Analytical characterization of spontaneous firing in networks of developing rat cultured cortical neurons

Takashi Tateno,[1] Akio Kawana,[2,*] and Yasuhiko Jimbo[2]

[1]*Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka-shi, Osaka 560-8531, Japan*
[2]*NTT Basic Research Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-0198, Japan*

We have used a multiunit electrode array in extracellular recording to investigate changes in the firing patterns in networks of developing rat cortical neurons. The spontaneous activity of continual asynchronous firing or the alternation of asynchronous spikes and synchronous bursts changed over time so that activity in the later stages consisted exclusively of synchronized bursts. The spontaneous coordinated activity in bursts produced a variability in interburst interval (IBI) sequences that is referred to as "form." The stochastic and nonlinear dynamical analysis of IBI sequences revealed that these sequences reflected a largely random process and that the form for relatively immature neurons was largely oscillatory while the form for the more mature neurons was Poisson-like. The observed IBI sequences thus showed changes in form associated with both the intrinsic properties of the developing cells and the neural response to correlated synaptic inputs due to interaction between the developing neural circuits.

PACS number(s): 87.18.Sn

I. INTRODUCTION

Spontaneous electrical activity that is correlated across large numbers of neurons occurs in the brain. It has been suggested that such correlated activity plays critical roles in sleeping and waking [1], in the planning and performance of motor actions [2,3], and in the formation of the neural circuits of developing sensory systems [4,5]. Several authors have reviewed the issues of correlated activity across neurons for neural network operation, in particular, at the cortical level (for example, Fujii *et al.* [6] and references within) and in the visual system [7,8]. Although such spontaneous coordinated activity may have physiological importance, little is known about its dynamics.

The fine structure of spatiotemporal activity in the intact brain is not, however, easy to measure. To characterize a system, one generally needs a long-term recording of its stationary state, and external stimuli and interactions of connectivity with other regions of the brain can make it hard to obtain such a recording *in vivo*. Neural network models subjected to biological constraints of structural and dynamical plausibility have extensively been constructed and can shed light on the ways a complex assembly of units behaves (see, for example, several relevant books [9–11]). However, most such studies have relied on theoretical models that mimic some of the known properties of neurons. One may thus question the extent to which such models may be relied upon to provide a sufficiently realistic representation of correlated activity in networks of large numbers of neurons.

The similarity between the pharmacological effects on excitability and synaptic transmission that are observed in cultures and the effects observed in cortical slice preparation

and *in vivo* suggests that the preparation of a cultured network provides a reasonable model system for examining network activity [12–14]. Cultures of neurons thus provide a useful approximation of such truly physiological preparations as slices and intact cortices. Rather than rely on theoretical models of neural networks, we have here used networks of rat cortical neurons cultured on the surface of a multielectrode array.

A dissociated neural culture should also be useful for studying the electrical activity associated with the development of a neural network. Immature dissociated neurons reconnect in a seemingly random fashion and reconstruct networks. Even in a culture, however, the structure and connectivity of a network as a whole are hard to estimate. In this study, the spread of neurite as measured was several hundreds of micrometers and the density of neurons in the cultures was less than several hundreds of cells/mm^2. Our estimate was that it was possible for each single neuron to be connected to approximately 600 others (see Sec. II for more details).

Recent work in dynamical analysis has provided mathematical tools that are capable of evaluating nonlinear deterministic properties in time series of recorded data. Investigators have applied such tools to finding nonlinear structure in the time series of neurophysiological data [15–19]. To what extent does firing in a network behave the way a deterministic process behaves or to what extent does it behave as an effectively stochastic process? This has been a basic question in such studies. Evidence from electroencephalogram (EEG) and single-unit cross-correlation studies suggests that coordinated activity may involve a nonlinear determinism of low dimension [20,21]. However, the relatively few observations of *in vitro* neural preparations have tended to show clear evidence of nonlinear deterministic mechanisms [15–19].

Both linear and nonlinear analyses of neurobiological time series have provided useful insights into the possibilities for synaptic cording. In recent studies, in particular, Segundo and co-workers have applied nonlinear analysis to the

---

*Present address: Department of Electronics and Systems, Faculty of Engineering, Takushoku University, 815-1 Tatemachi, Hachioji-shi, Tokyo 193-8585, Japan.

sequences of interspike intervals in spike trains recorded from synaptically inhibited crayfish pacemaker neurons [22,23]. Different postsynaptic spike-train "forms," each of which denotes a set of timings with shared properties, are generated by varying the average rate and/or dispersion of the intervals of presynaptic spike trains. They reported that it was possible to characterize each form in terms of its dimension, nonlinearity, and predictability. They indicated that different dimensions, nonlinearities, and predictabilities were classified as forms that pertain to universal behavior categories and assigned names to these forms, for example, "noisy," "periodic," "quasiperiodic," and "intermittent" [24,25]. These methods of linear and nonlinear analyses are broadly applicable to any train of spikes that may be described as an ordered set of intervals, including those from synaptic arrangements, which are composed of sets of numerous weakly convergent terminals.

In this study, we started by observing the spatial "patterns" of the spontaneous activity that arises in cultured networks of developing cells. This observation was concentrated on whether or not there were correlations between cell activities. Here, a spatial "pattern" is taken to imply synchrony or asynchrony of the electrical activity of cells. We then characterized the "forms" of burst sequences, where a burst denotes a tendency for firing within a group of cells to be clustered in time. We use forms to imply types of burst-interval dispersions and sequences of bursts, which are independent of overall averages. These were much like the forms of spike trains, such as "periodic," "bursting," and "Poisson-like" used by, for example, Segundo et al. [26] and Tuckwell [27]. To find out whether the firing "timings" (descriptions of bursts as a point process) generated in a population of neurons are distinguishable from a random process, we statistically evaluated the burst-sequence data. We also applied nonlinear analysis to find out if some form of nonlinear determinism was present and to obtain some measure of complexity for the sequences.

Neurons themselves are complex systems with a physical structure that indicates many variables. At the cellular level, for instance, these include the conductances of the ion channels, the postsynaptic potentials, and the release of neurotransmitters. Networks of neurons are, of course, even more complex. The experiments in this study demonstrated that a variety of characteristic forms are imposed on the timing of bursts of neural activity. We have used the approaches outlined above to address the changes in form that arise during the concerted activity seen in the development of large numbers of cells to form networks and discuss the underlying physiological factors.

This paper is organized as follows. The materials and methods we used are described in Sec. II. One cluster process and the way we treated the recorded data in which it was apparent are described in Sec. III. The spatial "patterns" observed in the earlier developmental stage of the neural networks are described in the first part of Sec. IV A. The characteristics of the burst sequences observed in the networks of their more mature stages are then explained, and histograms of interburst intervals, autocorrelation histograms, and power spectra are used to classify the sequences

into three forms (Sec. IV B). Nonlinear analysis is then used to detect the nonlinear determinism and to estimate complexity measures of the sequences, such as correlation dimensions and degrees of nonlinear prediction error (Sec. IV C). We conclude in Sec. V with a discussion of our results.

II. MATERIALS AND METHODS

A. Cell culture

We used a slight modification of the cell-culture method used by Muramoto et al. [28]. The cortices of 17-day-old Wistar rat embryos were chopped into small pieces, digested in 0.02% papain (Boehringer), and then mechanically dissociated by trituration. The cells were then resuspended in Dulbecco's modified Eagle's medium (Sigma) supplemented by 5% heat-inactivated fetal bovine serum (HyClone), 5% heat-inactivated horse serum (Gibco), 2.5 μg/ml of insulin (Sigma), and penicillin-streptomycin (5–40 U/ml, Sigma). Electrode arrays that had been coated with poly-D-lysine (Sigma) and laminin (Sigma) were then plated with the cells. These neurons were then cultured in an incubator at 37 °C with an H_2O-saturated atmosphere, consisting of 90% air and 10% CO_2.

B. Estimation of connectivity

The degree of connectivity in an intact cortex is high and it is estimated that each neuron makes several thousands of connections to other neurons [29]. Unlike a cortex in vivo, however, a cultured cortical network is intrinsically localized. It is thus difficult to precisely characterize connectivity within a particular preparation. We estimated cell connectivity in the following way. First, we administered an intracellular injection of Lucifer Yellow (Sigma) and then measured the arborization of the neurons by visual inspection; we found this to be 1.2 ± 0.5 mm^2 [mean $\pm$ SD (standard deviation), $n=12$ neurons]. The density D of our cultures was $1.4\pm0.6\times10^2$ neurons/mm^2 ($n=12$ cultures). We then used the following method to infer the number of cells (N_{cell}) that may be connected to a given cell. We assume that two neurons are, respectively, located on disks of radius r_1 and r_2 and that the disks have a dense covering of axons and dendrites. If the distance l between the centers of the two disks is less than r_1+r_2 (i.e., $l<r_1+r_2$), a connection between the neurons is possible. Furthermore, if the axons and dendrites of every neuron in the culture are assumed to have exactly the same length ($2r$), it is then possible for each neuron to have connections with all other neurons that have centers within a circle of radius $2r$ around the center of the first neuron's disk. As a result, $N_{cell}=4\pi r^2 D$ provides an estimate of the number of connections per cell. It would thus be possible for each neuron of the cultures in this study to be connected to about 633 ($1.2^2\pi\times1.4\times10^2$) other neurons.

C. Recording

We used an electrode-array substrate [30–32] that had 64 recording sites within a 1.6×1.3-mm^2 area to make extracellular recordings of the activity of neurons in 44 cultures.
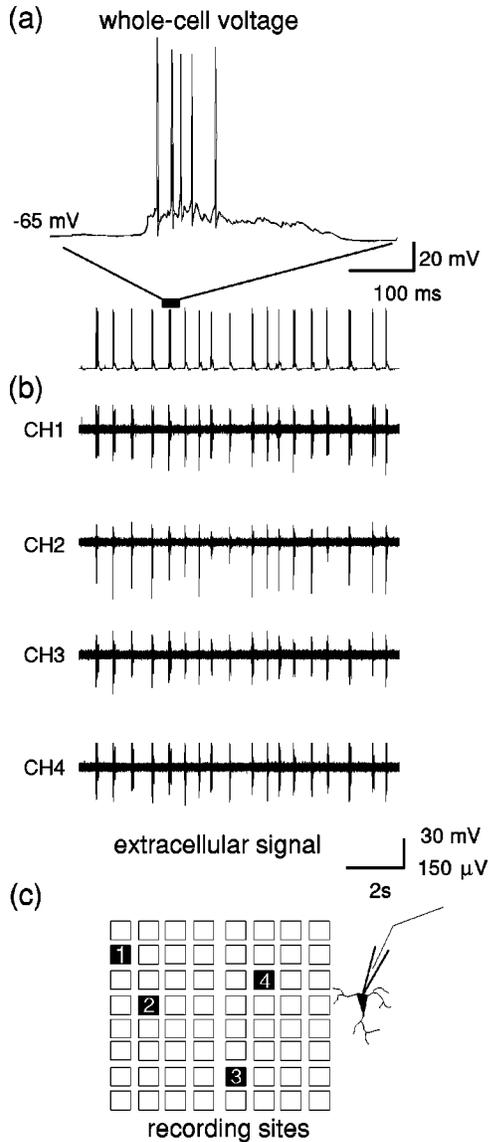
FIG. 1. Spontaneous synchronized firing in a culture at 25 days *in vitro* (DIV). (a) Action potentials of a cell under a whole-cell voltage clamp to the resting level (-65 mV). (b) Extracellular voltages recorded at four sites, along with the simultaneously recorded whole-cell voltage. (c) The four selected recording sites.

Each recording site covered an area of 30×30 μm^2, and the distance between the centers of adjacent sites was 180 μm [33] [Fig. 1(c)]. For this experiment, we used a special incubator with a measuring system of the same design as one that had been built by Jimbo *et al.* [33].

Recording sessions lasted for 5–8 h during which the cells were kept at 37 °C, and the session took place after the cells had been cultured for 3–65 days *in vitro* (DIV). The day before each recording session, the culture medium was exchanged for a medium that contained neither insulin nor penicillin-streptomycin, but was otherwise the same. A digital signal processor stored the amplitudes and widths of those spikes where the recorded signal rose above a threshold set at five times the standard deviation of the baseline noise. The amplitudes and widths recorded at each of the 64 measuring

sites were sorted and classified by using the hierarchical clustering method [34] and then converted into pulse trains, each of which was assumed to represent the response of a single cell [35]. Trains of spikes identified in this way exhibited autocorrelation functions with low values around the origin: this is compatible with their having been generated by individual cells that become refractory after firing.

Standard single-cell statistics such as average intervals, standard deviations, coefficients of variation, histograms, and average firing rates were also routinely computed. We principally used these, however, to assist in the interpretation of the other measures. Basic statistics (i.e., sample sizes, average intervals, standard deviations, and coefficients of variation) for spike trains of individual neurons obtained from nine cultures are shown in Table I and the sets of spike trains are labeled as neuron number 1–9.

D. Detecting bursts

We identified bursts in each spike train by applying the method described in Sec. III. Each was considered as an individual event; a series of bursts was considered to be a realization of a point process with a timing that was described by the sequences of interburst intervals. In short, the diagnosis of a train as "bursty" is based on the predominance of two clearly different time scales in the interspike intervals (ISIs), which in turn led to clearly separate modes in the ISI histogram (not shown here). We took the longer of the intervals or the sum of the intervals as the interburst intervals. Figure 1(b) shows the synchronization of firing between cells in a culture at 25 DIV.

III. THE CONCEPT OF THE CLUSTER PROCESS

In this study, bursts imply tendencies to fire synchronously and in mathematical parlance constitute "clusters" [36,37]. Clusters are used to indicate a set of firing events that are closely spaced in time. Our method for identifying clusters (bursts) in each of the spike trains we analyzed is explained bellow. Each cluster was taken to be an individual event; each series of clusters was considered to be the realization of a point process with a timing described by the

TABLE I. Summary of statistics on spike trains.

		Statistics on spike trains			
Neuron no.	Days *in vitro*	Sample size	Average interval (s)	Standard deviation (s)	Coefficient of variation
1	5	5949	22.0	83.1	3.77
2	7	4753	4.58	7.18	1.57
3	12	10092	3.53	5.03	1.42
4	15	3888	11.3	14.6	1.29
5	16	7197	10.1	15.7	1.56
6	17	8214	3.43	11.3	3.29
7	28	6447	2.42	5.49	2.27
8	30	6582	2.25	2.66	1.18
9	58	5578	3.72	7.55	2.03

sequences of intercluster intervals (ICIs) or interburst intervals (IBIs). In such a cluster process, a primary process determines the timing of the clusters. Each of the primary events triggers a secondary series. Figure 2(a) is a simple illustration of a cluster process.

Grüneis *et al.* have proposed a method for estimating the statistical parameters of a cluster process when the primary process is a Poisson process [37]. In the paper presented here, we have adopted their model as a framework. We start by applying their parameter-estimation method for a Poisson cluster process to the experimental data. After that, we estimated the times at which the primary process had created events.

Here is a brief explanation of the model. Suppose that $N_{tot}(t)$ is a stochastic variable that denotes the number of events in a time interval $[0,t]$. Let $\langle n_c \rangle$ denote the mean rate of the primary process (Poisson process). The cluster consists of N random events where N is a stochastic variable with the distribution function $p_m = \text{Prob}\{N=m\}$. The model assumes that the clustered distribution is described by

$$p_m = m^z \Big/ \sum_{\hat{m}=1}^{N_0} \hat{m}^z \quad (m=1,\dots,N_0), \tag{1}$$

where z is a real number. The function p_m is called a clustered distribution and m takes on values in the range $1,2,\dots,N_0$, where N_0 is the maximum value of m. The time interval between the kth and $(k+1)$th events in the cluster is

(a)

clusters

(b)
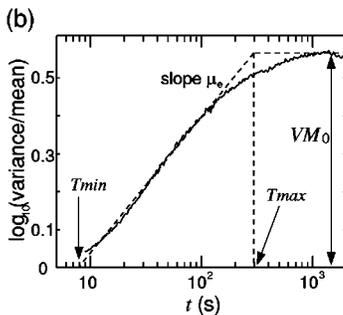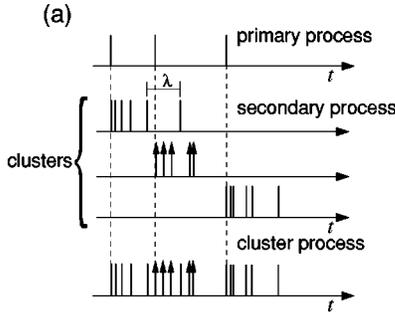
FIG. 2. (a) Illustration of a cluster process. In such a process, each of a series of primary events is assumed to form a so-called renewal process. Each of the primary events triggers a secondary series of events. Each series of secondary processes is called a "cluster." (b) The variance/mean (Fano factor) curve for numbers of spikes $N(t)$ versus period of observation t (s) on a log vs log scale. The experimental parameters (T_{min}, T_{max}, VM_0, and μ_e) that are derived from the curve are indicated. The culture had been 17 DIV. The neuron is listed as neuron number 6 in Tables I and II.

denoted by λ_k. As a definition, the first event in the cluster to be considered is called the primary event. A cluster containing i events thus exhibits $(i-1)$ intervals and the time over which i successive events occur is

$$\Lambda_i = \sum_{k=1}^{i-1} \lambda_k \quad (i=2,3,\dots). \tag{2}$$

The distribution of Λ_i is expressed as

$$F_{i-1}(t) = \text{Prob}\{\Lambda_i \leqslant t\} \quad (i=2,3,\dots). \tag{3}$$

In this study, the intervals between events within the clusters are assumed to be (νth order) gamma distributed with a distribution density function expressed as

$$w_1(\lambda) = \frac{\nu\beta}{(\nu-1)}(\nu\beta\lambda)^{\nu-1}\exp(-\nu\beta\lambda) \quad (\nu=1,2,\dots), \tag{4}$$

where $\beta = E[\lambda]^{-1}$ and $E[\cdot]$ represents expectation. When $\nu=1$, the process is Poissonian and the intervals are exponentially distributed. We thus have the probability distribution function of the intervals t of i successive events

$$w_{i-1}(t) = \frac{\nu\beta}{\{\nu(i-1)-1\}}(\nu\beta t)^{\nu(i-1)-1}\exp(-\nu\beta t)$$
$$(i=2,3,\dots). \tag{5}$$

The distribution of Λ_i is given by

$$F_0(t) = 1, \tag{6}$$

$$F_{i-1}(t) = 1 - \sum_{k=1}^{i-1} w_k(t) \quad (i=2,3,\dots). \tag{7}$$

We then obtain

$$\text{Prob}\{N_{tot}(t) \geqslant i\} = \text{Prob}\{N_{tot}(t) > i-1\}\text{Prob}\{\Lambda_i \leqslant t\}$$
$$= R_N(i-1)F_{i-1}(t) \quad (i=2,3,\dots), \tag{8}$$

where

$$R_N(i) = \text{Prob}\{N_{tot}(t) > i\} = \sum_{m=i+1}^{N_0} p_m. \tag{9}$$

Since the cluster constitutes a renewal process (Poisson process), we obtain the following expression for the probability density of $N_{tot}(t)$:

$$\text{Prob}\{N_{tot}(t) = i\} = \text{Prob}\{N_{tot}(t) \geqslant i\} - \text{Prob}\{N_{tot}(t) \geqslant i+1\}$$
$$= R_N(i-1)F_{i-1}(t) - R_N(i)F_i(t)$$
$$(i=2,3,\dots). \tag{10}$$

Next, we briefly explain the procedure we followed to fit the data to the model. The Poisson cluster process with intervals λ that are exponentially distributed is sufficiently described by four parameters: z, N_0, $\langle n_c \rangle$, and β. The additional parameter ν is needed when the intervals are gamma

Neutron no.	$\langle N_c \rangle$	$\langle N \rangle$	$\langle \lambda \rangle$ (s)	N_0	z	β (units of $1/\langle \lambda \rangle$)
1	9.63[-3]	4.74	8.09[-1]	5	6.31	1.23
2	2.55[-2]	8.56	4.15[-2]	18	−2.18	2.40
3	2.94[-2]	9.62	2.16	45	−1.05	0.464
4	6.04[-2]	2.36	5.89[-2]	10	−2.81	17.0
5	2.97[-3]	3.34	1.39[-2]	14	−2.13	72.2
6	1.24[-1]	2.35	3.23[-2]	8	−1.43	31.0
7	1.82[-1]	2.27	1.50[-1]	15	−1.89	6.68
8	2.20[-1]	2.03	1.72[-2]	48	−2.30	58.0
9	8.97[-2]	2.88	2.02[-1]	15	−1.57	4.96

distributed. We assumed that the intervals were exponentially distributed and that it was, in fact, possible to approximate the distribution functions of the intervals in our recorded data by assuming $\nu = 1$. We thus have to derive four parameters from the experimental data. For this purpose, we used counting statistics on the Poisson cluster process. That is, the number of events $[N_{tot}(t)]$ counted in a time interval t is used as a random variable in obtaining the estimate. The variance-to-mean values (also known as a Fano factor curve) for $N_{tot}(t)$ were then calculated and plotted on a log vs log graph to obtain the variance/mean curve. As is shown in Fig. 2(b), parameters T_{min}, T_{max}, and VM_0 are readily obtained from such a plot, where VM_0 is the saturated value of the variance/mean curve. In addition, a redundant parameter

$$\mu_e = \frac{\log_{10}(VM_0)}{\log_{10}(T_{max}/T_{min})} \qquad (11)$$

is used to represent the line of fit to the slope μ of the variance/mean curve under the condition that $E[N^2]/E[N] = VM_0$. If this condition is satisfied, the value of μ depends on two parameters, N_0 and z. Best-fit values for N_0 and z are thus obtained for $\mu \approx \mu_e$. In practice, since N_0 is an integer and z is a real number, it is convenient to start by finding N_0 first and then obtain z. We obtain β by comparing either T_{min} or T_{max} with the abscissa of the variance/mean curve. Finally, we derive $\langle n_c \rangle$ from $\langle n_c \rangle = E[N_{tot}]/E[N]$, where the expectation $E[N]$ of N is calculated from known values N_0 and z. Examples of parameters estimated from the experimental data are given in Table II.

We now estimate the timing of the primary process. Suppose that $N_W(t)$ is a stochastic variable that denotes the number of events in some time interval $[t, t+W](t, W \geq 0)$ within the cluster. The probability density of the number of spikes in that interval is then given by

$$\text{Prob}\{N_W(t) = m\} = \text{Prob}\{N_{tot}(t+W) - N_{tot}(t) = m\}$$

$$= \text{Prob}\{N_{tot}(W) - N_{tot}(0) = m\}$$

$$= \text{Prob}\{N_{tot}(W) = m\}$$

$$= R_N(m-1)F_{m-1}(W) - R_N(m)F_m(W)$$

$$(m = 2, 3, \dots), \qquad (12)$$

because $N_{tot}(t)$ is stationary and $\text{Prob}\{N(0) = 0\} = 1$. Thus, we estimated the timing of the primary process in the following way. If the total recording time T is divided into small bins, each with a time width of Δt, the number of spikes within each bin obeys the probability distribution $\text{Prob}\{N_{\Delta t}(t) = m\}$ under the condition that no two secondary series overlap. The number of spikes, $U_{\Delta t}(i)(i = 1, \dots, M)$, in each bin was then counted and $U_{\Delta t}(i)$ was smoothed by using a Hanning window. In practice, bins with widths from 20 ms to 100 ms were used. The times at which local extrema exceeded a threshold were regarded as indicating events of the primary process. The threshold was set as $\langle N_{\Delta t} \rangle - \sigma$, where $\langle N_{\Delta t} \rangle$ and σ are, respectively, the mean and standard deviation of $N_{\Delta t}$.

IV. RESULTS

A. Spontaneous firing patterns in networks of developing cortical neurons

The cultures on the electrode arrays had areas in the range from 10 to 50 mm² and contained 4000–10 000 cells. Spontaneous activity varied from culture to culture and according to the stage of development. Some typical spontaneous firing patterns were, however, apparent: (i) continual asynchronous firing of cells, (ii) alternation between asynchronous spikes and synchronous bursts (local synchronization), and (iii) exclusively synchronized bursts [14].

Figure 3 shows how the spike-train patterns observed in one culture changed from 3 to 7 DIV. At 3 DIV, both asynchronous firing and locally synchronized bursts with a long spike-propagation delay (3–5 ms/μm) are visible. At 5 DIV, globally synchronized bursts with a shorter spike-propagation delay represent the dominant pattern of activity. At 7 DIV, the spike trains observed at all sites of the electrodes show a much tighter coincidence than they did at 5 DIV. Average numbers of spikes within each of the bursts recorded at each electrode had also increased.

It was not possible to discern any changes in spatial patterns of firing by naked-eye observation after synchronized firing had begun to occur. The average SD and coefficient of variation ($\mathcal{C}$) of ICI (IBI) sequences for 176 cells from all 44 cultures are shown in Fig. 4. Both the average and SD of the ICIs decreased as the number of days *in vitro* increased, while the $\mathcal{C}$ values varied greatly from culture to culture and showed no tendency to change.

Since some of the analysis in this section involves the assumption that the process is stationary, it is desirable to start by verifying that the sequences of clusters are stationary. A point process is (strictly) stationary according to the mathematical definition when the joint distribution of the number of events in any k fixed intervals is invariant under translation [38], that is, when the distribution is the same in any pair of intervals. However, it is difficult to use the experimental data to satisfy this strict definition. A more practical definition (i.e., of weak stationarity) requires only that certain first- and second-order statistical properties remain invariant. In practice, some criteria that discriminate one data
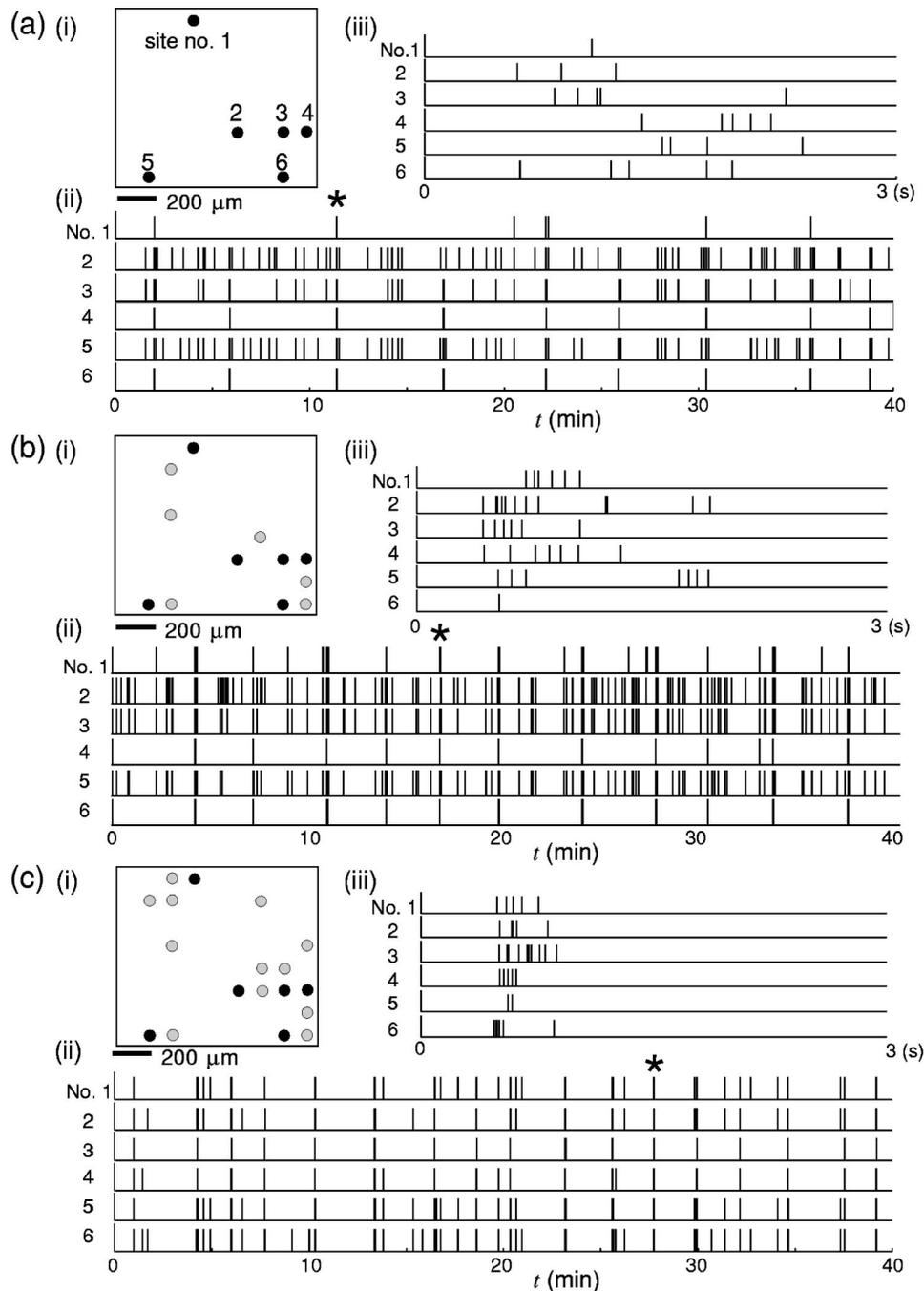
FIG. 3. Changes in spontaneous firing (spike trains) in a developing culture at (a) 3 DIV, (b) 5 DIV, and (c) 7 DIV. (i) Spatial configuration of electrodes by which signals were detected. (ii) Spontaneous firing over 40 min. (iii) An expanded view of the 3-s period marked with an asterisk in (ii).

set from another are necessary in order to recognize stationarity. In the present work, stationarity was equated with freedom from trends. This was judged by Kendall rank-correlation tests, a practical application of the Kendall coefficient τ [39]. The tests (used at the 0.05 confidence level) measure the correlation between the orders and magnitudes of the values. In particular, the results of the tests for the "Poisson" form are discussed in Sec. IV B 2.

B. Stochastic analysis of cluster sequences

In this section, we investigate the "forms" of the series produced by the cluster process (primary process) during the development of cultured neural networks. This is based on analysis of the "forms" as stochastic point processes. It is possible to deduce many forms on an *a priori* basis, and the main idea of doing so is that the formal features of each would carry implications as to the subjacent physiology. Ac-
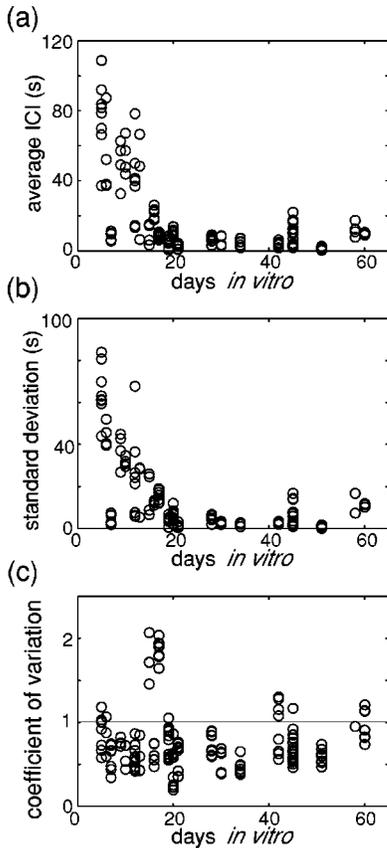
FIG. 4. Statistics on sequences of intercluster intervals (ICIs) for 176 cells from the 44 cultures, at 7–65 days in vitro (DIV). (a) Average ICI versus DIV. (b) Standard deviation of ICIs versus DIV. (c) Coefficient of variation versus DIV.

cording to Tuckwell's classification [27], it is possible to distinguish ten different major types of ISI distribution for sequences of spike trains. Here, three forms found for ICI densities in the experimental data were used. Distributions of ICIs that are close to a quite narrow Gaussian, bimodal (or multimodal), and exponential distribution are referred to as "delta-like" (periodic), "bimodal" (bursting), and "Poisson" forms, respectively [27]. Bursting as used here refers to a burst of clusters (e.g., a doublet, triplet, or short train of high-frequency clusters) and this differs from our usage elsewhere in this paper. Figure 5A shows typical examples of the three types of ICI histograms.

In this study, however, "classification" indicates generalizations found by study of the cluster sequences. Most histograms of ICIs are clearly of one type or another, but the separation of types is not absolute. The Poisson process is, for instance, a theoretical construct, so we do not intend to strictly verify that a cluster sequence is produced by a Poisson process, but rather to find a form that we refer to as "Poisson-like." Even though the separation of types is not always absolute and there is a continuous gradation from one type to another, we need some criteria for classification of the types in order to characterize the forms of the cluster sequences. Thus, in identifying the Poisson-like form, for example, the relevant comparison is with a flat autocorrela-

tion histogram of ICIs and an autocorrelation function of the Poisson point process.

1. Tendencies to periodicity in cluster sequences

First, we were interested in finding tendencies for ICIs to repeat periodically. This was important for the "delta-like" form, so we explored and measured the correlational properties of the series. A typical way of evaluating repetition is to compute the correlation functions or histograms (CFs) for series of events. The CFs for the cluster series were calculated on the basis of the methods described by Bryant *et al.* [40]. In particular, the autocorrelation histogram or autocorrelogram provides an estimate of the mean probability of firing (or a cluster in this study) for a neuron as a function of time subsequent to the beginning of a spike (cluster). Since autocorrelation histograms (autocorrelograms) are symmetrical and always have a peak at a lag of zero, we have only plotted the values for positive time and thus excluded the uninteresting delta-like function at the origin (Fig. 5B).

The "delta-like" histogram for one set of ICIs that is shown in Fig. 5A(a) has a single sharp mode with a small amount of spread in both directions. Figure 5B(a) shows that the corresponding autocorrelation histogram is highly "periodic" with several modes at integer multiples of the ICI histogram's mode and that such modes extend throughout the range of the autocorrelation histogram. Figure 5B(b) shows an "early mode" type, which has an obvious peak that corresponds to the modal value of the ICI histogram [Fig. 5A(b)]. Such peaks are occasionally followed by additional and smaller peaks. This type suggests a weak tendency toward periodic firing in clusters or bursting-type firing. The third is a "flat" type of autocorrelation histogram [Fig. 5B(c)] that shows only small deviations from the average value. This indicates that the firing times are independent of each other, except, of course, for the early effect of the refractory period of the cells. Thus, the ICI histogram is a special type of asymmetrical histogram in which the mode is close to the extreme left of the range and the decay to the right is nearly exponential, as is suggested by Fig. 5A(c).

Another way of evaluating the correlational properties of a series is to compute its power spectrum. This is generally achieved by estimating the power spectrum of the correlational properties across different shifts of time lags. We used this standard method to estimate the power spectrum [41] and found tendencies to periodicity in the cluster sequences that were within the criteria for assigning significance. This was achieved by testing whether or not deviations of the oscillatory components from the average of the power spectrum were statistically significant. Power spectra of the cluster sequences are plotted in Fig. 6 and the three types (i.e., periodic, bursting, and Poisson forms) shown are the same as those in Fig. 5.

The criterion used was three times the standard deviation (3σ) from the average on a logarithmic scale. If the oscillatory components in the frequency domain are beyond the criterion (the dashed lines shown in Fig. 6), we regarded the sequences as having a strong tendency to repeat periodically [e.g., Fig. 6(a)] and classified the sequence as being of the
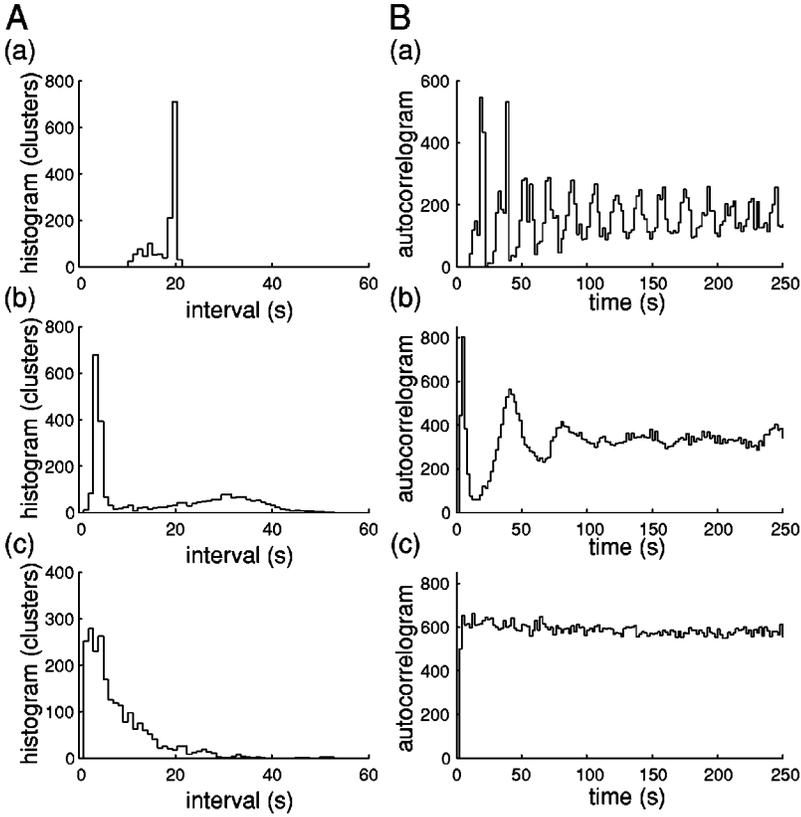
FIG. 5. Intercluster interval (ICI) histograms (A) and autocorrelation histograms (B, autocorrelograms) for single cells. For the ICI histograms and autocorrelograms, 1-s and 2-s bin widths were, respectively, used. (a) A culture at 12 DIV. $N_c = 1395$, $\mu_c = 17.3$ (s), $\sigma_c = 2.64$ (s), and $\mathcal{C}_c = 0.1524$. (b) At 18 DIV. $N_c = 2676$, $\mu_c = 16.4$ (s), $\sigma_c = 13.9$ (s), and $\mathcal{C}_c = 0.852$. (c) At 58 DIV. $N_c = 2283$, $\mu_c = 7.76$ (s), $\sigma_c = 7.58$ (s), and $\mathcal{C}_c = 0.977$. N_c is the number of clusters and μ_c, σ_c, and $\mathcal{C}_c$ are the ICI mean, standard deviation, and coefficient of variation, respectively.

delta-like'' form. Oscillatory tendencies were visible in the cluster sequences from many neurons at 15–20 DIV. However, the number of neurons that exhibit no oscillatory tendency increases at above 21 DIV. All of the results are summarized in Table III.

2. Tests of ICI series to identify the Poisson-like form

In this section, we investigate the characteristics of the cluster sequences in terms of the Poisson process. We investigated the "delta-like" (periodic) form of the cluster sequences in the preceding section. Discriminating the "Poisson" and "bursting" forms from each other in a way that may be inferred from the recorded data, in particular, turns out to be the main issue in this section. However, as stated earlier, verifying the Poisson property of an actually recorded sequence generally requires that many tests verifying the property must be satisfied; we are thus seeking a form we call Poisson-like, rather than the actual Poisson form, in the cluster sequences. Thus, to investigate whether or not cluster sequences displayed a "Poisson-like" form, we tested (i) the goodness of the fit of the distribution function and (ii) the independence of the intervals, on the basis of the method of nonparametric statistical inference [39].

If N sample observations are the values of a random variable, a comparison of the observed and expected cumulative relative frequencies is possible for each of the different observed values. Hence, for a given N-sample observation, several statistical measures (tests) of goodness of the fit are functions of the deviation of the observed cumulative distribution (empirical distribution) from the corresponding cumulative probability as expected under the null hypothesis. A variety of functions of the deviations are used in such tests. We have used the Kolmogorov-Smirnov one-sample statistic, which we briefly explain below.

We let $F_X(x)$ be an unknown cumulative distribution function and $S_N(x)$ be the empirical distribution function of $F_X(x)$. In practice, $S_N(x)$ is calculated as the sample repartition of ICI sequences $\{I_i\}$ ($i = 1, \dots, N$). That is, it is the proportion of sample observations that are less than or equal to x for some real number x,

$$S_N(x) = \frac{(\text{number of } I_i \leq x)}{N}. \quad (13)$$

Let $F_0(x) = 1 - \exp(-\lambda x)$ be the cumulative distribution function of a Poisson process that has a rate $\lambda = 1/(\Sigma_{i=1}^N I_i / N)$. The null hypothesis H_0 we tested was thus

$$H_0: \quad F_X(x) = F_0(x) \quad \text{for all} \quad x, \quad (14)$$

which is called a two-sided test of goodness of the fit. $S_N(x)$ is the statistical image of the distribution produced by $F_X(x)$. If the null hypothesis is true, there should be little difference, beyond sampling variation, between $S_N(x)$ and $F_0(x)$, for any x. To test the empirical distribution function for a Poisson process, we used the two-sided Kolmogorov-Smirnov (KS) statistics D_N,

$$D_N = \sup_x |S_N(x) - F_X(x)|. \quad (15)$$

This statistics is based on the maximal absolute measure of the "distance" between $F_X(x)$ [$F_0(x)$ under the null hypoth-
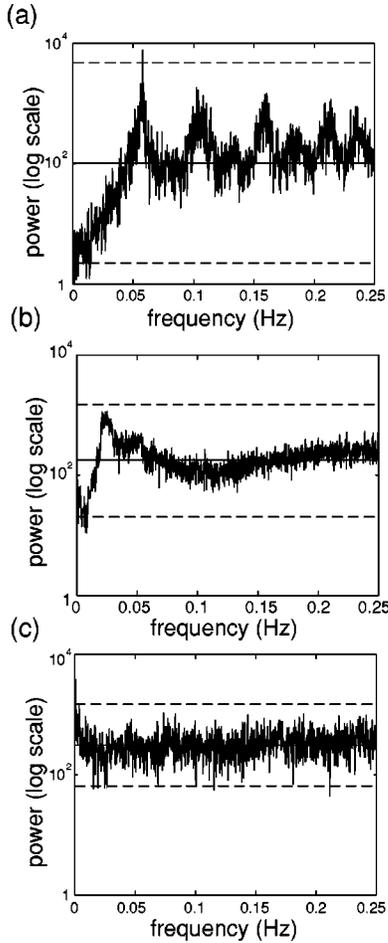
FIG. 6. Power spectra of cluster sequences. The abscissa scale is cycles per second (Hz) and the ordinate is power, on a logarithmic scale. The horizontal solid and dashed lines on each plot, respectively, indicate the mean power and the mean $\pm$ three standard deviations, in logarithmic form. The original data are the same as shown in Fig. 5, and the plots in (a), (b), and (c) similarly correspond to the "periodic," "bursting," and "Poisson" forms, respectively.

esis] and the empirical distribution function $S_N(x)$. In general, the distribution of D_N is not asymptotically normal, but a convenient approximation to the sampling distribution of D_N has been derived [39,38]. Let $d_{N,\alpha}$ be a value such that

$$\text{Prob}\{D_N > d_{N,\alpha}\} = \alpha. \tag{16}$$

TABLE III. Tendencies to periodicity displayed by the cluster sequences.

Days *in vitro*	Number of cells		(%)	
	Over 2σ	Over 3σ	Sample number	
1–10	6	(21.4)	4	(14.3)	28	(100)
11–20	25	(41.7)	11	(18.3)	60	(100)
21–40	13	(32.5)	0	(0)	40	(100)
41–65	2	(4.17)	0	(0)	48	(100)
Total	46	(26.1)	15	(8.52)	176	(100)
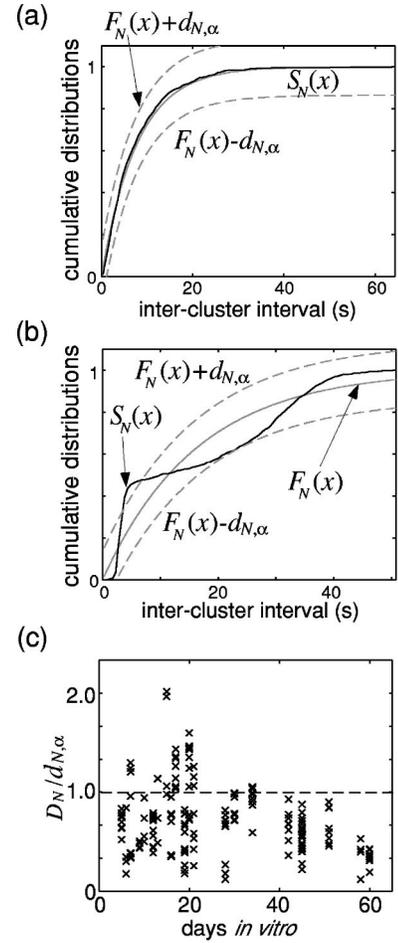
FIG. 7. Results of testing distributions for Poisson-like characteristics. (a) Comparison, for data on ICI sequences from a cell in a culture at 45 days *in vitro* (DIV), of the empirical distribution function $S_N(x)$ with the expected Poisson distribution function $F_N(x)$. The basic statistics: $N=3183$, $\mu=7.56$ (s), $\sigma=7.43$ (s), and $\mathcal{C}=1.02$. N is the number of clusters and μ, σ, and $\mathcal{C}$ are the ICI mean, standard deviation, and coefficient of variation, respectively. (b) A similar comparison for the data on sequences from a cell in an 18-DIV culture. The basic statistics: $N=2676$, $\mu=16.4$ (s), $\sigma=13.9$ (s), and $\mathcal{C}=0.852$. (c) Ratios between the two-sided Kolmogorov-Smirnov (KS) statistics D_N and $d_{N,\alpha}$ versus DIV for 161 cells from 41 cultures. The data sets excluded those ICI sequences classified as being of the "delta-like" form. The 95% confidence level is used in the test: $\alpha=0.05$.

The two-sided test of goodness of the fit using the statistics D_N then rejects the null hypothesis at a level α if the observed value of D_N is greater than $d_{N,\alpha}$. We used the 95% confidence level: $\alpha=0.05$. The hypothesis H_0 was rejected if the empirical distribution function passed outside the band bounded above by the smaller of $S_N(x)+d_{N,\alpha}$ and 1 and bounded below by the larger of $S_N(x)-d_{N,\alpha}$ and 0.

Figure 7 shows the empirical distribution functions $S_N(x)$ and the corresponding Poisson distribution functions $F_0(x)$. In Fig. 7(a), there is a good match between the observed and expected distributions of ICIs from a cell cultured for over 45 DIV. In Fig. 7(b), the sequences from a cell cultured for 18 DIV deviate markedly from the corresponding Poisson

process. The results of the KS statistics D_N for all of the data sets, except for those cluster sequences classified as being of the ''delta-like'' form, are summarized in Fig. 7(c). In the plot, the ordinate represents the ratio of D_N to $d_{N,\alpha}$. Small values of the ratio (less than 1) indicate good agreement between the distributions of the sequences and those produced by Poisson processes. In terms of the distributions, the proportion of data sets with cluster sequences of the ''Poisson'' form increases with the number of days *in vitro*.

The statistical test of the independence of ICI sequences $\{I_i\}$ that we have used here is based on Sperman's coefficient of rank correlation [39]. Let us draw a sample of n pairs $\{(I_i, I_{i+k})\}$ $(i=1,\ldots,n)$ for some positive integer k and denote the respective ranks of the variables in the sample as

$$U_i = \mathrm{rank}(I_i) \quad \text{and} \quad V_i = \mathrm{rank}(I_{i+k}) \quad (i=1,\ldots,n),$$

(17)

where $\mathrm{rank}(X_i)=j$ if the ith element X_i is the jth smallest in the sample. The derived sample observations of n pairs are thus $\{(u_i, v_i)\}$ and $u_i, v_i = 1, \ldots, n$ for $i=1,\ldots,n$. The (kth-order) Sperman's coefficient R_k of rank correlation is defined as

$$R_k = \frac{12\sum_{i=1}^{n}(u_i - \bar{U})(v_i - \bar{V})}{n(n^2-1)},$$

(18)

where

$$\bar{U} = \bar{V} = \frac{n+1}{2}.$$

(19)

$R_k\sqrt{n-1}$ is known to be asymptotically normal, i.e., to approach a normal distribution when n is large ($n>10$). Independence of the intervals is thus tested by a null hypothesis $H_0(k)$: $R_k=0$ $(k=1,2,\ldots)$. The hypothesis is rejected at the confidence level α $(\alpha=0.05)$ if $|R_k|\sqrt{n-1}>z_{\alpha/2}$, where $z_{\alpha/2}$ is read from a standardized normal table [42].

Values of the first-order and second-order Sperman's coefficients of rank correlation R_1 and R_2 are plotted against DIV in Fig. 8. At 10–13 and 21–40 DIV, the values are in good agreement with the hypothesis. The values of R_k for higher values of k show similar tendencies to those shown by R_k for $k=1$ and 2. Furthermore, the absolute values of R_k for values of k greater than 8 are all less than $z_{\alpha/2}$ and are thus not significantly different from 0. All results of testing for goodness of the fit to the distribution function and the independence of the intervals are summarized in Table IV. Deviations from the null hypothesis are visible in most of the R_1 and R_2 values for data sets at 14–20 DIV (the result is not shown here because the values are scattered over a wider range of scale than those for 10–13 and 21–40 DIV). These negative results for the null hypothesis are easily explained by the oscillatory tendencies of the data sets of ICI sequences from cultures at 14–20 DIV, as was discussed in the preceding section.

In some sets of the ICI sequences we found a trend, which was practically determined to exist by Kendall rank-
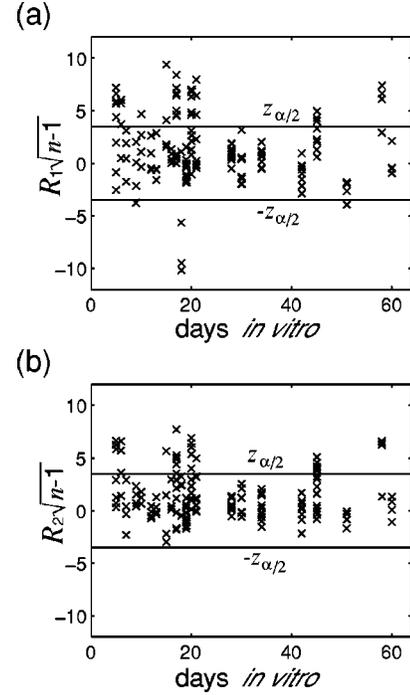
FIG. 8. Results for independence of the test of the ICI sequences. Values of the first-order and second-order Sperman's coefficients of rank correlation R_1 and R_2 are plotted against DIV. (a) The first-order Sperman's coefficients of rank correlation R_1. For comparison, since $R_k\sqrt{n-1}$ $(k=1,2,\ldots)$ approaches an asymptotically normal distribution for increasingly large values of n, the values are multiplied by a constant $\sqrt{n-1}$, where n is the length of each data set. (b) The second-order correlation coefficient R_2. Independence of the intervals is thus tested against a null hypothesis $H_0(k)$: $R_k=0$ $(k=1,2,\ldots)$. The hypothesis is rejected at the confidence level α $(\alpha=0.05)$ if $|R_k|\sqrt{n-1}>z_{\alpha/2}$, where $z_{\alpha/2}$ is read from a standardized normal table and $z_{\alpha/2}\approx 3.481$.

correlation tests using the Kendall coefficient τ [39]. The tests (administered at the 0.05 confidence level) measured the correlation between the orders and the magnitudes of the ICI sequences. In such cases, we were unable to treat the data as representing a homogeneous Poisson process, so we repeatedly divided such data sets into two sets, up to a maximum of four times (as a result, 2–16 subsets of the original data were obtained). We then tested the null hypothesis on the subsets thus obtained.

C. Nonlinear analysis of ICI series

Although linear modeling techniques are only able to partially represent the underlying system because they do not take the nonlinear contribution into account, they do allow us to deal with simplified forms of problems. A large number of nonlinear algorithms for characterizing real-valued data have thus been developed [43]. In the work reported on in this section, we used correlation dimensions and a method of nonlinear forecasting.

In general, the dynamics of a system are simplified considerably when it is sufficiently stationary to be governed by its attractor. The trajectory of the state point in phase space is

TABLE IV. Results of tests for Poisson processes.

Days *in vitro*	Number of cells	(i) No trend	(ii) KS test	(iii) Independence	(i), (ii), and (iii)	
1–10	24	21	15	12	10	(41.7%)
11-20	49	45	19	14	7	(14.3%)
21–40	40	38	25	32	20	(50.0%)
41–65	48	45	36	29	27	(56.3%)
Total	161	149	95	87	64	(39.8%)

able to portray the essential features of the system's dynamics as long as it provides a good approximation of the attractor. Grassberger and Procaccia have described a method for using correlation integrals to determine the attractor's correlation dimension in a simple way [44]. These correlation integrals seem to be a good tool, since they are effective and their calculation usually requires fewer data points than other methods. We begin in the usual way by constructing a delay embedding of the data. The τ-dimensional embedding yields vectors of the form $x_\tau(i) = (I_i, \ldots, I_{i-\tau+1})$. In the Grassberger and Procaccia (GP) method, we start by counting the normalized number $C(r, \tau)$ of pairs of vectors in τ-dimensional phase space that have vector differences less than r. For small values of r, the correlation integral function $C(r, \tau)$ is known to behave according to a power law,

$$C(r, \tau) \propto r^\nu. \tag{20}$$

If the correlation exponent ν approaches some limit d as τ is increased, the correlation dimension of the attractor is d. Although the correlation dimension d is defined in terms of a limit as the radius r approaches zero, it is in practice numerically estimated from the slope of the curve of the log of the correlation integral versus the log of radius. If the graph of $\log_{10} C(r, \tau)$ versus $\log_{10} r$ has an apparently linear region, this is called the scaling region. In the GP method, it is assumed that most of the information about dimension is contained in this scaling region. We estimated the slope of the curve by a least-squares fit of a straight line to points on the curve. Eckmann and Ruelle have shown that we need more than $10^{d/2}$ data points to accurately estimate the correlation dimension d from the GP algorithm [45]. Since there were 1000–3500 points in each data set of this work, the results of estimation are only accurate if the correlation dimension is less than the Eckmann-Ruelle bound $d_a = 7.09$.

The correlation integrals of a data set of ICIs at 30 DIV are shown in Fig. 9(a). When the slopes of the curves are plotted against the embedding dimension τ, they seem to converge to a value ($d = 6.82$) as τ increases [Fig. 9(b)]. All of the results are summarized in Table V. In the immature stages (1–10 DIV), the correlation dimensions are quite uniformly scattered over the interval from 1 to 7. At the more mature stages (21–65 DIV), however, few correlation dimensions below 4 appear. The number of higher-order ($d \geqslant 5$) correlation dimensions correspondingly increases.

For the data we used, however, to obtain the values of the dimension less than the bound d_a, we need to go to embedding dimensions of 20 or more before they start seeing a saturation as shown in Fig. 9(b). This is even beyond the conservative embedding dimension limit of $2D+1$ [46], which should be about $14-15$ if the dimension D is really around 7. This by itself suggests that the data are high dimensional, or that the low dimensionality cannot be resolved with these small data sets.

To confirm this suspicion, therefore, for those data sets of ICI sequences that have correlation dimensions below d_a, we applied surrogate data analysis. That is, to test the null hypothesis (H_0) that our results are explicable as the result of nondeterministic (linear stochastic) processes, we compared the data with sets of mathematical controls [47,48] that are called "surrogate data." Following the method proposed by Theiler *et al.* [47], we selected the residual probability α (0.04) of a false rejection, corresponding to a level of significance of $(1-\alpha) \times 100\%$ (96%). For a two-sided test, we then generate $M = 2/\alpha - 1$ ($M = 49$ when $\alpha = 0.04$) surrogate sequences, resulting in a probability of α that the original data would give either the smallest or largest value of the correlation dimension. We thus used two methods to make M sets of surrogates for each testing set $\mathcal{T}$: one (i) was a phase-randomized method based on the discrete Fourier transform,
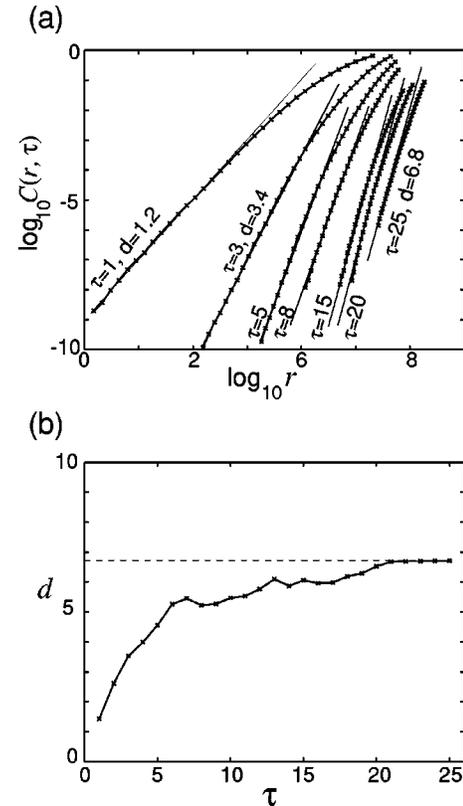
FIG. 9. (a) Correlation integrals for a data set from a culture at 31 DIV. The basic statistics: $N = 2487$, $\mu = 17.6$ (s), $\sigma = 14.3$ (s), and $\mathcal{C} = 0.812$. N is the number of clusters and μ, σ, and $\mathcal{C}$ are the ICI mean, standard deviation, and coefficient of variation, respectively. (b) Slope of the curve of the correlation integral d versus the embedding dimension τ.

TABLE V. Summary of the results for estimated correlation dimension.

Days *in vitro*	$d<3$	$3\leqslant d<4$	$4\leqslant d<5$	$5\leqslant d<6$	$6\leqslant d<7$	$d>7$	Total
			Number of cells (%)				
1–10	4	5	5	7	4	3	28
	(10.7)	(17.9)	(17.9)	(25.0)	(14.3)	(10.7)	(100)
11–20	3	4	10	12	18	13	60
	(5.0)	(6.7)	(16.7)	(20.0)	(30.0)	(21.7)	(100)
21–40	0	8	7	10	8	7	40
	(0)	(20.0)	(17.5)	(25.0)	(20.0)	(17.5)	(100)
41–65	0	4	11	9	7	17	48
	(0)	(8.3)	(22.9)	(18.8)	(14.6)	(35.4)	(100)
Total	7	21	33	38	38	30	176
	(4.0)	(11.9)	(18.8)	(21.6)	(26.7)	(17.1)	(100)

while the other (ii) was a random-shuffle method [47]. The same procedure as outlined above was used to calculate the correlation integrals of the original and surrogate data. We thus obtained the correlation dimensions d_0 for the original data and $d_1, \ldots, d_M$ for the surrogates. Figure 10A shows the results of analysis using surrogate data produced by the phase-randomized method to check the null hypothesis (H_0) for those examples of ICI sequences, where the correlation dimensions of the original data sets are less than d_a. The analysis gave the negative result that the null hypothesis (H_0) should not be rejected for four [Figs. 10A(a), 10A(c), 10A(f), and 10A(g)] of the seven data sets. However, although we rejected the hypothesis H_0 for three of the data sets [Figs. 10A(b), 10A(d), and 10A(e) we note that the correlation dimensions d_0 (longer vertical lines marked with circles) of the original data sets for which there were positive results are very close to those of the corresponding surrogates (shorter vertical line without a circle). Analysis using the random-shuffle surrogate method also showed the negative result that the null hypothesis (H_0) should not be rejected, as plotted in Figs. 10B(a) and 10B(f).

We also used a nonlinear forecasting method originally developed by Sugihara and co-worker [49,50] and extended by Sauer [51,52]. Sauer showed that a determinism that underlies sequences of neurophysiological data may be detected by analyzing the interspike intervals in the data. Since this method is capable of detecting a weak determinism, it has been used to evaluate many neurophysiological data sets [15–18].

In nonlinear forecasting, we divided each data set into halves, one of which was the training set $\mathcal{D}$, while the other provided the testing set $\mathcal{T}$. After selecting an interval I_i in $\mathcal{T}$ as an index point, the k nearest neighbors of the related τ-dimensional vector $x_\tau(m_j)(j=1,\ldots,k)$ in $\mathcal{D}$, where k is 3% of all embedded vectors, were found. Future sequences I_{i+h}, h steps ahead of the index point I_i, were estimated by averaging the sequences h steps ahead of the k nearest neighbors I_{m_j+h}. Here I_{m+h} denotes the estimated value of I_{i+h}. The accuracy of the prediction was evaluated by computing an h-step normalized prediction error (NPE) as the square of all prediction errors averaged over the whole testing set $\mathcal{T}$ and then dividing this mean square by the variance of the ICI

sequence of the mean (I_{mean}) of the predicted series,

$$\text{NPE}(h)=\frac{\sum_{i=1}^{N_\mathcal{T}}(I_{i+h}-I_{m+h})^2/N_\mathcal{T}}{\sum_{i=1}^{N_\mathcal{T}}(I_{i+h}-I_{mean})^2/N_\mathcal{T}}, \qquad (21)$$

where $N_\mathcal{T}$ is the number of elements in the testing set $\mathcal{T}$. A normalized prediction error of less than 1 means that it is possible to predict the series with an accuracy beyond that of the baseline prediction by the mean interval of the sequences. To test the null hypothesis that our results were explicable as having been produced by nondeterministic (linear stochastic) processes, we again applied surrogate analysis. We used the same two methods as we had in testing the correlation integrals to make 49 sets of surrogates for each testing set $\mathcal{T}$. The same procedure was again used to calculate the NPEs of the original and surrogate data.

Figures 11(a) and 11(b) show the NPE as obtained by forecasting the ICI sequences of a culture at 17 DIV. The one-step-ahead NPEs of the original data (thick lines) are less than 1 and are apparently separated from the surrogate data (thin lines) generated by both the phase-randomized [Fig. 11(a)] and random-shuffle [Fig. 11(b)] methods. However, a value of the one-step-ahead NPE, near 1.0, means that the degree of predictability is small and that this method of prediction is almost the same as substituting the original values with the mean value of the intervals in the sequences. These randomly shuffled and phase-randomized surrogates are not well suited for surrogate data analysis in these cases because NPEs greater than 1 imply that the algorithm provides a poorer prediction than the prediction from the mean value of the data set. This often happens with interval data, so that we have to be careful with the interpretation of the results as was mentioned by Schreiber and Schmitz [53]. There is, moreover, no clear difference between predictions more than two steps ahead from the original and from the surrogate data. The one- and two-step-ahead NPEs of ICI sequences from all of the data sets are shown in Figs. 11(c) and 11(d). The variance of the NPEs decreases as DIV increases, and most of the two-step-ahead NPEs are greater
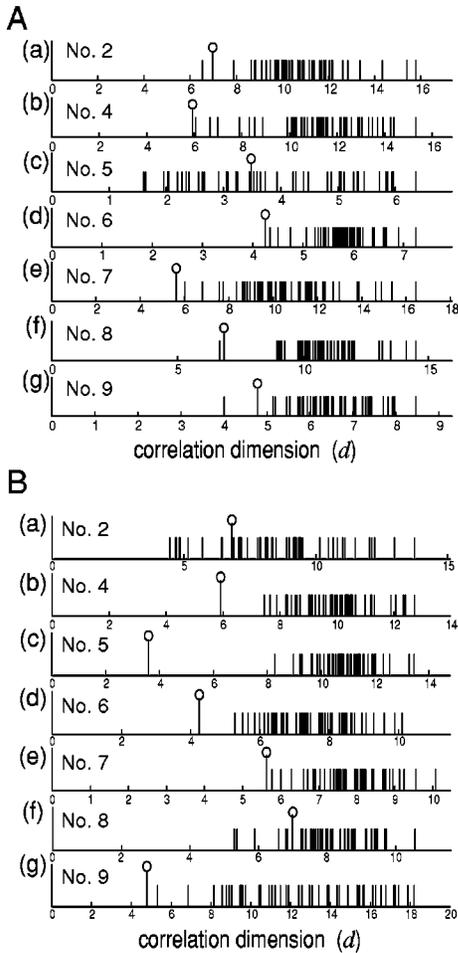
FIG. 10. Surrogate data tests of the seven data sets listed in Table I (for neuron numbers 2, 4, 5, 6, 7, 8, and 9). The correlation dimensions of the selected seven data sets are less than d_a ($=7.09$) while those of the other two (neuron numbers 1 and 3) are greater than d_a. Longer lines with circles indicate correlation dimensions (d_0) of the original data sets and shorter lines indicate those of the corresponding surrogates ($d_1, \ldots, d_{49}$), as calculated by using the phase-randomized (A) and random shuffle (B) methods. For (a) neuron number 2 listed in Table I, $d_0 = 6.98$, (b) number 4, $d_0 = 5.92$, (c) number 5, $d_0 = 3.48$, (d) number 6, $d_0 = 4.25$, (e) number 7, $d_0 = 5.63$, (f) number 8, $d_0 = 6.82$, and (g) number 9, $d_0 = 4.78$.

than 1. Thus, for two-step-ahead prediction, the deterministic hypothesis for the corresponding sequences was rejected. This implies that there is no two-step predictability of the sequences. Furthermore, no predictability may be expected for higher-order forecasting.

V. DISCUSSION

The role of spontaneous activity in the configuring of operative circuits has recently been recognized [54]. A typical example of this is the activity-dependent process involved in the establishment of cortical columns [8]. The spatiotemporal properties of spontaneous firing in networks of developing neurons undergo large changes, but little is known about the rules for evolution over time that govern these changes (i.e.,
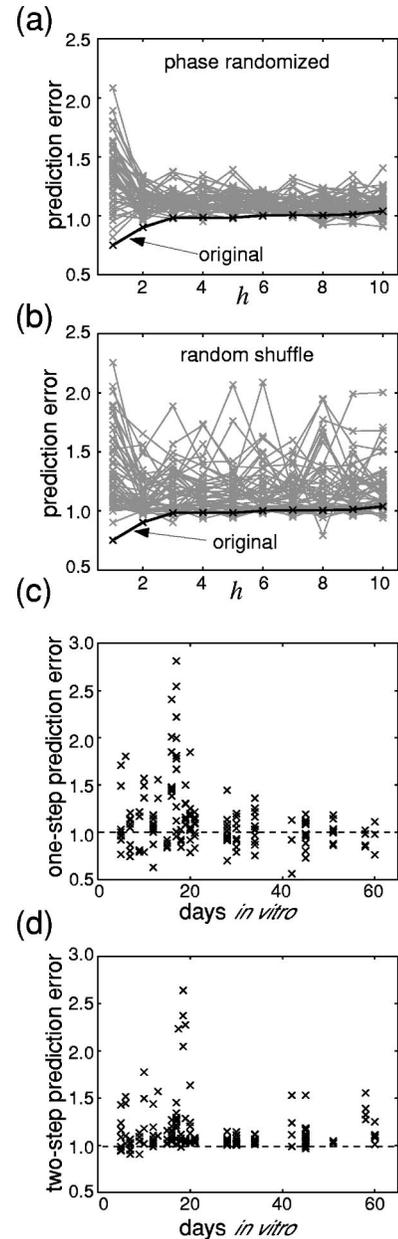
FIG. 11. (a) Nonlinear prediction error of the original data set (thick line) in a 31-DIV culture and 49 sets of phase-randomized surrogate data (thin lines) versus h, where h is the number of steps ahead of a index point. The embedding dimension $\tau = 5$. The basic statistics: $N = 2676$, $\mu = 16.3$ (s), $\sigma = 13.9$ (s), and $\mathcal{C} = 0.852$. N is the number of clusters and μ, σ, and $\mathcal{C}$ are the ICI mean, standard deviation, and coefficient of variation, respectively. (b) Nonlinear prediction error of the original data set (thick line) and 49 sets of random-shuffle-generated surrogate data (thin lines) versus h. The original data set is the same set as was used in (a). The embedding dimension $\tau = 5$. (c) One-step-ahead normalized nonlinear prediction error (NPE) versus DIV. (d) Two-step-ahead NPE versus DIV.

about their dynamics). As for spatial "patterns," some neural groups, in the early stages of their development, of the cultured cortical networks examined in this work were seen to form small groups in which activity was synchronized. Later,

globally synchronized bursts (tight coincidences of spike timing between the cells over the whole recording area during periods of several tens of milliseconds and more) were observed. The synchronization of spontaneous activity was stable over periods of several hours, and this stability can lead to activity-dependent changes in the network and in the dynamics of individual neurons.

Furthermore, the significance of the activity that is correlated between presynaptic and postsynaptic cells or the lack of thereof, i.e., synchrony or asynchrony, has long been known to be a critical point. This has been extensively discussed in terms of physiological concepts, mathematical argument, and simulation. In this paper, we have used "synchronization" to mean tight coincidences between cells in terms of spike timing. Understanding why there is synchrony or asynchrony of the activity of a cell in a network, in terms of both its own presynaptic and postsynaptic activity and of the interactions between the other cells of the network, is a critical point for understanding the activity of the developing cortical cell. We intended to further investigate this issue in future work. A number of experimental studies of such synchrony, both *in vivo* and *in vitro*, have recently been carried out. Introductions to the early work, which involves argument from mathematical bases and simulation, are available (e.g., Segundo and co-workers [26,55]). Reviews and summaries of recent issues to do with the cortex, in particular, are also available in, for instance, Fujii *et al.* [6].

Our results in this study have shown that, at least in cultured cortical neurons, concerted activity is capable of producing synchronized bursts (clusters) and that some of sequences of clusters are correlated in time [e.g., Figs. 5B(a) and 5B(b)]. The results also showed that the ICI sequences at all stages of development appeared to be primarily manifestations of random processes. Any nonlinear determinism that did appear was of high dimension of a dynamical system. The dimension analysis, in particular, suggests that the data we used are high dimensional, and low-dimensional determinism is not likely to be present, or that the low dimensionality cannot be resolved with these small data sets due to noise or nonstationarity. In the latter case, we need extremely large amounts of clean data to get any kind of accurate dimension estimate in such high embedding dimensions, which we did not have in these experiments. However, we think that the dimension and prediction analysis, along with surrogate analysis, is worthwhile even though the results were not clearly in favor of nonlinear determinism. The results of such analyses thus complement those of the cluster analysis.

One of the main empirical findings of this study was that neurons at 15–25 DIV (referred to in this work as the early stage of development) yield ICI sequences, which may be interpreted as the product of stochastic processes with a relatively high-order temporal correlation or of a nonlinear system with high dimension ($d \geq 5$). In this stage, in particular, oscillatory (periodic or bursting) "forms" are visible in the ICI sequences of many neurons. In contrast, the ICI sequences obtained from cultures at more than 30 DIV (referred to in this work as the later stages) may be attributed to stochastic processes with a relatively low-order temporal correlation or to a "Poisson-like" form. In addition, the change in the temporal correlation is associated with the transition of the network structure in developing cortical neurons.

Our results thus show that synchronized spontaneous activity, at least in cultured cortical networks, falls into a pattern of multicellular clusters of activity. The timing of these clusters is basically that produced by a stochastic process or by a deterministic system with a high-order dimension. A similar kind of activity may underlie the high variability of cortical spike trains seen *in vivo* [6].

Segundo and co-workers argued for the extrinsic properties of the spike-train "form" on the basis of mathematics and simulation, illustrating their work with examples that included bursts and putting their results in the context of neural "coding" [26,55]. The "code" implies the representation and transformation of information. As the pertinent conditions vary for a set of neurons (in the case of our study, with development), there are corresponding alternations in neural activity. Segundo and co-workers provide an extensive explanation of how the postsynaptic averages, variabilities, and "forms" depend on several presynaptic parameters, including precise tendencies towards firing or not firing simultaneously. The models they used are simple and are restricted to excitatory junctions and we have dealt with sequences of clusters rather than sequences of spikes. However, these earlier results provide a basis for speculation on the functional changes in networks of cultured neurons during development.

Segundo *et al.* summarized the main conclusions of their study on the output of the post synaptic cell [26]. We paraphrase these conclusions here (1) (i) When there is no interdependence between presynaptic terminals that impinge on the postsynaptic cell, presynaptic terminals are few, and excitatory postsynaptic potentials are large, the output varies with the presynaptic "form;" (ii) if terminals are numerous and weak, the presynaptic form ceases to have an effect and the postsynaptic cell generates the same output regardless of the detailed structure of the corresponding input; (2) when there is an interdependence between presynaptic terminals, which involves only a proportion of all terminals or only those sets of terminals within separate and independent groups, the activity of the postsynaptic cell is a function of the statistical form of its input channels, even if the presynaptic terminals are numerous and weak.

In the early stage (1–25 DIV) of a cultured network, a single cell is affected by a relatively small number of adjacent cells because its neurites are not extensive on the substrate. As the results of Segundo *et al.* show, a single strong synaptic influence on the spike train's form may come from a single powerful terminal or from several powerful terminals. That is, the postsynaptic potential has a relatively large effect on the activity of a cell. Hence, cases (1) (i) and (2) are the candidate mechanisms in the early states of its development for the spike train's form. In addition, neurophysiological and pharmacological studies have shown that γ-aminobutyric acid (GABA) receptor antagonists (e.g., picrotoxin and bicuculline methiodide) have marked effects on cultured cortical neurons in the latter part (15–20 DIV) of the early stage of development. These effects include the

induction of anomalous spontaneous activity and the kinetic modulation of synaptic currents [13,56]. In the earliest period (1–14 DIV), however, such antagonists have little effect on spontaneous activity. This suggests that inhibitory (GABAergic) transmission in the network may have a strong effect on the dynamics of concerted activity in populations of neurons. In the later part of the early stage, therefore, the clearly oscillatory properties of the ICI sequences may be because the synapses are arranged according to category (2) described above.

Another type of network, the simplest, is formed with numerous inputs, either independent or synchronous, from cells immediately presynaptic to those being recorded from. Stevens and Zador have recently reported that independent synaptic inputs are not able to account for the highly irregular spike trains observed from cortical neurons *in vivo*. They suggest that the high degree of variability in firing may be explained by a simple alternative model of synaptic drive in which inputs arrive synchronously [57]. Furthermore, their conclusion implies that synchrony may be an important aspect of the neural code in that it provides a way of encoding signals with a high degree if temporal fidelity over a population of neurons.

Another possible reason for the characteristic changes in the ICI sequences is the changes in the intrinsic properties of the developing cells. In many cases, clusters (bursts) of activity are based on the interaction between slow and fast dynamics. The slow dynamics is, in general, responsible for the triggering of clusters. The mechanism by which a slow depolarizing wave is generated differs according to the type of cell. A slow wave in a cortical culture is likely, for example, to reflect an intracellular influx of calcium and to involve nonactivating Na^+ channels or n-methyl-D-aspartate (NMDA) channels [13]. During the development of the cultured cells, the mechanism that underlies the generation of the slow wave becomes mature and the properties of the ICI sequences change accordingly. The existence of specialized mechanisms for cluster generation suggests that clusters are an important feature of neural signals in time. The significance of clusters in processing does not, however, depend on whether the clusters are simple network-driven mechanisms or are simply due to the physiological properties of individual cells.

[1] A. Destexhe, D. Contreras, and M. Steriade, J. Neurosci. **19**, 4595 (1999).

[2] E. Vaadia *et al.*, Nature (London) **373**, 515 (1995).

[3] A. Riehle *et al.*, Science **278**, 1950 (1997).

[4] W. R. Lippe, Brain Res. **703**, 205 (1995).

[5] M. Weliky and L. C. Katz, Science **285**, 599 (1999).

[6] H. Fujii *et al.*, Neural Networks **9**, 1303 (1996).

[7] M. Meister, R. O. L. Wong, D. A. Baylor, and C. J. Shatz, Science **252**, 939 (1991).

[8] C. J. Shatz, Neuron **5**, 745 (1990).

[9] *The Handbook of Brain Theory and Neural Networks*, edited by M. A. Arbib (MIT Press, Cambridge, 1995), pt. III.

[10] R. D. Traub, J. G. R. Fefferys, and M. A. Whittington, *Fast Oscillations in Cortical Circuits* (MIT Press, Cambridge, 1999).

[11] *Pulsed Neural Networks*, edited by W. Maass and C. M. Bishop (MIT Press, Cambridge, 1999).

[12] R. A. Deisz and H. J. Luhmann, in *The Cortical Neuron*, edited by M. J. Gutnick and I. Mody (Oxford University Press, Oxford, 1995).

[13] H. P. C. Robinson *et al.*, J. Neurophysiol. **70**, 1606 (1993).

[14] H. Kamioka *et al.*, Neurosci. Lett. **206**, 109 (1996).

[15] T. Chang *et al.*, Biophys. J. **67**, 671 (1994).

[16] S. J. Schiff *et al.*, Biophys. J. **67**, 684 (1994).

[17] R. E. Hoffman, W.-X. Shi, and B. S. Bunney, Biophys. J. **69**, 128 (1995).

[18] L. M. Prida, N. Stollenwerk, and J. V. Sanchez-Andres, Physica D **110**, 323 (1997).

[19] P. E. Rapp *et al.*, J. Neurosci. **14**, 4731 (1994).

[20] W. Singer and C. M. Gray, Annu. Rev. Neurosci. **18**, 555 (1995).

[21] W. J. Freeman, Int. J. Neural Syst. **7**, 473 (1996).

[22] J. P. Segundo, J.-F. Vibert, and M. Stiber, Neuroscience **87**, 15 (1998).

[23] J. P. Segundo *et al.*, Neuroscience **87**, 741 (1998).

[24] J. P. Segundo *et al.*, Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**, 549 (1991).

[25] J. P. Segundo *et al.*, Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**, 873 (1991).

[26] J. P. Segundo *et al.*, Kybernetik **4**, 157 (1968).

[27] H. C. Tuckwell, *Stochastic Processes in the Neurosciences* (SIAM, Philadelphia, PA, 1989).

[28] K. Muramoto *et al.*, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci. **64**, 319 (1988).

[29] V. Braitenberg and A. Schüz, *Cortex: Statistics and Geometry of Neural Connectivity* (Springer, Berlin, 1991), Chap. 33.

[30] G. W. Gross, IEEE Trans. Biomed. Eng. **26**, 73 (1979).

[31] J. Pine, J. Neurosci. Methods **2**, 19 (1980).

[32] Y. Jimbo, H. P. C. Robinson, and A. Kawana, IEEE Trans. Biomed. Eng. **40**, 804 (1993).

[33] Y. Jimbo, T. Tateno, and H. P. C. Robinson, Biophys. J. **76**, 670 (1999).

[34] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification* (Wiley, New York, 2001), Chap. 10.

[35] M. Meister, J. Pine, and D. A. Baylor, J. Neurosci. Methods **51**, 95 (1994).

[36] D. R. Cox and V. Isham, *Point Processes* (Chapman and Hall, Boca Raton, FL, 1980), Chap. 3.

[37] F. Grüneis, M. Nakao, and M. Yamamoto, Biol. Cybern. **62**, 407 (1990).

[38] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events* (Maehuen, London, 1966).

[39] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference* (Dekker, New York, 1992), Chaps. 4 and 12.

[40] H. Bryant, A. R. Marcos, and J. P. Segundo, J. Neurophysiol. **36**, 205 (1973).

[41] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, England, 1993), Chap. 6.

[42] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians* (Cambridge University Press, Cambridge, England, 1966), Vol. 1.

[43] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).

[44] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).

[45] J. K. Eckmann and D. Ruelle, Physica D **56**, 185 (1992).

[46] F. Takens, in *Dynamical Systems and Turbulence, Warwick 1980*, edited by D. Rand and L. S. Young (Springer, Berlin, 1981), pp. 366–381.

[47] S. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).

[48] S. Theiler and D. Prichard, in *Nonlinear Dynamics and Time Series: Building a Bridge Between the Natural and Statistical Sciences*, edited by C. Culter and D. Kaplan, Fields Institute Publications Vol. 11 (American Mathematical Society, Providence, RI, 1997), pp. 99–113.

[49] G. Sugihara and R. May, Nature (London) **344**, 734 (1990).

[50] G. Sugihara, Philos. Trans. R. Soc. London, Ser. A **348**, 477 (1994).

[51] T. Sauer, Phys. Rev. Lett. **72**, 3811 (1994).

[52] T. Sauer, in *Nonlinear Dynamics and Time Series: Building a Bridge Between the Natural and Statistical Sciences* (Ref. [48]), pp. 63–75.

[53] T. Schrieber and A. Schmitz, Physica D **142**, 346 (2000).

[54] L. C. Katz and C. J. Chatz, Science **274**, 1133 (1996).

[55] J. P. Segundo, in *The Neurosciences: Second Study Program*, edited by G. C. Wuarton, T. Melnechuk, and F. O. Schmitt (Rockefeller University Press, New York, 1970), pp. 569–586.

[56] S. Watanabe, Y. Jimbo, H. Kamioka, Y. Kirino, and A. Kawana, Neurosci. Lett. **210**, 41 (1996).

[57] C. F. Stevens and A. M. Zador, Nat. Neurosci. **1**, 210 (1998).