Orchestrated Excitatory and Inhibitory Learning Rules Lead to the Unsupervised Emergence of Up-states and Balanced Network Dynamics.

AUTHORS:

Saray Soldado-Magraner[1], Helen Motanis[2], Rodrigo Laje[3,*], Dean V. Buonomano[1,4,*]

AFFILIATIONS:

[1]Department of Neurobiology, University of California, Los Angeles, CA, USA.

[2]Department of Neurosurgery, University of California, Los Angeles, CA, USA.

[3]Present addresses: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Argentina, and Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina.

[4]Department of Psychology University of California, Los Angeles, CA, USA.

[*]RL and DVB are Joint Senior Authors on this work.

ABSTRACT

Self-sustaining dynamics maintained through recurrent connections are of fundamental importance to cortical function. We show that Up-states—an example of self-sustained network dynamics—autonomously emerge in cortical circuits across three weeks of *ex vivo* development, establishing the presence of unsupervised synaptic learning rules that lead to globally stable emergent dynamics. Computational models of excitatory-inhibitory networks have established that four sets of weights ($W_{E\leftarrow E}$, $W_{E\leftarrow I}$, $W_{I\leftarrow E}$, $W_{I\leftarrow I}$) cooperate to generate stable self-sustained dynamics, but have not addressed how a family of learning rules can operate in parallel at all four weight classes to achieve self-sustained inhibition-stabilized regimes. Using numerical and analytical methods we show that standard homeostatic rules cannot account for the emergence of self-sustained dynamics due to the paradoxical effect. We derived a novel family of homeostatic learning rules that operate in parallel at all four synaptic classes, which robustly lead to the emergence of Up-states and balanced excitation-inhibition.

INTRODUCTION

Synaptic learning rules refer to the biological algorithms that govern the strength of synapses. Learning rules such as Hebbian plasticity, spike-timing-dependent plasticity, and forms of homeostatic plasticity have vastly advanced our understanding of the mechanisms underlying neurodevelopment, learning, and cognition (Hebb, 1949; Abbott and Nelson, 2000; Turrigiano and Nelson, 2004; Feldman, 2012; Ahmadian and Miller, 2019). In most theoretical studies, however, a single learning rule is implemented in a single class of synapses, while other synapse classes are set at reasonable initial values and held constant. In contrast, experimental data reveals that over the course of development and learning, most classes of synapses are plastic (Buonomano and Merzenich, 1998; Feldman, 2009; Froemke, 2015; Chiu et al., 2019). Thus, even in a simplified cortical circuit with only a single class of excitatory and inhibitory neurons, four synaptic learning rules must govern four classes of weights: $W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, $W_{I \leftarrow I}$.

The importance of carefully regulating multiple weight classes in parallel is particularly evident in cortical circuits that are defined in part by the presence of recurrent excitation (Douglas et al., 1995; Shepherd, 1998), which if not carefully held in check by inhibition can lead to epileptiform activity (McCormick, 1989; Steriade and Contreras, 1998)—notably circuits such as the striatum and cerebellum, which do not possess recurrent excitation do not exhibit epileptic foci. It is well established that many experimentally observed regimes require recurrent excitation, including Up-states (Steriade et al., 1993; Timofeev et al., 2000), asynchronous states (van Vreeswijk and Sompolinsky, 1998; Renart et al., 2010) and persistent activity associated with working memory (Fuster and Jervey, 1981; Goldman-Rakic, 1995; Wang, 2001). Experimental and theoretical evidence indicates that the self-sustaining persistent neural activity observed during Up-states is held in check by strong inhibition (van Vreeswijk and Sompolinsky, 1996; Shu et al., 2003; Renart et al., 2010). More generally, that cortical circuits can function as an inhibition-stabilized network in which increases in the firing rate of excitatory neurons are rapidly counterbalanced by inhibition (Tsodyks et al., 1997a; Ozeki et al., 2009; Rubin et al., 2015; Sanzeni et al., 2020b).

At the computational level self-sustained activity and inhibition-stabilized networks are often modeled as a simplified circuit composed of excitatory (*E*) and inhibitory (*I*) subpopulations of neurons with four classes of synaptic weights $W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, $W_{I \leftarrow I}$. Critically, in order to generate stable persistent activity these weights must obey certain theoretically well-defined relationships (Tsodyks et al., 1997a; Brunel, 2000; Ozeki et al., 2009; Rubin et al., 2015; Jercog et al., 2017). Here we address the question of how weights at all four synapse classes can be set in a self-organizing manner. Consistent with previous results we first show that cortical networks seem to be ontogenetically programmed to homeostatically generate Up-states. Specifically, while Up-states occur spontaneously *in vivo* during anesthesia, slow-wave sleep, and quiet wakefulness (Steriade et al., 1993; Timofeev et al., 2000; Beltramo et al., 2013; Hromádka et al., 2013), and in acute slices (Sanchez-Vives and McCormick, 2000; Shu et al., 2003; Fanselow and Connors, 2010; Sippy and Yuste, 2013; Xu et al., 2013; Sadovsky and MacLean,

2014; Neske et al., 2015; Bartram et al., 2017), they also emerge in organotypic cultures over the course of *ex vivo* development (Plenz and Kitai, 1998; Seamans et al., 2003; Johnson and Buonomano, 2007; Kroener et al., 2009; Motanis and Buonomano, 2015; Motanis and Buonomano, 2020). Furthermore, Up-state frequency appears to be homeostatically regulated—e.g., optogenetically stimulating cortical circuits over the course of days decreases Up-state frequency (Motanis and Buonomano, 2015; Motanis and Buonomano, 2020).

Some computational models have incorporated synaptic learning rules operating in parallel at multiple synapse classes (Lazar et al., 2009; Bauer et al., 2014; Binas et al., 2014; Mackwood et al., 2020). However, the issue of how four key synapse classes ($W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, $W_{I \leftarrow I}$) are governed across development to lead to self-sustained dynamics has not been addressed. One possibility is that standard homeostatic forms of plasticity could underlie the emergence of inhibition-stabilized networks. In homeostatic learning rules, such as synaptic scaling it is generally assumed that excitatory weights are up-regulated (down–regulated) when neurons are below (above) some ontogenetically determined set-point of average neural activity, and conversely that inhibitory weights are regulated in the opposite direction (Turrigiano et al., 1998; van Rossum et al., 2000; Kilman et al., 2002; Turrigiano and Nelson, 2004; Peng et al., 2010). Here we show that such standard homeostatic rules are inherently unstable in the context of inhibition-stabilized networks. This is in part a consequence of the so-called paradoxical effect, in which an external *increase* in the excitatory drive to inhibitory neurons produces a net *decrease* in steady-state firing rate of inhibitory neurons (Tsodyks et al., 1997a; Ozeki et al., 2009; Rubin et al., 2015). We show that the existence of the paradoxical effect constrains the learning rules that are capable of leading to inhibition-stabilized networks, and suggests the need for either "paradoxical" or nonlocal learning rules. We develop a family of homeostatic learning rules operating in parallel at all four weight classes, which lead to the unsupervised emergence of Up-states in the inhibition-stabilized regime. Importantly, we predict that for the learning rules to be local the plasticity in the inhibitory neurons should be, paradoxically, anti-homeostatic.

RESULTS

Up-states represent a transition from a quiescent state to a network-wide regime in which both excitatory and inhibitory neurons are active (**Fig. 1A**). During Up-states the firing rate of excitatory neurons is relatively low (1-5 Hz) indicating that recurrent excitation is held in check by appropriately tuned inhibition (Neske et al., 2015; Jercog et al., 2017; Romero-Sosa et al., 2020). By recording in organotypic cortical circuits it is possible to show that Up-states develop over the course of *ex vivo* development (Plenz and Kitai, 1998; Seamans et al., 2003; Johnson and Buonomano, 2007; Kroener et al., 2009; Motanis and Buonomano, 2015; Motanis and Buonomano, 2020). Early in development, at 8 days-in-vitro (DIV-8) most of the neurons are silent, while at later stages (DIV-23) spontaneous Up-states are observed (**Fig. 1B**). Over the first four-weeks of ex vivo development there

is an increase and stabilization of Up-state frequency and duration (**Fig. 1C-D**) (Johnson and Buonomano, 2007; Motanis and Buonomano, 2015). The fact that Up-states emerge autonomously during *ex vivo* development indicates that synaptic learning rules are in place to orchestrate the unsupervised emergence of Up-states.
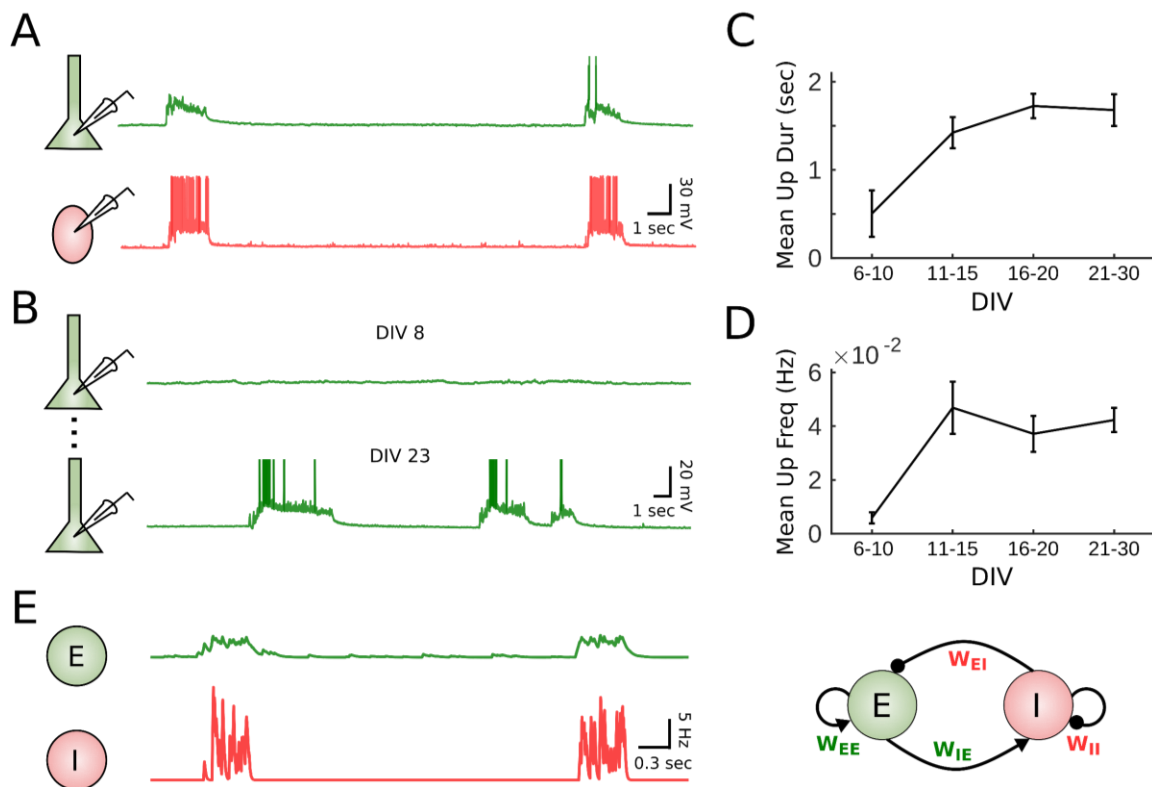

Figure 1. Up-states emerge autonomously over the course of *ex vivo* development.
   (A) Example of Up-states in simultaneously recorded pyramidal (green) and parvalbumin (PV) positive inhibitory neurons (red).
   (B) Spontaneous activity recording of a pyramidal neuron at 8 and 23 days *in vitro* development (DIV). Up-states are present only at later developmental stages.
   (C) Evolution of the mean Up-state duration over the course of *ex vivo* development.
   (D) Evolution of the mean Up-state frequency over the course of *ex vivo* development.
   (E) Two-population firing rate model of Up-states. The schematic of the model is shown on the right. The dynamics of the excitatory (green) and inhibitory (red) populations are governed by four synaptic weights, $W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, and $W_{I \leftarrow I}$. Traces correspond to the firing rate of each of the populations in the presence of external noise.

Computational studies have demonstrated that Up and Down states can be simulated as a bistable dynamical system composed of interconnected populations of excitatory (*E*) and inhibitory (*I*) neurons (**Fig. 1E**), in which Down-states represent a quiescent fixed-point, and Up- or asynchronous states represent a second non-trivial fixed-point attractor. In the Up regime recurrent excitation produces amplification, but the activity is held in check by rapid inhibition. The dynamics settles into a stable fixed-point attractor, and instantiates an example of an inhibition-stabilized network. The neural dynamics within two-population models is governed by four classes of synaptic weights $W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$,

$W_{I \leftarrow I}$. (**Fig. 1E,** right). Analytical and numerical studies have demonstrated that these four weights must obey certain "balanced" relationships in order to support the stable self-sustaining dynamics—e.g., if excitation is too strong, runaway (or saturated) excitation occurs, and if inhibition is too strong only the trivial quiescent fixed-point will be stable (Tsodyks et al., 1997b; van Vreeswijk and Sompolinsky, 1998; Brunel, 2000; Ozeki et al., 2009; Rubin et al., 2015; Jercog et al., 2017) (see Section 2.2 in the Supplementary Material).

In most computational models the set of four weights is determined analytically or through numerical searches. Since Up-states emerge autonomously over the course of development in *ex vivo* cortical networks, and because all four weight classes have been observed to undergo synaptic plasticity in experimental studies, we next asked how the stable self-sustained dynamics characteristic of Up-states might emerge in a self-organizing manner.

Standard homeostatic learning rules do not account for the emergence of Up-states

One attractive possibility is that cortical neurons are homeostatically programmed to generate Up-states. Specifically, that both excitatory and inhibitory neurons exhibit ontogenetically programmed firing rate setpoints, and they homeostatically adjust their excitatory and inhibitory weights to reach these target setpoints (Motanis and Buonomano, 2015; Motanis and Buonomano, 2020). Numerous homeostatic synaptic learning rules have been proposed based on experimental findings. Such rules are traditionally defined by changes in synaptic weights that are proportional to an "error term" defined by the difference between the setpoint and the average activity levels (Turrigiano et al., 1998; van Rossum et al., 2000; Kilman et al., 2002; Turrigiano and Nelson, 2004; Liu and Buonomano, 2009b; Peng et al., 2010; Vogels et al., 2011; Mackwood et al., 2020), e.g., $\Delta W_{E \leftarrow E} \propto E_{set} - E_{avg}$.

To determine if homeostatic learning rules can lead to stable self-sustained dynamics we started with the two-population model proposed by (Jercog et al., 2017) (see Methods), but instead of setting the weights manually, we initialized weights at random values and applied a family of homeostatic learning rules to all four weights classes (**Fig. 2A**). We initially assumed, based on experimental data, that during Up-states the *E* and *I* populations exhibited target setpoints at 5 and 14 Hz, respectively (Romero-Sosa et al., 2020). We first examined whether a set of four learning rules can lead to a transition from a quiescent Down-state to a stable self-sustained dynamic regime (representing a permanent Up-state) in response to a brief external input (initially low levels of noise were used to avoid spontaneous Up↔Down transitions).
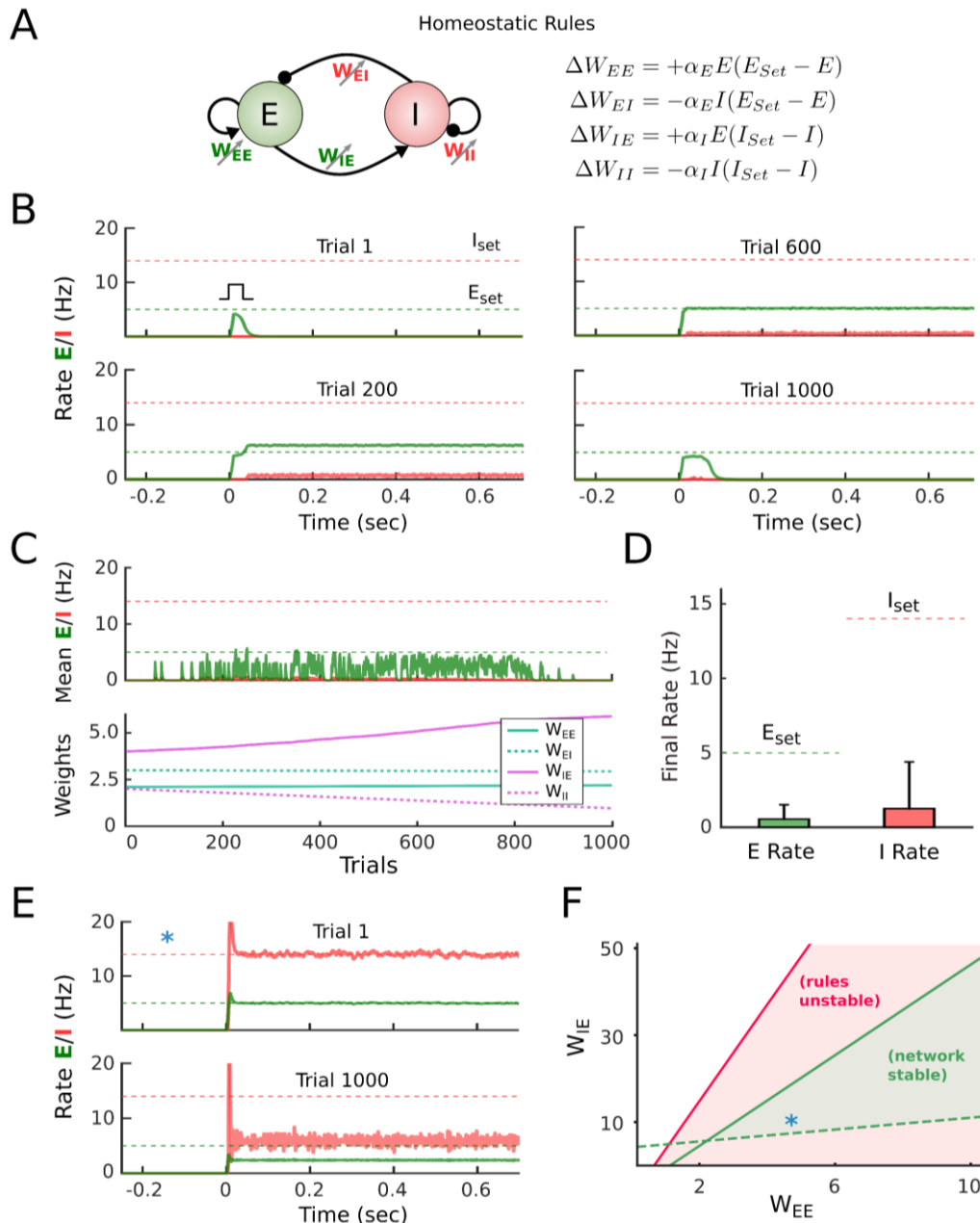
Figure 2. Standard homeostatic rules are unstable and do not lead to the emergence of self-sustained activity.

(A) Schematic (left) of the population rate model in which the four weights are governed by standard homeostatic learning rules (right)

(B) Example simulation of the network over the course of simulated development. The evolution of the firing rate of the excitatory and inhibitory population within a trial in response to a brief external input is shown in every plot. $E_{set}=5$ and $I_{set}=14$ represent the target homeostatic setpoints. Weights are initialized to $W_{EE}=2.1$ $W_{EI}=3$ $W_{IE}=4$ $W_{II}=2$. Note that while an evoked Up-state emerges by Trial 200 the firing rates do not converge to their setpoints, and by Trial 1000 the Up-state is now longer observed.

(C) Average rate across trials (upper plot) for the excitatory and inhibitory populations for the data shown in (B). Weight dynamics (bottom plot) produced by the homeostatic rules across trials for the data shown in (B).

(D) Average final rate for 100 independent simulations with different weight initializations.

(E) Simulation of a network starting with weights that generate Up-states that match the $E_{set}=5$ and $I_{set}=14$ Hz (Trial 1, top). After 1000 trials the network has diverged from its setpoints, indicating the synaptic learning rules are unstable. Weights were initialized to $W_{EE}=5$ $W_{EI}=1.09$ $W_{IE}=10$ $W_{II}=1.54$.

(F) Numerical solution for the analytical condition of stability of the neural system and learning rule system as a function of W_{EE} and W_{IE}. Blue asterisk corresponds to the initial conditions shown in Panel E (top).

Although the rules are homeostatic in nature (e.g., if I is below I_{set}, an increase of $W_{I \leftarrow E}$ and a decrease in $W_{I \leftarrow I}$ would be induced), they fail to lead the network to converge to stable Up-states (**Fig. 2B-D**). At the beginning of the experiment (Trial 1), an external input to the excitatory population does not engage recurrent activity because $W_{E \leftarrow E}$ is too weak. By Trial 200 the weights have evolved and although the external input now generates an apparently stable Up-state, the activities E and I do not match the corresponding setpoints and thus the weights keep evolving —specifically, the network is in a nonbiologically observed regime in which $E > I$. By Trial 600 $E = E_{set}$ but $I < I_{set}$, and rather than converging to I_{set}, the network returns to a regime without an Up-state by Trial 1000. A that point both setpoint "error terms" have increased, keeping the weight dynamics going. Results across 100 simulations with different weight initializations (see Methods) further indicates that the standard homeostatic rules are ineffective at driving E and I towards their respective setpoints and generating stable self-sustained dynamics (**Fig. 2D**).

In order to support our numerical simulations, we analytically characterized the stability of the family of standard homeostatic learning rules. Overall, a network in which the weights undergo plasticity can be characterized as a dynamical system composed of two subsystems: the neural subsystem composed of the two differential equations that define E and I dynamics, and the synaptic learning rule subsystem defined by the four learning rules (see Supplementary Material). We make use of the two very different time scales of the neural (fast) and learning rule (slow) subsystems to perform a quasi-steady state approximation of the neural subsystem; then we compute the eigenvalues of the four-dimensional learning rule subsystem, and finally get an analytical expression for the stability condition of the learning rules (see Section 2.3 in the Supplementary Material). For the entire system to be stable, both the neural and learning rules subsystems have to be stable. We found that the standard homeostatic rules are unstable for biologically meaningful parameter values in which the neural system is stable. An example of this result is shown in **Fig. 2E-F** for a particular set of parameter values (see Methods). Here the network is initialized at a condition satisfying the homeostatic setpoints E_{set} and I_{set} (**Fig. 2E**). Although the neural subsystem is stable at this condition (Trial 1), the standard homeostatic rules drive the weight values and the average activity of the network away from the setpoints (Trial 1000). This simulation is represented by the blue asterisk in **Fig. 2F** where the theoretical regions of stability are shown (see Supplementary Material for a detailed explanation). Critically, **Fig. 2F** shows that the stability region of the neural subsystem, i.e., an inhibition-stabilized network as in (Ozeki et al., 2009; Jercog et al., 2017), is entirely within the region where the standard homeostatic learning rules are unstable. We therefore conclude that the standard homeostatic rules do not display a dynamical regime that would lead to the emergence of stable Up states in an inhibition-stabilized network. Similarly, a combination of analytical and numerical methods also indicates that variants of these homeostatic rules, such as Synaptic Scaling (Turrigiano et al., 1998; van Rossum et al., 2000; Sullivan and de Sa, 2006) are also unstable (see Supplementary Material, Section 2.6). These results suggest that, in the context of inhibition-stabilized networks, "standard" homeostatic learning rules are inherently unstable and cannot account for the emergence of Up states.

The paradoxical effect constrains the learning rules that can lead to Up-states.

The inability of the homeostatic learning rules to generate stable Up-states is in part a consequence of a counterintuitive, yet well described, property of the two-population models of Up-states and inhibition-stabilized networks: the so-called paradoxical effect (Tsodyks et al., 1997a; Rubin et al., 2015; Sanzeni et al., 2020a). Specifically, if during an Up-state one increases the external drive to the inhibitory population, the net result will counterintuitively be a decrease in the firing rate of the inhibitory units. This paradoxical effect can be understood in terms of the $I{\to}E{\to}I$ loop: the increased inhibitory drive should lead to a lower steady-state rate for E, but this new steady-state value also relies on a given E/I balance, thus there must be a parallel decrease in the I firing rate (in other words, the decrease in E decreases the drive to I by more than the external increase to I). This paradoxical effect has profound consequences for learning rules that attempt to drive excitatory and inhibitory weights to their setpoints.
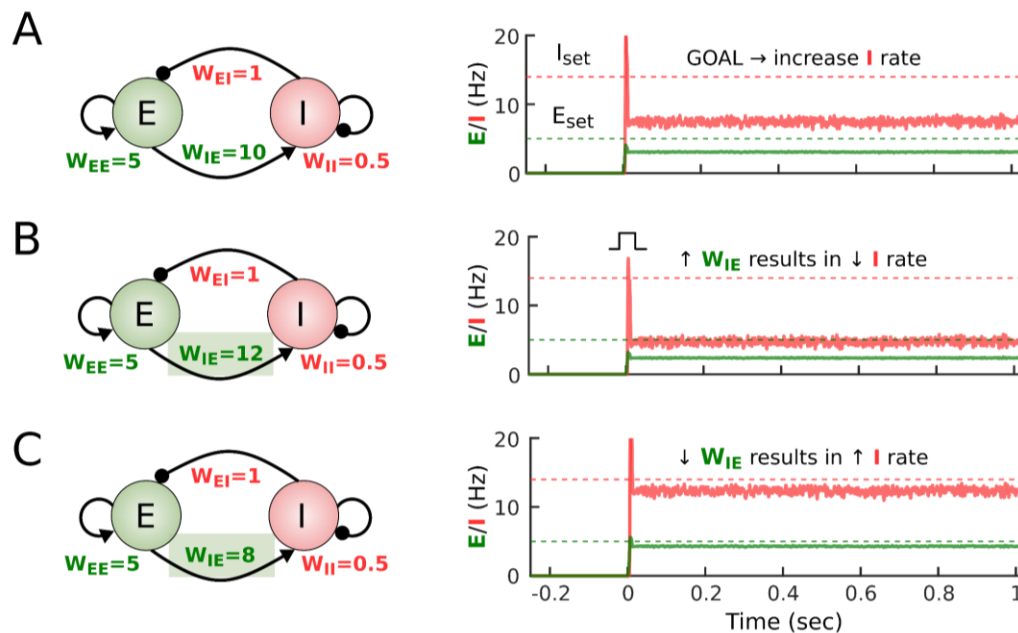

Figure 3. The paradoxical effect constrains the learning rules that can lead to Up-states.
(A) Example of the self-sustained dynamics of a two-population model with weight values shown in the diagram. Both the E and I firing rates fall below their respective setpoints. The objective is to adjust the weights so that the E and I activity match their setpoints.
(B) An increase of W_{IE} from 10 to 12 results in a paradoxical decrease of the I rate.
(C) Paradoxically, an effective way to increase the steady-state I firing rate is to decrease its excitatory drive, i.e., W_{IE}.

Consider a network state in which the I rate falls significantly below its setpoint, and the E rate is close to its setpoint (**Fig. 3A**). In order to reach the I setpoint a homeostatic manipulation on the inhibitory neuron would intuitively result in an increase of $W_{I{\leftarrow}E}$. However, as it can be seen if we increase $W_{I{\leftarrow}E}$ (**Fig. 3B**), the inhibitory population actually decreases because of the paradoxical effect—thus increasing the error term $I_{set}{-}I$. To

increase the steady-state inhibitory rate we can decrease the excitatory weight onto the inhibitory neurons (**Fig. 3C).** This simple example shows the complexity of designing a coherent set off rules in such a coupled system (see an analysis of the paradoxical effect in Section 2.2.3 of the Supplementary Material).

A novel cross-homeostatic rule robustly leads to the emergence of self-sustained Up-states

Given that a standard set of homeostatic learning rules does not account for the development of Up-states, we attempted to identify alternative learning rules. By defining a loss function based on the difference between E and I and their respective setpoints we derived a set of learning rules using gradient descent. This approach led to mathematically complex and biologically implausible rules; however, approximations and simulations inspired a simple class of learning rules that we will refer to as cross-homeostatic (see Methods). The main characteristic of this set of rules is that the homeostatic setpoints are "crossed" (**Fig. 4A)**. Specifically, the weights onto the excitatory neuron ($W_{E \leftarrow E}$ and $W_{E \leftarrow I}$) are updated to minimize the inhibitory error while weights into the inhibitory neuron ($W_{I \leftarrow E}$ and $W_{I \leftarrow I}$) change to minimize the excitatory error. Although apparently non-local, these rules can be interpreted as the total inhibitory and excitatory input current into the cell. Such input could be read by the cell via activation of metabotropic mGlu and GABA$_B$ receptors. Indeed, similar local rules have been derived for $W_{E \leftarrow I}$ and $W_{I \leftarrow E}$ weights in models with only two plastic weights (Mackwood et al., 2020).

An example of the performance of the cross-homeostatic rules is shown in **Fig. 4B-C**. After an initial phase with no sustained firing rate (Trial 1), recurrent activity reaches a stable Up-state (Trial 25), whose average rate continues to converge towards its defined setpoints (Trial 300) until the learning rule system reaches steady state (Trial 1000). The average E and I rates of the network evolve asymptotically towards the defined setpoints, as the weights evolve and converge (**Fig. 4C)**. Across different weight initializations the rules proved effective in driving the mean Up-state activity of the network to the target E and I setpoints (**Fig. 4D**). These results establish that this family of cross-homeostatic rules leads to inhibition-stabilized networks and is robustly stable.
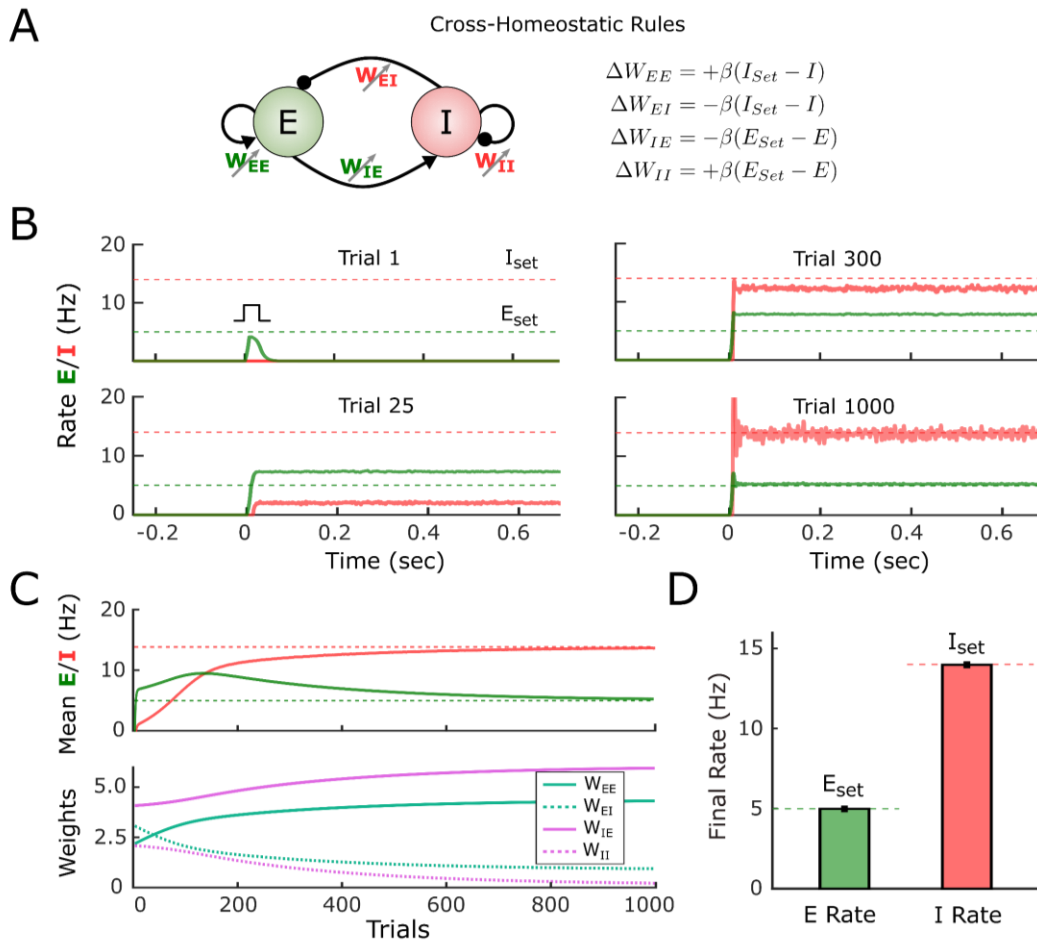
Figure 4. A family of cross-homeostatic learning rules lead to self-sustained dynamics at E_{set} and I_{set}.

(A) Schematic of the model and the family of cross-homeostatic learning rules.

(B) Example network dynamics across simulated development. The network is initialized with weights that do not lead to self-sustained dynamics in response to an external input (Trial 1). By Trial 25 a stable Up-state is observed, but at firing rates far from the target setpoints (dashed lines). By Trial 1000 the network has converged to an Up-state in which E and I firing rate match their respective setpoints. Weights are initialized to W_{EE}=2.1 W_{EI}=3 W_{IE}=4 W_{II}=2.

(C) Average rate across trials (upper plot) for the excitatory and inhibitory populations for the data shown in (B). Weight dynamics (bottom plot) induced by the homeostatic rules across trials for the data shown in (B)

(D) Average final rate for 100 independent simulations with different weight initializations (see Methods).

To contrast the mechanisms underlying the instability of the standard homeostatic rules and stability of the cross-homeostatic rules, we compared the trajectories of the weights across different experiments for both rules (**Fig. 5A-D**). The weight trajectory from its initial value to its final one is shown for 100 different simulations. Each line corresponds to individual experiments with different weight initializations. Circles indicate the final values of the weights. Independently of the initial conditions, the weights converge to a line attractor (actually a 2D plane attractor in 4D weight space) for the cross-homeostatic rules (**Fig. 5B-D**), but not for the standard homeostatic rules (**Fig. 5A-C**). Note that this attractor refers to the sets of weights that generate Up-states where E and I activity matches E_{set} and I_{set} respectively. That is, for a given pair of setpoints (E_{set}, I_{set}) the final values of the weights $W_{E \leftarrow I}$ and $W_{I \leftarrow I}$ are linear functions of $W_{E \leftarrow E}$ and $W_{I \leftarrow E}$. This is a

direct consequence of the steady state conditions for the nontrivial fixed-point of the two-population model (Tsodyks et al., 1997a; Ozeki et al., 2009; Jercog et al., 2017), where the slope of the line is defined by the setpoints E_{set}/I_{set} (**Fig. 5C-D**, see Methods). Numerical simulations confirm that the cross-homeostatic, but not the standard, homeostatic rule robustly guides Up-states to different E_{set} and I_{set} setpoints (**Fig. 5E-F**).
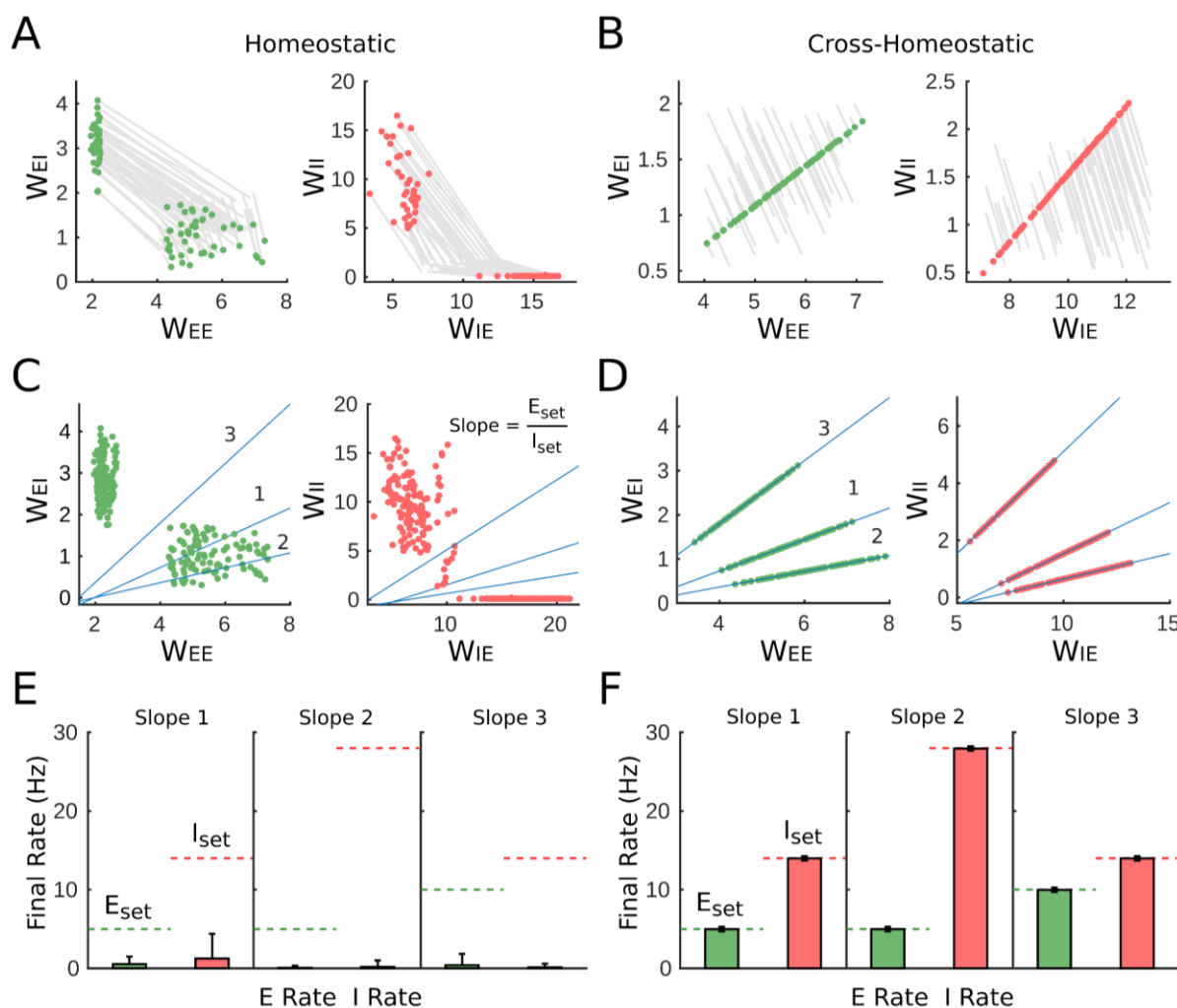

Figure 5. Cross-homeostatic rules lead to stable balanced dynamics while homeostatic rules fail to do so.
    (A) Weight trajectories for 100 different simulations with random weight initializations for the homeostatic rules. Lines show change from initial to final (circles) weight values.
    (B) Trajectories shown for the cross-homeostatic rules for the same initial conditions as in (A).
    (C) Final weight values for homeostatic plasticity simulations with same starting conditions as in (A) and (B) but for additional homeostatic pairs of setpoints. Line 1: $E_{set}=5$, $I_{set}=14$; Line 2: $E_{set}=5$, $I_{set}=28$; Line 3: $E_{set}=10$, $I_{set}=14$. Data shown in (A) and (B) corresponds to Line 1. Blue lines correspond to the theoretical linear relationship between the excitatory and inhibitory weights at a fixed point obeying E_{set} and I_{set}. The slope of the line is defined by the ratio of the setpoints (see Methods).
    (D) Final weight values for cross-homeostatic plasticity simulations under the same initial conditions as in (C).
    (E) Final rate for the excitatory and inhibitory population after homeostatic learning. The different bar plots correspond to the data shown in (C), where 3 different pairs of setpoints are tested.
    (F) Final rate for the excitatory and inhibitory population for the cross-homeostatic learning from data in (D).

To further validate the effectiveness and stability of the cross-homeostatic rule we again used analytic methods to determine the eigenvalues of the 4-dimensional learning-rule dynamical system governed by the four cross-homeostatic rules. As above, stability is determined by the sign of the real part of the eigenvalues of the system. It can be shown (see Section 2.4 in the Supplementary Material) that this learning rule is stable for any set of parameter values, provided that the stability conditions of the neural subsystem are satisfied.

Therefore, it is possible to formally establish that the cross-homeostatic learning rules are inherently stable and can account for the emergence and maintenance of inhibition-stabilized self-sustained regimes and Up-states.

A locally balanced homeostatic rule also accounts for the emergence of Up-states

The solutions achieved by the cross-homeostatic rule highlight the fundamental relationship between the $W_{E \leftarrow E}/W_{E \leftarrow I}$ and $W_{I \leftarrow E}/W_{I \leftarrow I}$ ratios, which is analytically imposed by the relationship among weight values and firing rates for Up-states to exist (whether stable or not; see Methods). For example, to satisfy $\frac{dE}{dt} = 0$ in the Up-state fixed point, the net excitation and inhibition must obey a specific "balance", meaning that once $W_{E \leftarrow E}$ or $W_{E \leftarrow I}$ is determined, the other is analytically constrained for a given set of setpoints and parameters. This implies that the 4-dimensional system defined by the four learning rules can be reduced to 2-dimensions—because $W_{E \leftarrow I} \propto W_{E \leftarrow E}$ and $W_{I \leftarrow I} \propto W_{I \leftarrow E}$. We thus exploited this constraint to design a simple set of local rules that would lead to the emergence of self-sustained Up-states. Here $\triangle W_{E \leftarrow E}$ follows the standard homeostatic learning rule, while $\triangle W_{I \leftarrow E}$ follows a local anti-homeostatic rule (**Fig. 6A**). $W_{E \leftarrow I}$ and $W_{I \leftarrow I}$ converge to linear functions of $W_{E \leftarrow E}$ and $W_{E \leftarrow I}$, respectively, so their change is always "locked" to the evolution of the plastic weights. Because of this imposed relationship, we will refer to this as the balanced-homeostatic rule. We interpret the plasticity of the $W_{E \leftarrow I}$ and $W_{I \leftarrow I}$ to rely on ontogenetically programed ratios of $W_{E \leftarrow E}$ and $W_{I \leftarrow E}$ weights, respectively. This interpretation is supported by observations that after the induction of excitatory plasticity the inhibitory weights can slowly "rebalance" (Froemke et al., 2007; Froemke, 2015)

As shown in **Fig. 6B**, the balanced-homeostatic rule can drive the activity of the network to a stable Up-state. Starting in a weight regime that does not support self-sustained activity (Trial 1), the rule initially leads to a self-sustained Up-state in which E is close to E_{set} but I is well below I_{set} (Trials 70 and 600). This initial developmental stage is followed by asymptotic convergence to I_{set}. (Trial 1000). Note that due to the linear relationship imposed on the weights, a minimum weight must be assigned to $W_{E \leftarrow E}$ and $W_{I \leftarrow E}$ in order to prevent negative weights (see Methods)—which is why we initialize $W_{I \leftarrow E}$ at higher values than in **Fig. 2** and **4**. Across 100 simulations all networks converged to the target setpoints (**Fig. 6D**). As expected, the balanced-homeostatic rule also drove the ($W_{E \leftarrow E}$, $W_{E \leftarrow I}$) and ($W_{I \leftarrow E}$, $W_{I \leftarrow I}$) weight pairs to the stable line attractors across a range of setpoints (**Fig. S1**).
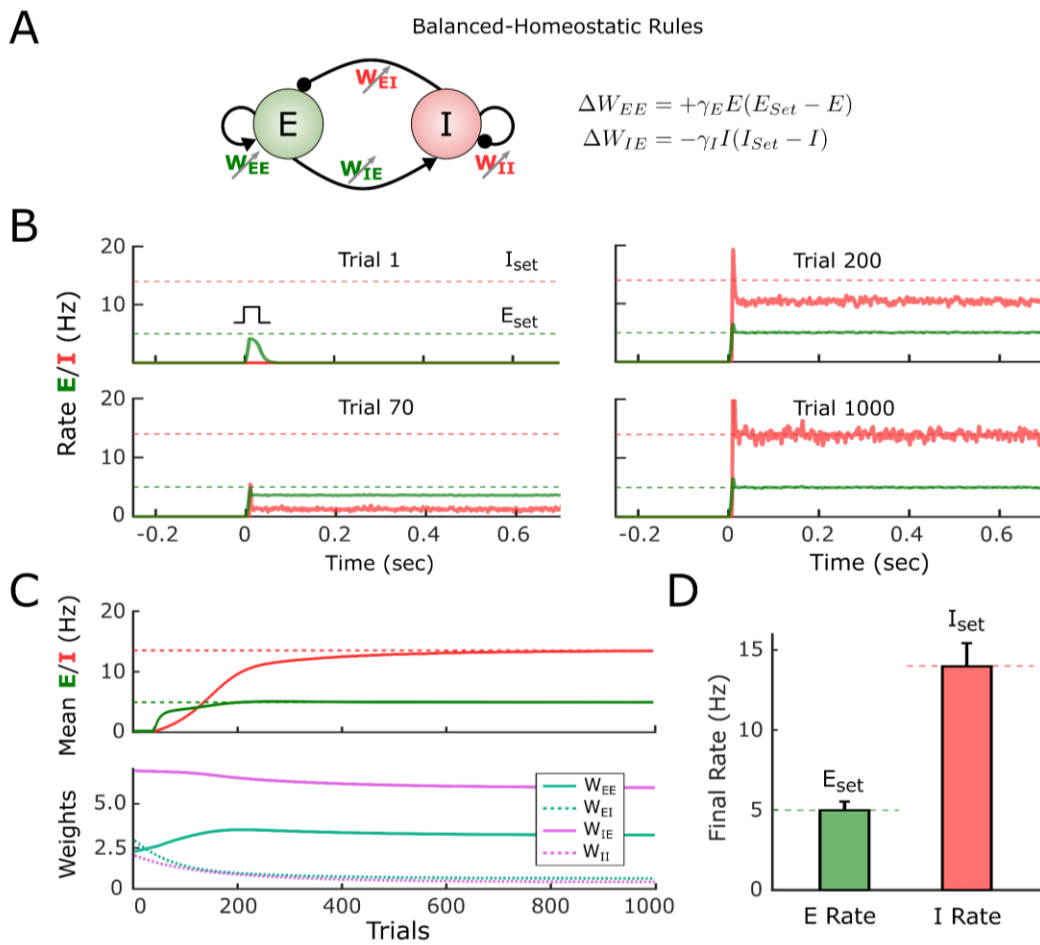
Figure 6. Balanced-homeostatic rules also lead to self-sustained dynamics at E_{set} and I_{set}.

(A) Schematic of the model and the family of cross-homeostatic learning rules.

(B) Example network dynamics across development. The network is initialized with weights that do not lead to self-sustained dynamics in response to an external input (Trial 1). By Trial 70 a stable Up-state is observed, but at firing rates far from the target setpoints (dashed lines). By Trial 1000 the network has converged to an Up-state in which the E and I firing rates match their respective setpoints. Weights are initialized to W_{EE}=2.1 W_{EI}=3 W_{IE}=7.5 W_{II}=2.

(C) Average rate across trials (upper plot) for the excitatory and inhibitory populations for the data shown in (B). Weight dynamics (bottom plot) governed by the balanced-homeostatic rules trials for the data shown in (B)

(D) Average final rate for 100 independent simulations with different weight initializations (see Methods).

As above, we developed analytic expressions for the stability of the Balanced-homeostatic rules. This analysis demonstrated that this set of rules is stable for biologically meaningful parameter values, and that its stability region overlaps with the stability region of the neural subsystem for infinitely many values of the two free weights $W_{E \leftarrow E}$ and $W_{I \leftarrow E}$ (see Section 2.5 in the Supplementary Material).

Simulating the empirically observed development of Up-states

In the previous simulations the inhibitory and excitatory populations converge to a permanent Up-state (i.e., an asynchronous state) because in the absence of significant noise there were no spontaneous Up↔Down transitions. We thus next simulated the developmental emergence of spontaneous Up-states in the model with the cross-homeostatic learning rules, and contrasted the performance of the model with the experimentally observed developmental time-course of Up-state frequency and duration (**Fig.1 C-D**). Following Jercoq et al. (2017) we increased the noise levels and included an adapting current to the excitatory population to capture Up→Down transitions. With initial weights that correspond to early developmental stages which lack any spontaneous Up-state, the network undergoes a progressive increase in spontaneous Up↔Down transitions (**Fig. 7A**), with the average Up rate converging to the defined setpoints (**Fig. 7B**). The network converged to the setpoints across multiple simulations (**Fig. 7C**) and the development of Up-state frequency and duration paralleled the experimentally observed developmental time-course (**Fig. 7D**).
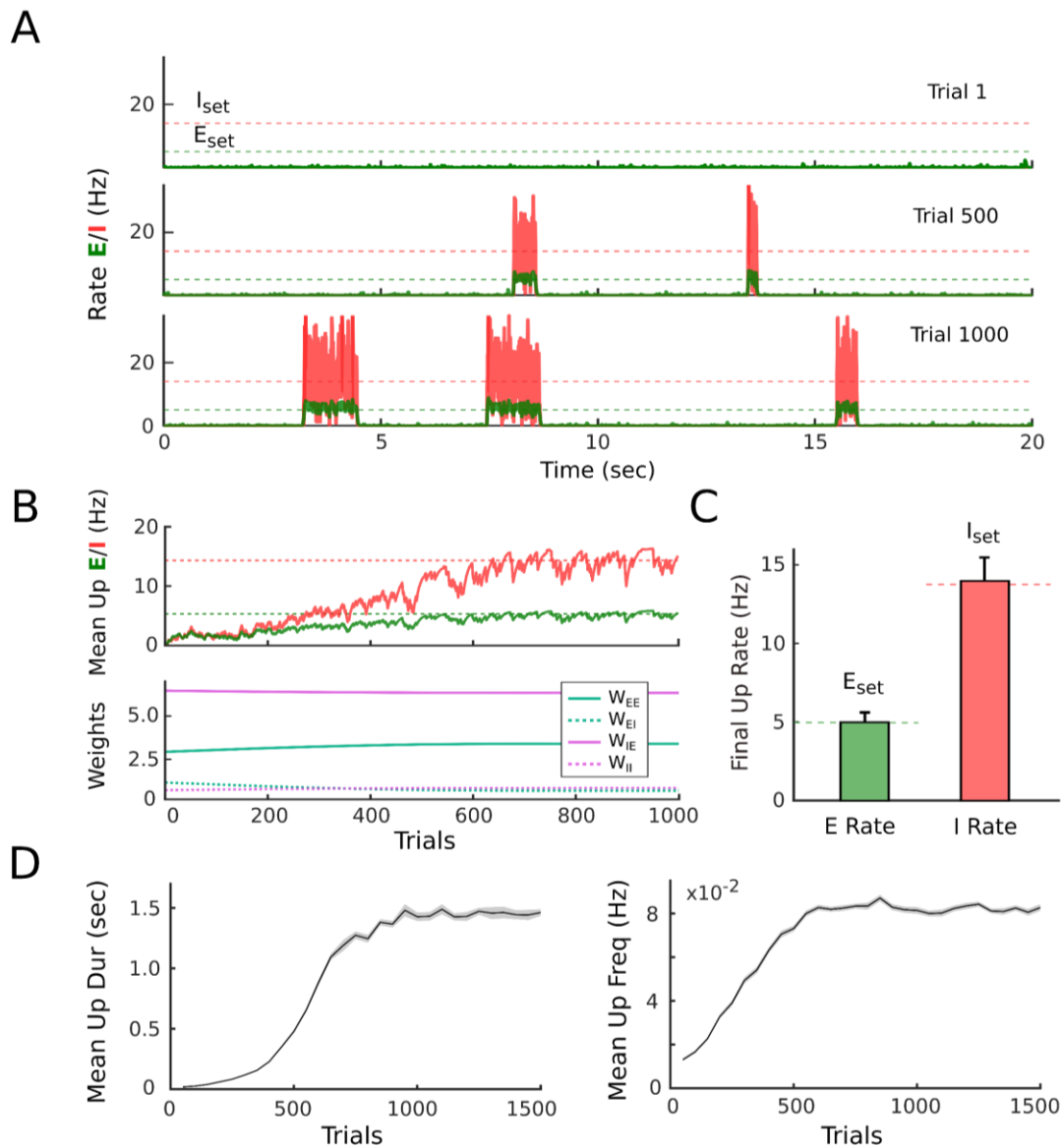
Figure 7. Orchestrated excitatory and inhibitory rules lead to the emergence of stable Up↔Down transitions.

(A) Simulation of the development of spontaneous Up-states. The weights of the two-population model evolve following cross-homeostatic rules and under the presence of high external noise and intrinsic excitatory adaptation. Under these conditions, stable Up↔Down transitions emerge (Trial 1000) over the course of simulated development from a quiescent stage (Trial 1). Initial weights are W_{EE}=3 W_{EI}=1 W_{IE}=7 W_{II}=0.5.

(B) Average rate during the Up-states for both neural populations (upper plot) for the simulation shown in (A). Bottom plot shows the evolution of the weights.

(C) Final mean Up rate for 20 independent simulations. Initial weights are specified in (A).

(D) Average Up duration and frequency over the course of simulated development. Averages are computed every 50 trials for 20 independent simulations. Shaded area shows ± SEM.

DISCUSSION

Elucidating the learning rules that govern the connectivity of neural circuits represents a critical goal in neuroscience, because learning rules provide a link between the cellular, systems, and computational levels of analyses. For example, elucidation of Hebbian associative synaptic plasticity linked simple computations at the level of single proteins (the NMDA receptor) with higher-order computations at the level cortical networks and cognition (Hebb, 1949; Miller et al., 1989; Buonomano and Merzenich, 1998; Martin et al., 2000; Song et al., 2000). To date however, most learning rules have been studied primarily in the context of a single synapse class—e.g., how associative and homeostatic plasticity operating at $W_{E \leftarrow E}$ synapses (while holding other weights constant) contribute to cortical map formation. Yet, it is widely accepted that virtually all synapse classes undergo plasticity governed by class-specific learning rules (Abbott and Nelson, 2000; Froemke, 2015; Hennequin et al., 2017; Chiu et al., 2019). Thus, it is necessary to understand how many learning rules operating in parallel lead to functional dynamical regimes and computations. Here we have taken a first step towards addressing this problem by examining how independent learning rules operating at four different synapse classes can capture the experimentally observed emergence of Up-states in cortical networks.

Here we focused on the emergence of Up-states because: 1) they represent one of the best examples of both self-sustained dynamics and of an inhibition-stabilized network regime; and 2) Up-states emerge autonomously during *ex vivo* development, and are homeostatically modulated by external activity (Plenz and Kitai, 1998; Seamans et al., 2003; Johnson and Buonomano, 2007; Kroener et al., 2009; Motanis and Buonomano, 2015; Motanis and Buonomano, 2020)—thus suggesting that there are homeostatic learning rules in place to guide cortical networks to exhibit Up-states.

Our results first demonstrated that standard homeostatic rules that attempt to drive excitatory and inhibitory neurons to target setpoints during Up-states, are ineffective at generating stable self-sustained dynamics. Both numerical simulations (**Fig. 2**) and analytical analyses (see Supplementary Material) reveal that a system of four homeostatic learning rules is unable to reliably and stably drive excitatory and inhibitory neurons to their respective setpoints. Indeed, we showed that, despite the fact that an Up-state can exist under these rules (i.e., a set of weight values making $I=E_{set}$ and $I=I_{set}$ in the absence of noise), the standard homeostatic family of learning rules cause the network to diverge away even at the smallest amount of noise because the Up-state is unstable under these rules (**Fig. 2E-F**). These results emphasize the need to understand the experimentally described forms of homeostatic plasticity in specific contexts, and whether they are symmetrically present in the excitatory and inhibitory branches of cortical circuits.

The paradoxical effect and paradoxical learning rules.

One of the reasons standard homeostatic rules operating at all four synapses struggle with generating stable self-sustained dynamics is related to the paradoxical effect (**Fig.**

3). Up-states represent a nontrivial fixed-point in which recurrently generated excitation is counterbalanced by the net inhibition through the $E{\rightarrow}I{\rightarrow}E$ branch. As a result of this coupled negative feedback, if the excitatory drive to the inhibitory population is increased there will be a decrease in excitatory firing rate which in turn will decrease the firing of the inhibitory population. The paradoxical effect establishes that the decrease excitatory input to the inhibitory population that arises from the decreased rate of excitatory neurons must be larger than the initial increased drive of the inhibitory population (Tsodyks et al., 1997a; Ozeki et al., 2009). This effect has profound consequences for the unsupervised emergence of Up-states, e.g., if there is an ontogenetically determined Up-state I_{set}, it is not sufficient to simply homeostatically adjust $W_{I{\leftarrow}E}$ and $W_{I{\leftarrow}I}$ to reach this setpoint (**Fig. 3,** Section 2.2.3 of the Supplementary Material).

A family of cross-homeostatic rules provide a solution to the paradoxical effect and can robustly drive networks to self-sustained dynamic regimes (**Fig. 4-5**). As implemented here, this family of rules has the counterintuitive and nonlocal property of having "cross-homeostatic setpoints"—e.g., the error terms in $W_{E{\leftarrow}E}$ and $W_{E{\leftarrow}I}$ correspond to I_{set}-I. As has been pointed out, such nonlocal cross-homeostatic rules can be implemented as a local rule in which the excitatory neurons are programmed to seek a target amount of inhibitory synapse activation (Mackwood et al., 2020). Although this hypothesis has not been directly experimentally tested, it seems inconsistent with experimental studies that have used bicuculline to increase network activity, and have observed weaker putative $W_{E{\leftarrow}E}$ synapses (Turrigiano et al., 1998; Turrigiano and Nelson, 2004).

Therefore, we further considered a balanced homeostatic rule that proposes that the excitatory weights onto excitatory and inhibitory neurons are governed by local homeostatic and anti-homeostatic rules, respectively. Again, as a consequence of the paradoxical effect anti-homeostatic regulation of the inhibitory population is more effective than homeostatic regulation. This rule proposes that there are ontogenetically determined balances which determine the strength of $W_{E{\leftarrow}I}$ and $W_{I{\leftarrow}I}$ synapses in relation to the $W_{E{\leftarrow}E}$ and $W_{I{\leftarrow}E}$ synapses, respectively.

Conclusions and experimental predictions

The balanced-homeostatic rule is consistent with current experimental data with respect to the $W_{E{\leftarrow}E}$ weight (Turrigiano et al., 1998; Burrone et al., 2002; Turrigiano and Nelson, 2004; Goel and Lee, 2007). While homeostatic plasticity of $W_{E{\leftarrow}I}$, $W_{I{\leftarrow}E}$, and $W_{I{\leftarrow}I}$ is more complex and dependent on inhibitory cells types (Hartman et al., 2006; Xue et al., 2014; Joseph and Turrigiano, 2017), to the best of our knowledge there is little evidence for the notion of anti-homeostatic plasticity of the $W_{I{\leftarrow}E}$ class. But it is important to stress that there are likely multiple forms of homeostatic plasticity being driven by multiple homeostatic setpoints (Liu et al., 1998; O'Leary et al., 2013; O'Leary et al., 2014). Specifically, in the context of the current work, it is critical to distinguish between the Up-state firing rate setpoint, and the more conventional average activity setpoint. Whereas these two setpoints clearly interact, it remains unclear if neurons have independent setpoints for average activity and Up-state firing rate. Thus, it remains to be determined

if the existence of anti-homeostatic plasticity is observed in inhibitory neurons in response to shifts in Up-state activity levels.

There is, however, substantial experimental evidence that support the prediction that plasticity of excitatory synapses drives a "rebalancing" of inhibitory weights—consistent with the notion of an ontogenetically determined balance (Froemke et al., 2007; Froemke, 2015; Antoine et al., 2019; Chiu et al., 2019). For example, after induction of excitatory plasticity inhibitory synapses onto excitatory neurons "rebalance" over the course of minutes (Froemke et al., 2007). Furthermore, the same NMDA-dependent protocols that induce potentiation of EPSPs can also result in potentiation of IPSPs (Chiu et al., 2018; Field et al., 2020), and on the anatomical level there is a correlation between the number of inhibitory and excitatory anatomic synaptic contacts on every dendritic compartment (Iascone et al., 2020).

Together our findings demonstrate that even if a given learning rule in isolation generates beneficial properties, it is not necessarily the case that the same rule will be effective in the context of a recurrent excitatory/inhibitory network governed by multiple rules operating in parallel. Furthermore, because emergent neural dynamic regimes are highly nonlinear, and in particular, that stable self-sustained dynamic regimes exhibit a paradoxical effect in these networks, it is likely that the brain exhibits "paradoxical" or anti-homeostatic learning rules, such as those proposed here, to generate self-sustained dynamic regimes.

METHODS

Ex vivo slice preparation and Electrophysiology

The experimental results shown in Figure 1 relied on the reanalysis of previously published data (Motanis and Buonomano, 2015; Romero-Sosa et al., 2020). All animal experiments followed guidelines established by the National Institutes of Health (NIH) and were approved by the Chancellor's Animal Research Committee at the University of California, Los Angeles.

Organotypic slices were prepared using the interface method (Stoppini et al., 1991). Briefly, five to seven day-old WT mice were anesthetized with isoflurane and decapitated. The brain was removed and placed in chilled cutting media. Coronal slices (400 µm thickness) containing auditory cortex were sliced using a vibratome (Leica VT1200) and placed on filters (MillicellCM, Millipore, Billerica, MA, USA) with 1 mL of culture media. Culture media was changed at 1 and 24 hours after cutting and every 2-3 days thereafter. Cutting media consisted of EMEM (MediaTech cat. #15-010) plus (final concentration in mM): $MgCl_2$, 3; glucose, 10; HEPES, 25; and Tris-base, 10. Culture media consisted of EMEM plus (final concentration in mM): glutamine, 1; $CaCl_2$, 2.6; $MgSO_4$, 2.6; glucose, 30; HEPES, 30; ascorbic acid, 0.5; 20% horse serum, 10 units/L penicillin, and 10 µg/L streptomycin. Slices were incubated in 5% $CO2$ at 35°C.

Whole cell recordings were made from LII/III pyramidal regular spiking neurons, identified using infrared differential interference contrast visualization. Recordings were performed at 7-30 days *in vitro* (DIV) in artificial cerebrospinal fluid composed of (in mM): 125 NaCl, 5.1 KCl, 2.6 $MgSO_4$, 26.1 $NaHCO_3$, 1 NaH_2PO_4, 25 glucose, and 2.6 $CaCl_2$ (ACSF was formulated to match the standard culture media). The temperature was maintained at 32-33 °C, and the perfusion rate set at 4.5-5 mL/min. The internal solution for whole-cell recordings contained (in mM): 100 K-gluconate, 20 KCl, 4 ATP-Mg,10 phospho-creatine, 0.03 GTP-Na, and 10 HEPES and was adjusted to pH 7.3 and 300 mOsm. In order to be considered for analysis, cells had to have a resting potential less than -55 mV and not change by more than 15% over the course of the recording. The criteria for input and series resistance were 100-300 MΩ and <25 MΩ, respectively.

For the data shown in Figure 1A, PV-Cre mice were used. The slices were transfected at 5-7 DIV with 1 µL of LSL-tdTomato (Addgene#: 100048) and recordings were performed 16-18 days after transfection. Fluorescent PV neurons were targeted using a 430 nm LED to visualize tdTomato.

Computational model

A two-population firing-rate model was implemented based on Jercoq et al (2017). The firing rate of the excitatory (E) and inhibitory (I) population obeyed Wilson and Cowan dynamics (Wilson and Cowan, 1972):

$$(1) \quad \tau_E \frac{dE}{dt} = -E + f_E\big(W_{EE}E(t) - W_{EI}I(t) - a + \eta_E(t)\big)$$

$$(2) \quad \tau_I \frac{dI}{dt} = -I + f_I\big(W_{IE}E(t) - W_{II}I(t) + \eta_E(t)\big)$$

where W_{XY} represents the weight between the presynaptic unit Y and postsynaptic unit X. τ_X and η_X represent a time constant and an independent noise term, respectively. The time constants were set to $\tau_E = 10\ ms$ for the excitatory and $\tau_I = 2\ ms$ for the inhibitory population. The noise term was an Ornstein-Uhlenbeck process with a time constant of $1\ ms$ and standard-deviation σ_X. All simulations had a $\sigma_X = 0.1\ Hz$ except for Figure 7 where $\sigma_X = 1.2\ Hz$ was used. To elicit Up-states a step current was injected at the beginning of each trial on the excitatory population, except for Figure 7, where Up-state transitions were spontaneously elicited by the higher levels of noise. For Figure 7 an intrinsic spike frequency adaptation term was included for the excitatory population (Jercog et al., 2017):

$$(3) \quad \tau_a \frac{da}{dt} = -a(t) + \beta E(t)$$

where β represents the adaptation gain. For the data show in Figures 1 to 6 no adaptation was present ($\beta=0$). On Figure 7 we set $\beta=0.4$.

The function $f_Y(x)$ represents the intrinsic excitability of the neurons, and it is modeled as a threshold-linear (ReLU) function with threshold θ_Y and gain g_Y.

$$(4) \quad f_Y(x) = \begin{cases} 0 & if\ x < \theta_Y \\ g_Y\ (x - \theta_Y) & if\ \ x \geq \theta_Y \end{cases}, Y = \{E, I\}$$

This function models the F-I curve, or activation function, characteristic of nerve cells. The values for θ_Y and g_Y obey experimentally grounded relations (Jercog et al., 2017; Romero-Sosa et al., 2020). The threshold for the excitatory population is set to $\theta_E = 4.8$ while the inhibitory population has a greater threshold $\theta_I = 25$. The gain of the inhibitory population $g_I = 4$ is also greater than the excitatory one $g_E = 1$ as confirmed by experiments (Romero-Sosa et al., 2020)

Synaptic plasticity

The four weights ($W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, $W_{I \leftarrow I}$) follow homeostatic and homeostatic-like synaptic plasticity rules driven by the deviation of the excitatory and inhibitory population rates to their target setpoints (E_{set} and I_{set}). Three different learning rules are presented in this study.

Standard homeostatic family of rules:

$$(5) \quad \Delta W_{EE} = +\alpha_E E(E_{set} - E)$$

$$\Delta W_{EI} = -\alpha_E I(E_{set} - E)$$

$$\Delta W_{IE} = +\alpha_I E(I_{set} - I)$$

$$\Delta W_{II} = -\alpha_I I(I_{set} - I)$$

Where the learning rates are $\alpha_E = g_E \alpha$ and $\alpha_I = g_I \alpha$, with $\alpha = 0.00002$. The setpoints follow experimental values previously characterized in (Romero-Sosa et al., 2020) ($E_{set} = 5$ and $I_{set} = 14$). For Figure 5 two alternative pairs of setpoints were explored ($E_{set} = 5$ and $I_{set} = 24$) and ($E_{set} = 10$ and $I_{set} = 14$).

The exact formulation of the presented homeostatic rules can be recovered after a zeroth order approximation of a gradient descent derivation on the following loss function:

$$(6) \quad L = \frac{1}{2}(E - E_{set})^2 + \frac{1}{2}(I - I_{set})^2$$

The mathematical derivation can be found in the Supplementary Material (Section 3). The configuration of setpoints follows a classic homeostatic formulation (Turrigiano et al., 1998; Rossum et al., 2000; Liu and Buonomano, 2009a; Vogels et al., 2011), where every neural population adapts its input weights homeostatically in order to minimize its error term. It is proved in the Supplementary Material (Section 2.3) that these rules are inherently unstable for biologically meaningful parameter values in the context of inhibition-stabilized networks.

Cross-homeostatic family of rules:

$$(7) \quad \Delta W_{EE} = +\beta(I_{set} - I)$$

$$\Delta W_{EI} = -\beta(I_{set} - I)$$

$$\Delta W_{IE} = -\beta(E_{set} - E)$$

$$\Delta W_{II} = +\beta(E_{set} - E)$$

Where the learning rate is set to $\beta = 0.001$. These rules represent non-standard forms of homeostatic plasticity, where each neural population adjusts its incoming weights to minimize the target setpoint of the presynaptic population. However, considering our population model, this formulation could be interpreted in a local manner: as the neurons minimizing the total presynaptic input current to a target setpoint. We note that an equivalent rule for W_{IE} has been recently derived by (Mackwood et al., 2020) and successfully implemented in its input current form in a multi-unit rate model. Different to (Mackwood et al., 2020), the cross-homeostatic rules were inspired by a first order approximation of a gradient descent derivation of the loss function presented in (6) (see Supplementary Material, Section 3). We further prove that these rules are inherently

stable (Supplementary Material, Section 2.4) and lead to balanced dynamics (Figure 5) in the context of inhibition-stabilized networks.

Balanced-homeostatic family of rules:

$$(8) \quad \Delta W_{EE} = +\gamma_E E(E_{set} - E)$$

$$\Delta W_{EI} = \frac{1}{\tau_P}(f(W_{EE}) - W_{EI})$$

$$\Delta W_{IE} = -\gamma_I I(I_{set} - I)$$

$$\Delta W_{II} = \frac{1}{\tau_P}(g(W_{IE}) - W_{II})$$

Where:

$$(9) \quad f(W_{EE}) = \frac{(E_{Set}W_{EE} - \theta_E)g_E - E_{Set}}{I_{Set}g_E}$$

$$(10) \quad g(W_{IE}) = \frac{(E_{Set}W_{IE} - \theta_I)g_I - I_{Set}}{I_{Set}g_I}$$

Where the learning rates are $\gamma_E = g_E\alpha_1$ and $\gamma_I = g_I\alpha_2$, with $\alpha_1 = 0.002$ and $\alpha_2 = 0.00002$. W_{EE} and W_{IE} obey classic homeostatic forms of plasticity (neurons update their input weight based on their own setpoint), with W_{IE} being anti-homeostatic. The presynaptic term for W_{IE} , which should be *E* on the classic homeostatic formulation, was also changed to *I* to make the rule even more local. W_{EI} and W_{II} slowly evolve towards a hardwired function of the homeostatic weights with a plasticity time constant $\tau_P = 100$ (trial units). The hardwired relationship (9) and (10), corresponds to the steady-state solution of the neural system when the inhibitory and excitatory rates are its target setpoints. The solution can be obtained by setting the left side of equations (1) and (2) to zero, and substitute the steady state E and I values by E_{set} and I_{set}. It is thus a necessary condition for the system to be at its setpoints that a linear relationship (or balance) between the excitatory and inhibitory weights is obeyed. The slope of the line corresponds to the relation E_{Set}/I_{Set}. The linear constraint is exploited to reduce the number of free weights following homeostatic learning rules.

We prove that these rules are stable for a biologically meaningful set of parameter values, provided the neural subsystem satisfies its stability conditions (Supplementary Material, Section 2.5).

All rules:

For all simulations, the weights were updated after the completion of every trial. The trials last 2 seconds, except for for Figure 7 in which a 20 second trial period was implemented.

For our numerical simulations, E and I on every rule are implemented as average firing rates. The average of E and I is computed after every trial and then is low pass filtered by a process with a time constant τ_{trial}. We set $\tau_{trial} = 2$ for the homeostatic and cross-homeostatic rules and $\tau_{trial} = 10$ for the balanced rule.

In Figure 7, the average rate is linearly rectified by a hypothetical activity sensor (e.g., a calcium sensor) with threshold 1 Hz and gain 1. In this manner, only the rate during active periods was integrated.

A minimum weight of 0.1 was set for all weights, except with the balanced rule, whose W_{EE} and W_{IE} minimum obeys the value set by the linear relationship of the rule when W_{EI} and W_{II} are at its minimum 0.1 weight (eq. 9 and 10).

A saturation to the excitatory and inhibitory firing rate was added for practical reasons, and set at 100 and 250 Hz respectively. It aided the visualization of the results of Figure 5, where many experiments would have diverged towards infinity due to the instability of the homeostatic rule. Note the saturation is not necessary for the cross-homeostatic rule to be inherently stable as proved in the Supplementary Material.

Data analysis and Software availability

Data are represented by the mean ± SEM.

In Figures 2, 3 and 6D, Figure 5 and Supplementary Figure 1 we randomly initialize the weights in between the following ranges:

W_{EE}[4,7], W_{EI}[0.5,2], W_{IE}[7,13], W_{II}[0.5,2]

Experimental and computational analysis were performed in custom-written MATLAB R2020a software. SageMath was used for the analytical proofs (see Supplementary Material). The MATLAB source code that reproduces Figures 2, 4, 6 and 7, and the Jupyter notebooks with SageMath code will be available at:

https://github.com/BuonoLab/UpDev2020.git

ACKNOWLEDGEMENTS
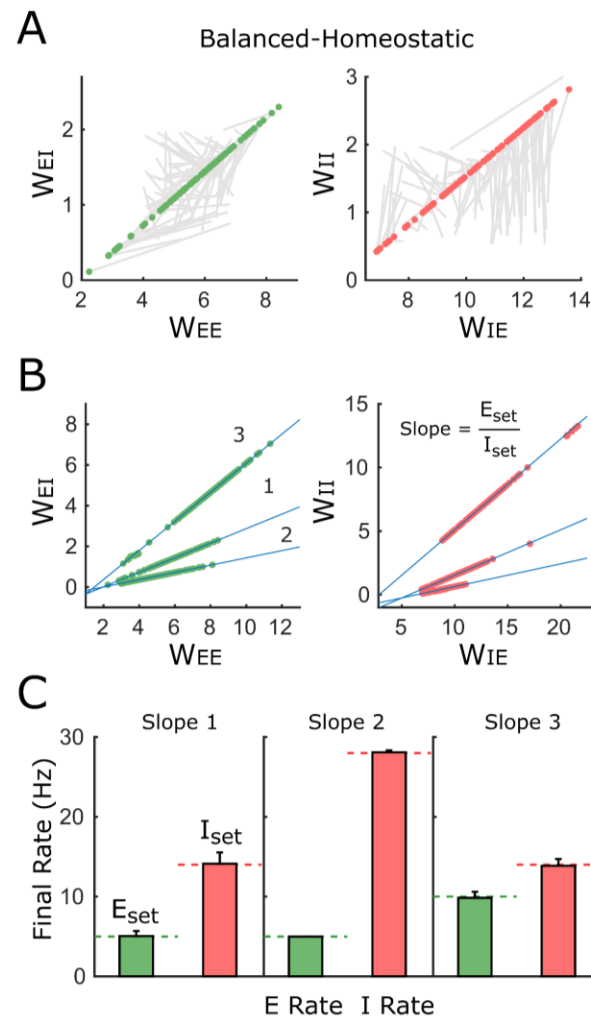
Supplementary Figure 1. Balanced-homeostatic rules lead to stable balanced dynamics at the setpoints.

(A) Weight trajectories for 100 different simulations with random weight initializations for the balanced-homeostatic rules. Lines show trajectories from initial to final values (circles).

(B) Final weight values for simulations with same starting conditions as in (A) but for additional homeostatic pairs of setpoints. Line 1: $E_{set}=5$, $I_{set}=14$; Line 2: $E_{set}=5$, $I_{set}=28$; Line 3: $E_{set}=10$, $I_{set}=14$. Data shown in (A) corresponds to Line 1. Blue lines correspond to the theoretical linear relationship between the excitatory and inhibitory weights at a fixed point obeying E_{set} and I_{set}. The slope of the line is defined by the ratio of the setpoints.

(C) Final rate for the excitatory and inhibitory population after balanced-homeostatic learning. The different bar plots correspond to the data shown in (B), where 3 different pairs of setpoints are tested.

REFERENCES

Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. Nat Neurosci 3:1178-1183.

Ahmadian Y, Miller KD (2019) What is the dynamical regime of cerebral cortex? arXiv:190810101 [q-bio].

Antoine MW, Langberg T, Schnepel P, Feldman DE (2019) Increased Excitation-Inhibition Ratio Stabilizes Synapse and Circuit Excitability in Four Autism Mouse Models. Neuron 101:648-661.e644.

Bartram J, Kahn MC, Tuohy S, Paulsen O, Wilson T, Mann EO (2017) Cortical Up states induce the selective weakening of subthreshold synaptic inputs. Nature Communications 8:665.

Bauer R, Zubler F, Pfister S, Hauri A, Pfeiffer M, Muir DR, Douglas RJ (2014) Developmental self-construction and-configuration of functional neocortical neuronal networks. PLoS Comput Biol 10:e1003994.

Beltramo R, D'Urso G, Dal Maschio M, Farisello P, Bovetti S, Clovis Y, Lassi G, Tucci V, De Pietri Tonelli D, Fellin T (2013) Layer-specific excitatory circuits differentially control recurrent network dynamics in the neocortex. Nat Neurosci 16:227-234.

Binas J, Rutishauser U, Indiveri G, Pfeiffer M (2014) Learning and stabilization of winner-take-all dynamics through interacting excitatory and inhibitory plasticity. Frontiers in computational neuroscience 8:68.

Brunel N (2000) Dynamics of networks of randomly connected excitatory and inhibitory spiking neurons. Journal of Physiology-Paris 94:445-463.

Buonomano DV, Merzenich MM (1998) Cortical plasticity: from synapses to maps. Annual Rev Neuroscience 21:149-186.

Burrone J, O'Byrne M, Murthy VN (2002) Multiple forms of synaptic plasticity triggered by selective suppression of activity in individual neurons. Nature 420:414-418.

Chiu CQ, Barberis A, Higley MJ (2019) Preserving the balance: diverse forms of long-term GABAergic synaptic plasticity. Nature Reviews Neuroscience 20:272-281.

Chiu CQ, Martenson JS, Yamazaki M, Natsume R, Sakimura K, Tomita S, Tavalin SJ, Higley MJ (2018) Input-Specific NMDAR-Dependent Potentiation of Dendritic GABAergic Inhibition. Neuron 97:368-377.e363.

Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH (1995) Recurrent excitation in neocortical circuits. Science 269:981-985.

Fanselow EE, Connors BW (2010) The Roles of Somatostatin-Expressing (GIN) and Fast-Spiking Inhibitory Interneurons in up-down States of Mouse Neocortex. Journal of Neurophysiology 104:596-606.

Feldman DE (2009) Synaptic Mechanisms for Plasticity in Neocortex. Annual Review of Neuroscience 32:33-55.

Feldman Daniel E (2012) The Spike-Timing Dependence of Plasticity. Neuron 75:556-571.

Field RE, D'amour JA, Tremblay R, Miehl C, Rudy B, Gjorgjieva J, Froemke RC (2020) Heterosynaptic Plasticity Determines the Set Point for Cortical Excitatory-Inhibitory Balance. Neuron 106:842-854.e844.

Froemke RC (2015) Plasticity of Cortical Excitatory-Inhibitory Balance. Annual Review of Neuroscience 38:195-219.

Froemke RC, Merzenich MM, Schreiner CE (2007) A synaptic memory trace for cortical receptive field plasticity. Nature 450:425-429.

Fuster JM, Jervey JP (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. Science 212:952-955.

Goel A, Lee H-K (2007) Persistence of Experience-Induced Homeostatic Synaptic Plasticity through Adulthood in Superficial Layers of Mouse Visual Cortex. J Neurosci 27:6692-6700.

Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14:477-485.

Hartman KN, Pal SK, Burrone J, Murthy VN (2006) Activity-dependent regulation of inhibitory synaptic transmission in hippocampal neurons. Nat Neurosci 9:642-649.

Hebb DO (1949) Organization of behavior. New York: Wiley.

Hennequin G, Agnes EJ, Vogels TP (2017) Inhibitory Plasticity: Balance, Control, and Codependence. Annual Review of Neuroscience 40:557-579.

Hromádka T, Zador AM, DeWeese MR (2013) Up states are rare in awake auditory cortex. Journal of Neurophysiology 109:1989-1995.

Iascone DM, Li Y, Sümbül U, Doron M, Chen H, Andreu V, Goudy F, Blockus H, Abbott LF, Segev I (2020) Whole-neuron synaptic mapping reveals spatially precise excitatory/inhibitory balance limiting dendritic and somatic spiking. Neuron.

Jercog D, Roxin A, Barthó P, Luczak A, Compte A, de la Rocha J (2017) UP-DOWN cortical dynamics reflect state transitions in a bistable network. eLife 6:e22425.

Johnson HA, Buonomano DV (2007) Development and Plasticity of Spontaneous Activity and Up States in Cortical Organotypic Slices. J Neurosci 27:5915-5925.

Joseph A, Turrigiano GG (2017) All for One But Not One for All: Excitatory Synaptic Scaling and Intrinsic Excitability Are Coregulated by CaMKIV, Whereas Inhibitory Synaptic Scaling Is Under Independent Control. The Journal of Neuroscience 37:6778-6785.

Kilman V, van Rossum MC, Turrigiano GG (2002) Activity deprivation reduces miniature IPSC amplitude by decreasing the number of postsynaptic GABA(A) receptors clustered at neocortical synapses. J Neurosci 22:1328-1337.

Kroener S, Chandler LJ, Phillips PEM, Seamans JK (2009) Dopamine Modulates Persistent Synaptic Activity and Enhances the Signal-to-Noise Ratio in the Prefrontal Cortex. PLoS ONE 4:e6507.

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

Liu JK, Buonomano DV (2009a) Embedding Multiple Trajectories in Simulated Recurrent Neural Networks in a Self-Organizing Manner. J Neurosci 29:13172-13181.

Liu JK, Buonomano DV (2009b) Embedding Multiple Trajectories in Simulated Recurrent Neural Networks in a Self-Organizing Manner. J Neurosci 29:13172-13181.

Liu Z, Golowasch J, Marder E, Abbott LF (1998) A model neuron with activity-dependent conductances regulated by multiple calcium sensors. J Neurosci 18:2309-2320.

Mackwood O, Naumann LB, Sprekeler H (2020) Learning excitatory-inhibitory neuronal assemblies in recurrent networks. In: Neuroscience.

Martin SJ, Grimwood PD, Morris RG (2000) Synaptic plasticity and memory: an evaluation of the hypothesis. Annu Rev Neurosci 23:649-711.

McCormick DA (1989) GABA as an inhibitory neurotransmitter in human cerebral cortex. J Neurophysiol 62:1018-1027.

Miller KD, Keller JB, Stryker MP (1989) Ocular dominance column development: analysis and simulation. Science 245:605-615.

Motanis H, Buonomano DV (2015) Delayed in vitro Development of Up States but Normal Network Plasticity in Fragile X Circuits. Eur J Neurosci 42:2312-2321.

Motanis H, Buonomano D (2020) Decreased reproducibility and abnormal experience-dependent plasticity of network dynamics in Fragile X circuits. Scientific Reports 10:14535.

Neske GT, Patrick SL, Connors BW (2015) Contributions of Diverse Excitatory and Inhibitory Neurons to Recurrent Network Activity in Cerebral Cortex. The Journal of Neuroscience 35:1089-1105.

O'Leary T, Williams AH, Caplan JS, Marder E (2013) Correlations in ion channel expression emerge from homeostatic tuning rules. Proceedings of the National Academy of Sciences 110:E2645.

O'Leary T, Williams Alex H, Franci A, Marder E (2014) Cell Types, Network Homeostasis, and Pathological Compensation from a Biologically Plausible Ion Channel Expression Model. Neuron 82:809-821.

Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D (2009) Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression. Neuron 62:578-592.

Peng Y-R, Zeng S-Y, Song H-L, Li M-Y, Yamada MK, Yu X (2010) Postsynaptic Spiking Homeostatically Induces Cell-Autonomous Regulation of Inhibitory Inputs via Retrograde Signaling. The Journal of Neuroscience 30:16220-16231.

Plenz D, Kitai ST (1998) Up and down states in striatal medium spiny neurons simultaneously recorded with spontaneous activity in fast-spiking interneurons studied in cortex-striatum-substantia nigra organotypic cultures. J Neurosci 18:266-283.

Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD (2010) The Asynchronous State in Cortical Circuits. Science 327:587-590.

Romero-Sosa JL, Motanis H, Buonomano DV (2020) Differential excitability of PV and SST neurons results in distinct functional roles in inhibition stabilization of Up-states. bioRxiv:2020.2011.2026.395343.

Rossum MCWv, Bi GQ, Turrigiano GG (2000) Stable Hebbian Learning from Spike Timing-Dependent Plasticity. J Neurosci 20:8812-8821.

Rubin Daniel B, Van Hooser Stephen D, Miller Kenneth D (2015) The Stabilized Supralinear Network: A Unifying Circuit Motif Underlying Multi-Input Integration in Sensory Cortex. Neuron 85:402-417.

Sadovsky AJ, MacLean JN (2014) Mouse Visual Neocortex Supports Multiple Stereotyped Patterns of Microcircuit Activity. The Journal of Neuroscience 34:7769-7777.

Sanchez-Vives MV, McCormick DA (2000) Cellular and network mechanisms of rhythmic recurrent activity in neocortex. Nat Neurosci 3:1027-1034.

Sanzeni A, Akitake B, Goldbach HC, Leedy CE, Brunel N, Histed MH (2020a) Inhibition stabilization is a widespread property of cortical networks. eLife 9:e54875.

Sanzeni AH, Akitake B, Goldbach HC, Leedy CE, Brunel N, Histed MH (2020b) Inhibition stabilization is a widespread property of cortical networks. bioRxiv:656710.

Seamans JK, Nogueira L, Lavin A (2003) Synaptic Basis of Persistent Activity in Prefrontal Cortex In Vivo and in Organotypic Cultures. Cerebral Cortex 13:1242-1250.

Shepherd GM (1998) The synaptic organization of the brain. . New York: Oxford University.

Shu Y, Hasenstaub A, McCormick DA (2003) Turning on and off recurrent balanced cortical activity. Nature 423:288-293.

Sippy T, Yuste R (2013) Decorrelating Action of Inhibition in Neocortical Networks. The Journal of Neuroscience 33:9813-9830.

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

Steriade M, Contreras D (1998) Spike-Wave Complexes and Fast Components of Cortically Generated Seizures. I. Role of Neocortex and Thalamus. Journal of Neurophysiology 80:1439-1455.

Steriade M, McCormick D, Sejnowski T (1993) Thalamocortical oscillations in the sleeping and aroused brain. Science 262:679-685.

Stoppini L, Buchs P-A, Muller D (1991) A simple method for organotypic cultures of nervous tissue. Journal of neuroscience methods 37:173-182.

Sullivan TJ, de Sa VR (2006) Homeostatic synaptic scaling in self-organizing maps. Neural Networks 19:734-743.

Timofeev I, Grenier F, Bazhenov M, Sejnowski TJ, Steriade M (2000) Origin of slow cortical oscillations in deafferented cortical slabs. Cereb Cortex 10:1185-1199.

Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL (1997a) Paradoxical Effects of External Modulation of Inhibitory Interneurons. The Journal of Neuroscience 17:4382-4388.

Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL (1997b) Paradoxical Effects of External Modulation of Inhibitory Interneurons. J Neurosci 17:4382-4388.

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

Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 391:892-896.

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

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

van Vreeswijk C, Sompolinsky H (1998) Chaotic balanced state in a model of cortical circuits. Neural Comput 10:1321-1371.

Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W (2011) Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. Science 334:1569-1573.

Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:455–463.

Wilson HR, Cowan JD (1972) Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. Biophysical Journal 12:1-24.

Xu H, Jeong H-Y, Tremblay R, Rudy B (2013) Neocortical Somatostatin-Expressing GABAergic Interneurons Disinhibit the Thalamorecipient Layer 4. Neuron 77:155-167.

Xue M, Atallah BV, Scanziani M (2014) Equalizing excitation-inhibition ratios across visual cortical neurons. Nature 511:596-600.