

[5] J. Rubin, D. D. Lee, and H. Sompolinsky, Phys. Rev. Lett. **86**, 364 (2001).
[6] A. A. Koulakov, T. Hromádka, and A. M. Zador, J. Neurosci. **29**, 3685 (2009).
[7] B. Barbour, N. Brunel, V. Hakim, and J. P. Nadal, Trends Neurosci. **30**, 622 (2007).
[8] *Handbook of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Weinheim, Germany, 2003).
[9] *The Structure and Dynamics of Networks*, edited by M. Newman, A.-L. Barabási, and D. J. Watts (Princeton University Press, Princeton, NJ, 2006).
[10] *Handbook of Large-Scale Random Networks*, edited by B. Bollobás, R. Kozma, and D. Miklós (Springer, Berlin, Germany, 2009).
[11] A. O. Stauffer and V. C. Barbosa, Theor. Comput. Sci. **355**, 80 (2006).
[12] M. Abeles, *Corticonics: Neural Circuits of the Cerebral Cortex* (Cambridge University Press, Cambridge, UK, 1991).
[13] V. C. Barbosa, *An Introduction to Distributed Algorithms* (The MIT Press, Cambridge, MA, 1996).
[14] L. F. Abbott and S. B. Nelson, Nat. Neurosci. **3**, 1178 (2000).
[15] G. Q. Bi and M. M. Poo, J. Neurosci. **18**, 10464 (1998).
[16] G. Q. Bi and M. M. Poo, Annu. Rev. Neurosci. **24**, 139 (2001).
[17] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, Trends Cogn. Sci. **8**, 418 (2004).
[18] O. Sporns, G. Tononi, and R. Kötter, PLoS Comput. Biol. **1**, 245 (2005).
[19] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, J. Neurosci. **26**, 63 (2006).
[20] D. S. Bassett and E. Bullmore, Neuroscientist **12**, 512 (2006).
[21] Y. He, Z. J. Chen, and A. C. Evans, Cereb. Cortex **17**, 2407 (2007).
[22] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, Proc. Natl. Acad. Sci. U.S.A. **104**, 10240 (2007).
[23] J. C. Reijneveld, S. C. Ponten, H. W. Berendse, and C. J. Stam, Clin. Neurophysiol. **118**, 2317 (2007).
[24] O. Sporns, C. J. Honey, and R. Kötter, PLoS ONE **2**, e1049 (2007).
[25] C. J. Stam and J. C. Reijneveld, Nonlinear Biomed. Phys. **1**, 3 (2007).
[26] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[27] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).
[28] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, Phys. Rev. Lett. **94**, 018102 (2005).
[29] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. E. Hulshoff Pol, Neuroimage **43**, 528 (2008).
[30] M. Kaiser and C. C. Hilgetag, Neurocomputing **58-60**, 297 (2004).
[31] M. Kaiser and C. C. Hilgetag, Phys. Rev. E **69**, 036103 (2004).
[32] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **64**, 025101(R) (2001).
[33] R. Ananthanarayanan and D. S. Modha, *Proceedings of the 2007 ACM/IEEE Conference on Supercomputing* (ACM, New York, NY, 2007), p. 3.
[34] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[35] A. O. Stauffer and V. C. Barbosa, IEEE/ACM Trans. Netw. **15**, 425 (2007).
[36] D. Balduzzi and G. Tononi, PLoS Comput. Biol. **4**, e1000091 (2008).
[37] W. J. Freeman, R. Kozma, B. Bollobás, and O. Riordan, in [10], pp. 277–324.
[38] S. Yu, D. Huang, W. Singer, and D. Nikolić, Cereb. Cortex **18**, 2891 (2008).
[39] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[40] W. J. Freeman, Scholarpedia **2**, 1357 (2007).
[41] N. Forbes, IEEE Intell. Syst. **15**, 2 (2000).