Combined effects of LTP/LTD and synaptic scaling in formation of discrete and line attractors with persistent activity from non-trivial baseline

Timothee Leleu · Kazuyuki Aihara

Received: 3 February 2012/Revised: 13 May 2012/Accepted: 28 June 2012/Published online: 14 July 2012
© Springer Science+Business Media B.V. 2012

Abstract In this article, we analyze combined effects of LTP/LTD and synaptic scaling and study the creation of persistent activity from a periodic or chaotic baseline attractor. The bifurcations leading to the creation of new attractors have been detailed; this was achieved using a mean field approximation. Attractors encoding persistent activity can notably appear via generalized period-doubling bifurcations, tangent bifurcations of the second iterates or boundary crises, after which the basins of attraction become irregular. Synaptic scaling is shown to maintain the coexistence of a state of persistent activity and the baseline. According to the rate of change of the external inputs, different types of attractors can be formed: line attractors for rapidly changing external inputs and discrete attractors for constant external inputs.

Keywords Persistent activity · Attractor network · LTP/LTD · Synaptic scaling · Chaos

Introduction

A number of recent studies support Hebb's hypothesis (Hebb 1949) that persistent activity in the working memory could be the result of reverberatory excitation, i.e. self-sustained excitation due to feedback connections, in local neuronal populations (Wang 2001; Major and Tank 2004).

These studies suggest that some consciousness related information could be encoded in the firing rates of neuronal groups, or cell assemblies (Hebb 1949; Buzsaki 2010). However, the characterization of these cell assemblies remains difficult (Buzsaki 2010), and requires reconsideration of their original definition in order to understand them within the framework of dynamical systems theory (Buonomano and Maass 2009).

Auto-associative networks (Amari 1972; Hopfield 1982) allow for the modeling of cell assemblies as local minima of a Lyapunov function (or wells with lower energy) rather than modeling them directly by detailing synaptic connectivity. This is possible when the weight matrix in a fully connected network is symmetric. These associative models are well-understood (Hertz et al. 1991), and the assemblies created have simply fixed point attractors, which can be recalled by providing the corresponding (incomplete) external inputs.

However, many properties of working memory do not fit in this framework. Among such properties, we focus on the following ones:

- (1) Recorded persistent activity has low levels of irregular firing (see Compte et al. 2003), and shows time-dependency (Romo et al. 1999). Moreover, it has been argued that baseline activity may not be a fixed-point attractor, but a chaotic one as observed in the olfactory bulb notably (Skarda and Freeman 1987), and squid giant axons (Matsumoto et al. 1987; see review Korn 2003), modulated by some frequency; theta oscillations in the hippocampus (Buzsaki 2002), gamma waves (Bartos et al. 2007), and further perturbation by noise (Faisal et al. 2008).
- (2) Single neurons can reach not only one level as described using fixed point attractor models, but

T. Leleu (✉) · K. Aihara
Graduate School of Engineering, The University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
e-mail: timothee@sat.t.u-tokyo.ac.jp

K. Aihara
Institute of Industrial Science, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

several levels of persistent activity (Romo et al. 1999; Funahashi et al. 1989; Ranck 1985). In the management of analog rather than discrete states, some memories can be better understood as “line” attractors (Compte et al. 2000; Brody et al. 2003), which involve the encoding of a continuum of states.

To explain these observations, various models have been proposed. Some recent models focus on the formation of assemblies rather than the characterization of their existing states (Brunel 1996, 2003; Renart et al. 2007). Long-term potentiation and depression (LTP/LTD) have been identified as important biological mechanisms underlying the creation of assemblies (Sejnowski 1999). LTP/LTD have been shown to lead the network to a bistable regime (Brunel 1996) (baseline and persistent activity) through a saddle-node bifurcation (Brunel 2003). The formation of pair-associative memory and prospective activity (Mongillo et al. 2003) could thus be modeled by considering the effects of plasticity on the dynamics of the network.

Homeostatic plasticity (Turrigiano et al. 1998; Turrigiano and Nelson 2004; Turrigiano 2008; Abbott and Nelson 2000) may act in combination with LTP/LTD and limit the growth of weights. This mechanism could explain the selectivity of neurons, i.e., the shape of the tuning curves in the visual cortex (Bienenstock et al. 1982). Further, this mechanism could stabilize synaptic connectivity that allows the formation of line attractors, and avoid drifts due to small deviations in this distribution of weights (Renart et al. 2003). The combined effects of LTP/LTD and synaptic scaling could thus explain both the continuum of memories and discrete memories.

To understand the formation of these complex memories (parametric (Romo et al. 1999), spatial (Funahashi et al. 1989), pair-associated (Miyashita 1988; Sakai and Miyashita 1991; Naya et al. 1996), sequential (Mushiake et al. 2006), assemblies encoding higher cognition processes such as language (Pulvermüller 1996, etc.), we must consider various biological mechanisms (Durstewitz et al. 2000) that can act in combination at different time scales. The change in the dynamics that characterize the creation of assemblies can then be described by formulating canonical models to explain the underlying bifurcations (Hoppensteadt and Izhikevich 1997).

The previous work on the subject primarily relied on simulations or treated the case of fixed-point attractors (Brunel 1996, 2003; Dauce et al. 1998; Mongillo et al. 2003; Renart et al. 2003; Compte 2006). Few models provide a detailed analytical description of the effects of plasticity from the viewpoint of general attractors (Hoppensteadt and Izhikevich 1997; Siri et al. 2008) (periodic/chaotic, modulated, or disrupted by noise), and

formulate mechanisms that result in the creation of discrete or continuous attractors because of neural plasticity (Renart et al. 2003; Pool and Mato 2010; Hansel and Sompolinsky 1998). The one aspect that all these studies have in common is that the time scale of neural plasticity is a determining factor in the creation of cell assemblies. Intuitively, the mechanisms of plasticity should be tuned properly so that the characteristics of the input patterns can be learnt as opposed to the variations in the patterns due to the chaotic baseline attractor, modulation, or noise. Neural systems must indeed be able to manage these time scales adequately.

We propose to analyze the dynamics associated with the creation of cell assemblies for a general case of a chaotic baseline attractor by implementing LTP/LTD and synaptic scaling and taking into account the different time scales of neural plasticity. The combined effects of these two plasticity principles are shown to result in the creation of both discrete and continuous (periodic/chaotic) attractors that coexist with the (periodic/chaotic) baseline in accordance with the relation between the characteristic time of neural plasticity and the rate of change of the external inputs. We thus illustrate the dependence of memory formation on the relation between the time scales of plasticity phenomena and the external stimuli.

The remainder of this article is organized as follows. First, we quantify the “smooth” deformation of the chaotic baseline attractor induced by learning, by approximating the combined effects of LTP/LTD and synaptic scaling in the case of non-trivial trajectories of activity. The conditions under which the baseline activity does not reach the LTP/LTD threshold (i.e., the system does not learn the baseline attractor characteristics when no external inputs are applied) are given. Then, the bifurcations induced due to this change in neural connectivity are studied. These bifurcations can be involved in the creation of persistent activity (tangent, generalized period-doubling (Kuznetsov 1998), and crisis bifurcations when the baseline is a chaotic attractor). Finally, the relationship between the time scales of neural plasticity and the rate of the external inputs is clarified so as to understand the conditions under which discrete or continuous attractors can be created.

The system described in this paper is very high-dimensional, hybrid, and nonlinear. When analysis is possible, it relies on first order approximations and mean field approximations, and exploits the properties (persistence under perturbations) of normal hyperbolicity of attractors, bifurcation analyses, etc. In other cases, simulations are conducted. Preliminary results can be found in (Leleu and Aihara 2011). The numerical parameters typically used in simulations in this study are given in Appendix 1.

Methods

Network model

The network considered in the following is a local group of M fully connected neurons, possibly as a “hypercolumn” (Buxhoeveden and Casanova 2002). Each unit (or neuron) in the model is the sum of inhibitory interneurons and excitatory pyramidal cells with synapses that share the same characteristic time scale,¹ so that synaptic connections can be either positive or negative.

Chaos in neurons has been reported in squid giant axons (Matsumoto et al. 1987), in which spikes need not necessarily to be all-or-none. This motivated the development of the chaotic neural network model (Aihara et al. 1990), which models chaos in neurons. The latter is based on the Nagumo-Sato model (1972) Caianiello (1961), which originates from the McCulloch-Pitts equations (McCulloch and Pitts 1943) and considers the influence of the refractoriness effect. It can be interpreted as the discretization of the leaky integrate-and-fire neurons (Ibarz et al. 2011), with a reset level after spiking that depends on the state of the neuron before firing. In the following, this chaotic neural network model is used to implement an initial chaotic baseline attractor. Because this study focuses on a canonical description of neural activity, rather than the reproduction of quantitative aspects, the use of this model affords good balance between analytical simplicity and realism.²

The equations describing the time-evolution of unit i are given as follows:

$$y_i(n+1) = F_i(y(n)), \quad (1)$$

$$F_i(y(n)) = ky_i(n) + V_i(n) - \alpha x_i(n) + \theta_i + a_i(n), \quad (2)$$

$$V_i(n) = \sum_{j \neq i} \omega_{ij}(n) x_j(n), \quad (3)$$

$$x_j(n) = f(y_j(n)), \quad (4)$$

where y_i denotes the internal state of the neuron i ; y , the vector of internal states; V_i , the local interaction term;

¹ It is important to note that phenomena that are related to spike timing (such as synchrony of spikes) are not taken into account here; however, they may play an important role in cognition studies (a temporal-coding hypothesis). Synchronous firing may disrupt persistent activity (Compte 2006) because of the different time scales for the receptors (primarily AMPA and GABA_A), and create oscillations (Compte et al. 2003). All receptors are assumed to have an identical time scale for simplicity; thus only asynchronous persistent activity is considered in the following.

² Note that the chaos observed in this network is of a different nature than the one observed in Sompolinsky et al. (1988), Doyon et al. (1994). Here, the average activity over the population of neurons, which is given by the mean field approximation, is chaotic (see Appendix 4).

ω_{ij} , the connection weight from neuron j to i ; $\alpha > 0$, the refractoriness factor; θ_i , the threshold; x_i , the activity of neuron i ; f , the activation function; and a_i , the external inputs. The map f used here is the cumulative distribution function of the standard normal law of mean zero and variance $1/p$ (p represents the steepness parameter). The system is initialized at the step $n = n_0$.

Note that the system can also be described using activity x_i as a variable as follows:

$$x_i(n+1) = G_i(x(n)), \quad (5)$$

$$G_i(x(n)) = f(F_i(y(n))). \quad (6)$$

The neurons are also supposed to be weakly connected (Hoppensteadt and Izhikevich 1997). Although LTP/LTD result in a change of the synaptic efficacies, we observe that synaptic scaling can maintain the weights to bounded values; therefore, the approximation of weakly connected neurons remain true. Initially, the weights and thresholds have a normal random distribution of means μ_ω, μ_θ and standard deviations $\sigma_\omega, \sigma_\theta$, respectively. Moreover, means and standard deviations of weights decrease with the number of neurons M , i.e., $\mu_\omega = \mathcal{O}(\frac{\omega}{M})$ and $\sigma_\omega = \mathcal{O}(\frac{1}{M})$, respectively. Note that the refractoriness factor α should have the same order of magnitude as that of μ_ω .

Long-term synaptic potentiation/depression

It is observed in vivo that LTP/LTD act on a slower time scale than the time scale of neuronal dynamics. When the baseline is irregular (chaotic, modulated, noisy, etc.), the above mentioned property is fundamental. If LTP/LTD depend on an average over the window T_0 of internal states, then the change in weights depends on the following properties:

1. The structure of the chaotic attractor by use of its invariant measure and ergodicity;
2. the frequency f_m of the modulation when T_0 is of the order of magnitude of $\frac{1}{f_m}$;
3. the expected noise value by use of the law of large numbers.

The change in weights depends on the time-averaged internal states over the period T_0 , given as $Y_i(n) = \frac{1}{T_0} \sum_{l=1}^{T_0} y_i(n-l)$. The time-averaged activity over the window T_0 is given as $X_i(n) = \frac{1}{T_0} \sum_{l=1}^{T_0} x_i(n-l)$. In the following, T_1 is the slower time scale at which the weights $\omega_{ij}(n)$ increase ($T_1 \gg T_0$). The following generalized learning is exploited, considering both LTP and LTD (Hoppensteadt and Izhikevich 1997; Siri et al. 2008):

$$\omega_{ij}(n+1) = e\omega_{ij}(n) + \Lambda_{ij}(n), \quad (7)$$

$$\Lambda_{ij}(n) = \Lambda(Y_i(n), Y_j(n)), \quad (8)$$

where e is the decay factor which is set to be $e = 1$ for simplicity, and Λ_{ij} is a function of the time average of the internal states of post-synaptic neuron i and pre-synaptic neuron j . When both i and j have strong averaged internal states, then Λ_{ij} is positive (long-term potentiation). When i alone has strong averaged internal state, Λ_{ij} is negative (long-term depression); otherwise, Λ_{ij} is equal to zero.

For numerical simulations, Λ is chosen as

$$\Lambda(Y_i(n), Y_j(n)) = \lambda \operatorname{sgn}(Y_j(n) - d)h(Y_i(n) - d), \quad (9)$$

where d is the threshold that controls LTP/LTD; h , the Heaviside step function; sgn , the signum function; and λ , a positive parameter.

To model the slow change in weights, with characteristic time T_1 , Λ is considered to decrease with T_1 . Thus, its order of magnitude λ is given as follows because the neurons are also weakly connected:

$$\lambda = \mathcal{O}\left(\frac{\epsilon}{MT_1}\right), \quad (10)$$

where ϵ is a positive constant with magnitude $\epsilon \gg \frac{1}{T_0}$.

This Hebbian-like learning can also be seen as an averaging of the effects of STDP in spiking networks (Burkitt et al. 2004); similar rules are commonly used in the literature (Brunel 1996, 2003; Artola and Singer 1993; Bienenstock et al. 1982; Siri et al. 2008). The ABS model (Artola and Singer 1993) uses two different thresholds for LTP and LTD. The BCM model (Bienenstock et al. 1982) implements homeostatic plasticity by using a suitable adaptation of the threshold controlling LTP/LTD; however, in the following, synaptic scaling is considered.

Homeostatic plasticity

Homeostatic plasticity is the change of synaptic connectivity onto each post-synaptic neuron based on its activity level (Turrigiano et al. 1998). Homeostatic plasticity acts at a slower time scale than LTP/LTD. In the following, only synaptic scaling is considered, under which synapses to a particular neuron are changed by an amount proportional to their strength (Turrigiano et al. 1998). We consider that the target firing level of each neuron is regulated so that it remains within a bounded interval (a “target zone” (Turrigiano and Nelson 2004)). When the time-averaged internal state, over a window of size $T_2 \gg T_1$, of a neuron i reaches the lower or upper boundaries, s_i^- and s_i^+ , respectively, the weights are scaled towards their original magnitudes. Similar to the notation described in section “Long-term synaptic potentiation/depression”, the time-averaged internal state over the window T_2 is given as $\Upsilon_i(n) = \frac{1}{T_2} \sum_{l=1}^{T_2} y_i(n-l)$. It is considered that the change

in weights, when scaling is active³ ($\Upsilon_i(n) > s_i^+$ or $\Upsilon_i(n) < s_i^-$), is as follows:

$$\omega_{ij}(n+1) = \omega_{ij}(n_0) + r(\omega_{ij}(n) + \Lambda_{ij}(n) - \omega_{ij}(n_0)). \quad (11)$$

This equation can also be written as follows:

$$\omega_{ij}(n+1) = \omega_{ij}(n) + \Psi_{ij}(n), \quad (12)$$

$$\Psi_{ij}(n) = \omega_{ij}(n_0)(1-r) + r\Lambda_{ij}(n) + (r-1)\omega_{ij}(n), \quad (13)$$

where $r \approx 1$ is the synaptic scaling ratio. The term $r\Lambda_{ij}(n)$ in Eq. (13) corresponds to the effects of LTP/LTD. To ensure that the modeled synaptic scaling is much slower than LTP/LTD, r is chosen such that the term $(r-1)\omega_{ij}(n)$ is negligible as compared to $r\Lambda_{ij}(n)$. Thus, the following condition is required:

$$\frac{(r-1)\bar{\omega}}{M} \ll \lambda. \quad (14)$$

The magnitude of r is thus given by the following equation:

$$(r-1) \ll \frac{\epsilon}{\bar{\omega}T_1}. \quad (15)$$

External inputs

The considered local group of neurons also receive long-range connections from other neuronal groups. These external inputs a_i are of short duration, i.e., the term a_i is non-null only for a few steps n . In the following, we assume that external inputs have a time scale approximately equal to the neural dynamics, are repeated many times at a frequency $f_i \gg \frac{1}{T_0}$ during the learning, and change at a certain frequency f_c , i.e., the external input changes every T_c steps, with $T_c \gg T_0$. The average external input per window T_0 at step n is given as $A_i(n) \approx a_i(n)f_i$, and is the “perceived” external input. Therefore, here we consider external inputs that change at an intermediate time scale between those of LTP/LTD and synaptic scaling.

The effects of plasticity rules are summarized in Fig. 1.

The external inputs used for the simulations are Mexican hat functions of fixed amplitude, which simultaneously excite and inhibit different populations of neurons. This choice is motivated by the observation of lateral inhibition in the cortex. The external inputs are sufficiently strong to force the internal states $Y_i(n)$ over the threshold d . The shapes of the stimuli are not an important factor in the current analysis, and similar results would be obtained with Gaussian, sinusoidal, or even Heaviside-step functions. The center of the Mexican hat of each external input varies

³ Otherwise, only LTP/LTD are active, see Eq. (7).

Fig. 1 Long-term plasticity occurs if the time-averaged internal state of the post-synaptic unit i is greater than d . If the time average of the pre-synaptic unit j is greater than d , LTP occurs; otherwise, LTD occurs. Moreover, synaptic scaling occurs when the time-averaged internal state of a unit i is above or under the boundaries of the target zone s_i^- and s_i^+ , respectively. The target zones for the units i and j are shown in the figure

randomly (uniformly distributed) over the population of neurons.

Results

Deformation of baseline attractor

The initial baseline attractor can be periodic or chaotic (see Fig. 2 for examples of temporal neuronal activity). When the connectivity of the system is changed slowly, LTP/LTD act to deform the baseline attractor. In the following, we approximate the “smooth” deformation of the attractor by quantifying the changes in the time-averaged internal states $Y_i(n)$ to extract the dependence of $Y_i(n)$ on system parameters. These approximations hold for the generalized LTP/LTD rule described in Eq. (7), but only under particular conditions of the scaling parameters. These approximations can easily be adapted to describe the “smooth” deformation of the baseline attractor due to external inputs.

Approximations of effects of LTP/LTD alone

Because the changes in the synaptic efficacies $\Lambda_{ij}(n)$ depend on $Y_i(n)$, the time average over a window T_0 , it is

necessary to describe the evolution of $Y_i(n)$. However, the interdependence of Y_i 's is complex, and thus, they cannot simply be explicated without considering some approximations.

The change in connectivity over an interval T_1 is of the order of a very small ϵ . We consider that for a small number of steps n , $Y_i(n)$ is approximately time-averaged over the trajectory of a quasi-static dynamical system $\tilde{F}(n)$ ⁴ as follows:

$$\tilde{F}_i(n) = \{F_i \setminus \Lambda_{ij}(m) \approx 0, \forall j, |n - m| \ll T_1\}, \forall i, \quad (16)$$

where $F_i \setminus \Lambda_{ij}(m) \approx 0$ means that, for the step m , the terms $\omega_{ij}(m)$ in the definition of the map F_i (see Eqs. (2) and (3)) are approximatively constant.

Similar to Eq. (16), quasi-static dynamical systems $\tilde{G}_i$ are associated with G_i , $\forall i$. For steps m far from n with $|n - m| \gg T_1$, the cumulative change in a weight ω_{ij} cannot be omitted. It is then convenient to consider the cumulative change in weights between n and $n - T_1$ as a small perturbation of the size $\mathcal{O}(\frac{\epsilon}{M})$ (see Eq. (10)). The quasi-static dynamical system $\tilde{F}(n)$ at step n can thus be modeled as a perturbed dynamical system of $\tilde{F}(n - T_1)$. The perturbed system is defined from Eqs. (1)–(3) as follows:

$$\begin{aligned} y_i(n+1) &= \tilde{F}_i(n)(y(n)), \\ &= ky_i(n) + \sum_{i \neq j} \left(\omega_{ij} + \frac{\epsilon}{M} v_{ij} \right) x_j(n) - \alpha x_i(n) + \theta_i, \\ &= \tilde{F}_i(n - T_1)(y(n)) + \epsilon H_i(y(n)), \end{aligned} \quad (17)$$

where $H_i(y(n)) = \sum_{i \neq j} \frac{v_{ij}}{M} x_j(n)$ and $v_{ij} = \mathcal{O}(1)$ represent the weights of perturbation.

Similarly, $\tilde{G}_i$, the perturbed system of G_i , is defined as follows:

$$\begin{aligned} x_i(n+1) &= \tilde{G}_i(x(n)), \\ &= f(\tilde{F}_i(n)(y(n))), \\ &= f(\tilde{F}_i(n - T_1)(y(n)) + \epsilon H_i(y(n))), \\ &= \tilde{G}_i(n - T_1)(x(n)) + \epsilon H_i(y(n)) f'(\tilde{F}_i(y(n))) + \mathcal{O}(\epsilon^2), \end{aligned} \quad (18)$$

where f' is the derivative of f with $f' = \mathcal{O}(p)$.

Generally, the attractor of the perturbed system with perturbation size ϵ can be completely different than that of the unperturbed system. Under certain conditions, however, it can be expected that the perturbed attractor may lie close (at an ϵ distance) to that of the unperturbed system: (1) if the bifurcations induced by the perturbation are

⁴ An infinite set of maps indexed by the step n are thus considered.

Fig. 2 Examples of trajectories of $x_i(n)$ and $x_j(n)$ for two different neurons i and j of the initial baseline attractor

“soft,”⁵ (2) if the attractor is a normally hyperbolic invariant manifold.⁶

In the case of the system described in this paper, this approximation was motivated by two arguments: (1) LTP/LTD usually lead to a change in the complexity of the attractor of neural dynamics (Siri et al. 2008; Wang 2007). This change often occurs through a period-doubling route to chaos, and successive bifurcations are therefore soft. (2) The chaotic attractor is assumed here to be normally hyperbolic. It is noteworthy here that assumptions weaker than those that structural stability are required.

Given that T_0 is sufficiently large for $x_i(n)$ to approximately describe the entire attractor of the system $\tilde{G}(n)$, under ergodicity, it can be considered that $X_i(n)$ depends on

the structure of the attractor of the quasi-static dynamical system $\tilde{G}(n)$. Then, by considering two dynamical systems $\tilde{G}(n)$ and $\tilde{G}(n - T_1)$, we make the following assumption:

$$\Delta^{T_1} X_i(n) = \mathcal{O}(p\epsilon), \quad (19)$$

where $\Delta^{T_1} X_i(n) = X_i(n) - X_i(n - T_1)$.⁷

Similarly, we can also consider the systems $\tilde{F}_i(n)$ and $\tilde{F}_i(n - T_1)$ and make the following assumption:

$$\Delta^{T_1} Y_i(n) = Y_i(n) - Y_i(n - T_1) = \mathcal{O}(\epsilon). \quad (20)$$

When these conditions are verified for a limited number of steps n , i.e., $n\lambda \ll \bar{\omega}$, and $\bar{\omega}p \ll 1$ (and $\eta \gg \frac{1}{T_0}$), the following approximation holds (see Appendix 2):

$$\Delta^{T_1} Y_i(n) = \frac{1}{1-k} \sum_{i \neq j} \Delta^{T_1} \omega_{ij}(n) X_j(n - T_1) + \mathcal{O}(\eta), \quad (21)$$

where $\eta = \bar{\omega}p\epsilon$.

We verify in Appendix 2 that the approximation indeed holds for small changes to the weight matrix. We will see in section “Approximations of combined effects of LTP/LTD and synaptic scaling” that synaptic scaling assures

⁵ A soft bifurcation induces new stable attractors in a small neighborhood of the old one. For example, tangent bifurcations far from the cusp and subcritical pitchfork bifurcations are hard, whereas supercritical pitchfork bifurcations are soft (Hoppensteadt and Izhikevich 1997).

⁶ For a normally hyperbolic invariant manifold (NHIM), the contraction vectors orthogonal to the manifold are stronger than those along the manifold (Hoppensteadt and Izhikevich 1997). NHIM can be interpreted as the generalization of a hyperbolic fixed point to non-trivial attractors. Formal definitions can be found in Fenichel (1972); Hirsch and Shub (1977), Pesin (2004). A fundamental property of these manifolds is that they are persistent under perturbations, i.e., the perturbed invariant manifold has normal and tangent subspaces that are close to the original manifold. This structural stability assures that an attractor of the perturbed system lies near an attractor of the unperturbed system.

⁷ Note that the assumption in Eq. (19) does not hold for $\Delta^{T_s} X_i(n)$, $T_s \ll T_0$. Indeed, the terms $X_i(n)$ and $X_i(n - T_s)$ average almost identical trajectories, and therefore do not describe the attractors of two different dynamical systems (perturbed $\tilde{G}(n)$ and unperturbed $\tilde{G}(n - T_1)$), but the difference between consecutive steps of the same attractor. In that case, these consecutive steps could be at a distance that is equal to the size of the attractor.

that the change in weights remains small; thus, the approximation remains true for all n . Note that if “hard” types bifurcations or interior crises (Grebogi et al. 1983) occur during the route to chaos, then the approximation may eventually be wrong for some bifurcation points, but true otherwise. We therefore consider only the “smooth” deformation of the baseline.

Approximations of combined effects of LTP/LTD and synaptic scaling

When synaptic scaling is taken into account, a novel boundary where switching (Heemels et al. 2009) can occur is added to the dynamical system. In particular, we consider the time steps n_1 for which the time-averaged internal state $Y_i(n_1)$ crosses either s_i^+ or s_i^- . Because $Y_i(n)$ varies very slowly, we can expect that the internal states cross the boundaries only once, and then be scaled (recurrent modes at the boundaries s_i^+ and s_i^- are not considered here). In this case, it is shown, in Appendix 3, that the approximation given by Eq. (21) can be used to arrive at the following approximation:

$$Y_i(n) = r^{T_1} Y_i(n - T_1) + q(n) + (1 - r^{T_1}) Y_i(n_0) + \mathcal{O}(\eta), \quad (22)$$

where,

$$q(n) = \frac{1}{1-k} \sum_{i \neq j} \left[\sum_{t=1}^{T_1} r^t \Lambda_{ij}(n-t) \right] X_j(n - T_1). \quad (23)$$

The deformation in the direction i is thus dependent on the cumulative correlation between neurons j and i over T_1 time steps and is scaled by a “forgetting” ratio r . Because weights reach a limit value due to synaptic scaling, it can be considered that the approximation of Eqs. (22) and (23) hold for large steps n . In section “Distribution of weights”, we verify this approximation in the case of a constant external input.

Conditions for stability of baseline

If the external inputs have a low frequency f_i (compared to the inverse of the time scale T_0 of LTP/LTD), their time average A_i over T_0 may be small. When the increase in $Y_i(n)$ due to external inputs is smaller than d , i.e., the LTP/LTD threshold, there are no changes in synaptic connectivity. In this case, the “perceived” external input is too weak, or varies too slowly to induce any change in the baseline attractor, i.e., there is no learning.

On the other hand, external inputs repeated at a higher frequency may change synaptic connectivity until deforming the baseline attractor such that, even if no external inputs are applied in future trials, $Y_i(n)$ become larger than

the threshold d of LTP/LTD. In this case, synaptic plasticity is activated even in the absence of an external input and weights are strongly increased, which induces a high activation of the neurons. The condition under which this does not occur is expressed as follows:

$$Y_i(\infty) = \lim_{n \rightarrow \infty} Y_i(n) < d, A_i(n) = 0, \forall n, \quad (24)$$

where $Y_i(n)$ is approximated by Eqs. (22) and (23).

Creation of persistent activity: case of constant input

Distribution of weights

In this subsection, the external input is assumed to be constant. When there are no external inputs, if the condition described in Eq. (24) holds, the synaptic weights are constant, i.e., learning does not happen. The proportion of units activated by an external input is noted P_+ . When an external input is applied to the network at n_0 , $M P_+$ neurons are excited for $Y_i(n_0) > d$, and the other neurons are inhibited (see Fig. 3). According to Eq. (9), self-connections to the excited group (noted C^+) are strengthened, and connections to the inhibited group (noted C^-) are weakened. Then, before synaptic scaling is activated, the approximation of Eq. (21) shows that $Y_i(n)$ increases for $i \in C^+$ and decreases for $i \in C^-$. Therefore, in the next step $n_0 + 1$, the changes in weights will be identical. Thus, using mathematical induction, it can be concluded that the change in synaptic weight is constant in this case; namely $\Delta^{T_1} \omega_{ij}(n) = T_1 \Lambda(Y_i(n_0), Y_j(n_0)) = T_1 \Lambda_{ij}(n_0)$, $\forall n > n_0 + T_1$.

The deformation of the baseline attractor due to LTP/LTD only, as approximated in Eq. (21), can be simplified to the following expression:

$$\begin{aligned} \Delta^{T_1} Y_i(n) &= \frac{1}{1-k} \sum_{i \neq j} T_1 \Lambda_{ij}(n_0) X_j(n - T_1) + \mathcal{O}(\eta), \\ &= \frac{1}{1-k} \sum_{i \neq j} T_1 \Lambda_{ij}(n_0) X_j(n_0) + \mathcal{O}(\eta). \end{aligned} \quad (25)$$

For $n > n_1$, synaptic scaling deforms the baseline attractor as follows (see Eq. (22)):

$$Y_i(n) = r^{T_1} Y_i(n - T_1) + q + (1 - r^{T_1}) Y_i(n_1) + \mathcal{O}(\eta), \quad (26)$$

where,

$$q = \frac{1}{1-k} \frac{1-r^{T_1}}{1-r} \sum_{i \neq j} \Lambda_{ij}(n_0) X_j(n_1). \quad (27)$$

The time-averaged internal states are thus scaled at rate r^{T_1} for T_1 steps towards the limit value $Y_i(\infty)$, which is given as follows when the external input is constant:

Fig. 3 Time average of internal states $Y_i(n)$ when temporally-constant external input a_i is applied to the network. Two groups C^+ and C^- are distinguished according to the level of $Y_i(n)$ in comparison with that of d . The group C^+ corresponds to the neurons excited by the external input. Other neurons constitute the group C^-

$$\begin{aligned}
 Y_i(\infty) &= \frac{q}{1-r^{T_1}} + Y_i(n_1) + \mathcal{O}(\eta), \\
 &= \frac{1}{1-k} \frac{1}{1-r} \sum_{i \neq j} \Lambda_{ij}(n_0) X_j(n_1) + Y_i(n_1) + \mathcal{O}(\eta).
 \end{aligned}
 \quad (28)$$

Note that if the external input is removed, the time-averaged internal state $Y_i(n)$ is scaled at rate r^{T_1} per T_1 steps until it reaches s^+ if the neuron was previously excited over this boundary (or s^- if inhibited).

To verify the validity of Eq. (28), Fig. 4 shows the progression of the states $Y_i(mT_1)$, $m \in \{1, 2, \dots\}$, when a constant external input is applied. After a transitory linear deformation of the baseline attractor, the internal states are indeed scaled by ratio r^{T_1} and converge approximately at the limit of Eq. (28).

Figure 5 shows the deformation of the baseline attractor into two groups C^+ and C^- . These two groups are not distinguishable when learnt external inputs are not applied. Learning results in an increase in the variance of the internal states of neurons.

Mean field approximation

The monotonous growth in weights eventually induces a change in the dynamics of the network owing to the creation of reverberatory patterns with excitations and inhibitions. To reduce the dimensionality of the system and analyze the changes occurring in terms of the bifurcation theory, a macroscopic mean field approximation (Amari 1971) is used. The approximation serves to describe the two populations of

excited and inhibited neurons⁸ using a system of coupled equations. This approximation can be obtained using several methods like spin-glass model (Sompolinsky et al. 1988) and large deviation theory (Moynot and Samuelides 2002). We simply choose to consider that weights are randomized at each time-step n because of noise like chemical noise (Faisal et al. 2008). Appendix 4 contains a description of the considerations from which the coupled equations are derived. The variables of the coupled system are given as $y(n) = \{y_+(n), y_-(n)\}$, where y_+ and y_- are the macroscopic internal states of the populations C^+ and C^- , respectively. The coupled system is given as follows, where F represents the map of the macroscopic mean field approximation:

$$y(n+1) = F(y(n)), \quad (29)$$

$$F(y(n)) = ky(n) + \Omega x(n) + \mu_\theta, \quad (30)$$

where $x(n) = \{g(y_+(n)), g(y_-(n))\}$, $g(u) = f\left(\frac{u}{\sqrt{p^2 + (\frac{\sigma_\theta}{1-k})^2}}\right)$,

$$\Omega = \mu_\omega M \begin{pmatrix} P_+ & (1-P_+) \\ P_+ & (1-P_+) \end{pmatrix} + MP_+ \begin{pmatrix} \gamma^+ & 0 \\ \gamma^- & 0 \end{pmatrix}, \quad (31)$$

and γ^+ and γ^- are the cumulated changes in weights due to learning.

Bifurcation analysis of mean field approximation

To understand the creation of persistent activity under the effect of LTP/LTD, the set of bifurcation points of the

⁸ These populations should not be confused with excitatory and inhibitory populations which are common in neuroscience; for example in the Wilson-Conwan model (1972).

Fig. 4 Verification of the approximation of Eq. (28) for $r^{T_1} = 0.94$. **a** Progression of $L_1(m) = \log(1 - \frac{Y_i(mT_1)}{Y_i(\infty)})$ when m is varied. The slope of the linear regression is -0.067 corresponding to $r^{T_1} \approx 0.9352$. **b** Values of $Y_i(\infty)$ as a function of i and approximation of $Y_i(\infty)$ as given in Eq. (28)

mean field approximation are drawn in the parameter space $\{\gamma, \mu_\theta\}$.⁹ The system given by Eqs. (29)–(31) cannot be solved analytically. However, it is possible to find solutions based on numerical calculation. First, both parameters γ and μ_θ can be expressed as functions of y_+^* and

⁹ For simplicity, we assume $\gamma^+ = -\gamma^- = \gamma > 0$. Similar dynamics to the ones described in the following are also observed for other choices of γ^- and are not restrained to this set of parameters.

y_-^* , respectively, the fixed points of the system ($y^* = F(y^*)$). Then, the eigenvalues λ_1 and λ_2 of the Jacobian matrix of the coupled equations are computed analytically. Calculated at the fixed points y^* , the eigenvalues are functions of only y_+^* and y_-^* , based on the parameterization of γ and μ_θ . Finally, the following conditions for bifurcations are solved in $\{y_+^*, y_-^*\}$ using a numerical solver:

Fig. 5 Deformation of baseline attractors into two groups. **a** Monotonous variation of the time-averaged internal states $Y_i(m T_1)$ under the effect of LTP/LTD only. Each black line or grey dashed line represents the time-averaged internal state of a neuron. According to the initial levels of activity, two groups C^+ and C^- are created (black lines and grey dashed lines, respectively). **b** Resulting weight matrix ω_{ij} with i as the abscissa and j as the ordinate

$$\lambda_1 = 1 \text{ or}$$

$$\lambda_2 = 1 \text{ (a tangent, transcritical or pitchfork bifurcation),} \quad (32)$$

$$\lambda_1 = -1 \text{ or } \lambda_2 = -1 \text{ (a period-doubling bifurcation),} \quad (33)$$

$$\|\lambda_1\| = \|\lambda_2\| = 1 \text{ with complex conjugate } \lambda_1 \text{ and } \lambda_2 \text{ (a Naimark-Sacker bifurcation).} \quad (34)$$

Moreover, to determine the exact nature of the bifurcations in the system and to verify whether the transversality conditions are satisfied, the following two equations are also solved separately, using the given sets of parameters γ and μ_θ :

$$y_+^1 = F_+(y_+^1, y_-^1), \quad (35)$$

$$y_-^2 = F_-(y_+^2, y_-^2), \quad (36)$$

Fig. 6 Sets $\{\gamma, \mu_\theta\}$ for which bifurcations of the system given by Eqs. (29)–(31) occur. The line of crosses denotes the set of period-doubling bifurcation points PD (or PF2, pitchfork bifurcation of the second iterate of the map F), the line of dots denotes the tangent bifurcation points of the first iterate of the map (TB). Pitchfork bifurcations of the second iterate can be subcritical (sub PF2, i.e., involving an unstable period-two cycle) or supercritical (sup PF2, i.e., involving a stable period-two cycle). At the asterisks *, the bifurcation curves change between sub PF2 and sup PF2. The number of fixed points for the second iterate of the map F are represented by different variations of gray. The parameter space is divided into 5 regions: (A) two stable/unstable fixed points for the

second iterate of the map and an unstable fixed point for the first iterate coexist (3 fixed points), i.e., there is only a periodic or a chaotic baseline attractor; (B) two unstable fixed points, two stable/unstable fixed points, and a stable fixed point coexist (5 fixed points), i.e., there is an unstable period-two cycle, a periodic/chaotic attractor (the baseline attractor), and a stable fixed point (the persistent activity); (C) an unstable period-two cycle, two periodic/chaotic attractors (the baseline and persistent activity), and an unstable fixed point (7 fixed points); (D) there is a unique stable fixed point (a non-periodic/non-chaotic baseline attractor); and (E) two stable fixed points and an unstable fixed point for the first iterate of F coexist (3 fixed points)

i.e.,

$$y_+^1 = ky_+^1 + M\mu_\omega(P_+g(y_+^1) + (1 - P_+)g(y_-^1)) + MP_+\gamma g(y_+^1) + \mu_\theta, \quad (37)$$

$$y_-^2 = ky_-^2 + M\mu_\omega(P_+g(y_+^2) + (1 - P_+)g(y_-^2)) - MP_+\gamma g(y_+^2) + \mu_\theta. \quad (38)$$

The solutions of Eqs. (35) and (36) give two curves parameterized by $\{y_+^1, y_-^1\}$ and $\{y_+^2, y_-^2\}$, which intersect at the fixed points of the system.¹⁰ The bifurcation parameters can then be varied at the appearance of fixed points determined by the set of bifurcation points so as to decide on the nature of the bifurcations. Moreover, the number of intersections of the curves described in Eqs. (35) and (36) is the number of fixed points of the system described in Eqs. (29)–(31).

¹⁰ These two curves are the equivalent of nullclines for a system of differential equations.

The values of γ and μ_θ for which a bifurcation occurs (the set of bifurcation points) are detailed in Fig. 6. When μ_θ is fixed and γ is varied (deformation of the baseline attractor), due to the effects of LTP/LTD and synaptic scaling, there are several possible cases (see arrows at the left of Fig. 6). (case 1 with $\mu_\theta = -0.03$): From a fixed point attractor (the baseline), the persistent activity is created via a tangent bifurcation of the first iterate of the map. In an interval in γ , there is bistability between the fixed point attractor of the baseline and that of the persistent activity. The baseline attractor is finally destroyed by a second tangent bifurcation. (case 2 with $\mu_\theta = 0.03$): From a chaotic/periodic baseline attractor, a fixed-point attractor (the persistent activity) is created via a sub-critical pitchfork bifurcation of the second iterate of F . For an interval in γ , the periodic/chaotic baseline coexists with the stable fixed point. Tangent bifurcations of the second iterate of F finally occur. An unstable period-two

cycle and a stable period-two cycle (the baseline attractor) collapse. This case corresponds to a generalized period doubling (Kuznetsov 1998). (case 3 with $\mu_\theta = 0.247$): From a periodic/chaotic baseline attractor, novel stable and unstable periodic orbits appear via tangent bifurcations of higher iterates of F . The stable periodic or chaotic attractors coexist until tangent bifurcations occur or a boundary crisis (Grebogi et al. 1983) induces the destabilization of the baseline attractor when the basin boundary touches the baseline attractor. Note that occurrences of crises cannot be determined by local stability, and require consideration of global stability. Figure 6 shows the number of fixed points

for the second iterate of the map F to illustrate the tangent bifurcations of F^2 leading to such cases. Note that near the point labeled “cusp2” in Fig. 6, two successive tangent bifurcations of the second iterate of F occur.

The evolution of the steady state curves obtained by solving Eqs. (35) and (36) for case 1 is shown in Fig. 7. The first iterate of the map F does not give any information about the bifurcations related to cases 2 and 3, because the bifurcations involved concern the second iterate F^2 of the map. Similar to Eqs. (35) and (36), the steady-state curves are obtained for the second iterate F^2 under cases 2 and 3 as shown in Figs. 8 and 9, respectively. The intersections of

Fig. 7 Evolution of steady-state curves of the first iterate of map obtained by solving Eqs. (35) and (36) for case 1. **a** Initially, there is only a single stable fixed point, corresponding to the baseline activity. **b** Just after a tangent bifurcation, there are three fixed points. The two stable fixed points correspond to the baseline and persistent activity,

respectively. **c** Just after a second tangent bifurcation, there is again only a single stable fixed point because of very strong reverberatory excitation, the baseline activity has disappeared and only the attractor of persistent activity remains

Fig. 8 Evolution of steady-state curves of the second iterates of map for case 2. **a** There are initially three fixed points. Two of these fixed points correspond to the chaotic baseline, and constitute an unstable cycle. The middle fixed point is unstable. **b** The unstable cycle is now a stable period-two cycle. A subcritical pitchfork bifurcation occurs at the middle fixed point, which then becomes stable. There are thus five

*These curves concern the second iterate of the map F .

fixed points; the two newly created fixed points are unstable and correspond to an unstable period-two cycle. The periodic attractor corresponds to the baseline, whereas the stable fixed point to the persistent activity. **c** The stable and unstable period-two cycles collide via tangent bifurcations. Only the stable fixed point remains

Fig. 9 Evolution of steady-state curves of the second iterates of map F for case 3. **a** Initially, there are three fixed points. Two of them correspond to the periodic baseline, and are stable. The middle one is unstable. **b** The two stable fixed points become unstable because of period-doubling bifurcations. Two tangent bifurcations then occur and lead to the creation of a novel attractor. There are thus seven fixed points, including two novel stable and two novel unstable fixed

points. The two novel stable fixed points finally become unstable because of period-doubling bifurcations. There are thus seven unstable fixed points for $\gamma = 0.0044$. **c** An unstable period-two cycle becomes stable by period-doubling bifurcations. Stable and unstable period-two cycles then collide via tangent bifurcations. Only three fixed points remain for $\gamma = 0.006$

these curves give the fixed points for the first and second iterates of the map.

The bifurcation diagrams of the internal state $y_i(n)$, given by microscopic equations (not the mean field approximation), under cases 2 and 3 shown in Fig. 6, are shown in Fig. 10. Only the effects of LTP/LTD are taken into account in this figure. The microscopic states $y_i(n)$ bifurcate as predicted by the macroscopic mean field approximation.

For a given value of μ_θ , the interval of values of γ for which there exists a bistable region with the baseline attractor and the persistent activity is denoted as $K(\mu_\theta)$. It is interesting to note that under case 3 of Fig. 6, when the attractor of persistent activity is created for $\gamma \approx 0.0058$ via a boundary crisis, there is very long chaotic transient (Grebogi et al. 1983).

Growth of the basins of attraction

As we have seen in the bifurcation analysis, if LTP/LTD act for long enough to create very strong self-excitatory connections, the previously created attractors with the persistent activity are destroyed until there is a change in the baseline attractor. This can also be observed by considering the growth of the basins of attraction for all cases shown in Fig. 6. In the case of tangent bifurcations (case 1, see fig 11a), the basin of attraction of the persistent activity grows as γ increases. The loss of the baseline state corresponds to the growth of the basin of the new attractor over the whole phase space. The growth is also observed in

the case of generalized period doubling (case 2). However, the basin boundary is irregular. The basins of attraction are not simply connected and “islands” grow as learning progresses. Under case 3, the basin boundary is fractal around the region where the attractor of persistent activity is located, and becomes regular through a basin boundary metamorphosis (Grebogi et al. 1986). When the baseline attractor touches its basin boundary, it loses its stability. Thus the basin of the attractor of persistent activity does not simply grow as in case 1, but can disappear suddenly.

We have seen in this section that LTP/LTD alone induce a bifurcation that allows for persistent activity. If the effects of long-term plasticity are not controlled, the basin grows until the baseline attractor is destroyed. Synaptic scaling is thus necessary to limit the growth of weights and settle the regime of the network in the bistable region.

Stabilization effect of synaptic scaling

Synaptic scaling causes a convergence of the time-averaged internal states Y_i , and thus of the weights. The limit values of the internal states depend on the relative strength of LTP/LTD compared to that of synaptic scaling, as given in Eq. (28). The parameters can be selected such that the weights of the network converge to the critical values described in the previous section, near the point of a bifurcation to a bistable region (see fig. 12). Thus, the time-averaged internal states $\{Y_i(\infty)\}_i$ are chosen such that the corresponding γ of the mean field analysis is in $K(\mu_\theta)$,

Fig. 10 Bifurcation diagram of the internal state $y_i(n)$ of a neuron i for constant $\omega_{ij}(n)$. The changes in weights are indexed using the bifurcation parameter γ and only LTP/LTD are considered. **a** Case 2 of Fig. 6. The baseline attractor is deformed until a subcritical pitchfork bifurcation of the second iterates of the map (sub PF2) makes the system bistable between a periodic attractor and a fixed point. Through a tangent bifurcation of the second iterates of the map F , the network settles to a stable fixed point. **b** Case 3 of Fig. 6. Due to the effects of LTP/LTD, tangent bifurcations occur and result in the creation of another attractor (persistent activity). When this novel attractor is chaotic, it becomes unstable because of a boundary crisis. The resulting unstable manifold becomes stable again via an interior crisis as the existing attractor touches the stable manifold

namely $\gamma \in K(\mu_\theta)$. This limit deformation is denoted as $\{Y_i^c(\gamma, \mu_\theta)\}_i$. The combined effects of LTP/LTD and synaptic scaling for the corresponding scaling ratio r can then drive the network to this bistable region. This region is stable and feedback connections do not diverge to infinity due to the control afforded by synaptic scaling. Moreover,

the deformation of the baseline attractor persists even after the external input is removed. Thus, the condition for a neural system to exhibit memory through the creation of attractors is as follows:

$$Y_i(\infty) \approx Y_i^c(\gamma, \mu_\theta), \forall i, \quad (39)$$

Fig. 11 Evolution of basins of attraction in space $\{y_+(1), y_-(1)\}$ during learning under a constant external input, from left to right for increasing γ . The basin of persistent activity is colored white, whereas the basin of baseline is colored black. **a** Under case 1, the basin boundary is regular. **b** Under case 2, the basin boundary is irregular at first, and then becomes more irregular in the region of the baseline

attractor as γ increases. **c** Under the case $\mu_\theta = 0.022$ (similar to case 3), the basin boundary is irregular in the region of the persistent activity, and becomes regular as γ increases. The attractors of baseline and persistent activity are denoted by circles and asterisks, respectively. When the baseline attractor touches its basin boundary, it becomes unstable

Fig. 12 Bifurcation diagram of $y_i(n)$ for constant $\omega_{ij}(n)$. The changes in weights are indexed using the bifurcation parameter m . From the baseline chaotic attractor, LTP/LTD lead to a bifurcation (here corresponding to period-doubling). When the external input is fed to the network, the network temporarily reaches the bistable region. When the external input is removed, due to the effect of synaptic scaling, the state of the network converges back near the bifurcation state, to the bistable region. This configuration is stable even if further training is applied. Thus, the basin of attraction of the learnt state remains thus small (see section “Growth of the basins of attraction”), and previously learnt memories are not erased

where $Y_i(\infty)$ represents the limit deformation of the baseline attractor. Figure 12 shows the case where the combined effects of LTP/LTD and synaptic scaling first result in an activity level over $\{Y_i^c\}_i$. After the external input is removed, the synaptic scaling drives the system towards the state $\{Y_i^c\}_i$. The bifurcation diagram shows that the system settles at this final bistable region.

Creation of persistent activity: case of rapidly changing external inputs

It is considered now that external inputs vary rapidly compared to the changes in synaptic efficacies.

Distribution of weights

If external inputs change rapidly in a random manner, the combined effects of LTP/LTD and synaptic scaling can result in very different distributions of weights from that in the case described in section “Creation of persistent activity: case of constant input”. If the centers of the external inputs (Mexican hat functions) are distributed uniformly over a certain interval,¹¹ then the resulting synaptic couplings strengths decrease with the distance to the preferred cue

¹¹ Other distribution functions that induce multiple discrete attractors could be considered. This article focus only on the uniform distribution to show the creation of line attractors.

(Compte et al. 2000; Wang 2001; Pool and Mato, 2010) (see Fig. 13). Thus, the network implements “line” attractors (Hansel and Sompolinsky 1998), which, for example, may encode continuous stimuli in the spatial working memory (Wang 2001; Funahashi et al. 1989).

Similar to section “Distribution of weights”, each external input I always results in an increase in the weights of connections within the group of currently excited neurons (for $i \in C^+(I)$), a decrease otherwise (for $i \in C^-(I)$). Note that because all external inputs are considered identical, P_+ M neurons are excited by an external input I . Considering the distribution of external inputs at a time step n ,

- $P_+ M - 1$ or 1 different external inputs can increase or decrease ω_{ij} due to LTP or LTD, with $li - jl = 1$;
- ...
- $P_+ M - l$ or l different external inputs can increase or decrease ω_{ij} due to LTP or LTD, with $li - jl = l$;
- ...
- No or $P_+ M$ external input can increase or decrease ω_{ij} due to LTP or LTD, with $li - jl \geq P_+ M$;

After many external inputs in a window T_1 , synaptic scaling is not yet activated ($T_2 \gg T_1$). The expected change in weights due to LTP/LTD only, $E[\Delta_{ij}(n)] = \frac{1}{T_1} \sum_{t=1}^{T_1} \Delta_{ij}(t)$, is thus given as follows by applying the law of large numbers and considering the effects of both LTP and LTD:

Fig. 13 Resulting distribution of weights after n steps when external inputs change rapidly. **a** Connection matrix ω_{ij} in terms of i and j represented in grayscale. **b** Section (according to the dashed black

line in the left figure) of the distribution of weights, for $|i - j| = l$ in terms of l for a given i

$$E[\Lambda_{ij}(n)] = \left(P_+ - \frac{2l}{M}\right)\lambda, |i - j| = l, l < P_+M, \quad (40)$$

$$E[\Lambda_{ij}(n)] = -P_+\lambda, |i - j| = l, l \geq P_+M. \quad (41)$$

Note that if the frequency of changes in the external inputs is small, the average weight change does not reach its expected value during the window T_1 . In this case, the continuous attractor usually decomposes into a set of discrete attractors. We consider in the following a case that the deviation around the expected value is very small. When synaptic scaling is taken into account, i.e., $n \gg T_1$, the expected value of the weights for a time step n is given as follows (see Eq. (11)):

$$E[\omega_{ij}(n+1)] = rE[\omega_{ij}(n)] + (1-r)\omega_{ij}(n_0) + rE[\Lambda_{ij}(n)]. \quad (42)$$

Therefore, the limit synaptic connections are given as follows using Eqs. (40) and (41) (see Fig. 13b):

$$\begin{aligned} \lim_{n \rightarrow \infty} E[\omega_{ij}(n)] &= \omega_{ij}(n_0) + \frac{r}{1-r} \left(P_+ - \frac{2l}{M}\right)\lambda, |i - j| \\ &= l, l < P_+M. \end{aligned} \quad (43)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} E[\omega_{ij}(n)] &= \omega_{ij}(n_0) - \frac{r}{1-r} P_+\lambda, |i - j| \\ &= l, l \geq P_+M. \end{aligned} \quad (44)$$

Creation of persistent activity

If the “border effects” of the network are ignored (i.e. i varies from $-\infty$ to $+\infty$) or equivalently if a ring network is considered (Hansel and Sompolinsky 1998), the approximation of Eq. (22) shows that the change in $Y_i(n)$ is identical for all neurons.¹² Contrary to the case of a

constant external input (in section “[Creation of persistent activity: case of constant input](#)”), all neurons here belong to the same group.

In this case, the synaptic weights are not independent, and therefore, a mean field approximation cannot be applied. Simulations show that there is a critical value for which persistent activity appears (see fig. 14). The bifurcation involved can be studied by considering the first two terms in the Fourier decomposition of the synaptic efficacy ω_{ij} (see Hansel and Sompolinsky 1998).

Similar to the case of the constant external input given in section “[Creation of persistent activity: case of constant input](#)”, the proper proportion between LTP/LTD and synaptic scaling allows for excitatory connections to maintain persistent activity. However, the system has a continuum of stable states and the chaotic baseline attractor. Figure 15 shows that an external input produces a “bump” (Wang 2001) of activity around the center of the input (the center of the Mexican hat function), which is persistent, from the chaotic baseline attractor.

Creation of persistent activity: intermediate case

If external inputs change in a random manner, but only at an intermediate rate between the cases of constant input (section “[Creation of persistent activity: case of constant input](#)”) and rapidly changing inputs (section “[Creation of persistent activity: case of rapidly changing external inputs](#)”), a finite set of attractors is created. The distribution of weights is irregular as shown in Fig. 16 compared to the distribution of weights for continuous attractors. The continuous attractor breaks down into a discrete set of attractors (Wang 2001; Pool and Mato 2010). Trajectories which initial conditions are between attractors drift until

¹² Simply considering the symmetry of the change in weights, all neurons are similarly affected by synaptic plasticity.

Fig. 14 Bifurcation diagram of the baseline attractor $y_i(n)$ for constant $\omega_{ij}(n)$. The changes in weights are indexed using the bifurcation parameter m . At the critical step m_c , a novel line attractor is created (not shown here). For $m > m_c$, the system is bistable between a chaotic attractor and a line attractor

Fig. 15 Activity of neuron $x_i(n)$ in terms of i and time steps n in the case of the line attractor for two different inputs in **a** and **b**. The activity level $x_i(n)$ is indicated by levels of gray. The activity profile $x_i(n_f)$ at the final step of the trial n_f is also plotted. Initially, neurons

they settle to the nearest discrete “bump” of activity (see Fig. 17).

Discussion

In this article, we studied the effects of LTP/LTD and synaptic scaling with respect to a non-trivial baseline attractor of

have chaotic activity. After the external input, neurons respond in the shape of a persistent “bump” centered around the center position of the external input on the line attractor

activity. When the baseline is not a fixed point, learning depends upon the instantaneous structure of the attractor. In the case of a chaotic baseline attractor, learning is thus a priori sensitive to initial conditions (Siri et al. 2008; Hoppensteadt and Izhikevich 1997). However, it is known that synaptic connections are modified at a much longer time scale than that of membrane dynamics. Therefore, LTP/LTD and synaptic scaling act with information on the history of network states

Fig. 16 Resulting distribution of weights after n steps when external inputs change at an intermediate rate. **a** Connection matrix ω_{ij} in terms of i and j represented in grayscale. The distribution is more

irregular than in the case of Fig. 13. **b** Section of the distribution of weights along the dashed black line in the left figure, for $|i - j| = l$ in terms of l for a given i

Fig. 17 Activity of neuron $x_i(n)$ in terms of i and time steps n when external inputs change at an intermediate rate for two different inputs in **a** and **b**. The continuous attractor breaks down into a set of discrete attractors. The activity level $x_i(n)$ is indicated by levels of gray. The activity profile $x_i(n_f)$ at the final step of the trial n_f is also plotted.

Initially, neurons have chaotic activity. After the external input, neurons respond in the shape of a persistent “bump” centered around the center position of the external input. This bump shifts to the nearest discrete attractor

(i.e. the structure of the attractor). We have shown that the larger is the averaging window T_0 , the smaller is the uncertainty related to the observable $Y_i(n)$. For states separated over a large time scale T_1 , a large enough window T_0 allows for approximation of the baseline attractor deformation. Under certain conditions, the averaged internal state $Y_i(n)$ is modified in the same direction irrespective of the nature of the attractor (fixed point, chaotic, etc.). The change in weights due to LTP/LTD ultimately deforms the baseline attractor until persistent activity is created due to reverberation.

Synaptic scaling enables the transition of the system state to its bistable region and prevents growth of the basin of attraction of the learnt state so that the basin growth cannot

result in the erasure of other memories. According the relation between the final deformation of the baseline $Y_i(\infty)$ and the threshold for activation of LTP/LTD d , two cases are possible. If $Y_i(\infty) > d$, the steady state of the network after learning is the result of the continual effect of LTP/LTD and synaptic scaling. If $Y_i(\infty) < d$, scaling is only transiently activated after an external input has induced a growth of the time-averaged internal state over d .

Our results suggest that synaptic scaling controls the time-averaged internal state within a target zone (Turri-giano and Nelson 2004) rather than to a target state (Turri-giano 2008) in order for the neural network to reach the bistable regime. Moreover, when the external input

changes slowly, the basin of attraction of the learnt state can grow momentarily before synaptic scaling acts (see Fig. 12), and thus promote the activation of this memory on time scales of the size of T_2 . Synaptic scaling creates a forgetting effect for variations at a time scale larger than that of T_1 at a rate of r^{T_1} per T_1 time units. The model described in this article is thus related to a fading memory. This mechanism could promote recently learnt memories at time scales much larger than short-term plasticity. These hypotheses could be tested by observing experimentally if the mechanisms of synaptic scaling are still active long after learning. In the case of an abnormal ratio between the time scales of LTD/LTD and synaptic scaling, this transient increase in synaptic weights is likely to induce imbalance between excitation and inhibition and possibly result in epileptic seizures (Turrigiano et al. 1998).

We have shown that under a uniform distribution of external inputs, two types of memories (discrete and continuum) can result from the combined effects of LTP/LTD and synaptic scaling. The relation between the time scales of LTP/LTD and synaptic scaling could thus be tuned in normal conditions to promote a specific type of attractor. We have shown that if the frequency of the change in external inputs f_c is small compared with that for LTP/LTD, discrete attractors are created. On the contrary, if f_c is very large, a uniform distribution of external inputs is learnt by means of a line attractor. In the intermediate case, the continuous attractor breaks down into a set of discrete attractors. Trajectories drift to the nearest attractor. This case is more realistic than line attractors since it does not require fine tuning of the system parameters.

It can be considered that the regions of the brain that receive sensory information [visual cortex (Funahashi et al. 1989), central, arcuate sulcus and principal sulcus (Romo et al. 1999), limbic system (Ranck 1985), etc.] usually encode information from sensory organs in the form of a continuum. The stimuli are thus “perceived,” without being classifiable into discrete states (i.e., “fuzzy” states), and change rapidly as new information flows. Contrariwise, higher cognitive functions such as sequence of motor actions (Mushiake et al. 2006) or cell assemblies encoding words (Pulvermüller 1996) tend to be discrete events. They thus might participate in the process of abstraction and be maintained in memory for longer periods to allow for their combination with previously stored memories. If these two memories are the result of the same learning process, it raises the fundamental problem to determine the conditions under which one or the other type is formed. The answer to this question is important for understanding what constitutes “abstraction,” and how the brain can transform a continuum of stimuli into higher cognitive signals.

We also discussed in this article how novel attractors of persistent activity are created when the baseline either is

periodic or chaotic. In the case of a constant external input, the novel attractor encoding for persistent activity can appear via tangent bifurcations (case 1), period-doubling bifurcations (case 2), and tangent bifurcations of the second iterates, or crises (case 3).

Moreover, the attractor of persistent activity can be more complex,¹³ or irregular, than the baseline activity. This attractor can be either periodic or chaotic. It is known that cortical neurons do not fire according to a Poisson process (Softky and Koch 1992) and it has been argued that neural activity can be chaotic (Korn 2003). In this case, the transition from baseline to persistent activity can modify the dynamical regime from low to higher complexity (Rajan et al. 2010). The increase in variability of the interspike interval observed during mnemonic activity in vivo (Compte et al. 2003) and at the onset of a stimulus (Churchland et al. 2010) could be explained by a change in dynamics. We have proposed in the article a putative mechanism for these observations: the transition of the mean field dynamics to an attractor with higher or lower complexity. In a chaotic active state, the network generates reproducible patterns of activity, as observed in cortical networks (Ikegaya 2004). Here, these patterns can result from the intrinsic dynamics in local populations of neurons. Chaotic patterns of activity can be exploited to achieve motor tasks such as motion control (Yoshida et al. 2010) and tracking (Li and Nara 2008).

The chaotic active states are usually robust to noise. However, we have shown that in certain parts of the state space, transitions between attractors are possible due to the fractal nature of the basin boundaries. In case 1, the basin of attraction is regular and the new attractor can be interpreted to be a well of lower potential. This case is usually described in the attractor network formalism (Brunel 2003; Wang 2001). In cases 2 and 3, the basin can be irregular and fractal.¹⁴ The image of a well of energy is thus not representative of the underlying dynamics. If the initial condition of a trial is taken into a certain ball when the basin boundary is fractal, the trajectory has a certain probability of ending in either of the coexisting attractors [final state sensitivity (Ott 1993)]. This phenomenon could introduce non-reproducibility of neural response as observed in recordings of single neurons and allow some transitions between attractors. It is expected that a chaotic active state produces a deterministic pattern of activity until it switches to another attractor when the trajectory approaches a fractal basin tongue (Ott 1993). Switching between neural states can be triggered by noise (Mongillo et al. 2003; Wennekers and Palm 2009), oscillations (Ozaki et al. 2012) or other mechanisms such as short-

¹³ “Complex” refers here to the Kolmogorov-Sinai entropy.

¹⁴ The creation of fractal basin boundaries is also observed for systems with continuous time (Ott 1993), and is not an artifact due to the discrete time used in the current article.

term plasticity (Katori et al. 2011). The existence of fractal basins could be tested experimentally by recording the final state sensitivity of a cortical network.

Further work is needed to determine whether the bifurcations described in this article can also occur in more realistic models such as leaky integrate and fire neurons, quadratic integrate and fire neurons with adaptation which exhibit chaotic solutions (Zheng and Tonnelier 2008) and in biological networks. The effects of short-term plasticity (Abbott and Regehr 2004) or neuro-modulators (Durstewitz et al. 2000) should also be taken into account for modeling both short-term memory and working memory (Durstewitz et al. 2000). To model long-term memory and simulate networks that can create several successive assemblies, consolidation mechanisms such as late LTP (Abbott and Nelson 2000) should also be considered.

Studying the interactions between the time scales of neural plasticity is important for understanding the creation of complex memories. Spatial memory is likely to be a compromise between discrete and continuum attractors; sequential memory, a compromise between stability and instability of attractors (Tsuda 2001).

Acknowledgments This research is supported by the Global Center of Excellence (G-COE) “Secure-Life Electronics” sponsored by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), and by the Aihara Project, the FIRST program from the Japan Society for the Promotion of Science (JSPS), Japan.

Appendix 1: Parameters typically used for simulations

See Appendix Table 1

Table 1 Parameters typically used in the simulations

Notation	Detail	Value
M	Number of neurons	200
α	Refractoriness	−0.1
p	Steepness	33
μ_ω	Mean of weights	−0.58
σ_ω	Standard dev. of weights	0.05
μ_θ	Mean of thresholds	0.03
σ_θ	Standard dev. of thresholds	0.01
T_0	Averaging window for LTP/LTD	200
T_1	Time scale of LTP/LTD	500
T_2	Time scale of synaptic plasticity	1,000
λ	Increment in LTP/LTD	1×10^{-4}
d	Threshold for LTP/LTD	0.3
s_i^-	Lower boundary of the target zone	−0.02
s_i^+	Upper boundary of the target zone	−0.12
$1 - r^{T_1}$	Synaptic scaling ratio	0.0345
P_+	Ratio of excited neurons	0.15

Appendix 2: Approximation of the effects of LTP/LTD alone

In the following, changes due to LTP/LTD are considered. From Eqs. (1)–(3), by taking the time average over $\{n - T_0, \dots, n - 1\}$, we obtain the following expression:

$$\begin{aligned}
 Y_i(n) &= \frac{y_i(n - T_0) - y_i(n)}{T_0} \\
 &+ kY_i(n) + \frac{1}{T_0} \sum_{l=1}^{T_0} \sum_{i \neq j} \omega_{ij}(n - l)x_j(n - l) \\
 &- \alpha X_i(n) + \theta_i + A_i, \\
 &= kY_i(n) + \frac{1}{T_0} \sum_{l=1}^{T_0} \sum_{i \neq j} \omega_{ij}(n - l)x_j(n - l) \\
 &- \alpha X_i(n) + \theta_i + A_i \\
 &+ \mathcal{O}\left(\frac{1}{T_0}\right),
 \end{aligned} \tag{45}$$

where $A_i = a_i f_i$ is the “perceived” external input (see section “External inputs”), i.e., the level of external input averaged over window T_0 .

The changes in connectivity in the window T_0 can be considered to be very small (quasi-constant). Note that for $l < T_0$ and $e = 1$ (see Eq. (7)),

$$\begin{aligned}
 \omega_{ij}(n - l) &= \omega_{ij}(n) - \sum_{t=1}^l \Lambda_{ij}(n - t), \\
 &= \omega_{ij}(n) + \mathcal{O}(T_0 \lambda).
 \end{aligned} \tag{46}$$

By using Eq. (46) in Eq. (45), noting that $\frac{1}{T_0} \sum_{i \neq j} \sum_{l=1}^{T_0} \sum_{t=1}^l \Lambda_{ij}(n - t) = \mathcal{O}(T_0 \lambda M)$, we rewrite Eq. (45) as follows ($T_1 \gg T_0$, thus $\frac{1}{T_0} \gg T_0 \lambda M$):

$$\begin{aligned}
 Y_i(n) &= kY_i(n) + \sum_{i \neq j} \omega_{ij}(n)X_j(n) - \alpha X_i(n) \\
 &+ \theta_i + A_i + \mathcal{O}\left(\frac{1}{T_0}\right).
 \end{aligned} \tag{47}$$

By considering the changes in weights and activity over window T_1 , we rewrite Eq. (47) as follows:

$$\begin{aligned}
 Y_i(n) &= kY_i(n) + \sum_{i \neq j} (\omega_{ij}(n - T_1) + \Delta^{T_1} \omega_{ij}(n))(X_j(n - T_1) \\
 &+ \Delta^{T_1} X_j(n)) - \alpha(X_i(n - T_1) + \Delta^{T_1} X_i(n)) \\
 &+ \theta_i + A_i + \mathcal{O}\left(\frac{1}{T_0}\right).
 \end{aligned} \tag{48}$$

The above equation can also be written as follows:

$$Y_i(n) = Y_i(n - T_1) + \frac{1}{1 - k}(t_1 + t_2 + t_3 + t_4) + \mathcal{O}\left(\frac{1}{T_0}\right), \tag{49}$$

where:

$$t_1 = \sum_{i \neq j} \omega_{ij}(n - T_1) \Delta^{T_1} X_j(n), \quad (50)$$

$$t_2 = \sum_{i \neq j} \Delta^{T_1} \omega_{ij}(n) X_j(n - T_1), \quad (51)$$

$$t_3 = \sum_{i \neq j} \Delta^{T_1} \omega_{ij}(n) \Delta^{T_1} X_j(n), \quad (52)$$

$$t_4 = -\alpha \Delta^{T_1} X_i(n). \quad (53)$$

When the changes in weights are considered over a limited number of steps n such that $n\lambda \ll \bar{\omega}$ and the assumption of Eq. (19) is verified, the following approximations hold, $\forall n$:

$$\omega_{ij}(n) = \mathcal{O}\left(\frac{\bar{\omega}}{M}\right), \quad (54)$$

$$\Delta^{T_1} \omega_{ij}(n) = \mathcal{O}(T_1 \lambda), \quad (55)$$

$$\Delta^{T_1} X_j(n) = \mathcal{O}(p\epsilon). \quad (56)$$

Then the terms of Eqs. (50)–(53) are approximated as follows:

$$t_1 = \mathcal{O}(\bar{\omega} p \epsilon), \quad (57)$$

$$t_2 = \mathcal{O}(\epsilon), \quad (58)$$

$$t_3 = \mathcal{O}(p\epsilon^2), \quad (59)$$

$$t_4 = \mathcal{O}\left(\frac{\bar{\omega} p \epsilon}{M}\right). \quad (60)$$

Finally, because $\bar{\omega} p \ll 1$ and $\bar{\omega} p \epsilon \gg \frac{1}{T_0}$, the approximation of Eq. (21) holds.

To verify the validity of the approximation, the terms t_1, t_2, t_3 and t_4 are compared for several steps $mT_1, m = \{1, 2, \dots\}$, as shown in Fig. 18. Simulations confirm that as long as the change in weights remains small, t_2 is dominant. Note that because $t_2 = \mathcal{O}(\epsilon)$, this approximation is consistent with the assumption of Eq. (20) as applied to the change in $Y_i(n)$.

Appendix 3: Approximation of the combined effects of LTP/LTD and synaptic scaling

When the weights are scaled by synaptic scaling (see Eqs. (12) and (13)), the change in synaptic efficacy $\Delta^{T_1} \omega_{ij}(n) = \omega_{ij}(n) - \omega_{ij}(n - T_1)$ is given as follows:

$$\begin{aligned} \Delta^{T_1} \omega_{ij}(n) &= (1 - r) \omega_{ij}(n_0) + r \Lambda_{ij}(n - 1) + r \omega_{ij}(n - 1), \\ &= (1 - r^{T_1}) \omega_{ij}(n_0) + \sum_{t=1}^{T_1} r^t \Lambda_{ij}(n - t) \\ &\quad + (r^{T_1} - 1) \omega_{ij}(n - T_1). \end{aligned} \quad (61)$$

Fig. 18 Comparison of terms t_1, t_2, t_3 and t_4 of Eqs. (50)–(53), respectively, with the solid, dashed, dotted, and dash-dot lines in terms of steps $mT_1, m = \{1, 2, \dots\}$. So long as the change in weights remains small (i.e., for a small number of time steps m in this figure), the term t_2 is dominant

The change in weights $\Delta^{T_1} \omega_{ij}$ when synaptic scaling is active is of the same magnitude as that when only LTP/LTD modify the weights (see Eq. (14)). Therefore, the approximation of Eq. (21) still holds in this case. Then, the time-averaged internal state can be written as follows using Eqs. (21) and (61) as follows:

$$\begin{aligned} Y_i(n) &= Y_i(n - T_1) \\ &+ \frac{1 - r^{T_1}}{1 - k} \sum_{i \neq j} \omega_{ij}(n_0) X_j(n - T_1) \\ &+ \frac{1}{1 - k} \sum_{i \neq j} \left[\sum_{t=1}^{T_1} r^t \Lambda_{ij}(n - t) \right] X_j(n - T_1) \\ &+ \frac{r^{T_1} - 1}{1 - k} \sum_{i \neq j} \omega_{ij}(n - T_1) X_j(n - T_1) \\ &+ \mathcal{O}(\eta). \end{aligned} \quad (62)$$

By using Eq. (47), we rewrite this equation as follows:

$$\begin{aligned} Y_i(n) &= Y_i(n - T_1) \\ &+ \frac{1 - r^{T_1}}{1 - k} \sum_{i \neq j} \omega_{ij}(n_0) X_j(n - T_1) \\ &+ \frac{1}{1 - k} \sum_{i \neq j} \left[\sum_{t=1}^{T_1} r^t \Lambda_{ij}(n - t) \right] X_j(n - T_1) \\ &+ \frac{r^{T_1} - 1}{1 - k} ((1 - k) Y_i(n - T_1) + \alpha X_i(n - T_1) - \theta_i - A_i) \\ &+ \mathcal{O}(\eta). \end{aligned} \quad (63)$$

Finally, when $X_j(n - T_1)$ remains small due to the scaling, $\omega_{ij}(n_0) X_j(n - T_1) = \omega_{ij}(n_0) X_j(n_0) + \mathcal{O}(\eta)$ (see Eq. (57)). Using this approximation and Eq. (47) once more, we can rewrite Eq. (63) as follows:

$$\begin{aligned} Y_i(n) &= Y_i(n - T_1) \\ &+ \frac{1 - r^{T_1}}{1 - k} ((1 - k) Y_i(n_0) + \alpha X_i(n_0) - \theta_i - A_i) \\ &+ \frac{1}{1 - k} \sum_{i \neq j} \left[\sum_{t=1}^{T_1} r^t \Lambda_{ij}(n - t) \right] X_j(n - T_1) \\ &+ \frac{r^{T_1} - 1}{1 - k} ((1 - k) Y_i(n - T_1) + \alpha X_i(n - T_1) - \theta_i - A_i) \\ &+ \mathcal{O}(\eta). \end{aligned} \quad (64)$$

By developing Eq. (64), we obtain Eq. (22).

Appendix 4: Macroscopic mean field approximation

A single population of neurons is considered for the following demonstration. Generalization to multiple

populations is immediate. For simplicity, the refractoriness is included in the distribution of weights ($\alpha = \omega_{ii}$, and $V_i(n) = \sum_j \omega_{ij} x_j(n)$) and is thus assumed to be normally distributed. This approximation is possible so long as the refractoriness is of the same order of magnitude as μ_ω . Because weights are randomized at each time step n , for a given m , the local chaos hypothesis considered in this study is given as follows:

1. $x_j(n)$ and ω_{ij} are independent $\forall i, j$;
2. $x_j(n)$ and $x_i(n)$ are independent $\forall i \neq j$;
3. $\theta_i(n)$ and $V_i(n)$ are independent $\forall i$.

Internal states and activity are also random variables denoted by y_i and x_i , respectively, such that the system is a stochastic process. In the following, the mean of a random variable X belonging to the probability density function p_X is denoted by $m(X)$, the variance, $v(X)$, covariance of X and Y , $cv(X, Y)$:

$$m(X) = \int x p_X(x) dx; \quad (65)$$

$$v(X) = \int x^2 p_X(x) dx - m(X)^2; \quad (66)$$

$$cv(X, Y) = \int x y p_{X,Y}(x, y) dx dy - m(X) m(Y). \quad (67)$$

The mean of the internal states y_i and activity x_i are simply denoted y and x , respectively. By taking the mean over the neurons i in Eqs. (1)–(3), the mean of internal activity y is given as follows (using the hypotheses of independence described above and the law of large numbers):

$$\begin{aligned} y(n+1) &= ky(n) + \sum_j m(\omega_{ij} x_j(n)) + \mu_\theta, \\ &= ky(n) + Mm(\omega_{ij})x(n) + \mu_\theta. \end{aligned} \quad (68)$$

The mean activity is given as follows:

$$x(n) = \int u p_x(u) dx, \quad (69)$$

$$= \int f(t) p_y(f(t)) dt, \text{ by a change of variable,} \quad (70)$$

$$= f\left(\frac{y(n)}{\sqrt{p^2 + v(y(n))}}\right), \text{ for the particular } f \quad (71)$$

chosen (Amari 1971).

The variance of y_i is given as follows (using the hypotheses of independence described above):

$$\begin{aligned} v(y(n+1)) &= k^2 v(y(n)) + Mv(\omega_{ij} x_j(n)) + \sigma_\theta^2 \\ &+ 2kcv(y_i(n), \theta_i). \end{aligned} \quad (72)$$

Note that by induction, $y_i(n+1)$ is also given as follows:

$$y_i(n+1) = k^{n+1}y_i(0) + \sum_j \sum_{p=0}^n \omega_{ij}x_j(n-p)k^p + \frac{1-k^{n+1}}{1-k}\theta_i. \quad (73)$$

When the number of steps n is very large, this can be expressed as

$$y_i(n+1) = \sum_j \omega_{ij}\tilde{x}_j(n) + \frac{\theta_i}{1-k}, \quad (74)$$

where $\tilde{x}_j(n) = \sum_{p=0}^n x_j(n-p)k^p$. This term is also considered to be non-random as in Amari (1971). Therefore, Eq. (72) can be written as follows:

$$v(y(n+1)) = k^2v(y(n)) + M(\mu_\omega^2v(x(n)) + x(n)^2\sigma_\omega^2 + \sigma_\omega^2v(x(n))) + \sigma_\theta^2 + 2k\frac{\sigma_\theta^2}{1-k}. \quad (75)$$

Because the mean and standard deviation of the weights decrease proportionally to $\frac{1}{M}$ when $M \rightarrow \infty$, the variance is given as follows:

Fig. 19 Mean field approximation. **a** The standard deviation $\sqrt{v}$ of the internal states $y_i(n)$ is proportional to σ_θ ; the coefficient of proportionality (given by the slope of the linear regression) is equal to $\frac{1}{1-k}$, as described in Eq. (77). **b** Distance S between the mean of the microscopic internal states and the macroscopic mean field approximation for increasing number of neurons M . As M increases, the approximation becomes more accurate

$$v(y(n+1)) = k^2 v(y(n)) + \sigma_\theta^2 \frac{1+k}{1-k}. \quad (76)$$

After a few time steps, the variance is constant, and is given as follows:

$$v(y) = \sigma_\theta^2 \frac{1}{(1-k)^2}. \quad (77)$$

To validate the mean field approximation, we first verified the dependence of the variance of $y(n)$ on σ_θ , as shown at the top of Fig. 19. Different from Amari (1971) and Cessac et al. (1994), the mean field can also exhibit periodic/chaotic dynamics. There are thus two types of variations: one due to the mean-field dynamics (“macroscopic”) and another due to the randomness in the connectivity and threshold (“microscopic”). The two types of variation are dependent on the factor σ_θ . To verify the validity of the mean field, we must quantify the distance between the trajectory of the mean of microscopic states and that of the macroscopic mean field approximation. In the bottom of Fig. 19, it is shown that as the number of neurons increases, the accuracy of the mean field approximation increases.

The mean field approximation is generalized to two populations of neurons having the same variances of weights and thresholds, but different mean weights due to LTP/LTD, which leads to Eq. (29).

References

- Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. *Nat Neurosci* 3:1178–1183
- Abbott LF, Regehr WG (2004) Synaptic computation. *Nat Biotechnol* 431:796–803
- Aihara K, Takabe T, Toyoda M (1990) Chaotic neural networks. *Phys Lett A* 144(6–7):333–340
- Amari S-I (1971) Characteristics of randomly connected threshold-element networks and network systems. *Proc IEEE* 59(1):35–47
- Amari S-I (1972) Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans Comput* C-21(11):1197–1206
- Artola A, Singer W (1993) Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci* 16:480–487
- Bartos M, Vida I, Jonas P (2007) Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat Rev Neurosci* 8(1):45–56
- Bienenstock EL, Cooper LN, Munro PW (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci* 2:32–48
- Brody CD, Romo R, Kepecs A (2003) Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr Opin Neurobiol* 13(2):204–211
- Brunel N (1996) Hebbian learning of context in recurrent neural networks. *Neural Comput Appl* 8:1677–1710
- Brunel N (2003) Dynamics and plasticity of stimulus-selective persistent activity in cortical network models. *Cereb Cortex* 13(11):1151–1161
- Buonomano DV, Maass W (2009) State-dependent computations: spatiotemporal processing in cortical networks. *Nat Rev Neurosci* 10(2):113–125
- 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 Appl* 16(5):885–940
- Buxhoeveden DP, Casanova MF (2002) The minicolumn hypothesis in neuroscience. *Brain Behav Evol* 125:935–951
- Buzsaki G (2002) Theta oscillations in the hippocampus. *Neuron* 33:325–340
- Buzsaki G (2010) Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron* 68(3):362–385
- Caianiello ER (1961) Outline of a theory of thought-processes and thinking machines. *J Theor Biol* 1:204–235
- Cessac B, Doyon B, Quoy M, Samuelides M (1994) Mean-field equations, bifurcation map and route to chaos in discrete time neural networks. *Phys D* 74:24–44
- Churchland M et al (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13: 369–378
- Compte A (2006) Computational and in vitro studies of persistent activity: edging towards cellular and synaptic mechanisms of working memory. *Neurosci Behav Physiol* 139:135–151
- Compte A, Brunel N, Goldman-Rakic PS, Wang X-J (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10(9): 910–923
- Compte A, Constantinidis C, Tegner J, Raghavachari S, Chafee MV, Goldman-Rakic PS, Wang X-J (2003) Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J Neurophysiol* 90:3441–3454
- Dauce E, Quoy M, Cessac B, Doyon B, Samuelides M (1998) Self-organization and dynamics reduction in recurrent networks: stimulus presentation and learning. *Neural Networks* 11(3): 521–533
- Doyon B, Cessac B, Quoy M, Samuelides M (1994) On bifurcations and chaos in random neural networks. *Acta Biotheor* 42(2): 215–225
- Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. *Nat Neurosci* 3:1184–1191
- Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9(4):292–303
- Fenichel N (1972) Persistence and smoothness of invariant manifolds for flows. *Indiana Univ Math J* 21(3):193–226
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey’s dorsolateral prefrontal cortex. *J Neurophysiol* 61:331–349
- Grebogi C, Ott E, Yorke JA (1983) Crises, sudden changes in chaotic attractors, and transient chaos. *Physica D* 7(1–3):181–200
- Grebogi C, Ott E, Yorke JA (1986) Metamorphoses of basin boundaries in nonlinear dynamical systems. *Phys Rev Lett* 56:1011–1014
- Hansel D, Sompolinsky H (1998) Modeling feature selectivity in local cortical circuits. In: *Methods in neuronal modeling: from synapse to networks*. MIT Press, Cambridge
- Hebb D (1949) *The organization of behavior*. Wiley, New York
- Heemels W, Lehmann D, Lunze J, De Schutter B (2009) Introduction to hybrid systems. In: Lunze J, Lamnabhi-Lagarigue F (eds) *Handbook of hybrid systems control—theory tools applications*. Cambridge University Press, Cambridge
- Hertz J, Krogh A, Palmer RG (1991) *Introduction to the theory of neural computation*. Addison-Wesley Longman Publishing Co, Inc, Boston
- Hirsch MW, Pugh CC, Shub M (1977) *Invariant manifolds*. Springer, New York

- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Nat Acad Sci* 79(8):2554–2558
- Hoppensteadt FC, Izhikevich EM (1997) Weakly connected neural networks. Springer, New York
- Ikegaya Y (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Sci Agric* 304:559–564
- Ibarz B, Casado J, Sanjuan M (2011) Map-based models in neuronal dynamics. *Phys Rep* 501:1–74
- Katori Y, Sakamoto K, Saito N, Tanji J, Mushiaki H, Aihara K (2011) Representational switching by dynamical reorganization of attractor structure in a network model of the prefrontal cortex. *PLoS Comput Biol* 7(11):e1002266
- Korn H (2003) Is there chaos in the brain? ii. Experimental evidence and related models. *CR Biol* 326(9):787–840
- Kuznetsov YA (1998) Elements of applied bifurcation theory. 2nd edn. Springer, New York
- Leleu T, Aihara K (2011) Sequential memory retention by stabilization of cell assemblies. In: Post-conference proceedings of the 3rd international conference on cognitive neurodynamics, 9–13 June 2011 (in press)
- Li Y, Nara S (2008) Novel tracking function of moving target using chaotic dynamics in a recurrent neural network model. *Cogn Neurodyn* 2(1):39–48
- Major G, Tank D (2004) Persistent neural activity: prevalence and mechanisms. *Curr Opin Neurobiol* 14(6):675–684
- Matsumoto G, Aihara K, Hanyu Y, Takahashi N, Yoshizawa S, Nagumo J-I (1987) Chaos and phase locking in normal squid axons. *Phys Lett A* 123(4):162–166
- McCulloch W, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 7:115–133
- Miyashita Y (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nat Biotechnol* 335(6193): 817–820
- Mongillo G, Amit DJ, Brunel N (2003) Retrospective and prospective persistent activity induced by hebbian learning in a recurrent cortical network. *Eur J Neurosci* 18(7):2011–2024
- Moynot O, Samuelides M (2002) Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probab Theory Relat Fields* 123:41–75
- Mushiaki H, Saito N, Sakamoto K, Itoyama Y, Tanji J (2006) Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* 50(4):631–641
- Nagumo J, Sato S (1972) On a response characteristic of a mathematical neuron model. *Biol Cybern* 10:155–164
- Naya Y, Sakai K, Miyashita Y (1996) Activity of primate infero-temporal neurons related to a sought target in pair-association task. *Proc Nat Acad Sci* 93:2664–2669
- Ott E (1993) Chaos in dynamical systems. Cambridge University Press, New York
- Ozaki TJ et al (2012) Traveling EEG slow oscillation along the dorsal attention network initiates spontaneous perceptual switching. *Cogn Neurodyn* 6(2):185–198
- Pesin YB (2004) Lectures on partial hyperbolicity and stable ergodicity. *Zurich Lect Adv Math*
- Pool RR, Mato G (2010) Hebbian plasticity and homeostasis in a model of hypercolumn of the visual cortex. *Neural Comput Appl* 22:1837–1859
- Pulvermiller F (1996) Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology* 33(4): 317–333
- Rajan K, Abbott LF, Sompolinsky H (2010) Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys Rev E* 82:011903
- Ranck JB Jr (1985) Head direction cells in the deep cell layer of dorsal presubiculum in freely moving rats. In: Electrical activity of the archicortex. Publishing House of the Hungarian Academy of Sciences
- Renart A, Moreno-Bote R, Wang X-J, Parga N (2007) Mean-driven and fluctuation-driven persistent activity in recurrent networks. *Neural Comput Appl* 19:1–46
- Renart A, Song P, Wang X-J (2003) Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron* 38(3):473–485
- Romo R, Brody CD, Hernandez A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nat Biotechnol* 399:470–473
- Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. *Nat Biotechnol* 354:152–155
- Sejnowski T (1999) The book of hebb. *Neuron* 24(4):773–776
- Siri B, Berry H, Cessac B, Delord B, Quoy M (2008) A mathematical analysis of the effects of hebbian learning rules on the dynamics and structure of discrete-time random recurrent neural networks. *Neural Comput Appl* 20(12):2937–2966
- Skarda C, Freeman W (1987) How brains make chaos in order to make sense of the world. *Behav Brain Sci* 10:161–195
- Softky WR, Koch C (1992) Cortical cells should fire regularly, but do not. *Neural Comput Appl* 4:643–646
- Sompolinsky H, Crisanti A, Sommers HJ (1988) Chaos in random neural networks. *Phys Rev Lett* 61(3):259–262
- Tsuda I (2001) Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behav Brain Sci* 24: 793–810
- Turrigiano GG (2008) The self-tuning neuron: synaptic scaling of excitatory synapses. *Cell (Cambridge, MA, US)* 135(3):422–435
- Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nat Biotechnol* 391(6670):892–896
- Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci* 5(2):97–107
- Wang L (2007) Interactions between neural networks: a mechanism for tuning chaos and oscillations. *Cogn Neurodyn* 1(2):185–188
- Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* 24(8):455–463
- Wennekers T, Palm G (2009) Syntactic sequencing in Hebbian cell assemblies. *Cogn Neurodyn* 3(4):429–441
- Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12(1):1–24
- Yoshida H, Kurata S, Li Y, Nara S (2010) Chaotic neural network applied to two-dimensional motion control. *Cogn Neurodyn* 4(1): 69–80
- Zheng G, Tonnelier A (2008) Chaotic solutions in the quadratic integrate-and-fire neuron with adaptation. *Cogn Neurodyn* 3(3): 197–204