A MODEL OF NEOCORTEX

Elie Bienenstock

Division of Applied Mathematics
Brown University
Providence RI 02912, USA
e-mail: elie@dam.brown.edu

and CNRS, Paris, FRANCE

December 1994

[0]Supported by Office of Naval Research contract N00014-91-J-1021, National Science Foundation contract DMS-9217655, and ARL contract MDA972-93-1-0012.

Abstract

Prompted by considerations about (i) the compositionality of cognitive functions; (ii) the physiology of individual cortical neurons; (iii) the role of accurately-timed spike patterns in cortex; (iv) the regulation of global cortical activity, we suggest that the dynamics of cortex on the 1-ms time scale may be described as the activation of circuits of the *synfire-chain* type (Abeles 1982, 1991). We suggest that the fundamental computational unit in cortex may be a wave-like spatio-temporal pattern of synfire type, and that the binding mechanism underlying compositionality in cognition may be the accurate *synchronization* of synfire waves that propagate simultaneously on distinct, weakly-coupled, synfire chains. We propose that Hebbian synaptic plasticity may result in a *superposition* of synfire chains in cortical connectivity, whereby a given neuron participates in many distinct chains. We investigate the behavior of a much-simplified model of cortical dynamics devised along these principles. Calculations and numerical experiments are performed based on an assumption of randomness of stored chains, in the style of statistical physics. It is demonstrated that: (i) there exists a *critical value* for the total length of stored chains; (ii) this storage capacity is linear in the network's size; (iii) the behavior of the network around the critical point is characterized by the self-regulation of the number of synfire waves coactive in the network at any given time.

1 Introduction

Mathematical/computational models of the dynamics of cerebral cortex, inspired in particular from statistical physics (e.g. Little 1974; Hopfield 1982; Amit 1989), attempt to account for what appear to be fundamental properties of mental *representations,* or *entities:*

1. **Stability** Mental entities are persistent, or *stable,* over a time span characteristic of working memory (of the order of 1 s);

2. **Long-term memory (reproducibility)** A given entity can be evoked, or *retrieved,* reliably and reproducibly, with all or part of the specific features associated with it, at different times;

3. **Learnability** Learning allows the *storage* of new entities in long-term memory;

4. **Large storage capacity** The brain has the capacity to store and retrieve *large numbers* of distinct entities.

Another aspect of the dynamics of mental representations, fundamental yet less amenable to quantitative study, is *compositionality,* related to the property of *dynamical binding:*

5. **Compositionality** New entities can be constructed by composing with each other, partly in a recursive manner, entities that are already stored—each one with its individual associations—in our cognitive system. Compositionality is virtually limitless, although allowable constructions respect various domain-specific constraints. Composition relies on dynamical, i.e., reversible, *binding,* whereby the same constituents can be bound into different constructs at different times. A dynamically-bound entity can also be stored as such in long-term memory.

A model of cortex typically establishes a correspondence between mental entities and a collection of *states* of a suitably-defined dynamical system; these states are meant to describe spatio-temporal patterns of activity in cortex. Stability (Property 1) is often accounted for by the existence of *fixed points* in the dynamics; other stable states, e.g. limit cycles, can also be used. The retrieval of a given entity upon stimulation of the system by an appropriate input (Property 2) is interpreted as the convergence of the state towards a fixed point, or *attractor.* This operation is sometimes viewed as the *completion* of a partly-elicited entity, exploiting stored *associations* between its features; hence the frequently-used term *associative-memory model.* Learning (Property 3) is envisioned as a modification of the coupling constants between the neurons, i.e., the synaptic weights, resulting in an appropriate modification of the dynamics, e.g. the creation of an attractor at a prescribed point in state space. The storage capacity of the system (Property 4) is defined as the maximal number of distinct patterns—generally chosen at random—that can be stored and retrieved without mutual destructive interference.

In most models, the attractors corresponding to mental representations are defined in terms of the *firing rates* of a population of neurons; such firing rates are typically measured on the 1-s time scale. Thus, the *fine temporal structure* of spike trains on the 1-ms time scale, and in particular the accurate correlation or synchrony between spikes emitted by different neurons, is ignored. However, the idea that this fine temporal structure of neural activity could be used by the brain, in particular in binding and compositionality (Property 5), has become a focus of interest in the last few years (von der Malsburg 1981, 1987; see also von der Malsburg and Bienenstock 1986). Research in this area is still exploratory (for a short review see Bienenstock and Geman 1993).

We note at this point that *rate-coded* associative-memory models, which postulate the existence of attractors defined in terms of the firing rates of neurons, predict a *bimodal* distribution of firing rates across the neural population under consideration; moreover, the higher of the two modes should be near saturation ($\approx 10^3$ spikes/s). A discrepancy exists between these predictions and biological data (see however, e.g., Amit and Tsodyks 1991).

In this paper we propose that the fundamental unit of computation in cortex is a spatio-temporal activity pattern of the *synfire* type (Abeles 1982, 1991); such a pattern is characterized by the propagation of volleys of *nearly-synchronous* spikes—as measured on the 1-ms time scale—along a *synfire chain*, i.e., a sequence of *pools* of neurons connected in a feedforward way (Figure 1). We propose that the microstructure of cortical connectivity, shaped by Hebbian plasticity, is a *superposition* of synfire chains, whereby a neuron participates in many distinct chains. At any given time, a large number of synfire chains are simultaneously active. We propose to distinguish between two modes of synfire activity: a *background,* or *spontaneous,* mode, and a *computation-related* mode, as takes place for instance in perception or language-related behavior. We suggest that the latter is characterized by the *synchronous* propagation of synfire patterns along a number of distinct chains. Such synchronization, or dynamical binding, is made possible by weak synaptic coupling between chains; it results in several chains behaving, for a period of time, like a single broader chain. We propose that this mechanism may underlie compositionality in cortex. In background activity, synfire patterns along coactive chains are *asynchronous.*

We have argued elsewhere on the basis of numerical studies (Bienenstock and Doursat 1995) that neural circuits of the synfire type will develop *spontaneously* in an initially unstructured network as a result of Hebbian synaptic plasticity. This self-organization of synfire chains may be viewed a process of natural selection (Changeux and Danchin 1976; Edelman 1988), where neural circuits are selected according to a simple 'fitness' criterion related to the cooperativity of parallel multisynaptic pathways (von der Malsburg 1987). This provides a mechanism for incorporating both environmental and endogenous factors in the epigenesis of the nervous system (Section 9.3).

We shall present in this paper a much-simplified model of cortical dynamics devised along the principles outlined above. The main features of the model are as follows.

(a) The state of the system describes the spiking activity of a population of excitatory cortical neurons on the 1-ms time scale.

(b) State update is synchronous.

(c) The total activity level in the network is approximately constant.

(d) The network supports reproducible spatio-temporal patterns of activity, of the synfire type, defined on the 1-ms time scale. The length of a synfire chain determines the duration of the pattern it carries; this duration is measured on the 1-s time scale. Another important parameter is the *width* of the chain.

(e) Storing a given synfire pattern takes the form of a Hebbian reinforcement of *synfire links,* where a synfire link is defined as the collection of all synapses from one pool of neurons to the next pool in a given chain. Synfire patterns are superimposed in the network in such a way that each neuron participates in many stored patterns.

(f) Stable and reproducible synfire activity takes place if the *distribution of membrane potentials* across the network is, at any given time on the 1-ms time scale, bimodal. The distribution of *spiking rates,* measured on the 1-s time scale, is unimodal, and centered around a low value.

(g) The condition mentioned in (f) ensuring stability of synfire activity is satisfied provided a certain *storage capacity* is respected. This storage capacity is linear in the size of the network.

(h) *Dynamical binding* is achieved by the synchronization of synfire waves propagating along distinct chains; it is induced by weak synaptic coupling between these chains. (Dynamical binding cannot be studied under the strictly-synchronous update scheme—item (b) above.)

The motivation behind items (a) and (b) is twofold. On the one hand, there is reason to believe that the integration of afferent signals by the somato-dendritic membrane of a given cortical neuron is highly sensitive to the precise timing of these signals; some authors describe the function of cortical neurons as *coincidence detection* (Abeles 1982, 1991). Electrophysiological data moreover support the notion that the dynamics of cortex generates reproducible accurately-timed spatio-temporal patterns (Abeles et al. 1993a; see also Softky and Koch 1992), and that these play a role in behavior (Abeles et al. 1993b, 1993c). These issues will be discussed in Sections 3 and 4, and then again in Section 9.5, after the model is presented.

On the other hand, the use of the accurate time structure of synfire activity makes it possible to envisage a neural mechanism for binding and composition (Property 5), consisting of synchronizing the propagation of activity across different constituent synfire patterns (Bienenstock 1991, Abeles et al. 1993b). In Sections 5 and 6 we shall study a strictly-synchronous version of the model. Such an update scheme does not account for binding and compositionality, since it synchronizes all coactive waves by construction. We propose in Appendix B a refined version—which we call an *almost-synchronous* dynamics—motivated by more detailed biological considerations; dynamical binding is studied under this model in Appendix C.

The motivation behind item (c) is the existence in cortex of a system of inhibitory neurons which controls the global level of activity; such a system is *implicit* in the model. The regulation of firing rates is an important issue in the modeling of the physiology of cortex; we shall argue (Sections 2 and 7) that connectivity of the synfire-superposition type (item e) allows *robust* regulation of firing rates, provided the two modes in the distribution of membrane potentials (item f) are well-separated. In Appendix D we contrast this situation with the case of a *random* connectivity graph, where the distribution of membrane potentials is unimodal; we argue that in such a graph robust regulation of global cortical activity is practically ruled out.

The quantitative part of this paper (Sections 6 and 8) is concerned with storage capacity, defined as the number p of synfire links (item (e)) that can be stored without mutual destructive interference. As is customary, we shall base our estimations on an assumption of *randomness* of stored patterns. Analysis as well as computer simulations (Section 6.3) show that there exists a *critical value* p_c for p; p_c is the storage capacity of the network, and we shall see that it is linear in the size of the network. When p is of the order of p_c, each neuron participates in many stored patterns. The network then exhibits a non-trivial behavior, characterized by the occurrence of *spontaneous* synfire activity, and by the *self-regulation* of the total number of synfire waves present in the system at any given time.

Storage capacity is large but finite; the composition of stored patterns with each other is however expected to provide a virtually limitless repertoire of new entities (Section 8). We shall demonstrate (Appendix C) that the binding of synfire patterns into a composite pattern requires only *weak* synaptic coupling between chains, somewhat in the same way as phase locking of oscillators is induced by weak coupling between them.

One word of caution is in order at this point. Although the spatio-temporal patterns described in items (d) and (e) exhibit the properties of mental representations listed at the beginning of this section (Properties 1-5), they should not be viewed as standing in strict correspondence with such mental entities. The very notion of a mental *entity* is a problematic one, inasmuch as it attempts to put boundaries, both in semantic space and in time, where there need not be. Describing mental activity as a succession of evocations—internally or externally triggered—of separate entities does not do justice to the semantic connectedness, as well as to the permanent unfolding, of cognitive processes. A more satisfactory, if less quantifiable, account—in particular where language is concerned—is one where a mental process is construed as a flow, or scanning, through a highly connected cognitive domain, with shifts of emphasis—continuous or discrete—along a sequence of structures; the latter are part of a broad domain, the semantic background of the process taking place—cf. the notion of *profiling* in Langacker (1987).

On such a view, the use of *strictly stable* states to model the persistence of mental representations in working memory (Property 1) is not desirable; it may also be futile to count with precision how many entities fit within a given network. In this paper we suggest that the notion of stability as used to characterize *wave propagation* along a synfire chain (Abeles 1991) provides, in most cases, an appropriate model for the persistence of mental entities in working memory: we shall say that synfire propagation is stable if it is *reproducible*, and *robust* against perturbations. The length of the pathway, hence the temporal extent of stability, is finite. Thus, although Property 1 is not strictly met, Property 2 is: *partial* activation of any pool of a chain will reliably elicit the full activation of the next pool in that chain. In this sense, synfire chains can be said to display the property of associative memory. When chains include feedback links, the activity may reverberate indefinitely within a given chain. Such a situation provides a transition from the model proposed here to conventional associative-memory models (Section 9.4).

We shall in Section 8 derive a rough estimate for the *minimal* width (item d) of a chain required for stable synfire propagation; we shall term a chain of minimal width a *narrow* chain. We suggest that isolated narrow synfire patterns have little semantic content; any such pattern may convey as much 'meaning' as does, in visual perception, a single part of an object boundary, say a straight line, taken in isolation from its context (see Section 9.6). Storage capacity (item g), which is linear in the size of the network (Section 6), is expressed in terms of narrow chains.

We shall argue in Section 8 that at any given time a large number of narrow chains are simultaneously active in the brain. As mentioned above, we propose that the synchronization of synfire patterns provides the mechanism for compositionality in cortex (Bienenstock 1991). Composite, or broad, synfire patterns inherit from their narrow constituents the fundamental properties mentioned at the beginning of this section. From a mere combinatorial viewpoint, the number of composite patterns is virtually infinite (Section 9.7). However, as mentioned, (weak) synaptic coupling is required to bind synfire patterns with each other in a stable configuration. These couplings implement the domain-specific constraints that are manifest in all forms of compositionality (Property 5). It is these temporarily-bound broad synfire patterns that, we propose, provide a neural counterpart to mental entities as described in the beginning of this section.

In sum, the assumption put forward in this paper, namely that our brains compute with synfire patterns rather than with *fixed-point attractors* (Hopfield 1982; Amit 1989) or with *cardinal cells* (Barlow 1972), offers four main advantages:

- it provides a neural framework for compositionality that appears more plausible, in several ways, than the currently popular solution based on the phase locking of neural oscillators (see Section 9.7);
- it is consistent with, and further amplifies, the view that brains develop according to principles of natural selection;
- it accounts for the finding that some accurately-defined firing patterns in cortex reproduce more often than by chance and correlate with behavior (see discussion in Section 9.1);
- it suggests a mechanism for robust regulation of cortical activity, predicting a distribution of firing rates centered around low values.

The plan of the paper is as follows. The issue of low firing rates is discussed in Section 2. Section 3 discusses briefly the relative merits of synchronous and asynchronous update schemes in models of cortical dynamics. In Section 4 we provide a succinct review of the dynamics of synfire chains, based on the work of Abeles and collaborators. Section 5 presents the model and Section 6 describes its behavior, first qualitatively and then in more detail with the help of computer simulations. In Section 7 we briefly discuss the tolerance of the model to noise. Section 8 discusses plausible ranges for the various parameters of the model. In the general discussion (Section 9) we address again most of the above-mentioned topics; we consider briefly the issue of selection vs. instruction in brain epigenesis (Section 9.3); we outline a synfire-superposition model of cortical computation, focusing on perception and emphasizing the hierarchical/recursive aspects of compositionality; finally, we propose a metaphor which likens complex cortical spatio-temporal patterns constructed from elementary synfire waves to complex proteins obtained from linear chains of amino acids by a process of *folding* in three-dimensional space.

In Appendix A we consider an extension of the notion of a synfire chain, whose connections all have identical conduction time, into that of a *synfire braid;* the latter accommodates more realistic cortical circuits, with non-uniform conduction times. In Appendix B we describe the 'almost-synchronous' update scheme (see above) and discuss again the biological plausibility of the strictly synchronous scheme used in the paper. Based on this almost synchronous dynamics, Appendix C shows that weak synaptic coupling induces synchrony across chains. Finally, in Appendix D, we study the case of a random connectivity graph; we argue that such a graph does not allow robust regulation of global activity and does not support *long-range order* in accurately-defined activity patterns.

2 The regulation of global cortical activity

In this section we discuss, in view of some elementary physiological and anatomical findings, the problem of low firing rates mentioned in Section 1. Accounting for these observations is one of the goals of the model proposed in this paper.

The fact that the global activity in cortex is fairly stable around low yet non-zero levels is somewhat of a puzzle, in view of the following two observations. First, the overwhelming majority of neurons and of synapses in cortex—about 90%—are excitatory. Most excitatory neurons are pyramidal cells, producing local connections as well as long-distance axons that travel through the white matter. In this paper, we shall be interested in the dynamics of a large population of excitatory cortical neurons with dense feedback connectivity, typically a collection of pyramidal cells distributed across *several* areas (see discussion in Appendix A).

Second, the theoretical limit for the firing frequency of individual neurons, set by their absolute refractory period, is close to 10^3 spikes/s. In spite of this, sustained high firing rates are seldom observed in cortex. Although short high-frequency bursts do occur, cortical cells in general have low firing rates: the mean of the distribution of firing rates in cortex is of the order of 5 spikes/s, with a somewhat smaller median.

As mentioned in Section 1, these facts are not easily accounted for by rate-coded associative-memory models; an attractor in such a model is typically a bimodal distribution of firing rates, stable over a period of the order of 1 s, with the higher mode of the distribution located near the saturation frequency for cortical neurons (see however, e.g., Amit and Tsodyks 1991).

In all likelihood, inhibitory neurons play an important role in the regulation of global cortical activity. It is worth noting that inhibitory neurons are distinguished from excitatory ones in cortex by several anatomical and morphological features. In particular, inhibitory neurons in neocortex apparently never project to remote targets via the white matter, whereas pyramidal (excitatory) cells always or almost always do (Braitenberg 1977). Inhibitory neurons also tend to establish contacts in proximal parts of the postsynaptic cell's soma or dendrites; on dendritic shafts rather than spines; or, in the case of chandelier cells, on the initial segments of pyramidal-cell axons. Finally, the characteristic time of the feedback mediated by long-distance excitatory projections is longer than that of inhibitory feedback, since the latter is exclusively local.

These data are suggestive of the existence of a control over global cortical activity exerted by a system of relatively fast-acting and powerful inhibitory neurons. One may consider that the role of these neurons is to set the *effective firing thresholds* of the excitatory neurons. On this view, inhibitory neurons would play little direct part in the representation of entities of specific functional or semantic content; the connections to and from these neurons would be fairly diffuse. The regulatory function mediated by inhibitory neurons could nevertheless be complex and elaborate in several ways: different levels of activity could be imposed on different subpopulations of pyramidal cells, or on different areas of cortex, or at different times; these levels could also be made to depend in various ways on the outcome of ongoing computations, which might themselves involve the manipulation of specific mental representations.

There are essentially two alternatives for the modeling of cortical dynamics under the assumption that global activity is controlled by a dedicated system of inhibitory neurons. The first is to explicitly include in the model such a regulatory system, consisting of non-specific, i.e., diffusely-connected, inhibitory neurons. We shall briefly review in Section 4 a model that illustrates this approach (Abeles et al. 1993b). The second alternative is to leave out the inhibitory neurons, and to 'hard-wire' in the dynamics the assumption of conservation—or near conservation—of total activity. This is the approach adopted here (Section 5).

We stress that this device is a mere convenient way to model the function of a regulation mechanism. It is a legitimate simplification under the condition that this mechanism acts efficiently and *robustly.* This issue is an important one, as delayed-feedback systems often display chaotic behavior. We shall argue in Section 7 that the operating conditions of the model proposed are precisely such as to create favorable conditions for robust regulation. Under normal functioning of the model, the distribution of *membrane potentials* across the excitatory population at any given time is bimodal. Regulation can be expected to be effective and robust if the two modes of this distribution are well-separated and if the firing thresholds of the excitatory neurons, set by the diffuse inhibitory system, are kept at all times at intermediate values, between the two modes.

If we now consider the operation of the network from the perspective of *firing rates,* which are measured on a time scale about three orders of magnitude coarser than the one where synfire activity is observed, we are in effect averaging out the spatio-temporal structure characteristic of such activity. As mentioned in Section 1 (item f), synfire propagation, where activity does not dwell on the same collection of neurons, results in a *unimodal* distribution of firing rates, centered around *low* values (see Section 8). This accords well with experimental data.

In Appendix D we contrast the behavior just described, which takes place if the network is a superposition of synfire chains respecting the storage-capacity limitation (items e-g in Section 1), with that of a *randomly connected* one, where regulation is expected to generate chaotic behavior.

3 Update schemes: sequential or parallel

Models of cortical dynamics fall in two large categories, depending on whether they use an *asynchronous,* also called sequential, update scheme, or a *synchronous* scheme, also called parallel. We shall argue here (see also Section 9.5) that none of the two alternatives is fully satisfactory. However, depending upon the connectivity of the network under study, one update scheme may be more appropriate than the other.

In models inspired from statistical mechanics, asynchronous Glauber-type dynamics is the method of choice (Glauber 1963); this dynamics is easy to implement numerically, and, under suitable conditions, yields convergence to an analytically tractable equilibrium distribution. To justify the use of such a dynamics one may argue that it is unlikely that two given neurons will update their states at precisely the same instant. The asynchronous dynamics can then be viewed as appropriate if the time interval between two updates is interpreted as a very small fraction of a millisecond, roughly $\Delta t/N$, where Δt is a synaptic delay (about 1 ms), and N is the number of neurons in the network.

By doing so however one assumes that the information about a given update is made immediately available to other neurons, which is unrealistic. Evidently, many neurons in the brain fire at *virtually* identical times, and it would thus be more appropriate to make the update at least partly synchronous. However, it is not *a priori* clear which neurons will actually fire in synchrony and when. Equally importantly, it is not clear *a priori* when synchrony matters and when it does not. Stated differently, the artificially fast transmission of information implicit in the asynchronous dynamics may or may not make a difference. If neurons are believed to be activated by large numbers of postsynaptic potentials summed over durations significantly longer than the typical synaptic delay, the asynchronous dynamics could be viewed as appropriate.

Synchronous update schemes pay attention to the precise relative timing of the firing of neurons, in particular to temporal *coincidences* among spikes carried by different afferents to a given neuron. Such coincidences are taken by many researchers to be of crucial importance to cortical dynamics. The reason for this is that the integration of afferent signals by the postsynaptic membrane is likely to be a highly *non-linear* operation in time (it is also non-linear in space, but we shall not be concerned with this aspect). Thus, Abeles (1991, Section 7.1) argues that, while the average increase in a neuron's firing rate resulting from the arrival of n spikes in n different synapses is roughly n times the effect of a single spike if the spikes are *asynchronous,* the combined effect of these n spikes may be considerably larger if the spikes are *synchronous* or nearly so. Synchrony for that matter means that the EPSPs (excitatory postsynaptic potentials) caused by the n spikes overlap in time; the duration of most EPSPs in normal conditions in cortex is of the order of 5 ms, and Abeles estimates that on the order of 25 overlapping EPSPs are required to trigger the firing of a cortical neuron (with probability .5).

Whether the use of synchronous update is appropriate or not may depend on assumptions about the *connectivity* configuration of cortex. It is sometimes argued that synchronous update requires a 'central clock,' of doubtful existence. In this paper we shall focus our attention on networks of the *synfire-chain* type (Abeles 1982, 1991). As we shall see in the next section and in more detail in Appendix B, the activity in a synfire chain organizes itself spontaneously in volleys of nearly-synchronous spikes propagating along the sequence of pools of neurons that defines the chain. Synchronous update is, in a first approximation, appropriate for the modeling of this propagation. We shall argue in Appendix B that it is also appropriate for a network containing a *superposition* of synfire chains, under the conditions studied in this paper.

4 A brief review of the dynamics of synfire chains

A synfire chain is an idealized network, unlikely to occur in pure form in cortex. In its simplest version, a synfire chain (Abeles 1982, 1991) is a feedforward network including a large number of *pools,* i.e., layers, of neurons, with a fixed number n of neurons in each pool; n is called the *width* of the chain, and the number of pools p is its *length.* Connectivity from pool i to pool $i+1$ is either complete, i.e., n-to-n, or random with a large-enough probability of presence of each synapse. Figure 1 is a schematic diagram of a fragment of a synfire chain, with incomplete connectivity.

(Insert Figure 1 around here)

Unlike feedforward networks used for statistical decision making or pattern recognition (e.g. backpropagation nets), synfire chains do not perform transformations or mappings. We view them as a mechanism used by the brain to generate spatio-temporal patterns meeting the properties listed at the beginning of Section 1.

Note first that the pathways that make up a synfire chain *cooperate* with each other to make a neuron fire. Consider for instance, in Figure 1, the two pathways $i \longrightarrow j \longrightarrow l$ and $i \longrightarrow k \longrightarrow l$. Any spike emitted by i at time t increases the probability of both j and k emitting a spike at time $t+1$. This in turn increases the probability of having two synchronous spikes impinge on l at time $t+2$. In view of the property of coincidence detection of neurons (Section 3), this connectivity graph is more favorable to the firing of a target neuron—such as l—than a graph with an identical number of randomly-arranged connections.

This mechanism, when operating in a strictly feedforward chain with substantial *diverging/converging*—e.g. complete—connectivity from one pool to the next, generates *waves* of synchronized activity. These waves are synchronous volleys of spikes, which, when initiated in a given pool of the chain—at its beginning or at another position—travel down reliably and reproducibly to the end of the chain. In this *synfire-transmission* mode, all or almost all neurons in a given pool are active at nearly the same instant in time.

Several approaches can be used to study synfire transmission. In Abeles (1991, Chapter 7) synchronous update is used. It is demonstrated, under a simple noise model for membrane potentials, that if all neurons in pool i are activated simultaneously at a given time t, then, with high probability, all or almost all neurons in pool $i+1$ will be active at time $t+1$. Moreover, it is enough to activate a given fraction, called the *ignition threshold,* of the neurons in pool i to reliably trigger synfire activity in pool $i+1$. In this sense, synfire transmission is *stable.* The fluctuations of the postsynaptic membrane potentials result in *temporal jitter* in the firing times of neurons. It is demonstrated that in spite of this jitter the firing times of the neurons in a given pool i remain tight, that is, their scatter does not increase with i. Again, this is a property of *stability* of the synfire pattern as a whole.

A more elaborate model, consisting of a system of coupled differential equations simulating the evolution of each neuron's membrane potential and spiking activity, is studied numerically in Abeles et al. (1993b), and in more detail in Diesmann et al. (in preparation). The coupling is brought about by the excitatory synapses making up the chain structure, and also by a diffuse inhibitory system which regulates the overall activity level; biologically plausible values are used for the different parameters, including the durations of EPSPs and IPSPs (inhibitory postsynaptic potentials). Again, the non-linearity of temporal summation by the postsynaptic membrane resulting from the short duration of EPSPs is shown to ensure the *coherency* of the wave of activity throughout its propagation along the chain.

A synfire chain is an 'amplifier of synchronism' (Abeles et al. 1993b) in the following sense. If one stimulates the first few pools in the chain by increasing their global activity levels without specifically imposing coincident firing, a synchronized wave will form from time to time and reliably propagate down the chain. This will happen every time the number of coincident spikes in one of the stimulated pools exceeds the ignition threshold (see above). If the chain is *complete* (n-to-n connectivity), the ignition threshold is of the order of 25 (see Section 3).

We stress again at this point that synfire transmission is not *perfectly* reproducible; although stability of transmission as well as coherency of firing within each pool are ensured, the exact timing of firing of a given pool is subject to a small *jitter,* that is, variability from one occurrence of a wave to another. The jitter in the timing of the firing of a pool i is due to the randomness in the firing of individual neurons in i, which, as said, results chiefly from fluctuations of individual membrane potentials. However, due to the temporal *averaging* performed by the diverging/converging connectivity, the jitter in the time of firing of a given pool is smaller than the variability in the firing times of individual neurons. Parameters (e.g. the level of noise) can of course be chosen such that synfire transmission is made unreliable. Temporal jitter plays an important role in the model proposed (Section 6.4); a more detailed discussion of stability and jitter in synfire transmission is given in Appendix B.

In sum, a synfire chain is a feedforward network including a number of pools of neurons with diverging/converging connectivity from one pool to the next. Synfire transmission, triggered by suitable stimulation, consists of the stable propagation of a synchronized wave of activity along the chain. This spatio-temporal pattern is reproducible but not perfectly so. Random activity is present, and may interfere to various degrees with synfire activity.

5 The model

In this section we define a synchronous-update dynamics which complies with the requirement of conservation of total activity (Section 2); we also introduce a simple Hebbian-type storage prescription.

Consider a network of N neurons with binary-valued activity states: $x_i(t) \in \{0,1\}$. Our assumption is that, at all t, $\sum_{i=1}^{N} x_i(t) = r$ (or $\approx r$), for a given $r \ll N$. Specifically, we saw in Section 2 that the average activity rate of a cortical cell is 5 spikes/s; therefore, with a time unit of 1 ms, $r/N \approx 5 \times 10^{-3}$.

The *input* to neuron i is defined by:

$$V_i(t+1) = \sum_{j=1}^{N} w_{ij} x_j(t), \tag{1}$$

where w_{ij} is the synaptic weight from neuron j to neuron i. We shall also refer to V_i as the *membrane potential* of neuron i.

A straightforward synchronous conservative dynamics is obtained by activating at time $t+1$ those r neurons that have highest potentials $V_i(t+1)$. We shall term this an 'r-Winners-Take-All' (r-WTA) dynamics. A literal realization of this dynamics in the brain is hardly possible; however, it can be interpreted as resulting, in a first approximation, from the regulatory mechanism discussed in Section 2 (see also Section 7). Indeed, if the inhibitory system acts fast enough and at all times maintains the number of active neurons within certain limits, one should expect that these active neurons will be, for the most part, the neurons with highest excitatory input. The *strict* r-WTA dynamics, where the active neurons at a given time are *precisely* the ones with highest input, shall be investigated in Section 6; it may be viewed as corresponding to a 'zero-temperature' situation. The effect of adding noise to the membrane potentials will be examined briefly in Section 7.

We now introduce an important definition for the study of memory storage and retrieval in the model. A *synfire link* L, or simply *link* ('diverging/converging link' in Abeles 1991), is a collection of synapses from a set of neurons B to a set of neurons A. We shall assume that each of these two sets—not necessarily disjoint—is of size n, with $n \ll N$. In the following section we shall deal only with *complete* synfire links; these include a synapse from every neuron $j \in B$ to every neuron $i \in A$. Incomplete links, where a synapse from a neuron in B to a neuron in A is present with a given probability $\pi < 1$, will be considered in Section 7; such links are said to have *multiplicity* m (Abeles 1991) if every neuron in A receives contacts from at least m neurons in B and every neuron in B sends contacts to at least m neurons in A. We shall use the notation $L = (A, B)$, and we shall sometimes say that L is *active* at time t if all neurons in B are active at t and all neurons in A are active at $t+1$.

We *store* a synfire link $L = (A, B)$ by incrementing all synaptic weights w_{ij}, $i \in A$, $j \in B$. We shall investigate the dynamics of the network under various connectivity configurations resulting from *superpositions* of links. The rule for superimposing a collection of p links $L^\mu = (A^\mu, B^\mu)$, $\mu = 1, \ldots, p$, is:

$$w_{ij} = C \sum_{\mu=1}^{p} \mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}}, \tag{2}$$

where $\mathbf{1}_{\{E\}}$ is the indicator function of event E, and C is a positive constant which shall be set equal to 1. Equation 2 can be rewritten in the more familiar Hebbian form

$$w_{ij} = C \sum_{\mu=1}^{p} \xi_i^{\mu+} \xi_j^{\mu-}, \tag{3}$$

where, for any μ and any i, $\xi_i^{\mu+} = 1$ if $i \in A^\mu$, 0 if $i \notin A^\mu$, and $\xi_i^{\mu-} = 1$ if $i \in B^\mu$, 0 if $i \notin B^\mu$. This rule is sensitive to the *time structure* of stored activity patterns; in this respect, it is actually closer than symmetric storage rules to the letter—and probably the spirit—of Hebb's principle of synaptic plasticity (Hebb 1949). Note also that using $\{0, 1\}$ variables—rather than $\{-1, +1\}$ variables as is customary in models inspired from statistical physics—results in all weights being non-negative; inhibitory synapses however are implicit in the conservative dynamics.

We finally remark that the establishment of a synfire chain in cortex is envisaged here as a simple memory-storage operation, from the perspective of instructive learning. This picture shall be revised in Section 9.3.

6 Behavior of the model

This section contains the mains results of the paper. We shall describe the behavior of the r-WTA dynamics as a function of the *load* of the system, that is, the number of stored links. Although the calculations we shall make are relatively straigthforward, some readers may want to skip them. We therefore provide here a brief summary of results.

6.1 Summary of results

The situation of main interest to us is one where a number of *randomly constructed* synfire chains are stored. By this we mean that the pools that make up these chains are all chosen at random, independently of each other. As mentioned in the previous section, we assume that all chains are of width n, i.e., that there are exactly n neurons in each pool. The chief result is the following: if the *total length* p of stored chains, i.e., the total number of pools (equivalently, links), does not exceed a certain critical value p_c, storage/retrieval is efficient. This means that whenever a given pool in one of the stored chains is active at time $t = 1$ the next pool in the chain will be, with high probability, active at time $t = 2$. Assuming, somewhat artificially, that the last pool in the chain feeds back to the first pool, the life time of a synfire wave is, essentially, infinite. If this cyclic boundary condition does not hold and the chain is strictly feedforward, the wave dissolves, i.e., spreads randomly in the network, when it reaches the end of the chain.

Note that at any given time more than one pool will generally be active. The largest possible number of coactive pools, hence of coactive synfire waves, is determined by the constraint on total activity. Specifically, if we denote by h the number of coactive pools, then, since $r \ll N$ (see above, Section 5) and since all pools are random, the h coactive pools are essentially disjoint; hence an upper limit is $h \approx r/n$. Numerical values will be discussed in Section 8; we shall see that, typically, $1 \ll r/n \ll p$. Thus, if at time t the activity is concentrated on h of the p stored pools it remains so at subsequent times: synfire transmission takes place as described in Section 4, which means that the h packets of activity propagate along the stored chains. This propagation is robust against noise and reproducible, which corresponds to Properties 1 and 2 of Section 1.

We shall estimate the storage capacity p_c (Property 4) to be about $N^2/25r$. As mentioned above, $r/N \approx 5 \times 10^{-3}$; thus, p_c is of the order of $8N$. This calculation pertains to the chains we called *narrow* in Section 1; such chains are not supposed to convey much meaning when activated separately, which is typically the case in background activity. It is the synchronized activation of a number of such synfire patterns that, we propose, corresponds to mental representations as described in Section 1. A binding mechanism, ensuring the stability of such composite patterns (Property 5 of Section 1), is discussed in Appendix C. No quantitative estimate of the number of composite patterns can be achieved in the present framework. However, from a mere combinatorial point of view, this number is exceedingly large.

6.2 Superposition of synfire links

We store, according to equation 2, a collection of p links $L^\mu = (A^\mu, B^\mu)$, $\mu = 1, \ldots, p$, where all sets A^μ and B^μ are drawn at random independently of each other, uniformly over all subsets of size n of the set of N neurons. We fix a set of indices $M \subset \{1, \ldots, p\}$, of size $|M| = h$, and denote $\mathcal{B}_M = \bigcup_{\mu \in M} B^\mu$ and $\mathcal{A}_M = \bigcup_{\mu \in M} A^\mu$. We assume that at a given time t all neurons in $\mathcal{B}_M$ are active. The size of this union set $\mathcal{B}_M$ is thus at most r. To satisfy the constraint on total activity exactly, we activate at time t another $r - |\mathcal{B}_M|$ randomly-chosen neurons.

We define storage/retrieval to be *efficient* if the following is true: for any M as above with all neurons in $\mathcal{B}_M$ active at time t, all neurons in $\mathcal{A}_M$ are, with high probability,[1] active at time $t+1$. This requires of course that $|\mathcal{A}_M| \leq r$; it is realized if the potential $V_i(t+1)$ is larger for any $i \in \mathcal{A}_M$ than it is for any $i \notin \mathcal{A}_M$.

The calculation we shall now make underlies all statements on storage capacity made in this paper. For a given neuron i in the network and a given set M as above, we compute the input $V_i(t+1)$ to neuron i at time $t+1$. We denote this potential $V_{i,M}$, to indicate the dependency on M. Due to the randomness of A^μ and B^μ, $\mu = 1, \ldots, p$, the quantity $V_{i,M}$ is a random variable. We decompose it in two terms as follows:[2]

$$\begin{aligned} V_{i,M} &= \sum_{j=1}^{N} w_{ij} \mathbf{1}_{\{j \in \mathcal{B}_M\}} \\ &= \sum_{j=1}^{N} \sum_{\mu=1}^{p} \mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}} \mathbf{1}_{\{j \in \mathcal{B}_M\}} \\ &= U_{i,M} + W_{i,M}, \end{aligned} \tag{4}$$

where

$$U_{i,M} = \sum_{j=1}^{N} \sum_{\mu \in M} \mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}} \mathbf{1}_{\{j \in \mathcal{B}_M\}}, \tag{5}$$

and

$$W_{i,M} = \sum_{j=1}^{N} \sum_{\mu \in \{1,\ldots,p\} \setminus M} \mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}} \mathbf{1}_{\{j \in \mathcal{B}_M\}}. \tag{6}$$

[1] The source of randomness is the choice of the sets A^μ and B^μ, $\mu = 1, \ldots, p$.

[2] For simplicity, we shall assume that $|\mathcal{B}_M| = r$, i.e., that all active neurons at time t are in $\mathcal{B}_M$. The computation is the same if $|\mathcal{B}_M| < r$.

Consider the situation where neuron i occurs in exactly one of the A^μs, $\mu \in M$. It is easily seen that the random variable $U_{i,M}$, when conditioned on this event, is deterministically equal to $|B^\mu| = n$. If on the other hand we condition on the event $i \notin A^\mu$, $\mu \in M$, that is, $i \notin \mathcal{A}_M$, we evidently get $U_{i,M} = 0$.

Thus, the first component $U_{i,M}$ of the input to neuron i is *deterministically* equal to n or to 0, depending on whether neuron i does or does not belong to one of the A^μs, $\mu \in M$.

The second component, $W_{i,M}$, is a *crosstalk* term, and we wish to estimate its first- and second-order moments.[3] The mean of $W_{i,M}$ is easily calculated:

$$\begin{aligned}
E[W_{i,M}] &= E\left[\sum_{j=1}^{N} \sum_{\mu \in \{1,\dots,p\}\setminus M} \mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}} \mathbf{1}_{\{j \in \mathcal{B}_M\}}\right] \\
&= \sum_{j=1}^{N} \sum_{\mu \in \{1,\dots,p\}\setminus M} E\left[\mathbf{1}_{\{i \in A^\mu\}} \mathbf{1}_{\{j \in B^\mu\}} \mathbf{1}_{\{j \in \mathcal{B}_M\}}\right] \\
&= \sum_{j=1}^{N} \sum_{\mu \in \{1,\dots,p\}\setminus M} E[\mathbf{1}_{\{i \in A^\mu\}}] E[\mathbf{1}_{\{j \in B^\mu\}}] E[\mathbf{1}_{\{j \in \mathcal{B}_M\}}], \qquad (7)
\end{aligned}$$

for the three terms in each product are statistically independent. From 7 we see that $E[W_{i,M}] = r(p-h)n^2/N^2$. Let us denote $\eta_1 = rpn^2/N^2$. Since $h \ll p$, we have $E[W_{i,M}] \approx \eta_1$.

Computing the second-order moment of $W_{i,M}$ is not quite as straightforward, for the terms in the double sum defining $W_{i,M}$ (equation 6) are not statistically independent (if they were, $W_{i,M}$ would be a binomial variable, well-approximated by a Gaussian). An approximation however can be derived:[4] $\mathrm{var}(W_{i,M}) \approx r(p-h)(n^2/N^2)(1+nr/N) = E[W_{i,M}](1+nr/N)$. Thus, we have $\mathrm{var}(W_{i,M}) \approx \eta_2$, with $\eta_2 = \eta_1(1+nr/N)$. For biologically plausible values of n and r (see Section 8), $nr/N \approx .5$; hence the variance of the crosstalk η_2 is of the same order as its mean η_1.

We conclude that, provided $\sqrt{\eta_2} \ll n$, the distribution of membrane potentials $V_i(t+1)$ across the network is *bimodal.* The lower mode, centered around η_1, is contributed by those neurons i which do not belong to any A^μ, $\mu \in M$; the higher mode, centered around $n + \eta_1$, is contributed by those neurons i which belong to one A^μ, $\mu \in M$.[5] The width of each of the two modes is $\sqrt{\eta_2}$. Suppose for instance that the separation between the two modes is ten times their width: $\sqrt{\eta_2} \approx n/10$. There is then virtually no overlap between the two modes, and the potentials of neurons in $\mathcal{A}_M$ are all, with high probability, larger than the potentials of neurons outside of $\mathcal{A}_M$; in other words the deterministic component of the input dominates. With high probability, the set of active neurons at time $t+1$ then includes $\mathcal{A}_M$. If on the other hand $\sqrt{\eta_2} > n$, the crosstalk component dominates, and we should expect the activity to spread randomly throughout the network at time $t+1$.

[3] Although $W_{i,M}$ is not exactly Gaussian, its distribution is bell-shaped.

[4] The proof is left to the reader. Hints: define $X^\mu = \sum_{j=1}^{N} \mathbf{1}_{\{j \in \mathcal{B}_M\}} \mathbf{1}_{\{j \in B^\mu\}}$ and note that X^μ is a hypergeometric random variable, of mean nr/N and variance less than and almost equal to nr/N; note also that, conditioned on $|\mathcal{B}_M| = r$, the random variables X^μ, $X^{\mu'}$, $\mu, \mu' \notin M$, $\mu \neq \mu'$, are statistically independent.

[5] Neurons belonging to more than one A^μ are rare, since $n \ll N$. We may thus ignore the higher modes.

Going back to the mechanism implicit in the r-WTA dynamics, namely the regulation of firing thresholds, we see that the condition for the activity to reliably propagate along a given synfire link L^μ, $\mu \in M$, is that the firing thresholds be set appropriately, that is, at a value intermediate between the two modes of the distribution of membrane potentials.

In sum, storage/retrieval is efficient if and only if the standard deviation of the crosstalk, $\sqrt{\eta_2}$, is small with respect to the link width n. Let us now denote $\kappa = n/\sqrt{\eta_2}$; κ measures the *relative separation* between the two modes in the distribution of inputs; note that κ is a decreasing function of p: $\kappa = n/\sqrt{rp(n^2/N^2)(1 + nr/N)}$. If we ignore the factor $1+nr/N$ in the denominator, this simplifies into $\kappa \approx N/\sqrt{rp}$. We may predict that there is a *critical value* κ_c, probably smaller than 10, such that if $\kappa > \kappa_c$ the two modes are well separated and retrieval is efficient. With the above approximation, the condition $\kappa > \kappa_c$ is equivalent to $p < N^2/\kappa_c^2 r$.

We thus obtain a critical, or *maximal,* load: $p_c \approx N^2/\kappa_c^2 r$. This maximal load, or storage capacity, is the number of synfire links $L^\mu = (A^\mu, B^\mu)$ that can be stored and retrieved without destructive interference. If for instance the critical value κ_c is 5 and assuming that $r/N = 5 \times 10^{-3}$, we obtain $p_c = 8N$. We shall return soon to the issue of storage capacity (Section 6.3), and confirm this estimate by numerical methods.[6]

6.3 Superposition of chained synfire links (synfire chain)

To justify the assumption made in Section 6.2 that the set of active neurons at time t has a large overlap with a set of the form $\mathcal{B}_M$, we need to make further assumptions on the stored synfire links L^μ, $\mu = 1, \ldots, p$.

One may want to consider the case where stored links are *reflexive,* i.e., $A^\mu = B^\mu$, $\mu = 1, \ldots, p$. Note that synaptic weights are then *symmetric,* as in most associative-memory models. The analysis carried out in the previous section still applies:[7] storage/retrieval is efficient if $p < p_c$. If this condition is met, the activation of any set of the form $\mathcal{A}_M = \bigcup_{\mu \in M} A^\mu = \bigcup_{\mu \in M} B^\mu$ and of size r is now a *point attractor.*[8] Retrieving a stored link now takes the meaning of having the state converge to a point attractor, as in conventional associative-memory models. The number of different point attractors is the number of subsets of size $h \leq r/n$ of the set $\{1, \ldots, p\}$.

However, we argued in Section 2 that point attractors on the 1-ms time scale are biologically unrealistic, since cortical neurons fire at rates much lower than saturation. Therefore, the assumption $A^\mu = B^\mu$, $\mu = 1, \ldots, p$, appears to have little biological interest. We shall in Section 9.4 discuss another, more realistic, situation in which the present model comes close to classical associative-memory models.

Rather than taking $A^\mu = B^\mu$, we shall in this section assume that the p links are *chained* with each other as follows: $B^{\mu+1} = A^\mu$, $\mu = 1, \ldots, p-1$. The links are otherwise

[6] As mentioned, the calculations of the present section are performed for complete links only. Incomplete links will be dealt with in Section 7.

[7] The probability that a given pair (i, j) belongs to a given link $L^\mu = (A^\mu, A^\mu)$, that is, $i \in A^\mu$ and $j \in A^\mu$, is now $n(n-1)/N(N-1)$ rather than n^2/N^2; for parameter values of interest $(N, n \gg 1)$, these are not significantly different.

[8] The activation of a set $\mathcal{A}_M$ of size smaller than r is also a stable state, if the rest of the activity—which is random—is disregarded.

random, as in Section 6.2. This results in the storage of a *synfire chain,* of width n and length p. In cases of interest, $np \gg N$, and each neuron occurs many times in the chain.

Based on the results of Section 6.2, the analysis of the r-WTA dynamics in the synfire-chain case is straightforward: if the variance of the crosstalk term, $\eta_2 = rp(n^2/N^2)(1 + nr/N)$, is small enough with respect to the square of the chain width n^2, equivalently if $\kappa > \kappa_c$, equivalently $p < p_c$, the probability of failure of transmission along *any given link* of the chain is small. Thus, if a given pool B^μ is fully activated at time $t = 1$ (all its n neurons are active), all neurons in $B^{\mu+1}$ will, with high probability, be active at time $t = 2$. If, at $t = 1$, a number h of the p links of the chain are fully activated, these h activity packets will, from then on, propagate reliably along the chain, following its linear topology.

If on the other hand storage capacity is well exceeded, i.e., if κ is substantially smaller than κ_c, the activity, even when initially concentrated in a number of fully-active pools, spreads rapidly throughout the network and subsequently reverberates in a near-random fashion. The system is then overloaded: the random crosstalk component W_i in V_i dominates, because the two modes in the distribution of V_i largely overlap. This is essentially a case of *random* connectivity. We argue in Appendix D that the r-WTA dynamics in a random graph is characterized by lack of long-range order. Specifically, the *correlation range* of the dynamics on a random graph is short. In contrast, synfire transmission, which takes place if $\kappa > \kappa_c$, is characterized by *long-range order.*

When the system is around the critical point κ_c, one should expect waves to travel coherently for a positive but finite time and distance. This situation is similar to the one described in Abeles et al. (1993b) under the name *reverberating synfire chain,* i.e., 'synfire chains in which neurons participate in several positions.' A reverberating synfire chain displays a behavior that is intermediate between perfectly stable synfire transmission and random activity. In the reverberating chain described in Abeles et al. (1993b), a given neuron occurs on average *twice* in the chain. In the present model with $p \approx p_c$, each neuron occurs about $p_c n/N \approx 8n$ times. The overlap between links is thus considerable. For instance, with chain width $n = 100$ (see Section 8) and complete links, the overlap is 800.

6.3.1 Numerical experiments

Figures 2–5 show the results of simulations performed on a network of $N = 10000$ neurons, with chain width $n = 10$ and total activity $r = 500$. Note that the 'biological' relationship $r/N = 5 \times 10^{-3}$ does not hold here. The maximal number of coactive pools is $r/n = 50$. In Figure 2, the load, p, is 5700, corresponding to $\kappa = N/\sqrt{rp(1 + nr/N)} = 4.84$. For this value of κ there is a slight overlap between the two modes of the distribution of inputs V_i, so that synfire propagation is not *perfectly* stable: the system is slightly overloaded. In the simulation corresponding to Figure 2, $h = 27$ waves were triggered in the network at time $t = 1$. The figure shows the amplitudes of 10 of these waves as they travel along the chain for a period of time of length 100. Figure 3 shows the same simulation over 1000 time steps; by the end of this period, all of these 10 waves have died out.

(Insert Figures 2–4 around here)

As noted, a good indicator of the stability of synfire transmission is the correlation range of waves. Equivalently, one may measure the *average duration* of waves along the chain. To achieve a robust measurement of this duration, we first perform a moving (time and space) average on the amplitude of each wave (with a window of length 10); we then apply a threshold of $n/2$ to decide whether or not a wave is present at a given point of the chain, at a given time. All waves that occur during a long period of time are recorded in this way, and a histogram of wave durations is constructed. Figure 4 shows one such histogram, for the same load $p = 5700$, i.e., $\kappa = 4.84$. As expected, wave duration, which is then finite (because the system is slightly overloaded), obeys an exponential distribution. Figure 5 shows the mean of this distribution as a function of κ, when p is varied. One observes a sharp transition at a *critical value* $\kappa_c \approx 5.1$, corresponding to $p_c \approx 5000$. When $\kappa > \kappa_c$, i.e., $p < p_c$, wave duration is virtually infinite; synfire propagation then takes place in a perfectly reliable way.

(Insert Figures 5–6 around here)

Figure 6 shows the behavior of the network under the same load conditions as in Figures 2–4 ($\kappa = 4.84$), but this time with $h = 50$ waves (the maximal number) at time $t = 1$. These 50 *triggered* waves have all died out by the time $t = 2000$. However, they are partly replaced by other waves that form *spontaneously* at random positions along the chain. This results in a 'thermodynamic' equilibrium (remember that the dynamics is actually deterministic and that the only randomness so far is in the connections), whereby the number of waves fluctuates around a constant value.

(Insert Figure 7 around here)

The dependence of the average size of the equilibrium wave population on κ is shown in Figure 7 (the number of waves at time $t = 1$ in these simulations is $h(1) = 1$). Again, one observes a sharp transition at the critical value κ_c. Note that even when waves are perfectly stable, i.e., when κ is well above the critical value κ_c, the population size does not reach its theoretical upper bound $r/n = 50$. The reason for this is that a permanent loss of waves takes place at the end of the chain.[9]

Figure 8 illustrates this situation for $p = 3600$, i.e., $\kappa = 6.09$. This κ is well above the critical value, so that waves are perfectly stable and disappear only when they reach the end of the chain. In the simulation shown in Figure 8, only one wave is active at time $t = 1$; this wave, triggered at the beginning of the chain, is exactly in the middle of the chain at time $t = 1800$. Note also the very rapid increase of the number of *spontaneous* waves in the first few iterations. The total number of waves eventually fluctuates around 45, the combined result of a death process occurring at the end of the chain and a birth

[9]If the chain is cyclic—see below—the upper limit r/n *is* reached when $\kappa > \kappa_c$.

process occurring at random positions along the chain.

(Insert Figure 8 around here)

In sum, provided $\kappa > \kappa_c$, synfire transmission is a *stable* state of the dynamics (Property 1 in Section 1). As mentioned in Section 1, this notion of stability is somewhat different from the one used in traditional associative-memory models, for a wave that reaches the end of a chain dissolves there. Stability in the usual sense requires feedback. For instance, one may impose the *cyclic* boundary condition $A^p = B^1$; a biologically more plausible situation will be discussed in Section 9.4. Another important observation is that random activity in the network organizes itself spontaneously into synfire waves, providing a form of *background* synfire activity. As we shall shortly see (Section 6.3.2), the occurrence of this phenomenon depends somewhat on the setting of parameters. When this form of self-organization takes place starting from a *random* initial state, the distribution of inputs V_i, initially unimodal, progressively becomes bimodal. The synfire regime then acts as an *attractor*, whose attraction basin is the whole state space.

Note that it is in principle possible to have a number h of coactive waves larger than r/n if these waves are *partial*, i.e., if only r/h $(< n)$ neurons, on average, are active in any given pool. However, numerical simulations show that this situation is unstable. The r-WTA dynamics introduces strong *competition* between partially activated waves, so that, up to small fluctuations, waves are always fully activated.

6.3.2 Stability analysis in the mean

In order to gain further insight into the behavior of the network, we now consider the following argument, inspired from Abeles (1991, Chapter 7). Synfire propagation along a given chain is stable provided the activation of $n_1 < n$ neurons in each of a collection of pools B^μ, $\mu \in M$, elicits with high probability the activity of *more* than n_1 neurons in each of the pools $B^{\mu+1}$, $\mu \in M$. This condition can be analyzed fairly easily in the mean. Such an analysis consists in computing the *average*[10] number n_2 of neurons active in the pools $B^{\mu+1}$, $\mu \in M$, as a function of n_1; we shall term this function the *synfire transmission function,* and denote it $n_2 = \phi(n_1)$. If $\phi(n_1) > n_1$ we know that, in the mean, the amplitude of the waves will grow; propagation is then, on average, stable. The larger $\phi(n_1) - n_1$, the more reliable the propagation.

Denote by $F(v)$ the (cumulative) distribution function of the crosstalk term $W_{i,M}$ (Section 6.2). The input to a neuron i chosen randomly in the network obeys a *mixture* of two[11] probability distributions, corresponding to the cases $i \notin \mathcal{A}_M$ and $i \in \mathcal{A}_M$. The distribution function for this mixture is $G(v) = (N - hn)F(v) + hnF(v - n_1)$, since the deterministic component of the input has now amplitude n_1 rather than n. The r-WTA dynamics will, in the mean, activate the neurons whose inputs are larger than a *threshold* value v_0, satisfying $G(v_0) = N - r$. The fraction of these active neurons that fall in $\mathcal{A}_M$ is, in the mean, $hn(1 - F(v_0 - n_1))$. Therefore, each pool A^μ in $\mathcal{A}_M$ contains, on average,

[10] As usual, averaging is performed with respect to the random connectivity of the network, resulting from the randomness of the sets B^μ; the dynamics itself is deterministic, except for tie breaking.

[11] Ignoring again neurons i that belong to more than one pool $A^\mu = B^{\mu+1}$, $\mu \in M$.

$n_2 = n(1-F(v_0-n_1))$ active neurons at time $t+1$. We thus get: $\phi(n_1) = n(1-F(v_0-n_1))$, where v_0 is the number that satisfies $G(v_0) = N - r$, with $G(v)$ as defined above.

(Insert Figure 9 around here)

Figure 9 shows the curve $n_2 = \phi(n_1)$ for the usual parameter values $N = 10000, n = 10, r = 500, h = 50$, and κ ranging from 1 to 10 by increments of 1. To derive these curves, the distribution of the crosstalk term $W_{i,M}$ was approximated by a Gaussian of mean $\eta_1 = n^2/\kappa^2$ and variance $\eta_2 = \eta_1(1 + nr/N)$. As noted above, stability of synfire propagation requires that a large-enough portion of the curve $n_2 = \phi(n_1)$ be well above the identity line $n_2 = n_1$. We see from Figure 9 that this condition is satisfied, roughly, for $\kappa > 5$, confirming the results of the numerical experiments (Section 6.3.1). We also see that, for this setting of parameters N, n, r and h, as soon as κ is large enough for stable propagation to take place, the curve $n_2 = \phi(n_1)$ is above the identity line. This is true even for *small* values of n_1; small fluctuations are therefore able to grow into fully activated waves, which corresponds to the formation of spontaneous synfire waves described above.

(Insert Figure 10 around here)

Figure 10 shows a similar family of curves, this time with $h = 1$ rather than $h = 50$. These curves are the synfire transmission functions in the case of a *single* synfire wave; this single wave accounts for only 2% of the total activity in the network, the rest being randomly distributed. Figure 10 shows that the single-wave case is more favorable than the 50-wave case (Figure 9): all curves, in particular those for κ around κ_c, are shifted upwards. The reason for this lies in the *competition* inherent in the r-WTA dynamics; when less waves are active, less neurons compete.

As a consequence, if the system is slightly overloaded, it will 'react' to a large number of coactive waves by lowering this number until the remaining waves are more stable (by assumption they can never be *perfectly* stable, since the system is overloaded). This is observed for instance in Figure 6, where the size of the total wave population (the sum of the two curves), equal to 50 at time $t = 1$, decreases rapidly and eventually fluctuates around 27 (from about $t = 500$).

Since on the other hand waves form spontaneously if their number is 'too small' (see for instance Figure 8), we see that the operation of the system around the critical point $\kappa \approx \kappa_c$ is characterized by the *self-regulation* of the number of synfire waves present at any given time. There is an asymptotic $h^* = h^*(N, n, r, p)$ such that the system maintains the number of waves h approximately equal to h^*. The variation of h^* as a function of κ is precisely the curve shown in Figure 7 (in the case of a strictly feedforward chain). If the system is overloaded ($\kappa < \kappa_c$), h^* is less than the theoretical limit r/n ($h^* = r/n$ for $\kappa > \kappa_c$, if the chain is cyclic).

Figure 11 shows the synfire transmission curves for a different parameter setting. The curve $\phi(n_1)$ is now *below* the identity line for small values of n_1. In a case like this, spontaneous waves do not form as easily; spontaneous synfire activity may require additive noise (see Section 7). Alternatively, external stimulation may be used to trigger synfire activity.

(Insert Figure 11 around here)

Calculations were performed to investigate the stability conditions and estimate the critical value κ_c for a broad range of parameters. It was found that, for a domain that largely includes all parameter values of biological interest, the critical value κ_c does not vary much and stays around 5. As mentioned above, this results in a storage capacity $p_c \approx 8N$, under the assumption $r/N = 5 \times 10^{-3}$.

6.4 Multiple synfire chains

We just saw that with plausible parameter values p is of the order of N. This is a very large number (see Section 8) and the existence of a chain of such length in cortex is unlikely. More plausibly, the p stored links form a number of *separate* chains, which amounts to say that the relationship $B^{\mu+1} = A^{\mu}$ holds for most but not all μ, $\mu = 1, \ldots, p-1$. Let us assume, for simplicity, that p_1 chains are stored, all of same length p_2, so that $p \approx p_1 p_2$. We may again assume cyclic topology for each of the p_1 chains, if we wish to obtain true stability.

The condition for efficient storage is, as before, $p < p_c$, i.e., $\kappa > \kappa_c$. Storage capacity measured in terms of chains of length p_2 rather than links is p_c/p_2. When this capacity is respected, a stable state, or *attractor,* consists of a number $h \leq r/n$ of waves propagating along the p_1 chains, in any possible combination. In particular, a given chain may carry *several* waves at different positions (see Section 8 for numerical considerations).

A network consisting of a superposition of synfire chains respecting storage capacity meets Property 2 of Section 1, which is to say that the synfire patterns it supports are reliably reproducible; an appropriate *stimulation* of any of the stored chains at a given time will elicit stable synfire activity in that particular chain. As described in Abeles (1991, Section 7.5; see also Abeles et al. 1993b), stimulation can consist of increasing for a certain duration the *firing rate* of neurons in a few pools of the chain; this increase should be strong enough to produce, from time to time and with high enough probability, a *synchronous* activation of enough cells in any of the stimulated pools.

We proposed elsewhere (Bienenstock 1991) that synchronization of synfire propagation in coactive chains may be a mechanism for binding. An important requirement is that binding be dynamical, i.e., reversible: neural representations of mental entities should be allowed to co-occur either in a bound or an unbound state, or in various different binding configurations. In the present model however, coactive chains are synchronized *by construction.* Specifically, if pool i in one chain and pool j in another chain are both active at a given time t, pool $i+k$ in the first chain and pool $j+k$ in the second chain will both be active at the same time $t+k$ for $k > 0$. This model does *not* account for dynamical binding.

We saw however in Section 4 that in reality synfire transmission is reproducible only up to a certain amount of *temporal jitter* (Abeles 1991). Thus, the propagation in coactive chains should *not* be synchronized, unless appropriate synaptic coupling induces such synchronization. In Appendix B we modify the synchronous update scheme in such a way as to allow for temporal jitter in the firing of each neuron. The synchronization of coactive chains will then depend on the existence of synaptic coupling between them. We demonstrate then in Appendix C that *weak* coupling is enough to induce reliable synchronization, that is, to ensure the *stability* and *reproducibility* of composite patterns.

7 Behavior under noisy conditions

We saw in Section 6.2 that if storage capacity is respected and if activity at time t is concentrated on a union of stored links $\mathcal{B}_M = \bigcup_{\mu\in M} B^\mu$, the distribution of $V_i(t+1)$ is *bimodal;* the lower mode is contributed by cells i not in $\mathcal{A}_M = \bigcup_{\mu\in M} A^\mu$, while the higher mode is contributed by cells in $\mathcal{A}_M$.

These are favorable conditions for a *robust* operation of the regulatory system postulated in Section 2. Remember that this inhibitory system, implicit in the r-WTA dynamics, is likely to operate with a range of time constants somewhat shorter than the time constants involved in the excitatory feedback.[12] If the distribution of $V_i(t+1)$ is bimodal as described above, the task of the regulatory system is merely to keep the effective firing thresholds of all excitatory neurons within an interval, corresponding to the separation between the two modes of V_i. Large fluctuations in the total activity rate will be avoided.

This reasoning is valid only in the vicinity of a synfire attractor, when all the activity in the network is *already* concentrated on a set of neurons of the form $\mathcal{B}_M = \bigcup_{\mu\in M} B^\mu$ for some stored B^μs. If on the other hand activity is distributed randomly in the network, the distribution of V_i is unimodal, and one should expect somewhat chaotic behavior. Thus, the fairly rapid convergence of a *random* activity state to a synfire attractor, as described in Section 6.3 under the idealized r-WTA scheme, would, in reality, be considerably slower. However, our proposal (see Section 8) is that at any given time in the life of a normal functioning brain cortical activity is essentially all of synfire type.[13] Thus, in practice, the problem of 'bootstrapping' the system into the synfire regime does not arise.

In accord with these remarks (see also Abeles 1991, Section 7.4), we may conclude that, under normal operating conditions, the main source of perturbations in synfire propagation is the fluctuation of the individual membrane potentials, resulting in particular from input through synapses that are *not* part of stored chains. To investigate the effects of such perturbations on the dynamics, a random term is added to the input $V_i(t)$, independently for all i and t.

Figure 12 shows the results of such a simulation, with the usual parameters (see Section 6.3) but with $p = 3000$, hence $\kappa = 6.66$. In the absence of noise, this value of κ is well within the stability domain. Noise is uniformly distributed over the interval $[0, b]$, and the figure shows, as a function of t and for various values of b, how many of a collection of 50 waves present at time 1 are still present at time t. As expected, one observes a progressive degradation of the reliability of synfire transmission with increasing noise levels. Note that

[12] The term 'feedback' is used here because globally the population of excitatory neurons feeds back on itself: each neuron occurs many times in the collection of stored chains. This does not contradict the 'microstructural' fact that chains are strictly feedforward.

[13] The proportion of 'computation-related' waves as compared to spontaneous waves may however vary from one instant to another.

synfire transmission is perfectly stable for $b = 2$, corresponding to a situation where the amplitude of the fluctuations of the membrane potential of any neuron reaches 20% of the amplitude of the 'synfire input' to that neuron, coming from the previous pool in the chain. In Apendix D we contrast this situation with the high sensitivity to noise in case the network is a random graph—or for that matter an overloaded superposition of chains.

(Insert Figure 12 around here)

A different kind of noise that may affect the behavior of the system is 'quenched' noise, in the form of a *dilution* of the connectivity graph. Thus, consider a randomly-diluted connectivity graph, with dilution coefficient $\pi < 1$. Under the assumptions of Section 6.2, both the deterministic and the crosstalk components in $V_i(t+1)$ are now smaller by a factor π. Therefore, since storage capacity is computed by comparing $\sqrt{\eta_2}$, the square root of the average crosstalk, to the deterministic component, the dilution of the connectivity graph results in a drop of storage capacity by a factor $\sqrt{\pi}$. This prediction is confirmed by numerical experiments.

8 Numerical considerations

In this section we discuss various aspects of the model in view of a biologically plausible specification of parameter values. We stress again that these numerical values are mere orders of magnitude, as precise estimates would have little biological significance (see Section 1). Anatomical and physiological data reported in this section are widely accepted figures, quoted for instance in Abeles (1991).

Note first that stored chains (or braids—see Appendix A) could be arbitrarily wide; however, a *lower bound* for n can be derived. We saw in Section 3 that it takes on the order of 25 temporally overlapping EPSPs to trigger the firing of a cortical cell. The dilution coefficient within a *local* population of cortical pyramidal cells (see below) is estimated to be $\pi \approx 0.25$. Including in the model a number of patches of cortical tissue from distant but mutually interconnected areas (see Appendix A) should not result in a much different value for π. Thus, the lower bound for the width of a *chain,* where all afferents to a given neuron originate from a single pool of coactive neurons, is of the order of $n = 100$; this results in a multiplicity $m \approx 25$, as required. In the case of a *braid* (see Appendix A), where the afferents to a given neuron originate from neurons that fired at different times, n could be smaller. In the rest of this section and unless otherwise specified, the term *chain* will be short for chain *or* braid.

We shall say that a chain (or link, or braid) is *narrow,* if its width is equal to the lower bound $n = 100$: this is the smallest width that supports stable synfire activity. *Broad* synfire patterns will be obtained from narrow ones by composition, that is, synchronized coactivation (see below).

The number of coactive synfire links at any given time (Section 6.2) is at most r/n; an upper bound for this number is then $r/100 = 5 \times 10^{-3}N/100 = 5 \times 10^{-5}N$. The storage capacity (Section 6.3), corrected for the finite dilution π (Section 7), is $p_c \approx 8N\sqrt{\pi} \approx 4N$. Thus, a very small fraction of all stored links—about 1 in 10^5—are active at any given time. With these parameters and when $p = p_c$, the ratio of the *total* synaptic weight afferent to a given neuron as compared to the synaptic weight contributed by a *single* link on this neuron is on average $p_c n/N \approx 400$. Equivalently, one neuron participates in about 400 distinct stored synfire links. There is thus substantial overlap between stored memories.

A numerical value for p_2, the length of stored chains, could be taken to be the life time of a 'mental entity' in working memory, inasmuch as one wishes to specify such a numerical value (see Section 1 for reservations concerning such attempts). If we set $p_2 = 10^3$, corresponding to an entity lasting 1 s, the number of narrow synfire patterns of length p_2 that can be stored in a network of size N is βN, with $\beta \approx 4 \times 10^{-3}$.

In all likelihood, a large fraction of the cortical activity at any given time is *background* activity, not directly related to any specific task or mental process in which the brain may currently be engaged. The figure of 5×10^{-3} spikes/ms used here as mean cortical spiking rate should be viewed as resulting mostly from this background activity. Elevated firing rates related to specific computations or processing probably involve, at any given instant of time, a small fraction of all cortical neurons and therefore contribute little to the total activity.

We propose that background activity takes the form of synfire waves propagating randomly along stored chains. We moreover propose that these background synfire waves are independent Poisson-distributed events; there is little *synchrony* between background waves that propagate simultaneously along different chains, and there is little temporal *repetition* of synfire activity along any given chain. We shall propose in Section 9.3 that background, or *spontaneous,* activity may play an important role in the establishment and stabilization of narrow synfire connectivity.

We posit that the neural activity associated with a specific computation or processing at a given time differs from background activity in all three respects mentioned above. First, it involves only a small fraction of all chains that are active at any moment; as seen above however, a chain may be distributed on a number of distinct, spatially distant, cortical areas. Second, as proposed in Section 1, it is characterized by the *synchronization* of the synfire patterns that take part in it. Finally, it involves the *repeated* activation of chains, so that a chain that takes an active part in a computation is likely to carry, at any given time during this computation, *several* waves at different positions.

This latter proposal is not logically necessary; it is motivated merely by the observation that during the performance of certain sensory-motor tasks certain cortical cells fire reproducibly at rates higher than background (see discussion in Section 9.4). Data from Abeles et al. (1993a) support the assumption that all spikes that contribute to such elevated firing are part of synfire patterns. If we assume that cells rarely occur more than once in a chain of length $p_2 = 10^3$, the firing rate of a cell in an activated chain will be roughly equal to the number of waves currently traveling on this chain. Accordingly, we may expect the number of waves carried by an activated chain to be widely distributed, say between 1 and 50. Such repeated activation may in particular result from stimulating the chain by increasing the *rates* of activity of its first few pools (see Section 6.4); as long as this stimulation is in effect, waves form at random times—under constraints imposed by neuron refractoriness.

If we take, somewhat arbitrarily, 20 to be the average number of waves on such an activated chain of length 1 s, and if we assume, even more arbitrarily, that a given computation at a given time involves about 1% of all of cortical activity (of which 99% is then background), we obtain a number of computation-related coactive chains equal to $r/(n \times 20 \times 100) \approx 2.5 \times 10^{-8}N$. This is about 10^{-5} times the total number of stored chains of length 1 s, since the latter is βN with $\beta \approx 4 \times 10^{-3}$.

If we now set $N = 10^{10}$, which corresponds roughly to the collection of *all* cortical pyramidal cells, the following picture emerges.

A collection of randomly-overlapping *narrow* chains is stored in the network. Each such narrow chain carries a stable, i.e., reproducible, spatio-temporal pattern; this pattern is a synfire wave, defined on the 1-ms time scale, of variable duration, of the order of 1 s. Each pattern activates about 10^5 neurons per second, and the total store of narrow patterns is worth about 4×10^7 seconds, implying an overlap of about 400 patterns per neuron. Background activity takes the form of isolated waves, traveling randomly on narrow chains independently of each other. About 5×10^5 such waves are present in cortex at any given time; they result in a global mean activity rate of 5 spikes per second and per neuron; each stored narrow pattern is 'rehearsed' in this way on average once every 80 s. The activity related to a specific computation performed at a given time takes the form of the composition, or synchronized coactivation, of a highly-specific collection of about 250 narrow chains. (This last figure in particular is quite imprecise.) Such a coactivation of 250 narrow synfire chains in a synchronized, or *bound,* mode can be viewed as the elicitation of a *broad* synfire pattern, of width 2.5×10^4, involving 2.5×10^7 cells per second, i.e., on the order of one thousandth of all pyramidal cells. These cells fire at various rates, with an average of 20 spikes/s, which contributes no more than 1% to global cortical activity.

Attempting to quantify the behavior of the model beyond this point would hardly be meaningful. The number of composite entities that are possible from a mere *combinatorial* point of view is, for all practical purposes, infinite (see Property 5, Section 1).

There are however specific constraints on composition, which come from the requirements of stability and reproducibility of a composite construct, i.e., of a dynamically-bound broad chain. The stability and reproducibility of a composite entity, i.e., of a pattern of synchrony between waves (Appendix C), results from synaptic coupling between the chains that support these waves. Such coupling is established in the course of episodes of learning, when dynamically-bound composite constructs are inscribed in long-term memory. Only a very small fraction of the compositions that are possible *a priori,* i.e., from a mere combinatorial point of view, are actually stable and reproducible (Property 5, Section 1).

Note that saturating the network in terms of narrow synfire links entails the reinforcement of a total number of synapses of the order of $4Nnm$, that is, an average of $4nm$ synapses per neuron. With parameter values as above, $4nm \approx 4 \times 100 \times 25 = 10000$, roughly of the order of the number of synapses made or received by a pyramidal cell.

The distinction between narrow and broad synfire patterns is somewhat arbitrary. The model merely predicts the existence of an *optimal width* for stability of synfire activity. Patterns that are too narrow will be unstable because they do not meet the requirement of temporal overlap of converging EPSPs. Wide patterns are *composite* constructs, which are less stable, occur less frequently, but convey more meaning, i.e., interact with higher specificity with other patterns.

Note finally that we assumed in all of the above that cortical connectivity can be well-approximated by a randomly diluted graph. However, connectivity across large cortical distances is far from random. It is estimated that about one half of the cortical synapses are produced locally, while the other half originate from long-distance axons—mostly of cortical origin—that reach cortex from the white matter. Further, the connections established by long-distance axons are by no means scattered uniformly throughout cortex. On a gross anatomical scale, a given cortical area establishes mutual connections with only a limited number of other cortical areas; on a finer resolution, nearly all corticocortical projections display some spatial periodicity, or segregation into small patches—generally interdigitated with other projections.

We could of course limit the range of application of the model to small volumes of cortical tissue where connectivity can reasonably be considered random. For instance, the *local* axonal arborization emanating from a given pyramidal cell covers a volume of about .5 mm^3, which contains about $N = 2 \times 10^4$ pyramidal cells. The assumption of random dilution is probably a good approximation for such a small volume, and the dilution coefficient for this local population is, as mentioned above, $\pi \approx 0.25$. However, as explained in Appendix A, the model is meant to account for brain functions that are likely to involve multiple, mutually-distant, cortical sites. This requirement is not easily reconciled with the assumption that different synfire links are *independent* random subsets of the network. Similarly, stabilizing a given composite synfire pattern is possible only if the required circuitry is available.

These remarks however are chiefly quantitative. The qualitative predictions made in this paper are likely to be essentially correct, if the assumptions listed in Section 1 (items a-e) hold true for cortex.

9 Discussion

This paper proposes that the microstructure of cortical connectivity is a superposition of synfire chains. We presented a model based on this proposal, which shares several features with other models of cortex: persistence (stability) of states over the 1-s time scale; long-term memory; learnability; large storage capacity. Three main additional considerations motivate the synfire-superposition assumption: the compositionality of cognitive functions; the existence of reproducible tightly-timed spike patterns in cortex; the need for robust regulation of global cortical activity around low values.

9.1 Complex cortical events

We first discuss the role played by accurately-timed spike patterns in cortex. For the purpose of this discussion, we shall use the term *cortical event* to refer to any pattern of spikes that exhibits the following two features: (i) it is accurately timed on the 1-ms scale; (ii) it tends to reproduce. Further, we define—somewhat loosely—an *elementary* cortical event as one whose total duration does not exceed a few milliseconds, the typical conduction time in cortex; cortical events that extend over longer time spans will be termed *complex.* Based on evidence that there exist complex events in cortex, we shall argue that the cortical network is closer to a superposition of synfire chains than to a random graph. We shall also argue that random connectivity does not allow for robust regulation of global activity rates, whereas synfire connectivity does.

Cortex is distinguished by a rich recurrent connectivity, whereby a given pyramidal cell may be presynaptic to 10^4 or more other pyramidal cells. Of those, many may be physically distant from the presynaptic neuron, and from each other. Moreover, long-distance axons nearly always travel through the white matter and are myelinated, hence conduct with relatively high velocity and little time jitter (exceptions are Martinotti and superficial horizontal-axon cells, whose roles however are unclear). If one neglects the possibility of transmission failures at axonal bifurcations, it may be concluded from these observations that any spike emitted by one given pyramidal neuron will produce an accurately-timed constellation of EPSPs on a specific collection of cells, which are typically distributed across several cortical areas.

One may then entertain the hypothesis that allowing the delivery of signals with high precision in space and time constitutes one of the design principles of cortex. Although this is suggestive of the existence of cortical events as defined above, it does not strictly imply that such events should exist, for it takes many summed EPSPs to trigger the activity of a cortical cell.

The experimental approach used by physiologists to find cortical events, whether elementary or complex, consists of dividing the time axis into 1-ms steps and looking for reproducible events in this discretization. Elementary events, involving pairs or triplets of neurons, manifest themselves as small peaks in appropriately defined cross-correlation histograms. Complex events may be conveniently defined as sequences of synchronous activations—in the 1-ms discretization—of pools of neurons. Any such sequence, recorded at any time from any collection of neurons, is a legitimate candidate for a complex event; the sequence should however be *reproducible* in order to qualify as a cortical event. Reproducibility should be assessed by taking into consideration the statistical properties of the individual spike trains being recorded (see e.g. Abeles and Gerstein 1988). There is evidence, e.g. in the frontal cortex of awake monkeys, for the existence of complex cortical events (Abeles et al. 1993a), as well as for their possible role in behavior (Abeles et al. 1993b, 1993c). By conventional physiological standards, these data are scarce. However, it would be a misconception to infer from this scarcity that complex cortical events are unlikely to exist in any significant numbers. Simple statistical considerations (Abeles 1991) show that such events *should* be difficult to reveal experimentally, more so than elementary events, let alone than the elevated firing rates studied by single-cell methods.

Synfire chains provide, until otherwise demonstrated, the most parsimonious explanation for complex events in cortex. Note however that synfire chains may actually exist in cortex simply because the cortical connectivity graph is rich enough to include them as randomly-occurring *subgraphs.* One may then ask whether it is possible to account for the complex events observed in cortex *without* our rather strong assumption that cortical connectivity is a superposition of synfire chains. We consider in Appendix D, as a simple alternative to the synfire-superposition proposal, the hypothesis that cortex is a completely random graph, rich enough to include many synfire chains of given width and given multiplicity (Abeles 1991, Section 6.2). We argue that under this random-connectivity model *robust* regulation of the global activity level is essentially ruled out: the activity should be largely chaotic (see also Sompolinsky et al. 1988). We also argue that even under the unlikely assumption that there existed a mechanism that could stabilize the global activity level at all times, there would be little reproducibility in individual neural activities over any period of time significantly longer than a synaptic delay. Thus, even though cortex, under the assumption that it were randomly connected, might include, as subgraphs, large numbers of synfire chains of multiplicity equal to or larger than the firing threshold of cortical neurons, complex cortical events, as defined above, would not be observed; the only cortical events one would expect to see would be of the elementary type, i.e., short-lived.

9.2 The synfire-superposition hypothesis

These considerations—in addition to the ones pertaining to the compositionality of cognitive functions—lead us to seek a model of cortex in which connectivity departs from randomness in such a way as to (i) favor the reproducibility of transmission along specific multisynaptic pathways over durations well above a synaptic delay, and (ii) favor the regulation of global activity levels.

We proposed in this paper to envision the cortical network as a *superposition of synfire links* $L^\mu = (A^\mu, B^\mu)$, $\mu = 1, \ldots, p$. Under these conditions, if a given storage capacity is respected, i.e., if the load p is less than a given critical value p_c, *and* if the set of neurons active at time $t = 1$ has large enough overlap with a set $\mathcal{B}_M$ of the form $\mathcal{B}_M = \bigcup_{\mu \in M} B^\mu$, the distribution of membrane potentials at time $t = 2$ is *bimodal;* the lower peak is contributed by cells $i \notin \mathcal{A}_M = \bigcup_{\mu \in M} A^\mu$, while the higher peak is contributed by cells $i \in \mathcal{A}_M$ (Section 6.2).

If we further assume that the stored synfire links are *chained* with each other, thereby forming a collection of synfire chains, reliable and reproducible synfire transmission follows at once (Sections 6.3 and 6.4). We also argued (Section 7) that the bimodality of the distribution of membrane potentials creates favorable conditions for a robust operation of the regulation mechanism postulated in Section 2, thereby justifying the use of the conservative r-WTA dynamics.

We further demonstrated in Section 6.3.2 that synfire activity when the system is near criticality is characterized by the *self-regulation* of the population of waves: the network maintains the size of this population near a number h^*. This number itself undergoes a sharp transition at the critical load p_c.

The proposal that the microstructure of the cortical network is a superposition of synfire links is not unreasonable. However, the assumption that these links are chained together into long, regular, strictly feedforward chains of uniform width should be taken with a grain of salt. We shall argue shortly (Section 9.3) that synfire chains are most likely to develop in cortex as a result of a process of *natural selection*, where random spontaneous activity plays a central role. If so, one should expect chains to be irregular, both in their width and in the strength of their connections; these parameters were taken to be uniform in Section 5. However, it is readily seen that, provided there exist lower and upper bounds for these two quantities, the analysis performed in Sections 6.2 and 6.3 is still essentially valid. More debatable is the assumption that stored circuits are precisely *chains,* that is, feedforward one-dimensional networks. We shall touch upon this issue in the next two sections.

9.3 Hebbian development and learning: instructive or selective

We mentioned in Section 5 that the synfire-superposition assumption could be viewed as deriving from a form of Hebbian learning sensitive to the accurate time structure of patterns to be stored. We shall now discuss this proposal in more detail.

Synaptic modification is likely to proceed by small increments. If this is the case, many episodes of synaptic changes will be required to stabilize a given synfire link $L^\mu = (A^\mu, B^\mu)$. We derived in Section 8 a figure of 80 sec for the average time separating two spontaneous, or background, activations of a given synfire link. Even if this figure is inaccurate by one or two orders of magnitude, it indicates that each link is activated spontaneously quite frequently. This is consistent with the general observation that unused synaptic contacts in the central nervous system lose their efficacy rapidly, and eventually degenerate. An important implication is that learning in the proposed model cannot be envisaged as a purely *instructive* process, where the structure of a given stored chain is strictly determined by interactions with the environment. Rather, it should be construed as an *epigenetic* process, incorporating both environmental factors and self-organization mechanisms, the latter relying largely on spontaneous neural activity.

Self organization in this context means that the progressive *selection* of a given synfire pattern from the virtually infinite range of possible patterns stands in a relation of mutual positive feedback with the progressive reinforcement and stabilization of the synfire links that support this particular pattern. The more frequently a pattern is activated, the stronger the corresponding circuit will grow relative to competing circuits, and vice versa. It has been demonstrated elsewhere on the basis of a numerical study (Doursat 1991; Bienenstock 1991; Bienenstock and Doursat 1995) that such a process of self organization will take place reliably and under broad conditions as a result of accurate-time Hebbian plasticity complemented by a mechanism of *competition* for the making of synaptic contacts among converging and diverging fibres. When these two mechanisms are enforced in a network whose initial connectivity is random, synfire chains develop in a process of growth[14] by *accretion,* not unlike the growth of crystals.

This model of slow synfire-chain maturation by Hebbian plasticity is consistent with the view that the epigenesis of the brain follows principles of *natural selection,* as has been advocated by several authors (Changeux and Danchin 1976; Edelman 1988). When brain epigenesis is considered from this evolution-theoretic perspective, a simple notion of 'fitness' for cortical circuits emerges. Specifically, cortical pathways are selected for their ability to generate synchronous spikes in the fibres converging onto any given neuron (von der Malsburg 1981, 1987). This selection principle is a straightforward consequence of Hebbian plasticity acting with the dynamics of coincidence detection of neurons (Section 3) and with a mechanism of competition or constraint in synaptic growth. Anatomically, the fitness criterion for cortical circuits may be expressed as a degree of *cooperativity* between parallel multisynaptic pathways (see Section 4); it can be measured by the faithfulness of the circuit to a spatio-temporal consistency rule ensuring such cooperativity (equation 8 in Appendix A—see also Figures 13 and 14).

Chains are by no means the only circuits to meet this fitness criterion. Virtually any

[14] In this 'growth' process, the number of neurons and the total synaptic strength from and onto each neuron may remain nearly constant. Synapses are reinforced at the expense of other, competing, synapses.

graph with strong topological structure in low dimension[15] and with consistent conduction times along its connections is fitter, according to this criterion, than, say, a random graph (von der Malsburg 1987; see also von der Malsburg and Bienenstock 1986; Bienenstock and von der Malsburg 1987). Synfire chains however enjoy a particular status for at least two reasons: (a) being one-dimensional, they are the simplest possible type of such topological graphs; (b) they can be shown to develop spontaneously by a process of self-organization in an initially unstructured network. One-dimensional chains may therefore be thought of as 'building blocks' in the dynamics of neocortex. They may serve as 'units of selection' (Edelman 1988) on the time scale of ontogeny, and as 'units of computation' on the time scale of mental processing. We shall in Section 9.6 illustrate this proposal with an outline of a synfire-superposition model for visual cortical processing.

9.4 Synfire dynamics with strictly stable states

Although, as mentioned in Section 1, mental activity has often the character of a continually unfolding process rather than of a punctuated sequence of elicitations of stable states, it is nevertheless the case that in some situations a discrete item may be kept in working memory for a dozen of seconds or more. Much of recent electrophysiological research in the cortex of awake monkeys is carried out on animals trained to perform tasks involving such memory retention. The most striking finding in such experiments is the alteration—generally the elevation—of *firing rates* of cortical neurons (e.g. Miyashita and Chang 1988). These data are often taken to provide support to rate-coded associative-memory models (e.g. Amit and Tsodyks 1991). However, as noted in Section 8, there is evidence (Abeles et al. 1993a) that nearly all the spikes that contribute to such an elevation of firing rates during memory retention are part of various *long* complex cortical events, as defined in Section 9.1.

It is therefore of interest to investigate the conditions under which a model based on synfire superposition would also display classical associative-memory properties defined in terms of elevated firing rates (Hopfield 1982). This would clearly require that stored chains, instead of being strictly feedforward, include feedback links. Thus for instance, cyclic synfire chains, as described in Section 6.4, can be used to create stable activity patterns, specifically periodic attractors. However, the existence of cortical circuits of such regularity is doubtful. More plausibly, one may envisage connectivity graphs intermediate between feedforward synfire chains and random graphs with feedback. Such graphs would contain multiple, irregularly-arranged, feeddback loops, leading to a *reverberating-synfire* dynamics (see Section 6.3). A network including a superposition of such graphs would qualify as a classical associative-memory model, with attractors defined in terms of firing rates, and would also retain important features of the synfire model, namely the existence of complex events and the separation of the distribution of membrane potentials into two modes.

[15]This topology may or may not be consistent with the 3-D topology of the physical space in which the cortical network is embedded.

9.5 Issues of timing

We mentioned in Section 3 that, from the perspective of a general modeling of cortex, neither a strictly synchronous nor a completely asynchronous scheme would be fully satisfactory. As an example of an asynchronous *conservative* dynamics, consider the pair-exchange scheme used in lattice gas models. In such models, a $\{0,1\}$-valued variable $x_i(t)$ specifies whether or not a molecule is present at site i; a random pair of sites (i,j) is picked at time $t+1$, and the values of variables $x_i(t)$ and $x_j(t)$ are exchanged depending on whether a certain condition is met. A similar dynamics can be used for neural networks: at each time $t+1$ pick a random pair (i,j) such that $x_i(t) = 1$ and $x_j(t) = 0$ and swap their activities, i.e., set $x_i(t+1) = 0$ and $x_j(t+1) = 1$ with probability $P = f(V_j(t+1) - V_i(t+1))$, where f is a given increasing sigmoid-shaped function; as usual, the slope of the sigmoid may be viewed as an inverse temperature.

It is easily seen that under this asynchronous dynamics synfire propagation does *not* take place. Thus, suppose that the activity at time $t = 0$ is concentrated in h pools of a given chain; if these pools are fully active, $h \leq r/n$. Suppose for instance that pool B^2 is fully active at $t = 0$, while B^1, B^3 and B^4 are inactive. Exchanges first take place between B^2 and $A^2 = B^3$, yet as soon as B^3 has some activity in it, which occurs while B^2 is still partly active, exchanges start taking place between B^2 and $A^3 = B^4$ as well. Eventually an equilibrium distribution is reached, where the activity is spread uniformly over the entire chain.

We shall describe in Appendix B an *almost-synchronous* update scheme, where the neurons in pool μ of the chain fire almost simultaneously, around a time T^μ, $\mu = 1, \ldots$, a stochastic process which takes on a different set of values for each new wave that propagates on a given chain; we shall call T^μ the *intrinsic time* of the wave. In the case of a *single* chain, using the strictly synchronous update instead of the almost synchronous one boils down to the fairly innocuous assumption that T^μ is deterministically equal to μ, for $\mu = 1, \ldots$. We shall argue in Appendix B that the synchronous model is still legitimate in the case of multiple chains, provided synaptic interactions between stored chains are totally random. The intrinsic times of waves propagating on distinct chains are then statistically independent, which amounts to say that these waves are not synchronized.

We suggest here that with the introduction of Hebbian plasticity in the network the assumption that synaptic interactions between chains are totally random ceases to be valid. Specifically, one should expect to see synaptic interactions of the form studied in Appendix C, which tend to consistently *synchronize* the activity between chains. This is in line with the self-organization mechanism discussed in Section 9.3: when acting on chains that are already individually stabilized, this mechanism will bind these chains into broader, if less stable, chains, using whatever synaptic material is still available. Since the connectivity graph of cortex is diluted and because of additional timing constraints (the spatio-temporal consistency rule—see Appendix A), most couplings between chains will be weak, resulting in situations like the one studied in Appendix C.

9.6 Hierarchical/recursive cortical processing in the synfire-superposition model

In this section we outline a tentative synfire-superposition model for cortical processing, emphasizing the recursive/hierarchical aspect of compositionality (von der Malsburg 1981; Bienenstock and Geman 1993). We use visual perception to illustrate our proposal, which however applies to other cognitive functions as well. In general, compositional theories of perception—as contrasted for instance with models based on the use of prototypes—may be viewed as attempts to address in a principled way the fundamental issues of segmentation and invariance. Segmentation is achieved in the process of making explicit relationships at the lower levels of representation, while the interaction with stored higher-level compositions provides a mechanism for the 'labeling' of sensory input at various levels of abstraction.

Physiologically, the activity of neurons in V1, the first stage of visual cortical processing, is by all accounts strongly driven by visual input. Anatomically however, the cortical input to any neuron in V1—from V1 or from other cortical areas—by far outweighs direct visual input fed by geniculocortical afferents. We suggest that this abundant recurrent and descending connectivity subserves a synfire-type dynamics which may remain, with current investigating tools, largely hidden from the experimenter. Specifically, we suggest that each and every neuron in V1—as well as in the rest of neocortex—takes part in a large number of synfire chains. The task of these chains is to piece together the responses of single units into more complex patterns. In a first approximation, the response of an individual unit may be construed, in accord with the classical view, as an elevated firing rate, triggered by an elementary stimulus such as a small contour in appropriate retinal position and orientation. We propose that elevated firing in cortex is most of the time—perhaps always—part of synfire activity, along the lines indicated in Section 8. Synfire activity associated with the elevated firing of a cell in V1 may extend along multiple alternative routes, within V1 as well as into other cortical areas; it uses recurrent, ascending, as well as descending pathways.

As mentioned, there is considerable overlap between chains: each neuron in V1 may belong to hundreds of different chains.[16] These chains, having evolved and stabilized during earlier functioning of the system thanks to the epigenetic mechanism outlined in Section 9.3, support synfire patterns that are the neural counterparts of *mental representations* (Section 1) of various levels of complexity and abstraction.[17] At the lowest level, synfire waves in V1 may represent simple parts of visual stimuli such as straight or curved contour lines. These lines may themselves be parts of object boundaries, and may be assembled with each other to form a boundary description of a scene. Other local activity patterns may relate to surface patches and their attributes. The latter may be assembled into 2-D topologically-structured graphs (constructed from more elementary chain-like graphs—see Section 9.3) representing extended surfaces with smoothly-varying features. The two types of representations may be further combined, in a manner consistent with boundary-determined discontinuities, to form 3-D representations of objects.

[16] The figure of 400 derived in Section 8 can be taken as a coarse estimate for this number; just like other numerical estimates in this paper, it could be wrong by one or several orders of magnitude.

[17] The width of a synfire pattern may be a measure of its complexity.

In this model, there is a high degree of overlap and recursive embedding betwen representations of different levels. In much the same way as each neuron participates in many ‘narrow’ chains, each such chain is part of many broader chains and so on. This hierarchy provides for an essentially unbounded, recursive, combinatorics. Synfire chains of lower levels may be entirely contained within V1. Higher levels however require additional circuits, to inscribe in long-term memory specific combinatorial arrangements between lower-level synfire waves. This additional neural machinery may be located in cortical areas traditionally regarded as *integration* zones, but aptly renamed *convergence zones* by Damasio:

> “The two critical structures in the proposed architecture are the fragment record of feature-based sensory or motor activity, and the convergence zone, an amodal record of the combinatorial arrangements that bound the fragment records as they occurred in experience... There is no localizable single store for the meaning of a given entity within a cortical region. Rather, meaning is reached by widespread multiregional activation of fragmentary records pertinent to a stimulus, wherever such records may be stored within a large array of sensory and motor structures, according to a combinatorial arrangement specific to the entity.” (Damasio 1989)

9.7 Further speculations

As seen in Section 9.3, the dynamics of synfire waves on the relatively fast time scale of mental processing cannot be easily dissociated from the intertwined Hebbian dynamics of chains on the slower time scale of epigenesis and learning. To simplify, one may nevertheless envisage that in a first strictly self-organizational stage stable narrow chains start to grow independently of each other in large numbers, while in a second epigenetic stage complex patterns of couplings of various strengths develop between these essentially stabilized chains. These patterns of couplings will be determined by the history of functioning of the system, and will thus reflect events that took place in relation with interactions with the environment. Due to the sheer number of chains, to the dilution of the cortical connectivity graph, to the quenched randomness in this graph and in the conduction times, and to the near-simultaneous development of many couplings between chains, these couplings will be partly *frustrated,*[18] i.e., will contain constellations that stand in mutual conflict. Under these conditions, one should expect a high diversity of metastable patterns of binding. Such bindings will be highly sensitive to perturbations, or input, resulting in a behavior that could be described as *dynamical* binding.

Dynamical binding of synfire patterns may also rely, at least partly, on fast and reversible modification of synaptic transmission, as postulated by von der Malsburg (1981).

Remark that binding mechanisms proposed in the last few years generally rely on the phase-locking of periodic or almost periodic neural oscillators (e.g. Gray et al. 1989; Eckhorn et al. 1988; Shastri and Ajjanagadde 1993). The solution proposed here provides an alternative to the oscillation model, yet is not incompatible with it. Strictly

[18]The notion of frustration (Toulouse 1977) is used in statistical physics to describe a system including local interactions that cannot be satisfied simultaneously.

periodic oscillations in the synfire model would take place on cyclic chains (Section 6.4). We mentioned however that these highly-regular graphs are unlikely to develop in cortex. More plausible are partly-organized chains with some degree of feedback (Section 9.4). The findings about pseudo-periodic spiking activity in cortex, triggered in particular by appropriate visual stimulation (Gray et al. 1989; Eckhorn et al. 1988), might be reinterpreted as a manifestation of reverberating synfire activity in such circuits. It would also be of interest to apply to early visual cortical areas the techniques used by Abeles and Gerstein (1988) to reveal possible synfire patterns (see Section 9.6).

Note that the synfire solution to the binding problem provides, in the form of accurately-defined spatio-temporal patterns, a neural substrate richer than that afforded by mere temporal patterns (periodic oscillators). An immediate consequence, investigated in this paper, is that the firing of a given neuron can be part of many distinct synfire patterns, each of which is a stable reproducible entity with specific interactions with other such entities. A flexible medium of this sort is probably necessary for our brains to carry out the compositional operations that are manifested in perception, language, reasoning, metaphorical thinking, etc.

To further elaborate, remark that by introducing additional synfire links *within* a circuit that was initially a one-dimensional chain, a complex structure can be created. This process may be likened to that of protein folding: starting from a strongly-linked one-dimensional chain of amino acids, a complex three-dimensional protein—with highly specific affinities for other proteins—is created by introducing additional weak links in the chain, resulting in its folding in space. Similarly, an initially one-dimensional synfire chain may be thought of as 'folded' on itself as a result of the introduction of weaker links between neurons or pools that are distant from each other in the chain (see Sereno 1991 for related metaphors). The binding with each other of two or more synfire chains that have been folded in this way may be viewed as the establishment of a *structure-preserving correspondence* between complex entities, a type of operation that is characteristic of most—perhaps all—high-level cognitive functions.

Acknowledgements: Thanks to Stuart Geman for many insigthful suggestions, and to both Stuart Geman and Moshe Abeles for a careful critical reading of an early version of this paper.

References

Abeles, M. 1982. *Local cortical circuits: An electrophysiological study,* Springer-Verlag, Berlin.

Abeles, M., and Gerstein, G.M. 1988. Detecting Spatiotemporal Firing Patterns Among Simultaneously Recorded Single Neurons. *J. Neurophysiol.*, 60(3):909–924.

Abeles, M. 1991. *Corticonics: Neuronal circuits of the cerebral cortex.* Cambridge University Press, Cambridge UK.

Abeles, M., Bergman, H., Margalit, E., and Vaadia, E. 1993a. Spatiotemporal Firing Patterns in the Frontal Cortex of Behaving Monkeys. *J. Neurophysiol.*, 70(4):1629–1638.

Abeles, M., Vaadia, E., Bergman, H., Prut, Y., Haalman, I., and Slovin, H. 1993b. Dynamics of Neuronal Interactions in the Frontal Cortex of Behaving Monkeys. *Concepts in Neuroscience,* 4:131–158.

Abeles, M., Prut, Y., Bergman, H., Vaadia, E., and Aertsen, A. 1993c. Integration, Synchronicity and Periodicity, in Brain Theory—Spatio-Temporal Aspects of Brain Function, A. Aertsen, Ed., Elsevier Science, Amsterdam, pp. 149–181.

Amit, D.J. 1989. *Modeling Brain Function.* Cambridge University Press.

Amit, D.J.,and Tsodyks, M.V. 1991. Quantitative study of attractor neural networks retrieving at low spike rates, *Network,* 2:259–274.

Barlow, H.B. 1972. Single units and sensation: a neuron doctrine for perceptual psychology? Perception, 1:371–394.

Bienenstock, E. 1991. Notes on the growth of a composition machine, in *Proceedings of the Royaumont Interdisciplinary Workshop on Compositionality in Cognition and Neural Networks—I,* D. Andler, E. Bienenstock, and B. Laks, Eds., pp. 25–43.

Bienenstock, E. 1992. Suggestions for a Neurobiological Approach to Syntax, in *Proceedings of the Royaumont Interdisciplinary Workshop on Compositionality in Cognition and Neural Networks—II,* D. Andler, E. Bienenstock, and B. Laks, Eds., pp. 13–21.

Bienenstock, E., and Doursat, R. 1995. *The Hebbian Development of Synfire Chains.* In preparation.

Bienenstock, E., and Geman, S. 1993. Compositionality in Neural Systems. Preprint, to appear in The Handbook of Brain Theory and Neural Networks, M.A. Arbib Ed., Bradford Books/MIT Press.

Bienenstock, E., and von der Malsburg, C. 1987. A Neural Network for Invariant Pattern Recognition *Europhys. Lett.* 4(1):121–126.

Braitenberg, V. 1977. *On the Texture of Brains.* Springer, Berlin.

Burns, D.B., and Webb, A.C. 1978. The correlation between discharge times of neighboring neurons in isolated cerebral cortex. *Proc. R. Soc. Lond. [Biol.]* 203:347–360.

Changeux, J.P., and Danchin, A. 1976. Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature* 264:705–712.

Damasio, A. R. 1989. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition, *Cognition* 33:25–62.

Diesmann, M., Gewaltig, M.O., Aertsen, A., and Abeles, M. (in preparation). Conditions for stable propagation of synchronous activity in model cortical networks

Doursat, R. 1991. *Contribution à l'étude des représentations dans le système nerveux et dans les réseaux de neurones formels.* Ph.D. Thesis, Université Paris VI, Paris, France.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H.J. 1988. Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biol. Cybernetics*, 60:121–130.

Edelman, G.M. 1988. *Neural Darwinism: The Theory of Neuronal Group Selection.* Basic Books, New York.

Glauber, R.J. 1963. Time-dependent Statistics of the Ising Model. *J. Math. Phys.* 4:294–307.

Gray, C M., König, P., Engel, A. K., and Singer, W. 1989. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties, *Nature*, 338:334–337.

Hebb, D.O. 1949. *The Organization of Behavior.* Wiley, New York.

Hopfield, J.J. 1982. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proc. Natl. Acad. Sci. U.S.A.,* 79:2554–2558.

Langacker, R.W. 1987. *Foundations of Cognitive Grammar, Vol. 1, Theoretical Prerequisites.* Stanford University Press.

Little, W.A. 1974. The existence of persistent states in the brain. *Mathematical Biosciences*, 19:101–120.

Miyashita, Y. and Chang, H.S., 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature,* 331:68–70.

Sereno, M.I. 1991. Four Analogies Between Biological and Cultural/Linguistic Evolution, *J. Theor. Biol.* 151:467–507.

Shastri, L., and Ajjanagadde, V. 1993. From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings, *Behavioral and Brain Sciences*, 16:417–494.

Softky, W.R., and Koch, C. 1992. Cortical cells should fire regularly, but do not. *Neural Computation,* 4:643–646.

Sompolinsky, H., Crisanti, A., and Sommers, H.J. 1988. Chaos in random neural networks. *Phys Rev. Lett.* 61:259–262.

Toulouse, G. 1977. Theory of the frustration effect in spin glasses: I *Commun. Phys.* 2:115–119.

von der Malsburg, C. 1981. *The correlation theory of brain function.* Internal report 81-2, Max-Planck Institute for Biophysical Chemistry, Dept. of Neurobiology, Göttingen, Germany.

von der Malsburg, C. 1987. Synaptic Plasticity as Basis of Brain Organization. In *The Neural and Molecular Bases of Learning,* J.P. Changeux and M. Konishi, Eds., Wiley and Sons, pp. 411–432.

von der Malsburg, C., and Bienenstock, E. 1986. Statistical coding and short-term synaptic plasticity: A scheme for knowledge representation in the brain. In *Disordered Systems and Biological Organization,* E. Bienenstock, F. Fogelman, and G. Weisbuch, Eds., Springer-Verlag, Berlin, pp. 247–272.

APPENDIX

A Synfire braids

Here we reexamine the definition of a synfire chain from the point of view of the functional anatomy of cortex, bearing in mind the five properties of mental representations listed in Section 1. We extend the notion of a synfire chain, whose connections all have identical conduction times, into that of a *synfire braid,* whose connections have non-uniform conduction times. We argue—without providing detailed proofs—that the analysis carried out in the paper for a superposition of synfire chains extends to a superposition of synfire braids.

Remark first that neural activity related to the performance of a cognitive function often manifests itself simultaneously in *several* anatomically defined cortical areas. Inasmuch as specific functional roles can be assigned to these different areas—different sensory modalities for instance are localized in distinct regions of cortex—the question arises of how such distributed neural activity, corresponding to different fragments, aspects, or features of a given 'entity,' is assembled into a coherent whole. A given entity that is distributed in this anatomical sense may have a stable and reproducible semantic content in the sense of Section 1, as evidenced in its manifestations in conscious mental processes. In the present model we are therefore dealing with populations of neurons extending over several distinct cortical areas (see also Damasio 1989; Bienenstock 1991).

We also mentioned in Sections 2 and 7 that the regulation of global activity levels, which is implicit in the r-WTA dynamics, is likely to be efficient and robust provided inhibitory feedback acts somewhat faster than excitatory feedback. We argued that this was likely to be the case, since about half of the excitatory feedback is mediated by long-distance axons, while inhibitory connections are strictly local. There is moreover evidence that long-distance cortico-cortical excitatory connections traveling through the white matter are necessary to maintain a non-zero level of activity in cortex (Burns and Webb 1979; see discussion in Abeles 1991, Section 5.4.3). Again, this reasoning is valid only if the network under consideration is distributed over several distinct cortical areas.

Yet synfire transmission as described in Section 5 requires that axonal conduction times be *identical* for all the connections from a given pool to the next. This condition is met if the N neurons are all contained in a small volume of cortical tissue; connections are then local, and conduction times are significantly shorter than synaptic delays, which are all of the order of 1 ms. The condition however is violated if both local and long-distance connections are included, since conduction times plus synaptic delays are then scattered across a broad range of values, say from 1 to 10 ms.

On the basis of these considerations, Abeles (1991, Section 7.4.5) suggests that "it is possible that synchronous transmission is carried only locally, whereas transmission across large distances is carried by asynchronous volleys." Thus, synfire chains as proposed in Abeles (1982, 1991) are meant to be *local* cortical circuits.

Here however (see also Doursat 1991; Bienenstock 1991), we propose to *extend* the notion of a synfire chain so as to accommodate a range of different conduction times along different connections. As mentioned in Section 4, the mechanism of synfire transmission along a given chain relies on the temporal overlap of a large-enough number of EPSPs on the somato-dendritic membrane of every neuron in the chain. Such overlap occurs if the converging impulses giving rise to the EPSPs on a given neuron are well-synchronized. Synfire chains as described in Abeles (1982, 1991), consisting of discrete sequences of pools of neurons with feedforward connections of identical conduction times between successive pools, are the simplest of all neural circuits that guarantee such synchrony of converging spikes on each neuron. We shall now define a broader class of networks, with the same property of synchrony of converging spikes, but where conduction times are distributed over a range of values.

Consider a strictly feedforward network, i.e., a network without loops, with different conduction times on different connections. We assume that the conduction times are real-valued, distributed over an interval $[1, \Delta t_M]$, where Δt_M is, say, 10 ms. The network is called a *synfire braid* (Doursat 1991; Bienenstock 1991) if it satisfies the following spatio-temporal *consistency rule:* given any two neurons i and j in the network and given several alternative paths from j to i, the *total* conduction time from j to i is nearly the same along each of these paths. The paths under consideration are generally multisynaptic, and the condition should hold only for a limited synaptic depth, say of the order of 5. A simple example, of synaptic depth 2, is as follows. If j is presynaptic to k and l and if the latter are both presynaptic to i, then—denoting by Δt_{xy} the conduction time of the connection from neuron y to neuron x—the following relationship should hold:

$$\Delta t_{ik} + \Delta t_{kj} \approx \Delta t_{il} + \Delta t_{lj}. \tag{8}$$

Equation 8 may actually be taken as the definition of a synfire braid.

Consider first the idealized case of a braid in which equation (8) is satisfied *exactly* rather than approximately; the consistency rule then holds for paths of arbitrary synaptic depth. In this case, the Δt_{ij} may as well be taken to be multiples of a basic time unit, say 1 ms. Conduction times thus take their values in the set of integers $\{1, \ldots, \Delta t_M\}$, and the braid is actually, like a chain, a discrete sequence of pools of neurons. To each neuron i in the braid there corresponds a discrete time t_i, measured with respect to an arbitrary origin, and the consistency rule can be written simply in terms of these t_i as follows:

$$t_i = \Delta t_{ij} + t_j. \tag{9}$$

We shall refer to a synfire braid of this type as a *discrete-time* braid; a fragment of such a braid is shown schematically in Figure 13.

(Insert Figure 13 around here)

Note that equation 9 (hence equation 8) is trivially satisfied in the case of a synfire chain, with $\Delta t_{ij} = 1$ for any pair i and j such that there is a connection from j to i. Thus, a synfire chain is a particular instance of a discrete-time braid. We shall say that a braid has (converging) multiplicity m if every neuron in it receives at least m connections. It is readily seen that if we adapt either of the two synfire-chain models mentioned in Section

4 (Abeles 1991; Abeles et al. 1993b) to accommodate inhomogeneous discrete conduction times, we obtain essentially the same behavior for a discrete-time braid as for a synfire chain with same multiplicity. A triggering pattern in such a braid is distributed over Δt_M adjacent pools; it consists of the activation of a number of neurons at least equal to a given *ignition threshold,* with the proper timing relationships between these activations.

The r-WTA dynamics in a network with discrete conduction times is defined as before (Section 5), the input to neuron i being now:

$$V_i(t+1) = \sum_{j=1}^{N} \sum_{t'=t-\Delta t_M+1}^{t} \mathbf{1}_{\{t'+\Delta t_{ij}=t+1\}} w_{ij} x_j(t'). \tag{10}$$

Equation 10 is a straightforward generalization of Equation 1 to the case of inhomogeneous integer-valued conduction times.

The case of a network made of a superposition of discrete-time braids can be studied under the r-WTA dynamics along the same lines as the analysis performed in Section 6 for a superposition of chains. A *random* discrete-time braid of width n, length p and multiplicity m in a network of N neurons can be defined by the following two-step procedure: (i) draw independently p random sets $B^1, \ldots, B^p$ of size n each, and assign time $t_i = \mu$ to every neuron $i \in B^\mu$, $\mu = 1, \ldots, p$; (ii) to every neuron in B^ν, connect m randomly drawn neurons in $\bigcup_{\mu=\nu-\Delta t_M}^{\nu-1} B^\mu$, and define the conduction times for these connections according to equation 9. The usual cyclic boundary condition (Section 6.3) can be obtained, if desired, by taking μ and ν modulo p.

Just like for a chain (Section 5), the storage of a synfire braid results from the Hebbian modification of the synapses between the neurons that take part in the synfire pattern being stored. The plasticity rule is sensitive to the accurate time structure of the synfire pattern: only those connections whose conduction times happen to match this structure are reinforced.

With these straightforward adaptations, the behavior for superimposed braids is the same as for superimposed chains. Specifically, storage/retrieval is efficient provided the network is not overloaded. Note that the deterministic component of $V_i(t+1)$ is now m instead of n; the crosstalk component is easily seen to be $\eta_1 = rpnm/N^2$ instead of $\eta_1 = rpn^2/N^2$. As a result, storage capacity—when compared to the expression obtained for a chain—becomes multiplied by a factor m/n.

Consider now the more realistic case of a braid in which equation 8 is only *approximately* satisfied. The condition for synfire transmission to take place is that the scatter of the arrival times of converging spikes on any neuron in the braid be smaller than the duration of EPSPs; the latter will be denoted by τ. The braid can no more be partitioned into a discrete sequence of pools of neurons; the time t_i assigned to neuron i is now a *real* number, and the consistency condition reads:

$$|t_i - t_j - \Delta t_{ij}| \ll \tau, \tag{11}$$

for each pair of neurons (i, j) with a connection from j to i. This is illustrated in Figure

14.

(Insert Figure 14 around here)

The discrete-time r-WTA dynamics can still be used in a network with real-valued conduction times, provided we define the input to neuron i appropriately. One possible definition is:

$$V_i(t+1) = \sum_{j=1}^{N} \sum_{t'=t-\Delta t_M}^{t} \mathbf{1}_{\{t'+\Delta t_{ij} \in [t,t+1]\}} w_{ij} x_j(t'). \tag{12}$$

In this equation, the allowed scatter in the arrival times of converging spikes is one time unit, i.e., 1 ms, which is smaller than τ, the duration of an EPSP. One could allow a broader scatter of arrival times, yet this would make it necessary to include a refractoriness mechanism.

A *random* continuous-time braid may be defined as follows: (i) choose at random a collection of q_1 neurons from the N neurons in the network and assign to each neuron i in this collection a random time t_i distributed uniformly over a given interval $[0,T]$; (ii) choose at random, from this set of q_1 neurons, q_2 ordered pairs (i,j) such that $0 \leq t_i - t_j \leq \Delta t_M$ and connect j to i with a random conduction time satisfying equation (11). This results in a synfire braid of multiplicity $m \approx q_2/q_1$. Note that T is the equivalent, for a continuous-time braid, of the length p of a chain or discrete-time braid; q_1/T is the average number of active neurons per unit of time, we may denote $n = q_1/T$, and call n the *width* of the braid.

With these definitions, the analysis of the dynamics of a network made of a superposition of continuous-time braids is essentially the same as for discrete-time braids, with same storage capacity.

In sum, both the notion of a synfire chain and the r-WTA dynamics can be extended to accommodate inhomogeneous conduction times; the results derived in Section 6 are still valid under this extension. As a result, the network under study can be distributed over several interconnected cortical areas.

B Almost-synchronous dynamics

We first provide a brief summary of the discussion of the stability of synfire transmission presented in Abeles (1991, Section 7.4). We then describe, based on this discussion, a refinement of the synchronous discrete-time dynamics, in a spirit of a linear-stability analysis of perturbations of the firing times of individual neurons.

Consider a synfire chain $B^0, B^1, \ldots,$ of width n and multiplicity m. During transmission along a link $L^\mu = (B^{\mu+1}, B^\mu)$ in the chain, a given cell i in $B^{\mu+1}$ discharges when its membrane potential reaches the *firing threshold,* under the effect of a volley of m temporally overlapping EPSPs, received from neurons in B^μ active about 1 ms earlier. This summed EPSP is superimposed on *fluctuations* of the membrane potential of i, due mainly to synaptic inputs from neurons that are not part of the chain under consideration. These fluctuations induce variability in the time of firing of i. However, provided the multiplicity m is somewhat higher than the number of EPSPs it takes to bring the membrane potential to threshold, such fluctuations will *not* accumulate to such an extent as to disrupt synfire transmission; this is true irrespective of the length of the chain. Indeed, if the condition just mentioned is satisfied, the timing of the firing of cell i in $B^{\mu+1}$ is determined, roughly, by the *mode* of the arrival times of the m EPSPs. The mode is nearly identical for all neurons in $B^{\mu+1}$, and its variance is considerably smaller than the variance of the individual arrival times.

Another source of temporal jitter in synfire transmission is the stochasticity of transmitter release at each synaptic junction. This however generally contributes significantly less than the fluctuations of the postsynaptic membrane potentials.

In accord with this short discussion, the firing time of a neuron in $B^{\mu+1}$ can be viewed as the sum of two random variables: the first, common to all neurons in $B^{\mu+1}$, is the *mean* of the firing times of neurons in B^{μ} plus the common delay, taken to be 1 ms; the second is an independent random variable, resulting from the fluctuations of individual membrane potentials in $B^{\mu+1}$.

Denote by $t_1^{\mu}, \ldots, t_n^{\mu}$ the firing times of the n neurons in B^{μ} and assume that synfire transmission took place up to pool μ, i.e., that the scatter of these times is small with respect to 1 ms. The firing times of neurons in $B^{\mu+1}$ are then given by:

$$t_i^{\mu+1} = 1 + \frac{1}{n}\left(\sum_{j=1}^{n} t_j^{\mu}\right) + \Omega_i(\mu+1), \tag{13}$$

for $i = 1, \ldots, n$, where the $\Omega_i(\mu+1)$ are independent identically distributed (i.i.d.) random variables, of mean 0 and standard deviation σ, with $\sigma \ll 1$.

Equivalently, if we define $T^{\mu} = 1 + (1/n)\sum_{j=1}^{n} t_j^{\mu-1}$, the n neurons in B^{μ} fire at times $t_i^{\mu} = T^{\mu} + \Omega_i(\mu)$, where the $\Omega_i(\mu)$ are i.i.d. random variables as specified above. It is readily seen that: the stochastic process T^{μ} is Markov; for a given μ the standard deviation of T^{μ} is $\sigma\sqrt{\mu/n}$; if we assume that all neurons in B^0 fired at time 0, i.e., $t_i^0 = 0$, $i = 1, \ldots, n$, the expected value of T^{μ} is μ.

This model accounts for the two main characteristics of synfire transmission (Section 4): (i) *stability* of transmission, or 'resynchronization' (Abeles 1991), which means that the scatter in the firing of the n neurons in any pool μ is constant and does not increase with μ (the standard deviation is equal to σ in this model); (ii) *jitter* in T^{μ}, the global firing time of pool μ (the variance of T^{μ} grows linearly with μ).

Synfire transmission along a given chain is thus *almost-synchronous.* In the present paper, we used a *strictly-synchronous* dynamics, where the states of the neurons in pool μ are updated precisely at time μ, where μ is an integer. This is legitimate if it is understood that μ is an *approximation* of a real-valued stochastic process T^{μ} of mean μ. Different waves, along different chains or along the same chain, will produce different *realizations* of the stochastic process $\{T^{\mu}\}_{\mu=1,\ldots}$, which we shall call the *intrinsic time* of the wave.

We argue now that this approximation is legitimate not only in the case of a single chain, as modeled by equation 13, but also in the more general case of a network including a number of randomly-overlapping chains, as studied in Sections 5 and 6. In such a network, the interactions between the different chains are *random;* such interactions are accounted for by the noise term $\Omega_i(\mu+1)$ in equation 13. This random variable, *of mean zero,* represents fluctuations in the firing time of neuron i resulting from *crosstalk* EPSPs, which are statistically independent from the volley of EPSPs originating from cells in B^μ. As long as storage capacity is respected, i.e., $p < p_c$, the number of such crosstalk EPSPs impinging on neuron i shortly before time $T^{\mu+1}$ is small with respect to n (m if the chain is incomplete). In other words the size of the fluctuation of the membrane potential of i is small with respect to the summed EPSP caused by the volley of synchronous spikes emitted from pool B^μ at time T^μ; therefore, the assumption $\sigma \ll 1$ is legitimate.

In short, the strictly-synchronous update scheme used in the paper provides a first-order approximation to the more realistic almost-synchronous dynamics described here, provided the relationships between stored chains are random, as is the case in the situations studied in Section 6, and provided the storage capacity of the network is respected. Synfire waves propagating simultaneously along distinct chains are then, to a first approximation, independent of each other; in particular, the intrinsic times T^μ relative to distinct waves propagating on distinct chains are independent. These waves are *not* synchronized.

C Dynamical binding

We study here, under the almost-synchronous model described in Appendix B, a situation where the synaptic interactions between two different chains are *not* random, hence induce a perturbation in the firing time that can no more be accounted for by the sole zero-mean noise term $\Omega_i(\mu)$. Specifically, we consider two chains $B^0, B^1, \ldots$ and $C^0, C^1, \ldots$, both of width n, such that, for all μ, synaptic contacts exist from B^μ to $C^{\mu+1}$ and from C^μ to $B^{\mu+1}$. This regular pattern of synaptic interaction tends to induce *synchrony* between waves propagating simultaneously on the two chains.[19] The biological underpinnings of this situation are discussed in Section 9.5.

We are interested in situations where the number of synaptic inputs received by any neuron in one chain from neurons in the other chain is small with respect to the chain width, n. For simplicity, we take this number to be equal to 1 for every neuron. To further simplify the calculations, we assume that neuron i in pool B^μ makes a synaptic contact with neuron i in pool $C^{\mu+1}$ and that neuron i in pool C^μ makes a synaptic contact with neuron i in pool $B^{\mu+1}$. We wish to demonstrate, using a simple asymptotic argument in the length of the chains, that such weak synaptic coupling is enough to ensure the stability of the synchrony of synfire waves propagating along the two chains.

As in Appendix B, denote by $\{t_i^\mu\}_{i=1,\ldots,n}$ the firing times of neurons in B^μ, and let $\{s_i^\mu\}_{i=1,\ldots,n}$ be the firing times of neurons in C^μ. Under the assumptions made above, and if we further assume that the first pools of the two chains are activated simultaneously, the firing times obey the following equations:

$$t_i^{\mu+1} = 1 + \frac{1}{n+1}\left(s_i^\mu + \sum_{j=1}^{n} t_j^\mu\right) + \Omega_i(\mu+1), \tag{14}$$

$$s_i^{\mu+1} = 1 + \frac{1}{n+1}\left(t_i^\mu + \sum_{j=1}^{n} s_j^\mu\right) + \Omega_i'(\mu+1), \tag{15}$$

[19] One could have studied an equivalent situation where it is a non-random *overlap* between chains that induces synchrony. For instance, pools B^μ and C^μ may, for each μ, share a small number of neurons.

for $i = 1, \ldots, n$, where the $\Omega_i(\mu+1)$ and $\Omega'_i(\mu+1)$ are i.i.d. random variables of mean 0 and standard deviation $\sigma \ll 1$.

As in Appendix B, define $T^{\mu+1} = 1 + (1/n)\sum_{i=1}^{n} t_i^{\mu}$ and similarly let $S^{\mu+1} = 1 + (1/n)\sum_{i=1}^{n} s_i^{\mu}$; T^{μ} and S^{μ} are the intrinsic times of waves along the B chain and the C chain. Summing equations 14 and 15 over i and dividing by n shows that T^{μ} and S^{μ} form a joint Markov process; for a given μ, the expected values of T^{μ} and S^{μ} are μ if all neurons in B^0 and C^0 fired at time 0. Due to the coupling between equations 14 and 15, these intrinsic times are *not* independent. Subtracting the summed equations from each other, one obtains:

$$T^{\mu+2} - S^{\mu+2} = \frac{n-1}{n+1}(T^{\mu+1} - S^{\mu+1}) + \Omega''(\mu+1), \tag{16}$$

where the random variable $\Omega''(\mu+1)$ is Gaussian (or nearly so) of mean zero and standard deviation $\sigma\sqrt{2/n}$. The Markov process $T^{\mu} - S^{\mu}$ is a classic Ornstein-Uhlenbeck process. This process achieves an asymptotic *equilibrium* distribution, which depends on σ: the smaller σ, the closer $|T^{\mu} - S^{\mu}|$ remains to the origin. This demonstrates that synchronization is *stable.*

In conclusion, weak coupling is enough to ensure the stability of synchronous propagation on two chains. This provides a *binding* mechanism: provided the two chains are coactivated with appropriate initial timing, the intrinsic times of the waves will be nearly equal. A similar situation is studied numerically in Abeles et al. (1993b). In Section 9.7 we argue informally that in a network containing a large number of chains that are weakly bound with each other in partly conflicting ('frustrated') patterns, binding will manifest itself as a *dynamical* property.

D The random-graph model of cortical connectivity

Here we envisage the random-graph model for cortical connectivity (see Section 9.1) as a simple alternative to the synfire-superposition model; we ask whether the random-graph model can plausibly account for the existence of *complex* cortical events. Specifically, assuming that the connectivity graph of cortex, while random, is rich enough to include as subgraphs numerous synfire chains of given width and given multiplicity, we ask the following questions: (a) How should a chain be *defined,* i.e., what criterion—functional or anatomical—should be used to assign a given neuron to a particular chain? (b) To what extent does the analysis of an isolated chain apply to a chain embedded in a larger random network? (c) In particular, does the latter carry *stable* synfire activity?

In the random-graph model (Abeles 1991, Section 6.2), N neurons are assumed to form excitatory contacts with each other at random, with a given fixed probability π—the dilution coefficient—for each ordered pair of neurons. The likelihood of existence of a *subgraph* that constitutes a chain of width n and multiplicity m within this random graph of N neurons is assessed for various values of n. Given a set B of size n and given a neuron i, let $m_B(i)$ be the *B-multiplicity* of i, that is, the number of neurons in B that make contacts on i. One then computes, for a pool B^1 of n cells and for a number m, the *expected* number $x = x(N, \pi, n, m)$ of cells i in the network such that $m_{B^1}(i) \geq m$. If

$x \gg n$, the network almost certainly contains at least one set B^2 of size n such that the link $L = (B^2, B^1)$ is of multiplicity at least m. A chain of width n and multiplicity m can be defined recursively in this manner. This is one possible answer to question (a) above.

As a numerical example, consider the following values for the random graph: $N = 2 \times 10^4$, $\pi = 0.25$ (see Section 8). Set $n = 50$ and $m = 15$. The expected number of cells i with $m_{B^1}(i) \geq 15$ for a given set B^1 of size 50 is $x(N, \pi, n, m) = 5,038$ (Abeles 1991, Table 6.2.1). Therefore, there almost certainly exists a set B^2 of size $n = 50$ such that the multiplicity of the link $L = (B^2, B^1)$ is at least 15. Using this argument repeatedly one can find in the random network a chain $(B^\mu)_{\mu=1,\ldots,p}$ of width 50 and multiplicity 15, and of virtually any length.[20]

We shall now argue that, although many such chains are likely to exist in cortex under the random-connectivity model, these chains will *not* carry stable reproducible synfire activity[21].

With the above numbers and if we assume that the firing threshold—the number of synchronous converging spikes required to activate a neuron—is 15, activating a set of $n = 50$ neurons at time $t = 1$ will result in about $5,000$ neurons being active at time $t = 2$. If the width of the transmitted pattern is to be kept about equal to 50, a *regulation* mechanism—as discussed in Section 2—is required. Note that with the above numbers there almost certainly exist in the network chains of same multiplicity $m = 15$ but of width $n = 40$, $n = 60$, etc. Any particular value, such as 50, should then be viewed as imposed by the regulation mechanism. Thus, upon activation of a set B of size $n = 50$ at time $t = 1$, fifty neurons have to be chosen at time $t = 2$ from about $5,000$ neurons that are all equivalent if the firing threshold is 15. The only plausible way to make this choice a reproducible one is to activate the 50 cells with largest m_B. Regulation can then be viewed as adjusting the firing threshold to a value θ such that the number of cells i with $m_B(i) \geq \theta$ is about 50. This can only be done *in the mean:* the best choice for θ is the value given by $x(N, \pi, 50, \theta) \approx 50$; for instance, with the above values for N and π, the adjusted threshold is $\theta \approx 22$. This is precisely the r-WTA dynamics described in the present paper, with $r = n$.

Stated differently, the *definition* of a synfire chain of width n in a random graph (question (a) above) is now as follows. For any set B^1 of size n, let $\psi(B^1)$ be the set of n neurons with maximal B^1-multiplicity m_{B^1}. The synfire chain that has B^1 as first pool is obtained by iterating the map ψ: $B^2 = \psi(B^1)$, $B^3 = \psi(B^2)$, etc. Note that the chain is not uniquely defined, because ψ generally is not; ties in the definition of ψ can only be broken at random. We now face the following question (question (c) above): is synfire propagation on such a chain stable and reproducible?

Assume that a given set of neurons B of size n is active at time $t = 1$, and let a neuron i be active at time $t = 2$ if and only if $m_B(i) \geq \theta$, with θ such that $x(N, \pi, n, \theta) = n$. Due to the randomness of the graph, the number $m_B(i)$ is a random variable; its distribution is *unimodal,* specifically a binomial, well-approximated, in the limit of large n, by a Gaussian

[20] Here, the choice of m and n is arbitrary. A biologically motivated choice would be $m = 25$ and $n = 100$ (see Section 8). We chose somewhat smaller values for m and n in order to be able to use Table 6.2.1 of Abeles (1991). The conclusions of the discussion to follow are however valid for larger values of m and n, even more so than for small values.

[21] This problem is touched upon in Abeles (1991, Section 6.3.2).

of mean $n\pi$ and variance $n\pi(1-\pi)$. The value of θ such that $x(N,\pi,n,\theta)=n$ falls in the upper tail of this Gaussian, since n/N is a very small number. Denote by n' the number of neurons i such that $m_B(i) \geq \theta$, that is, the number of neurons active at time $t=2$; this number is itself a Gaussian variable, of mean n and standard deviation $\sqrt{n}$.

Thus, even if θ were perfectly adjusted, the regulation of activity would be quite noisy. Moreover, it is unreasonable to assume that θ can be adjusted very accurately. Yet n' happens to be *extremely sensitive* to the setting of θ. For instance, with parameter values as above, Table 6.2.1. in Abeles (1991) shows that decreasing m by as little as 1 results in an increase of $x(N,\pi,n,m)$ by a *multiplicative* factor of at least 3. This is true for any n between 10 and 50. For instance, for $n=50$, this 300% variation of the expected number of neurons active at time $t=2$ is caused by a mere 5% variation ($1/\theta$ with $\theta=22$) of the firing threshold. Under these conditions, the regulation would also be overwhelmed by *membrane-potential fluctuations* of postsynaptic neurons (compare to the robustness of the synfire-superposition model—see Section 7).

Finally, let us assume, for the sake of argument, that a perfectly robust regulation mechanism existed, causing the activity to follow *exactly* the chain B^1, $B^2=\psi(B^1)$, $B^3=\psi(B^2),\ldots$ constructed above. It is easily demonstrated numerically that, even in that case, synfire transmission would be highly unstable. Indeed, the k-th iterate of the map ψ on a random graph, with parameters in the same range as above and $k \gg 1$, is highly sensitive to small perturbations. Figure 15 shows the results of a computation performed with $N=5000, n=50$, and $\pi=.25$. The map ψ was iterated *twice,* starting from the *same* set B^0 of size n, but using two different seeds for tie breaking (see above). This yielded two realizations of the iterates B^k, and the curve shows the Hamming distance between the two realizations as a function of k (averaged over five trials). It is observed that within five or six iterations the two sets are virtually disjoint. Note that it is not even necessary to add noise (corresponding to membrane-potential fluctuations) to reveal the instability of synfire transmision in the random-graph model; the mere tie-breaking inherent in the definition of the chain suffices to overwhelm any reproducibility of transmission.

(Insert Figure 15 around here)

In conclusion, under the random-graph model for cortical connectivity, one should expect total activity in cortex to be a largely chaotic process, and one should not expect to observe reproducible complex events as defined in Section 9.1. Otherwise stated, the *correlation range* should be short: for $t_0 \gg 1$ and for any i and j, $E[x_i(t)x_j(t+t_0)] = E[x_i(t)]E[x_j(t+t_0)] = r^2/N^2$. In contrast, synfire transmission as studied in Section 6.3 is characterized by *long-range order:* if we assume cyclic boundary conditions for the stored chain, then, for any $i \in B^\mu$ and $j \in B^\nu$ and for any delay t_0, $E[x_i(t)x_j(t+t_0)] = r/N$ if $t_0 \equiv \nu - \mu \pmod{p}$, r^2/N^2 otherwise.

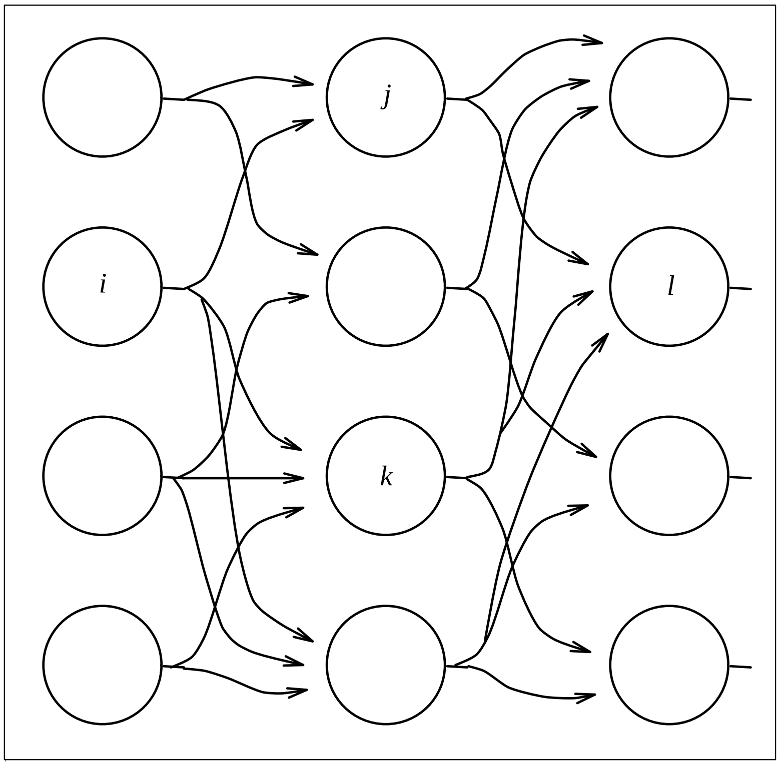

Figure 1: A fragment of a synfire chain. The length (number of *pools*) of the fragment is $p = 3$ and its width (number of neurons in each pool) is $n = 4$. Connections are feedforward, are allowed only between adjacent pools, and have identical conduction times (1 ms); the connectivity shown is incomplete. In view of the property of coincidence detection of neurons, parallel alternative pathways such as $i \longrightarrow j \longrightarrow l$ and $i \longrightarrow k \longrightarrow l$ *cooperate* in triggering the firing of target neurons (neuron l).

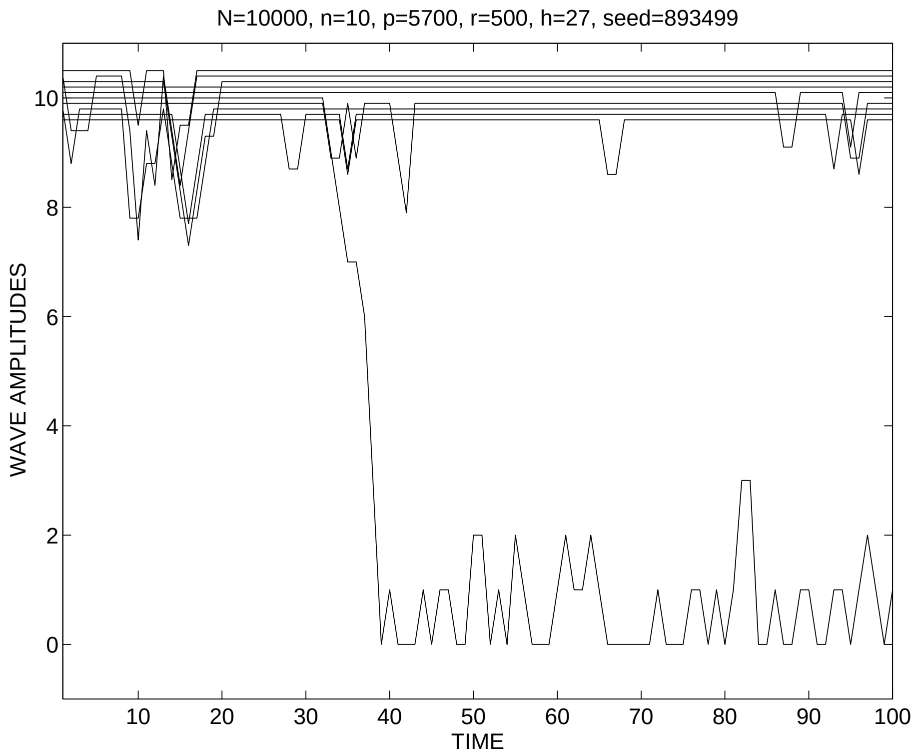

Figure 2: Synfire-wave transmission in a slightly overloaded system. A chain of width $n = 10$ and length $p = 5700$ is stored in a network of $N = 10000$ neurons, with $r = 500$ neurons active at any given time. The curve shows the amplitudes of 10 waves (the 10 traces are slightly offset from each other), out of 27 waves present at time $t = 1$. Under these parametric conditions, synfire transmission is not perfectly reliable; one of the 10 waves dies around time $t = 35$. The corresponding trace from $t = 40$ to $t = 100$ shows an uncorrelated binomial random process, the summed activity of a collection of $n = 10$ neurons where each neuron is active with probability $r/N = .05$.

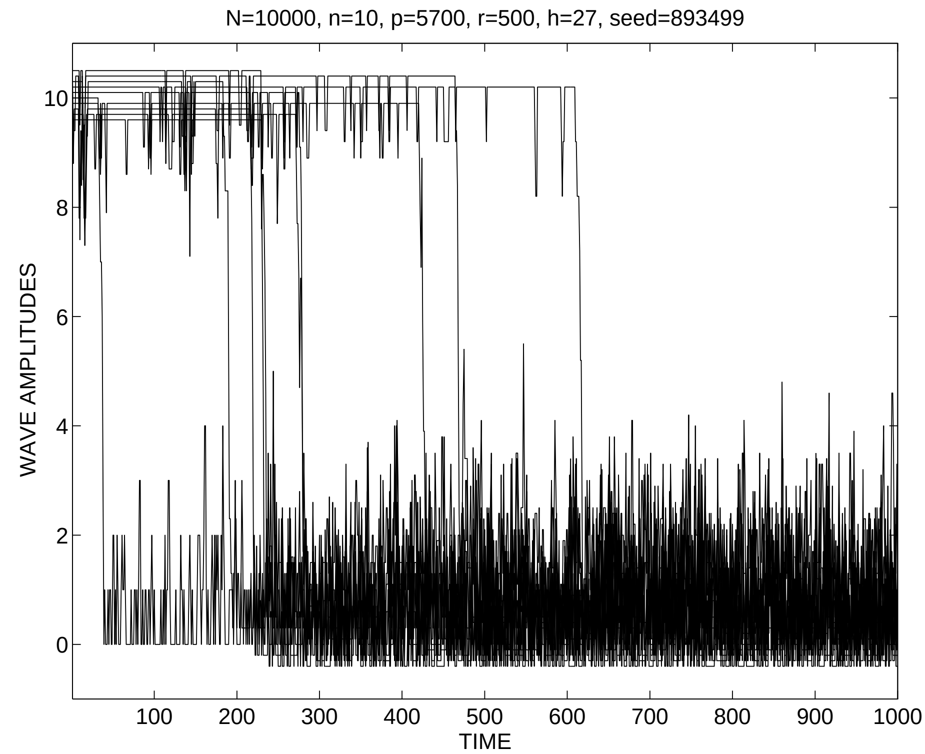

Figure 3: Same simulation as in Figure 2, but over 1000 time steps. All 10 waves have died out by time $t = 620$.

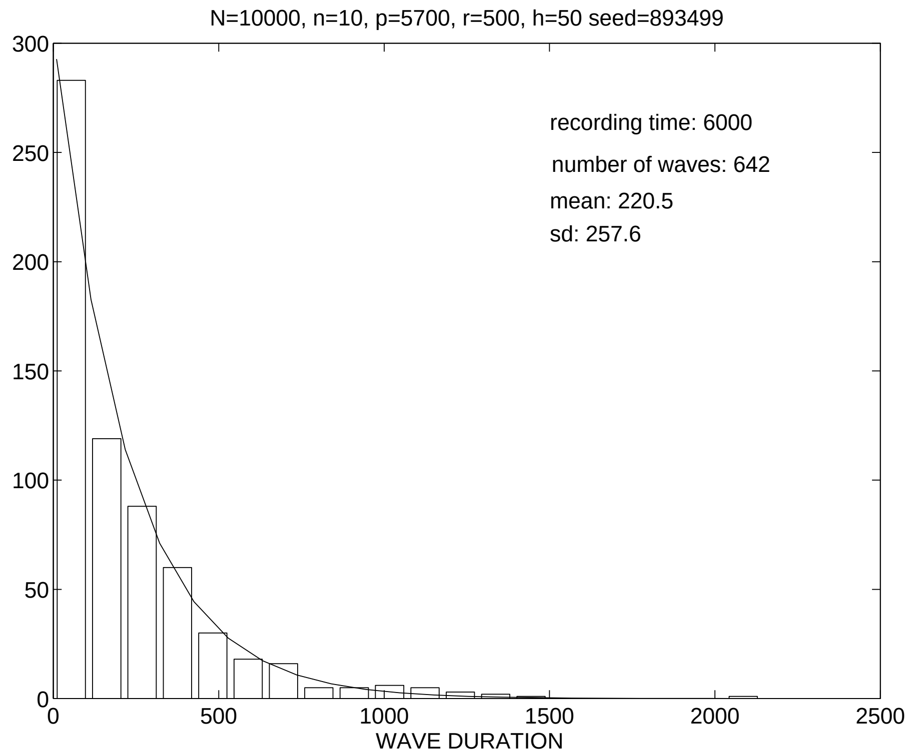

Figure 4: Histogram of durations of synfire waves. Waves are recorded during a 6000-step period, with parameters as in Figures 2 and 3. The solid line shows the fit by an exponential distribution.

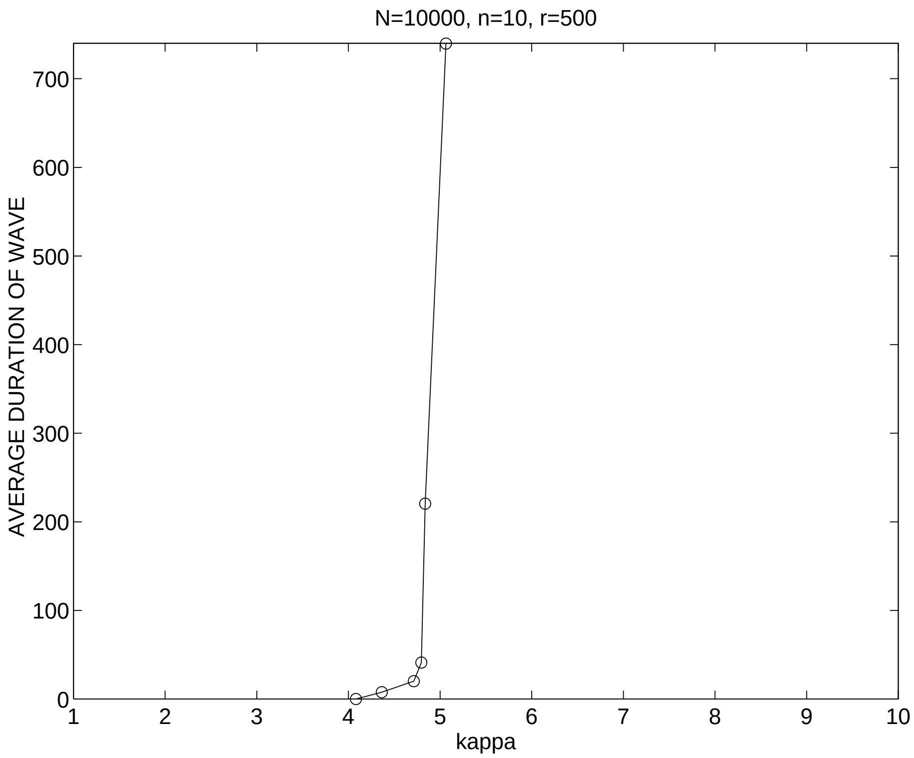

Figure 5: *Critical behavior* of the synfire-superposition model. The curve shows the mean duration of synfire waves (i.e., the mean of the distribution shown in Figure 4) as a function of κ, a parameter that varies inversely with the load, i.e., the length of the chain stored. κ relates to the distribution of membrane potentials in the network: it measures the separation between the modes corresponding to *crosstalk* and *crosstalk plus synfire input.* The curve shows a sharp transition at a critical value κ_c. With parameters as indicated on top of the figure, $\kappa_c \approx 5.1$; this corresponds to a chain length of about 5000. The critical value of κ is always near 5.

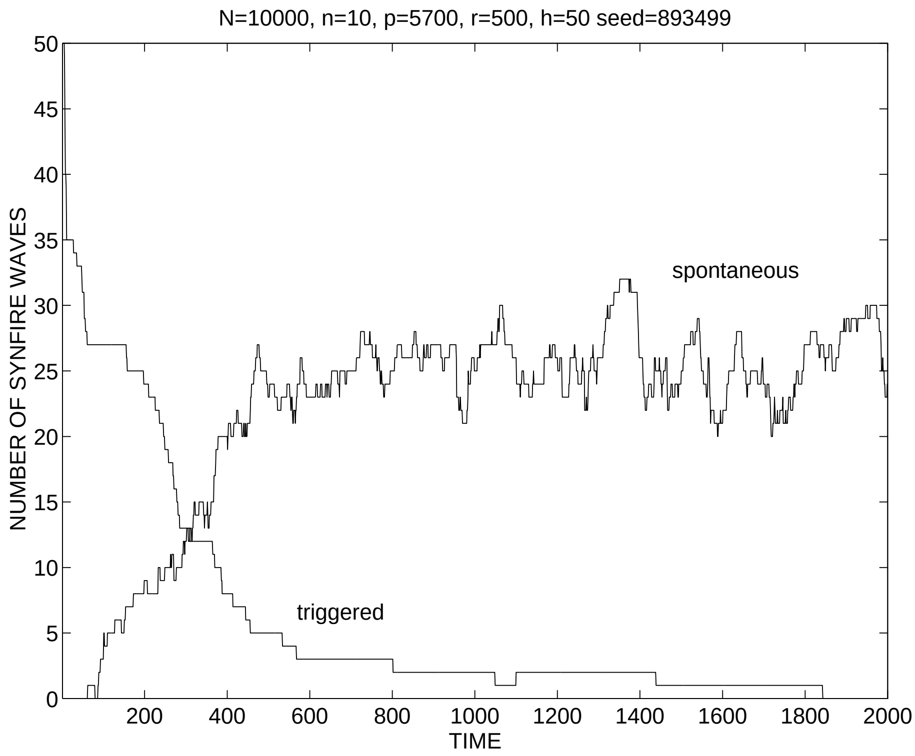

Figure 6: Dynamics of waves in a slightly overloaded system (same parameters as in Figures 2–4; κ is slightly below κ_c). Of an initial collection of 50 waves triggered at time $t = 1$, all have died out by time $t = 2000$. Other waves form spontaneously at random positions and times, resulting in an equilibrium population around 27.

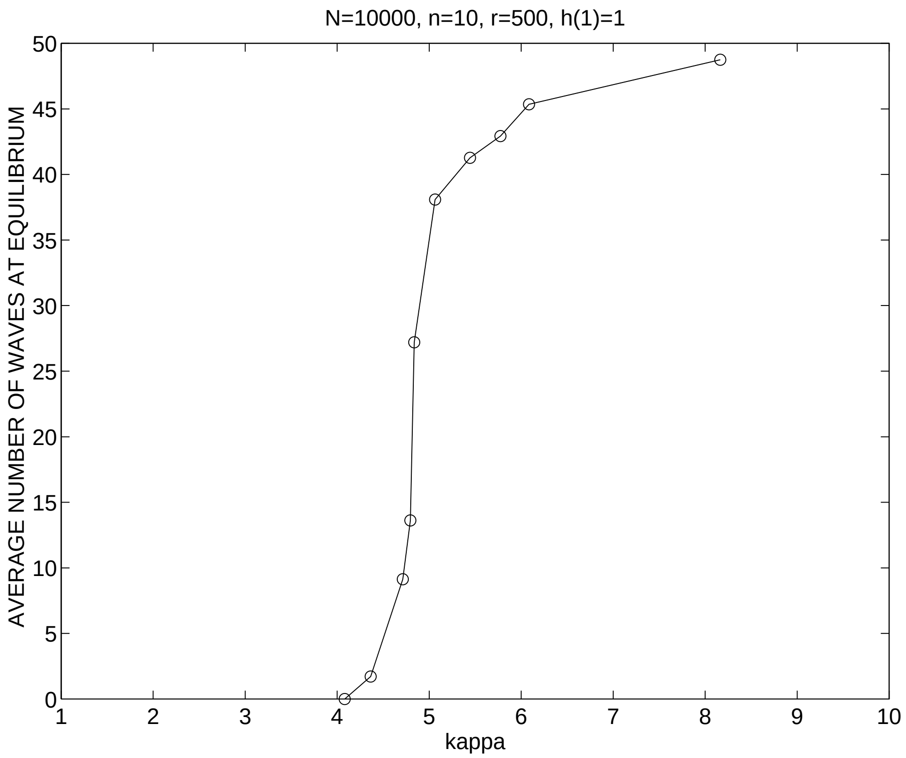

Figure 7: Average size of population of spontaneous synfire waves, as a function of κ. The curve shows a critical behavior, also manifest in wave duration (Figure 5).

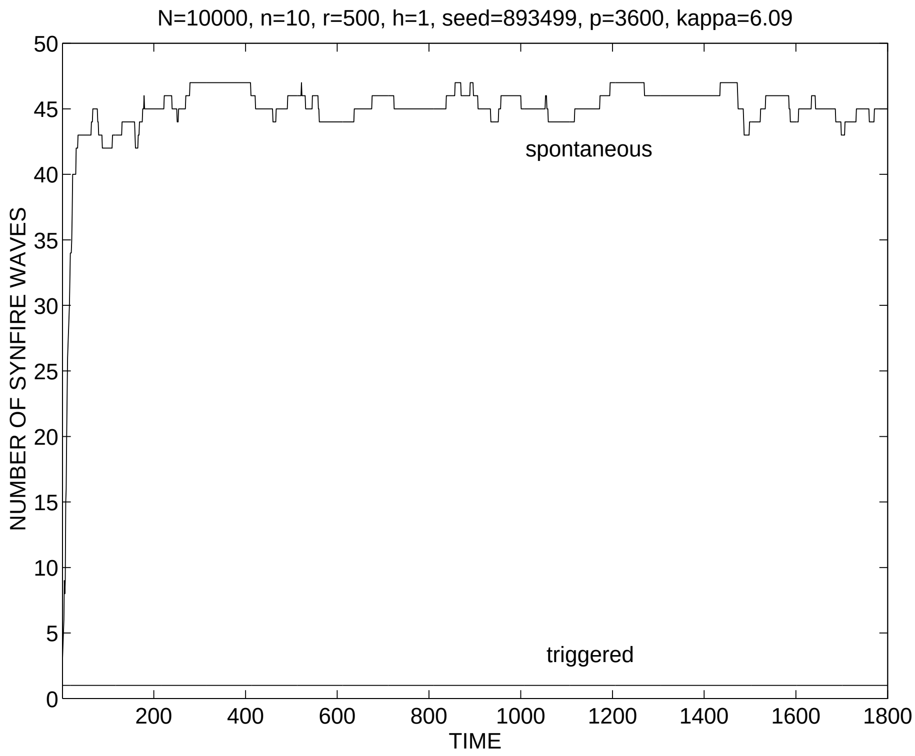

Figure 8: Dynamics of waves under stable synfire transmission. The length of the stored chain is $p = 3600$, resulting in $\kappa = 6.09$, well above the critical value. Synfire transmission is therefore *stable*, i.e., perfectly reliable. A single wave triggered at the beginning of the chain at time $t = 1$ is still alive at time $t = 1800$. Other waves form spontaneously at various positions along the chain, resulting in an equilibrium population of about 45; waves die when they reach the end of the chain. Under the conservative dynamics used in the model, the activity corresponding to a defunct wave (10 active neurons) spreads randomly throughout the network; it eventually is reassembled into a new wave.

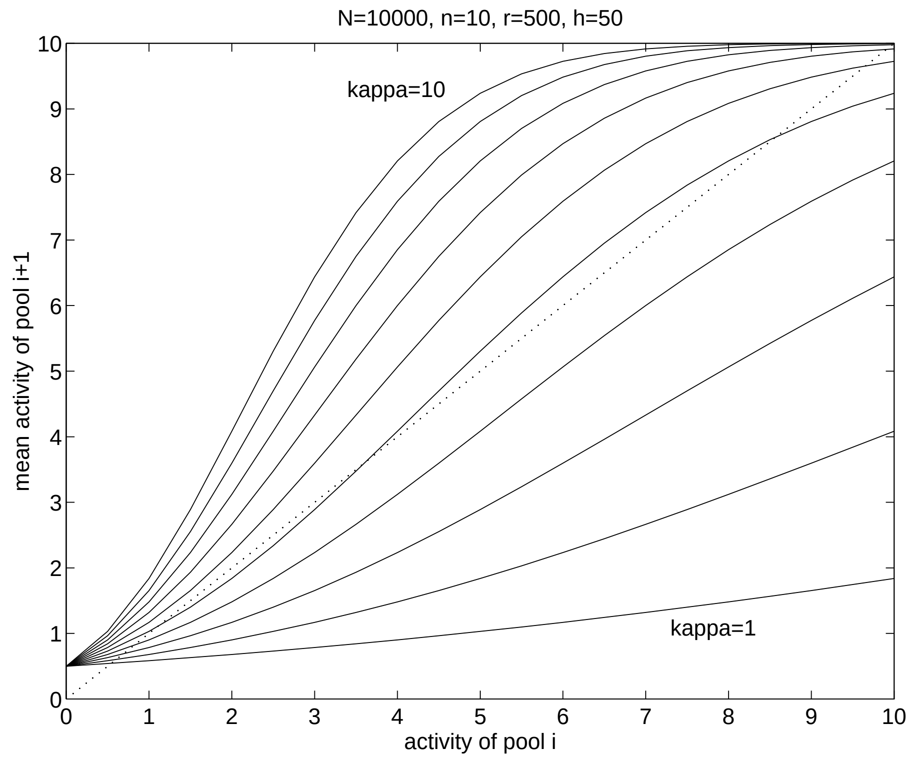

Figure 9: Synfire transmission functions. The curves show, for κ ranging from 1 to 10 by increments of 1 (high κ corresponds to low load), the expected number of neurons active in a given target pool of the chain as a function of the number of neurons active in the corresponding sending pool. If the upper part of the curve is well above the identity (dotted line) the average amplitude of the wave grows as it propagates along the chain, hence synfire transmission is stable. This diagram shows that κ should be larger than 5 for stable synfire transmission to occur, confirming the simulation results (Figures 5 and 7).

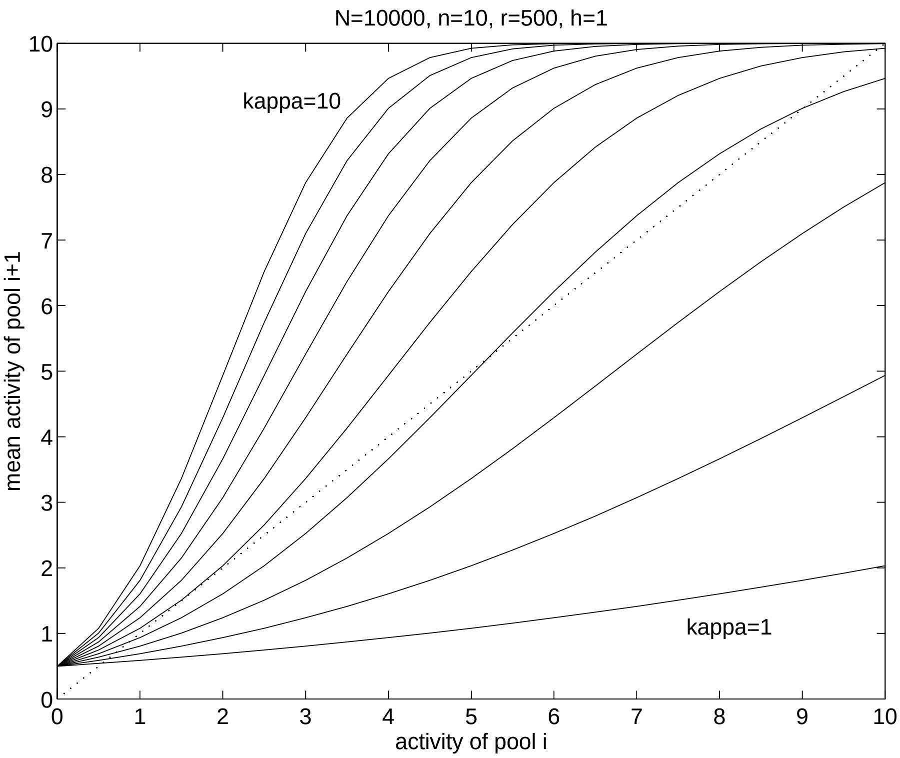

Figure 10: Synfire transmission functions for a single wave. All parameters are the same as in Figure 9, except that the number of waves on the chain is now $h = 1$ instead of $h = 50$. All curves are shifted upwards, resulting in more stable propagation.

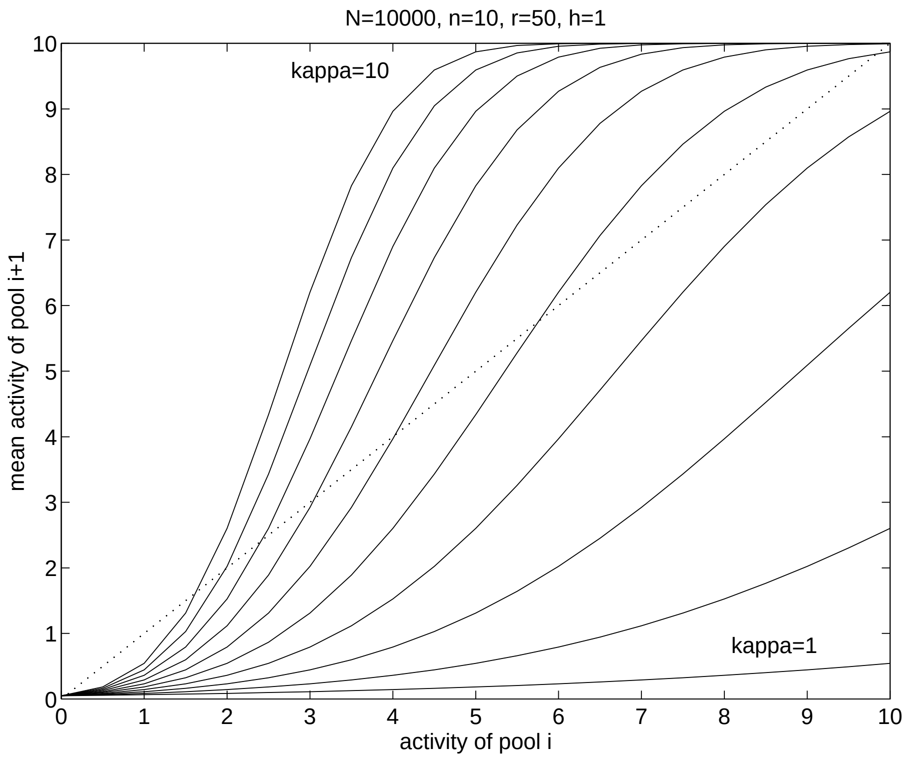

Figure 11: Synfire transmission functions whose lower part remains under the identity line. Parameters are the same as in Figure 10, except that the total activity in the network is now $r = 50$ instead of $r = 500$. With this parameter setting all curves start out below the identity line. This makes it more difficult for a small fluctuation to develop into a full-fledged synfire wave.

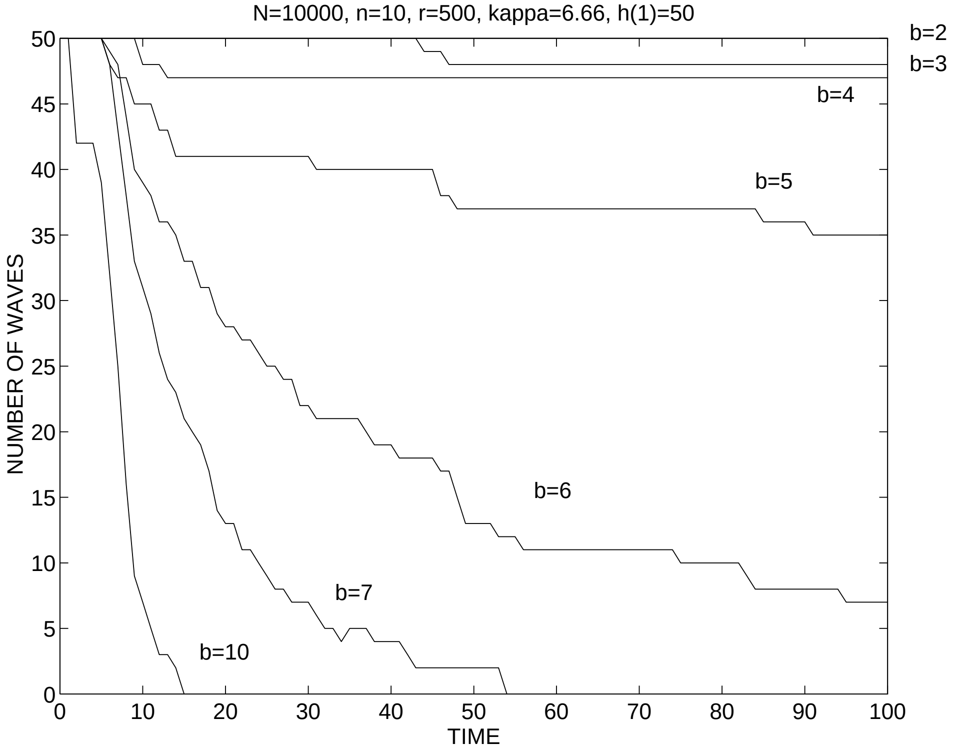

Figure 12: Synfire transmission under noisy conditions. White noise, uniformly distributed in the interval $[0, b]$, is added to the membrane potentials of all neurons. The curves show, for various values of b, the number of synfire waves still present at time t, out of a collection of 50 waves triggered at time 1.

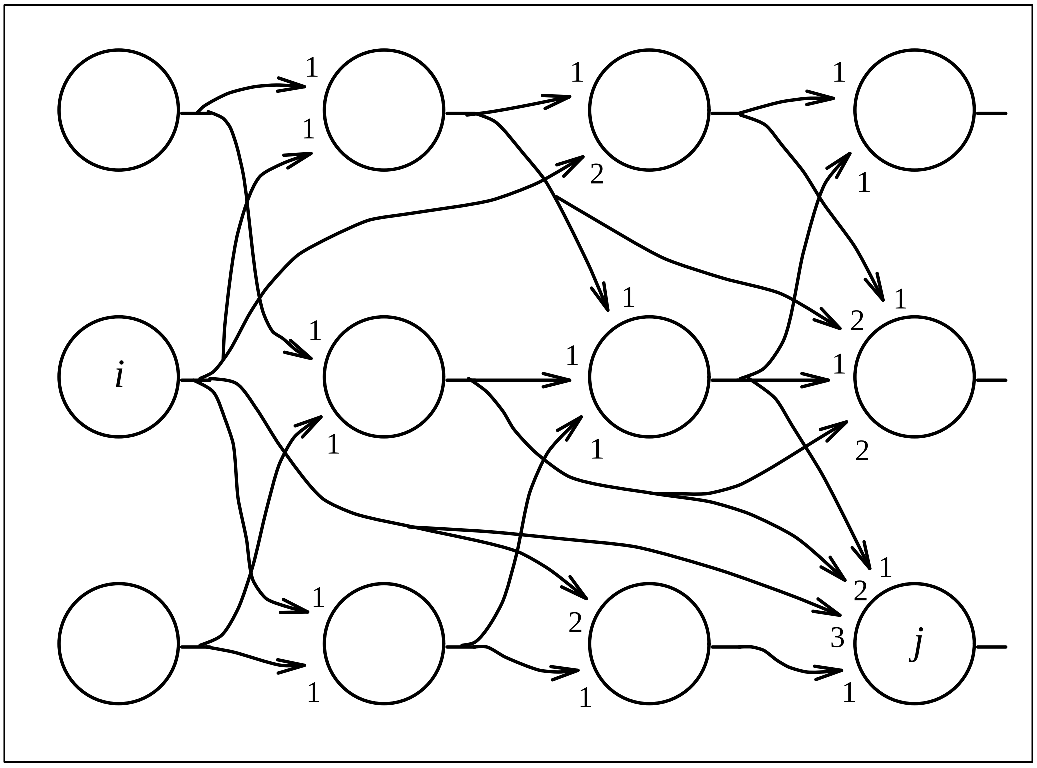

Figure 13: A fragment of a discrete-time synfire braid (Appendix A). Conduction times are indicated next to each synapse. As in a synfire chain, neurons are arranged in a discrete sequence of pools, with identical delay (1 ms) between any two adjacent pools. Unlike in a synfire chain, there are connections between non-adjacent pools; the corresponding conduction times (axonal plus synaptic delay) are multiples of 1 ms. The spatio-temporal consistency rule (equation 8) states that the summed conduction times along alternative pathways between two given neurons should be identical. For instance, the five different pathways from neuron i to neuron j all have a total conduction time of 3 ms, hence *cooperate* in triggering the activity of j. As a result, the braid behaves like a synfire chain.

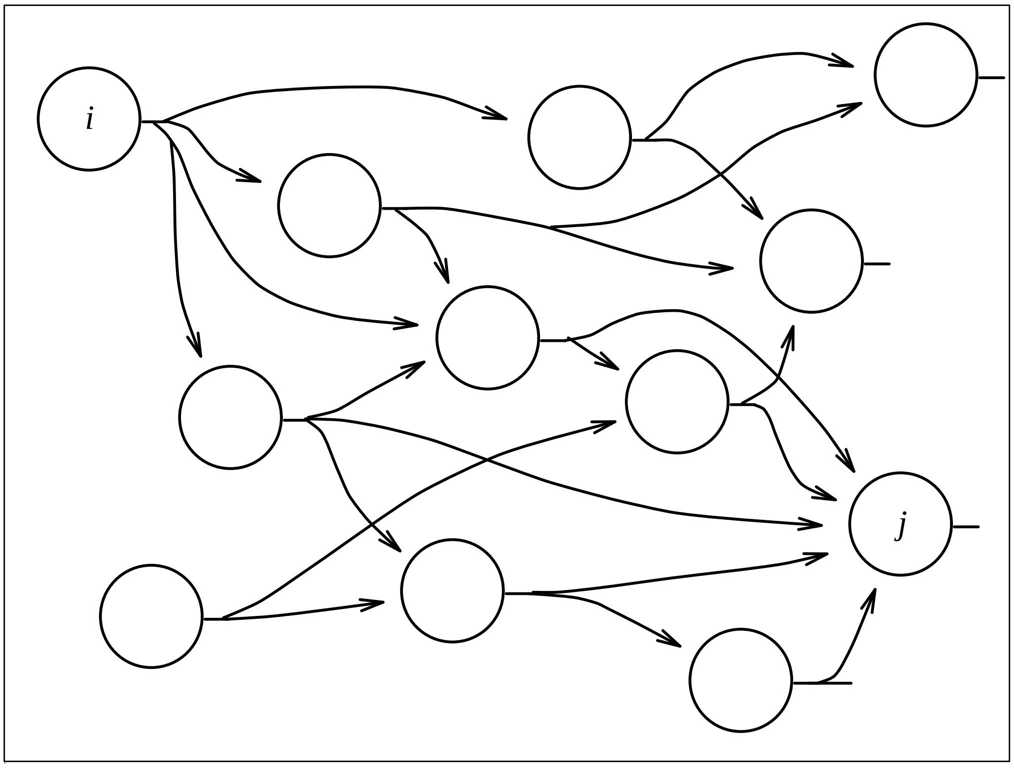

Figure 14: A fragment of a continuous-time synfire braid (Appendix A). The braid is a loop-free directed graph with real-valued conduction times (not indicated in the figure). Unlike in chains or discrete-time braids, neurons do *not* form a discrete sequence of pools. However, the spatio-temporal consistency rule (equation 8) still applies, requiring for instance that the 9 different pathways from neuron i to neuron j all have nearly identical total conduction times. Again, this consistency/cooperativity property yields a behavior similar to that of a synfire chain.

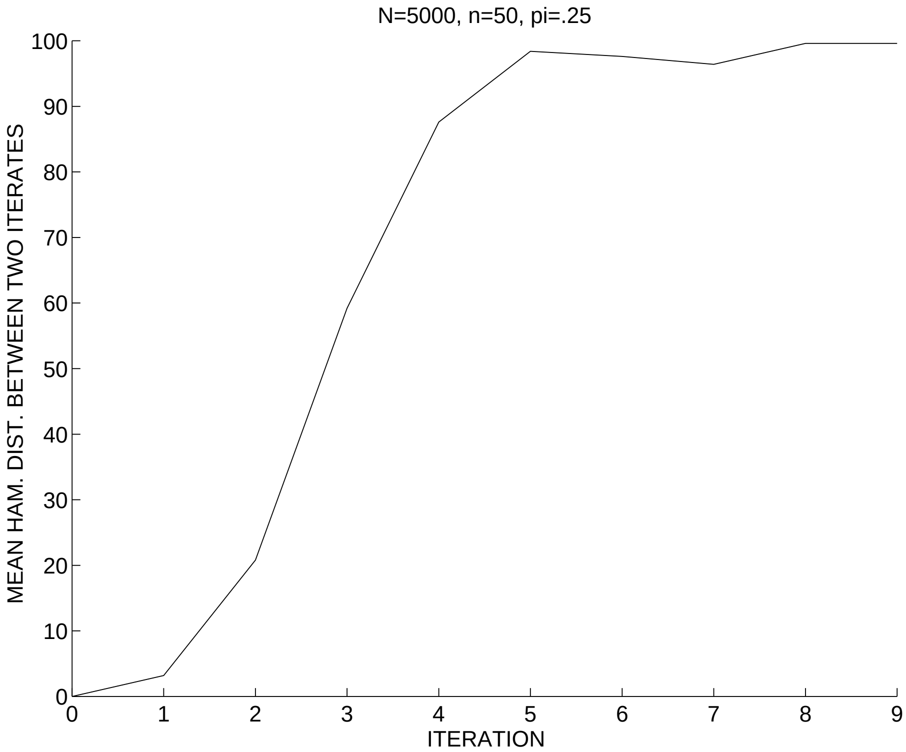

Figure 15: Lack of reproducibility of synfire transmission in the random-graph model (Appendix D). In a random graph, one could envisage the propagation of a synfire wave as follows. Upon activation of a set B of size n at time t, the n neurons that receive *maximal input* from B become active at time $t+1$; this n-WTA rule requires some tie breaking. In the simulation shown, the size of the graph is $N = 5000$ and the probability that a node i is connected to a node j is $\pi = .25$. Two 'synfire waves' of amplitude $n = 50$ are propagated, starting from the same initial set B^0 but using different seeds for tie breaking. The Hamming distance between the resulting two sets is shown as a function of time (averaged over five such experiments). It is seen that the two sets rapidly become disjoint (Hamming distance equal to $2n$). This demonstrates that the tie breaking inherent in the definition of a synfire chain in the random-graph model suffices to rule out long-range order: there is no reproducible synfire propagation in the random-graph model.