How spiking neurons give rise to a temporal-feature map: From synaptic plasticity to axonal selection

Christian Leibold, Richard Kempter, and J. Leo van Hemmen

Physik Department, Technische Universität München, D-85747 Garching bei München, Germany

A temporal-feature map is a topographic neuronal representation of temporal attributes of phenomena or objects that occur in the outside world. We explain the evolution of such maps by means of a spike-based Hebbian learning rule in conjunction with a presynaptically unspecific contribution in that, if a synapse changes, then all other synapses connected to the same axon change by a small fraction as well. The learning equation is solved for the case of an array of Poisson neurons. We discuss the evolution of a temporal-feature map and the synchronization of the single cells' synaptic structures, in dependence upon the strength of presynaptic unspecific learning. We also give an upper bound for the magnitude of the presynaptic interaction by estimating its impact on the noise level of synaptic growth. Finally, we compare the results with those obtained from a learning equation for nonlinear neurons and show that synaptic structure formation may profit from the nonlinearity.

I. INTRODUCTION

Nerve cells that are tuned to *temporal* features of a stimulus have been found in a number of auditory brainstem nuclei, where neuronal activity conserves the temporal structure of a sound stimulus to a certain degree [1–4]. A well-known example of a temporal feature of a sound stimulus is the so-called interaural time difference (ITD), i.e., the difference in the arrival time of a sound between both ears. The ITD is a measure of the spatial position of a sound source relative to the head, and in the brain of many animals neurons with specific ITD tuning are quite common. ITD-tuned neurons are characterized by a best ITD at which their firing rate is maximal.Moreover, ITD-sensitive neurons have been found to be spatially ordered according to their best ITD in at least two species [5,7].This is an example of a temporal-feature map. In contrast to maps of spatiotemporal features, which are, in principle, well understood [42], the explanation of how maps of merely temporal features can arise was an open problem until now. In particular, it was unclear how precision at a time scale of 10 μs can be achieved, because previously analyzed interaction mechanisms are too slow. Here we present a comprehensive theoretical analysis of a mechanism [8] that allows map formation as it occurs in the laminar nucleus of the barn owl, the first stage of the ascending auditory pathway receiving input from *both* ears.

The barn owl is a nocturnal predator, able to catch mice in complete darkness. Its resolution of the azimuthal position of a sound stimulus is 2° comparable to that of humans. But in contrast to humans, the barn owl's distance between both ears is only 5 cm so that the azimuthal localization task is much more difficult since the interaural time differences are the only relevant cue [10,11]. The lack of physical distance between a barn owl's ears is compensated by a higher temporal precision of phase locking of about 40 μs along the auditory pathway up to the laminar nucleus [5,12,13]. The best ITD of laminar neurons gradually changes along a specific direction within the laminar nucleus [5], cf. Fig. 1.

For neurons to form a map, i.e., an in some sense ordered arrangement of their best ITD, they need an interaction between each other. Here we analyze a map-formation mechanism [8] and show analytically how spike-based Hebbian learning [14–32] in conjunction with "axon-mediated spike-based learning" [33–37] leads to map formation. This is the first explicit example of a process leading to a temporal-feature map. If we identify an ITD with the azimuthal position of the acoustic stimulus, the map is an orderly representation of the owl's spatial environment.

In Sec. II we introduce the mathematical framework that is used to describe synaptic plasticity in the laminar nucleus. We define the Poisson neuron, explain the dynamics govern-
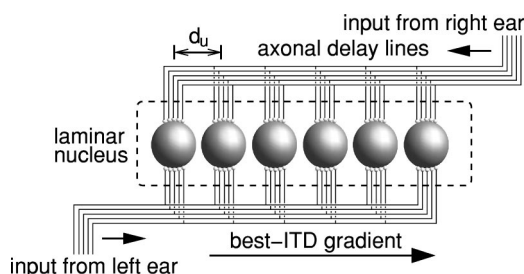

FIG. 1. Sketch of the neuronal anatomy in the barn owl's laminar nucleus [5]. Neuronal activity from one auditory frequency channel is conveyed (small arrows) by phase-locked spike trains in axon bundles (thin lines) that come from the left and right ear, run in parallel (to the dorsoventral direction, indicated by the long horizontal arrow), and contact neurons (large gray spheres) through synapses (small white balls). Measuring firing rates of neurons along this direction, one finds that the neuronal site where the firing rate is maximal varies continuously with the azimuthal location of the stimulus. Neurons are taken to be equidistant with d_u. In order to preserve the temporal structure of sound in the firing patterns of the afferent axons, temporal dispersion among the hundreds of axons has to be small. It has been shown that in the young animal, temporal dispersion is high, whereas in the adult owl, neuronal activity arrives at the laminar neurons temporally highly correlated [6].

ing synaptic modifications, and show how to adapt it to biological reality. We formulate a linear differential equation for the time evolution of the synaptic weights. Its solution is presented in Secs. III and IV. A quantitative measure of a map's quality is proposed in Sec. V. In Sec. VI we then discuss the amplification of spike noise by unspecific axonal learning. Finally, we show in Sec. VII that nonlinearities do not modify the major behavior of our model. To this end we prove that there exists a unique fixed point. Linearizing the dynamics about the fixed point, we find that for generic initial conditions the synaptic efficacies are first attracted towards it before the system evolves into the direction of the eigenvector whose eigenvalue has the largest positive real part.

II. THE MODEL

The present section gives a mathematical description of a biological system that exhibits the formation of an ITD map. We extend the learning equation for a single neuron [17,24,25] to an equation describing the development of synaptic couplings in a network of neurons.

A. The Poisson neuron

For simplicity, the analysis of the dynamics of synaptic transmission is at first performed by means of a linear stochastic neuron model, the Poisson neuron [17,24]. We will show in Sec. VII that taking into account nonlinearities yields almost identical results.

Given the membrane potential $v(t)$ of the linear Poisson neuron at time t, its firing probability is defined to be

$$\lim_{\delta t \to 0} \frac{\text{Prob}\{\text{neuron fires in } [t, t+\delta t]\}}{\delta t} := p_F(v(t)), \quad (1)$$

where p_F is a *linear* function

$$p_F(v) = \beta^{(0)} + \beta^{(1)} v, \quad (2)$$

$\beta^{(0)}$ and $\beta^{(1)}$ being positive constants. The occurrence of more than one spike in a time interval of length δt is $o(\delta t)$. Disjoint intervals being independent, the model is an inhomogeneous Poisson process [24,38].

To obtain an expression for the membrane potential, we use a simple model of synaptic transmission where a synapse is described by only a single parameter J_{mn} that weightens the excitatory postsynaptic potentials (EPSPs) induced by the nth input line ($1 \le n \le N$) to the neuron m ($1 \le m \le M$). It is called synaptic weight or efficacy. Let us assume that the fth spike at neuron n occurs at time $t_n^{(f)}$. After an axonal transmission delay Δ_{mn} from neuron n to neuron m, the spike evokes an EPSP $J_{mn}\epsilon(t - t_n^{(f)} - \Delta_{mn})$, see Fig. 2. We assume the response kernel ϵ (cf. Fig. 3) to be normalized so that $\int ds\, \epsilon(s) = 1$, causal, i.e., $\epsilon(s) = 0$ for $s \le 0$, and positive meaning simply $\epsilon(s) \ge 0$ for all s. The linear superposition of EPSPs evoked at the mth neuron by a set of N presynaptic cells yields the postsynaptic membrane potential
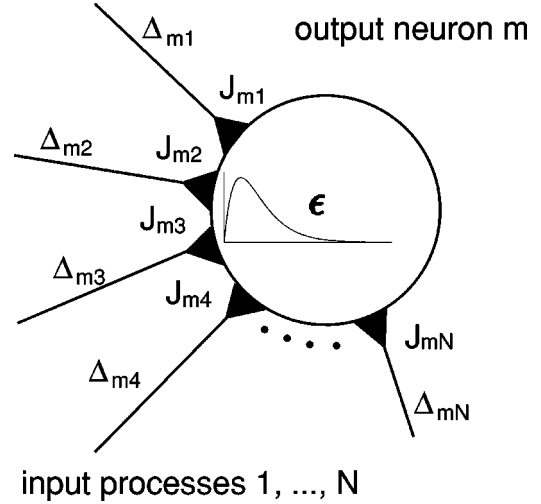

FIG. 2. Synaptic model. Each of N synapses connecting to output neuron m is described by a single variable J_{mn}, which weights the linearly superimposed spike response kernels ϵ that are retarded by the axonal transmission delay Δ_{mn}. The sum of responses yields the postsynaptic membrane potential as it is described by Eq. (3).

$$v_m(t) = \sum_{n=1}^{N} J_{mn} \sum_{\{t_n^{(f)}\}} \epsilon(t - t_n^{(f)} - \Delta_{mn}). \quad (3)$$

The second sum in Eq. (3) is meant to run over all firing times $t_n^{(f)} < t$ of neuron n.

B. Homosynaptic Hebbian learning

The synaptic weight J_{mn} is assumed to change in dependence upon the timing of presynaptic and postsynaptic spikes. As a starting point for the subsequent description of a
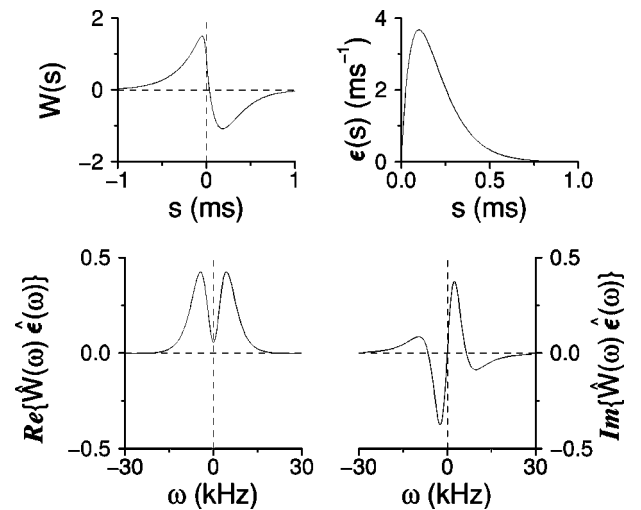

FIG. 3. Generic choice of the learning window W and the response kernel ϵ (upper panels). The convolution of W and ϵ, as they have been used in [8], acts as a bandpass filter, which is reflected by the real and imaginary parts of the Fourier transform $\hat{W}(\omega)\hat{\epsilon}(\omega)$ (lower panels).

synaptic weight's dynamics, we use a generalization of a local Hebbian rule for one postsynaptic cell [8,15,17,25].

An input spike arriving at synapse n of neuron m changes the weight J_{mn} by a constant amount ηw^{in}, where η is a small, positive scaling factor. Similarly, each output spike of neuron m results in a weight change ηw^{out} of all synapses of that neuron. Pairs of input and output spikes at times $t_n^{(f)}$ and $t_m^{(f)}$, respectively, lead to a weight change $\eta W(t_n^{(f)} - t_m^{(f)})$, where W is called the "learning window" [15].

For Poissonian input spike trains, it is shown in [17] that the alterations of J_{mn} according to the above procedure can be expressed as a differential equation solely depending on the time averages (see below) of ensemble-averaged input and output firing rates ν_n^{in} and ν_m^{out} (defined in Sec. II D 1), and the time-averaged correlation function between presynaptic spikes from neuron n and postsynaptic spikes at neuron m,

$$
C_{mn}(t+r,t) = \mathcal{T}^{-1} \int_{t-\mathcal{T}}^{t} dt' \lim_{\delta t \to 0} (\delta t)^{-2}
$$

$$
\times \text{Prob}\{\text{output cell m fires in } [t'+r, t'+r
$$

$$
+ \delta t) \text{ and input cell n fires in } [t', t'+\delta t)\}.
$$

$$
(4)
$$

The dynamical equation from [17] then reads

$$
\left(\frac{d}{dt} J_{mn} \right)_{\text{local}} = \eta \left[w^{\text{in}} \nu_n^{\text{in}}(t - \Delta_{mn}) + w^{\text{out}} \nu_m^{\text{out}}(t) \right.
$$

$$
\left. + \int_{-\infty}^{\infty} ds\, W(s) C_{mn}(t, t+s-\Delta_{mn}) \right]. \quad (5)
$$

This learning rule can account for synaptic structure formation at the level of a single neuron.

The so-called learning window W is a function of the time difference between a presynaptic and postsynaptic spike [15]. One can define its temporal width $\mathcal{W}$ as the interval of time differences s, where $W(s)$ is not negligibly small. It can be shown [17] that Eq. (5) is valid, if the averaging time $\mathcal{T}$ greatly exceeds $\mathcal{W}$. Furthermore, the scaling factor $\eta > 0$ should be small in the sense that the alteration of the efficacy on the averaging time scale $\mathcal{T}$ is far below the efficacy J_{mn} itself. As a result, the upper limit of $\mathcal{T}$ is the typical time scale of the dynamics of synaptic weights.

In addition to [17], we have included the axonal transmission delays Δ_{mn} that shift the time-averaged rate ν_n^{in} and the correlation function C_{mn}. The shift of ν_n^{in} can be neglected (see Sec. II D 1). The shift in the second argument of the correlation function, however, will turn out to be the essential step for temporal map formation.

C. Axon-mediated spike-based learning

In order to explain map formation we have to take into account an interaction between synapses of different neurons that coordinates the development of synaptic weights across the postsynaptic index m. By means of computer simulations

[8,9] it has been shown that synaptic changes propagating along axons are necessary for the evolution of a map of interaural time differences in the barn owl's laminar nucleus.

The presynaptic spread of weight modifications is implemented in a way that every local alteration $(dJ_{mn}/dt)_{\text{local}}$ of a synaptic weight in Eq. (5) is propagated to all other synapses at the same axon n. This yields the learning equation

$$
\frac{d}{dt} J_{mn} = \sum_{m' \in \text{axon } n} (\delta_{mm'} + \rho b_{mm'}) \left(\frac{d}{dt} J_{m'n} \right)_{\text{local}}, \quad (6)
$$

where $\delta_{mm'}$ denotes the Kronecker delta. The axonal coupling matrix $\rho b_{mm'}$ also accounts for the spatial range of the presynaptic interaction. The positive scaling factor ρ determines the overall strength of the interaction between neurons. We note that $b_{mm+1} \approx 1$ for all m and assume $b_{mm'} \geq 0$ for all m, m'. Below we will often set $b_{mm'} = 1$ for all $m \neq m'$.

D. Biological constraints

The solution to Eq. (6) depends strongly on the topology of the network as well as on the statistics of the presynaptic input activity. Specifying the axonal coupling matrix $\rho b_{mm'}$ and the input process p^{in}, in order to mimic the anatomy and physiology of the laminar nucleus of the barn owl, provides us with a noteworthy simple example of temporal map formation.

1. Poissonian input

According to the frequency decomposition in the ear, the barn owl's auditory brainstem is tonotopically organized. Neurons, therefore, belong to a specific frequency layer and carry spike trains phase locked to the acoustic input within the respective frequency band. We intend to model map formation within one isofrequency layer. We, therefore, suppose that all afferent axons are carrying similar temporal information in a way that spikes at $N \gtrsim 100$ presynaptic neurons are generated by identical Poisson processes with intensity $p^{\text{in}}(t)$. Presynaptic neurons are assumed to be statistically independent. Axons to postsynaptic neurons may have different delays. Therefore spikes from the presynaptic neuron n arrive at the postsynaptic neuron m with delay Δ_{mn}, i.e., they are generated by means of the Poisson intensity $p^{\text{in}}(t - \Delta_{mn})$.

We model phase-locked afferent activity by an inhomogeneous Poisson process with periodical intensity $p^{\text{in}}(t) = p g(t)$, where $g \geq 0$ is a T_p-periodic function with $\int_0^{T_p} d\tau\, g(\tau) = 1$. The parameter $p = \nu T_p$ is adjusting the process to a mean firing rate ν.

We have assumed that the averaging time $\mathcal{T}$ greatly exceeds T_p. As a result, temporal averages are translationally invariant in time. The time-averaged firing rate $\nu^{\text{in}}(t) = \overline{p^{\text{in}}}(t) = \mathcal{T}^{-1} \int_{t-\mathcal{T}}^{t} dt'\, p^{\text{in}}(t') = \nu$ is constant and identical for all input cells. Here, and elsewhere, an overbar denotes a time average over $\mathcal{T}$. The time-averaged correlation function between input spike trains at synapses n and n' (a detailed deduction can be found in [17], Appendix A),
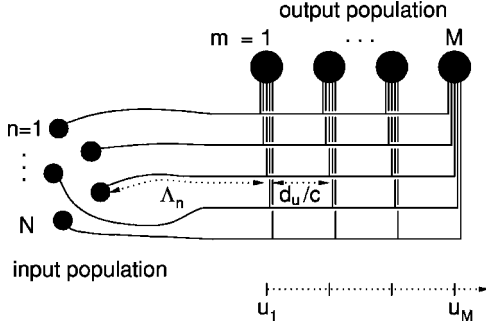
FIG. 4. Feed-forward topology. N input cells projected onto M output cells. The axonal conduction latencies Λ_n ($n=1,\ldots,N$) between the N input cells and the first output unit ($m=1$) are distributed over a temporal range that is at least as broad as the period T_p of the input process. As a constant axonal conduction velocity c is assumed, the spatial distance $u_m=(m-1)\,d_u$ between the mth and the first output neuron defines the total axonal delay $\Delta_{mn}=\Lambda_n+(m-1)d_u/c$ between input neuron n and output neuron m.

$$C^{\text{in}}_{nn'}(t',t)=\overline{p^{\text{in}}(t')p^{\text{in}}(t)}+\nu\,\delta_{nn'}\delta(t-t'), \qquad (7)$$

is of the form $C^{\text{in}}_{nn'}(t',t)=C^{\text{in}}_{nn'}(t'-t)$. This can be shown, if the T_p-periodic function g is represented by a Fourier series. Defining $\omega_p:=2\pi/T_p$, we write $g(t)=T_p^{-1}\Sigma_\mu\hat{g}_\mu\exp(i\mu\omega_p t)$ with Fourier coefficients $\hat{g}_\mu=\int_0^{T_p}d\tau\,g(\tau)\exp(-i\mu\omega_p\tau)$. Applying the Wiener-Khintchin theorem [38], we then find an expression for $\overline{p^{\text{in}}(t')p^{\text{in}}(t)}$ in Eq. (7) in terms of a Fourier series,

$$\overline{p^{\text{in}}(t')p^{\text{in}}(t)}=\nu^2\sum_{\mu=-\infty}^{\infty}|\hat{g}_\mu|^2\exp[2\pi i\mu(t-t')/T_p]. \quad (8)$$

As a consequence, the temporal average on the right-hand side of Eq. (7) is a function of $t-t'$ only, which considerably simplifies the following analysis.

2. Axonal topology

Within the laminar nucleus, axons run in parallel and contact postsynaptic cells. The network topology addressing this issue is sketched in Fig. 4, which will be the basis of the ensuing analysis. Each of the N axons is thought to contact all M output cells so that we can neglect the restriction $m'\in$ axon n in Eq. (6) and sum over all the postsynaptic cells.

In accordance with experimental data [5], the axonal conduction velocity c within the laminar nucleus is taken to be constant, cf. Fig. 4. The axonal delays Δ_{mn} are then unambiguously defined by

$$\Delta_{mn}=\Lambda_n+(m-1)d_u/c, \qquad (9)$$

where Λ_n denotes the axonal latency between the input neuron n and the first contacted output neuron ($m=1$). The spatial distance between output cells m and 1 is called $u_m=(m-1)d_u$.

3. Distance-dependent interaction

We further assume that the coupling matrix $b_{mm'}$ only depends on the spatial distance $d_u|m-m'|$ of the postsynaptic neurons [33–36]. In other words, the presynaptic interaction is described through a symmetric and translationally invariant linear operator.

E. Linear learning equation

To explicitly write down a dynamics for the synaptic weights, one has to specify expressions for the time-averaged firing rate $\nu_m^{\text{out}}(t)$ and the time-averaged correlation function $C_{mn}(t,t')$ in Eq. (5). Both depend on the synaptic weights that may change during learning. For Poisson neurons described by Eq. (2), we obtain the output rate

$$\nu_m^{\text{out}}(t)=\beta^{(0)}+\beta^{(1)}\nu\sum_{n=1}^{N}J_{mn}(t) \qquad (10)$$

and the correlation function [17]

$$C_{mn}(t',t)=\beta^{(0)}\nu+\beta^{(1)}\sum_{n'=1}^{N}J_{mn'}$$
$$\times\int_0^\infty ds\,\epsilon(s)C^{\text{in}}_{n'n}(t'-s-\Delta_{mn'},t). \quad (11)$$

We insert Eqs. (7), (10), and (11) as well as the learning equation (5) into Eq. (6) and obtain a linear dynamics for the synaptic weights,

$$\frac{d}{dt}J_{mn}=\sum_{m'}(\delta_{mm'}+\rho\,b_{mm'})$$
$$\times\left[k_1+\sum_{n'=1}^{N}(k_2+\delta_{nn'}k_3+Q_{nn'})J_{m'n'}\right]. \quad (12)$$

The constants k_1, k_2, and k_3 are defined to be

$$k_1=\eta\{\beta^{(0)}[w^{\text{out}}+\hat{W}(0)\nu]+w^{\text{in}}\nu\},$$

$$k_2=\eta\,\beta^{(1)}\nu[w^{\text{out}}+\hat{W}(0)\nu],$$

and

$$k_3=\eta\,\beta^{(1)}\nu\int ds\,W(s)\epsilon(-s), \qquad (13)$$

where $\hat{W}(\omega)=\int ds\,W(s)e^{-i\omega s}$ is the Fourier transform of the learning window. We will neglect the k_3 term in Eq. (12) later because it is of order N^{-1} compared to k_2 and $Q_{nn'}$, and N has been assumed to be large. In Eq. (12), the temporal structure of the input is hidden in

$$Q_{nn'}:=\eta\int ds\,W(s)q[s-(\Lambda_n-\Lambda_{n'})], \qquad (14)$$

where

$$q(r) := \beta^{(1)} \int ds' \, \epsilon(s') \overline{[p^{\text{in}}(t-s') - \nu][p^{\text{in}}(t+r) - \nu]}. \tag{15}$$

In the case of a periodic input density $p^{\text{in}}(t) = p \, g(t)$ as introduced in Sec. II D 1, applying Eq. (8) to Eqs. (15) and (14) leads to an expression of $Q_{nn'}$ in terms of a Fourier series,

$$Q_{nn'} = \sum_{\mu=-\infty}^{\infty} \hat{Q}_\mu \exp[i\mu\omega_p(\Lambda_n - \Lambda_{n'})] \tag{16}$$

with coefficients

$$\hat{Q}_\mu = \eta \, \beta^{(1)} \nu^2 |\hat{g}_\mu|^2 \hat{W}_\mu \hat{\epsilon}_\mu [1 - \delta_{\mu 0}]. \tag{17}$$

Here $\hat{W}_\mu$ and $\hat{\epsilon}_\mu$ denote the Fourier transforms of the learning window W and the EPSP function ϵ, taken at frequency $\omega = \mu \, \omega_p$, whereas $\hat{g}_\mu$ and $\hat{Q}_\mu$ are coefficients of a Fourier series. The linear driving force $\hat{Q}_\mu$ for temporal structure formation is the power spectrum $|\hat{g}_\mu|^2$ of the input process filtered through the learning window and the postsynaptic potential. Their temporal extent defines the bandpass performance of this filter. In Fig. 3 we show generic specimens of W and ϵ, as they have been used in [8], and their Fourier-transformed convolution $\hat{W} \, \hat{\epsilon}$. Furthermore, it should be noticed that $\hat{Q}_0$ vanishes. It will turn out that $Q_{nn'}$ has, therefore, no effect on the average synaptic weight, see Sec. IV C 1.

Typical values of the latencies Λ_n are assumed to be much smaller than the averaging time $\mathcal{T}$, which makes $Q_{nn'}$ dependent on a latency difference only, cf. Sec. II D 1. Equation (12) is fully specified so that we can proceed to analyzing its general structure and solution.

III. SEPARABILITY OF THE LEARNING EQUATION

Equation (12) is an autonomous linear differential equation for the set of coupling strengths J_{mn} ($1 \leq m \leq M, 1 \leq n \leq N$) and hence its solution can be given explicitly. In this section we show that the dynamics of the weights of a single neuron can be separated from the dynamics of map formation. Before doing so in Sec. IV we briefly discuss the general strategy for solving Eq. (12).

A. General structure and solution

Once the coupling vector $\mathbf{J}$, the inhomogeneity $\mathbf{j}$, and the linear operator $\mathcal{L}$ are identified, the differential equation (12) reads

$$\frac{d}{dt}\mathbf{J} = \mathcal{L}\mathbf{J} + \mathbf{j}. \tag{18}$$

The general solution of Eq. (18) is given by Duhamel's formula

$$\mathbf{J}(t) = \mathbf{J}^{\text{fix}} + e^{t\mathcal{L}}[\mathbf{J}(t=0) - \mathbf{J}^{\text{fix}}], \tag{19}$$

where fixed points ($d\mathbf{J}/dt = 0$) are defined by

$$\mathcal{L} \mathbf{J}^{\text{fix}} = -\mathbf{j} \tag{20}$$

and $e^{t\mathcal{L}} := \sum_{k=0}^{\infty} (t\mathcal{L})^k/k!$. The fixed point $\mathbf{J}^{\text{fix}}$ is unique if $\mathcal{L}$ is invertible. Otherwise Eq. (19) has to be extended by contributions from eigenspaces with eigenvalue 0 that grow polynomially with t.

The coupling vector $\mathbf{J}$ and the inhomogeneity $\mathbf{j}$ are both elements of an $(M \times N)$-dimensional linear space $\mathcal{R}$. The operator $\mathcal{L}$ is, therefore, an endomorphism over $\mathcal{R}$. We suspend a more thorough specification of $\mathcal{R}$ until Sec. III B since all equidimensional linear spaces are isomorphic. Let us suppose $\mathcal{L}$ is diagonalizable with eigenvalues λ and eigenvectors $\boldsymbol{\phi}_\lambda$, which constitute a basis of the vector space $\mathcal{R}$. Hence the deviation $\boldsymbol{\iota} := \mathbf{J}(t=0) - \mathbf{J}^{\text{fix}}$ of the initial couplings $\mathbf{J}(t=0)$ from the fixed point $\mathbf{J}^{\text{fix}}$ can be written as a linear combination $\boldsymbol{\iota} = \Sigma_\lambda a_\lambda \boldsymbol{\phi}_\lambda$ of these eigenfunctions. The time-dependent part $e^{t\mathcal{L}} \boldsymbol{\iota}$ of Eq. (19), therefore, reads $\Sigma_\lambda e^{t\lambda} a_\lambda \boldsymbol{\phi}_\lambda$. The above consideration can also be generalized to nondiagonalizable operators, but this will turn out to be dispensable for our example.

Summarizing, solving the fixed-point equation (20) and the eigenvalue problem,

$$\mathcal{L} \boldsymbol{\phi} = \lambda \boldsymbol{\phi} \tag{21}$$

is necessary and sufficient for analyzing the time course of the linear synaptic dynamics. An explicit solution for the special case of map formation in the barn owl's laminar nucleus will be given in Sec. IV. The remainder of this section shows how the dynamics of weights of a single neuron (also called the "temporal part") can be separated from the dynamics of map formation, i.e., the synchronization of all neurons' weight distributions across the array ("spatial part").

B. Spatiotemporal separability

We will show that both the linear operator $\mathcal{L}$ and the inhomogeneity $\mathbf{j}$ separate in space (m) and time (n). We, therefore, take the linear space $\mathcal{R}$ to be the direct-product space $\mathcal{R} = \mathbb{R}^M \otimes \mathbb{R}^N$ and specify an isomorphism $\mathcal{I}$ that is to be used for embedding the set of $M \times N$ synaptic weights J_{mn} into $\mathcal{R}$.

1. Embedding weights into tensor space

Let $\{\mathbf{e}_m ; 1 \leq m \leq M\}$ and $\{\mathbf{f}_n ; 1 \leq n \leq N\}$ be a basis of $\mathbb{R}^M$ and $\mathbb{R}^N$, respectively. Consequently $\{\mathbf{e}_m \otimes \mathbf{f}_n\}$ is a basis of $\mathcal{R}$. If we introduce scalar products so that

$$\mathbf{e}_m \cdot \mathbf{e}_{m'} = \delta_{mm'} \quad \text{and} \quad \mathbf{f}_n \cdot \mathbf{f}_{n'} = \delta_{nn'}, \tag{22}$$

we also obtain a scalar product on the tensor space $\mathcal{R}$ with the property that

$$(\mathbf{e}_m \otimes \mathbf{f}_n) \cdot (\mathbf{e}_{m'} \otimes \mathbf{f}_{n'}) = (\mathbf{e}_m \cdot \mathbf{e}_{m'})(\mathbf{f}_n \cdot \mathbf{f}_{n'}) = \delta_{mm'} \delta_{nn'}. \tag{23}$$

Our choice for the isomorphism $\mathcal{I}$ is now as simple as it can be,

$$\mathcal{I}: \ \mathbb{R}^{M \times N} \rightarrow \mathcal{R},$$

$$\mathcal{I}\{J_{mn}; 1 \le m \le M, 1 \le n \le N\} \mapsto \mathbf{J} = \sum_{m=1}^{M} \sum_{n=1}^{N} J_{mn} \mathbf{e}_m \otimes \mathbf{f}_n, \tag{24}$$

and thus

$$\mathcal{I}^{-1} \mathbf{J} = \{(\mathbf{e}_m \otimes \mathbf{f}_n) \cdot \mathbf{J}\} = \{J_{mn}\}. \tag{25}$$

The benefits of this transformation become clear after realizing that both the operator $\mathcal{L}$ and the inhomogeneity $\mathbf{j}$ are tensor products,

$$\mathcal{L} = \mathcal{L}^{\mathfrak{S}} \otimes \mathcal{L}^{\mathfrak{T}}, \tag{26}$$

$$\mathbf{j} = \mathbf{j}^{\mathfrak{S}} \otimes \mathbf{j}^{\mathfrak{T}}, \tag{27}$$

where the superscript indices $\mathfrak{S}$ and $\mathfrak{T}$ denote spatial and temporal components. The two parts of the linear operator read

$$\mathcal{L}^{\mathfrak{S}} \mathbf{e}_m = \sum_{m'} (\delta_{m'm} + \rho b_{m'm}) \mathbf{e}_{m'}, \tag{28}$$

$$\mathcal{L}^{\mathfrak{T}} \mathbf{f}_n = \sum_{n'} (k_2 + \delta_{n'n} k_3 + Q_{n'n}) \mathbf{f}_{n'}, \tag{29}$$

while the two parts of the inhomogeneity can be written as

$$\mathbf{j}^{\mathfrak{S}} = \mathcal{L}^{\mathfrak{S}} \mathbf{1}^{\mathfrak{S}}, \tag{30}$$

$$\mathbf{j}^{\mathfrak{T}} = k_1 \mathbf{1}^{\mathfrak{T}}, \tag{31}$$

with $\mathbf{1}^{\mathfrak{S}} = \Sigma_m \mathbf{e}_m$ and $\mathbf{1}^{\mathfrak{T}} = \Sigma_n \mathbf{f}_n$. Then the action of $\mathcal{L}$ onto the coupling vector $\mathbf{J}$ is calculated as

$$\mathcal{L} \mathbf{J} = \sum_{mn} J_{mn} (\mathcal{L}^{\mathfrak{S}} \mathbf{e}_m) \otimes (\mathcal{L}^{\mathfrak{T}} \mathbf{f}_n). \tag{32}$$

In this way we obtain a fully separable version of Eq. (18),

$$\frac{d}{dt} \mathbf{J} = \mathbf{j}^{\mathfrak{S}} \otimes \mathbf{j}^{\mathfrak{T}} + (\mathcal{L}^{\mathfrak{S}} \otimes \mathcal{L}^{\mathfrak{T}}) \mathbf{J}. \tag{33}$$

We note that although this differential equation is fully separable, its solution is generally not, i.e., $J_{mn} \neq J_m^{\mathfrak{S}} J_n^{\mathfrak{T}}$ does not factorize. Separable solutions occur if and only if the initial values $J_{mn}(t=0)$ are separable. The next two paragraphs demonstrate that a separable linear dynamics yields both a separable fixed point and separable eigensystems.

2. Separable fixed point

If we apply Eqs. (26) and (27), the fixed-point equation (20) yields

$$\sum_{mn} J_{mn}^{\text{fix}} (\mathcal{L}^{\mathfrak{S}} \mathbf{e}_m) \otimes (\mathcal{L}^{\mathfrak{T}} \mathbf{f}_n) = -\mathbf{j}^{\mathfrak{S}} \otimes \mathbf{j}^{\mathfrak{T}}. \tag{34}$$

The right-hand side of Eq. (34) is separable, as has to be the left-hand side. This requirement can only be fulfilled if also the fixed point $J_{mn}^{\text{fix}} = J_m^{\mathfrak{S}\text{fix}} J_n^{\mathfrak{T}\text{fix}}$ is separable. We then can read the spatial and temporal part separately so that we obtain two independent fixed-point equations. With Eq. (30) we find a spatial fixed-point equation

$$\mathcal{L}^{\mathfrak{S}} \mathbf{J}^{\mathfrak{S}\,\text{fix}} = \mathcal{L}^{\mathfrak{S}} \mathbf{1}^{\mathfrak{S}} \tag{35}$$

and with Eq. (31) a temporal one

$$\mathcal{L}^{\mathfrak{T}} \mathbf{J}^{\mathfrak{T}\,\text{fix}} = -k_1 \mathbf{1}^{\mathfrak{T}}, \tag{36}$$

where $\mathbf{J}^{\mathfrak{S}\,\text{fix}} = \Sigma_m J_m^{\mathfrak{S}\,\text{fix}} \mathbf{e}_m$ and $\mathbf{J}^{\mathfrak{T}\,\text{fix}} = \Sigma_n J_n^{\mathfrak{T}\,\text{fix}} \mathbf{f}_n$. These equations are unequivocal besides a constant multiplicative factor $\alpha \neq 0$. Multiplying $\mathbf{J}^{\mathfrak{T}\,\text{fix}}$ by α leads to a multiplication of $\mathbf{J}^{\mathfrak{S}\,\text{fix}}$ by α^{-1}. The position of the minus sign in Eqs. (35) and (36) is, therefore, arbitrary.

3. Separable eigensystems

We now focus on the eigenvalue problem (21) for the operator $\mathcal{L} = \mathcal{L}^{\mathfrak{S}} \otimes \mathcal{L}^{\mathfrak{T}}$ from Eq. (26). If we insert a separable ansatz for the eigenfunctions $\boldsymbol{\phi} = \boldsymbol{\phi}^{\mathfrak{S}} \otimes \boldsymbol{\phi}^{\mathfrak{T}}$ we obtain

$$\mathcal{L}^{\mathfrak{S}} \boldsymbol{\phi}^{\mathfrak{S}} \otimes \mathcal{L}^{\mathfrak{T}} \boldsymbol{\phi}^{\mathfrak{T}} = \lambda \boldsymbol{\phi}^{\mathfrak{S}} \otimes \boldsymbol{\phi}^{\mathfrak{T}}. \tag{37}$$

Again, separating coordinates, we find independent spatial and temporal eigenvalue problems, viz.,

$$\mathcal{L}^{\mathfrak{S}} \boldsymbol{\phi}^{\mathfrak{S}} = \lambda^{\mathfrak{S}} \boldsymbol{\phi}^{\mathfrak{S}}, \tag{38}$$

$$\mathcal{L}^{\mathfrak{T}} \boldsymbol{\phi}^{\mathfrak{T}} = \lambda^{\mathfrak{T}} \boldsymbol{\phi}^{\mathfrak{T}}, \tag{39}$$

with the total eigenvalue $\lambda = \lambda^{\mathfrak{S}} \lambda^{\mathfrak{T}}$. Let us suppose the eigenvalue problems (38) and (39) have been solved. Then the ansatz $\boldsymbol{\phi} = \boldsymbol{\phi}^{\mathfrak{S}} \otimes \boldsymbol{\phi}^{\mathfrak{T}}$ yields $M \times N$ eigenfunctions and hence we have found the complete eigensystem of Eq. (21).

IV. SPATIOTEMPORAL DELAY SELECTION

We separately solve the "spatial" and "temporal" parts of the differential equation (33) as outlined in the preceding section and discuss the biological relevance of the solutions.

A. Spatial solution

The solution of the spatial part of Eq. (33) is simple because the linear operator $\mathcal{L}^{\mathfrak{S}}$ as given by Eq. (28) is a cyclic, or circulant, matrix. We assume $|\rho| \ll 1$, cf. Sec. VI. Therefore $\mathcal{L}^{\mathfrak{S}}$ is invertible and has the unique fixed point in Eq. (35),

$$\mathbf{J}^{\mathfrak{S}\,\text{fix}} = \mathbf{1}^{\mathfrak{S}}. \tag{40}$$

With equally spaced output cells and a translationally invariant axonal coupling (see Sec. II D 3), $\mathcal{L}^{\mathfrak{S}}$ is translationally invariant as well. Translationally invariant, i.e., cyclic, operators are very common in quantum mechanics and solid state physics and their spectral theory is completely understood [39]. The eigenfunctions $\boldsymbol{\phi}^{\mathfrak{S}}(l)$ of $\mathcal{L}^{\mathfrak{S}}$ are plane waves with wavelength $l \in \mathbb{Z}$ so that

$$\boldsymbol{\phi}^{\mathfrak{S}}(l) := \sum_{m=1}^{M} e^{2\pi i m l/M} \mathbf{e}_m \tag{41}$$

for $1 \leq l \leq M$. The eigenvalues are

$$\lambda^{\mathfrak{S}}(l) = 1 + \rho \tilde{b}_l, \tag{42}$$

where

$$\tilde{b}_l = \sum_{m=-M/2+1}^{M/2} b_{0m} e^{-(2\pi i/M)ml} \tag{43}$$

is the spatial Fourier transform of the coupling matrix. This only holds exactly for $M \to \infty$ or periodic boundary conditions, but as shown by Ledermann [40], for finite M the alignment of the spectrum of $\mathcal{L}^{\mathfrak{S}}$ remains unchanged. Since we have assumed symmetric axonal coupling, $\lambda^{\mathfrak{S}}(l)$ is simply a real number. We also note that the scalar product of the eigenvectors is $\boldsymbol{\phi}^{\mathfrak{S}}(l) \cdot \boldsymbol{\phi}^{\mathfrak{S}}(l') = M\delta_{l,-l'}$, cf. Eq. (22).

At this stage, two important aspects ought to be stressed. First, in order to guarantee map formation, i.e., synchronization of synaptic structures along the neuron array, we have to ensure that all synapses at a specific axon n either grow or decay, cf. Figs. 1 and 4. This means that synaptic modifications are to be associated with the axon index n rather than with the postsynaptic index m. The eigenvector $\boldsymbol{\phi}^{\mathfrak{S}}(0) = \Sigma_m \mathbf{e}_m$ is made up of equal contributions from all postsynaptic cells $1 \leq m \leq M$, and, thus, represents a homogenous weight change along the axons. All other eigenvectors lead to combined strengthening and weakening of the synaptic strengths at one axon and, hence, disturb the formation of an ITD map. Second, if we interpret an eigenvalue as the velocity of growth of an eigenvector, Eq. (42) tells us that the eigenvector $\boldsymbol{\phi}^{\mathfrak{S}}(0)$ exhibits significantly faster growth than eigenvectors with $l \neq 0$, if $\rho \tilde{b}_0 \gtrsim 1$ and $|\rho \tilde{b}_l| \ll 1$ for $l \neq 0$. In our example of the laminar nucleus of the barn owl we will argue that $\tilde{b}_0 \approx M$, and, thus, $\rho \approx 1/M$ is already large enough for map formation.

B. Temporal solution

Handling $\mathcal{L}^{\mathfrak{T}}$ as defined in Eq. (29) is, in general, impossible unless we specify the distribution of the latencies Λ_n. We assume a uniform distribution of latencies in an interval of length T_p,

$$\Lambda_n = n T_p/N, \tag{44}$$

which does not favor any phase *a priori*. Generality is not restricted by introducing delays that only cover an interval of length T_p, because the dynamical equations are strictly periodic, cf. Fig. 5.

The definition of $\mathcal{L}^{\mathfrak{T}}$ in Eq. (29) contains the matrix $Q_{nn'}$, which acts as the driving force for structure formation as defined by Eqs. (14)–(17). We insert the latency distribution (44) into Eq. (16) and obtain

$$Q_{nn'} = \sum_{\mu=-\infty}^{\infty} \hat{Q}_\mu \exp[2\pi i \mu(n-n')/N]. \tag{45}$$

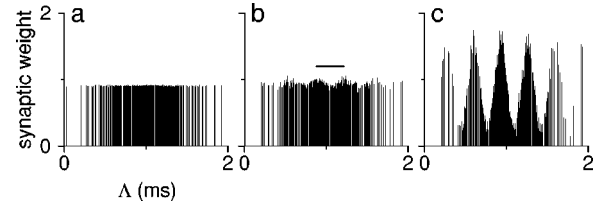FIG. 5. Structure formation at a single postsynaptic neuron. $N = 600$ synapses with Gaussian distributed latencies Λ (mean 1 ms, width 0.3 ms) are contacting output cell m. (a) Initially, the weights J_{mn} ($1 \leq n \leq N$) are equally distributed with mean 0.9 and standard deviation 0.01. The prominent contribution is yielded by the temporal eigenvector $\boldsymbol{\phi}^{\mathfrak{T}}(0)$ that describes the average weight, whereas all other eigenspaces participate upto a small portion. (b) After 15 s of learning, the prominent eigenspace [$T_p = 1/(3 \text{ kHz})$, horizontal bar] is already noticeable. (c) Temporal synaptic structure after 50 s. Parameters are taken from [8].

Because $Q_{nn'}$ only depends on the difference $n-n'$, the operator $\mathcal{L}^{\mathfrak{T}}$ is cyclic. The next steps consist of examining the eigensystems ($\boldsymbol{\phi}^{\mathfrak{T}}, \lambda^{\mathfrak{T}}$) of $\mathcal{L}^{\mathfrak{T}}$ and calculating the fixed-point solution of Eq. (36).

1. Temporal eigensystems

The eigenvectors of $\mathcal{L}^{\mathfrak{T}}$ are plane waves,

$$\boldsymbol{\phi}^{\mathfrak{T}}(\mu) = \sum_{n=1}^{N} \mathbf{f}_n e^{2\pi i \mu n/N} \quad \text{for} \quad 0 \leq \mu \leq N-1. \tag{46}$$

The vectors $\boldsymbol{\phi}^{\mathfrak{T}}(\mu)$ are normalized in such a way that $\boldsymbol{\phi}^{\mathfrak{T}}(\mu) \cdot \boldsymbol{\phi}^{\mathfrak{T}}(\mu') = N\delta^N_{\mu,-\mu'}$, where we define a modified Kronecker delta $\delta^N_{\mu\mu'}$ to equal 1 for $\mu = \mu' + \mathbb{Z}N$ and 0 otherwise. The eigenvectors $\boldsymbol{\phi}^{\mathfrak{T}}$ can be defined for all $\mu \in \mathbb{Z}$ since $\boldsymbol{\phi}^{\mathfrak{T}}(\mu) = \boldsymbol{\phi}^{\mathfrak{T}}(\mu+N)$ and $\boldsymbol{\phi}^{\mathfrak{T}}(-\mu) = \boldsymbol{\phi}^{\mathfrak{T}}(N-\mu)$.

To calculate the eigenvalues $\lambda^{\mathfrak{T}}(\mu)$ of eigenvectors $\boldsymbol{\phi}^{\mathfrak{T}}(\mu)$, we exploit $\Sigma_n \exp[2\pi i(\mu-\mu')n/N] = N\delta^N_{\mu\mu'}$, apply Eqs. (29), (45), and (46) to the eigenvalue problem (39), and obtain

$$\lambda^{\mathfrak{T}}(\mu) = N(k_2 \delta_{\mu,0} + k_3/N + \mathcal{Q}_\mu), \tag{47}$$

where $\mathcal{Q}_\mu := \Sigma_{\mu'} \hat{Q}_{\mu'} \delta^N_{\mu,\mu'}$. From now on, k_3/N will be neglected, since N has been assumed to be large, see Sec. II D 1.

2. Temporal fixed-point solution

To calculate the temporal fixed point, we take a closer look at Eq. (36). We already know from Eq. (46) that $\boldsymbol{\phi}^{\mathfrak{T}}(0) = \mathbf{1}^{\mathfrak{T}}$. Therefore $\mathbf{1}^{\mathfrak{T}}$ on the right-hand side of Eq. (36) is an eigenvector of $\mathcal{L}^{\mathfrak{T}}$ with eigenvalue $\lambda^{\mathfrak{T}}(0) = N(k_2 + \mathcal{Q}_0)$, cf. Eq. (47). Then the temporal fixed point is

$$\mathbf{J}^{\mathfrak{T} \text{ fix}} = \frac{-k_1}{N(k_2 + \mathcal{Q}_0)} \mathbf{1}^{\mathfrak{T}}. \tag{48}$$

For a biologically reasonable choice of W and ϵ with exponentially decaying slopes as in Fig. 3, we can set $\mathcal{Q}_0 = 0$ because $\hat{Q}_0 = 0$ [cf. Eq. (17)] and the coefficients $\hat{Q}_\mu$ decay

at a polynomial rate of at least μ^{-2} and thus $\hat{Q}_N = O(N^{-2})$. With parameters taken from [8] we even find $\hat{Q}_N = O(N^{-3})$

A prerequisite for structure formation is that synaptic weights should neither all decay to zero nor increase to infinity, see also Fig. 1. Therefore stabilization of the zeroth eigenvector $\boldsymbol{\phi}^{\mathfrak{T}}(0)$ due to a negative eigenvalue $\lambda^{\mathfrak{T}}(0)$ is necessary for structure formation. We will argue that the growth of a synaptic structure is possible, if a few eigenvectors with $\mu \neq 0$ have large positive eigenvalues. The following section shows how both constraints can be achieved.

C. Synthesis of spatial and temporal solution

The discussion of the general solution (19) of an autonomous linear differential equation, e.g., Eq. (12), can be reduced to an analysis of the prominent directions of temporal evolution, viz., the eigenspaces with largest eigenvalues, if all the other eigenvalues are significantly smaller [24,41,42]. From Eqs. (41) and (46) we find the product of spatial and temporal eigenfunctions to be plane waves,

$$\boldsymbol{\phi}(l,\mu) := \boldsymbol{\phi}^{\mathfrak{G}}(l) \otimes \boldsymbol{\phi}^{\mathfrak{T}}(\mu)$$

$$= \sum_{mn} \mathbf{e}_m \otimes \mathbf{f}_n \exp[2\pi i (ml/M + n\mu/N)] \quad (49)$$

with eigenvalues given by Eqs. (42) and (47),

$$\lambda(l,\mu) := \lambda^{\mathfrak{G}}(l)\lambda^{\mathfrak{T}}(\mu) = (1 + \rho \tilde{b}_l) N(k_2 \delta_{\mu 0} + \mathcal{Q}_\mu). \quad (50)$$

Let us first discuss the evolution of the average synaptic weight ($l = \mu = 0$).

1. Stabilization of the average weight

The average synaptic weight is defined as

$$J^{\text{av}} := (NM)^{-1} \sum_{mn} J_{mn} = (NM)^{-1} \mathbf{J} \cdot (\mathbf{1}^{\mathfrak{G}} \otimes \mathbf{1}^{\mathfrak{T}}), \quad (51)$$

where the second equality follows from Eqs. (23) and (24). Requiring the average weight to be stable and positive, we obtain conditions for the parameters k_1 and k_2 from Eq. (13). Equation (49) shows that $\mathbf{1}^{\mathfrak{G}} \otimes \mathbf{1}^{\mathfrak{T}}$ in Eq. (51) equals $\boldsymbol{\phi}^{\mathfrak{G}}(0) \otimes \boldsymbol{\phi}^{\mathfrak{T}}(0)$. The temporal evolution of J^{av} is, therefore, determined by $\lambda(0,0) = (1 + \rho \tilde{b}_0) N k_2$, cf. Eq. (50). If the learning dynamics is required to stabilize J^{av} to a finite value, then $\lambda(0,0)$ has to be negative. $\lambda^{\mathfrak{G}}(0) = 1 + \rho \tilde{b}_0$ is positive because the elements of the axonal coupling matrix $b_{mm'}$ in Eq. (6) have been defined to be non-negative. This means $\lambda^{\mathfrak{T}}(0) = N k_2$ must be *negative*. For a negative k_2 [see Eq. (13)], either the integral over the learning window $\hat{W}(0)$ or the postsynaptic contribution w^{out}, or both, have to be negative.

If we combine Eqs. (40) and (48), we find the overall fixed point

$$\mathbf{J}^{\text{fix}} = J^{\text{fix}} \mathbf{1}^{\mathfrak{G}} \otimes \mathbf{1}^{\mathfrak{T}}, \quad (52)$$

with the average weight at the fixed point

$$J^{\text{fix}} := -k_1/(N k_2). \quad (53)$$

The fixed point $\mathbf{J}^{\text{fix}}$ is a multiple of the eigenfunction $\boldsymbol{\phi}^{\mathfrak{G}}(0) \otimes \boldsymbol{\phi}^{\mathfrak{T}}(0)$ in Eq. (49), which is a special feature of the present dynamics. A negative eigenvalue $\lambda(0,0)$ then means that J^{av} is asymptotically governed by the fixed point, viz., $\lim_{t \to \infty} J^{\text{av}} = J^{\text{fix}}$. Therefore J^{fix} has to be positive so as to avoid vanishing of synaptic input. Equation (53) then implies that k_1 is restricted to *positive* values. For a positive k_1 and a negative k_2 in Eq. (13), the presynaptic contribution w^{in} must then be positive.

2. Structure formation

Synaptic strengths in the immature animal are assumed to be uniformly distributed around a multiple of $\mathbf{J}^{\text{fix}}$ in Eq. (52); compare Fig. 5. This initial distribution strongly contributes to the eigenspace $\boldsymbol{\phi}(0,0)$ with the negative real part of the eigenvalue, whereas all contributions to eigenspaces with the positive real part of the eigenvalue, i.e., all projections $\boldsymbol{\phi}(-\mu,-l) \cdot \mathbf{J}^{\text{fix}}$ are small, so that none of them has a head start. The eigenspaces with the largest eigenvalues thus dominate the dynamics. We are going to calculate its consequences for the time course of the membrane potential and predict the capability of temporal processing in our model network.

As we have seen in the preceding section, temporal eigenvalues with positive real parts are restricted to $\mu = 1,2,\ldots,N-1$. For a sufficiently large interaction width along axons we find spatial eigenvalues $\lambda^{\mathfrak{G}}(0) \gg \lambda^{\mathfrak{G}}(l \neq 0) = 1 + \rho \tilde{b}_l$, cf. Eq. (42). If, for instance, the interaction does not depend on the spatial distance between synapses, $b_{mm'} = 1$, we find $\lambda^{\mathfrak{G}}(l) = 1 + \rho M \delta_{l0}$. The relevant eigenvalues for structure formation are thus $\lambda(0,\mu) = (1 + \rho M) \mathcal{Q}_\mu$, cf. Eq. (50). The leading eigenfunction $\boldsymbol{\phi}(0,\mu)$ can now be interpreted as a synchronization of the phases of latency eigenfunctions $\boldsymbol{\phi}^{\mathfrak{T}}$ among the M output neurons. We have, therefore, arrived at a *selection of axons* rather than a selection of synapses.

Taking into account the bandpass property of our learning dynamics, as mentioned in the discussion of Eq. (16) (see also Fig. 3), it is likely that some $\mu = \mu^{\text{opt}}$ dominates the temporal eigenvalues. For sufficiently large input frequencies $\omega_p := 2\pi/T_p$, the optimal harmonic is the first one, $\mu^{\text{opt}} = 1$. Hence we expect the emerging structure to reflect the plane wave $\boldsymbol{\phi}(0,\pm 1)$.

Traveling excitation waves. What does the prominent eigenvector contribute to the postsynaptic potential? The answer is a plane-wave-like postsynaptic excitation traveling along the output array. As for the proof, we calculate the input-averaged ($\langle \cdots \rangle$) membrane potential (3) resulting from the synaptic efficacies $\mathbf{J} = \mathbf{J}^{\text{fix}} + a\,\boldsymbol{\phi}(0,\pm 1)$,

$$\langle v_m \rangle(t) = N\nu[J^{\text{fix}} + a\hat{g}_{\pm 1} \hat{\boldsymbol{\epsilon}}_{\pm 1} \exp\{\pm i\omega_p(t - c^{-1}u_m)\}] =: v^{(0)}$$

$$+ v^{(1)} \exp[\pm i\omega_p(t - c^{-1}u_m)] \quad (54)$$

with constants $v^{(0)}$ and $v^{(1)}$. The spatial wavelength, which is the difference between two isophase lines, is, therefore, given by $2\pi c \omega_p^{-1}$.

3. Independent input populations

Up to now, we have explored synaptic structure formation with only one set of independent input processes $p^{\text{in}}(t)$. In the case of the laminar nucleus, however, there are two sets of input populations representing the activity of the afferent fibers originating from *both* ears. Due to random variations of the spatial location of sound sources, these subpopulations are assumed to be independent, see below.

To deduce the synaptic dynamics of the laminar nucleus (that is driven by two independent input populations) from synaptic structure formation for one input population, as discussed above, we show that X independent input populations, denoted by x, with $1 \leq x \leq X$, develop independent synaptic structures.

Two sets of input cells are defined to be independent, if their spike trains are temporally uncorrelated on the time scale $\mathcal{T}$. To be concrete, we require that, in extension to Eq. (7), the time-averaged presynaptic correlation function reads

$$C^{\text{in}}_{x'x,n'n}(t',t) = \nu^2 + \delta_{x'x}[\overline{p^{\text{in}}_x(t')p^{\text{in}}_x(t)} - \nu^2]$$
$$+ \nu \delta_{x'x}\delta_{n'n}\delta(t'-t). \quad (55)$$

This kind of correlation can be achieved, for example, by random phase changes in periodically firing populations with equal rate ν. The periods $T_{p,x} = 2\pi/\omega_{p,x}$ of the subpopulations do not have to be identical but must be significantly shorter then the averaging time $\mathcal{T}$. The typical time difference between two phase changes, however, has to be of the same order of magnitude as $\mathcal{T}$, so as to not destroy the temporal correlations within one population.

We are now dealing with X input populations. The weights $J_{mn,x}$ of input line n, therefore, obtain a population index x and, in analogy to Eq. (14), we define $Q_{nn',x}$ by using the attributes of input population x. These substitutions again yield a spatiotemporally separable learning equation,

$$\frac{d}{dt}J_{mn,x} = \sum_{m' \in \text{axon}} (\delta_{mm'} + \rho b_{mm'})$$
$$\times \left[k_1 + \sum_{x'n'} J_{m'n',x'}\{k_2 + \delta_{xx'}(k_3\delta_{nn'} + Q_{nn',x})\} \right]. \quad (56)$$

The sum over x' in Eq. (56) is obtained since all populations contribute to the postsynaptic potential.

The spatial component of the inhomogeneity and linear operator remain unchanged when compared to the equivalent equation for only one input population (12), whereas the temporal coordinates are embedded into $\oplus^X_{x=1} \mathbb{R}^{N_x}$ with N_x input lines in population x. Then the temporal part of the linear operator reads

$$\mathcal{L}^{\mathcal{T}}\mathbf{f}_{n,x} = \sum_{n'x'} \mathbf{f}_{n',x'}[k_2 + \delta_{x'x}(k_3\delta_{n'n} + Q_{n'n,x})].$$

Since only the k_2 term mixes up different population indices one may suspect already at this stage, and in analogy to Sec. IV C 1, that different populations do not interfere in structure formation.

Structure formation. All eigenspaces $\boldsymbol{\phi}(l,\mu)$, with $\mu \neq 0$ remain unchanged as compared to those of the dynamics with only one population. The ansatz

$$\boldsymbol{\phi}(l,\mu,x) := \sum_{mn} \mathbf{e}_m \otimes \mathbf{f}_{n,x}\exp[2\pi i(lm/M + \mu n/N)] \quad (57)$$

yields the eigenvalues $\lambda(l,\mu,x) = (1 + \rho\tilde{b}_l)(N_x\mathcal{Q}_{\mu,x})$. In other words, each population x contributes to its own structure with its own driving force $\mathcal{Q}_{\mu,x}$ weighted with the respective number of input cells N_x.

Fixed point and normalization. In analogy to Eq. (48), the temporal part of the inhomogeneity $\mathbf{j}^{\mathcal{T}} = -k_1\Sigma^X_{x=1}\mathbf{1}^{\mathcal{T}}_x$ is an eigenvector of $\mathcal{L}^{\mathcal{T}}$ and, with $\mathcal{Q}_{0,x} = 0$ (cf. Sec. IV B 2), we find the fixed point to be

$$\mathbf{J}^{\text{fix}} = \frac{-k_1}{k_2\displaystyle\sum_{x=1} N_x}\mathbf{1}^{\mathfrak{S}} \otimes \sum_{x=1}^{X} \mathbf{1}^{\mathcal{T}}_x. \quad (58)$$

Its eigenvalue is $\lambda(0,0,0) = (1 + \rho\tilde{b}_0)(k_2\Sigma_{x'}N_{x'})$.

In contrast to the single population case, $\mathcal{L}^{\mathcal{T}}$ has $X-1$ eigenvectors $\mathbf{Y} = \Sigma_{xn}Y_x\mathbf{f}_{n,x}$ with eigenvalue 0, defined by the condition $\Sigma_x N_x Y_x = 0$. Since $\mathbf{1}^{\mathcal{T}}_x \cdot \mathbf{1}^{\mathcal{T}}_x = N_x$, the vectors $\mathbf{Y}$ are orthogonal to $\mathbf{j}^{\mathcal{T}}$ and, hence, do not influence the dynamics.

4. Standing waves and temporal feature maps

To see how a temporal-feature map can arise, we return to the example of a map of interaural time differences (ITDs) in the barn owl's laminar nucleus [8,43]. Two sets of input lines that correspond to the two ears are assumed to be independent as defined in Sec. IV C 3 but otherwise identical. The two periodic input processes generate spike trains that are phase locked to the acoustic stimulation of either of the two ears. Auditory processing in the laminar nucleus is subdivided into isofrequency layers, where the reciprocal of this frequency corresponds to the considered periodicity $T_p = 2\pi/\omega_p$. The axons from both ears run along the array of M output cells in opposite directions, interdigitate and, in doing so, contact the laminar neurons, see also Fig. 1.

After independent structure formation in both populations, the system faces two contrarily traveling excitation waves, $v^{\text{left}}_m(t)$ and $v^{\text{right}}_m(t)$ with $c^{\text{left}} = -c^{\text{right}}$, cf. Eq. (54). Their linear superposition leads to a standing wave phenomenon,

$$\langle v^{\text{left}}_m(t) + v^{\text{right}}_m(t)\rangle$$
$$= 2v^{(0)} + 2v^{(1)}\cos(\omega_p u_m/c + \mathbf{\Phi})\cos(\omega_p t). \quad (59)$$

The places u^{max} of the interference maxima are then defined by the phase offset $\mathbf{\Phi}$ between both classes of input pro-
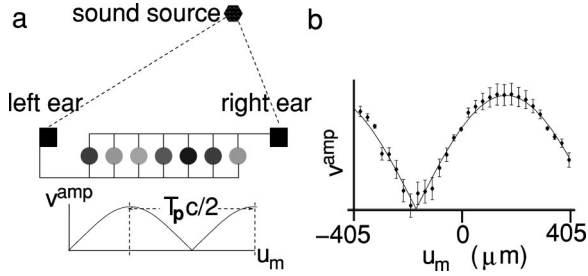
FIG. 6. Map of interaural time difference (ITD). (a) Schematic drawing of the coincidence detector array (gray shaded disks) as proposed by Jeffress [44]. Acoustic stimulation is delayed by (i) the spatial distance between the sound source and the ears (dashed lines) and (ii) the physiological transduction from the ear to the cell (solid lines). After delay tuning, the theoretically predicted average membrane potential $\langle v(t) \rangle$ is a standing wave across the laminar nucleus with period length $T_p c/2$. Its phase offset is determined by the interaural time difference. The azimuthal position of the sound source in auditory space is thus mapped onto the position u_m of maximal amplitude of the membrane potential v^{amp} within the laminar nucleus (the brighter the disk, the higher the membrane potential amplitude). (b) For comparison, we show the simulated amplitude of the membrane potential (dots with error bars) and the theoretical prediction [solid line, see Eq. (59)] for an ITD of 50 μs and synaptic weights after 875 s of formal learning. The interaction strength in the simulation is $\rho = 0.7/30$; the other parameter values have been taken from [8]. Each data point is obtained from a best fit of the average time course of the neurons' membrane potentials to Eq. (59).

cesses: $u^{\mathrm{max}} = \mathbb{Z} c T_p/2 - c\Phi/\omega_p$. We have thus obtained a *place code* representing the time difference Φ/ω_p between both classes. The time difference is determined by the hardwired conduction delay between the auditory organ and the laminar cell, and by the azimuthal position of the sound source, i.e., the ITD; cf. Fig. 6 and [44]. Since the conduction delay is an anatomic constant for each neuron, we have actually obtained a representation of the ITD in terms of the positions of interference maxima in the laminar nucleus.

These interference maxima cannot be translated into firing rate maxima with linear neurons, since the mean output rate $\nu_m^{\mathrm{out}} = \beta^{(0)} + \beta^{(1)} \overline{\langle v^{\mathrm{left}} + v^{\mathrm{right}} \rangle} = \beta^{(0)} + \beta^{(1)} 2v^{(0)}$ is independent of m. In Sec. VII, however, we show that nonlinear neurons lead to an identical synaptic structure and, therefore, to identical standing waves of the membrane potential. Because of the nonlinearity, these standing waves are then transformed into firing rates, i.e., an ITD map.

V. ORDER PARAMETERS

We introduce two order parameters serving as quality measures for the synaptic development in a single population of input lines.

A. Local and global order

The "local" order parameter is to describe the average extent of delay selection at the single cells, and the "global" one will indicate the synchronization of selected delays be-

tween the cells. We deduce the order parameters from a quantity called vector strength. The vector strength of a positive, T_p-periodic function is defined [45] to be the amplitude of its first Fourier component divided by the Fourier component of order zero. This yields a quantification of the development of the leading eigenfunction $\phi(0, \pm 1)$.

The vector strength of weights J_{mn} of neuron m is $|\Sigma_n e^{-i\omega_p \Lambda_n} J_{mn}/\Sigma_n J_{mn}|$. Then the *average* vector strength of M cells is defined to be

$$V^{\mathrm{avg}} = \left(\frac{1}{M} \sum_{m=1}^{M} \left| \frac{\sum_n e^{-i\omega_p \Lambda_n} J_{mn}}{\sum_n J_{mn}} \right|^2 \right)^{1/2}, \quad (60)$$

which indicates the systems' local order. We define the "axonal weight" as the sum of all weights of some axon, e.g., $\Sigma_m J_{mn}$ for axon n. The *axonal* vector strength then is the vector strength of the "axonal weights,"

$$V^{\mathrm{axon}} = \left| \frac{\sum_{mn} e^{-i\omega_p \Lambda_n} J_{mn}}{\sum_{mn} J_{mn}} \right|. \quad (61)$$

It is as a measure of the systems' global order.

B. Dynamics

With the results from Sec. IV it is possible to analytically specify the time course of the above order parameters. The sums in Eqs. (60) and (61) can be identified with projections of the weight vector on eigenspaces of the linear operator $\mathcal{L}^{\mathfrak{T}}$. Due to Eq. (46) and $J_{mn} = (\mathbf{e}_m \otimes \mathbf{f}_n) \cdot \mathbf{J}$ from Eq. (25), we find

$$\sum_n e^{-i\mu\omega_p \Lambda_n} J_{mn} = [\mathbf{e}_m \otimes \boldsymbol{\phi}^{\mathfrak{T}}(-\mu)] \cdot \mathbf{J}. \quad (62)$$

Exploiting the fact that the eigenvectors $\boldsymbol{\phi}^{\mathfrak{T}}(\mu)$ in Eq. (46) and $\phi^{\mathfrak{S}}(l)$ in Eq. (41) are a basis of $\mathbb{R}^N$ and $\mathbb{R}^M$, respectively, we can write $\mathbf{J}$ as a linear combination of eigenvectors $\boldsymbol{\phi}(l, \mu)$ of the operator $\mathcal{L}$,

$$\mathbf{J}(t) = \sum_{l,\mu} a_{l\mu}(t) \boldsymbol{\phi}(l, \mu),$$

where the time course of the coefficients

$$a_{l\mu}(t) = \frac{1}{MN} \boldsymbol{\phi}(-l, -\mu) \cdot \mathbf{J}(t)$$

$$= \frac{1}{MN} \sum_{mn} J_{mn}(t) \exp[-2\pi i(lm/M + \mu n/N)]$$

$$(63)$$

is governed by the differential equation

$$\frac{d}{dt}a_{l\mu}=\frac{1}{MN}\phi(-l,-\mu)\cdot\frac{d}{dt}\mathbf{J}$$

$$=\frac{1}{MN}\phi(-l,-\mu)\cdot(\mathcal{L}\mathbf{J}+\mathbf{j})$$

$$=a_{l\mu}\lambda(l,\mu)-\delta_{l0}\delta_{\mu0}\lambda(0,0)J^{\text{fix}},\qquad(64)$$

where we have used $\mathbf{j}=-\mathcal{L}\mathbf{J}^{\text{fix}}=\lambda(0,0)J^{\text{fix}}\phi(0,0)$, cf. Eq. (52). As a result, the $a_{l\mu}$ evolve independently of each other,

$$a_{00}(t)=J^{\text{fix}}+e^{\lambda(0,0)t}[a_{00}(0)-J^{\text{fix}}],$$

$$a_{l\mu}(t)=e^{\lambda(l,\mu)t}a_{l\mu}(0)\quad\text{for }(l,\mu)\neq(0,0).\qquad(65)$$

The projection defined in Eq. (62) then yields

$$[\mathbf{e}_m\otimes\phi^{\mathfrak{T}}(-\mu)]\cdot\mathbf{J}=N\sum_l e^{2\pi ilm/M}a_{l\mu}.\qquad(66)$$

If we assume the temporal evolution of the average weight to be much faster than structure formation or, in other words, the temporal eigenvalues (47) comply with $\lambda^{\mathfrak{T}}(0)<0$ and $|\lambda^{\mathfrak{T}}(0)|\gg\lambda^{\mathfrak{T}}(\mu^{\text{opt}})$, then we are allowed to set the average synaptic weight (51) equal to the fixed point (52). That is to say, in terms of Eq. (65), we substitute $a_{l0}=\delta_{l0}J^{\text{fix}}$. Consequently, the denominators of Eqs. (60) and (61) are independent of m since, according to Eqs. (62) and (66),

$$\sum_n J_{mn}=N\sum_l e^{2\pi ilm/M}a_{l0}=Na_{00}=NJ^{\text{fix}}.\qquad(67)$$

The numerator in Eq. (60) can be calculated similarly if we also use $\sum_m e^{2\pi ilm/M}=M\delta_{l0}$,

$$\sum_m|[\mathbf{e}_m\otimes\phi^{\mathfrak{T}}(-\mu)]\cdot\mathbf{J}|^2=M\sum_l|Na_{l\mu}|^2,$$

whereas the numerator in Eq. (61) is obtained directly as

$$\left|\sum_m[\mathbf{e}_m\otimes\phi^{\mathfrak{T}}(-\mu)]\cdot\mathbf{J}\right|=|\phi(0,-\mu)\cdot\mathbf{J}|=|MNa_{0\mu}(t)|.$$

Summarizing the above calculations we find

$$V^{\text{avg}}(t)=\left[\sum_l|a_{l1}(0)e^{\lambda(l,1)t}|^2\right]^{1/2}\Big/J^{\text{fix}}$$

and

$$V^{\text{axon}}(t)=|a_{01}(0)e^{\lambda(0,1)t}|/J^{\text{fix}}.$$

For the sake of simplicity we assume the axonal coupling matrix (6) to be $b_{mm'}=1$ for all m,m'. Due to Eq. (42) this yields eigenvalues $\lambda(l,1)=\lambda^{\mathfrak{T}}(1)(1+\rho M\delta_{l0})$ and, with $\lambda^{\mathfrak{T}}=\text{Re}[\lambda^{\mathfrak{T}}(\pm1)]$, we obtain

$$V^{\text{avg}}(t)=\gamma_0 e^{t\lambda^{\mathfrak{T}}(1+\rho M)}\left[1+(M-1)\left(\frac{\gamma_1}{\gamma_0}\right)^2 e^{-2\rho Mt\lambda^{\mathfrak{T}}}\right]^{1/2}\qquad(68)$$

and

$$V^{\text{axon}}(t)=\gamma_0 e^{t\lambda^{\mathfrak{T}}(1+\rho M)},\qquad(69)$$

where

$$\gamma_0:=|a_{01}(0)|/J^{\text{fix}},$$

$$\gamma_1:=\left[\sum_{l;l\neq0}|a_{l1}(0)|^2/(M-1)\right]^{1/2}/J^{\text{fix}}.\qquad(70)$$

The initial values $a_{l1}(0)$ are reduced to the two numbers γ_0 and γ_1. To estimate their order of magnitude, we calculate expectation values. As initial conditions, we assume $J_{mn}(0)=J^{\text{fix}}+\xi_{mn}$, the ξ_{mn} being uncorrelated white noise with mean zero and correlation function

$$\langle\xi_{mn}\xi_{m'n'}\rangle=d^2\delta_{mm'}\delta_{nn'}.$$

With Eqs. (49) and (25) we obtain the expectation value

$$\langle|a_{l1}(0)|^2\rangle=d^2/(MN).$$

Replacing $|a_{l1}|^2$ in Eq. (70) by $\langle|a_{l1}|^2\rangle$ we find

$$\gamma_0\approx\gamma_1\approx d/(J^{\text{fix}}\sqrt{MN}).\qquad(71)$$

These approximations have been used for plotting the time evolution of order parameters, viz., Eqs. (68) and (69), in Fig. 7 for the set of parameters from [8]. A comparison between Figs. 7 and 8 shows that the theoretical predictions are a reasonable approximation of the data obtained by numerical simulations with the same set of parameters.

C. Optimal coupling

The global order is always below the local order since

$$\frac{V^{\text{axon}}(t)}{V^{\text{avg}}(t)}=\left[1+\frac{(M-1)(\gamma_1/\gamma_0)^2}{e^{2\rho Mt\lambda^{\mathfrak{T}}}}\right]^{-1/2}<1.\qquad(72)$$

As $t\to\infty$, the ratio of vector strengths in Eq. (72) approaches 1 since both ρ and $\lambda^{\mathfrak{T}}$ are positive. So, if we wait sufficiently long, we should gain a perfect globally ordered map. The final quality of the map, therefore, seems to have only little dependence upon ρ. These considerations, however, only hold if the dynamics remains linear. For biological systems this assumption is generally not fulfilled, as the resources available for synaptic modification are restricted. This restriction can be mimicked by introducing an upper bound for the local order parameter V^{avg}. Once the boundary is reached (at $t=t^{\text{freeze}}>0$) the whole dynamics is assumed to be frozen. The difference between local and global order at freezing time, however, has a strong dependence upon ρ, cf. Fig. 7.

For fixed $t=t^{\text{freeze}}>0$, Eq. (72) shows that the bigger the axonal coupling ρ, the more the difference between the values of both order parameters diminishes. So one could argue that map formation becomes more and more effective, the larger the axonal coupling strength. In Sec. VI we will show
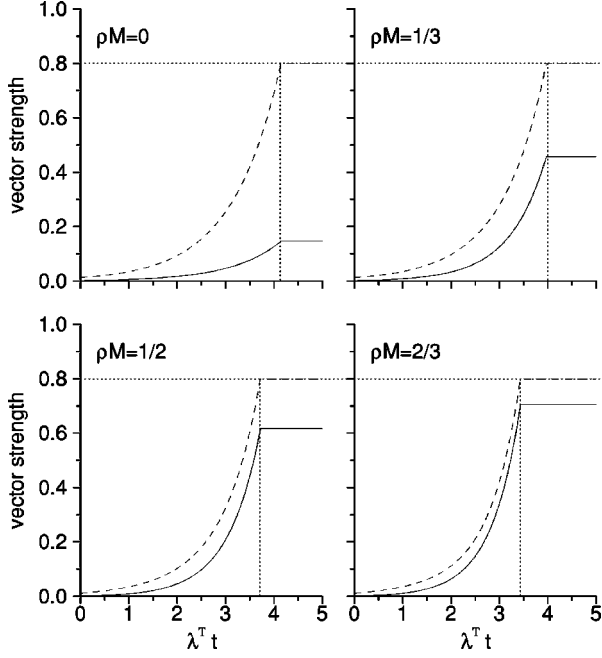
FIG. 7. Dynamics of order parameters. The global order parameter V^{axon} (solid line) and the local one V^{avg} (dashed line) are plotted as functions of time in units of the inverse real part of the leading eigenvalue $(\lambda^{\mathfrak{T}})^{-1}$ [see Eqs. (68), (69) and (71)]. The interaction strength ρM is varied systematically from $\rho M = 0$ (no interaction) to $\rho M = 2/3$ (strong interaction). Once the local order parameter reaches a saturation threshold, say vector strength 0.8 (dotted horizontal line), synapse growth stops. This can be achieved, e.g., by introducing an upper bound for the single synaptic weights. Hence at the same time (dotted vertical line) also the global order parameter stops increasing. Thus, the difference between global and local saturation depends on ρM. In accordance with [8], the following parameters have been used, $d/J^{\text{fix}} = 0.2$, $N = 250$, $M = 30$.

that this is not the case and that, due to noise, the upper limit for the axonal coupling strength $\rho \lesssim \mathcal{O}(M^{-1/2})$.

VI. INPUT NOISE

Up to now all results are based on Eq. (5). That is to say, we have been dealing with a time-averaged mean-field description, where the inherent temporal noise of the input processes is neglected. Calculating the variance of the temporal evolution of one synaptic weight (see [17], Appendix E) justifies the mean-field approximation for sufficiently small values of η: the diffusion constant is $\mathcal{O}(\eta^2)$ and the eigenvalues are $\mathcal{O}(\eta)$, which makes the time scale of synaptic modification due to input noise longer by a factor of $\mathcal{O}(\eta^{-1})$ than the time scale of structure formation.

If we set up an axonal interaction like in Eq. (6), the diffusion constant increases. Therefore, the time scale associated with input noise decreases. Hence we can gain a constraint for the coupling strength ρ, so as to avoid an increase of the order of magnitude of the input-noise evoked jitter.

The variance $\text{Var}(J_{mn})(t) := \langle J_{mn}^2 \rangle(t) - \langle J_{mn} \rangle^2(t)$ of a synaptic weight for $b_{mm'} = 1$ is governed by the dynamics (see also Secs. II B and II C)
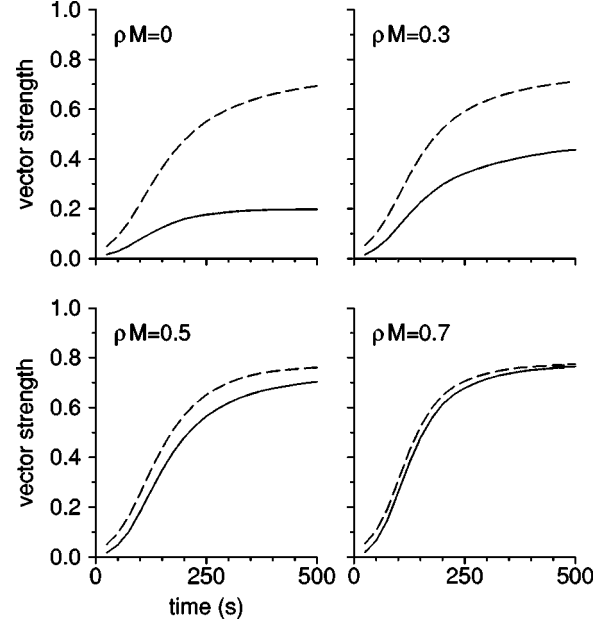

FIG. 8. Simulated dynamics of order parameters. The global (solid line) and local (dashed line) order parameters are plotted as functions of time, compare Fig. 7. As the interaction strength ρM is increased, the difference between the saturation values of global and local order diminishes. Saturation is realized by an upper bound of 2 for the individual synaptic weights. Again, parameters are chosen in accordance with [8].

$$\frac{d}{dt}J_{mn} = \eta \sum_{m'=1}^{M} (\delta_{mm'} + \rho) \left[w^{\text{in}} S_n(t - \Delta_{m'n}) + w^{\text{out}} S_{m'}(t) \right. $$
$$\left. + \int ds\, W(s) S_n(t + s - \Delta_{m'n}) S_{m'}(t) \right]. \quad (73)$$

Here $S_n(t) := \sum_{\{t_n^{(f)}\}} \delta(t - t_n^{(f)})$ denotes the spike train originating from the hillock of neuron n. Its ensemble average is, therefore, identified with intensity of the input process $p^{\text{in}}(t)$. Supposing that the efficacies are all approximately equal to J^{fix} at the beginning of learning, we obtain (see Appendix)

$$\text{Var}(J_{mn})(t) = t\{[1 + \rho M]^2 D_1 + [1 + 2\rho + \rho^2 M] D_2\}$$
$$+ \mathcal{O}(N^{-1/2}), \quad (74)$$

where

$$D_1 = \eta^2 [(w^{\text{in}})^2 \nu + 2 w^{\text{in}} \nu \nu^{\text{out}} \hat{W}(0) + \nu(\nu^{\text{out}})^2 \hat{W}(0)^2] \quad (75)$$

and

$$D_2 = \eta^2 \left[(w^{\text{out}})^2 \nu + 2 w^{\text{out}} \nu \nu^{\text{out}} \hat{W}(0) + \nu^2 \nu^{\text{out}} \sum_{\mu} |\hat{g}_{\mu}|^2 |\hat{W}_{\mu}|^2 \right.$$
$$\left. + \nu \nu^{\text{out}} \int ds\, W(s)^2 \right]. \quad (76)$$

Since the variance turns out to be linear in t, we can define a diffusion constant $D := t^{-1} \text{Var}(J_{mn})$. For $M \gg 1$, two scenarios exist.

(i) If $D_1 \gg D_2$ or $D_1 \approx D_2$, the leading power of the interaction terms in Eq. (74) is $(\rho M)^2$. For $\rho \lesssim M^{-1}$, D remains of the same order of magnitude when compared to $\rho = 0$. A number of 30 postsynaptic cells like in [8] then yields $\rho \lesssim 3.3\%$.

(ii) If $D_1 \ll D_2$, the leading power in Eq. (74) is $\rho^2 M$, which yields an upper limit $\rho \lesssim M^{-1/2}$. For $M = 30$ we get $\rho \lesssim 18\%$.

As shown in Figs. 7 and 8, an interaction strength $\rho = 1/2M = 1.7\%$ already suffices for an aligned development of local and global order with $M = 30$, and ρ is small enough not to violate the upper limits in either scenario. Furthermore, calculating the fraction D_1/D_2 with the parameters used in [8] yields $D_1/D_2 \approx 4 \times 10^{-3}$, which corresponds to scenario (ii), and so ρ is about ten times smaller than its possible limiting value.

To summarize, we state that in the limit $\rho \to \infty$ map formation is prevented because a high value of ρ amplifies the shot noise of spikes. For $\rho = 0$ there is no map formation either. As a consequence, there must be a value $\rho > 0$, where the synchronized development of temporal receptive fields is optimal. Already small values, e.g., $\rho \lesssim 1/M$ can be sufficient.

VII. MAP FORMATION WITH NONLINEAR NEURONS

So far we have dealt with linear Poisson neurons, cf. Eq. (2). We now extend the theory of spike-based Hebbian learning and map formation from neurons with a linear firing probability to those with an exponential firing probability [46]

$$p_F(v) = \nu_0 \exp[\beta v]. \tag{77}$$

As already announced in Sec. II A, we will derive a linear learning equation equivalent to Eqs. (12) and (18) for the deviation of the synaptic weight vector $\mathbf{J}$ from a fixed-point solution $\mathbf{J}^{\text{fix}}$ of the nonlinear dynamics. The idea is simple. We linearize Eq. (6) with respect to $\mathbf{J}^{\text{fix}}$ and will show that certain eigenvalues λ_i of the linearization $\mathbb{D} := \mathbb{D}(\mathbf{J}^{\text{fix}})$ may, and usually do, have a positive real part. The one with the largest real part $\text{Re}\{\lambda_i\} > 0$ will asymptotically dominate the time evolution $\exp(t\mathbb{D})$. One, therefore, may wonder: Why study an *unstable* fixed point?

The rationale underlying the above analysis is that the evolution operator $\exp(t\mathbb{D})$ is exactly soluble while the eigenspaces of $\mathbb{D}$ with $\text{Re}\{\lambda_i\} < 0$ indicate the "domain of attraction" of $\mathbf{J}^{\text{fix}}$. Given the initial position $\mathbf{J}(0)$, the system state $\mathbf{J}(t)$ first approaches the fixed point $\mathbf{J}^{\text{fix}}$ before the "expanding" eigenvalues with $\text{Re}\{\lambda_i\} > 0$ take over. Avoiding any special assumption regarding $\mathbf{J}(0)$, we assume its components to be independent random variables that are equidistributed in a certain interval (as in Fig. 5) and because $\mathbf{J}(0)$'s part in the "expanding" eigenspaces is negligible (see Sec. IV C 2), it is bound to first approach $\mathbf{J}^{\text{fix}}$. Hence the eigenvalue λ_i of $\mathbb{D}$ with the largest real part $\text{Re}\{\lambda_i\} > 0$ pro-

vides us with important information concerning pattern formation in the system of synaptic efficacies.

A. Linearizing the learning equation

In contrast to the linear learning Equation (18), the above activation function (77) yields a nonlinear synaptic dynamics,

$$\frac{d}{dt}\mathbf{J} = \mathcal{N}\mathbf{J}, \tag{78}$$

where $\mathcal{N}$ denotes a nonlinear operator mapping $\mathcal{R} \to \mathcal{R}$. Suppose we have found the fixed point $\mathcal{N}\mathbf{J}^{\text{fix}} = 0$ of the differential Equation (78). We then expand Eq. (78) in deviations $\boldsymbol{\iota} := \mathbf{J} - \mathbf{J}^{\text{fix}}$ so as to obtain the linearized dynamics

$$\frac{d}{dt}\boldsymbol{\iota} = \mathbb{D}\boldsymbol{\iota} + \mathcal{O}(\boldsymbol{\iota}^2). \tag{79}$$

Here $\mathbb{D}$ is the total derivative of $\mathcal{N}$ at position $\mathbf{J}^{\text{fix}}$.

B. Rate and correlation

To obtain expressions for $\mathcal{N}$ and its total derivative, we have to find a way of calculating the time-averaged rate $\nu_m^{\text{out}}(t) = \langle p_F(v_m(t)) \rangle$ and correlation function $C_{mn}(t+r,t) =: c_{mn}(t+r,t)$, cf. Eqs. (4) and (5). The correlation c_{mn} can be rewritten using Bayes' formula (for a proof, see [47], Appendix A),

$$c_{mn}(t+r,t) = \langle p_F(v_m(t+r)) \rangle|_{n,t} \, p^{\text{in}}(t). \tag{80}$$

The notation $\langle \cdots \rangle|_{n,t}$ indicates an average over input spike trains given a spike of input cell n at time t. For the time being, we leave the temporal averages aside.

Both ν_m^{out} and c_{mn} can be calculated by means of the Poissonian probability measure μ on the set of Poissonian input spike trains $\Omega = \cup_{F_1,\ldots,F_N} \Omega_{F_1 \cdots F_N}$, where $\Omega_{F_1 \cdots F_N}$ denotes the set of all spike trains with axon 1 conducting F_1 spikes, axon 2 conducting F_2 spikes, and so on; all this is to happen during a period of duration $\mathcal{T}$. Since the subsets $\Omega_{F_1 \cdots F_N}$ are disjoint, we can write

$$\langle p_F(v_m) \rangle = \int_\Omega d\mu \, p_F(v_m) = \sum_{F_1,\ldots,F_N} \int_{\Omega_{F_1 \cdots F_N}} d\mu \, p_F(v_m). \tag{81}$$

Though their occurrence is correlated through p_F, the individual spike events of a Poisson process are independent and, hence, the probability measure μ can be restricted to the subsets $\Omega_{F_1 \cdots F_N}$ so as to give

$$\int_{\Omega_{F_1 \cdots F_N}} d\mu \, p_F(v_m(t)) = \left(\prod_{n=1}^N \prod_{f=1}^{F_n} \int_{-\infty}^t dt_j^{(f)} \right) \times P_{\{F_n\}}(t, \{t_j^{(f)}\}) p_F(v_m(t)),$$

where, given the individual Poisson intensities p^{in},

$$P_{\{F_n\}}(t,\{t_j^{(f)}\})=\exp\left[-\sum_n\int_{-\infty}^t ds\, p^{\text{in}}(s)\right]$$
$$\times\left(\prod_{n=1}^N\frac{1}{F_n!}\prod_{f=1}^{F_n}p^{\text{in}}(t_j^{(f)})\right)$$

is the probability density for N Poissonian input spike trains with $\{F_n\}$ spikes at times $\{t_n^{(f)}\}$, cf. Appendix A 4 in [48].

The dependence of v_m on the input spike times can be found in Eq. (3). Substituting $y_{mn}(t-t_n^{(f)})=\beta J_{mn}\,\epsilon(t-t_n^{(f)}-\Delta_{mn})$ we find from Eq. (77),

$$p_F(v_m(t))=\nu_0\prod_{n=1}^N\prod_{f=1}^{F_n}\exp[y_{mn}(t-t_n^{(f)})]$$

and, thus,

$$\int_{\Omega_{F_1\cdots F_N}}d\mu\, p_F(v_m(t))$$
$$=\nu_0\exp\left[-N\int_{-\infty}^t ds\, p^{\text{in}}(s)\right]\prod_{n=1}^N\frac{1}{F_n!}$$
$$\times\left(\int_{-\infty}^t ds\, p^{\text{in}}(s)e^{y_{mn}(t-s)}\right)^{F_n}.$$

Taking advantage of Eq. (81) we end up with

$$\langle p_F(v_m(t))\rangle=\nu_0\exp\left\{\sum_{n=1}^N\int_{-\infty}^t ds\, p^{\text{in}}(s)[e^{y_{mn}(t-s)}-1]\right\}$$
$$=\nu_0\exp\left\{\sum_{n=1}^N\int_0^\infty ds\, p^{\text{in}}(t-s-\Delta_{mn})\right.$$
$$\left.\times[e^{\beta J_{mn}\epsilon(s)}-1]\right\}.\qquad(82)$$

The calculation of the conditional mean can be executed analogously. We find

$$\langle p_F(v_m(t+r))\rangle|_{n,t}=\langle p_F(v_m(t+r))\rangle e^{y_{mn}(r)}.\qquad(83)$$

The total derivative $\mathbb{D}$ in Eq. (79) is determined by its partial derivatives with respect to J_{mn},

$$\frac{\partial\langle p_F(v_m(t))\rangle}{\partial J_{mn}}=\beta\langle p_F(v_m(t))\rangle$$
$$\times\int_0^\infty ds\,\epsilon(s)\, p^{\text{in}}(t-s-\Delta_{mn})e^{\beta J_{mn}\epsilon(s)}\qquad(84)$$

and

$$\frac{\partial\langle p_F(v_m(t+r))\rangle|_{n',t}}{\partial J_{mn}}=\beta\langle p_F(v_m(t+r))\rangle\left[\int_0^\infty ds\,\epsilon(s)\right.$$
$$\times p^{\text{in}}(t+r-s-\Delta_{mn})e^{\beta J_{mn}\epsilon(s)}$$
$$+\delta_{nn'}\epsilon(r-\Delta_{mn})\Bigg]$$
$$\times\exp[\beta J_{mn'}\,\epsilon(r-\Delta_{mn'})].\qquad(85)$$

The temporal averages of Eqs. (82)–(85) will turn out to be simple, once we have found the correct fixed-point couplings J_{mn}^{fix}.

C. Fixed point of the nonlinear dynamics

We will prove under very general conditions that Eq. (78) has the fixed-point solution $\mathbf{J}^{\text{fix}}=J^{\text{fix}}(\mathbf{1}^{\mathfrak{S}}\otimes\mathbf{1}^{\mathfrak{T}})$. We, therefore, insert $J_{mn}=J^{\text{fix}}$ into Eq. (82) and due to

$$\sum_n p^{\text{in}}(t-\Delta_{mn})=\sum_n p^{\text{in}}(t-u_m/c-n\,T_p/N)$$
$$=p\sum_n g(t-u_m/c-n\,T_p/N)$$
$$=\nu\sum_{n\mu}\hat{g}_\mu\exp[2\pi i\mu(t-u_m/c$$
$$-n\,T_p/N)/T_p]$$
$$=\nu\sum_\mu N\delta_{\mu 0}\hat{g}_\mu=N\nu,$$

the mean firing probability (82) is constant,

$$\nu^{\text{out}}=\langle p_F(v_m(t))\rangle=\nu_0\exp\left\{N\nu\int_0^\infty ds\,[e^{\beta J^{\text{fix}}\epsilon(s)}-1]\right\}.\qquad(86)$$

Consequently, also the correlation function is independent of t,

$$C_{mn}(t+r,t)=\overline{p^{\text{in}}(t)\langle p_F(v_m(t+r))\rangle|_{nt}}$$
$$=\nu^{\text{out}}\nu\exp[\beta J^{\text{fix}}\,\epsilon(r-\Delta_{mn})].\qquad(87)$$

We insert Eqs. (86) and (87) into Eq. (5) and obtain from $\mathcal{N}\mathbf{J}^{\text{fix}}=\mathbf{0}$,

$$\gamma:=\frac{-w^{\text{in}}}{\nu\int ds\, W^{\text{eff}}(s;\beta J^{\text{fix}})+w^{\text{out}}}=\frac{\nu^{\text{out}}}{\nu},\qquad(88)$$

where the effective learning window is $W^{\text{eff}}(s;x)$ $:=W(s)\,e^{x\epsilon(-s)}$. In biological neural networks the ratio γ between the input and output rates is of the order of one. Any arbitrary value of $\gamma>0$ can be obtained by adjusting w^{in} and w^{out} on the left-hand side of Eq. (88) for any given values of ν and βJ^{fix}. Moreover, since it will turn out in Sec. VII D that the denominator $\nu\int ds\, W^{\text{eff}}(s;\beta J^{\text{fix}})+w^{\text{out}}$ has to be

negative in order to guarantee stabilization of the output rate, $\gamma > 0$ again requires $w^{\text{in}} > 0$. We thus postulate a fixed positive value γ and due to Eq. (86) we can write

$$\ln(\gamma\nu/\nu_0)/(N\nu) = \int_0^\infty ds\,[e^{\beta J^{\text{fix}}\epsilon(s)} - 1] =: \psi(\beta J^{\text{fix}}). \tag{89}$$

Here ψ is a monotonically increasing function of βJ^{fix} with $\psi(0) = 0$ and hence, for $\ln(\gamma\nu/\nu_0) > 0$, there is a unique fixed-point solution βJ^{fix}.

D. Linearized dynamics

The discussion of the eigensystems of the linearized dynamics (79) requires temporally averaged expressions of the partial derivatives in Eqs. (84) and (85) at $J_{mn} = J^{\text{fix}}$. We, therefore, introduce an effective spike response kernel $\epsilon^{\text{eff}}(s;x) := \epsilon(s)\,e^{x\epsilon(s)}$ and find

$$\overline{\frac{\partial\langle p_F(v_m(t))\rangle}{\partial J_{mn}}} = \beta\nu^{\text{out}}\,\nu\hat{\epsilon}^{\text{eff}}(0;\beta J^{\text{fix}}), \tag{90}$$

where $\hat{\epsilon}^{\text{eff}}(\omega;x)$ is the Fourier transform of $\epsilon^{\text{eff}}(s;x)$ with respect to the first argument. The temporally averaged partial derivative of the ensemble-averaged correlation function c_{mn} from Eq. (80) is

$$\overline{\frac{\partial C_{mn'}(t+r,t)}{\partial J_{mn}}} = \beta\,\nu^{\text{out}}\exp[\beta J^{\text{fix}}\epsilon(r - \Delta_{mn'})]$$

$$\times\left[\delta_{nn'}\,\nu\,\epsilon(r - \Delta_{mn}) + \int ds\,\epsilon^{\text{eff}}(s;\beta J^{\text{fix}})\right.$$

$$\left.\times\overline{p^{\text{in}}(t+r-s-\Delta_{mn})p^{\text{in}}(t)}\right]. \tag{91}$$

This yields a familiar form of the linearized dynamics (79),

$$\frac{d}{dt}\boldsymbol{\iota} = \sum_{mn}(\mathbf{e}_m \otimes \mathbf{f}_n) \times \sum_{m'n'}(1 + \rho\,\delta_{mm'})$$

$$\times[k_2 + \delta_{nn'}k_3 + Q_{nn'}]\boldsymbol{\iota}_{m'n'}, \tag{92}$$

where the constants k_2 and k_3 equal

$$k_2 = \eta\beta\nu^{\text{out}}\nu\,\epsilon_0^{\text{eff}}[\hat{W}_0^{\text{eff}}\nu + w^{\text{out}}],$$

$$k_3 = \eta\beta\nu^{\text{out}}\nu\int ds\,\epsilon(-s)W^{\text{eff}}(s;\beta J^{\text{fix}})$$

and the Hebbian matrix $Q_{nn'}$ is

$$Q_{nn'} = \eta\beta\nu^{\text{out}}\nu^2\sum_{\mu=-\infty}^{\infty}|\hat{g}_\mu|^2\,\hat{\epsilon}_\mu^{\text{eff}}\hat{W}_\mu^{\text{eff}}[1 - \delta_{\mu,0}]$$

$$\times\exp[i\omega_p(\Lambda_n - \Lambda_{n'})]. \tag{93}$$

Here we have introduced the abbreviations $\hat{W}_\mu^{\text{eff}} = \hat{W}^{\text{eff}}(\mu\omega_p;\beta J^{\text{fix}})$ and $\hat{\epsilon}_\mu^{\text{eff}} = \hat{\epsilon}^{\text{eff}}(\mu\omega_p;\beta J^{\text{fix}})$. Proceeding as
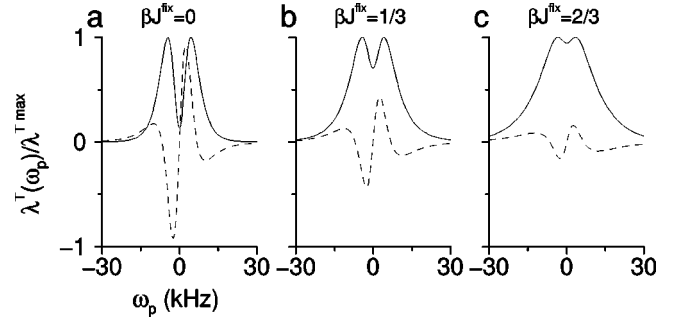

FIG. 9. Spectrum of temporal eigenvalues in dependence on the nonlinearity βJ^{fix}. Real parts are drawn as solid lines, dashed lines are imaginary parts. The eigenvalues are normalized with respect to the maximum value of their real part. (a) $\beta J^{\text{fix}} = 0$. The leading eigenvalue is equivalent to that of the linear model neurons. (b) $\beta J^{\text{fix}} = 1/3$, (c) $\beta J^{\text{fix}} = 2/3$. As βJ^{fix} increases from (a) to (c), the bandpass (for $\beta J^{\text{fix}} = 0$) becomes a low pass, and the imaginary parts approach zero.

in the case of the linear Poisson neuron we find that the temporal evolution of synaptic couplings, as described by the linear model, Eq. (12), is equivalent to a linearization of a nonlinear dynamics around the fixed-point solution. Moreover, also the fixed-point couplings are equivalent.

In Eq. (50) for a linear Poisson neuron, the spectrum of the latency dynamics is ruled by $\hat{W}\hat{\epsilon}$. In contrast, the Hebbian matrix $Q_{nn'}$ in Eq. (92) contains the factors $\hat{W}^{\text{eff}}\hat{\epsilon}^{\text{eff}}$. Its spectrum is shown in Fig. 9 for different values of βJ^{fix}. For $\beta J^{\text{fix}} = 0$ we obtain the same result as in the linear case, see Fig. 3. The reader may notice that the imaginary part of $\hat{W}\hat{\epsilon}$ in Fig. 3 is not small enough for biologically relevant frequencies to neglect a drifting of the nascent structure. This is to be expected, since the learning window W is taken from [8], where map formation is investigated with nonlinear cells. However, in the above discussion we have shown that nonlinearities change eigenvalues and, hence, for simulations with linear neurons, other choices for W and ϵ have to be made in order to produce nondrifting maps.

An increase of the nonlinearity βJ^{fix} has two consequences: (i) a smaller imaginary part so that the drifting tendency of the structure decreases and (ii) an increasing low-pass edge of the spectrum so that the sensitivity of the learning rule to higher frequencies is enhanced, cf. Fig. 9. In order to grasp how this works, we note that postsynaptic firing is preferentially induced by presynaptic spikes *preceding* output spikes by about the rise time of ϵ. In Fig. 10(a) we have plotted the effective learning window $W^{\text{eff}}(s)$ for $\beta J^{\text{fix}} = 0.1$, 0.25, and 0.5. The maximum of W^{eff} is shifted to the left (towards negative s) by the rise time of ϵ. The larger the positive peak, the more effective is the synaptic structure formation, since eigenvalues of the synaptic dynamics are provided by the Fourier transform of Q [see Eq. (47)] which is a convolution of W^{eff} with ϵ^{eff}, see Fig. 10(b). The maximum of Q is close to $s = 0$, and the magnitude of the peak is strongly affected by the nonlinearity. The more distinct the peak of Q at $s = 0$ the larger the enhancement of its real parts in Fourier space. This enhancement of the real part compared to the imaginary part as a function of βJ^{fix} is summarized in
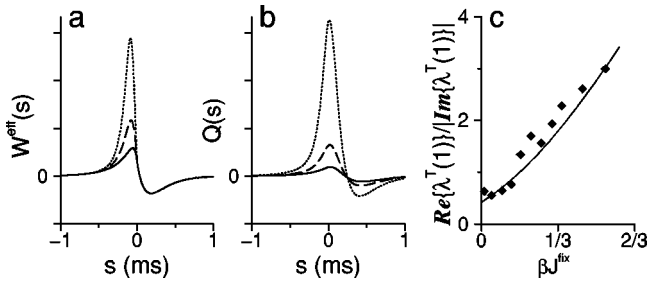
FIG. 10. Nonlinearities can stabilize structure formation. (a) The effective learning window W^{eff} for three values of $\beta J^{\text{fix}} = 0.1$ (solid line), 0.25 (dashed), and 0.5 (dotted). The larger βJ^{fix}, the more pronounced is the positive peak of W^{eff}. (b) The function Q as the convolution of W^{eff} and ϵ^{eff} has a maximum at about $s = 0$. (c) The ratio of real and imaginary parts of the leading eigenvalue $\lambda^{\mathfrak{T}}(1)$ from Eq. (47) for $T_p = 1/3$ ms (solid line) is a monotonically increasing function of βJ^{fix}. Diamonds are obtained from computer simulations with parameters taken from [8]. We have calculated the contribution of the leading eigenvector to the synaptic structure $a_1 := \boldsymbol{\phi}^{\mathfrak{T}}(-1) \cdot \mathbf{J}(t) \propto \exp[t \, \lambda^{\mathfrak{T}}(1)]$. Real and imaginary parts of the leading eigenvalue are computed as a linear fit to the amplitude $\ln|a_1(t)| = t \, \text{Re}\{\lambda^{\mathfrak{T}}(1)\} + \text{const}$ and the phase $\arg[a_1(t)] = t \, \text{Im}\{\lambda^{\mathfrak{T}}(1)\} + \text{const}$ of the synaptic structure, respectively. Fits have been performed at times before the amplitude has reached twice its initial value.

Fig. 10(c). We have plotted the ratio of the real and the imaginary part of the leading eigenvalue as predicted by the present theory and as extracted from computer simulations. This result confirms that nonlinearities may stabilize the Hebbian structure formation.

VIII. DISCUSSION

Synaptic plasticity plays an important role in the development and ongoing modification of neuronal maps [49]. Earlier theories on map formation [50–54] due to synaptic plasticity deal with rate-based linear neuron models. In spite of their success in explaining, e.g., visual orientation maps, ocular dominance columns, and even spatiotemporal maps of visual directional selectivity, they are not capable of explaining maps of stimulus features that are solely temporal.

Recently, spike-based learning rules have become more and more prevalent [8,14–32]. The present work is an attempt to understand the formation of temporal-feature maps by means of a spike-timing dependent synaptic learning rule. As a facinating example, we have studied maps of interaural time differences (ITDs) in the laminar nucleus of the barn owl.

Synaptic learning at the level of a single laminar neuron can result in the development of ITD tuning. The best ITD where a neuron has its maximum response depends on random initial conditions. If we set up an array of neurons, the best ITDs can be ordered systematically through presynaptic nonspecific propagation of a local Hebbian learning along an axon, called axon-mediated spike-based learning (AMSL), viz., Eq. (5). We have modeled this presynaptic interaction and have shown that the resulting learning equation separates into a temporal and a spatial part. The temporal part of the

learning equation represents the dynamics of delay selection at a single neuron.

The spectrum of temporal eigenvalues shows that the learning rule acts as a low-pass filter, causing the evolution of higher-order harmonics of the weight structure to be slower than the increase of the principal harmonic. Similar results have been found in [19]. The spatial part of the learning equation mirrors the presynaptic interaction. Synchronization of the individual neurons' best ITDs needs a long-range presynaptic interaction.

If the Fourier-transformed coupling matrix $\tilde{b}_l$ as given by Eq. (42) has the property $\tilde{b}_0 \approx \tilde{b}_{\pm 1}$, for instance, because of a spatially restricted interaction, then synaptic efficacies arise with spatially oscillating strengths. Thus, axonally mediated coordination of synaptic strengthening may break down at smaller distances, which leads to a patchy pattern of afferent innervation and, thus, to a representation of stimulus diversity along the direction of arborization. We note that in the case of the Jeffress map, diversity is a consequence of the arrangement of laminar neurons along antiparallel axons and a T_p-periodic tuning of their delays (Fig. 6).

The quality of the temporal-feature map has been shown to increase with the magnitude $\rho > 0$ of the presynaptic interaction along the axon. Because the spatial eigenvalues scale with ρM and the number M of presynaptically coupled neurons might be quite high, even small values of $\rho \lesssim M^{-1}$ suffice for map formation. This may be the reason for why there are only a few papers [33–37] experimentally verifying presynaptic propagation of synaptic learning. There is also a computational reason for a small ρ. Presynaptic propagation of synaptic learning supports map formation as long as ρ is small enough to guarantee the validity of the temporally and ensemble-averaged mean-field learning equations (5) and (6). Both are violated, if the inherent noise of the stochastic processes gets too big. Sources of noise are the shot noise of spike input and output and the temporal jitter of the stochastic input process. Each synaptic weight, therefore, performs a random walk around the mean trajectory. If its standard deviation approaches the same order of magnitude as the mean, the mean-field description breaks down. We found that the presynaptic interaction amplifies noise because fluctuations of synaptic weights are propagated to neighboring neurons. We have obtained an estimate of the order of magnitude of the axonal coupling ρ (Sec. VI), where the standard deviation is small enough to keep the mean-field assumption intact. As a consequence, the exact value of ρ is not critical to map formation.

Most of our calculations have been performed using a (linear) Poisson neuron. Real neurons are neither Poissonian nor linear. We have, therefore, shown that introducing nonlinearity into the Poissonian dynamics leads to an almost identical learning equation (92) and very similar behavior of the synaptic dynamics, see also [8]. Furthermore, map formation can *benefit* from the nonlinearity, which, in our case, tends both to inhibit drifting of the synaptic structure and to increase the temporal precision of the learning rule. In order to discuss stabilization of the output rate in an even more general case (cf. Sec. IV C 1), it has been shown elsewhere [25] that nonlinear threshold dynamics and also refractori-

ness yield similar learning equations.

For map formation, there are various synaptic interaction mechanisms, such as recurrent synaptic coupling [51–54], gap junctions, and diffusing messengers. Trying to explain *temporal* map formation by means of these mechanisms, we have to bear in mind that the difference between best ITDs of neighboring cells in the laminar nucleus and in our model is less than 10 μs. Then the abovementioned alternative interaction mechanisms are unsuitable to achieve such an ITD map. Gap junctions, though they cannot be ruled out in young animals, have time constants of about the membrane time constant ($\approx$100 μs). Extracellular diffusion is spatially isotropic and, therefore, cannot produce diversity. Recurrent synaptic coupling, besides its nonexistence in the laminar nucleus, is at least as slow as 100 μs, a drawback that also applies to the time scale of population responses [55]. In short, the temporal precision of ITD tuning at single neurons is a result of the local spike-based Hebbian learning rule, whereas the even higher temporal precision of the map, i.e., the gradient of best ITDs, is warranted by the axonal topography and the presynaptic nonspecific propagation of local learning.

We note that, although our learning rule is capable of forming structures that suffice for temporal information processing in a submillisecond range, the actual weight changes, induced by temporally precise interactions between presynaptic and postsynaptic activity, may occur at a time scale of minutes. The time scale of structure formation depends on the learning parameter η, which, e.g., in computer simulations [8] is supposed to be of the order of 10^{-4}. In reality η may even be smaller.

In summary, we suggest that presynaptic unspecific propagation of synaptic learning (AMSL) is a key mechanism to the formation of maps whenever there is a temporal aspect. The simple reason behind this suggestion is that temporal correlations of input activity are best preserved along the afferent axons.

APPENDIX: CALCULATION OF THE VARIANCE

To estimate the effect of shot noise due to spike trains, we calculate the variance of a single synaptic weight $\mathrm{Var}(J_{mn})(t) := \langle J_{mn}^2 \rangle(t) - \langle J_{mn} \rangle^2(t) = \langle [J_{mn}(t) - J_{mn}(0)]^2 \rangle - \langle J_{mn}(t) - J_{mn}(0) \rangle^2$, at the beginning of learning, i.e., when all strengths J_{mn} are randomly distributed around the fixed-point value J^{fix}. Here $\langle \cdots \rangle$ denotes the ensemble average addressing the stochastic nature of both the input and the output spike trains $S(t)$. Referring to the stochastic differential equation (73), we find the variance to be

$$
\begin{aligned}
\mathrm{Var}(J_{mn})(t) = \eta^2 \sum_{m'm''} [\delta_{mm'} + \rho][\delta_{mm''} + \rho] \Bigg\{ & (w^{\mathrm{in}})^2 \int_0^t dt' \int_0^t dt'' [\langle S_n(t')S_n(t'') \rangle - \nu^2] \\
& + 2w^{\mathrm{in}}w^{\mathrm{out}} \int_0^t dt' \int_0^t dt'' [\langle S_n(t')S_{m'}(t'') \rangle - \nu\,\nu^{\mathrm{out}}] + (w^{\mathrm{out}})^2 \int_0^t dt' \int_0^t dt'' [\langle S_{m'}(t')S_{m''}(t'') \rangle - (\nu^{\mathrm{out}})^2] \\
& + 2w^{\mathrm{in}} \int_0^t dt' \int_0^t dt'' \int_{-\infty}^{\infty} ds\, W(s)[\langle S_n(t')S_n(t''+s-\Delta_{m'n})S_{m'}(t'') \rangle - \nu^2\nu^{\mathrm{out}}] \\
& + 2w^{\mathrm{out}} \int_0^t dt' \int_0^t dt'' \int_{-\infty}^{\infty} ds\, W(s)[\langle S_n(t'+s-\Delta_{m'n})S_{m'}(t')S_{m''}(t'') \rangle - (\nu^{\mathrm{out}})^2\nu] \\
& + \int_0^t dt' \int_0^t dt'' \int_{-\infty}^{\infty} ds'\, W(s') \int_{-\infty}^{\infty} ds''\, W(s'')[\langle S_n(t'+s'-\Delta_{m'n})S_n(t''+s''-\Delta_{m''n}) \\
& \times S_{m'}(t')S_{m''}(t'') \rangle - \nu^2(\nu^{\mathrm{out}})^2] \Bigg\}.
\end{aligned}
\tag{A1}
$$

Here we assumed the output rate ν^{out} to be independent of the postsynaptic location m, which we will show to be a good approximation for randomly distributed weights, provided the number of input lines N is very large.

Let us now perform the six ensemble averages in Eq. (A1) step by step by recursively applying the definition of the conditional probability, $\mathrm{Prob}\{\cup_{i=1}^I B_i\}$

$= \mathrm{Prob}\{B_1\}\,\mathrm{Prob}\{B_2|B_1\}\cdots\mathrm{Prob}\{B_I|\cup_{i=1}^{I-1}B_i\}$, where $\mathrm{Prob}\{B_i|B_j\}$ means $\mathrm{Prob}\{B_i$ under condition $B_j\}$. Since we are dealing with Poisson processes, the conditional probability densities can be computed according to [17], Appendix A.

(i) $\langle S_n(t')S_n(t'') \rangle = p^{\mathrm{in}}(t')[p^{\mathrm{in}}(t'') + \delta(t'-t'')]$. In addition, if we assume $t \gg T_p$, we obtain $t^{-1}\int_0^t dt'\,p^{\mathrm{in}}(t') \approx \nu$

and, therefore, the first term inside the curly brackets of Eq. (A1) equals $(w^{\text{in}})^2 \nu t$.

(ii) $\langle S_n(t') S_{m'}(t'') \rangle = p^{\text{in}}(t') \lim_{\delta t \to 0} [\delta t^{-1} \text{Prob}\{ m' \text{ fires}$ in $[t'', t'' + \delta t)$ if n fires at $t'\}]$. Applying expansion (2) to the conditional firing probability, we find

$$\langle S_n(t') S_{m'}(t'') \rangle = p^{\text{in}}(t') \Bigg[\beta^{(0)} + \beta^{(1)} \sum_n J_{m'n}$$

$$\times \int ds' \, \epsilon(s') p^{\text{in}}(t'' - s' - \Delta_{m'n})$$

$$+ \beta^{(1)} J_n \, \epsilon(t'' - t' - \Delta_{mn}) \Bigg]$$

$$=: p^{\text{in}}(t')[p_{m'}^{\text{out}}(t'') + O(1/N)].$$

The last term of the first equality is $O(1/N)$ as compared to the summation over all N synapses. The output intensity $p_m^{\text{out}}(t)$ can be calculated explicitly if we plug in the Fourier series of the periodic input process $p^{\text{in}} = p \, g$ introduced in Sec. II D 1,

$$p_m^{\text{out}}(t) = \beta^{(0)} + \beta^{(1)} \nu \sum_\mu \hat{g}_\mu \hat{\epsilon}_\mu \sum_n J_{mn}$$

$$\times \exp[i \mu \omega_p(t - \Delta_{mn})].$$

We consider the system to be in the initial state where synaptic weights are randomly distributed around J^{fix}. We, therefore, apply the central limit theorem to $\sum_n J_{mn} = N J^{\text{fix}}[1 + O(1/\sqrt{N})]$, and due to $\sum_n e^{i \mu \omega_p \Lambda_n} \approx N \delta_{\mu, 0}$, we obtain

$$p_m^{\text{out}}(t) = \beta^{(0)} + \beta^{(1)} \nu N J^{\text{fix}} =: \nu^{\text{out}}. \tag{A2}$$

This also justifies the above assumption of the output rate being independent of the spatial position m. We thus end up with

$$\langle S_n(t') S_{m'}(t'') \rangle = p^{\text{in}}(t') \nu^{\text{out}} + O(1/\sqrt{N}).$$

The second term on the right of Eq. (A1) is, therefore, of order $O(1/\sqrt{N})$.

(iii) $\langle S_{m'}(t') S_{m''}(t'') \rangle = p^{\text{out}}(t')[p^{\text{out}}(t'') + \delta(t' - t'') \delta_{m'm''}]$. With Eq. (A2), the third term, therefore, equals $(w^{\text{out}})^2 \nu^{\text{out}} \delta_{m'm''} t$.

(iv) $\langle S_n(t') S_n(t'' + s - \Delta_{m'n}) S_{m'}(t'') \rangle = p^{\text{in}}(t')\{[p^{\text{in}}(t'' + s - \Delta_{m'n}) + \delta(t'' - t' + s - \Delta_{m'n})][p_{m'}^{\text{out}}(t'') + O(1/N)]\}$. Equation (A2) and both time integrals yield the fourth term, $2 w^{\text{in}} \nu \nu^{\text{out}} \hat{W}(0) t$. Here we assumed the integration time t to be much larger than the width of the learning window $\mathcal{W}$ and the typical values of Δ_{mn}.

(v)

$$\langle S_n(t' + s - \Delta_{m'n}) S_{m'}(t') S_{m''}(t'') \rangle$$

$$= p^{\text{in}}(t' + s - \Delta_{m'n})\{[p_{m'}^{\text{out}}(t') + O(1/N)]$$

$$\times [p_{m''}^{\text{out}}(t'') + O(1/N) + \delta_{m'm''} \delta(t' - t'')]\}.$$

Applying again Eq. (A2) we find that both temporal integrations are straightforward. So we write the fifth term $2 w^{\text{out}} \nu \nu^{\text{out}} \hat{W}(0) \delta_{m'm''} t$.

(vi)

$$\langle S_n(t' + s' - \Delta_{m'n}) S_n(t'' + s'' - \Delta_{m''n}) S_{m'}(t') S_{m''}(t'') \rangle$$

$$= p^{\text{in}}(t' + s' - \Delta_{m'n})[p^{\text{in}}(t'' + s'' - \Delta_{m''n}) + \delta(t' - t''$$

$$+ s' - s'' + \{u_{m''} - u_{m'}\}/c)][p_{m'}^{\text{out}}(t') + O(1/N)]$$

$$\times [p_{m''}^{\text{out}}(t'') + O(1/N) + \delta_{m'm''} \delta(t' - t'')].$$

We split up the above expression into four parts, apply Eq. (A2), and insert it into Eq. (A1). We then end up with four integrals over time (a)–(d),

(a)

$$(\nu^{\text{out}})^2 \int_0^t dt' \int_0^t dt'' p^{\text{in}}(t' + s' - \Delta_{m'n}) p^{\text{in}}(t'' + s'' + \Delta_{m''n})$$

$$= (\nu^{\text{out}})^2 \nu^2 t^2.$$

(b)

$$(\nu^{\text{out}})^2 \int_0^t dt' \int_0^t dt'' \, p^{\text{in}}(t' + s' - \Delta_{m'n}) \delta(t' - t'' + s' - s''$$

$$+ [u_{m''} - u_{m'}]/c) = (\nu^{\text{out}})^2 \nu \, t.$$

(c)

$$\nu^{\text{out}} \delta_{m'm''} \int_0^t dt' \int_0^t dt'' \, p^{\text{in}}(t' + s' - \Delta_{m'n}) p^{\text{in}}(t'' + s''$$

$$+ \Delta_{m''n}) \delta(t' - t'')$$

$$= \nu^{\text{out}} \delta_{m'm''} \int_0^t dt' \, p^{\text{in}}(t' + s' - \Delta_{m'n})$$

$$\times p^{\text{in}}(t' + s'' - \Delta_{m''n})$$

$$= \nu^{\text{out}} \delta_{m'm''} \nu^2 \sum_{\mu \mu'} \hat{g}_\mu \hat{g}_{\mu'}$$

$$\times \int_0^t dt' \exp\{i \omega_p[\mu(t' + s' - \Delta_{m'n})$$

$$+ \mu'(t' + s'' - \Delta_{m''n})]\}$$

$$= \nu^{\text{out}} \nu^2 t \sum_\mu |\hat{g}_\mu|^2 \exp[i \mu \omega_p(s' - s'')] \delta_{m'm''}.$$

Here we have exploited $\int_0^t dt' \, e^{i \mu \omega_p t'} = t \, \delta_{\mu \, 0}$ for $t \gg T_p$.

(d)

$$\nu^{\text{out}} \delta_{m'm''} \int_0^t dt' \int_0^t dt'' \, p^{\text{in}}(t' + s' - \Delta_{m'n})$$

$$\times \delta(s' - s') \delta(t' - t'')$$

$$= \nu^{\text{out}} \nu \, t \, \delta(s' - s'') \delta_{m'm''}.$$

Executing both integrals over the learning window (s' and s'') yields the sixth term of Eq. (A1), $t\nu^{\text{out}}\nu\{\hat{W}(0)^2[\nu^{\text{out}}+\nu]+\delta_{m'm''}[\nu\Sigma_\mu|\hat{W}_\mu|^2|\hat{g}_\mu|^2 +\int ds\, W(s)^2]\}$.

Adding (i)–(vi) leads to a much simpler expression for Eq. (A1),

$$\text{Var}(J_{mn})=t\sum_{m'm''}[D_1+\delta_{m'm''}D_2][\delta_{mm'}+\rho][\delta_{mm''}+\rho],$$
$$(A3)$$

where

$$D_1=\eta^2[(w^{\text{in}})^2\nu+2\,w^{\text{in}}\nu\nu^{\text{out}}\hat{W}(0)+\nu(\nu^{\text{out}})^2\hat{W}(0)^2],$$

$$D_2=\eta^2\left[(w^{\text{out}})^2\nu+2\,w^{\text{out}}\nu\nu^{\text{out}}\hat{W}(0)+\nu^2\nu^{\text{out}}\sum_\mu|\hat{g}_\mu|^2|\hat{W}_\mu|^2\right.$$
$$\left.+\nu\nu^{\text{out}}\int ds\, W(s)^2\right].$$

The double sum in Eq. (A3) then yields

$$\text{Var}(J_{mn})(t)=t\{[1+\rho M]^2D_1+[1+2\rho+\rho^2M]D_2\},$$
$$(A4)$$

as announced in Eq. (74).

[1] J.O. Pickles, *An Introduction to the Physiology of Hearing* (Academic Press, London, 1988).

[2] M. Konishi, T.T. Takahashi, H. Wagner, W.E. Sullivan, and C.E. Carr, in *Auditory Function. Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Wall, and W. M. Cowan (Wiley, New York, 1988), pp. 721–746.

[3] B. Grothe, T.J. Park, and G. Schuller, J. Neurophysiol. **77**, 1553 (1997).

[4] H. Agmon-Snir, C.E. Carr, and J. Rinzel, Nature (London) **393**, 268 (1998).

[5] C.E. Carr and M. Konishi, J. Neurosci. **10**, 3227 (1990).

[6] C.E. Carr, in *Advances of Hearing Research*, edited by G.A. Manley, G.M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore, 1995), pp. 24–30.

[7] E.M. Overholt, E.W. Rubel, and R.L. Hyson, J. Neurosci. **12**, 1698 (1992).

[8] R. Kempter, C. Leibold, H. Wagner, and J.L. van Hemmen, Proc. Natl. Acad. Sci. U.S.A. **98**, 4166 (2001).

[9] C. Leibold, R. Kempter, H. Wagner, and J.L. van Hemmen, Neurocumputing **38-40**, 1401 (2001).

[10] A. Moiseff and M. Konishi, J. Neurosci. **1**, 40 (1981).

[11] A. Moiseff, J. Comp. Physiol., A **164**, 629 (1989).

[12] W.E. Sullivan and M. Konishi, Proc. Natl. Acad. Sci. U.S.A. **83**, 8400 (1986).

[13] C. Köppl, J. Neurosci. **17**, 3312 (1997).

[14] A.V.M. Herz, B. Sulzer, R. Kühn, and J.L. van Hemmen, Biol. Cybern. **60**, 457 (1989).

[15] W. Gerstner, R. Kempter, J.L. van Hemmen, and H. Wagner, Nature (London) **383**, 76 (1996).

[16] B. Ruf and M. Schmitt, in *Proceedings of the Seventh International Conference on Artificial Neural Networks (ICANN '97)*, edited by W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud (Springer, Heidelberg, 1997), pp. 361–366.

[17] R. Kempter, W. Gerstner, and J.L. van Hemmen, Phys. Rev. E **59**, 4498 (1999).

[18] P.D. Roberts, J. Comput. Neurosci. **7**, 235 (1999).

[19] C.W. Eurich, K. Pawelzik, U. Ernst, J.D. Cowan, and J.G. Milton, Phys. Rev. Lett. **82**, 1594 (1999).

[20] S. Song, K.D. Miller, and L.F. Abbott, Nat. Neurosci. **3**, 919 (2000).

[21] W.M. Kistler and J.L. van Hemmen, Neural Comput. **12**, 385 (2000).

[22] M.C.W. van Rossum, G.-q. Bi, and G.G. Turrigiano, J. Neurosci. **20**, 8812 (2000).

[23] W. Senn, H. Markram, and M. Tsodyks, Neural Comput. **13**, 35 (2001).

[24] J. L. van Hemmen, in *Handbook of Biological Physics*, edited by F. Moss and S. Gielen (Elsevier, Amsterdam, 2001), Vol. 4, pp. 771–823.

[25] R. Kempter, W. Gerstner, and J.L. van Hemmen, Neural Comput. **13**, 2709 (2001).

[26] W.B. Levy and D. Stewart, Neuroscience **8**, 791 (1983).

[27] C.C. Bell, V.Z. Han, Y. Sugawara, and K. Grant, Nature (London) **387**, 278 (1997).

[28] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann, Science **275**, 213 (1997).

[29] L.I. Zhang, H.-z.W. Tao, C.E. Holt, W.A. Harris, and M.-m. Poo, Nature (London) **395**, 37 (2000).

[30] D. Debanne, B.H. Gähwiler, and S.M. Thompson, J. Physiol. (Paris) **507**, 237 (1998).

[31] G.-q. Bi and M.-m. Poo, J. Neurosci. **18**, 10 464 (1998).

[32] D.E. Feldman, Neuron **27**, 45 (2000).

[33] T. Bonhoeffer, V. Staiger, and A. Aertsen, Proc. Natl. Acad. Sci. U.S.A. **86**, 8112 (1989).

[34] T. Bonhoeffer, A. Kossel, J. Bolz, and A. Aertsen, in *Cold Spring Harbor Symposia on Quantitative Biology* (Cold Spring Harbor Laboratory Press, New York, 1990), Vol. LV, pp. 137–146.

[35] S. Cash, R.S. Zucker, and M.-m. Poo, Science **272**, 998 (1996).

[36] R.M. Fitzsimonds, H.-J. Song, and M.-m. Poo, Nature (London) **388**, 439 (1997).

[37] H.-z.W. Tao, L.I. Zhang, G.-q. Bi, and M.-m. Poo, J. Neurosci. **20**, 3233 (2000).

[38] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981), Sec. III 3; C. Kittel, *Elementary Statistical Physics* (Wiley, New York, 1958), Sec. 28.

[39] A.C. Aitken, *Determinants and Matrices*, 9th ed. (Oliver and Boyd, Edinburgh, 1956), Sec. 51.

[40] W. Ledermann, Proc. R. Soc. London, Ser. A **182**, 362 (1944).

[41] D.J.C. MacKay and K.D. Miller, Network **1**, 257 (1990).

[42] S. Wimbauer, W. Gerstner, and J.L. van Hemmen, Network **9**, 449 (1998).

[43] C.E. Carr and R.E. Boudreau, J. Comp. Neurol. **334**, 337 (1993).

[44] L.A. Jeffress, J. Comp. Physiol. Psychol. **41**, 35 (1948).

[45] J.M. Goldberg and P.B. Brown, J. Neurophysiol. **32**, 613 (1969).

[46] W. Gerstner and J. L. van Hemmen, in *Models of Neural Networks II*, edited by E. Domany, J.L. van Hemmen, and K. Schulten (Springer, New York, 1994), pp. 1–93.

[47] R. Kempter, W. Gerstner, J.L. van Hemmen, and H. Wagner, Neural Comput. **10**, 1987 (1998).

[48] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).

[49] D.V. Buonomano and M.M. Merzenich, Annu. Rev. Neurosci. **21**, 149 (1998).

[50] K.D. Miller, J.B. Keller, and M.P. Stryker, Science **245**, 605 (1989).

[51] K.D. Miller, J. Neurosci. **14**, 409 (1994).

[52] K.D. Miller, in *Models of Neural Networks III*, edited by E. Domany, J.L. van Hemmen, and K. Schulten (Springer, New York, 1996), pp. 55–78.

[53] S. Wimbauer, O.G. Wenisch, K.D. Miller, and J.L. van Hemmen, Biol. Cybern. **77**, 453 (1997).

[54] S. Wimbauer, O.G. Wenisch, J.L. van Hemmen, and K.D. Miller, Biol. Cybern. **77**, 463 (1997).

[55] W. Gerstner, Neural Comput. **12**, 43 (2000).