

# Neuronal Assemblies

GEORGE L. GERSTEIN, PURVIS BEDENBAUGH, AND AD M. H. J. AERTSEN

(Invited Paper)

**Abstract**—This paper examines the concept of neuronal assembly as it has appeared in selected portions of the literature. The context is experimental access to real neuronal assemblies in working brains, as made possible by recent technological progress. One current measure of assembly organization is based on correlation of firing among neurons; recent observations show that such correlations can vary rapidly. In this paper, we demonstrate that dynamic firing correlation can be caused either by dynamic changes in neuronal connection strengths or, alternatively, by the effects of an unobserved (large) pool of other neurons. The static connectivity within the pool appears to be important in determining these effects.

## INTRODUCTION

It is probably fair to say that the bulk of our knowledge about the nervous system is on the level of single neurons and their molecular and membrane mechanisms. Over the last 35 years much has been learned about firing patterns of single neurons in various parts of the nervous system under a variety of stimulus and behavioral conditions. More recently, a large body of work on receptors, channels, and modulators has served to explain and to emphasize the importance of endogenous properties of individual neurons. In addition, anatomical studies with modern tracer methods have provided a great deal of information about gross connectivity between different nuclei and brain regions. In vertebrate nervous systems, however, detailed local anatomical connectivity has only been worked out in a few structures like cerebellum.

Long before emergence of all the current complexities of single neuron anatomy and physiology, it was recognized that the computational processes constantly carried out by the nervous system could not be explained by the properties of single neurons alone. Sherrington, in the 1930's, verbalized the idea that neurons must cooperate in fulfilling a complex task. This concept has been completely assimilated into neurobiological thought and explanation, but only in the most general way. Experimental access to the properties of real neuronal assemblies has largely been impossible until quite recently. Thus, the development of the assembly concept has been somewhat unfettered. It is the purpose of this paper to examine and compare the various suggested concepts of neuronal as-

sembly, and to address the problems of experimental verification.

## CONCEPTUAL MODELS OF ASSEMBLIES

### A. Motor Pools

Sherrington [10] suggested that the motor neurons serving a particular muscle form a pool of potentially active cells, and that there would be appropriate recruitment of neurons within this pool as the excitatory input (presumably from various sources) increased. The detailed distribution of such input, or of its effects, was assumed to differ across the motor pool, so as to "fractionate" the motor neuron population into a "discharge zone" of firing cells and a "subliminal fringe" of cells that were not sufficiently excited to fire. Inputs which excited (spatially) overlapping populations of neurons would be able to interfere or add nonlinearly. Thus, "occlusion" of discharge would result from two inputs that caused overlapping discharge zones, producing a combined output that is smaller than the sum of the separate outputs. Similarly, "facilitation" of discharge would result from two inputs which caused overlapping subliminal fringes, producing a combined output that is larger than the sum of the separate outputs.

The accuracy of this conceptual model was tested by Lloyd [36], [37] with experiments in which two dissected dorsal root twigs were stimulated separately or simultaneously at various strengths. The resulting discharge was observed as a slightly delayed potential on the corresponding ventral root. Under the assumption that amplitude of this ventral root potential was proportional to the number of neurons in the motor pool that were brought to discharge, Lloyd showed all the phenomena that Sherrington had proposed.

For our present purposes, we note that the neurons forming or having membership in a motor pool are defined exclusively through commonality of output projection (i.e., to a particular muscle). There is also an implied nonuniform distribution of the input effects. There is no requirement for connections or interactions among the member neurons in the pool, although these are of course present [38]. For an illuminative example of assembly coding in the context of the motor system, we refer the reader to the recent literature on coding of saccadic eye movements in the superior colliculus (e.g., model oriented: Gisbergen *et al.* [22]; experiment oriented: Lee *et al.* [35]).

Manuscript received June 7, 1988; revised July 25, 1988. This work was supported by Grant SDF 0013 from the System Development Foundation and Grant N00014-87-K-0766 from the Office of Naval Research.

G. L. Gerstein and P. Bedenbaugh are with the Department of Physiology, University of Pennsylvania, Philadelphia, PA 19104.

A. M. H. J. Aertsen is with the Max-Planck-Institute for Biological Cybernetics, Tuebingen, West Germany.

IEEE Log Number 8824412.

### B. Hebbian Assemblies

In 1949, in his seminal book *The Organization of Behavior*, Hebb [28] introduced the concept of the cell assembly and postulated mechanisms of both perception and learning: “It is proposed first that a repeated stimulation of specific receptors will lead slowly to the formation of an ‘assembly’ of association-area cells which can act briefly as a closed system after stimulation has ceased; this prolongs the time during which the structural changes of learning can occur and constitutes the simplest instance of a representative process (image or idea)” [28, ch. 4]. The changes mentioned here refer to “the structural change at the synapse which has been assumed to take place in learning. . . The general idea is an old one, that any two cells or systems of cells that are repeatedly active at the same time will tend to become ‘associated,’ so that activity in one facilitates activity in the other.” This “neurophysiological postulate” later came to be known as “Hebb’s rule”; similar formulations, however, can already be found in McCulloch and Pitts [40].

Regarding the formation of the cell assembly, Hebb stresses the statistical nature of both the timing of impulses and the connections to be enhanced, and how “in a larger system a statistical constancy might be quite predictable”; on anatomical considerations Braitenberg [5], [7] observes that the cerebral cortex contains all the necessary equipment and therefore would seem a good place for cell assemblies.

An important property of the Hebbian cell assembly in view of what is known about human perception, pattern recognition, and the like, is the capacity for “pattern completion”: because of the strengthened connectivity, excitation of only part of the assembly will result in an “ignition” of the entire assembly [5]. The cell assembly, as a postulated cooperative group of neurons, formed by learning and sustained by later use, thus presents an alternative to the hierarchically organized scheme of the “cardinal” neuron [4]. In order to describe the temporal sequences of assembly activity involved in perception, Hebb introduces the “phase sequence”; this point was further developed by Braitenberg [5] in the form of a threshold control mechanism, which in its dynamic effects might be called a “pump of thoughts.”

The notions of cell assemblies and modifiable synapses were later formalized [9], [8], [44], [39] and recently led to an outburst of theoretical work on neuronal networks with learning properties. In this context, it is interesting to note that many contemporary workers make a distinction between the Hebb rule and a “teacher” system [46]. However, in rereading the original books [28], [29] we noted that these two concepts were very much intermixed. In fact, every illustration that Hebb gives involves a convergence of synapses (so as to make the postsynaptic cell fire), although he did not explicitly separate the pupil and teacher synapse.

It should be stressed that both cell assembly and the modifiable synapse were introduced on theoretical grounds. Over the years, indirect evidence had accumu-

lated regarding the actual presence of Hebbian synapses in the brain. Only recently, however, has there been a direct demonstration of a Hebbian-type modifiable synapse in hippocampus slices [27] and (Bonhoeffer *et al.* in preparation).

### C. Cortical Columns

The notion of the “cortical column” as a basic functional module of the neocortex was introduced by Mountcastle [42] on the basis of physiological experiments in the cat’s somatic sensory cortex. He observed that the (static) receptive field properties (RF location on body surface and modality, i.e., nature of adequate driving stimulus) of single neurons were distributed over the cortex in a most characteristic fashion: the cortex appeared to be organized in vertical columns, running through all six cortical layers, such that neurons in a column share the same receptive field properties. Each local column is thus specified by RF location and modality; rows of columns (or slabs) with the same RF location and those with the same modality specification are arranged in more or less orthogonal fashion. Comparable observations were made for the visual cortex ([31]; reviewed [32]): in area 17 there is an orderly map of visual space onto the cortical surface; within this map there appears to be a segregation into a columnar system, with the columns characterized by ocular dominance and by the orientation of the preferred bar stimulus. Within a single ocular dominance column (width approximately 400  $\mu\text{m}$ ), there is a gradual progression of preferred bar orientation, with 1 mm displacement along the cortex surface covering the full range of 180° orientation shift (“orientation hypercolumn”). There is evidence that auditory cortex is organized into columns associated with sound frequency and localization of the sound source in space [41], [33].

In a review of the columnar organization of the neocortex, Mountcastle [43] discusses a number of general characteristics, which for our current purposes we summarize as follows: 1) the basic functional module for input–output processing in cortex is the cortical column: a vertically arranged group of cells, heavily interconnected along the vertical axis, sparsely in horizontal directions (however, see [21]); 2) the columnar system allows for an orderly representation of several variables simultaneously in a 2-D map—the depth dimension is occupied with the task of processing and distribution of inputs and outputs; 3) the order parameters of the columnar organization vary with the particular cortex area, and are determined, on the one hand by the specific thalamo–cortical input (e.g., ocularity), on the other hand (mainly) by intracortical processing (e.g., orientation tuning); 4) ordered sets of columns in different cortical areas (as well as modules in subcortical structures) are interconnected specifically, thus allowing for preservation of topological relations; 5) the columnar functional organization allows for a partially-shifted overlap across a topographical representation–dynamic isolation of the active elements of a column is achieved by lateral, pericolumnar inhibition.

Mountcastle [43] both further subdivides the column into smaller subunits, the "minicolumns," and groups columns for different order parameters into larger entities, the "macrocolumns." These levels of organization and the above-mentioned horizontal interactions among columns with the same cortical area are not undisputed, however [30], [6], [20]. Note that the functional columnar organization scheme as sketched here is not trivially connected to the modular architectonics of the neocortex as described by Szentagothai [50], [51]. That module is a honeycomb-like cortical structure with diameter around 300  $\mu\text{m}$ , anatomically defined by the arborization of cortico-cortical afferents rather than by the termination of specific thalamo-cortical afferents.

#### D. Correlation of Firing

Suppose that we are able to observe simultaneously the electrical activity of some number of neurons which are embedded in a functioning nervous system—the multineuron experiment. We may now define membership in an assembly on the basis of correlated firing, i.e., whether there are any preferred timing relationships between and among the firings of the observed neurons. The simplest such comparison we can make is of the relationship between two spike trains. The appropriate tool is cross correlation, which is a measure of delayed coincidences between spikes in the two trains. If the correlogram is flat, then no values of delayed coincidence are favored. This uniform level of correlation corresponds to the null hypothesis of independent neuron firing; its value can be simply calculated from the two firing rates, the observation interval and the bin width used in the correlogram. When peaks or valleys appear in the correlogram, there are favored values of delayed coincidence, and we may infer some relationship between the two neurons. Note that the cross-correlation measurement is an average statement over the entire piece of data being analyzed. Such averaging is necessitated by the stochastic and variable nature of neuronal firings.

Preferred values of delayed coincidence are potentially important to the operations of the brain. If two neurons converge synaptically upon a third, the probability that the postsynaptic neuron will fire will be much higher for near coincident, rather than individual, input spikes. Temporal summation in a dendrite is the likely underlying mechanism, i.e., an endogenous property of the single neuron. Thus, many neurons are effective detectors of (slightly) delayed coincidence of input, making available a powerful mechanism for computation and coding.

For the experimenter, however, the measurement of correlation among spike trains can be used to abstract the logical relationships between the observed neurons. By appropriate manipulations and control calculations we may parse the observed "raw" correlogram into contributions caused by direct stimulus effects on the two neurons and contributions caused by neural interactions. The latter can, on the basis of peak width and timing, usually be parsed into contributions from reasonably direct con-

nection between two neurons (possibly mono- and/or polysynaptic) and contributions from shared input that originates in an unobserved, but not stimulus related source. The final result of all this massage is a highly simplified block diagram—the "effective connectivity." We stress that this is an abstraction: it is the simplest neuron-like circuit that would produce an equivalent correlogram. In particular, the "effective connectivity" may be only a subset of the actual anatomical connectivity since it deals only with connections and relations that are active during the time of measurement. It is useful to describe the "effective connectivity" with a connectivity matrix  $W$  of "effective" synaptic weights. Matrix element  $W_{ij}$  would represent the "effective" influence exerted by neuron  $i$  on neuron  $j$ . Note that for real brains, in contrast to most physically oriented models (e.g., spin-glasses) the connectivity matrix  $W$  is in general not symmetric. In subsequent portions of this paper we will no longer use the double quote around effective connectivity; the meaning, however, is intended to be explicitly as above.

An excellent review of the mathematics and procedures surrounding such uses of cross correlation can be found in the book by Glaser and Ruchkin [23], along with many references to the original literature.

In developing and using such procedures over the years, our laboratory has shown that correlated firing of simultaneously recorded neurons can indeed be clearly demonstrated in a number of tissues, and that various interesting effective connectivities may be inferred [49], [11], [12]. A number of other laboratories have made similar observations; reviews can be found in Abeles [1], Gerstein *et al.* [17], and Kruger [34].

Thus, cooperativity of neuron firing is a readily demonstrated phenomenon, and can be used as a defining indicator of neural assemblies and assembly processes.

#### E. Other Definitions of Assembly

We obviously do not intend to give an exhaustive listing here. However, two additional approaches are important.

1) Palm [44] has written a book concerned with some of the mathematics of neuronal assemblies. His definition is largely in accord with that of Hebb, except that it is applied most particularly to circuit arrangements for associative memory. Such assemblies are characterized by heavy feedback upon themselves, and are able to learn, store, and differentiate a large number of input stimuli in a particularly efficient manner.

2) Shaw [47], [48] proposed to examine neuronal assemblies using the idea of ergodicity. Consider the firing pattern of a single neuron in response to some repeated short stimulus. Even though the stimulus is constant for each repetition, the response of the neuron will vary widely, presumably because of the stochastic properties of neuron firing. The usual solution to this variability has been to compute the PST (peri-stimulus-time) histogram, an average measure of the firing probability relative to the

instant(s) of stimulus presentation. In order to attain a PST histogram that is reasonably smooth, it may often be necessary to average over some 30–50 (or more) repeated stimulus presentations. Now, it is perfectly obvious that the organism owning the neurons does not require 50 repetitions of the stimulus in order to detect and perhaps react in some appropriate way; a single presentation is usually enough. Since the requisite information is not present in the firing of a single neuron during the single stimulus presentation, we must assume that ergodicity holds: there should be a neural assembly whose summed firing for a single stimulus carries the same information as does the PST histogram of a single neuron after many stimulus presentations.

Using this idea, Shaw considers the PST histogram of a single neuron's response firing to be the indirect measure of an assembly response. As a further step, he then makes cross correlations between different (more or less simultaneously) measured PST histograms (Shaw *et al.*, in preparation). Features in such correlograms are interpreted as interassembly processes.

#### COMPARISONS

The several views of neuronal assemblies sketched out above have originated over a period of 50 years. Yet they share a number of properties.

Both neuron pools and cortical columns emerged directly from experimental observation. In both, membership in the assembly is based on commonality of individual neuron properties. In the case of the motor pool it is a shared output purpose: all the neurons that can contribute to the activity of a single muscle. In the case of the cortical column it is a shared response property: all the neurons that respond to a particular stimulus parameter specification. Neither concept of assembly directly addresses possible interactions among the members of the assembly; rather, the neurons involved are treated as strictly parallel in their functions.

Both the Hebbian and correlational concepts originated theoretically, and were eventually validated by experiment. In both, the primary defining property is interaction of individual neurons. Coactivation of neurons is the direct observable in both assembly concepts, and this, of course, includes coactivation from shared stimulus response properties as in cortical columns. But alternate sources of coactivation (i.e., nonstimulus shared input or direct (synaptic) interaction) do exist, so that these concepts of assembly go far beyond cortical columns. Experimental access to assembly properties through measurement of neuronal coactivation or firing correlation necessarily requires parallel rather than serial observation of the neuron population. This has only become possible in the relatively recent past after a great deal of technological development [17], and is carried on in relatively few laboratories.

The ergodic model of assembly makes no direct assumptions and allows no inference about interactions among the member neurons. The PST histogram of one

neuron is set equal to a population response and the assembly is considered to move between active and inactive stable states [48]. There is no specification of how this population response is produced or how it is assessed in the nervous system, however.

It should be obvious that the several conceptual models of neuronal assemblies overlap only partially in what they address. Considerable further exploration is needed to allow conversion between the models.

#### CONCEPTS OF ASSEMBLY DYNAMICS

In all the above, we have considered basically static or only slowly varying assemblies. A considerable body of work exists for slow modification of cortical columns by developmental manipulations like deprivation. Spatial extent (relative size) of columns can be profoundly affected, as well as some of the response properties of the individual neurons. On a similar time scale, the Hebbian assembly first comes into being through its training period, the initial stimulations or experiences that mold the internal connectivity. In terms of the correlation model of assembly, we would therefore expect to see only constant effective connectivity in a mature assembly. The matrix of effective synaptic weights  $W_{ij}$  should be constant on time scales of (approximately) minutes to hours.

Such a static picture of assembly organization is intuitively unsatisfactory, given the rapidity with which the nervous system deals with its constantly varying environment. How could we determine experimentally whether assembly organization is really so static, or whether there are rapid changes in reaction to varying context or purpose?

The correlation concept of assembly, together with the implicit requirement for simultaneous observation of many neurons can be used to examine dynamics of assembly organization. However, new, additional analytic tools are required. Since we remain faced by the stochastic or noisy characteristics of neuron firing, average measures remain inevitable, but it is possible to divide the data periods over which the averages are taken in order to get at dynamics that are temporally related to stimulus presentation. Two relevant tools that we have recently developed are described briefly here.

#### TOOLS FOR DETECTION OF ASSEMBLY DYNAMICS

*Tool 1:* The gravity transformation. The basic difficulty in using cross-correlation measures to examine assembly properties among some number of observed neurons is purely combinatorial. Since the basic measurement is for a pair of spike trains, there is an enormous proliferation of correlograms to be examined should even a modest number like ten neurons be simultaneously recorded [the number grows as  $0.5N(N - 1)$ ]. The gravity tool is a transformation that allows the entire group of neurons under study to be evaluated for correlation of firing in a single computation, rather than in a (large) set of pair computations.

In the gravity transformation, each of the  $N$  neurons is

represented by a particle in an  $N$  space; all particles are initially equidistant. Each particle carries a charge which is incremented by every firing of the corresponding neuron, and then subsequently decays with an appropriate time constant. The charge history of a particle thus looks like a low-pass filtered version of the corresponding neuron's spike train. We now set operating rules that make particles attract each other with a force proportional to the product of their charges, and specify a viscous medium. Then particles corresponding to neurons with appreciable correlation of firing will move towards each other with a velocity proportional to the degree of firing correlation. Should there be more than one group of neurons showing (separate) firing correlations, the particle system will condense into a corresponding set of clusters. Particles corresponding to independently firing neurons will continue to wander around the  $N$  space in a way reminiscent of Brownian motion. As in the simple cross-correlation calculation, part of the result, i.e., of the clustering, will be caused by direct modulation of the individual neuron firing rates by the stimulus. Appropriate correction is possible, so that we can study the firing coordination that is not directly caused by stimulus.

Thus, the original problem of temporal correlation structure among  $N$  spike trains has been transformed into a spatial clustering problem which is easily followed and interpreted. Considerably more detailed descriptions of the gravity transformation [19], [16] and modifications of it [2] have been published. The computation is in use at several laboratories, and its particularly effective in screening blocks of data to identify those spike trains which should be studied in more detail.

The gravity transformation is effective for examining dynamics of the correlation structure at several time scales. Within a particular condensation, which may cover a period where the stimulus was presented many times, there may be changes in aggregation speed of some clusters. To the extent that this can be made significant compared to statistical fluctuations, such changes of aggregation speed correspond to changes in the underlying neuronal interaction strengths.

At a somewhat different time scale, we may compare condensations for two different stimulus conditions. The best experimental design for this purpose would be to present the two stimuli alternately and repeatedly or in a pseudorandom order. The two stimulus conditions are thus interleaved in a way which eliminates slow drift (of anything) as a confounding factor in the analysis. The recorded data must be edited by "cutting and pasting" so as to produce two continuous strips, each involving presentation of only one stimulus condition. The aggregations and the speed with which they are attained in the two condensations provide average measures that allow comparison of the assembly organizations in the two stimulus conditions. Although these are measures averaged over many repeated presentations of stimulus, the interleaved design means that any organizational changes that

are detected occur on a time scale of the interstimulus time.

*Tool 2:* The joint PST histogram. We return our attention to a single pair of spike trains. Suppose these two neurons correspond to clustering particles in the gravity representation or that there is a significant peak in the correlogram, i.e., evidence for one or more of the several types of neuronal interaction that can be represented by effective connectivity. Are the delayed coincidence events that contributed to the correlation peak distributed randomly in time? Or, in other words, is the effective connectivity constant or varying? One approach to this class of problem is to treat these particular delayed coincidence events as just another spike train, and then to use the usual tools of spike train analysis to examine the temporal structure of this higher order spike train [13], [18].

A related approach to the problem, particularly useful to examine possible stimulus effects on the effective connectivity is construction of the joint PST histogram. This is a recently improved version of the old joint PST scatter diagram [18], [3]. The basic construction is shown in Fig. 1. We cut the data at the beginning of each stimulus presentation, and then plot each of the two spike trains along a cartesian axis, placing the time of stimulus at the origin. On the plane we plot points at the locations corresponding to the logical AND of the two spike train segments. This operation is repeated for each stimulus presentation.

The points in the plane are now binned at some appropriate resolution. Bins falling along (near) the diagonal correspond to near coincidental events in the two spike trains, i.e., to events that are signatures of the several forms of neuronal interaction. Bin counts for the diagonal region [see Fig. 1(b)] represent the time locked average for near coincidences in the same sense that the PST histogram represents the time locked average of the single firings. We label these counts the PST Coincidence histogram, and present them as in the right portion of Fig. 4 along the diagonal. It can also be shown that summation over each stripe of paradiagonal bins in the joint PST histogram, normalized for the number of bins in each sum, gives the cross correlogram of the two spike trains. This is plotted at the upper right of the right portion of Fig. 4. In addition, for time reference, the ordinary PST histograms for each spike train are presented at the left edge and twice across the bottom of the display; all the relevant time relationships are simultaneously open for inspection and comparison.

As in ordinary cross correlation, the "raw" joint PST histogram includes the effects of direct stimulus modulation of the individual firing rates of the two neurons. We have shown that an optimal correction for this factor is to subtract the bin by bin product of the two PST histograms, and then to divide by the bin by bin product of the standard deviations of the two PST histograms [3], [45]. After this correction the residual correlation effects (direct connection and shared nonstimulus input) are uncovered and we can examine their stimulus-locked time course.

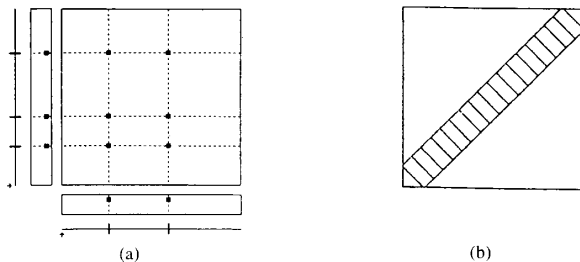


Fig. 1. Joint peri-stimulus-time histogram. (a) Method of construction. Time of occurrence of impulses from the two units relative to a stimulus presentation are shown schematically on the bottom and left axes. The contribution of each spike to its cell's PST histogram is shown in the rectangular boxes, and their joint contribution to the JPST histogram is shown in the square box. A JPST histogram like that shown in Fig. 4 gradually emerges. (b) Diagonal region of JPST. Joint events that fall into the boxed region occur at about the same time and are the "nearly coincident" events of the sort that would lead to a gravity condensation. When the firing of the horizontal axis cell significantly precedes the firing of the vertical axis cell, the event falls into the upper left triangle. Similarly, when the firing of the vertical axis cell significantly precedes the firing of the horizontal axis cell the event falls into the lower right triangle (modified from [3]).

Such a corrected JPST histogram corresponding to Fig. 4 is shown in Fig. 5.

#### MEASURED ASSEMBLY DYNAMICS

Using the above tools, we have examined various appropriate simulations as well as a number of recordings from real multineuron experiments. The simulations involved several model neurons coupled by synapses that either were constant in time or alternatively (in other cases) were modulated in strength by the stimulus. In both of these situations, the joint PST histogram procedure was correctly able to recover the known effective connectivity and its stimulus-time locked structure [3].

Real recordings, when examined with both the gravity tool and with the joint PST histogram in some cases show constant effective connectivity, while in other cases there were dynamic changes. When they appear, the modulations of effective connectivity seem to occur on two different time scales. The gravity tool as described above gives a short-term moving average measure of the correlation structure among the observed neurons over many repeated presentations of the same stimulus; we interpret this as a short-term moving average value of the effective connectivity for the particular stimulus condition. The gravity condensations for the same neurons under two different stimulus conditions are often different, suggesting different effective connectivities. However, since the experiments are run with alternate repeated presentation of the two stimulus conditions, the assembly organization must switch back and forth approximately every half second, i.e., the stimulus presentation interval [14], [15].

Stimulus effects on the effective connectivity measured by the joint PST histogram, when they occur, are on a much shorter time scale, around tens of milliseconds [15], [3], [24].

#### INTERPRETATIONS

Given that real neurons often show dynamic changes in correlation structure in association with stimulus, we have the problem of interpreting such changes in terms of assembly organization. In the course of developing the gravity and joint PST histogram tools, we used simulation to verify that they could detect and quantify modulation of connections if these had been included in the simulation. Conversely, when connections were constant in those simple simulations, the gravity and joint PST histogram measured constant correlation and hence constant effective connectivity. Thus, one possible interpretation of stimulus dependent correlation structure among real neurons would be to suggest that the connections among the neurons are changing their effective value: the  $W_{ij}$  are dynamic. Although the Hebbian scheme for synaptic change is too slow to explain such observations, mechanisms like presynaptic inhibition have an adequately fast time scale.

A different possible level of interpretation of the correlation dynamics is in the following analogy [15]. Assume that each stimulus sets up a disturbance in the neuronal mass that spreads in the manner of ripples in a pond. Different stimuli would correspond to ripples starting from different locations. As these disturbances travel through, they will cause correlated activity. After the corrections for direct stimulus modulation of firing rates we would still expect to find different residual correlation for different stimuli because the disturbances have traveled somewhat different paths, so that particular connections are differentially effective. It is possible that optical detection methods [25], [26] could be used to further examine such possibilities among large masses of neurons.

Recently, we have found another class of possible interpretations of dynamic residual correlations, i.e., a new set of mechanisms which can cause correlation dynamics without explicitly invoking changes in synaptic effectiveness.

We demonstrate this result by examining spike trains from a simulation of 100 model neurons in a square array, but with anisotropic connections so as to allow different propagation paths for stimuli applied at different locations. Specifically, the simulation was organized to mimic "synfire chains" [1] arranged to carry excitation from the top toward the bottom of the neuron array. Each horizontal row of neurons received input only from the row above, and transmitted only to the row below. By means of additional independent input, each neuron was set to a spontaneous firing rate of about five spikes per second. Stimuli were applied either to the top row of neurons, or to the leftmost column of neurons in the square array. The stimuli were 100 ms duration independent impulse trains drawn from identical exponential interval distributions. A total of 2499 stimuli were presented one every 400 ms. During the stimulus the cells in the network fired at about 15 spikes/second.

The parameter of interest for the present purposes is the probability of interrow connections, i.e., how many neu-

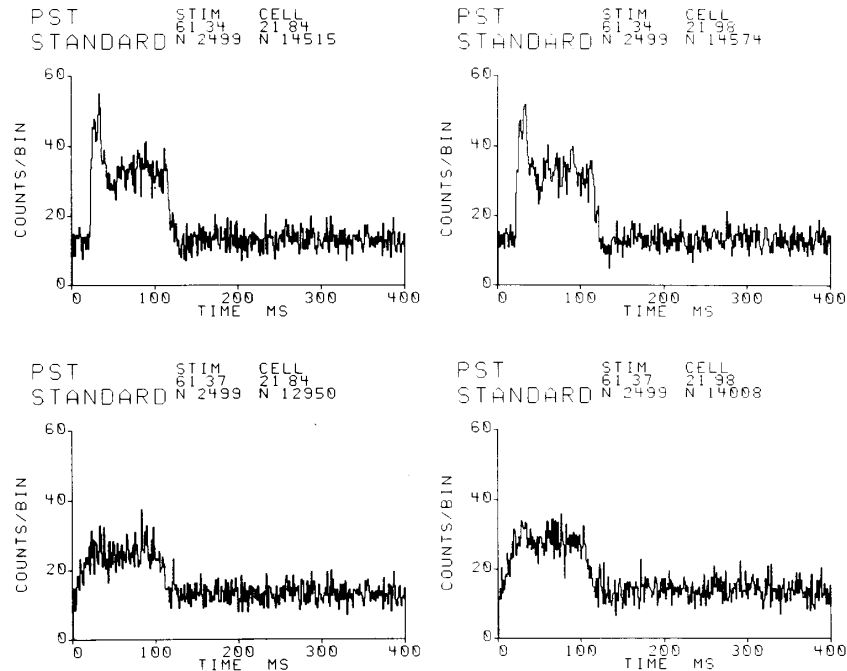


Fig. 2. PST histograms of pre- and postsynaptic cells under two different stimulus conditions. Top, stimulus presented to the top of the network; bottom, stimulus presented to the side of the network. Left is presynaptic and right is postsynaptic. See text for organization of the network.

rons in the row above feed a particular neuron. This quantity can be termed the convergence.

The figures shown here are drawn from a simulation with a high degree of convergence. Typical PST histograms for two neurons near the bottom of the net are shown in Fig. 2. The neurons have similar, but not identical PST histograms for a particular stimulus, and there is a qualitative difference for the two stimuli: stimulation at the top of the array produces a rapid rise with overshoot in the time-locked firing probability, while stimulation at the left edge of the array produces a much more gradual rise and fall in the firing probability.

Corrected cross correlograms for the same two neurons are shown in Fig. 3, including the two stimulus conditions defined above as well as spontaneous activity. It is clear that the three cross-correlogram peaks are quantitatively different, the largest corresponding to the top stimulus, the smallest to spontaneous activity. Thus, the residual correlation is a function of the stimulus condition; we would classify this situation as stimulus dependent effective connectivity, although the underlying synaptic connectivity is known to be quite constant.

The raw joint PST histogram for the same data is shown in Fig. 4. As expected, the raw PST coincidence histogram (diagonal in the right part) shows a stimulus-locked modulation, some of which can presumably be explained by the individual PST histograms. The corresponding normalized joint PST histogram is shown in Fig. 5. Note

that the normalized PST coincidence histogram still shows a fairly strong stimulus-locked modulation. Again, we would classify such findings as showing stimulus dependent effective connectivity with constant synaptic connectivity.

If, however, the convergence parameter in the simulation is taken to be smaller, so that fewer neurons converge onto a given neuron in the row below, the same group of calculations show little effects of stimulus condition. We do not show these results explicitly here, but summarize them as follows. The PST histograms are very similar to those shown in Fig. 2. The corrected cross correlograms (unlike Fig. 3) are almost the same for all stimulus conditions. The raw joint PST histogram looks very similar to Fig. 4, while the normalized joint PST histogram shows a much flatter PST coincidence histogram than does Fig. 5. More detailed simulations as well as analytic calculations (Boven *et al.*, in preparation) confirm this effect of convergence and the activity carried by it on the modulation of effective connectivity.

We have therefore demonstrated at least one situation in which a network with known constant connections among the neurons will mimic the effects of dynamic effective connectivity. The relevant conditions are a high degree of convergence from a population of neurons that are all activated by the stimuli. On the other hand, when the degree of convergence is lower, we would correctly recover the actual constant synaptic connections.

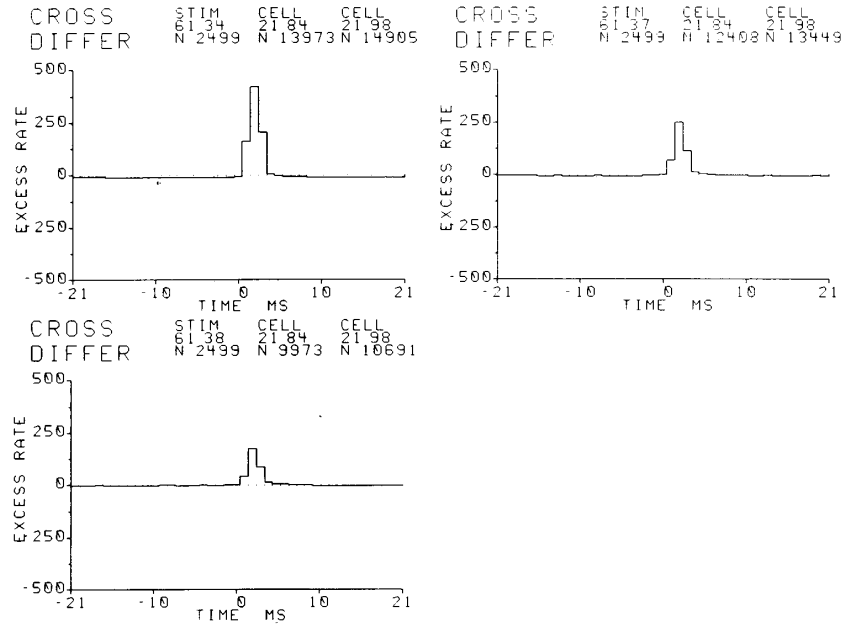


Fig. 3. Cross correlograms under two different stimulus conditions and spontaneous activity. The correlograms were corrected for stimulus modulation of the cells' firing rates by subtracting the cross correlation of the PST histograms. These differences were then normalized (divided) by the product of the number of presynaptic spikes and the bin width. This process yields the excess firing rate (in spikes per second) of the post synaptic cell per presynaptic spike. The 99 percent significance levels for these data at the scale used are invisibly small; i.e., these peaks are extremely significant. The corrected correlogram is largest for stimulation with the grain of the connections (top stimulus—top left), smaller for stimulation orthogonal to the grain of the connections (side stimulus—top right), and smallest for no stimulus (spontaneous activity—bottom left). The cross correlograms shown are for the two cells whose PST histograms are shown in Fig. 2. See text for organization of the network.

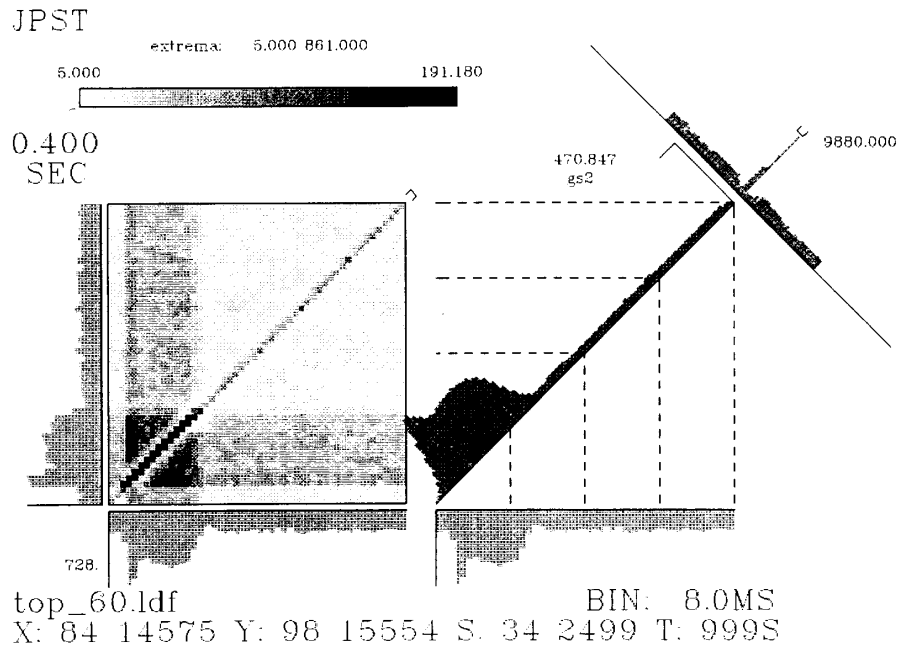


Fig. 4. Uncorrected "raw" joint peri-stimulus-time histogram. The brightness of small squares in the box at left code the number of delayed coincident spike events occurring at various poststimulus times [see Fig. 1(a)]. To the left and below the box are the PST histograms for each of the two spike trains. At right center diagonal is the smoothed PST histogram for the higher order train of near coincident spikes as produced by summation in the bins shown in Fig. 1(b). At the upper right is the cross correlogram constructed by summation within paradiagonal stripes of the JPST in the box at left and normalizing for the number of small squares along that diagonal.



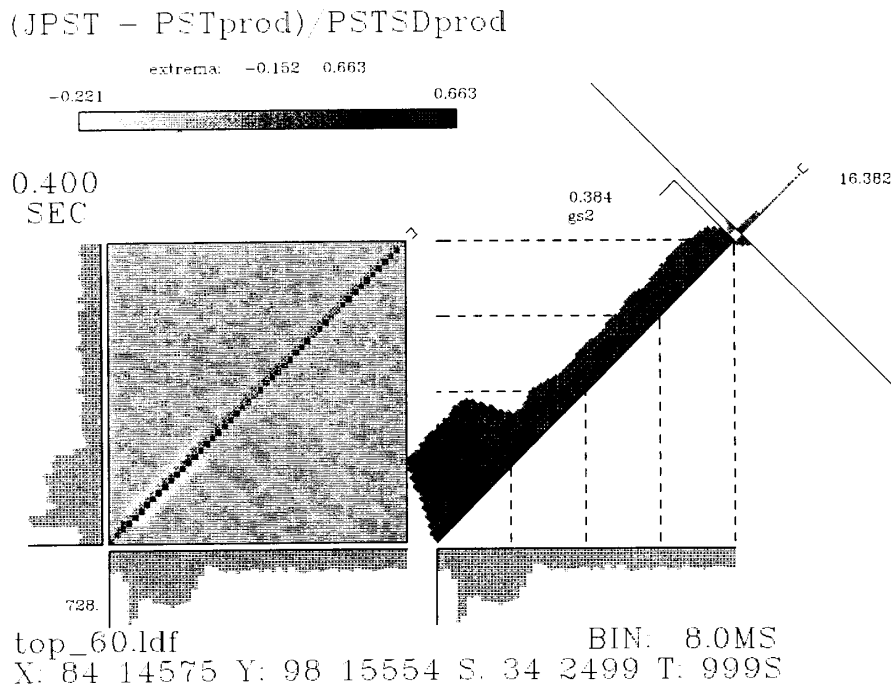


Fig. 5. Normalized joint peri-stimulus-time histogram. This is a redisplay of the information in Fig. 4 after correcting for the effects of stimulus modulation of the firing rates of the two neurons. The binwise product of the two individual PST histograms was subtracted from each small square of the JPST of Fig. 4. The value of each small square was then divided by the appropriate bin by bin product of the standard deviations of the PST histograms. Note that even after this correction the apparent connection strength shown along the diagonal at right center is modulated in response to the stimulus, whereas all traces of the single unit firing rate modulations (broad horizontal and vertical bands in Fig. 4) have disappeared.

#### DISCUSSION AND CONCLUSION

In this paper, we have examined the concept of neuronal assembly in a number of different incarnations. It turns out that each of these versions of the concept is sufficiently different that it is hard to make logical bridges.

Neither the motor pool concept or the well-known cortical column concepts of assembly depend on or imply anything about the internal structure of information flow. They are defined entirely through properties related to external input or output. In contrast, both the Hebbian and correlation models are defined mainly through interactions among the member neurons. Both also depend on input, the Hebbian model for its original development, the correlation model in relation to parsing the interactions.

All four assembly concepts allow overlaps of subassemblies. In each case, a single neuron can be a member of several different assemblies, and can participate in each of their activities. As in the previous paragraph, this multiple allegiance is defined in terms of internal circuitry or relations only in the Hebbian and correlation models.

Dynamics, stimulus related or otherwise, are implicit

in all four assembly models. However, they can be detected explicitly only in the Hebbian and correlation models. Time scales for the Hebbian model are long, while for the correlation model they are unspecified, i.e., long or short.

In the last part of this paper we have examined the extent to which a system with constant connections can appear to be dynamically organized. We have here demonstrated that at least one class of connections, i.e., highly convergent systems, can give all the measured signatures of dynamic organization. This result can be viewed as a failure of the measurements to be sufficiently sensitive, and hence as a call for further development of tools that could perhaps do the job better. On the other hand, we have shown that dynamics in the relations of two neurons can be the reflection of group activity that is elsewhere in the net, and hence not under direct observation. This, then, is a potential tool for measurement of emergent, large population properties.

Finally, none of the four assembly models explicitly deals with a way to define "state." This concept has proven vital for the successful description, understanding

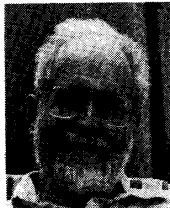
and prediction of physical systems, and there is no reason to believe it any less crucial for neuronal systems. Unfortunately, the types of state descriptors that are used on physical systems are not directly applicable to the neuronal situation. Additional work on both real and artificial neuronal assemblies is essential.

## REFERENCES

- [1] M. Abeles, *Local Cortical Circuits—Studies of Brain Function*, Vol. 6. New York: Springer, 1982.
- [2] A. Aertsen, T. Bonhoeffer, and J. Kruger, "Coherent activity in neuronal populations: Analysis and interpretation," in *Physics of Cognitive Processes*, E. R. Caianiello, Ed. Singapore: World Scientific Publishing, 1987, pp. 1-34.
- [3] A. M. H. J. Aertsen, G. L. Gerstein, M. Habib, G. Palm, P. Gochin, and J. Kruger, "Dynamics of neuronal firing correlation: Modulation of 'effective connectivity'," *J. Neurophysiol.*, to be published.
- [4] H. B. Barlow, "Single units and sensation: A doctrine for perceptual psychology?," *J. Theor. Biol.*, vol. 1, pp. 371-394, 1972.
- [5] V. Braitenberg, "Cell assemblies in the cerebral cortex," in *Theoretical Approaches to Complex Systems—Lecture Notes in Biomathematics*, Vol. 21, R. Heim and G. Palm, Eds. New York: Springer, 1978, pp. 171-188.
- [6] —, "Charting the visual cortex," in *Cerebral Cortex*, Vol. 3, A. Peters, and E. G. Jones, Eds. New York: Plenum, 1985, pp. 379-414.
- [7] —, "Two views of the cerebral cortex," in *Brain Theory*, G. Palm and A. Aertsen, Eds. New York: Springer, 1986, pp. 81-96.
- [8] G. S. Brindley, "The classification of modifiable synapses and their use in models for conditioning," *Proc. Roy. Soc. London.*, vol. B 168, pp. 361-376, 1967.
- [9] E. R. Caianiello, "Outline of a theory of thought processes and thinking machines," *J. Theor. Biol.*, vol. 2, pp. 204-235, 1961.
- [10] R. S. Creed, D. E. Denny-Brown, J. C. Eccles, E. G. T. Liddell, and C. S. Sherrington, *Reflex Activity of the Spinal Cord*. London: Oxford Univ., 1932.
- [11] J. W. Dickson and G. L. Gerstein, "Interactions between neurons in the auditory cortex of the cat," *J. Neurophysiol.*, vol. 37, pp. 1239-1261, 1974.
- [12] I. Espinosa and G. L. Gerstein, "Cortical auditory neuron interactions during presentation of 3-tone sequences: Effective connectivity," *Brain Res.* vol. 450, pp. 39-50, 1988.
- [13] R. D. Frostig, Y. Gottlieb, E. Vaadia, and M. Abeles, "The effects of stimuli on the activity and functional connectivity of local neuronal groups in the cat auditory cortex," *Brain Res.*, vol. 272, pp. 211-221, 1983.
- [14] G. L. Gerstein, "Interactions within neuronal assemblies: Theory and experiment," in *Brain Organization and Memory: Cells, Systems and Circuits*, J. L. McGaugh, N. M. Weinberger, and G. Lynch, Eds. London: Oxford Univ., 1988.
- [15] —, "Information flow and state in cortical neural networks: Interpreting multi-neuron experiments," in *Organization of Neural Networks*, W. v. Seelen, G. Shaw, and R. Leinhos, Eds. Weinheim: VCH Verlagsgesellschaft, 1987, pp. 51-75.
- [16] G. L. Gerstein and A. M. H. J. Aertsen, "Representation of cooperative firing activity among simultaneously recorded neurons," *J. Neurophysiol.*, vol. 54, pp. 1513-1528, 1985.
- [17] G. L. Gerstein, M. J. Bloom, I. E. Espinosa, S. Evanczuk, and M. R. Turner, "Design of a laboratory for multineuron studies," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-13, pp. 668-676, 1983.
- [18] G. L. Gerstein and D. H. Perkel, "Mutual temporal relationships among neuronal spike trains," *Biophysical J.*, vol. 12, pp. 453-473, 1972.
- [19] G. L. Gerstein, D. H. Perkel, and J. E. Dayhoff, "Cooperative firing activity in simultaneously recorded populations of neurons: Detection and measurement," *J. Neurosci.*, vol. 5, pp. 881-889, 1985.
- [20] C. D. Gilbert, "Horizontal integration in the neocortex," *TINS*, vol. 8, pp. 160-165, 1985.
- [21] C. D. Gilbert and T. N. Wiesel, "Clustered intrinsic connections in cat visual cortex," *J. Neurosci.*, vol. 3, pp. 1116-1133, 1983.
- [22] J. A. M. van Gisbergen, A. J. van Opstal, and A. A. Tax, "Collicular ensemble coding of saccades based on vector summation," *Neurosci.*, vol. 21, pp. 541-555, 1987.
- [23] E. M. Glaser and D. S. Ruchkin, *Principles of Neurobiological Signal Analysis*. New York: Academic, 1976.
- [24] P. M. Gochin, G. L. Gerstein, and J. A. Kaltenbach, "Dynamic temporal properties of synaptic interaction in rat dorsal cochlear nucleus" to be published.
- [25] A. Grinvald, "Real-time optical mapping of neuronal activity: From single growth cones to the intact mammalian brain," *Ann. Rev. Neurosci.*, vol. 8, pp. 263-305, 1985.
- [26] A. Grinvald, B. M. Salzberg, V. Lev-Ram, and R. Hildesheim, "Optical recording of synaptic potentials from processes of single neurons using intracellular potentiometric dyes," *Biophys. J.*, vol. 51, pp. 643-651, 1987.
- [27] B. Gustaffson, H. Wigstrom, W. C. Abraham, and Y.-Y. Huang, "Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials," *J. Neurosci.*, vol. 7, pp. 774-780, 1987.
- [28] D. O. Hebb, *The Organization of Behavior—A Neuropsychological Theory*. New York: Wiley, 1949.
- [29] —, *A Textbook of Psychology*. 3rd ed. Philadelphia: Saunders, 1972.
- [30] D. H. Hubel, "Columns and their function in the primate visual cortex," in *Theoretical Approaches in Neurobiology*, W. E. Reichardt, T. Poggio, Eds. Cambridge, MA: M.I.T., 1981, pp. 109-115.
- [31] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, vol. 195, pp. 215-243, 1968.
- [32] D. H. Hubel and T. N. Wiesel, "Functional architecture of macaque monkey visual cortex" (Ferrier lecture), *Proc. Roy. Soc. London*, vol. B. 198, pp. 1-59, 1977.
- [33] T. J. Imig and H. O. Adrian, "Binaural columns in the primary field (A1) of cat auditory cortex," *Brain Res.*, vol. 138, pp. 241-257, 1977.
- [34] J. Krüger, "Simultaneous individual recordings from many cerebral neurons: techniques and results," *Rev. Physiol. Biochem. Pharm.*, vol. 98, pp. 177-233, 1983.
- [35] C. Lee, W. H. Rohrer, and D. L. Sparks, "Population coding of saccadic eye movements by neurons in the superior colliculus," *Nature*, vol. 332, pp. 357-360, 1988.
- [36] D. P. C. Lloyd, "Reflex action in relation to pattern and peripheral source of afferent stimulation," *J. Neurophysiol.*, vol. 6, pp. 111-119, 1943.
- [37] —, "On the relation between discharge zone and subliminal fringe in a motoneuron pool supplied by a homogeneous presynaptic pathway," *Yale J. Biol. Med.*, vol. 18, pp. 117-121, 1945.
- [38] —, "Facilitation and inhibition of spinal motoneurons," *J. Neurophysiol.*, vol. 9, pp. 421-438, 1946.
- [39] C. von der Malsburg, "Nervous structures with dynamical links," *Ber. Bunsenges. Phys. Chem.*, vol. 89, pp. 703-710, 1985.
- [40] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, pp. 115-133, 1943.
- [41] J. Middlebrooks, J. C. Dykes, and M. M. Merzenich, "Binaural response-specific bands in primary auditory cortex (A1) of the cat: topographical organization orthogonal to isofrequency contours," *Brain Res.*, vol. 181, pp. 31-48, 1980.
- [42] V. B. Mountcastle, "Modality and topographic properties of single neurons of cat's somatic sensory cortex," *J. Neurophysiol.*, vol. 20, pp. 408-434, 1957.
- [43] —, "An organizing principle for cerebral function: the unit module and the distributed system," in *The Mindful Brain*, G. M. Edelman and V. B. Mountcastle, Eds. Cambridge, MA: M.I.T., 1977, pp. 7-50.
- [44] G. Palm, *Neural Assemblies—Studies of Brain Functions*, Vol. 7. New York: Springer, 1982.
- [45] G. Palm, A. M. H. J. Aertsen, and G. L. Gerstein, "On the significance of Correlations among neuronal spike trains," *Biol. Cybern.*, vol. 59, pp. 1-11, 1988.
- [46] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing. Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: M.I.T., 1986, pp. 318-362.
- [47] G. L. Shaw and D. J. Silverman, "Simulations of the trion model and the search for the code of higher cortical processing," in *Computer Simulation in Brain Science*, R. Cotterill, Ed. Cambridge, UK: Cambridge Univ. Press, 1988.
- [48] G. L. Shaw, E. Harth, and A. B. Scheibel, "Cooperativity in brain

function: Assemblies of approximately 30 neurons," *Exp. Neurol.*, vol. 77, pp. 324-358, 1982.

- [49] J. K. Stevens and G. L. Gerstein, "Interactions between cat lateral geniculate neurons," *J. Neurophysiol.*, vol. 39, pp. 309-326, 1976.
- [50] J. Szentagothai, "The neuron network of the cerebral cortex: a functional interpretation" (Ferrier lecture), *Proc. Roy. Soc. London*, vol. B. 201, pp. 219-248, 1978.
- [51] J. Szentagothai, "The modular architectonic principle of neural centers," *Rev. Physiol. Biochem. Pharm.*, vol. 98, pp. 11-61, 1983.
- [52] D. H. Perkel, G. L. Gerstein, and G. P. Moore, "Neuronal spike trains and stochastic point processes: II—Simultaneous spike trains," *Biophys. J.*, vol. 7, pp. 419-440, 1967.



**George L. Gerstein** was born in Berlin, Germany, in 1933 and emigrated to the United States in 1938. He received the B.A., M.A., and Ph.D. degrees in physics from Harvard University, Cambridge, MA, in 1952, 1954, and 1958, respectively.

He switched interests to neurobiology during a postdoctoral stay at the Massachusetts Institute of Technology, Cambridge, MA. Since 1964 he has been a Professor in the Physiology Department at the University of Pennsylvania, Philadelphia, PA.

He has worked on the neurobiology of the auditory and visual systems, with particular emphasis on the development of spike-train analysis as a way to study coding and information transfer. During the past decade he has been involved in experimental and theoretical investigation of neuronal assemblies as elements of computational processes in the brain.



**Purvis Bedenbaugh** was born on July 11, 1961. He received the B.S.E. degree in biomedical engineering from Duke University, Durham, NC in 1983 and the M.S. degree in bioengineering from Clemson University, Clemson, SC in 1984. He is currently in the Ph.D. program in bioengineering at the University of Pennsylvania, Philadelphia, PA.

He became interested in neuroscience while working in the laboratory of B.S. Nashold, Jr., of the Division of Neurosurgery, Duke University

School of Medicine. His current interests are in quantitative study of neural assemblies and integrative activity in the nervous system.



**Ad M. H. J. Aertsen** was born in Oosterhout, The Netherlands, in 1948. He received the M.Sc. degree in physics from the Utrecht State University, The Netherlands, in 1973, and the Ph.D. degree in physics from the Nijmegen University, The Netherlands, in 1981.

From 1983 to 1984 he spent a postdoctoral year with George Gerstein at the Department of Physiology, University of Pennsylvania School of Medicine, Philadelphia. Since 1984 he has been a Research Associate in Valentino Braitenberg's

Department at the Max-Planck-Institute for Biological Cybernetics, Tuebingen, Federal Republic of Germany. His current research interests focus on the analysis of neurobiological signals (single- and multineuron spike trains, optical recordings), primarily from the auditory and visual system, in relation to experimental and theoretical aspects of neural information processing.