(2)

FINAL REPORT

PERIOD COVERED: 8/1/89 - 12/31/91

Defense Projects Agency Grant N00014 - 89 - J- 3187

"Modeling of Learning-Induced Receptive Field Plasticity in Auditory Neocortex"

Principal Investigators - Norman M. Weinberger[1]

Jack Sklansky[2]

[1] Center for the Neurobiology of Learning and Memory

and

[2] Department of Electrical and Computer Engineering

University of California

Irvine, CA 92717

(714) 856-5512

(714) 856-6726

1. Introduction, Background and Goals

A motivating assumption of this research is that insights into the structure and operation of the brain can inspire important ideas for the design of intelligent machines, in particular machines that detect and classify auditory and visual signals buried in noise.

In this project we focus on the ability of the auditory cortex to adapt its sensitivity to pure tones. Our objective is a model of learning that will enable us to predict changes in the receptive fields of pyramidal cells in the auditory cortex in response to conditioning. This model must account for the receptive fields before conditioning as well as for the conditioning. To do this our model is constructed in two major parts: a local process that accounts for the preconditioned states of the pyramidal neurons, and a global process that accounts for the dispersed impact of conditioning across the pyramidal neurons.

In earlier models of sensory system function the processing of sensory information is accomplished by a hierarchical organization of fixed feature detectors. However, recent findings in awake behaving animals have shown that neuronal tuning to acoustic features, e.g., frequency, is systematically altered in the auditory cortex as a result of learning. Responses to training signals are increased whereas responses to other stimuli are decreased, often enough to make the training signal become the most potent stimulus for a cell (Diamond and Weinberger, 1986; 1989; Bakin and Weinberger, 1990). This *adaptive filtering* appears to be a fundamental property of auditory signal processing.

Adaptive filtering in the auditory cortex represents the impact of behavioral training on the receptive fields of neurons. The receptive field of a neuron is determined by the stimulus parameters to which a cell responds. A frequency receptive field typically is a bell-shaped function of frequency, similar to that of an electronic band-pass filter, centered at a "best" frequency. *Adaptive filtering is said to have occurred when conditioning with one frequency results in a systematic change (plasticity) of the receptive field that is highly specific to the conditioning signal, often moving the best frequency to or toward the conditioning frequency.*

The goal of this project is to formulate a mathematical model of conditioning-induced adaptive filtering in the auditory cortex. In the following sections we briefly report how the relevant neurophysiological data were obtained

and then present in detail our "global-local model" of adaptive auditory filtering. This model successfully accounts for the *adaptive filtering operating simultaneously at two sites of the auditory cortex, and accounts for the effect of the interaction between these sites on their receptive fields.*

2. Neurophysiological Protocol

2.1. Preparation

Recordings of neuronal responses to tones were obtained from the auditory cortex of adult male guinea pigs (Cavia Porcellus). All procedures involving subjects were conducted in strict accordance with approved protocols under the supervision of the University Veterinarian and Animal Research Committee. Neuronal discharges were obtained from fine wire microelectrodes which were implanted 1-2 weeks prior to training when the subjects were under general anesthesia. Following routine post-operative care, the subjects were adapted to head stabilization in a hammock, within an illuminated acoustic chamber for 3-4 days (3 hour sessions). Head stabilization is essential to insure constancy of acoustic stimuli at the ear during the determination of receptive fields. Subjects readily adapted to this protocol as determined by heart rate adaptation In many cases, subjects were trained while awake but receptive fields were obtained before and after training when subjects were under general anesthesia (sodium pentobarbital or ketamine-xylezine). Adaptive filtering was expressed under anesthesia as well as in the waking state. For training, subjects received a tone (the "training tone" or "conditioned stimulus") (six sec.) followed by a brief (0.5 sec) very mild (1-2 ma) footshock (unconditioned stimulus, US). Learning that the tone predicts the US is very rapid. Only twenty pairings were presented, over a forty minute session and learning was indexed by the development of a change in heart rate or a change in ongoing behavior when the training tone was presented (Bakin and Weinberger, 1990; Edeline and Weinberger, 1991; 1992).

2.2 Acoustic Stimulation and Neuronal Recording

Neuronal discharges were recorded by conventional neurophysiological amplifiers (bandpass 0.3-3.0 kHz) and collected by a Brainwave Workstation (Brainwave Systems, Denver, Co.). Receptive fields were obtained by presenting tones of calibrated intensity and frequency (0-90 db sound pressure level [SPL] , 0.5-45.0 kHz, 50 -100 msec.) to the ear contralateral to the recording electrodes,

under the control of a Brainwave Systems Digital Signal Processer. Stability and reliability checks were routinely performed. All receptive fields were based on twenty repetitions of the tone by intensity protocol.

2.3. Determination of Frequency Receptive Fields

Quantified frequency receptive fields were obtained by determining the number of discharges to each tone stimulus using programs developed in this laboratory. The pre-tone (background) rate of discharge was subtracted from the evoked discharge to prevent spurious effects on receptive fields due to random changes in background rate. Data presented in this report were obtained from single units or small clusters which displayed the same characteristics as those of single unit responses. Dominant tuned responses were those occurring 10-50 ms. after tone onset.

Receptive fields were determined immediately before and immediately following behavioral training (i.e., "conditioning") as well as at various retention times up to several weeks following training. Because the effects of training were similar regardless of the time at which receptive fields were determined after training, the mathematical model pertains to all training effects. A quantitative analysis of the effects of training was obtained for each case by subtracting the pre-training receptive field from the post-training receptive field(s). These functions are hereafter referred to as "RF difference functions" (see Section 3.0).

3.0. Neurophysiological Findings

3.1. Introduction

As noted in Section 1.0, the empirical basis for the formulation of a mathematical model of adaptive filtering were the findings that conditioning "retunes" the auditory cortex such that frequency receptive fields are altered to "favor" the processing of behaviorally important stimuli. Specifically, responses to the frequency of the training tone (conditioned stimulus, CS) are increased whereas responses to other tones, including the pre-training most effective tone (the "best frequency", BF) are reduced. In many cases, these opposite changes are sufficiently large so that the CS becomes the new BF, i.e., the cells are completely retuned to the CS frequency. Within a discussion of our Local-Global Model (Section 4), this is referred to as "complete" learning. In other cases, there is either a shift of the BF toward but not completely to the CS frequency. If the CS

frequency is at a large frequency distance from the BF (e.g., > 1.0 octaves), or if the pre-training response to the CS frequency is very weak (i.e., the CS frequency is outside of the receptive field of the cells being studied), then there may be a decrease at the BF and an increase at the CS without any actual shift of the BF.

Whatever the detailed expression of adaptive filtering, the processing of the training tone is facilitated with reference to other acoustic frequencies. Thus, the receptive field or "neuronal filter" is modified in a highly specific way by conditioning to emphasize behaviorally important acoustic frequencies.

3.2. Source of Adaptive Filtering in the Auditory Cortex

The source of adaptive filtering in the auditory cortex is an issue which lies within the scope of this project because the mathematical model should account for how adaptive filtering is produced by the brain.

Of particular relevance, the Model assumes that adaptive filtering in the auditory cortex does not simply arise subcortically and is then projected to the auditory cortex, where it is detected by recording electrodes. Rather, the Model assumes that adaptive filtering actually arises in the auditory cortex. Moreover, it does so by combining two subcortical sources of input: (1) a highly specific and unchanging acoustic frequency input which specifies the current frequency in the acoustic environment, as transduced by the cochlea; (2) an auditory non-frequency specific input which serves as a "training signal" to the cortex, indicating the behavioral importance (i.e., acquired signal value) of the current acoustic frequency.

This architectural substrate is well substantiated by previous empirical research from many laboratories (reviewed in Weinberger et al, 1984). Thus, the auditory cortex receives *dual* projections from the auditory thalamus: (1) the auditory lemniscal line from the ventral medial geniculate (MGv); (2) the auditory non-lemniscal path,from the magnocellular medial geniculate (MGm). As previously emphasized (Weinberger et al, 1990a,b), the MGv does not alter its response to tones regardless of their acquired signal value; in contrast, the MGm rapidly increases its responses to tones as they become behaviorally important due to training. During CS-US pairing trials, the MGv provides essentially unaltered, detailed frequency input to the auditory cortex. In contrast, the MGm, which receives both CS and US input, is probably the first site of associative plasticity. However, its neurons are very broadly tuned and so cannot provide detailed frequency information to the auditory cortex.

In addition to these findings, it was necessary to determine whether any plasticity of receptive fields occurs with the MGv and the MGm and if so, whether it accounts for cortical adaptive filtering. Consequently, we performed appropriate experiments. Within the MGv, adaptive filtering is weak, transient or non-existent and could not account for adaptive filtering in the auditory cortex (Edeline and Weinberger, 1991). Within the MGm, receptive field plasticity does occur but its characteristics are very different from those which characterize the auditory cortex and could not account for adaptive filtering in the cortex (Edeline and Weinberger, 1992). Accordingly, both recordings obtained during training trials and recordings of receptive fields obtained after training trials establish that adaptive filtering in the cortex is not simply a reflection of adaptive filtering subcortically. Rather, a new process does develop at the cortical level. The MGv provides detailed frequency information and the MGm provides information about the behavioral importance of an acoustic stimulus. It is within the auditory cortex that these two inputs are combined to produce cortical adaptive filtering.

Our mathematical model is based closely on these established neurobiological findings. Thus, within the Model, the MGv input to the cortex serves to provide detailed but unchanging frequency information and the MGm input serves as the training input (see Section 4).

3.3. The Importance of Simultaneous Recordings from Different Cortical Sites

Although our discovery of adaptive filtering in the auditory cortex provided the impetus for mathematical modeling, the prior data were deemed insufficient to develop a powerful model which would provide deep insights into the processes which produce adaptive filtering.

We considered the prior data insufficient because they were obtained from only one site within the auditory cortex within a single training session. There were compelling reasons to hypothesize that adaptive filtering did not develop only at the randomly-selected recording site within the frequency representation in the auditory cortex, but rather developed across the representational cortical mantle. Indeed, we specifically predicted that the area of representation for the training frequency would increase (Weinberger et al, 1990a,b), a prediction which subsequently has received strong empirical support (Recanzone et al, 1991; Scheich & Simonis, 1991).

Given the likelihood that adaptive filtering is a fundamental process which encompasses the frequency representation of the entire primary auditory cortex,

it was evident that cortical interactions ("global learning", see Section 4) which are likely to exist could not be detected if recordings were confined to a single cortical site. Therefore, in the current project, we undertook to obtain recordings simultaneously from two sites within different parts of the frequency representation for which training-induced interactions within the cortical network would capable of detection and characterization.

This proved to be a formidable task because this had to be accomplished in waking, behaving subjects; moreover, the simultaneous recordings had to be maintained continually from the pre-training period, through behavioral training (often in moving subjects) to the completion of post-training receptive field determinations. Nevertheless we were successful, and so were able to provide the needed quantitative neurophysiological receptive field data which enabled the formulation and testing of our global-local model.

3.4. Simultaneous Observation of Adaptive Filtering at Two Sites

In this section, we describe an example of adaptive filtering at two sites where receptive fields were observed simultaneously before and after conditioning. At both sites, modifications of receptive fields met the definition of adaptive filtering, i.e., facilitation of the processing of the conditioning frequency vs. non-conditioning frequencies.

Figure 3.1 presents quantified receptive fields both pretraining and posttraining (upper panels) and the receptive field difference functions (post minus pretraining receptive fields, lower panels) for electrodes 2 and 3 from subject BW07; of five implanted electrodes, these were the only recording sites which yielded adequate recordings throughout the session. In this training session, the training frequency (CS) was 2.5 kHz.

Pretraining, electrode #2 responded best to frequencies below 1 kHz; its BF was 0.75 kHz. However, posttraining its tuning shifted drastically; its BF changed to 2.5 kHz, that is, its new BF was at the frequency of the CS ("complete learning"). The lower panel, showing the RF difference function for this electrode, depicts how training altered the tuning at this recording site. It is clear that training caused increased responses at some frequencies and decreased responses at other frequencies. The largest increases in response were centered on the CS frequency, and indeed the absolute largest increased response was at the CS frequency itself. The largest decreases were centered on the pretraining BF, and in fact the largest absolute decrease was exactly at the pretraining BF.

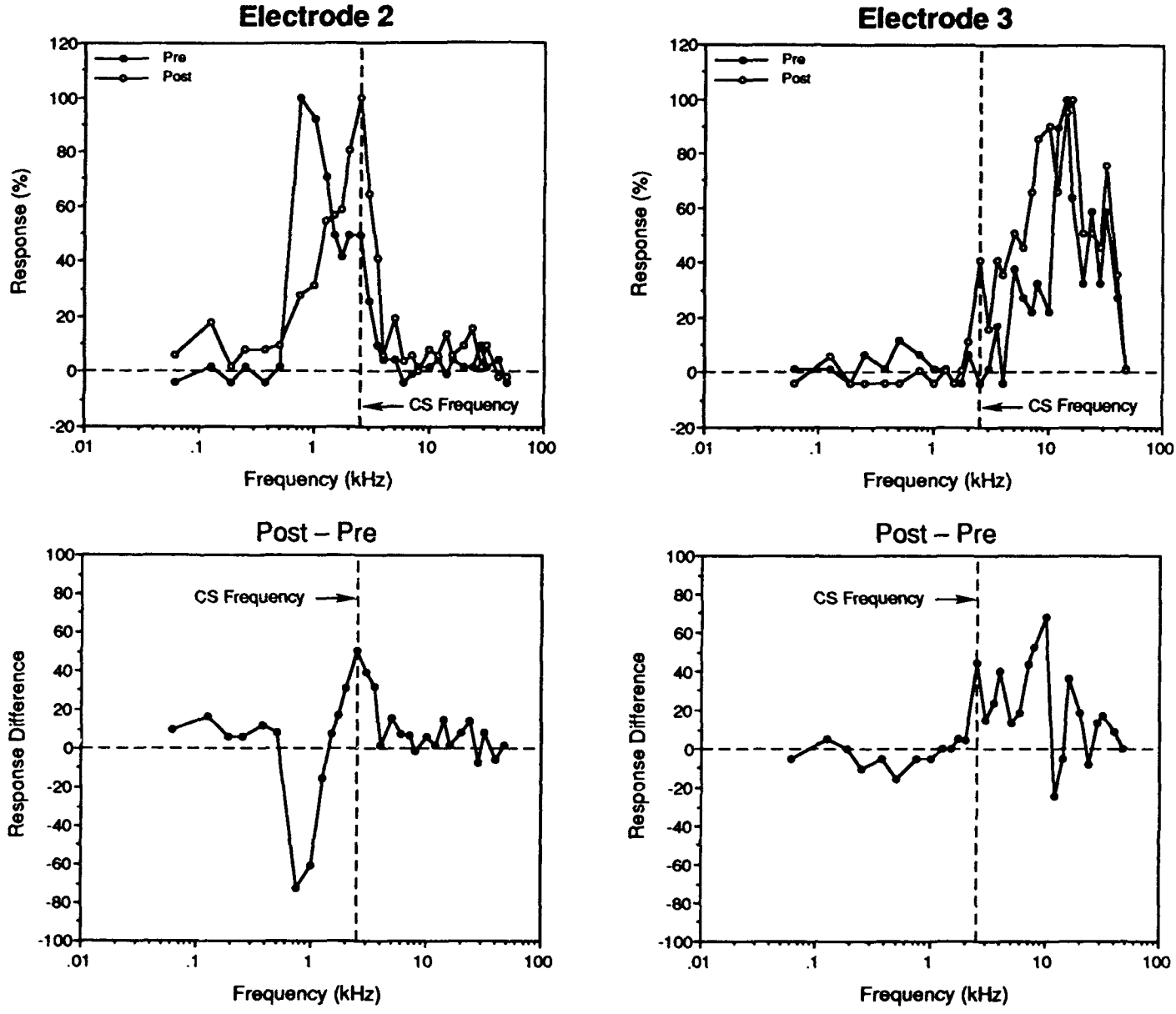


figure 3.1 An example of simultaneous recordings from two loci with the auditory cortex.

Electrode #3 was located anterior to electrode #2 and so was tuned to higher frequencies. Pretraining, its BF was 16.0 kHz . There was no response at the CS frequency of 2.5 kHz;indeed, 2.5 kHz produced a slight suppression of activity at this recording site. Thus, the BF was more than 2.5 octaves above the CS frequency. At this large octave (and anatomical) distance, one would not expect a shift of tuning to or even toward the CS, because the CS frequency was outside of the receptive field. In fact, there was no large change of the BF posttraining; it actually shifted slightly higher, to 18.0 kHz. Nonetheless, conditioning did have a frequency specific effect. Posttraining, the CS frequency produced a very clear excitatory response. Moreover, responses to frequencies between the CS frequency (2.5 kHz) and 10.0 kHz increased (see the RF difference function in the lower panel). The overall result was to increase the bandwidth of tuning such that it included the lower frequencies, down to the CS frequency ("partial learning"). In summary, conditioning altered the tuning so that responses at or on the high frequency side of the CS frequency were facilitated.

Overall, this example shows adaptive tuning which developed simultaneously at two widely separated recording sites within the orderly frequency representation of the primary auditory cortex. When the CS frequency is within the receptive field (electrodes #2), then a large shift in tuning can occur, even complete retuning so that the training frequency becomes the BF. When the CS frequency is not within the receptive field (electrode #3), then the BF does not shift to or toward the CS frequency, but responses to the CS frequency can still increase (indeed, be converted from suppression to clear excitation) and frequencies between the CS and the BF also exhibit facilitated responses. In short, it seems that adaptive filtering occurs across the frequency representation. The challenge, then, is to formulate a mathematical model which can account for such findings. We now consider our global-local model, the major focus of this project.

4. The Global-Local Model of Adaptive Filtering

4.1. Introduction

In this section, we present our "global-local model of adaptive filtering". We will first describe the model (Section 4.1), then consider in detail "global conditioning" (Section 4.2), followed by "local conditioning" (Section 4.3). We conclude the presentation of our model with a consideration of its neurophysiological

interpretation. Our evaluation of the model's ability to account for the neurophysiological findings of adaptive filtering is presented in Section 5.0.

4.1.1. Global and Local Processes

Our global-local model represents the receptive fields of the auditory cortex as a set of parallel signal processors or, equivalently, as a single-input multiple-output system. This is illustrated in Figure 4.1.

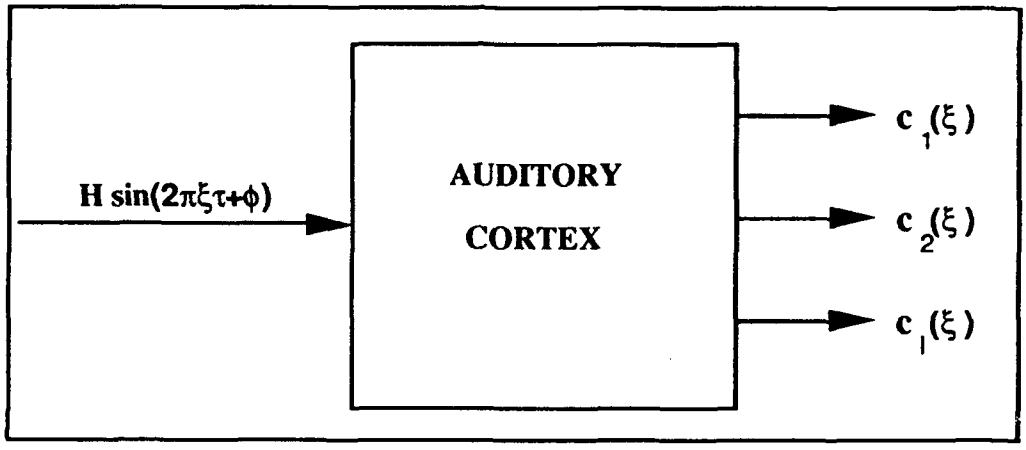


figure 4.1 Receptive fields in the auditory cortex.

Here each cortical output $c_i(\xi)$ represents the number of spikes per second of the i^{th} pyramidal cell as a function of the amplitude H and frequency ξ of the stimulus tone. Each such output, viewed as a function of H and ξ, is a receptive field. Conditioning affects each receptive field in a manner that depends partly on $c_i(\xi)$ and its neighboring cells (a local effect) before conditioning, and partly on pyramidal-cell-independent aspects of conditioning (a global effect).

Thus our model consists of two major components: a local process and a global process. This is reflected in the structure of the model: a neural network consisting of three layers. The first layer is a set of n acoustic resonators. The second is a global layer, and the third is a local layer. The details of this structure are described in the next section.

4.1.2. The Structure of the Model

Our model is structured as a three layer neural network, illustrated in Figure 4.2. The first layer is a set of resonators, the k^{th} of which produces an output $v_k(\xi)$ in response to the signal H sin $(2\pi\xi\tau + \varnothing)$. In this model the phase ϕ is ignored by all of the layers. The present model does not account for H, but since we didn't study the effect of H in this project period, we omit H from our symbols for the current model. We expect to introduce H in later developments of our model.

The outputs of the global and local layers are represented as functions of t, where t denotes the stage of conditioning. The preconditioned state of the system occurs at t = 0. The postconditioned state of the system occurs at t = 1.

The global layer consists of two components: a global feature extractor and a global trainer. The global feature extractor receives all of the $v_k(\xi)$'s (there are n of them) to produce a set of m features $\{y_r(0,\xi)\}$ by a set of m linear summators, each summator formed by a set of weights $\{w_{kr}(0)\}$. Usually m<<n. This structure is illustrated in Figure 4.3. We denote these weights by a matrix **W** (0).

Global training adjusts **W** (0) in response to conditioning. This adjustment is represented compactly by a small set of global training parameters $\{e_r\}$. During conditioning the auditory cortex receives a strong sinusoidel signal at the conditioning frequency ξ, followed shortly thereafter by an input which signals that the reinforcement has occurred. In our model the conditioning process is enabled only when g = 1, where g is a conditioning control signal as indicated in Figure 4.2. When there is no conditioning, g = 0.

The local layer consists of l components, each component associated with a single pyramidal cell. Each component consists of a local discriminant and a local trainer, and produces the output $c_i(\xi)$. The local discriminant receives all of the features $\{y_r(0),\xi)\}$ and the global training parameters $\{e_r\}$ (there are m of them) to produce the response $c_i(t,\xi)$ by a linear summator formed by a set of weights $\{A_{ir}(t)\}$ (see Figure 4.3). We denote these weights by a vector **A** (t). Each receptive field $c_i(t,\xi)$ output is the average number of spikes per second in response to a tone of amplitude H and frequency ξ. The local trainer produces a local scaling effect represented by local scaling parameter h_i and adjusts **A** (0) in response to a) the conditioning control signal g, b) the set of inputs $\{y_r(0,\xi)\}$, c) the global training parameters $\{e_r\}$, and d) the receptive field $c_i(0,\xi)$.

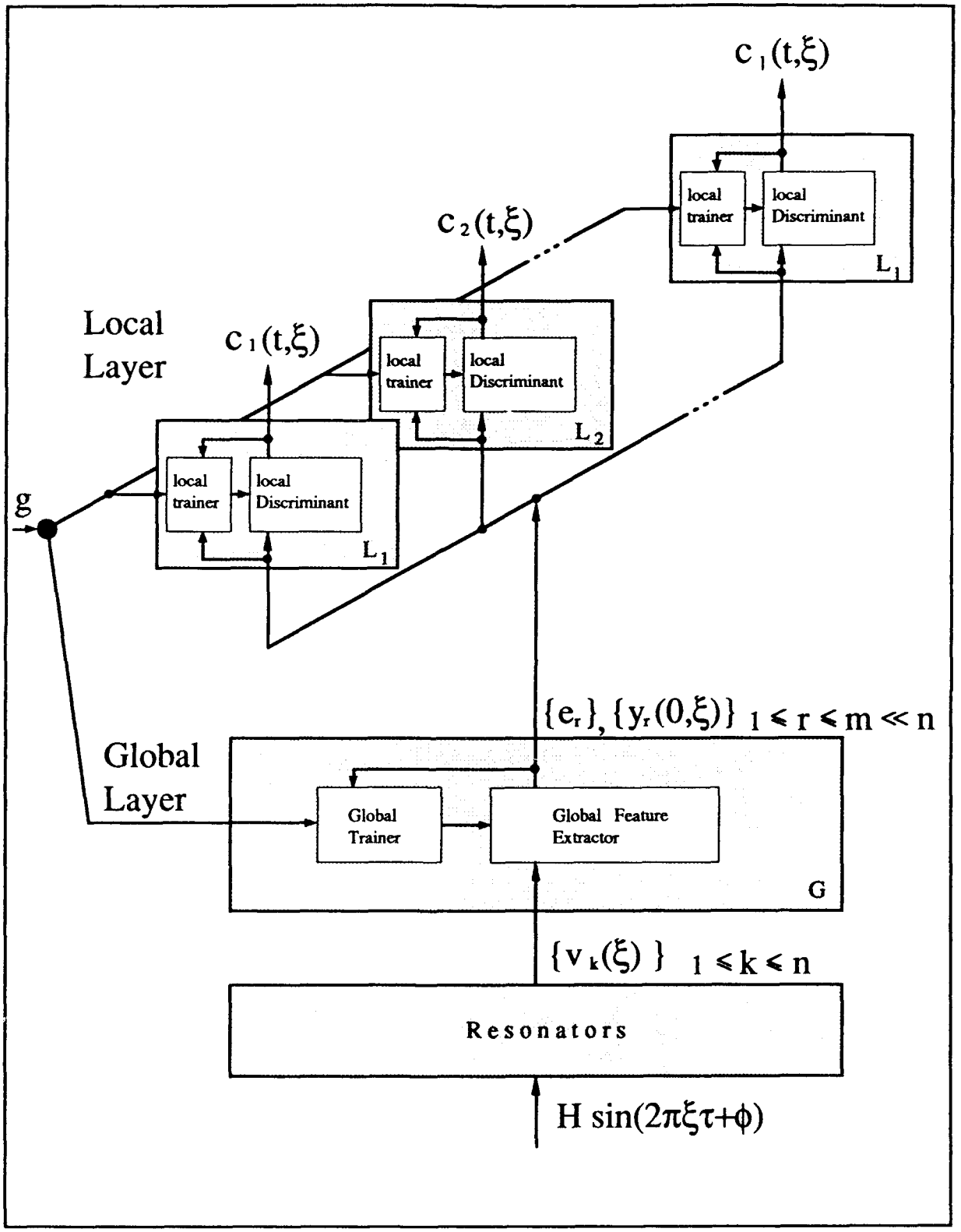


figure 4.2 : The Global - Local Model.

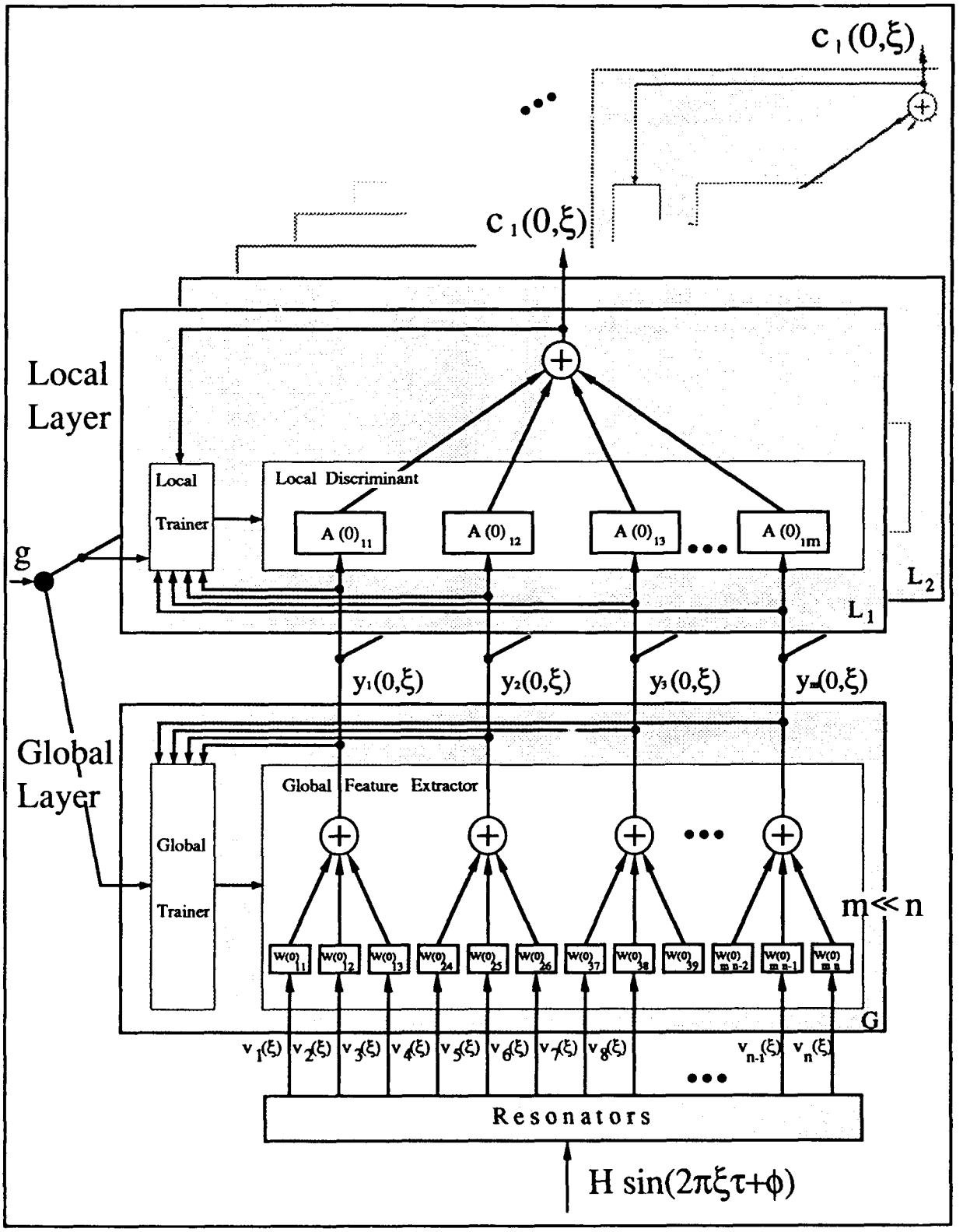


figure 4.3 : Details of the Global - Local Model , at t = 0.

4.1.3. The Layers of the Model

Below we describe each of the three layers.

4.1.3.1. The Resonators

The k^{th} resonator produces an output $v_k(\xi) = v(\xi - k\xi_o)$, where $v(\xi)$ is a positive even function of ξ. In our simulations we assumed that

$$v(\xi) = \begin{cases} 0.5\left(1 + \cos\frac{\pi|\xi|}{b}\right) & \text{for } \frac{|\xi|}{b} < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

Figure 4.4 describes the general kernel shape. Figure 4.4(b) describes the kernels in a range of frequencies. We denote the number of kernels by n. This number may be large -- perhaps several hundred.

4.1.3.2. The Global Layer

This layer consists of a global feature extractor and a global trainer. At t = 0 the outputs of the global layer formed in accordance with the following equations:

$$y_r(0,\xi) = \sum_{k=1}^{n} w_{kr}(0) v_k(\xi) \qquad (r=1,\ldots,m), \tag{4.2}$$

$$e_r = 1 \text{ for all r.}$$

At t = 1 e_r is revised to reflect the tuning to the conditioning frequency ξ_c.

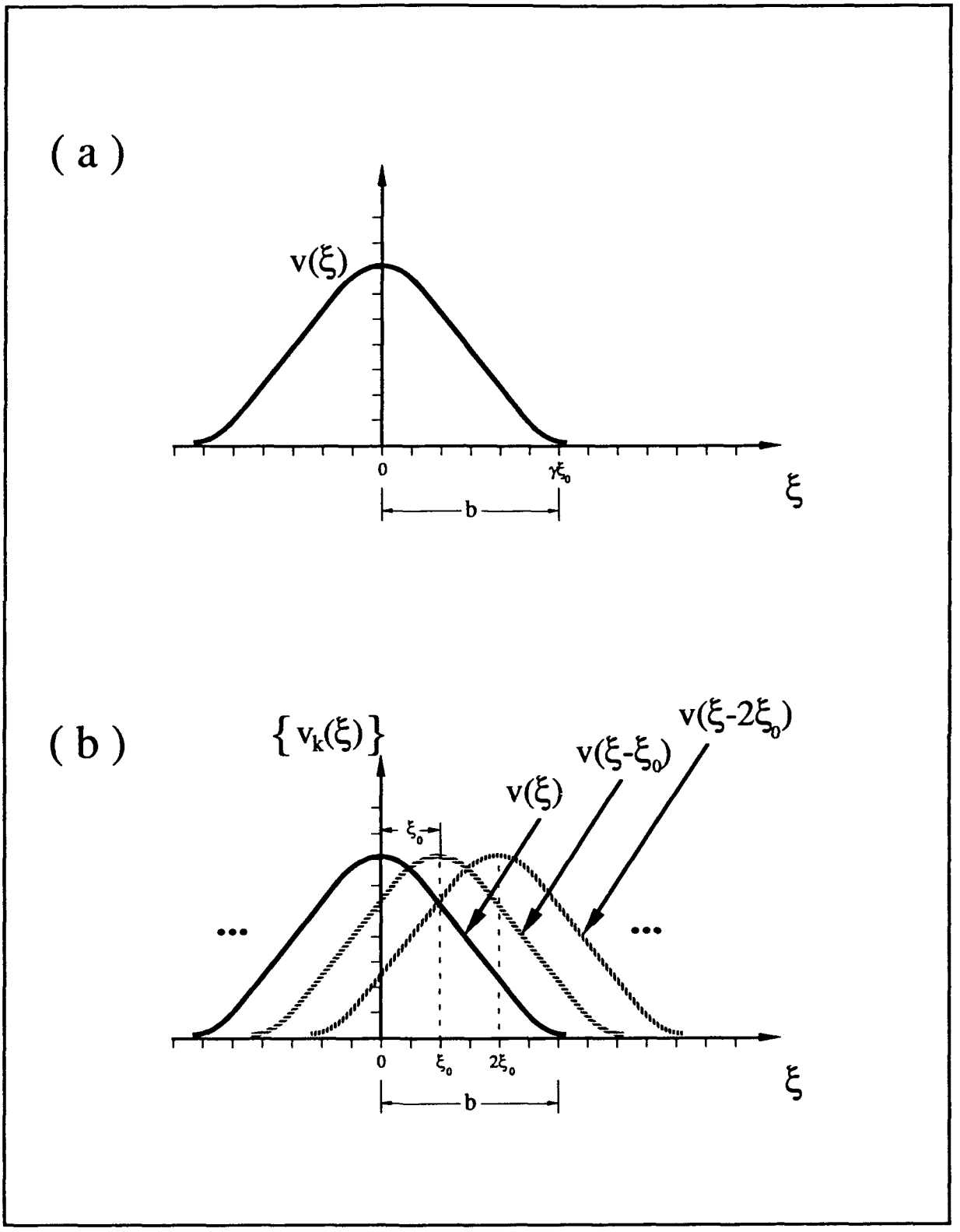


figure 4.4 : The kernels $v(\xi-k\xi_0)$.

4.1.3.3. The Local Layer.

This layer consists of a local trainer and a set of m local discriminants for each of the l output signals $\{c_k(t,\xi)\}$ (k = 1,...,1). This is illustrated in Figure 4.3 for t =0. The i^{th} local discriminant transforms $\{y_r(t,\xi)\}$ (r = 1,...,m) into $c_i(t,\xi)$ by the following linear equation:

$$c_i(0,\xi) = \sum_{r=1}^{m} A_{ir}(0) y_r(0,\xi) \qquad i = 1,...,l. \tag{4.3}$$

4.1.4. Operation of the Network

In this section we describe how the model operates during the conditioning process.

Preconditioning stage (t = 0)

Step 1 : Initialize the network parameters and weights.

Step 2 :

$$y_r(0,\xi) = \sum_{k=1}^{n} w_{kr}(0) v_k(\xi) \qquad r = 1,...,m. \tag{4.4}$$

Step 3 :

$$c_i(0,\xi) = \sum_{r=1}^{m} A_{ir}(0) y_r(0,\xi) \qquad i = 1,...,l. \tag{4.5}$$

Postconditioning stage (t = 1)

In this stage the weights of $A_{ir}(0)$'s and $w_{kr}(0)$'s are adjusted in response to the conditioning frequency ξ to predict the postconditioning stage. First the $w_{kr}(0)$'s are adjusted by the global trainer in terms of the global training

parameters $\{e_r\}$. Then the local trainer produces a local scaling effect on the features $\{y_r(0,\xi)\}$, and adjusts the $A_{ir}(0)$'s.

Step 4 (global training):
Estimate the global training parameters $\{e_r\}$, by the procedure described in Section 4.2.2.

Step 5 (local training begins):
Compute $w_{kr}(1)$ by the equation:

$$w_{kr}(1) = w_{kr}(0) + \Delta w_{kr} = w_{kr}(0) = h_i e_r w_{kr}(0) \quad \text{for} \quad k = 1,...,n;\ r = 1,...,m \tag{4.6}$$

where h_i is a local scaling parameter, determined by a procedure described in Section 4.3.2.

Compute the features $\{y_r(1,\xi)\}$ by the equation:

$$y_r(1,\xi) = y_r(0,\xi) + h_i \sum_{k=1}^{n} e_r w_{kr}(0)\, v_k(\xi) \qquad r = 1,...,m. \tag{4.7}$$

Step 6 (continuation of local training):

$$A_{ir}(1) = A_{ir}(0) + \Delta A_{ir} \qquad \text{for } r = 1,...,m;\ i = 1,...,l. \tag{4.8}$$

The determination of ΔA_{ir} is described in Section 4.3.3

Step 7 (computation of the modeled receptive fields of the pyramidal neurons):

$$c_i(1,\xi) = \sum_{r=1}^{m} A_{ir}(1) y_r(1,\xi) \quad \text{for } i = 1,\cdots, l\,. \tag{4.9}$$

4.2. Global training

In this section we describe the details of global conditioning in our model. The global layer a) reduces the relatively large number of resonator inputs $v_k(\xi)$ to a small number of features $y_r(0,\xi$ and e_r and b) tunes the $\{e_r\}$ in response to

conditioning, suppressing some of them and strengthening others. This tuning is a form of competitive learning. Both of these activities of the global layer are in accord with current understanding of the neural architecture and adaptive processes in the auditory cortex (Weinberger et al, 1990b). The features $\{y_r(0,\xi\}$ and $\{e_r\}$ are inputs to all of the cortical segments in the local layer simultaneously. The training of the $\{e_r\}$ takes place only during conditioning. The following section describes how to achieve these purposes.

4.2.1. Convolution

In this section we describe the operation of the global feature extractor. This subsystem implements Equation 4.2. It transforms the n inputs $\{v_k(\xi\}$ to the m outputs $\{y_r(\xi)\}$, where m << n, thereby achieving a reduction of the input data by a factor n/m. The global feature extractor carries out m convolutions simultaneously, each convolution yielding a feature $y_r(\xi)$.

Equation 4.2 includes a variable t (t = 0), denoting the stage of conditioning. In the rest of the section we will often omit the argument t to simplify our notation. With this omission Equation 4.2 becomes:

$$y_r(\xi) = \sum_{k=1}^{n} w_{kr}\, v_k(\xi) \tag{4.10}$$

Below we show that Equation 4.2 (and hence Equation 4.10) represents a convolution operation with respect to x. Suppose $v_1(\xi,)$ $v_2(\xi),\ldots,v_n(\xi)$are even functions, as illustrated on Figure 4.4, and suppose $v_k(\xi) = v(\xi - k\xi_o)$. The weights $\{w_{kr}\}$ multiply the kernels $\{v_k(\xi)\}$ in accordance with Equation 4.10. Let γ denote reduction ratio from kernels to features:

$$\gamma = \frac{n}{m} = \frac{\text{number of kernels}}{\text{number of features}}. \tag{4.11}$$

The weights $\{w_{kr}\}$ may be viewed as a function $w_r(p\xi_0)$, where

$$k = r\gamma + p. \tag{4.12}$$

This equivalence between $\{w_{kr}\}$ and $w_r(p\xi_0)$ is illustrated in Figure 4.5. Here we see that $w_{kr} = w_r(p\xi_0)$. The feature $y_r(\xi)$ is formed by $w_r(p\xi_0)$ and $v(\xi - k\xi_0)$ as follows:

$$y_r(\xi) = \cdots + w_r(-\xi_0)\, v(\xi - (r\gamma - 1)\xi_0) + w_r(0)\, v(\xi - \gamma r\xi_0) + w_r(\xi_0)\, v(\xi - (r\gamma + 1)\xi_0) + \cdots .$$

This equation may be written in the following form:

$$y_r(\xi) = \sum_{p=1-r\gamma}^{n-r\gamma} w_r(p\xi_0)\, v(\xi - (r\gamma + p)\xi_0)$$

This is equivalent to

$$y_r(\xi) = \sum_{p=1-r\gamma}^{n-r\gamma} w_r(p\xi_0)\, v[(\xi - r\gamma\xi_0) - p\xi_0] \qquad (4.13)$$

Equation 4.13 represents a convolution between $w_r(\xi)$ and $v(\xi - r\gamma\xi_0)$ over integer multiples of ξ_0. We denote this symbolically by

$$y_r(\xi) = w_r(\xi) * v(\xi - r\gamma\xi_0) . \qquad (4.14)$$

This convolution is represented by the block diagram in Figure 4.6.

This diagram shows $y_r(\xi)$ as the response of a linear system with impulse response $w_r(\xi)$ to the input signal $v(\xi - r\gamma\xi_0) = v_{r\gamma}(\xi)$.

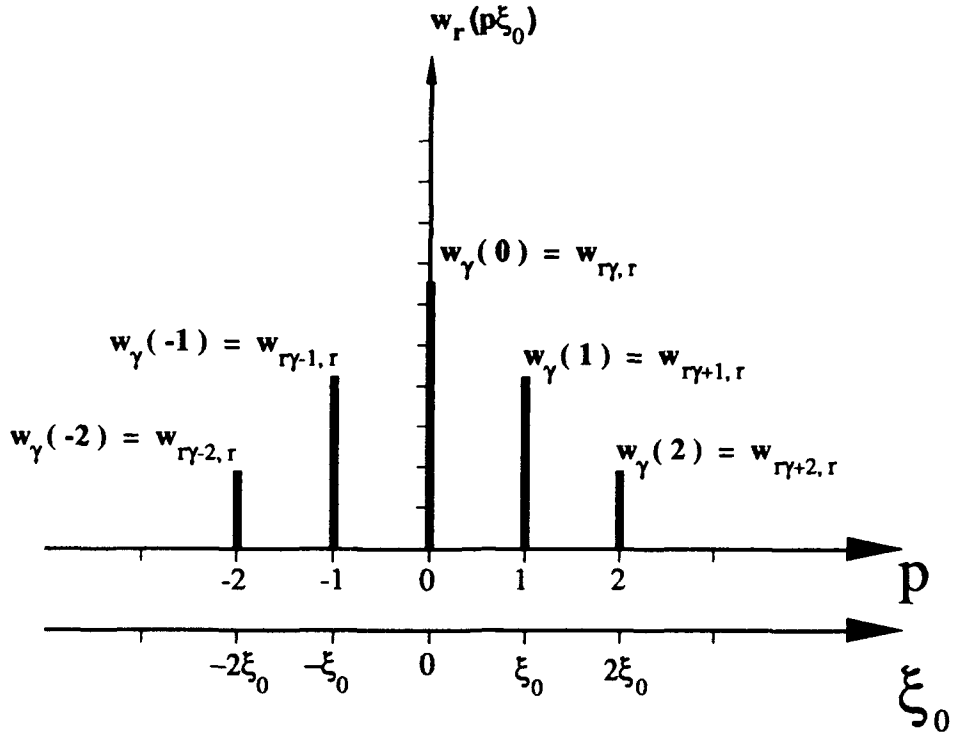


figure 4.5 : w_{kr} may be viewed as a function of ξ .

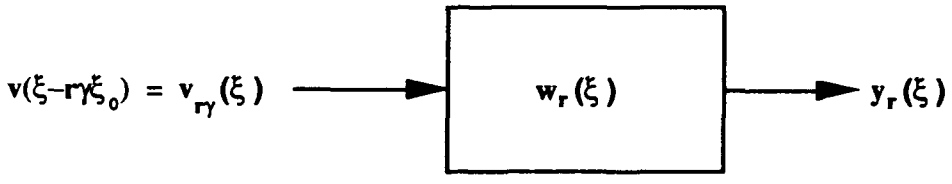


figure 4.6 : The Global feature extractor as a convolution system .

We now show that Equation 4.13 and Equation 4.10 are equivalent. Note that

$$w_r(p\xi_0) = w_{kr},$$

$$v\left[(\xi - r\gamma\xi_0) - p\xi_0\right] = v(\xi - (r\gamma + p)\,\xi_0) = v(\xi - k\xi_0) = v_k(\xi).$$

Substituting the right members of these equations into Equation 4.13 when replacing p by k - rγ, we obtain Equation 4.10.
We assumed in our experiments that the kernel v(ξ) which is used in Equation 4.14 has the following form:

$$v(\xi) = \begin{cases} 0.5\left(1 + \cos\dfrac{\pi|\xi|}{b}\right) & \text{for } \dfrac{|\xi|}{b} < 1 \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(4.1)}$$

We assume that before conditioning all of the $w_r(\xi)$'s have the form $a\xi_0 w(\xi)$, where a is a scaling constant that is determined experimentally.

$$w_r(\xi) = a\xi_0 w(\xi) \quad \textit{for all } r \qquad \textbf{(4.15)}$$

In our experiments we assumed that w(ξ) is triangular as follows:

$$w(\xi) = \begin{cases} 0.5\left(1 - \dfrac{|\xi|}{G}\right) & \text{for } \dfrac{|\xi|}{G} < 1 \\ 0 & \text{otherwise.} \end{cases} \qquad \textbf{(4.16)}$$

Figure 4.7 illustrates $v_k(\xi)$ and $w_r(\xi)$. The variables b and G denote the "bandwidths" of these functions. The constant a is strongly affected by the choice of the bandwidth b and G. The value of a varies approximately inversely as b and G.

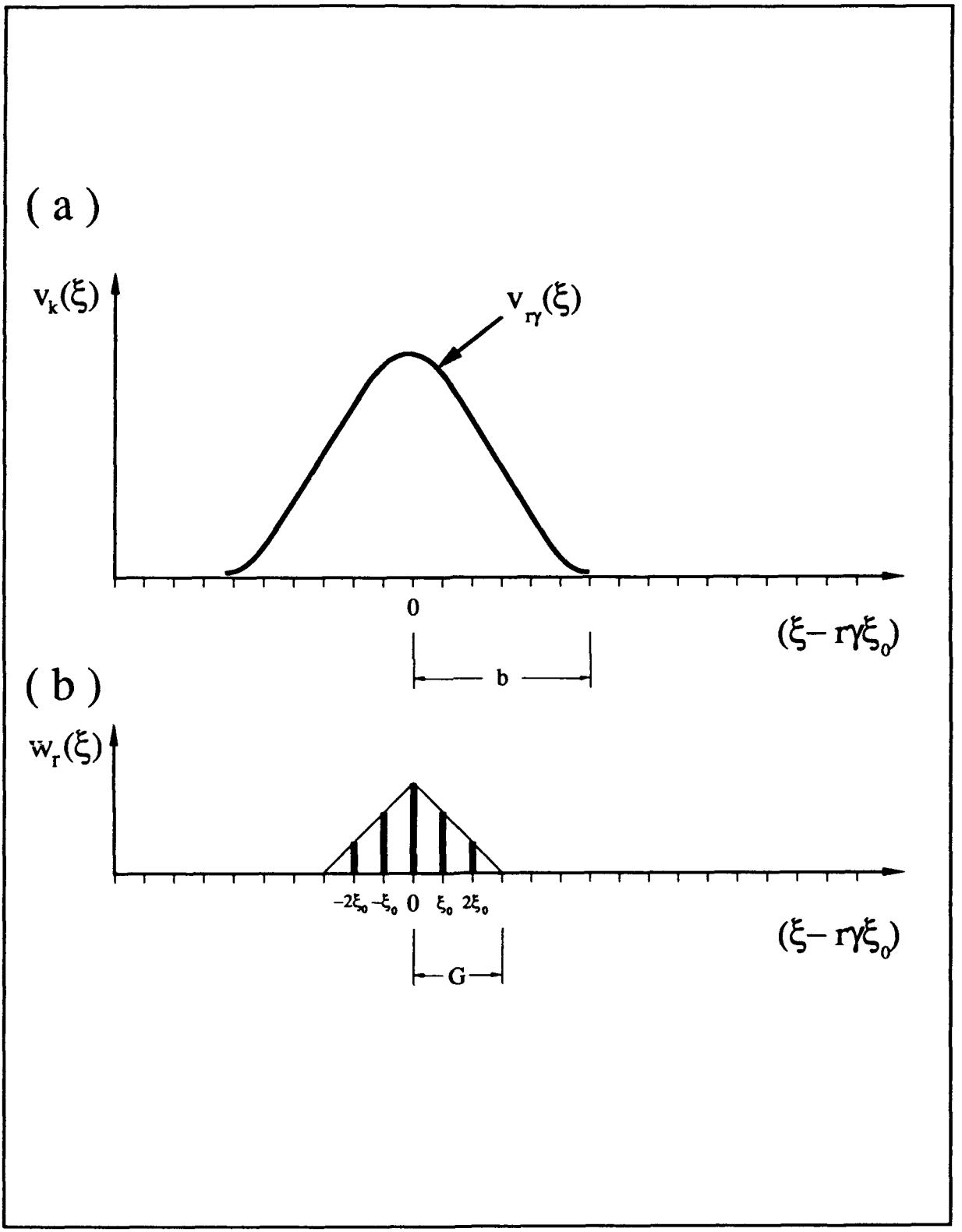


figure 4.7 : The two functions to be convolved .

Figure 4.8(a) shows the set of kernels $v_k(\xi)$ as determined by Equation 4.1, and Figure 4.8(b) shows the results of convolving these kernels with the weight functions $w_r(\xi)$ as determined by Equation 4.16.

Now we describe how the parameters ξ_0, m, n, γ, b, G and a, should be chosen. Recall that ξ_0 represents the difference between the center frequencies of every pair of adjacent kernels. m denotes number of features. n denotes number of kernels. γ denotes the reduction ratio defined in Equation 4.11. b and G are the bandwidths of $v(\xi)$ and $w(\xi)$ respectively. a is a scaling constant.

Choosing m.

The value of γ is the ratio of m to n. In Section 4.3.1 we show that the observed values of ξ in $c_i(\xi)$ are spaced by $\gamma\xi_0$ and that the number of these values is m. Therefore first we choose m equal to the number of observed values of ξ in $c_i(\xi)$. (If the observed values of $c_i(\xi)$ are not uniformly spaced then $\gamma\xi_0$ is the smallest spacing between observed values of ξ.)

Choosing γ

We choose the value of γ to be a large as possible restricted only by practical computational considerations. We believe that a physiologically based choice of γ may be greater than 1000. But for practical purposes we have chosen $\gamma = 5$.

Choosing n.

The value of n is determined by Equation 4.11.

Choosing ξ_0 and b.

The value of ξ_0 depends on the number of nonzero samples of $v(\xi)$ used in the digital convolution in Equation 4.13. A practical choice of the number of nonzero samples is 10: more than 10 yields small improvements in the smoothing in the convolution but at the cost of increased computational complexity; less than 10 increases substantially the aliasing produced by infrequent sampling of $v(\xi)$ in Equation 4.13. Let b denote half of the range of values of ξ where $v(\xi)$ is significantly greater than zero, as illustrated in Figure 4.7. We refer to b as the bandwidth of $v(\xi)$. Thus a practical choice of ξ_0 is approximately b/10. (This choice may vary in practical circumstances by factor of 2). The choice of b is determined by a trade-off between the smoothness of modeling $c_i(\xi)$ and the aliasing. In Section 4.3.1 we present a detailed explanation of how to model $c_i(\xi)$

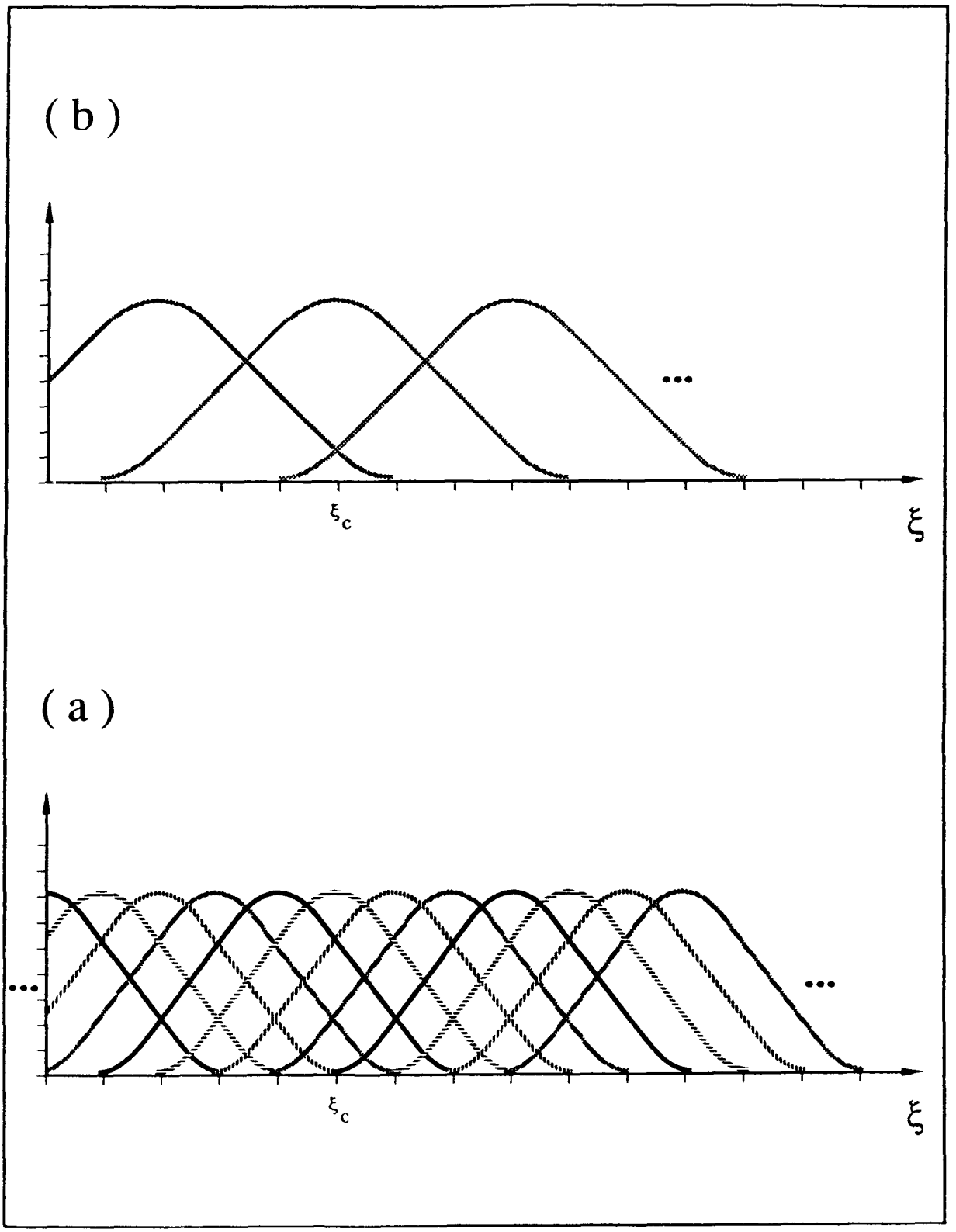


figure 4.8 : The features $\{y_r(0,\xi)\}$ as a result of the convolution operation .

and how the choice of b influences the quality of the model. We found in our experiments that a practical choice of b is :

$$b = \gamma\xi_0$$

Choosing G

Let G denote one-half of the range of values of $w(\xi)$ where $w(\xi)$ is significantly greater than 0, as indicated in Figure 4.7(b) and in Equation 4.15. Our experiments have indicated that a practical choice of G is G = b/2.

Choosing a

The parameter a is chosen experimentally so to satisfy Equations 4.15 and 4.16.

We can represent the operation of the global feature extractor in matrix form. In particular we can represent 4.10 as a matrix multiplication:

$$\mathbf{Y} = \mathbf{VW}$$

where

$$y_{jr} = \sum_{k=1}^{n} v_{jk} w_{kr} .$$

$$y_{jr} = y_r(jx_1),$$
$$v_{jk} = v_k(jx_1),$$
$$w_{kr} = w_r(px_0),$$
$$x_1 = \text{sampling interval},$$
$$j = \text{sampling index}.$$

Y,V,W, are matrices. The kernel k and the feature number r range over {1,...,n} and {1,...,m} respectively. Matrix **W** represents the set of weights. Its dimensions are n by m.

Figure 4.9 illustrates the shape of matrix **W** viewed as a function of k and r. This figure shows the triangular shape of $\{w_r(\xi)\}$. In the preconditioning stage

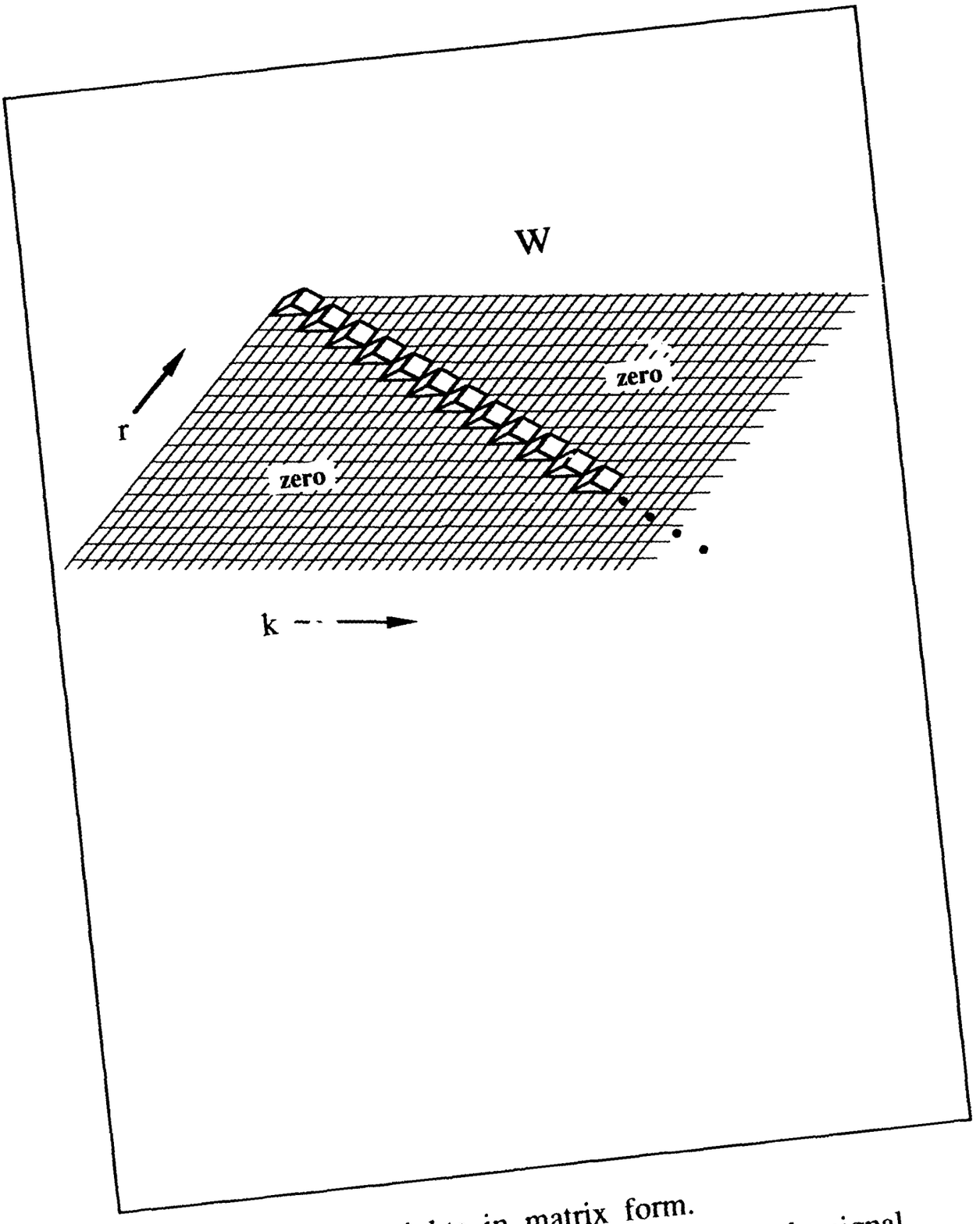


figure 4.9 : The Weights in matrix form.
The triangle shape represents triangle signal.

all the $\{w_r(\xi)\}$ have the same shape. The matrix form may be useful in simulating the network on a digital computer.

Summarizing, we have 1) assumed that the distribution of the weights within the matrix **W** are triangular functions, 2) $y_r(\xi)$ is formed by a convolution between $w_r(p\xi_0)$ and $v(\xi - r\gamma\xi_0)$, and 3) the number of features m are much smaller than the number of kernels n. The reduction ratio n/m is denoted by γ.

4.2.2. Competitive Learning

In this section we describe how the process of global conditioning is represented by the global-local model. After the conditioning there are global changes which have their biggest effects around the conditioning frequency ξ_c. The effect is weakened when ξ is far from ξ_c. During conditioning described in Steps 4,5 in Section 4.1.4 the weights are changed. (The initial values represent the preconditioning stage.) Our global-local model provides a simple rule for changing these weights so as to achieve the tuning of the global feature extractor during conditioning. This tuning is achieved as a form of competitive learning, described below.

We define the global training parameters $\{e_r\}$ for all r. Each e_r (and each local scaling parameter h_i) multiplies $w_{kr}(0)$ to produce Δw_{kr}:

$$\Delta w_{kr} = h_i e_r w_{kr}(0) = h_i e_r w_r(p\xi_0) \tag{4.17}$$

Recall from Equation 4.6 that $w_{kr}(0)$ = the value of $w_{kr}(t)$ for t = 0.
by Equation 4.7 and 4.17,

$$y_r(1,\xi) = \sum_{k=1}^{n} w_{kr}(0) v_k(\xi) + h_i \sum_{k=1}^{n} e_r w_{kr}(0) v_k(\xi) \quad r = 1, \cdots, m .$$

In our experiments we have found that the $\{e_r\}$ are usually zero for all values of r except in a small neighborhood where $r \cong \xi_c / (\gamma\xi_0)$. This is illustrated in Figure 4.10. h_i is the local scaling parameter which is explained in Section 4.3.2.

This result can be expressed in matrix form as follows:

$$\Delta \mathbf{W} = \mathbf{h_i E W(0)},$$

$$\mathbf{W(1)} = \mathbf{W(0)} + \Delta \mathbf{W} = (\mathbf{I} + \mathbf{h_i E})\mathbf{W(0)}$$

where **E** is a diagonal m by m matrix whose r^{th} element is e_r. Figure 4.11 illustrates a typical form of the matrix $\Delta \mathbf{W}$.

Thus global training in the auditory cortex may be represented by $\{e_r\}$ or, equivalently, by **E**. A great advantage of the use of $\{e_r\}$ is that these parameters represent the global training process in a very compact form. In many of our experiments only three nonzero values of e_r were sufficient to represent the global training.

We have developed a primitive trial-and-error procedure for finding the e_r's. This procedure is closely connected with our procedure for finding the local scaling parameters $\{h_i\}$. We discuss both of these procedures at the end of Section 4.3.2.

Figure 4.12 illustrates the results of the convolution operation in postconditioning stage. This figure should be compared to Figure 4.8(a), which illustrates the results of the convolution in the preconditioning stage. The functions $\{y_r(1,\xi)\}$ do not have a same shape as $\{y_r(0,\xi)\}$. The amplitude of the feature corresponds to ξ_c is bigger than the others.

$$y_r(1,\xi) > y_r(0,\xi) \quad \text{for } r = \frac{\xi_c}{\gamma \xi_0},$$

$$y_r(1,\xi) < y_r(0,\xi) \quad \text{for } r = \frac{\xi_c}{\gamma \xi_0} \pm 1$$

Summary of Section 4.2

We have shown how to design the global feature extractor by so as to reduce the large number of kernels to a small number of features, and how to represent global training compactly by the parameters $\{e_r\}$.

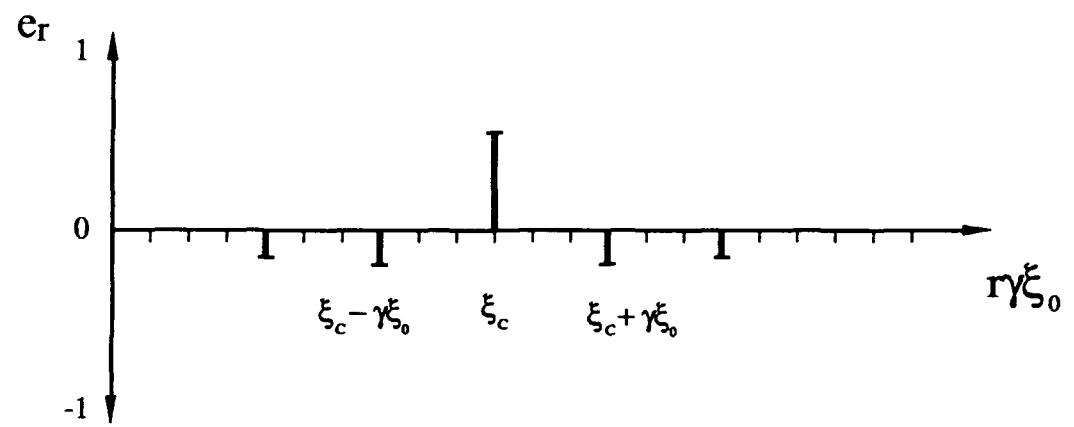


figure 4.10 : Example of $\{e_r\}$.

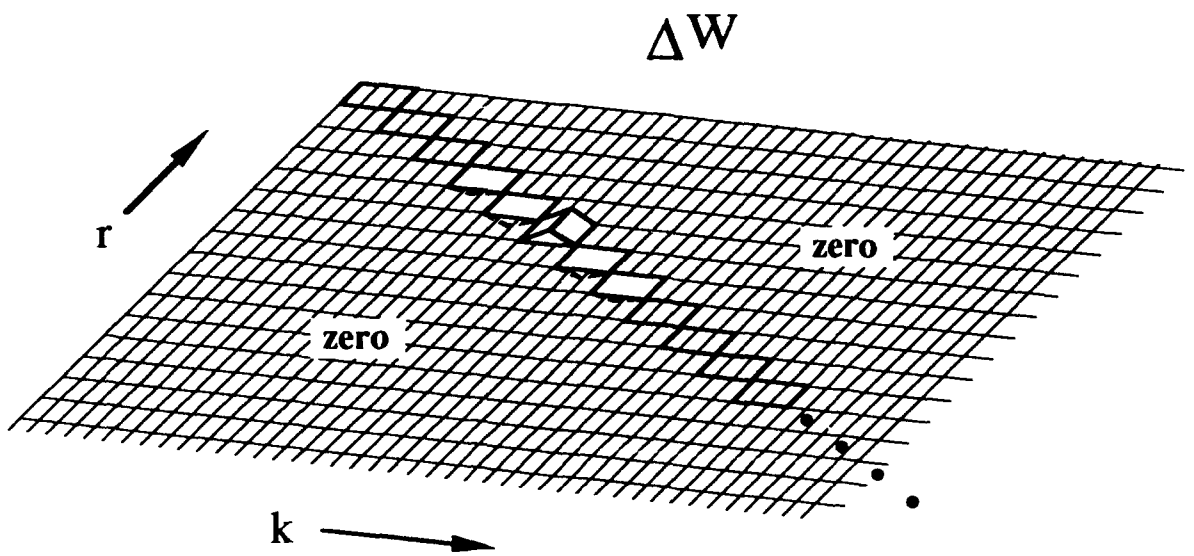


figure 4.11 Matrix Δw

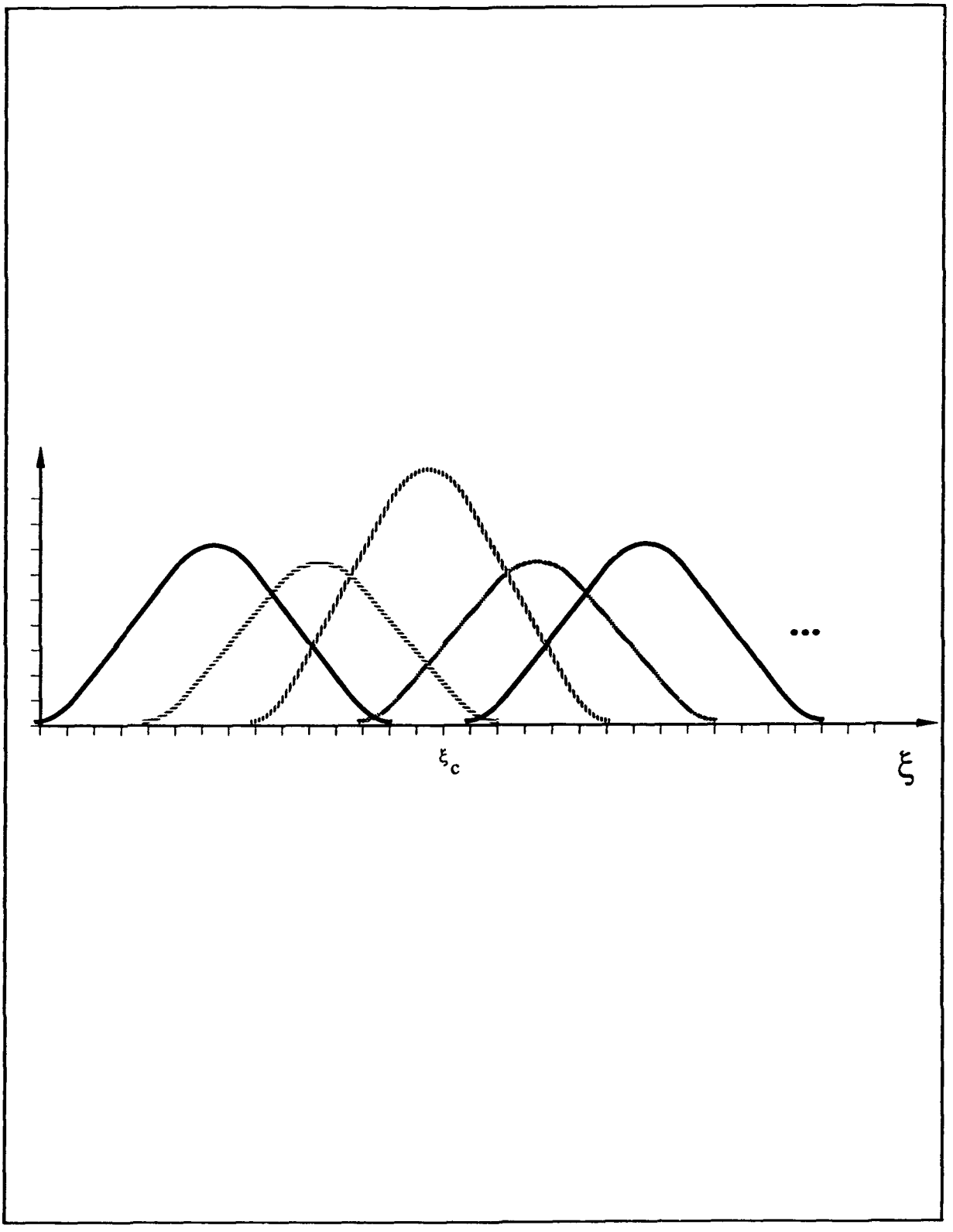


figure 4.12 : Results of convolution in postconditioning stage.

4.3. Local Conditioning

We describe in this Section the details of local training. In the global-local model training is implemented by the local layer. The local layer consists of a set of components $\{L_i\}$, where L_i produces the receptive field $c_i(t,\xi)$. Each L_i consists of a local trainer and a local discriminant, as illustrated in Figure 4.3 All of the L_i's receive the same set of features $\{y_r(t,\xi\}$ and the global training parameters $\{e_r\}$. The features $\{y_r(0,\xi)\}$ and the parameters $\{e_r\}$ are produced by the global layer as described in Section 4.2.

The purposes of each L_i are : a) to represent the preconditioned receptive field $c_i(0,\xi)$, b) to account for local scaling in the conditioning process, c) to account for local tuning in the conditioning process. Local scaling depends on the overall scale of $c_i(1,\xi) - c_i(0,\xi)$, the difference between the postconditioned and preconditioned receptive fields. Local tuning depends primarily on the local best frequency and ξ_c. Let ξ_{bi} denote the local best frequency of , i.e. the value of ξ where $c_i(0,\xi)$ is a maximum. The local tuning affects primarily the amplitude of $c_i(1,\xi)$ for values of ξ between ξ_{bi} and ξ_c.

Local tuning consists primarily of two phenomena: a shift of the preconditioning best frequency toward the conditioning frequency and a suppression of the receptive field at the preconditioning best frequency.

The local tuning implemented by each L_i can exhibit either complete learning or partial learning of the conditioning frequency ξ_c. Complete learning consists of a full shift of the best frequency to the conditioning frequency ("full shift") and significant suppression of the receptive field at the preconditioning best frequency ("suppression"). These changes in the receptive field are illustrated in Figure 4.13(a). Partial learning may be a result of any of the following effects: a) no shift or only a partial shift of the postconditioning best frequency toward the conditioning frequency ("no shift" or "partial shift"); or b) no suppression of the receptive field at the preconditioning best frequency ("no suppression"); or c) both no shift (or partial shift) and no suppression. Figure 4.13(b) illustrates a combination of partial shift and suppression.

In the following section we describe how the model implements the above properties, namely a) representation of the preconditioning stage, b) local scaling, and c) local tuning.

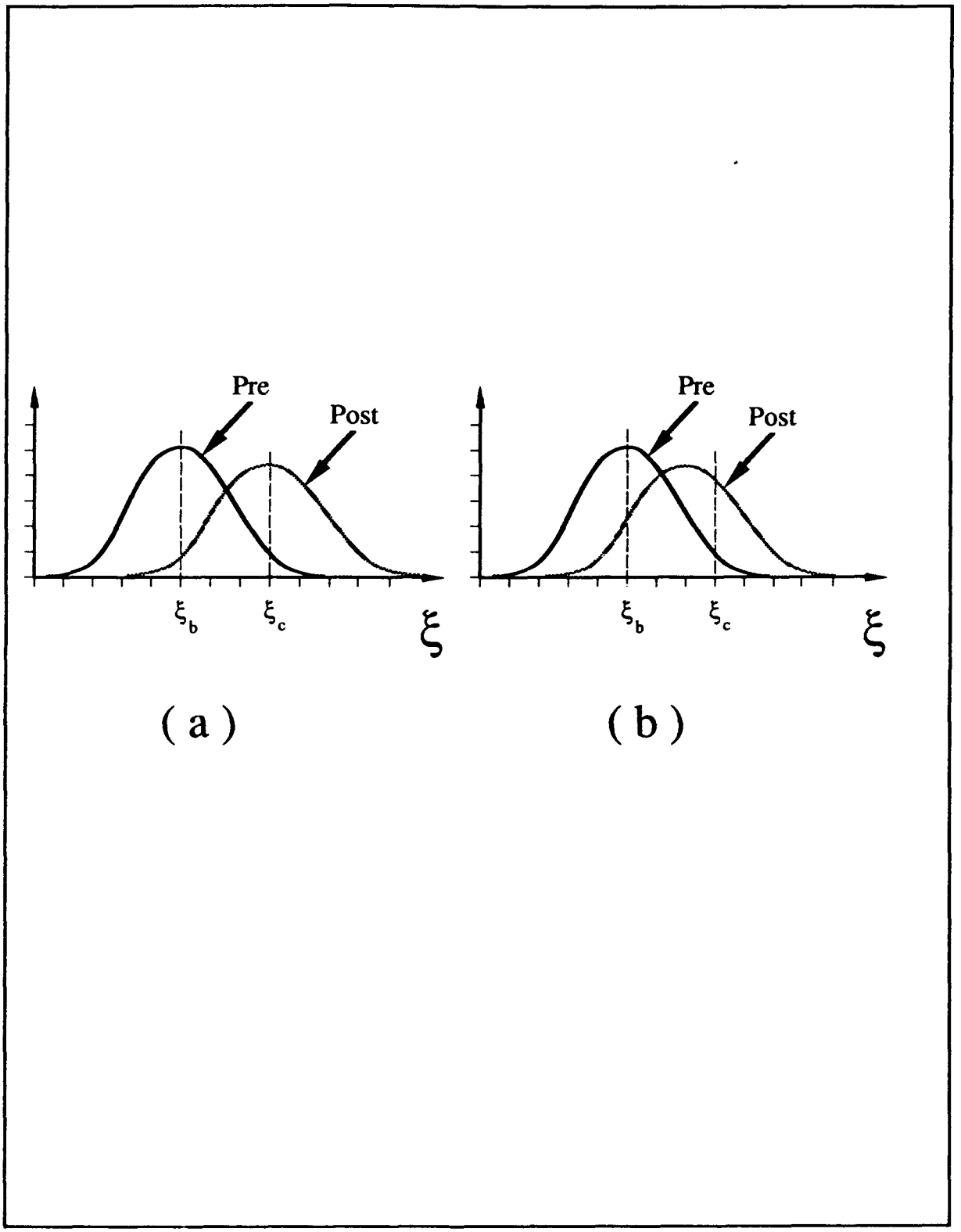


figure 4.13 : (a) Complete learning. (b) Partial learning.

4.3.1. Modeling of Preconditioning

As explained in Section 4.1.4 the i^{th} local discriminant computes the preconditioning curve $c_i(0,\xi)$ by Equation 4.3. The $A_{ir}(0)$'s are the experimental data points of the receptive field observed before conditioning. The $y_r(0,\xi)$'s are the features which are computed by the global layer before conditioning. We can view the $y_r(0,\xi)$'s as interpolation functions. (An interpolation function is usually bell shaped, as illustrated in Figure 4.8.) The $A_{ir}(0)$'s are discrete samples of a function to be interpolated and the $\{y_r(0,\xi)\}$ provide a means of interpolating the values of $c_i(0,\xi)$ for values of ξ between samples. Thus we need a function $y_r(0,\xi)$ which is effective for interpolation. Let $\hat{c}_i(0,\xi)$ denote the values of the receptive field observed at the i^{th} cortical pyramidal neuron. (This is to be distinguished from $c_i(0,\xi)$ produced by the model.)

In our model each $A_{ir}(0)$ is set equal to $\hat{c}_i(0\ ,\ r\gamma\xi_0)$, where $r\gamma\xi_0$ is the r^{th} observed value of ξ. Thus in our model $c_i(0\ ,\ r\gamma\xi_0) \equiv \hat{c}_i(0\ ,\ r\gamma\xi_0)$ for all r. Our model interpolates $c_i(0,\xi)$ for values of ξ not equal to $r\gamma\xi_0$ by the equation:

$$\begin{aligned} c_i(0,\xi) &= \sum_{r=1}^{m} A_{ir}(0)\, y_r(0,\xi) \\ &= \sum_{r=1}^{m} A_{ir}(0)\, y\left(\xi - r\gamma\xi_0\right), \end{aligned} \tag{4.18}$$

where $A_{ir}(0) = \hat{c}_i(0\ ,\ r\gamma\xi_0)$.

The choice of the "bandwidth" b (discussed in Section 4.2.1) affects the modeling of $c_i(0,\xi)$. Too large a value of b may introduce too much smoothing of $c_i(0,\xi)$ thereby suppressing important variations in $c_i(0,\xi)$. Too small a value of b may introducing aliasing in the form of false oscillations in the modeled receptive field. These effects are illustrated in Figures 4.14, 4.15. Figure 4.14 (a) and (b) show the aliasing effect. Figure 4.15(a) shows good modeling. Figure 4.15(b) shows effect of too much smoothing.

4.3.2. Local Scaling

Local training imparts a scaling effect on the $y_r(t,\xi)$'s. We represent this effect by the coefficient h_i. In order to estimate the value of h_i we assume :

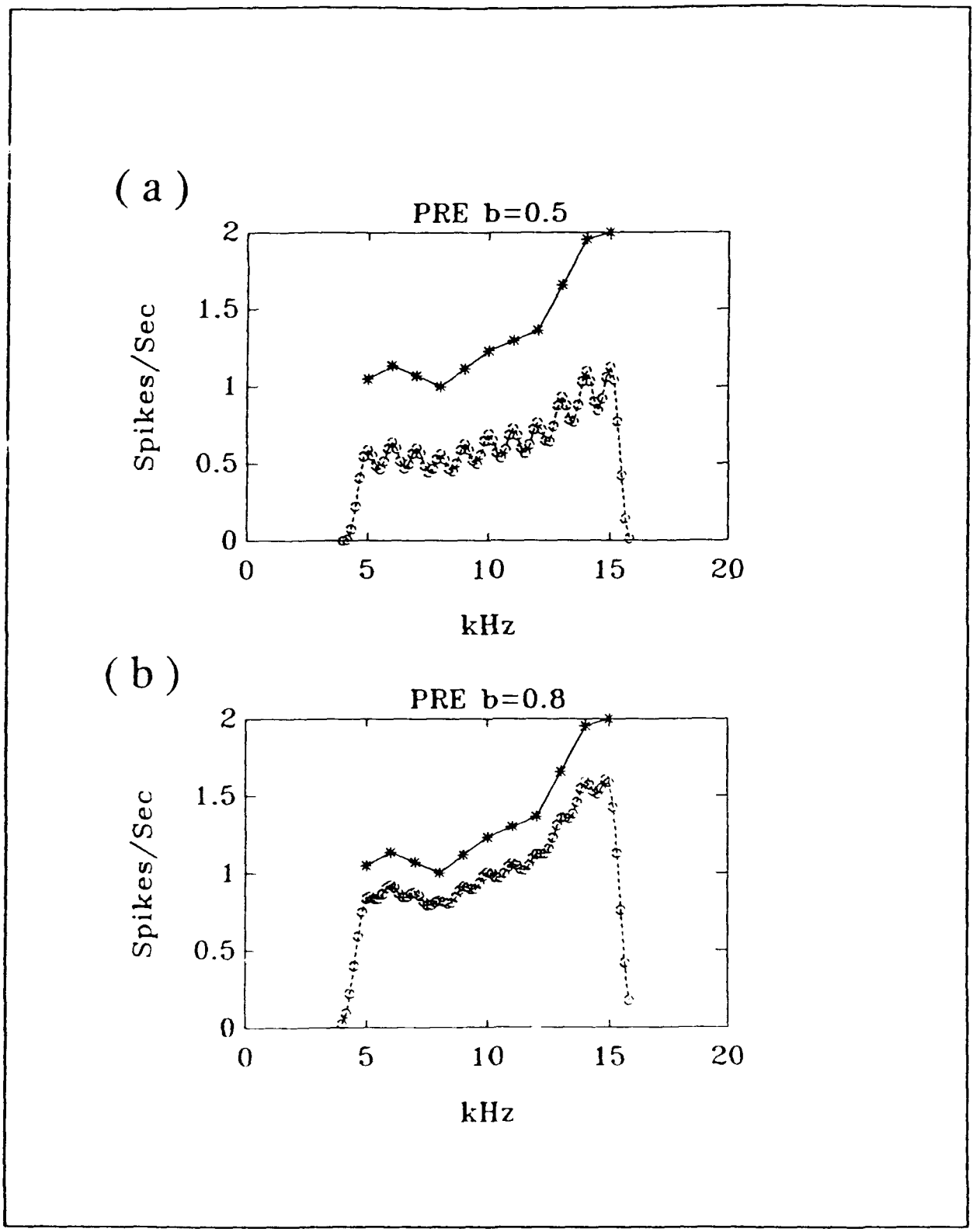
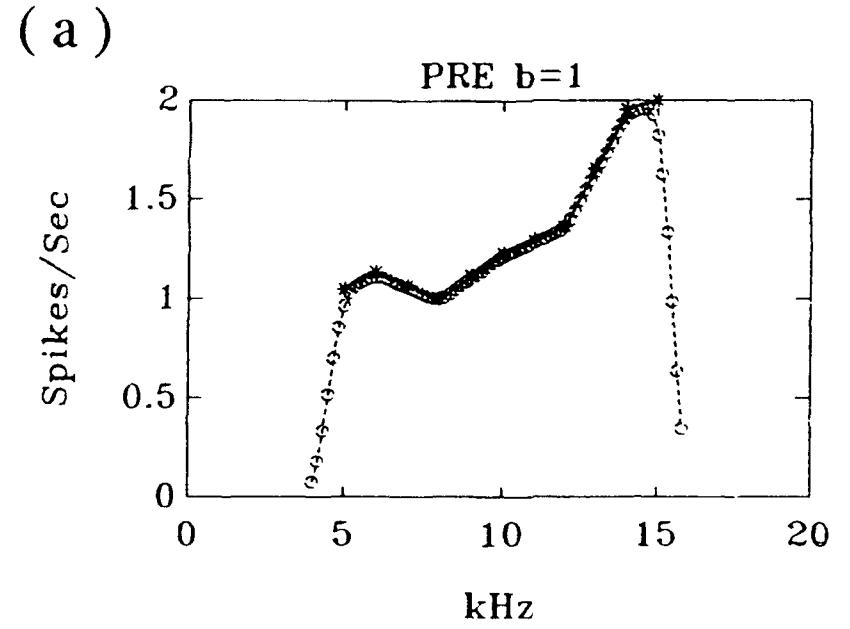


Figure 4.14: The choice of b affects the modeling of $c_1(0,\xi)$. Aliasing effect for different values of b.



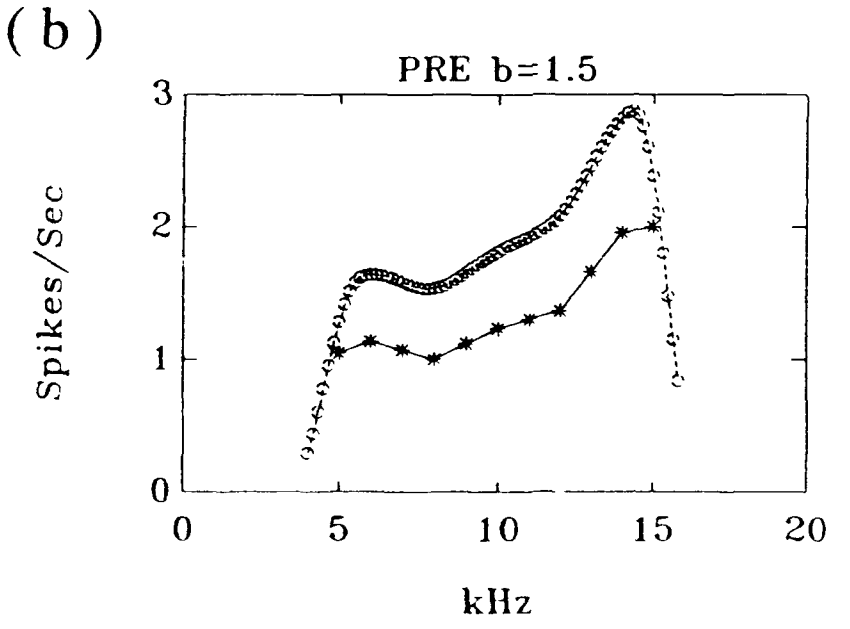


Figure 4.15: The choice of b affects the modeling of $c_1(0,\xi)$. (a) good modeling (b) effect of too much smoothing.

$$c_i(1,\xi_c) \equiv \sum_{r=1}^{m} A_{ir}(0)\, y_r(0,\xi_c) + h_i \sum_{r=1}^{m} e_r A_{ir}(0)\, y_r(0,\xi_c) \qquad \textbf{(4.19)}$$

Thus h_i accounts for the change of $c_i(t,\xi)$ for values of ξ near ξ_c.

We have developed a primitive trial-and-error procedure for finding the h_i's and the $\{e_r\}$. Let I denote that value of i such that $|\dot{\hat{c}}_I(1,\xi) - \dot{\hat{c}}_I(0,\xi)| \geq |\dot{\hat{c}}_i(1,\xi) - \dot{\hat{c}}_i(0,\xi)|$ for all i. First we assign $h_I = 1$. Then the $\{e_r\}$ are determined so that $c_I(1,\xi) \equiv \hat{c}_I(1,\xi)$ for ξ lying in a small neighborhood of ξ_c. For $i \neq I$, the $\{e_r\}$ are the same but h_i is adjusted so that $c_I(1,\xi) \equiv \hat{c}_I(1,\xi)$ for $\xi \equiv \xi_c$. This procedure was not so difficult in our experiments because usually only three nonzero members of $\{e_r\}$ seemed to be sufficient to model the observed change in the receptive fields.

4.3.3. Local Tuning

As explained at the beginning of Section 4.3, the local tuning affects the value of $c_i(t,\xi)$ for values of ξ primarily between ξ_c and ξ_{bi}. We developed a systematic method to change the weights $A_{ir}(0)$'s according to the local tuning which occur during conditioning. This method enables us to predict both complete learning and partial learning. Below we describe our procedure for modeling local tuning in terms of the $A_{ir}(t)$'s for t=0 and t=1.

We define a tuning function $T_i(\xi)$. This function is constructed by the following Steps:

Step 1:
By subtracting Equation 4.5 from Equation 4.19 we obtain:

$$c_i(1,\xi) - c_i(0,\xi) = h_i \sum_{r=1}^{m} e_r A_{ir}(0)\, y_r(0,\xi)\,,$$

Let

$$\Delta c_i(\xi) = c_i(1,\xi) - c_i(0,\xi). \tag{4.20}$$

Figure 4.16 shows an idealized example of the function $\Delta c_i(\xi)$.

Step 2:

Let $b_d = |\xi_c - \xi_b|$. This quantity is the "bandwidth" of a bell shaped function $d_i(\xi)$, defined as follows:
q_c is a coefficient which determines the amplitude of $d_i(\xi)$.

$$d_i(\xi) = \begin{cases} 0.5q_c\left(1 + \cos\dfrac{\pi|\xi - \xi_c|}{b_d}\right) & \text{for } \dfrac{|\xi - \xi_c|}{b_d} < 1 \\ \\ & \text{otherwise} \end{cases} \tag{4.21}$$

Step 3:

We define a negative bell shaped function $p_i(\xi)$. This function expresses the local negative effect for values of ξ near ξ_{bi}. We assume in this model that

$$p_i(\xi) = \begin{cases} -0.5q_b\left(1 + \cos\dfrac{\pi|\xi - \xi_b|}{b_b}\right) & \text{for } \dfrac{|\xi - \xi_b|}{b_b} < 1 \\ \\ 0 & \text{otherwise} \end{cases} \tag{4.22}$$

there q_b is a coefficient which determines the amplitude of $p_i(\xi)$, and b_b is the "bandwidth" of $p_i(\xi)$. This "bandwidth" needs to be small to restrict the effect of $p_i(\xi)$ on $c_i(1,\xi)$ to values of ξ close to ξ_{bi}.

there q_b is a coefficient which determines the amplitude of $p_i(\xi)$, and b_b is the "bandwidth" of $p_i(\xi)$. This "bandwidth" needs to be small to restrict the effect of $p_i(\xi)$ on $c_i(1,\xi)$ to values of ξ close to ξ_{bi}.

Step 4:
The tuning function $T_i(\xi)$ is obtained by

$$T_i(\xi) = q_t[\Delta c_i(\xi) - d_i(\xi) + p_i(\xi)]. \quad \textbf{(4.23)}$$

Figure 4.17 shows an idealized $T_i(\xi)$.

The parameters q_c and q_b, which determine the amplitudes of $d_i(\xi)$ and $p_i(\xi)$ have to be adjusted in order to obtain the best predictions by the model. The parameter q_t is a scaling parameter which is usually equal to 1.
Unfortunately our estimates of these parameters suffered for lack of sufficient experimental data.

The function $T_i(\xi)$ displays the negative effect of local tuning near the best frequency ξ_{bi} and shows the lack of an effect near the conditioning frequency ξ_c. We found that $T_i(\xi)$ enables the model to predict the postconditioning best frequency in both complete learning and partial learning.

Let ΔA_{ir} denotes the difference between $A_{ir}(1)$ and $A_{ir}(0)$ as indicated in Equation 4.8. Our model computes ΔA_{ir} by 3

$$\Delta A_{ir} = A_{ir}(0)\, T_i(r\gamma\xi_0). \quad \textbf{(4.24)}$$

The weights $A_{i1}(0),\ldots,A_{im}(0)$ are set equal to experimental data points $\hat{c}_i(0, r\gamma\xi_c)$ as described in the second paragraph of Section 4.3.1.

This computation enables the determination of $A_{ir}(1)$ by Equation 4.8. Using these $A_{ir}(1)$'s the i^{th} receptive field is computed by combining Equation 4.7 and 4.9: we can express $c_i(1,\xi)$

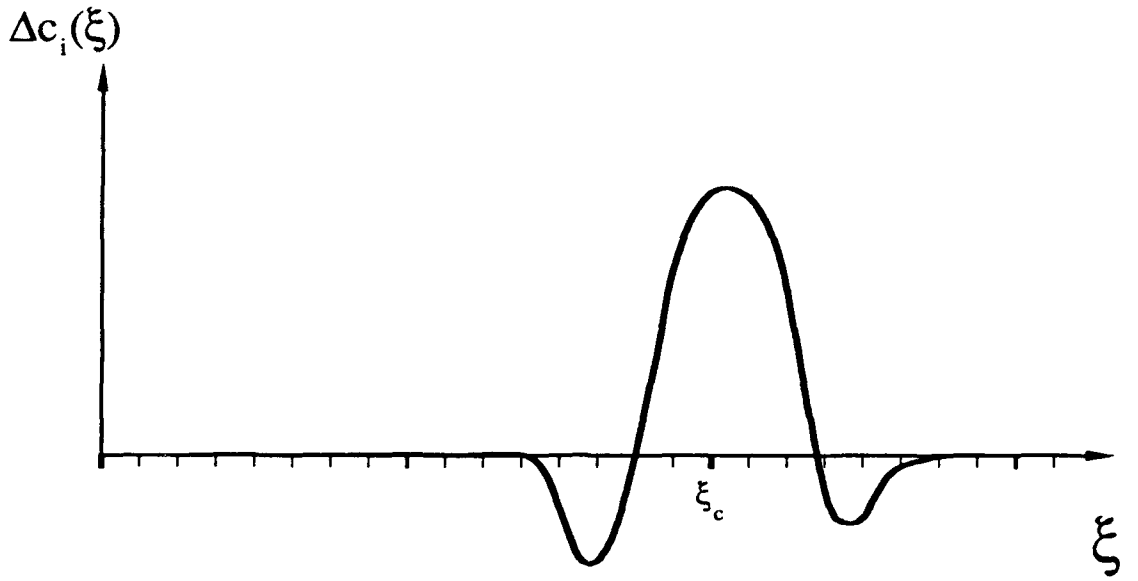


figure 4.16 : Δc_i (ξ).

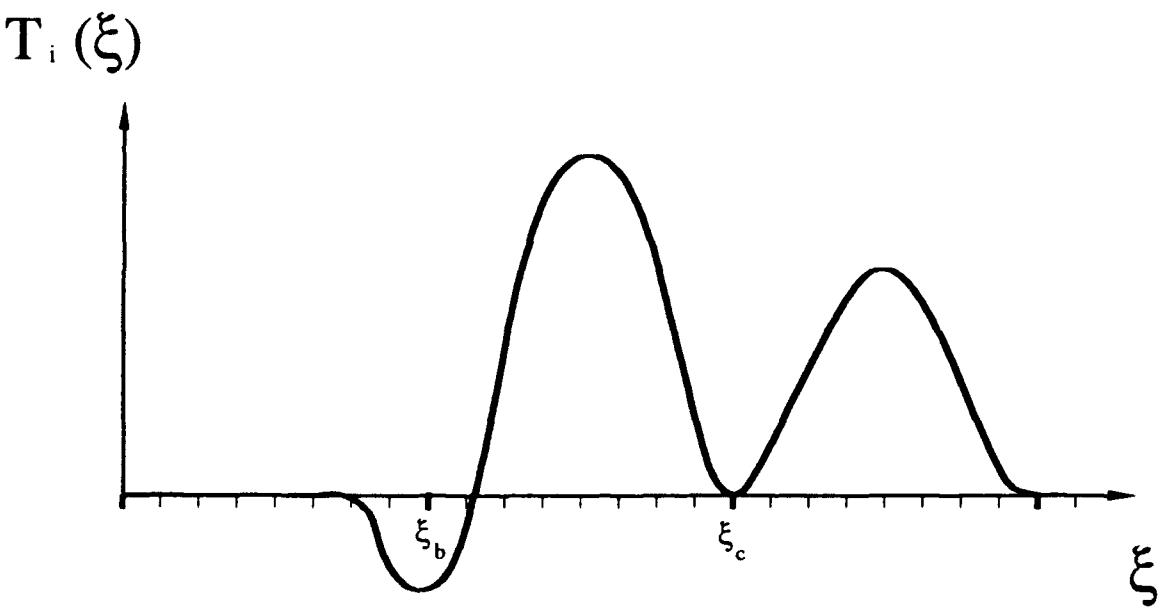


figure 4.17 : The local learning function.

$$c_i(1,\xi) = A_{ir}(1)\, y_r(0,\xi) + h_i \sum_{r=1}^{m} e_r A_{ir}(1)\, y_r(0,\xi)\,. \tag{4.25}$$

Here we see that $c_i(1,\xi)$ is expressed in terms of the global training parameters $\{e_r\}$, the local scaling parameter h_i, and the adjusted local discriminant weights $\{A_{ir}(1)\}$.

There is a possible generalization to the tuning function $T_i(\xi)$. To account for conditioning induced suppression of $c_i(0,\xi)$ at more than one local maximum of $c_i(0,\xi)$, let $p_{ik}(\xi)$ denote functions of the form:

$$p_{ik}(\xi) = \begin{cases} -0.5 q_k \left(1 + \cos \dfrac{\pi |\xi - \xi_k|}{b_k}\right) & \text{for } \dfrac{|\xi - \xi_k|}{b_k} < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.26}$$

where q_k, b_k, and ξ_k are associated with the k^{th} local maximum in $c_i(0,\xi)$. Then the generalized $T_i(\xi)$ will have the form

$$T_i(\xi) = q_t\left[\Delta c_i(\xi) - d_i(\xi) + p_{il}(\xi) + \cdots + p_{ik}(\xi)\right]. \tag{4.27}$$

An idealized example of generalized $T_i(\xi)$ is shown in Figure 4.18.

Summary of Section 4.3

We have shown how to design the local layer so as to account for local scaling and local tuning -- primarily by the local scaling parameter h_i and the local tuning function $T_i(\xi)$.

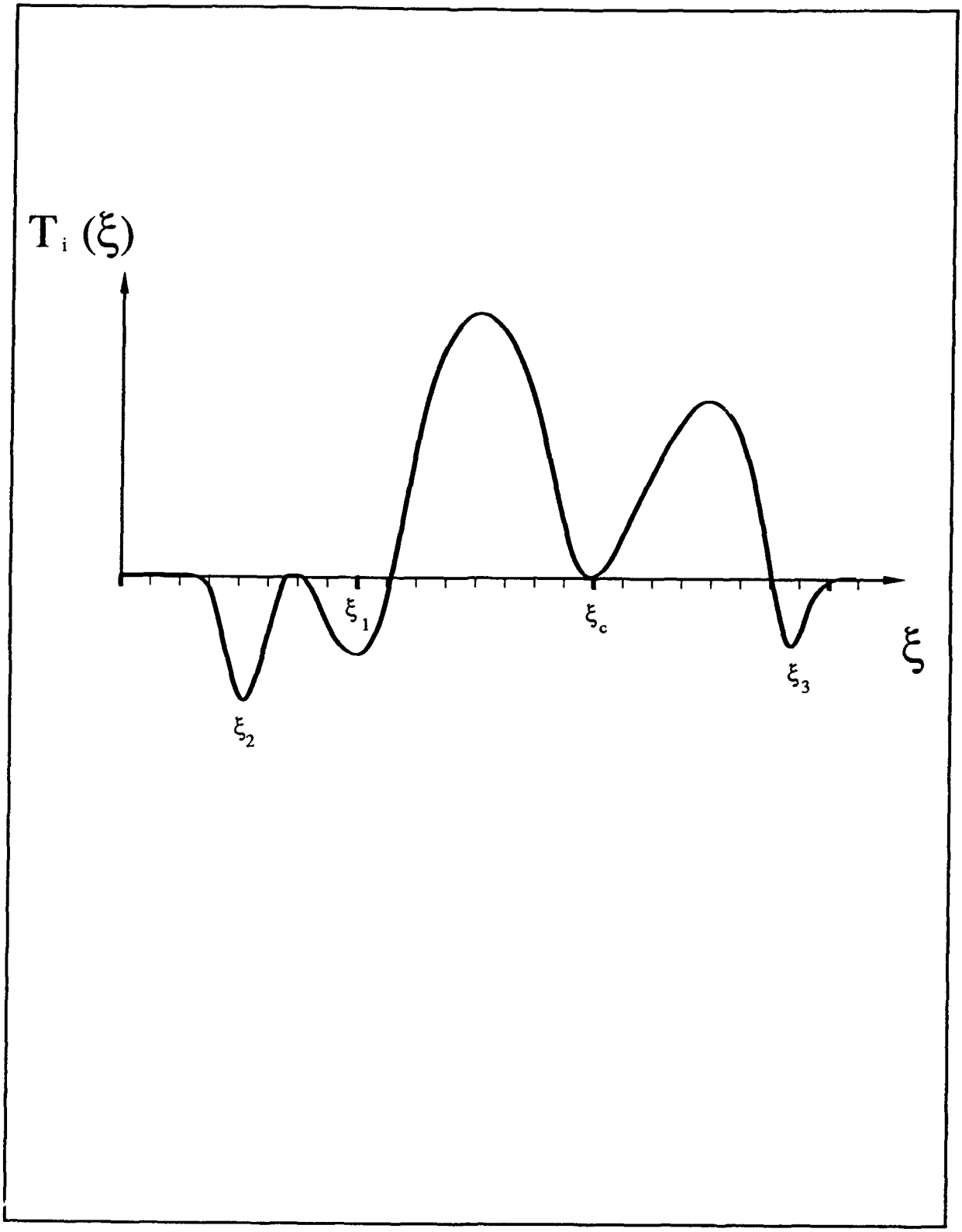


figure 4.18 : The local learning function
with local maxima negative behavior.

4.4. Neurophysiological Interpretation of the Global-Local Model.

Below we suggest a neurophysiological interpretation of each component of the global local model. These components are illustrated in Figure 4.2. See also Section 3.2. for a summary of relevant neurophysiological findings.

Resonators

The resonators represent cells of the ventral medial geniculate nucleus of the thalamus (MGv).

Global Feature Extractor.

The elements of the weight matrix in the global feature extractor represent synapses from the MGv onto cortical pyramidal cells.

Local Discriminant.

The tuning effect carried out by the local discriminant may reflect intracortical effects on the synapses of the MGv to pyramidal cells which originate from other cortical pyramidal cells or from cortical interneurons. Another possibility is that this represents processes inside of the cortical pyramidal cells themselves which modulate the global effect.

Conditioning Control Signal.

The conditioning control symbol g represents a signal from the magnocellular medial geniculate nucleus (MGm) to the cortical pyramidal cells which indicates that reinforcement has followed presentation of the conditioning tone (i.e., that the CS and the UCS are temporally paired).

5.0. Preliminary Evaluation of the Model

We have tested our global local model for its ability to predict postconditioned receptive fields from preconditioning receptive fields and a knowledge of the amplitude and frequency of the conditioning tone, using data from our two electrode experiments. In addition we used data from single-electrode experiments to help us evaluate the validity of local tuning in our model. (At least two electrodes are needed in an evaluation of global conditioning.)

The results of these tests support the proposed model, and encourage us to develop it further.

5.1. Observed Forms of Local Learning

We observed in the data the following phenomena related to local learning.

1. Local scaling: the range of amplitudes of $|c_i(1,\xi) - c_i(0,\xi)|$ may vary from one pyramidal cell to another.

2. Local tuning: consists of a shift of the preconditioning best frequency toward ξ_c and a suppression of $c_i(1,\xi)$ at the preconditioning best frequency.

Local tuning may display either

a) complete learning of ξ_c or

b) partial learning of ξ_c

5.1.2. Predictions

Our model succeeded in predicting the effects of global training, local scaling, local tuning with both complete learning and partial learning, and it succeeded in accounting for multiple maxima in preconditioned receptive fields.

We illustrate these predictions for data obtained in Experiment 9a. We refer to the two electrodes in this experiment as "e1" and "e2".

The observed and predicted receptive fields for electrode e1 in Experiment 9a are shown in Figures 5.1 and 5.2 and for the electrode e2 in Figures 5.3 and 5.4. Note that the best frequency for electrode e1 is greater than ξ_c, and that the best frequency for electrode e2 is less than ξ_c. Nonetheless, conditioning increased responses at the training frequency and decreased responses at the pre-training best frequency in both cases. The curves of the difference between the postconditioned and preconditioned receptive fields, $c_i(1,\xi) - c_i(0,\xi)$, in Figures 5.2 and 5.4 best display the effect of conditioning. The predicted difference curves in these figures display the effect of local tuning including the suppression at the best frequencies, and increased responses at ξ_c.

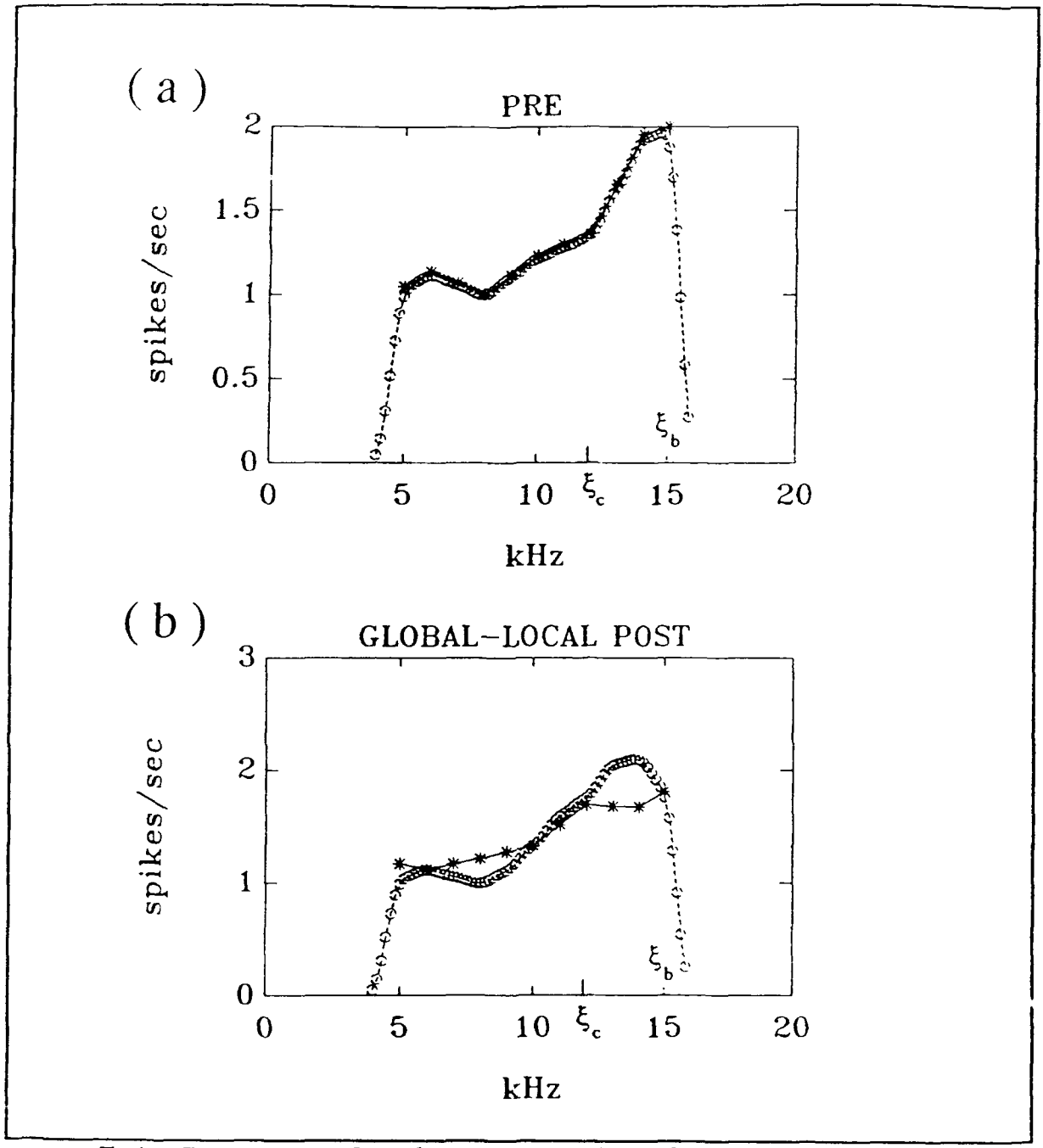


figure 5.1 Experiment 9a, electrode e1, global-local training. (a) Observed points of processed preconditioning receptive field are marked by *, predicted points by o. (b) Similarly for post observed and predicted points.

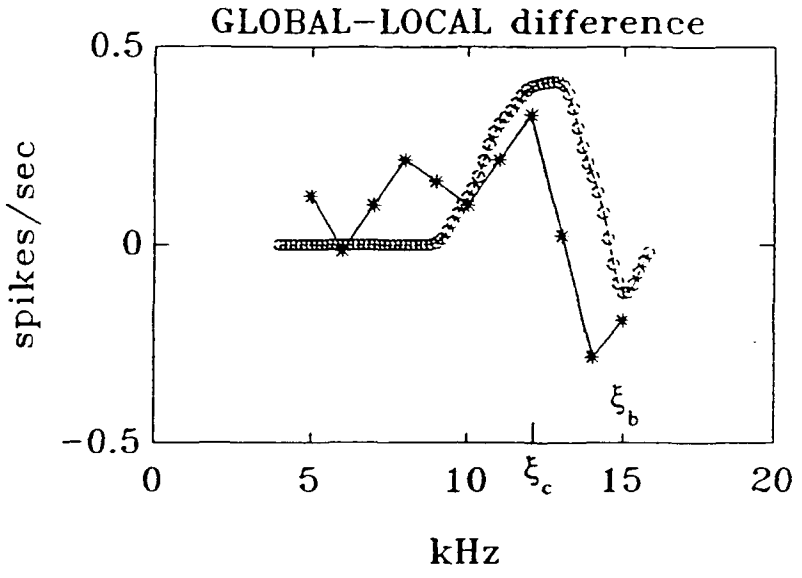


figure 5.2 Experiment 9a, Electrode e1, global-local training: processed difference curve. Observed points are marked by *, predicted points are marked by o.

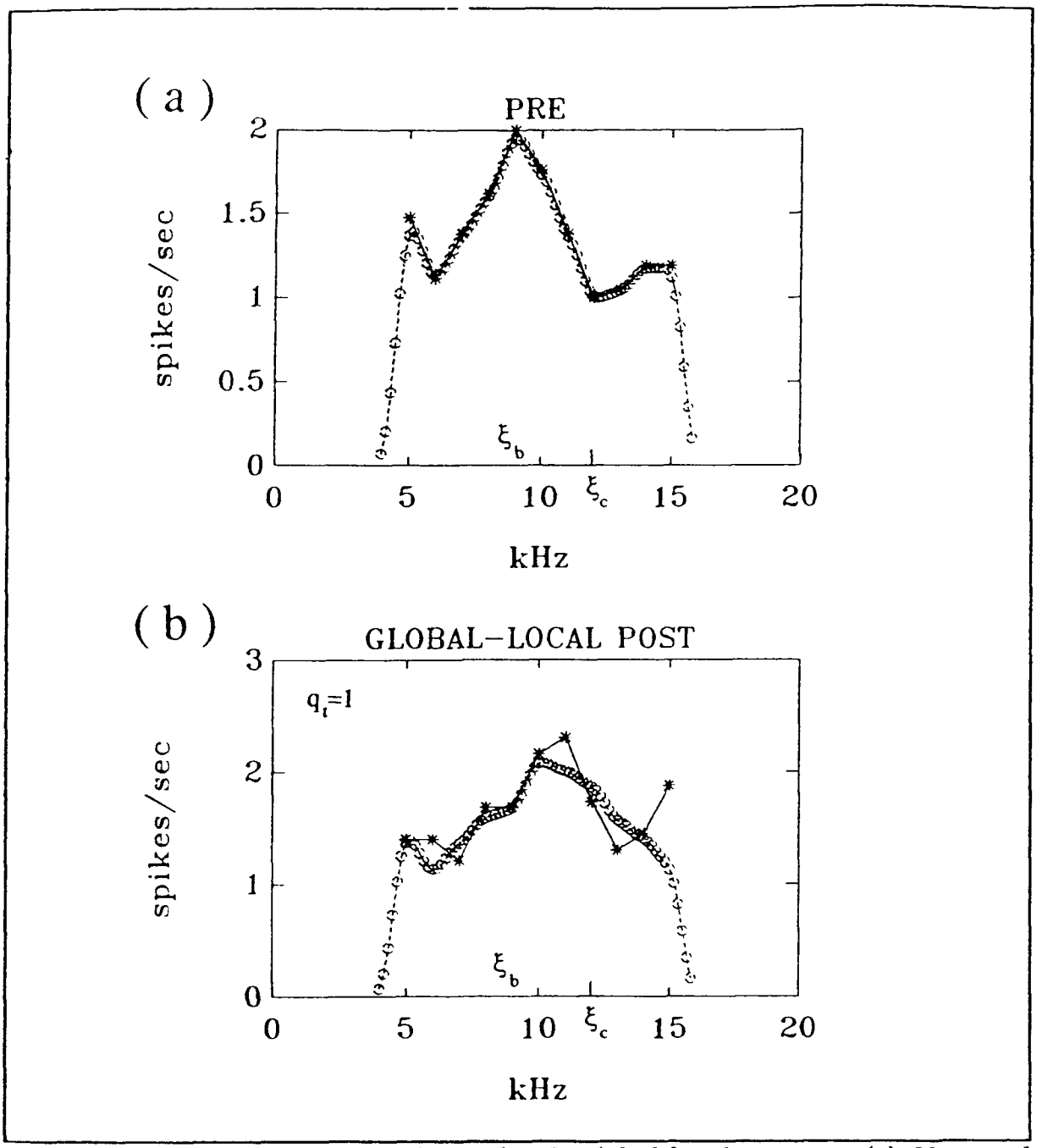


figure 5.3 Experiment 9a, electrode e2, global-local training. (a) Observed points of processed preconditioning receptive field are marked by *, predicted points by o. (b) Similarly for post observed and predicted points.

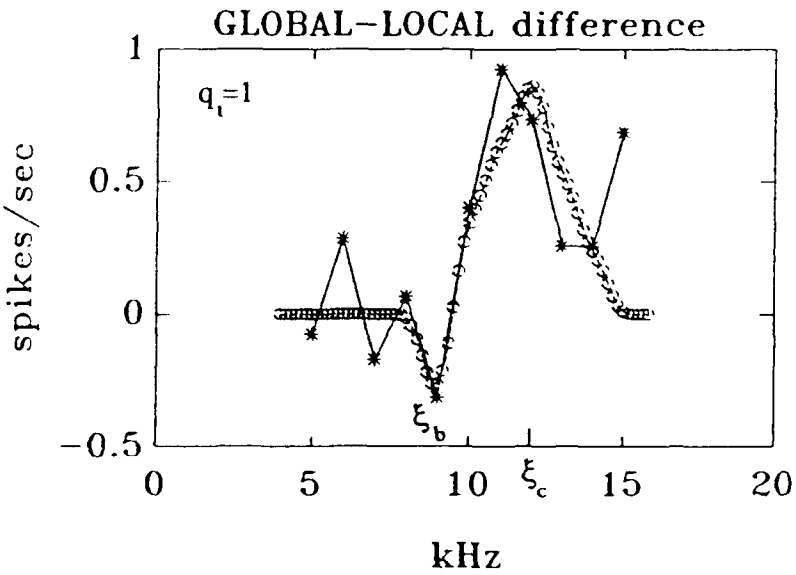


figure 5.4 Experiment 9a, Electrode e2, global-local training: processed difference curve. Observed points are marked by *, predicted points are marked by o.

6. Summary and Conclusions

Evolution has, by natural selection, produced a wide variety of exquisitely sensitive and highly specific biological mechanisms that produce adaptive behavior and species survival. Our recent discovery that sensory neocortex, specifically the auditory cortex, operates on the principle of adaptive filtering, while surprising and unanticipated in the field of sensory neurobiology, fits within the Darwinian framework at two levels.

First, natural selection has provided brains which can 'fine tune" an organism for the environment within which it develops and lives. In this sense, learning processes begin where information specified by the genome ends. because information specified by the gene cannot be sufficiently specific to the varied environments within which an animal finds itself.

Second, while "fine tuning" is a metaphor for learning, it is also an actual description of the operation of the auditory cortex. That is, the cortex is finely tuned by learning to match not simply the physical environment, but also the behaviorally-significant environment. Moreover, the process of adaptive filtering, by which this occurs, appears to be selectionistic in a real competitive sense. Behaviorally important stimuli gain at the expense of less important stimuli in the responses of cortical cells and undoubtedly in the area of the cortex which they command. In short, adaptive filtering seems to be a fundamental process, well honed in evolution.

Although our global-local model is the first which has addressed adaptive filtering in the auditory cortex, it has surprising accuracy and power. While we do not claim that this model provides a complete and exhaustive account of adaptive filtering, nonetheless its successes should not be minimized. The model is both highly structured and quantitative yet sufficiently flexible to account for a wide variety of individual expressions of adaptive plasticity. A unique feature of the model is its incorporation of both global and local learning processes. Either one alone cannot account for adaptive filtering. A hallmark of our global-local model is that it is based upon anatomical and functional architectures, that is, it operates within the biological constraints of brain operation. This contrasts sharply with models in which networks are based on random connectivity.

It is almost always the case that more research and refinement are needed; the results of the present relatively brief project are no exception. Yet, the global-

local model is sufficiently well developed to serve as an impetus for the creation of prototype auditory learning systems.

***

References

Bakin, J.S. and Weinberger, N.M. (1990). Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. Brain Research, 536, 271-286.

Diamond, D.M. and Weinberger, N.M. (1986). Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. Brain Research, 372: 357-360

Diamond, D.M. and Weinberger, N.M. (1989). The role of context in the expression of learning-induced plasticity of single neurons in auditory cortex. Behavioral Neuroscience, 103: 471-494.

Edeline, J-M. and Weinberger, N.M. (1991a). Subcortical adaptive filtering in the auditory system: Associative receptive field plasticity in the dorsal medial geniculate body. Behavioral Neuroscience, 105, 154-175.

Edeline, J-M and Weinberger, N.M. (1991b). Thalamic short term plasticity in the auditory system: Associative retuning of receptive fields in the ventral medial geniculate body. Behavioral Neuroscience, 105, 618-639

Edeline, J-M. and Weinberger, N.M. (1992). Associative retuning in the thalamic source of input to the amygdala and auditory cortex: Receptive field plasticity in the medial division of the medial geniculate body. Behavioral Neuroscience, 106, 81-105

Racanzone, G.H., Schreiner, C.E., Hradek, G.T., Sutter, M.L., Beitel, R.E. & Merzenich, M.M. (1991). Functional reorganization of the primary auditory cortex in adult owl monkeys parallel improvements in performance at an auditory frequency discrimination task. Society for Neuroscience Abstracts, 17, 534.

Scheich, H. & Simonis, C. (1991). Conditioning changes frequency representation in gerbil auditory cortex. Society for Neuroscience Abstracts, 17, 450.

References (cont'd)

Weinberger, N.M., Diamond, D.M. and McKenna, T.M. (1984). Initial events in conditioning: Plasticity in the pupillomotor and auditory systems. In: G. Lynch, J.L. McGaugh, and N.M. Weinberger, (Eds.), Neurobiology of Learning and Memory, New York: The Guilford Press, pp. 197-227.

Weinberger, N.M., Ashe, J.H., Metherate, R., McKenna, T.M., Diamond, D.M., and Bakin, J.S. (1990a). Retuning auditory cortex by learning: A preliminary model of receptive field plasticity. Concepts in Neuroscience, 1: 91-132.

Weinberger, N.M., Ashe, J.H., Metherate, R., McKenna, T,M., Diamond, D.M., Bakin, J.S., Lennartz, R.C., and Cassady, J.M. (1990b). Neural adaptive information processing: A preliminary model of receptive field plasticity in auditory cortex during Pavlovian conditioning. In: Neurocomputation and Learning: Foundations of Adaptive Networks, M. Gabriel and J. Moore (Eds.) Bradford Books/MIT Press, Cambridge, Mass, pp. 91-138.

Publications Supported by N00014-89-J-3187

Edeline, J-M and Weinberger, N.M. (1991). Thalamic short term plasticity in the auditory system: Associative retuning of receptive fields in the ventral medial geniculate body. Behavioral Neuroscience, 105, 618-639

Edeline, J-M. and Weinberger, N.M. (1992). Associative retuning in the thalamic source of input to the amygdala and auditory cortex: Receptive field plasticity in the medial division of the medial geniculate body. Behavioral Neuroscience, 106, 81-105

Weinberger, N.M. (In press). Beyond neuronal excitability: Receptive field analysis reveals that association specifically modifies the representation of auditory information. In: Neuropsychology of Memory (2nd ed.) L. Squire & N. Butters (Eds.) New York: Guilford Press. In press.

Weinberger, N.M., Ashe, J. & Edeline, J-M. (In press). Learning-induced receptive field plasticity in the auditory cortex: Specificity of information storage. In: Neural Bases of Learning and Memory. J. Delacour (Ed.) World Scientific Publishing: Singapore. In press.