Systems/Circuits

Effects of Cellular Homeostatic Intrinsic Plasticity on Dynamical and Computational Properties of Biological Recurrent Neural Networks

Jérémie Naudé,[1] Bruno Cessac,[2] Hugues Berry,[3] and Bruno Delord[4]

[1]UMR CNRS 7102, Université Pierre et Marie Curie, 75005 Paris, France, [2]NeuroMathComp Project Team, INRIA, 06902 Sophia Antipolis Cedex 2, France, [3]BEAGLE, INRIA Rhône-Alpes, Université de Lyon, LIRIS, UMR5205, 69100 Villeurbanne, France, and [4]Institut des Systèmes Intelligents et de Robotique, UMR CNRS 7222, Université Pierre et Marie Curie, 75005 Paris, France

Homeostatic intrinsic plasticity (HIP) is a ubiquitous cellular mechanism regulating neuronal activity, cardinal for the proper functioning of nervous systems. In invertebrates, HIP is critical for orchestrating stereotyped activity patterns. The functional impact of HIP remains more obscure in vertebrate networks, where higher order cognitive processes rely on complex neural dynamics. The hypothesis has emerged that HIP might control the complexity of activity dynamics in recurrent networks, with important computational consequences. However, conflicting results about the causal relationships between cellular HIP, network dynamics, and computational performance have arisen from machine-learning studies. Here, we assess how cellular HIP effects translate into collective dynamics and computational properties in biological recurrent networks. We develop a realistic multiscale model including a generic HIP rule regulating the neuronal threshold with actual molecular signaling pathways kinetics, Dale's principle, sparse connectivity, synaptic balance, and Hebbian synaptic plasticity (SP). Dynamic mean-field analysis and simulations unravel that HIP sets a working point at which inputs are transduced by large derivative ranges of the transfer function. This cellular mechanism ensures increased network dynamics complexity, robust balance with SP at the edge of chaos, and improved input separability. Although critically dependent upon balanced excitatory and inhibitory drives, these effects display striking robustness to changes in network architecture, learning rates, and input features. Thus, the mechanism we unveil might represent a ubiquitous cellular basis for complex dynamics in neural networks. Understanding this robustness is an important challenge to unraveling principles underlying self-organization around criticality in biological recurrent neural networks.

Introduction

Homeostatic intrinsic plasticity (HIP) stabilizes nervous systems functions against environmental or pathological perturbations, through a negative feedback of neuronal firing on intrinsic excitability (Schulz, 2006). For instance, HIP maintains stereotyped network activity patterns in invertebrates (LeMasson et al., 1993; Marder and Goaillard, 2006). In vertebrates, HIP was proposed to regulate the average level of neuronal activity (Maffei and Turrigiano, 2008; Remme and Wadman, 2012). However, perceptual, motor, or cognitive processes rely on neural dynamics; its complexity largely exceeds simple features such as target levels or stereotyped patterns of activity (Vogels et al., 2005). HIP was thus suggested to regulate the complexity of activity dynamics in vertebrate recurrent networks (e.g., neocortical columns, the hippocampus; Lazar et al., 2007, 2009; Siri et al., 2007, 2008; Wardermann and Steil, 2007; Marković and Gros, 2012) with possibly important functional consequences. However, the situation remains largely obscure for several reasons.

First, the effect of HIP on network dynamics is unclear as it was found to either increase or decrease the complexity of the dynamics, depending on conditions (Lazar et al., 2007; Wardermann and Steil, 2007; Marković and Gros, 2012). Second, HIP was suggested as a mechanism preventing the pathological saturation of firing frequency and synchronization that is typical of Hebbian synaptic plasticity (SP; Turrigiano, 1999; Abraham, 2008; Watt and Desai, 2010). Yet, the proposition (Siri et al., 2007, 2008) that HIP may oppose these pathologies, by counteracting the underlying decrease in the complexity of the dynamics, has remained unexplored. Third, recurrent networks display improved recognition performances at the critical boundary between irregular and regular dynamics (the "edge of chaos"; Bertschinger and Natschläger, 2004; Siri et al., 2007, 2008). However, in networks with HIP, critical dynamics does not imply improved performance (Lazar et al., 2007), which generally arises at subcritical (regular) regimes (Steil, 2007; Wardermann and Steil, 2007; Lazar et al., 2009), urging for a better understanding of these relations (Lazar et al., 2007, 2009; Steil, 2007; Schrauwen et al., 2008).

Here, we focus on the unsolved issue of the transmission of HIP effects along scales, i.e., how can one explain HIP effects at the global network level, based on its original action, which actually operates at the level of individual neurons? To evaluate the biological relevance of this transmission, we model cardinal architectural and plasticity features characterizing vertebrate recurrent networks. In particular, we derive a generic biophysical HIP rule modeling the ubiquitous activity-dependent kinase/phosphatase regulation of sub-threshold conductances (Delord et al., 2007), which sets the current threshold for firing (Naudé et al., 2012), in contrast to previous models based on rules that are phenomenological or derived to optimize information transfer. Using a dynamical mean-field approach and numerical simulations, we unravel a new cellular mechanism of HIP action, the transduction of inputs by large derivative ranges of the transfer function, which ensures increased dynamics complexity, robust interaction with SP at the edge of chaos, and improved pattern separability at the network scale.

Materials and Methods

Neuronal dynamics. We study the effects of HIP on the activity dynamics of random recurrent neural networks. Our neural network model includes N firing-rate neurons with discrete dynamics. To account for the slower timescale (approximately minutes to days) of synaptic and intrinsic plasticity, compared with neuronal activity (~10 ms; Siri et al., 2007, 2008), we update the synaptic weights and firing thresholds at a timescale slower than that of neuronal activities. We call a "learning epoch" the period of τ successive updates of neuron activities separating two updates of the synaptic weights and thresholds. In real neurons, the induction of plasticity (through signaling pathways) can be quite fast (approximately seconds to minutes), but its expression (as effective receptor or voltage-gated channels changes) operates at lower timescales (approximately minutes to days). Here, a learning epoch (the time step for learning) corresponds to 5.10^3 or 10^4 activity time steps of 10 ms, i.e., ~ 1 min, encompassing the timescale of induction as well as a minimal estimate of the timescale of expression of plastic processes. In simulations, learning typically operates within about $T = 1000$ epochs, i.e., ~15 h. We could have considered higher numbers of time steps for the learning epoch, to encompass larger timescales for the expression of plastic processes. However, this would have dramatically slowed down simulations, which were already quite demanding. Actually, the present results are independent of the exact timescale of the learning epoch, as long as it is sufficiently slower, compared with the timescale for neuronal activation, so that neuronal dynamics can converge toward its attractor within each learning epoch (which was the case here). In the following, t denotes the update step of neuronal activation, and T the update step of synaptic weights and neuronal thresholds (i.e., the learning epoch). Let $x_i^{(T)}(t)$ be the firing frequency of neuron i at time, during learning epoch T, ranging between 0 (quiescence) and 1 (saturation). The dynamics of the model is given by the following:

$$x_i^{(T)}(t + 1) = f(u_i^{(T)}(t) - \theta_i^{(T)}) \qquad (1)$$

where $f(u) = (1 + \tanh(Gu))/2$ is a sigmoidal transfer function, G, the gain of the transfer function, and $\theta_i^{(T)}$ the neuronal threshold of neuron i during the epoch T. The quantity $u_i^{(T)}(t)$ is called the local field:

$$u_i^{(T)}(t) = \sum_{j=1}^{N} w_{ij}^{(T)} x_j^{(T)}(t) + \xi_i, \qquad (2)$$

i.e., the weighted sum of synaptic inputs due to recurrent connectivity in the network (where $w_{ij}^{(T)}$ is the synaptic weight from the presynaptic neuron j to the postsynaptic neuron i), plus ξ_i, a constant external input (the vector $\xi = \{\xi_i\}$ is the input pattern applied to the network). The timescale separation between neuron activity and plasticity updates allows the analysis of network dynamics based on numerical estimation of the largest Lyapunov exponent $L^{(T)}$. The activity $x_i^{(T)}(t)$ of each neuron i

during each learning epoch T is thus computed for each time t using constant values for the synaptic weights $w_{ij}^{(T)}$ and firing threshold $\theta_i^{(T)}$. Neuronal activities, averaged over τ time steps, are used to update the synaptic weights and the neuronal thresholds for the next learning epoch, according to the plasticity rules (see below). The activity at the end of a learning epoch is then used to set the initial conditions of the next learning epoch $x_i^{(T+1)}(1) = x_i^{(T)}(\tau)$.

Network architecture. The network connectivity is set at random using a biologically realistic setting with sparse connectivity and segregation of excitatory and inhibitory neurons (following Dale's principle; Siri et al., 2007). Each neuron is inhibitory (with probability p_I) or excitatory ($p_E = 1 - p_i$) and projects to $N_c = p_c N$ randomly chosen postsynaptic neurons with uniform probability, where p_c is the connection probability, independent of the polarity of the synapse. We chose p_I and p_c to reflect typical cortical architectural constraints. The initial synaptic weights $w_{ij}^{(1)}$ of connections are drawn from a gamma distribution, which ensures that excitatory (respectively, inhibitory) neurons only project to synapses with positive (respectively, negative) weights. If neuron j is inhibitory, $w_{ij}^{(1)}$ follows $\Gamma(-\mu_w/n_I; \sigma_w/n_I)$ where $\Gamma(m,s)$ is a gamma distribution with mean m and SD s, with $n_I = p_I p_c N$ being the total number of synapses from an inhibitory neuron. Similarly, if neuron j is excitatory, then $w_{ij}^{(1)}$ follows $\Gamma(-\mu_w/n_E; \sigma_w/n_E)$ with $n_E = p_E p_c N$. Normalization terms (n_E and n_I) guarantee that during the first epoch, the expected total excitation received by a postsynaptic neuron equals the expected total inhibition (balanced network). This initial global balance between excitation and inhibition can eventually be broken by SP. Note that with this setting, the initial weight matrix $W(1)$ is asymmetric, i.e., $w_{ij}(1)$ differs from $w_{ji}(1)$ (with probability 1). Autapses were prohibited throughout learning epochs ($w_{ii} = 0$ for all T).

SP. We used a Hebbian SP rule to update synaptic weights at the end of each epoch (Siri et al., 2007, 2008), depending on the average firing frequencies of presynaptic and postsynaptic neurons, as follows:

$$w_{ij}^{(T+1)} = \lambda w_{ij}^{(T)} + S_{ij} \frac{\alpha}{N} m_i^{(T)} m_j^{(T)} H(m_j^{(T)}), \qquad (3)$$

where α is the Hebbian plasticity rate, and $\lambda \epsilon [0;1]$ is a passive forgetting coefficient for weights' evolution. In this equation $m_i^{(T)}$ denotes the reduced average activity of neuron i during epoch T as follows:

$$m_i^{(T)} = \frac{1}{\tau} \sum_{t=1}^{\tau} x_i^{(T)}(t) - d \qquad (4)$$

where $d = 0.1$ is the threshold from which the neuron i is considered as active during epoch T, i.e., when its average activity is >10% of its maximal value. $H()$ is the Heaviside function ($H(m) = 0$ if $m < 0$, 1 otherwise), and prevents heterosynaptic plasticity. The term $m_i^{(T)} m_j^{(T)} H(m_j^{(T)})$ represents the correlative principle of Hebb's rule: weights increase when both the presynaptic and postsynaptic neurons are active above d. The Heaviside function requires that the presynaptic neuron is active to induce synaptic weights' changes, preventing heterosynaptic depression and potentiation when both presynaptic and postsynaptic neurons are inactive. Finally, s_{ij} represents the sign of the synapse: $s_{ij} = +1$ for presynaptic excitatory neurons, $s_{ij} = -1$ for inhibitory ones. The excitatory or inhibitory nature of the presynaptic neuron also affects the Hebbian plasticity rate α. Indeed, to ensure the global balance of excitation and inhibition in our setting, where excitatory neurons are three times more numerous than inhibitory ones, we used $\alpha_I p_I = \alpha_E p_E$, where α_E (respectively, α_I) stands for excitatory (inhibitory) SP rates. This global excitation/inhibition balance has dramatic consequence on the effects of intrinsic plasticity on network dynamics (see Results).

HIP. In the framework of firing rate coding, intrinsic plasticity affects neuronal excitability by modifying the transfer function, i.e., the relationship between the input and the resulting firing frequency; here, we consider the homeostatic plasticity of the maximal conductance of a subthreshold voltage-gated ionic current (Desai et al., 1999; Gibson et al., 2006; van Welie et al., 2006), i.e., an ionic current activating below the

threshold for action potential (Naudé et al., 2012). This includes, e.g., the leak, persistent and slowly inactivating sodium, low-threshold calcium, or muscarinic, and slowly inactivating potassium currents, which affect the current threshold of the neuronal transfer function (Bekkers and Delaney, 2001; Brickley et al., 2001; Vervaeke et al., 2006; O'Leary et al., 2010). We consider an intrinsic learning rule preserving essential features of intrinsic plasticity: (1) HIP modifies the threshold $\theta_i^{(T)}$ between learning epochs, shifting the transfer function and modifying the input range to which the neuron responds, and (2) $\theta_i^{(T)}$ evolves as a function of the neuron mean firing frequency $<x_i^{(T)}>$, through the activation of intracellular biochemical signaling pathways.

In the model, the threshold of the function transfer is modified by HIP according to the following:

$$\theta_i^{(T)} = \bar{\theta}(A_\theta + B_\theta F_i^{(T)}), \qquad (5)$$

where $\bar{\theta}$ is the maximal threshold value and $F_i^{(T)}$ the phosphorylated fraction of the plastic subthreshold conductance, assumed to constitute the functional pool of channels affecting excitability. Initially, all neuron thresholds are set to 0, i.e., $F_i^{(0)} = -A_\theta/B_\theta$. Parameters A_θ and B_θ are set so as to ensure a homeostatic regulation of the threshold by neuronal activity: an increase (respectively, a decrease) of the firing frequency leads to a decrease (respectively, an increase) of neuronal excitability. The phosphorylated fraction of the subthreshold plastic conductance $F_i^{(T)}$, which determines the threshold value, is regulated by activity-dependent kinases and phosphatases (Ganguly et al., 2000; Cudmore and Turrigiano, 2004; Zhang et al., 2004). We model these intracellular pathways using the activity-dependent kinase and phosphatase (aKP) framework (Delord et al., 2007), adapted to discrete time dynamics as follows:

$$F_i^{(T+1)} = F_i^{(T)} + K_i^{(T)}(1 - F_i^{(T)}) - P_i^{(T)}(1 - F_i^{(T)}) \qquad (6)$$

where $K_i^{(T)}$ and $P_i^{(T)}$ represent mean activation levels of protein kinases and phosphatases K and P at learning epoch T (Delord et al., 2007). This biochemical pathway is activated by intracellular calcium concentration, as observed experimentally (Ganguly et al., 2000; Cudmore and Turrigiano, 2004; Zhang et al., 2004) as follows:

$$K_i^T = k_{max} \frac{Ca_i^{(T)p}}{Ca_i^{(T)p} + K_K^p} \qquad (7a)$$

and

$$P_i^T = p_{max} \frac{Ca_i^{(T)p}}{Ca_i^{(T)p} + K_P^p} \qquad (7b)$$

where p is the Hill constant, k_{max} and p_{max} are kinetic coefficients, K_p and K_K are the phosphatase and kinase calcium half-activations, respectively. $Ca_i^{(T)}$ represents the mean concentration of the intracellular calcium during learning epoch T. At the equilibrium state, intracellular calcium concentration linearly depends on the mean firing rate (Helmchen et al., 1996; Wang, 1998), which we simulate with the following:

$$Ca_i^{(T)} = \frac{1}{\tau} \sum_{t=1}^{\tau} x_i^{(T)}(t) + Ca_0 \qquad (8)$$

with basal concentration Ca_0 (Wang, 1998; Delord et al., 2007). A non-null basal calcium concentration ensures non-null rate of plasticity for very low neuronal activities (Delord et al., 2007). This HIP model is sufficiently realistic to take into account essential dynamical properties of intrinsic plasticity, i.e., the gradation of threshold modifications and the activity-dependent time constant governing its dynamics (Delord et al., 2007).

Input patterns. To assess the ability of the network to learn and discriminate inputs, we used n_ξ arbitrary input patterns $\xi_k(k = \{1, \ldots, n_\xi\})$, defined as $\xi_{i,k} = 0.2 f_{k,1}\left(2\pi\frac{i}{N}\right) f_{k,2}\left(8\pi\frac{i}{N}\right)$, where i is the neuron index, N the total number of neurons, and $f_{k,1}$ and $f_{k,2}$ two trigonometric functions (see below). We used either $n_\xi = 2, 3,$ or 4 input patterns to assess how separability behaves with increasing num-

bers of inputs. In a given condition, the network was trained by repeating the sequence formed by the n_ξ inputs at each successive learning epoch (e.g., $\xi_1, \xi_2, \xi_1, \xi_2, \ldots$ for $n_\xi = 2$). At the end of every learning epoch, synaptic weights and thresholds were modified according to SP and HIP rules, respectively, using the neuronal activities evoked by the specific input presented. Besides, at each learning epoch, the network was presented off-line with the $n_\xi - 1$ other input patterns. These responses were not used to update weights and thresholds, but only to compute the input pattern separability of the network. Specifically, the distance between the response of the network to two given inputs ξ_k and ξ_l was computed, at each learning epoch T, as $\Delta x_{\xi k, \xi l}^T = \left[\frac{1}{N}\sum_{i=1}^{N}(\langle x\rangle_{i,\xi k}^T - \langle x\rangle_{i,\xi l}^T)^2\right]^{1/2}$, where $\langle x\rangle_{i,\xi}^T = \frac{1}{\tau}\sum_{t=1}^{\tau} x_{i,\xi}^T(t)$ was the average activity of neuron i at epoch T during presentation of the input pattern ξ. This distance was normalized by the distance between the corresponding inputs $\Delta\xi_{\xi k, \xi l} = \left[\frac{1}{N}\sum_{i=1}^{N}(\xi_{i,k} - \xi_{i,l})^2\right]^{1/2}$, and averaged over input pattern pairs, to compute the input pattern separability of the network $S = \binom{n_\xi}{2}^{-1}\sum_{k=1}^{n_\xi}\left(\sum_{l=k+1}^{n_\xi}\left(\frac{\Delta x_{\xi k, \xi l}}{\Delta\xi_{\xi k, \xi l}}\right)\right)$, in a given learning condition. In Figure 6D–F, S was averaged over simulations run for 10 random realizations of the synaptic weight matrix and computed at the end of the simulation, once learning was achieved. The functions $\{f_{k,1}, f_{k,2}\}$ were chosen as {sin, cos}, {cos, sin}, {cos, cos}, and {sin, sin} for the four inputs $\xi_1, \xi_2, \xi_3, \xi_4$, respectively.

Numerical estimation of the Lyapunov exponent from numerical simulations. Lyapunov exponents measure the local tendency of nearby trajectories living on an attractor of the dynamics: (1) to get closer, on average, in some directions of the phase space, corresponding to a negative Lyapunov exponent and (2) or to move apart in other directions, corresponding to a positive Lyapunov exponent. There are as many Lyapunov exponents as the phase space dimensions (N in our model). The sum of all Lyapunov exponents gives the local average volume contraction in the tangent space of an attractor. This sum is always negative from the definition of an attractor. When all Lyapunov exponents are negative the attractor is a fixed point, or, in the discrete time case, a finite union of points constituting a stable periodic orbit. If at least one exponent is positive, then the attractor is chaotic. The largest Lyapunov exponent gives an efficient and rapid way of characterizing the system dynamics: fixed point or periodic orbit when negative, chaos when positive. When it is zero the attractor has some neutral direction neither expanding nor contracting. This is the case, e.g., when the attractor is a closed invariant curve (typically a distorted circle where dynamics acts as a rotation with an irrational winding number), or an invariant torus.

For a given neural network, both synaptic weights and neuronal thresholds evolve through learning epochs, according to Hebbian SP and HIP rules, respectively, from their initial values (random synaptic weight matrix, null thresholds; see above). The largest Lyapunov exponent, $L_1^{(T)}$, was estimated numerically from simulations of the neural network models with a QR-based method (Von Bremen et al., 1997) at each learning epoch T, i.e., using the synaptic weights $w_{ij}^{(T)}$ and the neuronal thresholds $\theta_i^{(T)}$ values. Employing these values as fixed parameters for the learning epoch T, a random initial condition of the network activity was drawn and the network dynamics was first iterated for 2000 activity update steps to allow convergence to the (fixed-point, limit-cycle, quasiperiodic, or chaotic) attractor. The largest Lyapunov exponent was then computed over successive periods of 1000 activity update steps, until reaching an absolute convergence criterion of 10^{-3} between successive Lyapunov exponents. This procedure was repeated at every learning epoch T, to compute the learning-induced evolution of the largest Lyapunov exponent, $L_1^{(T)}$. To ensure robustness of this computation, $L_1^{(T)}$ was averaged over 10 realizations of the random synaptic weight matrix (see Results). When relevant, the steady-state value of the largest Lyapunov exponent, $L_1^{(\infty)}$, was estimated by considering the value of $L_1^{(T)}$ at large learning epochs ($T = 2000$). Note that both $L_1^{(T)}$ and $L_1^{(\infty)}$ differ from the theoretical largest Lyapunov exponent, L_1^{MF}, for two reasons. First, whereas $L_1^{(T)}$ and $L_1^{(\infty)}$ are obtained from the standard biological model, L_1^{MF} is derived

from a simplified network model characterized by synaptic homogeneity, i.e., devoid of Dale's principle and sparse connectivity. Second, $L_1^{(T)}$ and $L_1^{(\infty)}$ are estimated numerically from simulations of finite-sized networks, while L_1^{MF} is computed through an analytical expression obtained from the dynamical mean-field theory valid at the thermodynamical limit ($N \to +\infty$). Thus, a direct quantitative comparison of $L_1^{(T)}$ or $L_1^{(\infty)}$ with L_1^{MF} is not relevant. However, qualitative forms of agreement (e.g., similar parameters dependence) indicate the ability of the theory to account for the mechanisms underlying HIP effects.

Standard network model. In the present study, we define the standard model as a recurrent network endowed with balanced excitatory/inhibitory drive, Dale's principle, sparse connectivity, and no plasticity. In the different sections, either HIP, SP, or both forms of plasticity are added to study their effects on emerging properties. Unless stated, the parameters used are $N = 200$, $G = 5$, $\tau = 10^4$, $p_i = 0.25$, $p_c = 0.15$, $\mu_W = 50$, $\sigma_W = 1$, $A_\theta = 1.1$, $B_\theta = -3.635$, $Ca_0 = 0.1$, $p = 4$, $p_{max} = k_{max} = 10^{-3}$, $K_p = 1/3$, $K_K = 2/3$, $\bar{\theta} = 20$.

Results

Mean-field analysis of HIP effects

In this first section, we study the effects of the sole HIP, i.e., in the absence of Hebbian SP, to assess the mechanism by which it affects neuronal dynamics at the global scale of the network. To this aim, we use a dynamic mean-field theory formerly developed for this type of network (Cessac, 1995; Doyon et al., 1995). This method allows the characterization of the dynamics averaged over the synaptic weight distribution. Here, in absence of SP, we suppose that synaptic weights are independent and identically distributed, as usually considered in dynamic mean-field theories with random interactions (Sompolinsky et al., 1988; Arous and Guionnet, 1995; Faugeras et al., 2009). In addition, HIP is a local plasticity rule that does not affect the correlation between synaptic weights. Hence, it does not correlate initially uncorrelated weights and it is possible to extend the results obtained in (Moynot and Samuelides, 2002). We denote $E()$ the expectation over synaptic weight matrix realizations. Our mean-field analysis assumes synaptic homogeneity, i.e., we considered a fully connected network composed of a single population of neurons, in the thermodynamical limit ($N \to +\infty$). In this case, the local field (i.e., the input − the threshold) of neurons, u_i, are Gaussian (Cessac, 1995; Moynot and Samuelides, 2002), with the following expectation:

$$\mu_i^{(T)} \equiv E(u_i^{(T)}) = E(\theta_i^{(T)}) + \xi_i, \quad (9)$$

where the neuronal threshold expectation is obtained from Equation 5 as follows:

$$E(\theta_i^{(T)}) = \bar{\theta}(A_\theta + B_\theta E(F_i^{(T)})) \quad (10)$$

The expectation of the threshold is itself altered by modifications of the plastic variable F, which evolves according to (Eq. 6, see Materials and Methods) the following:

$$F_i^{(T+1)} = F_i^{(T)} + K(M_i^{(T)})(1 - F_i^{(T)}) - P(M_i^{(T)}) F_i^{(T)} \quad (11)$$

where the expectation of neuronal activity is given by (Moynot and Samuelides, 2002) the following:

$$M_j^{(T)} \equiv E(x_j^{(T)}) = \int_{-\infty}^{\infty} f(\sqrt{V^{(T)}} h + \mu_j^{(T)}) Dh \quad (12)$$

with

$$Dh = (2\pi)^{-1/2} \exp\left(-\frac{h^2}{2}\right) dh \quad (13)$$

and where $V^{(T)}$ is the variance of the local field. Let $V_i^{(T)}$ be the variance of $u_i^{(T)}$ at the epoch T. The mean-field theory establishes that $V_i^{(T)}$ is given, in the thermodynamic limit, by the following:

$$V_i^{(T)} = \lim_{N \to +\infty} \frac{J^2}{N} \sum_{j=1}^{N} \int_{-\infty}^{+\infty} f(\sqrt{V^{(T)}} h + \mu_j^{(T)})^2 Dh \equiv V^{(T)} \quad (14)$$

A specific case holds when $\mu_j^{(T)}$ does not depend on j (the input is independent of j). Then, $V^{(T)} = J^2 \int_{-\infty}^{\infty} f(\sqrt{V^{(T)}} h + \mu_j^{(T)})^2 Dh$. Equations 9–14 are self-consistent and can be used to compute the largest Lyapunov exponent, which characterizes the complexity of the network dynamics (Cessac, 1995) as follows:

$$L_1^{MF(T)} = \frac{1}{2} \log(J^2 \sum_{j=1}^{N} \int_{-\infty}^{\infty} f'(\sqrt{V^{(T)}} h + \mu_j^{(T)})^2 Dh) \quad (15)$$

Thus, the mean-field approach brings important insights about the mechanism by which HIP expresses at the network's scale. First, HIP affects the mean of the local field of neurons $\mu_j^{(T)}$ (Eq. 9) which in turn determines its variance $V^{(T)}$ (Eq. 14). Second, both $\mu_i^{(T)}$ and $V^{(T)}$ affect the complexity of the network dynamics (Eq. 15). Third, these variables act through the derivative of the transfer function of neurons, f', which determines neurons' sensitivity to their inputs (Eq. 15). According to Equation 15, decreasing the absolute value of the expectation $\mu_i^{(T)}$ should reduce the network's dynamics complexity. However, it is not straightforward to predict from the mere equation of how the combined effect of $\mu_i^{(T)}$ and $V^{(T)}$ of local fields should affect this complexity. To evaluate this effect, we computed the maximum Lyapunov exponent predicted by the dynamical mean-field theory. That is, we iterated Equations 9–14 until they reached a stable fixed point, corresponding to a steady-state distribution of the network activity (Cessac, 1995). Then, from Equation 15, we were able to compute the predicted maximal Lyapunov exponent, L_1^{MF}, as the steady-state value of $L_1^{MF(T)}$. Specifically, we computed L_1^{MF} for a large range of the external input ξ (i.e., $[-\bar{\theta}, \bar{\theta}]$, i.e., $[-20, 20]$) in the presence or absence of HIP, to evaluate the robustness of HIP on the dynamics of the network.

We observed that, in the absence of HIP, L_1^{MF} reaches negative values when increasing the amplitude of the input ξ. A negative maximal exponent corresponds to a regular dynamics (stable fixed-point or periodic orbit; Fig. 1A, red). Actually, without the homeostatic compensation provided by HIP, large inputs (Fig. 1B, vertical red band) are transformed through the transfer function (plain black curve) into large outputs ($x_i = f(u_i) \approx 1$; horizontal red stripe). However, the derivative of the transfer function (dashed black curve) at large inputs is small ($f'(u_i) \approx 0$). Therefore, fluctuations of the input (vertical red band) have little impact on postsynaptic neurons because they are contracted into small output fluctuations (horizontal red band). Hence, fluctuations of activity are dampened in the network (low variance of the activity; Fig. 1C, red) and regular dynamics are promoted. In the dynamic mean-field theory, Equation 15 confirms that this reduced derivative is indeed the factor decreasing L_1^{MF}, bringing the network's activity into more regular regime, where dynamics complexity is lower.
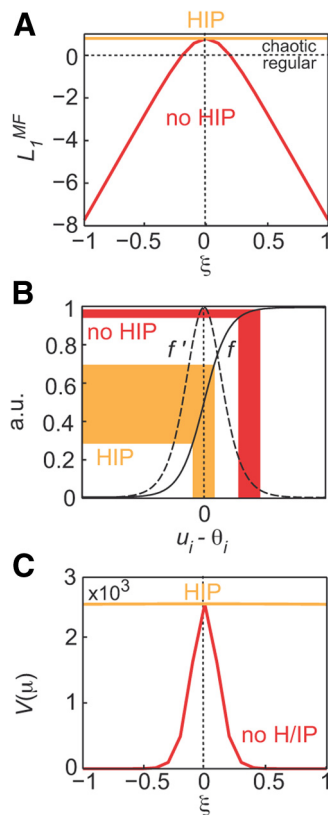
Figure 1. Mean-field analysis of HIP effects on activity dynamics. **A**, Largest Lyapunov exponent computed with the dynamical mean-field approach (see Materials and Methods), L_1^{MF}, as a function of the input amplitude ξ, in the absence (red) or presence (yellow) of HIP. The border $L_1 = 0$ (black dotted line) separates irregular (chaotic) and regular (limit-cycle, fixed-point) dynamics. **B**, The neuronal transfer function, f (black curve) and its derivative f' (black dashed curve), as a function of the local field, i.e., the difference between the input u_i and the neuron's threshold θ_i. In the absence of HIP, input fluctuations (red bands) give rise to contracted output fluctuations. In contrast, in the presence of HIP, output fluctuations are amplified (yellow bands; see text). **C**, Local field variance at the end of the simulation in the absence (red curve) or presence (yellow curve) of HIP. **B**, **C**, The results depicted were obtained by iterating Equations 9–15 from the theoretical analysis. For parameters, see Materials and Methods.

In contrast, in the presence of HIP, L_1^{MF} was positive, i.e., dynamics was chaotic, and virtually constant, for a large range of input amplitude (of the order of $[-\bar{\theta}, \bar{\theta}]$, the range of possible threshold modifications; Fig. 1A, yellow). In this case, because of the homeostatic compensation provided by HIP, even large inputs are compensated by a shift of the threshold, resulting in a null expectation of the local field distribution (Fig. 1B, vertical yellow band). As a consequence, the output is balanced ($x_i = f(u_i) \approx 0.5$; horizontal yellow band), and the derivative is maximal ($f'(u_i) \approx G/2$). Therefore, fluctuations of the input are amplified in the output and thus maintained in the network. The activity presents a large variance (Fig. 1C, yellow), which should favor complex dynamics. Again, the mean-field theory we have developed evidences that the maximal derivative observed in the presence of HIP is indeed the factor increasing L_1^{MF}, bringing the network's dynamics into a chaotic regime.

Thus, the dynamic mean-field theory suggests that within the range of possible threshold plastic modifications $[-\bar{\theta}, \bar{\theta}]$ HIP (1) exerts a compensation to external inputs that is robust to their exact amplitude and (2) guarantees complex dynamics operating through a mechanism relying on a large derivative of the transfer function, and translating as large fluctuations of neuronal activity. In the following sections, we compare the predictions of the
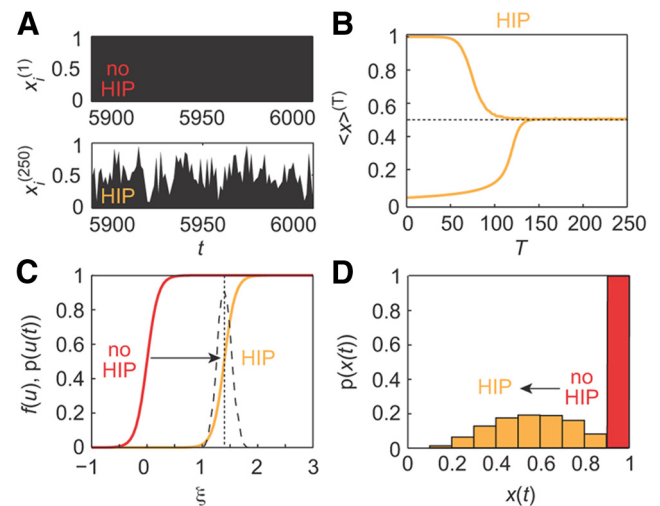

Figure 2. Cellular effects of HIP. Response of a neuron endowed with threshold HIP and receiving synaptic inputs from a chaotic recurrent network devoid of plasticity. **A**, Firing rate histograms as a function of the time t within learning epochs $T = 1$ [i.e., saturated activity, HIP has not yet operated (upper histogram)] and $T = 250$ (lower histogram). **B**, Averaged activity during learning with HIP, $\langle x \rangle^{(T)}$, as a function of learning epochs, from two initial conditions. **C**, Shift in the transfer function in the absence of HIP (red curve) and following its effect (yellow curve), in response to the distribution of network inputs (dotted black curve). Note that this distribution does not change during the simulation. **D**, Distributions of instantaneous activities in the absence (large initial activity, red) and presence (yellow) of HIP. For parameters, see Materials and Methods.

dynamic mean-field theory, which are exact in the thermodynamical limit $N \to +\infty$, to simulations in finite sized models.

Cellular effects of HIP

Here, we first checked that in a single neuron model endowed with HIP, homeostatic compensation of the local field indeed maintains the activity variability characterizing the mechanism we have unveiled. To test this proposal, we consider a single plastic neuron receiving inputs from a simulated standard recurrent model network (i.e., Dale's principle, sparse connectivity, balanced excitatory/inhibitory drive, no plasticity, see Materials and Methods), with chaotic dynamics. We observed that, starting from an initial state where the neuron was saturated due to a particular realization of its incoming synaptic weights (Fig. 2A, upper histogram, firing rate as a function of the time t within learning epoch $T = 1$), the presence of HIP was able to restore a large variability of activity (Fig. 2A, lower histogram, learning epoch $T = 250$). Actually, when the activity of the plastic neuron was initially large, HIP caused a progressive increase of the firing threshold, leading to a decrease of the time-averaged activity toward a median value ($\langle x(t) \rangle^{(T)} = 0.5$; Fig. 2B, upper trace). Symmetrically, when the initial activity was low, HIP decreased the firing threshold, balancing activity (Fig. 2B, lower trace). In each case, HIP adapted the threshold and canceled the neuron's local field so that the neuron exploited the linear part of its transfer function (Fig. 2C), raising firing variability (Fig. 2D).

Network effects of HIP

We next considered the effect of HIP on the standard recurrent network model, addressing the joint effect of HIP and SP in the next section. This standard model was designed as a biologically realistic recurrent network with sparse connectivity and segregation between excitatory and inhibitory neurons (Dale's princi-
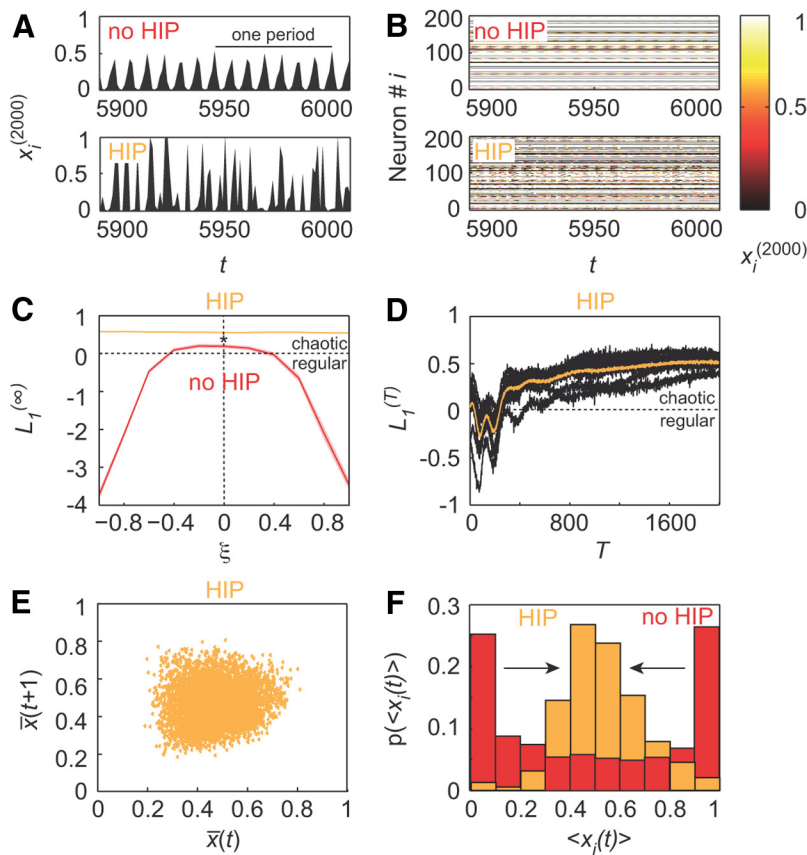
Figure 3. Network effects of HIP. Characterization of the recurrent neural network activity in the presence of HIP (no Hebbian SP). ***A***, Firing rate histograms as a function of the time t at the end of the simulations ($T = 2000$) in a static network (absence of plasticity; upper, the bar indicates one period of oscillation) and with HIP (lower). ***B***, Network activity rasters in the static network (upper) and with HIP (lower). ***C***, Steady-state mean value of the largest Lyapunov exponent, $L_1^{(\infty)}$ at the end of the simulations ($T = 2000$), as a function of the external input amplitude, in the absence (yellow) or presence of HIP (red). Values are averaged over 10 network realizations. SEM of $L_1^{(\infty)}$ in the presence or absence of HIP is represented as light yellow and pink areas, respectively (note that the SEM is very small in the absence of HIP while it is not even visible in its presence). Values are averaged over 10 realizations. ***D***, Evolution of $L_1^{(T)}$ during simulations: individual traces (black) and average over 10 network realizations (yellow), with HIP (no external input, i.e., $\xi = 0$). ***E***, Bidimensional reconstruction of a dynamical attractor from a sample network in ***D*** at the end of the simulation, using the Takens method (see Materials and Methods). ***F***, Distribution of time-averaged individual activities at the beginning (red) and at the end (yellow) of simulations, averaged over 10 realizations. The shape of the strange attractor obviously depends on initial realizations of the synaptic weights and thresholds, and on model parameters. ***C, D***, Horizontal black dotted lines: $L_1 = 0$.

ple). When no plasticity was present, we observed that, at the end of the simulations ($T = 2000$), the dynamics of the network was regular, e.g., a periodic orbit, as illustrated by the firing rate histogram of a individual neuron (Fig. 3A, upper, the bar indicates one period of the oscillation) and the raster of network activity (Fig. 3B, upper). In contrast, the activity in the presence of HIP led to chaotic dynamics (Fig. 3A, lower; B, lower).

Using this network and a constant external input applied to all neurons, we tested the theoretical predictions made by the dynamical mean-field approach developed above. We performed these simulations with $N = 200$ and computed numerically the maximal Lyapunov exponent (see Materials and Methods), averaging its value over 10 realizations of the synaptic weight matrix, for each external input amplitude tested. In the following, we denote $L_1^{(T)}$ the maximal Lyapunov exponent computed in the simulations at the learning epoch T, and $L_1^{(\infty)}$ its steady-state value at large simulation times ($T = 2000$). Since the mean-field theory is exact for a fully connected model in the limit $N \to +\infty$, one cannot expect a perfect match between the Lyapunov exponent

derived from the theory (L_1^{MF}) and those estimated in the simulations ($L_1^{(T)}$ or $L_1^{(\infty)}$). However, qualitative agreement indicates the heuristic capacity of the dynamical mean-field theory.

We observed that the steady-state largest Lyapunov exponent $L_1^{(\infty)}$ was higher with HIP (Fig. 3C, yellow) than without (red). On average, $L_1^{(T)}$ increased over time in the presence of HIP and stabilized at positive values, indicating a chaotic behavior (Fig. 3D, yellow). Moreover, HIP led the network to chaotic dynamics independently of the initial network dynamics (chaotic ($L_1^{(T)} > 0$) or not ($L_1^{(T)} < 0$); Figure 3D, black traces), which depended on the random realizations of synaptic weights (that possibly induced strong variations of individual neuronal local fields). The resulting chaotic behavior lives on a strange attractor (Eckmann and Ruelle, 1985). It is possible to visualize attractors by plotting the average network activity during consecutive simulation times (t and $t + 1$), using the Takens method (i.e., with time delay 1 here; Takens, 1981; Doyon et al., 1994). Such a representation in the plane $\{\bar{x}(t), \bar{x}(t + 1)\}$ reveals the topological structure of the attractor underlying network dynamics. For instance, a single point reveals that dynamics is stuck at a fixed-point attractor, separate points indicate a periodic orbit attractor within which the system migrates, and a closed 1D dense variety (curve) indicates a quasiperiodic attractor, while a cloud of points reveals a chaotic attractor, as found here (Fig. 3E). Note that HIP increased the dynamics complexity of biological networks regardless of the value of the external input (i.e., even with $\xi = 0$; Fig. 3C), whereas the dynamical mean-field approach predicted identical L_1^{MF} values at $\xi = 0$ (Fig. 1A).

Despite this discrepancy, attributable to the different hypotheses considered (sparse vs full connectivity, Dale's principle or not, and finite or infinite size), we found a good qualitative agreement between the theory and the simulations. Indeed, the similar dependence upon input amplitude and HIP presence demonstrated that the theory correctly accounts for the mechanisms underlying HIP effects.

For the range of parameters we used, we found that, independently of (1) the amplitude of the input current considered and (2) the random realization of the weight matrix, HIP was able to strongly decrease the saturation of neuronal activities (Fig. 3F, yellow), even when an important part of the network was initially saturated (red). The results obtained with network models (i.e., in the present and following sections) were robust to large variations of the network size (data not shown). Thus, HIP was able to maintain dynamical and chaotic fluctuations in the network and provided large variability of neuronal activity, confirming the mechanism unveiled by the dynamical mean-field approach.

HIP and SP interaction in the recurrent network

We addressed the important issue of whether HIP could prevent runaway effects of SP, such as pathological levels of firing frequency and synchronization resulting from the positive feedback between neuronal activity and synaptic connectivity induced by SP (Abraham, 2008; Watt and Desai, 2010). In recurrent network models, these "runaway" effects manifest as a decrease of dynamics complexity toward regular dynamics (i.e., from positive to negative maximal Lyapunov exponents (Siri et al., 2007, 2008). It was proposed that HIP might oppose this dead-end evolution (Siri et al., 2007, 2008). However, this issue has remained unexplored hitherto. Unfortunately, the issue of HIP and SP interaction cannot be addressed in the realm of dynamic mean-field theory. Indeed, SP creates correlations between synaptic weights, whereas mean-field theory assumes statistical independence between synaptic weights. To our knowledge, no current theoretical tool yet exists to handle this case, so we relied on numerical simulations to investigate these issues.

We typically found that the dynamics of the network corresponded to a fixed point at the end of the simulations ($T = 2000$) when SP was present, as illustrated for the activity of an individual neuron (Fig. 4A, upper) and of the network (Fig. 4B, upper). In contrast, the activity in presence of HIP in addition to SP led to complex patterns of activity (Fig. 4A, lower firing rate histogram; B, lower raster). To quantify this apparent increase in complexity, we computed the largest Lyapunov exponent in standard recurrent networks endowed with SP or SP and HIP (see Materials and Methods). We followed the evolution of the maximal Lyapunov exponent (averaged over 10 realizations of the network) as a function of the learning epoch T. We first observed that in the presence of the sole SP, the steady-state value of the largest Lyapunov exponent, $L_1^{(\infty)}$, was negative (Fig. 4C, purple). This occurred whatever the value of $\alpha/(1 - \lambda)$, which quantifies the magnitude of SP (Siri et al., 2007, 2008). Note that the largest Lyapunov exponent increased with $\alpha/(1 - \lambda)$. This arose because at lower values of this ratio, passive forgetting dominates the Hebbian term (see Eq. 3), so that synaptic weights strongly decrease and neurons tend to be disconnected. In such a situation, the network encountered the most static situation possible, i.e., the lowest largest Lyapunov exponent. Above the range illustrated ($\alpha/(1 - \lambda) < 200$), the Hebbian rate became so large that the network was unstable, i.e., the synaptic matrix did not converge anymore to a steady state and the Lyapunov exponent exhibited oscillations. We therefore limited our exploration to this range. Even when initial conditions gave rise to an initially positive largest Lyapunov exponent $L_1^{(1)}$, the SP rule led to a decrease of $L_1^{(T)}$ to negative steady-state
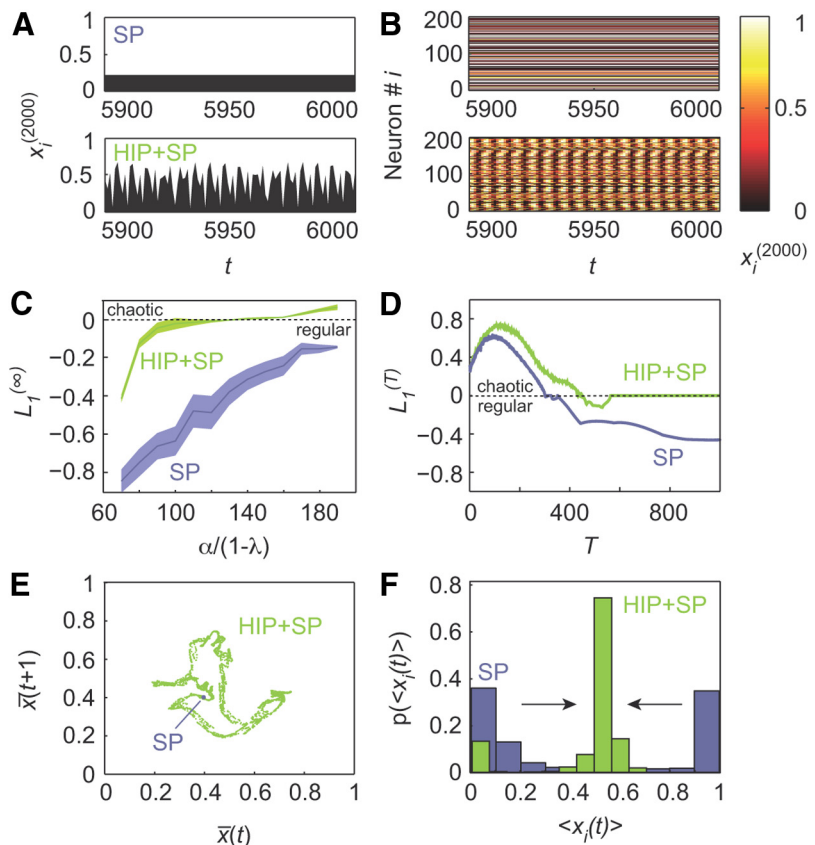

Figure 4. HIP and SP interaction in the recurrent network. **A**, Firing rate histograms as a function of the time t within learning epoch ($T = 2000$) when SP is present, in the absence (upper) or presence (lower) of HIP. **B**, Network activity rasters corresponding to **A**. **C**, Steady-state mean value of the largest Lyapunov exponent, $L_1^{(\infty)}$, at the end of the simulations ($T = 2000$) for a balanced network with SP, Dale's principle, and sparse connectivity, in the presence (green) or absence (purple) of HIP, as a function of the SP rate, $\alpha/(1 - \lambda)$. Values are averaged over 10 network realizations. SEM of $L_1^{(\infty)}$ in the presence or absence of HIP is represented as light green and light purple areas, respectively. **D**, Evolution of $L_1^{(T)}$ during network learning in the presence (green) or absence (purple) of HIP. **E**, Bidimensional reconstruction of the dynamical attractor from a sample network in **D** at the end of the simulation, using the Takens method in the presence (green) or absence (purple) of HIP. **F**, Distribution of time-averaged individual activities in the presence (green) or absence (purple) of HIP at the end of the simulations ($T = 2000$), averaged over 10 realizations. **C**, **D**, Horizontal black dotted line: $L_1 = 0$. Parameters: **A–F**, $\bar{\theta} = 20$; **A**, **B**, $\alpha/(1 - \lambda) = 160$, $\xi = 2$; **C–F**, $\alpha/(1 - \lambda) = 140$; **D–F**, $\xi = 0$.

values $L_1^{(\infty)}$ (Fig. 4D, purple). Accordingly, the network attractor observed at long learning times was a fixed point (Fig. 4E, purple). Moreover, our simulations indicated that this fixed point corresponds to a divergence of neuronal activities toward saturation (Fig. 4F, purple histogram), an illustration of the "runaway" effects induced by SP.

When HIP was added to SP, the largest Lyapunov exponent $L_1^{(T)}$ converged toward steady-state values ($L_1^{(\infty)}$) close to zero or even positive in a large domain of synaptic rates tested (Fig. 4C; D, green). Hence, the presence of HIP led the networks to chaotic dynamics ($L_1^{\infty} > 0$) or to the so-called edge of chaos ($L_1^{\infty} \approx 0$; Fig. 4E, green dots). To evaluate the influence of SP on the effects of HIP in the network model, we ran simulations, in which we varied systematically the rate of SP, $\alpha/(1 - \lambda)$. For a large range of this parameter, network dynamics reached the boundary between irregular and regular dynamics (i.e., near-zero largest Lyapunov exponent; Fig. 4C, green) and the system maintained itself in this regime (Fig. 4D, green). Finally, in these networks endowed with both SP and HIP, neuronal activities were closer to 0.5, with large variability (Fig. 4F, green). Hence, even in the presence of SP, the effects of HIP rely on the mechanism unveiled in the dynamical mean-field theory.

Moreover, our results support the hypothesis of HIP as an effective regulator of the runaway effects of SP, since it counteracts the divergence of neuronal activities and the subsequent decrease in dynamics complexity (Siri et al., 2007, 2008; Abraham, 2008; Watt and Desai, 2010).

Biological robustness of HIP effects

We assessed the effect of several important biological constraints, related to network architecture and SP balance, on the ability of HIP to control the network dynamics through its homeostatic influence.

We first compared different network architectures to investigate the impact of Dale's principle (i.e., the separation of excitatory-only and inhibitory-only neurons) and of the sparseness of the standard recurrent network (see Materials and Methods). We built the surface map of the average steady-state largest Lyapunov exponent at the end of simulations, $L_1^{(\infty)}$, as a function of the rates of SP ($\alpha/(1 - \lambda)$) and HIP ($\bar{\theta}$), for the standard network model (synaptic balance, Dale's principle, and sparse connectivity). This surface was essentially composed of a flat region that spanned large ranges of the rates space, at $L_1^{(\infty)}$ values close to 0 (Fig. 5A). Periodic orbit (PO), critical (C; quasiperiodic, i.e., marginally stable), and chaotic dynamics all occupied significant domains in this region, while fixed-point (FP) dynamics was confined to its borders, at low plasticity rates (Fig. 5B). This illustrates the strong robustness of HIP effects on dynamics complexity to large variations of plastic rates. Discarding both Dale's principle and sparse connectivity in the network did not qualitatively affect the planar shape of the $L_1^{(\infty)}$ surface, indicating that the robustness of HIP effects extended to the presence or absence of these two important architectural determinants of recurrent networks (compare Fig. 5A,B).

We then investigated whether the balance between excitatory and inhibitory synaptic drive was important for the action of HIP. In the standard network model (see Materials and Methods), the parameter α, which determines the rate for SP, is set to different values for excitatory and inhibitory synapses to compensate for the larger proportion of excitatory neurons $\left(\text{i.e., } \alpha_E = \alpha/\sqrt{3} \text{ and } \alpha_I = \sqrt{3}\alpha, \text{ thus } \frac{\alpha_I}{\alpha_E} = \frac{p_E}{p_I} = 3\right)$, so that the resulting drive is balanced (i.e., $p_E\alpha_E = p_I\alpha_I$). To destabilize this balance in these networks, we used identical values of the learning rate α for excitatory and inhibitory synapses, i.e., we set $\alpha_E = \alpha_I = \alpha$. In this case, the excitatory drive is three times larger than that of the inhibition. Contrarily to balanced networks, unbalanced networks failed to stabilize at chaotic dynamical regimes or at the edge of chaos, systematically displaying fixed-point dynamics (FP; Fig. 5 compare C with A, B). As a general rule, neuronal activities were largely saturated so that the variability of neuronal activity remained low. SP caused a divergence of local fields toward positive values that could not be compensated by HIP (Fig. 5D, bluish distribution), whereas this was possible in balanced networks (lime distribution).

Hence, these simulations specify some of the conditions necessary for a proper regulation of dynamical regime by HIP in recurrent networks. We pinpoint the negligible influence of static architectural constraints (i.e., Dale's principle and sparse connectivity), suggesting that HIP could serve a similar regulatory role in very different network types (e.g., cortical or not) in vertebrates. In contrast, HIP requires excitatory SP to be balanced by its inhibitory counterpart, which seems a ubiquitous property of vertebrate networks (Shadlen and Newsome, 1994; Tsodyks and Sejnowski, 1995; Gaiarsa et al., 2002). Hence, our modeling study stresses the necessary synergy between intrinsic and synaptic homeostatic processes to generate complex activity dynamics. Beyond that important requirement, our results indicate that the maintenance of complex dynamics by HIP is a very robust property that neither depends on (1) particular architectural constraints, (2) the presence of SP, and (3) specific rates of HIP and SP.

Functional impact of HIP on pattern recognition

Previous studies have shown that computational performance on the recognition of static or time-varying inputs is improved when recurrent networks operate at the edge of chaos (Bertschinger and Natschläger, 2004; Siri et al., 2007, 2008). Machine-learning studies have assessed whether HIP may improve the recognition of time-varying stimuli by stabilizing recurrent network dynamics complexity at this border. However, in these studies, critical dynamics and improved performance have proved rather disjoint (see Introduction; Lazar et al., 2007, 2009; Steil, 2007; Wardermann and Steil, 2007). In the present model embedded with a more biologically realistic rule based on the actual enzymatic signaling pathways responsible for intrinsic plasticity in real neurons, we evaluated the possible functional impact of HIP in learning to discriminate static input patterns. The results obtained were robust to large variations in the scaling or spatial structure of input patterns. To readout the network behavior in a very general manner (i.e., without training any particular downstream output layer as in reservoir computing), we simply considered the ability to separate input patterns, based on its activity. To do so, we computed input separability, S, the network, and time-averaged distance between individual neuronal activity in response to pairs of arbitrary input patterns $\{\xi_k, \xi_l\}$ (see Materials and Methods).

In the absence of HIP, input separability increased during learning, as exemplified in Figure 6A (three realizations of the initial synaptic weights, purple traces), being maximal when $L_1^{(T)}$ was close to zero (Fig. 4D, purple trace, i.e., around $T \approx 500$). At this edge of chaos, the network lost its stability with respect to small perturbations, exhibiting strong nonlinear responses to the presence of inputs, and was therefore more sensitive to their presentation. As the learning epoch T increased, input separability decreased after this transient behavior and stabilized below its maximal value in the present balanced network (Fig. 6A, purple traces). However, inspection of the reconstructed attractors of the dynamics at large T values, in response to the two input patterns showed that the average activity level was comparable (Fig. 6B; i.e., two fixed-point attractors). In the presence of HIP (in addition to SP), input separability was higher than in the sole presence of SP, especially at long learning times (Fig. 6A, green traces). The reconstructed attractors of the dynamics occupied a larger region in the phase space and were more discriminable from each other when HIP was present (Fig. 6C; a quasiperiodic attractor (input ξ_1, light green) and a periodic attractor (input ξ_2, dark green)). Therefore, input separability was enhanced in the presence of HIP.

To better assess this computational effect of HIP, we evaluated separability in different learning conditions, i.e., for increasing values of $\bar{\theta}$ (the magnitude of HIP) and when teaching the network with $n_\xi = 2, 3,$ or 4 input patterns. Consistent with our initial findings (Fig. 6A–C), we found that separability increased with the magnitude of HIP, compared with when SP only was present (Fig. 6D). Moreover, whereas separability was independent of the number of input patterns in the absence of HIP (Fig. 6E, purple), the enhanced separability due to HIP was remarkably

increased when larger numbers of input patterns were presented to train the network (Fig. 6E, green).

Finally, to address the issue of how input separability relates to dynamics complexity in the biologically realistic networks considered in the present study, we systematically plotted separability as a function of the maximal Lyapunov exponent $L_1^{(\infty)}$ evaluated at the end of simulations (i.e., after learning), for all the conditions and network realizations simulated (Fig. 6F, increasing HIP magnitude: graded color tints from purple to green; n_ξ = 2 (circles), 3 (squares), or 4 (triangles)). We found that increasing values of $\bar\theta$ led to more complex dynamics (larger $L_1^{(\infty)}$ values) after learning, up to the edge of chaos ($L_1^{(\infty)} \approx 0$; no strongly supracritical dynamics was found, in contrast to simulations run in the absence of external inputs, where $L_1^{(\infty)}$ could rise up to ~0.1; Fig. 5B). Linear regression of the data (Fig. 6F) indicated that, as an overall trend, input separability increased with the complexity of the dynamics (as a larger magnitude of HIP was considered). Thus, separability was enhanced as the complexity of dynamics approached criticality. In comparison, in realizations of the network devoid of any form of plasticity

Figure 5. Biological robustness of HIP effects. *A*, 2D map view of the steady-state value of the largest Lyapunov, $L_1^{(\infty)}$, exponent at the end of the simulations ($T = 2000$), as a function of the SP rate, $\alpha/(1 - \lambda)$, and the HIP rate, $\bar\theta$, for balanced networks with Dale's principle and sparse connectivity, with boundaries (purple curves) qualitatively separating FP (Fixed Point) and PO (Periodic Orbit) dynamics ($L_1 = -10^{-1}$), PO and C (Critical; i.e., quasiperiodic) dynamics ($L_1 = -10^{-2}$), and critical and chaotic dynamics ($L_1 = 0$). *B*, Same as *A* for balanced networks devoid of Dale's principle and sparse connectivity. *C*, Same as *A* for unbalanced networks with Dale's principle and sparse connectivity. *D*, Distribution of time-averaged activities, at the end of simulations, in networks with HIP, Dale's principle, and sparse connectivity, in the presence of unbalanced (bluish green) and balanced (lime) SP.

(Fig. 6F, black dots), we found that whereas the complexity of dynamics ranged in a large region encompassing criticality ($L_1^\infty \epsilon[-0.2, 0.4]$), the separability remained low ($s \approx 0.01$), compared with when plasticity was present ($S \epsilon [0.08, 0.18]$). Therefore, the dependence between the separability and the complexity of the dynamics was an acquired property of the network, which emerged from the interaction between synaptic and intrinsic and plasticity.

Together, these results indicate that criticality cannot be considered as a sufficient condition for improved computational performance (as found in some reservoir networks endowed with artificial plastic rules; Lazar et al., 2007). However, in networks endowed with realistic biophysical plastic processes, while SP alone increases separability at subcritical dynamics, separability significantly improves when dynamics approaches criticality, thanks to the combination of SP and HIP.

Discussion

To our knowledge, we propose the first network model including a biologically realistic HIP rule, which is neither phenomenological (Lazar et al., 2006, 2007, 2009; Remme and Wadman, 2012) nor designed to maximize information transmission (Stemmler and Koch, 1999; Steil, 2007; Wardermann and Steil, 2007; Schrauwen et al., 2008; Marković and Gros, 2012). Here, HIP regulates the neuronal threshold through modifications of subthreshold voltage-gated conductances (Naudé et al., 2012), as frequently observed following HIP induction (Desai et al., 1999; Gibson et al., 2006; van Welie et al., 2006). Moreover, HIP modifications operate through a realistic description of signaling pathways' kinetics regulating intrinsic conductances in real neurons (Delord et al., 2007).
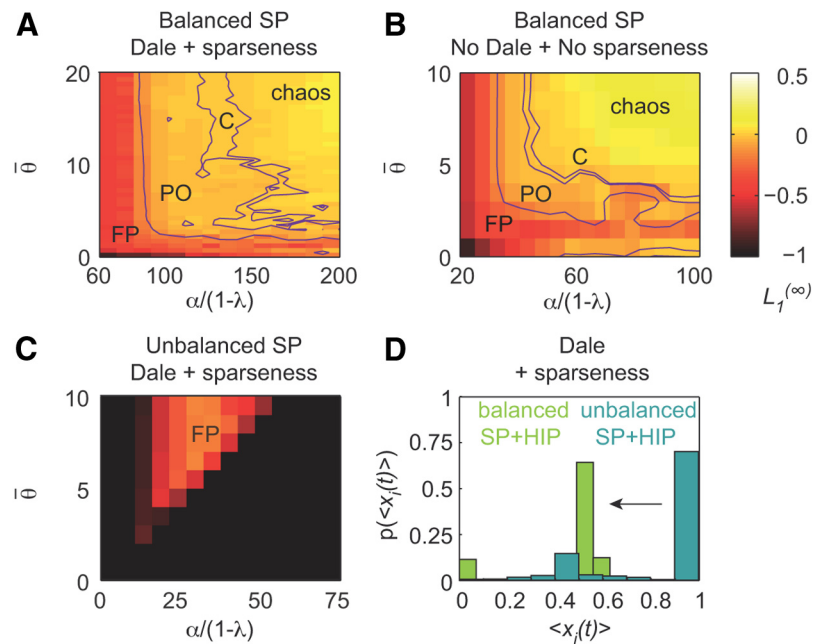
Functionally, it was initially proposed that HIP of the threshold is able to control the level of neuronal activity, avoiding neuronal saturation (Desai et al., 1999; Fig. 7A). Indeed, depending on its activity-dependence (e.g., calcium kinetics, kinase/phosphatase calcium dependence), a HIP rule can bring neurons to any arbitrary working point in the linear range of the transfer function. Besides, it was suggested that HIP actually maximizes information transmission (i.e., mutual information) between input and output distributions (Stemmler and Koch, 1999). Since, an artificial HIP rule was derived to implement this principle (Triesch, 2005) and improve reservoir learning (Steil, 2007; Schrauwen et al., 2008). In practice, such a rule adjusts the threshold so that low inputs are mapped evenly in the output distribution (Fig. 7B). In vertebrates, this principle makes sense at early sensory stages to transmit maximal information; however, in higher structures, neurons rather specifically respond to large input fluctuations, losing large amounts of information, to ensure neuronal selectivity (Ringach and Malone, 2007). HIP may either favor selection (Fig. 7C) or transmission in different neurons, depending on the HIP rule and transfer function (e.g., overall gain, symmetry).

The present study suggests that, beyond setting the overall level of activity or input processing, a prominent effect of HIP is to increase the complexity of activity dynamics, which is fundamental for higher order vertebrates' computations. This increase is consistent with that observed in reservoir networks, which share recurrent architecture with vertebrate networks (Lazar et al., 2007; Wardermann and Steil, 2007; Marković and Gros, 2012). In Lazar et al. (2007), HIP decreased complexity, but only when dynamics were spontaneous and SP present. Thus, our results strengthen the principle that increased dynamics com-
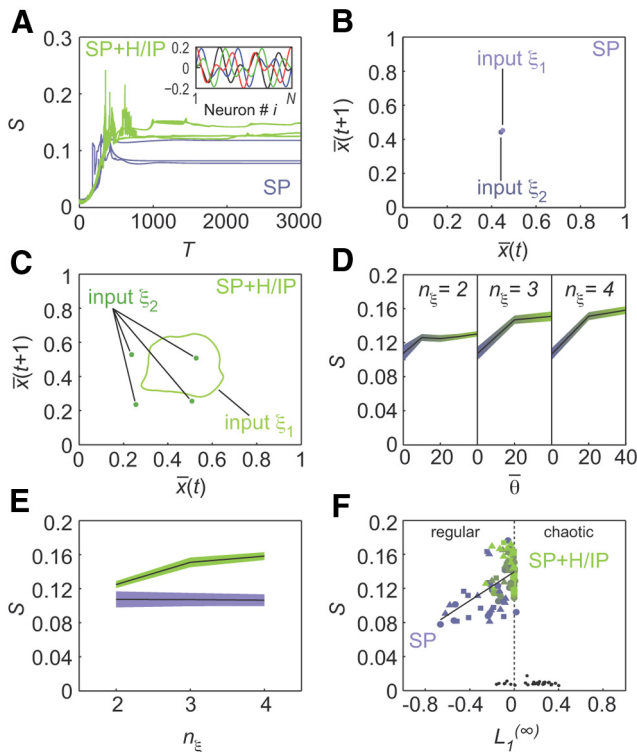
Figure 6. Functional impact of HIP on input recognition. The presence of HIP improves the network ability to discriminate input patterns, i.e., separability. **A**, Separability between two input patterns ξ_1 and ξ_2 alternatively presented to the network (see Materials and Methods), as a function of the learning epoch, in the absence (purple traces) or presence (green traces) of HIP ($\bar{\theta} = 20$), for three realizations of the initial synaptic matrix. Inset: input patterns ξ_1 (black), ξ_2 (blue), used here, and ξ_3 (red) and ξ_4 (green) used in addition to patterns ξ_1 and ξ_2 in **D–F**. **B**, Bidimensional reconstruction of the dynamical attractor for one realization in the sole presence of SP, when ξ_1 and ξ_2 are presented to the recurrent network at the end of the simulation ($T = 3000$). **C**, Same as in **B** when HIP is present in addition to SP. **D**, Separability as a function of $\bar{\theta}$, the magnitude of HIP, when the network is trained with $n_\xi = 2, 3,$ or 4 inputs. **E**, Separability as a function of n_ξ, the number of inputs used to train the network ($\bar{\theta} = 40$). **F**, Separability as a function of $L_1^{(\infty)}$, the maximal Lyapunov exponent, at the end of simulations for all the learning conditions and network realizations simulated ($\bar{\theta} = \{0, 10, 20, 30, 40\}$: graded color tints from purple to green; $n_\xi = 2$ (circles), 3 (squares), or 4 (triangles). Linear regression: $p < 10^{-6}$. Black dots correspond to realizations of the network devoid of plasticity. **A–F**, $\alpha/(1 - \lambda) = 160$.

plexity represents a general and robust (see below) consequence of HIP in biological recurrent networks. The mean-field approach we have devised indicates that increased complexity relies on a yet unexplored cellular effect of HIP: setting the neuronal working point so that input distributions are transduced by ranges of the transfer function with large derivative (Fig. 7D). Consequently, this principle may have important functional consequences, which we discuss now.

It has been suggested that HIP may compensate runaway effects of SP, i.e., destabilization to saturated activity levels arising from the positive feedback between activity and connectivity (Turrigiano, 1999; Watt and Desai, 2010). Previous studies from our group unraveled that SP-mediated pathological activity actually results from a general decrease of networks dynamics complexity due to SP (Siri et al., 2007, 2008). Here, we show that HIP increases dynamics complexity (through its effect on the transfer function derivative) and opposes the SP-mediated decrease, bringing network dynamics at the edge of chaos. This effect was only observed when SP preserved synaptic excitation/inhibition balance, an important determinant of network dynamics with
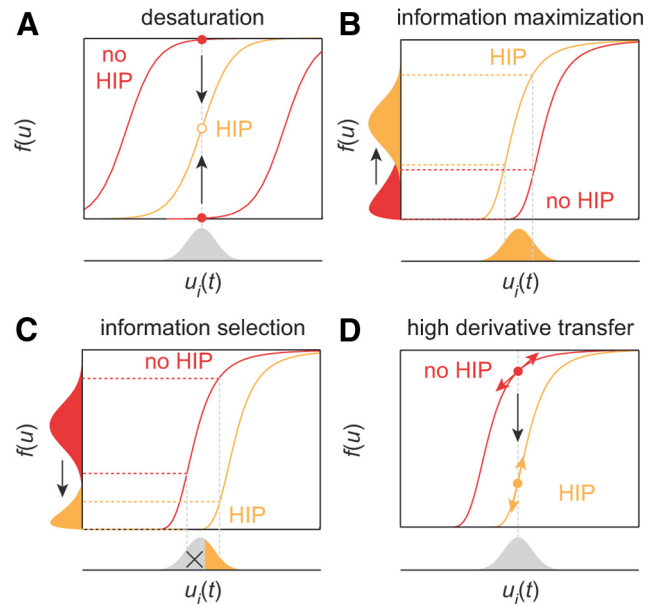

Figure 7. Proposed mechanisms for HIP effects. **A**, Desaturation. In response to a (gray) distribution of inputs $u_i(t)$, HIP can shift the transfer function (yellow curve) so that the average response (yellow dot) is not saturated or quiescent anymore, as in the absence of HIP (red curves and dots). **B**, Information maximization. HIP can shift the transfer function (yellow curve) so that the entire input distribution is mapped onto the output (yellow) distribution, whereas in the absence of HIP (red curve), low inputs are not mapped in the output (red) distribution. **C**, Information selection. HIP can subserve the opposite operation, to produce a selective mapping whereby the neuron responds only to large inputs fluctuations (yellow part of the input distribution). **D**, High derivative transfer. We propose that HIP exerts its prominent effect on recurrent network dynamics through shifts of the threshold, so that the input distribution is transduced by large derivative ranges of the transfer function.

uncorrelated synaptic weights (Shadlen and Newsome, 1994; Tsodyks and Sejnowski, 1995; van Vreeswijk and Sompolinsky, 1996). Our results thus suggest that HIP might act in concert with SP to control the dynamics complexity of recurrent networks, as was suggested for their activity level (Remme and Wadman, 2012). Such a synergistic regulation could arise from the coactivation of HIP and SP, through common signaling pathways (Misonou et al., 2004; van Welie et al., 2004; Fan et al., 2005; Karmarkar and Buonomano, 2006; Turrigiano, 2011). In two studies of the Triesch group, dynamics of the reservoir networks endowed with HIP and SP were clearly subcritical (Lazar et al., 2007, 2009). Such regular dynamics presumably emerges from the interaction between time-structured inputs and the spike timing-dependent SP rule considered in these studies. In contrast, here, inputs are static and the rules are not time dependent. The discrepancy might also possibly arise from the crude description of the inhibition (global k-winner-take-all dynamics) used in Lazar et al. (2007).

Moreover, with larger HIP magnitudes, performance increased as dynamics approaches the edge of chaos. In this region, the dynamics did not systematically exhibit exact criticality, but rather covered distinctive behaviors situated around this particular frontier, which included long periodic orbits and quasiperiodic or chaotic dynamics. The specific regime encountered depended on the input pattern, learning parameters, and initial conditions of the weights and thresholds. This result is consistent with previous studies of our group, where sensitivity to a single static input pattern was maximal at critical dynamics in recurrent networks endowed with SP (Siri et al., 2007, 2008). It also parallels the previous observation that in static reservoir recurrent

networks (no plasticity, trained readout), computational performances on time-varying inputs are improved at criticality (Bertschinger and Natschläger, 2004). However, it was shown that HIP-endowed recurrent networks with critical dynamics do not systematically perform better than static networks (Lazar et al., 2007). Moreover, HIP noticeably improves reservoir networks performances that clearly exhibit subcritical dynamics (i.e., regular; Steil, 2007; Wardermann and Steil, 2007; Lazar et al., 2009). This discrepancy between the present study and these studies (Lazar et al., 2007, 2009; Steil, 2007; Wardermann and Steil, 2007) may stem from the different underlying operation, i.e., steady-state dynamics upon static inputs here (activity level, attractor nature) versus transient dynamics (i.e., "fading memory") in response to time-varying inputs in reservoir networks. In particular, in reservoir models endowed with spike timing-dependent plasticity (Lazar et al., 2007, 2009), time-dependent SP and time-structured inputs presumably represent a common condition for both generating sub-critical dynamics (see above) and ensuring input recognition. Here, inputs are static and SP is not time dependent so that what is learned is the spatial structure of inputs. Thus, one might expect that best performances should be observed at very regular dynamics, e.g., fixed-point attractors. Performance is indeed enhanced at such regular regimes with SP. However, when HIP is added, performance culminates at criticality. This discrepancy might arise from the rules considered: actual kinetics of signaling pathways regulating excitability in real neurons (Delord et al., 2007; Naudé et al., 2012), here, versus phenomenological (Lazar et al., 2007, 2009) or optimized rules (Steil, 2007; Wardermann and Steil, 2007; Schrauwen et al., 2008). This divergence may require future efforts to be fully understood, as the mechanisms underlying improved performance have remained a constant question (Lazar et al., 2007, 2009; Steil, 2007; Schrauwen et al., 2008).

Here, we observed that input separability increases with the number of input patterns presented to train the network. However, for practical reasons, we restricted our study to limited pattern numbers (i.e., 4). In the future, it would be interesting to assess the effects of HIP on the memory capacity of networks for larger numbers of input patterns. Another appealing possibility would be to address possible improvements in the learning of complex activity patterns in FORCE networks that endow supervised network SP and a trained readout (Sussillo and Abbott, 2009), when incorporating HIP.

Our results indicate that the maintenance of complex dynamics by HIP in biological networks is a very robust property that neither depends on the size of the network, particular architectural constraints (Dale's principle and sparse connectivity), the presence of SP, specific learning rates, or the scaling and spatial patterning of inputs. In particular, the extreme robustness of HIP in bringing network dynamics at the edge of chaos in the presence of SP is stunning (Fig. 4C, 5 A, B). Understanding this robustness is an important challenge for future studies that might unravel principles underlying self-organization around criticality in recurrent neural networks. Recently, it was suggested that HIP of both the threshold and gain of the transfer function might bring complex activity patterns in recurrent networks (Markovíc and Gros, 2012). Unfortunately, these behaviors were not extremely robust to changes in architectural or HIP parameters and were underlain by fast kinetics bringing rapid oscillations of excitability with no current experimental support. However, examining how the robustness we have observed with threshold HIP would extend to both threshold and gain HIP using slow kinetics in biological networks appears as an exciting possibility.

Together, the present results suggest that HIP effects may represent a ubiquitous determinant of biological networks across neuronal systems. Beyond the impressive robustness we have observed in recurrent architectures, the attractive perspective emerges that HIP processes share a general role in setting the working point of neural networks, independently of their specific function, e.g., the organization of rhythmic activity in invertebrates or the maintenance of complex dynamics for higher order computations and nontrivial behavior formation in vertebrates.

References
Abraham WC (2008) Metaplasticity: tuning synapses and networks for plasticity. Nat Rev Neurosci 9:387. CrossRef Medline

Arous GB, Guionnet A (1995) Large deviations for Langevin spin glass dynamics. Probab Theory Related Fields 102:455–509. CrossRef

Bekkers JM, Delaney AJ (2001) Modulation of excitability by alpha-dendrotoxin-sensitive potassium channels in neocortical pyramidal neurons. J Neurosci 21:6553–6560. Medline

Bertschinger N, Natschläger T (2004) Real-time computation at the edge of chaos in recurrent neural networks. Neural Comput 16:1413–1436. CrossRef Medline

Brickley SG, Revilla V, Cull-Candy SG, Wisden W, Farrant M (2001) Adaptive regulation of neuronal excitability by a voltage-independent potassium conductance. Nature 409:88–92. CrossRef Medline

Cessac B (1995) Increase in complexity in random neural networks. J Phys I (France) 5:409–432. CrossRef

Cudmore RH, Turrigiano GG (2004) Long-term potentiation of intrinsic excitability in LV visual cortical neurons. J Neurophysiol 92:341–348. CrossRef Medline

Delord B, Berry H, Guigon E, Genet S (2007) A new principle for information storage in an enzymatic pathway model. PLoS Comput Biol 3:e124. CrossRef Medline

Desai NS, Rutherford LC, Turrigiano GG (1999) Plasticity in the intrinsic excitability of cortical pyramidal neurons. Nat Neurosci 2:515–520. CrossRef Medline

Doyon B, Cessac B, Quoy M, Samuelides M (1994) On bifurcations and chaos in random neural networks. Acta Biotheoretica 42:215–225. CrossRef

Doyon B, Cessac B, Quoy M, Samuelides M (1995) Mean-field equations, bifurcation map and chaos in discrete time, continuous state, random neural networks. Acta Biotheoretica 43:169–175. CrossRef Medline

Eckmann JP, Ruelle D (1985) Ergodic theory of chaos and strange attractors. Rev Mod Phys 57:617–656. CrossRef

Fan Y, Fricker D, Brager DH, Chen X, Lu HC, Chitwood RA, Johnston D (2005) Activity-dependent decrease of excitability in rat hippocampal neurons through increases in I(h). Nat Neurosci 8:1542–1551. CrossRef Medline

Faugeras O, Touboul J, Cessac B (2009) A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs. Front Comput Neurosci 3:1. CrossRef Medline

Gaiarsa JL, Caillard O, Ben-Ari Y (2002) Long-term plasticity at GABAergic and glycinergic synapses: mechanisms and functional significance. Trends Neurosci 25:564–570. CrossRef Medline

Ganguly K, Kiss L, Poo M (2000) Enhancement of presynaptic neuronal excitability by correlated presynaptic and postsynaptic spiking. Nat Neurosci 3:1018–1026. CrossRef Medline

Gibson JR, Bartley AF, Huber KM (2006) Role for the subthreshold currents ILeak and IH in the homeostatic control of excitability in neocortical somatostatin-positive inhibitory neurons. J Neurophysiol 96:420–432. CrossRef Medline

Helmchen F, Imoto K, Sakmann B (1996) Ca2+ buffering and action potential-evoked Ca2+ signaling in dendrites of pyramidal neurons. Biophys J 70:1069–1081. CrossRef Medline

Karmarkar UR, Buonomano DV (2006) Different forms of homeostatic plasticity are engaged with distinct temporal profiles. Eur J Neurosci 23:1575–1584. CrossRef Medline

Lazar A, Pipa G, Triesch J (2006) The combination of STDP and intrinsic plasticity yields complex dynamics in recurrent spiking networks. ESANN'2006 proceedings 647–652.

Lazar A, Pipa G, Triesch J (2007) Fading memory and time series prediction

in recurrent networks with different forms of plasticity. Neural Netw 20:312–322. CrossRef Medline

Lazar A, Pipa G, Triesch J (2009) SORN: a self-organizing recurrent neural network. Front Comput Neurosci 3:23. Medline

LeMasson G, Marder E, Abbott LF (1993) Activity-dependent regulation of conductances in model neurons. Science 259:1915–1917. CrossRef Medline

Maffei A, Turrigiano GG (2008) Multiple modes of network homeostasis in visual cortical layer 2/3. J Neurosci 28:4377–4384. CrossRef Medline

Marder E, Goaillard JM (2006) Variability, compensation and homeostasis in neuron and network function. Nat Rev Neurosci 7:563–574. CrossRef Medline

Marković D, Gros C (2012) Intrinsic adaptation in autonomous recurrent neural networks. Neural Comput 24:523–540. CrossRef Medline

Misonou H, Mohapatra DP, Park EW, Leung V, Zhen D, Misonou K, Anderson AE, Trimmer JS (2004) Regulation of ion channel localization and phosphorylation by neuronal activity. Nat Neurosci 7:711–718. CrossRef Medline

Moynot O, Samuelides M (2002) Large deviations and mean-field theory for asymmetric random recurrent neural networks. Prob Theory Relat Fields 123:41–75. CrossRef

Naudé J, Paz JT, Berry H, Delord B (2012) A theory of rate coding control by intrinsic plasticity effects. PLoS Comput Biol 8:e1002349. CrossRef Medline

O'Leary T, van Rossum MC, Wyllie DJ (2010) Homeostasis of intrinsic excitability in hippocampal neurones: dynamics and mechanism of the response to chronic depolarization. J Physiol 588:157–170. CrossRef Medline

Remme MW, Wadman WJ (2012) Homeostatic scaling of excitability in recurrent neural networks. PLoS Comput Biol 8:e1002494. CrossRef Medline

Ringach DL, Malone BJ (2007) The operating point of the cortex: neurons as large deviation detectors. J Neurosci 27:7673–7683. CrossRef Medline

Schrauwen B, Wardermann M, Verstraeten D, Steil JJ, Stroobandt D (2008) Improving reservoirs using intrinsic plasticity. Neurocomputing 71:1159–1171. CrossRef

Schulz DJ (2006) Plasticity and stability in neuronal output via changes in intrinsic excitability: it's what's inside that counts. J Exp Biol 209:4821–4827. CrossRef Medline

Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. Curr Opin Neurobiol 4:569–579. CrossRef Medline

Siri B, Quoy M, Delord B, Cessac B, Berry H (2007) Effects of Hebbian learning on the dynamics and structure of random networks with inhibitory and excitatory neurons. J Physiol Paris 101:136–148. CrossRef Medline

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 20:2937–2966. CrossRef Medline

Sompolinsky H, Crisanti A, Sommers HJ (1988) Chaos in random neural networks. Phys Rev Lett 61:259–262. CrossRef Medline

Steil JJ (2007) Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning. Neural Netw 20:353–364. CrossRef Medline

Stemmler M, Koch C (1999) How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. Nat Neurosci 2:521–527. CrossRef Medline

Sussillo D, Abbott LF (2009) Generating coherent patterns of activity from chaotic neural networks. Neuron 63:544–557. CrossRef Medline

Takens F (1981) Detecting strange attractors in turbulence. In: Dynamical systems and turbulence, Warwick, 1980. Lecture notes in mathematics (Rand A, Young LS, eds), pp 366–381. New York: Springer.

Triesch J (2005) A gradient rule for the plasticity of a neuron's intrinsic excitability. In: Artificial neural networks: biological inspirations—ICANN 2005 (Duch W, Kacprzyk J, Oja E, Zadrozny S, eds), pp 65–70. New York: Springer.

Tsodyks M, Sejnowski T (1995) Rapid state switching in balanced cortical network models. Network 6:111–124. CrossRef

Turrigiano G (2011) Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. Annu Rev Neurosci 34:89–103. CrossRef Medline

Turrigiano GG (1999) Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. Trends Neurosci 22:221–227. CrossRef Medline

van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science 274:1724–1726. CrossRef Medline

van Welie I, van Hooft JA, Wadman WJ (2004) Homeostatic scaling of neuronal excitability by synaptic modulation of somatic hyperpolarization-activated Ih channels. Proc Natl Acad Sci U S A 101:5123–5128. CrossRef Medline

van Welie I, van Hooft JA, Wadman WJ (2006) Background activity regulates excitability of rat hippocampal CA1 pyramidal neurons by adaptation of a K+ conductance. J Neurophysiol 95:2007–2012. CrossRef Medline

Vervaeke K, Hu H, Graham LJ, Storm JF (2006) Contrasting effects of the persistent Na+ current on neuronal excitability and spike timing. Neuron 49:257–270. CrossRef Medline

Vogels TP, Rajan K, Abbott LF (2005) Neural network dynamics. Annu Rev Neurosci 28:357–376. CrossRef Medline

Von Bremen H, Udwadia F, Proskuroswki W (1997) An efficient QR based method for the computation of Lyapunov exponents. Physica D 101:1–6. CrossRef

Wang XJ (1998) Calcium coding and adaptive temporal computation in cortical pyramidal neurons. J Neurophysiol 79:1549–1566. Medline

Wardermann M, Steil JJ (2007) Intrinsic plasticity for reservoir learning algorithms. ESANN'2007 proceedings 513–518.

Watt AJ, Desai NS (2010) Homeostatic plasticity and STDP: keeping a neuron's cool in a fluctuating world. Front Synaptic Neurosci 2:5. Medline

Zhang W, Shin JH, Linden DJ (2004) Persistent changes in the intrinsic excitability of rat deep cerebellar nuclear neurones induced by EPSP or IPSP bursts. J Physiol 561:703–719. CrossRef Medline