Energetics of stochastic BCM type synaptic plasticity and storing of accurate information

Jan Karbowski

Institute of Applied Mathematics and Mechanics,

University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland

Abstract

Excitatory synaptic signaling in cortical circuits is thought to be metabolically expensive. Two fundamental brain functions, learning and memory, are associated with long-term synaptic plasticity, but we know very little about energetics of these slow biophysical processes. This study investigates the energy requirement of information storing in plastic synapses for an extended version of BCM plasticity with a decay term, stochastic noise, and nonlinear dependence of neuron's firing rate on synaptic current (adaptation). It is shown that synaptic weights in this model exhibit bistability. In order to analyze the system analytically, it is reduced to a simple dynamic mean-field for a population averaged plastic synaptic current. Next, using the concepts of nonequilibrium thermodynamics, we derive the energy rate (entropy production rate) for plastic synapses and a corresponding Fisher information for coding presynaptic input. That energy, which is of chemical origin, is primarily used for battling fluctuations in the synaptic weights and presynaptic firing rates, and it increases steeply with synaptic weights, and more uniformly though nonlinearly with presynaptic firing. At the onset of synaptic bistability, Fisher information and memory lifetime both increase sharply, by a few orders of magnitude, but the plasticity energy rate changes only mildly. This implies that a huge gain in the precision of stored information does not have to cost large amounts of metabolic energy, which suggests that synaptic information is not directly limited by energy consumption. Interestingly, for very weak synaptic noise, such a limit on synaptic coding accuracy is imposed instead by a derivative of the plasticity energy rate with respect to the mean presynaptic firing, and this relationship has a general character that is independent of the plasticity type. An estimate for primate neocortex reveals that a relative

metabolic cost of BCM type synaptic plasticity, as a fraction of neuronal cost related to fast synaptic transmission and spiking, can vary from negligible to substantial, depending on the synaptic noise level and presynaptic firing.

Email: jkarbowski@mimuw.edu.pl

Keywords: energy cost of synaptic plasticity; accurate storing of synaptic information; bistability; memory lifetime; metabolism; thermodynamic limits on synaptic information.

1. Introduction

Information and energy are intimately related for all physical systems because information has to be written on some physical substrate which always comes at some energy cost (Landauer 1961; Bennett 1982; Leff and Rex 1990; Berut et al 2012; Parrondo et al 2015). Brains are physical devices that process information and simultaneously dissipate energy (Levy and Baxter 1996; Laughlin et al 1998) in the form of heat (Karbowski 2009). This energetic cost is relatively high (Laughlin et al 1998; Aiello and Wheeler 1995; Attwell and Laughlin 2001; Karbowski 2007), which is the likely cause for a sparse coding strategy in neural circuits (Balasubramanian et al 2001; Niven and Laughlin 2008). Experimental studies (Shulman et al 2004; Logothetis 2008; Alle et al 2009), as well as theoretical calculations based on data (Harris et al 2012; Karbowski 2012), indicate that fast synaptic signaling, i.e. synaptic transmission, together with neuron's action potentials are the major consumers of metabolic energy. This type of energy use is of electric origin, and is caused by flows of electric charge due to voltage and concentration gradients and subsequent pumping ions out to maintain these gradients (Attwell and Laughlin 2001; Karbowski 2009). This very pumping of electric charge requires large amounts of energy.

Brains are also highly adapting objects, which learn and remember by encoding and storing long-term information in excitatory synapses (dendritic spines) (Kasai et al 2003; Takeuchi et al 2014). These important slow processes are driven by correlated electric activities of pre- and post-synaptic neurons (Markram et al 1997; Bienenstock et al 1982; Miller and MacKay 1994, Song et al 2000; Van Rossum et al 2000) and cause

plastic modifications in spine's intrinsic molecular machinery, leading to changes in spine size, its conductance (weight) and postsynaptic density (PSD) (Kasai et al 2003; Bonhoeffer and Yuste 2002; Holtmaat et al 2005; Meyer et al 2014). Consequently synaptic plasticity and associated information writing and storing must cost energy, since spines require some energy for inserting and maintaining AMPA and NMDA receptors on spine membrane (Huganir and Nicoll 2013; Choquet and Triller 2013), as well as for powering various molecular processes associated with PSD (Lisman et al 2013; Miller et al 2005). In contrast to fast synaptic transmission and neuron discharges, which are of electric nature, the plastic slow synaptic processes are of chemical origin (interactions between spine proteins), and thus require the chemical energy (Karbowski 2019).

One of the empirical manifestations of the plasticity-energy relationship is present for mammalian cortical development, during which synaptic density can change several fold and strongly correlates with changes in glucose metabolic rate of cortical tissue (Karbowski 2012). Unfortunately, despite a massive literature on modeling synaptic plasticity (e.g. Bienenstock et al 1982; Miller and MacKay 1994, Song et al 2000; Van Rossum et al 2000; Billings and Van Rossum 2009; Clopath et al 2010; Pfister and Gerstner 2006; Tetzlaff et al 2011; Toyoizumi et al 2014; Costa et al 2015; Shouval et al 2002; Graupner and Brunel 2012; Ziegler et al 2015; Fusi et al 2005; Benna and Fusi 2016; Gutig et al 2003; Smolen et al 2012), our theoretical understanding of the energetic requirements underlying synaptic plasticity and memory storing is currently lacking. In particular, we do not know the answers to the basic questions, such as how

does energy consumed by plastic synapses depend on key neurophysiological parameters, and more importantly, whether energy restricts the precision of synaptically encoded information and its lifetime, and to what extent. Such a knowledge might lead to a deeper understanding of two fundamental problems in neuroscience: one related to the physical cost and control of learning and memory in the brain (Kasai et al 2003; Takeuchi et al 2014; Lisman et al 2013; Costa et al 2015; Kandel et al 2014; Chaudhuri and Fiete 2016; Zenke and Gerstner 2017), and another more practical related to dissecting the contribution of synaptic plasticity to signals in brain imaging (Attwell and Laughlin 2001; Logothetis 2008; Engl et al 2017; Magistretti et al 1999; Shulman et al 2004). A recent study by the author (Karbowski 2019) provided some answers to the above questions, by analyzing molecular data in synaptic spines and by modeling energy cost of learning and memory in a cascade model of synaptic plasticity (mimicking molecular interactions in spines). From that study it follows that the average cost of synaptic plasticity constitutes a small fraction of the metabolic cost used for fast excitatory synaptic transmission, about $4 - 11\%$, and that storing longer memory traces can be relatively cheap (Karbowski 2019). However, this study left open other questions, e.g., how does the energy cost of synaptic plasticity depend on neuronal firing rates, synaptic noise, and other neural characteristics, and what is the relationship between such energy cost and a precise storing of synaptic information?

The main goal of this study is to uncover a relationship between synaptic plasticity, its energetics, and a precise information storing at excitatory synapses for one of the best known forms of synaptic plasticity due to Bienenstock, Cooper, and Munro, the

so-called BCM rule (Bienenstock et al 1982). This is a different (more macroscopic) but a complementary level of modeling to the one (microscopic) in (Karbowski 2019). Specifically, we want to determine the energy cost associated with the accuracy of information encoded in a population of plastic synapses about the presynaptic input. Additionally, we want to find the relationship between plastic energy rate and memory duration about a single event at synapses. In other words, our goal is to find the metabolic requirement of maintaining an accurate information at synapses in the face of ongoing variable neural activity and thermodynamic fluctuations inside spines associated with variation in the number of membrane receptors. The phenomenological BCM rule has been shown to explain several key experimental observations (Cooper and Bear 2012), and it is equivalent to a more microscopic STDP rule (Markram et al 1997; Song et al 2000; Van Rossum 2000) under some very general conditions (Pfister and Gerstner 2006; Izhikevich and Desai 2003). Since, the BCM rule is believed to describe initial phases of learning and memory (Zenke and Gerstner 2017), the focus of this work is on the energy cost and coding accuracy of the early synaptic plasticity, i.e. early long-term potentiation (e-LTP) and depression (e-LTD), which lasts from minutes to several hours. We do not consider explicitly the effects of memory consolidation that operate on much longer time scales and which are associated with late phases of LTP and LTD (l-LTP and l-LTD) (Ziegler et al 2015; Benna and Fusi 2016; Redondo and Morris 2011). However, we do provide a rough estimate of the energetics of these late processes, and they turn out to be much less energy demanding than the early phase plasticity.

One can question whether the approach taken here, with the macroscopic BCM type model, is reasonable for modeling and calculating energy cost of synaptic plasticity? Maybe a more microscopic approach should be used with explicit molecular interactions between PSD proteins? However, the basic problem with such a microscopic more detailed approach is that we do not know most of the molecular signaling pathways in a dendritic spine, we do not know the rates of various reactions, and even the basic mechanism of encoding information at synapses is unclear. For example, for a long time it was thought that CaMKII persistent autophosphorylation provides a basic mechanism of information storage via bistability (Lisman et al 2013; Miller et al 2005). However, experimental data indicate that CaMKII enhanced activity after spine activation is transient and lasts only about 2 min (Lee et al 2009), which casts doubts on its persistent enzymatic activity and its role as a "memory molecule" (for a review see, Smolen et al 2019). Taking all these uncertainties into account, it seems that more macroscopic approach might be more reliable, at least partly.

Because synapses/spines are small, they are strongly influenced by thermal fluctuations (Kasai et al 2003; Choquet and Triller 2013; Statman et al 2014). For this reason, this paper uses universal methods of stochastic dynamical systems and non-equilibrium statistical mechanics (Nicolis and Prigogine 1977; Van Kampen 2007; Risken 1996; Lan et al 2012; Mehta and Schwab 2012; Tome 2006; Tome and de Oliveira 2010; Seifert 2012). The latter are generally valid for all physical systems, including the brain, operating out of thermodynamic equilibrium. Regrettably, the methods of non-equilibrium thermodynamics have virtually not been used in neuroscience despite their large po-

tential in linking brain physicality with its information processing capacity, with two recent exceptions (Goldt and Seifert 2017; Karbowski 2019). (This should not be confused with equilibrium thermodynamics, whose methods have occasionally been used in neuroscience, although in a different context, e.g., Balasubramanian et al 2001; Tkacik et al 2015; Friston 2010.)

General outline of the problem considered.

It is generally believed that long-term information in excitatory synapses is encoded in the pattern of synaptic strengths or weights (membrane electric conductance), which is coupled to the molecular structure of postsynaptic density within dendritic spines (Takeuchi et al 2014; Lisman et al 2013; Miller et al 2005; Kandel et al 2014, Zhu et al 2016). This study considers the energy cost associated with maintaining the pattern of synaptic weights. In particular, we analyze the energetics and information capacity of the fluctuations in the number of AMPA and NMDA receptors on a spine membrane, or equivalently, fluctuations in the synaptic conductance. Such a variability in the receptor number tends to spread the range of synaptic weights (affecting their structure and distribution) that has a negative consequence on the encoded information and can lead to its erasure. In terms of statistical mechanics, the receptor fluctuations increase the entropy associated with the distribution of synaptic weights, and that entropy has to be reduced to preserve the information encoded in the weights. This very process of reducing the synaptic entropy production is a nonequilibrium phenomenon that costs some energy, which has to be provided by various processes involving ATP generation (Nicolis and Prigogine 1977).

The BCM type of synaptic plasticity used here is a phenomenological model that does not relate in a straightforward way to the underlying synaptic molecular processes. Empirically speaking, a change in synaptic weight in e-LTP is caused by a sequence of molecular events, of which the main are: activation of proteins in postsynaptic density, which subsequently stimulates downstream actin filaments elongation (responsible for a spine enlargement), and AMPA and NMDA receptor trafficking (Huganir and Nicoll 2013; Choquet and Triller 2013). Therefore, it is assumed here that BCM-type rule used here macroscopically reflects broadly these three microscopic processes, especially the first and the last. (Spine volume related to actin dynamics is not explicitly included in the model, although it is known experimentally that spine volume and conductance are positively correlated (Kasai et al 2003).) Thus, it is expected that the synaptic energy rate calculated here is related to ATP used mainly for postsynaptic protein activation through phosphorylation process (Zhu et al 2016), and receptor insertion and movement along spine membrane. Obviously, there are many more molecular processes in a typical spine, but they are either not directly involved in spine conductance variability or they are much faster than the above processes (e.g. releasing Ca^{2+} from internal stores is fast). A detailed empirical estimation based on molecular data suggests that protein activation via enhanced phosphorylation is the dominant contribution to the energy cost (ATP rate) of synaptic plasticity (Karbowski 2019). Therefore, the theoretical energy rate of synaptic plasticity determined here should be viewed as a minimal but a reasonable estimate of energetic requirement of LTP and LTD, and it is strictly associated with the information encoded in synaptic weights.

Experimental data show that excitatory synapses can exist in two or more stable states, characterized by discrete synaptic weights or sizes (Kasai et al 2003; Montgomery and Madison 2004; Petersen et al 1998; O'Connor et al 2005; Loewenstein et al 2011; Bartol et al 2015). Data on a single synapse level indicate that synapses can operate as binary elements either with low or high electric conductance (Petersen et al 1998; O'Connor et al 2005). On the other hand, the data on a population level, more relevant to this work, show that synapses can assume more than two stable discrete states (Kasai et al 2003; Loewenstein et al 2011; Bartol et al 2015). In either case, the issue of bistability vs. multistability is not yet resolved. In this study, a minimal scenario is considered in which synapses together with their postsynaptic neuron can effectively act as a binary coupled system, characterized by a single variable, which is the mean-field postsynaptic current with one or two stable states. The bistability is produced here from an extended BCM model, which in principle allows for continuous changes in synaptic weights for individual synapses. The important point is that these continuous weights are correlated, due to plasticity constraints, and thus converge on a mean-field population level either to one or to a couple of stable values.

Synaptic plasticity processes are induced by a correlated firing in pre- and post-synaptic neurons, and thus a model of neuron activity is also needed. This study uses a firing rate neuron model of the so-called class one nonlinear firing rate curve, which is believed to be a good approximation to biophysical neuronal models (Ermentrout 1998; Ermentrout and Terman 2010); see the Methods for details.

The paper is organized as follows. First, we introduce and solve an extended model

of the classical BCM plasticity rule. Then, we derive an effective equation for the mean-field stochastic dynamics of the synaptic currents for that extended plasticity model. Next, we translate this effective equation into probabilistic Fokker-Planck formalism, and derive an effective potential for the mean-field synaptic current. With the help of the effective potential we find entropy production and Fisher information associated with the synaptic plasticity stochastic dynamics. Entropy production is related to the energy cost of the extended BCM plasticity, while the Fisher information is related to the accuracy of encoded information in a population of plastic synapses about the presynaptic input. Details of the calculations are provided in the Methods (and some in Supporting Information S1).

2. Results

Model of synaptic plasticity: stochastic BCM type.

We consider a sensory neuron with N plastic excitatory synapses (dendritic spines). We assume that synaptic weights w_i ($i = 1, ..., N$), corresponding to spine electric conductances, change due to two factors: correlated activity in presynaptic and postsynaptic firing rates (f_i and r, respectively), and noise in spine conductance ($\sim \sigma_w$). The noise is caused by two basic factors: an internal thermodynamic fluctuations in spines because of their small size (< 1 μm) and relatively small number of molecular components (Kasai et al 2003; Statman et al 2014), and by presynaptic fluctuations in the firing rates that drive the ionic and molecular fluxes in spines. The dynamics of synaptic weights is given by a modified BCM plasticity rule (Bienenstock et al 1982):

$$\frac{dw_i}{dt} = \lambda f_i r(r-\theta) - \frac{(w_i - \epsilon a)}{\tau_w} + \frac{\sqrt{2}\sigma_w(1+\tau_f f_i)}{\sqrt{\tau_w}}\eta_i \tag{1}$$

$$\tau_\theta \frac{d\theta}{dt} = -\theta + \alpha r^2, \tag{2}$$

where λ is the amplitude of synaptic plasticity controlling the rate of change of synaptic conductance, τ_w is the weights time constant controlling their decay duration, θ is the homeostatic variable the so-called sliding threshold (adaptation for plasticity) related to an interplay of LTP and LTD with time constant τ_θ, and α is the coupling intensity of θ to the postsynaptic firing rate r. The noise term in Eq. (1) is represented as Gaussian white noise η_i with zero mean and Delta function correlations, i.e., $\langle \eta_i(t) \rangle_\eta = 0$ and

$\langle \eta_i(t)\eta_j(t')\rangle_\eta = \delta_{ij}\delta(t-t')$ (van Kampen 2007). The amplitude of the noise in weights is proportional to the standard deviation σ_w (in units of conductance) due to basic thermodynamic fluctuations in spines, and to the factor $(1+\tau_f f_i)$ due to additional fluctuations in the presynaptic activities. The latter factor simply amplifies the basic thermodynamic fluctuations. The time scale τ_f of fluctuations in f_i was added in the noise term to maintain a unitless form of the amplifying factor. Finally, the product ϵa is the minimal synaptic weight when there is no presynaptic stimulation ($f_i = 0$), where the unitless parameter $\epsilon \ll 1$. There are two modifications to the conventional BCM rule: the stochastic term $\sim \sigma_w$, and the decay term of synaptic weights with the time constant τ_w, which is key for reproducing a binary nature of synapses (Petersen et al 1998; O'Connor et al 2005) and for determining energy used by synaptic plasticity.

The conventional BCM rule (i.e. for $\tau_w \mapsto \infty$ and $\sigma_w = 0$) describes temporal changes in synaptic weights due to correlated activity of pre- and post-synaptic neurons (both f_i and r are present on the right in Eq. (1)). These activity changes can either increase the weight, if postsynaptic firing r is greater than the sliding threshold θ (this corresponds to LTP), or they can decrease the weight if $r < \theta$ (corresponding to LTD). The interesting aspect is that θ is also time dependent, and it responds quickly to changes in the postsynaptic firing. In effect, when both dynamical processes in Eqs. (1-2) are taken into account, the synapse is potentiated for low r (LTP) and depressed for high r (LTD).

We assume, in accordance with empirical data, that presynaptic firing rates f_i change on a much faster time scale ($\tau_f \sim 0.1 - 1$ sec) than the synaptic weights w_i (changes

on time scale $\tau_w \sim 1$ hr). We further assume that each presynaptic firing rate f_i fluctuates stochastically around mean value f_o with a standard deviation σ_f, and that these fluctuations are uncorrelated. This implies that there is a time scale separation between neural activities and synaptic plasticity activities.

Numerical solution of the stochastic extended BCM plasticity model.
In this section we solve numerically the model represented by $N+1$ Eqs. (1-2).

We first consider the model without synaptic noise, i.e., $\sigma_w = 0$. This deterministic system can exhibit collective bistability, regardless of whether σ_f is 0 or finite. The critical factor in generating bistability is that the time constant τ_θ for the homeostatic variable θ is much smaller than synaptic plasticity time constant τ_w. That is, the variable θ must be much faster than the synaptic weights w_i. Typically, bistability is found for $\tau_\theta/\tau_w \leq 0.06$, and we work in this regime throughout the whole study (for the neurobiological validity of this regime, see the Discussion section). Collective bistability means that all synaptic weights can converge to two different fixed points depending on the initial conditions (Fig. 1). When all synapses start from sufficiently small weights, then they all converge into the same small synaptic weight ϵa. If the initial weights are much larger than ϵa, then all synapses become asymptotically strong. Thus, there is a strong collective behavior of synapses in the deterministic case.

Two main parameters that control the shape of the bifurcation diagram are the synaptic plasticity amplitude λ and the mean firing rate f_o (Fig. 2). For a fixed f_o, synapses can be either in monostable or bistable phase, depending on the value of λ (Fig. 2A). Generally, for small and sufficiently large λ there is monostability, while for

intermediate λ there is bistability. For a fixed λ, the picture is slightly more complex: bistability can emerge already for $f_o = 0$ (for intermediate λ), or for some finite f_o (for small λ), or bistability can never appear (for large λ) (Fig. 2A). The bifurcation diagram, i.e., the dependence of asymptotic value of w_i on f_o is presented in Fig. 2B.

Inclusion of synaptic noise, i.e. $\sigma_w > 0$, leads to stochastic fluctuations of individual synapses. In the monostable regime, fluctuations are around a given fixed point (either weak or strong weight). In the bistable regime, individual synapses fluctuate between weak and strong weights (Fig. 3A). Despite synaptic noise, the collective nature of synapses is statistically preserved, as all synapses have similar weight distributions (Fig. 3B, C). These distributions are much more spread in the bistable regime than in the monostable, and they seem to be almost uniform for bistability.

Dimensional reduction of the stochastic BCM model: dynamic mean field. The stochastic system of $N + 1$ equations described by Eqs. (1-2) is not tractable analytically, because it is the coupled nonlinear system. The coupling takes place via postsynaptic firing rate r, which depends on all synaptic weights w_i (in Eqs. 1-2). In this section an effective mean-field model corresponding to the extended BCM model in Eqs. (1-2) is presented and discussed that is amenable to analytical considerations. In this dynamical mean-field, we focus on a single dynamic variable, which is a population averaged synaptic current v defined by Eq. (25) in the Methods. The single variable v is sufficient to describe the global stochastic dynamics of the original model given by Eqs. (1-2), because together with postsynaptic firing r it forms a closed mathematical system of just two equations; see below. The practical reason behind introducing the

dynamic mean-field is that this approach enables us to obtain explicit formulae for synaptic plasticity energy rate and coding accuracy.

We can reduce the multidimensional system (1-2) into a single effective equation, primary because of the time scale separation between neural firing dynamics (changes typically on the order of seconds or less) and between synaptic plasticity (changes on the timescale of minutes/hours). Moreover, we assume that the two synaptic plasticity processes, described by Eqs. 1 and 2, have two distinct time scales, and τ_w dominates over τ_θ in duration. For the neurophysiological validity of this assumption, see the Discussion. Consequently, for times of the order of τ_w, we have $d\theta/dt \approx 0$, which implies that $\theta \approx \alpha r^2$. The details of the reduction procedure can be found in the Methods, in which we obtain a single plasticity equation for a population averaged excitatory postsynaptic current v per synapse, which is related to w_i and f_i by $v = (\beta/N)\sum_i f_i w_i$, where β depends on neurophysiological parameters and is defined in Eq. (25). The result of the reduction procedure is

$$\frac{dv}{dt} = hr^2(1-\alpha r) - (v - \epsilon c f_o)/\tau_w + \frac{\sqrt{2}\sigma_v}{\sqrt{\tau_w}}\overline{\eta}. \tag{3}$$

This equation essentially couples slow synaptic activities with fast neural activities, and gives a single equation describing the mean-field dynamics of the coupled system: synapses plus their postsynaptic neuron. In Eq. (3), the symbol h is the driving-plasticity parameter given by

$$h = \lambda\beta(f_o^2 + \sigma_f^2), \tag{4}$$

with f_o and σ_f denoting the mean and standard deviation of presynaptic firing rates. Mathematically, the driving-plasticity h is proportional to the product of plasticity amplitude λ and the presynaptic driving $(f_o^2 + \sigma_f^2)$, which implies that h grows quickly with the presynaptic firing rate. Physically, h is proportional to the electric charge that, on average, can enter the spine due to a correlated activity of pre- and post-synaptic neurons (h has a unit of electric charge). This means that the magnitude of h is a major determinant of the plasticity (driving force counteracting the synaptic decay), since larger h can experimentally correspond to more Ca^{+2} entering the spine and a higher chance of invoking a change in synaptic strength, which agrees qualitatively with the experimental data (Huganir and Nicoll 2013; Lisman et al 2012).

The rest of the parameters in Eq. (3) are $c = a\beta$, and $\overline{\eta} = (\sum_i \eta_i)/\sqrt{N}$, which denotes a new (population averaged) Gaussian noise with zero mean and delta function correlations. This population noise has the amplitude σ_v, which corresponds to a standard deviation of v when $h = 0$, and it is given by

$$\sigma_v = \frac{\beta\sigma_w}{\sqrt{N}}[f_o + \tau_f \left(f_o^2 + \sigma_f^2\right)]. \tag{5}$$

Note that σ_v scales as $1/\sqrt{N}$, and it is a product of the intrinsic synaptic conductance noise and of the presynaptic neural activity. The latter implies that a higher presynaptic activity amplifies the current noise.

In Eq. (3), the postsynaptic firing rate r assumes its quasi-stationary value (due to

time scale separation), and is related to v through (for details see the Methods):

$$r = \frac{1}{2}\left(-A^2\kappa + \sqrt{A^4\kappa^2 + 4A^2 v}\right), \tag{6}$$

where A is the postsynaptic firing rate amplitude, and κ is the intensity of firing rate r adaptation. Broadly speaking, the magnitude of κ reflects the strength of neuronal self-inhibition due to adaptation to synaptic stimulation (see Eqs. 21 and 22 in the Methods). Generally, increasing κ leads to decreasing postsynaptic firing rate r (Fig. 4A). For $\kappa = 0$, we recover a nonlinear firing rate curve (square root dependence on synaptic current v) that is characteristic for class one neurons (Ermentrout 1998; Ermentrout and Terman 2010), while for sufficiently large κ, i.e. for $\kappa \gg 2\sqrt{v}/A$, we obtain a linear firing rate curve $r(v) \approx v/\kappa$ (Fig. 4A). Equations (3) and (6) form a closed system for determining the stochastic dynamics of the postsynaptic current v.

Geometric steady state solution of the deterministic mean-field: emergence of bistability in v.

We can use geometric considerations to gain some intuitive understanding of the mean-field deterministic behavior represented by Eqs. (3) and (6). If we put $dv/dt = 0$ and $\sigma_v = 0$ in Eq. (3), we can rearrange it to obtain

$$v = g(v), \tag{7}$$

where the right hand side of this equation, $g(v) = \epsilon c f_o + \tau_w h r^2(1 - \alpha r)$, and it depends on v only through r as in Eq. (6) (see Fig. 4A). Moreover, the function $g(v)$ has a maximum with height proportional to h. When h is very small, Eq. (7) has only one solution $v \sim O(\epsilon)$ (i.e. one intersection point of the curves representing the functions on the right and on the left; Fig. 4B). This solution corresponds to weak synapses and monostable regime. Increasing h, by increasing f_0, causes an increase in the maximal value of the right hand side in Eq. (7), such that more solutions are possible (Fig. 4B). In particular, when h grows above a certain critical value h_{cr}, Eq. (7) generates 3 solutions (one $\sim O(\epsilon)$ and two other $\sim O(1)$), of which the middle one is unstable (Fig. 4B). This case corresponds to bistable regime with two stable solutions, representing weak and strong synaptic currents that can be called, respectively, "down" and "up" synaptic states. These two states could hypothetically be related to thin and mushroom dendritic spines, with small and large number of AMPA receptors, respectively (Bourne and Harris 2007). For very large driving-plasticity h the two lower solutions disappear and we have again a monostable regime with strong synapses only (Fig. 4B).

A geometrical condition for the emergence of bistability is when the function $g(v)$ in Eq. (7) first touches tangentially the line $y = v$, i.e. when $dg/dv = 1$ (Fig. 4B). Solving this condition together with Eq. (6) yields for $\epsilon \ll 1$ the critical value of the driving-plasticity parameter h_{cr} as

$$h_{cr} = \frac{\alpha\kappa}{\tau_w}\left(1 + \sqrt{1 + (\alpha\kappa A^2)^{-1}}\right)^2 + O(\epsilon). \tag{8}$$

Note that for very fast decay in Eq. (3), i.e. for $\tau_w \mapsto 0$, the bistability is lost, since then $h_{cr} \mapsto \infty$, and there is only one solution corresponding to weak synapses $v \sim O(\epsilon)$. Bistability is also lost in the opposite limit of extremely slow decay, $\tau_w \mapsto \infty$, but in this case the only stable solution corresponds to strong synapses. Interestingly, for very strong neural adaptation, $\kappa \mapsto \infty$, the bistability also disappears, since then $h_{cr} \mapsto \infty$. This case corresponds to extremely small postsynaptic firing rates, $r \approx v/\kappa \approx 0$ (Fig. 4A), and indicates the absence of a driving force capable of pushing synapses to a higher conducting state. On the other hand, when there is no adaptation, $\kappa \mapsto 0$, the critical $h_{cr} \mapsto (\tau_w A^2)^{-1}$, i.e. it is finite. This means that it is easier to produce synaptic bistability for neurons with stronger nonlinearity in their firing rate curves (see Eq. 6; Fig. 4A).

Analytic steady state solution of the deterministic mean-field.

The above geometric intuition can be supported by an analytic approach. In the deterministic limit, $\sigma_v = 0$ (which is obtained either for $N \mapsto \infty$ or for $\sigma_w = 0$), we can solve the mean-field model of Eq. (3) in the stationary state, i.e. we can find its fixed points by setting $dv/dt = 0$. To achieve this, it is more convenient to work with the postsynaptic firing rate r than with v variable, due to a nonlinear dependence of r on v, which is given by Eq. (6). Inverting Eq. (6), we find $v = \kappa r + (r/A)^2$, which can be used in the condition $dv/dt = 0$. This generates an equation for roots of the cubic polynomial in the r variable:

$$\alpha\tau_w h r^3 - (\tau_w h - A^{-2})r^2 + \kappa r - \epsilon c f_o = 0. \tag{9}$$

The discriminant Δ of this equation is

$$\Delta = \kappa^2(\tau_w h - A^{-2})^2 - 4\alpha\tau_w\kappa^3 h - 2\epsilon c f_o(\tau_w h - A^{-2})[2(\tau_w h - A^{-2})^2 - 9\alpha\tau_w\kappa h] \\ -27(\epsilon c f_o \alpha\tau_w h)^2. \tag{10}$$

The sign of Δ determines how many real roots Eq. (9) has. Specifically, if $\Delta < 0$, then Eq. (9) has one real root, whereas if $\Delta > 0$ then Eq. (9) has three distinct real roots. The former case corresponds to monostability, while the latter to bistability in the mean-field deterministic dynamics of Eq. (3). The transition between this two regimes takes place for $\Delta = 0$. The existence of these regimes obviously depends on the values of various parameters in the discriminant Δ.

The phase diagram of mono- vs. bi-stability in the parameter space of λ, f_o (plasticity amplitude vs. mean presynaptic firing) is shown in Fig. 5. A specific bifurcation diagram, computed numerically from Eq. (3) for $\sigma_v = 0$, in which a stationary value of v is plotted as a function of f_o is presented in Fig. 6. The phase and bifurcation diagrams for stationary v in Figs. 5 and 6 look qualitatively similar to the phase and bifurcation diagrams of stationary w_i in Fig. 2 for the full N-synaptic system of Eqs (1-2).

Equation (9) can be solved analytically for small ϵ, as a series expansion in ϵ. The detailed procedure is described in the Suppl. Information S1. Depending on the sign of $\Delta_o \equiv \lim_{\epsilon \mapsto 0} \Delta/\kappa^2$, there can be one or three fixed points, which have the following values

$$\begin{aligned} v_d &= c f_o \epsilon + \left(\frac{c f_o}{\kappa}\right)^2 h\tau\epsilon^2 + O(\epsilon^3) \\ v_u &= \kappa r_+ + (r_+/A)^2 + O(\epsilon) \\ v_{max} &= \kappa r_- + (r_-/A)^2 + O(\epsilon), \end{aligned} \tag{11}$$

where

$$r_\pm = \frac{h\tau_w - A^{-2} \pm \sqrt{\Delta_o}}{2\alpha\tau_w h}. \tag{12}$$

The value v_d is the fixed point for weak synapses (down state), while v_u is the fixed point for strong synapses (up state). The intermediate value v_{max} corresponds to an always unstable fixed point, which serves as a boundary between the domains of attraction for down and up fixed points. Thus all initial values of v in the $(0, v_{max})$ interval converge asymptotically into v_d, and all initial values of v in the (v_{max}, ∞) interval converge asymptotically into v_u. From Eqs. (11) and (12) it can be seen that the value of the intermediate point v_{max} decreases as f_o (or h) increases, from the value v_u (at the onset of bistability for $\Delta_o = 0$) to the value v_d. This mean that the domain of attraction for

the v_u fixed point increases at the expense of the domain of attraction for the v_d fixed point, which shrinks with increasing f_o in the bistable regime.

The critical value of the driving-plasticity parameter h_{cr} for the emergence of bistability in Eq. (9) can be also obtained directly from the discriminant Δ in the limit $\epsilon \mapsto 0$. We obtain h_{cr} by setting $\Delta_0 = 0$, and solving this equation for h.

Stochastic mean-field: numerics and effective potential for synaptic current. When the synaptic noise is present, $\sigma_v > 0$, the synaptic current v fluctuates. The distribution of v is unimodal for small mean firing rate f_o, and bimodal for sufficiently large f_o (Fig. 7). The bimodal distribution reflects the bistability found for the deterministic case, and it corresponds to synapses changing weights from weak to strong.

Stationary average synaptic current $\langle v \rangle$, which is a measure of synaptic weights, increases weakly with presynaptic firing rate mean f_o and its standard deviation σ_f (Figs. 8 and 9). Mean-field values of $\langle v \rangle$ (computed from Eq. (41)) start to deviate from the exact numerical values (computed from Eqs. 1-2) for larger levels of synaptic noise σ_w and for higher f_o (Figs. 8 and 9).

Stochastic Eq. (3) for the dynamics of v can be mapped into an equation for the dynamics of the probability distribution of v conditioned on f_o, i.e. $P(v|f_o)$, described by a Fokker-Planck equation (see Eqs. (31-34) in the Methods). In the stochastic stationary state, characterized by the stationary probability distribution $P_s(v|f_o)$, we can define a new and important quantity called an effective potential $\Phi(v|f_o)$, which is a function of the synaptic current v. The effective potential Φ is proportional to the amount of energy associated with the synaptic plasticity described by Eq. (3), and it is

equal to the integral of the right hand side of Eq. (3) with $\sigma_v = 0$ (Van Kampen 2007; see Eq. (36) in the Methods). The explicit form of the effective potential Φ is

$$\Phi(v|f_o) = \frac{v}{\tau_w}\left(\frac{1}{2}v - \epsilon c f_o\right) - h\left[\kappa r^3\left(\frac{1}{3} - \frac{\alpha r}{4}\right) + \frac{r^4}{A^2}\left(\frac{1}{2} - \frac{2\alpha r}{5}\right)\right]. \tag{13}$$

Note that the second term in Φ (with the large bracket) is proportional to the plasticity amplitude λ through h. This term depends on v through the firing rate r (see Eq. 6). In general, the functional form of the potential $\Phi(v|f_o)$ determines the thermodynamics of synaptic memory, and thus it is an important function.

The shape of the potential $\Phi(v|f_o)$ depends on the relative magnitude of the driving-plasticity h and the inverse of the decay time constant $1/\tau_w$ (Fig. 10A). In fact, there are two competing terms in Φ that are controlled by $1/\tau_w$ and h. The first term ($\sim 1/\tau_w$) maintains monostability, while the second ($\sim h$) promotes bistability. For h greater than the critical value h_{cr} (Eq. 8), there is bistability and Φ has two minima at v_u and v_d, corresponding to up (strong) and down (weak) synaptic states (Fig. 10A), similar to the result for the deterministic limit. For very large h, there is again only one minimum related to strong synapses (Fig. 10A). The two minima are separated by a maximum corresponding to a potential barrier at v_{max}. Metastable values of v, i.e. the minima and maximum of the potential, can be found from the condition $d\Phi/dv = 0$, which is equivalent to finding the fixed points of Eq. (3) in the deterministic limit.

If we use a mechanical analogy and treat v as a spatial coordinate, then synaptic plasticity can be visualized as a movement in v space (state transitions), which is

constrained by the energy related to Φ. This means that the shape of the function $\Phi(v)$ determines what kind of motions in v-space (state space) are possible or more likely. In particular, the binary nature of synaptic plasticity given by Eq. (3) can be described as transitions between two wells of the effective potential $\Phi(v|f_0)$, corresponding to weak and strong synapses, or down and up synaptic states (e.g. Billings and van Rossum 2009; Graupner and Brunel 2012). These transitions, caused by intrinsic synaptic noise (σ_w) and fluctuations in the presynaptic input (σ_f), can be thought as a "hill climbing" process in the v space, which requires energy due to a barrier separating the two wells (Fig. 10). The dwelling times in both states (T_u, T_d) can be found from the classic Kramers "escape" formula (Eq. 47; Van Kampen 2007), and they are generally much larger than the time constant τ_w (Fig. 10B).

We define the memory time T_m of the synaptic system as a characteristic time needed to relax synaptic weights to their stationary values following a brief perturbation, or single memory event. Mathematically, it is equivalent to finding a relaxation time of the probability distribution $P(v|f_o)$ to its steady state distribution $P_s(v|f_o)$ after a brief perturbation; see Eq. (50) in the Methods. The characteristic memory time T_m is strictly related to the dwelling times T_u and T_d by Eq. (51), and they mutual relationship is depicted in Fig. 10B. Generally, the memory lifetime T_m is very small in the monostable regime ($T_m \sim \tau_w$), i.e. for small presynaptic firing. However, it jumps by several orders of magnitude when synapses become bistable (i.e. when $h \approx h_{cr}$), but then T_m monotonically decreases with increasing f_o (Fig. 10B).

Energy rate of synaptic plasticity.

In this section we determine the energy rate, or metabolic rate, associated with stochastic BCM type synaptic plasticity. In a nutshell, energy is provided to the synaptic system to drive the plasticity related transitions between different synaptic weights associated mainly with the increase of synaptic weights. In the steady state this energy rate balances the energy dissipated due to the synaptic noise, which tends to decrease synaptic weights.

The plasticity related energy rate is determined both numerically, for the whole system of N synapses described by Eqs. (1-2), and analytically for the mean-field approximation described by a single Eq. (3). The numerical procedure for the whole system is described in the Methods (see section "Numerical simulations of the full synaptic system"), and the analytical results are described below.

The power dissipated by synaptic plasticity $\dot{E}$ in the mean-field approximation is proportional to the average temporal rate of the effective potential decrease, i.e. $-\langle d\Phi(v|f_0)/dt\rangle$, where $\langle ...\rangle$ denotes averaging with respect to the probability distribution $P(v|f_o)$. Since the potential $\Phi(v|f_0)$ depends on time only through v, after rearranging we get $\dot{E} \sim -\langle (dv/dt)(d\Phi/dv)\rangle$. Thermodynamically, this formula is equivalent to the entropy production rate associated with the stochastic process described by Eq. (3), and represented by the effective potential $\Phi(v|f_o)$ (Nicolis and Prigogine 1977; Lan et al 2012, Mehta and Schwab 2012; Tome 2006; Tome and de Oliveira 2010). The synaptic plasticity energy rate per synapse $\dot{E}$ can be found analytically using $1/N$ expansion, and in the steady state takes the form (see Methods):

$$\dot{E} = p_d \dot{E}_d + p_u \dot{E}_u \tag{14}$$

where $\dot{E}_d$ and $\dot{E}_u$ are the energy rates dissipated, respectively, in the down and up synaptic states, which have the occupancies p_d and p_u. The energy rates $\dot{E}_d$ and $\dot{E}_u$ are given by

$$\dot{E}_i = \frac{E_o D}{4(\Phi_i^{(2)})^2}[3(\Phi_i^{(3)})^2 + 2\Phi_i^{(2)}\Phi_i^{(4)}] + O(D^2), \tag{15}$$

where $i = d$ (down state) or $i = u$ (up state). The symbols Φ_i and $\Phi_i^{(n)}$ denote values of the potential $\Phi(v)$ and its n-th derivative with respect to v for $v = v_i$. The symbol E_o is the characteristic energy scale for variability in synaptic (spine) conductance, and it provides a link with underlying molecular processes (see the Methods). For convenience, we defined a new noise related parameter D, which is

$$D = \sigma_v^2/\tau_w. \tag{16}$$

D can be viewed as the effective noise amplitude, and D relates to the number of synapses N as $D \sim 1/N$.

Note that in Eq. (15) the terms of the order $O(1)$ disappear, and the first nonzero contribution to $\dot{E}$ is of the order $O(1/N)$, since $D \sim 1/N$. Moreover, to have nonzero power in this order, the potential $\Phi(v)$ must contain at least a cubic nonlinearity.

Eqs. (14) and (15) indicate that energy is needed for plasticity processes associated with the potential Φ "hill climbing", which is in analogy to the energy needed for a particle trapped in a potential well (of a certain shape) to escape. The energetics of such a "motion" in the v-space depends on the shape of the potential, which is mathematically accounted for by various higher-order derivatives of Φ. Thus, a fraction of synapses that were initially in the down state can move up the potential gradient to the up state by overcoming a potential barrier, but this requires the energy that is proportional to σ_v^2 and to the derivatives of the potential. By analogy, a similar picture holds for synapses that were initially in the up state. The prefactor D in Eq. (15) indicates that the transitions up$\leftrightarrow$down, as well as local fluctuations near these states, cost energy that is proportional to the intrinsic synaptic noise ($\sim \sigma_w$) and presynaptic activity (including its fluctuations f_o and σ_f). The important point is that if there is no intrinsic spine noise ($\sigma_w = 0$), then there are no transitions between the up and down states in the steady state, and consequently there is no energy dissipation ($\sigma_v = 0$), regardless of the fast presynaptic input magnitude. Likewise for very long decay of synaptic weights, $\tau_w \mapsto \infty$, corresponding to very slow synaptic plasticity (and the lack of the decay term in Eq. 1), there is no energy used. In such a noiseless stationary state, the plasticity processes described by Eq. (3) are energetically costless, since there are no net forces that can change synaptic weight, or mathematically speaking, that

can push synapses in the v-space. (This is not true under non-stationary conditions when there is some temporal variability in one or more parameters in Eq. 3, leading to dissipation, but the focus here is on the steady state). This situation resembles the so-called "fluctuation-dissipation" theorem known from statistical physics (Nicolis and Prigogine 1977; Van Kampen 2007; Risken 1996), where thermal fluctuations always cause energy loss. In our case, this fluctuation-dissipation relationship underlines a key role of thermodynamic fluctuations for the metabolic load of synaptic plasticity.

We can compare the energy rate coming from the mean-field (Eq. 3) to the energy rate computed numerically for the full synaptic system described by Eqs. (1-2). The results are presented in Figs. 11 and 12. Generally, a better agreement between mean-field and exact results is achieved for intermediate synaptic noise σ_w and also for intermediate values of mean presynaptic firing rate f_o. For larger σ_w, in the regions close to mono-bistability transitions, there are peaks in the mean-field $\dot{E}$ that are absent in the numerical $\dot{E}$ (Fig. 11). These peaks are the artifacts of the approximation methods used in the mean-field. Moreover, Fig. 11 shows that the energy rate $\dot{E}$ mostly increases steadily with f_o (the exact result). The exception is a narrow interval near the mono- to bi-stability regions, where $\dot{E}$ slightly decreases (Fig. 11). The energy rate also steady increases with the standard deviation in the presynaptic firing σ_f (Fig. 12).

A next interesting question is how the plasticity energy rate depends on the synaptic weights? In Fig. 13 we plot the dependence of $\dot{E}$ on the average synaptic current $\langle v \rangle$, which is proportional to the synaptic weights and spine size (Kasai et al 2003). It is clear that the synaptic energy rate related to plasticity grows nonlinearly with $\langle v \rangle$. For

small $\langle v \rangle$, the energy rate $\dot{E}$ depends weakly on $\langle v \rangle$, whereas for large $\langle v \rangle$ it increases strongly with $\langle v \rangle$ (Fig. 13).

Which dependence of $\dot{E}$ is stronger: on f_o or on $\langle v \rangle$? In Fig. 14 it is shown that the energy rate $\dot{E}$ increases nonlinearly both with f_o and with $\langle v \rangle$, but the dependence on the average synaptic current $\langle v \rangle$ is much steeper.

Energy cost of plastic synapses as a fraction of neuronal electric energy cost: comparison to experimental data.

In order to assess the magnitude of the synaptic plasticity energy rate, we compare it to the rate of energy consumption by a typical cortical neuron for its electric activities related to fast synaptic transmission, action potentials and maintenance of the resting potential (Attwell and Laughlin 2001). The neural spiking activity and synaptic transmission are known to consume the majority of the neural energy budget (Attwell and Laughlin 2001; Harris et al 2012; Karbowski 2012). The ratio of the total energy rate used by plastic synapses $N\dot{E}$ to the neuron's energy rate $\dot{E}_n$ (given by Eq. (63) in the Methods) is computed for different presynaptic firing rates f_o, various levels of synaptic noise σ_w, and for different cortical regions. The results for macaque and human cerebral cortex are shown in Figs. 15 and 16. These plots indicate that the synaptic plasticity contribution depends strongly on the level of synaptic noise σ_w; the higher the noise the larger the ratio $N\dot{E}/\dot{E}_n$. Higher firing rates f_o also tend to increase that ratio but not that strongly, and the dependence is nonmonotonic (Figs. 15 and 16). Generally, the value of $N\dot{E}/\dot{E}_n$ ranges from negligible ($\sim 10^{-4} - 10^{-3}$) for small/intermediate noise ($\sigma_w = 0.1$ nS), to substantial ($\sim 10^{-2} - 10^{0}$) for very large noise ($\sigma_w = 2$ nS). The

results are qualitatively similar across different cortical regions within one species, as well as between human and macaque cortex, despite large differences in the cortical sizes of both species (Figs 15 and 16). Small quantitative differences are the result of small differences in the synaptic densities between areas and species.

Information encoded by plastic synapses.

In our model, information or memory about the mean input f_o is written in the population of synapses, represented by the synaptic current v. In the stochastic steady state, the synaptic current v is characterized by probability distribution $P_s(v|f_0)$, which is related to the potential $\Phi(v|f_0)$. This means that information encoded in synapses also depends on the structure of the potential (Eq. 13).

The accuracy of the encoded information can be characterized by Fisher information I_F (Cover and Thomas 2006). In general, larger I_F implies a higher coding precision. Fisher information, related to synaptic current v, can be derived analytically (see the Methods). In the limit of small effective noise amplitude D we obtain:

$$I_F(f_o) = p_u p_d \left(\left(\frac{\Phi_u - \Phi_d}{D} \right)' + \frac{1}{2} \left[\frac{(\Phi_u^{(2)})'}{\Phi_u^{(2)}} - \frac{(\Phi_d^{(2)})'}{\Phi_d^{(2)}} \right] \right)^2 + \sum_{i=d,u} p_i \left(\frac{\Phi_i^{(2)}}{D} (v_i')^2 + \frac{1}{2} \left[\frac{(\Phi_i^{(2)})'}{\Phi_i^{(2)}} - \frac{D'}{D} \right]^2 \right) \tag{17}$$

where p_i denote the fractions of synapses in the up ($i = u$) and down ($i = d$) states, and the prime denotes a derivative with respect to f_o. Note that the effective noise

amplitude D depends on f_o, since σ_v depends on f_o and $D \sim \sigma_v^2$ (see Eq. (5)).

The first term in Eq. (17) proportional to $p_d p_u$ is of the order of $\sim 1/D^2 \sim N^2$, and it appears only in the bistable regime (both p_d and p_u must be nonzero). This term depends on the difference in the potentials between up and down states. The second term in Eq. (17) proportional to the weighted sums of p_d and p_u is of the order of $\sim 1/D \sim N$, and it is always present regardless of mono- or bistability. Thus, the first term is much bigger than the second for small D, which is the primary reason why I_F (and coding accuracy) is several orders of magnitude larger when synapses are bistable (see below). Because Fisher information $I_F(f_o)$ is either proportional to $1/D^2 \sim N^2$ (in the bistable regime) or to $1/D \sim N$ (in the monostable regime), it implies that many synapses are much better in coding the presynaptic firing f_o than a single synapse.

Equation for Fisher information (Eq. 17) indicates that there is no simple relationship between I_F and the synaptic current v. Rather, I_F depends in a nonlinear manner on the derivatives of synaptic currents in up and down states v'_u and v'_d. This follows form the fact that the potential Φ depends in a complicated way on v (see Fig. 10, and Eq. (13)).

Accuracy and lifetime of synaptically stored information vs plasticity energy rate.

How the long-term energy used by synapses relates to the accuracy and persistence of stored information? The above results indicate that $\dot{E}$ and I_F depend inversely on the synaptic noise σ_v (or D), suggesting that its lowering should be beneficial since gain in information is accompanied by a decrease in synaptic energy rate.

A more complicated picture emerges if other parameters are varied, notably driving presynaptic input f_o, at different regimes of mono- and bistability (Fig. 17). At the onset of bistability, Fisher information I_F and memory lifetime T_m both increase dramatically, whereas the plasticity energy rate $\dot{E}$ increases mildly. Approximate mean-field calculations of $\dot{E}$ provide a small peak at the transition point, but more exact numerical calculations of $\dot{E}$ based on Eqs. (1-2) indicate a smooth behavior (with a slight decrease), which suggest that the small peak in the mean-field is an artifact of the approximation (Fig. 17). Taken together, this implies that a high improvement in information coding accuracy and its retention, in the initial region of bistability, do not involve a huge amounts of energy. On the contrary, the corresponding energy cost is rather small.

For higher f_o, deeper in the bistability region, there is a different trend. In this coexistence region, $\dot{E}$ increases monotonically, while I_F and T_m decrease, which in turn indicates an inefficiency of information storing. However, even here the huge values in I_F and T_m overcome the growth in $\dot{E}$. For even higher f_o, in the monostable phase with strong synapses only, $\dot{E}$ still increases monotonically, whereas I_F and T_m further decrease to the levels similar for very small f_o. Consequently, the biggest gains in synaptic information precision and lifetime per energy used ($I_F/\dot{E}$ and $T_m/\dot{E}$) are achieved for the bistable phase only (Fig. 18). Interestingly, the gains in the information precision and lifetime depend nonmonotonically on the plasticity amplitude λ, and there are some optimal values of λ that are different for the gains $I_F/\dot{E}$ and $T_m/\dot{E}$ (Fig. 19).

Taken together, these results suggest that storing of accurate information in synapses

can be relatively cheap in the bistable regime, and thus metabolically efficient.

Precision of coding memory is restricted by sensitivity of synaptic plasticity energy rate on the driving input.

The above results suggest that synaptic energy utilization does not limit directly the coding precision of a stimulus, because there is no simple relationship between Fisher information and power dissipated by synapses. However, a careful inspection of the curves in Fig. 17 suggests that there might be a link between I_F and the derivative of $\dot{E}$ with respect to the driving input f_o. In fact, it can be shown that in the most interesting regime of synaptic bistability, in the limit of very weak effective noise $D \mapsto 0$, we have either (see the Methods)

$$I_F(f_o) = \frac{(\partial p_d^{(0)}/\partial f_o)^2}{p_u^{(0)} p_d^{(0)}} \left[1 + O(D)\right], \tag{18}$$

or equivalently

$$I_F(f_o) = \frac{(\partial \dot{E}/\partial f_o)^2}{p_u^{(0)} p_d^{(0)} (\dot{E}_u - \dot{E}_d)^2} \left[1 + O(D)\right], \tag{19}$$

where $p_d^{(0)}, p_u^{(0)}$ are the fractions of synapses in the down and up states (weak and strong synapses) in the limit $D \mapsto 0$. It is important to stress that simple formulas (18) and (19) have a general character, since they do not depend explicitly on the potential Φ, and thus they are independent of the plasticity type model. Equation (18) shows

that synaptic coding precision increases greatly for sharp transitions from mono- to bistability, since then $(\partial p_u^{(0)}/\partial f_o)^2$ is large. Additionally, Eq. (19) makes an explicit connection between precision of synaptic information and nonequilibrium dissipation. Specifically, the latter formula implies that to attain a high fidelity of stored information, the energy used by synapses $\dot{E}$ does not have to be large, but instead it must change sufficiently quickly in response to changes in the presynaptic input.

We can also estimate a relative error e_f in synaptic coding of the average presynaptic firing f_o. This error is related to Fisher information by a Cramer-Rao inequality $e_f \geq (f_o\sqrt{I_F})^{-1}$ (Cover and Thomas 2006). Using Eq. (19), in our case this relation implies

$$e_f \geq \frac{\sqrt{p_u^{(0)} p_d^{(0)}}}{f_o|\dot{E}'/(\dot{E}_u - \dot{E}_d)|}, \tag{20}$$

where the prime denotes the derivative with respect to f_o. The value of the product $p_u p_d$ is in the range from 0 to 1/4. In the worst case scenario for coding precision, i.e. for $p_u^{(0)} p_d^{(0)} = 1/4$, this implies that a 10% coding error ($e_f = 0.1$), corresponds to the relative sensitivity of the plasticity energy rate on presynaptic firing $f_0|\dot{E}'/(\dot{E}_u - \dot{E}_d| = 5$. Generally, the larger the latter value, the higher precision of synaptic coding. In our particular case, this high level of synaptic coding fidelity is achieved right after the appearance of bistability (Fig. 17).

3. Discussion

Summary of the main results.

In this study, the energy cost of long-term synaptic plasticity was determined and compared to the accuracy and lifetime of an information stored at excitatory synapses. The main results of this study are:

(a) Formulation of the dynamic mean-field of the extended BCM synaptic plasticity model (Eqs. 3-5).

(b) Energy rate of plastic synapses increases nonlinearly both with the presynaptic firing rate (Figs. 11 and 14) and with average synaptic current or weights (Figs. 13 and 14).

(c) Coding of more accurate information in synapses need not require a large energy cost (cheap long-term information). The accuracy of stored information about presynaptic input can increase by several orders of magnitude with only a mild increase in the plasticity energy rate at the onset of bistability (Figs. 17 and 18).

(d) The accuracy of information stored at synapses and its lifetime are not limited by the available energy rate, but by the sensitivity of the energy rate on the presynaptic firing. For very weak synaptic noise the coding accuracy at plastic synapses (Fisher information) is proportional to the square of derivative of the plasticity energy rate with respect to mean presynaptic firing (Eq. 19).

(e) Energy rate of synaptic plasticity, which is of chemical origin, constitutes in most cases only a tiny fraction of neuron's energy rate associated with fast synaptic transmission and action potentials, which are of electric origin. That fraction can be substantial

only for very large synaptic noise and presynaptic firing rates (Figs. 15 and 16).

Discussion of the main results.

The dynamic mean-field for synaptic plasticity was derived analytically by applying (i) the timescale separation between neural and synaptic plasticity activities, and (ii) dimensional reduction of the original synaptic system. The formulated mean-field of the synaptic current v seems to work reasonably well for average $\langle v \rangle$ if intrinsic synaptic noise σ_w is small and presynaptic firing rates f_o are not too high (Fig. 8). For larger σ_w and f_o, the mean-field value $\langle v \rangle$ diverges form the exact numerical average calculated from Eqs. (1-2).

The mean-field approximation to the synaptic energy rate $\dot{E}$ was additionally derived in the limit of small effective noise D (either large N or small σ_w, or both). Surprisingly, the mean-field approximation for $\dot{E}$ works better for intermediate noise σ_w than for its smaller values (Fig. 11). For those intermediate values of σ_w, the energy rate $\dot{E}$ calculated in the mean-field is close to that calculated numerically in the whole neurophysiological interval of f_o variability (Fig. 11, middle panel). It seems that the primary reason for the breakdown of the mean-field for $\langle v \rangle$ and $\dot{E}$ is the way the integrals in Eqs. (43) and (57) were approximated. In those integrals, it was assumed that x_{d1} (the lower limit of integration) tends to $-\infty$ for $D \mapsto 0$, which however is not always true, since $x_{d1} \sim v_d \sim \epsilon$, and v_d can be very small for very small value of ϵ (see Eq. 11), especially if f_o is small. As a consequence, the real value of x_{d1} can be in the range $-(0.1 - 1)$, even for very small D.

Comparing the mean-field $\dot{E}$ to its numerical values suggests that the peaks in the

mean-field approximation of $\dot{E}$ are artifacts (Fig. 11). They are the result of some (small) differences in the exact location of the transitions points mono/bistability between mean-field and the numerics. This causes certain errors in the relative magnitudes of p_d and p_u, which leads to over- or under-estimates in the mean-field values of $\dot{E}$ (Eq. 14).

Nonlinear increase of plasticity energy rate with presynaptic firing rate f_o and average synaptic current v (Figs. 11, 13, and 14) suggests that high presynaptic activities and large synapses/spines are metabolically costly (there exists a positive correlation between synaptic current and size; see Kasai et al (2003)). Consequently, it seems that large firing rates and synaptic weights (proportional to average synaptic current) should not be preferred in real neural circuits. This simple conclusion is qualitatively in line with experimental data for cortical neurons, showing low mean firing rates and weak mean synaptic weights, with skewed distributions (Buzsaki and Mizuseki 2014).

The most striking result of this study is that precise memory storing about presynaptic firing rate f_o does not have to be metabolically expensive. Strictly speaking, the information encoded at synapses, i.e., its accuracy and lifetime, do not have to correlate positively with the energy used by synapses (Fig. 17). Such a correlation is only present at the onset of synaptic bistability, where a large increase in information precision (I_F) and lifetime (T_m) is accompanied only by a mild increase in the energy rate. This suggests an energetic efficiency of stored information in the bistable synaptic regime, i.e., relatively high information gain per energy used (Fig. 18). Moreover, the results in Fig. 18 show that there exists an optimal value of the presynaptic firing rate for which

the information gain per energy, as well as memory lifetime per energy, are maximal. An additional support for the metabolic efficiency of synaptic information comes from the fact that energy used $\dot{E}$ and coding precision I_F depend the opposite way on the effective noise amplitude D (compare Eqs. (15) and (17)), and thus, I_F increases, while $\dot{E}$ decreases with decreasing D. Because $D \sim 1/N$, this also implies that $I_F \sim N^2$ in the bistable regime, i.e., that more synapses (large N) are much better at precise coding of mean presynaptic firing than a single synapse ($N = 1$). Taken together, these findings are compatible with a study by Still et al (2012) showing that abstract stochastic systems with memory, operating far from thermodynamic equilibrium, can be the most predictive about an environment if they use minimal energy.

Estimating an external variable is never perfect, and it is shown here that synaptic coding accuracy (I_F) relates to the derivative of the energy rate with respect to an average input. The fundamental relationship linking memory precision and synaptic metabolic sensitivity is present in Eq. (19), which is valid regardless of the specific plasticity mechanism, as long as synapses can exist in two metastable states, in the limit of very small synaptic noise D. This binary synaptic nature is a key feature enabling a high fidelity of long-term synaptic information (Petersen et al 1998), despite ongoing neural activity, which is generally detrimental to information storing (Fusi et al 2005). Specifically, for realistic neurophysiological parameters, it is seen from Fig. (17) that the relative coding error in synapses $e_f \sim (f_o\sqrt{I_F})^{-1}$ can be as small as $0.03 - 0.1$ (or $3 - 10\%$) near the onset of bistability. However, away from that point the error gets larger. Thus, again it seems that there exist an optimal firing rate f_o for which coding

accuracy is maximal and quite high, despite large fluctuations in presynaptic neural activities (large σ_f in relation to f_o).

Neural computation is thought to be metabolically expensive (Aiello and Wheeler 1995; Laughlin et al 1998; Attwell and Laughlin 2001; Karbowski 2007 and 2009; Niven and Laughlin 2008; Harris et al 2012), and it must be supported by cerebral blood flow and constrained by underlying microvasculature and neuroanatomy (Karbowski 2014 and 2015). It is shown here that an important aspect of this computation, namely long-term synaptic plasticity involved in learning and memory, constitutes in most cases only a small fraction of that neuronal energy cost associated mostly with fast synaptic transmission and spiking (Figs. 15 and 16). Specifically, for intermediate/large synaptic noise ($\sigma_w = 0.1$ and 0.5 nS), metabolic cost of synaptic plasticity can maximally be on a level of $1 - 10\%$ of the electric neuronal cost, both for human and macaque monkey (Figs. 15 and 16). Higher levels of synaptic plasticity cost (maximally 100% of electric cost) are possible, but only for very large synaptic noise, $\sigma_w = 2.0$ nS (Figs. 15 and 16). The latter value is, however, unlikely because it is 20 times larger than the mean values of synaptic weights w_i (see Fig. 2B), and thus, it seems that higher costs of synaptic plasticity are physiologically implausible. Taken together, these results suggest that a precise memory storing can be relatively cheap, which agrees with empirical estimates presented in (Karbowski 2019).

Discussion of other aspects of the plasticity model.

In this study, an extended BCM model of synaptic plasticity is introduced and solved. There are 3 additional elements in our model (Eq. 1) that are absent in the classical

BCM plasticity rule: weight decay term ($\sim 1/\tau_w$), synaptic noise ($\sim \sigma_w/\sqrt{\tau_w}$), and the nonlinear dependence of the postsynaptic firing rate on synaptic input (Eq. 6). Moreover, it is assumed here that presynaptic firing rates fluctuate stochastically and fast around a common mean f_o with standard deviation σ_f. These features make the behavior of our model significantly different from the behavior of the classical BCM model (Bienenstock et al 1982). In particular, due to the stochasticity of synaptic weights, our model does not exhibit an input selectivity, in contrast to the classical BCM rule. Input selectivity in the classical BCM means that the largest static presynaptic firing rate "selects" its corresponding synapse by increasing its weight, in such a way that the weights of all other synapses decay to zero. In our model this never happens, because all synapses are driven on average by the same input, and more importantly, synaptic noise constantly brings all synapses up and down in an unpredictable fashion. For these reasons, the mean-field approach proposed here, although mathematically correct, does not make sense for the classical BCM rule (no weights decay, no noise) if our goal is studying input selectivity, because in that model only one synapse is effectively present at the steady-state, and there is no need for large N approach.

The main reasons for choosing the mean-field approach, and constructing a single dynamical equation for the population averaged synaptic current v are: (i) we wanted to treat analytically the multidimensional stochastic model given by Eqs. (1-2), and (ii) the variable v emerges as a natural choice, since r in Eqs. (1-2) depends only on one variable, precisely on v (see Eq. 6). The feature (ii) makes Eqs. (3) and (6) a closed mathematical system of just two equations that can be handled analytically.

Another practical reason behind introducing the dynamic mean-field is that it enables us to obtain explicit formulae for synaptic plasticity energy rate and Fisher information (coding accuracy).

In deriving the dynamic mean-field we assumed that the time constant related to w_i dynamics, i.e. τ_w, is much larger than the time constant related to the sliding threshold θ, which is τ_θ. This is in agreement with empirical observations and estimations, since τ_w must be of the order of 1 hr to be consistent with slice experiments, showing wiping out synaptic potentiation after about 1 hr when presynaptic firing becomes zero (Frey and Morris 1997; Zenke et al 2013). (Note that τ_w refers to the decay of synaptic weights to the baseline value ϵa, and it should not be confused with a characteristic time of plasticity induction, which is controlled by the product $\lambda f_i r$ in Eq. (1) and which can be much faster, $\sim$ minutes (Petersen et al 1998; O'Connor et al 2005).) On the other hand, the time constant τ_θ must be smaller than about 3 min for stability reasons (Zenke and Gerstner 2017; Zenke et al 2013), and it even has been estimated to be as small as $\sim$ 12 sec (Jedlicka et al 2015).

Although individual synapses in the original model Eqs. (1-2) exhibit bistability (see Figs. 1 and 2), this bistability has a collective character. That is, if most synaptic weights are initially weak, then they all converge into a lower fixed point. On the other hand, if a sufficient fraction of synaptic weights is initially strong, then they all converge into an upper fixed point (Fig. 1). This means that the majority of synapses participate in a coordinated switching between up and down states, due to effective noise (internal and external). This mechanism is probably different from the

mechanism found in Petersen et al (1998) and O'Connor et al (2005), where bistability was reported on a level of a single synapse, independent of other synapses. (However, from these papers it is difficult to judge how long the potentiation lasts in the absence of presynaptic stimulation). Our scenario for bistability is conceptually closer to the model of synaptic bistability proposed by Zenke et al 2015 (Zenke et al 2015), which also emerges on a population level. Interestingly, both models, the one presented here and the one in (Zenke et al 2015), exhibit the so-called anti-Hebbian plasticity, in the sense that LTP (i.e. $\dot{v} > 0$) appears for low firing rates, instead of LTD as for classical BCM rule. However, in the present model the initial LTP window is very narrow, and appears for very small postsynaptic firing rates $r < (cf_0/\kappa)\epsilon \sim O(\epsilon)$. This feature is necessary for stable bistability, and does not contradict experimental results on BCM rule verification (Kirkwood et al 1996), showing LTD for low firing rates. The reason is that these experiments were performed for firing rates above 0.1 Hz, leaving uncertainty about LTP vs. LTD for very low activity levels (or very long times).

The cooperativity in synaptic bistable plasticity found here is to some extent similar to the data showing that neighboring dendritic spines interact and tend to cluster as either strong or weak synapses (Govindarajan et al 2006; Govindarajanet al 2011). These clusters can be as long as single dendritic segments, which is called "clustered plasticity hypothesis" (Govindarajan et al 2006; Govindarajan et al 2011). However, the difference is that in the present model there are no dendritic segments, and spatial dependence is averaged over, which leads effectively to one synaptic "cluster" either with up or down states.

Metabolic cost of synaptic plasticity in the mean-field: intuitive picture. The formula for the plasticity energy rate (Eqs. 15) contains various derivatives of the effective potential Φ, which encodes the plasticity rules for synaptic weights. In this scenario, the synaptic plasticity corresponds to a driven stochastic motion of the population averaged postsynaptic current v in the space constrained by the potential Φ, in analogy to a ball moving on a rugged landscape with a ball coordinate corresponding to v. Because our potential can exhibit two minima separated by a potential barrier, the plasticity considered here can be viewed as a stochastic process of "hill climbing", or transitions between the two minima (the idea of "synaptic potential" was used also in van Rossum et al (2000), Billings and van Rossum (2009), Graupner and Brunel (2012)).

The energy rate of plastic synapses $\dot{E}$ (or power dissipated by plasticity) is the energy used for climbing the potential shape in v-space, and it is proportional to the average temporal rate of decrease in the potential, $-\langle d\Phi/dt\rangle$, due to variability in v. In terms of thermodynamics, the plasticity energy rate $\dot{E}$ is equivalent to the entropy production rate, because synapses like all biological systems operate out of thermodynamic equilibrium with their environment and act as dissipative structures (Nicolis and Prigogine 1977). Dissipation requires a permanent influx of energy from the outside (provided by blood flow, see e.g. Karbowski (2014)) to maintain synaptic structure, which in our case is the distribution of synaptic weight. A physical reason for the energy dissipation in synapses in the steady state is the presence of noise (both internal synaptic $\sim \sigma_w$, and external presynaptic $\sim \sigma_f$), causing fluctuations, that tend to wipe

out the pattern of synaptic weights. Thermodynamically speaking, this means reducing the synaptic order and thus increasing synaptic entropy. To preserve the order, this increased entropy has to be "pumped out", in the form of heat, by investing some energy in the process, which relates to ATP consumption.

Thermodynamics of memory storing and bistability.

The general lack of high energetic demands for sustaining accurate synaptic memory may seem non-intuitive, given an intimate relation between energy and information known from classical physics (Leff and Rex 1990). For example, transmitting 1 bit of information through synapses is rather expensive and costs 10^4 ATP molecules (Laughlin et al 1998), and a comparative number of glucose molecules (Karbowski 2012), which energetically is much higher ($\sim 10^5 kT$) than a thermodynamic minimum set by the Landauer limit ($\sim 1kT$) (Landauer 1961). Additionally, there are classic and recent theoretical results that show dissipation-error tradeoff for biomolecular processes, i.e., that higher coding accuracy needs more energy (Lan et al 2012; Mehta and Schwab 2012; Barato and Seifert 2015; Bennett 1979; Lang et al 2014). How can we understand our result in that light?

First, there is a difference between transmitting information and storing it, primarily in their time scales, and faster processes generally need more power (see also below). Second, it is known from thermodynamics that erasing an information can be more energy costly than storing information (Landauer 1961; Bennett 1982), since the former process is irreversible and is always associated with energy dissipation, and the latter can in principle be performed very slowly (i.e. in equilibrium with the environment)

without any heat released. In our system, the information is maximal for intermediate presynaptic input generating metastability with two synaptic states (Fig. 2). If we decrease the input below a certain critical value, or increase it above a certain high level, our system becomes monostable, which implies that it does not store much information (entropy is close to zero). Thus, the transition from bistability to monostability is equivalent to erasing the information stored in synapses, which according to the Landauer principle (Landauer 1961; Berut et al 2012) should cost energy.

Third, the papers showing energy-error tradeoff in biomolecular systems (Lan et al 2012; Mehta and Schwab 2012; Lang et al 2014; Bennett 1979; Barato and Seifert 2015) use fairly linear (or weakly nonlinear) models, while in our model the plasticity dynamics is highly nonlinear (see Eqs. 1, 3, and 6). Additionally, we consider the prediction of an external variable (average input f_o), in contrast to some of the biomolecular models (Bennett 1979; Barato and Seifert 2015), which dealt with estimating errors in an internal variable.

Cost of synaptic plasticity in relation to other neural costs.

The energy cost of synaptic plasticity is a new and an additional contribution to the overall neural energy budget considered before and associated with fast signaling (action potentials, synaptic transmission, maintenance of negative resting potential), and slow nonsignaling factors (Attwell and Laughlin 2001; Engl and Attwell 2015). The important distinction between slow synaptic plasticity dynamics and fast signaling is that the former is of chemical origin (protein/receptor interactions), while the latter is of electric origin (ionic movement against gradients of potential and concentration).

Consequently these two, although coupled but to a large extent, separate phenomena have different characteristic time and energy scales, which results in rather small energy cost of synaptic plasticity in relation to the fast electric cost (Figs. 15 and 16).

The earlier studies of the neuronal energy cost (Attwell and Laughlin 2001; Engl and Attwell 2015) provided important order of magnitude estimates based on ATP turnover rates, but they had mainly a phenomenological character and cannot be directly applied to nonlinear dynamics underlying synaptic plasticity. Contrary, the current approach and the complementary approach taken in (Karbowski 2019) are based on "first principles" taken from non-equilibrium statistical physics and in combination with neural modeling can serve as a basis for future more sophisticated calculations of energy used in excitatory synapses, possibly with inclusion of some molecular detail (e.g. Lisman et al 2012; Miller et al 2005; Kandel et al 2014).

The calculations performed here indicate that the energy dissipated by synaptic plasticity increases nonlinearly with presynaptic firing rate (Fig. 11). The dependence on presynaptic firing is consistent with a strong dependence of CaMKII autophosphorylation level on Ca^{2+} influx frequency to a dendritic spine (De Koninck and Schulman 1998), which should translate to a similar dependence of ATP consumption rate related to protein activation on presynaptic firing. Moreover, these results raise the possibility of observing or measuring the energetics of synaptic plasticity for high firing rates. It is hard to propose a specific imaging technique for detecting enhanced synaptic plasticity, but nevertheless, it seems that techniques relying on spectroscopy, e.g., near-infrared spectroscopy with its high spatial and temporal resolution, could be of help.

Regardless of whether the energetics of synaptic plasticity is observable or not, it could have some functional implications. For example, it was reported that small regional decreases in glucose metabolic rate associated with age, and presumably with synaptic decline, lead to significant cognitive impairment associated with learning (Gage et al 1984).

A relatively small cost of plasticity in relation to neuronal cost of fast electric signaling (Figs. 15 and 16) is in some part due to relatively slow dynamics of spine conductance decay, quantified by $\tau_w \sim 1$ hr (Frey and Morris 1997; Zenke et al 2013), since $\dot{E} \sim 1/\tau_w$ in Eqs. 15 and 16. The time scale τ_w characterizes the duration of early LTP on a single synapse level. On a synaptic population level, characterized by synaptic current v, the duration of early LTP is given by T_m (memory maintenance of a brief synaptic event), which can be of the order of several hours.

Late phases of LTP and LTD, during which memory is consolidated, are much slower than τ_w and they are governed by longer timescales of the orders of days/weeks (Ziegler et al 2015; Redondo and Morris 2011). Consequently, one can expect that such plasticity processes, as well as equally slow homeostatic synaptic scaling (Turrigiano and Nelson 2004), should be energetically inexpensive. Nevertheless, there are experimental studies related to long-term memory cost in fruitfly that claim that memory in general is metabolically costly (Mery and Kawecki 2005; Placais and Preat 2013; Placais et al 2017). However, the problem with those papers is that they do not measure directly the energy cost related to plasticity in synapses, but instead they estimate the global fly metabolism, which indeed affects long-term memory (Mery and Kawecki 2005; Placais

and Preat 2013). In a recent paper by Placais et al (2017) it was found that upregulated energy metabolism in dopaminergic neurons is correlated with long-term memory formation. However, again, no measurement was made directly in synapses, and thus it is difficult to say how much of this enhanced neural metabolism can be attributed to plasticity processes and how much to enhanced neural and synaptic electric signaling (spiking and transmission). It is important to stress that the energy cost of protein synthesis, process believed to be associated with long-term memory consolidation (Kandel et al 2014), was estimated to be very small, on a level of $\sim 0.03 - 0.1\%$ of the metabolic cost of fast synaptic electric signaling related to synaptic transmission (Karbowski 2019; see also below for an alternative estimate). Consequently, it is possible that memory induction, maintenance, and consolidation involve a significant increase in neural activity and hence metabolism, but it seems that the majority of this energy enhancement goes for upregulating neural electric activity, not for chemical changes in plastic synapses.

The energetics of very slow processes associated with memory consolidation were not included in the budget of the energy scale E_o (present in Eq. 15, and estimated in the Methods), since we were concerned only with the early phases of LTP and LTD, which are believed to be described by BCM model (both standard and extended). Nevertheless, for the sake of completeness, we can estimate the energy cost of the late LTP and LTD, as well as energy requirement of mechanical changing of spine volume (also not included in the budget of E_o).

Protein synthesis, which is associated with l-LTP and l-LTD, underlines synaptic consolidation and scaling (Kandel et al 2014). There are roughly 10^4 proteins in PSD

including their copies (Sheng and Hoogenraad 2007), on average each with $\sim 400-500$ amino acids, which are bound by peptide bonds. These bonds require 4 ATP molecules to form (Engl and Attwell 2015), which is $4 \cdot 20kT$ of energy (Phillips et al 2012). This means that chemical energy associated with PSD proteins is about $(3.2-4.0) \cdot 10^8 kT$, i.e. $(1.6-2.0) \cdot 10^7$ ATP molecules, or equivalently $(1.4-1.75) \cdot 10^{-12}$ J. Given that an average lifetime of PSD proteins is 3.7 days (Cohen et al 2013), we obtain the energy rate of protein turnover as $\sim (4.6-5.8) \cdot 10^{-18}$ W, or $52-65$ ATP/s per spine. For human cerebral cortex with a volume of 680 cm^3 (Hofman 1988) and average density of synapses $3 \cdot 10^{11}$ cm^{-3} (Huttenlocher and Dabholkar 1997), we have $2 \cdot 10^{14}$ synapses. This means that the global energy cost of protein turnover in spines of the human cortex is $(9.2-11.5) \cdot 10^{-4}$ W, or equivalently $(1-1.3) \cdot 10^{16}$ ATP/s, which is extremely small ($\sim 0.01\%$) as human cortex uses about 5.7 Watts of energy (Karbowski 2009).

The changes in spine volume are related directly to the underlying dynamics of actin cytoskeleton (Honkura et al 2008; Cingolani and Goda 2008). We can estimate the energy cost of spine size using a mechanistic argument. Dendritic spine grows due to pressure exerted on the dendrite membrane by actin molecules. The reported membrane tension is in the range $(10^{-4}-1)$ kT/nm^2 (Phillips et al 2009) with the upper bound being likely an overestimate, given that it is close to the so-called rapture tension ($1-2$ kT/nm^2), when the membrane breaks (Phillips et al 2009). A more reasonable value of the membrane tension seems to be 0.02 kT/nm^2, as it was measured directly (Stachowiak et al 2013). Taking this value, we get that to create a typical 1 μm^2 of stable spine requires $2 \cdot 10^4 kT$ or 10^3 ATP molecules. Since the actin turnover rate

in spine is 1/40 $\sec^{-1}$ (Honkura et al 2008), which is also the rate of spine volume dynamics, we obtain that the cost of maintaining spine size is 25 ATP/s. This value is comparable but two-fold smaller than the ATP rate used for PSD protein turnover per spine ($52 - 65$ ATP/s) given above.

How do the costs of protein turnover and spine mechanical stability relate to the energy cost of e-LTP and e-LTD calculated in this paper using the extended BCM model? From Fig. 11, we get that the latter type of synaptic plasticity uses energy in the range $(10^{-3} - 10^{0})E_o$ (solid lines for exact numerical results) per second per spine, depending mainly on firing rate and synaptic noise. Since the energy scale $E_o = 2.3 \cdot 10^4$ ATP (see the Methods), we obtain that the energy cost of the plasticity related to e-LTP and e-LTD is $23 - 23000$ ATP/s, i.e., its upper range can be 400 times larger than the contributions from protein turnover and spine volume changes. This result strongly suggests that the calculations of the energetics of synaptic plasticity based on the extended BCM model provide a large portion of the total energy required for the induction and maintenance of synaptic plasticity.

4. METHODS

Neuron model.

We consider a sensory neuron with a nonlinear firing rate curve (so called class one, valid for most biophysical models) and with activity adaptation given by (Ermentrout 1998; Ermentrout and Terman 2010)

$$\tau_r \frac{dr}{dt} = -r + \bar{A}\sqrt{I_{syn} - s} \tag{21}$$

$$\tau_a \frac{ds}{dt} = -s + \bar{\kappa} r \tag{22}$$

where r is the instantaneous neuron firing rate with mean amplitude $\bar{A}$, s is the adaptation current (or equivalently self-inhibition) with the intensity $\bar{\kappa}$, τ_r and τ_a are the time constants for variability in neural firing and adaptation, and I_{syn} is the total excitatory synaptic current to the neuron provided by N excitatory synapses, i.e., $I_{syn} \sim \sum_i f_i w_i$. If $I_{syn} < s$ in Eq. (21), then this equation simplifies and becomes $\tau_r dr/dt = -r$. In order to ensure a saturation of the firing rate r for very large number of synapses N, and for s to be relevant in this limit, $\bar{A}$ and $\bar{\kappa}$ must scale as $\bar{A} = A/\sqrt{N}$ and $\bar{\kappa} = N\kappa$. In a mature brain N can fluctuate due to structural plasticity, but we assume in agreement with the data (Sherwood et al 2020; DeFelipe et al 2002) that there is some well defined average value of N.

We assume that the neuron is driven by stochastic presynaptic firing rates f_i ($i = 1, ..., N$) that change on a much faster time scale τ_f than the synaptic weights w_i.

Additionally, we assume that the fast variability in presynaptic firing rates is stationary in a stochastic sense, i.e., the probability distribution of f_i does not change in time. Consequently, for each time step t, in the stationary stochastic state we can write

$$f_i(t) = f_o + \sigma_f x_i(t) \tag{23}$$

where f_o is the mean firing rate of all presynaptic neurons and σ_f denotes the standard deviation in the variability of f_i. The variable x_i is the Gaussian random variable, which reflects noise in the presynaptic neuronal activity. For the noise x_i we have the following averages (Van Kampen 2007): $\langle x_i \rangle_x = 0$ and $\langle x_i x_j \rangle_x = \delta_{ij}$, where the last equality means that different x_i are independent, which also implies that fluctuations in different firing rates f_i are statistically independent. Eq. (23) allows negative values of f_i, which is not realistic. However, in analytical calculations this is not a problem, because we use only average of f_i and its standard deviation σ_f. In numerical simulations, we prevent the negative values of f_i by setting $f_i = 0$, whenever f_i becomes negative.

Given Eq. (23), one can easily verify the following average

$$\langle f_i(t)^2 \rangle_x = f_o^2 + \sigma_f^2, \tag{24}$$

Equation (24) indicates that presynaptic firing rates fluctuate around average value f_o with standard deviation σ_f. The important point is that these fluctuations are fast, on

the order of τ_f ($\sim 0.1 - 1$ sec), which is much faster than the timescale τ_w. Eq. (24) is also used below.

Definition of synaptic current per spine v.

The synaptic current I_{syn} has two additive components related to AMPA and NMDA receptors, $I_{syn} = I_{ampa} + I_{nmda}$, with the receptor currents

$I_{ampa} = qq_{ampa}|V_r|\tau_{ampa}g_{ampa}\sum_{i=1}^{N} f_i M_i^{ampa}$,

and

$I_{nmda} = qq_{nmda}|V_r|\tau_{nmda}g_{nmda}\sum_{i=1}^{N} f_i M_i^{nmda}$,

where q is the probability of neurotransmitter release, V_r is resting membrane potential of the neuron (we used the fact that the reversal potential for AMPA/NMDA is close to 0 mV; Ermentrout and Terman (2010)), g_{ampa} and g_{nmda} are single channel conductances of AMPA and NMDA receptors, q_{ampa} and q_{nmda} are probabilities of their opening with characteristic times τ_{ampa} and τ_{nmda}. The symbols M_i^{ampa} and M_i^{nmda} denote AMPA and NMDA receptor numbers for spine i. Data indicate that during synaptic plasticity the most profound changes are in the number of AMPA receptors M^{ampa} and opening probability of NMDA q_{nmda} (Kasai et al 2003; Huganir and Nicoll 2013; Matsuzaki et al 2004). We define the excitatory synaptic weight w_i as a weighted average of AMPA and NMDA conductances, i.e.,

$w_i = (\tau_{nmda}q_{nmda}M_i^{nmda}g_{nmda} + \tau_{ampa}q_{ampa}M_i^{ampa}g_{ampa})/(\tau_{nmda} + \tau_{ampa})$.

This enables us to write the synaptic current per spine, i.e. $v = I_{syn}/N$ (which is more convenient to use than I_{syn}), as

$$v = \frac{\beta}{N} \sum_{i=1}^{N} f_i w_i, \tag{25}$$

where $\beta = q|V_r|(\tau_{nmda} + \tau_{ampa})$. The current per spine v is the key dynamical variable in our dimensional reduction procedure and subsequent analysis (see below).

Dependence of the postsynaptic firing rate r on synaptic current v.

The time scales related to neuronal firing rates and firing adaptation τ_f, τ_r and τ_a are much faster than the time scale τ_w associated with synaptic plasticity. Therefore, for long times of the order of τ_w, firing rate r and postsynaptic current adaptation s are in quasi-stationary state, i.e., $dr/dt \approx ds/dt \approx 0$. This implies a set of coupled algebraic equations:

$$\begin{aligned} r &= A\sqrt{v - s/N} \\ s &= N\kappa r, \end{aligned} \tag{26}$$

which yields a quadratic equation for r, i.e., $r^2 + A^2\kappa r - A^2 v = 0$. The solution for r, which depends on v, is given by

$$r = \frac{1}{2}\left(-A^2\kappa + \sqrt{A^4\kappa^2 + 4A^2 v}\right), \tag{27}$$

Note that r depends nonlinearly on the synaptic current v. Additionally, s/N is always smaller than v in the steady state, which means that r in Eq. (26) is well defined.

Dimensional reduction of the extended BCM model: Dynamic mean-field model.

We focus on the population averaged synaptic current v (Eq. 25). Since v is proportional to weights w_i, and because r depends directly on v, it is possible to obtain a closed form dynamic equation for plasticity of v. Thus, instead of dealing with N dimensional dynamics of synaptic weights, we can study a one dimensional dynamics of the population average current v. This dimensional reduction is analogous to observing the motion of a center of mass of many particle system, which is easier than simultaneous observation of the motions of all particles. Such an approach is feasible for an analytical treatment where one can directly apply the methods of stochastic dynamical systems and thermodynamics (Van Kampen 2007).

The time derivative of v, given by Eq. (25), is denoted with dot and reads

$\dot{v} = (\beta/N) \sum_{i=1} (\dot{f}_i w_i + f_i \dot{w}_i) \approx (\beta/N) \sum_{i=1} f_i \dot{w}_i,$

where we used the fact that fluctuations in f_i are much faster than changes in weights w_i, and hence f_i are in stochastic quasi-stationary states. Now, using Eq. (1) for $\dot{w}_i$ and quasi-stationarity of θ, we obtain the following equation for $\dot{v}$:

$$\dot{v} = \frac{\lambda\beta}{N} r^2 (1 - \alpha r) \sum_{i=1}^{N} f_i^2 - \frac{1}{\tau_w} \left(v - \frac{\epsilon c}{N} \sum_{i=1}^{N} f_i \right) + \frac{\sqrt{2}\beta\sigma_w}{N\sqrt{\tau_w}} \sum_{i=1}^{N} (f_i + \tau_f f_i^2)\eta_i, \tag{28}$$

where $c = a\beta$.

The next step is to perform averaging over fast fluctuations in presynaptic rate f_i. We need to find the following three averages with respect to the random variable x_i: $\langle \sum_{i=1}^{N} f_i \rangle_x$, $\langle \sum_{i=1}^{N} f_i^2 \rangle_x$, and $\langle \sum_{i=1}^{N} f_i^2 \eta_i \rangle_x$.

From Eq. (23) it follows that $\langle f_i \rangle_x = f_o$, and thus the first average is

$$\langle \sum_{i=1}^{N} f_i \rangle_x = N f_o. \tag{29}$$

The second average follows from Eq. (24), and we have

$$\langle \sum_{i=1}^{N} f_i^2 \rangle_x = N(f_o^2 + \sigma_f^2). \tag{30}$$

The third average can be decomposed as $\langle \sum_{i=1}^{N} f_i^2 \eta_i \rangle_x = \sum_{i=1}^{N} \langle f_i^2 \rangle_x \, \eta_i =$
$= \sum_{i=1}^{N} (f_o^2 + \sigma_f^2)\eta_i$,
where we used the fact that the noise η is independent of the noise x, and again Eq. (24).

The final step is to insert the above averages into the equation for $\dot{v}$ (Eq. 28). As a result we obtain Eq. (3) in the main text, which is a starting point for determining energetics of synaptic plasticity and information characteristics.

Distribution of synaptic currents in the stochastic mean-field model: weak and strong synapses.

Stochastic Eq. (3) for the population averaged synaptic current v can be written in short notation as

$$\frac{dv}{dt} = F(v) + \sqrt{2D}\overline{\eta}, \tag{31}$$

where the function $F(v)$ is defined as

$$F(v) = hr^2(1 - \alpha r) - (v - \epsilon c f_o)/\tau_w, \tag{32}$$

and D is the effective noise amplitude (it includes also fluctuations in the presynaptic input) given by

$$D = \sigma_v^2/\tau_w. \tag{33}$$

Eq. (31) corresponds to the following Fokker-Planck equation for the probability distribution of the synaptic current $P(v|f_o;t)$ conditioned on f_o (Van Kampen 2007):

$$\begin{aligned}\frac{\partial P(v|f_o;t)}{\partial t} &= -\frac{\partial}{\partial v}\left(F(v)P(v|f_o;t)\right) + D\frac{\partial^2 P(v|f_o;t)}{\partial v^2} \\ &= -\frac{\partial J(v)}{\partial v}\end{aligned} \tag{34}$$

The function $J(v)$ in the last equality in Eq. (34) is the probability current, which is $J(v) = F(v) - D\partial P(v)/\partial v$.

The stationary solution of the Fokker-Planck equation (Eq. 34) is obtained for a constant probability current J (Gardiner 2004; Van Kampen 2007). For monostable systems, which have a unique steady state (fixed point) one usually sets $J(v) = 0$, which corresponds to a detailed balance (Gardiner 2004; Tome 2006). Such a unique steady state corresponds to thermal equilibrium with the environment (Tome 2006) and the solution is of the form (Van Kampen 2007)

$$P_s(v|f_o) \sim \exp\left(-\Phi(v|f_o)/D\right), \tag{35}$$

where $\Phi(v|f_o)$ is the effective potential for synaptic current v, and it is obtained by integration of $F(v)$ in Eq. (31), i.e.,

$$\Phi(v|f_o) = -\int_0^v dx\ F(x). \tag{36}$$

The potential can have either one (monostability) or two (bistability) minima, depending on f_o and other parameters. The explicit form of $\Phi(v|f_o)$ is shown in Eq. (13).

For bistable systems, for which there are two possible steady states (fixed points), the situation is more complicated. In the presence of nonzero, $D > 0$, effective synaptic noise, there can be noise induced jumps between the two fixed points. In this case the probability current J in the steady state must be a nonzero constant, because of the exchange of probabilities between the two fixed points, or equivalently, because of the stochastic jumps of v between two potential wells (Gardiner 2004). For small noise D, such jumps between the two potential wells happen on very long time scales. This long-time dynamics is the primary reason that the stationary state of such "driven" systems (by thermal and presynaptic fluctuations) is globally out of thermal equilibrium with the environment and it is called thermodynamic nonequilibrium steady state, in which the detailed balance is broken (Tome 2006). However, locally, close to each fixed point and for not too long times the system is in local thermal equilibrium. Thus, we can locally approximate the probability distribution of v by the form given in Eq. (35), by expanding the potential $\Phi_s(v)$ around v_d and v_u. Using a Gaussian approximation, which should be valid for small D (either for large N or for small σ_w or both), we can write $P_s(v)$ for v close to v_d as $P_s(v) \sim e^{-\Phi_d/D} \exp\left(-\frac{\Phi_d^{(2)}}{2D}(v - v_d)^2\right)$, and for v close to v_u we have $P_s(v) \sim e^{-\Phi_u/D} \exp\left(-\frac{\Phi_u^{(2)}}{2D}(v - v_u)^2\right)$, where $\Phi_i^{(2)}$ is the second derivative of the potential with respect to v at v_i, where the subscript i is either d (down state) or u (up state). For the sake of computations we have to extend these local approximations to

longer intervals of v, corresponding to the domains of attraction for two fixed points v_d and v_u. Consequently, we assume that the first approximation works for $0 \leq v \leq v_{max}$, and the second for $v > v_{max}$. In sum, we approximate the stationary probability density $P_s(v|f_o)$ as two Gaussian peaks centered at v_d and v_u:

$$P_s(v|f_o) = \begin{cases} P_d(v) = Z^{-1} e^{-\Phi_d/D} \exp\left(-\frac{\Phi_d^{(2)}}{2D}(v - v_d)^2\right); & \text{for} \quad 0 \leq v \leq v_{max} \\ P_u(v) = Z^{-1} e^{-\Phi_u/D} \exp\left(-\frac{\Phi_u^{(2)}}{2D}(v - v_u)^2\right); & \text{for} \quad v > v_{max} \end{cases} \tag{37}$$

where Z is the normalization factor, which can be written as a sum $Z = Z_d + Z_u$, with

$$\begin{aligned} Z_d &= e^{-\Phi_d/D} \int_0^{v_{max}} dv \ \exp\left(-\frac{\Phi_d^{(2)}}{2D}(v - v_d)^2\right) \\ &= e^{-\Phi_d/D} \sqrt{\frac{\pi D}{2\Phi_d^{(2)}}} [\text{erf}(|x_{d1}|) + \text{erf}(|x_{d2}|)] \end{aligned} \tag{38}$$

and

$$\begin{aligned} Z_u &= e^{-\Phi_u/D} \int_{v_{max}}^{\infty} dv \ \exp\left(-\frac{\Phi_u^{(2)}}{2D}(v - v_u)^2\right) \\ &= e^{-\Phi_u/D} \sqrt{\frac{\pi D}{2\Phi_u^{(2)}}} [1 + \text{erf}(|x_u|)] \end{aligned} \tag{39}$$

where erf(...) is the error function, $x_{d1} = \sqrt{\frac{\Phi_d^{(2)}}{2D}}(-v_d)$, $x_{d2} = \sqrt{\frac{\Phi_d^{(2)}}{2D}}(v_{max} - v_d)$, and $x_u = \sqrt{\frac{\Phi_u^{(2)}}{2D}}(v_{max} - v_u)$. Note that because the unstable fixed point v_{max} depends on

f_o, the arguments of the error functions in Z_d and Z_u change with changes in f_o. This influences the determination of p_d, p_u, as well as energy rate and Fisher information (see below).

Fractions of weak and strong synapses.

We define the fraction of synapses in the down state p_d (fraction of weak synapses) as the probability that synaptic current v is in the domain of attraction of the down fixed point in the deterministic limit. This takes place for v in the range $0 \leq v \leq v_{max}$, where v_{max} is the unstable fixed point separating the two stable fixed points v_d and v_u. By analogy, the fraction of synapses in the up state p_u is the probability that v is greater than v_{max}. We can write this mathematically as

$$p_d = \int_0^{v_{max}} dv \; P_d(v|f_o) \equiv Z_d/Z \tag{40}$$

and

$$p_u = \int_{v_{max}}^{\infty} dv \; P_u(v|f_o) \equiv Z_u/Z \tag{41}$$

where $P_d(v|f_o)$ and $P_u(v|f_o)$ are given by Eq. (37). Using the expressions for Z_d and Z_u, we find an explicit form of p_d as

$$p_d = \left(1 + e^{(\Phi_d - \Phi_u)/D} \sqrt{\frac{\Phi_d^{(2)}}{\Phi_u^{(2)}}} \frac{[1 + \mathrm{erf}(|x_u|)]}{[\mathrm{erf}(|x_{d1}|) + \mathrm{erf}(|x_{d2}|)]}\right)^{-1} \tag{42}$$

Note that p_d and p_u sum to unity, since $Z = Z_d + Z_u$.

Average values of v and r in the mean-field.

Average value of the synaptic current v in the mean-field is denoted as $\langle v \rangle$, and computed as

$$\langle v \rangle = \int_0^{v_{max}} dv\ P_d(v)\ v + \int_{v_{max}}^{\infty} dv\ P_u(v)\ v. \tag{43}$$

Execution of these integrals yields

$$\langle v \rangle = p_d v_d + p_u v_u + O(e^{-x_{d1}^2}, e^{-x_{d2}^2}, e^{-x_u^2}), \tag{44}$$

where $O(e^{-x_{d1}^2}, e^{-x_{d2}^2}, e^{-x_u^2})$ denotes small exponential terms in the limit of very small D.

Standard deviation of v can be found analogically, which yields

$$\sqrt{\langle v^2 \rangle - \langle v \rangle^2} = \sqrt{(v_u - v_d)^2 p_u p_d + D(\frac{p_d}{\Phi_d^{(2)}} + \frac{p_u}{\Phi_u^{(2)}})}. \tag{45}$$

The average value of the postsynaptic firing rate r, denoted as $\langle r \rangle$, is computed in the limit of large A (see Eq. 6). In this limit $r \approx (v/\kappa)\,[1 - v/(A\kappa)^2]$, and we find

$$\begin{aligned}\langle r \rangle \approx\, & p_d \left[\frac{v_d}{\kappa} - \frac{1}{A^2\kappa^3} \left(v_d^2 + D/\Phi_d^{(2)} \right) \right] \\ & + p_u \left[\frac{v_u}{\kappa} - \frac{1}{A^2\kappa^3} \left(v_u^2 + D/\Phi_u^{(2)} \right) \right], \end{aligned} \tag{46}$$

which means that the form of $\langle r \rangle$ is more complicated than $\langle v \rangle$.

Transitions between weak and strong synaptic states: Kramer escape rate. For cortical neurons the number of spines per neuron are very large, i.e. $N \sim 10^3 - 10^4$ (Elston et al 2001; DeFelipe et al 2002; Sherwood et al 2020), and thus one can expect that σ_v is small and consequently the fluctuations around the population average current v are rather weak. The results described below are obtained in the limit of small σ_v.

Plastic synapses can jump between down and up states due to effective synaptic noise σ_v or D. From a physical point of view, this corresponds to a noise induced "escape" of some synapses through a potential barrier. Average dwelling times in the up (T_u) and down (T_d) states can be determined from the Kramers's formula (Van Kampen 2007):

$$T_i = \frac{2\pi}{\sqrt{\Phi_i^{(2)} |\Phi_{max}^{(2)}|}} \exp\left(\frac{\tau_w}{\sigma_v^2} \Delta\Phi_i \right), \tag{47}$$

where the index $i = d$ or $i = u$, $\Phi_i^{(2)}$ and $\Phi_{max}^{(2)}$ are the second derivatives of the potential at its minima ($v = v_i$) and maximum ($v = v_{max}$), and the potential difference $\Delta\Phi_i = \Phi(v_{max}) - \Phi(v_i) > 0$. Note that for large number of synapses N, the exponential factor in Eqs. (47) can be large, which can lead to very long dwelling times that are generally much longer than any time scale in the original Eqs. (1-2). The fact that the times T_u and T_d are long but finite is an indication of metastability of "locally" stable up and down synaptic states.

There exist a relationship between fractions of weak/strong synapses and the Kramer's escape times T_d and T_u in the limit of very weak noise $D \mapsto 0$. Namely, it can be easily verified that in this limit $p_d^{(0)}/p_u^{(0)} = T_d/T_u$, and consequently, we can write

$$p_d^{(0)} = \frac{T_d}{T_d + T_u}, \tag{48}$$

where $p_d^{(0)}$ is the fraction of weak synapses for $D \mapsto 0$ given by

$$p_d^{(0)} = \left(1 + e^{(\Phi_d - \Phi_u)/D} \sqrt{\frac{\Phi_d^{(2)}}{\Phi_u^{(2)}}}\right)^{-1}. \tag{49}$$

Memory lifetime.

Synaptic memory lifetime T_m is defined as a characteristic time the synapses remember a perturbation to their steady state distribution. Mathematically, it means that we have to consider a time-dependent solution of the probability density $P(v|f_0;t)$ to the Fokker-Planck equation given by Eq. (34). This solution can be written as (van Kampen 2007; Risken 1996)

$$P(v|f_o;t) = P_s(v|f_o) + \sum_{k=0}^{\infty} e^{-\gamma_k t}\psi_k(v|f_o), \tag{50}$$

where γ_k and $\psi_k(v|f_o)$ are appropriate eigenvalues and eigenvectors. The eigenvalues are inverses of characteristic time scales, which describe a relaxation process to the steady state. The smallest eigenvalue, denoted as γ_0, determines the longest relaxation time $1/\gamma_0$, and we associate that time with the memory lifetime T_m. It has been shown that $\gamma_0 = 1/T_d + 1/T_u$ (van Kampen 2007; Risken 1996), which implies that

$$T_m = \frac{T_u T_d}{T_u + T_d}. \tag{51}$$

A similar approach, through eigenvalues, to estimating the memory lifetime was adopted also in (Fusi and Abbott 2007).

Entropy production rate, entropy flux, and power dissipated by plasticity. Processes underlying synaptic plasticity (e.g. AMPA receptor trafficking, PSD protein phosphorylation, as well as protein synthesis and degradation; see Huganir and Nicoll (2013), Choquet and Triller (2013)) operate out of thermodynamic equilibrium, and therefore require energy influx. At a stochastic steady state, this energy is dissipated as heat, which roughly corresponds to a metabolic rate of synaptic plasticity. The rate of dissipated energy is proportional to the average rate of decrease in the effective potential Φ, or equivalently to the entropy production rate (Nicolis and Prigogine 1977).

Given the above, we can write the energy rate for synaptic plasticity $\dot{E}$ as $\dot{E} \sim -\langle d\Phi(v|f_0)/dt\rangle = -\langle \Phi^{(1)}\dot{v}\rangle$, where $\Phi^{(1)}$ is the first derivative of Φ with respect to v, the symbol $\dot{v}$ is the temporal derivative of v, and the averaging $\langle ...\rangle$ is performed over the distribution $P(v|f_0)$. The second equality follows from the fact that v is the only variable in the potential that changes with time on the time scale τ_w. Next, we can use Eq. (3) or (31) in the equivalent form, namely $\dot{v} = -\Phi^{(1)} + \sqrt{2/\tau_w}\sigma_v\overline{\eta}$, and this equation resembles the motion of an overdamped particle (with negligible mass) in the potential Φ, with v playing the role of a spatial coordinate. After that step, we can write the energy rate as $\dot{E} \sim \langle[\Phi^{(1)}]^2\rangle - \sqrt{2/\tau_w}\sigma_v\langle\Phi^{(1)}\overline{\eta}\rangle$. The final step is to use the Novikov theorem (Novikov 1965) for the second average, i.e. $\langle\Phi^{(1)}\overline{\eta}\rangle = \frac{1}{2}\sqrt{2/\tau_w}\sigma_v\langle\Phi^{(2)}\rangle$. This leads to

$$\dot{E} \sim \frac{\sigma_v^2}{\tau_w}\left(-\langle\Phi^{(2)}\rangle + (\tau_w/\sigma_v^2)\langle[\Phi^{(1)}]^2\rangle\right).$$

We can obtain a similar result for $\dot{E}$ using a thermodynamic reasoning. The dynamics of synaptic plasticity is characterized by the distribution of synaptic currents per

synapse $P(v|f_o)$, which evolves in time according to Eq. (34). With this distribution we can associate the entropy $S(t)$, defined as $S(t) = -\int_0^\infty dv\ P(v|f_o) \ln P(v|f_o)$, measuring the level of order in a typical spine. It can be shown (Nicolis and Prigogine 1977; Tome 2006; Tome and de Oliveira 2010) that the temporal derivative of the entropy, dS/dt, is composed of two competing terms, $dS/dt = \Pi - \Gamma$, called entropy production rate (Π) and entropy flux (Γ), both per synapse. In the case of thermodynamic equilibrium, which is not biologically realistic, one has $dS/dt = \Pi = \Gamma = 0$, and there is neither energy influx to a system nor dissipated energy to the environment. However, for processes out of thermodynamic equilibrium, relevant for spine dynamics, we still can find a stationary regime where entropy of the spine does not change, $dS/dt = 0$, but entropy flux Γ and entropy production Π are nonzero and balance each other (Nicolis and Prigogine 1977; Tome 2006). It is more convenient to determine the stationary dissipated power by finding the entropy flux, which is given by (Tome 2006; Tome and de Oliveira 2010) (see Suppl. Infor.)

$$\Gamma = \frac{\tau_w}{\sigma_v^2} \langle [\Phi^{(1)}]^2 \rangle - \langle \Phi^{(2)} \rangle \tag{52}$$

Note that Eq. (52) is very similar in form to the energy rate $\dot{E}$ derived above; the two expressions differ only by the factor σ_v^2/τ_w, and none of them has the units of energy (Γ has the unit of the inverse of time). Thus, we need to introduce the energy scale in the problem. Generally, the stationary dissipated power per synapse $\dot{E}$ can be written as $\dot{E} = E_o \Pi = E_o \Gamma$ (Nicolis and Prigogine 1977), where E_o is the characteristic energy

scale associated with spine conductance changes, and its value is estimated next.

Estimation of the characteristic energy scale for synaptic plasticity.

As was said in the Introduction, the BCM model (either classical or extended) is only a phenomenological model of plasticity that does not relate directly to the underlying molecular processes in synapses. Consequently, a small, single, change of synaptic weight by Δw_i in Eq. (1) is accompanied in reality by many molecular transitions in synapse i. This means that a single degree of freedom related to w_i is in fact associated with many, hidden, molecular degrees of freedom. To be realistic in our energy cost estimates, we have to include those hidden degrees of freedom.

If we dealt with a process representing a single degree of freedom, then the energy scale E_o relating entropy flux Γ and energy rate $\dot{E}$, would be $E_o = kT$ (Nicolis and Prigogine 1977), where k is the Boltzmann constant and T is the tissue absolute temperature ($T \approx 310$ K). However, a dendritic spine is a composite object with multiple components and many degrees of freedom (Bonhoeffer and Yuste 2002; Holtmaat et al 2005; Meyer et al 2014; Choquet and Triller 2013), and hence the characteristic energy scale E_o is much bigger than kT. The changes in spine conductance on time scale of ~ 1 hr, i.e. for e-LTP and e-LTD, are induced by protein interactions in PSD (Lisman et al 2012; Kandel et al 2014) and subsequent membrane trafficking associated with AMPA and NMDA receptors (Borgdorff and Choquet 2002; Huganir and Nicoll 2013; Choquet and Triller 2013). Protein interactions are powered by phosphorylation process, which is one of the main biochemical mechanism of molecular signal transduction in PSD rele-

vant for synaptic plasticity (Bhalla and Iyengar 1999; Zhu et al 2016). Phosphorylation rates in an active LTP phase can be very fast, e.g., for CaMKII autophosphorylation they are in the range $60 - 600$ min^{-1} (Bradshaw et al 2002). Other processes in a spine, most notably protein turnovers in PSD (likely involved in l-LTP and l-LTD), are much slower ~ 3.7 days (Cohen et al 2013), and therefore their contribution to the energetics of the early phase of spine plasticity seems to be much less important (see, however Discussion for an estimate of the protein turnover energy rate).

The energy scale for protein interaction can be estimated as follows. A typical dendritic spine contains about 10^4 proteins (including their copies) (Sheng and Hoogenraad 2007). One cycle of protein phosphorylation requires the hydrolysis of 1 ATP molecule (Hill 1989; Qian 2007), which costs about $20kT$ (Phillips et al 2012). Each protein has on average 4-6 phosphorylation sites (Collins et al 2005; Trinidad et al 2012). If we assume conservatively that only about 20% of all PSD proteins are phosphorylated, then we obtain the energy scale for protein interactions roughly $2 \cdot 10^5 kT$, which is $8.6 \cdot 10^{-16}$ J.

Energy scale for receptor trafficking can be broadly decomposed into two parts: energy required for insertion of the receptors into the spine membrane, and energy related to their horizontal movement along the membrane to the top near a presynaptic terminal. The insertion energy for a typical protein is either about $3 - 17$ kcal/mol (Gumbart et al 2011) or $8 - 17kT$ (Grafmuller et al 2009), with the range spanning $4 - 25kT$, and is caused by a deformation in the membrane structure (Gumbart et al 2011). Since an average spine contains about 100 AMPA (Matsuzaki et al 2001;

Smith et al 2003) and 10 NMDA (Nimchinsky et al 2004) receptors, we obtain the total insertion energy in the rage $500 - 3200kT$. The second, movement contribution can be estimated by noting that typical forces that overcome friction and push macromolecules along membrane are about 10 pN, and they are powered by ATP hydrolysis (Fisher and Kolomeisky 1999). AMPA and NMDA receptors have to travel a spine distance of about 1 μm (Benavides-Piccione et al 2013), which requires the work of $110 \cdot 10^{-11} \cdot 10^{-6}$ N·m= $1.1 \cdot 10^{-15}$ J or $2.5 \cdot 10^5 kT$. The latter figure is 100 times larger than the insertion contribution, which indicates that the energy scale for receptor trafficking is dominated by the horizontal movement and is similar to the above for protein phosphorylation.

To summarize, the total energy scale E_o for spine conductance is about $E_o = 2 \cdot 10^{-15}$ J, or equivalently $4.6 \cdot 10^5 kT$ (or $2.3 \cdot 10^4$ ATP molecules).

Analytical approximation of the energy rate related to synaptic plasticity. It is not possible to find analytically the entropy flux Γ in Eq. (52) for an arbitrary probability distribution. However, Γ can be determined approximately for the probability distribution $P_s(v|f_o)$ in Eq. (37), by the saddle point method as a series expansion in the small noise amplitude D, which is proportional to $1/N$. We can write the entropy flux Γ in terms of the probability densities P_d and P_u appearing in Eq. (37) as

$$\Gamma = \langle \frac{(\Phi^{(1)})^2}{D} - \Phi^{(2)} \rangle_d + \langle \frac{(\Phi^{(1)})^2}{D} - \Phi^{(2)} \rangle_u \tag{53}$$

where

$$\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_d=\int_0^{v_{max}}dv\ P_d(v)\left(\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\right) \tag{54}$$

and

$$\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_u=\int_{v_{max}}^{\infty}dv\ P_u(v)\left(\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\right) \tag{55}$$

The essence of the saddle point method is in noting that for very small D, the probability distributions in Eq. (37) have two sharp maxima corresponding to two most likely synaptic currents v_d and v_u. This implies that the values of v that are the closest to v_d and v_u in $(\Phi^{(1)})^2/D-\Phi^{(2)}$ provide the biggest contributions to the integrals in Eqs. (54) and (55), and hence to the entropy flux Γ. Consequently, we have to expand the function $(\Phi^{(1)}(v))^2/D-\Phi^{(2)}(v)$ around v_d and v_u.

For v near v_d, the expansion is simpler if we introduce a unitless variable x, related to v such that $(v-v_d)=\sqrt{2D/\Phi_d^{(2)}}x$, where $\Phi_d^{(2)}$ is the second derivative of $\Phi(v)$ at $v=v_d$. Then to the order $\sim D$ we have:

$$\begin{aligned}\frac{(\Phi^{(1)}(x))^2}{D}-\Phi^{(2)}(x)&=\Phi_d^{(2)}(2x^2-1)+\sqrt{\frac{2D}{\Phi_d^2}}\Phi_d^{(3)}(2x^3-x)\\&+\frac{2D}{\Phi_d^2}\left[\Phi_d^{(4)}(\frac{2}{3}x^4-\frac{1}{2}x^2)+\frac{1}{2}\frac{(\Phi_d^{(3)})^2}{\Phi_d^{(2)}}x^4\right]+O(D^{3/2})\end{aligned} \tag{56}$$

A similar expression holds v near v_u, with a substitution $\Phi_d^{(n)}\mapsto\Phi_u^{(n)}$. Thus for $\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_d$ we have

$$\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_d = Z^{-1}\sqrt{\frac{2D}{\Phi_d^{(2)}}}e^{-\Phi_d/D}\int_{x_{d1}}^{x_{d2}} dx\; e^{-x^2}\left(\Phi_d^{(2)}(2x^2-1)+\sqrt{\frac{2D}{\Phi_d^2}}\Phi_d^{(3)}(2x^3-x)\right.$$
$$\left.+\frac{2D}{\Phi_d^{(2)}}\left[\Phi_d^{(4)}(\frac{2}{3}x^4-\frac{1}{2}x^2)+\frac{1}{2}\frac{(\Phi_d^{(3)})^2}{\Phi_d^{(2)}}x^4\right]\right) \quad (57)$$

where the limits of integration are $x_{d1} = \sqrt{\frac{\Phi_d^{(2)}}{2D}}(-v_d)$, and $x_{d2} = \sqrt{\frac{\Phi_d^{(2)}}{2D}}(v_{max}-v_d)$. Execution the above integrals yields

$$\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_d = Z^{-1}\sqrt{\frac{2D}{\Phi_d^{(2)}}}e^{-\Phi_d/D}\frac{\sqrt{\pi}}{8}[\mathrm{erf}(|x_{d1}|)+\mathrm{erf}(|x_{d2}|)]$$
$$\left(\frac{D}{(\Phi_d^{(2)})^2}[3(\Phi_d^{(3)})^2+2\Phi_d^{(2)}\Phi_d^{(4)}]+O(e^{-x_{d1}^2},e^{-x_{d2}^2})+O(D^2)\right), \quad (58)$$

where in the limit of small D, the exponential terms ($\sim e^{-x_{d1}^2}, e^{-x_{d2}^2}$) are small, and thus negligible. Next, it is easy to note that the prefactor in front of the large bracket simplifies, i.e, $Z^{-1}\sqrt{\frac{2D}{\Phi_d^{(2)}}}e^{-\Phi_d/D}\frac{\sqrt{\pi}}{8}[\mathrm{erf}(|x_{d1}|)+\mathrm{erf}(|x_{d2}|)] = \frac{Z_d}{4Z} = p_d/4$.

Applying the same procedure for $\langle\frac{(\Phi^{(1)})^2}{D}-\Phi^{(2)}\rangle_u$ gives us the total expression for the entropy flux Γ

$$\Gamma = \sum_{i=u,d} p_i\frac{D}{4(\Phi_i^{(2)})^2}[3(\Phi_i^{(3)})^2+2\Phi_i^{(2)}\Phi_i^{(4)}]+O(D^2), \quad (59)$$

where $i = d$ (down state) or $i = u$ (up state).

Having the entropy flux, we can determine analytically the power dissipated per synapse $\dot{E}$ due to synaptic plasticity. The result is

$$\dot{E} = E_o\Gamma \equiv p_d\dot{E}_d + p_u\dot{E}_u \tag{60}$$

where $\dot{E}_d$ and $\dot{E}_u$ are the energy rates dissipated in the down and up states, respectively. They take the form:

$$\dot{E}_i = \frac{E_o D}{4(\Phi_i^{(2)})^2}[3(\Phi_i^{(3)})^2 + 2\Phi_i^{(2)}\Phi_i^{(4)}] + O(D^2). \tag{61}$$

Note that the first nonzero contribution to the energy rate is of the order $\sim D$.

Neuron energy rate related to fast electric signaling.

We provide below an estimate of the energy used by a sensory neuron for short-term signaling for the sake of comparison with the energy requirement of synaptic plasticity. It has been suggested that the majority of neuronal energy goes to pumping out Na^+ ions (Na^+-K^+-ATPase), which accumulates mostly due to neural spiking activity, synaptic background activity, and passive Na^+ influx through sodium channels at rest (Attwell and Laughlin 2001). It has been shown that this short-term neuronal energy cost can be derived from a biophysical neuronal model, compared across species, and represented by a relatively simple equation (Karbowski 2009, 2012):

$$\mathrm{CMR}_{glu} = a_0 + a_1\langle r\rangle + b\rho_s f_o, \tag{62}$$

where CMR_{glu} is the glucose metabolic rate [in μmol/(cm$^3\cdot$ min)], ρ_s is the synaptic density, $\langle r\rangle$ is the average postsynaptic firing rate, and the parameters a_0, a_1, and b characterize the magnitude of the above three contributions to the neural metabolism, i.e. resting, firing rate, and synaptic transmission, respectively (Karbowski 2012). The average postsynaptic rate $\langle r\rangle$ is found from Eq. (46).

According to biochemical estimates, one oxidized glucose molecule generates about 31 ATP molecules (Rolfe and Brown 1997). In addition, 1 ATP molecule provides about $20kT$ of energy (Phillips et al 2012). This means that the short-term energy rate per neuron, denoted as $\dot{E}_n$, is given by

$$\dot{E}_n = 31\cdot 20\frac{N_A kT}{\rho_n}\mathrm{CMR}_{glu}, \tag{63}$$

where N_A is the Avogadro number, and ρ_n is the neuron density. We estimate the ratio of the synaptic plasticity power to neural power, i.e. $\dot{E}/\dot{E}_n$ across different presynaptic firing rates for three areas of the adult human cerebral cortex (frontal, temporal, and visual), and two areas of macaque monkey cerebral cortex (frontal and visual).

The values of the parameters a_0 and a_1 in Eq. (62) are species- and area-independent, and they read $a_0 = 2.1\cdot 10^{-10}$ mol/(cm^3 s), and $a_1 = 2.3\cdot 10^{-9}$ mol/cm^3 (Karbowski

2012). The rest of the parameters take different values for human and macaque cortex. Most of them are taken from empirical studies, and are given below. The parameter b, present in Eq. (62), is proportional to the neurotransmitter release probability and synaptic conductance, and it was estimated based on fitting developmental data for glucose metabolism CMR_{glu} and synaptic density ρ_s (which vary during the development) to the formula (62) (Karbowski 2012).

The following data are for an adult human cortex. The adult CMR_{glu} is 0.27 μmol/(cm^3·min) (frontal cortex), 0.27 μmol/(cm^3·min) (visual cortex), and 0.24 μmol/(cm^3·min) (temporal cortex) (Chugani 1998). The parameter b reads: $1.16 \cdot 10^{-20}$ mol (frontal), $0.63 \cdot 10^{-20}$ mol (visual), $0.17 \cdot 10^{-20}$ mol (temporal) (Karbowski 2012). Note that the value of b is 7 times larger for the frontal cortex than for the temporal, which might suggest that the product of neurotransmitter release probability and synaptic conductance is also 7 fold larger in the frontal cortex. This high difference may seem unlikely, however, it is still plausible, given that the release probability is highly variable and can assume values between 0.05-0.7 (Bolshakov and Siegelbaum 1995, Frick et al 2007; Volgushev et al 2004; Murthy et al 2001), and synaptic weights in the cortex are widely distributed (Loewenstein et al 2011). Neuron density ρ_n reads: $36.7 \cdot 10^6$ cm^{-3} (frontal), $66.9 \cdot 10^6$ cm^{-3} (visual), $59.8 \cdot 10^6$ cm^{-3} (temporal) (Pakkenberg and Gundersen 1997). Synaptic density ρ_s reads: $3.4 \cdot 10^{11}$ cm^{-3} (frontal), $3.1 \cdot 10^{11}$ cm^{-3} (visual), $2.9 \cdot 10^{11}$ cm^{-3} (temporal) (Huttenlocher and Dabholkar 1997).

The following data are for an adult (6 years old) macaque monkey cortex. The adult CMR_{glu} is 0.34 μmol/(cm^3·min) (frontal cortex), 0.40 μmol/(cm^3·min) (visual cortex)

(Noda et al 2002). The parameter b reads: $0.4 \cdot 10^{-20}$ mol (frontal), and $3.8 \cdot 10^{-20}$ mol (visual) (Karbowski 2012). Neuron density ρ_n reads: $9 \cdot 10^7$ cm^{-3} (frontal), $31.9 \cdot 10^7$ cm^{-3} (visual) (Christensen et al 2007). Synaptic density ρ_s reads: $5 \cdot 10^{11}$ cm^{-3} (frontal) (Bourgeois et al 1994), $6 \cdot 10^{11}$ cm^{-3} (visual) (Bourgeois and Rakic 1993).

Fisher information and coding accuracy in synapses.

Fisher information $I_F(f_o)$ about the driving input f_o is a good approximation of the mutual information between the driving presynaptic activity and postsynaptic current v (Brunel and Nadal 1998). It is also a measure of the coding accuracy and it is defined as (Cover and Thomas 2006)

$$I_F(f_o) = \langle \left(\frac{\partial \ln P_s(v|f_o)}{\partial f_o} \right)^2 \rangle. \tag{64}$$

Taking into account the form of probability density, Eq. (37), we can rewrite this equation as

$$I_F(f_o) = \langle \left(\frac{\partial \ln P_d(v|f_o)}{\partial f_o} \right)^2 \rangle_d + \langle \left(\frac{\partial \ln P_u(v|f_o)}{\partial f_o} \right)^2 \rangle_u. \tag{65}$$

where

$$\langle\left(\frac{\partial \ln P_i(v|f_o)}{\partial f_o}\right)^2\rangle_i = \left[Z'/Z + (\Phi_i/D)'\right]^2 p_i$$
$$+2\left[Z'/Z + (\Phi_i/D)'\right]\langle\Big(\Phi_i^{(2)}(v-v_i)^2/(2D)\Big)'\rangle_i$$
$$+\langle\left[\Big(\Phi_i^{(2)}(v-v_i)^2/(2D)\Big)'\right]^2\rangle_i \tag{66}$$

Our first goal is to express the factor Z'/Z in terms of the potential and its derivatives. To do it, we compute the following average:

$\langle\frac{\partial \ln P_s}{\partial f_o}\rangle = \langle\frac{\partial \ln P_d}{\partial f_o}\rangle_d + \langle\frac{\partial \ln P_u}{\partial f_o}\rangle_u$.

The left hand side of this equation is zero, since

$\langle(\ln P_s)'\rangle = \int_0^\infty dv(P_s)' = (\int_0^\infty dv P_s(v))' = 0$,

where a prime denotes a derivative with respect to f_o. Additionally,

$\langle(\ln P_i)'\rangle_i = -Z'/Z - (\Phi_i/D)' - \langle\Big(\Phi_i^{(2)}(v-v_i)^2/(2D)\Big)'\rangle_i$.

Combining the last two equations we obtain a relation between Z'/Z and the potentials:

$$Z'/Z = -(\Phi_d/D)'p_d - (\Phi_u/D)'p_u - \langle\Big(\Phi_d^{(2)}(v-v_d)^2/(2D)\Big)'\rangle_d$$
$$-\langle\Big(\Phi_u^{(2)}(v-v_u)^2/(2D)\Big)'\rangle_u. \tag{67}$$

After insertion of this expression into Eq. (66) and after some algebra, we arrive at the Fisher information

$$
\begin{aligned}
I_f(f_o) = \left[\left(\frac{\Phi_u - \Phi_d}{D}\right)'\right]^2 p_d p_u + 2\left(\frac{\Phi_u - \Phi_d}{D}\right)' \\
\times \left[p_d \langle \left(\Phi_u^{(2)}(v - v_u)^2/(2D)\right)' \rangle_u - p_u \langle \left(\Phi_d^{(2)}(v - v_d)^2/(2D)\right)' \rangle_d\right] \\
- \left(\langle \left(\Phi_d^{(2)}(v - v_d)^2/(2D)\right)' \rangle_d + \langle \left(\Phi_u^{(2)}(v - v_u)^2/(2D)\right)' \rangle_u\right)^2 \\
+ \langle \left[\left(\Phi_d^{(2)}(v - v_d)^2/(2D)\right)'\right]^2 \rangle_d + \langle \left[\left(\Phi_u^{(2)}(v - v_u)^2/(2D)\right)'\right]^2 \rangle_u
\end{aligned} \tag{68}
$$

The averages in the above equation can be computed to yield:

$$
\langle \left(\Phi_i^{(2)}(v - v_i)^2/(2D)\right)' \rangle_i = \frac{1}{2} p_i \left[\frac{(\Phi_i^{(2)})'}{\Phi_i^{(2)}} - \frac{D'}{D}\right] + O(e^{-1/D}) \tag{69}
$$

and

$$
\langle \left[\left(\Phi_i^{(2)}(v - v_i)^2/(2D)\right)'\right]^2 \rangle_i = 2 p_i (v_i')^2 \frac{\Phi_i^{(2)}}{2D} + \frac{3}{4} p_i \left[\frac{(\Phi_i^{(2)})'}{\Phi_i^{(2)}} - \frac{D'}{D}\right]^2 + O(e^{-1/D}) \tag{70}
$$

After insertion of these expressions into Eq. (68), and some algebraic manipulations, we arrive at Eq. (17) for I_F in the Results.

Relationship between synaptic energy rate and Fisher information in the limit $D \mapsto 0$.

Below we derive the relation given by Eqs. (18) and (19) in the limit of very weak synaptic noise, $D \mapsto 0$. In this limit, it can be noted that the product of the fractions of weak and strong synapses is

$$p_d^{(0)} p_u^{(0)} = \frac{e^{(\Phi_d - \Phi_u)/D} \sqrt{\Phi_d^{(2)}/\Phi_u^{(2)}}}{\left(1 + e^{(\Phi_d - \Phi_u)/D} \sqrt{\Phi_d^{(2)}/\Phi_u^{(2)}}\right)^2}. \tag{71}$$

This expression enables us to write in a compact form the derivative of $p_d^{(0)}$ with respect to f_o as

$$\frac{\partial p_d^{(0)}}{\partial f_o} = p_d^{(0)} p_u^{(0)} \left[\left(\frac{\Phi_u - \Phi_d}{D} \right)' + \frac{1}{2} \left[(\Phi_u^{(2)})'/\Phi_u^{(2)} - (\Phi_d^{(2)})'/\Phi_d^{(2)} \right] \right] \tag{72}$$

where the prime denotes differentiation with respect to f_o. On the other hand it can be noted that, in the bistable regime, the Fisher information in Eq. (17) in the leading order $1/D^2$ can be written as

$$I_F(f_o) = p_d^{(0)} p_u^{(0)} \left[\left(\frac{\Phi_u - \Phi_d}{D} \right)' + \frac{1}{2} \left[(\Phi_u^{(2)})'/\Phi_u^{(2)} - (\Phi_d^{(2)})'/\Phi_d^{(2)} \right] \right]^2 + O(1/D), \tag{73}$$

which is similar in form to the expression for $\partial p_d^{(0)} \partial f_o$. This suggests that we can combine the two equations, and arrive at

$$I_F(f_o) = \frac{\left(\partial p_d^{(0)}/\partial f_o\right)^2}{p_d^{(0)} p_u^{(0)}} \left[1 + O(D)\right], \tag{74}$$

which is Eq. (18) in the Results.

Next, we want to relate Eq. (74) for I_F to the energy rate. Energy rate $\dot{E} = p_d^{(0)} \dot{E}_d + p_u^{(0)} \dot{E}_u$ can be differentiated with respect to f_o, which yields

$$\frac{\partial \dot{E}}{\partial f_o} = \frac{\partial p_d^{(0)}}{\partial f_o}(\dot{E}_d - \dot{E}_u) + p_d^{(0)} \frac{\partial \dot{E}_d}{\partial f_o} + p_u^{(0)} \frac{\partial \dot{E}_u}{\partial f_o}, \tag{75}$$

where we used the relation $\partial p_d^{(0)}/\partial f_o = -\partial p_u^{(0)}/\partial f_o$, which follows from the fact that $p_d^{(0)} + p_u^{(0)} = 1$. Now it is crucial to note that the first term in $\partial \dot{E}/\partial f_o$, i.e. $\partial p_d^{(0)}/\partial f_o$, is of the order $1/D$, whereas the rest terms ($\sim p_d^{(0)}, p_u^{(0)}$) are of the order of one. This implies that the first term dominates in the limit $D \mapsto 0$. Thus, we can write approximately the expression for $\partial p_d^{(0)}/\partial f_o$, involving the energy rate as

$$\frac{\partial p_d^{(0)}}{\partial f_o} = \frac{\partial \dot{E}/\partial f_o}{(\dot{E}_d - \dot{E}_u)} \left[1 + O(D)\right] \tag{76}$$

Finally, if we combine Eqs. (74) and (76), we obtain Eq. (19) in the Results for the bistable regime.

Numerical simulations of the full synaptic system.

Numerical stochastic dynamics of the whole synaptic system given by Eqs. (1-2) were performed using a stochastic version of Runge-Kutta scheme (Roberts 2012).

Energy dissipated for plasticity by the full synaptic system (Eqs. 1-2) was computed numerically using the approach presented in (Tome 2006, Tome 2010). We can rewrite Eq. (1) in a more compact form as

$$\frac{dw_i}{dt} = F_{w,i} + \frac{\sqrt{2}\sigma_w(1+\tau_f f_i)}{\sqrt{\tau_w}}\eta_i \tag{77}$$

where

$$F_{w,i} = \lambda f_i r(r-\theta) - \frac{(w_i - \epsilon a)}{\tau_w}. \tag{78}$$

This enables us to write the entropy flux in the steady state (equivalent to entropy production rate) of the full synaptic system in a compact form. Consequently, the numerical entropy flux per synapse of the whole system Γ_{num} is

$$\Gamma_{num} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\tau_w}{\sigma_w^2}\langle\left(\frac{F_{w,i}}{(1+\tau_f f_i)}\right)^2\rangle + \langle F'_{w,i}\rangle\right) \tag{79}$$

where $F'_{w,i} = \partial F_{w,i}/\partial w_i$, and it is given by

$$F'_{w,i} = -\frac{1}{\tau_w} + \frac{\lambda\beta A^2 f_i^2(2r-\theta)}{N(2r+\kappa A^2)}. \tag{80}$$

The numerical energy rate $\dot{E}_{num}$ is $\dot{E}_{num} = E_o\Gamma_{num}$. The brackets $\langle ... \rangle$ in Eq. (79) denote averaging over fluctuations in synaptic noise and presynaptic firing rates (averaging over η and x stochastic variables). In numerical simulations, these averages are computed as temporal averages over long simulation time. This equivalence in averaging is guaranteed due to ergodic theorem. The minimal number of time steps for numerical convergence is of the order of $\sim 10^5$.

Parameters used in computations.

The following values of various parameters were used: $V_r = -65$ mV, $q = 0.35$ (Volgushev et al 2004), $\tau_{nmda} = 150$ msec (Nimchinsky et al 2004), $\tau_{ampa} = 5$ msec (Smith et al 2003), $\tau_f = 1.0$ sec, $a = 1.0$ nS, $\alpha = 0.3$ sec (Zenke et al 2013), $\epsilon = 3 \cdot 10^{-4}$, $A = 600$ Hz$/\sqrt{nA}$, $\tau_w = 3600$ sec (Frey and Morris 1997; Zenke et al 2013), $\sigma_f = 10$ Hz (Buzsaki and Mizuseki 2014), $N = 2 \cdot 10^3$ (average value for many species of primates; see Sherwood et al (2020), Elston et al (2001)). The amplitude of synaptic weight noise σ_w was taken in the range $0.02 \leq \sigma_w \leq 0.5$ nS, which is the range suggested in experimental studies (Matsuzaki et al 2001; Smith et al 2003). The two undetermined parameters are λ and κ, and two sets of values were used for them: (i) $\kappa = 0.001$ (nA·sec), $\lambda = 9 \cdot 10^{-7}$ (nS·sec^2), and (ii) $\kappa = 0.012$ (nA·sec), $\lambda = 10^{-5}$ (nS·sec^2), in

order to obtain a transition to the bistable regime for $f_o \sim 1 - 5$ Hz. The value of A was chosen to have postsynaptic firing rate in the range $0.1 - 10$ Hz. The value of κ was chosen to obtain v_u in the neurophysiological range ~ 1 pA (O'Connor et al 2006).

Supporting Information.

S1 Text. This file contains the details of some calculations. (PDF)

Acknowledgments

The work was supported by the Polish National Science Centre (NCN) grant no. 2015/17/B/NZ4/02600.

References

Aiello LC, Wheeler P (1995) The expensive-tissue hypothesis: The brain and the digestive-system in human and primate evolution. *Curr. Anthropology* **36**: 199-221.

Alle H, Roth A, Geiger JRP (2009) Energy-efficient action potentials in hippocampal mossy fibers. *Science* **325**: 1405-1408.

Attwell D, Laughlin SB (2001) An energy budget for signaling in the gray matter of the brain. *J. Cereb. Blood Flow Metabol.* **21**: 1133-1145.

Balasubramanian V, Kimber D, Berry MJ (2001) Metabolically efficient information processing. *Neural. Comput.* **13**: 799-815.

Barato AC, Seifert U (2015) Thermodynamic uncertainty relation for biomolecular processes. *Phys. Rev. Lett.* **114**: 158101.

Bartol TM, Bromer C, Kinney J, Chirillo MA, Bourne JN, et al (2015) Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife* **4**: e10778.

Benavides-Piccione R, Fernaud-Espinosa I, Robles V, Yuste R, DeFelipe J (2013) Age-based comparison of human dendritic spine structure using complete three-dimensional reconstructions. *Cerebral Cortex* **23**: 1798-1810.

Benna MK, Fusi S (2016) Computational principles of synaptic memory consolidation. *Nature Neurosci.* **19**: 1697-1706.

Bennett CH (1979) Dissipation-error tradeoff in proofreading. *BioSystems* **11**: 85-91.

Bennett CH (1982) The thermodynamics of computation - a review. *Int. J. Theor. Physics* **21**: 905-940.

Berut A, Arakelyan A, Petrosyan A, Ciliberto S, Dillenschneider R, Lutz E (2012)

Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* **483**: 187-190.

Bhalla US, Iyengar R (1999) Emergent properties of networks of biological signaling pathways. *Science* **283**: 381-387.

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.

Billings G, van Rossum MC (2009) Memory retention and spike-timing-dependent plasticity. *J. Neurophysiol.* **101**: 2775-2788.

Bolshakov VY, Siegelbaum SA (1995) Regulation of hippocampal transmitter release during development and long-term potentiation. *Science* **269**: 1730-1734.

Bonhoeffer T, Yuste R (2002) Spine motility: phenomenology, mechanisms, and function. *Neuron* **35**: 1019-1027.

Borgdorff AJ, Choquet D (2002) Regulation of AMPA receptor lateral movements. *Nature* **417**: 649-653.

Bourgeois JP, Rakic P (1993) Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. *J. Neurosci.* **13**: 2801-2820.

Bourgeois JP, Goldman-Rakic, PS Rakic P (1994) Synaptogenesis in the prefrontal cortex of rhesus monkeys. *Cereb. Cortex* **4**: 78-96.

Bourne J, Harris KM (2007) Do thin spines learn to be mushroom spines that remember? *Curr. Opin. Neurobiol.* **17**: 381-386.

Bradshaw JM, Hudmon A, Schulman H (2002) Chemical quenched flow kinetic stud-

ies indicate an intraholoenzyme autophosphorylation mechanism for Ca^{2+}/Calmodulin-dependent protein kinase II. *J. Biol. Chem.* **277**: 20991-20998.

Brunel N, Nadal JP (1998) Mutual information, Fisher information and population coding. *Neural Comput.* **10**: 1731-1757.

Buzsaki G, Mizuseki K (2014) The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* **15**: 264-278.

Chaudhuri R, Fiete I (2016) Computational principles of memory. *Nature Neurosci.* **19**: 394-403.

Choquet D, Triller A (2013) The dynamic synapse. *Neuron* **80**: 691-703.

Christensen JR, Larsen KB, Lisanby SH, Scalia J, Arango V, et al (2007) Neocortical and hippocampal neuron and glial cell numbers in the rhesus monkey. *Anatomical Record* **290**: 330-340.

Chugani HT (1998) A critical period of brain development: Studies of cerebral glucose utilization with PET. *Preventive Medicine* **27**: 184-188.

Cingolani L, Goda Y (2008) Actin in action: the interplay between the actin cytoskeleton and synaptic efficacy. *Nat. Rev. Neurosci.* **9**: 344-356.

Clopath C, Busing L, Vasilaki E, Gerstner W (2010) Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nature Neurosci.* **13**: 344-352.

Cohen LD, Zuchman R, Sorokina O, Muller A, Dieterich DC, Armstrong JD, et al. (2013) Metabolic turnover of synaptic proteins: kinetics, interdependencies and implications for synaptic maintenance. *PLoS ONE* **8**: e63191.

Collins MO, Yu L, Coba MP, Husi H, Campuzano I, Blackstock WP, Choudhary JS,

Grant SCN (2005) Proteomic analysis of in vivo phosphorylated synaptic proteins. *J. Biol. Chem.* **280**: 5972-5982.

Cooper LN, Bear MF (2012) The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nat. Rev. Neurosci.* **13**: 798-810.

Costa RP, Froemke RC, Sjostrom PJ, Rossum MCW (2015) Unified pre- and postsynaptic long-term plasticity enables reliable and flexible learning. *eLife* **4**: e09457.

Cover TM, Thomas JA (2006) *Elements of Information Theory.* Wiley: Hoboken, NJ.

DeFelipe J, Alonso-Nanclares L, Arellano JI (2002) Microstructure of the neocortex: Comparative aspects. *J. Neurocytol.* **31**: 299-316.

De Koninck P, Schulman H (1998) Sensitivity of CaM kinase II to the frequency of Ca^{2+} oscillations. *Science* **279**: 227-230.

Elston GN, Benavides-Piccione R, DeFelipe J (2001) The pyramidal cell in cognition: A comparative study in human and monkey. *J. Neurosci.* **21**: RC163 (1-5).

Engl E, Attwell D (2015) Non-signalling energy use in the brain. *J. Physiol.* **593**: 3417-3429.

Engl E, Jolivet R, Hall CN, Attwell D (2017) Non-signalling energy use in the developing rat brain. *J. Cereb. Blood Flow Metab.* **37**: 951-966.

Ermentrout GB (1998) Linearization of F-I curves by adaptation. *Neural Comput.* **10**: 1721-1729.

Ermentrout GB, Terman DH (2010) *Mathematical Foundations of Neuroscience.* Springer: New York, NY.

Fisher ME, Kolomeisky AB (1999) Molecular motors and the forces they exert. *Physica*

A 274: 241-266.

Frey U, Morris RGM (1997) Synaptic tagging and long-term potentiation. *Nature* **385**: 533-536.

Frick A, Feldmeyer D, Sakmann B (2007) Postnatal development of synaptic transmission in local networks of L5A pyramidal neurons in rat somatosensory cortex. *J. Physiol.* **585**: 103-116.

Friston K (2010) The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **11**: 127-138.

Fusi S, Drew PJ, Abbott LF (2005) Cascade models of synaptically stored memories. *Neuron* **45**: 599-611.

Fusi S, Abbott LF (2007) Limits on the memory storage capacity of bounded synapses. *Nat. Neurosci.* **10**: 485-493.

Gage FH, Kelly PAT, Bjorklund A (1984) Regional changes in brain glucose metabolism reflect cognitive impairments in aged rats. *J. Neurosci.* **4**: 2856-2865.

Gardiner CW (2004) *Handbook of Stochastic Methods.* Berlin: Springer.

Goldt S, Seifert U (2017) Stochastic thermodynamics of learning. *Phys. Rev. Lett.* **118**: 010601.

Govindarajan A, Kelleher RJ, Tonegawa S (2006) A clustered plasticity model of long-term memory engrams. *Nat. Rev. Neurosci.* **7**: 575-583.

Govindarajan A, Israely I, Huang SY, Tonegawa S (2011) The dendritic branch is the preferred integrative unit for protein synthesis-dependent LTP. *Neuron* **69**: 132-146.

Grafmuller A, Shillcock J, Lipowsky R (2009) The fusion of membranes and vesicles:

pathway and energy barriers from dissipative particle dynamics. *Biophys. J* **96**: 2658-2675.

Graupner M, Brunel N (2012) Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proc. Natl. Acad. Sci. USA* **109**: 3991-3996.

Gumbart J, Chipot C, Schulten K (2011) Free-energy cost for translocon-assisted insertion of membrane proteins. *Proc. Natl. Acad. Sci. USA* **108**: 3596-3601.

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

Harris JJ, Jolivet R, Attwell D (2012) Synaptic energy use and supply. *Neuron* **75**: 762-777.

Hill TL (1989) *Free Energy Transduction and Biochemical Cycle Kinetics.* New York: Springer-Verlag.

Hofman MA (1988) Size and shape of the cerebral cortex in mammals. II. The cortical volume. *Brain Behav. Evol.* **32**: 17-26.

Holtmaat AJ, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, et al (2005) Transient and persistent dendritic spines in the neocortex in vivo. *Neuron* **45**: 279-291.

Honkura N, Matsuzaki M, Noguchi J, Ellis-Davies GCR, Kasai H (2008) The subspine organization of actin fibers regulates the structure and plasticity of dendritic spines. *Neuron* **57**: 719-729.

Huganir RL, Nicoll RA (2013) AMPARs and synaptic plasticity: the last 25 years.

Neuron **80**: 704-717.

Huttenlocher PR, Dabholkar AS (1997) Regional differences in synaptogenesis in human cerebral cortex. *J. Comparative Neurology* **387**: 167-178.

Izhikevich EM, Desai NS (2003) Relating STDP to BCM. *Neural Comput.* **15**: 1511-1523.

Jedlicka P, Benuskova L, Abraham WC (2015) A voltage-based STDP rule combined with fast BCM-like metaplasticity accounts for LTP and concurrent 'heterosynaptic' LTD in the dentate gyrus in vivo. *PloS Comput. Biol.* **11**: e1004588.

Kandel ER, Dudai Y, Mayford MR (2014) The molecular and systems biology of memory. *Cell* **157**: 163-186.

Karbowski J (2007) Global and regional brain metabolic scaling and its functional consequences. *BMC Biol.* **5**: 18.

Karbowski J (2009) Thermodynamic constraints on neural dimensions, firing rates, brain temperature and size. *J. Comput. Neurosci.* **27**: 415-436.

Karbowski J (2012) Approximate invariance of metabolic energy per synapse during development in mammalian brains. *PLoS ONE* **7**: e33425.

Karbowski J (2014) Constancy and trade-offs in the neuroanatomical and metabolic design of the cerebral cortex. *Front. Neural Circuits* **8**: 9.

Karbowski J (2015) Cortical composition hierarchy driven by spine proportion economical maximization or wire volume minimization. *PloS Comput. Biol.* **11**: e1004532.

Karbowski J (2019) Metabolic constraints on synaptic learning and memory. *J. Neurophysiol.* **122**: 1473-1490.

Kasai H, Matsuzaki M, Noguchi J, Yasumatsu N, Nakahara H (2003) Structure-stability-function relationships of dendritic spines. *Trends Neurosci.* **26**: 360-368.

Kirkwood A, Rioult MG, Bear MF (1996) Experience-dependent modification of synaptic plasticity in visual cortex. *Nature* **381**: 526-528.

Lan G, Sartori P, Neumann S, Sourjik V, Tu Y (2012) The energy-speed-accuracy trade-off in sensory adaptation. *Nature Physics* **8**: 422-428.

Landauer R (1961) Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **5**: 183-191.

Lang AH, Fisher CK, Mora T, Mehta P (2014) Thermodynamics of statistical inference by cells. *Phys. Rev. Lett.* **113**: 148103.

Laughlin SB, de Ruyter van Steveninck RR, Anderson JC (1998) The metabolic cost of neural information. *Nature Neurosci.* **1**: 36-40.

Lee SR, Escobedo-Lozoya Y, Szatmeri eM, Yasuda R (2009) Activation of CaMKII in single dendritic spines during long-term potentiation. *Nature* **458**: 299-304.

Leff HS, Rex AF (1990) *Maxwell's Demon: Entropy, Information, Computing.* Princeton Univ. Press: Princeton, NJ.

Levy WB, Baxter RA (1996) Energy efficient neural codes. *Neural Comput.* **8**: 531-543.

Lisman J, Yasuda R, Raghavachari S (2012) Mechanisms of CaMKII action in long-term potentiation. *Nat. Rev. Neurosci.* **13**: 169-182.

Loewenstein Y, Kuras A, Rumpel S (2011) Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *J. Neurosci.* **31**: 9481-9488.

Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* **453**: 869-878.

Magistretti PJ, Pellerin J, Rothman DL, Shulman RG (1999) Energy on demand. *Science* **283**: 496-497.

Markram H, Lubke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**: 213-215.

Matsuzaki M, Ellis-Davies GCR, Nemoto T, Miyashita Y, Iino M, Kasai H (2001) Dendritic spine geometry is critical for AMPA receptor expression in hippocampal CA1 pyramidal neurons. *Nat. Neurosci.* **4**: 1086-1092.

Matsuzaki M, Honkura N, Ellis-Davies GCR, Kasai H (2004) Structural basis of long-term potentiation in single dendritic spines. *Nature* **429**: 761-766.

Mehta P, Schwab DJ (2012) Energetic costs of cellular computation. *Proc. Natl. Acad. Sci. USA* **109**: 17978-17982.

Mery F, Kawecki TJ (2005) A cost of long-term memory in Drosophila. *Science* **308**: 1148.

Meyer D, Bonhoeffer T, Scheuss V (2014) Balance and stability of synaptic structures during synaptic plasticity. *Neuron* **82**: 430-443.

Miller KD, MacKay DJ (1994) The role of constraints in Hebbian learning. *Neural Comput.* **6**: 100-126.

Miller P, Zhabotinsky AM, Lisman JE, Wang X-J (2005) The stability of a stochastic CaMKII switch: Dependence on the number of enzyme molecules and protein turnover. *PLoS Biol.* **3**: e107.

Montgomery JM, Madison DV (2004) Discrete synaptic states define a major mechanism of synaptic plasticity. *Trends Neurosci.* **27**: 744-750.

Murthy VN, Schikorski T, Stevens CF, Zhu Y (2001) Inactivity produces increases in neurotransmitter release and synapse size. *Neuron* **32**: 673-682.

Nicolis G, Prigogine I (1977) *Self-Organization in Nonequilibrium Systems.* Wiley: New York, NY.

Nimchinsky EA, Yasuda R, Oertner TG, Svoboda K (2004) The number of glutamate receptors opened by synaptic stimulation in single hippocampal spines. *J. Neurosci.* **24**: 2054-2064.

Niven B, Laughlin SB (2008) Energy limitation as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.* **211**: 1792-1804.

Noda A, Ohba H, Kakiuchi T, Futatsubashi M, Tsukada H, et al (2002) Age-related changes in cerebral blood flow and glucose metabolism in conscious rhesus monkeys. *Brain Res.* **936**: 76-81.

Novikov EA (1965) Functionals and the random-force method in turbulence theory. *Sov. Phys. JETP* **20**: 1290-1294.

O'Connor DH, Wittenberg GM, Wang SSH (2005) Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc. Natl. Acad. Sci. USA* **102**: 9679-9684.

Pakkenberg B, Gundersen HJG (1997) Neocortical neuron number in humans: Effect of sex and age. *J. Comp. Neurol.* **384**: 312-320.

Parrondo JMR, Horowitz JM, Sagawa T (2015) Thermodynamics of information. *Na-*

ture Physics **11**: 131-139.

Petersen CC, Malenka RC, Nicoll RA, Hopfield JJ (1998) All-or-none potentiation at CA3-CA1 synapses. *Proc. Natl. Acad. Sci. USA* **95**: 4732-4737.

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

Phillips R, Kondev J, Theriot J, Garcia H (2012) *Physical Biology of the Cell.* Garland Science: London.

Phillips R, Ursell T, Wiggins P, Sens P (2009) Emerging roles for lipids in shaping membrane-protein function. *Nature* **459**: 379-385.

Placais PY, Preat T (2013) To favor survival under food shortage, the brain disables costly memory. *Science* **339**: 440-442.

Placais PY et al (2017) Upregulated energy metabolism in the Drosophila mushroom body is the trigger for long-term memory. *Nature Comm.* **8**: 15510.

Qian H (2007) Phosphorylation energy hypothesis: Open chemical systems and their biological function. *Annu. Rev. Phys. Chem.* **58**: 113-142.

Redondo RL, Morris RGM (2011) Making memories last: the synaptic tagging and capture hypothesis. *Nat. Rev. Neurosci.* **12**: 17-30.

Risken H (1996) *The Fokker-Planck equation.* Springer: Berlin.

Roberts AJ (2001) Modify the improved Euler scheme to integrate stochastic differential equations. ArXiv:1210.0933 (www.arxiv.org/abs/1210.0933)

Rolfe DFS, Brown GC (1997) Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol. Revs.* **77**: 731-758.

Seifert U (2012) Stochastic thermodynamics, fluctuation theoremes and molecular machines. *Rep. Prog. Phys.* **75**: 126001.

Sheng M, Hoogenraad CC (2007) The postsynaptic architecture of excitatory synapses: A more quantitative view. *Annu. Rev. Biochem.* **76**: 823-847.

Sherwood CS et al (2020) Invariant synapse density and neuronal connectivity scaling in primate neocortical evolution. *Cerebral Cortex*, advance online publication.

Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. USA* **99**: 10831-10836.

Shulman RG, Rothman DL, Behar KL, Hyder F (2004) Energetic basis of brain activity: implications for neuroimaging. *Trends Neurosci.* **27**: 489-495.

Smith MA, Ellis-Davies GCR, Magee J (2003) Mechanism of the distance-dependent scaling of Schaffer collateral synapses in rat CA1 pyramidal neurons. *J. Physiol.* **548**: 245-258.

Smolen P, Baxter DA, Byrne JH (2012) Molecular constraints on synaptic tagging and maintenance of long-term potentiation: a predictive model. *PLoS Comput. Biol.* **8**: e1002620.

Smolen P, Baxter DA, Byrne JH (2019) How can memories last for days, years, or a lifetime? Proposed mechanisms for maintaining synaptic potentiation and memory. *Learning and Memory* **26**: 133-150.

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

Stachowiak JC, Brodsky FM, Miller EA (2013) A cost-benefit analysis of the physical

mechanisms of membrane curvature. *Nature Cell Biol.* **15**: 1019-1027.

Statman A, Kaufman M, Minerbi A, Ziv NE, Brenner N (2014) Synaptic size dynamics as an effective stochastic process. *PLoS Comput. Biol.* **10**: e1003846.

Still S, Sivak DA, Bell AJ, Crooks GE (2012) Thermodynamics of prediction. *Phys. Rev. Lett.* **109**: 120604.

Takeuchi T, Duszkiewicz AJ, Morris RGM (2014) The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Phil. Trans. R. Soc. B* **369**: 20130288.

Tetzlaff C, Kolodziejski C, Timme M, Worgotter F (2011) Synaptic scaling in combination with many generic plasticity mechanisms stabilizes circuit connectivity. *Front. Comput. Neurosci.* **5**: 47.

Tkacik G, Mora T, Marre O, Amodei D, Palmer SE, Berry MJ, Bialek W (2015) Thermodynamics for a network of neurons: Signatures of criticality. *Proc. Natl. Acad. Sci. USA* **112**: 11508-11513.

Tome T (2006) Entropy production in nonequilibrium systems described by a Fokker-Planck equation. *Braz. J. Physics* **36**: 1285-1289.

Tome T, de Oliveira MJ (2010) Entropy production in irreversible systems described by a Fokker-Planck equation. *Phys. Rev. E* **82**: 021120.

Toyoizumi T, Kaneko M, Stryker MP, Miller KD (2014) Modeling the dynamic interaction of Hebbian and homeostatic plasticity. *Neuron* **84**: 497-510.

Trinidad JC, Barkan DT, Gulledge BF, Thaalhammer A, Sali A, Schoepfer R, Burlingame AL (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at murine synapse. *Molecular & Cellular Proteomics* **11**: 215-229.

Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.* **5**: 97-107.

Van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry.* Elsevier: Amsterdam.

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

Volgushev M, Kudryashov I, Chistiakova M, Mukovski M, Niesman J, et al (2004) Probability of transmitter release at neocortical synapses at different temperatures. *J. Neurophysiol.* **92**: 212-220.

Zenke F, Hennequin G, Gerstner W (2013) Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. *PloS Comput. Biol.* **9**: e1003330.

Zenke F, Agnes EJ, Gerstner W (2015) Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat. Commun.* **6**: 6922.

Zenke F, Gerstner W (2017) Hebbian plasticity requires compensatory processes on multiple timescales. *Phil. Trans. R. Soc. B* **372**: 20160259.

Zhu J, Shang Y, Zhang M (2016) Mechanistic basis of MAGUK-organized complexes in synaptic development and signalling. *Nat. Rev. Neurosci.* **17**: 209-223.

Ziegler L, Zenke F, Kastner DB, Gerstner W (2015) Synaptic consolidation: from synapses to behavioral modeling. *J. Neurosci.* **35**: 1319-1334.

Figure Captions

Fig. 1

Asymptotic behavior of the deterministic extended BCM model. Temporal dependence of deterministic synaptic weights w_i from Eqs. (1-2). (Upper panel) When mean firing rate f_o is smaller than some threshold value, then all weights converge on the same asymptotic value regardless of the initial values (red lines for strong initial synapses, and blue lines for weak initial synapses). This case corresponds to monostability. (Lower panel) When f_o is in the intermediate interval, then synaptic weights can converge into two separate values, depending on initial conditions. In this regime, there are two different coexisting fixed points (bistability). Strong initial weights lead to large final value (red lines), and weak initial weights lead to low final value (blue lines). Parameters used: $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$.

Fig. 2

Phase and bifurcation diagrams of w_i for the deterministic extended BCM model. (A) Schematic phase diagram λ vs. f_o for monostable (mono) and bistable (bi) behavior of synaptic weights w_i. Bistability is lost for sufficiently large λ. Note that the critical values of λ and f_o, for which bistability emerges and disappears, are inversely related. (B) Bifurcation diagram of an asymptotic synaptic weight vs. f_o for a typical synapse. The bistable regime is indicated by dotted lines. Parameters used: $\lambda = 9 \cdot 10^{-7}$, and $\kappa = 0.001$.

Fig. 3

Distributions of synaptic weights w_i for the stochastic extended BCM model. (A) Temporal fluctuations of an individual synapse in the bistable regime. Note stochastic jumps between weak and strong weights. (B) Distribution of synaptic weights for an individual synapse (the same as in A) in the bi- ($f_o = 10$ Hz; blue solid line) and mono-stable ($f_o = 1$ Hz; red dashed line) regimes. (C) Cumulative distribution of synaptic weights for all N synapses in the bi- ($f_o = 10$ Hz; blue solid line) and mono-stable ($f_o = 1$ Hz; red dashed line) regimes. Note that both distributions in (B) and (C) are very similar. Parameters used: $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_w = 0.1$ nS.

Fig. 4

Firing rates and emergence of bistability in the mean-field model: theory. (A) Postsynaptic firing rates r as functions of the population averaged synaptic current v for different neuronal adaptations values κ (in nA·sec). Increasing κ causes decrease in r and makes the functional form $r(v)$ more linear. (B) Graphical solutions of Eq. (7) and multiple roots for stationary v. For small driving-plasticity h there is only one intersection of $g(v)$ and the line $y = v$ at $v \sim O(\epsilon)$, corresponding to v_d and monostability (dashed red line; $f_0 = 0.3$ Hz, $\sigma_f = 8$ Hz). For higher h ($h > h_{cr}$) there are three intersections, but the middle one corresponds to an unstable solution, which in effect yields two stable solutions, i.e. bistability (dashed-dotted yellow line; $f_0 = 5.0$ Hz, $\sigma_f = 10$ Hz). When h is very large, then there is only one intersection, and it occurs for large v, corresponding to monostability with strong synapses v_u only (dotted green line; $f_0 = 30.0$ Hz, $\sigma_f = 10$ Hz). In panel (B) $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$.

Fig. 5

Phase diagram of mono- vs. bi-stability in the deterministic mean-field model. Numerical solution of Eq. (3) and (6) for $\sigma_v = 0$. (A) $\kappa = 0.001$, (B) $\kappa = 0.012$. The solid and dashed lines are the boundaries between mono- (mono) and bistable (bi) regimes, and they correspond to the condition $\Delta = 0$ in the (λ, f_o) space.

Fig. 6

Bifurcation diagram of v in the deterministic mean-field model. Numerical solution of Eq. (3) and (6) for $\sigma_v = 0$. (A) Parameters used: $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$. (B) Parameters used: $\lambda = 1 \cdot 10^{-5}$ and $\kappa = 0.012$. Note that both diagrams in A and B are qualitatively similar, and there exists a critical value of f_o above which bistability emerges.

Fig. 7

Distribution of synaptic current v in the stochastic mean-field model. Numerical solution of Eq. (3) and (6) for $\sigma_v > 0$. In the monostable regime ($f_o = 1$ Hz) the distribution is unimodal with a sharp peak around $v_d \approx 0$. In the bistable regime ($f_o = 10$ Hz) the distribution is bimodal with two maxima around two fixed points v_d and v_u. (A) Parameters used: $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$. (B) Parameters used: $\lambda = 10^{-5}$ and $\kappa = 0.012$. In both (A) and (B) $\sigma_w = 0.1$ nS.

Fig. 8

Average synaptic current $\langle v \rangle$ as a function of mean presynaptic firing rate: comparison of exact results with mean-field. (A) Dependence of $\langle v \rangle$ on f_o for

$\sigma_w = 0.02$. (B) The same for $\sigma_w = 0.1$. (C) The same for $\sigma_w = 0.5$. For all panels solid lines correspond to exact result (from Eqs. 1-2), dashed lines to mean-field (from Eq. 44). The results are for $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$.

Fig. 9

Average synaptic current $\langle v \rangle$ as a function of standard deviation of the presynaptic firing rate: comparison of exact results with mean-field. (A) Dependence of $\langle v \rangle$ on σ_f for $f_o = 1$ Hz. (B) The same for $f_o = 10$ Hz. For all panels solid lines correspond to exact numerical result (from Eqs. 1-2), dashed lines to mean-field (from Eq. 44). The results are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_w = 0.1$.

Fig. 10

Effective synaptic potential, metastability, and memory lifetime: theory. (A) The metastable synaptic states can be described in probabilistic terms and correspond to minima of an effective potential $\Phi(v|f_o)$. For weak presynaptic driving input f_o the potential Φ has only one minimum at $v_d \sim O(\epsilon)$, related to weak synapses. If f_o is above a certain threshold, then the potential displays two minima, corresponding to bistable coexistence of weak and strong synapses (v_d and v_u). In the bistable regime, the synapses can jump between weak and strong states due to fluctuations in the input and/or synaptic noise. (B) Characteristic long times in the up (T_u), down (T_d) synaptic states, and memory lifetime T_m as functions of presynaptic firing f_o. Curves in (A) and (B) are for $\lambda = 9 \cdot 10^{-7}$ and $\kappa = 0.001$.

Fig. 11

Energy rate $\dot{E}$ as a function of mean presynaptic firing rate: comparison of exact results with mean-field. Results are for $\sigma_w = 0.02$ nS (upper panel), $\sigma_w = 0.1$ nS (middle panel), and $\sigma_w = 0.5$ nS (lower panel). Solid lines correspond to exact numerical results for the whole system of N synapses obtained from Eq. (79). Dashed lines correspond to the mean-field approach (Eqs. 14-15). The best agreement between exact and mean-field results is for the intermediate σ_w, and not too small f_o. Note two peaks in the mean-field result for $\dot{E}$ (corresponding to mono $\leftrightarrow$ bistability transitions) for larger noise σ_w, which are absent in the exact results. These peaks are the artifacts of the approximation method in the mean-field. All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_f = 10$ Hz.

Fig. 12

Energy rate $\dot{E}$ as a function of standard deviation of the presynaptic firing rate: comparison of exact results with mean-field. Results are for $f_o = 1$ Hz (upper panel), and $f_o = 10$ Hz (lower panel). Solid lines correspond to exact numerical results obtained from Eq. (79), and dashed lines correspond to the mean-field results obtained from Eqs. (14-15). A better agreement between exact and mean-field results is obtained for the higher f_o. All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_w = 0.1$ nS.

Fig. 13

Energy rate $\dot{E}$ as a function of average synaptic current $\langle v \rangle$: exact numerical results. Note a sharp increase in $\dot{E}$ for larger $\langle v \rangle$. Energy rate is calculated from Eq. (79), and $\langle v \rangle$ is calculated from Eqs. (1-2). All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$,

and $\sigma_f = 10$ Hz.

Fig. 14

Nonlinear dependence of energy rate on presynaptic firing rate and average synaptic current: exact numerical results. (A) The ratio $\dot{E}/f_o$ as a function of f_o. (B) The ratio $\dot{E}/\langle v \rangle$ as a function of $\langle v \rangle$. The energy rate $\dot{E}$ depends nonlinearly both on f_o and v, but the dependence on the synaptic current $\langle v \rangle$ is more steep. Energy rate is calculated from Eq. (79), and $\langle v \rangle$ is calculated from Eqs. (1-2). All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_f = 10$ Hz.

Fig. 15

Energy cost of synaptic plasticity as a fraction of neuron's electric energy cost in the cerebral cortex of macaque monkey. The ratio of the total energy rate used by plastic synapses $N\dot{E}$ (chemical energy) to neuron's energy rate $\dot{E}_n$ (electric energy used mainly for fast synaptic transmission and neural spiking) as a function of presynaptic firing rate f_o for different levels of synaptic noise σ_w, and different regions of the macaque cortex (visual and frontal). Note that the energy contribution of plastic synapses to the neuron's energy budget depends strongly on the synaptic noise level. For weak and intermediate noise this contribution is mostly marginal. For very large noise ($\sigma_w = 2$ nS) it can be substantial, but only for very large firing rates. The neuron's energy rate $\dot{E}_n$ was computed using Eq. (63), while the plasticity energy rate of all synapses $N\dot{E}$ was computed from Eqs. (14-15). All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_f = 10$ Hz.

Fig. 16.

Similar as in Fig. 15, but for human cerebral cortex. Overall, the ratio of the energy rate of plastic synapses to the neuron's electric energy rate for human cortex is very similar to the one for macaque cortex.

Fig. 17.

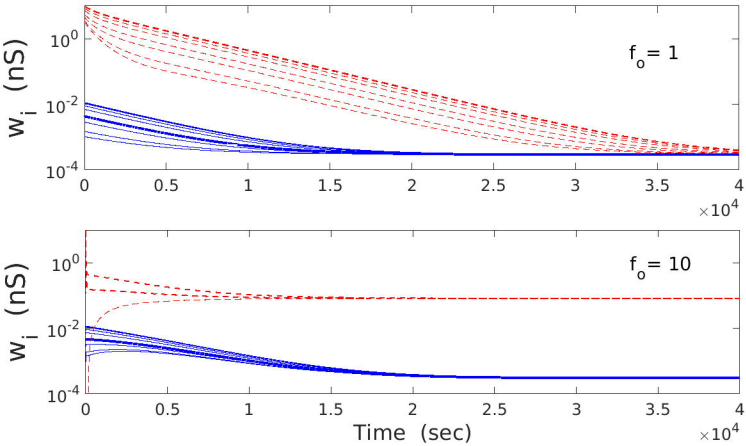
Comparison of synaptic plasticity energy rate with accuracy and lifetime of stored information as a function of presynaptic firing rate. Dependence of p_d, $\dot{E}$, I_F, and T_m on firing rate f_o. Fisher information I_F and memory lifetime T_m have large peaks at the onset of bistability, whereas the synaptic energy rate $\dot{E}$ increases only mildly. Beyond the transition point to bistability, I_F and T_m exhibit a different dependence on f_o than $\dot{E}$. The former two quantities decrease while the latter increases with f_o. In the dependence of $\dot{E}$ on f_o, the solid line corresponds to the mean-field approximation, and the dashed line to the exact numerical result. All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_w = 0.1$ nS.

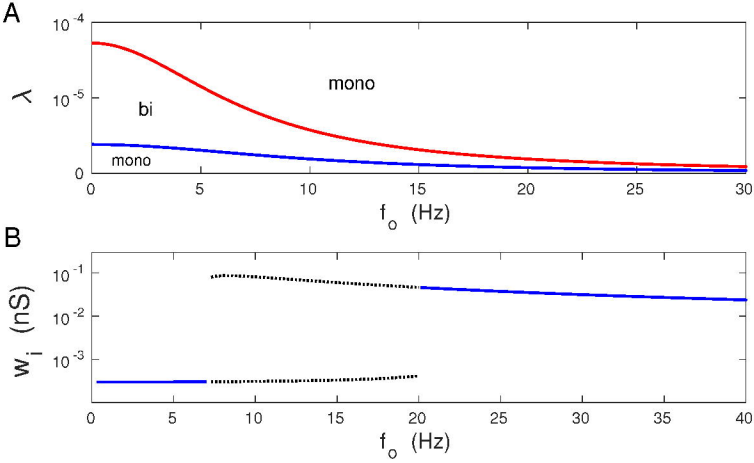
Fig. 18.

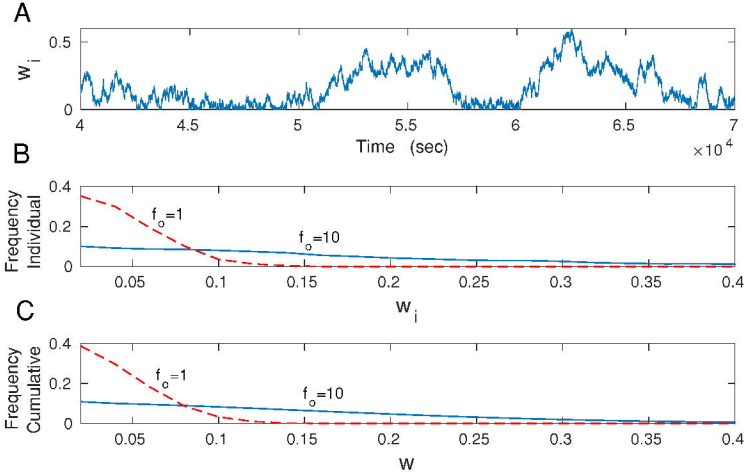
Gains in information accuracy and lifetime per synaptic energy used as functions of presynaptic firing rate. Fraction of synapses in the down state p_d (upper panel), the ratio $I_F/\dot{E}$ (middle panel) and $T_m/\dot{E}$ (lower panel) as functions of f_o. The biggest gains in information and memory lifetime are at the transition point to the bistability. Solid lines correspond to the energy rate calculated in the mean-field, and dashed lines to the energy rate calculated numerically. All plots are for $\lambda = 9 \cdot 10^{-7}$, $\kappa = 0.001$, and $\sigma_w = 0.1$ nS.

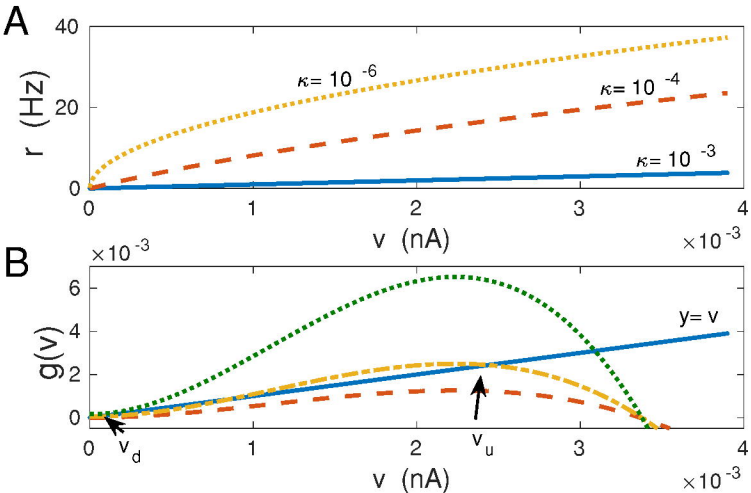
Fig. 19.

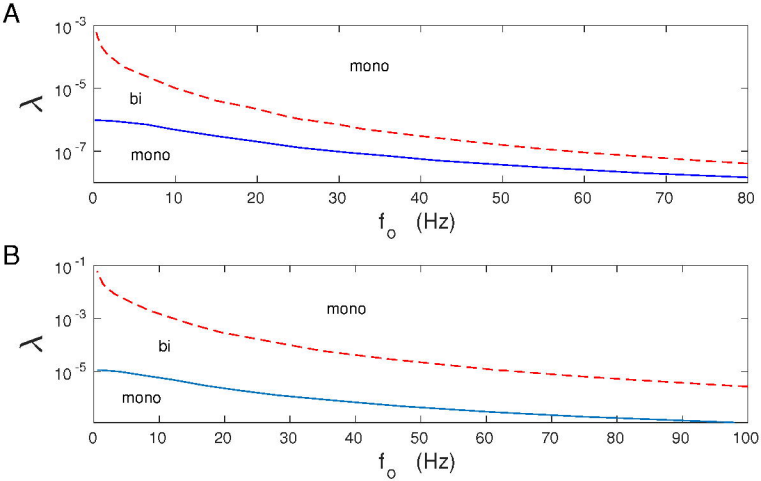
Gains in information accuracy and lifetime per synaptic energy used as functions of plasticity amplitude. Fraction of synapses in the down state p_d (upper panel), the ratio $I_F/\dot{E}$ (middle panel) and $T_m/\dot{E}$ (lower panel) as functions of λ. Note the sharp peaks for $I_F/\dot{E}$ at the transition point from mono- to bi-stability, and for $T_m/\dot{E}$ at the transition point from bi- to mono-stability. Solid lines correspond to the energy rate calculated in the mean-field approach. All plots are for $f_o = 10$ Hz, $\kappa = 0.001$, and $\sigma_w = 0.1$ nS.

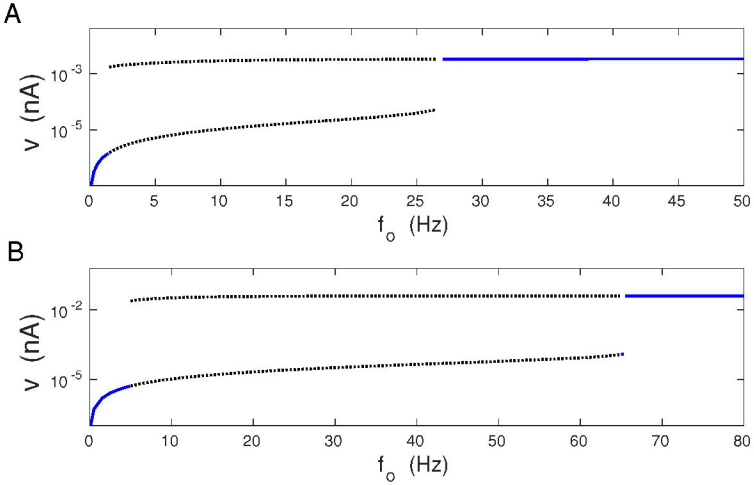

$f_o = 1$
w_i (nS)
10^0
10^{-2}
10^{-4}
0
0.5
1
1.5
2
2.5
3
3.5
4
$\times 10^4$
$f_o = 10$
w_i (nS)
10^0
10^{-2}
10^{-4}
0
0.5
1
1.5
2
2.5
3
3.5
4
$\times 10^4$
Time (sec)

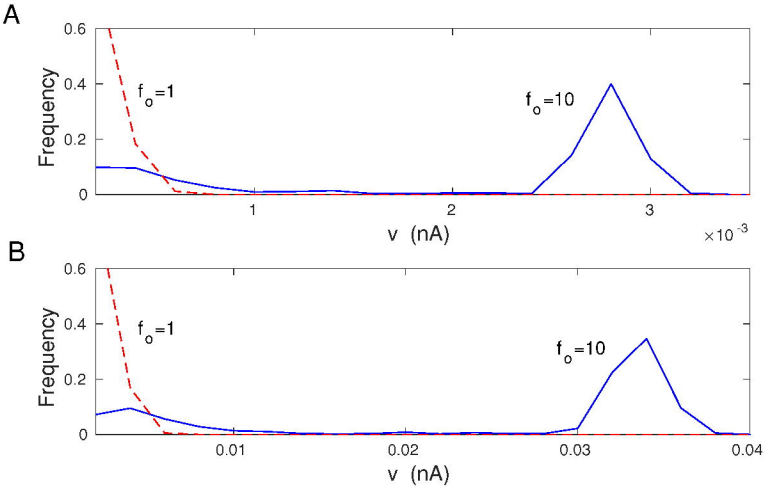

A
10^{-4}
10^{-5}
0
λ
mono
bi
mono
0
5
10
15
20
25
30
f_o (Hz)
B
10^{-1}
10^{-2}
10^{-3}
w_i (nS)
0
5
10
15
20
25
30
35
40
f_o (Hz)

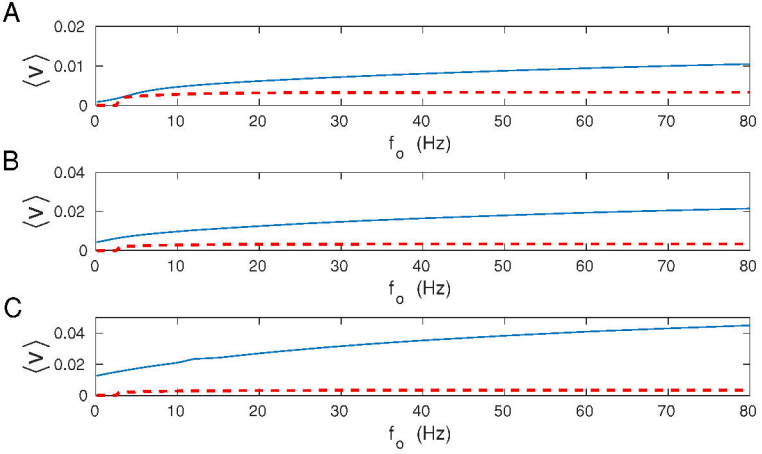

A
w_i
0.5
0
4
4.5
5
5.5
6
6.5
7
Time (sec)
$\times 10^4$
B
Frequency Individual
0.4
0.2
0
f_o=1
f_o=10
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
w_i
C
Frequency Cumulative
0.4
0.2
0
f_o=1
f_o=10
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
w

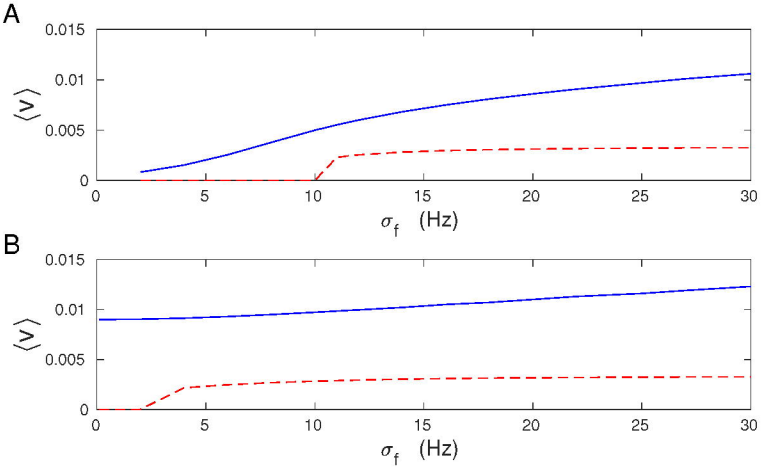

A
40
20
0
r (Hz)
κ= 10 -6
κ= 10 -4
κ= 10 -3
0
1
2
3
4
v (nA)
×10 -3
B
×10 -3
6
4
2
0
g(v)
y= v
0
v d
1
2
v u
3
4
v (nA)
×10 -3

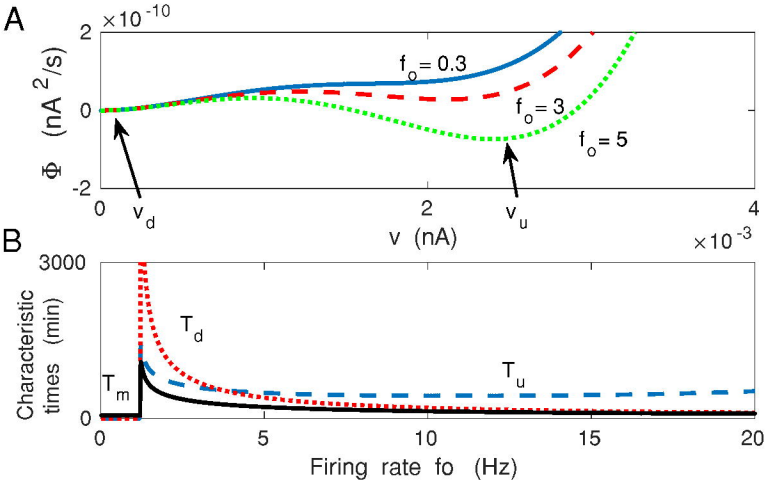

A
10^{-3}
10^{-5}
10^{-7}
λ
mono
bi
mono
0
10
20
30
40
50
60
70
80
f_o (Hz)
B
10^{-1}
10^{-3}
10^{-5}
λ
mono
bi
mono
0
10
20
30
40
50
60
70
80
90
100
f_o (Hz)

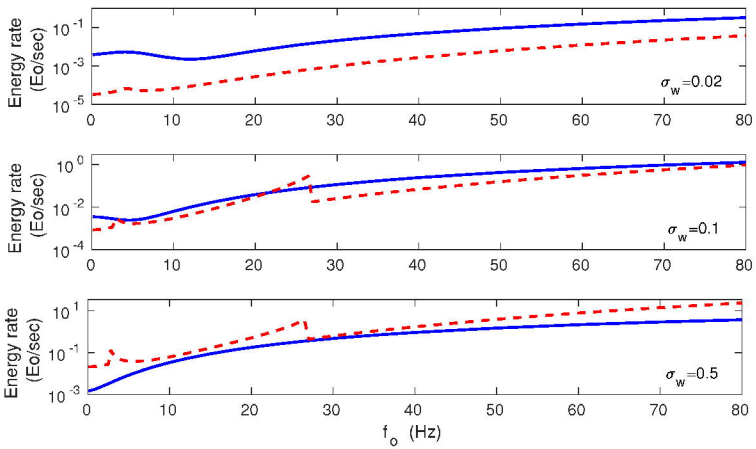

A
v (nA)
10^-3
10^-5
0
5
10
15
20
25
30
35
40
45
50
f_o (Hz)
B
v (nA)
10^-2
10^-5
0
10
20
30
40
50
60
70
80
f_o (Hz)

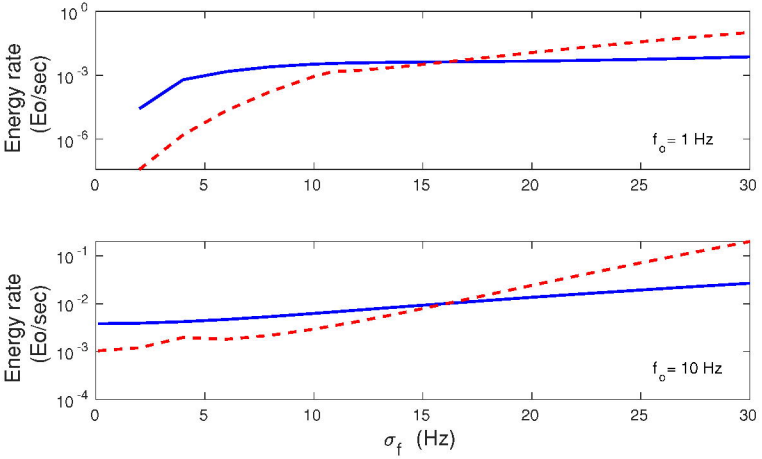

A
0.6
0.4
0.2
0
Frequency
f_o=1
f_o=10
1
2
3
v (nA)
×10^-3
B
0.6
0.4
0.2
0
Frequency
f_o=1
f_o=10
0.01
0.02
0.03
0.04
v (nA)

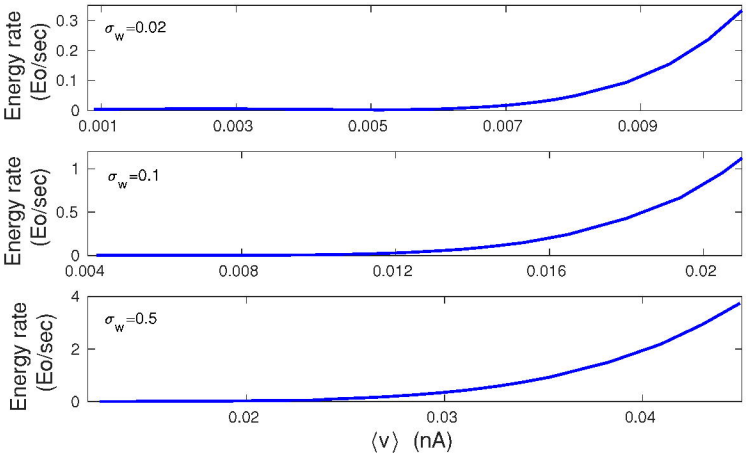

A
0.02
0.01
0
⟨v⟩
0
10
20
30
40
50
60
70
80
f_o (Hz)
B
0.04
0.02
0
⟨v⟩
0
10
20
30
40
50
60
70
80
f_o (Hz)
C
0.04
0.02
0
⟨v⟩
0
10
20
30
40
50
60
70
80
f_o (Hz)

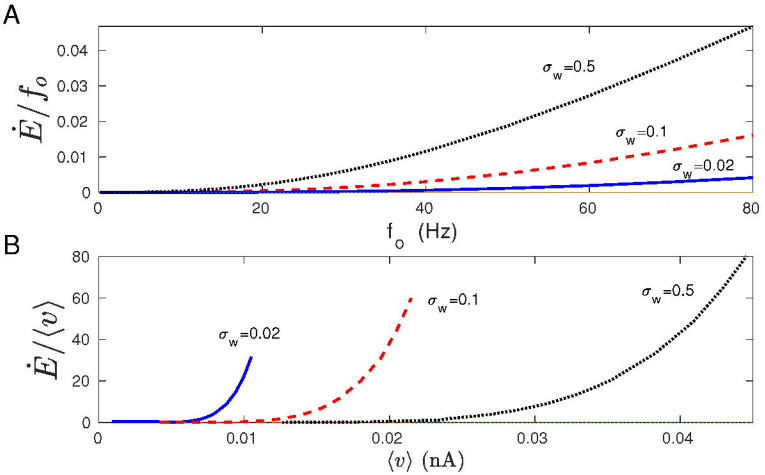

A
0.015
0.01
0.005
0
⟨v⟩
0
5
10
15
20
25
30
σf (Hz)
B
0.015
0.01
0.005
0
⟨v⟩
0
5
10
15
20
25
30
σf (Hz)

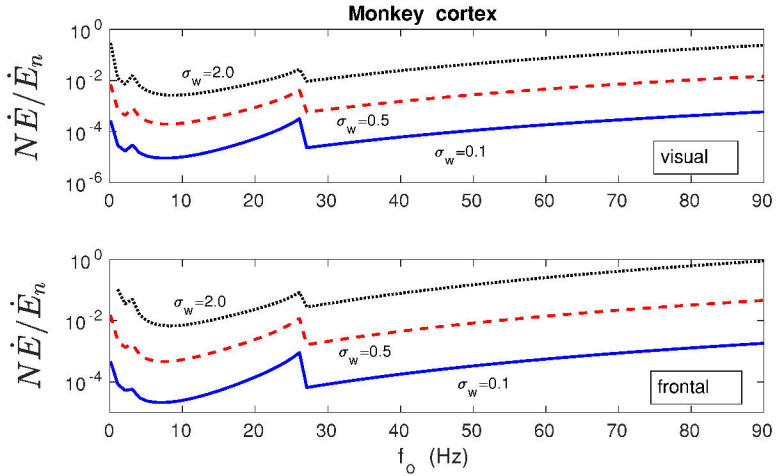

A
×10^-10
2
0
-2
Φ (nA^2/s)
f_o = 0.3
f_o = 3
f_o = 5
0
2
4
v_d
v_u
v (nA)
×10^-3
B
3000
0
Characteristic times (min)
T_d
T_u
T_m
0
5
10
15
20
Firing rate fo (Hz)

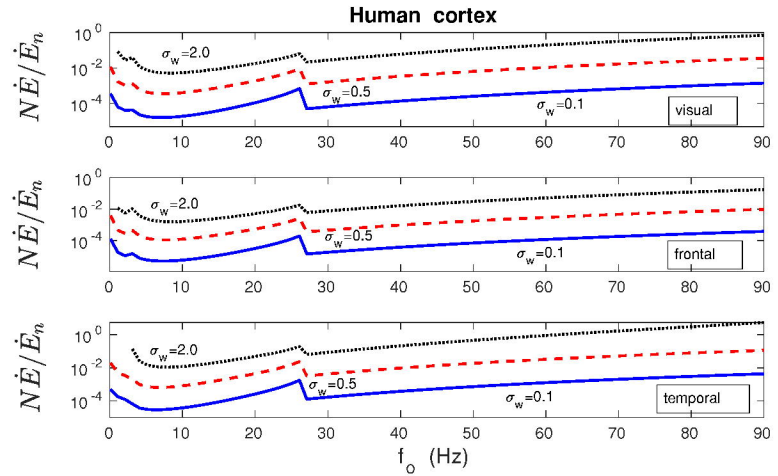

Energy rate
(Eo/sec)
10^{-1}
10^{-3}
10^{-5}
σ_w=0.02
0
10
20
30
40
50
60
70
80
Energy rate
(Eo/sec)
10^{0}
10^{-2}
10^{-4}
σ_w=0.1
Energy rate
(Eo/sec)
10^{1}
10^{-1}
10^{-3}
σ_w=0.5
f_o (Hz)

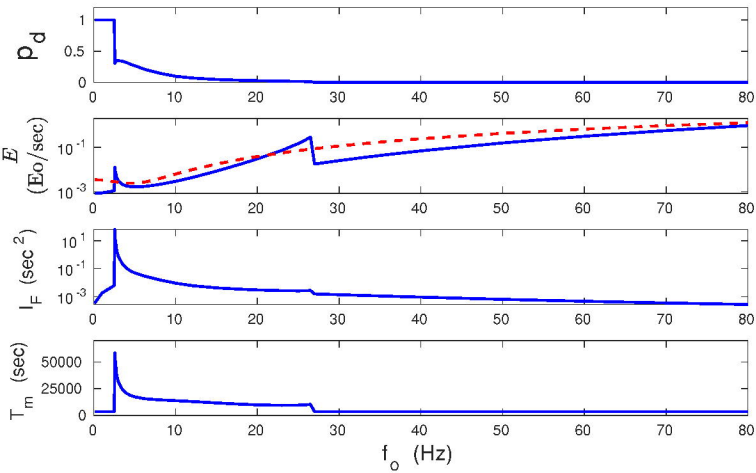

Energy rate
(Eo/sec)
10^{0}
10^{-3}
10^{-6}
f_o= 1 Hz
0
5
10
15
20
25
30
Energy rate
(Eo/sec)
10^{-1}
10^{-2}
10^{-3}
10^{-4}
f_o= 10 Hz
0
5
10
15
20
25
30
σ_f (Hz)

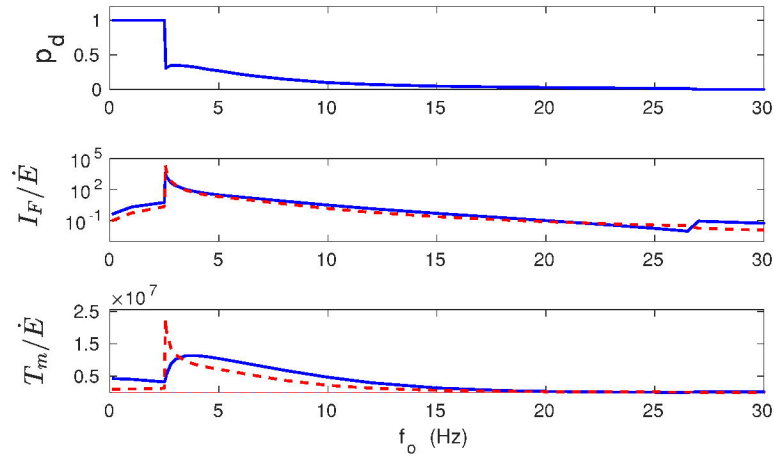

Energy rate
(Eo/sec)
σ_w=0.02
0.3
0.2
0.1
0
0.001
0.003
0.005
0.007
0.009
Energy rate
(Eo/sec)
σ_w=0.1
1
0.5
0
0.004
0.008
0.012
0.016
0.02
Energy rate
(Eo/sec)
σ_w=0.5
4
2
0
0.02
0.03
0.04
⟨v⟩ (nA)

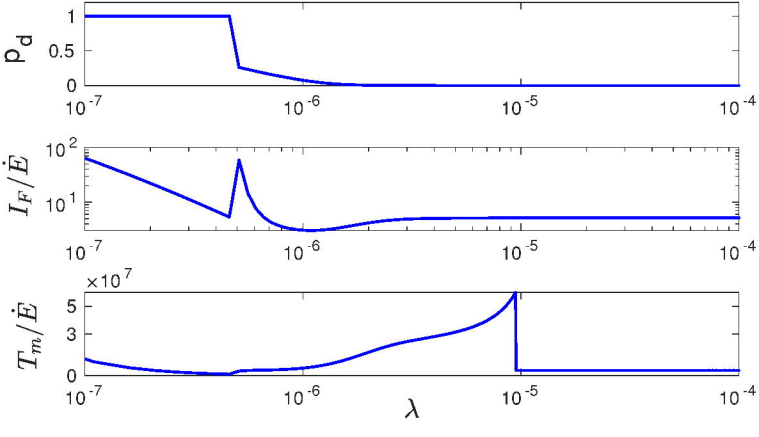

A
$\dot{E}/f_o$
0.04
0.03
0.02
0.01
0
σ_w=0.5
σ_w=0.1
σ_w=0.02
0
20
40
60
80
f_o (Hz)
B
$\dot{E}/\langle v \rangle$
80
60
40
20
0
σ_w=0.02
σ_w=0.1
σ_w=0.5
0
0.01
0.02
0.03
0.04
$\langle v \rangle$ (nA)


Monkey cortex
$N\dot{E}/\dot{E}_n$
10^0
10^{-2}
10^{-4}
10^{-6}
σ_w=2.0
σ_w=0.5
σ_w=0.1
visual
0
10
20
30
40
50
60
70
80
90
$N\dot{E}/\dot{E}_n$
10^0
10^{-2}
10^{-4}
σ_w=2.0
σ_w=0.5
σ_w=0.1
frontal
0
10
20
30
40
50
60
70
80
90
f_o (Hz)


Human cortex
$N\dot{E}/\dot{E}_n$
10^0
10^{-2}
10^{-4}
σ_w=2.0
σ_w=0.5
σ_w=0.1
visual
frontal
temporal
0
10
20
30
40
50
60
70
80
90
f_o (Hz)


p_d
1
0.5
0
$\dot{E}$ (Eo/sec)
10^{-1}
10^{-3}
I_F (sec^2)
10^{1}
10^{-1}
10^{-3}
T_m (sec)
50000
25000
0
0
10
20
30
40
50
60
70
80
f_o (Hz)


p_d
1
0.5
0
0
5
10
15
20
25
30
$I_F/\dot{E}$
10^5
10^2
10^{-1}
$T_m/\dot{E}$
$\times 10^7$
2.5
1.5
0.5
f_o (Hz)


p_d
1
0.5
0
10^{-7}
10^{-6}
10^{-5}
10^{-4}
$I_F/\dot{E}$
10^{2}
10^{1}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
$\times 10^{7}$
$T_m/\dot{E}$
5
3
0
10^{-7}
10^{-6}
10^{-5}
10^{-4}
λ