Scale-free and economical features of functional connectivity in neuronal networks

Jean-Philippe Thivierge*
School of Psychology and Center for Neural Dynamics, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5
(Received 20 June 2014; published 29 August 2014)

A form of activity that is highly studied in cultured cortical networks is the neuronal avalanche, characterized by bursts whose distribution follows a power law. While the statistics of neuronal avalanches are well characterized, much less is known about the neuronal interactions from which they arise. We examined statistical dependencies between pairs of cells in spontaneously active cultures of cortical neurons using an information measure of transfer entropy. We show that the distribution of transfer entropy follows a power law with a slope near 3/2. Using graph-theoretic approaches of weighted networks, we demonstrate that this power law maximizes a measure of global economy that accounts for both the efficiency of neuronal interactions as well as the overall traffic in the network. Finally, we describe a pairwise Poisson model that captures the statistics of information transfer in a population of spiking neurons. Using this model, we show that avalanches can occur in systems with weak pairwise interactions, and that strong pairwise interactions can arise without avalanches, suggesting that these two measures capture distinct properties of brain dynamics.

DOI: 10.1103/PhysRevE.90.022721 PACS number(s): 87.85.dq

I. INTRODUCTION

Neuronal networks of cortical neurons *in vitro* show complex patterns of coordinated activity [1,2]. A hallmark feature of this activity is the presence of network bursts, characterized by large network events whose activity propagates across a substantial portion of the neuronal tissue. The statistics of these network bursts has been a subject of considerable study in recent years. The term "neuronal avalanche" has been coined to describe bursts that follow a power law distribution [3,4]. The presence of avalanches with a slope near 3/2 has led to the possibility that networks of cortical neurons both *in vitro* and *in vivo* operate near a critical state.

Despite a growing body of theoretical models that capture the statistics of neuronal avalanches, there is no consensus on the nature of neuronal interactions from which they emerge. In fact, different computational models capture the statistics of avalanches despite widely different patterns of connectivity—while some models are fully connected [5], others are scale free [6], random [7], modular [8], or nearest neighbor [9].

Experiments on cortical cultures *in vitro* do not grant us access to the underlying synaptic connections across cells; however, it is possible to study the statistical patterns of neuronal interactions—also termed "functional connectivity"—between neurons [10–12]. These interactions are highest for cells whose activity arises in close temporal contiguity. Several approaches have been proposed to estimate neuronal interactions from the activity of a neuronal network, including (but not limited to) Granger causality [12], Bayesian approaches [11], partial directed coherence [13], and transfer entropy [10,14–17].

Recently, important insights into the organization of neuronal networks have been gained by combining analyses of functional connectivity with techniques from statistical physics and graph theory. One common approach is to estimate the strength of functional connectivity between neurons (also termed "weights"), then construct a binary graph that throws away strength values and retains only a yes or no assessment of functional connections. This approach has revealed interesting characteristics of functional connectivity, including a small-world organization, characterized by a short characteristic path length and high clustering coefficient [18].

An alternative approach is to retain strength values when examining the patterns of functional connections. Analyses of weighted networks have shown a scale-free distribution of neuronal interactions [19], and have characterized causal relationships between neurons [12] as well as the direction of information flow across a network [10].

While the distribution of weights likely plays a key role in routing information across a network, this exact role remains unclear. Here we examine this question by recording spontaneous activity from cultures of neocortical neurons *in vitro*. As reported previously, these recordings show avalanches characterized by a power law distribution of neuronal activity [3,10,20]. We estimated functional connectivity using an information-theoretic measure termed "transfer entropy" [10,14,15,17]. From this measure, we derived a weighted network of interaction between all pairs of neurons. We analyzed this network using graph-theoretic approaches that reveal a fundamental role of weight distributions in channeling information across a network of interconnected neurons. In addition, we describe a Poisson model that captures the statistics of neuronal avalanches and examines the role of pairwise interactions in generating this form of activity.

II. RESULTS

We recorded cell cultures using a 60-microelectrode array [10,21,22] [Fig. 1(a)]. Spontaneous activity was characterized by population bursts that recruited a large proportion of the network [Fig. 1(b)]. To quantify spontaneous activity, we examined the propagation of neuronal avalanches, defined as bursts of activity spreading in time and across neurons (see Methods).

*Present address: School of Psychology, Faculty of Social Sciences, University of Ottawa, 136 Jean-Jacques Lussier, Vanier Hall, office 2064, Ottawa, Ontario, Canada K1N 6N5; jthivier@uottawa.ca

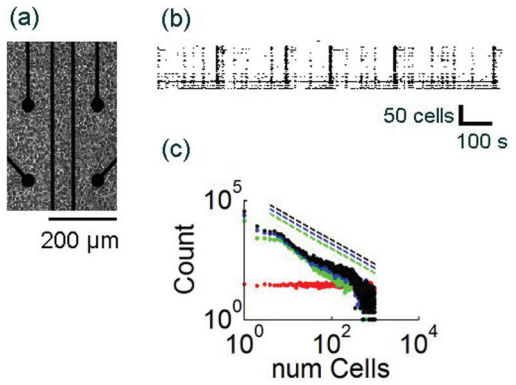

FIG. 1. (Color) Neuronal activity in cultures of neocortical neurons. (a) Culture of cortical neurons plated on microelectrode array (only a subset of array shown). Electrodes are spaced 200 μm apart. (b). Spike raster showing the activity of a population of neurons. (c) Distribution of the number of cells participating in avalanches, with time windows of 10 ms (black), 15 ms (blue), and 20 ms (green). Red, shuffled spikes (window of 10 ms). Dashed lines show best-fitting power law obtained with maximum likelihood estimation.

A. Neuronal avalanches

A total of 36 236 avalanches were extracted in recordings from eight cortical cultures. While these avalanches can be described using several measures, here we focused on the number of cells active per avalanche [20]. Most avalanches activated only a small subset of neurons; however, a small minority of avalanches recruited a large proportion of the population. The overall distribution of neurons recruited during avalanches resembled a truncated power law [Fig. 1(c)]. We employed a maximum likelihood method to compare the fit of avalanches to a power law versus exponential distribution. The estimation ($\hat{\alpha}$) of the slope α of a bounded discrete power law is obtained as follows [23]:

$$\hat{\alpha} = \arg\max_{\alpha} \left[-\alpha \left(\sum_{i=1}^{n} \ln x_i \right) - n \ln \zeta(\alpha, x_{\min}, x_{\max}) \right], \quad (1)$$

where x_i is the ith element from the vector of all data $X = \{x_1, x_2, \ldots, x_n\}$ and n is the size of the data vector. Here, the nth element of X corresponds to the number of neurons active during the nth avalanche recorded. The Hurwitz zeta function, $\zeta(\alpha, x_{\min}, x_{\max})$, is given by

$$\zeta(\alpha, x_{\min}, x_{\max}) = \zeta(\alpha, x_{\min}) - \zeta(\alpha, x_{\max}), \quad (2)$$

where

$$\zeta(\alpha, x) = \sum_{i=1}^{\infty} \frac{1}{(i+x)^{\alpha}}. \quad (3)$$

Techniques for estimating lower and upper bounds from raw data are covered elsewhere [23]; here, we set $x_{\min}$ and $x_{\max}$ to the minimum and maximum values of observed data, respectively. In related work, we have examined the robustness of estimates to missing data [20] and shown that the above method provides a more accurate estimate of power laws than other approaches [23,24].

The above power law estimate can be compared to that of an exponential function. The maximum likelihood estimator of parameter λ for the exponential function is

$$\hat{\lambda} = \arg\max_{\lambda} \left\{ n[\ln(\lambda) - \ln(e^{-\lambda x_{\min}} - e^{-\lambda x_{\max}})] - \lambda \sum_{i=1}^{n} x_i \right\}. \quad (4)$$

While estimates could be derived for cases other than the power law and exponential distributions, fitting functions such as a lognormal distribution is difficult and often inconclusive [24] and therefore not included here. In order to compare the fit of power law and exponential functions, we employed a Bayesian information criterion as opposed to the usual log-likelihood ratio that is typically used in this scenario [25,26]. In statistical analyses, the log-likelihood ratio is applicable only when one of the functions considered is a special case of the other [27]. The Bayesian information criterion (B) is obtained by calculating the log probability of the power law and exponential functions, respectively, given a dataset X [28]:

$$\begin{aligned} B_1 &= -L(\alpha|X) + \log(n)\phi_1, \\ B_2 &= -L(\lambda|X) + \log(n)\phi_2, \end{aligned} \quad (5)$$

where ϕ is the number of parameters in each function. The log-likelihood functions $L(\alpha|X)$ and $L(\lambda|X)$ are

$$\begin{aligned} L(\alpha|X) &= -n \ln \zeta(\alpha, x_{\min}, x_{\max}) + \sum_{i=1}^{n} \ln x_i^{-\alpha}, \\ L(\lambda|X) &= n[\ln(\lambda) - \ln(e^{-\lambda x_{\min}} - e^{-\lambda x_{\max}})] - \lambda \sum_{i=1}^{n} x_i. \end{aligned} \quad (6)$$

The Bayesian information criterion is asymptotically consistent, meaning that if a set of functions includes the true function from which the data were generated, the probability that the true function has the highest value of B will approach 1 as the number of observations increases to infinity. The relative merit (R_i) of each function i can be assessed by

$$R_i = \frac{\exp(-2\log[B_i])}{\sum_{j=1}^{F} \exp(-2\log[B_j])}, \quad (7)$$

where $F = 2$ is the number of functions considered, namely exponential and power law. Equation (7) is highest for the function that offers the best fit to the experimental data.

We calculated the above Bayesian information criterion [Eqs. (5)–(7)] to compare the fit of a power law and exponential function to the distribution of avalanches. The power law attained a markedly higher relative fit (0.969) compared to the exponential function (0.031). The value of the best-fitting power law slope was $\hat{\alpha} = 1.52$. The overall shape of the distribution was not markedly altered when using different time windows to compute avalanches [Fig. 1(c), black, blue, and green filled circles]. This distribution, however, was disrupted when spikes were randomly shuffled prior to computing avalanches [Fig. 1(c), red filled circles].

The presence of neuronal avalanches in cultured networks *in vitro* confirms previous reports and is suggestive of a system that operates near the critical state [4,25] (however, see the Discussion for controversies on this issue). Next, we quantify the amount of information that is exchanged between neurons and examine its relation to avalanches.

B. Transfer entropy

In order to estimate the amount of information exchanged between pairs of neurons, we employed an information-theoretic measure of transfer entropy (TE). This measure quantifies the amount of information in a neuron found in the recent past history of other neurons. Transfer entropy can be used to probe asymmetries in neural relations (i.e., neuron A influencing B, but B not influencing A) [10,14–17], and is argued to reveal information flow more precisely than correlation-based measures, which are limited to symmetrical interactions. Transfer entropy requires no *a priori* knowledge of either inputs or anatomical connectivity, and therefore can be applied to spontaneously active cultures of cortical neurons whose synaptic connections are too dense to yield a full anatomical characterization. Transfer entropy can track interactions under a broad distribution of delays; in addition, it is robust to noise, small errors in spike sorting, cross talk between signals, and spike jitter or randomness [10]. One known limit of transfer entropy, however, is its robustness to processing coarse-grained data [17], an issue that we address in the results below. We also note that transfer entropy is a strictly pairwise measure, and therefore does not control for all orders of interaction between neurons.

We employ a recent extension to transfer entropy that takes into account higher-order interactions between neurons as well as a range of temporal delays [15]. The amount of TE (denoted T) from neuron j to neuron i (measured in units of bits) is given by

$$T_{j\to i} = \sum p\left(i_{t+1}, i_t^{(k)}, j_{t+1-d}^{(l)}\right) \log_2 \frac{p\left(i_{t+1}\middle| i_t^{(k)}, j_{t+1-d}^{(l)}\right)}{p\left(i_{t+1}\middle| i_t^{(k)}\right)}, \quad (8)$$

where d is a time delay; i_t denotes the status of neuron i at time t, and could be either 1 or 0, indicating a spike or no spike, respectively; and $p(\cdot)$ denotes the empirical probability of having the status denoted in parentheses. The parameters k and l determine the order of Eq. (8), and refer to the number of past time bins considered, respectively, for neurons i and j. We computed Eq. (8) for combinations of $k = (1,2,3), l = (1,2,3)$, and delays of $d = (1,2,\ldots,10)$, beyond which entropy tends to drop markedly [10]. For simplicity, we employed the maximum value of entropy taken over all combinations of k, l, and d; alternatives are examined elsewhere [15]. Equation (8) is easier to grasp when decomposed as follows. First, $p(i_{t+1}, i_t^{(k)}, j_{t+1-d}^{(l)})$ denotes the joint probability of an event involving $\{i_{t+1} = 1, i_t^{(k)} = 1, j_{t+1-d}^{(l)} = 1,\}$, obtained empirically as the count of all such events divided by the total length of the recording in milliseconds. Second, $p(i_{t+1}|i_t^{(k)}, j_{t+1-d}^{(l)})$ denotes a conditional probability, obtained by summing all events involving $\{i_{t+1} = 1, i_t^{(k)} = 1, j_{t+1-d}^{(l)} = 1,\}$, and dividing by the sum of events involving $\{i_{t+1} = 0, i_t^{(k)} = 1, j_{t+1-d}^{(l)} = 1,\}$ and events involving $\{i_{t+1} = 1, i_t^{(k)} = 1, j_{t+1-d}^{(l)} = 1,\}$. Finally, $p(i_{t+1}|i_t^{(k)})$ is obtained by summing all events involving $\{i_{t+1} = 1, i_t^{(k)} = 1\}$ and dividing by the sum of events involving $\{i_{t+1} = 0, i_t^{(k)} = 1\}$ and events involving $\{i_{t+1} = 1, i_t^{(k)} = 1\}$. The sum in Eq. (8) is taken over all combinations of i_{t+1}, $i_t^{(k)}$ and $j_{t+1-d}^{(l)}$.

Details of the implementation of the algorithm can be found elsewhere along with free software [15]. Taken as a whole, Eq. (8) describes how the future activity of a neuron is influenced by its own past history. This raises a potential bias, as it makes TE dependent upon the firing rates of both neurons i and j. We address this issue using a statistical permutation test where we calculate TE after randomizing spike times for each neuron independently. Our method of randomization takes each spike and the interspike interval immediately following it, and moves it to a different time in the data. This method of shuffling preserves the distributions of both spike rates and interspike intervals. Thus, the permutation test constitutes a statistical control for the effect of firing rates on transfer entropy.

The permutation procedure was repeated 1000 times, and a value of transfer entropy for permuted data ($T_{j\to i}^{p}$) was obtained by averaging the values of transfer entropy obtained across all repeats of the permutation procedure. The final value of TE from neuron j to i ($T_{j\to i}^{f}$) was obtained by subtracting the value of transfer entropy obtained from permuted data ($T_{j\to i}^{p}$) from the value of transfer entropy obtained from the original data ($T_{j\to i}^{d}$), and normalizing by the entropy rate (H_i) (the conditional entropy of neuron i conditional on its past):

$$T_{j\to i}^{f} = h\left(\frac{T_{j\to i}^{d} - T_{j\to i}^{p}}{H_i}\right), \quad (9)$$

where the entropy rate is given by

$$H_i = -\sum p\left(i_{t+1}, i_t^{(k)}\right) \log_2 p\left(i_{t+1}\middle| i_t^{(k)}\right). \quad (10)$$

The Heaviside function h sets $T_{j\to i}^{f}$ to zero when $T_{j\to i}^{d} < TE_{j\to i}^{p}$, in which case the randomized data account for the observed TE value. To account for estimation errors in computing transfer entropy, we analytically derived the variance of Eq. (8) as follows:

$$V_{j\to i} = \frac{1}{\eta} \sum p\left(i_{t+1}, i_t^{(k)}, j_{t+1-d}^{(l)}\right)\left[1 - p\left(i_{t+1}, i_t^{(k)}, j_{t+1-d}^{(l)}\right)\right] \times \left\{\log_2\left[\frac{p\left(i_{t+1}\middle| i_t^{(k)}, j_{t+1-d}^{(l)}\right)}{p\left(i_{t+1}\middle| i_t^{(k)}\right)}\right] - T_{j\to i}^{d}\right\}^2, \quad (11)$$

where η is the temporal precision employed for calculating transfer entropy, set to the resolution of the data (1 ms) [10]. We computed values of transfer entropy [Eq. (8)] and variance [Eq. (11)] for all pairs of neurons in a given recording. Then, we eliminated (by setting to zero) all values where $T_{j\to i}^{d}$ was less than 3 standard deviations (s.d.) above $T_{j\to i}^{p}$ [where s.d. is the square root of the variance obtained in Eq. (11)].

We plotted the distribution of transfer entropy taken over all pairs of neurons [Fig. 2(a)]. Values ranged between 0 and 4.2 bits, and were rescaled on a log-log plot between 0 and 10^2. To characterize this distribution, we assessed the relative fit of a power law compared to an exponential function using a Bayesian information criterion [Eqs. (5)–(7)]. The fit of a power law was superior (power law, 0.5862; exponential, 0.4138). The best-fitting slope of power law was $\hat{\alpha} = 1.47$. A power law of transfer entropy suggests that most information transfer between cells was relatively low, with a few interactions having disproportionally large

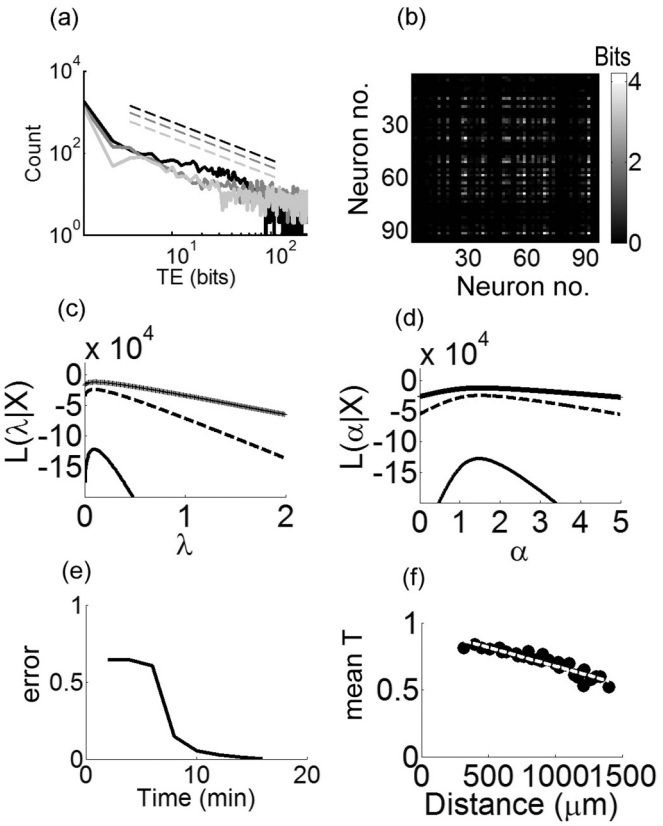

FIG. 2. Information transfer between pairs of neurons. (a) Distribution of transfer entropy obtained from all pairs of neurons, using windows of 1 ms (black), 2 ms (dark gray), and 4 ms (light gray). Dashed lines show best-fitting power law. (b). Matrix of transfer entropy for all pairs of neurons in a recording. (c) Log likelihood of the λ parameter estimated for the exponential function. Solid black, dashed, and gray lines are obtained from 100%, 50%, and 25% of the data, respectively. (d). Same as panel (c), but obtained for the α parameter of the power law function. (e). Mean square error between values of transfer entropy obtained for consecutive time windows of 1,2,..., 20 min in duration. (f) Relationship between mean transfer entropy (T) and physical distance between microelectrodes. Panels (a) and (c)–(f) are obtained after pooling data from all recordings.

values. An example of transfer entropy matrix derived from one recording is shown in Fig. 2(b). Of all (N^2–N) possible pairs of interactions in the network, only a small subset of 2.33% had a value of one bit or more. We insured that the maximum likelihood technique employed to estimate the slope of power law [Eq. (1)] was not strongly affected by the size of the dataset employed. Analyses that considered only 50% or 25% of all pairs of neurons led to similar peaks in the log-likelihood estimation of slope [Figs. 2(c) and 2(d)].

Because we performed an estimation of information transfer over a finite recording time (20 min), we sought to determine if this estimation was stable or whether it fluctuated with the length of the recording period. For this purpose, we began by calculating transfer entropy for the whole 20-min period on a cortical culture. Then, we calculated transfer entropy for periods of different duration (1 min, 2 min, ..., 19 min). We computed the mean squared error between entropy obtained for the whole 20-min period versus these shorter periods [Fig. 2(e)]. As the length of the time period increased, the mean square error decreased, indicating that the estimation of entropy converged to specific values, and that further changes due to increased recording times would likely be limited. We also confirmed, as in previous work, that values of transfer entropy decreased gradually as the physical distance between electrodes increased [$r^2 = 0.87$, $p < 9.1801\times10-^{14}$; Fig. 2(f)] [10].

In sum, our analyses show a power law distribution of information transfer with a slope near 3/2 between pairs of neurons. Next, we examine the role of this distribution on network communication by using graph-theoretic analyses of weighted networks.

C. Network economy

We aimed to determine whether the power law distribution of information transfer had any bearing on the efficiency of communication across a given network. To address this question, we calculated the shortest weighted path length r_{ij} between all pairs of nodes i and j,

$$r_{ij} = \sum_{a_{uv} \in gi \overset{w}{\leftrightarrow} j} w_{uv}^{-1}, \tag{12}$$

where w_{uv} is a connection weight and $gi \overset{w}{\leftrightarrow} j$ is the shortest path between nodes i and j. The connection between two given nodes u and v is denoted a_{uv}, and is set to $a_{uv} = 1$ when a link (u,v) exists, and $a_{uv} = 0$ otherwise. Then, we computed a measure of global network communication efficiency [29]:

$$E_g = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{r_{ij}}, \tag{13}$$

where N is the set of all nodes. In technical terms, Eq. (13) calculates the mean inverse path length in the network. Put differently, it calculates the communication over all possible pairs of neurons in the network. Based on Eqs. (12) and (13), higher weights of interaction w_{uv} lead to a lower path length r_{ij}, in turn leading to a higher global efficiency E_g.

To examine the influence of weights on global efficiency, we generated a series of synthetic datasets that followed a power law with different slopes [24]:

$$x = x_{\min}(1-\omega)^{-1/(\alpha-1)}, \tag{14}$$

where $x_{\min} = 1$ is the minimum value of x and $0 < \omega < 1$ is a uniform random variable [Fig. 3(a)]. For each dataset, we created a network of $N = 100$ neurons and k edges per neuron formed by random connections. Weights for these edges were obtained from Eq. (14). As expected, increasing the slope of the power law distribution in Eq. (14) led to a reduction in global efficiency [Fig. 3(b)]. This is not surprising given that, from Eqs. (12) and (13), networks with lower weights have lower efficiency. Thus, in principle, it would be possible to construct highly efficient networks simply by increasing the weights of edges. However, if these weights represent the strength of functional connections, increasing their value would translate to increasing the amount of communication between neurons in the network. By analogy to highway traffic, more vehicles on the road result in more goods being moved around, but also

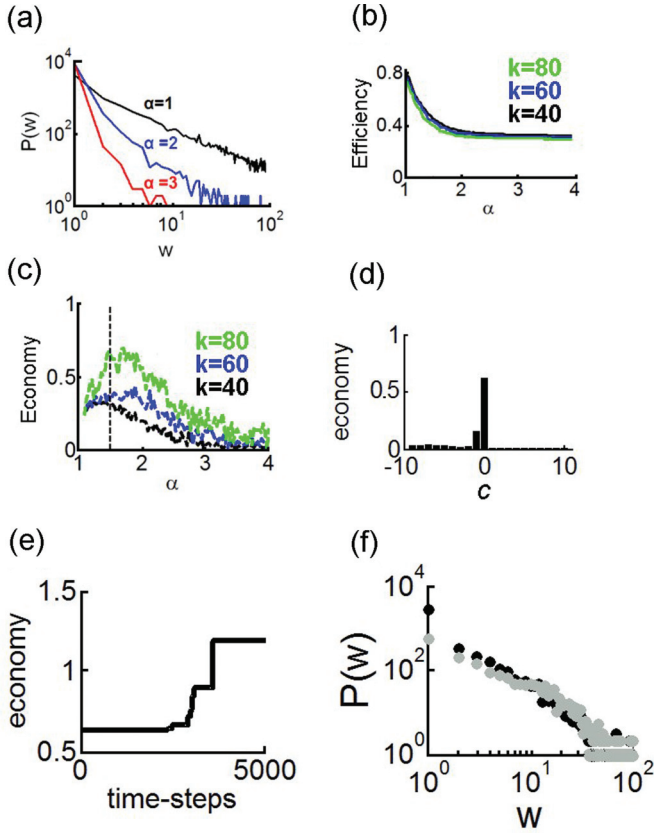

FIG. 3. (Color) Relation between power law distribution of information transfer and network economy. (a) Strength of information transfer (w) with different slopes (α). These distributions were obtained from synthetic data [Eq. (14)]. (b) Network efficiency decreases as the slope (α) of the power distribution of strength increases. Different colored lines correspond to networks where each node had k degree by forming random connections. (c) Economy is a measure that trades off connection strength and efficiency [Eq. (15)] and is highest when strength follows a power law with slope near 3/2 (dashed vertical line). (d) Network economy is diminished when network weights are increased or decreased by a constant value c. Data are averaged over all recordings. (e) Network economy improves when adjusting individual weights of a network with a search algorithm. (f) Power law distribution is largely unaltered after adjusting individual network weights. Black circles, initial distribution; gray circles, distribution following weight adjustments for 5000 time steps.

more congestion. In neuronal networks, this congestion comes at a metabolic cost associated with neuronal activity [30]. The strength of functional interactions, therefore, may benefit from a judicious trade-off between low communication on the one hand, and low metabolic cost on the other hand.

To account for both communication efficiency and the cost associated with strong functional weights, we propose a modified measure of global efficiency, termed global economy, that considers weighted path length [Eq. (12)] but penalizes for large average weights ($\bar{w}$):

$$E_p = \bar{w}^{-1} \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{r_{ij}}. \qquad (15)$$

By applying the above measure to synthetic networks generated with Eq. (14), we found that the distribution of connections across all pairs of nodes influenced global economy of the network. Network economy showed a prominent peak around a slope of $\alpha \approx 3/2$ [or slightly higher, Fig. 3(c)]. This value is comparable with empirical results obtained on cortical networks (see above).

Next, we examined global network economy [Eq. (15)] in our experimental recordings of cortical cultures. We calculated this measure for each experimental network. Then, we subtracted a small constant c from all of the weight values in the network and calculated the measure again. We repeated this procedure for a range of whole values $c = -10, -9, \ldots, 10$. Finally, we averaged the resulting value of global network economy over all experimental networks. We found that both the addition and subtraction of a small constant c led to lower values of global network economy [Fig. 3(d)]. On the one hand, low weight values reduce global network economy by driving down the value of global efficiency [Eq. (13)]. On the other hand, high weight values are penalized by the measure of global economy [Eq. (15)]. An optimal balance between these two poles was reached when $c \approx 0$, suggesting that weights attained a high value of global economy that could not be further improved by adding or subtracting a constant value to all weights in the network.

However, one possibility is that adjusting individual weights—as opposed to altering all weights in the network by a constant value—may lead to increases in global economy. To test this possibility, we took as a starting point one experimental network and calculated its global economy. Then, we employed a search algorithm to examine whether small random adjustments to individual weights could improve global economy. This search algorithm begins by selecting one weight $w > 0$ at random and modifies it by adding a value drawn from a Gaussian distribution with mean of zero and standard deviation of 0.1. Next, global economy is recalculated. If the resulting value was higher than in the original network, the new weight is maintained; otherwise, the weight is reverted back to its original value. We repeated this procedure for a total of 5000 time steps for all experimental networks. Overall, the search algorithm improved global network economy until ~4000 time steps when it began to stagnate and did not improve further [Fig. 3(e)]. Thus, it was possible to improve upon global network economy by altering individual weights in the network. The overall weight distribution, however, remained largely unaltered [Fig. 3(f)].

While other weight distributions (e.g., lognormal, exponential, etc.) may yield different results than the ones here, our goal was to examine the consequences of a weight distribution that follows a power law as observed experimentally. Overall, we found that weighted functional connections that followed a power law with a slope near 3/2 maximized the efficiency of network communication while minimizing the overall traffic burden, providing a possible consequence of scale-free interactions in neuronal networks.

D. Pairwise Poisson model

The above results show the presence of a power law distribution in two different aspects of our experimental data.

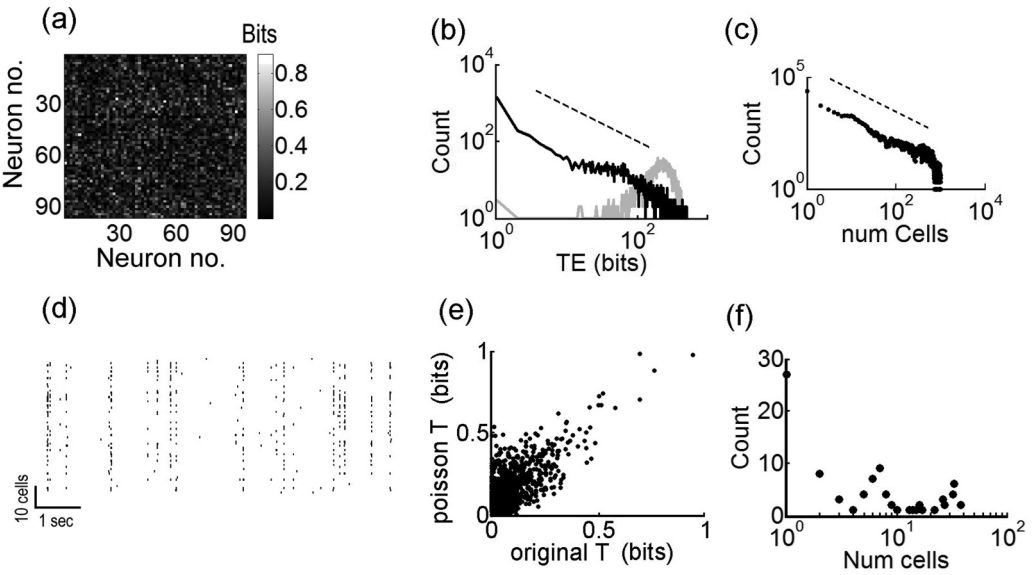

FIG. 4. Dissociation between avalanches and functional connectivity. (a) Matrix of transfer entropy for all pairs of neurons after constrained pairwise shuffling. (b) Distribution of transfer entropy for original (solid black line) and shuffled (solid gray line) activity. Dashed line, best-fitting slope of power law of original data obtained by maximum likelihood estimation. (c) Distribution of the number of cells active during avalanches obtained after shuffling. (d) Raster of spike data obtained from the pairwise Poisson model. (e). Transfer entropy (T) obtained with a pairwise Poisson model is strongly correlated with transfer entropy obtained with the original data. (f) Avalanche distribution obtained from the pairwise Poisson model. Panels (b), (c), (e), and (f) are obtained after pooling data from all recordings.

First, we reported, as in previous work, a power law of activity related to neuronal avalanches [Fig. 1(c)]. Second, we reported a power law of transfer entropy related to functional connectivity in the network [Fig. 2(a)]. It is worthwhile to consider whether these two power law distributions are related to each other or whether they constitute independent properties of network activity.

To address this question, our first step is to consider whether it is possible to artificially generate activity that preserves avalanches but disrupts functional connectivity between pairs of cells. Such "surrogate activity" can be obtained by randomizing experimental data using a technique of constrained pairwise shuffling [11]. In this technique, the spike times of two randomly chosen nodes i and j are switched, and the procedure is repeated until all spikes in a given recording have been randomized. This technique preserves the mean firing rate of all nodes as well as the number of spikes per time bin.

As a result of constrained pairwise shuffling, the matrix of weighted functional connections obtained via transfer entropy is disrupted [Fig. 2(b) versus Fig. 4(a)] and no longer follows a power law [Fig. 4(b)]. Notice that the matrix of transfer entropy obtained after shuffling contains approximately four times fewer bits of information than the original data. Neuronal avalanches, however, are entirely preserved [Fig. 4(c)]. Thus, it is possible to disrupt spike times such that functional connectivity is lost yet avalanches are preserved. Put differently, a network with weak pairwise interactions can nonetheless produce avalanches that match those observed experimentally.

Next, we address the inverse question, namely whether it is possible to disrupt avalanches while preserving functional connectivity. We developed a pairwise Poisson model where spike occurrences are influenced by both a random process and the transfer entropy calculated empirically. In this model, we begin with an $N \times N$ matrix of transfer entropy $\mathbf{P}$ obtained from a dataset where N is the number of neurons being recorded. The probability $\hat{p}_i$ of a spike occurring at time t is influenced by both a Poisson process and the matrix $\mathbf{P}$ [10],

$$\hat{p}_i = \frac{v_i - (1 - v_i\tau_r)\sum_j v_j p_{ij}}{1 - v_i\tau_r}, \tag{16}$$

where $v_i = 0.001$ is a baseline firing rate, $\tau_r = 3$ ms is a refractory period, and p_{ij} are elements of $\mathbf{P}$. We simulated 20 min of activity for 100 neurons [Fig. 4(d)], and analyzed the resulting activity using transfer entropy as above. We found a strong correspondence between the original transfer entropy (obtained from matrix $\mathbf{P}$) and the transfer entropy obtained from the pairwise Poisson model [Pearson's correlation, $r = 0.83$, $p < 0.0001$, Fig. 4(e)]. This result suggests that the Poisson model was able to capture the statistics of functional connectivity. Avalanches, however, no longer followed a power law distribution [Fig. 4(f)]. Therefore, it was possible to generate surrogate activity where functional connectivity was maintained but avalanches were disrupted.

In sum, analyses of surrogate datasets show that it is possible to maintain avalanches despite weak pairwise interactions and, conversely, to have strong pairwise interactions yet no avalanches. These results suggest that a power law distribution of avalanches and functional connectivity represent two properties of neuronal activity that may be distinguished from each other.

III. DISCUSSION

While neuronal avalanches have been reported for over a decade, their relationship to functional connectivity has

remained unclear. Here we estimated weighted functional connections in cortical cultures *in vitro* using a measure of transfer entropy, and showed that their distribution follows a power law. Using a graph-theoretic approach, we linked the slope of this power law to an efficient network communication that trades off a high volume of interactions against its cost in terms of increased traffic.

With the use of surrogate datasets, we showed that a power law of information transfer can arise independently of neuronal avalanches (and vice versa). It is therefore possible that different forms of functional connectivity support avalanches. Indeed, computational models with widely different forms of connectivity yield avalanches of neuronal activity [5–9]. Here, we add to these results by showing that even a model with weak pairwise interactions can produce prominent avalanches, and conversely that strong pairwise interactions can arise in the absence of avalanches. This result can guide the development of models that examine the mechanisms giving rise to avalanches, including both recurrent and feedforward networks [10].

Related studies have reported scale-free networks of functional connections in the brain [19,31]. An important distinction is that our work examines the distribution of functional weights and not the number of connections per node (i.e., incoming or outgoing degree). Degree distributions can be obtained by measuring the interactions between nodes in a network, and applying a cutoff to determine whether a connection is present or absent. This approach depends upon *ad hoc* thresholding, which contains several problems that are beyond the scope of the current work [32].

Related work has examined functional networks of the human brain [33], and showed that these networks attain a balanced trade-off between communication efficiency and the cost associated with a large number of connections. While it is difficult to reconcile our observations at the microscale with their results at the macroscale, our work opens the possibility of a novel trade-off that involves the strength of interactions rather than the number of links between networks nodes. Such a trade-off could in principle be examined in weighted macroscopic networks of the brain [34].

There are likely a number of principles describing the organization of functional networks of cortical neurons, including a small-world organization [11,18] and a heavy-tail degree distribution [35], as well as several other characteristics [36]. A unified framework incorporating these different requisites remains to be formulated. Importantly, this framework should link neuronal interactions with their underlying structural connectivity, known to follow a heavy-tail distribution [37,38]. Attempts at using transfer entropy for this purpose have been met with some success but definitive evidence on datasets of cultured cortical networks remains to be obtained [15].

Future work is aimed at assessing the link between avalanches and functional connectivity in other experimental networks at the micro- and macroscale. One challenge is the difficulty in evaluating avalanches [24,26], particularly in the awake brain where neuronal activity is not punctuated by periods of quiescence [39]. This issue calls into question the critical nature of brain activity, a question that is well beyond the scope of our current work.

In conclusion, networks of *in vitro* cortical neurons exchange information according to a power law distribution with a slope near $\alpha \approx 3/2$. Using surrogate datasets, we demonstrated that this distribution maximizes a trade-off between the amount of information exchanged in the network and the overall traffic burden. Because we obtained this result in generic networks with minimal biological details, it may generalize to other domains where the strength of network connections follows a heavy-tail distribution [40].

IV. METHODS

A. Tissue culture and recording

Here, we provide an overview of methods, and refer elsewhere for further details [10,21,22]. Cell cultures were recorded using a 60-microelectrode array (ALA Scientific, Germany) [Fig. 1(a)]. Dissociated cortical neuron cultures were prepared from embryonic day 18 Sprague Dawley rats (Charles River Canada, Quebec, Canada) and plated at 1.5 million cells per milliliter on polyethylinimine coated electrodes. Only cultures that exhibited a dense, homogeneous monolayer of healthy neurons were retained for recordings and data analysis. We performed all recordings using Multi Channel System (MCS) software for microelectrode arrays. Cultures grown between 23 and 28 days *in vitro* were recorded daily for a 20-min duration. Recording parameters included 1100.0 amplifier gain, input voltage range of -2048 to $+2048$ mV, and a sampling frequency of 5000 Hz. Low-frequency shifts in the raw signal were removed using a high-pass filter with a cutoff frequency of 200 Hz. We performed online extracellular spike detection using MC_RACK software (Multi Channel System), with a threshold set to 3 standard deviations above the mean of the signal at each channel. The resulting spike data were stored for offline spike sorting, performed with Plexon software (version 3.0, Plexon Inc., Texas). Spike data were analyzed using custom software written in MATLAB (Mathworks Inc., Natick, Massachusetts). We excluded from further analysis neurons whose firing rate was ± 5 s.d. above or below the mean of all neurons within a given microelectrode array.

B. Avalanches

Avalanches were identified by using nonoverlapping time bins of a fixed size, set to 10 ms by default (we also examined variations in the length of the time bin; see Results). This duration was chosen to reflect the time scale over which coordinated activity likely affects the response of downstream neurons [41]. We defined an avalanche as a series of consecutive bins where all bins possess at least one spike. In addition, an avalanche must have been preceded and followed by at least one time bin with no spikes. The number of neurons per avalanche was defined as the number of cells spiking at least once during a given avalanche.

ACKNOWLEDGMENTS

This study was funded by a grant to J.P.T. from the Natural Sciences and Engineering Council of Canada (Grants No. 210977 and No. 210989) as well as from the Canadian Institutes of Health Research (Grant No. 6105509). The author is thankful to Amy Aylsworth for technical assistance.

[1] S. Marom and G. Shahaf, Q. Rev. Biophys. **35**, 63 (2002).
[2] J. van Pelt, I. Vajda, P. S. Wolters, M. A. Corner, and G. J. A. Ramakers, Prog. Brain Res. **147**, 171 (2005).
[3] J. M. Beggs and D. Plenz, J. Neurosci. **23**, 11167 (2003).
[4] D. Plenz and T. C. Thiagarajan, Trends Neurosci. **30**, 101 (2007).
[5] A. Levina, J. M. Herrmann, and T. Geisel, Nat. Phys. **3**, 857 (2007).
[6] G. L. Pellegrini, L. de Arcangelis, H. J. Herrmann, and C. Perrone-Capano, Phys. Rev. E **76**, 016107 (2007).
[7] J. N. Teramae and T. Fukai, J. Comput. Neurosci. **22**, 301 (2007).
[8] M. Rubinov, O. Sporns, J. P. Thivierge, and M. Breakspear, PLoS Comput. Biol. **7**, e1002038 (2011).
[9] L. F. Abbott and R. Rohrkemper, Prog. Brain Res. **165**, 13 (2007).
[10] K. Vincent, J. S. Tauskela, and J. P. Thivierge, Front. Comput. Neurosci. **6**, 86 (2012).
[11] S. Pajevic and D. Plenz, PLoS Comput. Biol. **5**, e1000271 (2009).
[12] A. J. Cadotte, T. B. DeMarse, P. He, and M. Ding, PLoS ONE **3**, e3355 (2008).
[13] B. Gourevitch, R. L. Bouquin-Jeannes, and G. Faucon, Biol. Cybern. **95**, 349 (2006).
[14] B. Gourevitch and J. J. Eggermont, J. Neurophysiol. **97**, 2533 (2007).
[15] S. Ito, M. E. Hansen, R. Heiland, A. Lumsdaine, A. M. Litke, and J. M. Beggs, PLoS ONE **6**, e27431 (2011).
[16] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).
[17] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, J. Comput. Neurosci. **30**, 45 (2011).
[18] J. H. Downes, M. W. Hammond, D. Xydas, M. C. Spencer, V. M. Becerra, K. Warwick, B. J. Whalley, and S. J. Nasuto, PLoS Comput. Biol. **8**, e1002522 (2012).
[19] P. Bonifazi, M. Goldin, M. A. Picardo, I. Jorquera, A. Cattani, G. Bianconi, A. Represa, Y. Ben-Ari, and R. Cossart, Science **326**, 1419 (2009).
[20] D. Langlois, D. Cousineau, and J. P. Thivierge, Phys. Rev. E **89**, 012709 (2014).
[21] J. S. Tauskela, H. Fang, M. Hewitt, E. Brunette, T. Ahuja, J. P. Thivierge, T. Comas, and G. A. Mealing, J. Biol. Chem. **283**, 34667 (2008).
[22] K. Vincent, J. S. Tauskela, G. A. Mealing, and J. P. Thivierge, PLoS ONE **8**, e54478 (2013).
[23] H. Bauke, Eur. Phys. J. B **58**, 167 (2007).
[24] A. Clauset, C. R. Shalizi, and M. E. Newman, SIAM Rev. **51**, 661 (2009).
[25] A. Klaus, S. Yu, and D. Plenz, PLoS ONE **6**, e19779 (2011).
[26] J. Touboul and A. Destexhe, PLoS ONE **5**, e8982 (2010).
[27] S. Hélie, Tutorials Quant. Methods Psychol. **2**, 1 (2006).
[28] H. S. Bhat and N. Kumar, Technical Report, University of California, 2010, http://nscs00.ucmerced.edu/~nkumar4/BhatKumarBIC.pdf.
[29] V. Latora and M. Marchiori, Eur. Phys. J. B **32**, 249 (2003).
[30] S. Hallermann, C. P. de Kock, G. J. Stuart, and M. H. Kole, Nat. Neurosci. **15**, 1007 (2012).
[31] X. Liu, B. D. Ward, J. R. Binder, S. J. Li, and A. G. Hudetz, PLoS ONE **9**, e92182 (2014).
[32] F. Gerhard, G. Pipa, B. Lima, S. Neuenschwander, and W. Gerstner, Front. Comput. Neurosci. **5**, 4 (2011).
[33] S. Achard and E. Bullmore, PLoS Comput. Biol. **3**, e17 (2007).
[34] M. Ercsey-Ravasz, N. T. Markov, C. Lamy, D. C. Van Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy, Neuron **80**, 184 (2013).
[35] K. V. Srinivas, R. Jain, S. Saurav, and S. K. Sikdar, Eur. J. Neurosci. **25**, 3276 (2007).
[36] L. M. Bettencourt, G. J. Stephens, M. I. Ham, and G. W. Gross, Phys. Rev. E **75**, 021915 (2007).
[37] S. Song, P. J. Sjostrom, M. Reigl, S. Nelson, and D. B. Chklovskii, PLoS Biol. **3**, e68 (2005).
[38] G. Buzsaki and K. Mizuseki, Nat. Rev. Neurosci. **15**, 264 (2014).
[39] N. Dehghani, N. G. Hatsopoulos, Z. D. Haga, R. A. Parker, B. Greger, E. Halgren, S. S. Cash, and A. Destexhe, Front. Physiol. **3**, 302 (2012).
[40] A. Barrat, M. Barthelemy, and A. Vespignani, Phys. Rev. E **70**, 066149 (2004).
[41] M. R. Cohen and A. Kohn, Nat. Neurosci. **14**, 811 (2011).