Biol Cybern (2007) 96:533–546
DOI 10.1007/s00422-007-0148-2

ORIGINAL PAPER

Biological Cybernetics

Spike-timing-dependent plasticity for neurons with recurrent connections

A. N. Burkitt · M. Gilson · J. L. van Hemmen

Received: 2 May 2006 / Accepted: 3 March 2007 / Published online: 6 April 2007

Abstract The dynamics of the learning equation, which describes the evolution of the synaptic weights, is derived in the situation where the network contains recurrent connections. The derivation is carried out for the Poisson neuron model. The spiking-rates of the recurrently connected neurons and their cross-correlations are determined self-consistently as a function of the external synaptic inputs. The solution of the learning equation is illustrated by the analysis of the particular case in which there is no external synaptic input. The general learning equation and the fixed-point structure of its solutions is discussed.

1 Introduction

The hypothesis that changes in the efficacies of neuronal connections, both during and after development, depend upon the correlations in timing of pre- and postsynaptic action potentials (spikes) has received considerable experimental support (Bi and Poo 2001). A number of features of neural information processing have successfully been accounted for by models of such spike-timing-dependent synaptic plasticity (STDP). These include the astonishing precision of barn owl sound localization (Gerstner et al. 1996) and the development of a temporal-feature map in the avian laminar nucleus (Leibold et al. 2001, 2002).

The development of a general theory of synaptic plasticity that is based on the notions of Hebbian learning (Hebb 1949) and a "learning window", that relates the presynaptic input and postsynaptic output times to the corresponding change of the synaptic weight, has underpinned these developments (for a review see van Hemmen 2001). Specifically, a synaptic weight is potentiated if a presynaptic input precedes a postsynaptic spike, and is depressed otherwise (Markram et al. 1997). In this paper we generalize this theory of synaptic plasticity to analyze the situation where a neuron is part of a network that has recurrent synaptic connections. The recurrent dynamics introduces additional difficulties in analyzing the resulting pattern of synaptic strengths. Previous analyzes of synaptic plasticity have focussed on the situation in which the synapses are part of a feed-forward network structure in which recurrent connections do not occur.

We first present a derivation of the differential equation governing the evolution of both the feed-forward and the recurrent synaptic weights, which is given in terms of the timing relationships between the pre- and postsynaptic spikes, in Sect. 2. The evaluation of the spike-timing correlations is presented in Sect. 3 for the Poisson neuron and extends the earlier analysis for feed-forward networks (Kempter et al. 1998). In the feed-forward model the output spiking-rate of the postsynaptic neuron has a linear dependence upon the synaptic weight, which facilitates the analysis, whereas in the recurrently connected model investigated here there is a nonlinear dependence. In Sect. 4 the full system of five coupled equations is presented, consisting of four consistency

A. N. Burkitt (✉) · M. Gilson
The Bionic Ear Institute, 384–388 Albert Street,
East Melbourne, VIC 3002, Australia
e-mail: aburkitt@bionicear.org

A. N. Burkitt · M. Gilson
Department of Electrical and Electronic Engineering,
The University of Melbourne, Melbourne, VIC 3010, Australia

A. N. Burkitt
Department of Otolaryngology, The University of Melbourne,
Melbourne, VIC 3010, Australia

J. L. van Hemmen
Physik Department, TU München,
85747 Garching bei München, Germany

equations and two differential equations that describes the learning on the external and recurrent weights. The solution of the learning equation in the case where there are no external synaptic inputs is presented in Sect. 5.

2 Spike-timing-dependent synaptic plasticity (STDP)

In this section the "learning equation" is derived for the Poisson neuron model by using the explicit time relationships between synaptic inputs and spike outputs. We consider a network of N neurons, each of which receives synaptic input via both recurrent connections (denoted by J_{ij}, a $N \times N$ matrix with zeros on the diagonal in order to prevent self-connections) and feed-forward connections from M external inputs (the recurrent connections are denoted by K_{ik}, a $N \times M$ matrix). The evolution of both these sets of weights is considered.

2.1 Weight dynamics

Consider first the change in the synaptic strength of an excitatory recurrent synapse J_{ij} that connects the postsynaptic neuron i with the presynaptic neuron j. At synapse $\{ij\}$ (with $1 \le i, j \le N$) input spikes arrive at times t_j^n (n is a label representing the index of the sequence of spikes), and these spikes are the output spikes of neuron j in the network. Likewise, at synapse $\{ik\}$ (with $1 \le i \le N$ and $1 \le k \le M$) input spikes arrive at times $\widehat{t}_k^m$ (m represents the index of the spike-sequence), and these spikes are from *external* inputs. This is depicted in Fig. 1, which shows only the synaptic connections between the three illustrated neurons (two recurrently connected and one external).

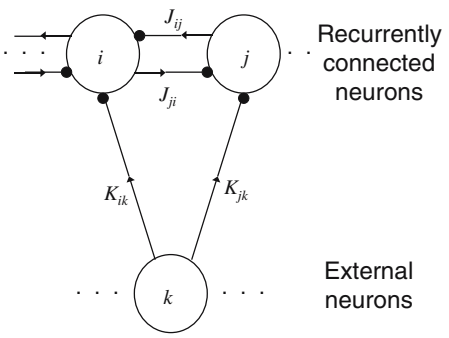

Fig. 1 Synaptic connectivity illustrated for two neurons within a network of recurrently connected neurons $\{i, j\}$ receiving external synaptic input (a single external neuron k is illustrated here). Neurons are indicated by a *large circle* and excitatory synapses by *small filled circles*. We study the development of both the recurrent synaptic weights J_{ij} ($1 \le i, j \le N$) and the weights from external neurons K_{ik} ($1 \le i \le N$, $1 \le k \le M$). The neuron i produces output spike flow denoted by $S_i(t)$ and the external neuron k produces a spike flow denoted by $\widehat{S}_k(t)$; cf. (6)

The set of excitatory synaptic efficacies $\{J_{ij}, K_{ik}\}$ ($1 \le i, j \le N,\ 1 \le k \le M$) determines the membrane potential of neuron i,

$$v_i(t) = V_r + \sum_{j,n} J_{ij}(t_j^n)\, \varepsilon(t - t_j^n) + \sum_{k,m} K_{ik}(\widehat{t}_k^m)\, \varepsilon(t - \widehat{t}_k^m), \tag{1}$$

where V_r is the reset potential after a spike (the voltage scale is chosen so that $V_r = 0$ in what follows), $J_{ij} = 0$ for $i = j$, and $\varepsilon(t)$ gives the time-course of an excitatory postsynaptic potential (EPSP); $\varepsilon(t)$ is also called the synaptic response function. The magnitude of the membrane potential determines the times t_i^n at which a postsynaptic neuron i will fire, i.e., the times $v(t_i^n) = \vartheta$ where ϑ is the spiking threshold. The firing times t_i^n of the postsynaptic neuron may, and in general will, depend on J_{ij} and K_{ik}. Once the neuron has fired, J_{ij} increases or decreases according to whether $t_j^{n'} - t_i^n < 0$ or > 0. More precisely, synaptic change is determined by the *learning window* W (Gerstner et al. 1996) through its value $W(t_j^{n'} - t_i^n)$. An example of such a learning window is illustrated in Fig. 2.

Given the input and output firing times, the change $\Delta J_{ij}(t) := J_{ij}(t) - J_{ij}(t - T_l)$ of the efficacy of synapse $\{ij\}$ (synaptic strength) during a learning session of duration T_l and ending at time t is governed by

$$\Delta J_{ij}(t) = \eta \left[\sum_{t - T_l \le t_j^{n'} < t} w^{\text{in}} + \sum_{t - T_l \le t_i^n < t} w^{\text{out}} + \sum_{t - T_l \le t_j^{n'},\, t_i^n < t} W(t_j^{n'} - t_i^n) \right], \tag{2}$$

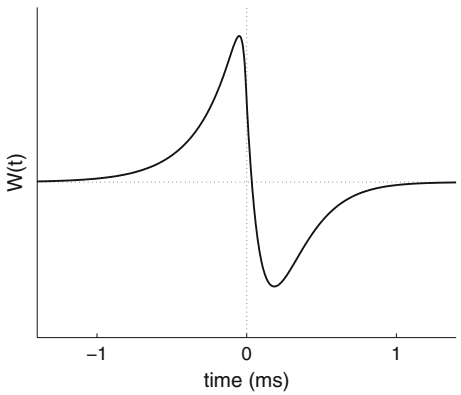

Fig. 2 Learning window $W(t)$ governing spike-timing-dependent synaptic plasticity (STDP). It is a function of the time difference s between pre- and postsynaptic spikes. For a generic excitatory synapse we have $W(s) > 0$ for $s < 0$, i.e., when a presynaptic spike arrives earlier than the postsynaptic one and contributes to spike generation, and $W(s) < 0$ for $s > 0$, i.e., those that come too late shall be punished. The time scale illustrated here is that of the barn owl (Kempter et al. 2001b)

where ηw^{in} and ηw^{out} are changes in the weight J_{ij} induced respectively by the arrival of a spike at neuron j or the generation of a spike by neuron i. The prefactor $0 < \eta \ll 1$ ensures explicitly that learning is slow on a neuronal time scale (milliseconds). Throughout what follows this condition is referred to as the *adiabatic hypothesis*. It holds in numerous biological situations and is a mainstay of the arguments below. The learning window W takes into account how pre- and postsynaptic *spikes* interact through temporal correlation, which is the new meaning of Hebbian learning, while w^{in} and w^{out} describe their respective effects separately. We now analyze three consequences of the adiabatic hypothesis $0 < \eta \ll 1$ in turn.

2.2 Input processes are self-averaging

Consider first the input processes generating the spike sequences $\{t_i^n, 1 \le i \le N\}$ and $\{\widehat{t}_k^m, 1 \le k \le M\}$. The spiking-rate is modelled as an inhomogeneous Poisson process (van Hemmen, 2001, Appendix B), viz.,

– There exists a density $\lambda_i(t)$ such that

$$\text{Prob}\{\text{one event in } [t, t+\Delta t)\} = \lambda_i(t)\Delta t\ . \tag{3}$$

– For $o(\Delta t)$ meaning $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$, we have

$$\text{Prob}\{\ge 2 \text{ events in } [t, t+\Delta t)\} = o(\Delta t)\ . \tag{4}$$

We note that this is somewhat analogous with neuronal refractoriness.

– Events in *disjoint* intervals are independent.

Because $0 < \eta \ll 1$, we can take T_l on the one hand so long that it greatly exceeds all neuronal time constants, including interspike intervals and the temporal width of the learning window W, and on the other hand so short that the J_{ij} have hardly changed. That is, T_l can be chosen so as to separate neuronal and synaptic timescales ($T_l \ll \frac{1}{\eta}$), a feature of the analysis that enables the present analytical treatment.

Spike generation is (nearly) always a *local* process in time and so are the $1 \le j \le N$ input processes generating the input spikes t_j^n from the recurrently connected neurons and $1 \le k \le M$ input processes generating input spikes $\widehat{t}_k^m$ from the external inputs.

Because of the independence of disjoint intervals, the sum in (2) is self-averaging over the randomness. The strong law of large numbers (Lamperti 1966) ensures that the average can be used in the sum (2), rather than a specific realization. This "ensemble average" is denoted by angular brackets $\langle\ldots\rangle$ and it comes for free (up to the limit that neurons within the network are weakly correlated, which can be considered realistic for more than 20 neurons). The error is of the order of the standard deviation and has a Gaussian distribution according to the central limit theorem (Lamperti 1966). This error is a "noise" that will not be discussed further here.

2.3 Time averaging

Introducing the spike flows associated with the neurons and the external inputs, respectively,

$$\begin{aligned} S_i(t) &= \sum_{t_i^n \le t} \delta(t - t_i^n), \quad i = 1, \ldots, N\ , \\ \widehat{S}_k(t) &= \sum_{\widehat{t}_k^m \le t} \delta(t - \widehat{t}_k^m), \quad k = 1, \ldots, M, \end{aligned} \tag{5}$$

where $\delta(t)$ is the Dirac delta function, we can rewrite (2)

$$\frac{\Delta J_{ij}(t)}{T_l} = \eta\Bigg\{\frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\,[w^{\text{in}}\langle S_j(t')\rangle + w^{\text{out}}\langle S_i(t')\rangle] + \frac{1}{T_l}\int_{t-T_l}^{t}\mathrm{d}t' \int_{t-T_l-t'}^{t-t'} \mathrm{d}u\, W(u)\langle S_i(t')S_j(t'+u)\rangle\Bigg\}. \tag{6}$$

Averages over the time scale of learning, T_l, are denoted by an overline $\overline{f(t)} := T_l^{-1}\int_{t-T_l}^{t}\mathrm{d}t'\,f(t')$. The mean spiking-rates are then defined as $\nu_i(t) := \overline{\langle S_i(t)\rangle}$. Averaging S_j over the randomness is a good approximation as T_l becomes large, which is a direct consequence of the strong law of large numbers (Lamperti 1966) and of the Poissonian events being independent in *disjoint* intervals. Note that the mean spiking-rates $\nu_i(t)$ are distinguished from the instantaneous spiking-rates $\lambda_i(t)$, which are just the ensemble average $\lambda_i(t) := \langle S_i(t)\rangle$. The mean spiking-rate is slowly varying and consequently is related to the instantaneous spiking-rate by $\nu_i(t) = \overline{\lambda_i(t)}$. The integrals in (6) depend on t, and the first and second terms in this equation can be substituted by $\nu_j(t)$ and $\nu_i(t)$.

The double integral in the last term in (6) explicitly correlates input and output, which is a distinguishing property of Hebbian learning. Let us consider a "typical" t', say, $t' = t - T_l + xT_l$ with $0 < x < 1$. Then the lower bound of the integral over u is effectively $-xT_l$, while the upper bound is $(1-x)T_l$. The learning window W is *local in time*, typically of the order of milliseconds for the auditory system and tens of milliseconds for most of the cortex, so that it is much shorter than T_l. Hence for a "typical" t' the lower bound of the integral over u can be replaced by $-\infty$, whereas the upper bound can be replaced by $+\infty$, so that up to a negligible error we are left with

$$\frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}u\, W(u)\langle S_i(t')S_j(t'+u)\rangle$$
$$= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u)\frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle S_i(t')S_j(t'+u)\rangle. \quad (7)$$

The key idea behind this is the method of averaging (Sanders and Verhulst 1985) as it is used in solving *non*autonomous differential equations. In agreement with the averaging philosophy we take the J_{ij} to be constant while evaluating the integrals in (7).

2.4 Deriving the learning equation

Due to the adiabatic hypothesis the change $\Delta J_{ij}(t) = J_{ij}(t) - J_{ij}(t - T_l)$ is still "small" so that we can replace $\Delta J_{ij}(t)/T_l$ in (6) by the differential quotient $\mathrm{d}J_{ij}/\mathrm{d}t$. Exploiting (7), we then obtain the *learning equation* (Kempter et al. 1999)

$$\frac{\mathrm{d}}{\mathrm{d}t}J_{ij}(t) = \eta\Bigg[w^{\mathrm{in}}\,\nu_j(t) + w^{\mathrm{out}}\,\nu_i(t) + \int_{-\infty}^{\infty} \mathrm{d}u W(u)\frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\langle S_i(t')S_j(t'+u)\rangle\Bigg]. \quad (8)$$

Except for the adiabatic hypothesis, which is a very weak assumption, the above equation is *universally valid* and exact.

It is a nice aspect of (8) that the final integral over t' is nothing but a time average of the correlation function $\langle S_i(t')\, S_j(t'')\rangle$. We may interpret it as the joint probability density of observing an input spike at the jth synapse of neuron i at time t'' and an output spike at time t'.

It is straightforward to derive the exactly analogous expression for the learning equation of the synapse K_{ik}, which connects neuron i with the kth external input, so as to get

$$\frac{\mathrm{d}}{\mathrm{d}t}K_{ik}(t) = \eta\Bigg[w^{\mathrm{in}}\,\widehat{\nu}_k(t) + w^{\mathrm{out}}\,\nu_i(t) + \int_{-\infty}^{\infty} \mathrm{d}u W(u)\frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\langle S_i(t')\widehat{S}_k(t'+u)\rangle\Bigg]. \quad (9)$$

The difficulty in evaluating these expressions is that the flow of spikes $S_i(t)$ from the neurons in the network depends upon the spikes on the inputs $\widehat{S}_k(t)$.

3 Learning dynamics with the Poisson neuron

The above highly nonlinear equations (8) and (9) that describe synaptic evolution can not in general be solved exactly. For a recurrent network such an exact solution would be highly desirable. We therefore take the *Poisson neuron* described by an inhomogeneous Poisson process with rate function, or intensity, $\rho_i(t) = \nu_0 + \nu_i(t) \geq 0$ with the membrane potential $\nu_i(t)$ given by (1) and ν_0 chosen so that $\rho_i(t) \geq 0$. Here we discuss the ensuing learning dynamics.

3.1 Poisson neuron

The process of generating an output spike is highly nonlinear since the weights (J_{ij}, K_{ik}) appear in a number of places, so that evaluating the correlation function or solving the system of differential equations (8) is in general not possible analytically. This is, however, not the case if we model spike generation by means of the *Poisson neuron* (Kempter et al. 1998), an inhomogeneous Poisson process with rate function, or intensity,

$$\rho_i(t) = \nu_0 + \nu_i(t) = \nu_0 + \sum_{j,n} J_{ij}(t_j^n)\varepsilon(t - t_j^n) + \sum_{k,m} K_{ik}(\widehat{t}_k^m)\varepsilon(t - \widehat{t}_k^m) \geq 0\,. \quad (10)$$

Here ν_0 is a spontaneous spiking-rate making the right-hand side positive, if necessary. The sum over $j = 1, \ldots, N$ is over the recurrent weights (i.e., the weights connecting the N neurons), and the sum over $k = 1, \ldots, M$ is over feed-forward weights (i.e., from the M external inputs). Though $\rho_i(t)$ is linear in the postsynaptic potential v_i, spike generation is not. It is just a point process assigning high spiking probability to times t with large values of $v_i(t)$, as it ought to. To obtain an exact solution, one may alternatively use $\rho_i^{\mathrm{clipped}}(t) = \nu_1\Theta[v_i(t) - \vartheta_1]$ (Kistler and van Hemmen 2000), where Θ is the Heaviside step function, $\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x < 0$.

3.2 Output process for Poisson neuron

Because of causality, $\varepsilon(t) = 0$ for $t < 0$. In addition, we take its integral $\int \mathrm{d}t\, \varepsilon(t) = 1$. The stochastic approximation entails that for the ith Poisson neuron, $\rho_i(t) = \lambda_i(t) =$

$\langle S_i(t)\rangle$, where the latter average is over the stochastic input processes. That is, using (10) we obtain the following expression for the spiking-rates of the neurons:

$$\begin{aligned}\langle S_i(t)\rangle &= \nu_0 + \sum_{j=1}^{N} J_{ij}(t) \int_0^\infty \mathrm{d}u\, \varepsilon(u)\lambda_j(t-u), \\ &\quad + \sum_{k=1}^{M} K_{ik}(t) \int_0^\infty \mathrm{d}u\, \varepsilon(u)\widehat{\lambda}_k(t-u), \qquad (11)\\ \text{i.e.,}\quad \lambda_i(t) &= \nu_0 + \sum_{j=1}^{N} J_{ij}(t)\Lambda_j(t) + \sum_{k=1}^{M} K_{ik}(t)\widehat{\Lambda}_k(t)\ ,\end{aligned}$$

where $\widehat{\lambda}_m(t)$ is the spiking-rate of the mth external synaptic input and

$$\begin{aligned}\Lambda_i(t) &= \int_0^\infty \mathrm{d}u\, \varepsilon(u)\lambda_i(t-u), \quad i = 1,\ldots,N, \\ \widehat{\Lambda}_k(t) &= \int_0^\infty \mathrm{d}u\, \varepsilon(u)\widehat{\lambda}_k(t-u), \quad k = 1,\ldots,M,\end{aligned} \qquad (12)$$

which can be written as the convolutions $\Lambda_i(t) = (\varepsilon * \lambda_i)(t)$ and $\widehat{\Lambda}_k(t) = (\varepsilon * \widehat{\lambda}_k)(t)$. The derivation of (11), which involves averaging over the Poisson distribution of spike times, follows exactly that given by van Hemmen (2001). Previous studies, in which recurrent connections were not considered, did not contain any dependence upon $\lambda_i(t)$ on the right of Eq. (11).

3.3 Recurrent spiking-rates for Poisson neuron

The central difficulty in analyzing recurrent networks is that the spiking-rates of the neurons, $\lambda_i(t)$, are a function of the recurrent weights, $J_{ij}(t)$. Since the weights vary on a much slower timescale than the network activation, they can be considered quasi-constant (due to the adiabatic hypothesis).

For quasi-constant weights an explicit expression for $\lambda_i(t)$ as a function of ν_0 and the $\widehat{\lambda}_k(t)$ may be obtained using the Laplace transform $f_{\mathrm{L}}(s) := \mathcal{L}\{f(t)\} = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st} f(t)$. Hence

$$\Lambda_{\mathrm{L}_i}(s) = \frac{\nu_0}{s} + \varepsilon_{\mathrm{L}}(s)\left[\lambda_{\mathrm{L}_i}(s) - \frac{\nu_0}{s}\right], \qquad (13)$$

since $\lambda_i(t) = \nu_0$ for $t < 0$. There is a similar expression for $\widehat{\Lambda}_{\mathrm{L}_k}(s)$. Consequently we solve the self-consistency equation (11) as

$$\begin{aligned}\lambda_i(t) &= \mathcal{L}^{-1}\{\lambda_{\mathrm{L}_i}(s)\}, \\ \lambda_{\mathrm{L}_i}(s) &= \sum_{j=1}^{N} [\mathbb{1}_N - J\varepsilon_{\mathrm{L}}(s)]^{-1}_{ij} \left\{ E_j \frac{\nu_0}{s} \right. \\ &\quad + \sum_{j'=1}^{N} J_{jj'} E_{j'} \frac{\nu_0}{s}\,[1-\varepsilon(s)] \\ &\quad \left. + \sum_{k=1}^{M} K_{jk}\left[\frac{\nu_0}{s} + \varepsilon(s)\right]\left[\widehat{\lambda}_{\mathrm{L}_k}(s) - \widehat{E}_k \frac{\nu_0}{s}\right]\right\}\ ,\end{aligned} \qquad (14)$$

where $\mathbb{1}_N$ is the $N \times N$ identity matrix, E is the N-vector $(1,\ldots,1)$, $\widehat{E}$ is the M-vector $(1,\ldots,1)$, and $\mathcal{L}^{-1}\{\}$ is the inverse Laplace transform. This expression (14) gives the explicit dependence of the output spiking-rate of each neuron in terms of the external synaptic inputs, the spontaneous spiking-rate, and the weights in the network, provided that $[\mathbb{1}_N - J\varepsilon_{\mathrm{L}}(s)]$ is invertible. This expression is crucial for determining the weight dependence of the learning equation in the general case, such as when the external input is oscillatory. However it can be simplified considerably for the case in which the spiking-rates are constant, as discussed in Sect. 3.6.

3.4 Spike-time correlations for Poisson neuron

It now remains to calculate the correlation function $\langle S_i(t)\, S_j(t+u)\rangle$. In order to do this, it is useful to define the following correlations:

$$\begin{aligned}Q_{ij}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle S_i(t')\, S_j(t'+u)\rangle, \\ D_{ik}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle S_i(t')\, \widehat{S}_k(t'+u)\rangle, \\ \widehat{Q}_{kl}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle \widehat{S}_k(t')\, \widehat{S}_l(t'+u)\rangle, \qquad (15)\\ R_{ij}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle (\varepsilon * S_i)(t')\, S_j(t'+u)\rangle, \\ F_{ik}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle (\varepsilon * S_i)(t')\, \widehat{S}_k(t'+u)\rangle, \\ \widehat{F}_{ik}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle S_i(t')\, (\varepsilon * \widehat{S}_k)(t'+u)\rangle, \\ \widehat{R}_{kl}(t,u) &:= \frac{1}{T_l}\int_{t-T_l}^{t} \mathrm{d}t'\, \langle (\varepsilon * \widehat{S}_k)(t')\, \widehat{S}_l(t'+u)\rangle.\end{aligned}$$

We obtain

$$\langle S_i(t)S_j(t+u)\rangle = \left\langle \left[\nu_0 + \sum_{j'=1}^{N} J_{ij'}(t)(\varepsilon * S_{j'})(t) + \sum_{k=1}^{M} K_{ik}(t)(\varepsilon * \widehat{S}_k)(t) \right] S_j(t+u) \right\rangle , \tag{16}$$

and hence,

$$Q_{ij}(t,u) = \nu_0 \nu_j(t+u) + \sum_{j'=1}^{N} J_{ij'}(t) R_{j'j}(t,u) + \sum_{k=1}^{M} K_{ik}(t)\widehat{F}_{jk}(t+u,-u) + \delta_{ij}\delta(u)\nu_j(t), \tag{17}$$

where δ_{ij} is the Kronecker delta function. The final term on the RHS results from the autocorrelation of the Poisson process (Hawkes 1971) and we neglect a term $J_{ij}(t)\varepsilon(u)\nu_j(t)$ due to the spike-triggering effect. Likewise

$$D_{ik}(t,u) = \nu_0 \widehat{\nu}_k(t+u) + \sum_{j=1}^{N} J_{ij}(t) F_{jk}(t,u) + \sum_{k'=1}^{M} K_{ik'}(t)\widehat{R}_{k'k}(t,u)\ . \tag{18}$$

This may conveniently be written in matrix notation as

$$\begin{aligned} Q(t,u) &= \nu_0 E \nu^T(t) + J(t)R(t,u) + K(t)\widehat{F}^T(t+u,-u) \\ &\quad + \delta(u)\mathrm{diag}\,[\nu(t)]\,, \\ D(t,u) &= \nu_0 E \widehat{\nu}^T(t) + J(t)F(t,u) + K(t)\widehat{R}(t,u)\ , \end{aligned} \tag{19}$$

where E is defined following (14), $\mathrm{diag}(X)$ is the diagonal matrix with the vector X on the diagonal and zero elsewhere, and the superscript T denotes transposition. Because of the compact support of $W(u)$ and large separation of time scales $T_l \gg u$, we have used $\nu(t+u) = \nu(t)$ and $\widehat{F}(t+u,-u) = \widehat{F}(t,-u)$.

For the learning equations (8) and (9) we need to define the following integrals of these functions over the learning window:

$$\begin{aligned} Q^W_{ij}(t) &:= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u)\, Q_{ij}(t,u), \\ Q^V_{ij}(t) &:= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u)\, Q_{ij}(t,-u), \end{aligned} \tag{20}$$

and likewise for each of the other variables on the LHS of (15) (note that the superscript V involves the argument $-u$ in the last term of the integrand). The expressions (19) then give

$$\begin{aligned} Q^W(t) &= \widetilde{W}\,\nu_0 E \nu^T(t) + J(t)R^W(t) + K(t)\widehat{F}^{V^T}(t), \\ D^W(t) &= \widetilde{W}\,\nu_0 E \widehat{\nu}^T(t) + J(t)F^W(t) + K(t)\widehat{R}^W(t), \end{aligned} \tag{21}$$

where the tilde denotes integration over time $\widetilde{W} := \int_{-\infty}^{\infty} \mathrm{d}t\, W(t)$ and the autocorrelation term in (17) vanishes since we choose the learning window $W(t)$ such that $W(0) = 0$. The functions $R^W_{ij}(t)$, $F^W_{ik}(t)$, $\widehat{F}^V_{ik}(t)$, and $\widehat{R}^W_{kl}(t)$ are related to the functions $Q_{ij}(t)$, $D_{ik}(t)$, and $\widehat{Q}_{kl}(t)$ by

$$\begin{aligned} R^W_{ij}(t) &= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u) \int_0^{\infty} \mathrm{d}r\, \varepsilon(r)\, Q_{ij}(t,u+r), \\ F^W_{ik}(t) &= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u) \int_0^{\infty} \mathrm{d}r\, \varepsilon(r)\, D_{ik}(t,u+r), \\ \widehat{F}^V_{ik}(t) &= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u) \int_0^{\infty} \mathrm{d}r\, \varepsilon(r)\, D_{ik}(t,-u+r), \\ \widehat{R}^W_{kl}(t) &= \int_{-\infty}^{\infty} \mathrm{d}u\, W(u) \int_0^{\infty} \mathrm{d}r\, \varepsilon(r)\, \widehat{Q}_{kl}(t,u+r)\ . \end{aligned} \tag{22}$$

These expressions (21) and (22) allow all the spike-timing cross-correlation functions to be solved in terms of the spike-timing cross-correlation of the external inputs $\widehat{Q}^W_{kl}(t)$.

3.5 The learning equations in matrix notation

To obtain the learning equations for the weights $J_{ij}(t)$ and $K_{ik}(t)$ we now substitute the above results into the learning equations (8) and (9).

A problem with using matrix notation here is that the weights J_{ij} of the missing diagonal connections must remain constant, since it is forbidden for a neuron to be connected onto itself, i.e., all the J_{ii} must remain zero and $\frac{\mathrm{d}}{\mathrm{d}t}J(t)$ must also remain zero on the diagonal. To achieve this, we use projectors on the matrix space in which J belongs: Φ_J is the projector that operates on $N \times N$ matrices and that forces the matrix elements corresponding to the missing connections to zero. For example, consider a network of $N = 3$ neurons that is fully connected except for self-connections

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix}$$

Such projection matrices could also be used for networks with a more complex pattern of missing connections (e.g., the above network with a missing connection from neuron #1 to neuron #2 would have associated with it a different projection matrix that resulted in the term a_{21} on the RHS of the above equation being replaced by zero). Although we have considered here a fully connected recurrent network, it is possible using such projection matrices Φ_J and Φ_K to model any architecture of individual connections within the network by this framework.

Consequently, using the above matrix notation the learning equation (8) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}J(t) = \Phi_J \left[w^{\mathrm{in}} E \nu^T(t) + w^{\mathrm{out}} \nu(t)\, E^T + Q^W(t) \right], \tag{23}$$

where time has been rescaled by a factor of η. Likewise

$$\frac{\mathrm{d}}{\mathrm{d}t}K(t) = w^{\mathrm{in}}\, E\, \widehat{\nu}^T(t) + w^{\mathrm{out}}\, \nu(t)\, \widehat{E}^T + D^W(t)\,, \tag{24}$$

where $\widehat{E}$ is the M-vector (1,…,1).

An interesting consequence is that the equilibria of the system are defined for null images of the projectors, i.e. when their argument is in their null space. Note that the null space (or kernel) Ker Φ_J is a vector subspace of $\mathbb{R}^{N\times N}$, and the dimension increases when the number of synaptic connections decreases. Hence, the more missing connections there are in the network, the richer the equilibria may be; reformulated, sparsely connected networks may be more interesting (with more possible equilibria) than fully connected networks that contain the same number of neurons.

3.6 The case of constant external spiking-rates

We now consider the situation where the spiking-rates of the external synaptic inputs, $\widehat{\nu}_k(t)$, are quasi-constant. Then the spiking-rates of the recurrently connected neurons, $\lambda_i(t)$, can be considered constant over the timescale of T_l, but they vary over the timescale of changes in the synaptic weights (cf. Sect. 2.2). Thus we can neglect transients in the dynamics, so that $\lambda_i(t) = \nu_i(t) = \Lambda_i(t)$ and the expression (11) for the spiking-rates becomes

$$\nu_i(t) = \nu_0 + \sum_{j=1}^{N} J_{ij}(t)\nu_j(t) + \sum_{k=1}^{M} K_{ik}(t)\widehat{\nu}_k, \quad i = 1,\ldots,N, \tag{25}$$

since the response kernel $\varepsilon(t)$ is normalized. The solution, which includes the recurrent connections, is given by

$$\sum_{j=1}^{N} \left[\delta_{ij} - J_{ij}(t)\right] \nu_j(t) = \nu_0 E_i + \sum_{k=1}^{M} K_{ik}(t)\widehat{\nu}_k\,, \tag{26}$$

which can be written in matrix notation as

$$\nu(t) = [\mathbb{1}_N - J(t)]^{-1} [\nu_0 E + K(t)\widehat{\nu}]\ , \tag{27}$$

provided that the matrix $[\mathbb{1}_N - J(t)]$ is invertible.

This constant spiking-rate approximation also results in a considerable simplification of the expression for the spike-timing cross-correlations (15) so that $Q^W_{ij}(t) = R^W_{ij}(t)$, $D^W_{ik}(t) = F^W_{ik}(t)$, $\widehat{Q}^W_{ij}(t) = \widehat{R}^W_{ij}(t)$ and $D^V_{ik}(t) = \widehat{F}^V_{ik}(t)$. The self-consistency relations (21) for the case of quasi-constant spiking-rates can be written in matrix notation as

$$\begin{aligned}
Q^W(t) &= \widetilde{W}\,\nu_0\, E\, \nu^T(t) + J(t)\, Q^W(t) + K(t)\, D^{V^T}(t),\\
D^W(t) &= \widetilde{W}\,\nu_0\, E\, \widehat{\nu}^T(t) + J(t)\, D^W(t) + K(t)\, \widehat{Q}^W(t),\\
D^V(t) &= \widetilde{W}\,\nu_0\, E\, \widehat{\nu}^T(t) + J(t)\, D^V(t) + K(t)\, \widehat{Q}^V(t)\,,
\end{aligned} \tag{28}$$

where we have included the expression for $D^V(t)$ in order to provide a closed system of equations (recall that $\widehat{Q}^W(t)$ and $\widehat{Q}^V(t)$ are determined by the external input). Provided again that the matrix $[\mathbb{1}_N - J(t)]$ is invertible, we have

$$\begin{aligned}
Q^W(t) &= [\mathbb{1}_N - J(t)]^{-1} \left[\widetilde{W}\nu_0 E\, \nu^T(t) + K(t)\, D^{V^T}(t) \right],\\
D^W(t) &= [\mathbb{1}_N - J(t)]^{-1} \left[\widetilde{W}\nu_0 E \widehat{\nu}^T(t) + K(t)\, \widehat{Q}^W(t) \right],\\
D^V(t) &= [\mathbb{1}_N - J(t)]^{-1} \left[\widetilde{W}\nu_0 E \widehat{\nu}^T(t) + K(t)\, \widehat{Q}^V(t) \right].
\end{aligned} \tag{29}$$

4 Dynamical system that characterizes the network activity

Consequently when the external spiking-rates are constant we obtain a system of six coupled matrix equations governing both the network activity and the time evolution of the recurrent synaptic weights J_{ij} and the input synaptic weights K_{ik}: The network activity is given by the self-consistency condition on the firing rates (26), the three network correlation self-consistency conditions (29), and the two learning equations (23) and (24). We note that in the situation where there are no recurrent weights ($J_{ij} = 0$) the equation for the feed-forward weights K_{ik} is exactly that given in earlier studies (Kempter et al. 1999; van Hemmen 2001).

An important issue is the invertibility of the matrix $[\mathbb{1}_N - J(t)]$. If we discard the case of no inputs and no spontaneous firing rates, the equation

$$[\mathbb{1}_N - J(t)]\ \nu(t) = \nu_0\, E + K(t)\, \widehat{\nu}(t) \tag{30}$$

implies that the norm of $\nu(t)$ diverges to $+\infty$ when $[\mathbb{1}_N - J(t)]$ tends to a non-invertible matrix. Thus, if we start from an invertible matrix (e.g., with suitable constant weights) and we discard the case of diverging activity (note

that in such a case many of the approximations we used would turn out to be invalid), then the matrix will remain invertible.

We recapitulate the system of four consistency equations (firing rates and correlations) and two learning differential equations using matrix notation:

$$
\begin{aligned}
\nu(t) &= [\mathbb{1}_N - J(t)]^{-1}\left[\nu_0\, E + K(t)\,\widehat{\nu}(t)\right], \qquad (31)\\
D^W(t) &= [\mathbb{1}_N - J(t)]^{-1}\left[\widetilde{W}\,\nu_0\, E\,\widehat{\nu}^T(t) + K(t)\,\widehat{Q}^W(t)\right],\\
D^V(t) &= [\mathbb{1}_N - J(t)]^{-1}\left[\widetilde{W}\,\nu_0\, E\,\widehat{\nu}^T(t) + K(t)\,\widehat{Q}^V(t)\right],\\
Q^W(t) &= [\mathbb{1}_N - J(t)]^{-1}\left[\widetilde{W}\nu_0 E\nu^T(t) + K(t)D^{V^T}(t)\right],\\
\frac{\mathrm{d}}{\mathrm{d}t}K(t) &= \left[w^{\mathrm{in}}\,E\,\widehat{\nu}^T(t) + w^{\mathrm{out}}\,\nu(t)\,\widehat{E}^T + D^W(t)\right],\\
\frac{\mathrm{d}}{\mathrm{d}t}J(t) &= \Phi_J\left[w^{\mathrm{in}}\,E\,\nu^T(t) + w^{\mathrm{out}}\,\nu(t)\,E^T + Q^W(t)\right].
\end{aligned}
$$

Note that time has been rescaled to eliminate η.

Now, using the four consistency equations we can express the the two learning equations as coupled differential equations in the weights J and K in terms of only the external inputs

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}K(t) &= \Big\{w^{\mathrm{in}}\,E\,\widehat{\nu}^T(t) + w^{\mathrm{out}}\,[\mathbb{1}_N - J(t)]^{-1}\\
&\quad\times[\nu_0\,E + K(t)\,\widehat{\nu}(t)]\;\widehat{E}^T + [\mathbb{1}_N - J(t)]^{-1}\\
&\quad\times\left[\widetilde{W}\,\nu_0\,E\,\widehat{\nu}^T(t) + K(t)\,\widehat{Q}^W(t)\right]\Big\}, \qquad (32)\\
\frac{\mathrm{d}}{\mathrm{d}t}J(t) &= \Phi_J\Big(w^{\mathrm{in}}E\,[\nu_0 E + K(t)\widehat{\nu}(t)]^T\,[\mathbb{1}_N - J(t)]^{-1\,T}\\
&\quad + w^{\mathrm{out}}\,[\mathbb{1}_N - J(t)]^{-1}\,[\nu_0\,E + K(t)\,\widehat{\nu}(t)]\;E^T\\
&\quad + [\mathbb{1}_N - J(t)]^{-1}\,\Big\{\widetilde{W}\,\nu_0\,E\,\Big[\nu_0\,E\,E^T\\
&\quad + E\,\widehat{\nu}^T(t)\,K^T(t) + K(t)\,\widehat{\nu}(t)\,E^T\Big]\\
&\quad + K(t)\,\widehat{Q}^{V^T}(t)\,K^T(t)\Big\}\,[\mathbb{1}_N - J(t)]^{-1\,T}\Big).
\end{aligned}
$$

The dynamics of the network is described entirely by these two coupled differential matrix equations.

5 Solution of the learning equations with no external synaptic input

In order to illustrate how the above equations describe the learning dynamics, we present here the solution in the situation where there is no external synaptic input. In this case the network is driven entirely by the spontaneous spiking-rate $\nu_0 > 0$ of the recurrently connected neurons. This reduces the complexity of the calculation considerably, since we have only three simultaneous equations to solve.

We can compute ν and Q in terms of J (the dependence in t is implicit here)

$$
\begin{aligned}
\nu &= (\mathbb{1}_N - J)^{-1}\,\nu_0\,E,\\
Q &= (\mathbb{1}_N - J)^{-1}\;\widetilde{W}\,\nu_0\,E\,\nu^T = \widetilde{W}\,\nu\,\nu^T\,,
\end{aligned}
\qquad (33)
$$

where we henceforth drop the superscript W. The above relations (33) lead to the learning equation in terms of ν alone

$$
\frac{\mathrm{d}}{\mathrm{d}t}J = \Phi_J\left(w^{\mathrm{in}}\,E\,\nu^T + w^{\mathrm{out}}\,\nu\,E^T + \widetilde{W}\,\nu\,\nu^T\right)\,, \qquad (34)
$$

where Φ_J is the projection matrix described in Sect. 3.5 that nullifies the diagonal terms.

Consider now the difference $J_{ij} - J_{ji}$ $(i \neq j)$ at the stationary solution of (34);

$$
\frac{\mathrm{d}}{\mathrm{d}t}\left(J^*_{ij} - J^*_{ji}\right) = (w^{\mathrm{in}} - w^{\mathrm{out}})(\nu^*_j - \nu^*_i) = 0\,, \qquad (35)
$$

where the stationary solution is denoted by an asterisk. We deduce that ν is homogeneous over the network at the equilibrium, provided $w^{\mathrm{in}} \neq w^{\mathrm{out}}$, with

$$
\nu^* = \mu\,E \quad\text{and}\quad \mu := -\frac{(w^{\mathrm{in}} + w^{\mathrm{out}})}{\widetilde{W}}\,, \qquad (36)
$$

which requires that $\mu \geq 0$. A fixed-point of the matrix system is given by

$$
\begin{aligned}
\nu^* &= \mu\,E,\\
Q^* &= \widetilde{W}\,\mu^2\,E\,E^T,\\
\left(\mathbb{1}_N - J^*\right)\,E &= \frac{\nu_0}{\mu}\,E\,.
\end{aligned}
\qquad (37)
$$

The space of solutions for J is defined by the last equation of (37) and corresponds to the invertible matrices of $\mathcal{M}_N$ for which the column vector E as eigenvector for the eigenvalue ν_0/μ. Here $\mathcal{M}_N$ is the linear subspace of matrices of $\mathbb{R}^{N\times N}$ with zeros on the diagonal ($\mathcal{M}_N$ has dimension $N(N-1)$).

For stability of the spiking-rates, all the eigenvalues of J should be in [0, 1) so that the spiking-rates remain bounded, which implies a condition on the learning parameters: $\nu_0 < \mu$.

We now assess the stability of the homogeneous solution of the spiking-rates and the speed of convergence towards the fixed-point. We define the mean weight over the neurons $J_{\mathrm{av}}(t) := [N(N-1)]^{-1}\sum_{i\neq j} J_{ij}(t)$, mean correlation $Q_{\mathrm{av}}(t) := N^{-2}\sum_{i,j} Q_{ij}(t)$, and mean spiking-rate $\nu_{\mathrm{av}}(t) := N^{-1}\sum_i \nu_i(t)$. The approach here is to solve the equations for the homogeneous spiking-rate situation and find the conditions under which the resulting solution is stable.

The equations for the mean network activity and correlation (33) become

$$
\begin{aligned}
[1 - (N-1)J_{\mathrm{av}}]\,\nu_{\mathrm{av}} &= \nu_0,\\
[1 - (N-1)J_{\mathrm{av}}]\,Q_{\mathrm{av}} &= \widetilde{W}\,\nu_0\,\nu_{\mathrm{av}}\,.
\end{aligned}
\qquad (38)
$$

The stationary solution of the learning equation (23) becomes

$$(w^{\text{in}} + w^{\text{out}})\,\nu_{\text{av}} + Q_{\text{av}} = 0\,. \tag{39}$$

This gives the solution

$$\begin{aligned} \nu^*_{\text{av}} &= \mu, \\ Q^*_{\text{av}} &= \widetilde{W}\,\mu^2, \\ J^*_{\text{av}} &= \frac{\mu - \nu_0}{(N-1)\,\mu} \end{aligned} \tag{40}$$

with μ given in (36). The homogeneous fixed-point spiking-rate is independent of ν_0, but this homogeneous solution exists only for $\mu \geq \nu_0$. Note that all the solutions for J^* (37) have the properties of (40).

The stability is obtained by expanding the learning equation around the fixed-point $\nu_{\text{av}} = \nu^*_{\text{av}} + \Delta\nu_{\text{av}}$ and $J_{\text{av}} = J^*_{\text{av}} + \Delta J_{\text{av}}$, where

$$\Delta\nu_{\text{av}} = \frac{(N-1)\,\mu}{[1-(N-1)\,J^*_{\text{av}}]}\Delta\,J_{\text{av}} + \text{o}(\Delta J_{\text{av}})\,. \tag{41}$$

Consequently the learning equation (34) gives

$$\begin{aligned} \frac{\text{d}}{\text{d}t}\Delta J_{\text{av}} &= (w^{\text{in}} + w^{\text{out}} + 2\mu\widetilde{W})\Delta\nu_{\text{av}} + \text{o}(\Delta\nu_{\text{av}}) \\ &= \frac{(N-1)\,\mu^2\,\widetilde{W}}{[1-(N-1)J^*_{\text{av}}]}\,\Delta J_{\text{av}} + \text{o}(\Delta J_{\text{av}})\,. \end{aligned} \tag{42}$$

Therefore the mean weight is stable for $\widetilde{W} < 0$, in agreement with the corresponding condition on the stability for a single neuron with STDP (Song et al. 2000). We can gain further insight into the stability from the full matrix analysis of the variation of the weights $\Delta J(t) = \left[J(t) - J^*\right]$ about their fixed-point $J^* = J^*_{\text{av}}\,\Phi_J\,\left(E\,E^T\right)$:

$$\begin{aligned} \frac{\text{d}}{\text{d}t}\Delta J(t) &= \mathbb{L}\,[\Delta J(t)] + \text{o}\,[\Delta J(t)]\,, \\ \mathbb{L}\,[\Delta J(t)] &= -\nu_0\,\Phi_J\Big[(\mathbb{1} - J^*)^{-1}\Big(w^{\text{in}}\,\Delta J\,E\,E^T \\ &\qquad + w^{\text{out}}\,E\,E^T\,\Delta J^T\Big)\,(\mathbb{1} - J^*)^{-1^T}\Big], \end{aligned} \tag{43}$$

where the matrix operator $\mathbb{L}$ operates on $\mathcal{M}_N$. The eigenvalues $(\lambda_0, \lambda_1, \lambda_2)$ of the operator $\mathbb{L}$ for the homogeneous fixed-point are (see Appendix A)

$$\begin{aligned} \lambda_0 &= 0, \\ \lambda_1 &= -\frac{\mu^2\,(N-1)\,\left[w^{\text{in}}(N-1) - w^{\text{out}}\right]}{N\mu - \nu_0}, \\ \lambda_2 &= -\frac{\mu^2\,(N-1)\,(w^{\text{in}} + w^{\text{out}})}{\nu_0}, \end{aligned} \tag{44}$$

which have multiplicities of $N(N-2)$, $(N-1)$ and 1, respectively. This structure is similar to that for other fixed-points (the eigenvalues and stability of the fixed-point manifold are discussed in Appendix B). Stability of any fixed-point requires $w^{\text{in}} > 0$ (required by $\lambda_1 < 0$) and $(w^{\text{in}} + w^{\text{out}}) > 0$ (required by $\lambda_2 < 0$) in the large N limit. The convergence of the mean variables J^*_{av}, ν^*_{av}, Q^*_{av} is a consequence of the convergence towards the fixed-point manifold. The high multiplicity zero eigenvalue in the spectrum of $\mathbb{L}$ reflects the fact that the space of fixed-points is a continuum within which there is no constraint upon J. In the case where λ_1 and λ_2 are negative the whole manifold acts like an attractor within the space $\mathcal{M}_N$. Note that the condition $(w^{\text{in}} + w^{\text{out}}) > 0$ is equivalent to the condition for the homeostatic stability $\widetilde{W}^2 < 0$, from the definition of μ (36). If the condition $w^{\text{in}} > 0$ is not satisfied then all the fixed-points will be saddle-points, but the homeostatic equilibrium is still satisfied. This would introduce some additional deterministic variance to the evolution of the individual weights.

The calculation of the variance of the distribution of weights follows exactly analogously to that performed by Kempter et al. (1999). As in their analysis, the variance initially grows linearly in time with coefficient $\mathcal{D}$ when the initial weight distribution is a delta-function, but there is a nonlinear contribution that eventually dominates. The variance of a single weight $J_{ij}(t)$ may be calculated as var $J_{ij}(t) := \langle J^2_{ij}\rangle(t) - \langle J_i\rangle^2(t)$ as a function of time (the angular brackets denote an ensemble average; cf. Sec. 2.2). The calculation of var$J_{ij}(t)$, starting from some weight $J_{ij}(t_0)$ at time t_0, follows exactly that given in (Kempter et al. 1999), with the result

$$\begin{aligned} \text{var}\;J_{ij}(t) &= (t - t_0)\,\mathcal{D} \quad \text{for } (t - t_0) \gg \mathcal{W}, \\ \mathcal{D} &= \mu\,\left[(w^{\text{in}})^2 + (w^{\text{out}})^2\right] + \mu^2\,\widetilde{W}^2\,, \end{aligned} \tag{45}$$

where $\mathcal{W}$ is the width of the learning window W. Consequently, each weight $J_{ij}(t)$ essentially undergoes a diffusion process with diffusion constant $\mathcal{D}$.

The results of numerical simulations confirm the above analysis. Simulation results of ν_{av} and J_{av} as a function of time are shown in Fig. 3 (solid lines) for four different initial sets of weights. The simulations indicate that the fixed-points ν^*_{av} and J^*_{av} given in (40) and indicated in the plots by the dotted line, are reached asymptotically and remain stable. The dashed lines associated with each solid line are a fit to the curves with a single exponential (see Appendix C)

$$\begin{aligned} J_{\text{av}}(t) &= J^*_{\text{av}} + (J_{\text{av}}(0) - J^*_{\text{av}})\;\text{e}^{-t/\tau_J}, \\ \tau_J &= \frac{\nu_0\,\widetilde{W}^2}{(N-1)\,(w^{\text{in}} + w^{\text{out}})^3}\,. \end{aligned} \tag{46}$$

The results indicate that the expression (41) and (42) provide a very good description of the dynamics of the average weight and spiking-rate near the fixed-point.

The individual weights, however, tend to evolve to a bimodal distribution at the maximum and minimum allowed,

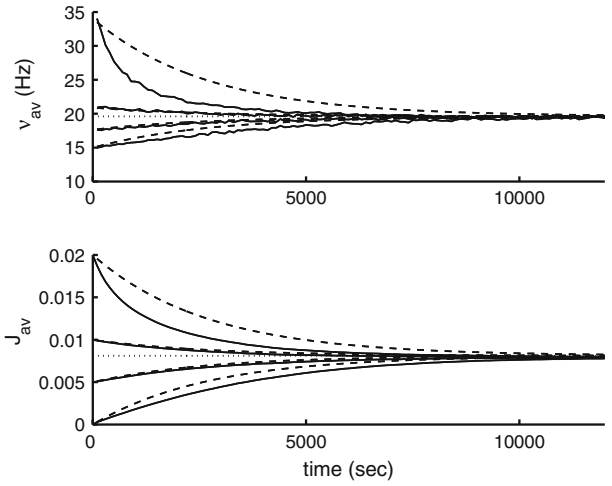

Fig. 3 Plots of the mean network spiking-rate $\nu_{\rm av}$ (upper plot) and weight $J_{\rm av}$ (lower plot) as a function of time for four different initial sets of weights. The theoretical value (40) for the fixed-point values $\nu^*_{\rm av}$ and $J^*_{\rm av}$ are shown by the dotted lines. The single exponential fit (46) is shown by the dotted lines (partially obscured by the solid lines for initial conditions near the fixed-point). Parameter values are N=30, $\eta = 10^{-7}$, $w^{\rm in} = 2$, $w^{\rm out} = 3$, and $\nu_0 = 15\,$Hz. The learning window $W(t)$ is given by $W(t) = c_{\rm D}{\rm e}^{-t/\tau_{\rm D}}$ for $t \geq 0$ and $W(t) = c_{\rm P}{\rm e}^{t/\tau_{\rm P}}$ for $t < 0$, with $c_{\rm D} = -10$, $c_{\rm P} = 5$, $\tau_{\rm D} = 34\,$ms, $\tau_{\rm P} = 17\,$ms. The time-step for the simulation is $10^{-4}\,$s and the time axis is given in seconds

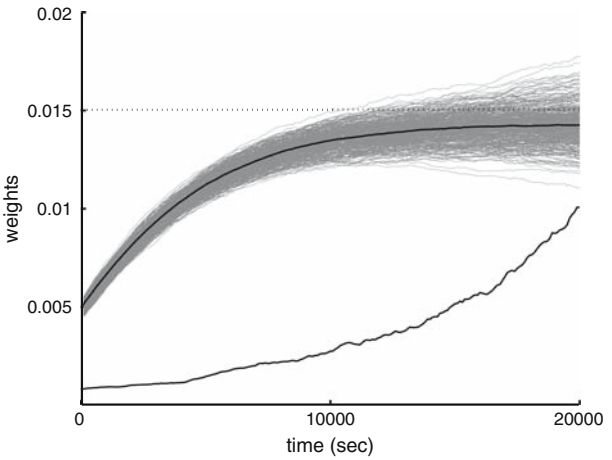

Fig. 4 Plot of the evolution of all the weights in a fully connected recurrent network of 20 neurons. The *thick line* is the mean of the weights, which is surrounded by the bundle of all the individual weight plots, which are in *grey*. The theoretical equilibrium value $J^*_{\rm av}$ is given by the *dotted line*. The *lower line* is a plot of the variance (multiplied by 10^4). The initial distribution of the weights is flat over an interval ± 0.0005 around its mean. Other parameter values are the same as in Fig. 3

although this does not affect the convergence of the average weight discussed above. Consequently "hard" bounds are required on the weights, as in the case of feedforward network connectivity (Kempter et al. 1999). The divergence of

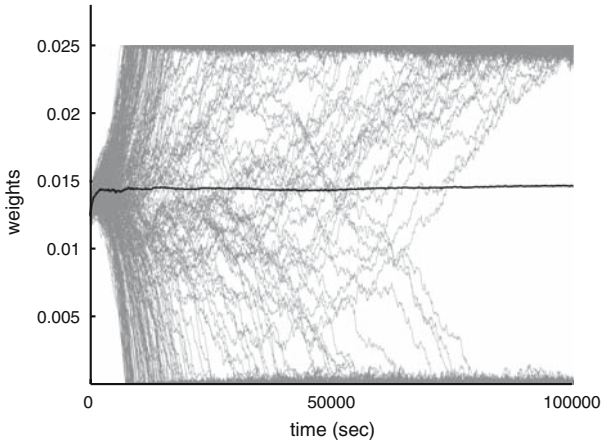

Fig. 5 Plot of the distribution of the weights after saturation in a fully connected recurrent network of 20 neurons. The *thick line* is the mean of the weights, which is surrounded by the bundle of all the individual weight plots, which are in *grey*. The theoretical equilibrium value $J^*_{\rm av}$ is given by the *dotted line*. The hard upper bound on the weights is set at 0.025. The learning-rate is $\eta = 10^{-6}$ and other parameter values are the same as in Fig. 3

the individual weights from the mean fixed-point weight $J^*_{\rm av}$ is consistent with the high multiplicity of the zero eigenvalue in (44). This weight evolution within the zero-eigenvalue manifold induces competition between the weights and eventually results in a bimodal distribution, i.e., each weight is either saturated or silent, as illustrated in Fig. 5.

In order to compare the theoretical value of the variance (45) with numerical simulations, it is more convenient to compute the variance of the distribution $\{J_{ij}\}$ of weights in a single learning trial

$$\mathrm{var}\{J_{ij}\}(t) = \frac{1}{[N(N-1)-1]} \sum_{\substack{i,j \\ i \neq j}}^{N} \left[J_{ij}(t) - J_{\rm av}(t)\right]^2 . \tag{47}$$

Figure 6 shows the comparison between the theoretical prediction (45) for the evolution of the variance $\mathrm{var}\{J_{ij}\}(t)$ with numerical simulation results for different values of N. The plots show that the variance initially grows in an approximately linear fashion with the theoretically calculated diffusion coefficient $\mathcal{D}$.

In contrast to the case where each neuron receives only feed-forward input (Kempter et al. 1999), each neuron here receives correlated inputs from the neurons within the network. Moreover, the correlation here is intrinsic to the network activity and not a parameter associated with the external inputs. However, the behavior has many similarities to that of the feed-forward case, since in both cases the weight dynamics causes the mean weight to approach a fixed-point value and the weight distribution to become bimodal.

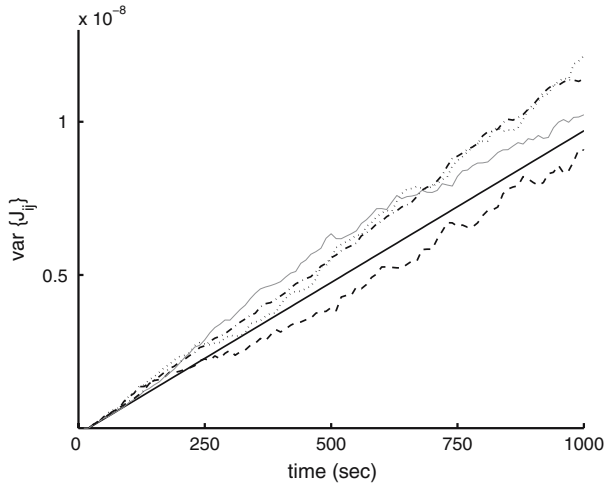

Fig. 6 Plot of the evolution of the variance of the weight distribution for the first 1,000 s. The *solid line* is the theoretical prediction (45). The *other lines* show the results from numerical simulations with different values of N: 20 (*dotted line*); 30 (*dashed line*); 50 (*dash-dotted line*); 100 (*grey line*). Other parameter values are the same as in Fig. 3

6 Discussion and conclusions

We have presented here a framework for the description of synaptic learning dynamics in networks of recurrently connected neurons. The analysis has employed the Poisson neuron (Kempter et al. 1998) since the linear dependence of the output spiking-rate upon the synaptic weights facilitates the analytic study of the solution. Just as for the case with feed-forward connectivity, this description encompasses the three main players, viz., the presynaptic neuron, the postsynaptic neuron, and the synapse that connects them. The learning equation embodies the concept of a *learning window* that describes the relationship between the input and output spike times and the associated changes in synaptic strength. Each pairing of input and output spikes through the learning window produces an incremental increase or decrease in synaptic efficacy.

This approximation allows us to find the explicit expression (26) for the spiking-rates of the recurrently connected neurons, which in turn allows the learning equations to be written explicitly. The learning equations (23) and (24) provide a rigorous mathematical framework to describe the resulting synaptic changes over time scales that are large in comparison with the duration of the learning window. An essential component of this stochastic analysis is the strong law of large numbers (Lamperti 1966), which ensures that the behavior of a large ensemble of synapses is essentially deterministic, even though individual synapses have considerable variability.

The effect of recurrent synapses is analyzed here in one particular case, namely where there is no external synaptic input. In future work we will analyze the asymptotic structure of the weights in more complex and interesting situations, such as where there are spontaneous homogeneous external inputs with no correlation and where the external synaptic inputs are partitioned into subgroups (Meffin et al. 2006), in order to understand both the nature of the information processing that the neurons carry out and the nature of the memories that they are capable of storing and recalling. In the case of feed-forward networks, in which there are no recurrent connections, the asymptotic pattern of evolution of the weights, as determined by the learning equation, is governed by the eigenvector of the matrix $\{Q_{ij}\}$ whose eigenvalue has the largest real part (Kempter et al. 1999; van Hemmen 2001; Burkitt and van Hemmen 2003). The goal of future studies will be to elucidate the effect of the recurrent connections upon this asymptotic pattern of weights.

In summary, the techniques presented here represent further stepping stones in the quest to understanding synaptic plasticity and thereby to further understand neuronal information processing capabilities in a network of interacting neurons.

Acknowledgements The authors cordially thank Walter Senn for his most constructive criticism. They also thank Hamish Meffin for a critical reading of an early version of the manuscript and detailed comments. They are at least equally indebted to David Grayden, Doreen Thomas, Iven Mareels, Chris Trengove, and Sean Byrnes for useful discussions. ANB is funded by the Australian Research Council (ARC Discovery Projects #DP0453205 and #DP0664271), the Brockhoff Foundation, and The Bionic Ear Institute. MG is funded by a scholarship from NICTA.

Appendix

A Calculation of eigenvalues of $\mathbb{L}$ for homogeneous fixed-point

This appendix gives the calculation of the eigenvalues λ (44) for the homogeneous fixed-point. The homogeneous solution provides some simplification of the matrix $(\mathbb{1}_N - J^*)$. We will now study the homogeneous solution where $J^* \equiv J^*_{\rm av} \Phi_J(E\,E^T)$. In this case, we have a simple expression for

$$\mathbb{1}_N - J^* = \left(1 + J^*_{\rm av}\right)\,\mathbb{1}_N \;-\; J^*_{\rm av}\,E\,E^T \tag{48}$$

and the inverse matrix is of the same form

$$\left(\mathbb{1}_N - J^*\right)^{-1} = (a-b)\,\mathbb{1}_N \;+\; b\,E\,E^T\;, \tag{49}$$

where a, b are defined by

$$\begin{aligned} a - (N-1)bJ^*_{\rm av} &= 1, \\ b - [(N-2)b + a]\,J^*_{\rm av} &= 0\;. \end{aligned} \tag{50}$$

Consequently

$$a = b + \frac{1}{1 + J^*_{\rm av}},$$
$$b = \frac{J^*_{\rm av}}{(1 + J^*_{\rm av})\left[1 - (N-1)J^*_{\rm av}\right]}\,. \tag{51}$$

For any matrix A in $\mathcal{M}_N$, $\mathbb{L}(A)$ can be written

$$\begin{aligned}\mathbb{L}(A) = -\nu_0 \Big\{ & b\ [a + (N-1)b]\ \left(w^{\rm in} + w^{\rm out}\right) \\ & \times \left(E^T\ A\ E\right)\ \Phi_J\left(E E^T\right) \\ & + (a-b)\ [a + (N-1)b]\ \Phi_J\left(w^{\rm in}\ A\ E\ E^T \right. \\ & \left. + w^{\rm out}\ E\ E^T\ A^T\right)\Big\}\,.\end{aligned} \tag{52}$$

Note that we used the fact that the matrix $\left(E^T\ A\ E\right)$ is actually a scalar and thus commutes with any other matrix (and can be removed from the argument of Φ_J). Likewise the following identities are useful

$$\begin{aligned} E^T\ E &= N, \\ E^T\left[(a-b)\ \mathbb{1}_N + b\ E\ E^T\right] &= [a + (N-1)\,b]\ E^T, \\ E^T\ \Delta J\ E &= \sum_{k,l} \Delta J_{kl}\,. \end{aligned} \tag{53}$$

Using Eq. (52), an eigenmatrix A related to the eigenvalue λ must satisfy

$$\begin{aligned} -\frac{\lambda}{\nu_0}\ A = {} & b\ [a + (N-1)\,b]\ \left(w^{\rm in} + w^{\rm out}\right) \\ & \times \left(E^T\ A\ E\right)\ \Phi_J\left(E\ E^T\right) \\ & + (a-b)\ [a + (N-1)b]\ \Phi_J\left(w^{\rm in}\ A\ E\ E^T \right. \\ & \left. + w^{\rm out}\ E\ E^T\ A^T\right]\,. \end{aligned} \tag{54}$$

We then multiply on the left by E^T and on the right by E to obtain a necessary condition on A and λ,

$$\begin{aligned} -\frac{\lambda}{\nu_0}\ & E^T\ A\ E \\ & = b\ [a + (N-1)b]\ \left(w^{\rm in} + w^{\rm out}\right) \\ & \quad \times \left(E^T\ A\ E\right)\ E^T\ \Phi_J\left(E E^T\right)\ E \\ & \quad + (a-b)\ [a + (N-1)b]\Big\{w^{\rm in} E^T\ \Phi_J\left(A\ E\ E^T\right)\ E \\ & \quad + w^{\rm out}\ E^T\ \Phi_J\left(E\ E^T\ A^T\right)\ E^T\Big\}\,. \end{aligned} \tag{55}$$

The following identity for any vector V in $\mathbb{R}^N$ is useful here:

$$\Phi_J\left(V\ E^T\right)\ E = (N-1)\ V\ , \tag{56}$$

which holds because the network topology is fully connected. Using this identity, we can rewrite the two last terms of (55) as

$$\begin{aligned} E^T\ \Phi_J\left(A\ E\ E^T\right)\ E &= (N-1)\ E^T\ A\ E, \\ E^T\ \Phi_J\left(E\ E^T\ A^T\right)\ E &= (N-1)\ E^T\ A^T\ E \\ &= (N-1)\ E^T\ A\ E\ , \end{aligned} \tag{57}$$

since $(E^T\ A\ E)$ is a scalar. We also have

$$E^T\ \Phi_J\left(E\ E^T\right)\ E = (N-1)\ E^T\ E = (N-1)\ N\ . \tag{58}$$

Consequently (55) becomes

$$\begin{aligned} \Big\{\frac{\lambda}{\nu_0} & + (w^{\rm in} + w^{\rm out})(N-1) \\ & \times [a + (N-1)b]^2\Big\}\ E^T\ A\ E = 0\,. \end{aligned} \tag{59}$$

Either the coefficient between the curly brackets on the left side is zero, which gives the eigenvalue

$$\lambda_2 \equiv -\nu_0 (w^{\rm in} + w^{\rm out})\ (N-1)\ [a + (N-1)\,b]^2 \tag{60}$$

or the sum of the coefficient of the matrix A are zero

$$E^T\ A\ E = \sum_{i,j} A_{ij} = 0\,. \tag{61}$$

In the second case, the condition on A to be an eigenmatrix of $\mathbb{L}$ becomes

$$\begin{aligned} -\frac{\lambda}{\nu_0}\ A = {} & (a-b)\ [a + (N-1)b]\Big[w^{\rm in}\ \Phi_J\left(A\ E\ E^T\right) \\ & + w^{\rm out}\ \Phi_J\left(E\ E^T\ A^T\right)\Big]\,. \end{aligned} \tag{62}$$

Multiplying on the right by E, we obtain

$$\begin{aligned} -\frac{\lambda}{\nu_0}\ A\ E = {} & (a-b)\,[a + (N-1)b]\Big[w^{\rm in}\Phi_J\left(A\ E\ E^T\right)\ E \\ & + w^{\rm out}\ \Phi_J\left(E\ E^T\ A^T\right)\ E\Big]\,. \end{aligned} \tag{63}$$

The following identity is useful for any V in $\mathbb{R}^N$ (proof by calculating the coefficients):

$$\Phi_J\left(E\ V^T\right)\ E = \left(E^T\ V\right)\ E\ -\ V\ . \tag{64}$$

Applied to $V = AE$, this relation leads to

$$\Phi_J\left(E\ E^T\ A^T\right)\ E = \left(E^T\ A\ E\right)\ E\ -\ A\ E\ . \tag{65}$$

Because here $\left(E^T\ A\ E\right) = 0$, (63) becomes

$$-\frac{\lambda}{\nu_0} A\ E = (a-b)\ [a+(N-1)b] \times \left[w^{\text{in}}(N-1)\ A\ E\ -\ w^{\text{out}}\ A\ E\right]. \tag{66}$$

Hence we have

$$\left\{-\frac{\lambda}{\nu_0}\ -\ (a-b)\ [a+(N-1)b] \times \left[w^{\text{in}}(N-1) - w^{\text{out}}\right]\right\} A\ E = 0\ . \tag{67}$$

Once again, either the coefficient between the curly brackets on the left side is zero, which gives the eigenvalue

$$\lambda_1 \equiv -\nu_0(a-b)\ [a+(N-1)b]\left[w^{\text{in}}(N-1)\ -\ w^{\text{out}}\right] \tag{68}$$

or the vector $A\ E$ is zero, i.e., for all i we have $\sum_j A_{ij} = 0$. The condition upon $A\ E$ implies $\mathbb{L}(A) = 0$, cf. (54), and hence the eigenvalue is determined

$$\lambda_0 \equiv 0\ . \tag{69}$$

Replacing the terms a, b by their values in (51) for each of the eigenvalues λ_0 (69), λ_1 (68), λ_2 (60), gives the expression (44) for the eigenvalues. The multiplicity of the eigenvalues follows from the analysis in Appendix B.

B Stability of the fixed-point manifold

The linearized operator $\mathbb{L}$ related to the learning equation for any fixed-point J^* on the manifold is defined by (43). We investigate here the spectrum of $\mathbb{L}$ for the matrices A in $\mathcal{M}_N$. Recall that the dimension of $\mathcal{M}_N$ is $N(N-1)$.

From the definition (43) of $\mathbb{L}$ it is clear that any matrix A for which $AE = 0$ is an "eigenmatrix" with eigenvalue zero: $\widetilde{\lambda}_0 = 0$. It can be shown that provided $w^{\text{in}} \neq w^{\text{out}}$ the converse is also true: any eigenmatrix A corresponding to the zero eigenvalue satisfies $AE = 0$ (which determines a linear subspace $\mathcal{A}_N$ of dimension $N(N-2)$ of $\mathcal{M}_N$). Note that this linear subspace $\mathcal{A}_N$ is parallel to the manifold of fixed-points (in the sense that the manifold is contained in the affine subspace defined by $A\ E = \frac{\mu-\nu_0}{\mu} E$). Note also that $\Phi_J\left(EE^T\right)$ is always an eigenmatrix of $\mathbb{L}$ with eigenvalue

$$\widetilde{\lambda}_2 = -\frac{\mu^2}{\nu_0}\left(w^{\text{in}}\ +\ w^{\text{out}}\right)\ . \tag{70}$$

So far the situation is exactly the same as for the homogeneous fixed-point.

B.1 Relationship between the spectra of $(\mathbb{1}_N - J^*)$ and $\mathbb{L}$

For any eigenvector V of $(\mathbb{1}_N - J^*)$ related to an eigenvalue γ ($\gamma \neq 0$ because $(\mathbb{1}_N - J^*)$ is invertible on the manifold), we can construct a matrix A defined by

$$A = \Phi_J\left(\alpha_1\ V\ E^T\ +\ \alpha_2\ E\ V^T\ +\ \alpha_3\ E\ E^T\right) \tag{71}$$

with

$$\{\alpha_1, \alpha_2, \alpha_3\} = \left\{w^{\text{in}},\ w^{\text{out}},\ \frac{-\frac{\mu}{\nu_0}\ (E^T V)}{(N-1)\left(\frac{\mu}{\nu_0} - \frac{1}{\gamma}\right)}\right\} \tag{72}$$

(with the extra condition $\gamma \neq \frac{\nu_0}{\mu}$) that satisfies

$$\mathbb{L}(A) = -\frac{\mu}{\gamma}\left[(N-1)\ w^{\text{in}}\ -\ w^{\text{out}}\right] A\ . \tag{73}$$

Such a matrix A is then an eigenmatrix of $\mathbb{L}$ with eigenvalue

$$\widetilde{\lambda}_1(\gamma) \equiv -\frac{\mu}{\gamma}\left[(N-1)\ w^{\text{in}}\ -\ w^{\text{out}}\right]\ . \tag{74}$$

Note that in Appendix A the eigenvalue $\lambda_1 = \widetilde{\lambda}_1(\gamma)$ with $\gamma = 1 + J^*_{\text{av}}$.

B.2 Case when $(\mathbb{1}_N - J^*)$ is diagonalisable

In this case we have a basis of N eigenvectors $\{V_1, \ldots, V_{N-1}, E\}$ corresponding to the eigenvalues $\{\gamma_1, \ldots, \gamma_{N-1}, \frac{\nu_0}{\mu}\}$ of $(\mathbb{1}_N - J^*)$. We can show that the construction of a family of matrices $\{A_k\}$ as in the previous section from a linearly independent family of vectors $\{V_k\}$ is still linearly independent provided $(w^{\text{in}} \neq w^{\text{out}})$. This follows because the construction involves an injective linear morphism: $V \mapsto \Phi_J\left(w^{\text{in}}\ V\ E^T + w^{\text{out}}\ E\ V^T\right)$. Moreover this family of $\{A_k\}$ together with $\Phi_J\left(EE^T\right)$ still forms a linearly independent family provided $(w^{\text{in}} + w^{\text{out}}) \neq 0$.

Note that in general (i.e., when $(N-1)\ w^{\text{in}} - w^{\text{out}} \neq 0$), the eigenvalues $\widetilde{\lambda}_1(\gamma_k)$ of $\mathbb{L}$ defined in the previous section are non-zero, which ensures the linear independence between these N matrices $\{A_k, \Phi_J\left(EE^T\right)\}$ $(k = 1, \ldots, N-1)$ and any basis of $\mathcal{A}_N$ with $N(N-2)$ elements. We thus obtain a basis of $\mathcal{M}_N$ formed by $N(N-1)$ eigenmatrices of $\mathbb{L}$. Consequently $\mathbb{L}$ is diagonalisable and the signs of its non-zero eigenvalues are given by the signs of $-\left[(N-1)\ w^{\text{in}} - w^{\text{out}}\right]$ and $-\left(w^{\text{in}} + w^{\text{out}}\right)$ (recall that $\gamma > 0$ and $\nu_0 > 0$).

Note that the signs do not depend on the fixed-point considered, and are constant over the whole manifold when we use hard bounds to keep $(\mathbb{1}_N - J^*)$ invertible at all times (the manifold is then a compact connected subset of a linear subspace of $\mathcal{M}_N$). This means γ^{-1} is bounded, and the zero eigenvalues are separated from the non-zero eigenvalues, provided the already stated conditions on the parameters hold: namely $\left[(N-1)\ w^{\text{in}} - w^{\text{out}}\right] \neq 0$, $\left(w^{\text{in}} + w^{\text{out}}\right) \neq 0$

and $w^{\text{in}} \neq w^{\text{out}}$. We say the fixed-points are "quasi-stable" when the non-zero eigenvalues are strictly negative.

B.3 Extension to the whole manifold

Because of the separation of the eigenvalues stated in Sect. B.2, there is no zero eigenvalue except for those along the direction of the manifold itself. Then, the "quasi-stable" property (i.e., along the orthogonal direction of the manifold) of any fixed-point would apply also in its surrounding. Because the eigenvalues of the linear operator related to them vary continuously along the manifold, we can use the fact that the diagonalisable matrices are dense in the manifold, in order to extent their dynamical properties to the whole manifold.

Consequently, the whole manifold acts like an attractor with respect to the rest of the space $\mathcal{M}_N$ if the following conditions on the parameters are satisfied

$$\begin{aligned} (N-1)\, w^{\text{in}} - w^{\text{out}} &> 0, \\ w^{\text{in}} + w^{\text{out}} &> 0\,. \end{aligned} \tag{75}$$

Along the manifold, due to the zero eigenvalues, there is no deterministic constraint (see discussion in Sect. 5 on the stochastic origin of the variance).

C Calculation of $J_{\text{av}}(t)$

The equations describing the evolution of $J_{\text{av}}(t)$ are straightforwardly derived from (31)

$$\begin{aligned} \nu_{\text{av}}(t) &= \nu_0\, [1-(N-1)J_{\text{av}}(t)]^{-1}\,, \\ Q^W_{\text{av}}(t) &= \widetilde{W}\,\nu_0\,\nu_{\text{av}}(t)\, [1-(N-1)J_{\text{av}}(t)]^{-1}\,, \\ \frac{\mathrm{d}}{\mathrm{d}t}J_{\text{av}}(t) &= \frac{\nu_0\,(w^{\text{in}}+w^{\text{out}})}{1-(N-1)J_{\text{av}}(t)} + \frac{\nu_0^2\,\widetilde{W}}{[1-(N-1)J_{\text{av}}(t)]^2}\,. \end{aligned} \tag{76}$$

This can be rewritten as

$$-\frac{1}{(N-1)}\,\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \frac{r}{x(t)} + \frac{s}{x^2(t)}\,, \tag{77}$$

where $x(t) = [1-(N-1)J_{\text{av}}(t)]$, $r = \nu_0\,(w^{\text{in}}+w^{\text{out}})$ and $s = \nu_0^2\,\widetilde{W}$. Hence by rearranging

$$\begin{aligned} -(N-1)\,\mathrm{d}t &= F(x)\,\mathrm{d}x, \\ F(x) &= \left[\frac{x}{r} - \frac{s}{r^2} + \frac{s^2}{r^2\,(rx+s)}\right]. \end{aligned} \tag{78}$$

Integrating from $\{t=0\,,\ x(0)=x_0\}$ to $\{t\,,\ x\}$ gives

$$t = \frac{1}{N-1}\left[F(x_0) - \frac{x^2}{2r} + \frac{sx}{r^2} - \frac{s^2}{r^3}\ln|rx+s|\right]. \tag{79}$$

Consequently the leading exponential term is given by

$$x(t) \sim e^{-(N-1)\,r^3\,t/s^2}\,, \tag{80}$$

which gives the time constant (46): $\tau_J = s^2/(N-1)\,r^3$. The deviations from this exponential behavior, as illustrated in Fig. 3, are due to the polynomial terms in (79) and are larger when the initial mean is further from the fixed-point J^*_{av}.

References

Bi GQ, Poo MM (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. Annu Rev Neurosci 24:139–166

Burkitt AN, van Hemmen JL (2003) How synapses in the auditory system wax and wane: theoretical perspectives. Biol Cybern 89:318–332

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. Nature 383: 76–78

Hebb DO (1949) The organization of behavior. Wiley, New York

Hawkes AG (1971) Point spectra of some mutually exciting point processes. J Roy Stat Soc B 33:438–443

Kempter R, Gerstner W, van Hemmen JL, Wagner H (1998) Extracting oscillations: neuronal coincidence detection with noisy periodic spike input. Neural Comput 10:1987–2017

Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. Phys Rev E 59:4498–4514

Kempter R, Leibold C, Wagner H, van Hemmen JL (2001b) Formation of temporal feature maps by axonal propagation of synaptic learning. Proc Natl Acad Sci USA 98(7):4166–4171

Kistler WM, van Hemmen JL (2000) Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic action potentials. Neural Comput 12:385–405

Lamperti J (1966) Probability, 2nd edn. Benjamin, New York. (1996) Wiley, New York

Leibold C, Kempter R, van Hemmen JL (2001) Temporal map formation in the barn owl's brain. Phys Rev Lett 87:248101

Leibold C, Kempter R, van Hemmen JL (2002) How spiking neurons give rise to a temporal-feature map: from synaptic plasticity to axonal selection. Phys Rev E 65:051915

Markram H, Lübke J, Forscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213–215

Meffin H, Besson J, Burkitt AN, Grayden DB (2006) Learning the structure of correlated synaptic subgroups using stable and competitive spike-timing-dependent plasticity. Phys Rev E73:041911

Sanders JA, Verhulst F (1985) Averaging methods in nonlinear dynamical systems. Springer, Berlin

Song F, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature Neurosci 3:919-926

Van Hemmen JL (2001) Theory of synaptic plasticity. In: Moss F, Gielen S (eds) Handbook of biophysics, vol 4. Elsevier, Amsterdam, pp 771–823