ELSEVIER

Physica D 75 (1994) 103–128

PHYSICA D

Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation

Ad Aertsen[a,1], Michael Erb[b], Günther Palm[c]

[a]*Institut für Neuroinformatik, Ruhr-Universität, PO Box 102184, W-44780 Bochum, Germany*
[b]*Abteilung Biophysik, Fakultät für Physik, Philipps-Universität, Renthof 7, W-35037 Marburg, Germany*
[c]*Abteilung Neuroinformatik, Fakultät für Informatik, Universität Ulm, Oberer Eselsberg, W-89069 Ulm, Germany*

Abstract

Electrophysiological studies of cortical function on the basis of simultaneous, separable multi-neuron recordings reveal interactions between cortical neurons which strongly depend on the stimulus and behavioral context. Moreover, these interactions exhibit dynamics on several different time scales, with time constants of modulation as low as tens of milliseconds. Possible mechanisms underlying such dynamic organization of the cortical network were investigated by simulations and analytic approaches. We review results from these different studies, concentrating on a comparison of the correlation dynamics in real cortical activity and in various neural network models. In particular, we discuss the influence of global network activity on the functional coupling between neurons embedded in such networks.

1. Introduction

Ever since the times of Sherrington [1] and Hebb [2], neurobiologists have pursued the idea that neurons do not act in isolation, but rather that they organize into assemblies for various computational tasks (see also [3] for an early formulation of this concept). Over the years, a number of different definitions of '*neural assembly*' have been proposed. Some of these were phrased in terms of anatomy, others in terms of shared function or of shared stimulus response (see [4] for a review). One operational definition for the cell assembly has been particularly influential: near-simultaneity or some other specific timing relation in the firing of the participating neurons. As, for instance, elaborated in the concept of the 'synfire chain' [5,6], the synaptic influence of multiple neurons converging onto others in the cortical network is much stronger if they fire in (near-)coincidence. Thus, *temporal coherence* or *synchronous firing*, postulated as a mechanism for perceptual integration [2], is in fact directly available to the brain as a potential neural code [7–9].

The notion that the functional organization of the cortex is based upon interactions within and among groups of cells in large neural networks is supported by the anatomical structure and, in particular, by the massive connectivity of this part of the brain [10]. Until recently, however, very few physiological data have directly addressed the cell assembly hypothesis. Neither the study of global activity in large populations of

[1] Present address: Center for Research of Higher Brain Function, Weizmann Institute of Science, Rehovot 76100, Israel.

0167-2789/94/$07.00
SSDI 0167-2789(93)00070-7

neurons, nor the recording of single neuron activity allow for a critical test of this concept. Rather, one seeks to observe the activities of many separate neurons simultaneously, preferably in awake, behaving animals, and to analyze these 'multi-neuron activities' for possible signs of (dynamic) interactions between the neurons. Results of such analyses may then be used to draw inference regarding the processes taking place within and between hypothetical cell assemblies. In this paper we will review some of the evidence produced in these experiments, and discuss their interpretation against the background of a theoretical study on the activity dynamics in biologically inspired neural network models.

2. Interactions between cortical neurons

The conventional method to analyze neural interactions is to cross-correlate the spike trains, usually from pairs of neurons, that were recorded simultaneously under some appropriate stimulus and/or behavioral conditions, and to inspect the results for departures of 'independence' [11,12]. Following [13], peaks and troughs in the cross-correlograms, after comparison with appropriate control measurements, are characterized by parameters describing their shape (symmetry, width, sign) and delay, and their possible dependence on stimulus or behavioral features. On the basis of such descriptors, the type and temporal acuity of the interaction exhibited by the neurons are then interpreted in terms of their 'functional coupling'. Recent developments in analysis methodology have considerably expanded the scope of these studies. It is now possible to examine the cooperativity in larger groups of neurons [14–16], to study the fine-temporal properties of firing correlation between two or three neurons [17,18], and to demonstrate the existence of accurate spatio-temporal firing patterns in the activity of single and multiple neurons [19–21].

In this paper we will focus on two of these approaches, addressing two different levels of spatio-temporal resolution. The first one is based on the *network correlation matrix*, and provides a condensed description of the functional couplings among an entire group of simultaneously recorded cortical neurons. The second one applies the *Joint-PSTH*, aiming at a fine-grained dynamic analysis of the interactions among two members of such a group. The salient result of such direct assembly observation is that the functional coupling among cortical neurons is *context-dependent* and *dynamic* on several different time scales. We will document these dynamic and context-dependent assembly properties using examples drawn from cat visual cortex recordings, made in the Krüger laboratory [22]; similar results have been reported for other cortical areas [23–26].

3. Context-dependence of coupling among groups of cortical neurons

When applied to the simultaneously recorded activity of larger groups of neurons, the pairwise analysis of cross-correlation inevitably leads to sizable numbers of correlograms. Thus, it was soon realized that there is a distinct need for a more concise representation of results. A compact way to present the coupling within a larger group of neurons is in the form of a '*connectivity matrix*', in which the value at a particular entry (i, j) denotes the strength of the coupling from neuron i to neuron j. This representation is, in fact, widely used in neural network theory. Note that the connection between any two nodes in a connectivity matrix needs not be symmetrical (neither does the cross-correlogram): the coupling from neuron i to neuron j is not necessarily the same as from neuron j to neuron i. A visual representation of the connectivity matrix is obtained by displaying the matrix values using a grey- or pseudo-color coding.

We have adopted this concept of the connec-

tivity matrix in the approach of the *'network correlation matrix'* [16], a modified realization of the 'gravitational clustering' algorithm [14,15]. Briefly, the algorithm 'learns' the coupling matrix among the neurons by analyzing their pairwise spike correlations in running time. Initially, all entries in the correlation matrix are set to zero; this expresses the lack of knowledge on the part of the experimenter at the onset of the analysis. As the recording proceeds, the neural activity is dynamically processed by the algorithm, and the various entries in the matrix are updated on the basis of (near) spike coincidences occurring among the corresponding neuron pairs. This updating takes the form of a Hebbian or a covariance-type 'learning rule'. Thus, a relative abundance of spike–spike and silence–silence coincidences among two neurons will give rise to a positive 'coupling', a net surplus of spike–silence combinations will lead to a negative value. In the course of time, the algorithm thereby determines a coupling matrix, which reflects the time-accumulated firing correlations among the group of neurons under the conditions prevailing during the recording. By comparing such correlation matrices for different experimental conditions (e.g. different stimuli or different modes of behavior), one obtains a compact description of the coupling among the entire group of observed neurons, and its possible dependence on stimulus and behavioral parameters.

Different types of normalization in the 'learning rule' of the correlation algorithm reveal different aspects of the network coupling: (1) *'raw' correlation* values (i.e. without any normalization) emphasize the full synchrony of firing, not subdivided into its possible constituents, such as co-modulation of firing rates through common stimulus or neuronal drive and/or spike correlation caused by inter-neuronal connectivity, (2) *'residual' correlation* (technically speaking, the cross-covariance), a differential measure obtained after subtraction of the contribution from firing rate co-variations, shows the net effect of functional coupling, still 'scaled' by the ongoing rates of the participating neurons, and (3) the *'normalized' correlation* (the normalized cross-covariance), a relative measure obtained after additional scaling for response modulations, reflects the strength of functional coupling, normalized for the time-varying rates of the associated neurons. Calibration runs for various types of simulated spike trains demonstrated that this network correlation algorithm (like the gravitational clustering) has superior sensitivity as compared to conventional cross-correlation analysis. Even under conditions of sparse firing and moderately weak connectivity (typical for cortical tissue), a mere few seconds of recording suffice to acquire a reliable estimate of the network coupling. (For a more extensive description of the correlation matrix algorithm, its mathematical formulation, and a discussion on the various rules to 'normalize' for firing rate modulations, we refer to [16].)

Let us consider a representative example of the assembly properties revealed by such network correlation matrices. Physiological data were taken from a set of experiments in a 12-electrode recording from cat visual cortex (area 17), made in the Krüger laboratory [22]. The electrodes were a 12-fold linear, organ pipe-like array of glass-coated Pt-Ir wires, with a spacing of 160 microns. These were introduced into the cortex such that their tips fell on a line perpendicular to the cortical surface, i.e. within a single cortical column. The top electrode was near the pial surface, the lower electrode was just into the white matter. 5 out of 12 electrodes (nrs. 3, 6, 7, 8, 9) each gave a reliable single unit recording, the remaining ones were judged to be probably single unit, with possibly additional spikes from one or more other units. The stimulus in these experiments consisted of a light bar (3° by 14′), moving at constant velocity and perpendicular to its orientation. The distance travelled was 3° in 1.8 s, arranged to cover the entire receptive fields of the neurons under study. At the end of the movement the bar remained stationary for

0.4 s, after which it moved in the opposite direction. Finally the bar was rotated by 22.5° (duration 1 s) after which another cycle of movement back and forth, perpendicular to the new orientation, was begun. This protocol was continued until after 8 cycles of 5 s, the original orientation was reached again. The complete stimulus sequence, lasting 40 seconds, was presented repeatedly; in the example shown, the first 16 trials (out of 22) were used for analysis (this restriction was caused by memory limitations of our computer).

The analysis of network coupling for the whole spectrum of stimulus orientations and directions of bar movement tested is summarized in Fig. 1. Fig. 1a shows the 'residual' spike correlations, i.e. the firing synchrony after subtraction of stimulus-driven firing fire correlations, Fig. 1b represents the 'normalized' correlations, properly scaled for the individual rates and their modulations. The layout of these figures reflects the different stimulus parameters used: they consist of a ring-like arrangement of 16 correlation matrices, each of which represents the network coupling exhibited for a particular bar orientation. The individual 12×12 matrices are positioned at the tip of a vector representing the associated direction of bar movement. Values in the ith column of a matrix (i numbered from left to right) represent weight factors of 'incoming connections', i.e. ending in neuron i; values in the jth row (j numbered from bottom to top) represent weight factors of 'outgoing connections', i.e. originating in neuron j. Values along the diagonal (entries (i, i)) were not updated and, thus, represent zero coupling. The strength of coupling is coded in a linear grey-scale, with larger values corresponding to darker grey. In each figure, the 16 matrices were plotted using the same scale, covering the entire range of matrix values in that figure; this enables a direct comparison of the network coupling under different experimental conditions.

Fig. 1a demonstrates the existence of a considerable degree of residual correlation among the observed group of neurons. This is revealed by the patch-like groupings of positive (i.e. darker grey) matrix entries 'inter-connecting' most

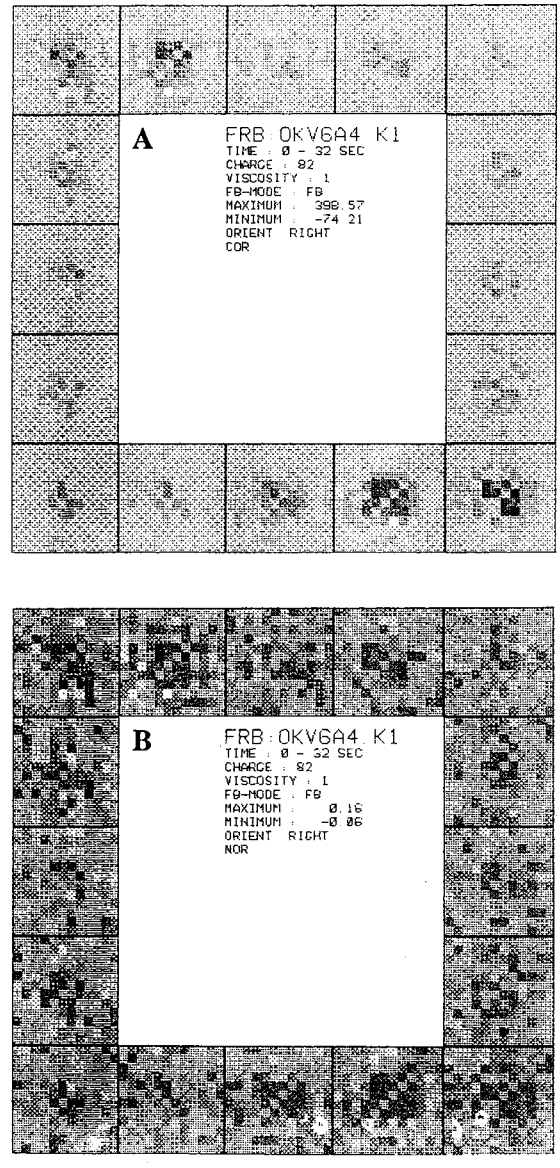


Fig. 1. Stimulus-dependence of network coupling in the cat visual cortex (area 17; data from Krüger [22]). Matrices of residual (A) and normalized (B) correlation between all 66 pairs of a 12-electrode recording for 16 directions of motion of a light bar stimulus. Each single network correlation matrix is positioned at the tip of a vector, indicating the direction of bar motion for which the matrix was computed. All matrices within a single panel are plotted with the same grey-scale.

(though not all, and not in a symmetrical fashion) of the neurons 4 to 9. Moreover, the degree of temporal coherence among these neurons is clearly stimulus-dependent: the strongest excess correlation is found for movements to the lower right, somewhat less strong for movements to the upper left, while the correlation decreases continuously the more the orientation and the direction of motion deviate from the optimal ones. Some degree of residual correlation, however, remains even in the weakest cases. Notice also that the distribution of high correlation values is not 'spatially' uniform over the matrix, but rather comes in the form of a compact block surrounding the diagonal. Recalling that the ordering of the neurons is according to their location in a linear, vertical array across the cortical depth, this strongly suggests that the coupling between neurons is not homogeneously distributed, but rather obeys the local neighborhood relations imposed by the cortical layering.

We should point out that some of the observed stimulus-dependence in Fig. 1a might be overestimated, even though the 'residual' correlations show the excess of spike correlation *after* subtraction of direct stimulus influences on the individual firing rates. The reason is that the cross-covariance presents an absolute measure of correlation, rather than a relative one. Thus, it is sensitive to the magnitude of the auto-covariances of the individual spike trains, which, in turn, might depend on the firing rate variations of the associated neurons. As we have demonstrated elsewhere [16–18], such dependence can be accounted for by proper normalization. Basically, this normalization amounts to a dynamical scaling of entries in the residual correlation matrix by the appropriate time-dependent auto-covariances of the spike trains involved. The result of this scaling operation is the 'normalized' correlation mentioned earlier. It is a relative measure of spike coherence, with values by definition restricted to the range between -1 and $+1$.

The normalized correlation matrices for the present example are shown in Fig. 1b. One observes that these reveal a much more differentiated picture of the network interactions, and especially of their stimulus-dependence, than Fig. 1a. Nevertheless, our findings from Fig. 1a are essentially confirmed. The dark patch of positive inter-group coupling among the recording sites 4 to 9 remains present, as well as its stimulus-dependence, with a preference for stimulus movements to the lower right, to a lesser degree for motion to the top left, and gradually decreasing for directions inbetween. In contrast to Fig. 1a, however, the coupling among this subgroup is not so dominant anymore. In addition one also observes interactions among other neuron pairs and/or for other stimulus orientations which so far escaped our attention, presumably because they were masked by coupling among more strongly firing neurons. Coupling on the whole, although being quite widespread, assumes only moderate values between -0.06 (white) and $+0.16$ (black). The asymmetry of this range with respect to the value zero may reflect a true asymmetry in the coupling, although we suspect that at least part of it is due to a preferential sensitivity of our algorithm for 'positive' coupling in the case of low firing rates (see a discussion on this issue in [12]).

We found the present example to be quite typical for these visual cortex experiments. In general, the network coupling as revealed by such correlation matrices involves a considerable fraction of the recorded neurons, is confined to moderately low values (usually between -0.1 and $+0.3$), and shows distinct stimulus-dependences. Taken together, these findings are not so much indicative of small, dedicated local cortical circuits, but rather point at relatively diffuse distributions of coupling, involving considerable fractions of the neurons in the cortical network. Note, however, that this observation may well depend on the geometrical arrangement of the recording sites: in all these experiments they were aligned linearly in a vertical cross section of

the cortex, presumably located within a single cortical column.

4. Dynamics of functional coupling between cortical neurons

So far our discussion focused on the global properties of network coupling, such as the degree of (dis)similarity of entire correlation matrices compared across different stimulus conditions. It should be born in mind, however, that this analysis addresses the interactions in the network only at the most general level. In fact, the temporal development of each individual grey pixel in each single matrix represents highly condensed information regarding the dynamics of interaction among a specific pair of neurons under certain experimental conditions. We will now turn to a much more detailed level of analysis, and study the modulations of functional coupling among pairs of neurons on a much faster time scale. There are a number of possible approaches to this problem; here we will use the Joint Peri Stimulus Time Histogram [18]. The Joint-PSTH is a temporal decomposition matrix of the ordinary cross-correlogram, and was specifically designed to describe the dynamical aspects of the firing correlation that are time-locked to a stimulus or behavioral event. Moreover, appropriate normalization of the Joint-PSTH enables us to distinguish contributions due to stimulus- or behavior-induced modulations of the individual neuron firing rates from these of inter-neuronal spike correlations, and to evaluate these differences for statistical significance [17]. Similarly to the case of the network correlation matrices, the normalization scales the cross-covariance for the associated pair of auto-covariances. For the Joint-PSTH this involves subtracting from each bin in the Joint-PSTH matrix the product of the two corresponding bins in the individual PST histograms, and dividing the result by the product of the standard deviations of these same PST bins. Subsequent summations over appropriate diagonal and para-diagonal bins in the matrix plane give us the appropriately normalized, 'neuronal interactions only', PST-coincidence histogram and cross-correlogram. Thus, we can examine fast, stimulus-related changes in the interactions between two observed neurons. (For a more detailed account and mathematical formulations the reader is referred to [17,18].)

Fig. 3 shows the results of Joint-PSTH analysis of the activity of two neurons, 7 and 8, taken from the interacting subgroup in Fig. 1. Spike train data, shown in Fig. 2, were taken from time sections of 5 seconds each (number of trials = 22), during which the light bar stimulus was moving back and forth in opposite directions, corresponding to "south–south–east" and "north–north–west" in Figs. 1a,b. The two panels in this figure show the 'raw' (Fig. 3a) and the normalized Joint-PSTH (Fig. 3b), together with the associated marginal distributions for the selected two neurons. Within each panel, the individual PST histograms, displaying the individual firing rates, are shown along the horizontal and vertical axes. The density of coincident firing is indicated in the Joint-PSTH matrix, according to the grey code above the matrix. At the right of each panel are respectively the PST-coincidence histogram (along the diagonal rising

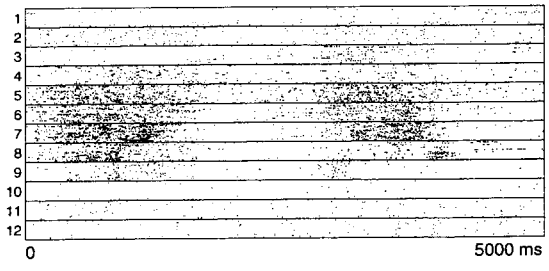


Fig. 2. Spike activity from a 12-electrode recording in the cat visual cortex. The dot display shows a selected portion of the spike train data used in Fig. 1. Data were taken from time sections of 5 seconds each (number of trials = 22), during which the light bar stimulus moved back and forth in opposite directions, corresponding to "south–south–east" (first half of the trial) and "north–north–west" (second half) in the matrix arrangement in Fig. 1.

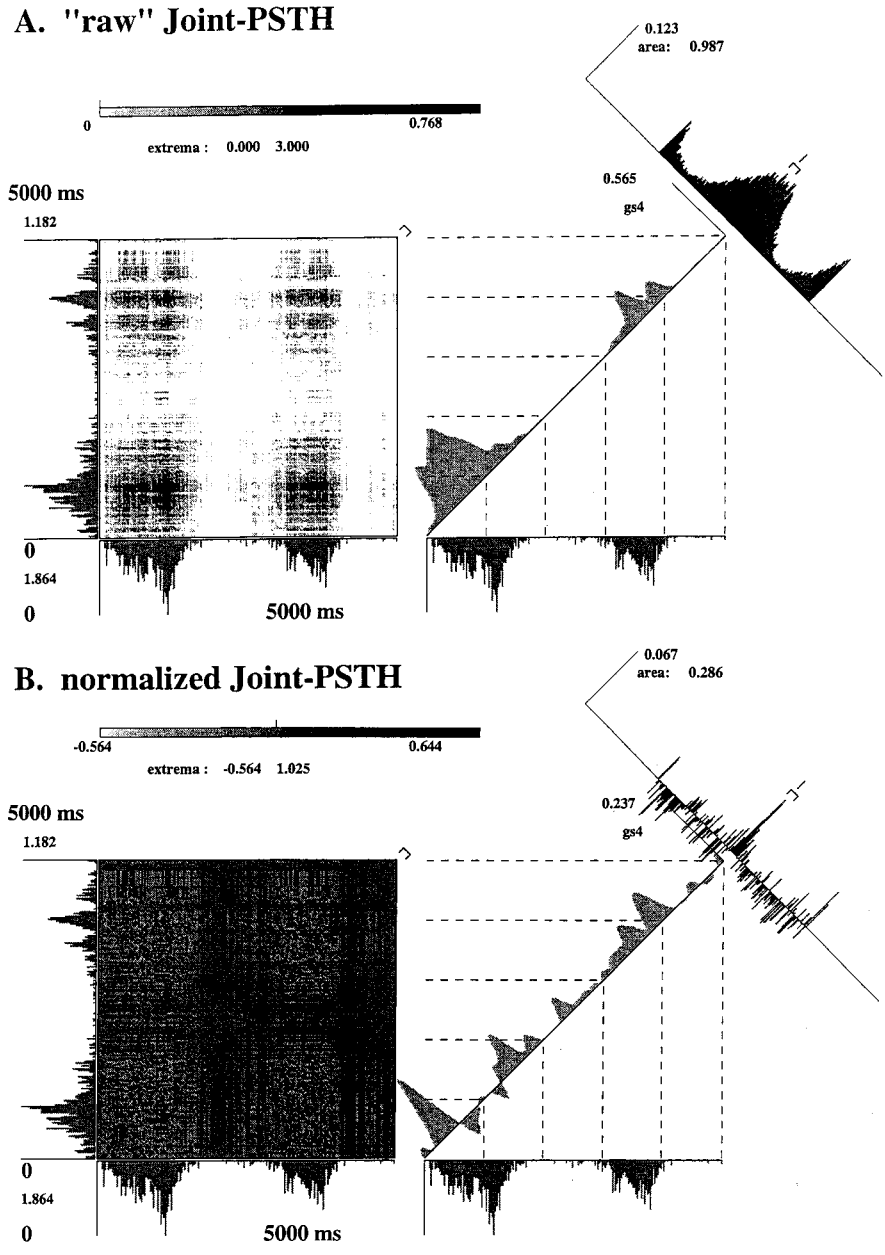


Fig. 3. Stimulus-locked dynamic correlation of firing for two single neurons (7 and 8), recorded simultaneously from the cat visual cortex during repeated presentation of a moving bar stimulus. The corresponding spike data are shown in Fig. 2. The upper panel (A) shows the 'raw' Joint-PSTH; the lower panel (B) shows the Joint-PSTH after normalization for stimulus-induced non-stationarities in the single neuron firing rates. The latter is calculated by subtracting from the 'raw' Joint-PSTH the cross-product matrix of the individual PST histograms, and by dividing the resulting difference matrix (bin by bin) by the cross-product matrix of the standard deviations of the PST histograms. The display format of both panels is the same. The left-hand half of the panel shows the Joint-PSTH matrix and the two ordinary PST histograms along its x- and y-axes (binwidth: 25 ms). Values in the Joint-PSTH matrix are coded in grey as indicated in the bar above the matrix. The tic mark above the grey bar corresponds to the value zero. All counts were divided by the number of stimulus presentations. The right-hand half of the panel shows the PST-coincidence histogram (running along the diagonal from lower left to upper right) and the conventional cross-correlation histogram (perpendicular to the diagonal, and running from upper left to lower right). The PST-coincidence histogram was smoothed using a Gaussian with a sigma of four bins; this particular value (gs4), as well as the location and width of the selected diagonal band are indicated in each panel. The position of true coincidence (zero delay) in the cross-correlogram coincides with the intersection point of the PST-coincidence histogram and the cross-correlogram; it is indicated by a tic mark above the diagonal band marker above the correlogram. Numbers of spikes: 1420 (neuron 7, x-axis), 788 (neuron 8, y-axis), recorded during 22 stimulus trials of 5 seconds each.

to the right) and the conventional cross-correlogram (perpendicular to the diagonal and descending to the right). The PST-coincidence histogram measures the counts near the diagonal of the Joint matrix and represents the time-locked rate of near-coincident firing of the two neurons, in the same sense as the ordinary PST histogram represents the stimulus time-locked rate of the individual neuron's firing.

In the 'raw' Joint-PSTH panel (Fig. 3a) we note that the individual PST histograms show a strong time-locked increase in firing for both neurons as the stimulus enters their receptive fields, with a clear preference for the first direction of motion (compare also Fig. 2). These are direction (as well as orientation) sensitive neurons. Similarly, the 'raw' Joint-PSTH matrix shows considerable hills at locations matching the PST peaks, and corresponding peaks are visible in the PST-coincidence histogram. In other words, individual, joint, and near-coincidence firings are all increased during portions of both directions of movement, but more so in the first direction. Note also that the time-averaged cross-correlogram (top right) is fully dominated by the co-modulation of firing rates, thus effectively masking any possible sign of neural interaction. The contribution of these stimulus time-locked modulations of the individual firing rates have been removed in the normalized Joint-PSTH, shown in Fig. 3b. Note first that after normalization, the time-averaged cross-correlogram shows a distinct peak, straddling the origin. This clearly points at the presence of spike synchronization exceeding the mere co-variation of rates. In order to find out the time course of this excess spike correlation, we concentrate our attention on the normalized PST-coincidence histogram along the diagonal. Note that during most of the time of the *second* direction of movement (roughly, between 3000 and 4000 ms), there is a positive level of excess near-coincidence firing. This level is not constant, though, but rather follows the time course of the individual neurons' firing rates. Note, however, the brief, but strong burst of excess co-firing around 2500 ms, just when the responses to the second direction of motion start to develop. The situation is much more complicated during the *first*, and opposite, direction of movement (roughly, between 0 and 2000 ms). Here we also observe a varying level of correlated firing, but its time course is altogether different from the individual neurons' responses. After an initial, large positive peak, coinciding with the rising phase of the individual responses, the correlation breaks down to zero and even becomes negative (i.e. there is a lower than expected amount of near-coincidence firing) as the individual firing rates reach their maximum. This stage of zero or negative interaction lasts for most of the duration of the peak response rates. Finally, the correlation again climbs to large positive values when the responses have reached their declining phase. (For a similar, even more pronounced, case of coupling with sign-reversal among visual cortex neurons see [23].)

In the example shown here, the switching from a condition of co-firing to one of anti-co-firing seems to be associated, at least roughly, with the transitions of low to high firing rate of the two observed neurons (compare the time course of the PST histograms with that of the normalized PST-coincidence histogram). This pattern is, in fact, very suggestive of a non-monotonic relation between the degree of spike synchrony and the amplitude of the single neuron responses. The spike correlation is positive when firing is weak to moderately strong (i.e. during the rising and falling stages of the first response and throughout most of the second response), whereas it becomes zero or even negative when the firing rates are high (as during the peak of the first response). For the first (and stronger) response, this non-monotonicity causes the interaction to behave essentially like a 'transient' version of the individual responses. Thus, it has the quality of a 'time derivative' (or, rather, its absolute value), 'signalling' the changes (rise and fall) of the responses, rather than their magnitude. Since the

second response does not reach such high values, the non-monotonicity does not take such strong effect and, correspondingly, the individual response waveforms are more closely reflected in the time course of the interaction (although some of the above described 'transiency' is also present here).

Summarizing, the example in Fig. 3 clearly demonstrates that near-coincidence firing can be strongly modulated by the stimulus presentation. These two neurons are repeatedly switching from a condition of co-incident firing to a stimulus-related period of uncorrelated or even anti-coincident firing. We have observed and reported similar phenomena in other recordings, both from the visual cortex [23] and other cortical areas, such as the prefrontal cortex [24–26]. (A review on experimental results, based on recordings from a variety of regions in the CNS of different animal species, made in several different laboratories, is currently in progress; Gerstein and Aertsen, in preparation.) These findings demonstrate that cortical neurons may exhibit rapid modulations of discharge synchronization that are related to stimulus context and behavioral state. Such modulations – which could not be inferred from single neuron observations – may switch the neurons' firing behavior from being mutually incoherent into a particular coherent state of joint synchrony, or, alternatively, from one particular pattern of mutual coherence into a different one. Each such pattern may last for only a few tens to hundreds of milliseconds. Finally, the observed modulations in synchronous firing may be, but need not be, associated with changes in either of the neurons' individual firing rates. In fact, the correlation of firing may follow a time course that deviates appreciably from that of the firing rate of either of the individual neurons. In general, therefore, rate coherence and spike correlation present two instances of firing synchrony at different levels of temporal acuity. At a coarse level of temporal resolution the correlation of firing is governed by the co-variation of the firing rates ('*rate coherence*'), at a more fine-grained level of time resolution, the detailed spike correlation ('*event coherence*') becomes the dominant term [27]. In any specific case, these different types of correlation may be, but need not be, correlated. These various phenomena appear to be robust across different regions in the brain, and across a variety of animal species. Thus, we conclude that dynamic cooperativity presents an emergent property of neuronal assembly organization in the brain.

5. Functional coupling in model neural networks

These findings in cortical multi-neuron recordings suggest that the usual concept of neurons with static inter-connections of fixed or only slowly changing efficacy (during learning, for example) is no longer appropriate. Instead, one should distinguish between *structural* (or anatomical) *connectivity* on the one hand and *functional coupling* (or effective connectivity) on the other [23,28]. Whereas the former can presumably be described as (quasi)stationary, the latter may be highly dynamic and context-sensitive. These findings raise a two-fold question: what is the nature of the underlying mechanisms, and what are the functional implications? In a number of theoretical studies we have sought to elucidate these issues. Several different mechanisms may be invoked to mediate the transition from static, anatomic connectivity to dynamic, functional coupling. On the one hand, the underlying mechanism may be *local*, as in von der Malsburg's proposal of rapid modulation of synaptic efficacy [29,30]. On the other hand, more *global* network effects might be involved, as for instance in Sejnowski's notion of the 'skeleton filter' [31].

In order to obtain more insight into these different alternatives, we investigated the activity dynamics in various physiology-oriented networks of model neurons, both of the spiking and of the analog type [32–34]. These networks were

designed to capture the typical features of real cortical networks, both in terms of anatomy and of physiology. Of special importance in this context is the physiological constraint of 'sparse' firing, i.e. that the activity in the network is required to be of low firing rate, such as typically observed in the neocortex [35]. As it turned out, maintaining stability in such a feedback system of sparsely firing neurons is by no means a trivial problem [36]. In fact, a stable solutions could only be obtained for very confined 'islands' in parameter space. Outside this restricted range, the network activity either died, exploded or developed strong, coherent oscillations, with frequency and amplitude dynamics governed predominantly by the inhibition parameters, and much less dependent of (1) the network architecture (uniform, random or structured), (2) the spiking or analog nature of the neural activity, and (3) the linear or nonlinear nature of the neural threshold function. A more extensive discussion on this stability issue and on the critical role that inhibition plays in the synchronization dynamics has been provided elsewhere [36]. In the present report we focus on the issues of stimulus-dependence and the dynamics of functional coupling. We took great care to select parameter values such that the network was operating in a regime of stable, sparse firing under non-stimulus conditions (low rate 'spontaneous' activity).

6. A feedback neural network model of spiking neurons

The model we investigated for the present purpose consisted of a feedback neural network of 100 spiking model neurons with fixed synaptic connections [36]. The layout of the network is shown in Fig. 4a, the corresponding equations are given in Fig. 4b. In this model, the neurons are inter-connected by excitatory synapses, inspired by neuroanatomical findings that about 90% of the cortical synapses are of this type [10]. The intra-cortical inhibition was modeled in a different way, as will be described later. The excitatory synapses are modeled as lowpass filters with delayed response, transforming the incoming spike activity into EPSPs. These are summated linearly on the dendritic tree to yield the instantaneous value of the membrane potential at the cell body. The probability of spike generation is modeled by a sigmoid function of this membrane potential; this probability, in turn, is modulated by a refractory mechanism, driven by the recent spike history of the neuron. The final spike output is obtained by using the firing probability as the instantaneous rate of a stochastic event generator.

In addition to this basic network of spiking neurons, the pyramidal cells, the model comprises a global inhibitory mechanism. This reflects our view that one of the principal tasks of intracortical inhibition is to keep the cortical network from "exploding into a condition of global, senseless activity" [10]. This inhibition is implemented in the form of two parallel, linear branches: a fast one, with a time constant in the range of that for the excitatory synapses (5 ms), and a slow one with a considerably larger time constant (200 ms). The fast branch is intended to mimic the action of the inhibitory non-pyramidal neurons, the slow one serves to regulate the overall activity in the network towards a preset value (threshold control [37,38]). In the experiments reported here, parameter values were chosen such that they obey physiological reality, while ensuring stability under sparse firing conditions, with an overall network activity in the order of 10–20 spikes per neuron per second on average (see [36] for more details).

For the present simulations we used a network with a fixed connectivity matrix, derived from a preceding study of associative memory and the performance of different types of learning rules [34,39–41]. In particular, we used a connectivity matrix in which were embedded the memory traces of a set of 10 sparsely occupied, randomly distributed input patterns. Briefly, the 'learning'

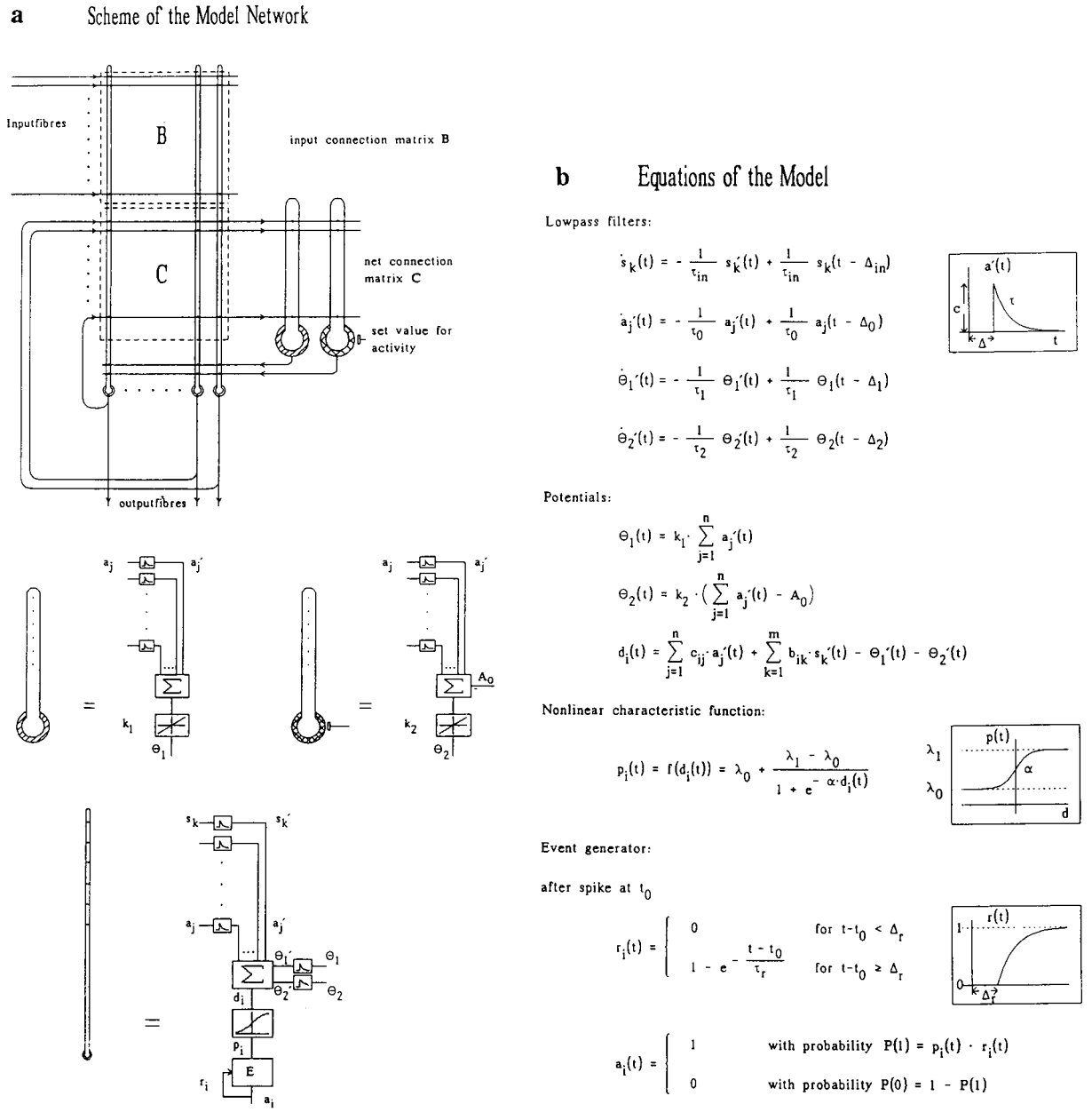


Fig. 4. Scheme (a) and equations (b) of the model neural network. Further explanation in text.

phase proceeded as follows. Initially, the connectivity of the network was given by a random matrix with connection strengths uniformly distributed between 0.25 and 0.35. The 10 input patterns, shown in the bottom panel of Fig. 5a, were presented to the network one by one in the order 1 to 10, each for a duration of 100 ms. Using 'analog' neurons and a covariance learning rule with clipping at connectivities 0 and 1, the initial random matrix gradually evolved during

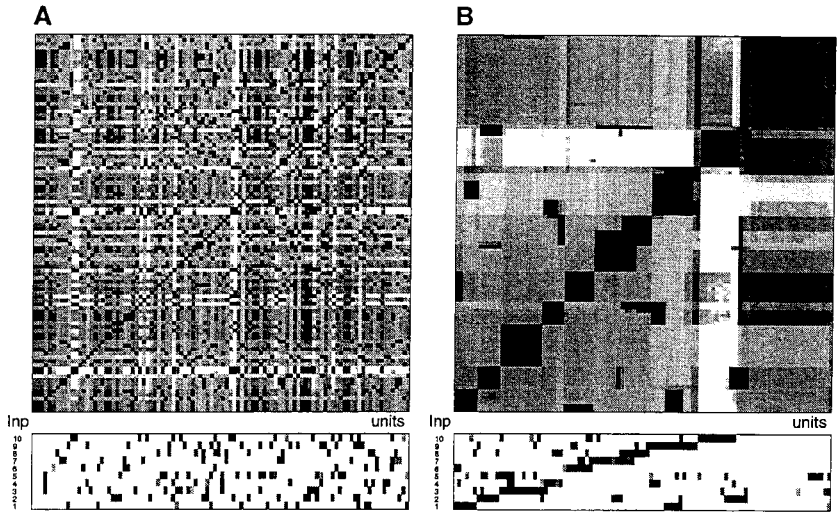


Fig. 5. 'Synaptic' connectivity matrix (top panels) of the model neural network after 'learning' 10 randomly generated input patterns (bottom panels) during 100 ms each. Initially, the connectivity of the 100 neuron network was a random matrix, with connection strength uniformly distributed between 0.25 and 0.35. In order to visualize the functional associations 'hidden' in the input pattern and the associated connectivity matrix (A), we re-sorted the neurons, such that the input patterns were more or less optimally segregated (B). Observe that after this permutation, groups of strongly interconnected model neurons ("cell assemblies") pop out, with (usually weaker) connections between the groups.

this 1 s of learning to the shape shown in the top panel in Fig. 5a. Clearly, the final connectivity has obtained a certain degree of structure, which, however, is hard to interpret in this random ordering. Unlike for the array of cortical recordings in Fig. 1, we now have no natural ordering scheme to guide our 'reading' of the connection matrix. In order to overcome this problem and to obtain a better overview of the functional associations hidden in the matrix, we re-sorted the neurons, such that the input patterns were more or less optimally segregated. In the present case, this sorting was done in two stages. First, we resorted all neurons, such that those 'belonging to' pattern 1 came first, followed by those that belonged to pattern 2 and were not already included in the first selection, then the same for those in pattern 3 and not already included in patterns 1 and 2, and so on for all 10 input patterns. The partial segregation of the connectivity matrix thus obtained was then improved 'by hand', by additional shuffling of individual neurons until the (largest) connection values were reasonably well concentrated in the form of compact blocks lying along the main diagonal of the matrix. The result of this permutation is shown in Fig. 5b, with the rearranged input patterns in the bottom panel, and the rearranged connectivity matrix above it. Notice the clear block-like structure along the diagonal, with each compact block (containing on the average some 7 neurons) denoting a heavily interconnected 'cell assembly', preferentially associated with one of the input patterns. Clearly, the segregation is not complete, as the input patterns are random, not orthogonal. Hence, associations do overlap, as is indicated by the (usually weaker) connections between the groups. Observe also that a fraction of almost 25% of the neurons (the outermost ones in Fig. 5b) is not dedicated to any one of the patterns in particular. They rather provide a diffuse background connectivity to each of the assemblies, while being driven by several of them. On

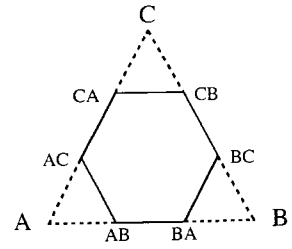
average, each assembly involves some 10 neurons, having an overlap of 2 to 3 neurons with one or more of the other assemblies.

7. Stimulus-dependence of coupling in the model neural network

In the following simulations, we proceeded with the spiking network with the fixed connectivity matrix of Fig. 5 (i.e. learning was terminated), and investigated the flow of activity upon dynamic manipulation of the input stimuli. These 'experiments' were designed to mimic the physiological situation. By adopting this approach, we could apply the same analytical tools as described in the foregoing sections and, thus, directly confront the outcome of the physiological experiments with that of the theoretical investigation. We refer to such experiments on theoretical constructs as '*in virtu*' recordings, to distinguish them from their *in vivo* and *in vitro* physiological counterparts.

In our experiments we continuously varied the stimulus to 'travel' between the stored input patterns. The associated excursions through the attractor landscape of the neural network were monitored by observing the variations in spike activity among a selected subgroup of neurons. The stimulus protocol is schematically illustrated in Fig. 6a. The symbols A, B and C denote any triplet from the stored input patterns; the continuous hexagon symbolizes the path taken by the stimulus. This path was designed to investigate the switching behavior of the network to varying mixtures of stored input patterns. To this end, we concentrated our attention on the boundary regions and avoided the attractors themselves. For instance, at any point in time along the branch AB–BA, the stimulus consists of a superposition of the two input patterns A and B, with weights linearly varying from a ratio 2:1 to 1:2. One such branch has a duration of 100 ms, sampled in steps of 1 ms. Thus, one 'round-trip' along the hexagon takes 600 ms. By

a



b

	A	B	C
SE1	9	4	2
SE2	8	7	10
SE3	4	2	10
SE4	9	8	7
SE5	9	2	4
SE6	8	10	7
SE7	4	10	2
SE8	9	7	8

Fig. 6. Composition of the stimulus ensembles, used to test the model neural network. The symbols A, B and C denote three different input patterns, selected from the stored set of ten (Fig. 5; bottom panels); the hexagon symbolizes the stimulus trajectory (a). Along each side of this hexagon, the stimulus consists of a superposition of two of these input patterns, with weights linearly moving from ratio 2:1 to 1:2. Each branch lasts 100 ms, sampled in steps of 1 ms; a complete 'round-trip' along the hexagon takes 600 ms. Four different choices of the triplet (A, B, C) lead to four different stimuli, each of which comes in two versions, depending on whether the hexagon is travelled 'clockwise' (SE1 to SE4) or 'counter-clockwise' (SE5 to SE8) (b).

making different selections for the triplet (A,B,C), we generated four different stimulus ensembles; by additional variation of the direction of the path taken ('clockwise' or 'counter-clockwise'), we arrived at a total of eight different stimulus ensembles (SE1 to SE4, and SE5 to SE8; cf. Fig. 6b).

The response of the network to these various stimuli was investigated by recording the spike activity of a subgroup of 16 of the 100 model neurons, in much the same way a physiological multi-neuron recording would pick up a fraction of the entire population of neurons. These 16 neurons were chosen such that each of them belonged to one or two assemblies, associated with the stored input patterns used for generating the stimulus ensembles SE1 to SE8. The 'synaptic' connectivity matrix for this subset of 16 neurons, a submatrix from the one in Fig. 5, is shown at the bottom left of Fig. 8. Notice again the block-like associations of the various cell assemblies and the overlap between them. This is summarized in the simplified membership diagram depicting three dominant assemblies

(numbered 1, 2, 3) within the group of 16. Note, however, that the actual assembly structure is more complex, and that more than three, partly overlapping, assemblies can be discerned in the synaptic connectivity matrix.

Fig. 7 shows the temporal variation of the input to the selected 16 neurons for two different stimuli (SE1 and SE3; top panels) and the associated spike activities, collected over 10 trials of 600 ms each (bottom panels). Notice the qualitative similarity of this picture with the physiological recording (Fig. 2) and, particularly, how the various neurons reflect the time course of their input in the temporal modulation of their (generally low) firing rates. We made simulations like these for each of the eight different stimulus ensembles SE1 to SE8. The multi-neuron spike trains 'recorded' during these simulations were analyzed for neural interactions using the network correlation matrix algorithm described in connection to Fig. 1. The 'normalized' correlation matrices, representing the time-averaged functional coupling among the neurons for each of these 8 different stimuli, are shown in the top two rows of Fig. 8, in a format identical to that of Fig. 1b.

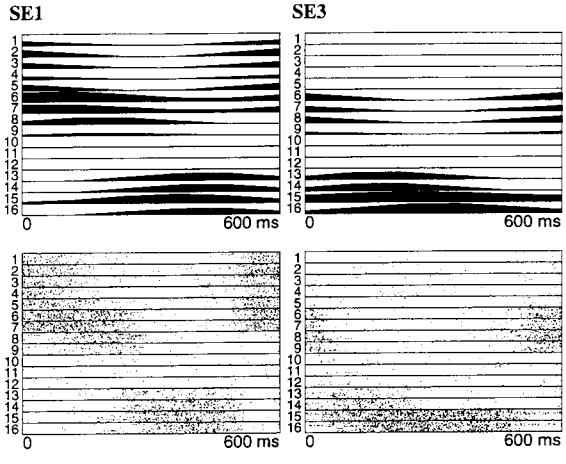


Fig. 7. '*In virtu*' recordings from 16 out of 100 neurons in the model network upon presentation of two different test stimuli (SE1 and SE3; Fig. 6). The top panels show the inputs to each of these 16 neurons, the bottom panels show the spike trains these neurons generate in response to 10 consecutive stimulus presentations. Compare these dot displays to the cortical recordings in Fig. 2.

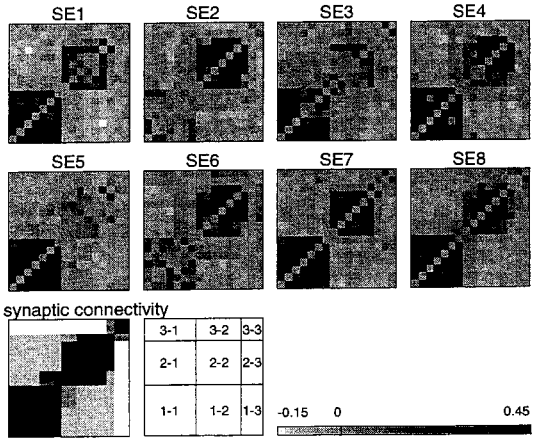


Fig. 8. Network correlation matrices (top eight panels), representing the functional coupling among a subset of 16 neurons in the model network in response to eight different test stimuli (SE1 to SE8; Fig. 6). All stimuli were presented at equal input strength of 0.25 (cf. Fig. 9). The synaptic connectivity matrix for these 16 neurons (a sub-matrix from the one in Fig. 5) is shown in the bottom panel (left), together with a simplified diagram of the three dominant cell assemblies (1, 2, 3), present in this group of 16 neurons. The format of the matrices is the same as in Figs. 1 and 5; the strength of coupling is coded in grey according to the grey bar (bottom right).

Two interesting comparisons can be made at this point. First, the network correlation matrices for these eight different stimulus ensembles reveal distinct stimulus-dependent differences in the degree of coupling among the neurons. Different subgroups of neurons more or less clearly pop out, depending on the stimulus used. These results are highly reminiscent of the stimulus-dependent coupling we observed in the physiological recordings in Fig. 1. In addition, however, we can also compare the correlation matrices with the synaptic connectivity matrix in the bottom panel. Note that this second comparison is usually not possible in the physiological context and, thus, presents a distinct advantage of such '*in virtu*' experiments. In particular, it demonstrates how, depending on the stimulus context, different portions of the underlying synaptic network express themselves

in the correlation matrix, whereas others, in spite of their strong and persistent synaptic connectivity, do not show up at all. For instance, the stimulus ensemble SE3 mainly reveals coupling among the first assembly (neurons 1 to 7), stimulus ensemble SE2 predominantly shows interactions in the second group (neurons 8 to 13), whereas both the first and second assembly (or, rather, different portions of the latter) are manifest with stimulus ensembles SE1 and SE4. None of these four stimuli is apparently able to reveal the third assembly. Observe also how the time-reversed stimulus ensembles SE5 and SE7 give rise to distinct differences as compared to their counterparts SE1 and SE3; such hysteresis effects are less visible for the stimuli SE6 and SE8. Finally, we note an interesting case of 'pattern completion', revealed by comparison of Figs. 7 and 8. Inspection of Fig. 7 shows that neurons 1 to 5 receive no direct input from stimulus SE3, whereas neurons 6 and 7 do. Accordingly, these first five neurons fire at very low rate, whereas neurons 6 and 7 more or less faithfully reflect the time course of the input in the elevation of their firing rates (cf. the dot display in Fig. 7). In spite of this difference in activation, however, the *entire* assembly lights up more or less uniformly coupled in the network correlation matrix in Fig. 8. Apparently, the synaptic interactions within cell group 1 suffice to recruit these first five neurons into coincident firing with the other members of the group, even though only part of the assembly was activated by the stimulus. Moreover, the pattern completion in this case is expressed in an increased correlation of firing, rather than in an increased rate of firing.

8. Dependence of coupling on stimulus strength

In a second series of simulations we investigated the statistical significance of individual values in the network correlation matrix. To this end we presented a stimulus ensemble repeatedly, and collected data from 10 such experiments, each one consisting of a sequence of 10 trials as before. For each of these experiments we measured the normalized correlation matrix and, thus, obtained 10 statistically independent realizations of the same coupling matrix. The degree of statistical variation was assessed for each individual matrix entry by determining the mean value and the standard deviation. By averaging these standard deviations for the entire matrix, we obtained the overall mean standard deviation. Assuming a Gaussian distribution, we defined a 5% significance level for deviations from zero coupling by setting a threshold at twice this mean standard deviation. This statistical analysis was performed for two different stimulus ensembles SE1 and SE2. In the same series of experiments we also studied the effect of changing the amplitude of stimuli presented to the network. The implicit assumption here is, of course, that such changes of input strength are monotonically related to changes in the intensity (or contrast, or some other appropriate parameter) of sensory stimuli in physiological experiments.

Fig. 9 presents a summary of our findings. For each of the stimuli SE1 and SE2, a triplet of panels in the top row shows superimposed plots of the mean values of selected matrix entries as a function of input strength. The left-hand panel of each triplet (Figs. 9a resp. 9d) contains the curves for all matrix entries within a single plot; the two pairs of right-hand panels (Figs. 9b,c, respectively 9e,f) consider only those matrix entries that correspond to coupling within each of the two dominant cell assemblies 1 and 2, i.e. those situated within the confines of the matrix blocks (1–1 and 2–2) around the diagonal in the bottom panel of Fig. 8. We found the mean standard deviation to be 0.04, only weakly dependent on the input strength and of the stimulus ensemble used. Hence, we set the 5% significance level for coupling at a constant correlation threshold of 0.08, indicated by the dashed line in the various plots. Several observations can be made from these results. We concen-

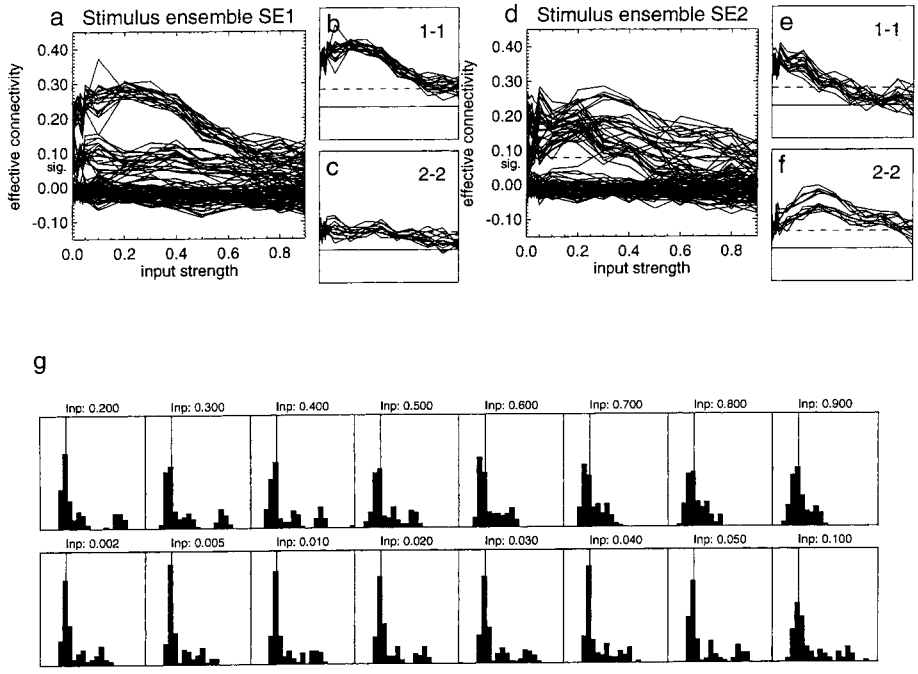


Fig. 9. Dependence of the functional coupling in the model neural network on stimulus composition and intensity. The top panels show the values of individual entries in the network correlation matrix as a function of the strength of the input stimuli SE1 (a–c) and SE2 (d–f). Each curve represents the mean of 10 statistically independent repetitions of the experiment. Panels (a) and (d) show the superimposed curves for all neuron pairs in the correlation matrix, panels (b,c) and (e,f) only for these pairs that belong to the two dominant assemblies (cell groups 1 and 2), as judged from the synaptic connectivity matrix (Fig. 8). The dashed line (sig) marks the 5% level of significance for deviation from zero coupling. The histograms in the bottom panels (g) show the distributions of values in the network correlation matrix for different values of input strength of the stimulus SE1 (scaling, as in (a–f) from −0.15 to +0.45; binwidth 0.025).

trate first on the separate sets of curves for the coupling within each of the assemblies 1 and 2. The narrow spread among the curves in each single panel (Figs. 9b,c and 9e,f) reveals that the coupling between the various members of an assembly is essentially uniform, and shows identical dependence on input strength. Generally, this dependence is non-monotonic, with a maximum at some intermediate value of stimulus intensity, and consistent fall-off at larger amplitudes. This suggests that there exists a restricted range of intermediate amplitude values in which the assembly organization can unfold most clearly (our examples of Fig. 8 were measured at an input strength of 0.25). Very weak or very strong input strengths, however, either do not adequately address the network or enslave it to such an extent that no room is left for the assemblies to organize. Apart from these common properties of the various sets of curves, there are also clear differences in the amplitudes and the shapes. This applies both when different assemblies are compared for the same stimulus (Fig. 9b vs. 9c; Fig. 9e vs. 9f), and for each single assembly compared across different stimuli (Fig. 9b vs. 9e; Fig. 9c vs. 9f). This differential dependence on the composition as well as on the strength of the input stimuli once more demonstrates how the network may drastically change its appearance, depending on the stimuli used to investigate its performance. Stated in more functional terms, it explains how the net-

work reorganizes the topology of its functional coupling, depending on the computational task it is faced with.

9. On 'reading' the network correlation matrix

Our interpretation of these differential dependences of spike correlations on stimulus properties was certainly facilitated by our knowledge of the synaptic connectivity matrix. Particularly, the ability to assign the individual neurons to different assemblies and to investigate their coupling in appropriately separated groups proved most helpful. Unfortunately, in the physiological setting such information is usually not available. The result is that a separation as in Figs. 9b,c resp. Figs. 9e,f is generally not possible. Rather, the experimenter is faced with the task to make sense out of entire collections of curves, such as shown in Figs. 9a and 9d. Clearly, the variance as well as the overlap of so many different shapes of curves make plots like these very hard to interpret, even more so as sampling at so many different intensity levels as done here is experimentally usually not feasible.

We see essentially two approaches to alleviate this dilemma. First, as illustrated in the permutation of the connectivity matrix in Fig. 5, considerable insight can be gained by taking into account the 'functional neighborhood' of the neurons and to use this to reorganize the 'spatial' arrangement of the coupling matrix. Such reorganization may provide useful hints how to subdivide the observed neurons into separate functional groups. We are currently investigating such re-sorting algorithms and their performance when applied to network correlation matrices of physiological multi-neuron recordings. A second approach is illustrated in Fig. 9g. This shows, in the form of a series of histograms, the distribution of values in the entire network correlation matrix, determined at different input strengths of the stimulus ensemble SE1. By inspecting these distributions for possible clustering of correlation values, one may obtain further evidence for functional subgroups in the population of observed neurons. Notice, for instance, that the collection of correlation values that significantly deviate from zero comes in two or three separate clusters for certain values of the stimulus amplitude (e.g. 0.03, 0.05), whereas for other amplitudes this distribution is more or less continuous. Thus, it appears that the combination of these two sources of evidence, one from the topology of the coupling matrix, the other from the distribution of coupling strengths, may indeed help in defining the functional grouping of the neurons and their individual alliances.

Summarizing, these simulations demonstrate that a network with fixed synaptic connectivity can exhibit strong context-dependence of functional coupling, expressed in the neuronal spike correlations. Thus, considerable and rapid changes in coupling may occur, without any associated changes in local connectivity. In contrast, our findings rather point at a more global mechanism, in which the instantaneous degree of coupling among the neurons is controlled by the stimulus-dependent flow of activity in the entire feedback network. Such global mechanisms do not have to invoke rapid modulations of synaptic efficacy [29,30], although they obviously do not exclude them as an additional mechanism either. Further experimental evidence is required to clear this issue.

10. Dynamics of functional coupling in the model neural network

Similarly to our approach in the physiological recordings, we now turn our attention to the dynamics of coupling between individual neurons in the model network using the Joint-PSTH. We will consider the interactions both within and across different cell assemblies. In order to obtain reliable statistics, we presented the stimulus ensemble SE1 to the network for a total of 250 consecutive trials. Fig. 10 shows the

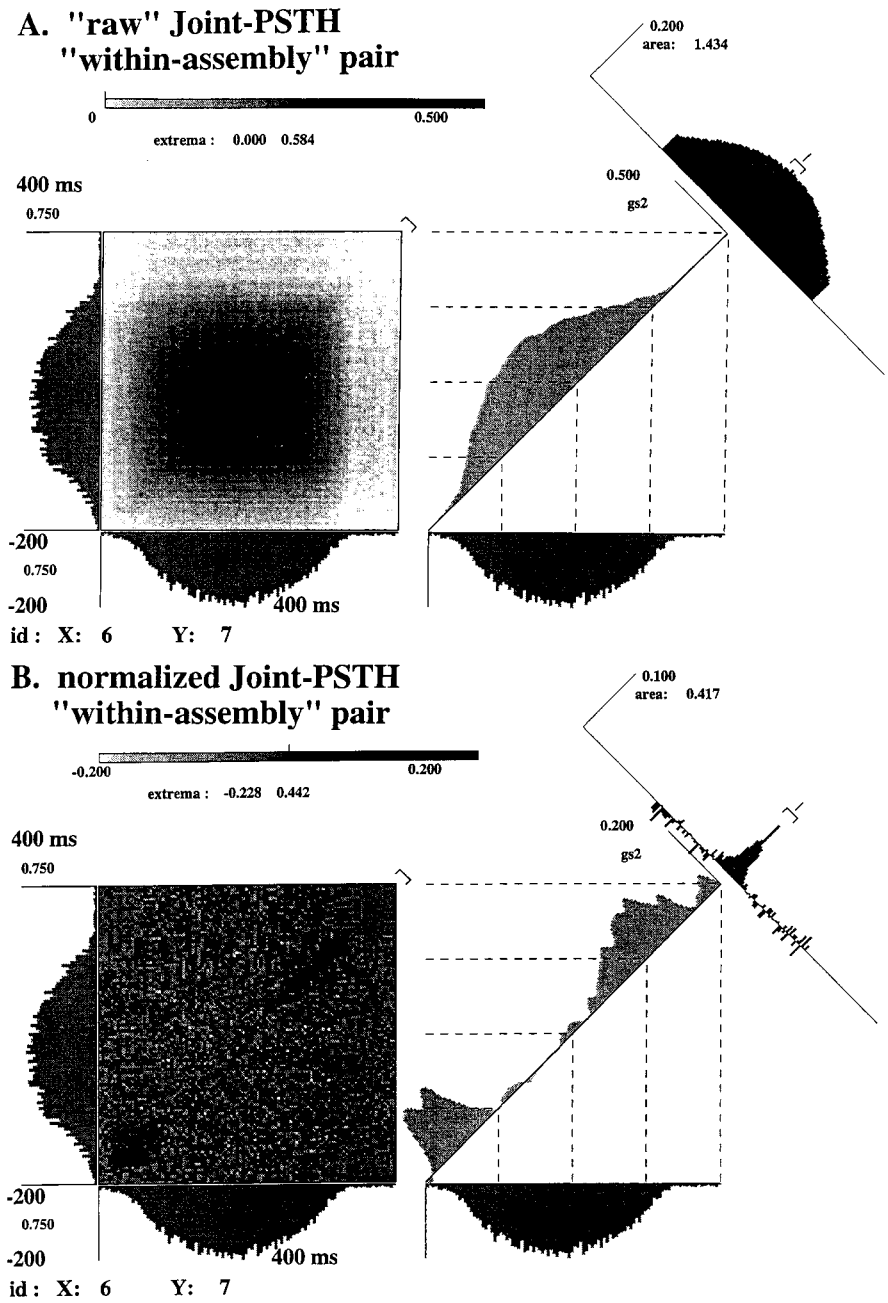


Fig. 10. Stimulus-locked dynamic correlation of firing between selected pairs of neurons from the model network during repeated presentation of the test stimulus SE1. The 'raw' (A) and the normalized (B) Joint-PSTH were determined for the pair (6,7), both members of the same assembly (cell group 1; Fig. 8). A similar result is obtained in the normalized Joint-PSTH (C) of another 'within-assembly' pair (8,9), selected from the assembly (2). Observe that quite different coupling dynamics are exhibited by the 'across-assembly' pair (6,8) (D). The format of these figures is the same as in Fig. 3, with a binwidth of 6 ms. Numbers of spikes: 9244 (neuron 6), 8426 (neuron 7), 4136 (neuron 8), 2390 (neuron 9), recorded during 250 stimulus trials of 600 ms each.

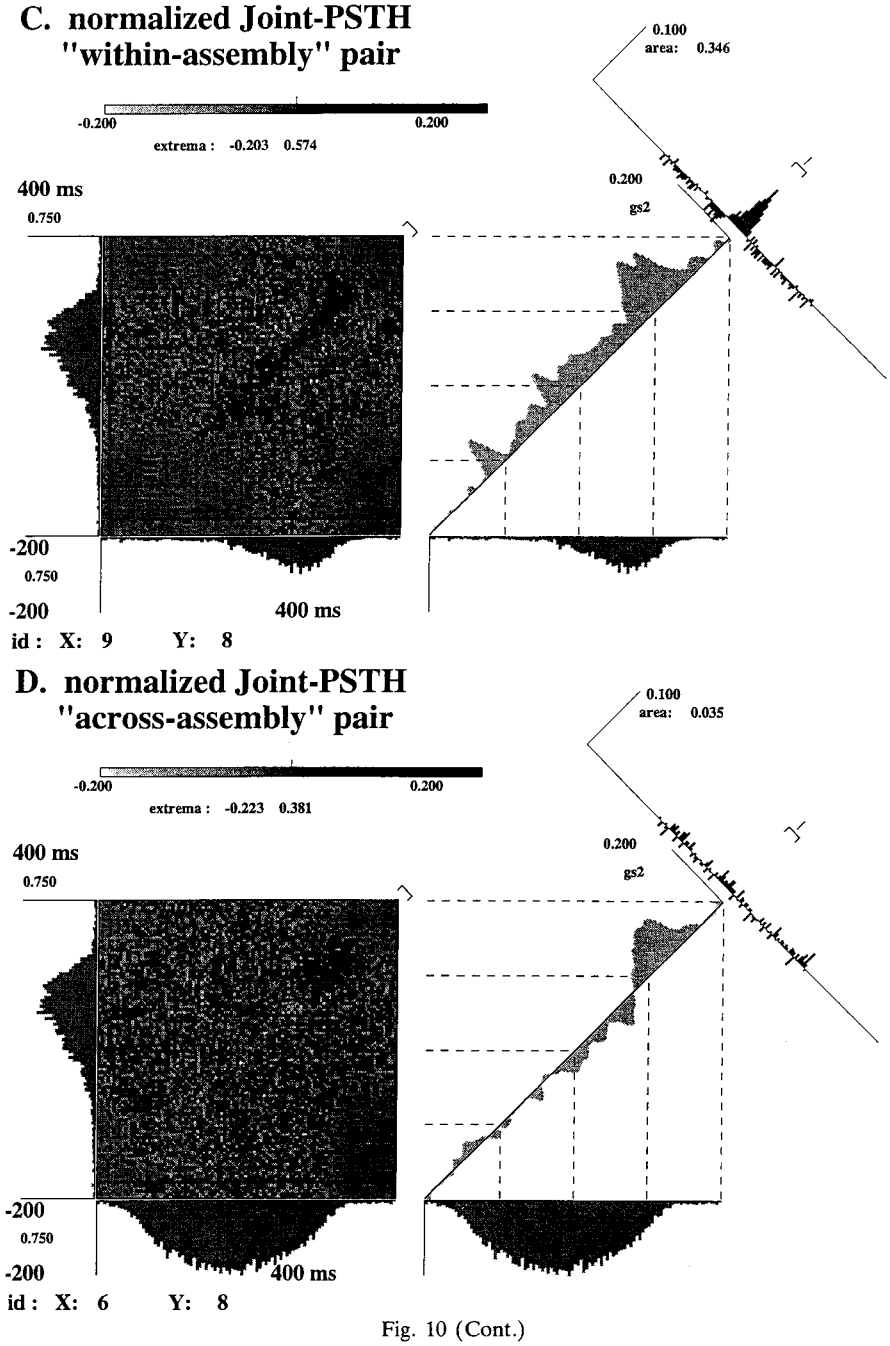


Fig. 10 (Cont.)

results of Joint-PSTH analysis of selected pairs of model neurons. The raw and the normalized Joint-PSTH for the pair (6, 7), both from cell assembly 1, are shown in Figs. 10a and 10b. Further normalized Joint-PSTHs are shown for the pair (8, 9) from cell group 2 in Fig. 10c, and, finally, for the across-group pair (6,8) in Fig. 10d. All three normalized Joint-PSTHs are displayed using the same scales, in order to enable comparison across neuron pairs.

Similarly to Fig. 3a, we observe that the individual PST histograms in the 'raw' Joint-PSTH (Fig. 10a) show a strong stimulus-locked modulation of firing for both neurons, and an accompanying increase in 'raw' correlation in the Joint-PSTH matrix and in the PST-coincidence histogram. Again, individual, joint, and near-coincidence firings all are increased in parallel during a portion of the stimulus trial. Also, as in Fig. 3a, the time-averaged cross-correlogram (top right) is dominated by the rate coherence, thereby virtually masking the signs of spike correlation. Inspection of the normalized Joint-PSTH, however, demonstrates that also here the normalization has an 'unmasking' effect, as is attested by the – hitherto invisible – symmetrical and moderately narrow positive peak in the time-averaged cross-correlogram. More interestingly, though, also the detailed time course of the excess spike correlation between the two model neurons is very similar to that between the cortical neurons. As in Fig. 3b, the normalized PST-coincidence histogram along the diagonal of Fig. 10b reveals positive correlation in the rising and falling phases of the responses, while the correlation breaks down to zero inbetween, and stays down for most of the duration of the peak response rates.

This behaviour is essentially reproduced by the pair (8, 9), taken from cell group 2 (Fig. 10c). Also these two model neurons co-vary in their response rates, albeit with a smaller dynamic range than the pair (6, 7). Similarly to the first pair, the normalized correlation reaches its maximum not during the peak response rates, but rather in the declining phase and, albeit to a smaller extent, in the rising phase. Comparison with further pairs showed that the behavior exhibited by these two pairs, each taken from a different cell group, is representative for the interactions taking place within a model assembly. The situation is quite different, however, for the interactions between model neurons that belong to different assemblies. This is illustrated for the pair (6, 8) in Fig. 10d. Observe that in this case the time-averaged normalized correlation (top right) equals zero. This explains the zero coupling among such across-assembly neurons in the network correlation matrix (Fig. 8, SE1), as this matrix essentially represents a time-averaged measure of excess correlation. Interestingly, this absence of time-averaged correlation does not imply that the coupling between the neurons is constantly zero throughout the stimulus trial. Instead, the normalized PST-coincidence histogram along the diagonal exhibits an alternating trajectory. It passes through a prolonged, shallow negativity during most of the elevated single neuron responses, before rising to a brief burst of positive spike correlation as the responses have almost decayed down to background level. This result once more demonstrates the dangers of taking the usual, time-averaged cross-correlogram as evidence for the presence or absence of neural interactions. Contributions of opposite sign may cancel in the time integral and, hence, leave no trace in the correlogram (see also [23]). More interestingly, though, are the functional implications of this type of interaction. Apparently the two model assemblies compete with each other during most of the time they are both strongly activated by the stimulus. This competition is presumably a consequence of the mean field coupling, mediated by the global inhibition in our model network. It only ceases to exist, or, rather, it is overridden by the positive cross-coupling between the assemblies, after the firing rates have decayed so much that the simultaneous co-existence of two active assemblies no longer poses a challenge to the activity regulation by the slow inhibition.

In summary, many of the features we encountered in our study of interactions among cortical neurons are reproduced in our simulations in remarkable detail. This is all the more surprising, since our model network was not designed with that particular purpose in mind. Rather, it was inspired by much more general considerations, originating from statistical neuro-

anatomy and basic physiology. Thus, we conclude that the observed consistency with the physiological interactions clearly speaks in favor of the biological plausibility of this type of neural network model. Moreover, it lends support to our thesis that most, if not all, of the dynamic context-dependence of cortical coupling can, in fact, be attributed to the global mechanisms embedded in such probabilistic feedback networks. In this context, it is interesting to note that remarkably similar phenomena of clustering and dynamic reorganization have been reported in other coupled nonlinear dynamical systems (see e.g. the contributions by Tsuda and Kaneko in these proceedings).

11. Modulation of functional coupling by background activity

In a final chapter we wish to quantify the influence of the overall network activity on the effective strength of an individual synapse. To this end we performed additional, more specifically designed simulations and analytical calculations [42]. The model system consisted of two neurons, with neuron 1 driving neuron 2 via a single synapse of fixed strength α (see Fig. 11a, inset). In addition, neuron 2 receives input from a pool of N independent, spontaneously firing neurons, each of them connected to 2 with strength β. All synaptic connections α and β have fixed and moderately weak strength, i.e. spikes arriving at any particular junction give rise to EPSPs with a constant, subthreshold magnitude. We studied the behavior of the efficacy of the connection between neurons 1 and 2 as a function of the background activity arriving at 2 from the neuron pool, for constant connectivity α and different values of pool coupling β.

This model system was investigated using equations and simulation algorithms adopted from [43]. The mean firing rate λ of all spontaneously active neurons was set to about 10 spikes/s ('sparse' firing). The pool activity could be varied by changing either the number of neurons N or the firing rate λ of each of the member neurons. The latter manipulation mimics the effect of driving the neuron pool with a stimulus S (Fig. 11a, inset). The effective coupling α' between neurons 1 and 2 was measured by cross-correlating the spike activity from the two: α' equals the number of correlated events N_c (i.e. the net area of the peak in the correlogram) divided by the number of presynaptic events N_1 [12]. Curves of the behavior of the effective connectivity α' as a function of pool activity $N\lambda$ for various choices of pool coupling β

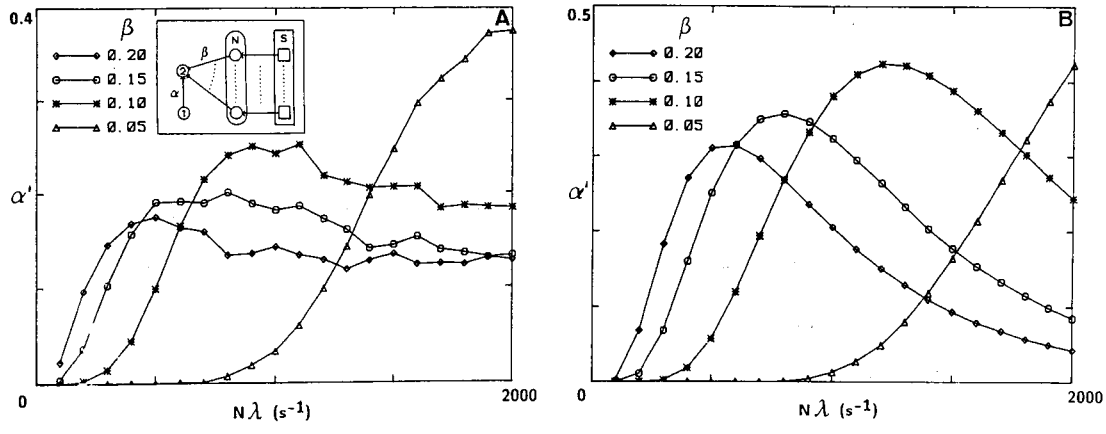


Fig. 11. Functional coupling α' as a function of background pool activity $N\lambda$, for different values of pool coupling β. The two panels show results from network simulations (A) and from analytic calculations for a simplified version of the network model (B). The structure of the network is shown in the inset of (A). Further explanation in text.

are shown in Fig. 11a. Two observations can be made:

(1) The efficacy α' varies in a non-monotonic fashion with pool activity $N\lambda$, even though the synapse itself is kept at a fixed strength α (=0.2) throughout all simulations. With increasing pool activity, the efficacy of the coupling initially increases strongly to reach a maximum, after which it slowly decays again. Dynamic changes of pool activity, reflected in excursions along the horizontal axis of Fig. 11a, thus would result in equally fast, but distorted excursions along the vertical axis, causing corresponding changes in effective coupling.

(2) The dynamic range of the efficacy α' increases with decreasing pool coupling β. Moreover, for smaller β the maximum of the effective coupling α' is actually higher, but is only reached at larger values of the pool activity. This finding is especially relevant for large, weakly coupled nets, of which the mammalian neo-cortex is a typical example.

The explanation of these phenomena is actually quite simple. Due to the subthreshold nature of the EPSPs, action potentials from neuron 1 by themselves will not suffice to make neuron 2 fire. For this they need the coincident arrival of spikes at the other synapses of neuron 2. These are provided by the diffuse background activity reaching 2 from the remainder of the network. The pool activity thus provides a background level, which, depending on its magnitude, will make activity from the driver neuron 1 more (or less) viable in eliciting activity from its target neuron 2. Hence, the background activity functions as a control parameter, determining the operating point of the receiver neuron, and, thereby, the efficacy of the otherwise subthreshold connections. This also explains our second observation, i.e. why the dynamic range increases with decreasing pool coupling. Due to the stochastic and pulse-like nature of the pool activity arriving at the target neuron, the background level of the membrane potential of 2 will fluctuate and, hence, occasionally reach threshold, causing neuron 2 to fire. Since these 'accidental' firings do not require coincident arrival of spikes from neuron 1, they do not contribute to the efficacy of the 1–2 connection. On the contrary, they are counterproductive since they cause spikes from 1 to arrive more often at its target neuron when this is in a non-responsive state ('refractoriness'). The size of the membrane potential fluctuations and, hence, the amount of occasional firing of 2 is proportional to the pool coupling β. This implies that the smaller this pool coupling, the closer one can regulate the operating point of the target neuron towards being liable to fire upon arrival of spikes from the driver neuron, without the danger of eliciting too many counteractive accidental firings. Rather than a 'noisy' control parameter, one thus obtains a smooth, and thereby more influential regulation of synaptic efficacy.

In order to investigate this background firing mechanism quantitatively, we analyzed a simplified model by analytic means [42,28]. An explicit expression for the effective coupling can be obtained by assuming (1) that the membrane potential at the site of spike initiation is the result of linear, spatio-temporal integration of incoming spike trains, and (2) that a spike is generated whenever this membrane potential crosses a fixed threshold level θ. The contribution to the membrane potential of neuron 2, due to the pool of N independent Poisson processes, each with constant rate λ, connected to neuron 2 with strength β can then be described as 'shot-noise' (i.e. a linearly filtered (time constant τ) Poisson process). For reasonably large N, the resulting distribution of membrane potential values can be approximated by a Gaussian distribution with mean value $\mu = N\lambda\beta$ and variance $\sigma^2 = N\lambda\beta^2/2\tau$. Following Abeles [5], the firing λ_2 of neuron 2 can then be calculated as

$$\lambda_2 = \frac{1}{\sqrt{2\pi}} \int_{(\theta-\mu)/\sigma}^{\infty} \exp(-\tfrac{1}{2}y^2)\,dy\,.$$

The effective coupling α' from neuron 1 to

neuron 2 equals the increase in firing rate λ_2 upon the arrival of a spike from neuron 1, i.e. upon a shift of the membrane potential over an amount α:

$$\alpha' = \frac{1}{\sqrt{2\pi}} \int_{(\theta-\mu-\alpha)/\sigma}^{(\theta-\mu)/\sigma} \exp(-\tfrac{1}{2}y^2)\, dy .$$

The resulting curves for the effective coupling α' are shown in Fig. 11b. Notice that they exhibit essentially the same non-monotonic behavior as the simulation results in Fig. 11a: an initial fast increase until a maximum is reached, followed by a slower decay. Also, the influence of the strength of pool coupling on the dynamic range of the effective coupling is reproduced, with smaller pool coupling leading to a larger modulation range. Finally, we note that the shape of the curves in Fig. 11 bears a surprising resemblance to the non-monotonicities we measured in the amplitude-dependence of coupling in the feedback model network (Fig. 9). Also, some of the non-monotonic behavior we observed in the dynamics of coupling in the Joint-PSTHs, both for the physiological (Fig. 3) and for the model spike trains (Fig. 10), fits these curves remarkably well. This, once more, suggests that rapid variations of background activity may indeed be instrumental in dynamically regulating the efficiency of synaptic coupling between cortical neurons.

The principal cause of these various dynamic effects resides in modulations of the rate of background activity projecting onto the target neurons. In another, related study it could be shown that very similar and equally rapid effects can be obtained by dynamically manipulating the internal correlation structure of the background activity, while keeping its magnitude constant [4,44,45]. We conclude that, even with the highly simplifying assumptions in this model, we have captured the essence of a nonlinear mechanism by which 'diffuse' background activity may govern the dynamic linking of neurons into functional groups. It is based on activity variations in the entire network, being 'projected down' onto the connections among the recipient neurons. As a result, rapid changes in activity in the entire network induce equally fast, nonlinearly distorted modulations of synaptic efficacy (see also [46,47]). These, in turn, are expressed in rapid modulations of spike correlation in the cortical activity, as can be observed in multi-neuron recordings.

12. Discussion

Whatever the mechanism underlying the stimulus-driven and behavior-related modulations of functional coupling ultimately turns out to be, these modulations form an interesting feature of cortical organization. In particular, they are the signature of an ongoing process of dynamical and activity-related 'linking' and 'unlinking' of neurons into varying, coherent groups. This process may have interesting functional implications at different levels of observation. At the single neuron level, it may explain how even little specificity in anatomical connections could be dynamically sorted out to yield the complex functional properties that have been observed for cortical neurons. Thus, it might provide a natural mechanism for the physiologically measured context-dependence and intrinsic dynamics of receptive fields in central sensory neurons [48–51]. At the multi-neuron level, dynamic coupling may account for coherence variations in a spatially distributed neural code. Several recently observed phenomena in cortical activity point at possible candidates for such a distributed code. One example is the observation of stimulus-specific oscillatory events in the cat visual cortex, with coherence properties that may extend over wide ranges of cortex [52–54]. A second is the relative exuberance of highly accurate and behavior-related spatio-temporal spike patterns in cortical activity, pointing at the presence of 'synfire reverberations' [55,56]. Dynamic coupling is, almost by definition, a natural

candidate to mediate the general process of 'temporal binding' [29,30]; an example is the task of object recognition by the visual nervous system [57–59]. Recent model work along these lines, e.g. [60–65], demonstrated that stimulus-related modulations of activity coherence, possibly differentiated into rate coherence and event coherence [27], may indeed subserve such binding functions (see also [9]). Finally, at the level of the organization of perception and action, modulation of functional coupling in inter-connected neural networks may provide a mechanism for the selection and successive ignition of neural assemblies within and across such networks. Spatio-temporal variation of input activity, carried onto the target networks by divergent–convergent projections, could modulate the activity levels in these networks and, hence, provide the means to select and dynamically switch from activation of one cell assembly to the next. Such 'threshold control'-like [37,38] mechanisms for the generation of 'phase sequences' [2] of cell assemblies have been invoked in recent theories of learning on the basis of 'effects of actions', such as presumably mediated by cortico-striatal interactions [66–68].

Summarizing, the highly dynamic interplay of activity and connectivity in the cortex gives rise to an ongoing process of functional reorganization. Everchanging groups of neurons, each one recruited for brief periods of time, become co-activated and again de-activated, following each other in rapid succession. It is our conjecture that this dynamic reorganization provides the substrate to implement the neural computations involved in 'higher brain function', including the capacity to perceive, to behave, and to learn.

Acknowledgements

This paper summarizes results from a number of ongoing collaborations. Contributions by Moshe Abeles, Karl-Heinz Boven, Valentino Braitenberg, George Gerstein, Jürgen Krüger, Israel Nelken, Almut Schüz, Eilon Vaadia and their colleagues are gratefully acknowledged. Multi-neuron spike train data were kindly made available by Jürgen Krüger. We thank Ichiro Tsuda and Kuni Kaneko for helpful discussions, and Martin Arndt for critically reading an earlier draft of the manuscript. This research was supported by the Max-Planck-Institute for Biological Cybernetics (Tübingen), the Neurobiology Program of the "Bundesministerium für Forschung und Technologie" (BMFT), the German–Israeli Foundation for Research and Development (GIF), the Volkswagen Foundation, and the Neural Network Program of the "Deutsche Forschungsgemeinschaft" (DFG).

References

[1] C. Sherrington, Man on his Nature. The Gifford Lectures, Edinburgh 1937–38 (Cambridge Univ. Press, Cambridge, 1941).
[2] D. Hebb, The Organization of Behavior. A Neuropsychological Theory (Wiley, New York, 1949).
[3] W. James, Psychology (Briefer Course) (Holt, New York, 1890).
[4] G.L. Gerstein, P. Bedenbaugh and A.M.H.J. Aertsen, IEEE Trans. Biomed. Engineer. 36 (1989) 4.
[5] M. Abeles, Local Cortical Circuits, An Electrophysiological Study (Springer, Berlin, 1982).
[6] M. Abeles, Corticonics. Neural Circuits in the Cerebral Cortex (Cambridge Univ. Press, Cambridge, 1991).
[7] D.H. Perkel and T.H. Bullock, Neural Coding, Neurosci. Res. Prog. Bull. Vol. 6 (1968).
[8] G.L. Gerstein and A. Michalski, in: Adv. Physiol. Sci. Vol. 30, Neural communication and control, Gy. Szekely, E. Labos and S. Damjanovich, eds. (Pergamon Press, Akademiai Kiado, Budapest, 1981) p. 93.
[9] P. Johannesma, A. Aertsen, H. van den Boogaard, J. Eggermont and W. Epping, in: Brain Theory, G. Palm and A. Aertsen, eds. (Springer, Berlin, 1986) p. 25.
[10] V. Braitenberg and A. Schüz, Anatomy of the Cortex Statistics and Geometry (Springer, Berlin, 1991).
[11] D.H. Perkel, G.L. Gerstein and G.P. Moore, Biophys. J. 7 (1967) 419.
[12] A.M.H.J. Aertsen and G.L. Gerstein, Brain Res. 340 (1985) 341.
[13] G.P. Moore, D.H. Perkel and J.P. Segundo, Ann. Rev. Physiol. 28 (1966) 493.
[14] G. Gerstein, D. Perkel and J. Dayhoff, J. Neurosci. 5 (1985) 881.

[15] G. Gerstein and A. Aertsen, J. Neurophysiol. 54 (1985) 1513.
[16] A. Aertsen, T. Bonhoeffer and J. Krüger, in: Physics of Cognitive Processes, E.R. Caianiello ed. (World Scientific, Singapore, 1987) p. 1.
[17] G. Palm, A.M.H.J. Aertsen and G.L. Gerstein, Biol. Cybern. 59 (1988) 1.
[18] A.M.H.J. Aertsen, G.L. Gerstein, M.K. Habib and G. Palm, J. Neurophysiol. 61 (1989) 900.
[19] J.E. Dayhoff and G.L. Gerstein, J. Neurophysiol. 49 (1983) 1334, 1349.
[20] M. Abeles and G.L. Gerstein, J. Neurophysiol. 60 (1988) 909.
[21] S. Grün, A. Aertsen, M. Abeles and G. Gerstein, in: Gen – Gehirn – Verhalten, N. Elsner and M. Heisenberg, eds. (Thieme, Stuttgart, 1993) p. 470.
[22] J. Krüger, J. Neurosci. Meth. 6 (1982) 347.
[23] A.M.H.J. Aertsen and G.L. Gerstein, in: Neuronal Cooperativity, J. Krüger, ed. (Springer, Berlin, 1991) p. 52.
[24] A. Aertsen, E. Vaadia, M. Abeles, E. Ahissar, H. Bergman, B. Karmon, Y. Lavner, E. Margalit, I. Nelken and S. Rotter, J. Hirnforsch. 32 (1991) 735.
[25] E. Vaadia, E. Ahissar, H. Bergman and Y. Lavner, in: Neuronal Cooperativity, J. Krüger, ed. (Springer, Berlin, 1991) p. 249.
[26] E. Vaadia and A. Aertsen, in: Information Processing in the Cortex: Experiments and theory, A. Aertsen and V. Braitenberg eds. (Springer, Berlin, 1992) p. 81.
[27] H. Neven and A. Aertsen, Biol. Cybern. 67 (1992) 309.
[28] A. Aertsen and H. Preissl, in: Nonlinear Dynamics and Neuronal Networks, H. Schuster ed. (VCH Verlag, Weinheim, 1991) p. 281.
[29] C. von der Malsburg, The correlation theory of brain function (Internal Report 81-2, Max-Planck-Institute for Biophysical Chemistry, Göttingen, 1981).
[30] C. von der Malsburg, in: Brain Theory, G. Palm and A. Aertsen, eds. (Springer, Berlin, 1986) p. 161.
[31] T.J. Sejnowksi, in: Parallel Models of Associative Memory, G.E. Hinton and J.A. Anderson, eds. (Lawrence Erlbaum, Hillsdale, 1981) p. 189.
[32] M. Erb, G. Palm, A. Aertsen and T. Bonhoeffer, in: Strukturbildung und Musteranalyse, Proc. 9th Cybernetics Congress (DGK, Göttingen, 1986) p. 23.
[33] M. Erb, A. Aertsen and G. Palm, in: Dynamics and Plasticity in Neuronal Systems, N. Elsner and W. Singer, eds. (Thieme, Stuttgart, 1989) p. 445.
[34] M. Erb, Simulation Neuronaler Netze: Stabilität, Plastizität und Konnektivität, Ph.D Thesis, Eberhard-Karls-Universität, Tübingen (1992) [in German].
[35] M. Abeles, E. Vaadia and H. Bergman, Network 1 (1990) 13.
[36] M. Erb and A. Aertsen, in: Information Processing in the Cortex: Experiments and theory, A. Aertsen and V. Braitenberg, eds. (Springer, Berlin, 1992) p. 201.
[37] V. Braitenberg, in: Theoretical Approaches to Complex Systems, Lecture Notes in Biomathematics, Vol. 21, R. Heim and G. Palm, eds. (Springer, Berlin, 1978) p. 171.
[38] G. Palm, Neural Assemblies, An Alternative Approach to Artificial Intelligence (Springer, Berlin, 1982).
[39] M. Erb, Computersimulation eines assoziativen Speichers als Modell für die Funktion der Großhirnrinde, MSc Thesis, Eberhard-Karls-Universität Tübingen, (1985) [in German].
[40] G. Palm, in: Brain Theory, G. Palm and A. Aertsen, eds. (Springer, Berlin, 1986) p. 211.
[41] G. Palm, in: Physics of Cognitive Processes, E.R. Caianiello, ed. (World Scientific, Singapore, 1987) p. 380.
[42] K.-H. Boven and A. Aertsen, in: Parallel Processing in Neural Systems and Computers, R. Eckmiller et al., eds. (Elsevier, Amsterdam, 1989) p. 53.
[43] R.J. MacGregor, Neural and Brain Modeling (Academic Press, San Diego, 1987).
[44] P.H. Bedenbaugh, G.L. Gerstein, K.-H. Boven and A.M.H.J. Aertsen, Soc. Neurosci. Abstr. 14 (1988) 651.
[45] P.H. Bedenbaugh, G.L. Gerstein and A.M.H.J. Aertsen, Soc. Neurosci. Abstr. 16 (1990) 1224.
[46] O. Bernander, R.J. Douglas, K.A.C. Martin and C. Koch, Proc. Nat. Acad. Sci. USA 88 (1992) 11569.
[47] M. Rapp, Y. Yarom and I. Segev, Neural Comp. 4 (1992) 518.
[48] A.M.H.J. Aertsen and P.I.M. Johannesma, Biol. Cybern. 43 (1981) 133.
[49] J.J. Eggermont, A.M.H.J. Aertsen, D.J. Hermes and P.I.M. Johannesma, Hearing Res. 5 (1981) 109.
[50] A. Aertsen, G. Gerstein and P. Johannesma, in: Brain Theory, G. Palm and A. Aertsen, eds. (Springer, Berlin, 1986) p. 7.
[51] H.R. Dinse, K. Krüger and J. Best, Concepts Neurosc. 1 (1990) 199.
[52] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk and H.J. Reitboeck, Biol. Cybern. 60 (1988) 121.
[53] C.M. Gray and W. Singer, Proc. Nat. Acad. Sci. USA 86 (1989) 1698.
[54] C.M. Gray, P. König, A. Engel and W. Singer, Nature 338 (1989) 334.
[55] M. Abeles, H. Bergman, E. Margalit and E. Vaadia, J. Neurophysiol. 70 (1993) 1629.
[56] M. Abeles, Y. Prut, H. Bergman, E. Vaadia and A. Aertsen, in: Brain Theory: Spatio-temporal aspects of brain function, A. Aertsen, ed. (Elsevier, Amsterdam, 1993) p. 149.
[57] P.M. Milner, Psychol. Rev. 81 (1974) 521.
[58] A.K. Engel, P. König, A.K. Kreiter, T.B. Schillern and W. Singer, Trends Neurosci. 15 (1992) 218.
[59] A. Aertsen and M. Arndt, Curr. Op. Neurobiol. 3 (1993) 586.
[60] H. Sompolinsky, D. Golomb and D. Kleinfeld, Proc. Nat. Acad. Sci. USA 87 (1990) 7200.
[61] H. Sompolinsky, D. Golomb and D. Kleinfeld, Phys. Rev. A 43 (1991) 6990.
[62] O. Sporns, J.A. Gally, G.N. Reeke Jr. and G.M. Edelman, Proc. Nat. Acad. Sci. USA 86 (1989) 7265.
[63] T.B. Schillen and P. König, Neural Comp. 3 (1991) 167.

[64] P. König and T.B. Schillen, Neural Comp. 3 (1991) 155.
[65] R. Eckhorn, H.J. Reitboeck, M. Arndt and P. Dicke, Neural Comp. 2 (1990) 293.
[66] R. Miller, in: Information Processing by the Brain: Views and hypotheses from a cognitive-physiological perspective, H. Markowitsch, ed. (Hans Huber, Bern, 1988) p. 179.
[67] J. Wickens, in: Information Processing in the Cortex: Experiments and theory, A. Aertsen and V. Braitenberg, eds. (Springer, Berlin, 1992) p. 271.
[68] D. Plenz and A. Aertsen, in: The Basal Ganglia IV. New ideas and data on structure and function, G. Percheron, J. S. McKenzie and J.S. Féger, eds. (Plenum Press, New York, 1994), in press.