Existence of high-order correlations in cortical activity

Andrea Benucci,* Paul F. M. J. Verschure, and Peter König
Institute of Neuroinformatics University & ETH Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
(Received 4 October 2002; revised manuscript received 19 May 2003; published 8 October 2003)

Neurons collect signals originating from a large number of other cells. The variability of this integrated population activity at the millisecond time scale is a critical constraint on the degree of signal integration and processing performed by single neurons. Optical imaging, EEG, and fMRI studies have indicated that cortical activity shows a high degree of variability at a time scale of hundreds of ms. However, currently no experimental methods are available to directly assess the variability in the activity of populations of neurons at a time scale closer to that of the characteristic time constants of neurons, i.e., around 10 ms. Here we integrate pertinent experimental data in one rigorous mathematical framework to demonstrate that (1) the high temporal variability in the spiking activity of individual neurons, (2) the second-order correlation properties of the spiking activity of cortical neurons, and (3) the correlations of the subthreshold dynamics, all impose high amplitude, fast variability in the population activity of cortical neurons. This implies that higher order correlations, a necessary condition for temporal coding models, must be a central feature of cortical dynamics.

DOI: 10.1103/PhysRevE.68.041905 PACS number(s): 87.19.La, 02.50.−r, 89.75.Fb, 95.75.Wx

I. INTRODUCTION

Neurons in the neocortex receive signals from thousands of other neurons [1–4]. The type of processing these circuits can perform depends on the cellular properties of the postsynaptic neurons and on the statistical properties of their presynaptic signals. Although knowledge about the former is quickly accumulating, our insight into the latter is more limited. In the visual system, stimulus properties change on a behavioral time scale of hundreds of milliseconds. Therefore, neuronal activity shows some degree of variability on this relatively long time scale [5]. As neurons with similar response properties have a higher probability to be connected [1,3], this variability does not average out. However, the temporal operations performed by neurons are characterized by time constants on the order of tens of milliseconds. Currently no experimental techniques are available to directly measure the dynamics of large numbers of neurons at this time scale. Hence, in order to understand the fast dynamics, i.e., measured at the millisecond time scale, of neuronal populations in the cortex, we resort to a theoretical approach. In particular, we want to assess whether the temporal variability in the activity of large neuronal ensembles averages out (zero-variability systems in the following), or whether, under general dynamical conditions, some degree of fast variability must emerge. Answering this question is of great importance to understand the operating conditions of cortical neurons and their encoding capabilities: in case of zero variability in the input dynamics to a given cortical neuron, it is exposed to a smooth sustained input current. In the opposite case of high-variability, it processes afferent signals that will vary strongly in both their total amplitude as well as their temporal properties.

*Corresponding author. Email address: ben@ini.phys.ethz.ch

A. Considerations from physiological data

The optimal solution to directly access the fast cortical variability would be to simultaneously record the spiking activity of individual neurons in a large ensemble of few thousands of cells. Currently, simultaneous recordings of only up to a maximum number of 10^2 neurons have been reported [6]. Hence, the available experimental results only provide indirect hints on the fast dynamics of cortical neurons. A number of relevant observations, however, are available: (a) the membrane potential (Vm) variability of single neurons at the millisecond time scale as revealed by intracellular studies [7]. (b) The reliability in the spike timing [8]. (c) The plausible presynaptic origin of the transition between so-called, up-down states [9] and (d) the high spike timing variability, as quantified by the coefficient of variation (CV) [10]. All of these have been interpreted as providing indirect hints for the existence of fast variability in the dynamics of populations of cortical neurons. However, there are many alternative, equally valid, interpretations to explain these phenomena and it remains unclear whether and under what conditions, fast variability should or should not occur.

B. Considerations from modeling studies

Sophisticated statistical analyses [11] have been used to identify patterns of activity in spike trains, interpreted as ''fingerprints'' of synchronous events involving a large fraction of neurons (i.e., high amplitude, fast variability) also called *higher order correlation events*.

In the literature, the concept of higher order correlations is used to indicate brief time intervals (a few ms) during which a large fraction of cells in the population produces a spike, thus generating what is also called a *population spike* or *unitary event* [12–14]. These spiking events represent a form of fast variability in the network's activity. Accordingly, in the present study, we will use the term ''higher order events'' (sometime omitting ''correlation'' for simplicity) to indicate the presence of fast, few ms, variability in the population

dynamics, as opposed to completely uniform activity with constant, zero variability. This former case can be associated with ''flat'' population PSTHs (i.e., the peri stimulus time histograms of the population's spiking activity), which indicate that, at any moment in time, a constant number of spikes is produced by the neuronal ensemble, and no higher order correlation events, or fast variability, appear.

The statistical significance of the results mentioned at the beginning of this section, however, is still hotly debated. Other theoretical studies either simply assumed the existence of higher order events [13], or focused on large-scale recurrent neural networks with simplified neuronal models [15–18] to demonstrate that fast variability in the form of population spikes is an emergent property of the systems studied [14].

In summary, also in interpreting the results of these theoretical approaches, we face the problem: it is not clear under which conditions, if ever, fast variability emerges.

On a positive side, the above-mentioned experimental and theoretical studies all contribute bits of information to this ''puzzle,'' i.e., the problem of the fast variability. In the remainder of the paper we will use a number of ''puzzle pieces,'' such as our knowledge of the pairwise correlations of neuronal activity and the CV of single cells. We will integrate these observations in a framework that will allow us to prove that under very general conditions, irrespectively of the preparation, cortical area, or species, etc., fast variability cannot cancel out. Subsequently we will consider the relevance of these results with respect to the information processing and coding strategies of cortical neurons.

C. Structure of the paper

After describing in detail three experimental results meant as important constraints of the neuronal dynamics (Sec. I D), Sec. II faces the question, using formal mathematical arguments, of whether it is possible to have zero fast variability in the activity of large ensembles of cells, given our experimental and theoretical knowledge of the cortical dynamics. We describe systems (ensembles of spiking neurons) without fast variability (Secs. I A and I B) and check their compatibility with physiological constraints (Sec. II C). We will make use of design theory, a branch of combinatorics, to provide examples of dynamical systems that are characterized by a total absence of fast variability and that are still fairly compatible with physiological constraints. This suggests that a more careful analysis is required to support the commonly held notion of nonzero variability in the short time scale. At the end of Sec. II, we describe the problem of classifying such systems in a formal and exhaustive way, which leads us to abandon the formalism of design theory in favor of a more classical statistical one. In Sec. III, in order to solve this ''classification problem,'' we introduce a statistical approach [random controller method (RC), Secs. III A and III B] to describe any systems with zero fast variability (including those introduced in the context of design theory in Sec. II) and verify again their compatibility with physiological constraints, Secs. III C and III D. Finally, in Sec. IV we prove the incompatibility of such systems with the cortical dynamics.[1] These steps allow us to formally prove that higher order correlations must exist in cortical dynamics. The implications of this observation will be further elaborated in the discussion.

D. Physiological constraints on neuronal processing

In order to assess whether a model accurately describes the cortical dynamics, it needs to satisfy a number of physiological constraints. The three constraints considered here are as follows:

(C1) High temporal variability observed in the spiking activity is quantified by the coefficient of variation (CV) [19], which is defined as the ratio between the standard deviation and the mean of the interspike intervals (ISI). The CV of cortical neurons is measured to be 1 or more, a variability that exceeds that of a pure Poisson process [20].

(C2) Second, the spiking activity of pairs of neurons in the primary visual cortex shows a correlation that is proportional to the similarity of their feature preference [21,22]. These experimentally observed correlations are ''weak'' in the sense that only very few spikes between pairs of spike trains are correlated. Typical correlation strength, measured as the peak amplitude of the cross-correlogram over the offset [23], is about 0.1 [21,22,24,25], implying that only 1 out of 10 spikes is in some specific temporal relationship [26]. This weak pairwise correlation has been observed in several cortical areas, in different species, and in anaesthetized as well as in awake behaving animals [22].

(C3) Paired intracellular recordings in primary visual cortex have detected correlations in the subthreshold membrane potentials of neighboring neurons [27,28]. It has been shown [27] that the normalized peaks of the cross-correlograms range around 0.4, which indicates that 40% of the time, the subthreshold membrane potential of neighboring cells in visual areas is correlated. This correlation is amplified during visual stimulation with normalized peaks as big as 0.8 [27].

Here we address the question how the above properties of cortical neurons, i.e., high CV and subthreshold and suprathreshold pairwise correlations, constrain the overall dynamics of cortical networks in the short time scale. We prove that in a neuronal ensemble, during high-input regimes, some degree of variability must exist.

II. ZERO-VARIABILITY SYSTEMS: COMPATIBILITY WITH CONSTRAINTS (C1) AND (C2)

In order to characterize the fast variability of cortical neurons, the second and higher order statistical properties (i.e., mean firing rate, pairwise correlation strength, and unitary events) of the input activity to any given target cell, compris-

[1]To help the reader follow the logical structure of the paper we will begin and end relevant sections, which are not pure formalisms or definitions, with ''Introduction'' and ''Summary'' paragraphs, respectively, each of them terminated by a ''•'' symbol. The reader who is not interested in the mathematical proofs can then easily follow the backbone of the logical structure, reading only those paragraphs.

ing 5000 to 20 000 input spike trains [4], need to be defined [3,21,24]. A qualitative observation gives an intuitive idea of a possible solution. Considering that the number of pairwise coincidences rises quadratically with the number of afferents, and that the number of spikes available to generate such coincidences rises only linearly with the number of afferents, it follows that spikes have to be ''used'' multiple times to generate coincidences, i.e., higher order correlations must appear. However, this intuition can be misleading and it raises the question whether the emergence of higher order events is a necessary consequence of the experimental constraints.

Introduction. Here we use systems (symmetric designs) that are well known in the field of combinatorics and design theory [29,30], but are not yet introduced in the study of cortical systems. For our very basic use of design theory, it is just necessary to know that this field of mathematics is extremely helpful in identifying features and emergent properties of large-scale complex systems constrained by ''some'' limitations. To provide an intuitive example of what we mean with a complex system, one can imagine a raster plot of the spiking activity of a large ensemble of cells. Design theory helps to understand the structure of such plots, and equivalently, the properties of the corresponding population PSTH.

In the fields of combinatorics and design theory [29] systems have been identified that have zero variability in their PSTH (''flat'' population PSTH) and nonzero pairwise correlations. For some of these systems the constraint of high CV is also satisfied. We now verify in detail the compatibility of these systems with the three physiological constraints introduced above individually. In the last paragraph of Sec. II B we will provide a clear example to give an intuitive idea of how symmetric designs can be generated and in which way they could relate to cortical dynamics.

A. Definitions and formalism: redefining the problem

It is useful to think of a spike train as a binary string where the ''ones'' represent the occurrence of a spike and the ''zeros'' a nonoccurrence. The length of the string is defined by the duration of the single ''bins,'' Δt_{bin}, which are used to discretize the spike train, and the number of bins N_{bin}. This is essentially the time resolution of the system: $\Delta t_{bin} N_{bin} = T$, with T being the total duration of the spike train. The string of time bins can be thought of as a group $S=\{s_1,\ldots,s_n\}$ of elements. Considering that the spikes can occur in any time bin, all elements between min(S) and max(S) must be included, and it is always possible to shift the elements such that $\min(S)=t_{start}$. In general, S can be thought of as a permutation of the elements of the set $\{\min(S),\ldots,\max(S)\}$. Therefore, the problem can be viewed in a finite field Z_n, for some integer n. Given these definitions, we assume that $S=\{0,1,2,\ldots,N_{bin}-1\}$, and that any spike train can be represented by a subset $C=\{c_1,\ldots,c_k\}$ of S. In order to faithfully represent the spike train we want every element in C to carry information about the spike timing. For example, $c_i=\lambda$ means that a spike occurs at a time $\Delta t_{bin}\lambda$. If the firing rate is constant then the size of every subset is of a fixed value k. Suppose we have $C_1,\ldots,C_m$ as distinct nonempty subsets (spike trains), then the weak pairwise correlation constraint (C2) translates to $|C_i \wedge C_j| = q$, for any pair of subsets, with q being a fixed parameter defining the number of overlaps (the strength of the correlation).

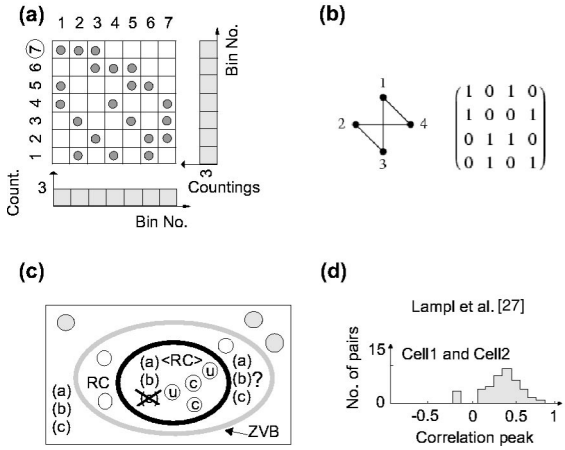

FIG. 1. Symmetric designs: (a) Representation of the Fano plane, projective plane of order 2, a symmetric (7,3,1) design. This ''squared grid'' can be associated with a raster plot of spike trains in the following way: every row represents the spiking activity of a single cell (seven in this example), where the occurrence of a spike is a filled circle and the nonoccurrence is simply an empty box. Note that the number of elements (filled circles, or ''spikes'' in the raster plot representation) along the columns is constant and equal to that along the rows (or spike trains), as indicated by the side histograms. (b) Example of a ''graph'' and its associated incidence matrix. (c) Schematic of the mathematical space where the described objects can exist. The light gray oval symbolizes the border of the group of zero-variability systems (ZVB), while the black oval is the border of the group of average RC systems ($\langle$RC$\rangle$). This latter group includes classified (C) and unclassified (U) symmetric designs (the circles with the corresponding letters), which satisfy the constraints of high CV, indicated with the letter (a), and of pairwise correlations for the spiking activity (b), but not the constraint for the subthreshold correlations, crossed (c). The space between the ovals, indicated simply with RC, represents in general the systems (open circles) produced by the RC method, which deviates from the average behavior, i.e., small N and duration time. For these cases, CV and correlation strengths cannot be fixed. (d) Histogram of the peak values of the subthreshold cross-correlated activity, adapted from Lampl *et al.* [27].

B. Symmetric designs

For the systems with a flat population PSTH described in design theory, the number of subsets $m=N_{bin}$, and these objects are called symmetric (N_{bin},k,q) designs [29,30].

One well studied example of symmetric design is the so-called Fano plane, Fig. 1(a): In this case $N_{bin}=7$. The Fano plane is a $p=2$ symmetric design, which means that $N_{bin}=p^2+p+1$, $k=p+1$ and $\lambda=1$, so blocks of different sizes can be created using higher p values. Other examples of large-scale symmetric designs are the so-called Hadamard designs [symmetric $(4t-1,2t-1,t-1)$ designs]; projective planes (symmetric $(p^2+p+1,p+1,1)$ designs), etc.

Symmetric designs interpreted as raster plots

We can interpret the Fano plane (and any symmetric design) as a raster plot, Fig. 1(a). If we assume that each of the seven columns represents an interval of 5 ms, i.e., the time resolution, the Fano plane represents the raster plot of seven neurons for a total trial duration of 35 ms. The filled circles represent the occurrence of a spike for the corresponding cell at a given time. For example, cell 4 emits three spikes, the first spike in a time between 0 and 5 ms, the second spike between 15 and 20 ms, and the third between 30 and 35 ms. All the cells have the same mean firing rate (every row has 3 spikes). Moreover, the neurons show a fixed pairwise correlation since, taken any of the possible pairs of cells, only one of the three spikes for each cell is located in the same column as for another cell. For example, both cells 2 and 6 emit a spike in a time between 10 and 15 ms and that is the only time overlap. The histogram at the bottom represents the population PSTH, which is characterized by zero variability. A large number of symmetric design systems have been reported in the literature [29]. Hence, to create a system that implements a cortical raster plot with an arbitrary number of cells and of any duration, characterized by variability at the hundreds of milliseconds time scale and no variability at the tens of milliseconds time scale, turns out to be very easy. Several symmetric designs (as the Fano plane), with duration of few tens of milliseconds, each with the same rate and correlation properties, can be vertically (temporally) aligned to obtain the desired number of rows (neurons in the population). This would form the first ''block of columns,'' i.e., the raster of the population for the first few tens of milliseconds. To provide the second and following blocks of columns (i.e., the complete temporal evolution of the system), symmetric designs with different mean rates can be used. This way the global population PSTH would show variability only in 100 of milliseconds time scale and not on a shorter time scale. Having this generating mechanism in mind, we will neglect in the following the global dynamics and focus on the ''building blocks,'' i.e., the symmetric designs.

Summary. Symmetric designs can be used as building blocks to produce raster plots (and from those PSTHs) characterized by a global variability at the hundreds of milliseconds time scale and no variability at the tens of milliseconds time scale. Such systems exhibit weak pairwise correlations, C2, as experimentally observed in the dynamics of cortical neurons. •

C. The coefficient of variation constraint (C1)

Introduction. To claim that symmetric designs can successfully describe the cortical dynamics, we need to verify their compatibility with the other constraints, (C1) and (C3). In this section we will focus on (C1), i.e., high temporal variability in the single cell spiking activity. •

In design theory, systems with high CV (coefficient of variation of interspike intervals) can be found when studying the so-called *cyclic difference sets* [29], an example using arithmetic mod(13) in the field Z_{13} is given in the supplementary material, Fig. 1(a), [42]. In this example, there are two remarkable features: the exponential that describes the distribution of the interspike intervals is necessarily incomplete and ''pathological'' cross-correlation structures, always strongly asymmetrical (i.e., nonzero abscissa for the main peak) are produced, supplementary material Fig. 1(b) [42].

None of the above zero-variability systems can describe the real cortical dynamics. They do not satisfy either the CV or cross-correlation constraint. Unfortunately, satisfactory examples cannot be found either in the context of graph theory [31], which focuses on the properties of the incidence matrix, Fig. 1(b), or advocating the so-called *balanced incomplete block designs* (BIBD, see Beth *et al.* [29] for a formal definition).

Summary. When detailing the compatibility of symmetric designs with constraint (C1) and (C2), problems arose in finding satisfactory examples. Nevertheless, no proof is available that at least one ''good'' system does not exist. We cannot proceed by enumerating examples, because ''classifying these objects is clearly impossible with present methods'' [32], i.e., we cannot make a list of all symmetric designs and exclude them one by one. Design theory has been useful to disprove the intuition that C1 and (C2) are sufficient to enforce higher order correlations. Nevertheless, to continue we have to utilize an additional approach, and return to symmetric designs further down. •

III. THE CLASSIFICATION PROBLEM AND THE RC METHOD

Introduction. To circumvent the above-mentioned classification problem and give a formal proof that zero-variability systems are not compatible with cortical dynamics at any time scale, we introduce a statistical method, the RC method (an intuitive idea of how the RC method works is given in the first paragraph of Sec. III A). This allows us to define statistical quantities, which can be directly compared with physiological and biophysical results, (C1), (C2) and (C3). This discussion will take all of Sec. III. Only in Sec. IV we will solve the above-mentioned classification problem. At the end of Sec. IV we will summarize the results and provide conclusions in relation to this statistical approach. •

A. An introduction to the RC method

To give a clear idea about what the RC method is, we could interpret again the incidence matrix associated with the Fano plane, Fig. 1(a), as a raster plot. At every time step a fixed number of cells k is specifically selected by what we call a *controller*, to make them spike, and the symmetric design structure is produced. More specifically, referring to Fig. 1(a), a generator of integers (i.e., the *controller*) would produce the sequences of triplets: (4, 5, 7), (1, 3, 7), and so on. These numeric sequences would simply mean that a filled circle (spike) should be assigned in the first column at row 4, 5, and 7, in the second column at row 1, 3, and 7, and so on, until the structure of the Fano plane is created. The raster plot of the Fano plane can then be considered as a possible subset of integers generated by a random-uniform generator of numerical sequences (RC method). Indeed, going back to the previous example, the specific sequence of integers that

would create the symmetric design, (4, 5, 7), (1, 3, 7), and so on, is a special choice between those that a random generator of triplets of integers can produce. Given the proper value for k, this method ensures that any pair of rows has, on average, the same number of intersections (fixed correlation strength) and all the rows contain the same number of spikes (fixed firing rate). From these considerations we can formulate the first important theorem.

Theorem 1. Any zero-variability system, with any pairwise correlation property, can be generated by the RC method.

Proof. Given a zero-variability system, however generated, it can be mapped into a binary matrix of zeros and ones as in Fig. 1(b), which shows its incidence matrix. The total number of elements along the columns must be constant by definition of zero variability. The row indices of the selected cells at every time step are ξ integers (4, 5, 7; 1, 3, 7; etc. in the previous example). The whole system can then be described by a sequence of integers of length $(T/\Delta t_{bin})\xi$, i.e., proportional to the total duration T, divided by the time resolution Δt_{bin}. There is always a nonzero probability that a random generator can produce such a sequence. ■

The concepts proposed here are schematized in Fig. 1(c).

B. Formalism for the RC method

In Sec. III A we said that the RC method can generate systems with fixed second order statistics, i.e., firing rate and pairwise correlations. We want now to better characterize this statement and to analyze the limitations of the method as well.

For fixed firing rate f and pairwise correlations strength c_s, a given cell a correlates with cell b on average, every $T_2=1/fc_s$ s; i.e., c_s determines the mean time between two correlated spikes. However, the mean time before two spikes correlate again is also determined by the probability P_r that the controller will reselect the cells a and b. From combinatorial analysis and by indicating the number of cells with $\gamma=2$, we have

$$P_r=\frac{\binom{N-\gamma}{n-\gamma}}{\binom{N}{n}},$$

where n is the number of spikes assigned at every step by the RC and N is the total number of spike trains. For $\gamma=1,2$, the previous expression can be approximated by $P_r=(n/N)^{\gamma}$, this way $T_1=(N/n)^2\Delta t_{bin}$, with T_1 the average time of coselection of a given pair of cells.

The method can produce systems according to the correlation and firing rate constraints, c_s and f, if the following condition is satisfied: the average frequency of occurrence of the intersections as determined by c_s, and the frequency of the random controller in selecting a given couple of cells must be the same. This results in the following condition, $T_2=T_1$, which is true if $n=N\sqrt{fc_s\Delta t_{bin}}$. This is what we call the *consistency* condition.

Theorem 2. The RC method satisfies the consistency condition if $f\Delta t_{bin}\geqslant c_s$.

Proof. If the process selects a given cell a with a frequency $1/T_1$ for the correlation constraints with cell b, this is also true in respect to the correlation constraints with all the other cells. Indeed, the frequency used by the controller to select a given cell can be computed from its probability of selection. For $\gamma=1$, $P_r=(n/N)^{\gamma}$, so the controller is selecting any given neuron with a frequency $f_{pr}=n/N\Delta t_{bin}$. Using the expression for n as above, we get the following expression, $f_{pr}=\sqrt{fc_s/\Delta t_{bin}}$. Since the selection of a cell corresponds to a spike insertion as well, it is necessary that $f_{pr}\leqslant f$ which means that $f\Delta t_{bin}\geqslant c_s$. ■

C. Consistency with cortical dynamics?

For typical values of $c_s=0.1$ and $\Delta t_{bin}=1$ ms [21], $c_s/\Delta t_{bin}=100$, so the consistency condition is achieved only for firing rates greater than 100 Hz, which is clearly not the mean rate of a cortical neuron [33,34]. Yet, increasing the value of Δt_{bin} or decreasing the value of c_s can lead to values of f that are within those reported, assuring the consistency. Thus, the two constraints applied so far are not sufficient to exclude a zero variability population PSTH.

D. Confronting the RC method with the subthreshold correlation constraint (C3)

We use now the third constraint, regarding the subthreshold correlations, to further narrow down the range of validity of zero-variability systems.

Theorem 3. The maximum and minimum fraction of time, T_f and T_u, respectively, during which two cells, whose spiking activities are produced by the RC method, can be correlated, is given by

$$T_f=1+\Delta t_{dec}\left[fc_s-2\sqrt{\frac{fc_s}{\Delta t_{bin}}}\right]$$

and

$$T_u=\Delta t_{corr}\left[fc_s+2\sqrt{\frac{fc_s}{\Delta t_{bin}}}\right]-2\Delta t_{dec}\sqrt{\frac{fc_s}{\Delta t_{bin}}}.$$

Proof. We first demonstrate the equality for T_f. As shown in Fig. 2(c) (left panel), in between two correlated spiking episodes of cells a and b (dark gray spikes on the sides), the average distribution in time of the uncorrelated spikes (light gray spikes) is such that between two consecutive spikes of cell b there will be a spike of cell a. This is because the RC method by definition minimizes the temporal overlaps between the spiking activities of the two cells. We call Δt_{dec} the time interval centered around the uncorrelated spikes of cells A and B, during which the membrane potential of the two cells cannot be correlated. This is simply due to the up-down swing of the action potential; if a cell is spiking and the other is not, the stereotyped shape of the action potential implies that the correlation coefficient between the two membrane potentials is necessarily very small, Fig. 2(a). From the consistency condition of Theorem 2, we know that the average

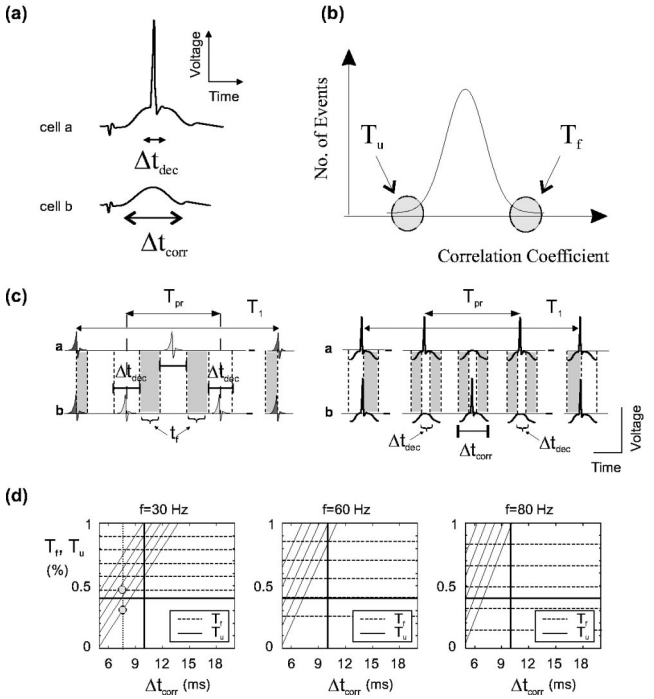

FIG. 2. The third constraint (C3). (a) When performing paired intracellular recordings from neighboring cells, a possible dynamical scenario could be characterized by common subthreshold fluctuations where in between correlated episodes, only one of the two cells spikes due to some fast noise, while the other does not. Because of the up-down swing of the action potential the correlation coefficient between the two traces computed in a time interval equal to the absolute refractoriness (Δt_{dec}) is low, but excluding such a brief event, the activity of the two cells could be on average correlated (Δt_{corr}). (b) This curve schematically reproduces the experimentally derived one as shown in Fig. 1(d). The parameters T_f and T_u can be related to the upper and lower tails of the distribution, respectively. (c) The left panel illustrates the method used to derive T_f (see text for the detailed derivation). Δt_{dec} refers to the duration of the up-down swing of the action potential (absolute refractoriness), T_{pr} is the interspike interval as derived by the RC method, and T_1 is the mean time of occurrence of correlated spikes. The shaded areas are the time intervals used to compute the ''free-time'' parameter. The right panel instead relates to the derivation of T_u. The main difference is that the shaded areas represent the intervals during which the two cells are for sure correlated ($\Delta t_{corr} - \Delta t_{dec}$). (d) The three panels show T_f and T_u as a function of Δt_{corr} for three different values of the mean firing rate: 30, 60, and 80 Hz, respectively. The five different horizontal and oblique lines refer to T_f and T_u while Δt_{dec} changes from 1 to 5 ms, respectively (1 for the top lines and 5 for the bottom lines). The two vertically aligned filled circles in the first panel indicate the values of T_f and T_u (from bottom to top, respectively), given a distribution whose mean is around 0.4, as in Fig. 1(d).

interspike interval is $T_{pr}=\sqrt{\Delta t_{bin}/fc_s}$ and the average distance in time between two correlated spikes is $T_1=T_2=1/fc_s$.

Given these definitions we compute the fraction of the total time available for possible subthreshold correlations between cell pairs, called ''free time,'' T_f. This parameter represents the upper tail of the distribution of the subthreshold correlation coefficient [Fig. 2(b)]. This is because if during the available free time the two cells are indeed always correlated, the maximum possible correlation coefficient would be achieved. This free time is schematically shown by the shaded areas in Fig. 2(c) (left panel). Its fractional value, obtained normalizing for a given interspike interval, is $T_f=[(T_{pr}-2\Delta t_{dec})/T_{pr}]+(\Delta t_{dec}/T_1)$, where the last term on the right-hand side comes from the consideration that correlated spikes add to the free time. Using the expressions for T_{pr} and T_1 as defined above, we get $T_f=1+\Delta t_{dec}[fc_s-2\sqrt{fc_s/\Delta t_{bin}}]$.

We proceed in a very similar way to compute the fraction of time during which two cells are necessarily correlated, T_u. This would represent the lower tail of the distribution of subthreshold correlation coefficients as in Fig. 2(b). First, we need to recall an experimental observation: when computing the cross-correlograms of the membrane potentials, Lampl *et al.* found that the widths of the peaks of these cross-correlograms were much broader than those computed for the spiking activity [22]. The mean value reported was around 40 ms for Vm vs <10 ms for the spiking activity [22,27]. This number indicates that when two cells are in a correlated state, they will on the average stay correlated for a duration comparable to the width of the cross-correlograms. We use this observation to introduce a correlation interval, Δt_{corr} ($\Delta t_{corr}>\Delta t_{dec}$), which is also centered around each spike. During Δt_{corr} the two cells are necessarily correlated [excluding the brief interval during the occurrence of the spike, Δt_{dec}; see Fig. 2(a)]. Δt_{corr} can be justified in the following way: Δt_{dec} has been defined before based on the stereotyped shape of action potentials. Hence, outside this time interval, we have no hypothesis about the correlation properties of the membrane potentials. Suppose instead that we have been able to measure the mean value of Δt_{dec}, then right outside this interval the two cells must on average, by definition, be correlated. An example of this scenario for the subthreshold dynamics is schematically shown in Fig. 2(a): both cells might depolarize towards threshold simultaneously, but only one emits a spike. During such an episode, the Vm of the two cells is strongly correlated, Δt_{corr}, except during the swing of the action potential, Δt_{dec}. The shaded areas in Fig. 2(c) (right panel) represent the time during which the two membrane potentials must be correlated for a given interspike interval. The fractional value is then $T_u=[(2\Delta t_{corr}-2\Delta t_{dec})/T_{pr}]+(\Delta t_{corr}/T_1)$, where the last term on the right side of the equation comes from the consideration that correlated spikes add to the time during which the two cells must be correlated. Using the expressions for T_{pr} and T_1 given above we get $T_u=\Delta t_{\mathrm{corr}}[fc_s+2\sqrt{fc_s/\Delta t_{bin}}]-2\Delta t_{dec}\sqrt{fc_s/\Delta t_{bin}}$. ■

We now want to check if there is a set of parameters $(\Delta t_{corr},\Delta t_{dec},f,c_s)$, such that T_f and T_u, i.e., the higher and lower tails of the distribution of correlation strengths, are compatible with the experimental data [Fig. 2(d)]. Every panel relates to a different average firing rate *f*, and it shows T_f and T_u as a function of Δt_{corr} for different values of Δt_{dec} with $c_s=$ set to 0.1. The thick horizontal and vertical lines at 0.4% and 10 ms indicate the mean correlation value as experimentally observed and the typical average width for

cross-correlograms of pairs of spike trains, respectively. The dotted horizontal lines correspond to different T_u values while Δt_{dec} is increasing from 1 to 5 ms (going from the top to the bottom, respectively). In the same way, the oblique lines show the behavior of T_u. The results demonstrate that there is *not* a set of values such that T_f and T_u would match the distribution found by Lampl *et al.* [27]. A very narrow distribution with a mean around 40% (as reported by Lampl *et al.*) can be obtained for $f \leqslant 30$ Hz, $\Delta t_{corr} < 10$ ms, and $\Delta t_{dec} = 5$ ms, as qualitatively indicated by the filled circles in the left part of panel *c*, corresponding to the filled circles shown in panel *b*. These values for *f*, Δt_{dec}, and Δt_{corr} are not compatible with the observed data and the dynamical conditions studied here, i.e., high input regimes. The conclusion is that the temporal distribution of the spikes produced by the RC method is not consistent with the subthreshold correlation properties observed in the experiments.

The incompatibility of the RC method with the physiological data results from its inability to capture the pertinent correlation statistics. This suggests that in order to satisfy the constraints (C1)–(C3) some degree of temporal alignments must occur. We evaluated this contention using a numerical simulation.

Numerical simulation. We show numerically that for systems that are statistically very close to those generated by the RC method, i.e., zero-variability systems, the T_f parameter is too low with respect to the experimentally reported values. Moreover, we show that only allowing variability in the population PSTH, it is possible to obtain values compatible with the experimental ones. We considered as an approximation of an RC system a group of uncorrelated Poisson spike trains. This is a simplification since Gaussian noise characterizes its dynamics, i.e., the nonflat population PSTH. We then designed an algorithm to produce weak pairwise correlated spike trains, allowing some degree of temporal alignments and high CV. Results for this simulation using a mean rate of 85 Hz and $\Delta t_{dec} = 4$ ms, are shown in Fig. 3. The lower plane (uncorrelated Poisson trains) shows a 35% value for the free-time parameter, while 55% is obtained for the correlated system, compatible with reported experimental values, Fig. 1(d). The two systems have the same CV and mean firing rate, but they differ in the correlation statistics. The important observation is that systems with some degree of temporal alignments are characterized by a free-time parameter that is compatible with (C3). This does not hold true for the uncorrelated Poisson case. The high variability close to the origin of the plot and the asymptotic behavior are important features, which will be discussed later on. Increasing the firing rate to 90 Hz leads to a decrease of T_f for the uncorrelated system to 30% and when widening Δt_{dec} to 5 ms, an additional reduction to 17% is observed. The algorithm allowed us to vary the degree of temporal alignments and accordingly the correlation strength: For $c_s = 0.2$ and unchanged rate and high CV, T_f reached 60%.

Note. Looking at the results there are two important observations: First, the expression for T_f, as soon as N is greater than 3, does not critically depend on N. Second, the variability associated with the uncorrelated Poisson system does not produce a free-time parameter compatible with

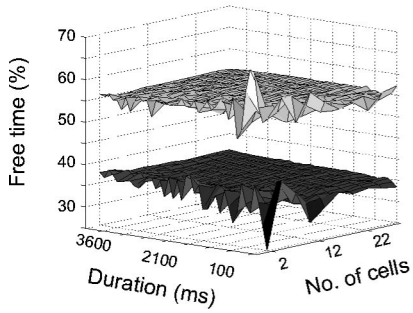

FIG. 3. Simulation results. The free-time parameter T_f is plotted against the total number of cells and the duration of the time window used to compute the mean value of the parameter for of each couple of spike trains. Such a mean has been computed averaging T_f values obtained from all possible combinations of two spike trains for a given number of cells in the sample. The upper trace refers to an ensemble of spike trains having weak pairwise correlations ($c_s = 0.1$). The lower plane instead refers to uncorrelated Poisson spike trains. In both cases the firing rate was fixed to 85 Hz. In the correlated case, characterized by temporal alignments in the spiking activity, the free-time parameter shows a higher value than for the uncorrelated case, 55% and 35%, respectively. Higher variability at the origin for both the planes is contrasted by a stable convergent behavior in ensembles having more than 2–3 cells and whose duration last longer than 100 ms.

(C3). This not only suggests that nonflat population PSTHs are necessary to match physiological conditions, but also that the required variability should exceed that of an uncorrelated Poisson system.

Summary. The RC method is a very general statistical approach to produce any system with zero fast variability, including the symmetric designs. When confronted with the physiological constraints (C1), (C2), and (C3), the systems generated by the RC method do not fulfill the subthreshold constraint (C3). Some degree of temporal alignments, or synchronized activity, is necessary. A numerical simulation supports these results. •

IV. EXCLUDING THE SYMMETRIC DESIGNS

Introduction. The average statistical behavior of systems generated by the RC method, including the symmetric designs, is not compatible with the three constraints described at the beginning. The third one, concerning the subthreshold correlations, plays a major role. We went around the ''classification problem'' and found a more general solution to create zero-variability systems and to verify their compatibility with the cortical dynamics. These results can be naturally extended to the symmetric designs.

The following *lemma* will answer the question whether or not zero-variability systems can describe cortical dynamics constrained by (C1), (C2), and (C3), thus solving the classification problem presented in Sec. III C. It comes simply as a *corollary* of Theorem 3.

Lemma 1. Symmetric designs cannot describe cortical dynamics as defined by (a) high CV at the single cell level; (b) pairwise correlations for the spiking activity; (c) pairwise correlations in the sub-threshold domain.

The proof is given in the supplementary material. Note that proving Lemma 1 is a central point to support the logic consistency of the paper: Symmetric designs have been here very useful to exemplify why constraints (C1) and (C2) cannot help in constraining the higher order correlation properties of the dynamics, which has been a controversial issue since the introduction of the synfire chain model by Abeles [35,36]. However, at the end of Sec. II, we could not solve the classification problem. If not resolved, there would have been the possibility that a ''special'' symmetric design, obeying constraints (C1)–(C3), could represent the cortical dynamics characterized by zero variability in the few millisecond-time scale. Lemma 1 proves this is not possible. •

From Theorem 3 and Lemma 1 we have the following.

Theorem 4. Dynamical systems that satisfy the constraints of subthreshold and suprathreshold correlations and high CV must display higher-order events in the population dynamics.

Proof. This is guaranteed by the validity of Theorem 3, Lemma 1, and from the observation that systems that are the complement of zero-variability ones (in mathematical terms) must have, by definition, some degree of variability, i.e., some degree of higher-order events. Moreover, the generality is guaranteed by the fact that the RC method can embrace all the systems with zero variability (Theorem 1). When the first two constraints are imposed, i.e., high CV and pairwise correlations for the spiking activity, any zero-variability system must behave according to the average behavior of the RC systems (see the proof in Lemma 1) and consequently be disregarded, since it cannot satisfy the third constraint. ■

Summary. To solve the classification problem we used the RC method. This allowed us to generate a larger group of systems characterized by a flat population PSTH, which included, as a subgroup, the symmetric designs. When checking for the compatibility between such systems and the constraints (C1), (C2), and (C3) (i.e., suprathreshold and subthreshold pairwise correlations and high CV) we found a negative answer, mainly for what concern the subthreshold correlations. Flat population PSTHs, and symmetric designs as well, cannot be compatible with a cortical dynamics constrained by (C1), (C2), and (C3). Under these conditions, some degree of variability must appear even in the tens milliseconds of ms time scale. •

V. DISCUSSION

Here, we showed that the available experimental data on the high temporal variability present in the spiking activity of individual cortical neurons, together with their pairwise correlation properties, enforce nonzero variability in the activity of populations of cortical neurons. Thus, the variability in the activity of a large number of neurons converging onto a common target does not average out. Although this was known for the slow cortical dynamics, at a time scale of hundreds of milliseconds, our results show that this is true also for the fast variability, in the few tens of milliseconds time scale. Although earlier theoretical work had suggested that this could be the case, a direct experimental or formal proof was missing. Hence, the contribution of our analysis is that by putting together fragmentary theoretical and experimental knowledge, it formally demonstrates under which general conditions fast variability must be observed. Whenever the three constraints (C1), (C2), and (C3) hold, fast variability must be observed irrespective of the preparation, type of cells, species, etc. To develop our proof we have made use of design theory. With this approach, we could identify zero variability systems that we subsequently matched against physiological constraints. We expect that this mathematical framework, new to computational neuroscience, might turn out to be a useful tool for further analyses of the complex dynamics of neuronal systems.

For our formal analysis, we relied on a number of experimental results, and we have to assess whether these apply under the relevant conditions. The numerical estimates we have used refer to values of the correlation strength, mean firing rate, and CV obtained in several species and cortical areas. However, many experiments are performed under anesthesia, and data on the neuronal dynamics in awake animals are practically not available. The situation with respect to the third constraint, subthreshold correlations, is even more problematic. Due to the immense technical difficulties of these recordings, only a few reports are available [27,28]. On the positive side, we have no reason to doubt that the results obtained in other species and cortical areas will be qualitatively different. Furthermore, the paradigms cited are those that form the backbone of much of experimental neurophysiology of the mammalian cortex. In this sense, the constraints we have considered reflect the *state of the art* in current neuroscience and we believe that our results are of general relevance.

Experimental data and theoretical studies can be put together to provide a comprehensive view and formally solve the problem of fast variability. Understanding this issue is of key relevance mainly for clarifying the consistency and feasibility of different neuronal coding mechanisms, which in one way or another do make assumptions on the fast variability. Indeed when considering widely used rate-based models, the key dynamical variable is the rate, i.e., the total count of inputs in a given time interval. Fast variability in the population activity, however, is a source of code degradation [37]. Population spikes induce high-amplitude ''noise'' in the inputs, thus degrading the reliability of its spike count. Moreover, they can elicit spikes by a target cell degrading its information transduction. The optimal working regime for such coding mechanisms would be a total absence of fast variability. In contrast, correlation-based models rely on higher-order correlation events, as a key dynamical feature used in information processing [38,36]. Neurons, acting in a coincidence detection mode, are supposed to be strongly responsive to such unitary events. These correlation phenomena are a central feature for temporal coding schemes and their computational relevance has been extensively explored in theoretical studies [35]. By providing general statements on when fast variability is necessarily observed in cortical dynamics, our study elucidates under which dynamical conditions the different coding mechanisms can work consistently and optimally.

At a cellular level, the results presented here contribute to

the hotly debated issue of cortical neurons acting as coincidence detectors or temporal integrators [39]. If the flow of signals converging onto cortical neurons is smooth in time, the average level of this input is the only available dynamical variable. In contrast, if cortical neurons act as coincidence detectors higher-order correlation events are necessary to explain, for example, the high CV in the spike timing [10,40], the bistability in the subthreshold dynamics [9,41], and spike timing reliability [8]. Thus, systems with a nonvanishing variability in the population activity are a necessary, minimum requirement for the neuron to act as a coincident detector.

The results presented here demonstrate that, in view of lack of direct experimental evidence, a rigorous mathematical analysis in combination with available physiological data can shed light onto a fundamental property of processing in neuronal circuits. Determining a set of general conditions under which variability must emerge in the population dynamics is an important step to identify the possible information processing strategies used by the cortex, and toward an understanding of the compatibility of cortical dynamics with correlation-based encoding.

ACKNOWLEDGMENTS

This work was supported by an ETH-Zürich grant, the EU IST-2000-28127/BBW 01.0208 grant, and the Swiss National Fund (SNF), Grant No. 31-61415.01.

[1] Z. F. Kisvarday, E. Toth, M. Rausch, and U. T. Eysel, Cereb. Cortex **7**, 605 (1997).
[2] P. L. Gabbott, K. A. C. Martin, and D. Whitteridge, J. Comp. Neurol. **259**, 364 (1987).
[3] R. J. Douglas and K. A. C. Martin, *The Synaptic Organization of the Brain*, edited by G. M. Shephered (Oxford University Press, Oxford, 1990), pp. 389–438.
[4] P. A. Salin and J. Bullier, Physiol. Rev. **75**, 107 (1995).
[5] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen, Science **27**, 1868 (1996).
[6] K. L. Hoffman and B. L. McNaughton, Science **297**, 2070 (2002).
[7] A. Destexhe and D. Paré, Neurocomputing **32-33**, 113 (2000).
[8] Z. F. Mainen and T. J. Sejnowski, Science **268**, 1503 (1995).
[9] C. J. Wilson and Y. Kawaguchi, J. Neurosci. **16**, 2397 (1996).
[10] C. F. Stevens and A. M. Zador, Nat. Neurosci. **1**, 210 (1998).
[11] L. Martignon, H. Von Hasseln, S. Rün, A. Aertsen, and G. Palm, Biol. Cybern. **73**, 69 (1995).
[12] S. M. Bothé, H. Spekreijse, and P. R. Roelfsema, Neural Comput. **12**, 153 (2000).
[13] S. Grün, M. Diesmann, and A. Aersten, Neural Comput. **14**, 43 (2001).
[14] A. Loebel and M. Tsodyks, J. Comput. Neurosci. **13**, 111 (2002).
[15] J. J. Hopfield and A. V. M. Herz, Proc. Natl. Acad. Sci. U.S.A. **92**, 6655 (1995).
[16] D. J. Amit and N. Brunel, Network **8**, 373 (1997).
[17] N. Brunel and V. Hakim, Neural Comput. **11**, 1621 (1999).
[18] D. J. Amit and N. Brunel, Cereb. Cortex **7**, 237 (1997).
[19] M. C. Teich, C. Heneghan, S. B. Lowen, T. Ozaki, and E. Kaplan, J. Opt. Soc. Am. **14**, 529 (1997).
[20] M. Rudolph and A. Destexhe, Phys. Rev. Lett. **86**, 3662 (2001).
[21] P. König, A. K. Engel, and W. Singer, Neural Comput. **7**, 469 (1995).
[22] W. Singer, Annu. Rev. Physiol. **55**, 349 (1993).
[23] P. König, J. Neurosci. Methods **54**, 31 (1994).
[24] E. Salinas and T. J. Sejnowski, Nat. Rev. Neurosci. **2**, 539 (2001).
[25] W. Singer and C. M. Gray, Annu. Rev. Neurosci. **18**, 555 (1995).
[26] P. König, A. K. Engel, and W. Singer, Proc. Natl. Acad. Sci. U.S.A. **92**, 290 (1995).
[27] I. Lampl, I. Reichova, and D. Ferster, Neuron **22**, 361 (1999).
[28] E. A. Stern, D. Jaeger, and C. J. Wilson, Nature (London) **394**, 475 (1998).
[29] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory* (Cambridge University Press, Cambridge, 1999).
[30] E. F. Assmus and J. D. Key, *Designs and their Codes* (Cambridge University Press, Cambridge, 1992).
[31] C. Berge, *The Theory of Graphs and its Applications* (Wiley, New York, 1962).
[32] D. Jungnickel (private communications).
[33] B. W. Connors and M. J. Gutnick, Trends Neurosci. **13**, 99 (1990).
[34] D. A. McCormick, B. W. Connors, J. W. Lighthall, and D. A. Prince, J. Neurophysiol. **54**, 782 (1985).
[35] A. Roskies, J. M. Wolfe, J. H. Reynolds, C. M. Gray, W. Singer, M. N. Shadlen, G. M. Ghose, M. Riesenhuber, C. Von der Malsburg, and A. Treisman, Neuron **24**, 7 (1999).
[36] M. Abeles, *Corticonics* (Cambridge University Press, Cambridge, 1991).
[37] M. E. Mazurek and M. N. Shadlen, Nat. Neurosci. **5**, 463 (2002).
[38] C. Von der Malsburg, Max-Plank-Institute für Biophysical Chemistry, Göttingen, Internal Report No. 81-2, 1981 (unpublished).
[39] W. R. Softky, Neuroscience **58**, 13 (1993).
[40] W. R. Softky and C. Koch, J. Neurosci. **13**, 334 (1993).
[41] J. Anderson, I. Lampl, I. Reichova, M. Carandini, and D. Ferster, Nat. Neurosci. **3**, 617 (2000).
[42] See EPAPS Document No. E-PLEEE8-68-024310 for an example of a high-CV system found using this method. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.