Biol Cybern (2009) 101:411–426
DOI 10.1007/s00422-009-0343-4

ORIGINAL PAPER

Biological Cybernetics

Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks III: Partially connected neurons driven by spontaneous activity

Matthieu Gilson · Anthony N. Burkitt · David B. Grayden · Doreen A. Thomas · J. Leo van Hemmen

Received: 23 April 2009 / Accepted: 19 October 2009 / Published online: 24 November 2009

Abstract In contrast to a feed-forward architecture, the weight dynamics induced by spike-timing-dependent plasticity (STDP) in a recurrent neuronal network is not yet well understood. In this article, we extend a previous study of the impact of additive STDP in a recurrent network that is driven by spontaneous activity (no external stimulating inputs) from a fully connected network to one that is only partially connected. The asymptotic state of the network is analyzed, and it is found that the equilibrium and stability conditions for the firing rates are similar for both full and partial connectivity: STDP causes the firing rates to converge toward the same value and remain quasi-homogeneous. However, when STDP induces strong weight competition, the connectivity affects the weight dynamics in that the distribution of the weights disperses more quickly for lower density than for higher density. The asymptotic weight distribution strongly depends upon that at the beginning of the learning epoch; consequently, homogeneous connectivity alone is not sufficient to obtain homogeneous neuronal activity. In the absence of external inputs, STDP can nevertheless generate structure in the network through autocorrelation effects, for example, by introducing asymmetry in network topology.

Keywords Learning · STDP · Recurrent neuronal network · Partial connectivity

M. Gilson (✉) · A. N. Burkitt · D. B. Grayden · D. A. Thomas
Department of Electrical and Electronic Engineering,
The University of Melbourne, Melbourne, VIC 3010, Australia
e-mail: mgilson@bionicear.org

M. Gilson · A. N. Burkitt · D. B. Grayden
The Bionic Ear Institute, 384-388 Albert St., East Melbourne,
VIC 3002, Australia

M. Gilson · A. N. Burkitt · D. B. Grayden · D. A. Thomas
NICTA, Victoria Research Lab, University of Melbourne,
Melbourne, VIC 3010, Australia

J. L. van Hemmen
Physik Department (T35) and BCCN–Munich, Technische Universität
München, 85747 Garching bei München, Germany

1 Introduction

Synaptic plasticity is central to understanding information processing in the brain, in that it is believed to structure neuronal circuits. It relies on changes at the molecular level, which take place locally at the synaptic site between neurons, which results in the strengthening (potentiation) or weakening (depression) of the synaptic weight. Recent studies have established the importance of the timing of individual spikes in synaptic plasticity (Gerstner et al. 1996; Markram et al. 1997; Bi and Poo 2001), which lead to the model of spike-timing-dependent plasticity (STDP). Previous theoretical studies have primarily investigated the weight dynamics induced by STDP for single neurons and the implications for feed-forward networks (Gerstner et al. 1996; Kempter et al. 1999; Gütig et al. 2003; Burkitt et al. 2004; Meffin et al. 2006). The cortex, however, is dominated by recurrent connections, and the effect of STDP has only begun to be addressed in such recurrent networks, mainly using numerical simulation (Song and Abbott 2001; Morrison et al. 2007). A few theoretical results on synaptic plasticity in recurrent architectures (Karbowski and Ermentrout 2002; Masuda and Kori 2007) illustrate the difficulty in incorporating the effect of feedback synaptic loops. The relationship between STDP and synchronous activity has been particularly studied in specific network configuration (Levy et al. 2001; Câteau et al. 2008; Lubenov and Siapas 2008). In a recent article (Burkitt et al.

2007), we developed a framework for the analysis of STDP in recurrent networks for general network topology and activity. This framework describes how the network activity, viz. firing rates and spike-time correlations, determines the evolution of the weights, which occurs on a slower time scale and the resulting development of structure within the network. In that cited article, the analysis was carried out only for network with full recurrent connectivity and no external inputs (Burkitt et al. 2007).

Now, we extend the analysis of the synaptic dynamics to the case of an arbitrary topology of partial recurrent connectivity. The focus of our analysis is laid on the stable asymptotic distributions of the firing rates and of the recurrent weights that take place after a sufficiently long learning epoch (emerged structure). This article is part of a series that analyzes the emergence of network structure induced by STDP. Previous articles in this series (Gilson et al. 2009a,b) investigated the case of a recurrently connected network stimulated by external inputs with plastic input synapses, while the recurrent connections are kept fixed. In this article, we focus on the case of plastic recurrent synapses with no external inputs; in other words, each neuron receives excitation from the same external background activity. This analysis is a first step toward investigating a network with plastic recurrent synapses and external inputs with a firing-rate and spike-time correlation structure (Gilson et al. 2009c), similar to that described by Gilson et al. (2009a). We constrain our study to additive STDP (Sect. 2.1), keeping in mind its limitations compared to more elaborate versions of STDP (van Rossum et al. 2000; Sjöström et al. 2001; Gütig et al. 2003; Pfister and Gerstner 2006; Meffin et al. 2006; Appleby and Elliott 2006; Morrison et al. 2007). Even though additive STDP is only one possible approximation to a more complete description of neuronal plasticity, it nevertheless is useful to initially study this case for a number of reasons. First, it is sufficiently simple that it allows tractable analytical methods to be used that are capable of providing considerable insight into the behavior of the system. In this sense, it provides a test bed for the development of new analytical tools that may lead to methods for investigating other forms of plasticity. The method for analyzing partially connected networks in this article could be more generally applied, for example, to weight-dependent (non-additive) STDP. Second, the results of the present analysis are still meaningful and allow us to gain a better understanding of some principles underlying more general forms of STDP. After presenting the Poisson neuron model (Sect. 2.2), we recapitulate the theoretical framework used to model the evolution of the recurrent weights in a partially connected network of neurons only driven by their spontaneous activity. We then analyze the equilibrium enforced on the weights by STDP (Sect. 3) and how a structure can emerge (Sect. 4) when there is no external stimulation.

2 Modeling learning and neuronal activity

2.1 Hebbian additive STDP

STDP describes the change in the synaptic weight induced by single spikes and pairs of spikes, and incorporates the precise timing of spikes. Here, we use the so-called Hebbian additive STDP model (Kempter et al. 1999), in contrast to the other STDP versions (van Rossum et al. 2000; Gütig et al. 2003). For two neurons in and out connected by a synapse in $\rightarrow$ out with weight J, the weight change δJ induced by a sole pair of pre- and post-synaptic spikes at times t^{in} and t^{out}, respectively, is given by the sum of three contributions:

$$\delta J = \eta \begin{cases} w^{\text{in}} & \text{at time} t^{\text{in}} \\ w^{\text{out}} & \text{at time} t^{\text{out}} \\ W(t^{\text{in}} - t^{\text{out}}) & \text{at time} \max(t^{\text{in}}, t^{\text{out}}). \end{cases} \quad (1)$$

The constant w^{in} (resp. w^{out}) accounts for the effect of each pre-synaptic (post-synaptic) spike, which occurs at time t^{in} (t^{out}). The STDP learning window function W determines the contribution of each pair of pre- and post-synaptic spikes in terms of the difference between the spike times $t^{\text{in}} - t^{\text{out}}$ (Gerstner et al. 1996; Kempter et al. 1999). Figure 1 illustrates a typical choice of W where pre-synaptic spikes that take part in the firing of post-synaptic spikes induce the potentiation (Hebb 1949). These three contributions are scaled by a learning parameter η, typically very small, so that learning occurs very slowly compared to the other neuronal and synaptic mechanisms. We chose η such that the weight change is three orders of magnitude below the corresponding upper bound. All spike pairs are involved in the weight modification; we do not consider any temporal restriction (Sjöström et al. 2001,

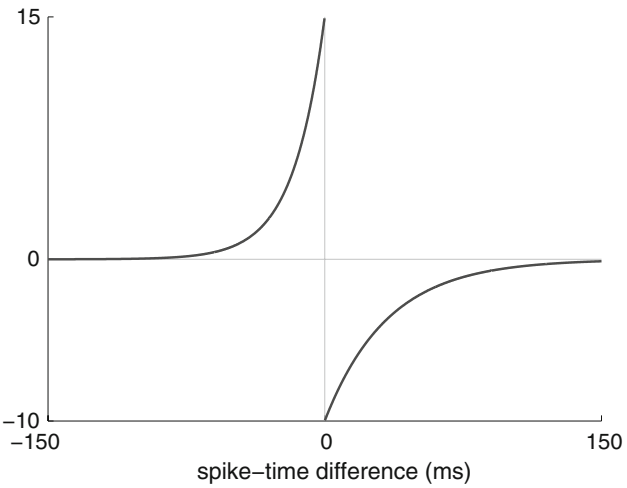


Fig. 1 Example of STDP window function W. It consists of one decaying exponential for potentiation (*left curve*) with time constant 17 ms and one for depression (*right curve*) with 34 ms. See Appendix 10 for details on the parameters

Burkitt et al. 2004). See Gilson et al. (2009a, Sect. 2.1) for more details.

2.2 Poisson neuron model

In the Poisson neuron model (Kempter et al. 1999), the spiking mechanism of a given neuron i is approximated by an inhomogeneous Poisson process driven by an intensity function $\rho_i(t)$ to generate an output spike-time series $S_i(t)$. The rate function $\rho_i(t)$ is to be related to the soma potential, and it evolves over time according to the excitation received from other neurons $j \neq i$ (self-connections are forbidden),

$$\rho_i(t) = \nu_0 + \sum_{j \neq i} \left[J_{ij}(t) \sum_n \epsilon\left(t - t_{j,n} - d_{ij}\right) \right]. \tag{2}$$

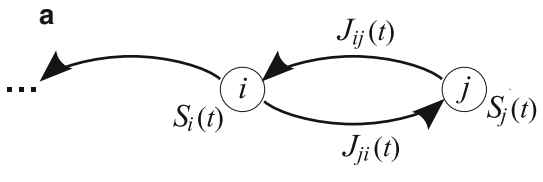
The constant ν_0 is the spontaneous firing rate (identical for all the neurons), which accounts for other incoming connections that are not considered in detail. Each pre-synaptic spike induces a variation of $\rho_i(t)$ taken care of by the post-synaptic potential (PSP), which is determined by the synaptic weights J_{ij}, the post-synaptic response kernel ϵ, and the delays d_{ij}. The kernel function ϵ models the PSP due to the current injected into the post-synaptic neuron as a consequence of one single pre-synaptic spike; $\epsilon(t)$ is normalized to one, i.e., $\int \epsilon(t)\, \mathrm{d}t = 1$, and to preserve causality, we have $\epsilon(t) = 0$ for $t < 0$. A PSP also incorporates a fixed delay, denoted by d_{ij} for the synaptic connection from neuron j. The overall synaptic influx is the sum of the PSPs over all spike times $t_{j,n}$ (related to the synapse from neuron j, and indexed by n). We only consider positive weights here, i.e., excitatory synapses. See Gilson et al. (2009a, Sect. 2.2) for details.

2.3 Dynamical system to model network activity

Previous articles developed a framework to describe the impact of STDP upon the neuronal dynamics (Kempter et al. 1999; Burkitt et al. 2007; Gilson et al. 2009a). A dynamical system of equations was derived to link the variables of importance for the network activity (firing rates and pairwise spike-time correlations) and the synaptic weights. Here, we recapitulate this framework that requires very slow synaptic plasticity compared to the neuronal activation dynamics (adiabatic hypothesis); thus the expectation values of the firing rates and pairwise correlations are quasi-constant in time for the network neurons.

2.3.1 Description of the network activity

We consider a recurrently connected network of N Poisson neurons as illustrated in Fig. 2a, where each neuron excites other neurons via connections that may form feedback loops in the network, but without self-connections. Typically N is large, and the network connectivity topology can be arbitrary with n^J recurrent connections. Partially connected networks are generated by randomly assigning neuron-to-neuron connections.



b

firing rates	$\nu_i(t)$
pairwise correlations	$Q_{ij}(t,u)$
weights	$J_{ij}(t)$
delays	d_{ij}

Fig. 2 Presentation of the network notation. **a** Schematic illustration of two of the N network neurons (*circles*, indexed by $1 \leq i, j \leq N$). The neurons are recurrently connected with plastic weights $J_{ij}(t)$ (*thick arrows*). The output spike train of neuron i is denoted by $S_i(t)$. **b** Table of the network variables that describe the neuronal activity (time-averaged firing rates $\boldsymbol{\nu}$ and correlation coefficients Q), and those related to the synaptic connections (weights J and delays d)

Spikes are considered to be instantaneous events. We denote by $S_i(t)$ the spike-time series (Dirac comb) of the neuron i, $1 \leq i \leq N$; its value is zero except at the times when a spike is fired and the spike train is described as a sum of Dirac delta-functions.

We define the time-averaged firing rate $\nu_i(t)$ of neuron i, over a duration T taken much larger than the time scale of the neuronal activation mechanisms but much smaller than the learning time scale related to η^{-1} (Kempter et al. 1999; Burkitt et al. 2007; Gilson et al. 2009a)

$$\nu_i(t) := \frac{1}{T} \int_{t-T}^{t} \left\langle S_i(t') \right\rangle \mathrm{d}t', \tag{3}$$

where the angular brackets denote the probabilistic ensemble average over the randomness so that $\langle S_i(t) \rangle$ is the instantaneous firing rate.

In order to take the delays into account, the correlation coefficients previously used by Burkitt et al. (2007) have been modified: Q_{ij}^W accounts for the correlation between neurons i and j, after the convolution with the STDP window function W and the delays d_{ij}

$$Q_{ij}^W(t) := \int_{-\infty}^{+\infty} W(u)\, Q_{ij}(t, u - d_{ij})\, \mathrm{d}u, \tag{4}$$

where Q_{ij} is the time-averaged correlation between neuron j and neuron i

$$Q_{ij}(t,u) := \frac{1}{T}\int_{t-T}^{t} \big\langle S_i(t')S_j(t'+u)\big\rangle \,\mathrm{d}t'. \tag{5}$$

The delay d_{ij} in (4) accounts for the transmission time required for a spike fired by neuron j to reach the synaptic site $j \to i$, which plays a role in the learning process.

2.3.2 The equations describing the dynamical system

The weights are modified due to STDP according to the activities of the pre- and post-synaptic neurons for each synapse. For a small time interval $[t, t+\delta t]$, the change in the input weight J_{ij} described in (1) can be expressed using the pre- and post-synaptic spike trains (Kempter et al. 1999; Burkitt et al. 2007)

$$\begin{aligned}\delta J_{ij}(t) = \eta \int_{t}^{t+\delta t} \Big[w^{\mathrm{in}}\, S_j(t'-d_{ij}) \;+\; w^{\mathrm{out}}\, S_i(t')\Big]\,\mathrm{d}t' \\ +\eta \int_{(t',u)\in\mathcal{I}(t)} W(u)\, S_i(t')\, S_j(t'-d_{ij}+u)\,\mathrm{d}u\,\mathrm{d}t'.\end{aligned} \tag{6}$$

Recall that $S_j(t'-d_{ij})$ is the delayed time series of the pre-synaptic spikes so that the time difference u at the synaptic site between the pre- and post-synaptic spikes (at respective times t^{pre} and t^{post} at the somas of both neurons) is $u = t^{\mathrm{pre}} + d_{ij} - t^{\mathrm{post}}$. The domain of integration $\mathcal{I}(t)$ is the subset $(t',u) \in \mathbb{R}^2$ satisfying the three conditions:

$$\begin{aligned} &t' \le t+\delta t;\\ &t'-d_{ij}+u \le t+\delta t;\\ &t \le t' \text{ or } t \le t'-d_{ij}+u. \end{aligned} \tag{7}$$

The first two lines ensure that the spikes occur before $t+\delta t$, and the last line that at least one of them is in the time interval $[t, t+\delta t]$.

When STDP slowly modifies the weights over a long time, the terms in (6) self-average and, consequently, they can be evaluated using the probabilistic ensemble average denoted by the angular brackets $\langle\cdots\rangle$; see Kempter et al. (1999); Burkitt et al. (2007); Gilson et al. (2009a) for details of similar derivations. The derivative of the expectation value of the weight J_{ij} denoted by $\dot{J}_{ij}(t)$ is approximated by the temporal average of the summed $\langle\delta J_{ij}(t')\rangle$ over a sufficiently long-time interval of duration T, which leads to the ensemble average taken on (6).

In the limit of large networks ($N \gg 1$), we can ignore the effects related to autocorrelation as a first approximation to evaluate the evolution of the expectation values for the recurrent weights and neuron firing rates. We keep in mind that such effects may have an impact on the dynamics, which will actually be discussed in Sect. 4; the general case will be the focus of a subsequent companion article (Gilson et al. 2009c). In this way, the spike trains are probabilistically quasi-independent for all pairs of neurons, i and j; the correlation coefficients, Q^W_{ij}, defined in (4) satisfy

$$Q^W_{ij}(t) \;\simeq\; \widetilde{W}\,\nu_i(t)\,\nu_j(t), \tag{8}$$

where $\widetilde{W}$ is the integral value of the function W

$$\widetilde{W} := \int_{-\infty}^{+\infty} W(u)\,\mathrm{d}u. \tag{9}$$

We finally obtain a system of matrix equations, which describes the dynamics of the firing rates and the weights.

$$\boldsymbol{\nu} = (\mathbb{1}_N - J)^{-1}\;\nu_0\,\mathbf{e}, \tag{10a}$$

$$\dot{J} = \Phi_J\left(w^{\mathrm{in}}\,\mathbf{e}\,\boldsymbol{\nu}^{\mathbf{T}} + w^{\mathrm{out}}\,\boldsymbol{\nu}\,\mathbf{e}^{\mathbf{T}} + \widetilde{W}\,\boldsymbol{\nu}\,\boldsymbol{\nu}^{\mathbf{T}}\right). \tag{10b}$$

Time has been re-scaled to remove η, and the time variable t is omitted from all the vectors and matrices that evolve over time. The vector $\mathbf{e}$ is the column vector with all N elements equal to one, i.e.,

$$\mathbf{e} := [1,\ldots,1]^{\mathbf{T}}. \tag{11}$$

The superscript, '$\mathbf{T}$', denotes the matrix transposition so that, e.g., $\mathbf{e}\,\mathbf{e}^{\mathbf{T}}$ is a $N \times N$ matrix. The projector Φ_J operates on the vector space of $N \times N$ matrices: it nullifies all the matrix components that correspond to non-existent connections in the network. In particular, since we forbid self-connection of the neurons, Φ_J nullifies all the diagonal terms. The system of equations (10a–10b) is the same system that was studied by Burkitt et al. (2007).

Considerations concerning the invertibility of the matrix $\mathbb{1}_N - J$ are discussed in Appendix 6. This invertibility property is required at all times to ensure that the firing rates remain bounded. We introduce bounds on the weights in numerical simulation for this purpose, since the weights tend to diverge due to the competition induced by STDP (Kempter et al. 1999; Burkitt et al. 2007). These bounds on the weights also induce bounds on the firing rates (see Appendix 6). All the simulation results presented in this article were run using the neuron and learning parameters listed in Appendix 10.

2.3.3 Analysis of the system dynamics

Our aim is to investigate the steady states of the network variables and of the weights, as well as their stability. Note that (10b) corresponds to pure rate-based learning, where synaptic delays do not play any role. The system of equations (10a–10b) describes the evolution of the expectation value of the weights, i.e., the first order of the stochastic process. In the remainder of the article, we refer to this leading order as the *drift* of the dynamics, in comparison to *higher orders*

ignored by (10a–10b). It is, however, possible to use our formalism to analyze further the latter (spike-based effects), as will be done in Sect. 4.

3 Equilibrium and stability

The study of the weight dynamics induced by STDP for a network with no external inputs was demonstrated in an earlier article for the case of full recurrent connectivity (Burkitt et al. 2007): we described the fixed points of the firing rates and of the recurrent weights, their stability, and the evolution of the weight variance. In particular, Burkitt et al. (2007) derived the stability conditions for the homeostatic equilibrium, namely, the situation where the mean firing rates and mean weights stabilize although individual firing rates and weights may continue to change, and the same analysis can be applied to the present case of partial connectivity. We recall these stability conditions: the mean firing rate and the mean weight over all neurons have stable equilibria provided

$$w^{\text{in}} + w^{\text{out}} > 0 \text{ and } \widetilde{W} < 0. \tag{12}$$

Here, we study the equilibria of the dynamical system (10a–10b) in terms of individual firing rates and mean weights (taking into account the network topology) and the corresponding stability conditions. The present analysis of the weight drift is similar to that of Burkitt et al. (2007), but here the projector Φ_J in (10b) is non-trivial and nullifies not only the diagonal elements, but also other elements according to the partial connectivity of the network. The matrix J belongs to the vector subspace of $\mathbb{R}^{N\times N}$ defined by $\mathbb{M}_J := \{X \in \mathbb{R}^{N\times N}, \Phi_J(X) = X\}$, whose dimension is the number of existing connections n^J.

3.1 Fixed point of the firing rates

We first find the equilibrium states of the network dynamics in terms of the firing rates. Setting $\dot{J} = 0$ in (10b) leads to the following condition on the firing rates for all existing connections $j \to i$:

$$w^{\text{in}}\nu_j + w^{\text{out}}\nu_i + \widetilde{W}\nu_j\nu_i = 0, \tag{13}$$

that is,

$$\nu_i = q(\nu_j) := -\frac{w^{\text{in}}\nu_j}{w^{\text{out}} + \widetilde{W}\nu_j}. \tag{14}$$

For any loop of synaptic connections from a given neuron i_0 back to itself through neurons $i_1, i_2, \ldots, i_{n-1}$, we have

$$\nu_{i_{m+1}} = q(\nu_{i_m}) \quad \text{for } m = 0, \ldots, n-1, \tag{15}$$

and thus $\nu_{i_0} = q^{\{n\}}(\nu_{i_0})$ where $q^{\{n\}}$ is the self-composition of q defined in (14) iterated n times. In other words, ν_{i_0} is a fixed point of $q^{\{n\}}$ whenever there exists a loop of length n in which neuron i_0 takes part. Similarly, for all the ν_{i_m} of the loop. Owing to the special form of q, the function $x \mapsto q^{\{n\}}(x)$ has only two fixed points, namely, those of q (see Appendix 7.1). This means that at the equilibrium of the learning weight dynamics, any neuron i with non-zero firing rate ν_i within a loop of arbitrary length must satisfy

$$\nu_i = \mu := -\frac{w^{\text{in}} + w^{\text{out}}}{\widetilde{W}}. \tag{16}$$

We discard silent neurons at the equilibrium since the consistency equation for the firing rates (10a) would then imply infinitely large weights.

Consequently, the firing rates at the steady state are homogeneous provided the network connectivity is such that each neuron is part of a loop. In a randomly connected network, the probability for the existence of such loops increases with the connectivity density and becomes almost equal to one above a certain value (i.e., for a sufficiently dense network). This follows since a Hamiltonian cycle (loop going through all neurons of the network) is almost certain to exist when there are more than $10\log(N)$ connections per neuron, in a random network of $N > 10$ neurons (Bondy and Murty 2008, Theorem 18.22). Typical cortical density involves 10^4 excitatory connections per neuron for 10^5 neurons per mm^3, which is compatible with these figures. The existence of loops for every neuron is thus a reasonable assumption for cortical-like neuronal networks, and this assumption will be made throughout this article.

3.2 Fixed points of the weights

A fixed point of the network dynamics denoted by $(\boldsymbol{\nu}^* = \mu\mathbf{e}, J^*)$ must satisfy the following condition on the weight matrix J^* according to (10a)

$$J^*\,\mathbf{e} = \frac{\mu - \nu_0}{\mu}\,\mathbf{e}. \tag{17}$$

Similar to the case of full connectivity (Burkitt et al. 2007), this equation characterizes an affine space of dimension $n^J - N$ (recall that n^J is the number of connections). For example, a fully connected network without self-connections corresponds to $n^J = N(N-1)$. A redistribution of the strengths of the incoming weights on each neuron, while keeping the sum of these weights constant at $(\mu - \nu_0)/\mu$, gives all the solutions J^*. In other words, the sum of the elements for each row of any J^* is equal to that constant.

In general, the dimension of the affine hyperplane of the J^* is non-zero, and there is a continuum of fixed points, except for the case where a single loop links all the neurons together (or several disjoint loops). In a single loop, there is only one incoming weight per neuron, which must be equal

to $(\mu - \nu_0)/\mu$ at the equilibrium; in other words, no redistribution is possible.

We recall that the matrix $\mathbb{1}_N - J^*$ must be invertible (cf. Appendix 6); this condition can be enforced by placing bounds on the weights, which we do in the numerical simulations. The weight matrix can then move on a manifold of fixed points denoted by $\mathcal{M}^*$, where the drift of the weights arising from the learning equation is zero and the weight evolution is only due to higher orders of the stochastic process (cf. Sect. 2.3.3), similar to the case of full connectivity (Burkitt et al. 2007).

3.3 Stability analysis

We derive from the learning equation (10b) the following linear operator that describes the evolution at the first order of the variation of the weight matrix $\Delta J := J - J^*$ around a given fixed point J^* (Burkitt et al., 2007, Sect. 5)

$$\begin{aligned}\dot{\Delta J} &\simeq \mathcal{L}(\Delta J)\\ &:= -\mu\,\Phi_J\Big[w^{\text{in}}\,\left(\mathbb{1}_N - J^*\right)^{-1}\,\Delta J\,\mathbf{e}\,\mathbf{e}^{\mathbf{T}}\\ &\quad + w^{\text{out}}\,\mathbf{e}\,\mathbf{e}^{\mathbf{T}}\,\Delta J^{\mathbf{T}}\,\left(\mathbb{1}_N - J^*\right)^{-1\,\mathbf{T}}\Big]. \end{aligned} \tag{18}$$

The matrices X such that $X\mathbf{e} = 0$ form a linear subspace of $\mathbb{M}_J$ of dimension $n^J - N$; for ΔJ in this subspace, (18) clearly gives $\dot{\Delta J} = 0$. This corresponds to a displacement along the fixed-point manifold $\mathcal{M}^*$ where the learning equation does not provide any constraint to the leading order, i.e., the drift term of the stochastic weight evolution is zero (Burkitt et al. 2007).

The eigenmatrices related to the linear operator defined in the rhs of (18) depend on the detail of the connectivity topology in a non-trivial way, as described in Appendix 7.2. In addition to $n^J - N$ eigenvalues equal to zero for the matrices $X \in \mathbb{M}_J$ such that $X\mathbf{e} = 0$, the spectrum of the linear operator $\mathcal{L}$ also contains all the eigenvalues of the following matrix L_{r}, as discussed in Appendix 7.2.1,

$$\begin{aligned} L_{\text{r}} &= w^{\text{in}} L_{\text{in}} + w^{\text{out}} L_{\text{out}},\\ L_{\text{in}} &:= -\mu\, R\,\left(\mathbb{1}_N - J^*\right)^{-1},\\ L_{\text{out}} &:= -\mu\,\Phi_J\left[\mathbf{e}\,\mathbf{e}^{\mathbf{T}}\right]\,\left(\mathbb{1}_N - J^*\right)^{-1} \end{aligned} \tag{19}$$

where R is the diagonal matrix whose ith element is the number of incoming connections for neuron i, namely,

$$R = \text{diag}\left[\Phi_J\left(\mathbf{e}\,\mathbf{e}^{\mathbf{T}}\right)\,\mathbf{e}\right]. \tag{20}$$

In the case of homogeneous recurrent connectivity, the spectrum of L_{r} for $|w^{\text{in}}| \gg |w^{\text{out}}|$ is almost the same as that of L_{in}, which lies in the left half-plane as illustrated in Fig. 3a. It follows that the condition $w^{\text{in}} \gg |w^{\text{out}}|$ ensures eigenvalues with large negative real parts for $\mathcal{L}$, as illustrated in Fig. 3c and e. On the contrary, if $|w^{\text{out}}| \gg |w^{\text{in}}|$, then the eigenvalues of $\mathcal{L}$ are almost those of L_{out}. Owing to the large number of almost-zero eigenvalues in the spectrum of L_{out} as shown in Fig. 3b, the eigenvalues of $\mathcal{L}$ are, then, more clustered around zero, and the sign of their real parts may vary according to w^{in}. For $w^{\text{out}} \gg w^{\text{in}} > 0$, the spectrum of $\mathcal{L}$ has eigenvalues with negative real parts (Fig. 3d). However, for $w^{\text{out}} \gg -w^{\text{in}} > 0$, the spectrum of $\mathcal{L}$ has eigenvalues with positive real parts (Fig. 3f). These conclusions on stability are independent of the fixed point J^* used to define $\mathcal{L}$ in (18), which means that the fixed-point manifold $\mathcal{M}^*$ is actually either attractive or repulsive as a whole, i.e., either all fixed points are attractive or none of them is attractive.

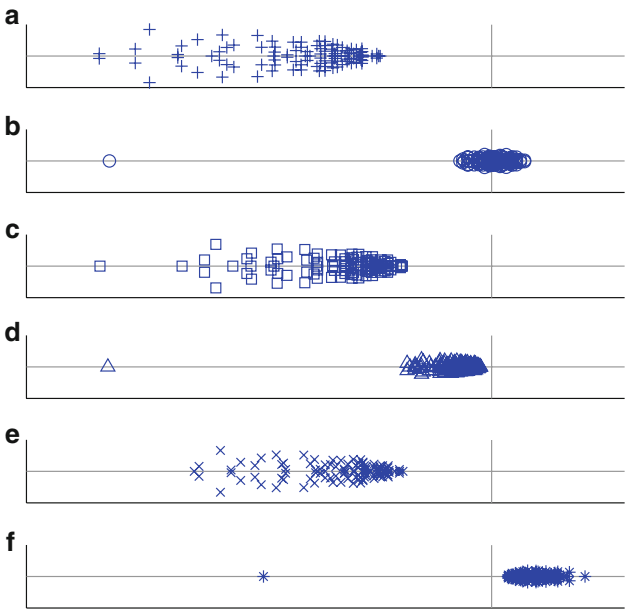


Fig. 3 Illustration of spectra for the linear operator $\mathcal{L}$. The scale is the same for the five plots, and the axes are the gray solid lines. **a** and **b** Spectra of the matrices L_{in} (each eigenvalue is a plus) and L_{out} (circle) defined in (19), for a network of $N = 100$ neurons with homogeneous partial connectivity (30%) and a randomly generated fixed point J^*. These two matrices were re-scaled by $w^{\text{in}} + w^{\text{out}}$. Spectrum of $\mathcal{L}$ for: **c** $w^{\text{in}} = 4$ and $w^{\text{out}} = 1$ (squares); **d** $w^{\text{in}} = 1$ and $w^{\text{out}} = 4$ (triangles); **e** $w^{\text{in}} = 4$ and $w^{\text{out}} = -1$ (crosses); **f** $w^{\text{in}} = -1$ and $w^{\text{out}} = 4$ (stars). The cases **c** and **e** show a spectrum similar to that of L_{in} in **a**, where the eigenvalues have larger negative real parts than in case **d**, whose spectrum is more similar to that of L_{out} in **b**. The case **f** shows many eigenvalues with positive real parts

It follows that in the case of random connectivity, where each neuron has the same number of incoming connections, the condition

$$w^{\text{in}} \gg |w^{\text{out}}| \tag{21}$$

is sufficient to ensure stability. On the other hand, the condition $w^{\text{out}} \gg |w^{\text{in}}|$ leads to weaker stability or even instability when $w^{\text{in}} < 0$. For inhomogeneous connectivity topology, we also expect $w^{\text{in}} \gg |w^{\text{out}}|$ to lead to stability. The condition (21) on the rate-based learning parameters is slightly

stronger than that derived in the case of full connectivity (Burkitt et al. 2007). In order to obtain a realizable equilibrium for the firing rates, i.e., $\mu \geq \nu_0 > 0$ in (16), the condition (21) implies that $\widetilde{W} < 0$, similar to the analysis by Burkitt et al. (2007). Last, the denser the recurrent connections are, the more attractive $\mathcal{M}^*$ is, when the stability conditions on w^{in} and w^{out} are satisfied (details are provided in Appendix 7.2.2).

4 Structural evolution of the recurrent weights

We now look in more detail at the evolution of the individual weights, when the stability conditions for the fixed-point manifold determined in Sect. 3, namely, (21) and $\widetilde{W} < 0$, are met. As shown previously for the case of full connectivity (Burkitt et al. 2007), the recurrent weights tend to individually diverge from each other due to the autocorrelation of the Poisson neurons; their variance increase relates to the fact that we are dealing with stochastic point processes. In this section, we show that this weight dispersion is affected by the connectivity density. Then, we investigate some properties of the asymptotic weight distribution.

4.1 Dispersion of the individual weights

In order to study the impact of recurrent connectivity upon the evolution of the recurrent weights, we use calculations involving the higher stochastic orders of the weight dynamics similar to those in a companion article (Gilson et al. 2009b). The connectivity density affects the weight dispersion, which stems from spike-triggering effects induced by the recurrent connections. This can be studied through the coefficients

$$\Gamma_{i,j,i',j}(t,t') := \left\langle \frac{\mathrm{d}J_{ij}^{\varpi}(t)}{\mathrm{d}t} \frac{\mathrm{d}J_{i'j}^{\varpi}(t')}{\mathrm{d}t} \right\rangle, \tag{22}$$

which are related to the second moment of the weight dynamics. The derivative $\mathrm{d}J_{ij}^{\varpi}(t)/\mathrm{d}t$ of the weight J_{ij} corresponds to one trajectory ϖ of the stochastic process (one realization of the network spiking history), and it consists of weight jumps for each spike and pair of spikes; see Appendix 8 and Gilson et al. (2009b, Sect. 2.3.3) for details. Over a homogeneously connected network with n^J connections, the sum of the contributions to $\sum \Gamma_{i,j,i',j}(t,t')$ of these spike-triggering effects is

$$(n_{\text{av}}^J)^3 J_{\text{av}}\, \mu\, w^{\text{in}} \left(w^{\text{in}} + w^{\text{out}} \right), \tag{23}$$

where $n_{\text{av}}^J = n^J/N$ is the mean number of incoming recurrent connections per neuron. Details are given in Appendix 8, (46); note that η^2 would be present if we had not re-scaled time. Under the stability conditions $w^{\text{in}} \gg |w^{\text{out}}|$ and $\widetilde{W} < 0$ (cf. Sect. 3.3), this sum is positive, which means limiting the increase of the weight variance; this effect is stronger when w^{in} is large and $\widetilde{W}$ is small.

The expression (23) is to be compared with the terms due to the first-order autocorrelation of the neurons. These first-order terms are independent of the connectivity and cause the variance of the recurrent weights J to increase, and, hence, the divergence of individual weights. For a homogeneously connected network, the sum of these terms over all the connections is given by (Burkitt et al., 2007, Eq. 45)

$$n^J \left\{ \mu \left[(w^{\text{in}})^2 + (w^{\text{out}})^2 \right] + \mu^2 \widetilde{W^2} \right\}. \tag{24}$$

The ratio of the expressions in (23) and (24), is given by

$$\frac{(n_{\text{av}}^J)^3 J_{\text{av}} w^{\text{in}} \left(w^{\text{in}} + w^{\text{out}} \right)}{n^J \left\{ \left[(w^{\text{in}})^2 + (w^{\text{out}})^2 \right] + \mu^2 \widetilde{W^2} \right\}} \sim \frac{n^J}{N^2} n_{\text{av}}^J J_{\text{av}}. \tag{25}$$

This ratio ignores the details of the STDP parameters to focus on the connectivity density. We consider $n_{\text{av}}^J J_{\text{av}} = (\mu - \nu_0)/\mu < 1$ to be of order one. The denser the recurrent connections, the closer to one the ratio is. In the case of full connectivity, the ratio is approximately $n_{\text{av}}^J J_{\text{av}}$. This indicates that the weight variance increases more slowly for a network with full connectivity (Fig. 4a) than that with partial connectivity (Fig. 4b).

4.2 Asymptotic pattern of recurrent weights

After a sufficiently long learning epoch, the recurrent weights J have evolved to either saturation or quiescence due to their increasing variance, while J remains on the fixed-point manifold $\mathcal{M}^*$ (Burkitt et al. 2007). The matrix of the weights J exhibits a constant number of saturated weights on each row, when J remains in the attractive manifold $\mathcal{M}^*$, because the sum of incoming weights is then constant for each neuron as discussed in Sect. 3.2. However, the number of potentiated weights on each column, i.e., the sum of the outgoing weights for each neuron, may vary depending on the initial weight distribution, as shown in Appendix 9 and illustrated in Fig. 5. The emerged weight structure in the recurrent network is thus strongly affected by the initial conditions. In other words, this asymptotic structure is not learned by the network in the sense of being constrained by STDP.

In the absence of external inputs, there is no weight structure to learn *per se*, but a structure may still emerge in addition to that remaining from the initial weight distribution, as described previously. For example, starting with full connectivity (except for self-connections), STDP tends to break the synaptic loops of length two between two neurons i and j, i.e., from a neuron to another one and then back to itself $j \to i \to j$. This can be explained by the second stochastic moment for two recurrent weights J_{ij} and J_{ji} in a similar manner to the calculation for the weight dispersion in Sect. 4.1. This second moment is related to $\Gamma_{i,j,j,i}(t,t')$, cf.

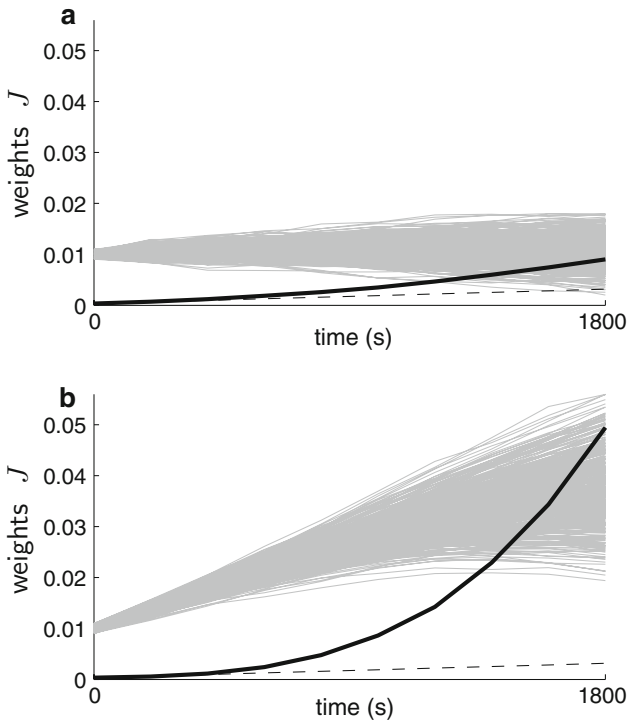


Fig. 4 Comparison of the evolution of the weight variance between two networks of $N = 50$ neurons each with **a** full connectivity and **b** 30% partial random connectivity. The weights were initialized to 0.01 ($\pm 10\%$) for both networks. The individual weights (*gray bundles*) of the fully connected network **a** tended to remain more clustered, whereas those of the partially connected network became more dispersed over time; note that the equilibrium values for the mean weight in the two cases are different because of the connectivity density. After an initial period of linear growth at the predicted rate given in (24) (*dashed line*), the nonlinear increase of the variance (*thick solid lines*, multiplied by a factor 1000 here) of the fully connected network **a** is slower than that for the partially connected network **b**. No weight saturated or became quiescent during this simulation.

(22), and its evaluation during the homeostatic equilibrium leads to a negative expression, namely,

$$-2\mu\left[(w^{\text{in}})^2 + (w^{\text{out}})^2 + w^{\text{in}}w^{\text{out}}\right] < 0, \tag{26}$$

whatever the values for the STDP parameters, as explained in Appendix 8.3. This means that STDP causes the weights J_{ij} and J_{ji} to diverge from each other due to the neuron autocorrelation effects in the network. Consequently, when the equilibrium value for the mean weight and the bounds are set such that half of the weights become saturated and half quiescent, the weight matrix J tends to become antisymmetric as illustrated in Fig. 6.

Other subtle constraints may be imposed on the recurrent weights by STDP, depending on the specific connectivity topology and/or the distribution of delays. However, these are not the primary aim of this article and will not be pursued further here.

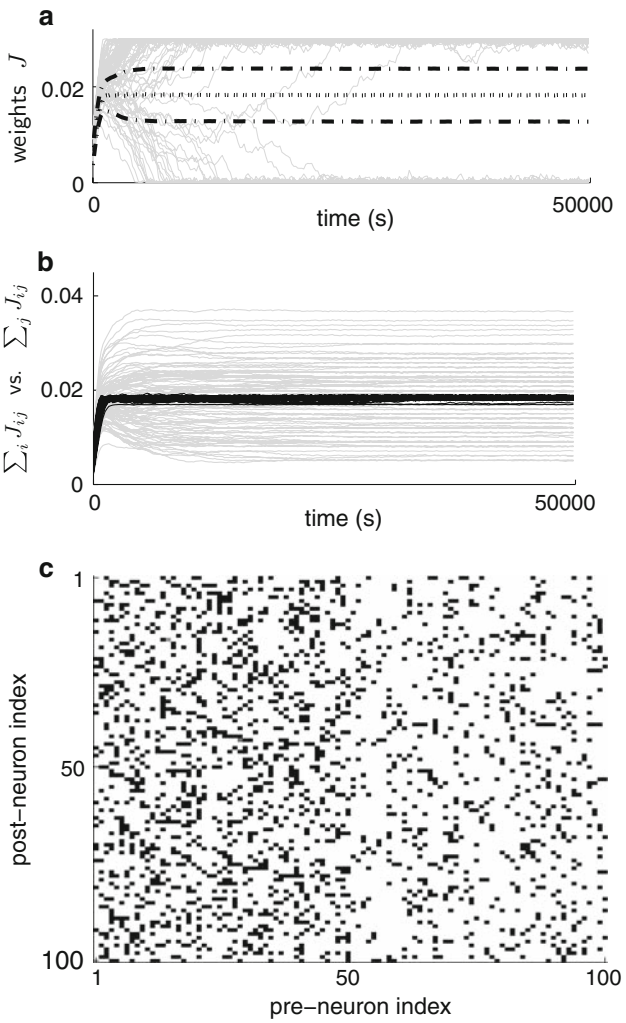


Fig. 5 Evolution of the distributions of incoming and outgoing weights for $N = 100$ neurons. The connectivity was random with 30% probability and the initial weights were respectively set with a spread of $\pm$ 10% around the following values: 0.01 from group 1 to 1; 0.005 from 1 to 2 and from 2 to 1; 0.0025 from 2 to 2. **a** The individual weights (*gray lines*, with only a representative portion plotted) diverged to the bounds. The incoming weight means for groups 1 and 2 (*dotted lines*, almost undistinguishable from each other) quickly converged to the same predicted equilibrium value. The outgoing weight means for groups 1 and 2 (*dotted-dashed lines*) stabilized to different values for a slightly longer period. **b** The means for each neuron of the incoming weights (*black traces*) converged toward the predicted equilibrium value and remained clustered together whereas the means of outgoing weights (*gray traces*) individually stabilized at different values eventually. **c** Matrix of J after 50,000 s of learning. Darker elements indicate potentiated weights. The left side corresponding to weights coming from group 1 had more potentiated weights than the right side (weights from 2) at the end of the simulation, similar to the initial conditions

5 Discussion

In this article, we have analyzed the learning dynamics induced by STDP in a recurrent neuronal networks with no external inputs, i.e., driven only by spontaneous activity related to the background excitation/inhibition ν_0, which

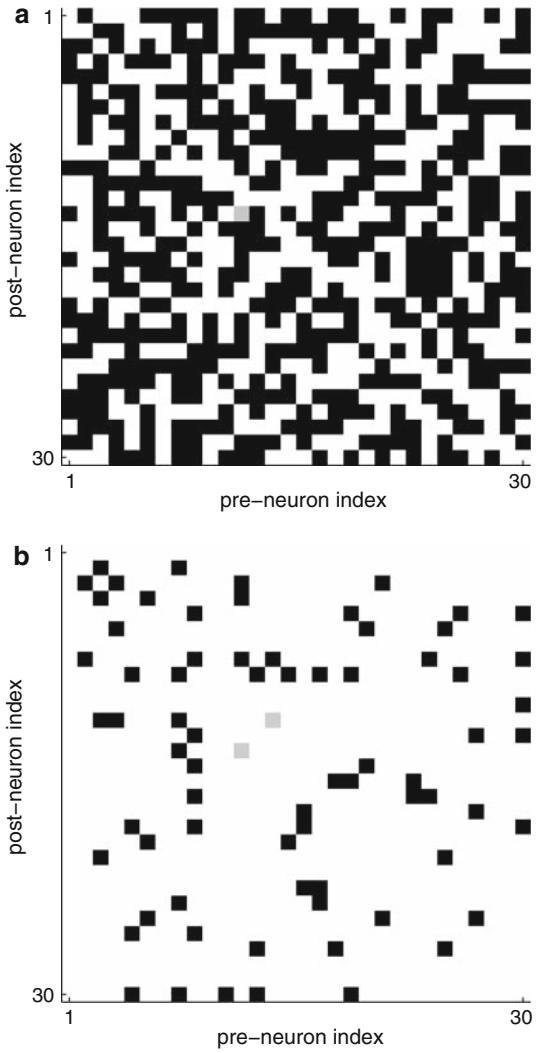


Fig. 6 Illustrative results of numerical simulations with $N = 30$ neurons. Dark pixels indicate potentiated weights. The spontaneous rate ν_0 has been set to 22 Hz to obtain an equilibrium $J_{\text{av}} = 0.0177$ roughly equal to half of the weight bound $\Theta = 0.3$ such that the asymptotic matrix J has almost as many saturated as quiescent weights. **a** Anti-symmetric pattern observed in the asymptotic matrix J when starting from initial full connectivity. For each pair of weights J_{ij} and J_{ji}, STDP almost always depressed one while potentiating the other, resulting in the breaking of most synaptic loops of length two in the network except for **b** 37 pairs of indices (i, j) and (j, i) (*dark pixels*). This corresponds to 4.25% of the total number of initial loops, which is to be compared with the expectation value of discrepancies $J_{\text{av}}/\Theta - 0.5 = 9\%$; note that, if the matrix J is completely antisymmetric, half of the weights will be saturated and the other half will be quiescent, and, hence, $J_{\text{av}} = \Theta/2$. The actual weight distribution is closer to an antisymmetric matrix than the theoretical prediction

is common to all neurons. We focused on the case of partial connectivity, extending our previous study for full connectivity (Burkitt et al. 2007). Using the Poisson neuron model (Kempter et al. 1999) in a framework developed previously by Burkitt et al. (2007), we carried out a fixed-point analysis on the expectation value of the firing rates and weights, as well as a study of higher stochastic orders of the weight dynamics to evaluate their dispersion due to the competition generated by STDP.

Stability for the neuron firing rates, and equivalently for the sum of incoming weights, can be obtained for a wide range of STDP parameters, in particular, when $w^{\text{in}} > |w^{\text{out}}|$ (21) and $\widetilde{W} < 0$. Under these conditions, a single pre-synaptic spike increases the weight by a greater amount than the effect of a single post-synaptic spike (either potentiation or depression), and STDP induces more depression than potentiation for uncorrelated inputs: rate-based anti-Hebbian learning rule. These stability conditions are similar to those obtained for full connectivity (Burkitt et al. 2007) and are in agreement with earlier numerical studies using recurrently connected networks of integrate-and-fire (IF) neurons (Song and Abbott 2001). The equilibrium values of the firing rates and of the sum of incoming weights are determined by the STDP parameters only. These conclusions hold for a broad range of neuronal and synaptic parameters (PSP kernel and delays); the discrepancies observed between prediction and simulation results are typically less than 5%. Synaptic delays do not appear in the equation system (10a–10b) and did not affect in numerical simulation the conclusions presented in this article. Note that relieving the restriction about no self-connections would not change the conclusions.

The individual weights exhibit a diverging behavior, which requires the use of bounds to constrain them to a given domain in numerical simulation. Our choice of STDP and neuron models is motivated by the desire to keep the equations tractable. A weight-dependent version of STDP may be used to implement soft bounds on the weights, i.e., a stable weight distribution, similar to the case of learning on feed-forward network architectures (van Rossum et al. 2000; Gütig et al. 2003; Meffin et al. 2006). We showed that the weight dispersion is stronger for lower density of recurrent connections; similar results are expected for other STDP versions. After a sufficiently long learning epoch with additive STDP, all weights are either saturated or quiescent. The asymptotic weight distribution is constrained by STDP through the sums of incoming weights, but it also strongly depends on the initial weight distribution and can be affected by higher stochastic orders when the firing-rate equilibrium is satisfied. For example, because of the autocorrelation of the neurons (independently of the neuron model), STDP tends to eliminate synaptic loops of length two in the network, which introduces asymmetry in the connectivity topology. It is also possible for an initial unbalanced weight distribution to survive the (partial) homogenization of STDP. In particular, for two neuron groups with inhomogeneous coupling (in terms of weights) within and between them, stronger within-group weights in one group do not induce sufficiently synchronized activity to strengthen the between-group connections to the other group. Such increased synchronization within one group does not necessarily lead to homogeneous

weights throughout the whole network. Consequently, the asymptotic weight structure is not completely constrained by learning but rather could partially reflect the initial weight distribution. When individual weights reach the bounds, the qualitative weight distribution tends to remain preserved. The results presented here contrast with the stable unimodal distribution of non-saturated weights observed for some versions of weight-dependent STDP (van Rossum et al. 2000; Gütig et al. 2003; Morrison et al. 2007). A more complete analysis of weight-dependent STDP remains to be carried out: it is beyond the scope of this article, but will be addressed in a subsequent companion article in this series.

The conclusions for the fixed point of the neuron firing rates described above are valid for any neuron model, provided the correlation structure between the neurons is sufficiently weak: the firing rates are then all constrained to the same equilibrium value due to the learning equation (10b). The stability conditions are expected to hold equally well for other neuron models with excitatory synaptic weights. This also implies the stabilization of the mean incoming weight, but their actual equilibrium value depends on the neuronal activation mechanism. The present analysis of the weight drift (rate-based learning) is a first step toward the study of network structure induced by STDP in recurrently connected neuronal networks (Gilson et al. 2009c).

Acknowledgments The authors are greatly indebted to Chris Trengove, Sean Byrnes, Hamish Meffin, Michael Eager, and Paul Friedel for their constructive comments. They are also grateful to Iven Mareels, Konstantin Borovkov, Dragan Nesic, and Barry Hughes for helpful discussions. MG is funded by scholarships from the University of Melbourne and NICTA. MG also benefited from an enjoyable stay at the Physik Department (T35) of the Technische Universität München. LvH gratefully acknowledges a most enjoyable stay at the Department of Electrical and Electronic Engineering at the University of Melbourne. LvH is partially supported by the BCCN–Munich. Funding is acknowledged from the Australian Research Council (ARC Discovery Project #DP0771815). The Bionic Ear Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastucture Support Program.

Appendix

6 Invertibility of $[\mathbb{1}_N - J(t)]$

In Sect. 2.3 and in the previous article (Burkitt et al. 2007), we required that the matrix $[\mathbb{1}_N - J(t)]$ is invertible for all times t since the contrary would imply a divergence of the firing rates, cf. (10a). Actually, the possibility of diverging firing rates is related to the properties of our Poisson neuron model. It can be illustrated with a single neuron with spontaneous rate $\nu_0 > 0$ connected to itself by a scalar weight J. In this case, the synaptic input is $\nu_0 + J\boldsymbol{\nu}$, and the resulting firing rate $\boldsymbol{\nu}$ is determined by

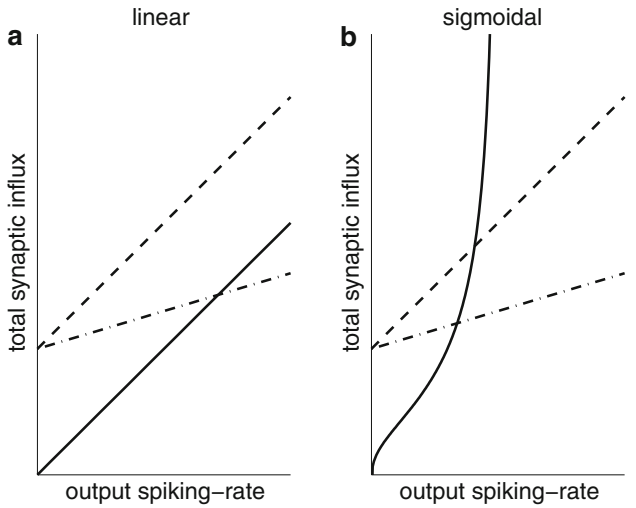


Fig. 7 Illustration of the impact of the activation function upon the weight constraint. This figure compares two Poisson neurons with **a** a linear and **b** a sigmoidal activation function σ. Each neuron is connected to itself with a scalar weight J. The dashed and the dashed-dotted lines correspond to different values of the weight J (resp. 1 and 0.3) in (28), while the solid lines correspond to the activation function. The intersection point determines the firing rate self-consistently constrained by the recurrent loop. For $J = 1$, the left plot has no solution, whereas the right plot does have a solution

$$\boldsymbol{\nu} = \frac{\nu_0}{1 - J}. \tag{27}$$

Provided $0 \leq J < 1$, the firing rate is finite and positive.

This constraint on the upper bound of J is relaxed if, instead of our version of the Poisson neuron model, we introduce an upper bound on the firing rate $\boldsymbol{\nu}$. For example, we may use a sigmoidal-like function σ such that the firing rate is defined by the self-consistency relation

$$\boldsymbol{\nu} = \sigma\,(\nu_0 + J\boldsymbol{\nu}). \tag{28}$$

This gives a solution of $\boldsymbol{\nu}$ for any value of $J \geq 0$, as illustrated in Fig. 7.

This can be extended to the case of several recurrently connected neurons, where J is a matrix. When σ is the identity function, the spectrum of J must then be within the unit circle. This condition on the spectrum relates to the expansion of $(\mathbb{1}_N - J)^{-1}$ in a power series, which is well defined for eigenvalues whose absolute values are strictly less than one.

A bounded activation function σ allows us to remove the upper bounds on the weights. However, the framework of this article exploits the linearity of the Poisson neuron model to make the analysis tractable. The qualitative behavior is expected to be the same for the Poisson neuron model with nonlinear activation function in terms of equilibria and stability, provided the activation function σ is continuous, increasing, and bounded. In this case, (28) always has a unique bounded solution $\boldsymbol{\nu}$ for any given value of J.

In simulations, an explicit bound on the weights J was introduced (generally around 0.9 for the sum of incoming recurrent weights). This ensured that $[\mathbb{1}_N - J(t)]$ remained invertible at all times. Consequently, for certain parameter values, the simulations showed some discrepancies from the analytical predictions.

7 Equilibrium induced by STDP

This appendix contains a number of derivations whose results are discussed in Sect. 3.

7.1 Fixed point of the firing rates in the presence of recurrent loops

We define the function q as in (14)

$$q(x) = -\frac{w^{\text{in}} x}{w^{\text{out}} + \widetilde{W} x}. \tag{29}$$

For a synaptic loop of length n, the firing rate ν_i of any neuron i within the loop satisfies $q^{\{n\}}(\nu_i) = \nu_i$ where $q^{\{n\}}$ denotes the nth iteration of the self-composition of q, i.e., $q^{\{n\}} := q \circ \ldots \circ q$.

For each $n \geq 0$, the function $q^{\{n\}}$ has a fractional form $ax/(b + cx)$ (proof by recurrence; a, b, and c depend on n). Thus, it has two fixed points at most, determined by the quadratic equation $ax - x(b + cx) = 0$. Since q has two fixed points, $q^{\{n\}}$ has the same two fixed points: 0 and $\mu := -(w^{\text{in}} + w^{\text{out}})/\widetilde{W}$.

7.2 Stability of the manifold of fixed points

We study the spectrum of the endomorphism related to the first-order derivative of the learning equation around a given fixed point J^*, defined in (18). In the following analysis, we fix J^* and denote by $\mathcal{L}$ the endomorphism that operates on matrices $X \in \mathbb{M}_J$

$$\mathcal{L}(X) = -\mu\, \Phi_J \Big[w^{\text{in}} \left(\mathbb{1}_N - J^*\right)^{-1} X\, \mathbf{e}\, \mathbf{e}^{\mathbf{T}} + w^{\text{out}}\, \mathbf{e}\, \mathbf{e}^{\mathbf{T}} X^{\mathbf{T}} \left(\mathbb{1}_N - J^*\right)^{-1\,\mathbf{T}} \Big]. \tag{30}$$

Recall that $\mathbb{M}_J$ is the space of $N \times N$ real matrices X such that $\Phi_J(X) = X$, i.e., matrices with non-zero elements only for indices (i, j) corresponding to an existing connection $j \to i$ in the network. The dimension of $\mathbb{M}_J$ is equal to the number of recurrent connections n^J. $\mathcal{L}$ has at least $n^J - N$ eigenmatrices related to the eigenvalue 0, since any matrix X such that $X\mathbf{e} = 0$ implies $\mathcal{L}(X) = 0$.

If the real parts of all eigenvalues in the spectrum of $\mathcal{L}$ are negative (i.e., in left half of the complex plane), then the fixed point J^* is stable. When all the J^* have negative real-part eigenvalues, then the fixed-point manifold $\mathcal{M}^*$ is attractive. Any J^* with one eigenvalue or more in the right half-plane will be unstable.

7.2.1 Decomposition of $\mathcal{L}$

We now study the N remaining eigenmatrices that do not correspond to the subspace $X\mathbf{e} = 0$. The columns of the matrix $(\mathbb{1}_N - J^*)^{-1}$ are denoted by the N-column vectors $\mathbf{g}_i$ for $1 \leq i \leq N$, namely, $\mathbf{g}_i = (\mathbb{1}_N - J^*)^{-1}\, \mathbf{x}_i$ with $\mathbf{x}_i$ the ith N-column vector of the canonical basis of $\mathbb{R}^N$ (with all elements equal to zero except that on the ith row, which is equal to one). We denote by A_{ij} the matrices of the canonical basis of $\mathbb{M}_J$ with all elements equal to zero except the element on the ith row and jth column. For each index i, all the matrices $\mathcal{L}(A_{ij})$ are identical, since $A_{ij}\mathbf{e} = \mathbf{x}_i$; thus, we fix an index $j_i = j(i)$ and one matrix $\breve{A}_i = A_{ij_i} \in \mathbb{M}_J$. For a given i, the identical images $\mathcal{L}(A_{ij})$ can be expressed in terms of the $\breve{A}_{i'}$ with $1 \leq i' \leq N$ and a matrix $Z(i) \in \mathbb{M}_J$ such that $Z(i)\mathbf{e} = 0$,

$$\begin{aligned}\mathcal{L}(A_{ij}) &= \mathcal{L}(\breve{A}_i) \\ &= -\mu\, \Phi_J \left[w^{\text{in}}\, \mathbf{g}_i\, \mathbf{e}^{\mathbf{T}} + w^{\text{out}}\, \mathbf{e}\, \mathbf{g}_i^{\mathbf{T}} \right] \\ &= -\mu \sum_{i'} \left\{ w^{\text{in}} \mathbf{x}_{i'}^{\mathbf{T}}\, \Phi_J \left[\mathbf{g}_i\, \mathbf{e}^{\mathbf{T}} \right] \mathbf{e} \right. \\ &\qquad \left. + w^{\text{out}} \mathbf{x}_{i'}^{\mathbf{T}}\, \Phi_J \left[\mathbf{e}\, \mathbf{g}_i^{\mathbf{T}} \right] \mathbf{e} \right\} \breve{A}_{i'} + Z(i). \end{aligned} \tag{31}$$

The matrix $Z(i)$ corresponds to the specific redistribution of the coefficients of $\mathcal{L}(A_{ij})$, where all the elements on each row i' are summed to form the coefficient of the element $\breve{A}_{i'} = A_{i'j_{i'}}$ (for a matrix X, the corresponding sum is $\mathbf{x}_{i'}^{\mathbf{T}} X \mathbf{e}$). This redistribution for each $\mathcal{L}(A_{ij}) = \mathcal{L}(\breve{A}_i)$ only depends on i and not on j. In other words, we reduce the dimensionality of $\mathbb{M}_J$ and work with classes of equivalent matrices $\Delta J \in \mathbb{M}_J$ that induce the first-order drift $\dot{\Delta J} \simeq \mathcal{L}(\Delta J)$, defined modulo the subspace $\{X \in \mathbb{M}_J, X\mathbf{e} = 0\}$.

Therefore, we can express the endomorphism $\mathcal{L}$ in the basis of $\mathbb{M}_J$ consisting of the N matrices $\breve{A}_i$, and a linearly independent family of $n_J - N$ matrices $X \in \mathbb{M}_J$ such that $X\mathbf{e} = 0$ to complete the basis

$$\mathcal{L} \sim \begin{pmatrix} L_{\text{r}} & 0 \\ L_Z & 0 \end{pmatrix}, \tag{32}$$

where we assimilate $\mathcal{L}$ with its matrix in the basis defined above. The $(n^J - N) \times N$ matrix L_Z is the expression of the $Z(i)$ of (31) in the subbase of $\{X \in \mathbb{M}_J, X\mathbf{e} = 0\}$. The $N \times N$ matrix L_{r} is given by

$$(L_{\text{r}})_{ij} = -\mu \left(w^{\text{in}} n_i^J (\mathbf{g}_j)_i + w^{\text{out}} \sum_{i' \to i} (\mathbf{g}_j)_{i'} \right). \tag{33}$$

The matrix element $(L_{\text{r}})_{ij}$ corresponds to the expression of $L_{\text{r}}(\breve{A}_i)$ in terms of the $\breve{A}_{i'}$ in (31); note that i and j in (33)

correspond to the indices i' and i resp. in (31). Note that $(\mathbf{g}_j)_i$ is the i^{th} element of the vector $\mathbf{g}_j$ defined above, i.e., the element of $(\mathbb{1}_N - J^*)^{-1}$ for indices (i, j). The sum $\sum_{i'\to i}$ is a sum over all i', such that there exists a connection from i' to i; n_i^J is the number of incoming connections of neuron i. This decomposition allows us to study the non-zero spectrum of $\mathcal{L}$, which coincides with that of L_{r} according to (32), excluding the $n^J - N$ eigenvalues equal to zero. From (33), we obtain (19), where R defined in (20) is the diagonal matrix with i^{th} element equal to n_i^J. Note that for the case of full connectivity except for self-connections, this links to the analysis by Burkitt et al. (2007).

7.2.2 Homogeneous connectivity topology

The matrix R in (20) can be approximated by $n_{\text{av}}^J \mathbb{1}_N$ in the case of random connectivity with roughly the same number $n_{\text{av}}^J = n^J/N$ of incoming connections per neuron. It follows that $L_{\text{in}} \simeq -\mu n_{\text{av}}^J (\mathbb{1}_N - J^*)^{-1}$. The spectrum of J^* is assumed to be in the unit circle at all times (cf. Appendix 6), which means that the spectrum of $(\mathbb{1}_N - J^*)^{-1}$ lies in the right half of the complex plane. Since $\mu > 0$, the spectrum of L_{in} (crosses in Fig. 3a) is in the left half-plane, i.e., its eigenvalues have negative real parts. The spectrum of L_{out} (circles in Fig. 3b) contains $N - 1$ eigenvalues roughly equal to zero due to the presence of $\Phi_J[\mathbf{e}\,\mathbf{e}^{\mathbf{T}}]$ (it is strictly zero for full connectivity except for self-connections), and one non-zero eigenvalue related to the eigenvector $\mathbf{e}$ given by

$$-\frac{\mathbf{e}^{\mathbf{T}}\, L_{\text{out}}\, \mathbf{e}}{N} \simeq -\frac{n_{\text{av}}^J \mu^2}{\nu_0} < 0, \tag{34}$$

which also lies in left half-plane. We have used the approximation $\Phi_J[\mathbf{e}\,\mathbf{e}^{\mathbf{T}}]\mathbf{e} \simeq n_{\text{av}}^J \mathbf{e}$.

The discussion about the spectrum L_{r} depending on the values of w^{in} and w^{out} is detailed in Sect. 3.3. For $w^{\text{out}} > 0$ and $w^{\text{in}} > 0$, we expect the spectrum to remain in the left half-plane, contained within the convex hull of the spectra of L_{in} and L_{out} expanded by the scale factor $w^{\text{in}} + w^{\text{out}}$. The conclusions on the stability are the same for all fixed points J^*, and hence they determine whether the whole fixed-point manifold $\mathcal{M}^*$ is attractive or not. Denser recurrent connectivity also gives larger positive values of n_{av}^J in $R \simeq n_{\text{av}}^J \mathbb{1}_N$ and in (34). This implies stronger stability of the fixed points J^* when the conditions on w^{in} and w^{out} are met.

8 Second order of the stochastic evolution of the weights

In this appendix, we provide details of calculations useful to evaluate the structural evolution of the recurrent weights due to STDP, which occurs after the fast convergence toward the homeostatic equilibrium described in Sect. 3. The weight dispersion can be related to the second moment of the stochastic evolution of the weight matrix J, through the multidimensional matrix $\Gamma(t, t')$ whose elements are defined in (22). We show how the connectivity is involved in the evaluation of this matrix, due to the autocorrelation of the neuron activity.

8.1 Analysis of the matrix $\Gamma(t, t')$

The trace of this matrix was used to evaluate the linear increase of the weight variance due to STDP near the beginning of the learning epoch for $t = t'$ and zero recurrent delays $d_{ij} = 0$ (Burkitt et al. 2007). The variance is the expectation value of the trace of the matrix product involving the derivative of J,

$$\begin{aligned}
&\text{Var}(J)(t) \\
&= \left\langle \frac{1}{n^J - 1} \sum_{j\to i} \left[J_{ij}(t) - J_{\text{av}}(t)\right]^2 \right\rangle \\
&= \frac{1}{n^J - 1} \Big\langle \text{trace} \Big\{ \Big[J(t) - J_{\text{av}}(t)\Phi_J(\mathbf{e}\mathbf{e}^{\mathbf{T}})\Big] \\
&\qquad \Big[J(t) - J_{\text{av}}(t)\Phi_J(\mathbf{e}\mathbf{e}^{\mathbf{T}})\Big]^{\mathbf{T}} \Big\} \Big\rangle,
\end{aligned} \tag{35}$$

where $\sum_{j\to i}$ is the sum over the existing connections. When the network is at the homeostatic equilibrium, the mean weight over the network (considered "deterministic") satisfies $J_{\text{av}}(t) = \text{const.}$; it follows that the growth rate of the weight variance is given by

$$\begin{aligned}
\frac{\text{d}\text{Var}(J)}{\text{d}t}(t) &= \frac{2}{(n^J - 1)t} \int_0^t \text{trace} \\
&\quad \times \left\langle \frac{\text{d}J^{\varpi}(t)}{\text{d}t} \left[\frac{\text{d}J^{\varpi}(t')}{\text{d}t}\right]^{\mathbf{T}} \right\rangle \text{d}t',
\end{aligned} \tag{36}$$

where $\frac{\text{d}J^{\varpi}(t)}{\text{d}t}$ denotes here the derivative of the weight matrix J for one stochastic trajectory; it is different from the drift $\dot{J}(t)$ (expectation value). Note that before the homeostatic equilibrium is reached, the variance will evolve both due to deterministic and stochastic contributions depending on the initial value of the variance if the weights are not homogeneous at the beginning of the learning. The stochastic part can then be evaluated using $\Gamma(t, t') - \dot{J}(t)\,\dot{J}^{\mathbf{T}}(t')$ instead of $\Gamma(t, t')$ alone in (36).

The non-diagonal elements of $\Gamma(t, t')$ can also be related to the stochastic dispersion of the weights J. The sign of $\Gamma_{i,j,i',j}(t, t')$ indicates whether the two incoming weights J_{ij} and $J_{i'j}$ of neurons i and i' evolve in the same direction (potentiation or depression): when positive, they tend to both either increase together or decrease together. Sets of weights for which $\sum \Gamma_{i,j,i',j}(t, t')$ (synaptic connections involving indices $i \neq i'$ and j) are more positive will exhibit a smaller dispersion. However, these terms do not directly relate to

the generation of the increasing variance described by (35) (Burkitt et al. 2007).

8.2 Autocorrelation effects on weight dispersion

We consider now the situation of network evolution at the equilibrium, i.e., the weight matrix $J(t)$ is on the manifold of fixed points $\mathcal{M}^*$ at all times without reaching the bounds and its drift $\dot{J}(t) = \langle \frac{\mathrm{d}J^{\varpi}(t)}{\mathrm{d}t} \rangle = 0$ with $\boldsymbol{\nu}(t) = \mu \mathbf{e}$. We want to evaluate the impact of the recurrent connectivity on the evolution of $\Gamma_{i,j,i',j}(t,t')$.

8.2.1 Impact of $J_{ii'}$ on $\Gamma_{i,j,i',j}(t,t')$

Here, we evaluate the effect of the presence of a single recurrent connection $i' \to i$ at the first order of the recurrence, by naively deriving the spike-triggering effects related to $J_{ii'}$ in $\Gamma_{i,j,i',j}(t,t')$. We use similar calculations to those by Gilson et al. (2009b, Appendix A) to evaluate the common evolution of input weights. Applying (6) to express the variation of the weights J_{ij} and $J_{i'j}$ for one stochastic trajectory, we obtain

$$\begin{aligned}
&\frac{\mathrm{d}J_{ij}^{\varpi}(t)}{\mathrm{d}t}\,\frac{\mathrm{d}J_{i'j}^{\varpi}(t')}{\mathrm{d}t} \\
&= \Big[w^{\mathrm{in}} S_j(t-d) + w^{\mathrm{out}} S_i(t) \\
&\qquad + \int W(u) S_i(t) S_j(t+u-d)\,\mathrm{d}u \Big] \\
&\quad \Big[w^{\mathrm{in}} S_j(t'-d) + w^{\mathrm{out}} S_{i'}(t') \\
&\qquad + \int W(u') S_{i'}(t') S_j(t'+u'-d)\,\mathrm{d}u' \Big],
\end{aligned} \tag{37}$$

where we assumed that all the recurrent delays are equal to d. Four terms induced by spike-triggering effects related to $J_{ii'}$ arise from (37) and contribute to $\Gamma_{i,j,i',j}(t,t')$ when taking the ensemble average on (37).

First, $(w^{\mathrm{out}})^2 S_i(t) S_{i'}(t')$ involves $J_{ii'}$ through the dependence of $S_i(t)$ on the past synaptic input history of $S_{i'}(t)$, according to (2), namely, for index $j' = i'$,

$$\rho_i(t) = \nu_0 + \sum_{j'} J_{ij'} \int \epsilon(r) S_{j'}(t-r-d)\,\mathrm{d}r. \tag{38}$$

This leads to an extra contribution due to the autocorrelation of $S_{i'}$ for $j' = i'$ and $t - r - d = t'$,

$$J_{ij'} \epsilon(t-t'-d) \left\langle S_{i'}(t') \right\rangle. \tag{39}$$

Note that this expression is a priori valid only for $t' < t$, but it actually holds in general since $\epsilon(t-t'-d) = 0$ for $t' \geq t$. Second, the term $[w^{\mathrm{out}} S_i(t)][\int W(u') S_{i'}(t') S_j(t'+u'-d)\,\mathrm{d}u']$ gives

$$\begin{aligned}
&w^{\mathrm{out}} J_{ii'} \epsilon(t-t'-d) \int W(u') \left\langle S_{i'}(t') S_j(t'+u'-d) \right\rangle \mathrm{d}u' \\
&\simeq w^{\mathrm{out}} J_{ii'} \epsilon(t-t'-d) \left\langle S_{i'}(t') \right\rangle \\
&\qquad \int W(u') \left\langle S_j(t'+u'-d) \right\rangle \mathrm{d}u',
\end{aligned} \tag{40}$$

where the spike trains $S_{i'}$ and S_j are taken to be independent (we only evaluate the leading order here). Third, the term $[\int W(u) S_i(t) S_j(t+u-d)\,\mathrm{d}u][w^{\mathrm{out}} S_{i'}(t')]$ gives

$$w^{\mathrm{out}} J_{ii'} \epsilon(t-t'-d) \left\langle S_{i'}(t') \right\rangle \int W(u) \left\langle S_j(t+u-d) \right\rangle, \mathrm{d}u. \tag{41}$$

Fourth and last, the term involving the function W twice, i.e., $[\int W(u) S_i(t) S_j(t+u-d)\,\mathrm{d}u][\int W(u') S_{i'}(t') S_j(t'+u'-d)\,\mathrm{d}u']$, gives

$$\begin{aligned}
&J_{ii'} \epsilon(t-t'-d) \left\langle S_{i'}(t') \right\rangle \int\int W(u) W(u') \\
&\quad \left\langle S_j(t+u-d) \right\rangle \left\langle S_j(t'+u'-d) \right\rangle \mathrm{d}u\,\mathrm{d}u'.
\end{aligned} \tag{42}$$

Summing the terms in (39), (40), (41), and (42), we obtain the total contribution to $\Gamma_{i,j,i',j}(t,t')$ due to the single weight $J_{ii'}$, at the leading order:

$$\begin{aligned}
&J_{ii'} \epsilon(t-t'-d) \left\langle S_{i'}(t') \right\rangle \\
&\quad \times \left[w^{\mathrm{out}} + \int W(u) \left\langle S_j(t+u-d) \right\rangle \mathrm{d}u \right]^2.
\end{aligned} \tag{43}$$

The coefficient of $J_{ii'}$ in (43) is positive, which tends to cause the weights J_{ij} and $J_{i'j}$ to evolve in the same direction, either potentiation or depression, cf. Appendix 8.1.

When the network is at the equilibrium, $\nu_i(t) = \mu$ for each neuron i and the time-averaged contribution to the weight coupling $\Gamma_{i,j,i',j}(t,t')$ due to $J_{ii'}$ given in (43) becomes

$$J_{ii'} \mu \left(w^{\mathrm{out}} + \widetilde{W} \mu \right)^2 = J_{ii'} \mu \, (w^{\mathrm{in}})^2, \tag{44}$$

where we have used the normalization of the PSP kernel function ($\int \epsilon = 1$) and the definition of μ in (16).

8.2.2 Impact of J_{ji} on $\Gamma_{i,j,i',j}(t,t')$

Similar to the calculation above, we now evaluate the effect of J_{ji} on $\Gamma_{i,j,i',j}(t,t')$ using (38) and examine the spike-triggering effect due to the autocorrelation of S_i. We find the equivalent to (44) for the time-averaged contribution at the equilibrium,

$$J_{ji} \mu \left(w^{\mathrm{out}} + \widetilde{W} \mu \right) \left(w^{\mathrm{in}} + \widetilde{W} \mu \right) = J_{ji} \mu \, w^{\mathrm{in}} w^{\mathrm{out}}, \tag{45}$$

where we have used (16). The sign of the coefficient of J_{ji} can either be positive or negative here. For example, our choice of parameters corresponds to $w^{\mathrm{in}} > 0$ and $w^{\mathrm{out}} < 0$ (cf.

Appendix 10), so the contribution in (45) is negative in this case.

8.2.3 Link to the density of recurrent connections

It follows from this analysis that the stronger the recurrent connections are in a neuron group, the more its weights tend to evolve together. Weakly connected sets of weights are more likely to exhibit individual weights that evolve in different directions (potentiation vs.depression). For homogeneous recurrent connectivity, where each connection has the probability $n^J/N(N-1)$ of existing (n^J is the number incoming recurrent connections), the lumped effect for the whole network corresponds to the sum of the terms in (44) and (45) for the two possible connections $i' \to i$ and $i \to j$, and for all triplets (i, i', j) when the connections $j \to i$ and $j \to i'$ exist:

$$\begin{aligned}
&\frac{1}{t}\int \sum_{j\to i}\sum_{j\to i'} \Gamma_{i,j,i',j}(t,t')\,\mathrm{d}t' \\
&\quad\simeq \frac{N(N-1)(N-2)(n^J)^3}{[N(N-1)]^3}\, J_{\mathrm{av}}\,\mu\, w^{\mathrm{in}}\left(w^{\mathrm{in}}+w^{\mathrm{out}}\right) \\
&\quad\simeq -(n^J_{\mathrm{av}})^3 J_{\mathrm{av}}\,\frac{w^{\mathrm{in}}\left(w^{\mathrm{in}}+w^{\mathrm{out}}\right)^2}{\widetilde{W}}, \qquad (46)
\end{aligned}$$

where we have used the definition of μ in (16) and taken the limit of a large network ($N \gg 1$ neurons). Recall that $n^J_{\mathrm{av}} = n^J/N$ is the mean number of incoming recurrent connections per neuron. Note that the triplets are ordered so that the triplet (i, i', j) accounts for the connections $j \to i$ and $j \to i'$; the connections $i \to i'$ and $i' \to j$ are taken into account by the triplet (i', i, j). The overall effect is positive provided w^{in} and $\widetilde{W}$ have opposite signs, which is the case where the fixed-point manifold $\mathcal{M}^*$ of the weights J is attractive (cf. Sect. 3.3 and 3.2): $w^{\mathrm{in}} > 0$ and $\widetilde{W} < 0$.

The contributions due to the recurrent connections at the first order of the recurrence are captured by the spike-triggering effects in (46). Similar to the expansion $(\mathbb{1}_N - J)^{-1} = \sum_n J^n$, it is possible to rigorously incorporate higher orders of autocorrelation induced by these spike-triggering effects. Since the network contains only positive weights, all of these effects are positive and they accumulate. The higher-order terms decay exponentially and consequently do not substantially change the result obtained in (46). This can be illustrated for a scalar J such that $1 - J < 1$ with a "safety" margin (J is not too close to 1), where $J/(1-J)$ and J are of the same order.

8.3 Weight evolution for a synaptic loop $j \to i \to j$

Now, we evaluate the effect of STDP on a given synaptic loop of length two between neurons i and j via the evolution of $\Gamma_{i,j,j,i}(t,t')$ defined in a similar manner to (22) with different indices. Similar to (37), we use (6) to express the relative evolution of the weights J_{ij} and J_{ji} for one stochastic trajectory, which relates to

$$\begin{aligned}
&\frac{\mathrm{d}J^{\varpi}_{ij}(t)}{\mathrm{d}t}\,\frac{\mathrm{d}J^{\varpi}_{ji}(t')}{\mathrm{d}t} \\
&\quad= \Big[w^{\mathrm{in}} S_j(t-d) + w^{\mathrm{out}} S_i(t) \\
&\qquad + \int W(u) S_i(t) S_j(t+u-d)\,\mathrm{d}u\Big] \\
&\qquad \Big[w^{\mathrm{in}} S_i(t'-d) + w^{\mathrm{out}} S_j(t') \\
&\qquad + \int W(u') S_j(t') S_i(t'+u'-d)\,\mathrm{d}u'\Big]. \qquad (47)
\end{aligned}$$

We consider the network to be at the homeostatic equilibrium to evaluate the effects due to the autocorrelation of the neurons, $\nu_i = \mu = -(w^{\mathrm{in}} + w^{\mathrm{out}})/\widetilde{W}$ for all i, cf. (16). In this case, the leading order of the terms that arise is negative, independent of the learning parameters,

$$\begin{aligned}
&2\left[w^{\mathrm{in}} w^{\mathrm{out}}\mu + (w^{\mathrm{in}} + w^{\mathrm{out}})\widetilde{W}\mu^2\right] \\
&\quad= -2\mu\left[(w^{\mathrm{in}})^2 + (w^{\mathrm{out}})^2 + w^{\mathrm{in}} w^{\mathrm{out}}\right] < 0, \qquad (48)
\end{aligned}$$

since the polynomial in x of the second order $x^2 + ax + a^2$ is always positive for any value of the coefficient a. Note that we did not use the Poisson neuron model here.

9 Dependence of the asymptotic weight distribution on initial conditions

We consider a specific example of evolution of the weights J with full connectivity except for self-connections, so that in this case, Φ_J only nullifies the diagonal terms of its matrix argument. The sums of the outgoing weights for each neuron are given by the elements of the row vector $\mathbf{e}^{\mathbf{T}} J$, which according to (10b) is

$$\begin{aligned}
\mathbf{e}^{\mathbf{T}}\dot{J} &= w^{\mathrm{in}}\,\mathbf{e}^{\mathbf{T}}\mathbf{e}\,\boldsymbol{\nu}^{\mathbf{T}} + w^{\mathrm{out}}\,\mathbf{e}^{\mathbf{T}}\boldsymbol{\nu}\,\mathbf{e}^{\mathbf{T}} + \widetilde{W}\,\mathbf{e}^{\mathbf{T}}\boldsymbol{\nu}\,\boldsymbol{\nu}^{\mathbf{T}} \\
&\quad - \left(w^{\mathrm{in}} + w^{\mathrm{out}}\right)\boldsymbol{\nu}^{\mathbf{T}} - \widetilde{W}\,\boldsymbol{\nu}^{\mathbf{T}}\,\mathrm{diag}(\boldsymbol{\nu}). \qquad (49)
\end{aligned}$$

We consider initial conditions for which the sums on the incoming weights are identical for each neuron, but the sums of the outgoing weights are inhomogeneous: $J\mathbf{e} \propto \mathbf{e}$ but $\mathbf{e}^{\mathbf{T}} J$ is not proportional to $\mathbf{e}^{\mathbf{T}}$. This implies homogeneous firing rates, i.e., $\boldsymbol{\nu} \propto \mathbf{e}$, since $\dot{J}\mathbf{e} \propto \mathbf{e}$ at all times using a similar equation to that above. Then, $\mathbf{e}^{\mathbf{T}}\dot{J}$ reduces to

$$\mathbf{e}^{\mathbf{T}}\dot{J} = (N-1)\nu_{\mathrm{av}}\left(w^{\mathrm{in}} + w^{\mathrm{out}} + \widetilde{W}\nu_{\mathrm{av}}\right)\mathbf{e}^{\mathbf{T}}.$$

Consequently, the sums of the outgoing weights (i.e., on each column of the matrix J) will evolve identically; hence, the

Table 1 Table of simulation parameters

Time step			10^{-4} s
Simulation duration			10^5 s
Poisson neurons			
Instantaneous firing rate	ν_0	=	5 Hz
Synapses			
Rise time constant	τ_A	=	1 ms
Decay time constant	τ_B	=	5 ms
Mean of recurrent delays	d	=	0.4 ms
Spread of recurrent delays			±0.2 ms
Mean of input delays	$\hat{d}$	=	7 ms
Spread of input delays			±1 ms
STDP			
Learning parameter	η	=	5×10^{-7}
Pre-synaptic rate-based coeff.	w^{in}	=	4
Post-synaptic rate-based coeff.	w^{out}	=	–0.5
Potentiation time constant	τ_P	=	17 ms
Potentiation scaling coefficient	c_P	=	15
Depression time constant	τ_D	=	34 ms
Depression scaling coefficient	c_D	=	10

initial discrepancies will remain after the learning stabilizes, when $\nu_{\text{av}} = \mu$.

This example illustrates that STDP does not reorganize the sums of the outgoing weights for each neuron, as it does for the sums of incoming weights to obtain homogeneous neuron firing rates. Similarly, for an initial inhomogeneous vector of firing rates $\boldsymbol{\nu}$, $\mathbf{e}^{\mathbf{T}} J$ will be modified until $\boldsymbol{\nu}$ converges to $\mu\mathbf{e}$, which may cause inhomogeneities to develop even if initially $\mathbf{e}^{\mathbf{T}} J$ is homogeneous. As a result, the asymptotic value of $\mathbf{e}^{\mathbf{T}} J$ is not constrained by STDP and this evolution does not relate to learning per se. A similar conclusion can be drawn for the case of partial connectivity.

10 Simulation parameters

The results in this article were obtained using Poisson neurons simulated in discrete-time with the parameters listed in Table 1, unless stated otherwise. The STDP window function W is given by

$$W(u) = \begin{cases} c_P \exp(u/\tau_P) & \text{for } u < 0 \\ -c_D \exp(-u/\tau_D) & \text{for } u > 0. \end{cases} \tag{50}$$

The PSP kernel ϵ is defined by

$$\epsilon(t) = \begin{cases} \frac{\exp(t/\tau_B) - \exp(t/\tau_A)}{\tau_B - \tau_A} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0. \end{cases} \tag{51}$$

These parameters are in the same range as those used in previous studies (Kempter et al. 1999; Burkitt et al. 2007).

References

Appleby PA, Elliott T (2006) Stable competitive dynamics emerge from multispike interactions in a stochastic model of spike-timing-dependent plasticity. Neural Comput 18(10):2414–2464

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

Bondy A, Murty USR (2008) Graph theory. Springer, Graduate Texts in Mathematics 244

Burkitt AN, Meffin H, Grayden DB (2004) Spike-timing-dependent plasticity: the relationship to rate-based learning for models with weight dynamics determined by a stable fixed point. Neural Comput 16(5):885–940

Burkitt AN, Gilson M, van Hemmen JL (2007) Spike-timing-dependent plasticity for neurons with recurrent connections. Biol Cybern 96(5):533–546

Câteau H, Kitano K, Fukai T (2008) Interplay between a phase response curve and spike-timing-dependent plasticity leading to wireless clustering. Phys Rev E 77(5):051909

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. Nature 383(6595):76–78

Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009a) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks I: input selectivity—strengthening correlated input pathways. Biol Cybern 101(2):81–102

Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009b) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks II: input selectivity—symmetry breaking. Biol Cybern 101(2):103–114

Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009c) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV: structuring synaptic pathways among recurrent connections. Biol Cybern. doi:10.1007/s00422-009-0346-1

Gütig R, Aharonov R, Rotter S, Sompolinsky H (2003) Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. J Neurosci 23(9):3697–3714

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Wiley, New York

Karbowski J, Ermentrout GB (2002) Synchrony arising from a balanced synaptic plasticity in a network of heterogeneous neural oscillators. Phys Rev E 65(3):031902

Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. Phys Rev E 59(4):4498–4514

Levy N, Horn D, Meilijson I, Ruppin E (2001) Distributed synchrony in a cell assembly of spiking neurons. Neural Netw 14(6–7)SI:815–824

Lubenov EV, Siapas AG (2008) Decoupling through synchrony in neuronal circuits with propagation delays. Neuron 58(1): 118–131

Masuda N, Kori H (2007) Formation of feedforward networks and frequency synchrony by spike-timing-dependent plasticity. J Comput Neurosci 22(3):327–345

Markram H, Lübke J, Frotscher M, Roth A, Sakmann B (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J Physiol (Lond) 500(2):409–440

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 E 73(4):041–911

Morrison A, Aertsen A, Diesmann M (2007) Spike-timing-dependent plasticity in balanced random networks. Neural Comput 19(6):1437–1467

Morrison A, Diesmann M, Gerstner W (2008) Phenomenological models of synaptic plasticity based on spike timing. Biol Cybern 98(6):459–478

Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity Source. J Neurosci 26(38): 9673–9682

Sjöström PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32(6):1149–1164

Song S, Abbott LF (2001) Cortical development and remapping through spike timing-dependent plasticity. Neuron 32(2):339–350

van Rossum MCW, Bi GQ, Turrigiano GG (2000) Stable Hebbian learning from spike timing-dependent plasticity. J Neurosci 20(23):8812–8821