Cooperation across timescales between and Hebbian and homeostatic plasticity

Friedemann Zenke\(^1\) and Wulfram Gerstner\(^2\)

April 13, 2016

1) Dept. of Applied Physics, Stanford University, Stanford, CA
2) Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne EPFL

Abstract

We review a body of theoretical and experimental research on the interactions of homeostatic and Hebbian plasticity, starting from a puzzling observation: While homeostasis of synapses found in experiments is slow, homeostasis of synapses in most mathematical models is rapid, or even instantaneous. Even worse, most existing plasticity models cannot maintain stability in simulated networks with the slow homeostatic plasticity reported in experiments. To solve this paradox, we suggest that there are both fast and slow forms of homeostatic plasticity with distinct functional roles. While fast homeostatic control mechanisms interacting with Hebbian plasticity render synaptic plasticity intrinsically stable, slower forms of homeostatic plasticity are important for fine-tuning neural circuits. Taken together we suggest that learning and memory relies on an intricate interplay of diverse plasticity mechanisms on different timescales which jointly ensure stability and plasticity of neural circuits.

Introduction

Homeostasis refers to a group of physiological processes at different spatial and temporal scales that help to maintain the body, its organs, the brain, or even individual neurons in the brain in a target regime where they function optimally. A well-known example is the homeostatic regulation of body temperature in mammals, maintained at about 37 degrees Celsius independent of weather condition and air temperature. In neuroscience, homeostasis or homeostatic plasticity often refers to the homeostatic control of neural firing rates. In a classic experiment, cultured neurons that normally fire at, say 5Hz, change their firing rate after a modulation of the chemical conditions in the culture, but eventually return to their target rate of 5Hz during the following 24 hours (Turrigiano and Nelson, 2000). Thus, the experimentally best-studied form of synaptic homeostasis happens on a slow timescale of hours or days. This slow form of synaptic homeostasis manifests itself as a
rescaling of the strength of all synapses onto the same neuron by a fixed fraction, for instance 0.78, a phenomenon called “synaptic scaling” (Turrigiano et al., 1998).

Mathematical models of neural networks also make use of homeostatic regulation to avoid runaway of the strength of synaptic weights or over-excitation of large parts of a neural network. In many mathematical models, homeostasis of synaptic weights is idealized by an explicit normalization of weights: if the weight (or efficacy) of one synaptic connection increases, weights of other connections onto the same neuron are algorithmically decreased so as to keep the total input in the target regime (Miller and MacKay, 1994). Algorithmic normalization of synaptic strength can be interpreted as an extremely fast homeostatic regulation mechanism. At a first glance, the result of multiplicative normalization (Miller and MacKay, 1994) is identical to “synaptic scaling” introduced above. Yet, it is fundamentally different from the experimentally studied form of synaptic homeostasis mentioned above, because of the vastly different timescales. While rescaling in the model is instantaneous, in biology the effects of synaptic scaling manifest themselves only after hours. Interestingly, researchers who have tested slow forms of homeostasis in mathematical models of plastic neural networks failed to stabilize the network activity.

Thus, we are confronted with a dilemma: Theorists need fast synaptic homeostasis while experimentalist have found homeostasis that is much slower. In this review we try to answer the following questions: Why do we need rapid forms of homeostatic synaptic plasticity? How fast does a rapid homeostatic mechanism have to be, hours, minutes, seconds or less? Furthermore, what are the functional consequences of fast homeostatic control? Can the combination of fast homeostasis with Hebbian learning lead to stable memory formation? And finally, if rapid homeostasis is a requirement, what is the role of slower forms of homeostatic plasticity?

Conceptually, automatic control of a technical system is designed to keep the system in the regime where it works well, which is called the “working point” or the “set point” of the system. Analogously, homeostatic regulation of synapses is important to maintain the neural networks of the brain in a regime where they function well. However, in specific situations our brain needs to change. For example, when we exit the subway at an unknown station, when we are confronted with a new theoretical concept, or when we learn how to play tennis, we need to memorize the novel environment, the new concept, or the unusual movement that we previously did not know or master. In these situations, the neural networks in the brain must be able to develop an additional desired “set point”, ideally without loosing the older ones.

In neuroscience it is widely accepted that learning observed in humans or animals at the behavioral level corresponds, at the level of neural networks, to changes in the synaptic connections between neurons (Morris et al., 1986; Martin et al., 2000). Thus, learning requires synapses to change. A strong form of synaptic homeostasis, however, keeps neurons and synapses always in the “old” regime, i.e. the one before learning occurred. In other words, strong homeostasis prevents learning. A weak or slow form of homeostasis, on the other hand, may not be sufficient to control the changes induced by memory formation. Before we can solve the riddle of stability during mem-
ory formation, and before we can answer the above questions, we need to classify known forms of synaptic plasticity.

Classification of Synaptic Plasticity

Synaptic plasticity can be classified using different criteria. We consider three of these: Functional relevance, synaptic specificity, and timescale of effect duration.

(i) Classification by potential function: Memory formation or homeostasis

Functionally, synaptic plasticity can be classified as useful either for homeostatic regulation or for memory formation and action learning. Many experiments on synaptic plasticity were based on Hebb’s postulate which suggests that synaptic changes, caused by the joint activity of pre- and postsynaptic neurons, should be useful for the formation of memories (Hebb, 1949). Inspired by Hebb’s postulate, classical stimulation protocols for long-term potentiation (Bliss and Lomo, 1973; Malenka and Nicoll, 1999; Lisman, 2003), long-term depression (Lynch et al., 1977; Levy and Steward, 1983), or spike-timing dependent plasticity (Markram et al., 1997; Bi and Poo, 1998; Sjöström et al., 2001), combine the activation of a presynaptic neuron, or a presynaptic pathway, with an activation, depolarization, or chemical manipulation of the postsynaptic neurons, to induce synaptic changes.

In parallel, theoreticians have developed synaptic plasticity rules (Willshaw and Von Der Malsburg, 1976; Bienenstock et al., 1982; Kempter et al., 1999; Song et al., 2000; Pfister and Gerstner, 2006; Clopath and Gerstner, 2010; Brito and Gerstner, 2016; Shouval et al., 2002; Graupner and Brunel, 2012), in part inspired by experimental data. Generically, in plasticity rules of computational neuroscience the change of a synapse from a presynaptic neuron $j$ to a neuron $i$ is described as

$$\frac{d}{dt}w_{ij} = F(w_{ij}, post_i, pre_j)$$

where $w_{ij}$ is the momentary “weight” of a synaptic connection, $post_i$ describes the state of the postsynaptic neuron (e.g. its membrane potential, calcium concentration, spike times, or firing rate) and $pre_j$ is the activity of the presynaptic neuron (Brown et al., 1991; Morrison et al., 2008; Gerstner et al., 2014). A network of synaptic connections with weights consistent with Hebb’s postulate can form associative memory models (Willshaw et al., 1969; Little and Shaw, 1978; Hopfield, 1982; Amit et al., 1985; Amit and Brunel, 1997), also called attractor neural networks (Amit, 1989). However, in all the cited theoretical studies, learning is supposed to have happened somewhere in the past, while the retrieval of previously learned memories is studied under the assumption of fixed synaptic weights. The reason is that a simple implementation of Hebb’s postulate for memory formation is not sufficient to achieve stable learning (Rochester et al., 1956).

Without clever homeostatic control of synaptic plasticity, neural networks with a Hebbian rule always move to an undesired state where all synapses go to their maximally allowed values and
all neurons fire beyond control — or to the other extreme where all activity dies out. Therefore, homeostasis in the sense of keeping the network in a desired regime, is a second and necessary function of synaptic plasticity. Candidate plasticity mechanisms for the homeostatic regulation of network function by synaptic plasticity include slow synaptic scaling (Turrigiano et al., 1998; Turrigiano and Nelson, 2000), rapid renormalization of weights (Miller and MacKay, 1994), or other mechanisms discussed below (Zenke et al., 2013, 2015; Chistiakova et al., 2015); but also synaptic plasticity of inhibitory interneurons (Vogels et al., 2011, 2013) or intrinsic plasticity of neurons (Daoudal and Debanne, 2003).

(ii) Classification by specificity of synaptic changes: Homosynaptic or heterosynaptic

In some of the classic induction protocols of long-term potentiation and depression, mentioned above, a presynaptic pathway (or presynaptic neuron) was stimulated together with an activation of the postsynaptic neuron. If a change was observed in those synapses that were stimulated during the induction protocol, the plasticity was classified as synapse-specific or “homosynaptic”. If a change was observed in other (unstimulated) synapses onto the same postsynaptic neuron, the plasticity was classified as unspecific or “heterosynaptic” (Lynch et al., 1977; Abraham and Goddard, 1983; Brown and Chattarji, 1994). Heterosynaptic plasticity has moved back in the focus recently, because of its potential role in homeostatic regulation of plasticity (Royer and Parè, 2003; Chen et al., 2013; Chistiakova et al., 2014, 2015; Zenke et al., 2015).

In the framework of Eq. (1), the change of a synapse may depend on the activity of the presynaptic neuron, the state of the postsynaptic neuron — or on both — as well as on the momentary value of the synaptic weight. Since we do not know the function \( F \) on the right-and-side of Eq. (1), we can imagine that it may contain different combinations of pre- and postsynaptic variables, such as

\[
F(w_{ij}, post_i, pre_j) = a_0(w_{ij}) + f_1(w_{ij}, post_i) + f_2(w_{ij}, pre_j) + a_1(w_{ij}) H(post_i, pre_j) \quad (2)
\]

In this picture, the term \( a_0(w_{ij}) \) would represent a drift of the synaptic strength that does not depend on the input or the state of the postsynaptic neuron, but only on the momentary value of the weight \( w_{ij} \) itself. More interesting is the term \( f_1(w_{ij}, post_i) \) which depends on the postsynaptic activity, but not on the presynaptic activity. Suppose the postsynaptic neuron went through a phase of rapid firing. Such a term will then induce changes that are not synapse-specific and would therefore be classified as heterosynaptic. Conversely, the term \( f_2(w_{ij}, pre_j) \) depends on the activity of the presynaptic pathway, but not on the state of the postsynaptic neuron. In the following we will call such a term transmitter-induced. If a presynaptic neuron is weakly stimulated for some time, the change caused by the term \( f_2(w_{ij}, pre_j) \) can be classified as homosynaptic, but not Hebbian. In equation Eq.(2) only the last term is able to account for changes that depend on the joint activity of pre- and postsynaptic neurons. This term is therefore homosynaptic and Hebbian, hence our choice of letter \( H \). The multiplicative factor \( a_1(w_{ij}) \) in front of \( H \) highlights that the amount of synaptic
plasticity may depend on the momentary state of the synaptic weight. For example, it might be
difficult or impossible to further increase a synapse that is already very strong, simply because strong
synapses require a big spine and therefore a lot of resources. The theoretical literature implements
resource limitations by keeping the weights $w_{ij}$ below a maximum value $w_{\text{max}}$. Depending on the
implementation details, the limit is called a “hard bound” or “soft bound”.

As an aside, we note that the term “heterosynaptic plasticity” is sometimes also used for synaptic
changes that are visible at the connection from a presynaptic neuron $j$ to a postsynaptic neuron $i$,
but induced by the activation of a third, typically modulatory neuron (Bailey et al., 2000). However,
in the following we do not consider this possibility.

Equations (1) and (2) that describe the change of synaptic weights as a function of neuronal
activity are called learning rules in the theoretical literature. For the rest of the paper, Eqs. (1) and
(2) provide the framework in which we analyze and compare different learning rules with each other.

(iii) Classification by duration of effects: Short-term versus long-term plasticity versus
consolidation

Synaptic changes induced by a sequence of four presynaptic spikes in rapid sequence typically decay
within a few hundred milliseconds (Abbott et al., 1997; Tsodyks and Markram, 1997; Tsodyks et al.,
1998) and are called short-term plasticity. The rapid decay of plasticity outcomes implies that the
changes are not useful for memory formation, but more likely involved in gain control (Abbott et al.,
1997), which implies a homeostatic role.

Synaptic changes induced by classic induction protocols (Bliss and Lomo, 1973; Malenka and
Nicoll, 1999; Lisman, 2003), however, last for several hours and are therefore potentially useful for
memory formation. We emphasize that the induction protocol itself often lasts less than a minute.
In other words, induction of synaptic plasticity is fast, but the changes persist for a long time, and
this is why the phenomenon is called long-term potentiation or long-term depression, which we will
abbreviate in the following as LTP or LTD.

Under suitable conditions the changes induced by a protocol of LTP or LTD are further consol-
idated after about an hour (Frey and Morris, 1998; Reymann and Frey, 2007; Redondo and Morris,
2011). In the rest of the paper we mainly focus on the phase of plasticity induction and neglect
consolidation and maintenance.

To summarize this part, we may state that the classification of synaptic plasticity is possible in
many different ways. A certain aspect of plasticity (or a certain mathematical term in a model)
can be treated at the same time as Hebbian, homosynaptic, and memory-building while another
term could be heterosynaptic and homeostatic, and a third term could be transmitter-induced,
homosynaptic, but non-Hebbian. Each of the terms can lead to a change that persists for a long
time or that decays within a few seconds.

Moreover, each of the terms could be strong so that its contribution is easily verified in exper-
iments; or it could be weak so that only a precise experiments with multiple repetitions would be able to isolate a given term. The problem of the effect size is intimately related to the question of the timescale for induction of plasticity, which is an important topic in itself to be discussed next.

Why do we need rapid forms of homeostatic synaptic plasticity?

With the definitions and distinctions made in the previous section, we are now ready to address the first of the questions posed in the introduction: Why do we need a fast form of homeostatic synaptic plasticity?

Intuitively, synaptic plasticity that is useful for memory formation must be sensitive to the present activation pattern of the pre- and postsynaptic neuron. Following Hebb's idea of learning and cell assembly formation, the synaptic changes should make the same activation pattern more likely to re-appear in the future, to allow contents from memory to be retrieved. However, the re-appearance of the same pattern will induce further synaptic plasticity. This leads to a positive feedback loop that can quickly run out of control. Anybody who was sitting in the audience when the positive feedback loop between the speaker's microphone and the loudspeaker resulted in an unpleasant shriek, knows what this means. Engineers have realized that positive feedback loops can only be stabilized by rapid control mechanisms. Similarly, the interaction between the firing activity in a neural network with synaptic plasticity for memory formation can lose stability, unless synaptic plasticity is controlled by rapid homeostasis.

While the plasticity rules derived from experiments look at most at one or two stimulation pathways at a time, neurons embedded in a network can receive inputs from thousands of different synapses. In this section, we will outline a simple, yet powerful framework which enables us to make quantitative predictions about the stability in large networks of spiking neurons. Synapses in the network will exhibit different forms of plasticity which serve specific functions for both memory formation and homeostasis.

Before we can give a precise answer to the question of “why do we need a fast form of homeostatic synaptic plasticity”, we have to consider what it means to be “fast” or “slow”. Hebbian synaptic plasticity manifests itself tens of seconds or a few minutes after plasticity induction (Petersen et al., 1998; O’Connor et al., 2005). This sets a first timescale of plasticity induction. Intuitively, we will call a homeostatic mechanism fast if it happens on the same timescale as Hebbian plasticity induction. It will be considered as slow, if it is much slower than induction of LTP, i.e., if it manifests itself over hours or days.

In the previous paragraph, we used the term “timescale”. To make the intuitive notion of timescale more precise, we need a short mathematical digression. The reader who is familiar with differential equations or dynamical systems can skip the next subsection.
Mathematical digression: The notion of a timescale

To formally introduce the notion of a timescale, we consider a linear system defined by the differential equation

$$\tau \frac{dy}{dt} = -(y - \theta)$$  \hspace{1cm} (3)

where $\theta$ and $\tau$ are two fixed positive parameters. The solution $y(t)$ of the differential equation relaxes exponentially to the value $y = \theta$, on the timescale $\tau$. Practically speaking that means that no matter how we initialize $y(t)$ at time $t = 0$, if we wait for a time $3\tau$, the difference $y(t) - \theta$ will have decreased by 95%. The exponential evolution is typical for a linear system. Moreover, the equation has a nice scaling behavior: If we know the solution for a parameter setting $\tau = 1$, we can get the solution for the parameter setting $\tau = 5$ by multiplying all times by a factor of five. This is the reason why we speak of $\tau$ as the timescale of the system.

The fact that a constant value of $y(t) = \theta$ for all $t$ is a solution means that $\theta$ is a fixed point of the system. To check this, remember that the derivative of a constant is zero (so that the left-hand side of Eq. (3) vanishes). And for $y = \theta$ the right-hand side is obviously zero, too. Moreover, this fixed point is stable because from any initial condition $y(t)$ will converge towards $\theta$.

Let us now consider the case where $\tau$ is negative. In this case, $y(t) = \theta$ is still a fixed point, but if we start with a value $y(t) > \theta$, then $y(t)$ will explode exponentially fast to large positive values while for an initial value $y(t) < \theta$ it will explode exponentially fast to large negative values. The timescale of explosion is again given by $\tau$.

However, our intuition for timescales breaks down for nonlinear systems. Therefore, when we consider a nonlinear system $\tau \frac{dy}{dt} = f(y)$, it does not suffice for $\tau$ to be large to be able to talk about a slow or fast timescale. We have to be specific about the behavior of $f$ in the regime that we are interested in. The mathematical trick to do this is to look for a fixed point of the equation, that is a value of $y$ with $f(y) = 0$. Suppose $y = y_0$ is a fixed point. We then study the derivative $df/dy$ at $y_0$. Let us denote this derivative by $f'$. In the neighborhood of $y_0$ (and only there!) the nonlinear equation is well approximated by a linear equation $\tau \frac{dy}{dt} = (y - y_0) f'$. Division by $f'$ brings $f'$ to the other side of the equation and enables us to identify the effective timescale $\tilde{\tau} = -\tau/f'$. If you are in doubt, compare your result with equation (3).

Another example of a linear system is a low-pass filter. Suppose we pass our variable $y(t)$ through a low-pass filter with time constant $\tau_d$ to yield

$$\tilde{y}(t) = \int_{-\infty}^{t} \exp\left[\frac{(t - t')}{\tau_d}\right] y(t') \, dt'.$$  \hspace{1cm} (4)

By taking the derivative on both sides of Eq. (4), we find that the low-pass filter $\tilde{y}$ is the solution of a differential equation $d\tilde{y}(t)/dt = -\tilde{y}(t)/\tau_d + y(t)$. This equation is similar to equation (3), if we replace the constant target value $\theta$ by a time-dependent target $y(t)/\tau_d$. 
Figure 1: **Most plasticity models can reproduce the notion of a plasticity threshold reported in experiments.** (a) The change in synaptic efficacy in many plasticity models is a function of variables related to postsynaptic activation. (b) Schematic illustration of the action of the homeostatic moving threshold in the BCM-model (Bienenstock et al., 1982). (c) Hypothetical evolution of synaptic efficacy as a function of time during an LTP induction experiment. The purple, linear time course is a common assumption underlying many plasticity models. However, since the data during plasticity induction are typically unobservable, many time courses and thus plasticity modes are compatible with the data.

**The induction timescale of Hebbian plasticity**

Because both learning rules and neurons are typically nonlinear, the mathematical definition of a timescale for Hebbian plasticity requires the existence of a fixed point. The reason is that for a nonlinear system we can define a timescale only in the vicinity of a fixed point, as we have seen above.

Even if a plasticity rule in isolation looks linear, the combination of plasticity with the neuronal dynamics typically makes the network as a whole nonlinear. Consider for instance a single neuron $i$ that is driven by $N$ inputs arriving at synapses $w_{ij}$ for $1 \leq j \leq N$. In a rate model, the state of the postsynaptic neuron is characterized by its firing rate $y_i = g \left( \sum_{j=1}^{N} w_{ij} x_j \right)$ where $x_j$ is the firing rate of the presynaptic neuron with index $j$. The function $g$ denotes the frequency-current relation of a single neuron and we assume that it is monotonically increasing, i.e., if we increase the input, the firing rate increases as well. In the theoretical literature $g$ is sometimes called the gain function.
of the neuron, hence our choice of letter $g$.

Let us first study a simple Hebbian learning rule

$$\frac{dw_{ij}}{dt} = \eta y_i x_j.$$  \hspace{1cm} (5)

Note that this is a special instance within the general framework of Eqs (1), and (2). To see this, consider a Hebbian term $H(\text{post}_i, \text{pre}_j) = y_i x_j$, and $a_1(w_{ij}) = \eta$ and set all other terms in Eq. (2) to zero. This learning rule is linear in $x_j$ and linear in $y_i$. However, if we insert $y_i = g\left(\sum_j w_{ij} x_j\right)$ into the learning rule, the learning dynamics become nonlinear in $x_j$. Nonlinearity implies that we cannot define a timescale of plasticity, unless we find a fixed point where $dw_{ij}/dt$ vanishes.

Are there fixed points of the dynamics? There is a fixed point if the postsynaptic or the presynaptic rate is zero. However, if the neuron is embedded in a network, it is reasonable to assume that several presynaptic neurons including the presynaptic neuron $j$ are active. Unless all weights $w_{ij}$ are zero, the postsynaptic neuron is therefore also active, and the weight $w_{ij}$ increases.

The argument can be formalized to show that $w_{ij} = 0$ is an unstable fixed point. If we increase $w_{ij}$ by just a little, the output $y_i$ also increases which increases $w_{ij}$ even further, which closes the positive feedback loop. More generally, models of Hebbian plasticity that are useful for memory formation all have an unstable fixed point. One important role of homeostatic plasticity in networks models with Hebbian plasticity is therefore to create (additional) stable fixed points for the learning dynamics as we will see later on in this section.

The simple example above only has a trivial fixed point at zero activity (zero weights). Moreover, it is lacking the notion of LTD. Plausible plasticity models have additional stationary points defined by the plasticity threshold between LTP and LTD (Fig. 1a). Inspired by experimental data (Artola et al., 1990; Dudek and Bear, 1992; Sjöström et al., 2001), the transition from LTP to LTD depends in models on the state of the postsynaptic neuron, e.g., its membrane potential, calcium level, inter spike interval (Artola and Singer, 1993; Shouval et al., 2002; Pfister and Gerstner, 2006; Clopath and Gerstner, 2010; Graupner and Brunel, 2012). As a paradigmatic example, which stands for a plethora of different plasticity rules with a postsynaptic threshold, we consider the following rate-based nonlinear Hebbian rule

$$\frac{dw_{ij}}{dt} = \eta x_j y_i (y_i - \theta)$$  \hspace{1cm} (6)

with a positive constant $\eta$. Whenever the postsynaptic firing rate $y_i$ equals the value of $\theta$, the weight $w_{ij}$ does not change. Thus this learning rule has a fixed point at the threshold $y_i = \theta$. For $y_i$ larger than $\theta$ the synaptic weights increase which corresponds to the induction of LTP in the model. For $y_i$ smaller than $\theta$ the synaptic weights decrease and LTD is induced.

Let us now embed this learning rule in a network. For the sake of simplicity we assume the gain function to be linear, $g\left(\sum_j w_{ij} x_j\right) = \sum_j w_{ij} x_j$. Moreover, we assume that (i) all weights $w_{ij}$ have the same value $w_{ij} = w$ and (ii) all $N$ neurons in the network fire with the same firing rate.
\( x_j = 1/N \), so that \( y = w \). Inserting these assumptions into Eq. (6) yields

\[
\frac{dw}{dt} = \eta w (w - \theta)
\]

which characterizes the change in neural activity \( y \) as governed by synaptic plasticity. We now linearize Expression (7) at the stationary point \( w = \theta \):

\[
\frac{dw}{dt} \approx \eta \theta (w - \theta)
\]

This is a linear differential equation with a solution \( w(t) \) that, for \( w(t) > \theta \), explodes exponentially fast on the timescale \( \tau = \frac{1}{\eta \theta} \) (cf. discussion of Eq. (3)). Thus, in this example, \( w = \theta \) is an unstable fixed point and the linearization procedure has enabled us to identify the timescale of plasticity induction. In practical implementations of a plasticity model, the exponential growth of the synaptic weights \( w_{ij} \) would stop when they obtain their maximal value \( w_{\text{max}} \). However, if all synapses onto a postsynaptic neuron, or even all synapses in a neural network sit at their upper bound, the network cannot function as a memory.

Note that the occurrence of the parameter \( \theta \) in the timescale \( \tau \) is a hallmark of the nonlinearity of the full system. Just like in the well-known Hodgkin-Huxley model where the timescale of the activation and inactivation variable depends on the voltage, the effective timescale \( \tau \) of a learning rule will depend on multiple factors such as the presynaptic activity, the slope of the gain function, or the threshold \( \theta \) between LTD and LTP.

Even though \( \tau \) in the example above is not the same as the induction timescale of long-term plasticity, in experiments where only a single presynaptic pathway is stimulated, the two are related. With our formalism we can also account for the strength of the recurrent feedback that is received by a neuron embedded into a network. Just repeat the above analysis under the assumption that the presynaptic and postsynaptic neurons are mutually connected, of the same type, and fire at the same rate (Zenke et al., 2013). Whatever you consider as a reasonable scenario, the timescale \( \tau \) characterizes how quickly a neuron or an entire recurrent network can generate positive feedback and is able to “run away”.

Instead of writing down a differential equation for the weights \( w \), as in Eq. (7), we could have written the system in terms of \( y \), too. A formulation in terms of the firing rates \( y \) highlights the fact that we have to think about run-away effects of synapses as being linked to run-away effects of neuronal activity. Note further, that in realistic neuronal network the presynaptic activity fluctuates and is different between one neuron and the next. Fluctuations give rise to a covariance matrix

\[
C = \frac{1}{n} \sum_{t=1}^{n} x_i(t)x_j(t)
\]

which may look complicated at a first glance. However, due to the symmetry of \( C \), there always exists a basis in which \( C \) is diagonal. When working in this basis, the plasticity equations decouple and take the shape of Eq. (7) with different values of \( \eta \).

In summary, by using the formalism introduced in this subsection, we can combine a plausible
data-based plasticity model and a given network model to find the effective timescale $\tau$ for this compound system. In doing so we take for instance explicitly into account that even if the change of a single synapse within one second might be small, the combined effect of millions of synapse of thousands of recurrently synapses could still be large. By distilling all these factors into a single number $\tau$ we can now return to the question of homeostatic plasticity. How can homeostatic control catch up with Hebbian plasticity that makes weights “explode” on the timescale $\tau$?

**Stability through homeostatic plasticity**

To stabilize Hebbian plasticity, theorists may select and mathematically formalize homeostatic mechanisms found in nature (Turrigiano, 2011). Two mathematical approaches to synaptic homeostasis are widespread. Either the Hebbian plasticity term is modified by a sliding threshold, or the Hebbian plasticity term for memory formation is complemented by an independent term for rate control. Although other choices are possible, we will now focus on those two mechanisms. Can either of these mechanisms guarantee stability of Hebbian learning? Are there fundamental restrictions with respect to how fast homeostatic plasticity has to be?

**Plasticity control through a sliding LTP-LTD threshold**

Several models assume that the threshold that marks the boundary between LTP and LTD is not fixed but may slide on a slow timescale (Bienenstock et al., 1982; Cooper et al., 2004; Pfister and Gerstner, 2006). We refer to this model class as BCM models in the following, where BCM is the acronym for Bienenstock-Cooper-Munro. The basic idea of that model class is that a sliding threshold could induce a homeostatic mechanism to locally control the amount of LTP or LTD in each synapse. The question then arises whether the sliding threshold can indeed lead to stable memory formation without the need for a further fast homeostatic mechanism.

To formalize this idea, we use the low-pass filter $\bar{y}_i$ of the postsynaptic firing rate (averaged over a time $\tau_d$) as a sensor that drives homeostatic plasticity; see (Eq. (4)). Intuitively, a homeostatic sliding threshold $\theta(\bar{y})$ increases the amount of LTD if the average postsynaptic activity $\bar{y}$ exceeds a given homeostatic set point $\kappa$ for extended periods of time (Fig. 1b). Conversely, when neuronal activity falls below this homeostatic activity target, the rate of LTP is increased at the expense of LTD. There are several different possible forms of such BCM models (Bienenstock et al., 1982; Cooper et al., 2004). The following is a simple, but representative example

$$\frac{dw_{ij}}{dt} = \eta x_j y_i (y_i - \theta(\bar{y}_i))$$

with a sliding threshold $\theta(\bar{y}_i)$. A common choice for the sliding threshold is $\theta(\bar{y}_i) = \frac{\bar{y}_i^2}{\kappa}$. Let us analyze this sliding threshold model. If we insert the expression for $\theta$ into Eq. (8) and choose a constant firing rate $y_i = \bar{y}_i = \kappa$, we see that the weight $w_{ij}$ does not change. Therefore
\( y_i = \kappa \) is a fixed point of the dynamics. The stability of this fixed point depends on the timescale of the low-pass filter \( \tau_d \) which in turn determines the timescale of metaplastic changes in Eq. (8) (Abraham (2008); Fig. 1b). It can be shown analytically that the system has a stable fixed point only if \( \tau_d \) is chosen smaller than some critical value \( \tau_d^{\text{crit}} \) (Cooper et al., 2004; Zenke et al., 2013). Thus, a slow homeostasis of the sliding threshold is, on its own, not sufficient to achieve stability.

**Rate control through synapatic scaling**

To introduce the second control mechanism, let us consider a neuron embedded in a large network of stochastically firing neurons. If many synaptic weights onto a given postsynaptic neuron increase, the postsynaptic firing rate \( y \) is expected to increase. If the firing rate remains high over some timescale \( \tau_d \), we may consider this as an indicator that the neuron runs out of its target firing regime. Let us therefore use the low-pass filtered rate \( \bar{y}(t) \) (Eq. (4)) as a driver of homeostasis (van Rossum et al., 2000).

Let us focus on a postsynaptic neuron \( i \) with rate \( y_i \). We adjust the weight \( w_{ij} \) from a presynaptic neuron \( j \) to our postsynaptic neuron according to the rule

\[
\frac{dw_{ij}}{dt} = \eta x_jy_i(y_i - \theta) + \tilde{\eta}(\kappa - \bar{y}_i)w_{ij}
\]

The first term on the right-hand side is the Hebbian plasticity term discussed earlier, now with a fixed threshold \( \theta \) that marks the transition from LTP to LTD. The multiplicative factors in the second term are \( \tilde{\eta} \) (a parameter) and the weight \( w_{ij} \) itself. The term \( \kappa \) inside the parenthesis of the second term corresponds to a target value of the postsynaptic activity and \( \bar{y}_i \) is a low-pass filtered version of the firing rate \( y_i \). A simple and biologically plausible way of how a neuron can sense \( \bar{y}_i \) on-line, is by reading out the concentration of a slow activity-related intracellular factor.

Three remarks are important. First, the homeostatic term corresponds to the heterosynaptic plasticity term \( f_1(w_{ij}, \text{post}_i) \) in Eq. (2). Second, the fact that the heterosynaptic term is proportional to the weight \( w_{ij} \) implies that, for \( \bar{y}_i > \kappa \), larger weights are depressed more than smaller ones. Thus this heterosynaptic term implements synaptic scaling. Third, in the language of control theory the heterosynaptic term implements homeostasis via a proportional controller. More sophisticated forms of heterosynaptic plasticity like a proportional-integral controller are possible (van Rossum et al., 2000), but this generally does not change the core of our argument that we will develop in the following.

How strong and how fast does the heterosynaptic term have to be to fulfill its homeostatic function? It can be shown analytically, that the requirement of stability in the above system (Eq. (9)) puts some tight constraints on the parameters \( \tau_d \) and \( \tilde{\eta} \) for a given value of \( \kappa \) (Zenke et al., 2013). First, \( \tilde{\eta} \) cannot be significantly different from \( \eta \). If \( \tilde{\eta} \) is much larger than \( \eta \) homeostatic control creates oscillations which are not observed in biological systems. On the other hand, if \( \tilde{\eta} \) is much
Figure 2: **The timescales of synaptic scaling or metaplasticity are faster in models than reported in experiments.** Here we plot the homeostatic timescale of either synaptic scaling or homeostatic metaplasticity as used in influential modeling studies (black). For comparison we plot the typical readout time for experimental studies on synaptic scaling and metaplasticity (red). Publications suffixed with * describe network models as opposed to the other studies which relied on single neurons. For reference we also indicate the typical timescales of short-term plasticity and long-term plasticity induction at the top. Note, that homeostatic timescales in models fall in the same range as the induction timescale of Hebbian long-term plasticity, which does not seem to be the case for experimental results on homeostatic plasticity.

smaller than $\eta$, the fixed point $\kappa$ system loses stability altogether. Similarly, we can derive stability conditions for $\tau_d$ and show that stability is only guaranteed if $\tau_d$ is smaller than some critical value $\tau_{d_{\text{crit}}}$ (which explicitly depends on $\eta$). The mathematical analysis thus shows that synaptic scaling needs to be “fast”, or at least faster than some reference. To further answer the question of how fast homeostatic plasticity has to be, we therefore have to quantify the speed of “normal” Hebbian plasticity.

Thus, our analytical approach has allowed us to answer the first question that we posed in the introduction. We need homeostatic synaptic plasticity which is faster than some critical value in order to achieve stability of learning dynamics. This statement is true for both the sliding threshold mechanism and the rate-control framework. However, what does “fast homeostatic plasticity” mean quantitatively? Is this critical timescale on the order of minutes, hours or days? To answer this question we need to quantify parameters such as $\eta$ with experimental data. In the following subsection we will address this question and translate the mathematical statements of this section into interpretable numbers.
How fast does a rapid homeostatic mechanism have to be?

Slow homeostatic control cannot stabilize the rapid positive feedback loop induced by Hebbian learning. However, in this statement, the meaning of “rapid” and “slow” is relative and based on a comparison of the timescale of homeostasis with that of induction of Hebbian plasticity. But what is the actual timescale of Hebbian plasticity?

To connect the framework of learning rules outlined in Eqs.(1–8) to data from plasticity experiments we can rely on a wealth of data on synaptic plasticity. Experiments which used STDP protocols to study induction of Hebbian plasticity are particularly informative (Sjöström et al., 2001; Senn et al., 2001). Theorists have made great strides in developing models which quantitatively capture a wide range of these data (Senn et al., 2001; Pfister and Gerstner, 2006; Froemke and Dan, 2002; Clopath and Gerstner, 2010). The parameters of the resulting spike-based plasticity models capture the timescale of plasticity induction in STDP experiments.

STDP models such as (Pfister and Gerstner, 2006) give a particular instantiation of the Hebbian term in Eq. (9) and make it possible to extract its parameters directly from experimental data. To bridge the conceptual gap between the spiking plasticity models and our simplified framework in the firing rate domain we have to consider neural firing statistics. Because neuronal spiking in vivo is highly irregular, we make the assumption that firing statistics can be approximated by a Poisson process. If both pre- and postsynaptic neurons fire stochastically and Poisson-like, the STDP models can be rewritten as firing rate models (Pfister and Gerstner, 2006). In particular, modern STDP models (Senn et al., 2001; Pfister and Gerstner, 2006; Clopath and Gerstner, 2010) exhibit a transition between LTD and LTD which depends on the firing rate, consistent with experimental data (Sjöström et al., 2001). We can exploit the connection between data, STDP models, and the effective Hebbian rate model so as to determine a plausible timescale for the induction of Hebbian plasticity.

(A short aside: In the context of STDP, it might seem paradoxical to assume Poisson firing, which explicitly gets rid of any temporal correlations between the input and output spike trains. However, as long as temporal correlations are as small as commonly seen in cortical activity or mimicked in simulations of balanced state networks (van Vreeswijk and Sompolinsky, 1996; Renart et al., 2010), the Poisson approximation gives accurate results.)

By plugging our estimate of the timescale of Hebbian plasticity into the analytical expressions for the parameters which characterize the critical limiting cases $\tau_d^{\text{crit}}$ and $\tilde{\eta}^{\text{crit}}$ (Zenke et al., 2013), we obtain a quantitative answer for how rapidly homeostatic mechanisms need to be. By analyzing the stability of a low-activity background state in a recurrent network model in which synaptic weights were evolving according to a plausible Hebbian plasticity model, we found that homeostatic control mechanisms need to act on a timescale in the range of seconds to minutes (Zenke et al., 2013). In the presence of network activity with temporal correlations or to provide robustness against perturbations, the timescale of homeostasis must be even shorter.
More specifically, and most importantly for the topic of this paper, the time constant \( \tau_d \) of the low-pass filter in the rate-control mechanism of Eq. (9) has to be short; otherwise the induced phase delay destabilizes the system. Possible choices for \( \tau_d \) include \( \tau_d \rightarrow 0 \), which means that \( \bar{y} = y \). In other words, heterosynaptic plasticity needs to be nearly instantaneous to achieve its aim to control of firing rates during ongoing Hebbian plasticity. Thus, if we want to achieve homeostasis by synaptic scaling, the synaptic scaling has to be extremely fast (Zenke et al., 2013).

Finally, this requirement cannot be loosened by more complicated homeostatic control mechanisms which rely on complicated molecular signalling cascades. Homeostatic control mechanisms which require the buildup of secondary or tertiary messengers, can be formalized in the language of control theory as integral controllers. Integral controllers can pick up small deviations from a target value and correct them. Therefore they are good for fine-tuning. However, they have to be “slow” (or more precisely, slower than some critical value), otherwise they become unstable even in the absence of Hebbian plasticity (Harnack et al., 2015).

In summary, the application of our general analytical theory to data on STDP experiments gives an answer to the second question posed in the introduction: Rapid homeostatic control implies a reaction of the controller on the timescale of seconds to minutes. This answer is in good agreement with most existing simulations of plastic network models (Fig. 2) — in each of these a rapid homeostatic control mechanisms on a timescale of seconds to minutes was implemented to maintain stability Pfister and Gerstner (2006); Lazar et al. (2009); El Boustani et al. (2012); Clopath and Gerstner (2010); Gjorgjieva et al. (2011); Litwin-Kumar and Doiron (2014).

What are functional consequences of rapid homeostatic control?

If we accept the notion that homeostatic plasticity acts on a short timescale, what are the functional consequences for plasticity and circuit dynamics? Neurons encode information in changes of their electrical activity levels. For instance, subsets of simple cells in the visual system fire spikes in response to specific edge-like features in the visual field (Hubel and Wiesel, 1962); cells in higher brain areas respond with high specificity to complex concepts and remain quiescent when the concept they are coding for is not brought to mind (Logothetis and Sheinberg, 1996; Hung et al., 2005; Quiroga et al., 2005); and finally certain neurons respond selectively with elevated firing rates over extended periods during working memory tasks (Miyashita, 1988; Funahashi et al., 1989; Wilson et al., 1993). The ability of neurons to selectively indicate through periods of strong activity the presence of specific features in the input or specific concepts in working memory is an important condition for computation.

Homeostatic plasticity, however, by definition tries to keep neuronal activity constant. In particular, most models of homeostatic plasticity have a single homeostatic target or set point. That means, that homeostatic plasticity tries to achieve some notion of constancy of one specific target variable over time. In many models this is the mean neuronal firing rate, but other target variables
Figure 3: The effect of slow and rapid homeostasis on the neuronal code. (a) Fluctuating external world stimulus over time. (b) Neural activity $A = w \times \text{input over time}$. A slow homeostatic mechanism adjusts $w$ to push $A$ towards a single homeostatic target (dashed line; Homeostasis is modelled as $\tau_{\text{slow}} \frac{dw}{dt} = (\text{Target} - A)$). The homeostatic target is indicated by the dashed line). With a slow homeostatic mechanism, a neuron can track input relatively accurately. (c) Same as b, but for a rapid homeostatic mechanism ($\tau_{\text{slow}} = 50\tau_{\text{fast}}$). If the homeostatic timescale becomes comparable to the timescale of the stimulus, homeostasis starts to interfere with the neuron’s ability to track the stimulus. (d) Rapid homeostatic mechanisms enforcing an allowed range (limits indicated by dotted lines). Even though the neuron cannot capture all the diversity of the input, it can capture some of it. Here we modeled homeostasis as the following nonlinear extension of the above model: $\tau_{\text{fast}} \frac{dw}{dt} = f(\text{Target} - A)$ with $f(x) = x$ for $\|x - \text{Target}\| > \text{Limit}$ and $f(x) = 0$ otherwise.
such as the bulk-synaptic conductance of all afferent synapses onto the same neuron are possible.

Is homeostatic control of activity compatible with the task of neurons to selectively respond to stimulation? If homeostatic control of firing rates is slow, neuronal firing can deviate substantially from the mean firing rates during short times and thus encode information (Fig. 3a,b). However, we already know that a slow homeostatic control mechanism cannot stabilize the ravaging effects of Hebbian plasticity. So what can we say about a rapid homeostatic mechanism? If the homeostatic mechanism acts on a short timescales (e.g. seconds to minutes) as it would be required to stabilize Hebbian plasticity, neural codes based on neuronal activity become problematic (Fig. 3c) because homeostatic plasticity starts to suppress activity fluctuations which might be encoding important information. Even more alarmingly, certain forms of homeostatic plasticity, like sliding threshold models of homeostatic metaplasticity, not only suppress high activity periods, but also “unlearn” previously acquired selectivity and delete previously stored memories (Fig. 4a–d). Therefore rapid homeostatic mechanisms which enforce a single homeostatic set point are not hardly desirable from a functional point of view. Thus the requirement of fast homeostatic control over Hebbian plasticity poses a problem.

It is important to appreciate that this problem arises from the combination of a single homeostatic target with the requirement to enforce the homeostatic constraint on a short timescale. However, there might be a simple solution to this conundrum. Suppose there are two (or more) homeostatic set points, or a target range (instead of a target point), implemented by multiple forms of rapid homeostatic control. For instance, one mechanism of homeostatic plasticity could activate above a certain activity threshold and ensure that neuronal activity does not exceed this threshold. Similarly, a second mechanisms could activate below a low-activity threshold. The combined action of the two mechanisms enforces neural activity to stay within an allowed range, but still permits substantial firing rate fluctuations inside that range (Fig. 3d).

When such a homeostatic mechanism is paired with a form of Hebbian plasticity which has its plasticity threshold within the limits of the allowed activity regime, the neural activity of the compound system naturally becomes bistable. Within the allowed range no homeostatic plasticity is active, but Hebbian plasticity is intrinsically unstable. Thus any value within the homeostatically allowed region will lead to either LTP or LTD until the system reaches the limits at which homeostatic plasticity rapidly intervenes by undoing any excess LTP or LTD from there on. The compound system exhibits therefore two stable equilibrium points one at low and one at high activity. If the high-activity fixed point corresponds to a memory retrieval state, it becomes becomes irrelevant whether the memory is recalled every other minute or if its only recalled once a year.

We can therefore answer the third of the questions raised in the introduction: *The functional consequences of rapid homeostatic control with a single set point are that neurons loose the flexibility that is necessary for coding.* The consequences are therefore undesirable. The proposed solution is to design rapid homeostatic control mechanisms that allow for several set points (bistability) or a target range of activity. In the next sections we will argue that the well-orchestrated combination of
Hebbian and non-Hebbian plasticity mechanisms naturally gives rise to such bistable dynamics. We will explain why these dynamics are useful for learning and memory and why, despite their stability, they still need to be accompanied by additional slow homeostatic mechanisms.

Can the combination of Hebbian and non-Hebbian plasticity lead to stable memory formation?

We now discuss a new class of plasticity models which combines a plausible model of Hebbian plasticity with two additional homeostatic plasticity mechanisms (Zenke et al., 2015). For sensible combinations these compound models do not suffer from the run-away effects of purely Hebbian plasticity and exhibit intrinsic bistability instead (cf. Fig. 3d).

The basic logic of bistable plasticity can be summarized as follows. At high activity levels a rapid form of heterosynaptic plasticity limits run-away LTP and creates synaptic competition. Similarly, at low activity levels a unspecific form of plasticity which only depends on presynaptic activity prevents run-away LTD. The well-orchestrated interplay between these adversarial plasticity mechanisms dynamically creates bistability of neuronal activity and prevents pathologic run-away effects.

Our approach is general and any Hebbian plasticity model can be stabilized through the addition of two non-Hebbian forms of plasticity. For illustration purposes, we will now focus on the triplet STDP model for which biologically plausible sets of model parameters exist (Pfister and Gerstner, 2006). To prevent run-away LTP in this model, we use a form of weight-dependent, multiplicative heterosynaptic depression (Chen et al., 2013; Chistiakova et al., 2015; Zenke et al., 2015) which has also been interpreted as a rapid form of depotentiation (Zhou et al., 2003). To prevent run-away LTD, we introduce hypothetical “transmitter-induced” plasticity which depends on presynaptic activity only and slowly increases synaptic efficacy in the absence of postsynaptic spiking (Zenke et al., 2015). These three plasticity mechanisms work in symphony to generate two stable levels of neuronal activity.

Let us consider the weight \( w_{ij} \) from a presynaptic neuron \( j \) to a postsynaptic neuron \( i \). Although the full model is an STDP model, we now express its core ideas in terms of the presynaptic firing rate \( x_j \) and the postsynaptic rate \( y_i \)

\[
\frac{dw_{ij}}{dt} = \delta \cdot x_j + \eta \cdot x_j y_i (y_i - \theta) - \beta \cdot (w_{ij} - \tilde{w}_{ij}) y_i^4
\]

(10)

Here, \( \delta \) and \( \beta \) are strength parameters for the two non-Hebbian components of the plasticity model; \( \eta \) is the strength parameter (learning rate) for Hebbian plasticity; and \( \tilde{w}_j \) serves as a reference weight that can be related to consolidation dynamics (Zenke et al., 2015; Ziegler et al., 2015). Note that, because the homeostatic mechanisms are “rapid”, no low-pass filtered variable \( \bar{y} \) appears in the
Figure 4: Intrinsically stabilized firing rates, synaptic weights and selectivity through rapid homeostatic mechanisms. (a) Schematic figure of single neuron setup with two distinct input pathways. The “active” pathway consists of 40 Poisson neurons firing at either 2 or 10Hz. The control pathway consists of 400 neurons all firing constantly at 2Hz. All weights are initialized at the same values. (b) Population firing rates of the input populations averaged over 2s bins. Firing rates in the active pathway (solid line) are temporally changed from 2Hz to 10Hz, whereas firing rates in the control pathway are constant at 2Hz. (c) Output firing rates of a single postsynaptic neuron with triplet STDP (Pfister and Gerstner, 2006) at all input synapses for two different time constants $\tau_d$ of the homeostatic sliding threshold (see Zenke et al. (2013); $\kappa=3\text{Hz}$). Purple: Slow homeostatic sliding threshold, $\tau_d = 1h$; Green: Fast sliding threshold, $\tau_d = 10s$. Top and bottom show the same firing rate plot for different y-axis ranges. (d) Relative weight changes for 10 randomly chosen weights from each pathway for the slow (purple) and the fast (green) sliding threshold. Darker colors correspond to active pathway weights and lighter colors to the control pathway. (e) Simplified sketch of the setup with two identical postsynaptic neurons which receive the same input as described in a. All input synapses are plastic (see Zenke et al. (2015)). The only difference between the two neurons in this setup are the different initial conditions. For Neuron 2 the active pathway weights are initialized at a lower value than for Neuron 1 which could be thought of being the outcome of previous heterosynaptic depression. (f) Output firing rates of the two neurons over time. Neuron 1 (black) responds selectively to the elevated inputs in the selective pathway (cf. b). Neuron 2 (orange) does not respond with elevated firing rates during stimulation. (g) Evolution of weights over time for Neuron 1. Active pathway weights are plotted in black and the control pathway in gray. For Neuron 1, the synapses in the active pathway initially undergo LTP during 10Hz stimulation of the active pathway synapses. However, weights quickly saturate. Synapses in the control pathway undergo heterosynaptic depression. (h) Same as g, but for Neuron 2. In contrast to that, Neuron 2 continues to fire with low firing rates and barely responds to the 10Hz input in the active pathway. The active pathway weights are slightly depressed during initial stimulation.
expression. Due to its rapid action and the high power, the heterosynaptic term in Eq. (10) acts as a burst detector which dominates at high activity levels and prevents LTP run-away dynamics (Fig. 4f).

For sensible choices of $\delta$ and $\beta$, neuronal firing rates remain in intermediate regimes (Fig. 4f) and synaptic weights in the model converge towards stable weights $\hat{w}$ whose values are dependent on the activation history and allow the formation of long-term memories. Importantly, the model preserves the plasticity threshold between LTD and LTP of the original triplet STDP model. The triplet model together with the non-Hebbian plasticity mechanisms, dynamically creates two stable equilibrium points. Overall, the synaptic weights converge rapidly towards one of two possible stable equilibrium states (Fig. 4g–h). First, there is a “selective” equilibrium state associated with high postsynaptic activity. In this state some weights are strong while other weights onto the same postsynaptic neuron remain weak. Thus the neuron becomes selective to features in its input (neuron 1 in Fig. 4f and g). Second, there is a low-activity fixed point without selectivity (Fig. 4f and h). Which fixed point a neuron converges to depends on initial conditions and the details of the activation pattern (Fig. 4g–h). Once weights have converged to one of the respective stable states, weights keep fluctuating, but do not change on average. Moreover, since the homeostatic mechanism does not impose a single set-point, activity patterns are not unlearned when a certain input is kept active for longer times (compare Fig. 4d with 4g).

There are several aspects worth noting about the model. First, heterosynaptic plasticity does not only stabilize Hebbian plasticity in the active pathway, it also introduces synaptic competition between the active and the control pathway (Fig. 4g). Unlike BCM-like models in which heterosynaptic depression of the inactive pathway depends on intermediate periods of background activity in between stimuli (Cooper et al., 2004; Jedlicka et al., 2015), here the heterosynaptic depression happens simultaneously to LTP induction (Fig. 4g). Second, although the learning rule effectively implements a rapid rescaling of the synaptic weights, it is still a fully local learning rule which only relies on information which is locally available at the synapse (cf. Eq. (10)). Third, in general the reference weight $\tilde{w}$ is not fixed, but follows its own temporal dynamics on a slower timescale ($\approx 20$min and more). Such complex synaptic dynamics are essential to capture experiments on synaptic consolidation (Frey and Morris, 1997; Redondo and Morris, 2011; Ziegler et al., 2015). Consolidation dynamics also play an important role in protecting memories from being overwritten by heterosynaptic plasticity. Finally, the stability properties of the learning rule Eq. (10) are not limited to simple feed-forward circuits, but generalize to more realistic scenarios. Specifically, it enables stable on-line learning and recall of cell assemblies in large spiking neural networks (Zenke et al., 2015).

In summary, orchestrating Hebbian and homeostatic plasticity on comparable timescales to dynamically create bistability, can reconcile the experimentally observed fast induction of synaptic plasticity with stable synaptic dynamics and ensure stability of learning and memory at the single neuron level. We now explore potential problems of intrinsically bistable plasticity?
Problems intrinsically bistable plasticity cannot solve

Consider a population of neurons with plastic synapses which follow intrinsically bistable plasticity dynamics such as the ones described in the last section. To encode and process information efficiently, neuronal populations need to create internal representations of the external world. Doing this efficiently requires the response to be sparse across the population. In other words, only a subset of neurons should respond for each stimulus. Moreover, different stimuli should evoke responses from different subsets of neurons within the population to avoid that all stimuli look “the same” to the neural circuit. Finally, individual neurons should respond sparsely over time. Imagine a neuron which is active for all possible stimuli. It would be as uninformative as one which never responds to any of the inputs. Therefore, to represent and process information in neural populations efficiently, different neurons in the population have to develop selectivity to different features.

Intrinsically, bistable plasticity alone does not prevent neurons from responding at low firing rates to all stimuli (see, for example, Neuron 2 in Fig. 4f). Moreover, with similar initial conditions both neuron 1 and 2 would have developed selectivity to the same input. Thus, in a large network where all synapses are changed by the intrinsically bistable plasticity rule introduced above, all neurons could end up responding to the same feature. The question then arises, whether it is possible to prevent such an undesired outcome.

To successfully implement network functions such as these examples, several network parameters and properties of the learning rules themselves need to be tuned to and maintained in sensible parameters regimes. To achieve this, additional forms of homeostatic plasticity need to be in place. However, due to the intrinsic stability of the learning rule, these additional homeostatic mechanisms can now safely act on much larger timescales, similar to the ones observed in biology.

Thus we have answered the fourth question posed in the introduction. Yes, the combination of Hebbian and rapid homeostatic plasticity can lead to stable memory formation, if homeostasis gives rise to intrinsic bistability. A combination of Hebbian plasticity with heterosynaptic LTD and transmitter-induced LTP gives rise to suitable bistable dynamics. However, on its own, intrinsic bistability of plasticity dynamics is not sufficient to induce complementary functionality across many neurons in a network. The question then arises whether slower forms of homeostasis could come into play in this context.

What is the role of slower homeostatic plasticity mechanisms?

Diverse homeostatic mechanisms exist in the brain at different temporal and spatial scales (Marder and Goaillard, 2006; Turrigiano, 2012; Davis, 2013; Gjorgjieva et al., 2016). We have argued that rapid homeostatic mechanisms are important for stability, but advantages do slow homeostatic mechanisms have and what is their computational role?

An advantage of slow homeostatic processes is that they can integrate activity over long timescales
to achieve precise regulation of neural target set points (Turrigiano and Nelson, 2000; van Rossum et al., 2000; Harnack et al., 2015). Longer integration times also allow to integrate signals from other parts of a neural network which take time to be transmitted as diffusive factors (Turrigiano, 2012; Sweeney et al., 2015). Slower homeostasis thus seems well suited for control problems which either require fine-tuning or the more global homeostatic regulation of functions at the network level.

There are at least two important dynamical network properties which are not directly controllable by the rapid homeostatic mechanisms in Eq. (10). First, temporal sparseness at the neuronal level. A neuron that never responds to any stimulus, will never do so under bistable plasticity alone unless the LTP threshold is lowered. Second, to ensure spatial sparseness at the network level lateral inhibition has to decorrelate neuronal responses. As excitatory synapses change during learning, the strength of lateral inhibition typically needs to be finely regulated.

The problem of temporal sparseness can be solved by any mechanism which ensures that a neuron which has been completely silent for very long, eventually “gets a chance” to activate above the LTP threshold. This can be achieved by either lowering the threshold as in the BCM theory (Bienenstock et al., 1982; Cooper et al., 2004; Zenke et al., 2015) or by slowly increasing the gain of either the neuron itself or the excitatory synapses through other forms of slow homeostatic plasticity (Daoudal and Debanne, 2003; Turrigiano, 2011; Toyoizumi et al., 2014). Finally, similar homeostatic effects could be achieved by decreasing inhibitory synaptic input through the action of neuron specific inhibitory plasticity (Vogels et al., 2013).

Decorrelating neural activity and enforcing sparse activity at the population level has typically been associated with lateral or recurrent inhibitory feedback in sparse learning paradigms (Olshausen and Field, 1996) and for models of associative memory (Willshaw et al., 1969; Little and Shaw, 1978; Hopfield, 1982; Amit et al., 1985; Amit and Brunel, 1997; Litwin-Kumar and Doiron, 2014; Zenke et al., 2015). However, to achieve sensible global network activity levels through recurrent inhibitory feedback requires fine-tuning (Amit and Brunel, 1997; Brunel, 2000).

Inhibitory synaptic plasticity (ISP) can achieve fine-tuning and homeostatic functions in spiking neural networks. Specifically, Hebbian forms of ISP, for which synchronous pre- and postsynaptic activation leads to a strengthening of inhibitory synapses, has been shown to be a powerful mechanism to regulate neural firing rates, to establish and maintain excitatory-inhibitory balance and to decorrelate neuronal firing in neural networks (Vogels et al., 2011; Zenke, 2014; Litwin-Kumar and Doiron, 2014). Existing Hebbian ISP models, however, increase inhibition between neurons which are frequently active together. This is useful to create excitatory inhibitory balance and prevent individual neurons from being synchronized. However, if neurons respond repeatedly together as part of the same cell assembly, these neurons will start to inhibit each other. Associative memory models, however, require the opposite situation: neurons in the same assembly should not inhibit each other while neurons of different cell assemblies inhibit each other so as to avoid that many cell assemblies are activated simultaneously. To achieve the latter, inhibition needs to be tuned globally.

Biological networks can achieve this type of global tuning through slow forms of homeostatic
plasticity (Turrigiano, 2012) which rely on diffusive signals (Sweeney et al., 2015). These signals could either globally scale synaptic strength at inhibitory synapses directly, or indirectly influence the levels of inhibition by modulating inhibitory learning rules through a third factor (Zenke et al., 2015; Frémaux and Gerstner, 2016).

It is important to realize that not all homeostatic mechanisms have to be permanently active. For instance, once the inhibitory feedback within a spiking neural network model is tuned to the “sweet spot” at which the network can operate, moving threshold mechanisms and ISP can safely be turned off while the network can still learn new stimuli (Zenke et al., 2015). It thus seems likely that in nature certain forms of homeostatic plasticity could be dormant for most of the time and spring into action only to prepare a network during the initial phase of development (Turrigiano and Nelson, 2004) or when an extreme external manipulation changes the network dynamics (Turrigiano et al., 1998).

Thus we are able to answer the final question as follows: 

**Slow homeostatic mechanisms tune parameters of plasticity rules and neurons to enable efficient use of the available resources in networks.**

For example, for the sake of efficiency, no neuron should be never active; no neuron should be always active; the number of neurons that respond to the exact same set of stimuli should stay limited.

Taken together with the results from the previous sections, these insights suggest two distinct roles of homeostatic plasticity mechanisms on different timescales. First, homeostatic plasticity on short timescales stabilizes Hebbian plasticity and makes synapses onto the same neuron compete with each other. Heterosynaptic plasticity is likely to play a major role for these functionalities. Second, homeostatic mechanisms on slower timescales achieve fine-tuning of multiple network parameters. A slow shift of the threshold between LTD and LTP, the slow rescaling of all synaptic weights, or a slow regulation of neuronal parameters are good candidates for these functionalities. Some of these slow mechanisms could be important only in setting up the network initially or after a strong external perturbation to the circuit.

**Discussion**

We discuss the relation of the above modeling approach to other models and to experimental data.

**Other intrinsically bi-stable plasticity models**

The notion of intrinsic stability (or bistability) has come up recently in other studies. Toyoizumi et al. (2014) proposed the following plasticity rule which formally fulfills the requirements of an intrinsically stable rule:

\[
\tau_w \frac{dw}{dt} = \begin{cases} 
[w_{\text{max}} - w]_+ [xy - \theta]_+ - [w - w_{\text{min}}]_+ [\theta - xy]_+ 
\end{cases} \quad \text{LTP} \quad \text{LTD}
\]  

(11)
The full rule is multiplicatively coupled with a slow homeostatic process which we have omitted in
the above equation. However, even in the absence of any additional form of homeostatic plasticity,
individual synaptic weights saturate at the soft bounds \( w_{\text{min}} \) and \( w_{\text{max}} \) respectively. Similarly, to
Eq. (10) the model possesses a plasticity threshold \( \theta \). However, here, the threshold is defined via the
product of pre- and postsynaptic activity, similar to (El Boustani et al., 2012), but in contrast to
most data-driven models in which the threshold is defined on postsynaptic quantities only (Artola
and Singer, 1993; Shouval et al., 2002; Pfister and Gerstner, 2006; Clopath and Gerstner, 2010;
Graupner and Brunel, 2012).

Another example of rapid stabilization of an otherwise unstable learning rule was devised by
Lim et al. (2015). Using an analytic framework the authors infer the following rate-based plastic-
ity model directly for experimental data on excitatory-to-excitatory connections in infero-temporal
cortex (ITC):

\[
\frac{dw}{dt} \propto x (y - \theta) .
\]

Because this model is unstable and leads to responses in a recurrent network model which are
much smaller than their ITC measurements, the authors added a rapid homeostatic mechanism
which maintains the average synaptic strength onto a neuron. This can be formalized as:

\[
\frac{dw}{dt} \propto (x - \bar{x}) (y - \theta) .
\]

(12)

It is interesting to note that in this model, weights grow with the rate \( \bar{x} \theta \) in the absence of
activity to oppose run-away LTD which is conceptionally similar to Eq. (10). However, in contrast
to Eq. (10), in the present model, run-away LTP is counterbalanced at the network level through
recurrent inhibitory feedback which pushes the activity \( y \) of most neurons below the plasticity
threshold \( \theta \) causing them to undergo LTD. This suggests that under some circumstances the rapid
modulation of excitatory plasticity through inhibitory feedback might be enough to avoid run-away
dynamics.

While there are several different ways to stabilize Hebbian plasticity, a plethora of theoretical
models seems to conceptually agree on two key factors: First, to achieve stability against run-away
LTP a model needs to ensure that negative feedback is rapid (Pfister and Gerstner, 2006; Lazar et al.,
2009; El Boustani et al., 2012; Clopath and Gerstner, 2010; Gjorgjieva et al., 2011; Litwin-Kumar
and Doiron, 2014). And second, the feedback needs to be strong enough to be able to compensate
for the leading power of Hebbian plasticity (Oja, 1982; Bienenstock et al., 1982; Cooper et al., 2004;
Tetzlaff et al., 2012; Zenke et al., 2013; Toyoizumi et al., 2014).

**Different models capture the same plasticity experiments**

A plausible plasticity model has to reproduce stable learning dynamics in simulations and it has
to capture a plethora of experimental data from rate-dependent (Bliss and Lomo, 1973), voltage-
dependent (Artola et al., 1990) and spike-timing-dependent (Markram et al., 1997; Bi and Poo, 1998; Sjöström et al., 2001) plasticity experiments. One salient feature captured by most models (Artola and Singer, 1993; Shouval et al., 2002; Pfister and Gerstner, 2006; Clopath and Gerstner, 2010; Graupner and Brunel, 2012) is the notion of a plasticity threshold which correlates with postsynaptic voltage, calcium concentration or other neuronal variables related to postsynaptic activation (Fig. 1a). However, most existing models are purely Hebbian and do not include the notion of a rapid homeostatic mechanism. If rapid homeostatic mechanisms exist — which is what we argue here — then how can it be that existing plasticity models without them can quantitatively capture the data from experiments?

There are presumably three main reason for this. First, STDP experiments typically manipulate a single pathway, either by stimulating a presynaptic neuron or a bundle of presynaptic axons. Sometimes a designated control pathway (i.e. a second presynaptic neuron) is missing, or, if it is not missing, the effect size in the control pathway is considered as weak. However, from a theoretical perspective, we expect that heterosynaptic effects caused by stimulation of one presynaptic pathway are weak when it is measured at one 'control' synapses. Even if it is weak on a per-synapses bases, its net effect is strong because it spreads out over thousands of synapses. Therefore even weak heterosynaptic plasticity at other synapses could implement a form of rapid homeostatic control (Chen et al., 2013; Chistiakova et al., 2015; Zenke et al., 2015).

Second, in an STDP experiments with 60 repetitions of pre-post-pairs, the total activation of the postsynaptic neuron is still in a reasonable regime. Therefore it is unclear whether the 'burst-detector' for heterosynaptic plasticity would be triggered (Chen et al., 2013; Chistiakova et al., 2015; Zenke et al., 2015).

Third, experiments typically rely on repeated pre- and postsynaptic activation. Moreover, during the induction protocol, synaptic efficacy changes are usually not observable. Plasticity models are thus fitted to pairs of initial and final synaptic strength. However, the unobserved intermediate synaptic dynamics could be qualitatively very different (Fig. 1c). These differences in the dynamics contain the answers to questions such as: Is the final synaptic strength stable or would it increase further with additional pairings? Is there a threshold of number of pairings that needs to be reached for an all or nothing effect?

Because the detailed dynamics of synapses during induction are not known, different plasticity models make different assumptions about the saturation of weights. Based on these assumptions, most existing plasticity can be coarsely classified into additive and multiplicative models. Additive plasticity models, for instance, assume a linear growth for plasticity induction (Fig. 1c; Pfister and Gerstner (2006); Clopath et al. (2010); Song et al. (2000)). In the absence of a hard upper bound, which is usually present in additive models, the weight would simply keep growing. Although additive models capture plasticity experiments for a fixed number of pairings, in simulated networks with ongoing pre- and postsynaptic activity, synaptic weights diverge and "get stuck" at the upper and lower bounds (Billings and van Rossum, 2009).
The equations of multiplicative models on the other hand depend on the synaptic weight explicitly which typically makes them converge to stable intermediate weight values (Gütig et al., 2003; Morrison et al., 2007; Gilson et al., 2010) and for ongoing spiking activity in network simulations gives rise to unimodal weight distributions (Morrison et al., 2007).

Both multiplicative and additive models generally do not lead to stability in network models (Morrison et al., 2007; Kunkel et al., 2011; Zenke et al., 2013) unless accompanied by additional homeostatic mechanisms. However, these mechanisms are typically added ad-hoc and they are not fitted to experimental data (e.g. Pfister and Gerstner (2006); Clopath et al. (2010)).

Due to the limited amount of experimental data, it is possible to construct a set of different models which are all consistent with experimental data. The model discussed in this paper is based on the triplet STDP model (and is therefore consistent with existing STDP data), but includes additional forms of activity-dependent non-Hebbian plasticity which give rise to intrinsic bistability (and which are harder to measure experimentally).

**Can we constrain plasticity models by experiments?**

There are multiple ways in which synaptic plasticity models could be constrained better through additional data. In the past, a large body of research has focused on homosynaptic associative plasticity, also called Hebbian plasticity, using pairing experiments with various protocols such as STDP. Here, we argue that heterosynaptic plasticity as well as transmitter-induced plasticity might be almost as important as Hebbian plasticity due to their possibly crucial role for network stability.

*Heterosynaptic plasticity* is a promising candidate to stabilize Hebbian plasticity models against run-away LTP is heterosynaptic plasticity (Chen et al., 2013; Zenke et al., 2013, 2015; Chistiakova et al., 2015; Jedlicka et al., 2015). While heterosynaptic plasticity has been observed in various experiments (Royer and Paré, 2003; Chistiakova et al., 2014), a conclusive picture and quantitative models are still missing. Is it possible to measure the timescale, frequency dependence, and weight-dependence of neuron wide heterosynaptic depression by manipulating the stimulation of the postsynaptic neuron? Another important question for the interpretation of heterosynaptic plasticity is whether heterosynaptic plasticity induces mostly synaptic depression similar or related to LTD or rather resets or prevents early LTP trough depotentiation at the unstimulated pathway (Zhou et al., 2003).

*Transmitter-induced plasticity* is important in models and might be present in many experiments, even though it has not been reported as such. Here transmitter-induced plasticity refers to potentially weak long-term potentiation that is caused by presynaptic firing in the absence of postsynaptic activity. Why is this form of plasticity potentially important? Suppose you have a network of neurons firing at low activity, so that any given neuron can be considered a weakly active postsynaptic neuron. Since low activity typically induces LTD, many plastic network simulations have the tendency to fall silent. To compensate for this theorists introduce either lower bounds on synaptic weights or add...
weak LTP triggered by presynaptic activity (Zenke et al., 2015; Lim et al., 2015). How realistic are these assumptions? Direct experimental evidence for such terms would for instance be the growth of synaptic efficacy during low activity “pre only” stimulation. It is possible that such a term would look like, and could easily be mistaken for, an unstable baseline (with positive drift) before the beginning of a plasticity induction experiment. Given the importance of such a term, even if very weak, from the theoretical perspective, it is worth explicit studies.

Consolidation of synapses is summarized in the present model by a reference weight $\tilde{w}$ (Zenke et al., 2015; Ziegler et al., 2015). Simulations predict that synaptic consolidation renders synapses inert against heterosynaptic plasticity. Intuitively, the measured synaptic weights become sticky and are always attracted back to their momentary stable state, i.e. weak or strong. This prediction requires future experimental clarification.

The path towards saturation of synaptic weights during a pairing experiment (Fig. 1c) is vital to building better plasticity models. Virtually any information which helps theorists to constrain how the synaptic weight increases would be helpful. Importantly, this also includes any information about conditions (or experimental protocols) which do not induce plasticity, despite the fact that either the presynaptic or the postsynaptic neuron or both have been activated.

Conclusion

One of the most striking differences between plasticity models and experimental data concerns the timescale. Hebbian plasticity can be induced within seconds to minutes (Bliss and Lomo, 1973; Artola et al., 1990; Markram et al., 1997; Bi and Poo, 1998). In simulated network models such a fast Hebbian plasticity leads to run-away activity within seconds unless Hebbian plasticity is complemented with rapid forms of homeostatic plasticity. Here, “rapid” means that homeostatic changes need to take affect on the timescale of seconds or at most a few minutes (Zenke et al., 2013). This, however, is much faster than most forms of homeostatic plasticity observed in experiments. One of the most extensively studied forms of homeostasis in experiments is synaptic scaling (Turrigiano et al., 1998) which proportionally scales up (down) synapses if the network activity is too low (high). However, even the fastest known forms of scaling take hours to days to cause measurable changes to synaptic weights (Fig. 2; Turrigiano and Nelson (2000); Ibata et al. (2008); Aoto et al. (2008)).

This apparent difference of timescales between homeostasis required for stability in models and experimental results is a challenge for current theories (Wu and Yamaguchi, 2006; Chen et al., 2013; Zenke et al., 2013; Toyoizumi et al., 2014). To reconcile plasticity models and stability in networks of simulated neurons, we need to reconsider models of Hebbian plasticity and how they are fitted to data.

In most plasticity induction experiments neither the time course of the manipulated synaptic weight nor changes of other synapses are observed during stimulation. Quantitative models of synaptic plasticity thus make minimal assumptions about these unobserved temporal dynamics and
generally ignore heterosynaptic effects entirely. In other words, missing experimental data makes it possible to build different models which all capture the existing experimental data, but make different assumptions about the unobserved dynamics. Importantly, some of these models become intrinsically stable (Oja, 1982; Miller and MacKay, 1994; Chen et al., 2013) or even bistable (Toyoizumi et al., 2014; Zenke et al., 2015). In most situations these models can be interpreted as compound models consisting of Hebbian plasticity and forms of rapid homeostatic plasticity. Importantly, all of the plasticity models cited in this paragraph only rely on quantities which are locally known to the synapse; i.e. the pre- postsynaptic activity as well as its own synaptic weight. Although such local forms of plasticity can solve the problem of stability at a neuronal level, in practice, most network models require additional fine-tuning of parameters to achieve plausible activity levels across a network of neurons. This role can be fulfilled slow homeostatic mechanisms which act on timescales of hours or days, consistent with experimental data on homeostatic plasticity.

In summary, based on theoretical arguments we suggest that Hebbian plasticity is intrinsically stabilized on extremely short timescales by fast homeostatic control, likely to be implemented by heterosynaptic plasticity while slow forms of homeostatic plasticity set the stage for stable learning. However, this hypothesis will now have to stand the test of time. It will thus be an important challenge for the next years to go beyond homosynaptic Hebbian plasticity and to gain a more complete understanding of the diverse interactions of Hebbian and homeostatic plasticity across timescales.

Acknowledgements. WG was supported for this work by the European Research Council under grant agreement number 268689 (MultiRules) and by the European Community’s Seventh Framework Program under grant no. 604102 (Human Brain Project). FZ was supported by the SNSF (Swiss National Science Foundation).

References


Ashok Litwin-Kumar and Brent Doiron. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat Commun*, 5, November 2014.


