Chapter 9

Pulse-Mode Network Modeling

§ 1. Pulse-coupled Neural Networks

Neural network models based on spiking proxy neuron population models occupy an intermediate position in the modeling hierarchy between the neural netlet and functional microcircuit models of chapter 7 and the map and network system models we discuss in chapter 11. Figure 9.1 illustrates the positions of these models in our hierarchy. Modeling at this level can be conveniently sub-classified into two levels: (1) proxy neuron population models (PNP sub-networks) built using pulse-coded models (e.g., Integrate & Fire, Eckhorn models); and (2) neural networks constructed from these population models. The latter are often called either pulse-coupled neural networks or pulse-coded neural networks (PCNNs in either case; the terminology has not yet been standardized). It is not uncommon for the distinction between PNP subnetwork level models and PCNN models to be blurred, in which case the neural network might be called a PNP (proxy neuron population) network because it is often given a crystalline neural network structure that does not distinguish functional subnetworks within it.

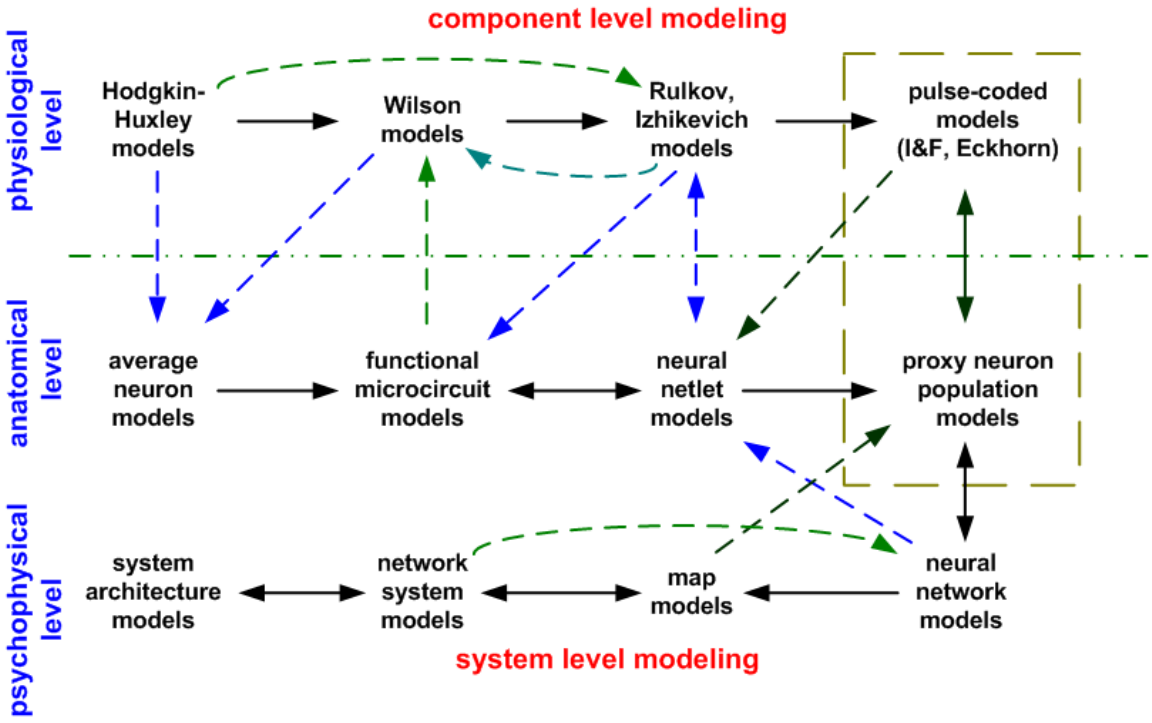

Figure 9.1: Hierarchy of modeling levels in the system doctrine roadmap.

PCNNs have an intuitive appeal that can quite easily be deceptive if one forgets that proxy population models such as the I&F or Eckhorn "neurons" do not in fact model individual neuron cells. Rather, they are models of the collective and functional behavior of large collections of tightly-coupled biological neurons having output tract projections that tend to fire or not fire synchronously with one another. Such a collection of tightly-coupled biological neurons is often called a **cell group**. This is an important fact to note because it means the input connections to, e.g., an Eckhorn node can legitimately be called **weights** but should not be assumed to model synapses. The input weights of an I&F or an Eckhorn model do reflect a measure of efficacy in terms of efferent (output) tract dependence on afferent signals converging on the cell group. But it is an error to assume these weights are direct models of synaptic efficacy because the actual *signaling* efficacy of afferents coming into the cell group will be a complicated function of the types of actual neurons in the cell group population, their interconnections, the types of neuro-transmitters employed, the ratio of excitatory to inhibitory cells within the cell group, and probably many other factors as well. The population proxy is a "black box" or "input-output" model of signaling behavior and it is a functional rather than a mechanistic model. If one wishes to "get inside" a cell group population and explore mechanisms, it is necessary to pull back one step in Figure 9.1 and model the system in terms of neural netlets and abstract neuron models such as the Rulkov or Izhikevich models.

Experimental evidence for the existence of cell groups, while still largely circumstantial owing to current instrumentation limitations, does nonetheless appear at present to strongly support the hypothesis that the central nervous system does contain them. However, there is also experimental evidence indicating that cell groups might not constitute permanent neural structures and that some dynamical "re-wiring" of biological signaling pathways does occur in the CNS. For example, evidence strongly indicates that the number of "functional columns" in the visual cortex of mammals is larger than the volume available to hold them if each functional column corresponded to a single biological column. If this is true, as it presently appears to be, the physiological situation limits the level of trust one can invest in the individual proxy nodes (I&F, Eckhorn "neurons") when modeling at this level.

The way this is dealt with is to assume that cell groups (at some biological population level) exist within more dynamically-variable populations called **cell assemblies**. Cell assemblies are postulated to be larger populations of interconnected cells in which cell-to-cell coupling is characterized by a higher degree of variability such that not every subpopulation within a cell assembly necessarily responds synchronously with all other subpopulations. There is almost no doubt that cell assemblies do exist within the CNS, and the greater functional flexibility they

provide allows the modeler to compensate for the apparently transitory nature of cell population phenomena. Cell assembly modeling is modeling at the level of the PNP network. PNP networks are modeled as interconnected *functional* assemblies of cell groups (each of which is modeled by a single population proxy such as the I&F or Eckhorn model). Under this hypothesis, a functional column would be a PCNN rather than a cell group or PNP network.

The picture that emerges from this cell group / cell assembly hypothesis is one where a cell group would consist of a tightly coupled population of neural netlets comprised of functional microcircuits. A cell assembly is then an aggregate of cell groups containing, in all likelihood, additional afferent inputs that serve a "control and configuration" rather than a "data path" functional role. It is at present an open question as to how large a biological population corresponds to a cell group within a cell assembly. In a great many ways, modeling at the neural network level in Figure 9.1 is currently the most difficult challenge for computational neuroscience, the difficulties arising not only from the complexity issue but also from the current lack of experimental tools and methods that would permit the finer structural details of the system to be identified and examined in the laboratory. The neural network level is in many ways the level within the systematic doctrine where biological mechanism and psychophysical phenomena must meet and merge. It is not unfair to say that at our present level of knowledge modeling at the neural network level is much more of an art than a science and that many published models are more Platonic than empirical.

Among the numerous shortcomings in current neural network theory, one of the most obvious and important is the lack of a generally agreed-to method for modeling adaptation at the pulse-coded neural network level. Adaptation in this context refers to models and algorithms for adjusting the weights within a PCNN model. We are in possession of functional models for adaptation at the map and network system levels of the hierarchy. We are likewise in possession of mechanistic models for short-term and long-term synaptic plasticity at the biological level. We do not yet possess a well developed theory for adaptation modeling at the PCNN level, although a few (mostly engineering-related) papers on this topic have appeared in the neural network literature. Currently these models do not yet command widespread acceptance in the computational neuroscience community and should presently be viewed as essays in the craft of modeling development. For that reason, this textbook does not discuss adaptation at the PCNN level although we will discuss adaptation in considerable depth at the map and network system levels in the later chapters. Almost all current PCNN models are fixed-weight models. Numerical values for these weights are most often obtained from *ad hoc* considerations (peculiar to the individual modeler-as-artist) or from organized functional search methods employing such

techniques as evolutionary computing algorithms and genetic algorithms [GOTS1].

§ 2. Activity and Moving Average Firing Rate in PNP Nodes

In chapter 7 we introduced Weinberg's three-regions characterization of systems and saw that neural systems fell into Region III, the region of organized complexity. Figure 9.2 illustrates the general layout of Weinberg's doctrine. Proxy neuron population models represent one approach to dealing with organized complexity on the (relatively) smaller scale. In essence, this technique can be regarded as an attempt to exploit relatively low randomness in cell group behavior by reducing the entire cell group to a single element, namely the PNP node. As we advance to the level of neural network modeling, these nodes are incorporated as elements in a network that frequently is given the form of a crystalline network (that is, a highly regular interconnected network of nodes). In such a network the node model is treated as a system with organized simplicity and complexity is to be captured by the network's overall behavior. It is at this level of modeling where we encounter the need to introduce and quantify the concept of the activity of a node and of a network comprised of multiple nodes.

Action is a vaguely defined generic term used by biologists, psychologists, and neuroscientists to denote the occurrence of some process. For example, it is used in "action potential" to denote the generation and propagation of an electrical impulse down the axon of a neuron. **Activity** is an equally ill-defined term used to denote the presence and, usually, the level of intensity of an action. To make these terms less vague, and therefore more meaningful, a science must be able to

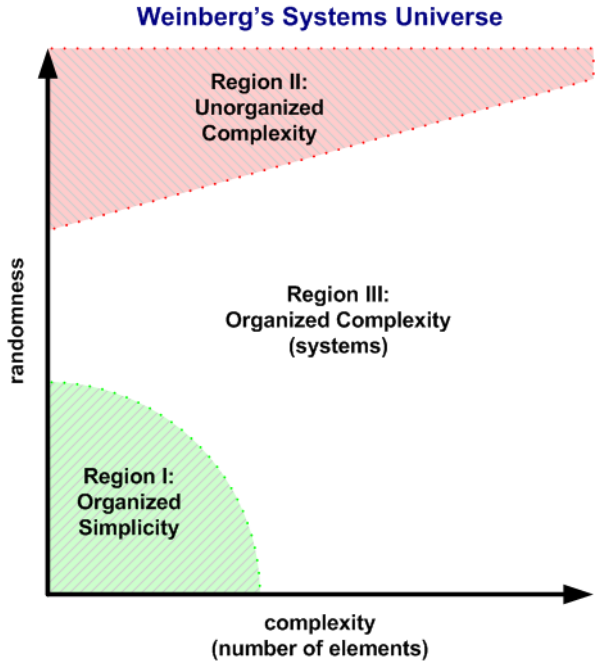

Figure 9.2: Weinberg's System Regions Model

define its technical terms in relationship to some experimentally determinable phenomenon. In the case of a population of neurons the most experimentally accessible type of measurable process is **metabolic rate**, the rate of oxygen consumption accompanying the enzymatic release of energy during cell metabolism. A measure of this can be obtained experimentally by brain imaging methods such as functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) imaging. In an fMRI, blood oxygen levels are detected by measuring changes in the ratio of oxyhemoglobin to deoxyhemoglobin, the levels of which vary with blood volume, flow, perfusion[1], and metabolism. In PET imaging changes in blood flow and glucose metabolism are detected. Both methods thereby provide quantitative indicators of metabolic rate over a cell population resolvable down to about 1 mm (for fMRI) to a few millimeters (PET). Because cell metabolic rate increases when a neuron produces action potentials, metabolic rate is an indicator of how much cell signaling is taking place in neural assemblies by neurons that produce action potentials. Thus, in this book **we define activity as any quantitative measure of metabolic rate in an assembly of neurons and glia**.

It is common for neuroscientists to describe the activity of a single spiking neuron cell in terms of the rate at which it produces a series of action potentials (APs). This is called the **firing rate** of the cell and is an easy quantity to measure if the cell produces action potentials at evenly spaced intervals of time. Unfortunately, many cells do not produce evenly-spaced action potentials and so the notion of an average firing rate is often introduced by theoreticians. (This is a practice going back to von Neumann). The introduction of the notion of an average firing rate introduces a number of other technical difficulties because it is defined by counting the number of APs produced over some time interval T and then dividing this number by T. This, however, is a rather faulty definition because for it to be physiologically meaningful one must assume that there is some psychophysical significance to the time interval T and that the statistics of cell signaling are stationary (that is, do not change) over the interval T. The former type of assumption is quite speculative and the latter is impossible to verify in any *physically* meaningful way. A third disadvantage of the firing rate model is that it cannot be applied to non-spiking neurons that signal by means of gap junction synapses, nor can it be applied to glial cells (the metabolic rates of which are also part of what is measured experimentally in an fMRI or a PET scan).

Nonetheless, when dealing with proxy neuron population models such as the I&F and the Eckhorn models, we must have some means of relating the firing actions of these nodes to the activity of the neural population. A less speculative way of doing so is to define the **instantaneous firing rate** of a PNP node as the reciprocal of the time between successive output

[1] An undefined term used to vaguely denote how much the blood "covers" the cells it is supplying.

pulses. This definition does not provide a one-to-one relationship to the population activity in the sense in which we defined this term above, but does provide *some* indication of the level of activity within the population. If T(t) is the time interval between an output pulse at time t and the immediately previous action potential, then $r(t) = 1/\mathrm{T}(t)$ is the instantaneous firing rate. We then *make the hypothesis* that the interval-average of $r(t)$,

$$\bar{r}[t-\chi, t] \equiv \begin{cases} \dfrac{\int_{t-\chi}^{t} r(x)\cdot p(x)dx}{\int_{t-\chi}^{t} p(x)dx}, & \int_{t-\chi}^{t} p(x)dx \neq 0 \\ 0, & \int_{t-\chi}^{t} p(x)dx = 0 \end{cases} \tag{9.1}$$

over time interval $\chi \gg$ T, is proportional to the population's time average activity over that same interval. Here $p(x)$ is defined to be equal to zero if no pulse occurs at time x and equal to a unit-strength Dirac delta function $\delta(x)$ if a pulse does occur at time x. Quantity $\int p(x)dx$ is merely the total number of pulses in interval χ. The quantity $\bar{r}$ is similar to von Neumann's pulse density representation [NEUM3] but is defined for cases where the instantaneous firing rate is variable.

To be experimentally meaningful, the interval χ must be of the same order as the temporal resolvability of the imaging method used to measure brain activity but with a *mathematical* lower bound for χ on the order of no less than several tens of milliseconds. The lower bound is needed to limit and control statistical variance in $\bar{r}$ since this variance constitutes an inherent "noise" in the statistic. One should note that $\bar{r}$ is defined over an interval. To incorporate it into computational analysis of a network we require a time-sampled quantity defined in terms of the computing interval. Such a quantity is said to be a discrete-time quantity. We therefore define the **moving average firing rate** of a PNP node as

$$\rho(k\cdot\Delta t) \equiv \bar{r}(k\cdot\Delta t - \chi,\, k\cdot\Delta t) \tag{9.2}$$

where Δt is the sampling interval of the computer simulation model and k is the sampling index. It has units of pulses per second. Computational accuracy dictates that Δt be very small compared to the interval χ.

Moving average firing rate ρ does not get entirely away from the shortcomings noted above; its usefulness ultimately comes down to linking it to experimentally determinable quantities. This linkage is conjectural at present because we presently lack the instrumentation tools necessary for identifying cell groups, making direct testing of the hypothesis presently impossible. Therefore ρ is a mathematical quantity rather than a physiological quantity. It is quantized since changes in its value are produced by integer changes in the number of pulses occurring within interval χ.

§ 3. Activity Propagation in Simple PNP Chains

Traditionally theoretical neuroscience has focused its almost-exclusive attention on action potentials and sequences of action potentials under the presupposition that this signaling must be the seat of what is called *the neural code*, wherein the long-sought secret of the phenomenon of mind is presumed to reside[2]. The presumption dates back to von Neumann's work [NEUM3] and has gone largely unquestioned by neuroscientists since that time. Although this is a metaphysical and materialist presupposition with many fundamental problems and issues confronting it[3], it is nonetheless correct to conclude there is a real relationship between cell *activity* and mental phenomena [WELL2, WELL3]. It goes farther than we are scientifically justified in saying that this relationship is restricted only to action potentials and action potential sequences because this neglects the existence of non-spiking signals (gap junctions) as well as the possibility that glial cells might also participate in the dynamics of information processing in brain function. However, non-spiking neuronal signals and, to a lesser extent, glial signaling processes (e.g. the propagation of calcium waves by glial cell networks) are relatively local functions and it is true that long-distance cell-to-cell communication takes place primarily by means of the propagation of action potentials[4].

Whatever the real relationship between brain function and the phenomenon of mind is, there is a mountain of empirical evidence today pointing to the strong hypothesis that this brain function is highly distributed within the structure of the brain. It is true that a great many functions that appear to be intimately related to specific mental phenomena are more or less highly localized in an individual's brain; it is no less true that such local pockets of activity by themselves are not sufficient to explain the overall link between brain function and mental phenomena (e.g. perception and cognition)[5]. To find the link between the psychological world and the biological world, we are led to consider and study the propagation and dependencies of neural activities to, from, and among different regions of the brain. This, in turn, gives rise to the idea of **processing**

[2] What most theoreticians today mean by the term "neural code" was called "the language of the brain" by von Neumann. Von Neumann drew a crucial distinction between the idea of a *language* (which is related to meanings and purposes) and the idea of a *code* (which is a symbolic representation of a language). This distinction is similar to the distinction information theorists draw between *information* and *data* (which is merely the representation of information). Von Neumann's distinction is one that tends to be forgotten by present day theoreticians.

[3] The problems and issues fall under the generic title of "the mind-brain problem."

[4] "Primarily" rather than "exclusively" because we cannot rule out the likelihood that signaling via the body's endocrine system also plays a role in the relationship between biological function and mental phenomena.

[5] Some neuroscientists are willing to at least entertain the notion that mind-brain relationship can be directly linked all the way down to the individual cell level. Malsburg critiqued this idea as the notion that a cell is an "atom of meaning" [MALS2]. It is a romantic but utterly ungrounded presupposition.

chains in brain structure modeled at the level of spiking-mode neural network organization.

The simplest form of a neuronal chain structure is illustrated by Figure 9.3 at the PNP level. It is a one-dimensional feedforward chain, i.e., activity propagation takes place unidirectionally and along a single signaling pathway (tract) from cell group to successive cell group. Spiking proxy neuron population models (e.g., Integrate & Fire or Eckhorn nodes) constitute PNP nodes making up the chain. Because each PNP node represents a cell group (and, therefore, each connection line in the diagram represents a tract of axons), transmission of activity from one node to the next is called an **activity wave**.

Until the late 1980s and early 1990s, theoretical neuroscience largely ignored wave-like signaling behavior in the brain and, instead, operated under a strict paradigm based on the average firing rate model. It furthermore regarded average firing rate exclusively in terms of action potential signaling rather than the view that these signals are merely indicative of activity; i.e., it was presupposed that the action potential spikes themselves (rather than the cell activity these signals stimulate) were the sole and proper carrier of information pertinent to mind-brain relationship. Again this presupposition can be traced back to von Neumann's work. It is a view that a computer theorist would call a "data path model" because the paradigm utterly ignores the existence of metabotropic signaling processes in the brain. (These processes were unknown in von Neumann's day). This view reflects what we might call a neuron-centric view of brain function; such a view is promoted by failure to appreciate the level of complexity existing in region III systems. Under this paradigm, the simple chain was viewed as performing the task of relaying information from place to place within the CNS. What remained unclear under this paradigm was how a simple relay operation could be related to the *mental* phenomenon of *knowledge* representation (since knowledge and information are not the same thing). Inherent in this paradigm was the vague idea that significance of signaling in terms of knowledge could be vested in the *location* of the PNP node. Roughly put, the meaning of the signal is presumed to depend upon the *localized* functions performed by the specific nodes where the signaling is taking place. That this is a metaphysical presupposition is downplayed by employing the terms "symbol" and "symbolic processing" in place of the more highly-charged word "meaning."

Beginning in the late 1980s, this "traditional" view began to be challenged under the weight of new experimental observations made possible by advances in instrumentation techniques. One of

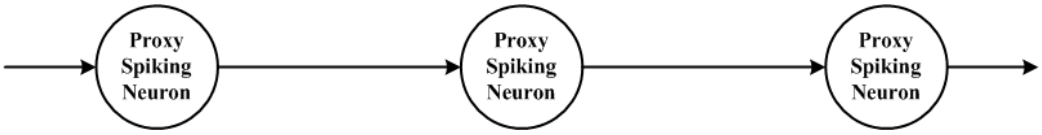

Figure 9.3: Simple one-dimensional chain of cell groups.

early critics of this dogma was Moshe Abeles:

> The prevailing view is that the firing times of a neuron are the realizations of a stochastic point process with variable underlying rate functions. The immediate outcome of this view are a host of procedures that estimate this putative underlying rate and its dependence on other observable events. . . There are several pieces of experimental evidence that do not conform to the view that all that there is to neuronal firing is the underlying firing rate. [ABEL5]

The evidence, Abeles argued, was pointing to a conclusion that, whatever it is that the brain is doing, brain function is not adequately described by this simple paradigm. This led Abeles to propose what is known as the synfire chain model of neural processing, which we discuss in the next section. As it happened, criticisms such as those of Abeles and others came to meld with an earlier criticism put forth in 1981 by Malsburg, leading to an increased appreciation that brain function (as opposed to mere signaling activity) presented a more complicated scientific problem than elementary computer-scientist-like habits of thinking had supposed:

> The purpose of this paper is to point out a specific deficiency in existing brain theory and to propose a way in which this gap could be filled. Although it leaves open a number of technical questions and presents more of a program than an accomplished theory, at least a language is developed to describe processes from the cellular to the cognitive level. . . There is every reason to believe in the existence of general principles governing the function of the brain. Cerebral anatomy is surprisingly homogeneous in spite of the diversity of functional modalities represented in its different parts. The rapid cultural development of man has created fields of mental activity for which the brain cannot have been prepared by phylogeny in any detailed way. Both arguments seem to force the conclusion that the brain is governed by general principles of organization. [MALS2]

Malsburg's "accomplished theory" has not yet been accomplished and a host of open questions still remain to be addressed by theoretical neuroscience. However, many of the current programs of research attempting to do so do employ the use of chain models and so an understanding of the basic properties of the simple chain is propaedeutic to any discussion of these more recent developments.

§ 3.1 Propagation in Integrate & Fire Chains

We first consider activity wave propagation in a simple chain comprised of identical I&F nodes. The basis for this analysis is the dynamics of the I&F node discussed in chapter 8 (§3.2). Each node will be presumed to receive or produce a unit impulse at each firing event. The connection weight of the input is w and the time constant for the leaky integrator is τ. Let the input to the chain consist of evenly-spaced unit impulses arriving at times $t_k = k \cdot \mathrm{T} \cdot \Delta t$ where Δt is the sampling interval of the simulation, k is a non-negative integer, and T is an integer denoting the number of sampling intervals between successive inputs to the chain. We assume the I&F nodes have initial conditions of zero prior to the first input pulse, which arrives at time $t_0 = 0$.

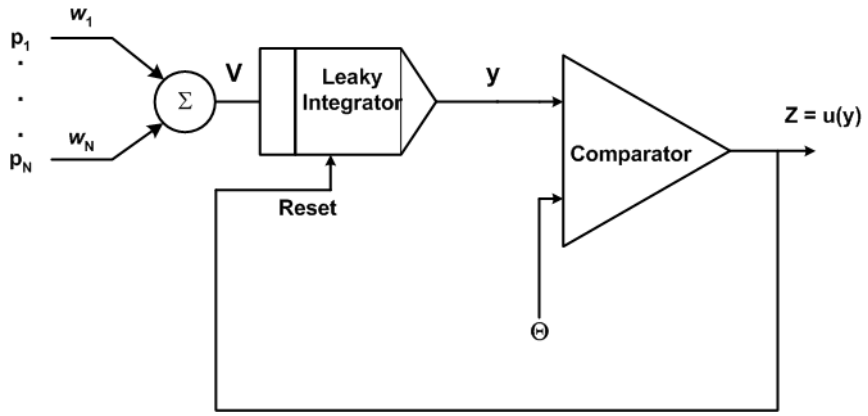

Figure 9.4: Block diagram of the Integrate & Fire node.

For convenience, the block diagram representation of the I&F node from Figure 8.10 is reproduced here as Figure 9.4. The first input pulse produces an input $V^{(1)}(0) = w$ to the leaky integrator (LI) of the first node. The response of the LI to this impulse is given by equation (8.5) as $y^{(1)}(\Delta t) = w/\tau$. We have two cases to consider according to whether (a) $y^{(1)} > \Theta$; or (b) $y^{(1)} < \Theta$.

For case (a) the LI output exceeds the firing threshold Θ and the first I&F node responds by producing a unit impulse at $z^{(1)}(\Delta t)$ and resetting the LI. We call this case the **all-pass firing mode** of the I&F node. Because we are assuming all the nodes in the chain are identical, firing by the first node produces an all-pass response in the subsequent nodes at $z^{(2)}(2\Delta t)$, $z^{(3)}(3\Delta t)$, etc. The same thing will happen for each successive input pulse, provided that T > 1, and so the chain will propagate this input as an activity wave for however long a chain we might have.

Wave propagation in this case occurs because the weight setting was sufficiently large to induce the all-pass firing mode. The function performed by the chain is nothing more than simple relay action, however, and the chain basically behaves in the same manner as an axon conveying an action potential. This is of rather limited usefulness in terms of implementing signal processing functions, however. If one allows multiple inputs from different sources to converge on each node in the chain some limited set of logic operations could be implemented by the chain, but, again, the types of the logic operations possible are rather limited.

We call case (b) the **highpass filter mode** because in this case multiple input pulses must be applied to the chain to evoke the first AP response at node 1. Its elementary dynamics were also discussed in §3.2 of chapter 8 and depend on the inter-pulse interval T of the input tetanus. The firing rate of the input tetanus is $R_{in} = 1/(T\Delta t)$ in this case.

In chapter 8 we saw that this rate must exceed a minimum value before any firing response will be evoked from the first node in the chain. This minimum input rate is found by solving equation (8.8) with the result

$$R_{in} > \frac{1}{\tau \cdot \ln\left[\frac{1}{1 - w/(\tau \cdot \Theta)}\right]}. \tag{9.3}$$

Provided the condition set by (9.3) is met, the first node in the chain will respond with an AP output on the $(M^{(1)} + 1)$st input pulse, resulting in an output firing rate $R^{(1)} = R_{in}/(M^{(1)} + 1)$. Here $M^{(1)}$ is the smallest integer satisfying equation (8.10) and is given by the condition

$$M^{(1)} \geq -1 - \frac{\tau}{T \cdot \Delta t} \cdot \ln\left[1 - \frac{\tau \cdot \Theta}{w}\left(1 - \exp\left[-T \cdot \Delta t/\tau\right]\right)\right] > 0. \tag{9.4}$$

$R^{(1)}$ now becomes the input firing rate to the second node in the chain. Because this rate is lower than R_{in}, the interval between successive impulses into the second node becomes $T^{(2)} = (M^{(1)} + 1)\cdot T$. In order to evoke a firing response from the second node in the chain, the chain's input tetanus must now satisfy the constraint

$$R_{in} > \frac{M^{(1)} + 1}{\tau \cdot \ln\left[\frac{1}{1 - w/(\tau \cdot \Theta)}\right]}$$

or the output of the second node's LI will not exceed its firing threshold and the activity wave will cease to propagate. If R_{in} satisfies this constraint, the second node will fire on its $(M^{(2)} + 1)$st input pulse and produce an output firing rate $R^{(2)} = R_{in}/(M^{(1)} + 1)(M^{(2)} + 1)$ where

$$M^{(2)} \geq -1 - \frac{\tau}{T^{(2)} \cdot \Delta t} \cdot \ln\left[1 - \frac{\tau \cdot \Theta}{w}\left(1 - \exp\left[-T^{(2)} \cdot \Delta t/\tau\right]\right)\right] > 0 \ .$$

Because we are assuming highpass filtering mode operation, the quantity $\tau\Theta/w > 1$ and so we have $M^{(2)} \geq M^{(1)}$.

We can continue this analysis by induction. In order for the third node in the chain to fire, the input tetanus must exceed the rate

$$R_{in} > \frac{\left(M^{(1)} + 1\right)\cdot\left(M^{(2)} + 1\right)}{\tau \cdot \ln\left[\frac{1}{1 - w/(\tau \cdot \Theta)}\right]}$$

and so on for each successive link in the chain until either the condition

$$M^{(L)} \geq -1 - \frac{\tau}{T^{(L)} \cdot \Delta t} \cdot \ln\left[1 - \frac{\tau \cdot \Theta}{w}\left(1 - \exp\left[-T^{(L)} \cdot \Delta t/\tau\right]\right)\right], \quad T^{(L)} = T \cdot \prod_{\ell=1}^{L-1}\left(M^{(\ell)} + 1\right), \tag{9.5}$$

$L > 1$, cannot be met by any integer value of $M^{(L)}$ or the input rate R_{in} fails to satisfy the constraint

$$R_{in} > \frac{\prod_{\ell=0}^{L-1} \left(M^{(\ell)} + 1\right)}{\tau \cdot \ln\left[\frac{1}{1 - w/(\tau \cdot \Theta)}\right]}, \quad M^{(0)} = 0 \ . \tag{9.6}$$

Unless both conditions are satisfied, activity wave propagation will cease after the (L – 1)st node in the chain. We call the propagated activity wave an **evanescent wave** because it will eventually die out for any sufficiently long chain of nodes. Thus, an I&F chain cannot propagate a replication of the input tetanus in highpass filter mode operation.

In principle at least, the simple I&F chain operating in evanescent wave mode can be regarded as performing a mathematical function, namely encoding a spatial representation of the input firing rate, since the length of the portion of the chain responding to the input tetanus depends on the input firing rate. The mapping operation involved in representing the input rate by means of the output signals of the chain is not particularly straightforward but presents no mathematical difficulty in its calculation. Because the number of links in the chain is always an integer, the mapping will be unable to uniquely resolve all possible R_{in} values; information theorists call such a situation an *information lossy code*. This is not necessarily a drawback, as the terminology might seem to imply, because *all logico-arithmetic mathematical functions are information lossy*. An operation that is information lossless merely relays information or transforms its data representation to another form but otherwise performs no mathematical function on it[6]. However, your author is not aware of any published examples where this property of the simple evanescent mode I&F chain has been exploited to date.

§ 3.2 Propagation in Eckhorn Chains

The absence of a reset after firing and the presence of a relative refractory period in the Eckhorn model produces differences in the response of a simple chain to an input tetanus in comparison with the I&F chain. Figure 9.5 reproduces the block diagram of an ENU (Eckhorn neural unit) earlier provided in chapter 8. The Eckhorn model is not commonly used in simple chain configurations. However, when it is the linking field input and associated leaky integrator are omitted because there is nothing in the chain structure that makes use of the linking field.

[6] Consider the elementary operation of integer addition. If the outcome of the addition is 7, we cannot tell from this result if the input was 0 + 7, 1 + 6, 2 + 5, or etc. To the information theorist this means there was information at the input to the adder that cannot be recovered from knowledge of the output alone and, therefore, the operation is information lossy. The mathematical function is information lossy but the nature of the information loss *is carefully constrained* such that a useful function, addition, results.

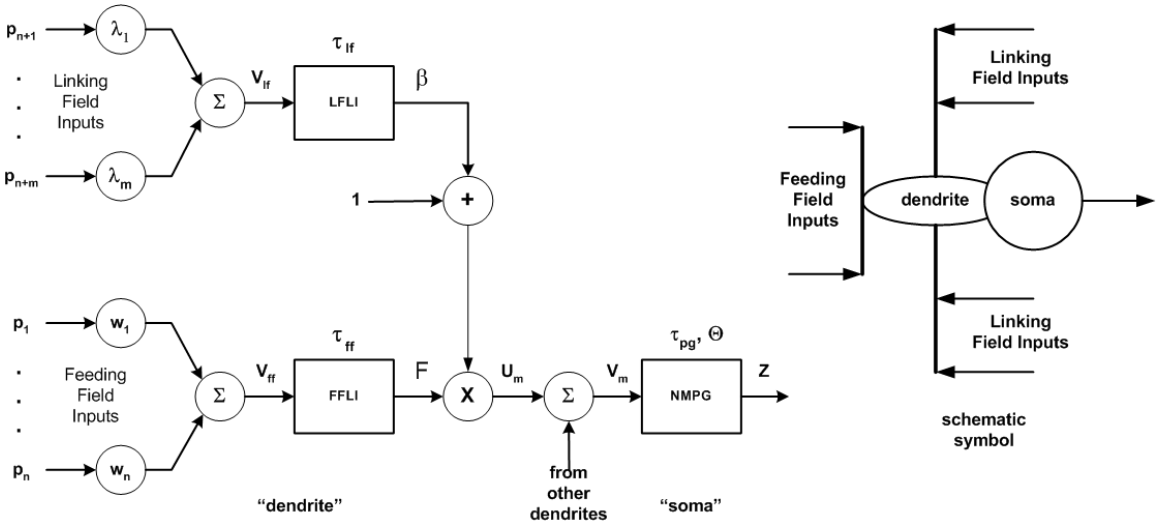

Figure 9.5: Block diagram (left) and schematic symbol (right) of an Eckhorn node. In a simple chain the linking field is omitted. Additional dendrites are used for inputs with differing time constants τ_{ff}.

We will again assume a single tract input with weight w driven by a succession of impulses occurring at times t_k as before. The neuromime pulse generator (NMPG) is described by equations (8.13) from chapter 8 (§3.4) and the feeding field leaky integrator action after discretization is described by

$$F(t+\Delta t) = F(t)\cdot \exp\left[-\Delta t/\tau_{ff}\right] + \frac{1}{\tau_{ff}}\cdot V_{ff}(t) \tag{9.7}$$

where $V_{ff} = w$ if there is an impulse at the feeding field input at time t and 0 otherwise. We will again assume a relaxed state, i.e. $F(t) = 0$ for $t < 0$, and the arrival of the first input at time $t = 0$. Owing to the absence of a linking field or other dendrites, we have $V_m = U_m = F$ in Figure 9.5.

§3.2.1 The Rate Multiplier and All-Pass Modes of Operation

We again divide our analysis on the basis of whether $\Theta_0 < w/\tau_{ff}$ or $\Theta_0 > w/\tau_{ff}$. In the first case, the NMPG responds to the input by firing a unit impulse and setting $\Theta(\Delta t) = \Theta_0 + A$. The next input pulse coming into the chain does not arrive until $t = \mathrm{T}\cdot\Delta t$ and what happens during the interval from $t = 0$ to $t = \mathrm{T}\cdot\Delta t$ depends on w, A, and the two time constants in the ENU. If during this interval the condition

$$F(t) = \frac{w}{\tau_{ff}}\cdot \exp\left[-t/\tau_{ff}\right] \geq \Theta_0 + A\cdot \exp\left[-t/\tau_{pg}\right] = \Theta(t) \tag{9.8}$$

the NMPG threshold Θ declines below the residual value of V_m and the NMPG fires again. We call this the **rate multiplier mode** of ENU operation because a single input pulse produces more

than one output pulse. Equation (9.8) is a nonlinear equation with no known closed-form solution and must be numerically or graphically analyzed to determine if rate multiplier mode (RMM) operation is effected by the ENU. RMM operation is not overly-common in applications of Eckhorn-based neural networks and has received very little study to date. As this is still for all practical purposes a latent field of research, we will not consider it any further in this book.

If $A > w/\tau_{ff}$ and $\tau_{ff} < \tau_{pg}$ RMM is usually not exhibited by the ENU and we have the Eckhorn version of all-pass mode operation. These conditions come from the following consideration. A sufficient condition for violating (9.8), and therefore preventing RMM operation, is

$$\frac{w}{\tau_{ff}} \cdot \exp\left[-t/\tau_{ff}\right] \leq A \cdot \exp\left[-t/\tau_{pg}\right] .$$

The maximum value for t in the interval between input pulses is $t = t = \mathrm{T} \cdot \Delta t$. After substituting this into the inequality above and some minor manipulation we obtain

$$\mathrm{T} \cdot \Delta t \cdot \left(\frac{1}{\tau_{pg}} - \frac{1}{\tau_{ff}} \right) < \ln\left(\frac{A \cdot \tau_{ff}}{w} \right).$$

With $A > w/\tau_{ff}$ the right-hand side of this inequality is positive and with $\tau_{ff} < \tau_{pg}$ the left-hand side is negative. Therefore RMM operation is prevented for any value of T as a response to the first input pulse into the ENU. This is not sufficient to ensure RMM will not be evoked by some later input pulse in the input tetanus because the feeding field LI is not reset by NMPG firing. However, in a great many cases the parameters of the system are such that RMM is not evoked and we next consider the all-pass mode of operation.

When the feeding field LI is driven by an input tetanus with input pulses occurring at fixed times $t_k = k \cdot \mathrm{T} \cdot \Delta t$, solution of difference equation (9.7) by induction results in

$$V_m\left(t_k + \Delta t\right) = \frac{w}{\tau_{ff}} \cdot \sum_{m=0}^{k} \exp\left[-m\mathrm{T}\Delta t/\tau_{ff}\right] = \frac{w}{\tau_{ff}} \cdot \frac{1 - \exp\left[-(k+1)\mathrm{T}\Delta t/\tau_{ff}\right]}{1 - \exp\left[-\mathrm{T}\Delta t/\tau_{ff}\right]} \tag{9.9}$$

with

$$V_m\left(t_k\right) = \frac{w}{\tau_{ff}} \cdot \frac{1 - \exp\left[-k\mathrm{T}\Delta t/\tau_{ff}\right]}{1 - \exp\left[-\mathrm{T}\Delta t/\tau_{ff}\right]} \cdot \exp\left[-(\mathrm{T}-1)\Delta t/\tau_{ff}\right] . \tag{9.10}$$

If the ENU fired once after the previous input pulse then its threshold will be approaching

$$\Theta\left(t_k + \Delta t\right) \to \Theta_0 + A \cdot \exp\left[-(T-1)\Delta t/\tau_{pg}\right] \tag{9.11}$$

with $\Theta(t_m + \Delta t) = \mathrm{Q}(t_k + \Delta t)$ for all indices $0 < m < k$. Now, for an input tetanus of impulses at

time indices t_m, the feeding field LI output will produce values $V_m(t_m + \Delta t) < V_m(t_k + \Delta t)$ throughout the series of inputs. Therefore, a sufficient condition to ensure firing of the ENU at each input pulse is

$$\Theta_0 + A \cdot \exp\left[-(\mathrm{T}-1)\Delta t/\tau_{pg}\right] \le \frac{w}{\tau_{ff}} \cdot \frac{1-\exp\left[-2\mathrm{T}\Delta t/\tau_{ff}\right]}{1-\exp\left[-\mathrm{T}\Delta t/\tau_{ff}\right]} \tag{9.12}$$

where we have merely set $k = 1$ in (9.9). Note, however, that V_m will asymptotically approach

$$\lim_{k\to\infty} V_m(t_k + \Delta t) \to \frac{w}{\tau_{ff}} \cdot \frac{1}{1-\exp\left[-\mathrm{T}\Delta t/\tau_{ff}\right]} \tag{9.13}$$

for a long tetanus and so (9.12) is not a sufficient condition to ensure RMM operation is not eventually evoked as the stream of input pulses continues indefinitely. Nonetheless, there is a broad range of parametric values for the ENU within which its operation remains all-pass and RMM is not evoked by long input sequences.

As the foregoing analysis shows, the equations governing operation of an ENU in all-pass mode are significantly more complicated than is the case for the I&F model operating in APM. There is little to be gained by employing the ENU rather than an I&F node in a simple APM chain unless one desires to allow RMM operation within this chain. For this reason it is much more common for the I&F node to be employed for APM chains. The same is not true, however, when a chain is desired that operates in highpass filter mode, which we discuss next.

§3.2.2 The Highpass Filter Mode of Operation

When $\Theta_0 > w/\tau_{ff}$ we have the highpass filter mode of operation for the ENU. Behavior of the ENU for the first evoked AP firing is identical to that of the I&F node in highpass filter mode. The first M inputs to the ENU evoke no output response and the ENU's first output pulse occurs on the (M + 1)st input pulse with M given by (9.4) with $\tau = \tau_{ff}$. Here is where the similarity between the two models ends because the feeding field LI (FFLI) is not reset by the firing action. For the m^{th} input pulse following the one that evoked the firing response the FFLI produces an input to the NMPG given by (9.9) with $k = (\mathrm{M} + m)\cdot\mathrm{T}\cdot\Delta t$. If the NMPG has not re-fired already, its threshold is given by (9.11) and if V_m exceeds Θ the ENU will fire again. The most important special case occurs when τ_{pg} is several times smaller than $\mathrm{T}\cdot\Delta t$ but still large enough to prevent RMM operation. In this case Θ will have decayed to approximately Θ_0 at $m = 1$ and, because V_m is larger than this value, the ENU will fire again at the (M + 2)nd input pulse and at every input pulse thereafter. Thus, after suppressing the first M input pulses, the ENU begins firing with no

decrease in the resulting output firing rate of the ENU relative to the input firing rate.

This means a simple chain of Eckhorn neural units in highpass filter mode can propagate an input tetanus for any chain of arbitrary length provided the input tetanus is maintained. There is, however, a cautionary note that must be made. If the input tetanus is of finite length, each ENU node in the chain will suppress the first M inputs it receives, resulting in an output pulse pattern equal in rate to the input tetanus but shorter in length by an additional amount M. As this pulse **packet** travels down the chain, the number of pulses it contains decreases at each succeeding link in the chain until eventually the number of output pulses in the packet falls below M. At this point the activity wave will propagate no further. The Eckhorn chain in highpass mode propagates **evanescent wave packets**. Figure 9.6 illustrates this behavior. Because M depends on T, the outputs of the chain can be regarded as an encoding of the input tetanus firing rate and duration.

This behavior for the simple Eckhorn highpass filter chain appears to be more consistent with current models of how signal processing is taking place in the brain (specifically, in the neocortex) than is the behavior of the simple I&F chain in highpass filter mode. It is currently conjectured that the mental phenomena of perception, memory, and consciousness is related to synchronous activity (possibly with phase delays) of multiple neuronal assemblies [BUZS], [STER]. In this conjecture, relatively low frequency activity in a neuronal region is thought to produce relatively short-range higher-rate signals that propagate over a small distance to other assemblies and serves to synchronize their low-frequency local activities ("oscillations") with each other. This action, if in fact this is what is happening, is called **binding** [BUZS], [ECKH2-3], [STER]. Propagation of evanescent wave packets at higher activity rates is more consistent with this conjectured binding mechanism than is the progressively slower-rate propagation by simple I&F chains. However, it is also clear that simple chains (or, for that matter, the chains we consider in the next section) cannot be the entire story if the binding conjecture is correct.

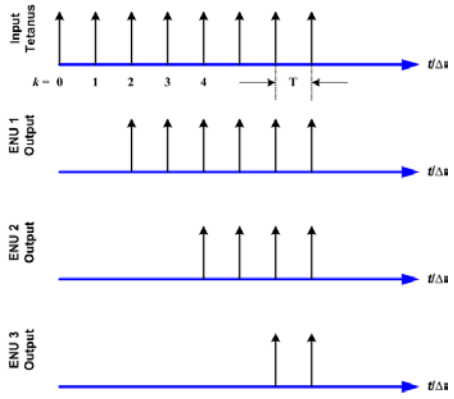

Figure 9.6: Illustration of simple Eckhorn highpass chain firing with M = 2.

The lowest activity rate, R_c, for which the incoming tetanus will evoke a firing response in highpass filter mode is called the **cut-in rate**. This is the rate for which the FFLI output is able to reach the baseline threshold Θ_0 for an infinitely long input tetanus. The cut-in rate is given by

$$R_c = \frac{1}{T_c \Delta t} = \frac{-1}{\tau_{ff} \ln\left(1 - w/\left(\tau_{ff}\Theta_0\right)\right)} \ . \tag{9.14}$$

The rate $R_p = 1/(T_p \Delta t) > R_c$ for which the ENU will respond with a designed value for M is called the **passband rate** for M. R_p is defined to be the input tetanus rate at which the output of the feeding field LI first achieves baseline threshold on the (M+1)st input pulse,

$$V_m = \frac{w}{\tau_{ff}} \frac{1 - \exp\left[-(M+1)T_p \cdot \Delta t/\tau_{ff}\right]}{1 - \exp\left[-T_p \cdot \Delta t/\tau_{ff}\right]} = \theta_0 \, . \tag{9.15}$$

(9.15) can be numerically solved for a specified value of M to obtain parameter T_p given the other parametric values or, contrariwise, both M and T_p can be specified to obtain design relationships for determining Θ_0, w, and τ_{ff}.

§ 4. Synfire Chains

Many neuroscientists find it hard to avoid the fallacy of thinking about population model structures, such as the simple chain just discussed, as if the nodes in the chain were neurons rather than complex structures composed of hundreds or thousands of neurons. Nonetheless, from time to time fruitful new ideas are produced even from such a faulty picture. One such fruitful idea is illustrated in Figure 9.7, the **synfire chain**.

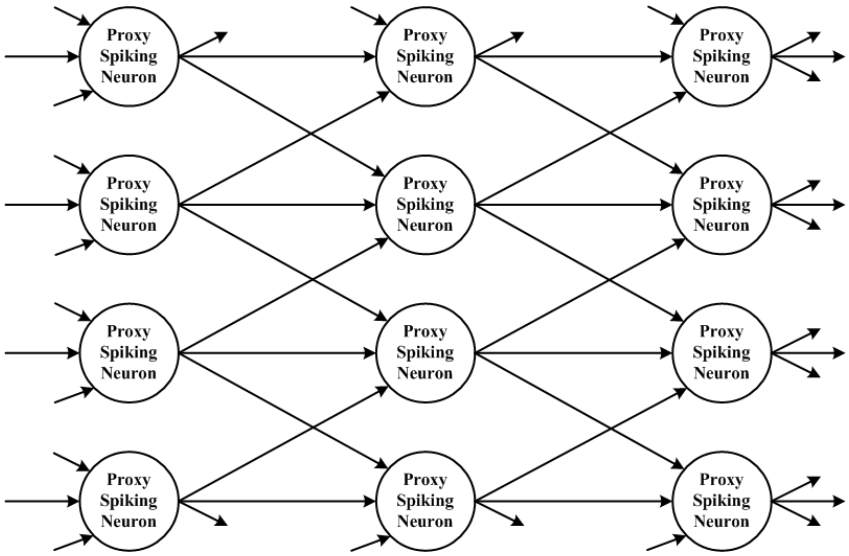

Figure 9.7: The synfire chain.

The term "synfire" was coined to mean a chained, layered structure of neurons in which firing activity was synchronized. Abeles first proposed the synfire chain hypothesis in 1982 as a means of explaining how precise spatiotemporal firing patterns could be generated in a noisy cortical environment. The main theoretical argument for the actuality of the existence of synfire chains in the CNS is as follows:

> The cortico-cortical excitatory synapses are weak and the firing rate is low. In this situation the neurons are insensitive to dispersed excitatory inputs. They act like coincidence detectors, not like integrators [ABEL2]. . . Thus, if a pool of neurons is activated by an asynchronous volley of spikes its output is more likely to be synchronized than its input. The weakness of cortico-cortical synapses also suggests that intracortical processing is carried by pools of neurons activating each other through multiple diverging and converging connections.
>
> Suppose we are confronted with a chain of pools of neurons connected by diverging/ converging connections as shown in [Figure 9.7]. Suppose also that the first pool is excited by a continuous depolarizing current which produces an unsynchronized elevation of firing rate in the pool. If this activity is to propagate at all through the network, it would do so in a synchronized mode due to the following process. The second pool would preferentially pick up instances at which several of the neurons in the first pool fired synchronously by chance. Thus the activity of the neurons in the second pool would show a larger degree of internal synchrony. This would amplify itself in the third pool, etc. Put in other words a chain of neuronal pools linked by diverging-converging connections acts like a spatio-temporal filter for synchronism. Such a chain is called a synfire chain. [ABEL5]

When proposed and for many years thereafter the synfire chain hypothesis, as well as the very existence of synfire chains, was quite controversial. It is not without controversy and criticism today. Nonetheless, more or less direct experimental evidence consistent with the existence of synfire-chain-like cortical organization has been reported:

> We find precise repetitions of spontaneous patterns of synaptic inputs in neocortical neurons in vivo and in vitro. These patterns repeat after minutes, maintaining millisecond accuracy. Calcium imaging of slices reveals reactivation of sequences of cells during the occurrence of repeated intracellular synaptic patterns. The spontaneous activity drifts with time, engaging different cells. Sequences of active neurons have distinct spatial structures and are repeated in the same order over tens of seconds, revealing modular temporal dynamics. Higher order sequences are replayed with compressed firing. . .
>
> Intracellular recordings and imaging experiments reveal the widespread existence of repeated dynamics in the spontaneous activity of neocortical circuits in vitro and in vivo. . . Our results demonstrate that the apparently stochastic cortical synapses can work with high reliability to produce stereotypical dynamics that are reduced in dimensional space. Our data agree with the prediction that cortical activity flows through chains of synchronized neurons (synfire chains), which are reactivated with high temporal precisions. Moreover, we describe a higher order grammar by which these chains themselves can be modules of larger temporal structures (cortical songs) defined by their sequence of activation and which can last for minutes. . . The mechanisms that generate and propagate synfire chains and cortical songs must be intrinsic to the cortical circuit because they are preserved in slices, and might reflect the faithful reactivation of specific circuits, mediated by stereotypical synaptic dynamics, and driven by pacemaker cells. . . We conclude that the neocortex can spontaneously generate precisely reverberating

> temporal patterns of activation, dynamical assemblies that could represent endogenous building blocks of cortical function. [IKEG]

Let us take note of the fact that what the experimental evidence brings to light is temporal activity and not anatomical layouts. Direct anatomical evidence for physical structures such as those depicted in Figure 9.7 has not yet been obtained. Let us further note that the existence of *repetitive* temporal patterns of activity implicates feedback connections from downstream links in a synfire chain back into its upstream links. This feedback is not depicted in the figure but has been part of the synfire chain hypothesis from its beginning. Abeles writes:

> Can such precise spatiotemporal firing sequences (motifs) serve as a code that can be recognized by other neurons? Probably not. According to the synfire chain hypothesis, what can be recognized is a wave of synchronous firing that travels through the synfire chain. These sparse firing sequences are only a signature of synfire activity . . . Although the synfire chain hypothesis is very appealing, it still remains to be shown that synfire chains exist in the intact brain and that they are relevant to the behavior of animals and humans. [ABEL6]

It must also be noted that synfire chains are neither the only nor the first model of chain-like activity propagation to be proposed. In 1969 Grossberg proposed a map/network-system-level model of chain-like neural processing, which he called an **avalanche network** [GROS11-12, 21]. Experiments such as those reported by Ikegaya et al. [IKEG] are equally consistent with the hypothesis that avalanche chains exist in brain structure. It is not improbable that something like a synfire chain structure at the PCNN modeling level might be a lower-level building block structure for the composition of Grossberg's avalanche chains. Avalanche chain networks are presently a very under-studied research topic but recent findings suggest they may provide one very important network-system-level building block for explaining neural system functions at the system architecture level that can be directly related to observable behavior [WELL5-6]. Your author regards synfire chains, avalanche chains, and the Eckhorn PCNN structures in the next section to be closely-coupled research topics.

Let us now look at the synfire chain structure. We will assume each node in the network receives L_{tot} afferent tract inputs. Let us call this the **fan-in** of the node. For the sake of convenience we will assume each input has an input weight of w. We will further assume each node is modeled by an I&F proxy. The central idea in the synfire chain hypothesis is that each layer in the chain responds with output pulses from some subpopulation in that layer only if the nodes within that subpopulation each receive some $L_n \le L_{tot}$ more or less synchronous input pulses in a single volley. In this context synchrony means the input pulses all arrive within a very short time-span of each other such that each node in the subpopulation gets a total input excitation

$$V = L_n \cdot w \equiv \mu \ge \tau \cdot \Theta . \tag{9.16}$$

We will call L_n the **critical stimulus level**. If for convenience we assume the pulses making up the input volley all occur at time t = 0, the LI output in response to the volley is

$$y(t) = \frac{\mu}{\tau} \cdot \exp(-t/\tau), \quad y(0) \geq \Theta . \tag{9.17}$$

Because this LI output exceeds the firing threshold of the I&F node, the node responds to it by producing an output pulse of its own and resetting the LI. This means the combination of synchronous inputs for the responding subpopulation puts each node in that subpopulation into the all-pass mode of operation. Note that it is not necessary for each node within the subpopulation to receive the same number $L \geq L_n$ of synchronous inputs. All that is required is each node in that subpopulation is excited into the all-pass mode by whatever particular volley it happens to receive. Once this is achieved, the particular value of μ for each subpopulation node is completely irrelevant because the I&F node then resets the LI to zero once more.

If a large enough subpopulation in the layer responds with an output volley of its own, it will evoke the all-pass response in another subpopulation (not necessarily of the same size) in the next layer in the synfire chain. If this second-layer subpopulation is large enough, it will evoke the all-pass response from a third-layer subpopulation, etc. and a traveling activity wave can ensue that propagates down the chain.

As can be seen in Figure 9.7, each node in a given layer projects its output to some N nodes in the next layer. We can call this the **fan-out** of a node. If N is less than the number of nodes in the next layer the synfire chain is said to be an **incomplete** chain [ABEL1, pg. 216]. Otherwise it is called a **complete** chain. Based upon various statistical approximations of cortical organization, Abeles argues that incomplete chains are more likely to be representative of cortical organization than are complete chains [ABEL1, pp. 216-221]. The argument is based on neuron-level modeling in randomly connected cortical networks and for these reasons is not particularly compelling at the level of the modeling hierarchy with which we are here concerned (because our nodes are cell groups, not neurons). However, we should also note that if not all the nodes in one layer fire synchronously, the effect at the next layer is more or less the same as we would have for an incomplete chain even if the network itself is modeled as a complete chain. For that reason we will consider only incomplete chain network structures (Figure 9.7-like connectivity) in this book.

The diagram of Figure 9.7 is drawn such that fan-in (L_{tot}) and fan-out (N) are equal for each node, but this is not a general restriction on a synfire net and Abeles' argument tends to imply that biologically realistic modeling would demand $L_{tot} < N$, although in practice synfire modelers have tended to set $L_{tot} = N < \mathbb{W}$ where $\mathbb{W}$ is the width of (number of nodes in) the layer [ABEL3], as

has been done in Figure 9.7. Abeles also argues that the source-destination connections from one layer to the next should be randomized, an argument that helps one try to apply region-II reasoning to this region-III system but is rather inconvenient insofar as a *quantitative* theory of the *signal processing function* performed by a synfire chain is concerned.

Let us assume $L_{tot} = N < \mathbb{W}$ but allow the layer-to-layer connections to be non-random and regular (that is, the same between each pair of layers). Examination of Figure 9.7 then illustrates an important first fact, namely that with equal *w* values for each connection the critical number of synchronous inputs L_n required to evoke firing response must be $L_n < L_{tot}$ because each node in the next layer receives inputs from a different collection of nodes in the previous layer. If $L_n = L_{tot}$ then at least some of the nodes in the next layer would not fire in response to the first layer volley (e.g. the nodes at the outer edges of the layer) even if all nodes in that first layer fired. In turn, *no* nodes in the *third* layer could receive a sufficient number of inputs to fire and the synfire wave would die out almost at once. This can be compensated for by allowing the weights at the "edge" nodes to be higher-valued (which in effect lowers L_n at these nodes), but this seems *prima facie* to be a rather artificial trick for the sole purpose of getting a synfire wave to propagate. If we wish to maintain equal *w* values for all input tracts (which is more or less the network-level counterpart to the average neuron of chapter 8) then we must require as a condition that $L_n < L_{tot}$. What may we then expect?

Wave propagation characteristics of the synfire chain depend not only on L_n but also on the connectivity patterns between layers and the distribution of firing nodes in the layers. At the time of this writing no mathematical theory for synfire node behavior has been reported other than the aforementioned various statistical (region-II-like) modeling approaches, although Hayon et al. have reported a mathematical "wave mechanics" model of synfire waves [HAYO]. What we will present here are some graphical examples that illustrate the principal qualitative features we wish to emphasize. We adopt the naming convention of referring to these examples as $\mathbb{W}$-J-L_n ("d") where $\mathbb{W}$ is the width of the chain, J is the number of nodes initially firing in the first layer, and ("d") is an optional descriptor of the distribution pattern of firing nodes in the first layer.

We will consider examples with $\mathbb{W} = 6$, $N = L_{tot} = 4$ and $L_n < L_{tot}$. Figures 9.8 illustrate our first example, a pair of 6-4-3 cases. In Figure 9.8A we have $J = L_n + 1$ and the feedforward connections from the firing subpopulation into the second layer are made to neighboring nodes, i.e. the output activities of the first layer project into a common second layer sub-pool. As the figure illustrates, the initial firing pattern evokes responses in $J_2 = 4$ nodes of the second layer, and these responses succeed in evoking a third-layer response, etc. A synfire wave is therefore successfully launched by the initial stimulation of layer 1.

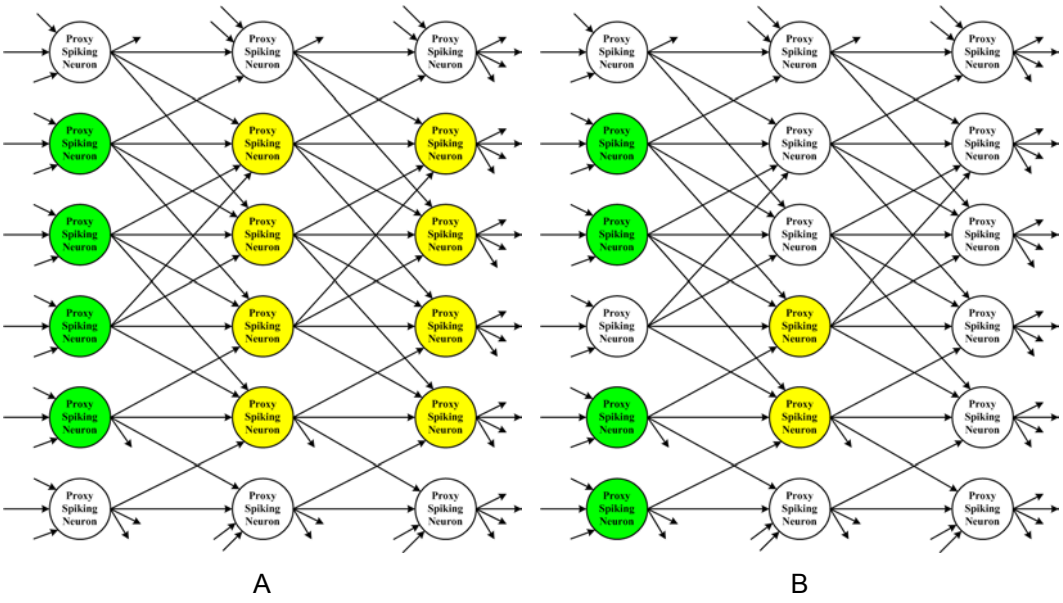

Figure 9.8: 6-4-3 examples. A: 6-4-3 case where nodes in the first-layer firing pool comprise a subpopulation making immediately neighboring forward projections to the second layer. B: 6-4-3 (split) case where the initial subpopulation of firing nodes in the first layer do not form a subpopulation of immediately neighboring forward projections.

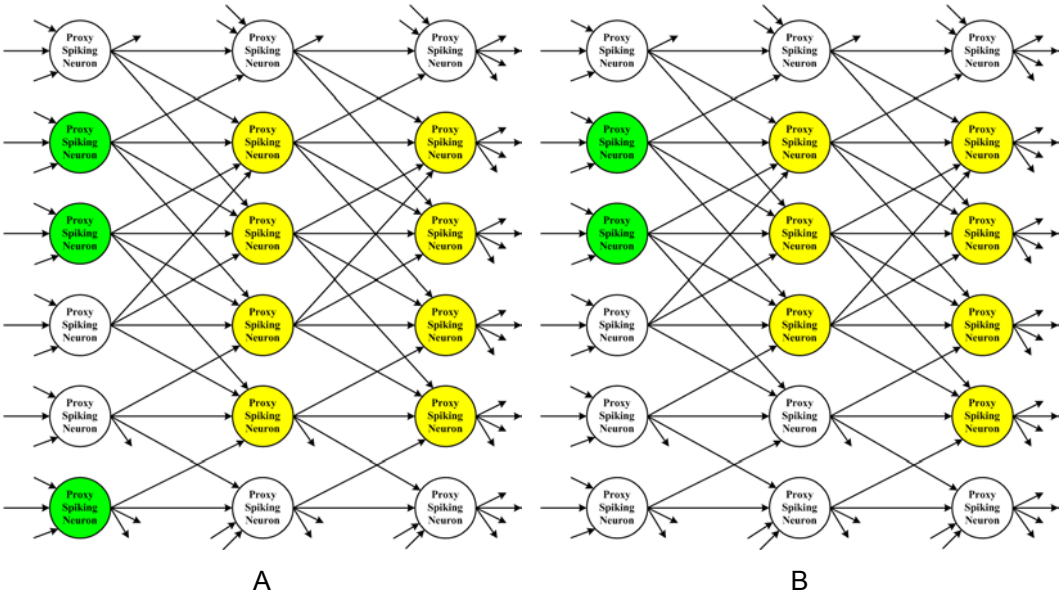

Figure 9.9: Examples of successful synfire wave stimulation. A: 6-3-2 (split) example where sufficient connectivity overlap into the second layer exists to evoke firing. In this case 4 nodes are stimulated and the wave thereafter propagates. B: 6-2-2 example where a synfire wave builds up over the course of the next two layers and thereafter successfully propagates.

Figure 9.8B is a 6-4-3 (split) case and here the outcome is very different. Outputs from the first layer are scattered among the population of the second layer such that only two nodes in the second layer have evoked responses. This falls below L_n and so a synfire wave fails to become established because of the connectivity pattern and distribution of activated first layer nodes.

That it is the combination of J, L_n, and the connection pattern between layers that determines whether or not a synfire wave is established, and not any one of these factors all by itself, is illustrated by the examples given in Figures 9.9. Figure 9.9A is a 6-3-2 (split) case where, again, we have $J = L_n + 1$ and a separation between activated nodes in the first layer population. There is, however, sufficient overlap of connection in the second layer to not only produce $J_2 > J > L_n$ firing responses in the second layer but also to immediately establish a propagating synfire wave. Figure 9.9B is a 6-2-2 case (i.e., $J = L_n$) illustrating $J_2 = 3 > J$ in the second layer followed by completing the establishment of a stable propagating synfire wave ($J_3 = 4$) in the third layer. Thus in this case we have a build-up transition from the stimulating layer to the eventual synfire wave pattern itself.

These graphical illustrations of our four examples offering a tantalizing hint for an approach to developing an algebraic-graph theoretic approach for *quantitative* mathematical modeling of crystalline neural network synfire chains. Figures such as 9.8 and 9.9 are called **trellis diagrams** by information theorists. A mathematical theory for analyzing trellis diagrams for constrained communication channels and for convolutional error correcting codes exists [WELL7, chapters 2 and 6]. A synfire chain is *prima facie* not the same as either of these types of systems, but the structure of a synfire chain nonetheless appears to be quite closely related to the structures of these two other types of systems. It may therefore be possible, with appropriate additions and modifications of a relatively minor nature, to extend the body of theory developed by information theorists to cover the case of the synfire chain. If it does turn out to be possible to obtain a mathematical description of synfire chains using trellis analysis techniques, then it will also be possible to extend the theory of synfire chains to take in the case where feedback loops exist within the chain, a case that at present has no general mathematical theory aiding in our quantitative understanding of synfire dynamics. At the time of this writing, this idea is no more than a conjecture and quantitative analysis of synfire chains remains an active research topic.

Why go to the trouble of constructing synfire chain models? After all, if all we wished to do was a relay-like propagation of activity down a chain, a simple chain of I&F nodes in all-pass mode would suffice. The neuron-level speculations that synfire chains exist in the brain is, at the present time, still no more than a conjecture and, in any event, the current synfire paradigm is placed at a point in the modeling hierarchy of Figure 9.1 far closer to biology than to psychology, whereas modeling at the level of this chapter is neural network modeling and lies closer to psychology than to biology. These criticisms are true so far as they go, but they overlook an important fact about synfire chains.

In the synfire chain diagrams presented above some of the nodes make projections *outside* the

synfire chain. Where do these go? One answer is *to other synfire chains*. Abeles et al. have shown interlocking synfire chains provides a possible model for feature binding [ABEL4]. He writes,

> [Bienenstock] posits that only a few cross-links between two synfire chains that obey appropriate timing constraints are necessary to assure that both synfire chains will act as a wider and more stable synfire chain. This is turn could lock into activities in other synfire chains, and so on. Thus a large structure of synfire chains can be dynamically generated to represent binding of many simple components into a meaningful composite mental representation. This type of binding is dynamic, so that a single synfire chain may, under different conditions, lock into a number of other synfire chains and therefore be part of numerous distinct composite representations. In Bienenstock's view, the most important aspect of such a dynamic binding is vertical binding, in which a higher-level concept binds to lower-level elements that compose it.
>
> Abeles et al. [ABEL7] have shown, by way of simulations, that two synfire chains with random connections may learn to lock in to each other if activated synchronously several times and if the synaptic modifications follow a time-asymmetric rule (i.e., synapses are strengthened if the presynaptic spike precedes the postsynaptic spike and weakened if the time order is reversed). . . Hayon has shown by way of simulation that hierarchical systems of synfire chains could be used to separate figure from ground, to implement a minimal description length (MDL) principle, and to exhibit the property of compositionality. However, these demonstrations were toy problems. It remains to be shown that synfire chains can efficiently solve such problems in real world situations. [ABEL3]

The possibility that systems of interlocking synfire chains might be capable of supporting binding (e.g., of "feature fragments") as well as representation functions that could be tied to higher-level *mental* phenomena is too important to ignore. Systems of cross-linked synfire chains represent an approach to dealing with cell *assembly* phenomena that, *prima facie*, are consistent with other conjectures, such as those of Damasio and of Malsburg, that have been deduced from entirely different (and fundamentally psychological) considerations. Further, if synfire-like-chains are building blocks for map and network system level models such as those of Grossberg's avalanche chains, a quantitative mathematical theory of synfire chains would likely prove to be a key link in the modeling hierarchy roadmap connecting biology and psychology.

We close this section with one final note. A system structure in which we have cross-linked synfire chains makes the definition of "a" synfire chain somewhat ambiguous. Put another way, in the language of set membership theory, the boundary between one synfire chain and another is "fuzzy" rather than "crisp." This is reflected in Abeles' remark about "wider" synfire chains above. That the boundary is fuzzy is what makes explicitly clear the fact that a quantitative general theory of synfire chain *systems* is a theory of cell assemblies. This arguably provides an opening for exploration of Malsburg's conjecture regarding dynamic link architectures [MALS3].

§ 5. Eckhorn Networks

With the exception of a network we will call the Johnson PCNN, PCNNs constructed with the

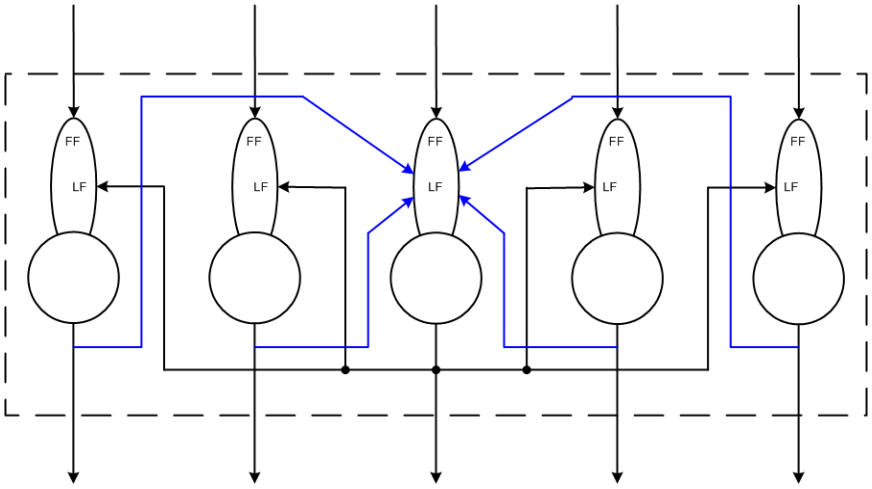

Figure 9.10: The basic Eckhorn PCNN cell assembly layer. It consists of N Eckhorn neural units laterally coupled to each other by linking field connections. For clarity, lateral connections are shown only for the central ENU in the figure. All ENUs in the layer make and receive similar connections with/from all other ENUs in the cell assembly. A layer of multiple overlapping assemblies is defined by the absence of linking field connections running laterally from some ENUs in one assembly with those of another.

Eckhorn neural unit are organized as chains of layered assemblies defined by linking field connections running laterally between all ENUs in the assembly. Figure 9.10 illustrates the basic Eckhorn cell assembly layer. The layer receives feeding field input tracts from a previous layer and projects its outputs to the feeding field inputs of a subsequent layer, thus forming an Eckhorn chain of assemblies. No general theory of Eckhorn PCNNs has yet been developed and published works to date largely focus on one- and two-layer arrangements of Eckhorn ENUs. Even here no general modeling (design) theory has yet been reported and the development of new Eckhorn network topologies (as well as parameter selection for reported topologies) is presently more art than science[7]. Our treatment must therefore be more qualitative than quantitative as we discuss the most common of these published topologies.

§ 5.1 The Basic Eckhorn Layer

The Eckhorn linking field exists to promote synchronization of the firing activities of the ENUs within an Eckhorn assembly. Typically all ENUs within a basic Eckhorn layer have identical ENU parameters with patterned linking field weights, λ. Let us assume that one ENU in the layer fires an output pulse at time $t = 0$. Referring to Figure 9.5 for our symbol definition, every other ENU within the assembly then receives a linking field input from this ENU, resulting in linking field response

[7] Most Eckhorn PCNN papers are published in the neuroscience literature, which has to date largely eschewed in-depth mathematical treatment and favored simulation studies accompanied by largely qualitative arguments.

$$\beta(t) = \frac{\lambda}{\tau_{lf}} \cdot \exp\left[-t/\tau_{lf}\right] . \tag{9.18}$$

If an ENU receives L simultaneous linking field inputs with average weight λ, this value is multiplied by L. The LFLI response produces a short-term modulation of the gain of the feeding field pathway, $G_{lf} = 1 + \beta$, as shown in Figure 9.5. Thus, $U_m(t) = [1 + \beta(t)] \cdot F(t)$. Typically τ_{lf} is much smaller than either the feeding field or NMPG time constants and so this gain modulation is short-lived. Note that linking field pulses cannot reach the NMPG input unless $F(t)$ is non-zero. Now suppose $F < \Theta_0$. If the linking field modulation boosts U_m such that $U_m \geq \Theta_0$, the ENU receiving linking field stimulation will now fire an output pulse of its own and also provide linking field stimulation to all other ENUs in the basic assembly layer (or layers if it belongs to more than one basic cell assembly layer). Because τ_{lf} is small, this promotes near-synchronous firing from all ENUs in the layer that have received feeding field inputs within a short time of the firing of the first ENU. Here a "short time" is typically on the order of a few milliseconds before or after the input to the ENU that fired first. This synchronizing window width depends on how many ENUs fire (which determines the values of $\beta(t)$ that result in the other ENUs).

With typical ENU parameters, the effect of the linking field is designed to briefly place the ENU in all-pass mode for its *feeding field* input tract. ENUs within the layer that do not receive feeding field inputs during this window do not fire. Thus, the output of the basic layer typically consists of some sub-pool of ENUs within it responding to poorly synchronized volley inputs with volley outputs that are more closely synchronized. If basic layers are connected in an Eckhorn version of a synfire chain, these outputs become better and better synchronized as the activity wave propagates down the chain. Note, too, that under certain input conditions and parameter values it is possible for the linking field effect to induce a rate-multiplier mode of operation for some or all of the ENUs in the layer. This RMM is a consequence of the *network* connectivity rather than of the individual ENU since the latter need not by itself be capable of RMM operation in the absence of its linking field connections. We can call this **network RMM**.

§ 5.2 The Eckhorn Spike Density Modulator

Eckhorn has reported some useful variations on the basic theme just presented [ECKH3]. Figure 9.11 illustrates one of these, the spike density modulator (SDM) [ECKH3-4]. The SDM was originally proposed as a network function for separating different feature representations that might be represented within the same set of input tracts arriving at the input of an Eckhorn layer. The underlying idea was that two adjacent objects represented by signaling via a common tract can be segmented by producing a robust phase shift in output signals representing these objects.

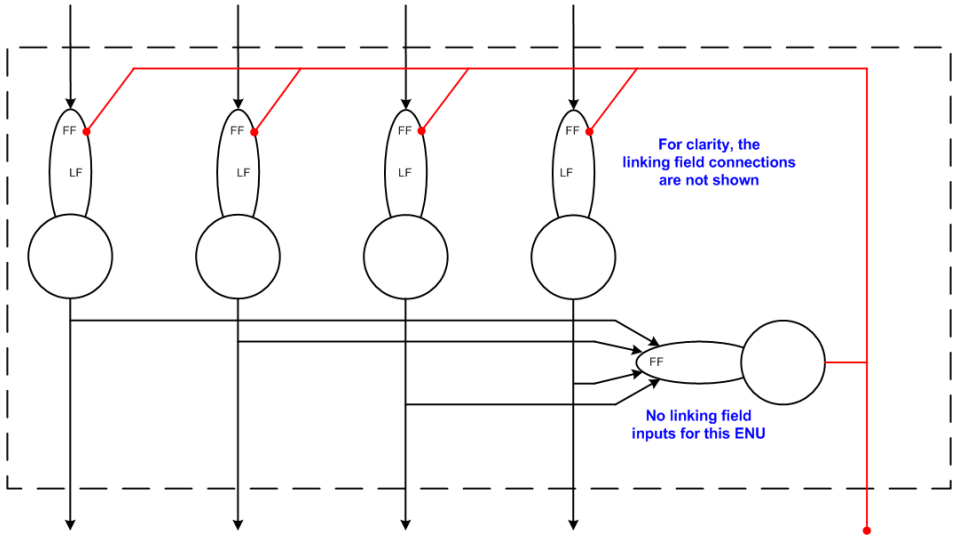

Figure 9.11: The spike density modulator (SDM). This assembly adds a single ENU receiving excitatory inputs from a basic Eckhorn layer and providing inhibitory feeding field input feedback to every ENU in the basic layer assembly. The inhibitory ENU typically has no linking field inputs, a larger feeding field time constant than the excitatory ENUs in the layer, and is set to operate in highpass filter mode. If the linking field weights among the excitatory ENUs are relatively weak and patterned in some non-uniform "footprint" (decreasing λ as the distance between ENUs increases), the layer can produce object segmentation among tract signals firing synchronously at the same rate by inducing relative phase shifts in the output tracts. If the linking field weights are uniform and not necessarily weak, the layer can produce pulse packeting of outputs based on their degree of lack of synchronization and/or differing firing rates across the tract inputs. This latter function gives this layer the name spike density modulator.

The notion of a "feature" is central in our current paradigm of the psychophysical relationship between biological activity and mental representation. In psychology a **feature** of a thing is an attribute of something (usually an object or event) that is critical to distinguishing that thing from other things. Nonetheless, this is a vague description (rather than a general definition) because the next question one asks is, "What identifies something in a signal or set of signals as representing a feature?" We will have more to say about this in chapter 15.[8] For now let us simply accept the idea that signals can possess some quality or qualities that can serve to represent a feature. For the SDM one such quality would be displacement in time (phase shift) between one vector of pulse volleys compared to others. In some models of brain function, e.g. Damasio's convergence zone model [DAMA2], objects and events are represented by binding together what may be called **feature fragments** represented by co-activities located in different parts of the brain. Thus the functions of feature segmentation and feature binding are central to neural system modeling.

[8] It is particularly important to realize that the idea of a "feature" is fundamentally the idea of a psychological object and is not an idea contained in and native to the definition of a signal. Rather, we may say a signal can *represent* a feature. This is not the same as saying a signal *is* a feature. Even the statement that a signal represents a feature is a *metaphysical*, not an empirical, assertion. The current mainstream psychophysical paradigm of neuroscience is materialist in character and does suffer from a number of fatal weaknesses. This is a topic discussed in depth in [WELL2-3].

Feature segmentation in the context of the SDM involves weakly-linked assemblies receiving tracts of input signals, arriving as synchronous volleys at the same activity rate, and outputting two (or more) sets of volley tracts separated in time (phase shifted) by some amount $0 < \Delta < T$ where T is the firing period of the incoming volleys. It is *not* currently known if the brain actually carries out feature segmentation in this way. To your author's knowledge, so far the phenomenon of object segmentation by phase shift has never been observed (although it is thought that object segmentation by means of differing firing *rates* has been observed in a few example cases). Nonetheless, the conjecture of temporal coding by phase shift is not an unreasonable one and there is no *a priori* reason to think the brain could only possess *one* kind of encoding scheme. To carry it out requires *patterned* weighting of relatively weak linking field connections among the ENUs in an assembly. A graph of λ vs. the spatial separation (or, for that matter, merely the numerical enumeration) of ENUs in the layer is called a weight **footprint**. An SDM segmentation layer typically is given a characteristic footprint in which λ decreases with increasing distance between ENUs according to some predefined function, often a Gaussian function. This has the effect of specializing different sub-pools (with a possibly fuzzy boundary between sub-pools) for "recognizing" (encoding) different features. This amounts to *encoding by spatial structure*. Such a PCNN structure is analogous to the well-known fact that the somatosensory cortex of the brain develops a precise **topographical map** of the body at the system architecture level of modeling. Grossberg and Levine proposed an analogous method of feature encoding/detection at the network system level in 1975 [GROS15].

An SDM with strong (and possibly non-uniform) linking weights is capable of performing a second type of signal processing function, and this is the one that gives it its name. This is to provide a partial and improved degree of synchronization for unsynchronized afferent volleys occurring at differing firing rates in different tracts but within a common range of signaling bandwidth. In this case the SDM produces **packets** of output signal volleys separated by relatively long periods of silence in the SDM output tracts. Such synchronized spike *densities*, along with the suppression of more transient activities, tend to be dominated by the most strongly activated tracts [ECKH4], [WELL4], and tend to be more effective in stimulating downstream networks with improved "signal-to-noise ratio" for the packeted signals. Figure 9.12 illustrates the packeting of signals by means of spike density modulation. Wells et al. used this property of spike density modulation to demonstrate a PCNN model of neocortical functional columns capable of spontaneously generating and propagating gamma-band (28-90+ Hz) waves in a chain structure (§5.5). Eckhorn et al. have used this packeting mechanism for the modeling of gamma-band-linked receptive field emergence [ECKH4].

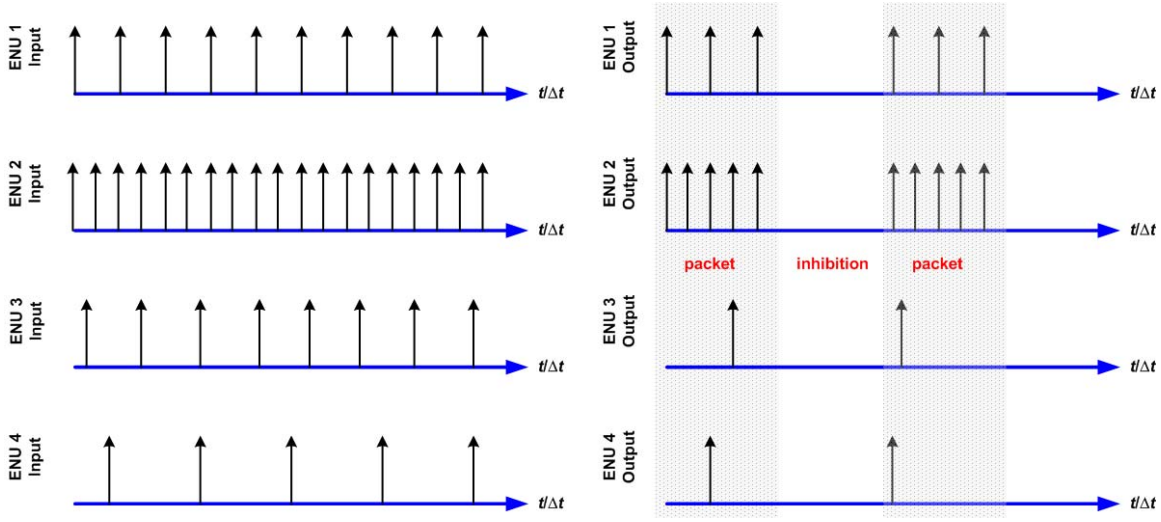

Figure 9.12: Illustration of packeting by an Eckhorn SDM.

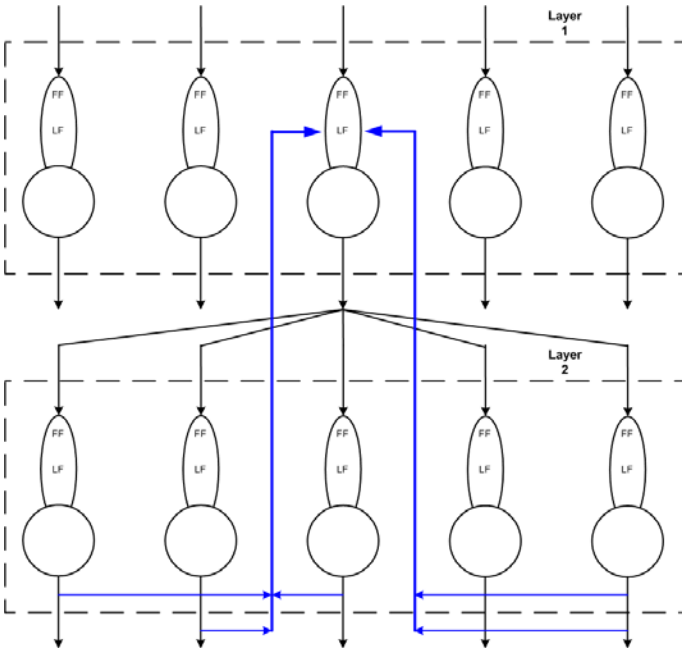

Figure 9.13: The two-layer Eckhorn network with top-down linking. The thick blue lines represent vector feedback of multiple linking field projections. For clarity only one feedforward projection into a layer-2 sub-pool is shown. ENUs in layer-2 receive multiple feedforward projections from ENUs in layer-1. Each layer also contains lateral linking field connections within the layer (not shown). ENUs in layer-2 provide linking field feedback to each ENU in layer-1 from which they receive feeding field inputs.

§ 5.3 Two-layer Eckhorn Network with Top-down Linking

Figure 9.13 illustrates a two-layer Eckhorn PCNN in which the second layer projects feedback via the linking field pathway into the first layer. The purpose of this top-down linking is to bind features represented by two activated subregions of the first layer into a single representation. Each node in the second layer receives feeding field inputs from several nodes in the first layer

and projects a linking field signal back to each first-layer node supplying it with a feeding field signal. In both layers, the nodes also make non-uniform lateral linking field projections to the other nodes in that layer with λ decreasing in value as the distance between nodes increases. The figure omits showing these connections to avoid cluttering the drawing with too much detail.

Typically the lateral linking field connections do not span the entire width of the layer, i.e., the layers are overlapping macro-assemblies of basic PCNN assemblies (Figure 9.10). Linking field weights decrease linearly with distance. Both layers make similar non-uniform lateral linking field connections with λ decreasing, e.g. linearly, with distance from source to destination node. For example, suppose layer-1 is composed of 50 nodes. Node 25 might make lateral linking field projections to layer-1 nodes 20, 21, 22, 23, 24, 26, 27, 28, 29, and 30. Five of these nodes, e.g. 23, 24, 25, 26, and 27, might project to a single layer-2 node. These five nodes make up the receptive field (RF) of the layer-2 node to which they project. The layer-2 node projects parallel linking field feedback back to nodes 23, 24, 25, 26, and 27 of layer-1, again in a patterned footprint.

Typically the number of nodes in layer-2 is less than the number of nodes in layer-1. Ratios on the order of 1:2 or smaller are typical. The layer-2 nodes are said to each have a larger receptive field (RF) than those of layer-1 because the of feeding field tracts converging from layer-1 at each layer-2 node. The RF of a layer-2 node is defined by the set of layer-1 nodes projecting to it. The RF of one layer-2 node is said to be *separate* from that of another because these nodes receive feeding field inputs from different sub-pools of layer-1 nodes. Top-down linking field feedback from layer-2 *indirectly* provides feedback across a wider span of layer-1 nodes because of the overlapping lateral linking field connections in this assembly. Top-down linking-field-induced firing of a layer-1 node can lead to the firing of other layer-1 nodes, which can in turn lead to firing of another layer-2 node; top-down feedback from this node can in turn lead to even more nodes firing in layer-1, etc. This would particularly be a likely dynamic if layer-1 operates in highpass filter mode in the absence of top-down feedback. Induced firing of layer-1 nodes would, of course, occur only in those nodes receiving sufficient nonzero activity inputs of their own; in the absence of feeding field signals, the linking field input of an ENU cannot reach the NMPG. Thus, firing activity can quickly spread across a wide breadth of layer-2, and this is the binding mechanism by which the RFs of layer-1 are said to be combined in a composite representation of feature fragments.

Eckhorn has used this network dynamic to model visual feature binding for the case where an object is divided into two parts by an obstruction (this is known as visual occlusion by vision and by image processing theorists) [ECKH3]. Despite the partial loss of data representing the visual object, the network is able to "fill in" the missing data and nonetheless "recognize" the occluded

object (assuming the object would have been recognized if the occlusion had not occurred). This phenomenon is a well-known capacity of the mammalian visual system. Although *data* is lost due to the occlusion, information theory tells us that the occlusion did not cause *information* loss (data and information are not the same thing [WELL7]), i.e. there was sufficient *redundancy* in the visual data to permit object recognition from the non-occluded data[9]. In this sense, the PCNN of Figure 9.13 can be called a "decoder" of information "encoded" by earlier signal processing in forming the RFs of the individual ENUs in layer-1. To date there has been very little mathematical theory developed for PCNNs such as that of Figure 9.13; it is an under-studied topic in computational neuroscience.

§ 5.4 Two-layer Eckhorn Network with Slow, Lateral Feeding Field Inhibition Feedback

Neuroscience currently hypothesizes that the strength of many types of stimuli are signaled in the early cortical networks (and in the peripheral nervous system) by modulations of spiking rates while features are characterized by the receptive field properties of neurons. At the modeling level we are dealing with here, the equivalent of this hypothesis is to say strength of stimulus is represented by activity ρ (9.2) and features by receptive field properties of population assemblies.

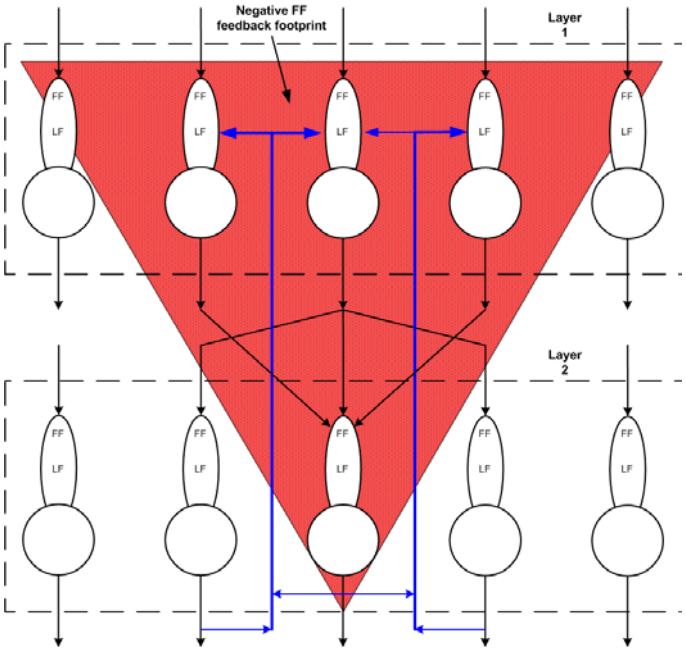

Figure 9.14: Two-layer PCNN of Figure 9.13 with additional negative feeding field feedback from layer-2 to layer-1. Negative feedback is from the layer-2 ENU to all layer-1 ENUs within the red-shaded triangle.

[9] It is a fundamental theorem of information theory that once information is lost it cannot be recovered by any means. Data, on the other hand, can be lost without information necessarily being lost if the information has been represented with redundancy in the data representation. This is the principle that makes error-detecting and error-correcting codes possible in communication systems.

In addition to the feature-fragment-binding function just discussed above, there is another representational function neural networks are called upon to perform. This is segmentation, the separation of signals into representations of individual objects present in the overall receptive field of the network system. The most famous example of this class of problem is the classical problem of "figure-ground" separation, the separation of one object in a background scene. A closely related example is provided by the so-called "cocktail party problem," in which you can "filter out" the hum of background conversations in order to listen to just one person who is speaking but, nonetheless, still remain cognizant enough of the background to, e.g., have your attention captured if someone else in the room speaks your name. Both problems call for the ability to bind feature fragments together to represent an object *and* to keep these objects separate enough that any one of them can be "brought into the foreground" as the others are "relegated to the background" using any sort of combination. The PCNN of Figure 9.14 is aimed at providing one key component for addressing compound problems of this sort.

In this network, the network of Figure 9.13 is modified by the addition of inhibitory *feeding field* feedback from each layer-2 ENU to both the subpopulation of layer-1 ENUs projecting to it *and* to neighboring ENUs *not* providing feeding field projections to the layer-2 node. The inhibitory feedback footprint is indicated by the red-shaded triangle shown in the figure. The inhibitory feeding field time constant τ_{iff} in the layer-1 ENUs is typically much larger (slower) than the excitatory feeding field time constant τ_{ff} by a typical factor of from 3- to 6-fold. (This makes the addition of a second "dendrite" desirable for the layer-1 ENUs because the inhibitory FFLI must now be different from the excitatory FFLI in each ENU; when this is the case, the inhibitory dendrite typically does not have a linking field and the inhibitory FFLI is summed in at the input of the NMPG as shown in Figure 9.5). Feedback weights are typically non-uniform across the inhibitory footprint, with a Gaussian footprint of weights, largest at the center and falling off rapidly toward the edges of the footprint, being commonplace.

How does this neural network help to accomplish the joint binding/segmentation problem? Eckhorn writes:

> There are many reasons why neighboring features belonging to the same object will evoke different spike rates in neighboring neurons. If such activities have to be linked by synchronized signal components this can obviously not be accomplished by a precise spike-to-spike synchronization where every spike is coincident with those in other neurons. This would unify the spike patterns and suppress the information about the individual contrasts carried by single neuron spike patterns. In addition, precise synchronization of entire spike patterns has never been observed in recordings. . . However, the coding of features as belonging together requires only a few bits of information and hence loose coupling of many spikes or tight coupling of a few spikes would suffice for it. Such coupling operations seem plausible because they would not interfere with the single neuron's rate coding (keeping rate coding and feature-binding-

by-synchronization nearly independent). [ECK3]

As a minor comment, note that we must read "neuron" as "ENU" in this quote at this level of our modeling hierarchy. Eckhorn's remarks go on to note that linking field signaling tends to synchronize firing activities but non-uniform inhibitory feeding field feedback signaling tends to separate tract signaling by causing small differences in firing rates. Referring specifically to Figure 9.14, inhibitory feedback to the layer-1 pool projecting to the layer-2 ENU is offset by the linking field feedback to these same ENUs. Inhibition to layer-1 ENUs outside of this pool, on the other hand, is not compensated for by linking field feedback from the layer-2 ENU supplying the inhibition. With the larger value of τ_{iff} employed in the feedback configuration, this leads to a gradual build-up of contrast enhancement across the layer-2 output vector and, thus, clearer separation of "figure" from "background" in the layer-1 activity patterns.

§ 5.5 Example: A Neocortical Column Model

Owing to the lack of a general theory for PCNNs, the modeling of neural systems by use of PCNNs tends to be rather more of an art than a science at present. Nonetheless, it is often not particularly difficult to obtain interesting models and model parameters. This author has noticed there is an intriguing similarity between PCNN modeling and the logical design of computer systems as practiced in computer engineering. Figure 9.15 presents one such example.

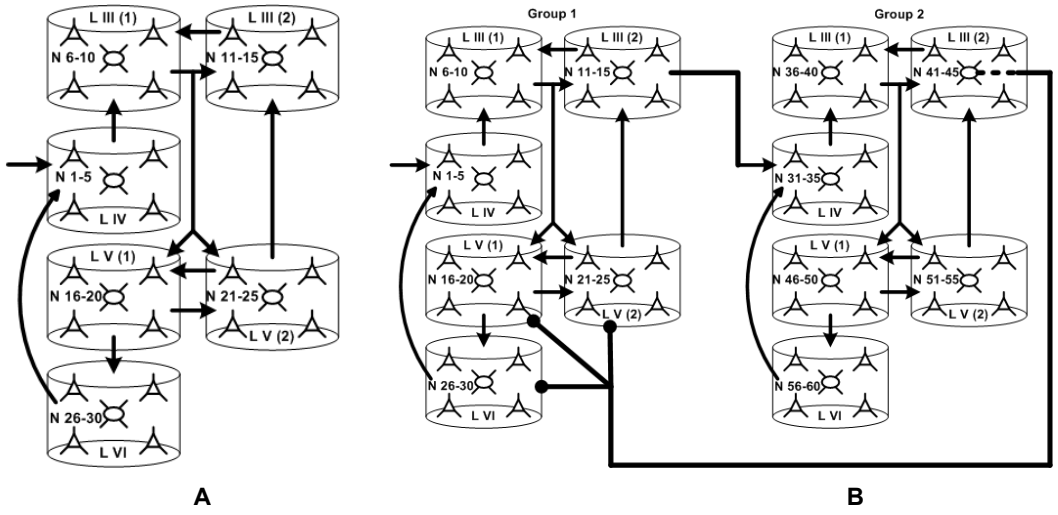

Figure 9.15: Neocortical column models constructed from Eckhorn PCNNs and based on known anatomical connectivity in the neocortex. (A) A column implementing a highpass filter function. (B) A column implementing a bandpass filter function by using two highpass subcolumns with the cut-in frequency of the second subcolumn higher than the cut-in frequency of the first. The second subcolumn projects inhibitory feedback into the first subcolumn when that subcolumn signals at a high rate. The assemblies represented by the small cylinders in the figure are implemented as spike density modulators. The two subcolumns each form large cell groups, denoted as group 1 and group 2, respectively. The highpass filters are each designed for M = 1 highpass operation with the cut-in frequency set by the choice of excitatory feeding field time constant. The ENUs have one excitatory and one inhibitory dendrite.

Wells, Lu, and Montoya [WELL4] proposed this PCNN in 2005 as a model for studying the generation and propagation of signals by neocortical columns in the low (0-3.5 Hz), alpha (7-14 Hz), beta (14-28 Hz), gamma (28-70 Hz), and high (> 70 Hz) frequency bands. The model structure was made to conform to known anatomical connectivities among the layers of the neocortex as reported by Douglas and Martin [DOUG1]. Our immediate purpose was to see if the Douglas and Martin cortical organization represented in Figure 9.15 could explain experimentally observed frequency domain and signaling correlation characteristics obtained from subdural recordings of an awake human subject as reported by Bruns and Eckhorn [BRUN].

The model is constructed from PCNN "kernels" composed of five ENUs. Each kernel is an Eckhorn spike density modulator (Figure 9.11). Full details about the model including all parameter and connection weight values are provided in [WELL4]. These SDM kernels are interconnected to represent neocortical layers II/III (L-III(1) and L-III(2) in the figure), IV (L-IV in the figure), V (L-V(1) and L-V(2) in the figure) and VI (L-VI in the figure). The overall cortical model is arranged in terms of cell groups of six kernels each (a total of 30 ENUs per group, Figure 9.15A). Parameters of the L-IV kernel were designed such that this kernel operates in highpass filter mode with M = 1 for synchronous volley inputs of a specified frequency band (alpha, beta, or gamma). Afferents are applied to the four excitatory ENUs that make up the L-IV SDM kernel. Within the highpass filter group, projections from each SDM kernel to the others are weighted such that a synchronous volley output from all four excitatory ENUs in the source kernel evokes an all-pass firing response in the destination kernel.

Wells et al. were able to show that two such highpass groups operating at different cut-in frequencies could be connected as shown in Figure 9.15B to form an overall column with bandpass signaling properties. For example, a group-1 subcolumn designed to pass beta and higher frequencies coupled with a group-2 subcolumn designed to pass gamma and higher frequencies forms a beta-band bandpass filter. Synchronous input volleys coming in below the beta band evoke no firing response from group-1. Volley inputs in the beta-band evoke a well synchronized cell group response from group-1 but no response from group-2. Input volleys in the gamma-band initially evoke a gamma-band response from both group-1 and group-2 but inhibitory feedback from group-2 into group-1 then suppresses firing by L-V and L-VI kernels in group-1. L-VI is taken as the output projection from group-1 to the next cortical column and so this feedback inhibition from group-2 prevents gamma- and higher-band synchronous input volleys from being propagated to other cortical columns by group-1. Projections from one cortical column to the next in a chain are made from L-VI of the source column to L-IV of the next column in the chain.

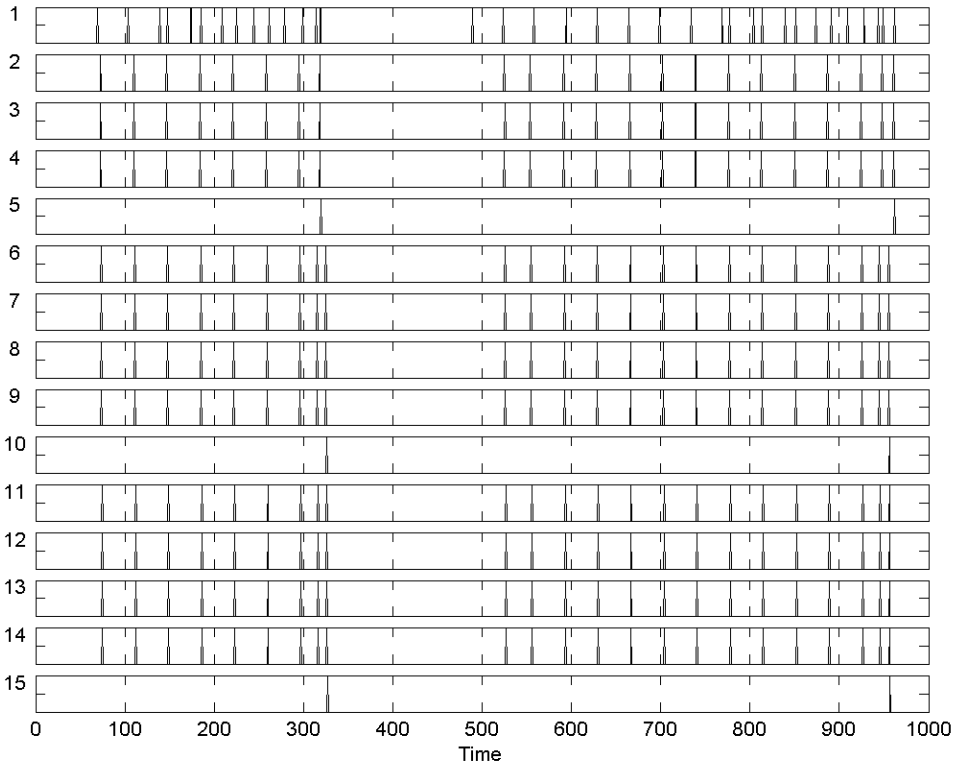

Figure 9.16: Synchronized response of gamma-band highpass group to unsynchronized tetanus inputs above and below its cut-in rate. ENU N1 of L-IV receives a tetanus of 28.57 Hz while ENUs N2-4 receive synchronized 27.03 Hz volley inputs. The group was designed to have a gamma-band cut-in rate of 28 Hz. The numbers at the left of each graph are ENU numbers in the group (see Figure 9.15A). ENUs 5, 10, and 15 are the inhibitory ENUs. SDM kernels L-IV, L-III(1), and L-III(2) responses are shown in this figure.

For well-synchronized input tracts the model group behavior is unremarkable. It behaves as one expects: passing signals above the designed cut-in rate, rejecting those below this rate, operating in bandpass mode (Figure 9.15B only) over the designed range of input signaling frequencies. The novel and interesting behaviors of this model show up when the input tracts are asynchronous. Figure 9.16 illustrates one such behavior, which Wells et al. termed "pull-in by secondary stimulation." A highpass group designed to pass gamma-band signals was given one gamma-band input tetanus (applied to ENU-1) at 28.57 Hz (designed cut-in rate = 28 Hz). The other three ENUs in L-IV received synchronous beta-band (27.03 Hz) inputs. Normally ENUs 2-4 would fail to respond to their inputs because the incoming firing rate is below their cut-in rate. However, because ENU-1 responds to its input (which is above its cut-in rate), its linking field projections to the other ENUs in the kernel evokes firing from them as well. The result is well-synchronized *beta-band* firing by the other kernels in the group. However, because the ENU-1 input is asynchronous to the others, there eventually develops a higher-than-beta firing response

beginning at around 300 ms, which in turn fires the inhibitory ENU and produces packeting of the group's firing response. The ability of this gamma-band highpass filter to respond to beta-band signaling near its cut-in due to the presence of a single gamma-band input shows the PCNN can "pull in" near-band-edge signals due to linking field action. Because the evoked response in the highpass cell group is packeted, this beta-band response travels as an evanescent wave in a chain of identical highpass columns, with the packet widths becoming shorter and shorter at each link in the chain. In this simulation, no signal propagated farther than the fifth link in the chain.

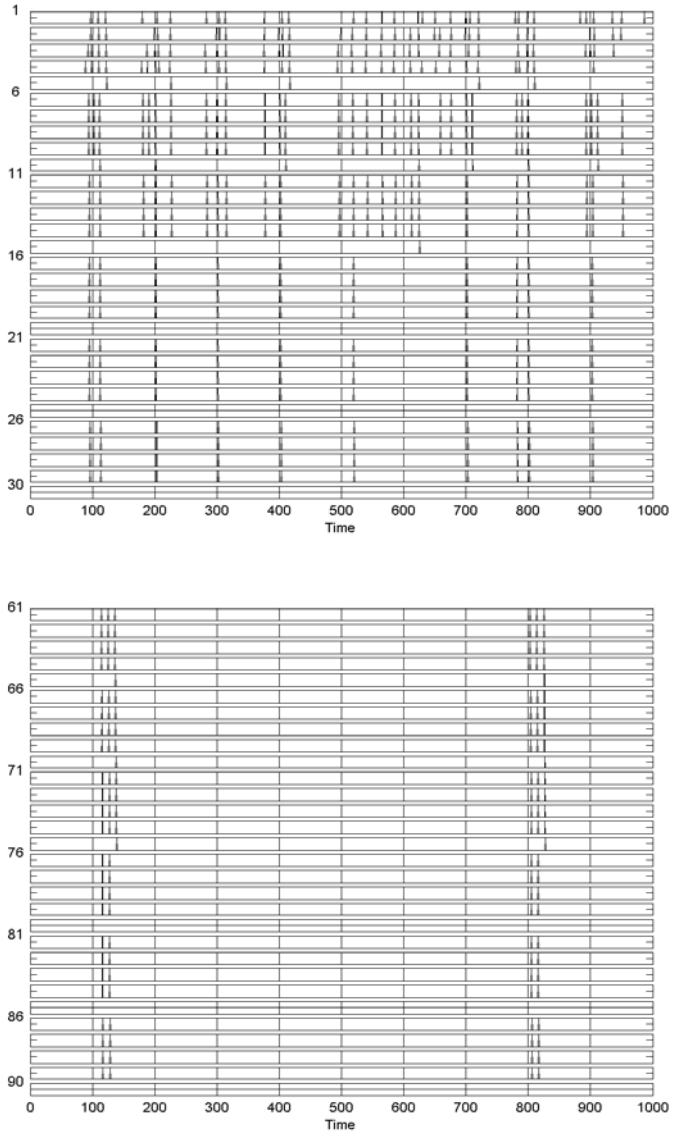

Figure 9.17: Spontaneous emergence of a traveling wave dipulse packet by asynchronous beta-band input signals. ENUs 1-30 belong to group-1 of the first column. ENUs 61-90 belong to group-1 of the second bandpass column connected in a chain to the first. Responses of subsequent columns in the chain is similar to ENUs 61-90. Afferent inputs to the first column in the chain were 19, 20, 21, and 22 Hz (mid-beta-band) applied to ENUs 1-4, respectively. The time axis is in milliseconds.

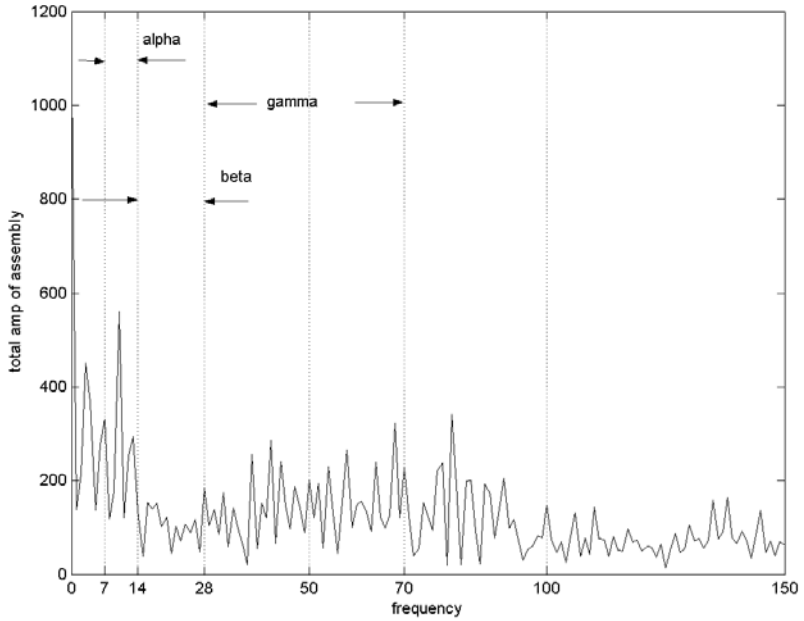

Figure 9.18: Measured spectrum modeled using an idealized subdural probe mimic for the chain response of the example of Figure 9.17. The frequency content of the firing patterns in the model columns shows a general attenuation of beta-band energy levels, the production of alpha- and low-band energies, and the spontaneous creation of gamma-band and high-band frequencies in response to asynchronous beta-band stimulation of the first link in the beta-band bandpass filter chain.

An even more interesting behavior is exhibited by beta-band bandpass columns in response to asynchronous beta-band input signals. This is illustrated in Figure 9.17, which illustrates two successive group-1 links in a chain of beta-band bandpass columns (Figure 9.15B). The first link received asynchronous L-IV beta-band input signals at rates of 19, 20, 21, and 22 Hz, respectively. Each L-IV excitatory ENU would by itself respond to its input tetanus because the incoming signal is in its passband. The linking field action of the L-III(1) kernel synchronizes the firing of this kernel in response to the first two ENU inputs it receives from L-IV, and this in turn stimulates firing in the other kernels in group-1. However, the firing of L-VI stimulates a pull-in response in L-IV and causes this entire kernel to re-fire *in a network-induced rate multiplier mode* at a firing rate much higher than beta-band. This firing rate is controlled by propagation delay down through the column. L-VI induces re-stimulation of L-III(1) but not the other kernels because these are still within their relative refractory periods. This *re*-firing of L-III(1) never reaches the L-VI kernel and so only one RMM-induced extra pulse is produced in L-IV and L-III(1).

However, at t = 103 ms ENU-1 receives its second external input pulse and re-fires when it comes out of its relative refractory period at t = 110 ms. Its linking field action then re-fires the

rest of kernel L-IV at t = 111 ms and this firing action propagates down to L-VI. It fires and again induces an RMM response in L-IV, which re-fires L-III(1). At this point the inhibitory ENU, N10 in L-III(1) fires and packets the response.

The network-stimulated RMM response of group-1 has produced a synchronized *dipulse volley* out of L-VI with a relative spacing between its two pulses of 17 ms (a 58.8 Hz, i.e. gamma-band, rate). This is sufficient to trigger the group-2 highpass part of the column, which responds by inhibiting L-V(1), L-V(2), and L-VI in group-1 until t = 200 ms. More significantly, however, this dipulse is transmitted to the second link in the chain (group-1 ENUs 61-90) where this dipulse packet stimulates the second column into its own network-induced rate multiplier mode of response. The net result is that the second link in the chain produces its own synchronized dipulse volley (spacing = 11 ms, a 90.9 Hz firing rate). This dipulse, transmitted to the third link in the chain, triggers network-induced RMM there as well, and the dipulse wave continued to propagate down the chain. This is not an evanescent propagation. The *dipulse* will continue to propagate down a chain of identical columns for however long the length of this chain is made.

What makes this dynamic particularly interesting is the mix of frequencies produced by this complex ensemble of firing activities. Figure 9.18 illustrates the spectrum of frequencies produced. Although the input stimuli were each a beta-band tetanus, relatively low beta-band energy is produced by the network. On the other hand, significant energy is produced in the low, alpha, gamma, and high bands by the network. This is at least *qualitatively* in agreement with the measured data of Bruns and Eckhorn. Furthermore, because these spectral energies are produced by a complex signaling ensemble rather than by a clean, well-defined tetanus of synchronized signals, standard instrumentation techniques (which work by employing electronic bandpass filters) would find relatively little time correlation between signals measured at one point in the neocortex and those measured at another. This, too, is consistent with a large body of established experimental findings. This behavior, and other interesting ones as well, emerged as a property of the *network*, the design of which only aimed at establishing the correct cut-in rates for its column groups. It is not known whether dipulses are actually produced and propagated in the neocortex *in vivo*, but the model raises interesting research questions in this regard.

§ 5.6 Johnson Networks

We close this chapter with a few brief remarks on an interesting artificial PCNN that has been popular with engineers working in the field of image processing. This network was first reported by Johnson [JOHN1-2], [RANG] and so we will call it a Johnson network. This early work led to a number of later refinements in engineering implementations of this artificial PCNN.

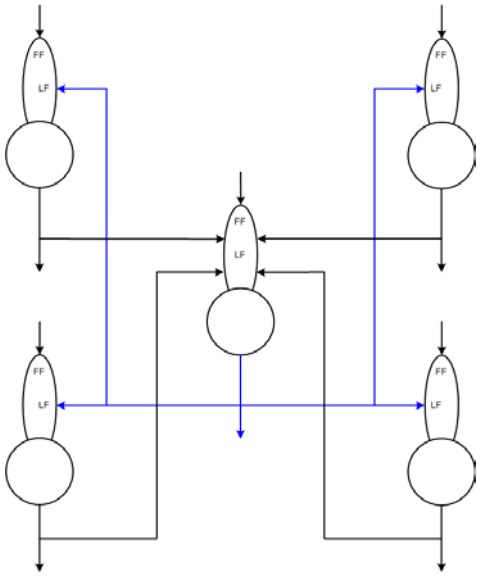

Figure 9.19: The basic cell of a Johnson PCNN.

Since its first publication in the mid-1990s, the Johnson network has seen some application by engineers and has had some interesting dynamics reported. However, it is not thought to be representative of actual biological organization, e.g. the visual system, and has had little note or importance in neuroscience. The Johnson PCNN (often called "*the* PCNN" by engineers) is arranged as a planar "crystalline" neural network of Eckhorn neural units, usually in a rectangular grid where all ENUs except those at the edges have four nearest neighbors. Figure 9.19 illustrates the basic crystalline cell structure of a Johnson PCNN. The center ENU defines the cell; the other ENUs are its nearest neighbors in a rectangular grid.

Little mathematical theory on the network-level dynamics of the Johnson network has been published to date, although a considerable number of papers and one book [LIND] have appeared in the engineering literature. These publications are entirely slanted toward the engineering aspects of the network as a signal processor for planar images. Every ENU in the network receives a single feeding field input from an image retina (usually presented as a gray-scale value rather than as a series of pulses). No ENU receives a feeding field input from any other ENU, although all project linking field signals to neighboring ENUs with connection weights generally decreasing with distance according to typically either a 1/distance variation or some version of a Gaussian footprint. The Johnson network is therefore a "flat" network in terms of input-output.

The most interesting feature of the Johnson PCNN is that its linking field connections produce the presence of wave-like firing actions across the network [JOHN2]. These have been termed "autowaves" by the engineers [LIND]. The source image tends to appear and disappear in a temporal sequence as the autowaves propagate in the network. The PCNN output signals are

typically processed by a post-processing image-processing algorithm of some sort. The autowave action of the network enables this post-processing to perform such image-enhancing operations as noise removal, smoothing, edge identification, etc.

In many ways, the Johnson network might be regarded as a kind of two-dimensional version of the basic Eckhorn layer of Figure 9.10. By itself, it seems to have very little relationship to biological signal processing; but this does not necessarily mean that multi-layer Johnson-like PCNN topologies, as two-dimensional versions of, e.g., Figures 9.12 or 9.13, have no relevance for biological processing. At the present time, though, this is mere conjecture.

§ 6. Summary

Pulse-coupled neural network models occupy a middle position in the modeling hierarchy of Figure 9.1. They have a number of interesting capabilities having to do with synchronization, segmentation, binding, foreground-background separation, and spectrally-rich spontaneous signal generation functions. Only very simple PCNN structures have been investigated and reported thus far and no large-scale, network-level general theory has yet appeared. They are capable of producing wave-like behaviors within the overall network activity profile, being capable of both evanescent wave, traveling wave, and autowave propagation. A small amount of theory attempting to describe PCNN wave-like behavior has been produced, e.g. [HAYO], but the theory is still in an incomplete state.

The basic unit of a PCNN models a putative cell group population. The neural network itself is capable of modeling more dynamical cell assemblies and is capable of exhibiting the formation as well as the disestablishment of such assemblies under differing signaling conditions. As a modeling-level abstraction, PCNNs provide a linkage between neural netlet models at the level of abstract neuron models (e.g. Rulkov, Izhikevich) and map-level models modeling much larger scale neural structures in the CNS. PCNNs have also been shown to be capable of biologically-plausible modeling of spino-neuromuscular systems [GOTS]. Although PCNN research has now been conducted for well over a decade, there is still much to learn about their network-level dynamics and the field is still a largely-open research area. PCNNs have a more direct link to experimental data obtained from probes (as opposed to brain imaging), but the state of the art in instrumentation has not yet advanced to the point where the more intriguing possibilities suggested by mathematical (speculative) network models can be put to the test. It remains an open possibility that PCNN dynamics might provide a more direct link to information theoretic questions regarding the putative "neural code" or codes the CNS is thought to use as its "von Neumann language."