# Orchestrated Excitatory and Inhibitory Learning Rules Lead to the Unsupervised Emergence of Self-sustained and Inhibition-stabilized Dynamics

**AUTHORS:**
Saray Soldado-Magraner[1], Rodrigo Laje[2,*], Dean V. Buonomano[1,3,*]

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

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

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

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

## ACKNOWLEDGEMENTS

We thank Juan Romero-Sosa and Helen Motanis for the sample traces shown in Figure 1. We thank Ben Liu, Shanglin Zhou, and Helen Motanis, for technical assistance and helpful discussions. We thank Joana Soldado-Magraner and Mike Seay for comments on the manuscript. This research was supported by NIH grant NS116589, and SSM was supported by the Swiss National Science Foundation (P2ZHP3-187943).

## ABSTRACT

Self-sustaining neural activity maintained through local recurrent connections is of fundamental importance to cortical function. We show that Up-states—an example of self-sustained, inhibition-stabilized network dynamics—emerge in cortical circuits across three weeks of *ex vivo* development, establishing the presence of unsupervised learning rules capable of generating self-sustained dynamics. Previous computational models have established that four sets of weights ($W_{E\leftarrow E}$, $W_{E\leftarrow I}$, $W_{I\leftarrow E}$, $W_{I\leftarrow I}$) must interact in an orchestrated manner to produce Up-states, but have not addressed how a family of learning rules can operate in parallel at all four weight classes to generate self-sustained inhibition-stabilized dynamics. Using numerical and analytical methods we show that, in part due to the paradoxical effect, standard homeostatic rules are only stable in a narrow parameter regime. In contrast, we show that a family of biologically plausible learning rules based on “cross-homeostatic” plasticity robustly lead to the emergence of self-sustained, inhibition-stabilized dynamics.

# INTRODUCTION

Self-sustained patterns of neural activity maintained by local recurrent excitation underlie many cortical computations and dynamic regimes, including the persistent activity associated with working memory (Fuster and Jervey, 1981; Goldman-Rakic, 1995; Wang, 2001), asynchronous states (van Vreeswijk and Sompolinsky, 1998; Renart et al., 2010), and Up-states (Steriade et al., 1993; Timofeev et al., 2000). Recurrent excitation, however, also has the potential to drive pathological and epileptiform regimes (McCormick, 1989; Douglas et al., 1995; Steriade and Contreras, 1998). Converging theoretical and experimental evidence indicate that cortical circuits that generate self-sustained dynamics operate in an inhibition-stabilized regime, in which positive feedback is held in check by recurrent inhibition (Tsodyks et al., 1997; Brunel, 2000; Ozeki et al., 2009; Rubin et al., 2015; Rutishauser et al., 2015; Jercog et al., 2017; Sanzeni et al., 2020).

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: $W_{E\leftarrow E}$, $W_{E\leftarrow I}$, $W_{I\leftarrow E}$, $W_{I\leftarrow I}$. Analytical and numerical analyses have shown that these weights must obey certain theoretically well-defined relationships in order to generate self-sustained, inhibition-stabilized dynamics (Tsodyks et al., 1997; Brunel, 2000; Ozeki et al., 2009; Rubin et al., 2015; Jercog et al., 2017). Yet, it is not known how the appropriate relationships between these four classes of weights could emerge in a self-organizing manner (Sadeh and Clopath, 2021). One possibility is that standard homeostatic forms of plasticity underlie the emergence of inhibition-stabilized networks. Homeostatic learning rules generally assume that excitatory weights are regulated in a manner proportional to the difference between some ontogenetically determined set-point and average neural activity (for both excitatory and inhibitory neurons)—and conversely that inhibitory weights onto excitatory neurons 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). However, it remains an open question whether homeostatic rules can lead to the self-organized emergence of self-sustained, inhibition-stabilized networks.

At both the experimental and computational level one of the simplest and best-studied examples of self-sustained activity are Up-states (Steriade et al., 1993; Timofeev et al., 2000). Up-states are characterized by network-wide regimes in which excitatory and inhibitory neurons transiently shift from a quiescent Down-state to a depolarized state with low to moderate firing rates (Sanchez-Vives and McCormick, 2000; Neske et al., 2015; Bartram et al., 2017). 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), 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), and 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).

Consistent with previous results we first demonstrate that Up-states emerge in both excitatory and inhibitory neurons over the course of the first few weeks of *ex vivo* development, suggesting that local cortical circuits are programmed to homeostatically generate Up-states. We next used computational models and analytical methods to explore families of homeostatic learning rules that operate in parallel at all four synapse classes and lead to self-sustained, inhibition-stabilized dynamics. We show that when driving the network towards a self-sustained, inhibition-stabilized regime, standard forms of homeostatic plasticity are only stable in a narrow region of parameter space. This is in part a consequence of the paradoxical effect—in which an *increase* in excitatory drive to inhibitory neurons produces a net *decrease* in the firing rate of those same inhibitory neurons (Tsodyks et al., 1997; Ozeki et al., 2009; Rubin et al., 2015). We next developed a family of homeostatic learning rules that include "cross-homeostatic" influences, and lead to the unsupervised emergence of Up-states in the inhibition-stabilized regime in a robust manner. These rules are consistent with experimental data and generate explicit predictions regarding the effects of manipulations of excitatory and inhibitory neurons on synaptic plasticity.

## RESULTS

### Up-states emerge autonomously during ex vivo development

Up-states represent a transition from a quiescent state to a self-sustained network-wide dynamic 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., 2021). 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. 1B**), 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. 1B,** inset). 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, whereas if inhibition is too strong only the trivial quiescent fixed point will be stable (Tsodyks et al., 1997; 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. In contrast, recordings in cortical organotypic cultures show that Up-states autonomously 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. 1C**). Here we further characterized the emergence of Up-states and asked whether the development of activity is in sync for both excitatory and inhibitory neurons. Using two-photon calcium imaging, we recorded the spontaneous activity in excitatory neurons and PV$^+$-inhibitory neurons by expressing GCamp6f under the CaMKII and Flex promoters in organotypic cultures of PV-Cre mice. Calcium imaging at DIV 12-13 revealed infrequent and short bouts of synchronous activity. By DIV 14-16 Up-states were observed, and over the entire four-weeks of *ex vivo* development there was an increase and stabilization of Up-state frequency and duration in both excitatory and inhibitory neurons (**Fig. 1D-F**)—suggesting the Up-states emerge in a co-dependent manner in both populations.

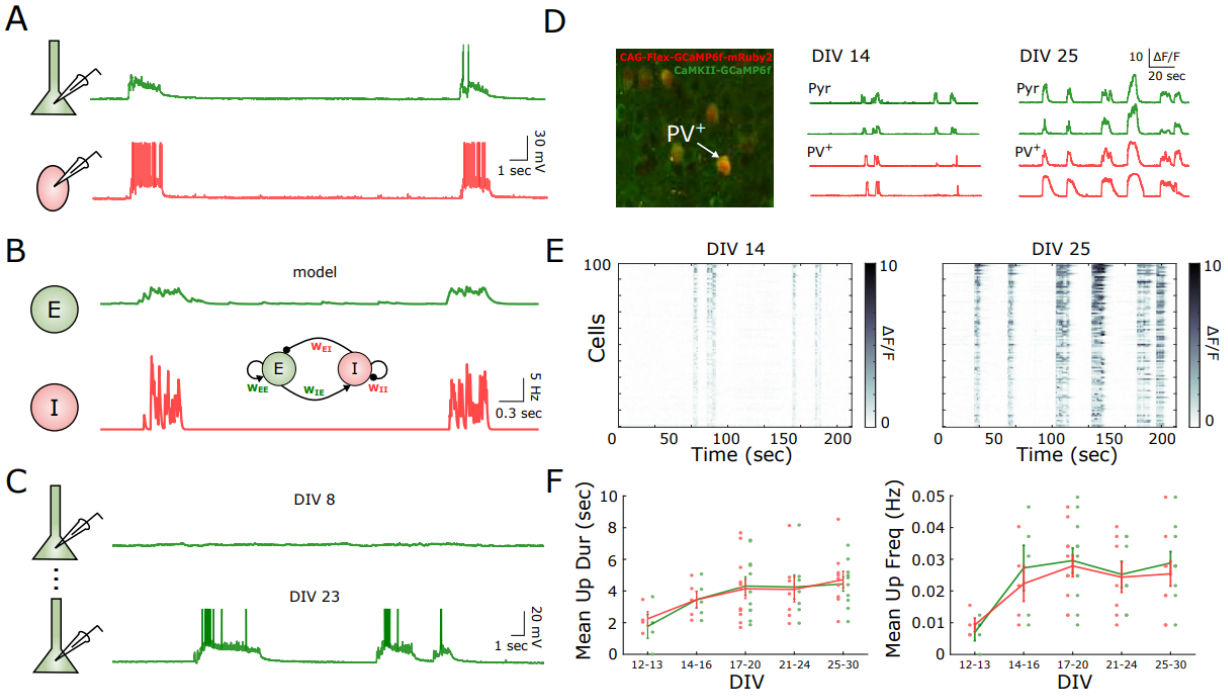


**Figure 1. Up-states emerge autonomously over the course of *ex vivo* development.**

**(A)** Example of Up-states in simultaneously whole-cell recordings of a pyramidal (green) and parvalbumin (PV) positive inhibitory neuron (red).

**(B)** Two-population firing rate model of Up-states. The schematic of the model is shown in the inset. 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.

**(C)** 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.

**(D)** Two-photon calcium imaging recording of excitatory and PV$^+$ neurons at different stages of development. Organotypic slices of Cre-PV mice were transfected with pAAV-CAG-Flex-GCamp6f-mRuby2 and pAAV-CaMKII-GCamp6f. Image shows an example slice with a PV$^+$ neuron expressing both, GCamp6f and the

mRuby2 red marker. Traces show the spontaneous calcium activity of 2 example $PV^+$ and excitatory cells at 14 and 25 DIV. Up-states can be observed more prominently at later stages (see Methods for the definition and quantification of Up-states).

**(E)** Spontaneous calcium activity of 100 example cells at 14 and 25 DIV. Synchronous activity events correspond to Up-states.

**(F)** Evolution of the mean Up-state duration and frequency over the course of *ex vivo* development for excitatory (green) and $PV^+$ neurons (red). A significant increase in mean Up duration (Two-way ANOVA: $F_{4,64} = 3.54$, $p = 0.011$) and frequency (Two-way ANOVA: $F_{4,64} = 4.75$, $p = 0.002$) was observed over developmental time, with no statistical effect of neuron type ($F_{4,64} = 0.13$, $p = 0.97$) or interaction effect ($F_{4,64} = 0.09$, $p = 0.98$).

The observation 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. 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 can only account for stable self-sustained activity in a narrow parameter regime

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 during Up-states, and they homeostatically adjust their excitatory and inhibitory weights to reach these target setpoints. Homeostatic learning 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 neurons average activity levels (Turrigiano et al., 1998; van Rossum et al., 2000; Kilman et al., 2002; Turrigiano and Nelson, 2004; Liu and Buonomano, 2009; Peng et al., 2010; Vogels et al., 2011), e.g., $\Delta W_{E \leftarrow E} \propto E_{set} - E_{avg}$ where any departure of the excitatory activity $E_{avg}$ from the setpoint $E_{set}$ would lead to a compensatory correction in the value of the weight $W_{E \leftarrow E}$.

We first asked is stable self-sustained dynamics can emerge in a standard two-population model (Jercog et al., 2017; see Methods) through homeostatic mechanisms. We initialized the four weights ($W_{E \leftarrow E}$, $W_{E \leftarrow I}$, $W_{I \leftarrow E}$, $W_{I \leftarrow I}$) of the model at random values and applied a standard family of homeostatic learning rules to all four weights classes (**Fig. 2A**). It is well established that $PV^+$-inhibitory neurons have higher firing rates than pyramidal neurons during Up-states (Neske et al., 2015; Romero-Sosa et al., 2021), thus based on experimental data we set the setpoints for the *E* and *I* populations during Up-states to 5 and 14 Hz, respectively (Romero-Sosa et al., 2021). We first asked whether the set of four standard homeostatic learning rules can lead to stable self-sustained dynamic regime (representing a permanent Up-state) in response to a brief external input (low levels of noise were used to avoid spontaneous Up↔Down transitions).

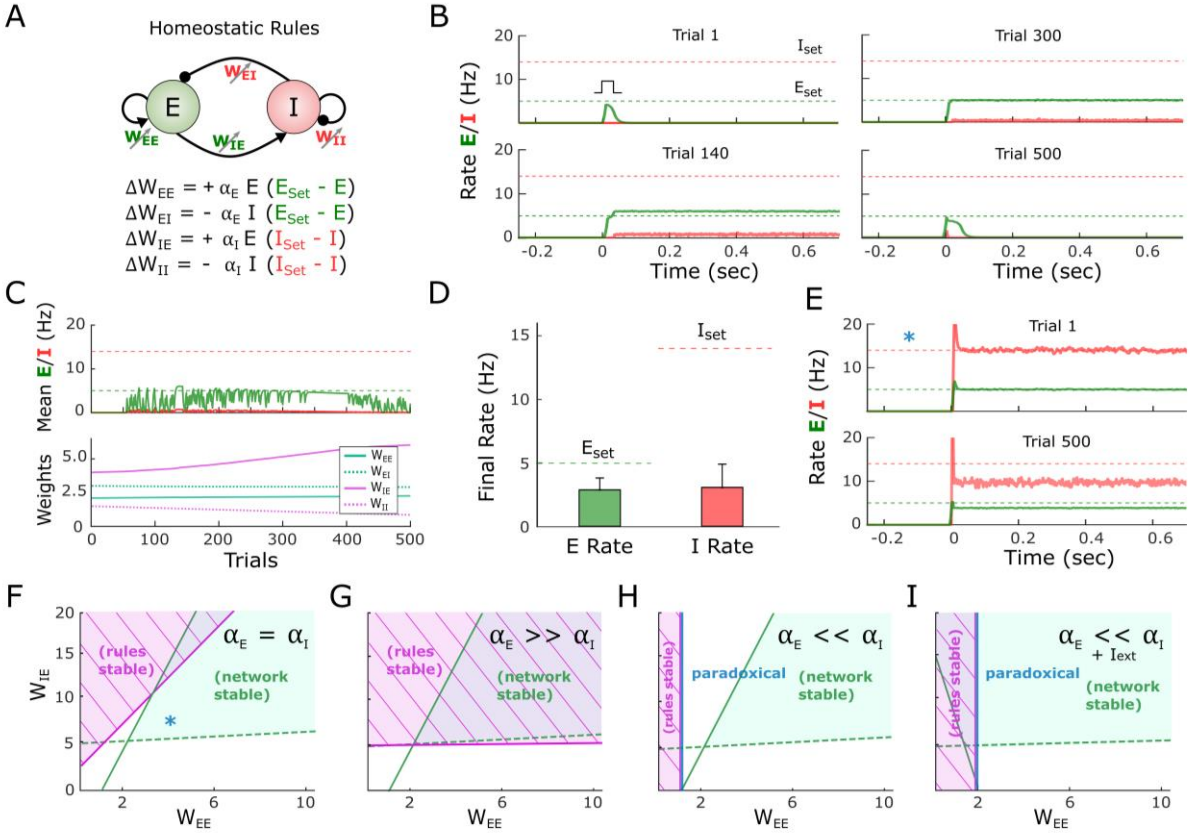


**Figure 2. Standard homeostatic rules are only stable in a narrow parameter regime.**

**(A)** Schematic (top) of the population rate model in which the four weights are governed by a family of homeostatic learning rules (bottom).

**(B)** Example simulation of the network over the course of simulated development. Each plot shows the firing rate of the excitatory and inhibitory population over the course of a trial in response to a brief external input. $E_{set}$=5 and $I_{set}$=14 represent the target homeostatic setpoints. Weights were initialized to $W_{EE}$=2.1, $W_{EI}$=3, $W_{IE}$=4, and $W_{II}$=2. Note that while an evoked Up-state emerges by Trial 200 the firing rates do not converge to their setpoints, and by Trial 500 the Up-state is no 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. Data represents mean ± SEM.

**(E)** Simulation of a network starting with weights that generate Up-states that match the $E_{set}$=5 and $I_{set}$=14 Hz setpoints (Trial 1, top). After 500 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)** Analytical stability regions of the neural and learning rule subsystems as a function of the free weights $W_{EE}$ and $W_{IE}$. (note that once a $W_{EE}$ and $W_{IE}$ are set to generate an Up-state with specific $E_{set}$ and $I_{set}$ values, $W_{EI}$ and $W_{II}$ are fully determined by $W_{EE}$ and $W_{IE}$, respectively). Here the stability plot is obtained by considering equal learning rates for all four learning rules (as used for panels B-E). Blue asterisk corresponds to the initial conditions shown in Panel D (top).

**(G)** Similar to F but with $\alpha_E >> \alpha_I$.

**(H)** Similar to F but with but with $\alpha_E << \alpha_I$. To the right of the blue line, the network is in a paradoxical regime (defined by the condition $W_{EE} g_E - 1 > 0$)

**(I)** Condition of stability of the neural system and learning rule system when the learning rate on the inhibitory neuron dominates and an external excitatory current is applied to the excitatory neuron. The

current produces an enlargement of the stability region of the neural subsystem. Right of blue line shows the area where the network is in a paradoxical regime.

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), in the example shown in **Fig. 2B-C** the network failed to converge to a stable Up-state (**Fig. 2B-C**). Initially (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 the brief external input triggers an Up-state, but the activities $E$ and $I$ do not match the corresponding setpoints—the network is in a nonbiologically observed regime in which $E > I$—so the weights keep evolving. 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. At that point both setpoint error terms have increased, leading to continued weight changes (**Fig. 2C)**. Results across 100 simulations with different weight initializations (see Methods) further indicate that the standard homeostatic rules are ineffective at driving $E$ and $I$ towards their respective setpoints and generating stable self-sustained dynamics (**Fig. 2D**).

To gain insights into why a family of homeostatic learning rules that might intuitively converge fails to do so, we can consider the case in which a network is initialized to a set of weights that already match $E_{set}$ and $I_{set}$ (**Fig. 2E**). Although the neural subsystem alone is stable at this condition (Trial 1), small fluctuations in $E$ and $I$ cause the homeostatic rules to drive the weight values and the average activity of the network away from the setpoints (Trial 500). It is possible to understand this instability by performing an analytical stability analysis. Specifically, a two-population 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 **Section 2.1** in the 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. For the results presented in **Fig. 2B-E** we assumed the learning rates driving plasticity onto the excitatory ($\alpha_E$) and inhibitory neurons ($\alpha_I$) to be equal. Under these conditions, the standard homeostatic rules are mostly unstable for biologically meaningful parameter values in which the neural system is stable. An example of this result is shown in **Fig. 2F** for a particular set of parameter values. Critically, **Fig. 2F** shows that the stability region of the neural subsystem, i.e., an inhibition-stabilized network (Ozeki et al., 2009; Jercog et al., 2017), is almost entirely within the region where the homeostatic learning rule system is unstable. Only when plasticity onto the excitatory neuron is significantly faster ($\alpha_E >> \alpha_I$) is there a substantial region of overlap between the stability of the neural and learning rules subsystems (**Fig. 2G,** see Supplementary Material, **Section 1.1**).

Because inhibitory neurons seem to undergo homeostatic plasticity as quickly or more quickly than excitatory neurons (Keck et al., 2011; Kuhlman et al., 2013; Gainey et al.,

2018; Ma et al., 2019) we conclude that standard homeostatic rules by themselves do not account for the emergence of stable self-sustained and inhibition-stabilized dynamics. 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 only stable in a narrow region of parameter space (see Supplementary Material, **Section 1.5**). We next show that the inherent instability of standard homeostatic learning rules is related to the paradoxical effect.

**The paradoxical effect hampers the ability of homeostatic rules to lead to self-sustained activity**

The inability of the homeostatic learning rules to generate stable Up-states is in part a consequence of the paradoxical effect, a counterintuitive, yet well described, property of two-population models of Up-states and inhibition-stabilized networks (Tsodyks et al., 1997; Ozeki et al., 2009). Specifically, if during an Up-state one increases the excitatory drive to the inhibitory population, the net result is a decrease in the firing rate of the inhibitory units. This paradoxical effect can be understood in terms of the $I \rightarrow E \rightarrow I$ loop: the increased inhibitory drive leads to a lower steady-state rate for *E,* but this new steady-state value requires a decrease in the *I* firing rate to maintain an appropriate E/I balance (in effect, 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.

The relationship of the paradoxical effect and the homeostatic rules performance is presented in **Fig. 2H.** The region of stability for the homeostatic learning rules is shown in a parameter regime where inhibitory plasticity is much faster ($\alpha_E << \alpha_I$). Contrary to when excitatory plasticity dominates, the region of stability is small, and there is no overlap with the region of stability of the neural subsystem. Crucially, the boundary of the stability region of the learning rule coincides with the condition for the paradoxical effect to be present (right of the blue line in **Fig.2H**, see Supplementary Material, **Sections 2.2 and 2.3**). Under these conditions, the rules can only be stable when the network is not in an inhibition-stabilized regime. If a network regime with non-zero *E* would be forced to exist in that region (for example, via tonic external current, **Fig. 2I**), it would only be stable in the non-paradoxical region with the learning rules in place (see **Section 2.5** in the Supplementary Material).

To understand the impact of the paradoxical effect on homeostatic learning rules 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, homeostatic plasticity in the inhibitory neuron would intuitively result in an increase of $W_{I \leftarrow E}$. However, because of the paradoxical effect an increase in $W_{I \leftarrow E}$ actually makes *I* decrease (**Fig. 3B**)—thus increasing the error term $I_{set} - I$. To increase the steady-state inhibitory rate we can "anti-homeostatically" decrease the excitatory weight onto the inhibitory neurons (**Fig. 3C).** This simple example shows the complexity of designing a coherent set of rules in such a coupled system (see an analysis of the paradoxical effect in **Section 2.2** of the Supplementary Material). This analysis also explains why homeostatic learning rules can

lead to self-sustained activity at the appropriated setpoints when $\alpha_E >> \alpha_I$. Essentially by allowing plasticity onto the *E* population to be faster one overcomes the counterproductive homeostatic plasticity associated with the paradoxical effect.

The interaction between the paradoxical effect and homeostatic plasticity in inhibitory neurons leads to the question of whether anti-homeostatic plasticity rules may be more effective that standard homeostatic rules—e.g., $\Delta W_{I \leftarrow E} \propto -(I_{set} - I_{avg})$. Thus, we also examined a number of hybrid families of learning rules with different combinations of homeostatic and anti-homeostatic rules. Indeed, some hybrid families exhibited large degrees of overlap between the stable regions of the network and learning rules subsystems. However, numerical simulations revealed that these rules were mostly ineffective in driving networks to self-sustained activity at the target setpoints (Supplemental Material, **Section 1.2**, and Supplementary **Fig.S1**). These two results are not inconsistent because the stability analysis speaks to cases when the network is initialized to weights that satisfy $E_{set}$ and $I_{set}$, not whether the rules will drive network activity into these stable areas from any initial state including a pre-developmental state. Thus, we interpret these results as meaning that while anti-homeostatic plasticity can contribute to stability of this dual dynamical system, anti-homeostatic plasticity is ineffective at driving the dynamics towards setpoints (in other words, that anti-homeostatic plasticity might allow for stable Up-state but does not necessarily generate sizable basins of attraction around Up-states).

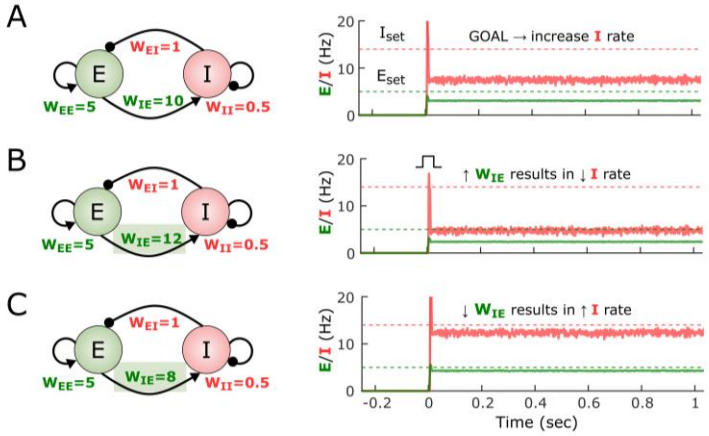


**Figure 3. The paradoxical effect constrains the learning rules that can lead to Up-states.**
**(A)** Example of the self-sustained dynamics of a two-population model with weight values shown in the diagram. Both the *E* and *I* firing rates fall below their respective setpoints. The objective is to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

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

Given that a standard set of homeostatic learning rules did not robustly lead to self-sustained dynamics we explored alternative learning rules. By defining a loss function based on the sum of the excitatory and inhibitory errors we analytically derived a set of learning rules using gradient descent (see **Section 3** in the Supplementary Material). 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, from the perspective of an excitatory neuron these rules can be interpreted as cells having a setpoints for the inhibitory input current onto the cell. Such input could be read by a cell as the activation of metabotropic receptors (e.g, $GABA_b$ and mGlu; see Discussion). Indeed, a similar cross-homeostatic rule has been derived for $W_{I\leftarrow E}$ weights (Mackwood et al., 2021).

An example of the performance of the cross-homeostatic rules is shown in **Fig. 4B-C**. After an initial phase with no self-sustained firing (Trial 1), recurrent activity reaches a stable Up-state (Trial 20), whose average rate continues to converge towards its defined setpoints (Trial 100) until the learning rule system reaches steady state (Trial 500). 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, and led to balanced dynamics (**Fig. 4D-E**). The weight trajectory from its initial value to its final one is shown for 100 different simulations (**Fig. 4D**). Each line corresponds to individual experiments with different 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; see **Section 2.1** in the Supplementary Material). 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 the "free" weights $W_{E\leftarrow E}$ and $W_{I\leftarrow E}$, respectively. This is a direct consequence of the steady state conditions for the nontrivial fixed-point of the two-population model (Tsodyks et al., 1997; Ozeki et al., 2009; Jercog et al., 2017), where the slope of the line is defined by the setpoints $E_{set}/I_{set}$ (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. Once the weights reach this specific relationship, the *E* and *I* rates reach their corresponding $E_{set}$ and $I_{set}$ values (**Fig. 4E**). Numerical simulations confirm that the cross-homeostatic rule robustly guides Up-states to different $E_{set}$ and $I_{set}$ setpoints (**Fig. 4F**), whose ratios define the slopes of the final relationship between the weights (**Fig. 4G**).

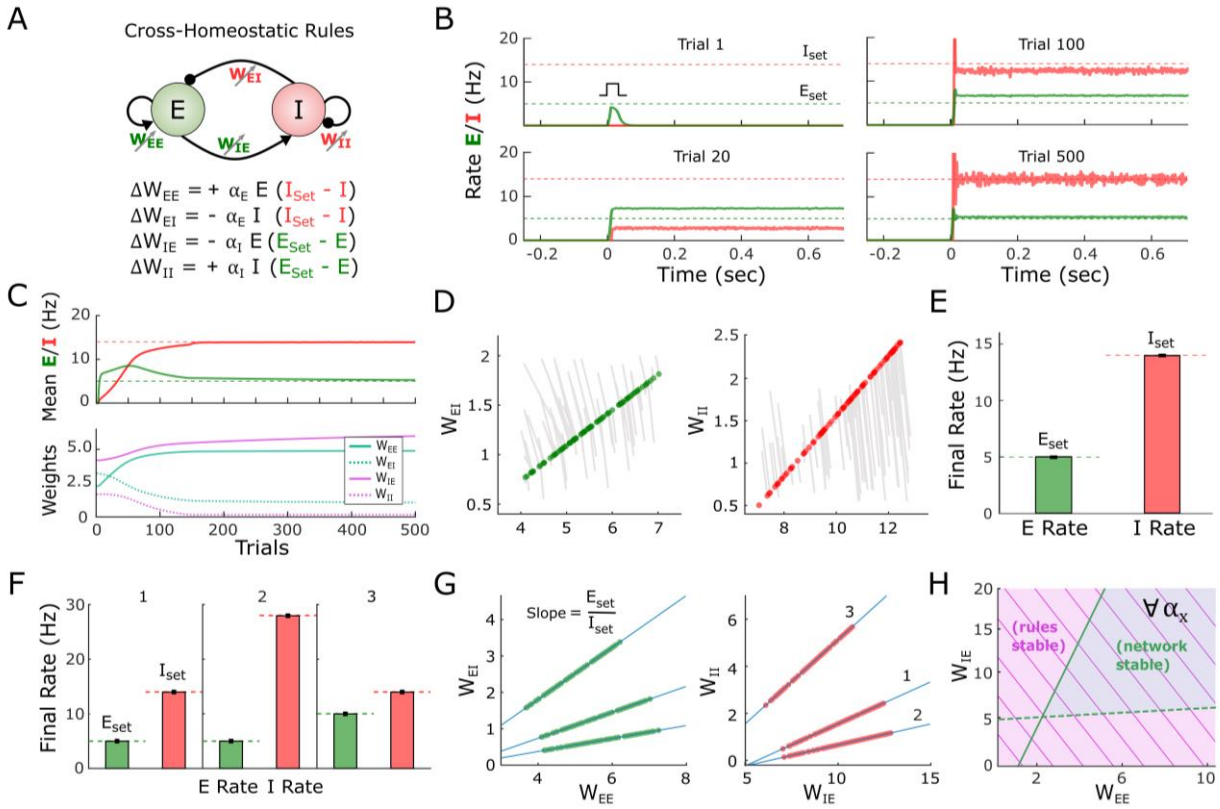


**Figure 4. A family of cross-homeostatic learning rules robustly lead to self-sustained dynamics at $E_{set}$ and $I_{set}$.**

**(A)** Schematic of the network 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, weights are initialized to $W_{EE}$=2.1 $W_{EI}$=3 $W_{IE}$=4 $W_{II}$=2). By Trial 20 a stable Up-state is observed, but at firing rates far from the target setpoints (dashed lines). By Trial 500 the network has converged to an Up-state in which *E* and *I* firing rate match their respective setpoints

**(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 cross-homeostatic rules across trials for the data shown in (B)

**(D)** Weight changes for 100 different simulations with random weight initializations (see Methods). Lines show change from initial to final (circles) weight values.

**(E)** Average final rates for 100 independent simulations with different weight initializations shown in (D). Data represents mean ± SEM.

**(F)** Final rates for the excitatory and inhibitory subpopulations after learning with same starting conditions as in (D) and (E) but for different setpoints. 1: $E_{set}$=5, $I_{set}$=14; 2: $E_{set}$=5, $I_{set}$=28; 3: $E_{set}$=10, $I_{set}$=14. Data shown in (D) and (E) corresponds to 1. Data represents mean ± SEM.

**(G)** Final weight values for homeostatic plasticity simulations for the three different pairs of setpoints shown in (F). 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).

**(H)** Analytical stability regions of the neural subsystem and learning rule subsystem as a function of $W_{EE}$ and $W_{IE}$. The stability condition holds for any possible combination of learning rates (see **Section 2.4** in the Supplementary Material)

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 family of 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 1.3** 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 (**Fig. 4H**). Therefore, it is possible to formally establish that the cross-homeostatic learning rules are inherently stable, and can robustly account for the emergence and maintenance of self-sustained inhibition-stabilized dynamics in the two-population model.

## Cross-homeostatic rules drive average activity in a multi-unit model to setpoints

The previous results demonstrate the robustness of the cross-homeostatic family of rules in driving a two-subpopulation rate model to a stable Up-state. We next examined if these rules are also effective when considering a multi-unit model in which there are many excitatory and inhibitory units. The firing-rate model was composed of 80 excitatory and 20 inhibitory recurrently connected neurons (**Fig. 5A**). In this case, individual neurons adjust their weights to minimize the average error of their presynaptic partners (see Methods). Starting with a random weight initialization, the network reaches stable self-sustained dynamics (**Fig. 5B-C**). However, individual units converge to different final rate values, satisfying the defined setpoints only as an average (green and red thick lines of **Fig. 5B**). This is a result of the nature of the cross-homeostatic rules: neurons adjust their weights to minimize the error of the *mean* activity of its presynaptic partners. For this reason, although the network is globally balanced, single units do not converge to the same balanced E-I line attractor (**Fig. 5D-E**), and little structure is observed in the connectivity matrix after learning (**Fig. 5F**). Simulations across 400 different initialization conditions demonstrate that the rules lead the average excitatory and inhibitory population activity to $E_{set}$ and $I_{set}$, respectively (**Fig. 5G-H)**. The cross-homeostatic rules are thus capable of driving a multi-unit model to a stable Up-regime, but they do not guide individual units to local setpoints.

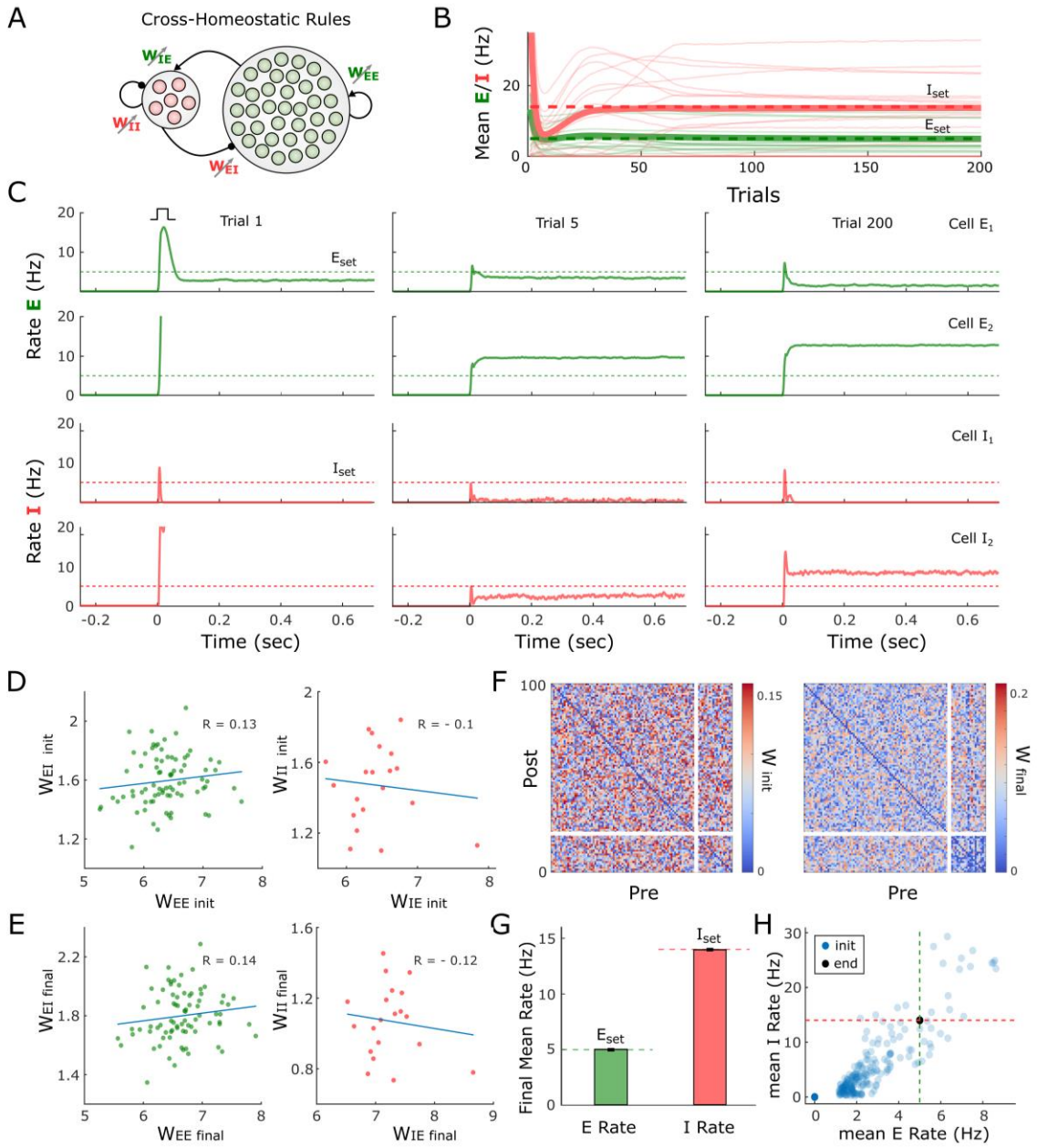


**Figure 5. Cross-homeostatic rules drive a multi-unit firing rate model to a global network balance.**

**(A)** Schematic (left) of the multi-unit rate model. The network is composed of 80 excitatory and 20 inhibitory units recurrently connected. The four weight classes are governed by cross-homeostatic learning rules (right). See Methods for a detailed explanation of the implementation.

**(B)** Evolution of the average rate across trials of 20 excitatory and inhibitory units in an example simulation. The network is initialized with random weights (see Methods) and so neurons present diverse initial rates. $E_{set}$=5 and $I_{set}$=14 represent the target homeostatic setpoints. Red and green lines represent the individual (thin lines) and average (thick lines) firing rate of inhibitory and excitatory population, respectively.

**(C)** Example of the firing rate of two excitatory and two inhibitory units at different points in (B). 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. Individual units converge to a stable Up-state but not to the defined setpoint.

**(D)** E-I weight relationships at the beginning of the simulation. Every dot represents the total presynaptic weight onto a single unit. Left excitatory neurons. Right inhibitory neurons.

**(E)** Same plot as in D at the end of the simulation.

**(F)** Weight matrix for the multi-unit model at the beginning (left) and end (right) of the simulation. First 20 neurons are inhibitory.
**(G)** Average firing rate of the units of the multi-unit model and for different initializations of weights (n=400). The network converges to the setpoints in average. Data represents mean ± SEM.
**(H)** Same data as in (G) but showing the average initial rate of the network for the multiple initializations (blue dots) and the average rate at the end (black). Target rates are shown in dotted lines (green, $E_{set}$=5, red $I_{set}$=14).

## A learning rule with cross-homeostatic and homeostatic terms leads to local convergence to setpoints.

The above results demonstrate a potential limitation of the cross-homeostatic family of rules: the target setpoints are only reached at the population level. An additional and potentially more serious limitation is that cross-homeostatic rules predict that artificially altering the activity of a small number of excitatory neurons within a large network would not directly produce homeostatic plasticity in these neurons, but directly produce plasticity in their postsynaptic inhibitory neurons. This prediction seems to conflict with homeostatic plasticity experiments that have targeted specific cell types rather than globally alter activity through pharmacological means (Burrone et al., 2002; Xue et al., 2014). We therefore assessed the scenario in which both cross-homeostatic and homeostatic rules operate in parallel, resulting in a “two-term cross-homeostatic” family of rules. These rules can actually be recovered after an approximation of a gradient descent derivation on a loss function that includes the difference between $E$ and $I$ and their respective setpoints (see **Section 3** in the Supplementary Material). In a two-subpopulation model, we first confirmed that this two-term cross-homeostatic family is stable—assuming that the learning rate of the homeostatic term does not dominate (see Supplementary Material, **Section 1.4**). Simulations with the same multi-unit model as **Fig. 5** show that with the two-term cross-homeostatic rule all individual units converge to their respective $E_{set}$ and $I_{set}$ (**Fig. 6A-C**). Importantly, in contrast to the single-term cross-homeostatic rule the total excitatory and inhibitory weight of each individual unit converged to the E-I balance of the line attractor predicted by the network equations (**Fig. 6D-E**), while more structure is also observed in the weight matrices (**Fig. 6F**)—i.e., there is less homogeneity between the four synapse classes. The convergence to the setpoints was stable across a wide range of initial states (Fig. **6G-H**). Thus, a hybrid family of learning rules that includes both cross-homeostatic and homeostatic forces provide global network stability, while also locally driving each unit to their setpoint and a balanced E-I regime.

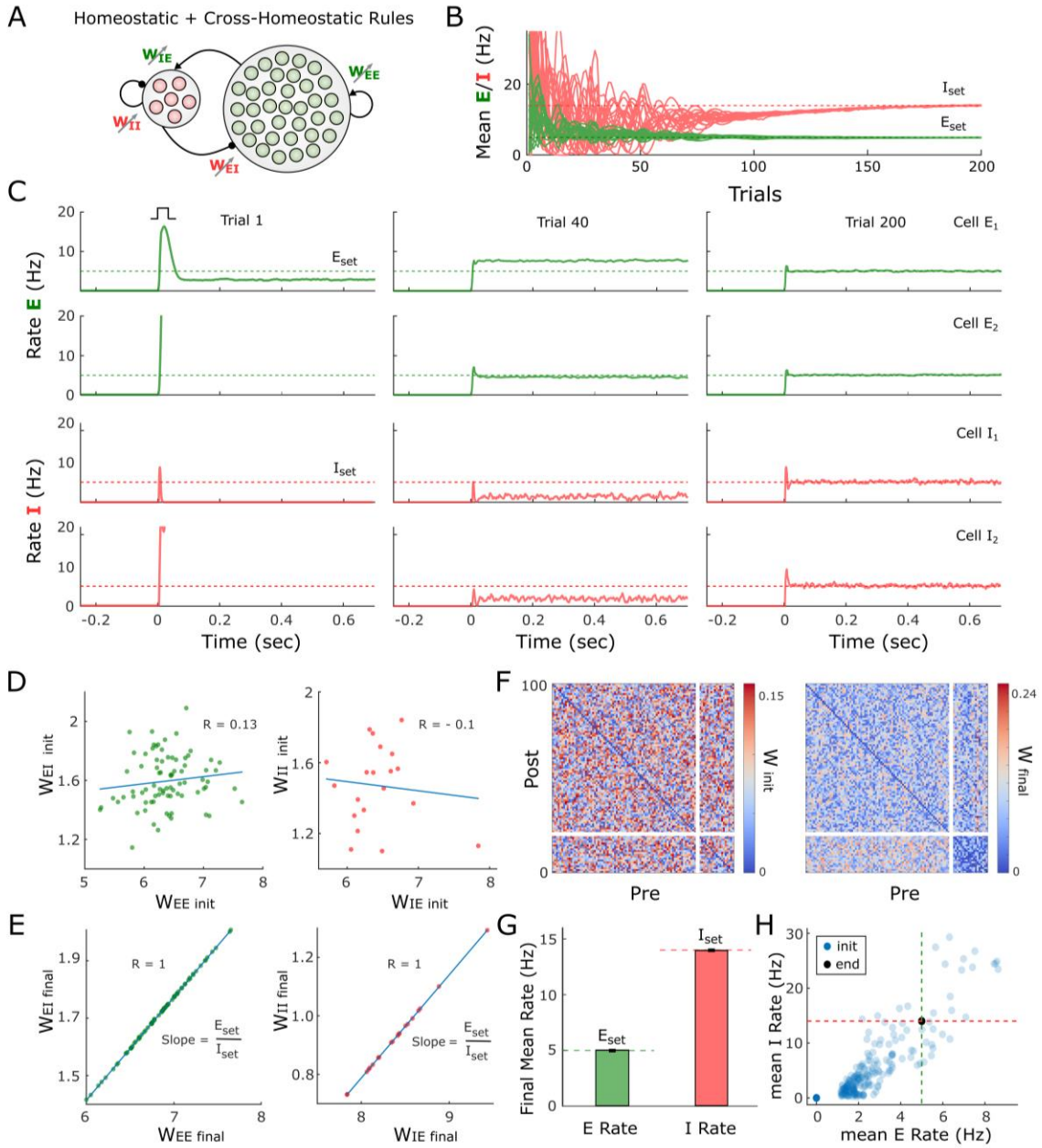


**Figure 6. Adding cross-homeostatic influences to homeostatic rules lead global and local convergence to setpoints.**

**(A)** Schematic (left) of the multi-unit rate model. The network is composed of 80 excitatory and 20 inhibitory units recurrently connected. The four weight classes are governed by homeostatic rules with cross-homeostatic influences (right). See Methods for a detailed explanation of the implementation.

**(B)** Evolution of the average rate across trials in an example simulation (20 excitatory and inhibitory units). The network is initialized with random weights (same as in **Fig. 5**, see Methods) and so neurons present diverse initial rates. $E_{set}$=5 and $I_{set}$=14 Hz represent the target homeostatic setpoints.

**(C)** Example of the firing rate of two excitatory and two inhibitory units at different points in (B). 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. Units converge to a stable Up-state and at an individual setpoint.

**(D)** E-I weight relationships at the beginning of the simulation. Every dot represents the total presynaptic weight onto a single unit. Left excitatory neurons. Right inhibitory neurons.

**(E)** Same plot as in D at the end of the simulation. The network has reached a stable state and weights converge to single E-I balance defined by a line attractor.

**(F)** Weight matrix for the multi-unit model at the beginning (left) and end (right) of the simulation. First 20 neurons are inhibitory.

**(G)** Average firing rate of the units of the multi-unit model and for different initializations of weights (n=400). Data represents mean ± SEM.

**(H)** Same data as in (G) but showing the average initial rate of the network for the multiple initializations (blue dots) and the average rate at the end (overlapping black circles). Target rates are shown in dotted lines (green, $E_{set}$; red, $I_{set}$).

## DISCUSSION

Elucidating the learning rules that govern the connectivity within neural circuits represents a fundamental goal in neuroscience, in part, because learning rules establish unifying principles that span molecular, cellular, systems, and computational levels of analyses. Elucidation of Hebbian associative synaptic plasticity, for example, linked simple computations at the level of single proteins (the NMDA receptor) with higher-order computations at the systems and computational levels (Hebb, 1949; Miller et al., 1989; Buonomano and Merzenich, 1998; Martin et al., 2000; Song et al., 2000). However, it remains the case that relatively little is known about the learning rules that give rise to complex neural dynamic regimes. Here we have taken steps towards exploring families of learning rules that operate in parallel at four different synapse classes and capture the experimentally observed emergence of Up-states in cortical networks.

Towards this goal we first confirmed that, in agreement with previous studies (Johnson and Buonomano, 2007; Motanis and Buonomano, 2020), Up-states emerge over the course of weeks of *ex-vivo* development. Because cortical organotypic cultures maintain much of their local and laminar architecture, and are mostly isolated from other cortical and subcortical inter-areal connectivity (Bolz, 1994; Echevarria and Albus, 2000; De Simoni et al., 2003), these results suggest the presence of local learning rules that lead to self-sustained, inhibition-stabilized activity in the absence of any supervisory, modulatory, or structured input signals from other brain areas.

We first explored whether standard formulations of homeostatic plasticity can account for the unsupervised emergence of Up-states—or more generally of self-sustained, inhibition-stabilized regimes. Based on experimental data we assumed that both excitatory and inhibitory neurons have an ontogenetically programmed activity setpoint during Up-states and that plasticity at the four weight classes is driven by homeostatic plasticity. Numerical simulations and analytical stability analyses revealed that while some initial conditions and parameter regimes led to self-sustained dynamics, they occupied a relatively narrow region of parameter space: when the rate of synaptic plasticity onto inhibitory neurons is much lower than that onto excitatory neurons (**Fig. 2**, and Supplementary Materials). When the rate of inhibitory and excitatory plasticity are comparable, analytical stability analyses confirmed that the region of stability of the network dynamics only overlapped in a narrow region. Such a narrow stability area seems incompatible with the robustness necessary in biological systems, and with experimental data showing that inhibitory neurons exhibit as much or more homeostatic plasticity than excitatory neurons (Keck et al., 2011; Kuhlman et al., 2013; Gainey et al., 2018; Ma et

al., 2019). We thus conclude that a family of standard homeostatic learning rules operating at all four synapse classes is not sufficient to account for the experimentally observed emergence of self-sustained dynamics in cortical circuits.

## Cross-homeostatic plasticity

Analyses of approximations of a gradient-descent-derived learning rule suggested, somewhat counterintuitively, that adjusting the *E* population based on the error of the *I* population (and vice-versa) may prove to be an effective family of learning rules. Indeed, numerical simulations and analytical stability analyses revealed that this cross-homeostatic rule was robustly stable (**Fig. 4**). The convergence to the excitatory and inhibitory setpoints, however, only occurred at the population level, not at the level of individual units. This observation, however, is not inconsistent with experimental data, which shows that *in vivo* neurons do exhibit a wide range of variability in their apparent setpoints (Hengen et al., 2016; Trojanowski et al., 2020). However, a significant concern with this single-term cross-homeostatic rule is that it predicts that selectively increasing activity in a subpopulation of excitatory neurons would first induce plasticity in inhibitory neurons ($W_{I \leftarrow E}$ and $W_{I \leftarrow I}$)—which could in turn lead to plasticity in the manipulated excitatory neurons ($W_{E \leftarrow E}$ and $W_{E \leftarrow I}$). Most homeostatic plasticity studies do not speak to this prediction because they have used pharmacological manipulations of both excitatory and inhibitory neurons. However, some studies have used cell-specific manipulations—e.g., cell-specific overexpression of potassium channels (Burrone et al., 2002; Xue et al., 2014)—that strongly support the notion that synaptic plasticity is guided at least in part on their own deviation from setpoint.

In our opinion, and although we have explored alternative rules (see Supplementary Material, **Section 1.6**), the most biologically plausible set of learning rules that lead to stable Up-states comprises a hybrid rule that includes both standard homeostatic and cross-homeostatic terms. Such a two-term cross-homeostatic rule robustly led to a self-sustained, inhibition-stabilized network, led to all units converging to their setpoints, and is directly consistent with current experimental data.

## Biological plausibility of cross-homeostatic plasticity

While the neural mechanisms underlying homeostatic plasticity remain to be elucidated, it is generally assumed that an individual neuron can maintain a running average of their firing rate over the course of hours as a result of $Ca^{2+}$-activated sensors. Based on the deviation of this value from an ontogenetically determined setpoint, neurons up- or down-regulate the density of postsynaptic receptors accordingly (Liu et al., 1998; Joseph and Turrigiano, 2017; Trojanowski et al., 2020). Two-term cross-homeostatic plasticity would require an additional, and apparently non-local information about the error in a given neuron's presynaptic partners. It is important to stress, however, that this rule is not necessarily a non-local rule, because any postsynaptic neuron has access to the mean activity of its presynaptic partners simply as a result of its postsynaptic receptor activation. Indeed, a plasticity rule for $W_{I \leftarrow E}$ weights with a similar cross-homeostatic error term has also been recently proposed and implemented based on the mean activation of

postsynaptic receptors—more specifically the net postsynaptic currents which provide a coupled measure of average presynaptic firing and synaptic weights (Mackwood et al., 2021).

Here we propose that cross-homeostatic plasticity could be implemented through metabotropic postsynaptic metabotropic receptors—e.g., mGlu and $GABA_b$. Such receptors would provide a mechanism for postsynaptic neurons to maintain a running average of the activity of its presynaptic partners that is decoupled from the synaptic weights. Metabotropic receptors are G-protein coupled receptors (GPRC) that provide a low-pass filtered measure of presynaptic activity and are involved in a large number of incompletely understood neuromodulatory roles (Blein et al., 2000; Niswender and Conn, 2010). Since metabotropic receptors appear to undergo less homeostatic and associative plasticity, they provide a measure of presynaptic activity that is naturally decoupled from the ionotropic receptors (e.g., AMPA and $GABA_a$) that are being up- and down-regulated.

Further support for the notion that individual neurons have access to global network activity emerges from studies suggesting that neurons might not homeostatically regulate activity at the individual neuron level, but rather at the global population level (Slomowitz et al., 2015). Such a global-level homeostasis could be achieved by non-synaptic paracrine transmission. Indeed, retrograde messenger systems are ideally suited for this role, as they have already been implicated in signaling mean activity levels to local capillaries, driving the activity-dependent vasodilation that underlies fMRI (Drew, 2019).

## The paradoxical effect and standard homeostatic rules

The paradoxical effect is one of the defining features of inhibition-stabilized networks, and a growing body of evidence suggests that Up-states and other self-sustained dynamic regimes are instantiations of inhibition-stabilized networks (Zucca et al., 2017; Mahrach et al., 2020; Sanzeni et al., 2020; Sadeh and Clopath, 2021). Here we show that the paradoxical effect applies important constraints to the potential learning rules that lead to the emergence of inhibition-stabilized networks. In the simplified case in which there is only homeostatic plasticity onto the inhibitory neurons, we can immediately see why the paradoxical effect renders standard homeostatic rules ineffective. If the *I* population is below its setpoint, the standard homeostatic rules would increase $W_{I \leftarrow E}$, which paradoxically would further decrease *I* (**Fig. 3**), thus further increasing the error instead of decreasing it (**Fig. 3**). This reasoning is related to why, when using the standard family of homeostatic rules, the rate of plasticity onto the inhibitory neurons has to be much smaller—in effect dampening the "paradoxical homeostatic plasticity effect". Furthermore, our analytical stability analyses show that in the limit of vanishingly small excitatory learning rates ($\alpha_{EE,EI} << \alpha_{IE,II}$) the stability region of the weight subsystem is bounded by the paradoxical condition. This means that the only allowed stable states with non-zero *E* activity will occur in the non-paradoxical regime, if any, and they will not be proper inhibition-stabilized Up-states.

## Future directions and experimental predictions

While we implemented homeostatic learning rules at all four synapses classes in our model, it is important to stress that we have omitted other well-characterized forms of synaptic plasticity. Of particular relevance, we did not include associative LTP or STDP. These forms of plasticity are generally considered to capture the correlation structure in networks which are driven by structured inputs. Arguably, because our circuits develop in the absence of any structured external input and because all excitatory and inhibitory neurons synchronously shift between Down↔Up states, it is possible that associative forms of plasticity do not contribute significantly to Up-state development. Nevertheless, future experimental and theoretical studies have to address the potential role for associative forms of synaptic plasticity in Up-state development.

An important implication of our results is that neuronal and network properties can operate in fundamentally different ways. That is, while homeostatic plasticity can lead to single neurons to reach their target setpoints in simple feedforward circuits, those same rules can be highly unstable when the neurons are placed even in the simplest of recurrent excitatory/inhibitory circuits with emergent dynamics. Furthermore, because emergent neural dynamic regimes are highly nonlinear, and in particular, that stable self-sustained dynamic regimes exhibit a paradoxical effect, it is likely that the brain exhibits “paradoxical” or counterintuitive learning rules to generate self-sustained dynamic regimes.

650 **METHODS**

651

652 ***Ex vivo slice preparation***

653 Organotypic slices were prepared using the interface method (Stoppini et al., 1991; Goel
654 and Buonomano, 2016). Briefly, five to seven day-old WT and PV-Cre mice were
655 anesthetized with isoflurane and decapitated. The brain was removed and placed in
656 chilled cutting media. Coronal slices (400 µm thickness) containing auditory cortex were
657 sliced using a vibratome (Leica VT1200) and placed on filters (MillicellCM, Millipore,
658 Billerica, MA, USA) with 1 mL of culture media. Culture media was changed at 1 and 24
659 hours after cutting and every 2-3 days thereafter. Cutting media consisted of EMEM
660 (MediaTech cat. #15-010) plus (final concentration in mM): $MgCl_2$, 3; glucose, 10;
661 HEPES, 25; and Tris-base, 10. Culture media consisted of EMEM plus (final
662 concentration in mM): glutamine, 1; $CaCl_2$, 2.6; $MgSO_4$, 2.6; glucose, 30; HEPES, 30;
663 ascorbic acid, 0.5; 20% horse serum, 10 units/L penicillin, and 10 µg/L streptomycin.
664 Slices were incubated in 5% CO2 at 35°C.

665

666 ***Two-photon Calcium Imaging***

667 Organotypic slices from PV-Cre mice (Jackson Laboratory #017320) were transfected at
668 1-3 DIV with pENN-AAV-CaMKII-GCaMP6f-WPRE-SV40 to selectively express
669 GCaMP6f in excitatory neurons, and pAAV-CAG-Flex-mRuby2-GSG-P2A-GCaMP6f-
670 WPRE-pA to visualize and selectively express GCaMP6f in $PV^+$ inhibitory neurons.
671 Transfection was achieved by gently delivering 1µL of each virus onto the slice using a
672 micropipette. Experiments were performed at least 12 days after transfection to allow for
673 robust expression.

674

675 Calcium imaging was performed with a galvo-resonant-scanning two-photon
676 microscope (Neurolabware) controlled by Scanbox acquisition software
677 (https://scanbox.org). A Coherent Chameleon Ultra II Ti:sapphire laser (Cambridge
678 Technologies) was used for GCaMP6f (920 nm) and mRuby excitation (1040 nm). A
679 16x water-immersion lens (Nikon, 0.8 NA, 3 mm working distance) was used. Image
680 sequences were captured using unidirectional scanning at a frame rate of ~ 15 Hz. The
681 size of the recorded imaging field was ~ 520 × 800 µm (512 × 796 pixels). Five min of
682 spontaneous activity was recorded at 920 nm at every developmental time point. Before
683 the recording a snapshot at 1040 nm was recorded in order to identify $PV^+$ neurons.
684 Regions of interests (ROI) for both excitatory and $PV^+$ neurons were established using
685 the imaging processing pipeline Suite2p (https://github.com/MouseLand/suite2p)
686 (Pachitariu et al., 2017). $\Delta F/F$ was calculated as $(F(t) - F0))/F0$, where F(t) was the raw
687 fluorescence filtered with a median filter with a window of 1 s. F0 was the running median
688 F(t) over the previous 20 s window. For each recorded slice and neural population
689 (excitatory or $PV^+$), potential Up-states were identified based on a threshold set at 1 above
690 the mean z-scored raw fluorencence F trace of all neurons. If these events remained
691 above threshold for at least one second, they were classified as Up-states. Mean Up

Frequency and Duration were computed over all Up-states detected within the 5 min spontaneous activity period.

***Computational model***

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

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

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

where $W_{XY}$ represents the weight between the presynaptic unit $Y$ and postsynaptic unit $X$. $\tau_X$ and $\eta_X$ represent a time constant and an independent noise term, respectively. The time constants were set to $\tau_E = 10ms$ for the excitatory and $\tau_I = 2ms$ for the inhibitory subpopulations. The noise term was an Ornstein-Uhlenbeck process with mean $\mu_x = 0$, a time constant $1/\theta_x = 1ms$, and a sigma parameter of $\sigma_x = 10$. To elicit Up-states a step current was injected at the beginning of each trial on the excitatory population.

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

$$f_Y(x) = \begin{cases} 0 \;\; if \;\; x < \theta_Y \\ g_Y(x - \theta_Y) \;\; if x \geq \theta_Y \end{cases}, Y = \{E, I\} \quad (3)$$

As in Jercog et al (2017) the thresholds were set to $\theta_E = 4.8$ and $\theta_I = 25$, and the gains to $g_E = 1$ and $g_I = 4$. The higher thresholds in PV neurons are consistent with experimental findings (Romero-Sosa et al., 2021).

The linear relationship between excitatory and inhibitory weights (**Fig. 4**) correspond to the steady-state solution of the neural subsystem when the inhibitory and excitatory rates are at its target setpoints. The solution can be obtained by setting the left side of equations (1) and (2) to zero, and substituting the steady state $E$ and $I$ values by $E_{Set}$ and $I_{Set}$.

$$W_{EI} = \frac{W_{EE}E_{Set}}{I_{Set}} - \frac{\theta_E g_E + E_{Set}}{I_{Set} g_E} \quad (4)$$

$$W_{II} = \frac{W_{IE}E_{Set}}{I_{Set}} - \frac{\theta_I g_I + I_{Set}}{I_{Set} g_I} \quad (5)$$

Thus, the slope of the E/I balance line in **Fig. 4** corresponds to $E_{Set}/I_{Set}$. See details and analytical results in **Section 2.2** of the Supplementary Material.

### *Synaptic plasticity*

Plasticity at all four weight classes ($W_{E\leftarrow E}$, $W_{E\leftarrow I}$, $W_{I\leftarrow E}$, $W_{I\leftarrow I}$) was governed by different families of homeostatic based learning rules, all driven by the deviation of the actual excitatory and inhibitory rates from their target setpoints ($E_{set}$ and $I_{set}$). Three different learning rules are presented in the main text of this paper.

*Standard homeostatic family of rules*:

$$(6) \qquad \begin{aligned} \Delta W_{EE} &= +\alpha_E E(E_{set} - E) \\ \Delta W_{EI} &= -\alpha_E I(E_{set} - E) \\ \Delta W_{IE} &= +\alpha_I E(I_{set} - I) \\ \Delta W_{II} &= -\alpha_I I(I_{set} - I) \end{aligned}$$

where $\alpha_E$ and $\alpha_I$ are the learning rates onto the excitatory and inhibitory units, respectively. All alphas are set to equal values in the simulation data shown in **Fig. 2** ($\alpha = 0.0001$). The setpoints were based on empirically measured values in *ex vivo* cortical circuits (Romero-Sosa et al., 2021): $E_{set} = 5$ and $I_{set} = 14$ Hz.

The configuration of setpoints follows a classic homeostatic formulation (Turrigiano et al., 1998; Rossum et al., 2000; Liu and Buonomano, 2009; Vogels et al., 2011), where every neural population adapts its input weights homeostatically in order to minimize its error term. As outlined in Supplementary Material (**Section 1.5**) we also examined variants of this formulation, such as standard synaptic scaling (which includes the weight as factor).

We prove that these rules are only stable in a narrow parameter regime (when excitatory plasticity dominates). See details and analytical results in **Section 2** of the Supplementary Material.

*Single-term cross-homeostatic family of rules*:

$$(7) \qquad \begin{aligned} \Delta W_{EE} &= +\alpha_E E(I_{set} - I) \\ \Delta W_{EI} &= -\alpha_E I(I_{set} - I) \\ \Delta W_{IE} &= -\alpha_I E(E_{set} - E) \\ \Delta W_{II} &= +\alpha_I I(E_{set} - E) \end{aligned}$$

All alphas are set to equal values in the simulation data shown in **Fig. 4** ($\alpha = 0.0001$), except on the example shown in **Fig. 4B-C**, where a rate of $\alpha = 0.0005$ was used. We note that an equivalent rule for $W_{IE}$ has been recently derived (Mackwood et al., 2020). For **Fig. 4** two alternative pairs of setpoints were explored ($E_{set} = 5$ and $I_{set} = 24$) and ($E_{set} = 10$ and $I_{set} = 14$).

We prove that these rules are stable for any set of parameters. See details and analytical results in **Section 1.3** of the Supplementary Material.

*Two-term cross-homeostatic family of rules*:

$$
\begin{aligned}
\Delta W_{EE} &= +\alpha_E E(E_{set} - E) + \alpha_E E(I_{set} - I) \\
\Delta W_{EI} &= -\alpha_E I(E_{set} - E) - \alpha_E I(I_{set} - I) \\
\Delta W_{IE} &= +\alpha_I E(I_{set} - I) - \alpha_I E(E_{set} - E) \\
\Delta W_{II} &= -\alpha_I I(I_{set} - I) + \alpha_I I(E_{set} - E)
\end{aligned} \tag{8}
$$

A single shared learning rate ($\alpha = 0.00001$) was used for **Fig. 6** (see section below on the multi-unit model). We prove that these rules are stable for a biologically meaningful set of parameter values, as long as the homeostatic part does not dominate (Supplementary Material, **Section 1.4**). The two-term rules combine homeostatic and cross-homeostatic terms. This formulation can be obtained after an approximation of a gradient descent derivation on the following loss function:

$$
L = \frac{1}{2}(E - E_{set})^2 + \frac{1}{2}(I - I_{set})^2 \tag{9}
$$

The mathematical derivation can be found in the Supplementary Material (**Section 3**).

An additional *Forced-Balance* learning rule, which exploits the steady-state solution of the neural subsystem at its target setpoints (equations 5 and 6), has also been explored (see **Section 1.6** of the Supplementary Material).

*All rules, numerical simulations:*
For all simulations, the weights were updated after the completion of every trial. The trials lasted 2 seconds. 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 $\tau_{trial} = 2$. The numerical integration time step was 0.1 ms. A minimum weight of 0.1 was set for all weights.

A saturation to the excitatory and inhibitory firing rate (100 and 250 Hz, respectively) was added to prevent the nonbiological scenario in which activity could diverge towards infinity under unstable conditions. Note the saturation is not necessary for the cross-homeostatic rule because it is inherently stable as proved in the Supplementary Material (**Section 1.3**).

In **Fig. 2D** and **4D-G** we initialize the weights uniformly in between the following ranges: $W_{EE}$[4,7], $W_{EI}$[0.5,2], $W_{IE}$[7,13], $W_{II}$[0.5,2]. Simulations were run for 3000 trials to assess stability and convergence.

*All rules, analytical stability analyses:*

We analyzed the entire dynamical system (composed of the neural subsystem and learning rule subsystem) for every synaptic learning rule considered in this work, and analyzed its stability. In every case, the general prescription is:

a) Take the combined neural and learning rule subsystems and nondimensionalize all variables, so that the two different time scales are evident (fast neural, slow synaptic plasticity). For the description of the learning rule subsystem we switch from discrete-time dynamics to continuous-time dynamics: $\Delta W \rightarrow \tau_0 \, dW/dt$

b) Make a quasi-steady state (QSS) approximation of the neural subsystem. This means we will consider the neural subsystem is fast enough so that it converges "instantaneously" (when compared to the synaptic plasticity subsystem) to its corresponding fixed point. For this we will require that the stability conditions of the neural subsystem are satisfied (see below).

c) Find the steady-state solution of the synaptic plasticity subsystem, i.e. the Up-state fixed point; compute the Jacobian of the synaptic plasticity subsystem at the Up-state; compute the eigenvalues of the Jacobian. Two out of the four eigenvalues are expected to be zero because the Up-state is not an isolated fixed point of the system but a continuous 2D plane in 4D weight space.

d) Address (linear) stability. If both nonzero eigenvalues have negative real parts, then the Up-state is stable under the learning rule; if at least one of the nonzero eigenvalues has positive real part, then the Up-state is unstable. (A note on abuse of notation: we might say indistinctly "the Up-state is stable/unstable" and "the learning rule is stable/unstable".)

See **Section 2** in the Supplementary Material.

***Multi-unit firing rate model***

A rate-based recurrent network model containing $N_e = 80$ excitatory and $N_i = 20$ inhibitory neurons was implemented with all-to-all connectivity (without self-connections). The activation of the neurons followed equations (1), (2) and (4). The same parameters as for the population model were used, where $W_{XY}$ represents now a matrix of synaptic weights from population $X$ to population $Y$. A minimum weight of $0.1/N_x$ for $W_{EI}$ and $W_{IE}$ and $0.1/(N_x\text{-}1)$ for $W_{EE}$ and $W_{II}$ was set for all weights.

The synaptic plasticity rules were implemented as follows.

*Cross-homeostatic family of rules:*

$$\Delta W_{ij}^{EE} = +\alpha E_j \sum_{k=1}^{N_I} (I_{set} - I_k)/N_I \qquad (10)$$

$$\Delta W_{ij}^{EI} = -\alpha I_j \sum_{k=1}^{N_I} (I_{set} - I_k)/N_I$$

$$\Delta W_{ij}^{IE} = -\alpha E_j \sum_{k=1}^{N_E} (E_{set} - E_k)/N_E$$

$$\Delta W_{ij}^{II} = +\alpha I_j \sum_{k=1}^{N_E} (E_{set} - E_k)/N_E$$

Where *i* and *j* represent the post- and presynaptic neurons, respectively, and *k* denotes the presynaptic inhibitory neurons targeting the excitatory neurons (or the presynaptic excitatory neurons targeting an inhibitory neuron). $N_E$ and $N_I$ denote the total number of excitatory and inhibitory neurons, respectively. The weights are therefore updated following the *average* presynaptic error of the crossed E/I population classes. Note as stated above that this formulation can be implemented in a local manner (see Discussion). A learning rate of $\alpha = 0.00002$ was used for all simulations.

*Two-term cross-homeostatic family of rules:*

$$\Delta W_{ij}^{EE} = +\alpha E_j (E_{set} - E_i) + \alpha E_j \sum_{k=1}^{N_I} (I_{set} - I_k)/N_I \quad (11)$$

$$\Delta W_{ij}^{EI} = -\alpha I_j (E_{set} - E_i) - \alpha I_j \sum_{k=1}^{N_I} (I_{set} - I_k)/N_I$$

$$\Delta W_{ij}^{IE} = +\alpha E_j (I_{set} - I_i) - \alpha E_j \sum_{k=1}^{N_E} (E_{set} - E_k)/N_E$$

$$\Delta W_{ij}^{II} = -\alpha I_j (I_{set} - I_i) + \alpha I_j \sum_{k=1}^{N_E} (E_{set} - E_k)/N_E$$

Here the first term represents the standard homeostatic rule, and the second term cross-homeostatic plasticity (as implemented above). A learning rate of $\alpha = 0.00001$ was used for all simulations.

In **Fig. 5G-H** and **6G-H** we initialize the mean weights of the population uniformly in between the following ranges: $W_{EE}$[1,6], $W_{EI}$[0.5,2], $W_{IE}$[5,7], $W_{II}$[0.5,2]. The weights

within each class were normally distributed around that mean (normalized by the number of neurons) with a variance of 0.1. Note that this initialization led to multiple initial conditions with exploding network rates (which were held in check by the saturation cutoff). Those initial rates are not displayed in **Fig. 5-6H** for visualization purposes, but the rules successfully brought all those cases to the corresponding setpoints. Simulations were run for 1000 trials to assess stability of the convergence. In the example shown in **Fig. 5A-F** and **6A-F** the weights were initialized uniformly in the interval [0 0.16] and the simulation was run for 200 trials.

***Statistics and Software availability***

Data are represented by the mean ± SEM. In **Fig. 1** a two-way ANOVA was performed to assess interaction of time and group (cell-type) on the development of Up-states.

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 **Fig. 2**, **4**, **5** and **6** is available at https://github.com/saraysoldado/UpDev2021. The Jupyter notebooks with SageMath code to reproduce all analytical results are available at: https://github.com/SMDynamicsLab/UpDev2021.

## SUPPLEMENTARY FIGURES

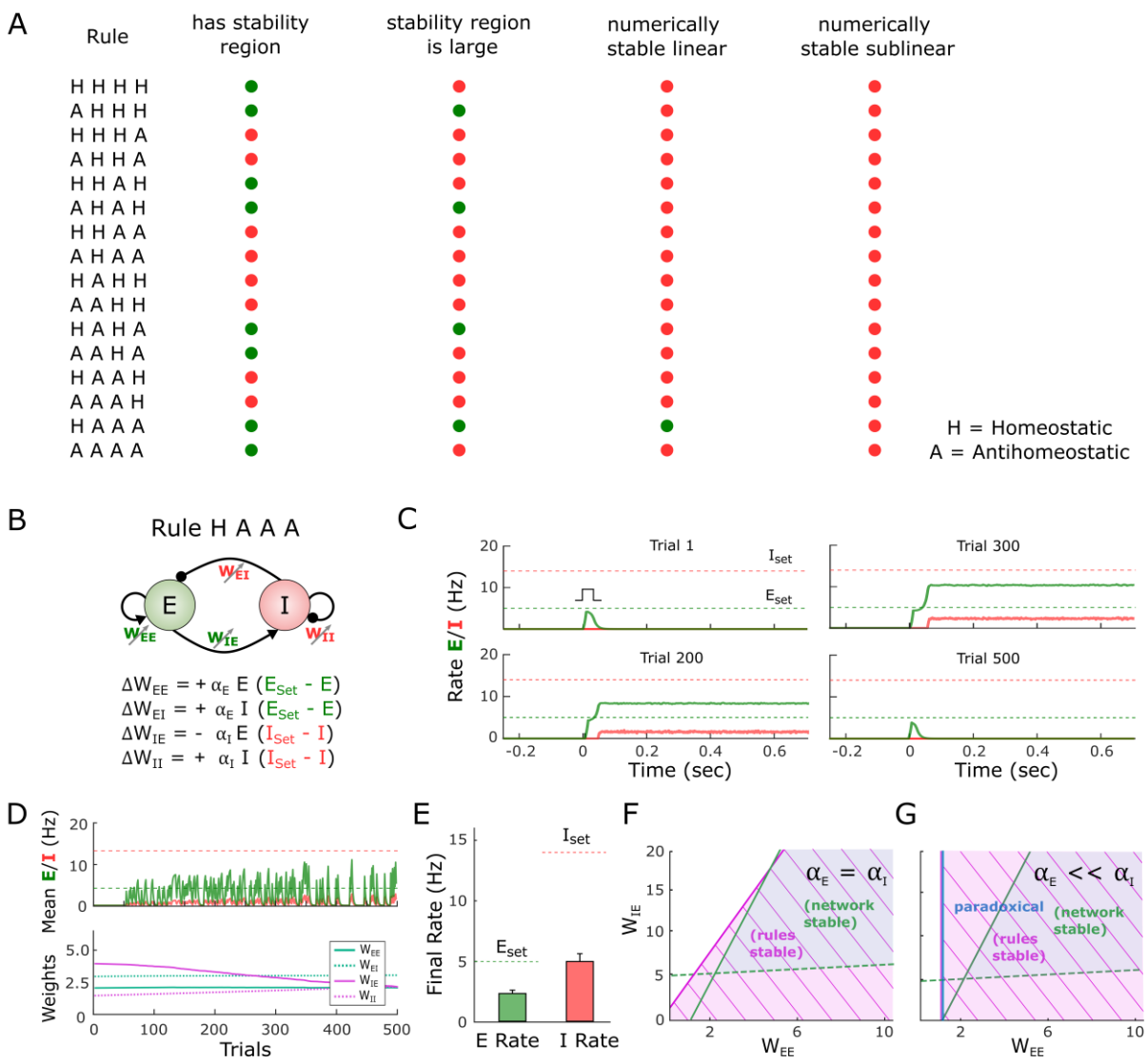


**Figure S1. Homeostatic and anti-homeostatic combinations of learning rules also fail to generate the emergence of self-sustained dynamics.**

**(A)** Sixteen variations of the standard homeostatic rules presented in **Fig. 2** were assessed for stability. The learning governing each four weight types, $W_{EE}$, $W_{EI}$, $W_{IE}$, $W_{II}$ was set to be either homeostatic (H) or antihomeostatic (A). The first rule on the table (HHHH) corresponds to the standard homeostatic rules presented in **Fig. 2**, where all weights obey homeostatic learning. All rules were tested for stability analytically and numerically. A red dot implies that the listed condition is not satisfied, while a green dot means that it does. The condition on the first column indicates whether a stability region for the learning rule is present. The second column indicates whether such region has a large overlap with the region of stability of the neural subsystem. The third column indicates whether the rule is successful, using numerical simulations, at driving the network to a stable Up-state when starting from regimes with self-sustained activity already present (meaning the network is initialized in the linear regime). The fourth column indicates the same as the former, but with the network initialized in the sub-linear regime, where activity is not initially present (e.g., as observed early in developmental conditions).

**(B)** Schematic (top) of the population rate model in which the four weights are governed by the HAAA rule in panel (A).

**(C)** Example simulation of the HAAA rule 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 were initialized to $W_{EE}$=2.1, $W_{EI}$=3, $W_{IE}$=4, and $W_{II}$=2 as in **Fig. 2**. Note that while an evoked Up-state emerges by Trial 200 the firing rates do not converge to their setpoints, and by Trial 500 the Up-state is no longer observed.

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

**(E)** Average final rate for 100 independent HAAA simulations with different weight initializations. Those initializations included cases in which the network starts in the sublinear regime (where the initial *E* firing rate was zero or very low). The weights were initialized uniformly between the following ranges: $W_{EE}$[1,3], $W_{EI}$[0.5,1.5], $W_{IE}$[4,8], $W_{II}$[0.2,0.8]. Data represents mean ± SEM.

**(F)** Analytical stability regions of the neural and HAAA learning rule subsystems as a function of the free weights $W_{EE}$ and $W_{IE}$. Here the stability plot is obtained by considering equal learning rates for all four learning rules (as used for panels C-E).

**(G)** Similar to F but with but with $\alpha_E << \alpha_I$. Right of blue line shows the area where the network is in a paradoxical regime (defined by the condition $W_{EE} * g_E - 1 > 0$). Contrary to standard homeostatic rules (**Fig. 2**), the HAAA rule is only stable in the paradoxical region of parameter space (i.e., $W_{EE} g_E - 1 > 0$; note white area to the left of the blue line). This may explain why the rule fails at driving the network to an Up-state when starting with developmental-like conditions.

## REFERENCES

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.

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.

Blein S, Hawrot E, Barlow P (2000) The metabotropic GABA receptor: molecular insights and their functional consequences. Cell Mol Life Sci 57:635-650.

Bolz J (1994) Cortical circuitry in a dish. Curr Opinion Neurobio 4:545-549.

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.

Buzsáki G, Mizuseki K (2014) The log-dynamic brain: how skewed distributions affect network operations. Nature Reviews Neuroscience 15:264-278.

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.

De Simoni A, Griesinger CB, Edwards FA (2003) Development of rat CA1 neurones in acute versus organotypic slices: role of experience in synaptic morphology and activity. J Physiol 550:135-147.

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

Drew PJ (2019) Vascular and neural basis of the BOLD signal. Current Opinion in Neurobiology 58:61-69.

Echevarria D, Albus K (2000) Activity-dependent development of spontaneous bioelectric activity in organotypic cultures of rat occipital cortex. Brain Res Dev Brain Res 123:151-164.

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.

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.

Gainey MA, Aman JW, Feldman DE (2018) Rapid Disinhibition by Adjustment of PV Intrinsic Excitability during Whisker Map Plasticity in Mouse S1. The Journal of Neuroscience 38:4749-4761.

Goel A, Buonomano DV (2016) Temporal interval learning in cortical cultures is encoded in intrinsic network dynamics. Neuron 91:320-327.

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.

Goold CP, Nicoll RA (2010) Single-Cell Optogenetic Excitation Drives Homeostatic Synaptic Depression. Neuron 68:512-528.

Haider B, Duque A, Hasenstaub AR, McCormick DA (2006) Neocortical Network Activity In Vivo Is Generated through a Dynamic Balance of Excitation and Inhibition. J Neurosci 26:4535-4545.

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.

Hengen Keith B, Torrado Pacheco A, McGregor James N, Van Hooser Stephen D, Turrigiano Gina G (2016) Neuronal Firing Rate Homeostasis Is Inhibited by Sleep and Promoted by Wake. Cell 165:180-191.

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.

Keck T, Scheuss V, Jacobsen RI, Wierenga CJ, Eysel UT, Bonhoeffer T, Hübener M (2011) Loss of sensory input causes rapid structural changes of inhibitory neurons in adult mouse visual cortex. Neuron 71:869-882.

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.

Kuhlman SJ, Olivas ND, Tring E, Ikrar T, Xu X, Trachtenberg JT (2013) A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex. Nature 501:543-546.

Liu JK, Buonomano DV (2009) 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.

Ma Z, Turrigiano GG, Wessel R, Hengen KB (2019) Cortical Circuit Dynamics Are Homeostatically Tuned to Criticality In Vivo. Neuron 104:655-664.e654.

Mackwood O, Naumann LB, Sprekeler H (2021) Learning excitatory-inhibitory neuronal assemblies in recurrent networks. eLife 10:e59715.

Mahrach A, Chen G, Li N, van Vreeswijk C, Hansel D (2020) Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation. eLife 9:e49967.

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.

Niswender CM, Conn PJ (2010) Metabotropic Glutamate Receptors: Physiology, Pharmacology, and Disease. Annual Review of Pharmacology and Toxicology 50:295-322.

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.

Okun M, Lampl I (2008) Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nature Neuroscience 11:535-537.

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 (2021) Differential excitability of PV and SST neurons results in distinct functional roles in inhibition stabilization of Up-states. Journal Of Neuroscience in press.

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.

Rudolph M, Pospischil M, Timofeev I, Destexhe A (2007) Inhibition Determines Membrane Potential Dynamics and Controls Action Potential Generation in Awake and Sleeping Cat Cortex. J Neurosci 27:5280-5290.

Rutishauser U, Slotine J-J, Douglas R (2015) Computation in dynamically bounded asymmetric systems. PLoS Comput Biol 11:e1004039.

Sadeh S, Clopath C (2021) Inhibitory stabilization and cortical computation. Nature Reviews Neuroscience 22:21-37.

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 (2020) Inhibition stabilization is a widespread property of cortical networks. eLife 9:e54875.

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.

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.

Slomowitz E, Styr B, Vertkin I, Milshtein-Parush H, Nelken I, Slutsky M, Slutsky I (2015) Interplay between population firing stability and single neuron dynamics in hippocampal networks. eLife 4:e04378.

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.

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.

Trojanowski NF, Bottorff J, Turrigiano GG (2020) Activity labeling in vivo using CaMPARI2 reveals intrinsic and synaptic differences between neurons with high and low firing rate set points. Neuron.

Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL (1997) 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 (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.

Zucca S, D'Urso G, Pasquale V, Vecchia D, Pica G, Bovetti S, Moretti C, Varani S, Molano-Mazón M, Chiappalone M, Panzeri S, Fellin T (2017) An inhibitory gate for state transition in cortex. eLife 6:e26177.