

Reviewed Preprint
v1 • August 3, 2026
Not revised

✉ For correspondence:

dylan.festa@gmail.com

gjorgjieva@tum.de

Competing interests: No competing interests declared

Funding: See page 43

Reviewing editor: Richard Naud, University of Ottawa, Canada

© 2026, Festa et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Structured stabilization in recurrent neural circuits through inhibitory synaptic plasticity

Dylan Festa^{1,2}✉, Claudia Cusceddu^{1,3}, Julijana Gjorgjieva^{1,3,4}✉

¹School of Life Sciences, Technical University of Munich, Freising, Germany • ²Max Planck Institute for Brain Research, Frankfurt am Main, Germany • ³School of Medicine and Health, Institute for Neuroscience, Technical University of Munich, Munich, Germany • ⁴Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

eLife Assessment

This work establishes a **valuable** theoretical finding about how the spike timing dependence of inhibitory plasticity shapes recurrent network connectivity. The combination of theoretical analysis and simulations provides **convincing** evidence that effective inhibitory connectivity forms a so-called Mexican-hat profile when multiple inhibitory neuron types follow distinct learning rules. These mechanisms are thought to be implicated in the contextual modulation of neuronal responses to stimuli.

<https://doi.org/10.7554/eLife.111666.1.sa2>

Abstract

In cortical circuits, inhibitory interneurons play a dual role: they regulate overall activity levels to prevent runaway excitation, and contribute to diverse computations. While unstructured inhibitory synaptic connections achieve the first role by homeostatically regulating firing rates, computational tasks often require structured excitatory-inhibitory (E/I) connectivity. Here, we consider a broad class of pairwise inhibitory spike-timing dependent plasticity (iSTDP) rules, demonstrating how inhibitory connections can self-organize to both stabilize excitation and generate functionally relevant connectivity motifs—a process we call “structured stabilization”. We show that in both small E/I circuits and large spiking recurrent neural networks the choice of iSTDP rule can lead to either mutually connected E/I pairs, or to lateral inhibition, where an inhibitory neuron connects to an excitatory neuron that does not directly connect back to it. In a one-dimensional ring network with two inhibitory subpopulations following these distinct iSTDP rules, the effective connectivity within the excitatory units self-organizes into a Mexican-hat-like profile, with excitatory influence in the center and inhibitory influence away from the center. This leads to emergent network responses such as contextual modulation effects as in the visual cortex and spatially modular activity characteristic of developmental spontaneous activity. Our theoretical work introduces a family of rules that retains the broad applicability and simplicity of spike-timing-based plasticity, while stabilizing activity and promoting specific connectivity motifs which support emergent network computations.

Significance Statement

Neural circuits are faced with the dual challenge of keeping activity stable while modifying synaptic strengths during development and learning. A prominent mechanism to keep excitation in check is inhibitory synaptic plasticity. In recurrent circuit models, we show that plasticity rules based on spike timing at synapses from inhibitory to excitatory neurons can stabilize activity while reinforcing specific connectivity motifs. The shape of these rules dictates the resulting motifs, and determines whether inhibitory neurons form strong reciprocal connections with their excitatory partners or lateral connections onto excitatory neurons that do not connect back. In

recurrent networks with local excitatory connectivity, these motifs generate surround suppression as in the visual cortex, and long-range spatial correlations characteristic of developmental spontaneous activity.

1 Introduction

In cortical circuits, GABA-ergic interneurons constitute approximately 10% to 20% of the total neural population, and can be further divided into subtypes that exhibit distinct characteristics in connectivity, plasticity and tuning properties (Jiang et al. 2015 [↗](#); Tremblay et al. 2016 [↗](#)). Fully understanding how these diverse interneuron subtypes influence neural circuit function remains a significant challenge. A key open question is how activity-dependent plasticity of inhibitory synapses in recurrent circuits sculpts this diversity into specific connectivity motifs while maintaining stable network activity.

The dynamic balance between excitation and inhibition is crucial for generating the irregular, Poisson-like firing patterns commonly observed *in vivo* in cortical circuits (Van Vreeswijk and Sompolinsky 1998 [↗](#); Brunel 2000 [↗](#)). This balance enhances the circuit's sensitivity to input, facilitating rapid information processing (Tsodyks and Sejnowski 1995 [↗](#)). Furthermore, it supports essential computations, such as receptive field emergence (Wallenstein and Hasselmo 1997 [↗](#); Froemke et al. 2007 [↗](#); Dornn et al. 2010 [↗](#); Kleberg et al. 2014 [↗](#); D'amour and Froemke 2015 [↗](#); Rubin et al. 2017 [↗](#)), neural assembly formation (Sadeh and Clopath 2021 [↗](#)), and the storage of precise temporal firing sequences (Maes et al. 2021 [↗](#); Riquelme et al. 2023 [↗](#); Gong and Brunel 2024 [↗](#); Shao et al. 2024 [↗](#)).

Excitatory-inhibitory (E/I) balance can be achieved without any specific inhibitory connectivity structures. Often circuit models consider inhibition simply as a “blanket” that regulates excitation, and incorporate it into a global normalization term or a correction to the synaptic weights between excitatory neurons (Jin 2002 [↗](#); Vogels et al. 2011 [↗](#); Ravid Tannenbaum and Burak 2016 [↗](#); Montangie et al. 2020 [↗](#)). While this simplified approach captures fundamental properties of neural dynamics, it overlooks the possible advantages of more specific E/I connectivity patterns. Incorporating structured inhibitory connections can further refine computations and circuit properties such as E/I co-tuning (Clopath et al. 2016 [↗](#); Mackwood et al. 2021 [↗](#); Miehl and Gjorgjieva 2022 [↗](#); Eckmann et al. 2024 [↗](#)), memory storage (Mongillo et al. 2018 [↗](#); Bergoin et al. 2023 [↗](#); Miehl et al. 2023 [↗](#); Agnes and Vogels 2024 [↗](#); Meissner-Bernard, Jenkins, et al. 2025 [↗](#); Meissner-Bernard, Zenke, et al. 2025 [↗](#)), predictive coding (Schulz et al. 2021 [↗](#)), and other aspects of neural dynamics (Mongillo et al. 2018 [↗](#); Dahmen et al. 2020 [↗](#); Sadeh and Clopath 2020 [↗](#); Hafizi et al. 2022 [↗](#)).

Experimental findings demonstrate that inhibitory neurons form non-random connectivity patterns that cannot be explained solely by cell distance or neuron morphology. Additional factors, including developmental factors such as spontaneous activity, interneuron subtype, synaptic dynamics and targeting preference significantly shape this connectivity structure (Peng et al. 2021 [↗](#); Udvary et al. 2022 [↗](#); Bollmann et al. 2023 [↗](#); Reimann et al. 2023 [↗](#); Wu et al. 2023 [↗](#); Kuan et al. 2024 [↗](#); Lakhera et al. 2024 [↗](#); Znamenskiy et al. 2024 [↗](#); Harth et al. 2025 [↗](#)). In the neonatal mouse cortex, even before the onset of external sensory inputs, inhibitory hub neurons regulate circuit development (Bollmann et al. 2023 [↗](#)). Moreover, theoretical studies suggest that spatial organization with a widespread Mexican-hat-like effective connectivity, consisting of local excitation and longer-range inhibition, can give rise to modular activity patterns with long-range spatial correlations propagating throughout the neonatal ferret cortex (Smith et al. 2018 [↗](#)). Such correlated activity is thought to serve as a scaffold for the emergence of specific stimulus feature tuning in the early stages of development (Mulholland et al. 2021 [↗](#); Lakhera et al. 2024 [↗](#); Trägenap et al. 2025 [↗](#)).

Despite these advances, the mechanisms and principles that enable the formation and maintenance of inhibitory connectivity patterns, especially in the context of excitatory connectivity, remain elusive. Genetic mechanisms play a fundamental role, particularly in early developmental stages (Duan et al. 2020 [↗](#); Hanganu-Opatz et al. 2021 [↗](#); Wu et al. 2023 [↗](#)). Equally important is synaptic plasticity driven by neural activity, which refines inhibitory connectivity

through a dynamic, activity-dependent process (Sreenivasan et al. 2022). This mechanism enables a closed-loop regulation of neural dynamics, allowing for precise tuning and adaptability in the face of change.

While the majority of experimental and theoretical studies investigate excitatory plasticity (e.g. Kempter et al. 1999; Song et al. 2000; Morrison et al. 2007; Shulz and Feldman 2013), far fewer examine inhibitory plasticity. Although some theoretical work has considered the impact of different forms of iSTDP in feedforward circuit architectures with a single postsynaptic neuron (Kleberg et al. 2014; Luz and Shamir 2014), there is still no general understanding of how pairwise iSTDP might influence the formation of specific connectivity motifs at the level of recurrently connected circuits. Theoretical work has proposed an iSTDP rule that guarantees the homeostatic regulation of neuronal rates (Vogels et al. 2011), becoming the “canonical” choice for plastic circuits. However, this rule tends to over-regulate rates in a homeostatic manner, producing narrow rate distributions in the excitatory population, which contradicts experimental findings (Buzsaki and Mizuseki 2014). In other work, inhibitory rules are specifically fit to experimental measurements (Lagzi et al. 2021; Bergoin et al. 2025), derived normatively from computation objectives (Mackwood et al. 2021), or are found through numerical optimization on a large parameter space (Confavreux et al. 2024). Here we ask how different spike timing-based rules at inhibitory synapses shape which excitatory and inhibitory neurons become strongly connected in recurrent circuits, and how the resulting excitatory-inhibitory motifs stabilize network firing rate, while supporting surround modulation and the generation of spatially correlated spontaneous activity.

To address these questions, rather than studying single rules, here we propose an integrated approach by analyzing a broad family of pairwise iSTDP rules and study their impact in small analytically tractable E/I circuits as well as in large-scale recurrent neuronal networks (RNNs) of spiking neurons. We identify mechanisms that stabilize excitatory activity while shaping circuit motifs, and show how, depending on the profile of the rule considered, inhibitory connectivity adjusts to the existing excitatory connectivity in distinct ways.

We first consider a recurrently connected two-neuron circuit of one excitatory and one inhibitory neuron, and derive how the choice of the iSTDP rule determines which features of neural activity drive the plasticity of inhibitory connections. At one extreme, rules dominated by firing rates are weakly influenced by existing connectivity structure in the two-neuron circuit, aiming primarily at the homeostatic regulation of firing rates, and thus promoting “blanket” inhibition. At the opposite extreme, we find covariance-dominated rules, strongly influenced by pairwise spike correlations that, in the absence of external inputs, arise from the underlying circuit connectivity. These rules stabilize neural activity, while promoting specific connectivity motifs, such as a mutually connected E/I pair or lateral inhibition, depending on the iSTDP features. We call the emergence of these specific connectivity motifs “structured stabilization”. This structured stabilization also generalizes to large-scale recurrently connected circuits where the same type of E/I connectivity motifs were observed under the different iSTDP rules. To further explore the functional consequences of covariance-dominated rules, we implement an RNN whose excitatory neurons follow a one-dimensional ring topology, and whose inhibitory neurons follow two distinct iSTDP rules. Here, inhibitory connectivity self-organizes into an effective Mexican-hat profile, generating long-range modular correlations of spontaneous activity. Finally, we demonstrate that the circuit implements contextual modulation, including suppression of responses to matching stimuli and amplification for mismatched stimuli. In sum, our study investigates how distinct inhibitory plasticity rules shape E/I connectivity motifs, from mutual to lateral inhibition, giving rise to network response properties such as surround modulation and long-range activity correlations observed in the adult and developing visual cortex.

2 Results

2.1 Inhibitory plasticity shapes specific E/I circuit motifs

We initially considered a general iSTDP rule that depends exclusively on local spike times applied to a synapse from an inhibitory presynaptic to an excitatory postsynaptic neuron (inh-to-exc) (Kempster et al. 1999 [\[1\]](#); Luz and Shamir 2014 [\[2\]](#); Akil et al. 2021 [\[3\]](#)). This rule contains two fixed components triggered by each pre- and post-synaptic firing events, denoted as α_{pre} and α_{post} , and a kernel function, $L(\tau)$ triggered by the co-occurrence of pre-post ($\tau > 0$) or post-pre ($\tau < 0$) spiking events (see Methods, Section 4.1 [\[4\]](#)). When the spiking activity is modeled as a stochastic process, the mean-field equation for the evolution of the average synaptic weight can be decomposed as follows (Kempster et al. 1999 [\[1\]](#)):

$$\frac{d\langle W \rangle}{dt} = \alpha_{\text{pre}} r_{\text{pre}} + \alpha_{\text{post}} r_{\text{post}} + B r_{\text{pre}} r_{\text{post}} + Q \quad (1)$$

with $B := \int_{-\infty}^{+\infty} d\tau L(\tau)$ and $Q := \int_{-\infty}^{+\infty} d\tau L(\tau) C(\tau)$

where r_{pre} and r_{post} are the pre- and post-synaptic firing rates, B is the net area of the iSTDP kernel $L(\tau)$, and Q captures the effect of the iSTDP kernel on the pre-post pairwise interactions, conveyed by the covariance density curve $C(\tau)$ (see Supporting Information, Section S1.1, or Kempster et al. 1999 [\[1\]](#); Morrison et al. 2008 [\[5\]](#); Luz and Shamir 2014 [\[2\]](#); Lagzi and Fairhall 2024 [\[6\]](#)). To understand how inhibitory plasticity influences circuit connectivity patterns, we first developed an analytical framework, in which rather than choosing a single rule, we defined a family of iSTDP rules. Using these rules, we aimed to derive the closed-form solution for a two-neuron recurrent E/I circuit, and link the shape of iSTDP to the pattern of learned inhibitory connections.

Several terms in Eq. 1 [\[1\]](#) depend solely on pre- and post-synaptic firing rates. Furthermore, in the general category of rules where B is high, the contribution of the $B r_{\text{pre}} r_{\text{post}}$ term can easily be an order of magnitude above Q , even in the presence of strong pre-post correlations (see Methods, Section 4.3 [\[4\]](#)). Therefore, when $Q \ll B r_{\text{pre}} r_{\text{post}}$, the plasticity dynamics of Eq. 1 [\[1\]](#) are strongly dominated by the firing rates. We call this plasticity regime “rate-dominated”. A prominent example of this is the previously established and well-studied rate-homeostatic iSTDP rule (Vogels et al. 2011 [\[7\]](#)). With $\alpha_{\text{post}} = 0$ and kernel integral $B > 0$, this rule drives the inhibitory weight to adapt continuously so that the postsynaptic rate converges to the target value r_{post^*} corresponding to a stable fixed point of Eq. 1 [\[1\]](#), namely: $r_{\text{post}} \approx -\alpha_{\text{pre}}/B$.

While a large value of B simplifies the analysis by making covariance-based terms negligible, it only represents a subset of possible plasticity rules. We therefore explored a broader range, including those for which the iSTDP kernel has zero or near-zero net area ($B \approx 0$). These rules amplify the contribution of the second-order correlations between pre- and postsynaptic neuron in the Q term. We refer to them as “covariance-dominated”. In choosing specific pairwise interaction kernels, we considered recent experimental findings that reported a symmetric iSTDP rule for parvalbumin-expressing (PV) interneurons, as well as an antisymmetric iSTDP rule for somatostatin-expressing (SST) interneurons, both projecting onto excitatory pyramidal neurons in the orbitofrontal cortex (Lagzi et al. 2021 [\[6\]](#)). Consequently, we focused on symmetric kernels with strong short-range potentiation and weak long-range depression (Fig. 1C [\[4\]](#)), or antisymmetric kernels in which pre-post and post-pre interactions balance each other with opposite signs (Fig. 1D [\[4\]](#)).

Analyzing the full dynamics of Eq. 1 [\[1\]](#) in a recurrently connected circuit is challenging because the covariance $C(\tau)$ generally depends on the synaptic weights themselves. We therefore considered a minimal two-neuron E/I recurrent circuit (Fig. 1A [\[4\]](#)) with fixed exc-to-inh connection (w_{exc}) but plastic inh-to-exc weight (w_{inh}). We modeled the neurons as time-varying Poisson processes, where each incoming spike changes the instantaneous spiking probability through an

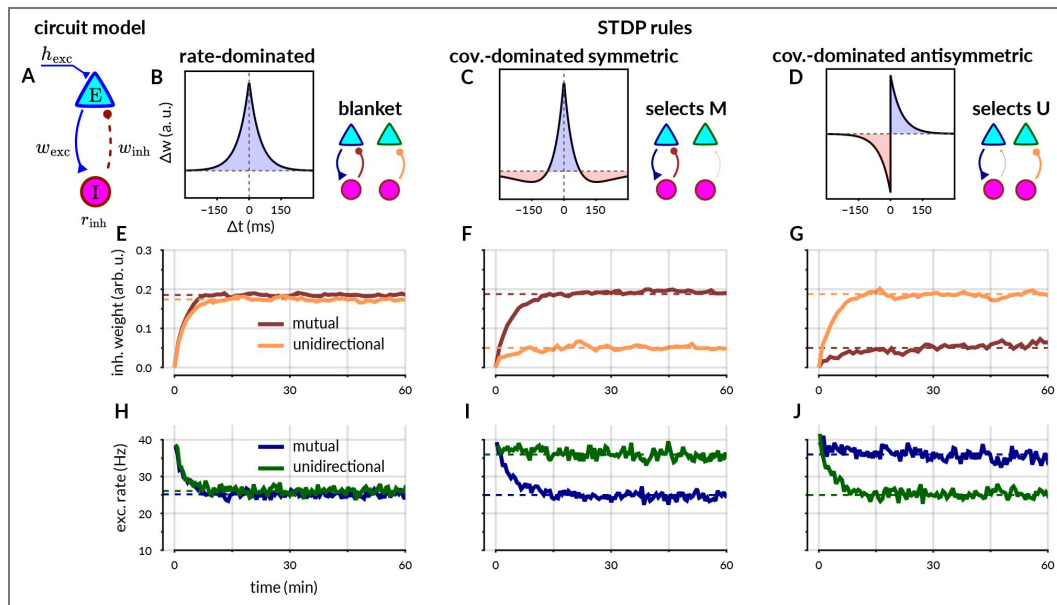


Figure 1. Effects of plasticity on E/I connectivity motifs in a two-neuron circuit.

A. Schematics of the two-neuron circuit model. Two time-varying coupled Poisson units, one excitatory and one inhibitory, are connected by a fixed excitatory weight w_{exc} and a plastic inhibitory weight w_{inh} . The instantaneous rates are computed as the convolution between each incoming spike and an exponential synaptic kernel that either increases (exc) or decreases (inh) the instantaneous firing probability. The input to the excitatory neuron h_{exc} is fixed, whereas the inhibitory input changes dynamically, keeping the inhibitory neuron at a fixed rate r_{inh} . Only the inhibitory synapse (dashed line) is plastic. The circuit is simulated in two configurations: unidirectional (U), with $w_{exc} = 0$ and mutual (M), with $w_{exc} = 0.5$. **B, C, D.** Pairwise interaction kernels of the three iSTDP rules tested (Eq. 11). We used the following numerical parameters. **B.** $B = 1$, $a_{pre} = -25$, $a_{post} = 0$, $\tau_+ = 50$ ms. **C.** $B = 0$, $a_{pre} = -0.12$, $a_{post} = 1.45$, $\tau_+ = 50$ ms, $\tau_- = 500$ ms. **D.** $B = 0$, $a_{pre} = -0.14$, $a_{post} = 7.95$, $\tau_+ = 30$ ms, $\tau_- = 30$ ms. **E, F, G.** Change in inhibitory synaptic weight over time for the M (brown) and U (light orange) circuit configurations, under the rules above each column. Dashed lines are analytic predictions for the fixed point (Methods, Eq. 11). **H, I, J.** Change in excitatory firing rate over time for the M (dark blue) and U (dark green) configurations, under the rules above. Dashed horizontal lines are the analytic predictions for the fixed-point (Methods, Eq. 11).

excitatory or inhibitory post-synaptic potential (EPSP/IPSP) kernel (internal traces shown in Fig. S1 [↗](#), see also Methods Section 4.4 [↗](#)). The inhibitory neuron's firing rate is kept constant by an external homeostatic mechanism. The excitatory neuron receives an excitatory fixed external input in addition to the inhibitory input. Using this simplified model, we derived analytically a closed-form expression for the fixed-point of the iSTDP dynamics (Methods, Eq. 11 [↗](#); Supporting Information). The condition of fixed r_{inh} is not required for a stable solution. Allowing r_{inh} to vary still leads to an analytic expression for the fixed point of the iSTDP dynamics, but the resulting closed-form involves the root of a quadratic equation and is therefore less intuitive (see Supporting Information, S1.3, Fig. S3 [↗](#)).

With this analytical framework in hand, we investigated how different iSTDP rules affect the resulting inhibitory synaptic strength and the excitatory postsynaptic activity. We considered two configurations of our E/I circuit: (1) a unidirectional (U) configuration, where the excitatory-to-inhibitory weight w_{exc} is set to zero, and (2) a mutual (M) configuration, featuring a strong excitatory-to-inhibitory connection. We studied the three distinct iSTDP rules in both configurations: a rate-dominated rule with a symmetric interaction kernel (Fig. 1B [↗](#)), a covariance-dominated rule also with a symmetric kernel (Fig. 1C [↗](#)), and a covariance-dominated rule with an antisymmetric kernel (Fig. 1D [↗](#)). As expected, the rate-dominated rule shows minimal sensitivity to the presence of the mutual connection. In both U and M configurations, the inhibitory synapse adjusts the excitatory firing rate toward the target rate dictated by the rate-dependent components, approximately given by $\tilde{r}_{\text{exc}} \approx -\alpha_{\text{pre}}/B$ (Fig. 1E [↗](#)). Minor deviations arise from the non-zero contribution of the Q term.

In contrast, the covariance-dominated rules are sensitive to both the circuit's connectivity structure and the shape of iSTDP kernel. With the symmetric kernel, the final inhibitory weight becomes significantly stronger in the M configuration compared to the U configuration (Fig. 1F [↗](#), compare brown and light orange). Conversely, with the antisymmetric kernel, the inhibitory weight is lower in the M configuration (Fig. 1G [↗](#)). Such specificity enables these rules not only to regulate excitatory rates, but also to shape specific E/I connectivity motifs—an effect we termed “structured stabilization”. This property emerges from the well-known interaction between the iSTDP kernel and second-order spiking statistics. While this has been previously studied for excitatory plasticity (Kempster et al. 1999 [↗](#); Song et al. 2000 [↗](#); Babadi and Abbott 2013 [↗](#); Shulz and Feldman 2013 [↗](#)), here we extend it to inhibitory plasticity in a minimally recurrent two-neuron E/I circuit and hence consider how recurrent inhibition regulates excitation. Specifically, the strong excitatory-to-inhibitory connection in the M configuration generates positively-correlated post-pre events. For the symmetric iSTDP, these events occur within the positive region of the kernel, leading to synaptic potentiation. Conversely, with the antisymmetric iSTDP rule, the same positively-correlated events interact with the negative region of the kernel, resulting in synaptic depression. Additionally, the inhibitory-to-excitatory weight generates a negative covariance density, localized instead in the pre-post region of the curve. As both rules are positive here, this introduces a negative feedback in both cases, limiting the value that w_{inh} can reach before plasticity switches into depression. Finally, adjusting the fixed components of the iSTDP rule, α_{pre} and α_{post} , is also crucial for reaching the desired excitatory firing rate (set to 25Hz) in the preferred circuit configuration.

Hence, our theoretical framework provides a link between different iSTDP rules and inhibitory connectivity in a reduced two-neuron E/I recurrent circuit motif.

2.2 The shape of inhibitory STDP affects inhibitory connectivity

Having established that different covariance-dominated rules can achieve structured stabilization, namely stabilize excitatory firing rates and shape specific E/I connectivity motifs, we next investigated how the specific form of the iSTDP kernel shapes the learned inhibitory connectivity. We considered four kernel shapes: symmetric, antisymmetric, and their “inverted” counterparts, obtained by reversing the sign of the kernel (Fig. 2A [↗](#)). We analyzed the behavior of these kernels

in the reduced two-neuron E/I circuit. To explore the parameter space, we varied the rate-dependent parameters B and α_{pre} , while keeping $\alpha_{\text{post}} = 0$. With this parametrization, the covariance-dominated regime corresponds to values of B and α_{pre} close to zero.

Our analysis found that both symmetric and antisymmetric kernels yield stable fixed-points in the covariance-dominated regime (Fig. 2A,B, top row, and Supporting Information Fig. S2C,D). These kernels, which we term “self-stabilizing” exhibit robustness to parameter fluctuations and reliably converge to a stable inhibitory weight. In contrast, the inverted antisymmetric and the inverted symmetric rule are unstable in the covariance-dominated regime, with inhibitory weights either decaying to zero or saturating depending on the initial conditions. We categorize these as “saturating” kernels (Fig. 2A, bottom panels, and Fig. S2E,F in Supporting Information). Saturating kernels can still reach stability with strong rate-dependent terms, which reduce the overall influence of the kernel shape on the final weight.

This difference in stability arises from how the kernels interact with changes in pre- and postsynaptic correlations. Consider a synapse operating under a covariance-dominated plasticity rule at its fixed point. At equilibrium, potentiation and depression events balance exactly. A perturbation that strengthens the inhibitory weight leads to a reduction in pre-post firing events due to the increased inhibition. For self-stabilizing kernels, which have a positive right branch (corresponding to pre-post events), this reduction in correlation reduces the occurrence of synaptic potentiation events, shifting the balance in favor of synaptic depression. The synapse therefore depresses, eventually compensating the initial perturbation from equilibrium. Conversely, saturating kernels show the opposite effect: since their kernel is negative in the pre-post region, a reduction of pre-post correlation results in fewer synaptic depression events, so the initial balance is further shifted in favor of synaptic potentiation, amplifying the perturbation and leading the synapse towards saturation. Finally, in rate-dominated rules, the specific shape of iSTDP kernel becomes less relevant. Weight fluctuations impact the dynamics through the $B r_{\text{pre}} r_{\text{post}}$ term rather than the correlation-dependent Q term. Consequently, these rules consistently converge to the classic rate-homeostatic solution $\tilde{r}_{\text{exc}} \approx -\alpha_{\text{pre}}/B$ (Vogels et al. 2011). In large networks, however, once these rules reach their homeostatic rate target, they can still promote additional structural organization at slower timescales (Lagzi and Fairhall 2024).

To further characterize the self-stabilizing rules, we extended our analysis beyond the U and M circuit configurations presented in Fig. 1, exploring a broader range of excitatory-to-inhibitory weights w_{exc} and input currents to the excitatory neuron. For the symmetric kernel, the learned inhibitory weight increases with w_{exc} (Fig. 2C), while for the antisymmetric kernel it decreases (Fig. 2D). Increasing the excitatory external input leads to stronger inhibitory weights in both cases (Fig. 2C,D), however the rate of the excitatory neuron still scales linearly with the input strength (Fig. 2F,G). In contrast, the rate-homeostatic rule exhibits a minimal dependence on w_{exc} (Fig. 2E) and a very tight control on the excitatory rate (Fig. 2H), which restricts the dynamic range of the excitatory neuron, limiting its ability to respond to changes in input. The E/I interactions in the two-neuron circuit allow for the weights to converge without saturation, differently from a two-neuron circuit with only excitatory neurons (Babadi and Abbott 2013).

In summary, our analysis of the reduced E/I circuit reveals a spectrum of iSTDP rules with distinct outcomes on circuit connectivity and dynamics. On the one hand, covariance-dominated rules with self-stabilizing kernels enable structured stabilization (stable activity and specific connectivity motifs). As a consequence, they selectively reinforce connectivity motifs based on pre- and post-synaptic correlations, and allow a broader range of excitatory firing rates. On the other hand, rate-dominated rules prioritize homeostatic control of excitatory activity over the formation of specific connectivity motifs. Thus, these rules ensure robust dynamical stability, potentially at the cost of computational flexibility. These results demonstrate a trade-off between the emergence of specific excitatory-inhibitory motifs and the tight regulation of excitatory rates in the class of iSTDP rules we considered.

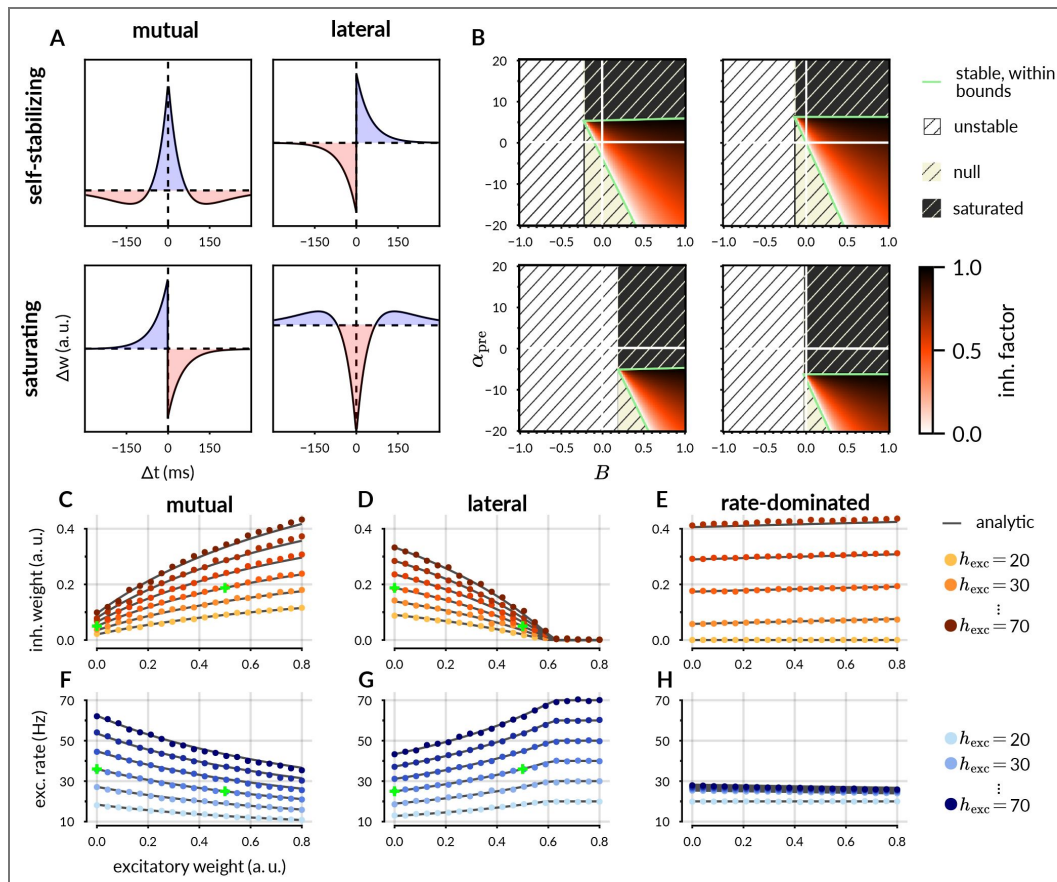


Figure 2. Parametric study of different shapes of iSTD kernel in the reduced 2D model.

A. Representation of four possible shapes of iSTD kernel, labeled by their effects in the covariance-dominated regime (i.e., with near-zero rate-dependent terms). **B.** Parametric study based on the analytic solution for each of the four iSTD kernels, varying α_{post} and the net area under the iSTD kernel, B (see Eq. 1). The color map represents the level of inhibition, where 1 indicates that the excitatory neuron is completely silenced, and 0 indicates no inhibition, so that the excitatory firing rate matches the input current **C,D,E.** Final inhibitory weight as a function of the incoming excitatory weight for three different iSTD rules. Dots are numerical simulations, gray lines are analytic results (Methods, Eq. 11), color shades indicate different input levels. **C.** Results for the symmetric, self-stabilizing rule. The green dots represent the same conditions shown in Fig. 1F. **D.** Results for the antisymmetric, self-stabilizing rule. The green dots represent the same conditions shown in Fig. 1G. **E.** Results for the symmetric rule in a rate-dominated regime. **F,G,H.** Corresponding plots showing the rate of the excitatory neuron after convergence.

2.3 Structured stabilization in randomly connected RNNs

How do our results on structured stabilization from two-neuron E/I circuits translate to large recurrent networks? We simulated a randomly connected RNN of conductance-based leaky integrate-and-fire (LIF) neurons, with 900 excitatory and 100 inhibitory units. We modeled the excitatory connections (exc-to-exc and exc-to-inh) as fixed and sparse, while the inh-to-exc connections were modeled as dense and initially small, but subject to iSTDP (Table 2). In this case the firing rate of each inhibitory neuron was not fixed as in the two-neuron E/I circuit, but was determined by the total amount of input into each neuron from its presynaptic partners, without any external constraint.

We first implemented a covariance-dominated plasticity rule with a symmetric kernel (Fig. 3A). Initially, the network produces a pathological high-firing regime, due to the low inhibitory weights. However, inhibitory plasticity quickly stabilizes the dynamics, generating an asynchronous irregular firing regime characteristic of balanced networks (Fig. 3B). To analyze the resulting inhibitory connectivity, we categorized inh-to-exc synaptic weights as either “mutual” or “unidirectional”. We called an inh-to-exc weight mutual when a reciprocal exc-to-inh connection exists between the same two neurons, reflecting a bidirectional link. Conversely, a weight is unidirectional in the absence of such excitatory reciprocal connection. Following the onset of plasticity, we observed that mutual inhibitory weights increase much more than unidirectional ones, up to the saturation threshold (Fig. 3C). This selectivity can be directly visualized through the matrix of synaptic weights after learning: the pattern of learned inh-to-exc weights mirrors the fixed exc-to-inh weights (Fig. 3E,F; see Supporting Information for full weight matrix).

Next, we investigated the impact of an asymmetric iSTDP rule within the same random RNN architecture and neural dynamics. As with the symmetric rule, this asymmetric iSTDP stabilizes the network dynamics (Fig. 3G). However, it promotes a markedly different connectivity pattern: purely lateral inhibition. In this configuration, inhibitory neurons strengthen their connections onto excitatory neurons lacking reciprocal connections back to them, thus forming exclusively unidirectional connections (Fig. 3H). This yields an inhibitory-to-excitatory connectivity pattern almost perfectly opposite to the previous configuration, so that existing exc-to-inh synapses carve gaps in the inh-to-exc connections (Fig. 3J,K). Achieving robust lateral inhibition with the asymmetric iSTDP rule requires the introduction of a small rate-dependent component ($B = 0.3$). This deviation from a purely covariance-dominated regime introduces characteristics of rate-homeostatic plasticity, such as rapid convergence to a narrow distribution of excitatory firing rates (Fig. 3G). This requirement is perhaps unsurprising given the challenges of parameter tuning in large RNNs, where closed-form solutions are unavailable (see Discussion for further considerations). Both the symmetric and the asymmetric rules converged to stable weight distributions, except for a very slow tendency for some mutual weights to reach either the upper bound (symmetric) or zero (asymmetric). These results are also robust to large variations in external input, and are compatible with network models where asynchronous irregular firing is generated intrinsically, without any external source of random noise (SI, Section S2 and Fig. S5; see also Mastrogiuseppe and Ostojic 2017).

In sum, our RNN simulations generalize the analytical predictions from the reduced E/I circuit model to larger networks of recurrently connected integrate-and-fire units. Symmetric iSTDP rules reinforce mutual inhibition, mirroring pre-existing excitatory connectivity, whereas asymmetric kernels favor lateral inhibition.

2.4 Response properties of a one-dimensional ring network

Having shown that different forms of iSTDP shape E/I connectivity in large-scale recurrent networks based on local correlations arising from direct connections, we next explored the functional consequences of structured stabilization in a recurrent network with a one-dimensional ring architecture. Ring networks are well-suited to model continuous attractor dynamics, supporting bumps of persistent activity (Amari 1977; Ben-Yishai et al. 1995; Khona

symbol	value	description
N_{exc}	1	number of excitatory neurons
N_{inh}	1	number of inhibitory neurons
τ_e	30 ms	EPSP kernel time constant
τ_i	30 ms	IPSP kernel time constant
h_e	40 Hz	excitatory input current
r_{inh}	80 Hz	inhibitory firing rate
<i>rate-dominated iSTDP rule</i>		
A_0	$5 \cdot 10^{-7}$	global learning rate
B	1	net area under the iSTDP kernel
α_{pre}	-25	constant dependent on presynaptic rate
α_{post}	0	constant dependent on postsynaptic rate
τ_+	50 ms	iSTDP kernel time constant
<i>covariance-dominated symmetric iSTDP rule</i>		
A_0	$3 \cdot 10^{-6}$	global learning rate
B	0	net area under the iSTDP kernel
α_{pre}	-25	constant dependent on presynaptic rate
α_{post}	0	constant dependent on postsynaptic rate
τ_+	50 ms	iSTDP kernel time constant for potentiation
τ_-	500 ms	iSTDP kernel time constant for depression
<i>covariance-dominated antisymmetric iSTDP rule</i>		
A_0	$3 \cdot 10^{-6}$	global learning rate
B	0	net area under the iSTDP kernel
α_{pre}	-25	constant dependent on presynaptic rate
α_{post}	0	constant dependent on postsynaptic rate
τ_+	50 ms	iSTDP kernel time constant for potentiation
τ_-	50 ms	iSTDP kernel time constant for depression

Table 1. Numerical parameters used for Fig. 1. The parameters in Fig. 2 match those of Fig. 1 unless otherwise indicated in the figure caption.

symbol	value	description
N_{exc}	900	number of excitatory neurons
N_{inh}	100	number of inhibitory neurons
C_{membrane}	200 pF	membrane capacitance
τ_{AMPA}	5 ms	glutamatergic synaptic time constant
τ_{GABA}	10 ms	GABAergic synaptic time constant
g_{leak}	10	leak conductivity
V_{L}	−60 mV	membrane leak reverse potential
V_{AMPA}	0 mV	AMPA reverse potential
V_{GABA}	−80 mV	GABA reverse potential
$V_{\text{threshold}}$	−50 mV	firing threshold
V_{reset}	−60 mV	reset membrane potential after firing
$p_{e \leftarrow e}$	0.2	exc-to-exc connection probability
$p_{i \leftarrow e}$	0.1	exc-to-inh connection probability
$p_{e \leftarrow i}$	1.0	inh-to-exc connection probability
$p_{i \leftarrow i}$	1.0	inh-to-inh connection probability
$w_{e \leftarrow e}$	1.0	exc-to-exc synaptic weight
$w_{i \leftarrow e}$	1.0	exc-to-inh synaptic weight
$w_{i \leftarrow i}$	0.3	inh-to-inh synaptic weight
<i>External input</i>		
N_{shared}	50	number of shared input neurons for exc and inh
$N_{\text{indep.,exc}}$	250	number of independent input neurons for exc
$N_{\text{indep.,inh}}$	0	number of independent input neurons for inh
r_{input}	50 Hz	rate of input neurons
$w_{\text{input,exc}}$	1	weight of input synapse to exc
$w_{\text{input,inh}}$	0.1	weight of (independent) input synapse to inh
<i>covariance-dominated symmetric iSTDP rule (panels A-F)</i>		
A_0	$2 \cdot 10^{-2}$	global learning rate
B	0	net integral of the STDP kernel
α_{pre}	−0.1	constant dependent on presynaptic rate
α_{post}	0	constant dependent on postsynaptic rate
τ_{+}	30 ms	iSTDP kernel time constant
τ_{-}	600 ms	iSTDP kernel time constant
<i>covariance-dominated asymmetric iSTDP rule (panels G-K)</i>		
A_0	$5 \cdot 10^{-3}$	global learning rate
B	0.3	net integral of the iSTDP kernel
α_{pre}	−0.7	constant dependent on presynaptic rate
α_{post}	0.2	constant dependent on postsynaptic rate
τ_{+}	60 ms	iSTDP kernel time constant for potentiation
τ_{-}	30 ms	iSTDP kernel time constant for depression

Table 2. Numerical parameters for the random RNN in Fig. 3 [↗](#).

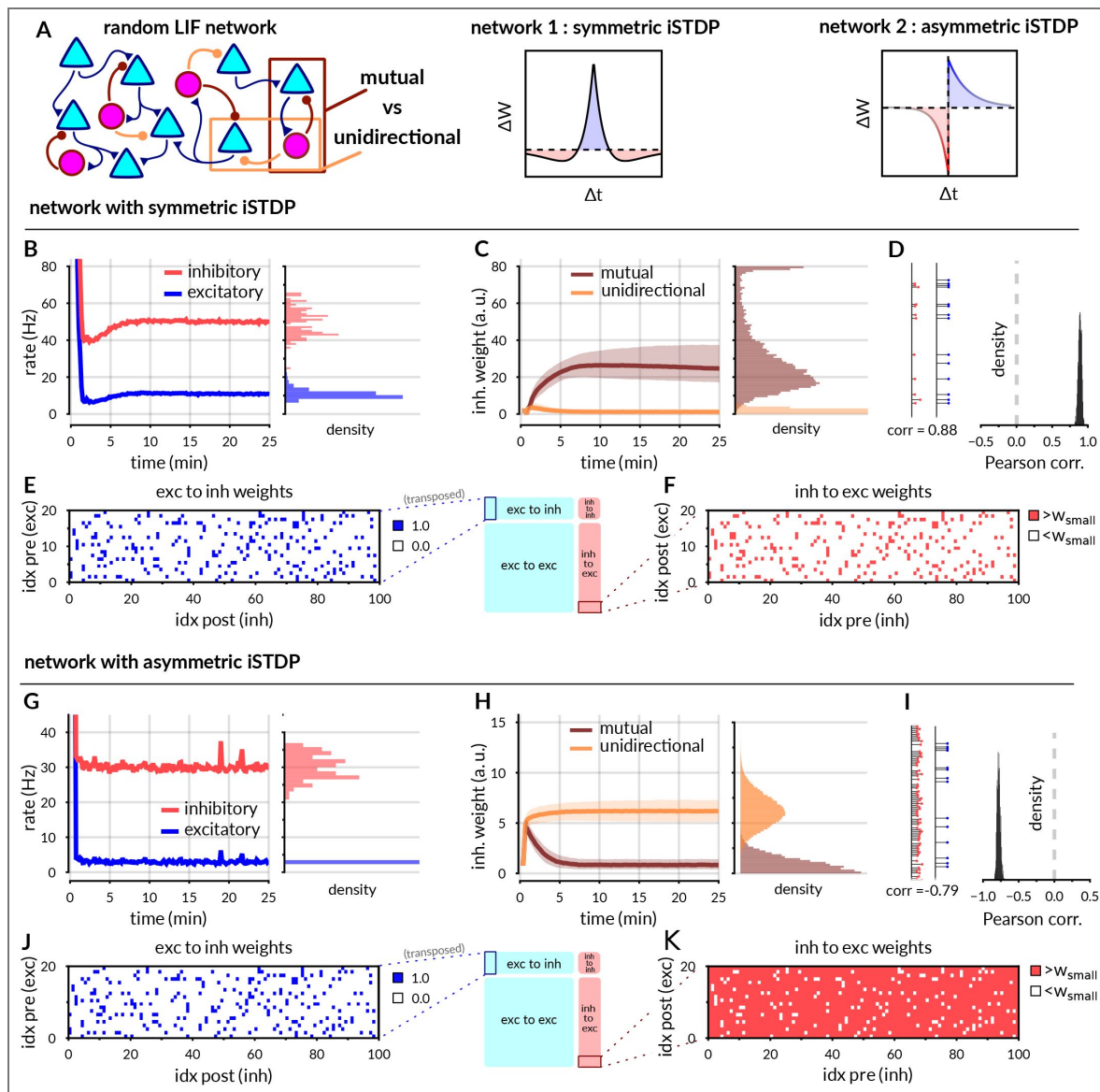


Figure 3. RNNs of conductance-based LIF neurons self-organize into specific E/I motifs under different iSTDP rules.

A left. Representation of the random network with ($N_{\text{exc}} = 900$, $N_{\text{inh}} = 100$) spiking neurons. The inh-to-exc weights form all-to-all connections and are plastic, all other weights are fixed. Exc-to-exc and exc-to-inh connections are sparse. **A middle.** Kernel of the symmetric, covariance-dominated iSTDP rule used in network 1 (panels B-F). **A right.** Kernel of the asymmetric, covariance-dominated iSTDP rule used in network 2 (panels G-K). Parameters in Table 2. **B.** Evolution of population mean firing rates over time, with final distribution of rates over the full populations. **C.** Evolution of inh-to-exc synaptic weights over time and final distribution. The weights are split into a “mutual” (brown) and a “unidirectional” (light orange) group. Only few weights reach the saturating value, set to 80. **D.** Example of incoming (blue) and outgoing (red) weights from a single inhibitory neuron, with correlation between them, and distribution of correlations for the full inhibitory populations across 10 random network realizations. **E.** Random portion of the fixed and sparse exc-to-inh weights (the full matrix, of size 900×100 is shown in SI, Fig. S4). **F.** Portion of the all-to-all and plastic inh-to-exc connections after learning, with indices matching panel E. For clarity the matrix has been binarized, with threshold-value w_{small} defined as 0.1 times the 0.9-quantile of the full weight distribution. See SI, Fig. S4, for a continuous representation of synaptic weights. **G-K.** Equivalent plots in a second network with identical network dynamics and initial conditions, but using the asymmetric iSTDP (see also SI, Fig. S6).

and Fiete 2022 [\[10\]](#)). This makes them relevant for studying neural computations such as head-direction encoding (Xie et al. 2002 [\[11\]](#); Kim et al. 2017 [\[12\]](#); Kim et al. 2019 [\[13\]](#); Petrucco et al. 2023 [\[14\]](#)) and orientation selectivity in the primary visual cortex (Goldberg et al. 2004 [\[15\]](#)).

We built a ring network of 800 excitatory and 200 inhibitory neurons arranged on a circle, with exc-to-exc and exc-to-inh weights decaying exponentially with the angular distance between neurons, following a rescaled Von Mises (Fig. 4A,B [\[16\]](#)). We further divided the inhibitory population into two subclasses, governed by distinct iSTDP rules, but sharing identical neural dynamics. Inspired by experimental results (Lagzi et al. 2021 [\[17\]](#)), one population of “PV” interneurons followed a symmetric kernel, and another population of “SST” interneurons followed an antisymmetric one, both rules operating in the covariance-dominated regime. Initially, we set both PV-to-exc and SST-to-exc synapses as all-to-all, weak, and unbiased (see Fig. 4A [\[16\]](#) for schematics, and Methods Table 3 [\[18\]](#) for numerical parameters).

As the network evolves under the influence of the two plasticity rules, the inhibitory populations develop specialized roles. PV interneurons, governed by symmetric iSTDP, form strong mutual connections with excitatory neurons, mirroring the distance-dependent excitatory connectivity (Fig. 4B [\[16\]](#) and Fig. 4C [\[16\]](#), red). Conversely, SST interneurons, driven by antisymmetric iSTDP, establish lateral inhibition across the ring (Fig. 4B [\[16\]](#) and Fig. 4C [\[16\]](#), orange). This structured stabilization, with both localized and broader components, raises the possibility of nontrivial effective interactions between excitatory neurons, potentially mediated by the different roles of the PV and SST interneurons.

To estimate the effective interaction profile between excitatory neurons, we used linear response theory, equivalent to a Taylor expansion of the network dynamics around its operating point (Sadeh and Clopath 2020 [\[19\]](#), see Methods for details). We found an effective connectivity profile resembling a Mexican-hat, with short-range excitation directly inherited from exc-to-exc connectivity and a broader inhibitory surround resulting from the combined action of PV and SST interneurons (Fig. 4D [\[16\]](#)). This effective connectivity profile motivated us to explore the resulting network dynamics and its response to external stimuli. We examined the emergent spontaneous activity and observed a complex spatial correlation structure, with periodic oscillations in correlations over distance (Fig. 4E [\[16\]](#), see Supporting Information Fig. S8 [\[20\]](#) for full correlation matrix). This emergent dynamics is reminiscent of the spatial correlation maps observed in the developing cortex of ferrets, and is considered a signature of modular cortical organization (Smith et al. 2018 [\[21\]](#); Powell et al. 2024 [\[22\]](#)).

After identifying aspects of emergent cortical dynamics in our network consistent with early developmental patterns, we investigated whether it also displays other fundamental computational properties classically measured in the visual system. A well-established nonlinear phenomenon in the visual cortex is surround suppression, in which gradually increasing the stimulus size initially amplifies the neuron’s response up to a peak, but then leads to reduced activation (DeAngelis et al. 1994 [\[23\]](#); Cavanaugh et al. 2002 [\[24\]](#)). After inhibitory plasticity self-organized network connectivity, we drove the network with patterned stimuli in the absence of subsequent plasticity. We delivered inputs of different strengths and sizes in the form of excitatory spike trains (Fig. 5A [\[25\]](#)). Each stimulus was presented 100 times in randomized order, with a 0.5 s presentation followed by a 0.5 s blank. We found that larger stimuli elicited lower peak firing rates in their center region, while the response of neurons with receptive field closer to the stimulus’ edge was larger (Fig. 5A [\[25\]](#), raster plot in SI, Fig. S8 [\[20\]](#)). Consequently, the center neuron’s response as a function of stimulus size exhibited a classic surround suppression profile (Fig. 5B [\[25\]](#), DeAngelis et al. 1994 [\[23\]](#)).

We carried out three control experiments to disentangle the contributions of the two interneuron populations with different plasticity rules to the dynamics of the ring network. A single inhibitory population with rate-dominated iSTDP can stabilize the system, although the emergent inh-to-exc connectivity does not seem patterned (Fig. S9B [\[26\]](#)). This network can still produce surround suppression, however the edge effect is reversed, with neurons on the stimulus edge firing at a lower rate compared to neurons in the center (Fig. S9C [\[26\]](#), note the convex shape of the curve for large stimuli). Moreover, the correlations do not exhibit any spatial structure. When considering

symbol	value	description
N_{exc}	800	number of excitatory neurons
N_{PV}	100	number of PV interneurons
N_{SST}	100	number of SST interneurons
sum $w_{\bullet \leftarrow e}$	145	total incoming exc weights to exc, PV, SST
$k_{e \leftarrow e}$	10	width of exc-to-exc connections
$k_{\text{PV} \leftarrow e}$	5	width of exc-to-pv connections
$k_{\text{SST} \leftarrow e}$	10	width of exc-to-sst connections
$p_{\text{PV,SST} \leftarrow \text{PV,SST}}$	0.9	connection probability for PV-PV, SST-SST and PV-SST
$w_{\text{PV} \leftarrow \text{PV}}$	2	PV-to-PV synaptic weight
$w_{\text{SST} \leftarrow \text{PV}}$	0.5	PV-to-SST synaptic weight
$w_{\text{SST} \leftarrow \text{SST}}$	2	SST-to-SST synaptic weight
$w_{\text{PV} \leftarrow \text{SST}}$	0.5	SST-to-PV synaptic weight
<i>External input</i>		
$h_{\text{in,exc}}$	1000 pA	constant input current to exc population
N_{input}	50	number of independent input neurons
w_{input}	1.0	weight of input synapse
$r_{\text{input,exc}}$	80 Hz	rate of input neurons to exc
$r_{\text{input,inh}}$	40 Hz	rate of input neurons to PV,SST
<i>symmetric iSTDP rule for the PV population</i>		
A_0	$2 \cdot 10^{-3}$	global learning rate
B	0	net integral of the iSTDP kernel
α_{pre}	-0.2	constant dependent on presynaptic rate
α_{post}	1	constant dependent on postsynaptic rate
τ_+	30 ms	STDP kernel time constant for potentiation
τ_-	600 ms	STDP kernel time constant for depression
<i>antisymmetric iSTDP rule for the SST population</i>		
A_0	$1 \cdot 10^{-3}$	global learning rate
B	0	net integral of the iSTDP kernel
α_{pre}	-0.2	constant dependent on presynaptic rate
α_{post}	1	constant dependent on postsynaptic rate
τ_+	30 ms	iSTDP kernel time constant for potentiation
τ_-	30 ms	iSTDP kernel time constant for depression

Table 3. Numerical parameters for the one-dimensional ring model. The parameters that regulate neural dynamics are the same as in Table 2.

Figure 4. Inhibitory self-organization in a spiking RNN topologically organized as a one-dimensional ring.

A. Representation of the model. 800 excitatory neurons are arranged in a one-dimensional ring topology, connected to each other and to their inhibitory counterparts with weights that depend on their distance in tuning space (Methods, Eq. 15). 100 “PV” inhibitory neurons connect all-to-all to the excitatory neurons with plasticity following a symmetric iSTDP rule, 100 “SST” neurons connect all-to-all to excitatory neurons with plasticity following an antisymmetric iSTDP. For both rules $B = 0$, $\alpha_{pre} = -0.1$, $\alpha_{post} = 0$. **B.** Full weight matrix after weight convergence. The excitatory weights (blue) are fixed, the inh-to-exc weights (red, lower right) are determined by plasticity. See SI, Fig. S4 for full weight matrix. **C.** Profile of the average outgoing weights as a function of tuning distance. For convenience, the profiles are normalized so that the maximum is either 1 or -1. **D.** Average effective interaction profile between the excitatory neurons, as a function of tuning (see Methods). **E.** Evolution of mean firing rate over time. **F.** Pearson correlation during spontaneous activity between an excitatory neuron and its neighbors, ordered by tuning distance and averaged over the entire population.

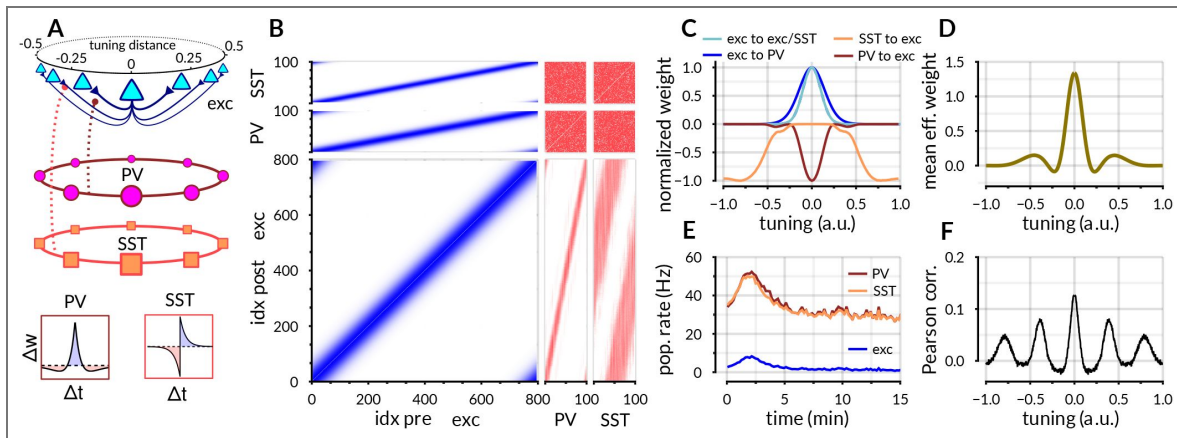
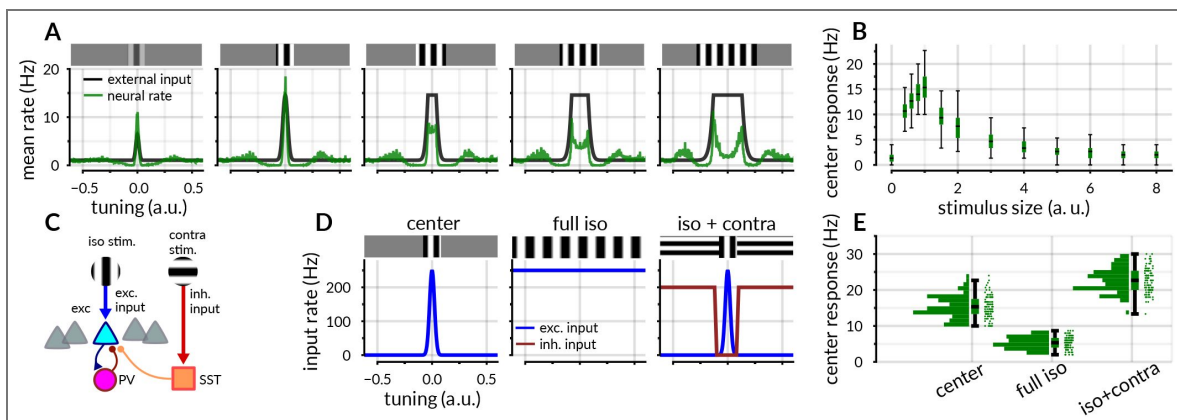


Figure 5. Response properties of the one-dimensional ring network.

A. Representation of the stimuli (black) and of the network’s response (green), as a function of the neuron’s location relative to the stimulus center, averaged over 100 trials. **B.** Response to the center neuron only as a function of stimulus size. **C.** Schematics of the iso-contra stimulation protocol. The *iso* direction corresponds to excitatory inputs to excitatory neurons with matching receptive field. The *contra* direction generates instead an inhibitory input on the SST population. **D.** Profiles of the center, full iso and contra stimuli to the network. **E.** Mean response of the center neuron, averaged over 100 trials, for the three stimuli.



networks with a single covariance-dominated PV-like (symmetric iSTDP kernel) or SST-like (asymmetric iSTDP kernel) plasticity rule, we found only pathological solutions that required careful parameter tuning. The PV-only network displayed pathological population-wide synchrony (Fig. S10E [↗](#)), while the SST-only network produced strongly clustered runaway excitation (Fig. S10J [↗](#)). This suggests that the two interneuron populations with different plasticity rules act synergistically to establish excitatory firing rate stability and specific E/I motifs.

We also investigated whether our model could also reproduce contextual facilitating effects, for example when a stimulus in the preferred orientation is surrounded by an orthogonal stimulus (Sillito and Jones 1996 [↗](#)). These effects are mediated by a third subclass of interneurons, the vasoactive intestinal peptide (VIP)-expressing cells, that inhibit the SST population and thereby disinhibit excitatory neurons (Keller et al. 2020 [↗](#)). We thus modeled the iso-oriented stimulus as excitatory spike trains delivered to the excitatory neurons, and introduced a contra-oriented stimulus as inhibitory inputs from VIP interneurons to the SST population at the corresponding location (Fig. 5C [↗](#)). We then compared the center-neuron response in three conditions: (1) an iso-oriented stimulus restricted to the neuron's receptive field; (2) a broader iso-oriented stimulus; (3) an iso stimulus in the receptive field combined with a contra stimulus in the distal regions (Fig. 5D [↗](#)). In agreement with experimental results (Keller et al. 2020 [↗](#)), the distal contra stimulus disinhibits the center neuron and facilitates its response in the iso+contra stimulus conditions (Fig. 5E [↗](#)).

These results demonstrate how structured stabilization shaped by the interplay of symmetric and antisymmetric iSTDP rules can give rise to network connectivity which generates activity consistent with both suppressive and facilitating contextual effects in the primary visual cortex, and to long-range correlation patterns consistent with spontaneous activity in the developing sensory cortex.

3 Discussion

In this work, we have shown that a family of inhibitory STDP rules can simultaneously stabilize excitation at the circuit level, and drive the formation of specific connectivity patterns among excitatory and inhibitory neurons in the form of E/I motifs. In the absence of correlated external inputs, or training signals, circuits self-organize purely through their spontaneous internal dynamics. We found that the shape of the iSTDP rule selectively reinforces different E/I motifs—blanket inhibition, reciprocal E/I connectivity and, lateral E/I connectivity—based on the correlations present in the network.

Prior computational studies addressing the self-organized formation of motifs in recurrent networks through STDP mostly focused on exc-to-exc plasticity (Ocker et al. 2015 [↗](#); Ravid Tannenbaum and Burak 2016 [↗](#); Triplett et al. 2018 [↗](#); Montangie et al. 2020 [↗](#); Manz and Memmesheimer 2023 [↗](#)). Previous work on iSTDP has been largely restricted to feedforward circuits and rules dominated by either firing rates or covariance (Kleberg et al. 2014 [↗](#); Luz and Shamir 2014 [↗](#); Akil et al. 2021 [↗](#); Vignoud and Robert 2022 [↗](#)). In contrast, our approach combines both: recurrence and iSTDP. This combination makes it possible to characterize how different timing-based rules at inhibitory synapses shape excitatory-inhibitory connectivity in recurrent circuits, and how the resulting E/I motifs influence circuit dynamics, including rate stabilization, surround modulation, and the structure of spontaneous activity.

We implemented a generalized iSTDP framework where individual contributions from spike pairs additively contribute to synaptic change, as in many models of excitatory synaptic plasticity. While many other factors may also contribute to inhibitory plasticity, including multiplicative effects (Morrison et al. 2007 [↗](#); Kunkel et al. 2011 [↗](#)), synaptic location (Agnes and Vogels 2024 [↗](#)) or plasticity on behavioral timescales (Gonzalez et al. 2023 [↗](#)), we found that additive iSTDP can still generate many interesting insights about E/I connectivity in recurrent networks. Our iSTDP framework combines both rate-dependent and covariance-dependent components in inducing inhibitory plasticity. The rate-dependent terms guarantee stable neural activity, whereas the covariance-dependent terms support the emergence of specific E/I motifs, motivated by recent

experimental evidence (Znamenskiy et al. 2024). Consequently, our framework provides deeper insight into both the regulatory and structural properties of generalized iSTDP rules, expanding our understanding of plasticity in closed-loop recurrent circuits.

One strength of our framework is that it clarifies how the iSTDP kernel interacts with other plasticity parameters. For example, in rate-dominated regime the weight evolution mainly depends on the net area of the STDP kernel and the rate-dependent terms, and not on the precise kernel shape. In the covariance-dominated regime, the weight evolution depends on specific spike covariance between pre- and postsynaptic spiking and the iSTDP kernel shape. The covariance-dominated regime, however, can pose greater challenges for parameter tuning. A promising way forward is to employ numerical methods, for instance using simulation-based inference and deep learning, which have already been applied to optimize both excitatory and inhibitory STDP rules in recurrent circuits (Confavreux et al. 2024). Our analytical approach can guide and complement these efforts, helping clarify parameter degeneracies and simplifying the parameter search process.

Here we derived analytic solutions for the two-neuron Poisson model, and relied on numerical simulations for larger networks. By contrast, a recent study showed that iSTDP can stabilize recurrent LIF networks and converge to stable solutions through a mean-field approach (Akil et al. 2021). To achieve analytical tractability, that work assumed a regime of weak interactions, where internally generated correlations become negligible as the size of the neural population grows, and plasticity was driven exclusively by externally-induced correlations. In our approach, finite network size, strong interactions, and a dominance of correlation-dependent terms together drive self-organization. It remains an open question whether these two analytic frameworks, based on different assumptions, can be integrated.

While our analytical approach primarily examined how small E/I circuits self-organize under different iSTDP rules, neural circuits can also form larger structures, including feedforward firing sequences and assemblies. Assemblies are groups of strongly connected neurons, often regarded as the fundamental building blocks of neural computations (Buzsáki 2010; Papadimitriou et al. 2020; Miehl et al. 2023). Previous work has examined how various iSTDP rules enable the learning, maintenance and self-repair of neural assemblies (Ocker and Doiron 2019; Lagzi et al. 2021; Bergoin et al. 2023; Bergoin et al. 2025). In particular, recent findings suggest that assemblies are formed and best preserved when inhibitory neurons are divided into subclasses that combine symmetric and antisymmetric iSTDP (Lagzi et al. 2021). Here, the first rule can maintain network stability while the second prevents assemblies from merging through lateral inhibition, as in our work. Related work has also found that a combination of positive symmetric (Hebbian) and negative symmetric (anti-Hebbian) covariance-based rules can form and maintain stable assemblies governed by similar principles (Bergoin et al. 2025). These results on assemblies agree with ours in underscoring the intuition that different rules serve different purposes: some promote E/I balance and internal network stability, while others introduce competition through lateral inhibition.

In this work we further combined multiple iSTDP rules in a network with a ring topology (Ben-Yishai et al. 1995) to demonstrate how their interaction stabilizes activity while also supporting specific network computations, such as contextual input modulation as observed in the visual cortex (DeAngelis et al. 1994; Cavanaugh et al. 2002), and long-range spatial correlations reminiscent of modular spontaneous activity observed in the developing cortex (Smith et al. 2018; Powell et al. 2024). The parametrization of iSTDP rules allowed us to explore the entire space of rules and link them to the emergence of specific E/I connectivity motifs beyond the choice of pre-set, fixed plasticity rules in existing work (Lagzi et al. 2021; Bergoin et al. 2023; Lagzi and Fairhall 2024; Bergoin et al. 2025).

Here, we addressed the spontaneous emergence of circuit connectivity at the sub-assembly level and in the absence of externally patterned inputs. It would be interesting to consider non-stationary inputs, such as input correlations varying in time and across neurons, or highly structured input such as assembly-activations or waves spreading across the circuit as observed during cortical development (Smith et al. 2018). Non-stationary input, jointly with excitatory

STDP, might help paint a more complete picture of prolonged cortical development especially after the onset of sensory experience (Trägenap et al. 2025 [\[3\]](#)). Excitatory plasticity can enable circuits to learn new representations, refine current ones, and recover after lesions.

Overall, our findings highlight how local inhibitory learning rules can both stabilize global network dynamics and produce specific connectivity motifs in a self-organized manner in recurrent circuits, with implications for understanding neural development and learning under structured stimuli.

4 Methods

4.1 General mean-field solution

We consider a class of pairwise iSTDP rules triggered by spiking events and purely based on spike-timing difference between pre- and postsynaptic neuron firing, defined by the following expression.

$$\begin{aligned} \text{on } t_k^{(\text{pre})} \rightarrow \Delta W^{(\text{pre})} &= \alpha_{\text{pre}} + \sum_{k' < k} L^{\Delta-} \left(t_k^{(\text{pre})} - t_{k'}^{(\text{post})} \right), \\ \text{on } t_k^{(\text{post})} \rightarrow \Delta W^{(\text{post})} &= \alpha_{\text{post}} + \sum_{k' < k} L^{\Delta+} \left(t_k^{(\text{post})} - t_{k'}^{(\text{pre})} \right); \end{aligned} \quad (2)$$

where $t_k^{(\text{pre})}$ and $t_k^{(\text{post})}$ represent pre- and postsynaptic spike events, respectively. In our numerical simulations the instantaneous change of weight is further regulated by a global learning rate A_0 , as in $\Delta W^{*(\text{pre,post})} = A_0 \Delta W^{(\text{pre,post})}$. The scaling by A is omitted in our equations for simplicity. Each presynaptic spike triggers a weight change consisting of a fixed component, denoted α_{pre} , and a pairwise interaction term. Similarly, each postsynaptic spike triggers a weight change with a fixed component α_{post} and a pairwise interaction term. The pairwise interaction terms capture the cumulative influence of joint spiking events, incorporating the contribution of all preceding spikes ($k' < k$) weighted by a kernel $L^{\Delta_{\pm}(\cdot)}$ that depends on the time differences between the spikes. The two functions $L^{\Delta-}(\cdot)$ and $L^{\Delta+}(\cdot)$ refer to the post-pre and the pre-post pairwise interaction, and are merged in a single function, known as the iSTDP pairwise interaction kernel, as follows:

$$L(\tau) := \begin{cases} L^{\Delta+}(\tau) & \text{if } \tau \geq 0 \\ L^{\Delta-}(-\tau) & \text{if } \tau < 0. \end{cases} \quad (3)$$

By convention, the $\tau > 0$ branch of the curve represents pre-post interactions (presynaptic spike preceding the postsynaptic spike), while the $\tau < 0$ branch represents post-pre interactions (postsynaptic spike preceding the presynaptic spike). In our framework, we consider $L(\tau)$ to be independent of α_{pre} and α_{post} . This is in contrast to previous formulations that incorporate these terms into the kernel representation (e.g., Vogels et al. 2011 [\[4\]](#) and Confavreux et al. 2024 [\[5\]](#)). We assume that synaptic plasticity operates on a much slower timescale than neural dynamics, allowing us to treat synaptic weights as constant when calculating first- and second-order firing statistics (Kempner et al. 1999 [\[6\]](#)). We further assume variability in the spike-train statistics, so that it can be effectively described as a stochastic process. Thus, we can write the evolution of the mean synaptic weights as:

$$\begin{aligned} \frac{d\langle W \rangle}{dt} &= \alpha_{\text{pre}} r_{\text{pre}} + \alpha_{\text{post}} r_{\text{post}} + B r_{\text{pre}} r_{\text{post}} + Q, \\ \text{with } B &:= \int_{-\infty}^{+\infty} d\tau L(\tau) \text{ and } Q := \int_{-\infty}^{+\infty} d\tau L(\tau) C(\tau); \end{aligned} \quad (4)$$

where B is the integral of the iSTDP kernel, and Q is the only term that depends on the pairwise covariance density between pre and post-synaptic spike trains, here denoted $C(\tau)$. For a complete derivation see Supporting Information or [Kempster et al. \(1999\)](#) and [Luz and Shamir \(2014\)](#).

4.2 Expressions for the symmetric and antisymmetric exponential kernel

[Eq. 4](#) can be applied to any form of iSTDP kernel, including, e.g., kernels with delayed interactions ([Babadi and Abbott 2013](#)). Here, we focus on two specific exponentially decaying shapes: symmetric and antisymmetric, defined as follows:

$$L^{\text{sy}}(\tau) = \frac{1}{2} \left(\frac{1}{\tau_+} A_+ e^{-|\tau|/\tau_+} - \frac{1}{\tau_-} A_- e^{-|\tau|/\tau_-} \right) \quad \text{and} \quad L^{\text{an}}(\tau) = \begin{cases} -\frac{A_-}{\tau_-} e^{\tau/\tau_-} & \text{if } \tau < 0 \\ \frac{A_+}{\tau_+} e^{-\tau/\tau_+} & \text{if } \tau \geq 0 \end{cases} \quad (5)$$

With this parametrization, $A_+ = A_-$ implies $B = 0$, corresponding to covariance-dominated rules, whereas for $A_- = 0$ we have $B = A_+$, a rate-dominated regime. This choice of $L(\tau)$ allows us to keep track of the entire spiking history through activity traces updated at every spike time and decaying in time with the same constants as the iSTDP kernel. These traces will determine the amount of potentiation (+) or depression (-) triggered by each spike. For the antisymmetric rule, the presynaptic and postsynaptic traces are respectively defined as follows:

$$\begin{aligned} \frac{d}{dt} T_{r+}(t) &= -\frac{1}{\tau_+} T_{r+}(t) \quad \text{and} \quad \frac{d}{dt} T_{o-}(t) = -\frac{1}{\tau_-} T_{o-}(t) \\ \text{update:} \quad \begin{cases} T_{r+}(t + \varepsilon) &= T_{r+}(t) + 1 \\ T_{o-}(t + \varepsilon) &= T_{o-}(t) + 1 \end{cases} \quad \text{on } t = t^{(\text{pre})} \end{aligned} \quad (6)$$

The symmetric rule requires, in addition to $T_{r+}(t)$ and $T_{o-}(t)$, the traces $T_{r-}(t)$ and $T_{o+}(t)$, defined analogously. The implementation of the asymmetric iSTDP rule is then given by:

$$\begin{cases} \Delta W^{\text{an}}(t) = \alpha_{\text{pre}} - \frac{A_-}{\tau_{A_+}} T_{o-}(t) & \text{on } t = t^{(\text{pre})} \\ \Delta W^{\text{an}}(t) = \alpha_{\text{post}} + \frac{\bar{A}_+}{\tau_+} T_{r+}(t) & \text{on } t = t^{(\text{post})}, \end{cases} \quad (7)$$

and for the symmetric rule:

$$\begin{cases} \Delta W^{\text{sy}}(t) = \alpha_{\text{pre}} + \frac{A_+}{2\tau_+} T_{o+}(t) - \frac{A_-}{2\tau_-} T_{o-}(t) & \text{on } t = t^{(\text{pre})} \\ \Delta W^{\text{sy}}(t) = \alpha_{\text{post}} + \frac{A_+}{2\tau_+} T_{r+}(t) - \frac{A_-}{2\tau_-} T_{r-}(t) & \text{on } t = t^{(\text{post})}. \end{cases} \quad (8)$$

4.3 Low impact of correlations in rate-dominated rules

To better understand the role of the net area B of the iSTDP kernel in the weight dynamics ([Eq. 4](#)), and how it determines whether a rule is rate-dominated or covariance-dominated, we will consider here a simplified scenario, where $L(\tau)$ prescribes a +1 increase in weight for coincident pre-post spikes and no increase otherwise. This choice yields $B = A_+ = 1$. Consider now a pre and a post-synaptic neuron that fire with Poisson statistics at rates $r_{\text{pre}} = r_{\text{post}} = r = 10$ Hz. The term $B r_{\text{pre}} r_{\text{post}}$ in [Eq. 4](#) takes a value of 100 s^{-2} regardless of the correlation between the two neurons. But how does it compare to the covariance term Q in the presence of strong correlations between pre and post-synaptic activity? We can consider a regime of (unnaturally) high correlations, such that half of the spikes of the two neurons are synchronous (Pearson correlation 0.5). The term Q , corresponding to the spike-count covariance, is 5 s^{-2} . So even for very strong correlations Q is 20

times smaller than $B r_{\text{pre}} r_{\text{post}}$. This illustrates that in regimes with strong B , the weight dynamics in Eq. 4 are primarily governed by first-order statistics, with — at best — second-order contributions from correlations in spiking activity.

4.4 The two-dimensional reduced model

Finding a general solution for Eq. 4 can be challenging. For our purposes, we consider a simplified case. First, we model an excitatory and an inhibitory neuron as mutually interacting Poisson processes (Hawkes 1971; Ocker et al. 2017). Each time the presynaptic neuron spikes, the firing probability of the postsynaptic neuron increases (or decreases) proportionally to the excitatory (or inhibitory) post-synaptic potential EPSP (IPSP) kernel:

$$E_e(\tau) = \frac{1}{\tau_e} e^{-\tau/\tau_e} \quad \text{and} \quad E_i(\tau) = -\frac{1}{\tau_i} e^{-\tau/\tau_i} \quad \text{with } \tau \geq 0 \quad (9)$$

We further assume that the input to the inhibitory neuron $h_{\text{inh}}(t)$ is regulated homeostatically to keep its rate r_{inh} constant. Namely,

$$h_{\text{inh}}(t) = r_{\text{inh}} - h_{\text{exc}} w_{\text{exc}} + r_{\text{inh}} w_{\text{exc}} w_{\text{inh}}(t). \quad (10)$$

In those conditions, Eq. 4 has an approximate fixed-point closed-form solution for the inhibitory connection w_{inh} :

$$\bar{w}_{\text{inh}} = \frac{\alpha_{\text{pre}} r_{\text{inh}} + h_{\text{exc}} \Psi}{r_{\text{inh}} (\Psi - M_{ei}^{(10)})}; \quad \bar{r}_{\text{exc}} = -\frac{\alpha_{\text{pre}} r_{\text{inh}} + M_{ei}^{(10)} h_{\text{exc}}}{\Psi - M_{ei}^{(10)}}; \quad (11)$$

with $\Psi = \alpha_{\text{post}} + B r_{\text{inh}} + M_{ie}^{(01)} w_{\text{exc}}$.

The coefficients $M_{ei}^{(10)}$ and $M_{ie}^{(01)}$ are found through the linear response theory (Hawkes 1971; Ocker et al. 2017), and depend on the iSTDP kernel and the EPSP / IPSP parameters (refer to the analytic derivation of the SI Text, Section S1.4).

For the symmetric iSTDP:

$$M_{ei}^{\text{sy},(10)} = \frac{A_+}{\tau_+ + \tau_i} - \frac{A_-}{\tau_- + \tau_i}; \quad M_{ie}^{\text{sy},(01)} = \frac{A_+}{\tau_+ + \tau_e} - \frac{A_-}{\tau_- + \tau_e} \quad (12)$$

For the antisymmetric iSTDP instead:

$$M_{ei}^{\text{an},(10)} = \frac{A_+}{\tau_+ + \tau_i}; \quad M_{ie}^{\text{an},(01)} = -\frac{A_-}{\tau_- + \tau_e} \quad (13)$$

Numerical parameters are in Table 1.

4.5 Randomly connected recurrent neural network

We built a random network of conductance-based leaky integrate-and-fire (LIF) neurons, with 900 excitatory units and 100 inhibitory units. The membrane potential obeys the following dynamics:

$$\frac{dV(t)}{dt} = \frac{1}{C_{\text{membrane}}} (-g_{\text{leak}} (V - V_L) - g_{\text{AMPA}} (V - V_{\text{AMPA}}) - g_{\text{GABA}} (V - V_{\text{GABA}}) + \text{input}) \quad (14)$$

$$\frac{dg_{\text{AMPA}}(t)}{dt} = -\frac{g_{\text{AMPA}}(t)}{\tau_{\text{AMPA}}}; \quad \frac{dg_{\text{GABA}}(t)}{dt} = -\frac{g_{\text{GABA}}(t)}{\tau_{\text{GABA}}}$$

With $g_{\text{AMPA}}(t)$ and $g_{\text{GABA}}(t)$ updated for each incoming signal proportionally to the excitatory and inhibitory synaptic weights. We assigned a probability of connection for all non-plastic connections (excitatory to excitatory, excitatory to inhibitory, inhibitory to inhibitory). The inhibitory to excitatory connectivity was all-to-all and plastic. The network received an external input in the form of Poisson spike trains at fixed rate. The excitatory neurons received a small number of shared spikes, along with spikes generated i.i.d. for each neuron, leading to a small overall correlated activity. The network displayed high robustness to variations in input strength (see SI, Fig. S11). Numerical parameters were determined through grid search (Table 2).

4.6 Ring model

The ring model was implemented with the same neural dynamics as the random network (Eq. 14). Here however both excitatory and inhibitory neurons received a baseline external input, independent for each unit. We used 800 excitatory units and 200 inhibitory units, further divided into a PV and an SST subclass. The synaptic weight between an excitatory neuron and its neighbors decayed with distance, following a bell-shaped curve modeled as a rescaled von Mises distribution:

$$w_{ij}^{(\bullet \leftarrow e)} := C_{\bullet \leftarrow e} (\exp(k_{\bullet \leftarrow e} \cos(\theta_i - \theta_j))); \quad (15)$$

where θ_i and θ_j represent the locations of the post- and presynaptic neuron on the ring, with each population equally spaced in the $[0, 2\pi)$ interval. The factor $k_{\bullet \leftarrow e}$ regulates the connection range. Finally the scaling factor $\sum w_{\bullet \leftarrow e}$. All inh-to-inh connections were fixed and dense. We considered two distinct inhibitory populations $C_{\bullet \leftarrow e}$ is tuned so that the sum of incoming weight⁺s to either exc, PV or SST population is at a desired value with identical dynamics, but different iSTDP rules. The first population had a symmetric iSTDP kernel, the second an antisymmetric one. The input was different than in the previous model, with a constant external current in the excitatory population, and independent excitatory spike trains for all units.

To measure the network response to external stimuli (Fig. 5), we froze plasticity after learning, and delivered stimuli as uncorrelated spike trains. Iso-oriented stimuli took the form of excitatory spikes delivered to the exc neurons, contra-oriented stimuli were modeled as inhibitory spikes delivered to the SST neurons. Numerical parameters, determined through grid search, are shown in Table 3.

Data availability

Circuit simulations are based on custom code written in Julia (Bezanson et al. 2017, <https://julialang.org>), also available as separate toolbox at <https://github.com/comp-neural-circuits/HawkesSimulator.jl>. For the large-scale network simulations we used Brian2 (Stimberg et al. 2019, <https://briansimulator.org>). Finally some of the analytic integrals were solved using Maple (<https://maplesoft.com>). The code is fully open-source and publicly available on GitHub under the Creative Commons Attribution 4.0 International (CC BY 4.0) License at the following link: <https://github.com/comp-neural-circuits/structured-stabilization-in-recurrent-neural-circuits>

Supporting Information

S1 Analytic results

S1.1 Mean-field equation for weight dynamics

In this section, we adapt the results introduced by [Kempster et al. \(1999\)](#) to our notation and our specific needs. We first describe a sufficiently general expression for the pairwise STDP rules. Given two neurons, one presynaptic and one postsynaptic, an excitatory or inhibitory STDP rule is a weight update triggered by either a pre- or a postsynaptic spike. The weight update solely depends on the time differences between the previous spike times, relative to the current spike time:

$$\begin{aligned} \text{on } t = t_k^{(\text{pre})} \rightarrow \Delta W^{(\text{pre})} &= \alpha_{\text{pre}} + \sum_{k' < k} L^{\Delta_-} \left(t - t_{k'}^{(\text{post})} \right) \\ \text{on } t = t_k^{(\text{post})} \rightarrow \Delta W^{(\text{post})} &= \alpha_{\text{post}} + \sum_{k' < k} L^{\Delta_+} \left(t - t_{k'}^{(\text{pre})} \right), \end{aligned} \quad (\text{S1})$$

Where t_k and t_k represent pre- and postsynaptic spike events, respectively. We can merge the functions L^{Δ_-} and L^{Δ_+} into a single STDP kernel defined as:

$$L(\tau) := \begin{cases} L^{\Delta_+}(\tau) & \text{if } \tau \geq 0 \\ L^{\Delta_-}(-\tau) & \text{if } \tau < 0. \end{cases} \quad (\text{S2})$$

Therefore $\tau > 0$ (the Δ_+ branch) represents pre-post interactions (presynaptic spike preceding the postsynaptic spike), while $\tau < 0$ (the Δ_- branch) represents post-pre interactions (postsynaptic spike preceding the presynaptic spike).

It is useful to express each spike train as a sum of Dirac delta functions centered at each spike time:

$$S^i(t) := \sum_k \delta(t - t_k^i) \quad (\text{S3})$$

where i denotes the neuron index. With this notation, the number of spikes in the $[t_a, t_b]$ interval is:

$$N_{[t_a, t_b]}^i = \int_{t_a}^{t_b} dt S^i(t) \quad (\text{S4})$$

Assuming stationary conditions, we can therefore define the firing rate of neuron i as:

$$r_i := \frac{1}{T} \int_0^T dt S^i(t) = \langle S^i(t) \rangle_t \quad (\text{S5})$$

Using the definition in Eq. S3, we can then rewrite the pairwise interaction terms in Eq. S1 as:

$$\sum_{k'} L^{\Delta_{\pm}}(t - t_{k'}^i) = \int_{-\infty}^t d\tau S^i(\tau) L^{\Delta_{\pm}}(t - \tau) = \int_0^{\infty} d\tau^* S^i(t - \tau^*) L^{\Delta_{\pm}}(\tau^*) \quad (\text{S6})$$

Where in the last step we performed the change of variable $\tau \leftarrow t - \tau^*$. This leads to the following expression for the instantaneous change in synaptic weights, numerically equivalent to Eq. S1:

$$\begin{aligned} \text{on } t = t_k^{(\text{pre})} &\rightarrow \Delta W^{(\text{pre})}(t) = \alpha_{\text{pre}} + \int_0^{\infty} d\tau S^{\text{post}}(t - \tau) L^{\Delta_{-}}(\tau) \\ \text{on } t = t_k^{(\text{post})} &\rightarrow \Delta W^{(\text{post})}(t) = \alpha_{\text{post}} + \int_0^{\infty} d\tau S^{\text{pre}}(t - \tau) L^{\Delta_{+}}(\tau) \end{aligned} \quad (\text{S7})$$

Now, using again the same definition of spike train in Eq. S3, an event-based update can be expressed as an integral over time over the pre- and post-synaptic spiking events. The total weight change occurring in an interval $[t_a, t_b]$ is:

$$\Delta W_{[t_a, t_b]} = \int_{t_a}^{t_b} dt S^{\text{pre}}(t) \Delta W^{(\text{pre})}(t) + \int_{t_a}^{t_b} dt S^{\text{post}}(t) \Delta W^{(\text{post})}(t) \quad (\text{S8})$$

We can then express the average weight change per second, considering a time interval $[0, T]$ large enough to ensure precise sampling in second-order firing statistics, but still small compared to the rate of changes in ΔW . Under this assumption of separation of timescales between neuronal dynamics and plasticity, we can substitute Eq. S7 into Eq. S8, and write:

$$\begin{aligned} \langle \Delta W(t) \rangle_t &= \frac{1}{T} \int_0^T dt S^{\text{pre}}(t) \alpha_{\text{pre}} + \frac{1}{T} \int_0^T dt S^{\text{post}}(t) \alpha_{\text{post}} + \\ &+ \frac{1}{T} \int_0^T dt \int_0^{\infty} d\tau S^{\text{post}}(t) S^{\text{pre}}(t - \tau) L^{\Delta_{+}}(\tau) \\ &+ \frac{1}{T} \int_0^T dt \int_0^{\infty} d\tau S^{\text{pre}}(t) S^{\text{post}}(t - \tau) L^{\Delta_{-}}(\tau) \end{aligned} \quad (\text{S9})$$

Since α_{pre} and α_{post} are constants, we can factor them out. The first two terms of the r.h.s. of Eq. S9 become:

$$\frac{1}{T} \int_0^T dt S^{\text{pre}}(t) \alpha_{\text{pre}} = r_{\text{pre}} \alpha_{\text{pre}} \quad \text{and} \quad \frac{1}{T} \int_0^T dt S^{\text{post}}(t) \alpha_{\text{post}} = r_{\text{post}} \alpha_{\text{post}} \quad (\text{S10})$$

The pairwise interaction terms instead can be simplified as follows:

$$\begin{aligned} &\frac{1}{T} \int_0^T dt \int_0^{\infty} d\tau S^{\text{post}}(t) S^{\text{pre}}(t - \tau) L^{\Delta_{+}}(\tau) \\ &= r_{\text{pre}} r_{\text{post}} \int_0^{\infty} d\tau L^{\Delta_{+}}(\tau) + \int_0^{\infty} d\tau C^{\Delta_{+}}(\tau) L^{\Delta_{+}}(\tau) \end{aligned} \quad (\text{S11})$$

and similarly for the last term in Eq. S9. In the above equation, $C^{\Delta+}(\tau)$ denotes the covariance density between the pre- and the postsynaptic neuron, following the convention that Δ_+ represents the pre-post order and Δ_- the post-pre order:

$$C^{\Delta+}(\tau) := \frac{1}{T} \int_0^T (S^{\text{post}}(t) - r_{\text{post}}) (S^{\text{pre}}(t - \tau) - r_{\text{pre}}) . \quad (\text{S12})$$

An analogous definition holds for $C^{\Delta-}(\tau)$, with the pre- and post-temporal order inverted. Therefore

$$\begin{aligned} \Delta W(t) \rangle_t = & r_{\text{pre}} \alpha_{\text{pre}} + r_{\text{post}} \alpha_{\text{post}} \\ & + r_{\text{pre}} r_{\text{post}} \int_0^\infty d\tau (L^{\Delta-}(\tau) + L^{\Delta+}(\tau)) + \\ & + \int_0^\infty d\tau C^{\Delta-}(\tau) L^{\Delta-}(\tau) + \int_0^\infty d\tau C^{\Delta+}(\tau) L^{\Delta+}(\tau) \end{aligned} \quad (\text{S13})$$

Now using Eq. S2, we can write:

$$r_{\text{pre}} r_{\text{post}} \int_0^\infty d\tau (L^{\Delta-}(\tau) + L^{\Delta+}(\tau)) = r_{\text{pre}} r_{\text{post}} \int_{-\infty}^\infty d\tau L(\tau); \quad (\text{S14})$$

and given that $C^{\Delta+}(\tau) = C^{\Delta-}(-\tau)$ we can perform the following variable substitution:

$$\int_0^\infty d\tau C^{\Delta-}(\tau) L^{\Delta-}(\tau) = \int_{-\infty}^0 d\tau C^{\Delta+}(\tau) L^{\Delta+}(\tau) \quad (\text{S15})$$

Finally, substituting Eq. S14 and Eq. S15 in Eq. S13, and defining $C(\tau) := C^{\Delta+}(\tau)$, we have:

$$\begin{aligned} \langle \Delta W(t) \rangle_t = & \alpha_{\text{pre}} r_{\text{pre}} + \alpha_{\text{post}} r_{\text{post}} + B r_{\text{pre}} r_{\text{post}} + Q \\ \text{with } B := & \int_{-\infty}^{+\infty} d\tau L(\tau) \text{ and } Q := \int_{-\infty}^{+\infty} d\tau L(\tau) C(\tau); \end{aligned} \quad (\text{S16})$$

Corresponding to Eq. 1 in the main text.

S1.2 Derivation of the closed-form solution for the E/I motif

In this section, we derive a closed-form solution for a reduced model of one inhibitory and one excitatory neuron (Eq. S16). We model the two neurons as mutually interacting Hawkes processes (Main Text, Fig. 1A, Hawkes 1971; Pernice et al. 2011). The excitatory unit receives a constant input current h_{exc} , a plastic inhibitory connection $w_{\text{inh}}(t)$ and it projects a fixed excitatory connection w_{exc} to the inhibitory neuron. For simplicity, we assume that the inhibitory unit is kept at a constant rate r_{inh} by an adaptive external current $h_{\text{inh}}(t)$ (10). In this framework, the average firing rates behave linearly, therefore, the excitatory rate can be written as:

$$r_{\text{exc}}(t) = h_{\text{exc}} - w_{\text{inh}}(t) r_{\text{inh}} . \quad (\text{S17})$$

As we focus on the connection $w_{\text{inh}}(t)$ from the inhibitory to the excitatory neuron, we rename $r_{\text{inh}} = r_{\text{pre}}$ and $r_{\text{exc}} = r_{\text{post}}$ in Eq. S16 and we get:

$$\begin{aligned} \left\langle \frac{dw_{\text{inh}}(t)}{dt} \right\rangle &= \alpha_{\text{pre}} r_{\text{inh}} + \alpha_{\text{post}} (h_{\text{exc}} - w_{\text{inh}} r_{\text{inh}}) \\ &\quad + Br_{\text{inh}} (h_{\text{exc}} - w_{\text{inh}} r_{\text{inh}}) + Q \\ &= \alpha_{\text{pre}} r_{\text{inh}} - w_{\text{inh}} r_{\text{inh}} (\alpha_{\text{post}} + Br_{\text{inh}}) + h_{\text{exc}} (\alpha_{\text{post}} + Br_{\text{inh}}) + Q. \end{aligned} \quad (\text{S18})$$

The only nontrivial quantity to compute is the term Q , as it requires the inhibitory-to-excitatory covariance density curve. To avoid ambiguities, we define the Fourier transform operator acting on a function $f(t)$ and the convolution operator acting on two functions $f(t)$ and $g(t)$ as follows:

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} dt \quad f(t) e^{-j\omega t} \quad \text{and} \quad (f * g)(\tau) = \int_{-\infty}^{+\infty} dt \quad f(t) g(\tau - t); \quad (\text{S19})$$

where j denotes the imaginary unit.

Note that the total integral Q of the product $C(\tau) L(\tau)$ (as defined in Eq. S16) is equivalent to its Fourier transform evaluated at $\omega = 0$. But for the inverse convolution theorem, the Fourier transform of the product $C(\tau) L(\tau)$ corresponds to the convolution of the two transforms $C(\omega)$ and $\hat{L}(\omega)$. In equations:

$$\begin{aligned} Q &= \int_{-\infty}^{\infty} d\tau C(\tau) L(\tau) = \hat{f}(C(\tau) L(\tau))(\omega) \Big|_{\omega=0} \\ &= \frac{1}{2\pi} (\hat{C}(\tilde{\omega}) * \hat{L}(\tilde{\omega}))(\omega) \Big|_{\omega=0} = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tilde{\omega} \hat{C}(\tilde{\omega}) \hat{L}(-\tilde{\omega}) \end{aligned} \quad (\text{S20})$$

The advantage of re-defining Q in the Fourier space is that, for Hawkes processes, the covariance $C(\omega)$ can be expressed in a closed form as a function of network connectivity and neural activity (Hawkes 1971; Pernice et al. 2011; Trousdale et al. 2012; Jovanović et al. 2015). More specifically, given the matrix of post-synaptic potential kernels (Eq. 9 in Methods)

$$\mathbf{E}(\tau) = \begin{pmatrix} \frac{1}{\tau_e} e^{-\tau/\tau_e} & 0 \\ 0 & -\frac{1}{\tau_i} e^{-\tau/\tau_i} \end{pmatrix} \quad (\text{S21})$$

and the connectivity matrix

$$\mathbf{W} = \begin{pmatrix} 0 & w_{\text{inh}} \\ w_{\text{exc}} & 0 \end{pmatrix}, \quad (\text{S22})$$

we can first write the vector of firing rates $\mathbf{r} = [r_{\text{exc}}, r_{\text{inh}}]^T$ as

$$r = \mathbf{h}_0 + \int_{-\infty}^{+\infty} d\tau \mathbf{G}(\tau) S(\tau - t) \quad (\text{S23})$$

where $\mathbf{h}_0$ is the baseline input and $\mathbf{G}(\tau) = \mathbf{E}(\tau) \mathbf{W}$ is referred to as the matrix of “interaction kernels”. At the equilibrium state,

$$r = \mathbf{h}_0 + \mathbf{G}r, \quad \text{with } \mathbf{G} := \int_{-\infty}^{+\infty} d\tau \mathbf{G}(\tau) \quad (\text{S24})$$

where the integral in the r.h.s. corresponds to the Fourier transform evaluated at zero, i.e., $\hat{G}(\omega)|_{\omega=0}$. Equivalently, the rate vector can be written in an explicit form:

$$r = (\mathbf{I} - \mathbf{G})^{-1} \mathbf{h}_0, \quad (\text{S25})$$

where $\mathbf{I}$ represents the identity matrix. Moreover, if we write the covariance matrix (cf. Eq. S12) as

$$\mathbf{C}(\tau) = \langle \mathbf{S}(t) \mathbf{S}(t - \tau)^T \rangle - \mathbf{r} \mathbf{r}^T \quad (\text{S26})$$

then we have the following expression in the Fourier domain, at the equilibrium state:

$$\hat{\mathbf{C}}(\omega) = (\mathbf{I} - \hat{\mathbf{G}}(\omega))^{-1} \mathbf{D} (\mathbf{I} - \hat{\mathbf{G}}^T(\omega))^{-1} \quad (\text{S27})$$

where $\mathbf{D}$ is a diagonal matrix with the firing rates on its diagonal. Finally, recall that, for a matrix $\mathbf{G}$ with spectral radius < 1 ,

$$\mathbf{I} - \mathbf{G} = \sum_{\alpha} \mathbf{G}^{\alpha}. \quad (\text{S28})$$

Therefore, the covariance matrix can be written as the following expansion

$$\hat{\mathbf{C}}(\omega) = \sum_{\alpha=0}^N \sum_{\beta=0}^N \hat{\mathbf{E}}^{\alpha}(\omega) (\hat{\mathbf{E}}^T)^{\beta}(-\omega) \left[(\mathbf{W}^{\alpha}) \mathbf{D} (\mathbf{W}^T)^{\beta} \right], \quad (\text{S29})$$

where the indices α, β indicate the number of connections between any spiking neuron in the network to the post- and to the presynaptic neuron, respectively. This captures the influence of each spike propagated in the network through existing synaptic connections on the correlated firing of the pre- and postsynaptic neurons. Within this formalism, in our reduced model we consider only pairwise interactions, that is, $\alpha + \beta = 1$. The inhibitory-to-excitatory covariance is the (1, 2) element of $\hat{\mathbf{C}}(\omega)$. We can compute it as:

$$\hat{C}(\omega) = \hat{E}_e(\omega) w_{\text{inh}} r_{\text{inh}} + \hat{E}_i(-\omega) r_{\text{exc}} w_{\text{exc}} \quad (\text{S30})$$

By combining Eq. S20 and Eq. S30, we get the following expression for Q :

$$Q = M_{ei}^{(10)} w_{\text{inh}} r_{\text{inh}} + M_{ie}^{(01)} w_{\text{exc}} r_{\text{exc}} \\ M_{jk}^{(\alpha\beta)} := \int_{-\infty}^{+\infty} d\omega \hat{E}_j^{\alpha}(\omega) \hat{E}_k^{\beta}(-\omega) \hat{L}(-\omega) \quad (\text{S31})$$

The $M_{ei}^{(10)}$, $M_{ie}^{(01)}$ coefficients can be computed analytically depending on the iSTDP rule (see next section). Considering them as constants, we can then write Eq. S18 as:

$$\left\langle \frac{dw_{\text{inh}}(t)}{dt} \right\rangle = \alpha_{\text{pre}} r_{\text{inh}} - w_{\text{inh}} r_{\text{inh}} (\alpha_{\text{post}} + B r_{\text{inh}}) \\ + h_{\text{exc}} (\alpha_{\text{post}} + B r_{\text{inh}}) + \\ M_{ei}^{(10)} w_{\text{inh}} r_{\text{inh}} + M_{ie}^{(01)} w_{\text{exc}} r_{\text{exc}} \quad (\text{S32})$$

or equivalently,

$$\left\langle \frac{dw_{\text{inh}}(t)}{dt} \right\rangle = \alpha_{\text{pre}} r_{\text{inh}} - w_{\text{inh}} r_{\text{inh}} \left(\Psi - M_{ei}^{(10)} \right) + h_{\text{exc}} \Psi \quad (\text{S33})$$

with $\Psi := \alpha_{\text{post}} + Br_{\text{inh}} + M_{ie}^{(01)} w_{\text{exc}}$.

We then find the stationary point (Eq. 11 in Methods)

$$\tilde{w}_{\text{inh}} = \frac{\alpha_{\text{pre}} r_{\text{inh}} + h_{\text{exc}} \Psi}{r_{\text{inh}} \left(\Psi - M_{ei}^{(10)} \right)} \quad (\text{S34})$$

The equilibrium point $\tilde{w}_{\text{inh}}$ is a stable fixed point attractor, and positive, that is $\tilde{w}_{\text{inh}} \in [0, w_{\text{max}}]$, if the following conditions hold:

$$\begin{aligned} \Psi - M_{ei}^{(10)} &> 0 \\ \alpha_{\text{pre}} r_{\text{inh}} + \Psi h_{\text{exc}} &> 0, \end{aligned} \quad (\text{S35})$$

If a solution for $\tilde{w}_{\text{inh}}$ exists, we can also express the steady-state postsynaptic rate as follows:

$$\tilde{r}_{\text{exc}} = h_{\text{exc}} - \tilde{w}_{\text{inh}} r_{\text{inh}} = - \frac{\alpha_{\text{pre}} r_{\text{inh}} + M_{ei}^{(10)} h_{\text{exc}}}{\Psi - M_{ei}^{(10)}}, \quad (\text{S36})$$

The solution is valid only for positive rates. This adds one more condition for a stable and positive solution. The constraints are therefore:

$$\begin{aligned} \Psi - M_{ei}^{(10)} &> 0 \\ \alpha_{\text{pre}} r_{\text{inh}} + \Psi h_{\text{exc}} &> 0 \\ r_{\text{inh}} &> \frac{M_{ei}^{(10)} h_{\text{exc}}}{-\alpha_{\text{pre}}} \text{ with } \alpha_{\text{pre}} < 0. \end{aligned} \quad (\text{S37})$$

S1.3 Circuit motif with unconstrained r_{inh}

In the 2D E/I motif presented in the main text, the value of r_{inh} was fixed both for in the numerical simulations and in the analytic expression of the dynamics (Eq. 10 in Methods). This led to a closed-form expression for the fixed point solution (Eq. 11 in Methods):

$$\begin{aligned} \tilde{w}_{\text{inh}} &= \frac{\alpha_{\text{pre}} r_{\text{inh}} + h_{\text{exc}} \Psi}{r_{\text{inh}} \left(\Psi - M_{ei}^{(10)} \right)}; \quad \tilde{r}_{\text{exc}} = - \frac{\alpha_{\text{pre}} r_{\text{inh}} + M_{ei}^{(10)} h_{\text{exc}}}{\Psi - M_{ei}^{(10)}}; \\ \text{with } \Psi &= \alpha_{\text{post}} + Br_{\text{inh}} + M_{ie}^{(01)} w_{\text{exc}} \end{aligned} \quad (\text{S38})$$

In this section, for completeness, we show how the condition of fixed r_{inh} is not required for a stable solution and can in fact be relaxed at the cost of simplicity in the analytic expression of the fixed-point solution.

Consider the model shown in Fig. S3A in Methods. Here both r_{exc} and r_{inh} are free parameters, given by the net input. Namely:

$$\begin{aligned} r_{\text{exc}} &= h_{\text{exc}} - w_{\text{inh}} r_{\text{inh}} \\ r_{\text{inh}} &= h_{\text{inh}} + w_{\text{exc}} r_{\text{exc}} \end{aligned} \quad (\text{S39})$$

where we dropped the time dependence in the notation for simplicity. We can rewrite the equations above as:

$$r_{\text{exc}} = \frac{h_{\text{exc}} - w_{\text{inh}} h_{\text{inh}}}{1 + w_{\text{exc}} w_{\text{inh}}}; \quad r_{\text{inh}} = \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}}. \quad (\text{S40})$$

Then, plugging the terms above into (S32), we obtain:

$$\begin{aligned} \left\langle \frac{dw_{\text{inh}}(t)}{dt} \right\rangle = & \alpha_{\text{pre}} \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}} \\ & - w_{\text{inh}} \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}} \left(\alpha_{\text{post}} + B \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}} \right) + \\ & h_{\text{exc}} \left(\alpha_{\text{post}} + B \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}} \right) \\ & + M_{ei}^{(10)} w_{\text{inh}} \frac{h_{\text{inh}} + w_{\text{exc}} h_{\text{exc}}}{1 + w_{\text{exc}} w_{\text{inh}}} + \\ & M_{ie}^{(01)} w_{\text{exc}} \frac{h_{\text{exc}} - w_{\text{inh}} h_{\text{inh}}}{1 + w_{\text{exc}} w_{\text{inh}}}. \end{aligned} \quad (\text{S41})$$

We can find a fixed point $\tilde{w}_{\text{inh}}$ in the equation above by setting $\left\langle \frac{dw_{\text{inh}}(t)}{dt} \right\rangle = 0$ and multiplying by $(1 + w_{\text{exc}} w_{\text{inh}})^2$. The equation becomes a quadratic equation for w_{inh} :

$$aw_{\text{inh}}^2 + bw_{\text{inh}} + c = 0, \quad (\text{S42})$$

With

$$\begin{aligned} a = & w_{\text{exc}} \left(M_{ei}^{(10)} (h_{\text{inh}} + h_{\text{exc}} w_{\text{exc}}) - M_{ie}^{(01)} h_{\text{inh}} w_{\text{exc}} - \alpha_{\text{post}} h_{\text{inh}} \right) \\ b = & (h_{\text{inh}} + h_{\text{exc}} w_{\text{exc}}) \left(\alpha_{\text{pre}} w_{\text{exc}} - B h_{\text{inh}} + M_{ei}^{(10)} \right) \\ & - \alpha_{\text{post}} h_{\text{inh}} (1 - h_{\text{exc}} w_{\text{exc}}) + M_{ie}^{(01)} w_{\text{exc}} (h_{\text{exc}} w_{\text{exc}} - h_{\text{inh}}) \\ c = & (h_{\text{inh}} + h_{\text{exc}} w_{\text{exc}}) (\alpha_{\text{pre}} + B h_{\text{exc}}) + M_{ie}^{(01)} w_{\text{exc}} h_{\text{exc}} + \alpha_{\text{post}} h_{\text{exc}} \end{aligned} \quad (\text{S43})$$

The solution is then given by

$$\tilde{w}_{\text{inh}} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (\text{S44})$$

as long as $b^2 - 4ac \geq 0$.

Figure S3 shows analytic and numeric results for unidirectional and mutually-connected motif configurations for different iSTDP rules.

S1.4 Derivation of first-order motif coefficients

To calculate the motif coefficients

$$M_{ei}^{(10)} = \int_{-\infty}^{+\infty} d\omega \hat{L}(-\omega) \hat{E}_i(\omega) \quad \text{and} \quad M_{ie}^{(01)} = \int_{-\infty}^{+\infty} d\omega \hat{L}(-\omega) \hat{E}_e(-\omega) \quad (\text{S45})$$

for the connection W , we need to compute the Fourier transform of the postsynaptic potential E_e , E_i and the Fourier transform of the iSTDP kernel.

The Fourier transforms of the excitatory and inhibitory postsynaptic potential (Eq. 9 in Methods) are respectively given by:

$$\hat{E}_e(\omega) = \frac{1}{1 + j\tau_e \omega}, \quad \text{and} \quad \hat{E}_i(\omega) = -\frac{1}{1 + j\tau_i \omega}, \quad (\text{46})$$

where j is the imaginary unit.

Consider first the Hebbian symmetric STDP kernel, defined via the function

$$L^{\text{Sy}}(\tau) = \frac{1}{2} \left(\frac{1}{\tau_+} A_+ e^{-|\tau|/\tau_+} - \frac{1}{\tau_-} A_- e^{|\tau|/\tau_-} \right) \quad (\text{S47})$$

with $A_+, A_- > 0$. The Fourier transform is given by

$$\hat{L}^{\text{sy}}(\omega) = \hat{L}^{\text{sy}}(-\omega) = \frac{A_+}{1 + \tau_+^2 \omega^2} - \frac{A_-}{1 + \tau_-^2 \omega^2}. \quad (\text{S48})$$

By substituting Eq. S46 and Eq. S48 in Eq. S45, we can calculate the first-order motif coefficients $M_{ei}^{\text{sy},(10)}$, $M_{ie}^{\text{sy},(01)}$:

$$M_{ei}^{\text{sy},(10)} = \frac{A_+}{\tau_+ + \tau_i} - \frac{A_-}{\tau_- + \tau_i} \quad ; \quad M_{ie}^{\text{sy},(01)} = \frac{A_+}{\tau_+ + \tau_e} - \frac{A_-}{\tau_- + \tau_e}, \quad (\text{S49})$$

and therefore find an analytical solution for the integral Q in Eq. S31. Consider now the antisymmetric iSTDP rule, with kernel:

$$L^{\text{an}}(\tau) = \begin{cases} -\frac{A_-}{\tau_-} e^{\tau/\tau_-} & \text{if } \tau < 0 \\ \frac{A_+}{\tau_+} e^{-\tau/\tau_+} & \text{if } \tau \geq 0 \end{cases} \quad (\text{S50})$$

whose Fourier transform is given by

$$\hat{L}^{\text{an}}(\omega) = \frac{A_+}{1 + j\tau_+\omega} - \frac{A_-}{1 - j\tau_-\omega}. \quad (\text{S51})$$

Again, we can calculate the first order motif coefficients $M_{ei}^{\text{an},(10)}$, $M_{ie}^{\text{an},(01)}$ for the antisymmetric rule by substituting Eq. S46 and Eq. S51 in Eq. S45:

$$M_{ei}^{\text{an},(10)} = \frac{A_+}{\tau_+ + \tau_i} \quad ; \quad M_{ie}^{\text{an},(01)} = -\frac{A_-}{\tau_- + \tau_e}. \quad (\text{S52})$$

These expressions can be included into Eq. S34, leading to the final expression for the symmetric rule:

$$\begin{aligned} \Psi &= \alpha_{\text{post}} + Br_{\text{inh}} + M_{ie}^{\text{Sy},(01)} w_{\text{exc}} \\ &= \alpha_{\text{post}} + Br_{\text{inh}} + \left(\frac{A_+}{\tau_+ + \tau_e} - \frac{A_-}{\tau_- + \tau_e} \right) w_{\text{exc}} \\ \tilde{w}_{\text{inh}} &= \frac{\alpha_{\text{pre}} r_{\text{inh}} + h_{\text{exc}} \left(\alpha_{\text{post}} + Br_{\text{inh}} + \left(\frac{A_+}{\tau_+ + \tau_e} - \frac{A_-}{\tau_- + \tau_e} \right) w_{\text{exc}} \right)}{r_{\text{inh}} \left(\alpha_{\text{post}} + Br_{\text{inh}} + \left(\frac{A_+}{\tau_+ + \tau_e} - \frac{A_-}{\tau_- + \tau_e} \right) (w_{\text{exc}} - 1) \right)} \end{aligned} \quad (\text{S53})$$

And to the final expression for the antisymmetric rule:

$$\begin{aligned} \Psi &= \alpha_{\text{post}} + Br_{\text{inh}} + M_{ie}^{\text{an},(01)} w_{\text{exc}} \\ &= \alpha_{\text{post}} + Br_{\text{inh}} - \frac{A_-}{\tau_- + \tau_e} w_{\text{exc}} \\ \tilde{w}_{\text{inh}} &= \frac{\alpha_{\text{pre}} r_{\text{inh}} + h_{\text{exc}} \left(\alpha_{\text{post}} + Br_{\text{inh}} - \frac{A_-}{\tau_- + \tau_e} w_{\text{exc}} \right)}{r_{\text{inh}} \left(\alpha_{\text{post}} + Br_{\text{inh}} - \frac{A_-}{\tau_- + \tau_e} w_{\text{exc}} - \frac{A_+}{\tau_+ + \tau_i} \right)} \end{aligned} \quad (\text{S54})$$

S2 Structured stabilization in spiking networks with intrinsically-generated fluctuating activity

Our main numerical results are based on conductance-based spiking RNNs, where excitatory neurons receive random Poisson inputs, with both a shared and an independent component. To further demonstrate the generality of our results, we consider a somewhat different but important class of spiking RNNs, where fluctuating activity is generated completely intrinsically and does not depend on external inputs (Van Vreeswijk and Sompolinsky 1998 [\[1\]](#); Mastrogiuseppe and Ostojic 2017 [\[2\]](#))

We adapted a current-based spiking RNN model that spontaneously generates irregular activity under constant input currents (Mastrogiuseppe and Ostojic 2017 [\[2\]](#)). The single-neuron dynamics is as follows:

$$\begin{aligned}\tau_m \frac{dV_i}{dt} &= -V_i + RI_i(t) + RI_0 \\ I_i(t) &= \sum_j W_{ij} \sum_k \delta(t - t_j^{(k)} - \Delta_{\text{delay}})\end{aligned}\quad (\text{S55})$$

(see Table S1 [\[3\]](#) for the full list of parameters). We selected initial conditions that resulted in a high-firing, nearly synchronous firing regime (Fig. S5A,B [\[4\]](#)). In line with our previous results, we found that our choices of iSTDP rules lead to stable dynamics, reducing excitatory firing rates. Moreover, depending on the profile of correlation kernel, the network converged to either a prevalence of reciprocal E/I connections Fig. S5E [\[5\]](#), or lateral connectivity, expressed by unidirectional connections Fig. S5F [\[6\]](#).

These observations underscore the robustness of our proposed iSTDP mechanisms in shaping network connectivity beyond specific network architectures. The consistent emergence of either reciprocal or lateral connections, depending on the iSTDP rule, highlights a general principle by which local plasticity can shape E/I circuit structure.

S3 Supporting Figures

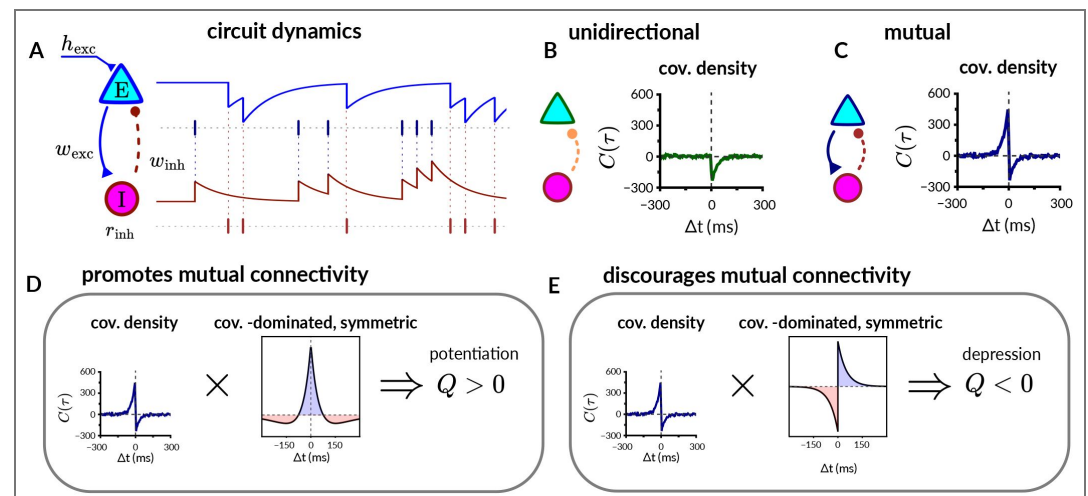


Figure S1. Schematics of the reduced circuit model and of interactions between covariance density and iSTDP kernel. **A.** Representation of the model used for Fig. 1 [\[1\]](#) in main text. Two time-varying coupled Poisson units, one excitatory and one inhibitory, are connected by a non-plastic excitatory weight w_{exc} and a plastic inhibitory weight w_{inh} . The instantaneous rates, here shown as traces, are computed as the convolution between each incoming spike and an exponential synaptic kernel that either increases (exc) or decreases (inh) the instantaneous firing probability. The input to the excitatory neuron h_{exc} is fixed, whereas the inhibitory input changes dynamically, keeping the inhibitory neuron at a fixed rate r_{inh} . Only the inhibitory synapse (dashed line)

symbol	value	description
N_{exc}	800	number of excitatory neurons
N_{inh}	200	number of inhibitory neurons
τ_m	20 ms	membrane time constant
$V_{\text{threshold}}$	20 mV	firing threshold
V_{reset}	10 mV	reset membrane potential
$\tau_{\text{refractory}}$	0.5 ms	refractory period
Δ_{delay}	0.55 ms	synaptic transmission delay
$R I_0$	0.23 mV	constant external input
$p_{e \leftarrow e}$	0.1	exc-to-exc connection probability
$p_{i \leftarrow e}$	0.1	exc-to-inh connection probability
$p_{e \leftarrow i}$	1.0	inh-to-exc connection probability
$p_{i \leftarrow i}$	1.0	inh-to-inh connection probability
$w_{e \leftarrow e}$	0.2	exc-to-exc weight
$w_{i \leftarrow e}$	0.2	exc-to-inh weight
$w_{i \leftarrow i}$	0.1	inh-to-inh weight
<i>symmetric iSTDP rule</i>		
A_0	$4 \cdot 10^{-3}$	global learning rate
B	0	net integral of the STDP kernel
α_{pre}	-0.05	constant dependent on presynaptic rate
α_{post}	0	constant dependent on postsynaptic rate
τ_+	30 ms	iSTDP kernel time constant
τ_-	600 ms	iSTDP kernel time constant
<i>asymmetric iSTDP rule</i>		
A_0	$5 \cdot 10^{-5}$	global learning rate
B	0.3	net integral of the iSTDP kernel
α_{pre}	-0.7	constant dependent on presynaptic rate
α_{post}	0.2	constant dependent on postsynaptic rate
τ_+	60 ms	iSTDP kernel time constant for potentiation
τ_-	30 ms	iSTDP kernel time constant for depression

Table S1. Numerical parameters for the spiking RNN with intrinsically-generated fluctuating activity (Eq. S55 [↗](#))

is plastic. **B.** Unidirectional configuration. Here $w_{exc} = 0$ and the covariance density between the pre (inh) and post (exc) neuron can only be negative in the pre-post ($\Delta t > 0$) section of the curve, due to the inhibitory interaction given by w_{inh} . **C.** Mutual configuration, $w_{exc} = 0.5$. Here the covariance density is positive in the post-pre ($\Delta t < 0$) section of the curve due to w_{exc} , and negative for ($\Delta t > 0$) for the effect of w_{inh} . **D.** Effect of a symmetric, covariance-dominated iSTDP kernel on a circuit in the M configuration: the positive left branch of the covariance density interacts with the kernel, resulting in a positive Q factor in the weight update. (Eq. 1 [↗](#)). **E.** Effect of an antisymmetric iSTDP kernel. Here the positive covariance density interacts with the negative branch of the kernel, resulting into synaptic depression. Hence, the Q term is negative, and mutual connections are discouraged.

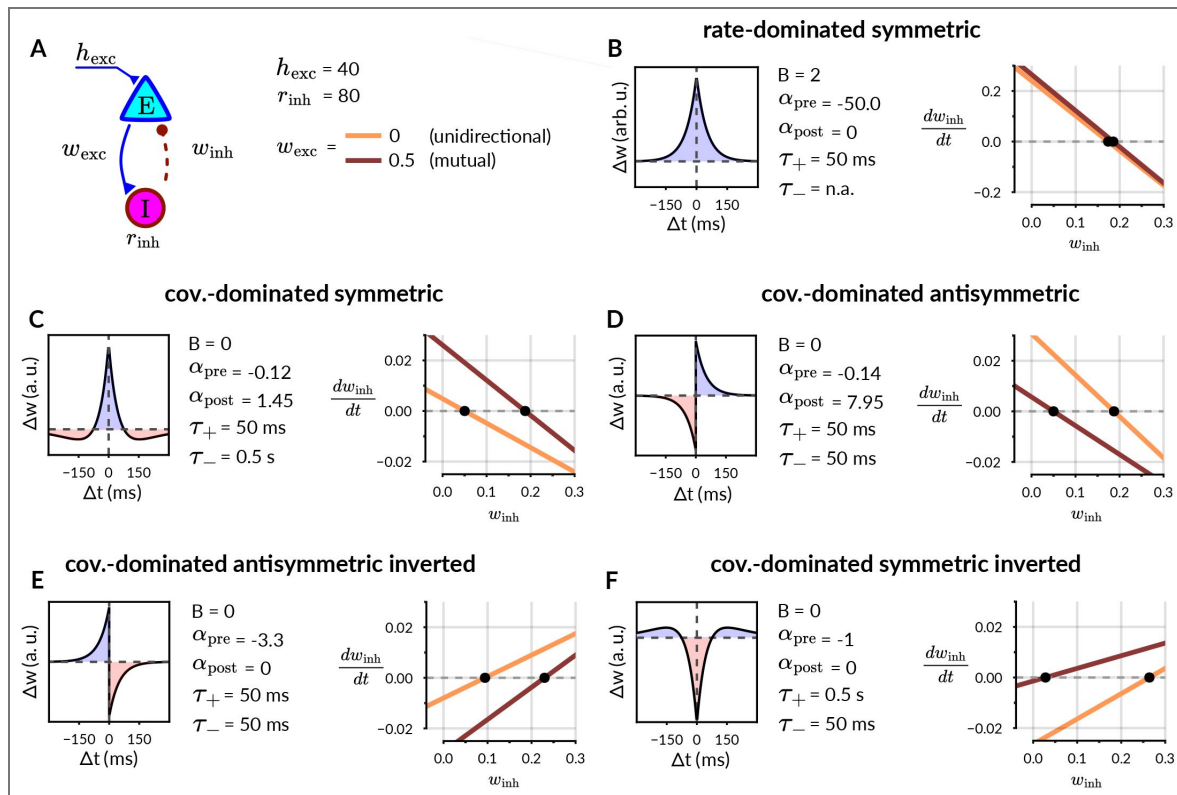


Figure S2. Analytic fixed point solutions for different iSTDP interaction kernels (complements Fig. 2 in main text).

A. Representation of the model (see also Fig. 1A). **B-E.** iSTDP rules, represented by their kernel and their additional parameters (left panels) and study of the associated weight dynamics as a one-dimensional differential equation (right panels), black dots represent fixed points. **B** Rate-dominated iSTDP rule. **C-E.** covariance dominated rules presented in the same order as in Fig. 2A, main text. Note that the configurations in **E** and **F** produce unstable fixed points.

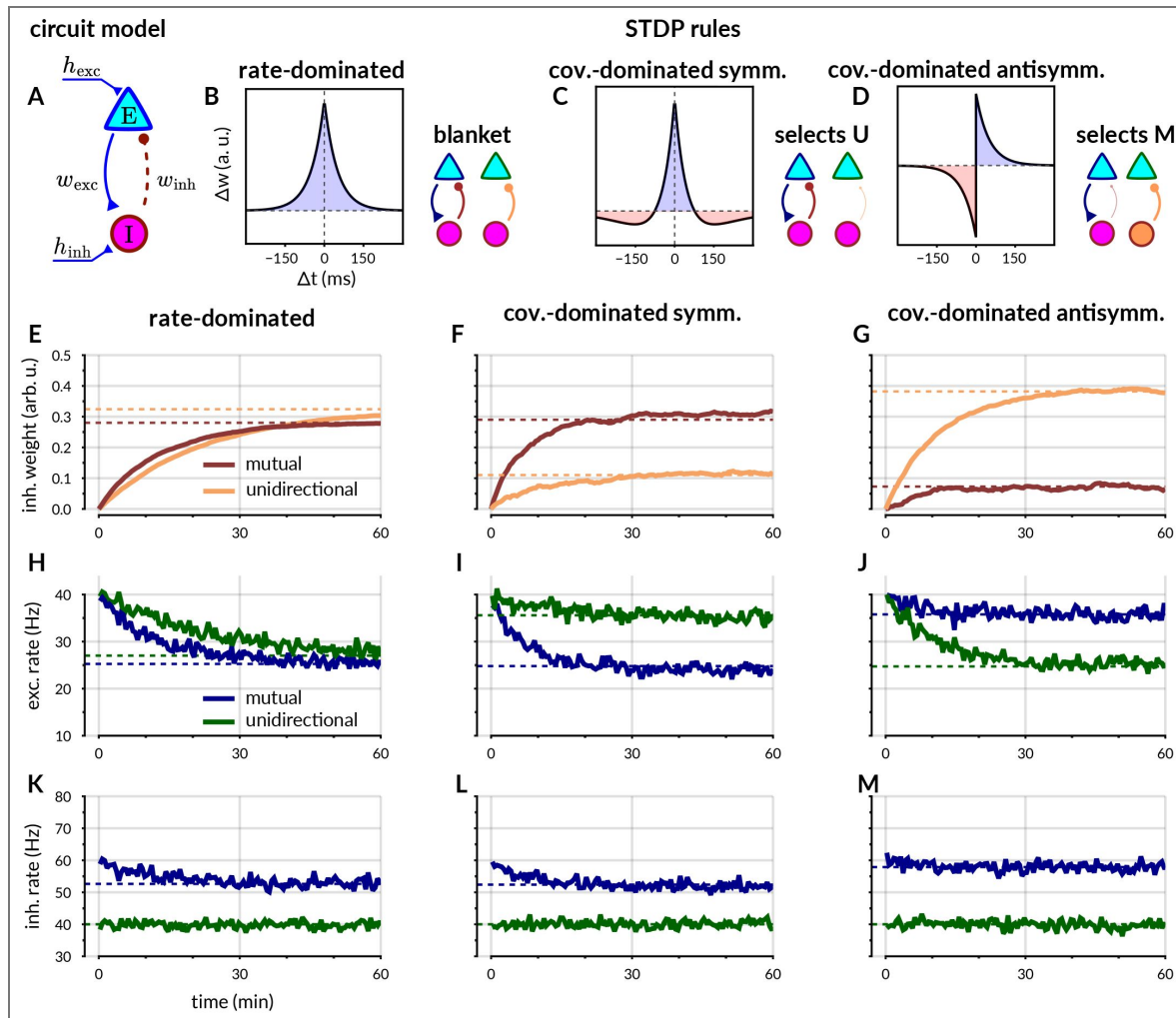


Figure S3. Effects of plasticity for E/I motifs in a reduced model with unconstrained r_{inh} .

See Section S1.3 for model description and details on the analytic calculation. **A.** Model schematics. Unlike the model in Fig. 1A, the inhibitory rate r_{inh} is unconstrained and determined by the net inhibitory input h_{inh} . **B-J.** Corresponding to the respective panels in Fig. 1, for the model with unconstrained r_{inh} . **K-L.** Value of r_{inh} during the simulation. Note that in the U configuration (dark green) r_{inh} is stationary and corresponds to h_{inh} .

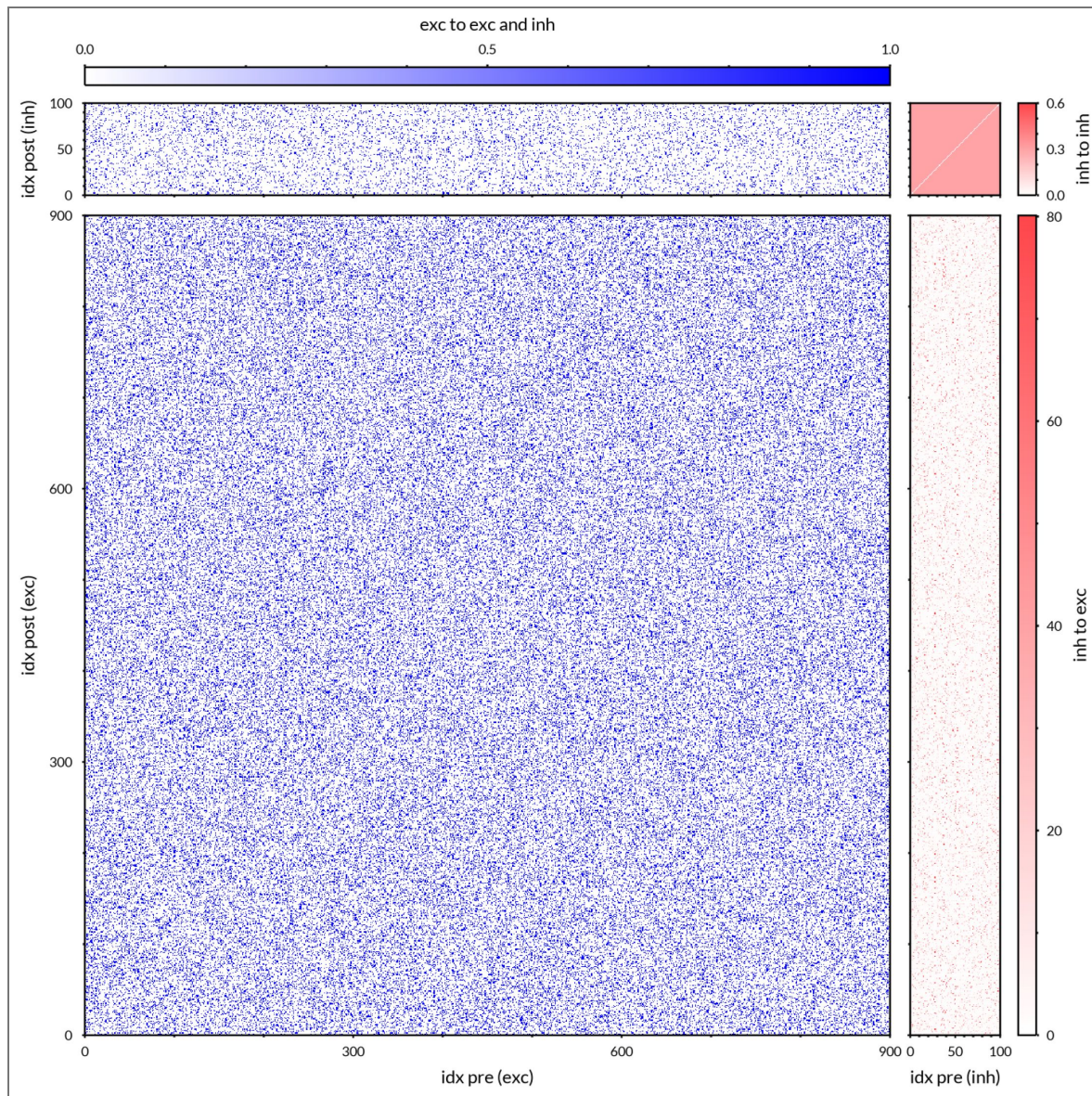


Figure S4. Full weight matrix for the random sparse RNN simulation with symmetric iSTDP (see Main Text, Fig. 3B-F).

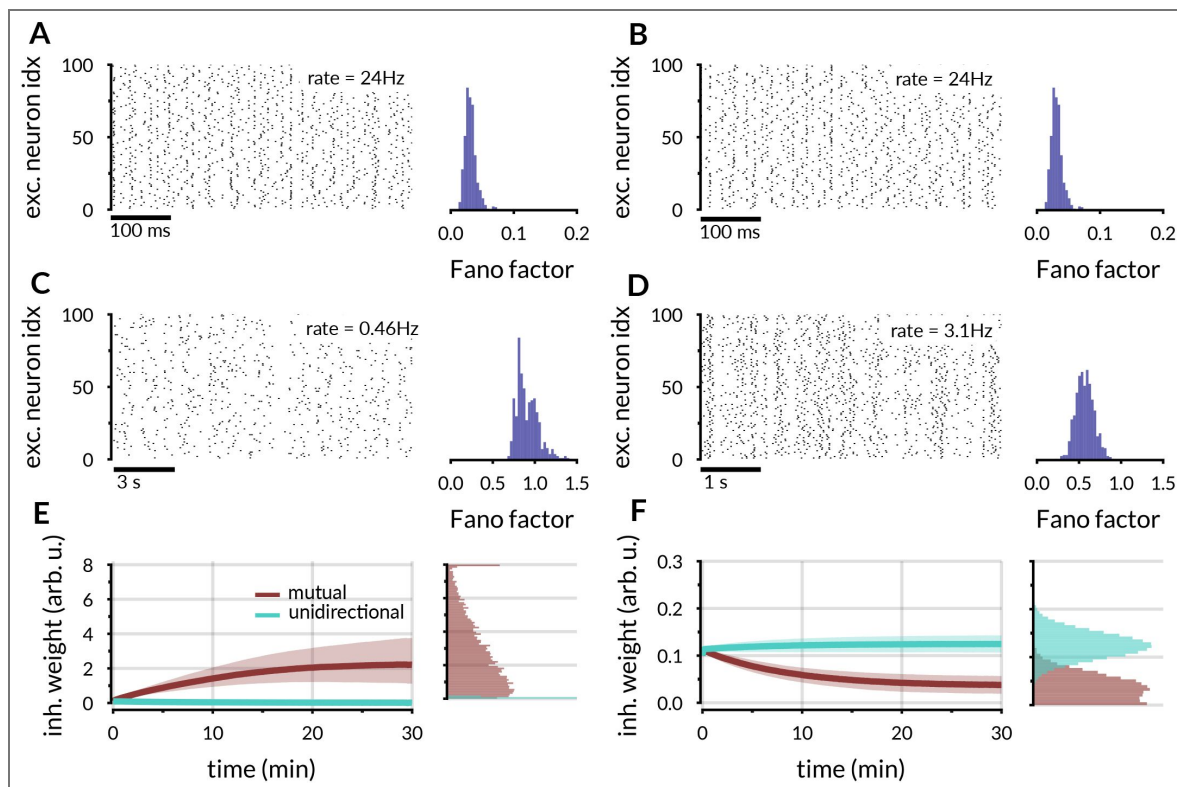


Figure S5. Randomly connected spiking RNN with intrinsically-generated fluctuating asynchronous irregular activity.

See Table S1 [\[link\]](#) for numerical parameters. **A-B.** Spiking activity of a subset 100 excitatory neurons (left) and distribution of Fano factors of excitatory neurons (right) before plasticity. **C-D.** Spiking activity and Fano factor distribution after weight convergence, using either a symmetric (**C**) or an antisymmetric (**D**) iSTDP kernel. **E-F.** Change of mutual and unidirectional inhibitory-to-excitatory weights during learning in the circuit. Equivalent to Fig. 3C [\[link\]](#) and H [\[link\]](#) in Main Text.

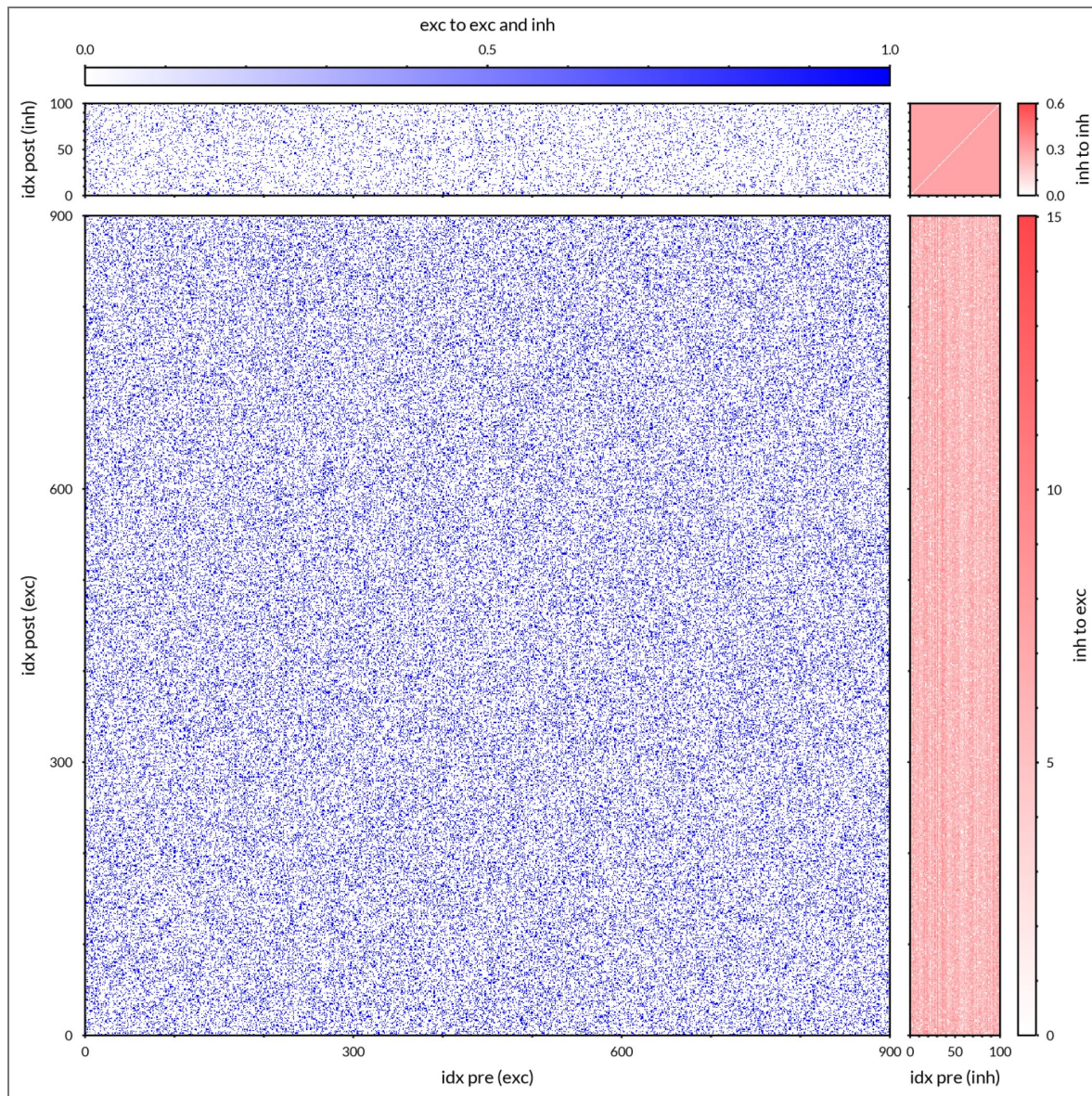


Figure S6. Full weight matrix for the random sparse RNN simulation with antisymmetric iSTDP (see Main Text, Fig. 3G-K [↗](#)).

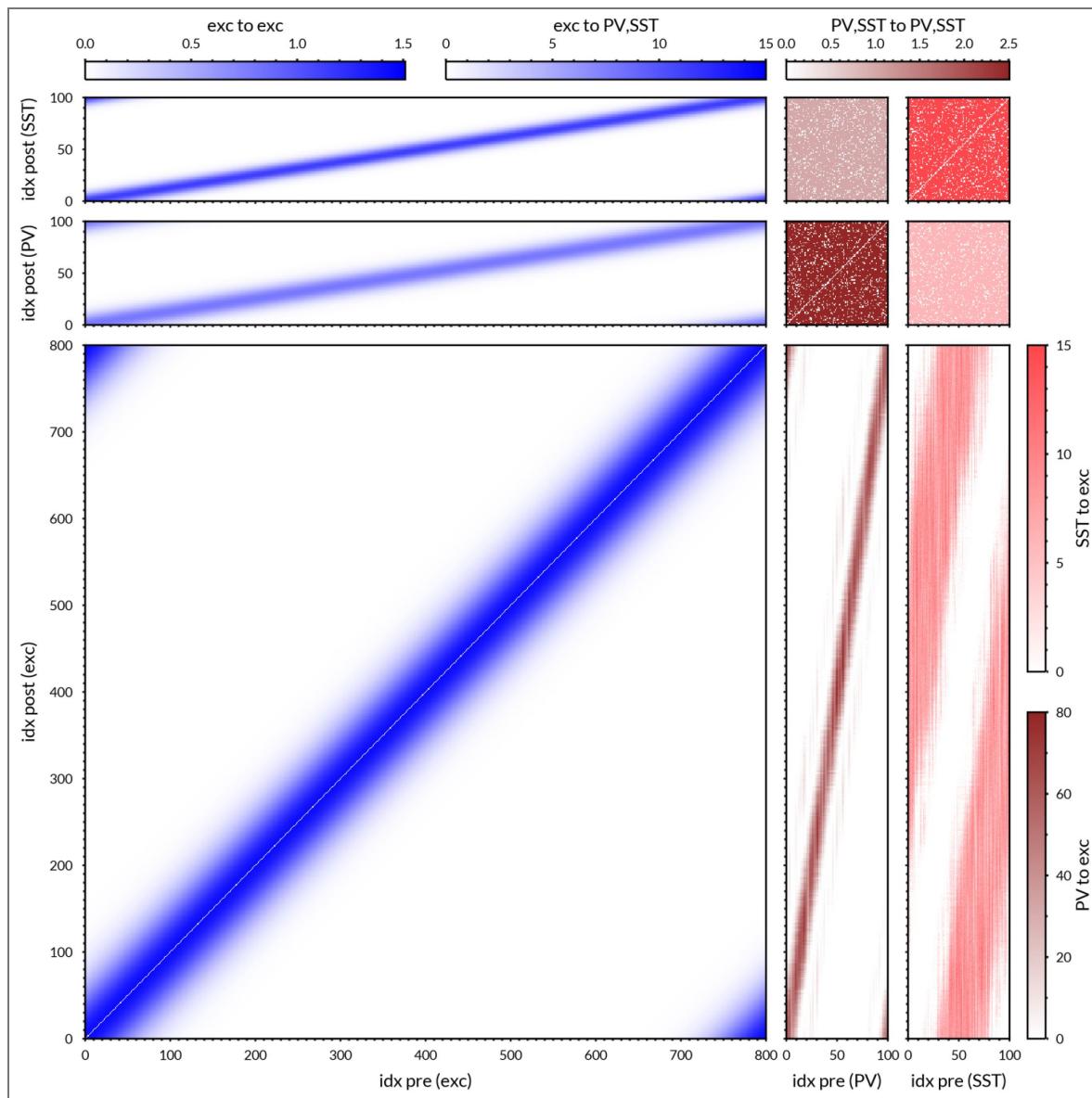


Figure S7. Full weight matrix for the recurrent network with 1D ring connectivity structure after weight convergence with two inhibitory populations.

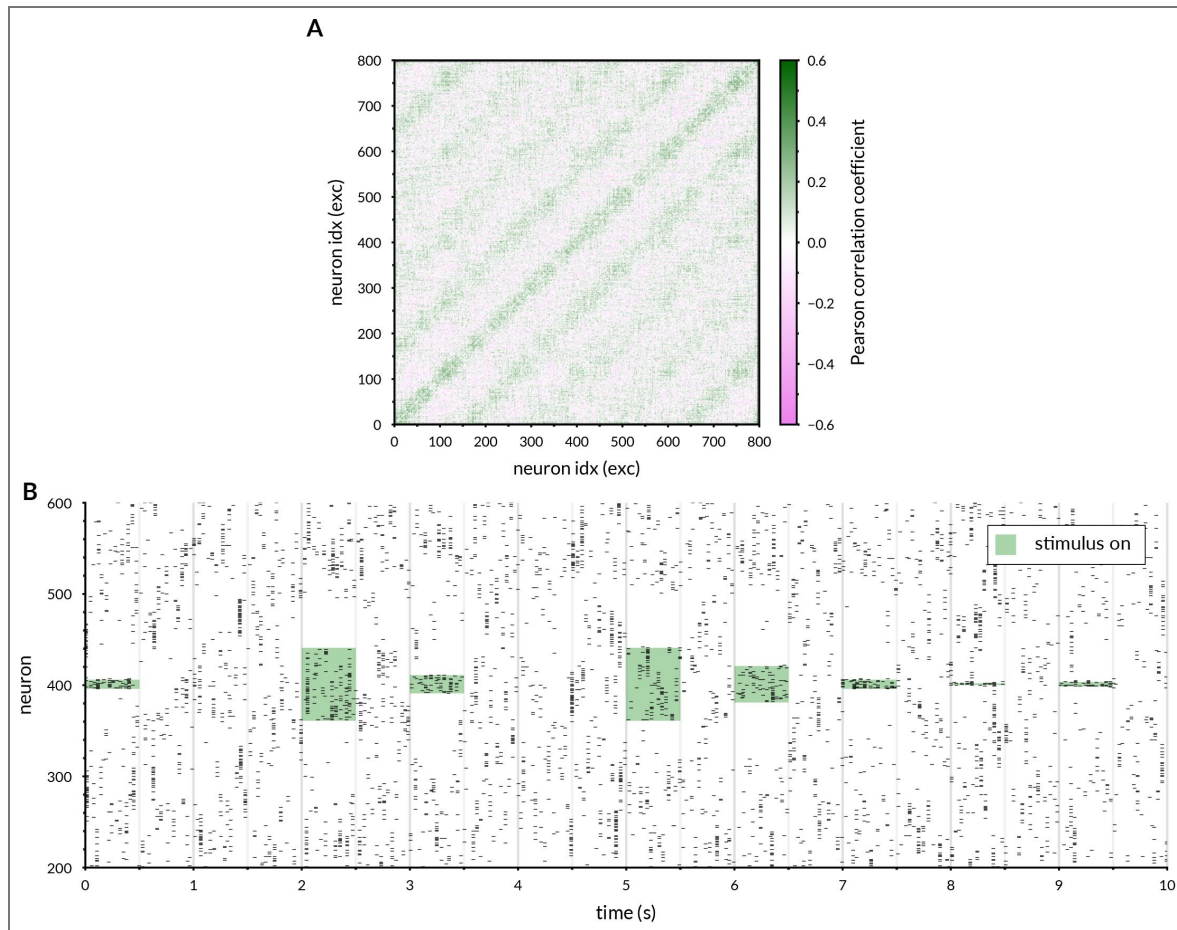


Figure S8.

A. Full matrix of Pearson correlations between excitatory neurons in the one-dimensional ring model during spontaneous activity, after convergence of plastic weights. **B** Spike raster showing the response of the 1D ring model to localized stimuli (see Fig. 4 in main text).

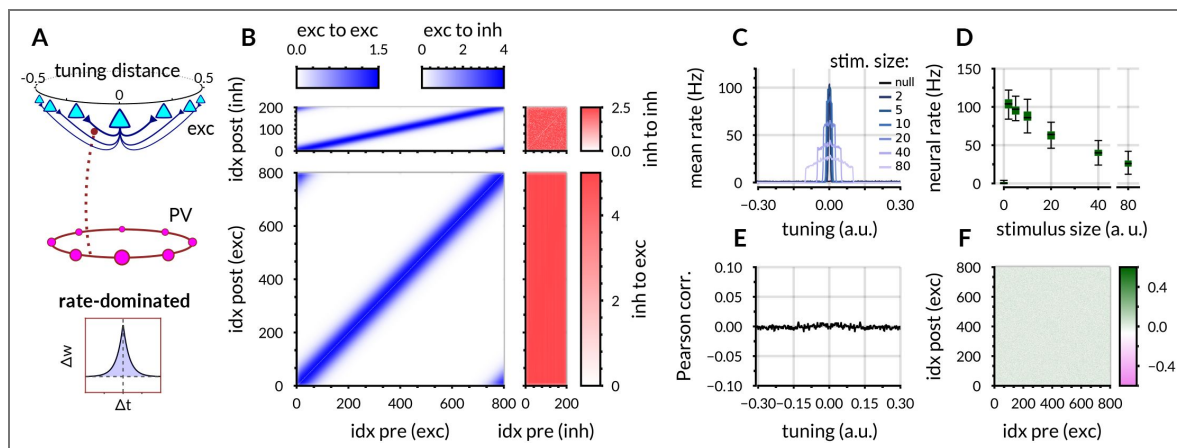


Figure S9. Spontaneous and evoked activity in a 1D ring structure with rate-dominated iSTDp.

A. Model schematics, compare with Fig. 4A. The excitatory component is identical to the previous case, but the inhibitory population follows a single, rate-dominated iSTDp. **B.** Full weight matrix after weight convergence, compare with Fig. 4B. **C.** Mean responses of the excitatory population to stimuli of different sizes, averaged over 100 trials (compare with Fig. 5A). **D.** Response of the center neuron to stimuli of different sizes (compare with Fig. 5B). **E.** Pearson correlation during spontaneous activity between an excitatory neuron and its neighbors, ordered by tuning distance and averaged over the entire population (compare with Fig. 4F). **F.** Pearson correlation of spontaneous activity for every pair of excitatory neurons (compare with Fig. S8A).

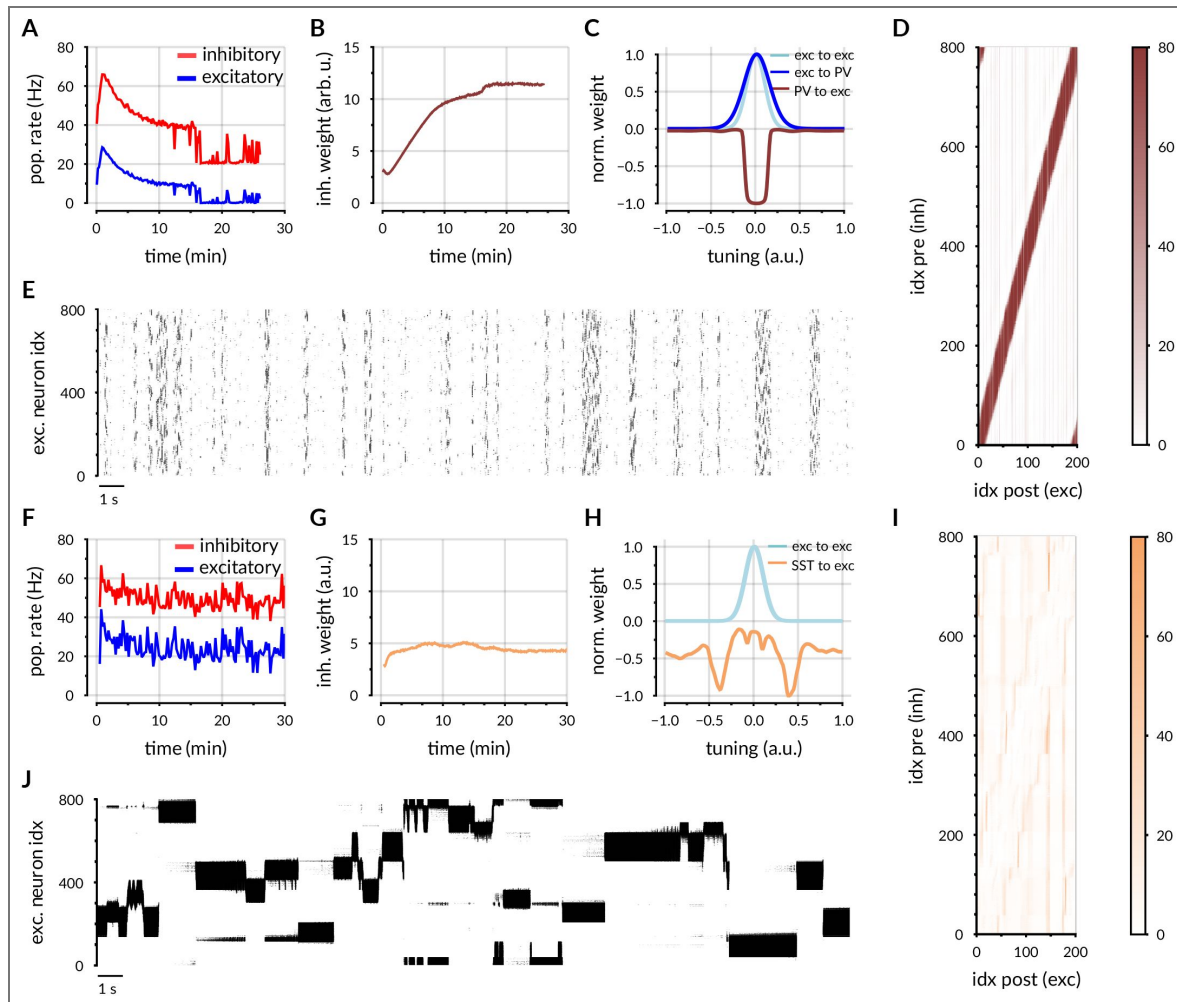


Figure S10. One-dimensional ring structure with only one inhibitory population (either PV or SST).

A-E. The RNN has 800 excitatory neurons, 200 “PV” neurons and no “SST” neurons. The network evolves under a PV-like covariance-dominated iSTDP rule with $B = 0$, $\alpha_{\text{pre}} = -1.0$, $\alpha_{\text{post}} = 1.0$. **A.** Evolution of population mean firing rates over time. **B.** Evolution of inh-to-exc synaptic weights over time. **C.** Profile of the average outgoing weights as a function of tuning distance. For convenience, the profiles are normalized so that the maximum is either 1 or -1. **D.** Inhibitory to excitatory weight matrix after weight convergence. **E.** Raster plot of excitatory neurons. **F-J.** Same but for a RNN with 800 excitatory neurons, 200 “SST” neurons and no “PV” neurons. The network evolves under a SST-like covariance-dominated iSTDP rule with $B = 0$, $\alpha_{\text{pre}} = -1.0$, $\alpha_{\text{post}} = 1.5$.

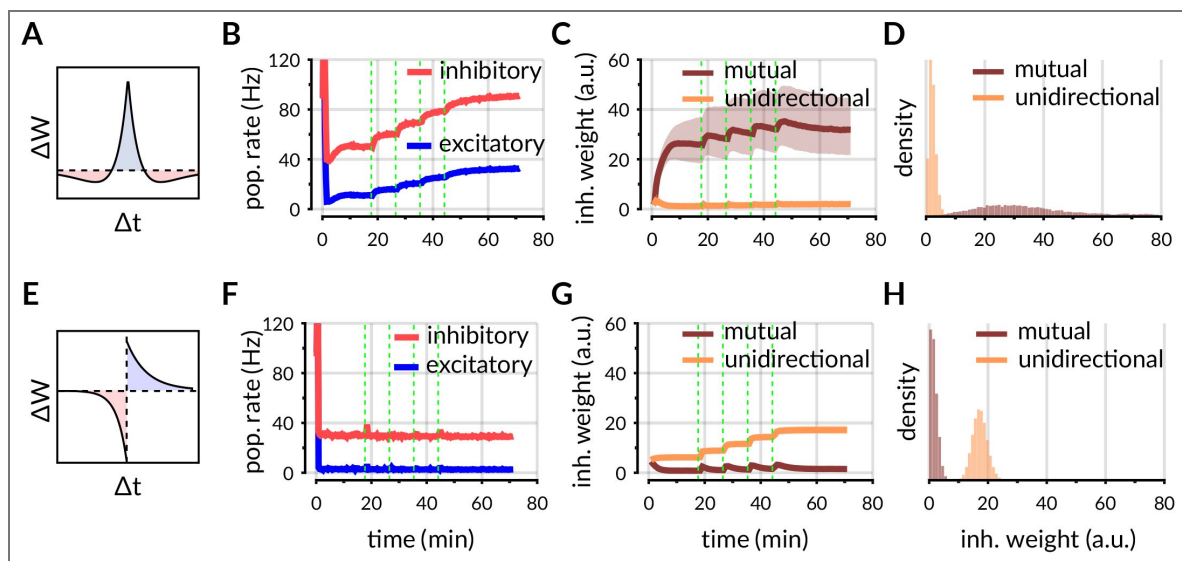


Figure S11. Effects of plasticity for different levels of mean inhibitory input.

A. Symmetric kernel. **B.** Evolution of population mean firing rates over time. The green vertical lines represent sequential increases of the mean input by a factor of 1.5, 2.0, 2.5, 3. **C.** Evolution of inh-to-exc synaptic weights over time. **D.** Final distribution of weights, split into a “mutual” (brown) and a “unidirectional” (light orange) group. **E-H.** Same as above but for asymmetric iSTDP.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (ERC CoG FeedbackCircuits, grant agreement no. 101170267), and under the Marie Skłodowska-Curie ITN SmartNets (grant agreement No. 860949) under the European Union's Horizon 2020 research and innovation program, by the International Human Frontier Science Program Organization (RGP0062/2021), and by the Technical University of Munich (TUM). Thanks to the "Computation in Neural Circuits" group for their useful insights and discussions. Special thanks to Alessio Quaresima for his valuable feedback and for independently testing our results.

Additional files

[Supplemental material](#) 

Additional information

Funding

Funder	Grant reference number	Author
EC H2020 PRIORITY 'Excellent science' H2020 European Research Council (ERC)	https://doi.org/10.3030/804824	Claudia Cusseddu Julijana Gjorgjieva Dylan Festa
EC Horizon Europe Excellent Science HORIZON EUROPE European Research Council (ERC)	https://doi.org/10.3030/101170267	Claudia Cusseddu Julijana Gjorgjieva
EC Horizon 2020 Framework Programme (H2020)	https://doi.org/10.3030/860949	Claudia Cusseddu Julijana Gjorgjieva
Technische Universität München (TUM)		Claudia Cusseddu Julijana Gjorgjieva Dylan Festa
Human Frontier Science Program (HFSP)	https://doi.org/10.52044/hfsp.rgp006 22021.pc.gr.163582	Dylan Festa Julijana Gjorgjieva

Author ORCID iDs

Dylan Festa:  <https://orcid.org/0000-0003-3803-1542>

Julijana Gjorgjieva:  <https://orcid.org/0000-0001-7118-4079>

References

- Agnes Everton J**, Vogels Tim P (2024) Co-Dependent Excitatory and Inhibitory Plasticity Accounts for Quick, Stable and Long-Lasting Memories in Biological Networks. *Nature Neuroscience* 1-11 <https://doi.org/10.1038/s41593-024-01597-4> | [PubMed](#)
- Akil Alan Eric**, Rosenbaum Robert, Josić Krešimir (2021) Balanced Networks under Spike-Time Dependent Plasticity. *PLOS Computational Biology* **17**:e1008958 <https://doi.org/10.1371/journal.pcbi.1008958> | [PubMed](#)
- Shun'ichi Amari** (1977) Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields. *Biological Cybernetics* **27**:77-87 <https://doi.org/10.1007/BF00337259> | [PubMed](#)
- Babadi Baktash**, Abbott LF (2013) Pairwise Analysis Can Account for Network Structures Arising from Spike-Timing Dependent Plasticity. *PLOS Computational Biology* **9**:e1002906 <https://doi.org/10.1371/journal.pcbi.1002906> | [PubMed](#)

- Ben-Yishai R, BarOr R L, Sompolinsky H** (1995) Theory of Orientation Tuning in Visual Cortex. *Proceedings of the National Academy of Sciences* **92**:3844-3848 <https://doi.org/10.1073/pnas.92.9.3844> | [PubMed](#)
- Raphaël Bergoin, Torcini Alessandro, Deco Gustavo, Quoy Mathias, Zamora-López Gorka** (2023) Inhibitory Neurons Control the Consolidation of Neural Assemblies via Adaptation to Selective Stimuli. *Scientific Reports* **13**:6949 <https://doi.org/10.1038/s41598-023-34165-0>
- Raphaël Bergoin, Torcini Alessandro, Deco Gustavo, Quoy Mathias, Zamora-López Gorka** (2025) Emergence and Maintenance of Modularity in Neural Networks with Hebbian and Anti-Hebbian Inhibitory STDP. *PLOS Computational Biology* **21**:e1012973 <https://doi.org/10.1371/journal.pcbi.1012973>
- Bezanson Jeff, Edelman Alan, Karpinski Stefan, Shah Viral B** (2017) Julia: A fresh approach to numerical computing. *SIAM review* **59**:65-98 <https://doi.org/10.1137/141000671>
- Bollmann Yannick, et al.** (2023) Prominent in Vivo Influence of Single Interneurons in the Developing Barrel Cortex. *Nature Neuroscience* **26**:1555-1565 <https://doi.org/10.1038/s41593-023-01405-5> | [PubMed](#)
- Brunel N** (2000) Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of computational neuroscience* **8**:183-208 <https://doi.org/10.1023/A:1008925309027> | [PubMed](#)
- Buzsáki György, Mizuseki Kenji** (2014) The Log-Dynamic Brain: How Skewed Distributions Affect Network Operations. *Nat Rev Neurosci* **15**:264-278 <https://doi.org/10.1038/nrn3687> | [PubMed](#)
- György Buzsáki** (2010) Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron* **68**:362-385 <https://doi.org/10.1016/j.neuron.2010.09.023>
- Cavanaugh James R, Bair Wyeth, Anthony Movshon J** (2002) Nature and Interaction of Signals From the Receptive Field Center and Surround in Macaque V1 Neurons. *Journal of Neurophysiology* **88**:2530-2546 <https://doi.org/10.1152/jn.00692.2001> | [PubMed](#)
- Clopath Claudia, Vogels Tim P, Froemke Robert C, Sprekeler Henning** (2016) Receptive Field Formation by Interacting Excitatory and Inhibitory Synaptic Plasticity. *bioRxiv* 066589 <https://doi.org/10.1101/066589>
- Confavreux Basile, Agnes Everton J, Zenke Friedemann, Sprekeler Henning, Vogels Tim P** (2024) Balancing Complexity, Performance and Plausibility to Meta Learn Plasticity Rules in Recurrent Spiking Net-works. *bioRxiv* <https://doi.org/10.1101/2024.06.17.599260>
- D'amour James A, Froemke Robert C** (2015) Inhibitory and Excitatory Spike-Timing-Dependent Plasticity in the Auditory Cortex. *Neuron* **86**:514-528 <https://doi.org/10.1016/j.neuron.2015.03.014> | [PubMed](#)
- Dahmen David, et al.** (2020) Strong Coupling and Local Control of Dimensionality across Brain Areas. *bioRxiv* 2020.11.02.365072 <https://doi.org/10.1101/2020.11.02.365072>
- DeAngelis GC, Freeman RD, Ohzawa I** (1994) Length and Width Tuning of Neurons in the Cat's Primary Visual Cortex. *Journal of Neurophysiology* **71**:347-374 <https://doi.org/10.1152/jn.1994.71.1.347> | [PubMed](#)
- Dorn Anja L, Yuan Kexin, Barker Alison J, Schreiner Christoph E, Froemke Robert C** (2010) Developmental Sensory Experience Balances Cortical Excitation and Inhibition. *Nature* **465**:932-936 <https://doi.org/10.1038/nature09119> | [PubMed](#)
- Duan Zhe Ran S, et al.** (2020) GABAergic Restriction of Network Dynamics Regulates Interneuron Survival in the Developing Cortex. *Neuron* **105**:75-92.e5 <https://doi.org/10.1016/j.neuron.2019.10.008> | [PubMed](#)
- Eckmann Samuel, Young Edward James, Gjorgjieva Julijana** (2024) Synapse-Type-Specific Competitive Hebbian Learning Forms Functional Recurrent Networks. *Proceedings of the National Academy of Sciences* **121**:e2305326121 <https://doi.org/10.1073/pnas.2305326121> | [PubMed](#)

- Froemke Robert C**, Merzenich Michael M, Schreiner Christoph E (2007) A Synaptic Memory Trace for Cortical Receptive Field Plasticity. *Nature* **450**:425-429 <https://doi.org/10.1038/nature06289> | [PubMed](#)
- Goldberg Joshua A**, Rokni Uri, Sompolinsky Haim (2004) Patterns of Ongoing Activity and the Functional Architecture of the Primary Visual Cortex. *Neuron* **42**:489-500 [https://doi.org/10.1016/S0896-6273\(04\)00197-7](https://doi.org/10.1016/S0896-6273(04)00197-7) | [PubMed](#)
- Gong Ziyi**, Brunel Nicolas (2024) Inhibitory Plasticity Enhances Sequence Storage Capacity and Retrieval Robustness. *bioRxiv* <https://doi.org/10.1101/2024.04.08.588573>
- Gonzalez Kevin C**, Negrean Adrian, Liao Zhenrui, Polleux Franck, Losonczy Attila (2023) Synaptic Basis of Behavioral Timescale Plasticity. *bioRxiv* <https://doi.org/10.1101/2023.10.04.560848>
- Hafizi Hadi**, et al. (2022) Inhibition-Dominated Rich-Club Shapes Dynamics in Cortical Microcircuits. *bioRxiv* <https://doi.org/10.1101/2021.05.07.443074>
- Hanganu-Opatz Ileana L**, et al. (2021) The Logic of Developing Neocortical Circuits in Health and Disease. *Journal of Neuroscience* **41**:813-822 <https://doi.org/10.1523/JNEUROSCI.1655-20.2020> | [PubMed](#)
- Harth Philipp**, et al. (2025) Dissecting Origins of Wiring Specificity in Dense Reconstructions of Neural Tissue. *bioRxiv* <https://doi.org/10.1101/2024.12.14.628490>
- Hawkes Alan G** (1971) Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika* **58**:83-90 <https://doi.org/10.2307/2334319>
- Jiang Xiaolong**, et al. (2015) Principles of Connectivity among Morphologically Defined Cell Types in Adult Neocortex. *Science* **350**:aac9462 <https://doi.org/10.1126/science.aac9462> | [PubMed](#)
- Jin Dezhe** (2002) Fast Convergence of Spike Sequences to Periodic Patterns in Recurrent Networks. *Physical Review Letters* **89** <https://doi.org/10.1103/PhysRevLett.89.208102> | [PubMed](#)
- Keller Andreas J**, et al. (2020) A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex. *Neuron* <https://doi.org/10.1016/j.neuron.2020.11.013> | [PubMed](#)
- Kempter Richard**, Gerstner Wulfram, Leo van Hemmen J (1999) Hebbian Learning and Spiking Neurons. *Physical Review E* **59**:4498-4514 <https://doi.org/10.1103/PhysRevE.59.4498>
- Khona Mikail**, Fiete Ila R (2022) Attractor and Integrator Networks in the Brain. *Nature Reviews Neuroscience* <https://doi.org/10.1038/s41583-022-00642-0> | [PubMed](#)
- Kim Sung Soo**, Hermundstad Ann M, Romani Sandro, Abbott LF, Jayaraman Vivek (2019) Generation of Stable Heading Representations in Diverse Visual Scenes. *Nature* **576**:126-131 <https://doi.org/10.1038/s41586-019-1767-1> | [PubMed](#)
- Kim Sung Soo**, Hervé Rouault, Druckmann Shaul, Jayaraman Vivek (2017) Ring Attractor Dynamics in the Drosophila Central Brain. *Science* **356**:849-853 <https://doi.org/10.1126/science.aal4835> | [PubMed](#)
- Kleberg Florence I**, Fukai Tomoki, Gilson Matthieu (2014) Excitatory and Inhibitory STDP Jointly Tune Feedforward Neural Circuits to Selectively Propagate Correlated Spiking Activity. *Frontiers in Computational Neuroscience* **8** <https://doi.org/10.3389/fncom.2014.00053> | [PubMed](#)
- Kuan Aaron T**, et al. (2024) Synaptic Wiring Motifs in Posterior Parietal Cortex Support Decision-Making. *Nature* 1-7 <https://doi.org/10.1038/s41586-024-07088-7> | [PubMed](#)
- Kunkel Susanne**, Diesmann Markus, Morrison Abigail (2011) Limits to the Development of Feed-Forward Structures in Large Recurrent Neuronal Networks. *Frontiers in Computational Neuroscience* **4** <https://doi.org/10.3389/fncom.2010.00160> | [PubMed](#)
- Lagzi Fereshteh**, Bustos Martha Canto, Oswald Anne-Marie, Doiron Brent (2021) Assembly Formation Is Stabilized by Parvalbumin Neurons and Accelerated by Somatostatin Neurons. *bioRxiv* <https://doi.org/10.1101/2021.09.06.459211>

- Lagzi Fereshteh**, Fairhall Adrienne L (2024) Emergence of Co-Tuning in Inhibitory Neurons as a Network Phenomenon Mediated by Randomness, Correlations, and Homeostatic Plasticity. *Science Advances* **10**:eadi4350 <https://doi.org/10.1126/sciadv.adi4350> | PubMed
- Lakhera Shreya**, Herbert Elizabeth, Gjorgjieva Julijana (2024) Modeling the Emergence of Circuit Organization and Function during Development. *Cold Spring Harbor Perspectives in Biology* a041511 <https://doi.org/10.1101/cshperspect.a041511> | PubMed
- Luz Yotam**, Shamir Maoz (2014) The Effect of STDP Temporal Kernel Structure on the Learning Dynamics of Single Excitatory and Inhibitory Synapses. *PLOS One* **9**:e101109 <https://doi.org/10.1371/journal.pone.0101109> | PubMed
- Mackwood Owen**, Naumann Laura B, Sprekeler Henning (2021) Learning Excitatory-Inhibitory Neuronal Assemblies in Recurrent Networks. *eLife* **10**:e59715 <https://doi.org/10.7554/eLife.59715> | PubMed
- Maes Amadeus**, Barahona Mauricio, Clopath Claudia (2021) Learning Compositional Sequences with Multiple Time Scales through a Hierarchical Network of Spiking Neurons. *PLOS Computational Biology* **17**:e1008866 <https://doi.org/10.1371/journal.pcbi.1008866> | PubMed
- Manz Paul**, Memmesheimer Raoul-Martin (2023) Purely STDP-based Assembly Dynamics: Stability, Learning, Overlaps, Drift and Aging. *PLOS Computational Biology* **19**:e1011006 <https://doi.org/10.1371/journal.pcbi.1011006> | PubMed
- Mastrogiuseppe Francesca**, Ostojic Srdjan (2017) Intrinsically-Generated Fluctuating Activity in Excitatory-Inhibitory Networks. *PLOS Computational Biology* **13**:e1005498 <https://doi.org/10.1371/journal.pcbi.1005498> | PubMed
- Meissner-Bernard Claire**, Jenkins Bethan, et al. (2025) Computational Functions of Precisely Balanced Neuronal Microcircuits in an Olfactory Memory Network. *Cell Reports* **44** <https://doi.org/10.1016/j.celrep.2025.115330> | PubMed
- Meissner-Bernard Claire**, Zenke Friedemann, Friedrich Rainer W (2025) Geometry and Dynamics of Representations in a Precisely Balanced Memory Network Related to Olfactory Cortex. *eLife* **13**:RP96303 <https://doi.org/10.7554/eLife.96303> | PubMed
- Miehl Christoph**, Gjorgjieva Julijana (2022) Stability and Learning in Excitatory Synapses by Nonlinear Inhibitory Plasticity. *PLOS Computational Biology* **18**:e1010682 <https://doi.org/10.1371/journal.pcbi.1010682> | PubMed
- Miehl Christoph**, Onasch Sebastian, Festa Dylan, Gjorgjieva Julijana (2023) Formation and Computational Implications of Assemblies in Neural Circuits. *The Journal of Physiology* **601**:3071-3090 <https://doi.org/10.1113/JP282750> | PubMed
- Mongillo Gianluigi**, Rumpel Simon, Loewenstein Yonatan (2018) Inhibitory Connectivity Defines the Realm of Excitatory Plasticity. *Nature Neuroscience* **21**:1463-1470 <https://doi.org/10.1038/s41593-018-0226-x> | PubMed
- Montangie Lisandro**, Miehl Christoph, Gjorgjieva Julijana (2020) Autonomous Emergence of Connectivity Assemblies via Spike Triplet Interactions. *PLOS Computational Biology* **16**:e1007835 <https://doi.org/10.1371/journal.pcbi.1007835> | PubMed
- Morrison A**, Diesmann M, Gerstner W (2008) Phenomenological Models of Synaptic Plasticity Based on Spike Timing. *Biological cybernetics* **98**:459-478 <https://doi.org/10.1007/s00422-008-0233-1> | PubMed
- Morrison Abigail**, Aertsen Ad, Diesmann Markus (2007) Spike-Timing-Dependent Plasticity in Balanced Random Networks. *Neural Computation* **19**:1437-1467 <https://doi.org/10.1162/neco.2007.19.6.1437> | PubMed
- Mulholland Haleigh N**, Hein Bettina, Kaschube Matthias, Smith Gordon B (2021) Tightly Coupled Inhibitory and Excitatory Functional Networks in the Developing Primary Visual Cortex. *eLife* **10**:e72456 <https://doi.org/10.7554/eLife.72456> | PubMed

- Ocker Gabriel K, et al. (2017) From the Statistics of Connectivity to the Statistics of Spike Times in Neuronal Networks. *Current Opinion in Neurobiology* **46**:109-119 <https://doi.org/10.1016/j.conb.2017.07.011> | PubMed
- Ocker Gabriel Koch, Doiron Brent (2019) Training and Spontaneous Reinforcement of Neuronal Assemblies by Spike Timing Plasticity. *Cerebral Cortex* **29**:937-951 <https://doi.org/10.1093/cercor/bhy001> | PubMed
- Ocker Gabriel Koch, Litwin-Kumar Ashok, Doiron Brent (2015) Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses. *PLOS Comput Biol* **11**:e1004458 <https://doi.org/10.1371/journal.pcbi.1004458> | PubMed
- Papadimitriou Christos H, Vempala Santosh S, Mitropolsky Daniel, Collins Michael, Maass Wolfgang (2020) Brain Computation by Assemblies of Neurons. *Proceedings of the National Academy of Sciences* **117**:14464-14472 <https://doi.org/10.1073/pnas.2001893117> | PubMed
- Peng Yangfan, et al. (2021) Spatially Structured Inhibition Defined by Polarized Parvalbumin Interneuron Axons Promotes Head Direction Tuning. *Science Advances* **7**:eabg4693 <https://doi.org/10.1126/sciadv.abg4693> | PubMed
- Petrucchio Luigi, et al. (2023) Neural Dynamics and Architecture of the Heading Direction Circuit in Zebrafish. *Nature Neuroscience* 1-9 <https://doi.org/10.1038/s41593-023-01308-5> | PubMed
- Powell Nathaniel J, et al. (2024) Common Modular Architecture across Diverse Cortical Areas in Early Development. *Proceedings of the National Academy of Sciences* **121**:e2313743121 <https://doi.org/10.1073/pnas.2313743121> | PubMed
- Tannenbaum Neta Ravid, Burak Yoram (2016) Shaping Neural Circuits by High Order Synaptic Interactions. *PLOS Computational Biology* **12**:e1005056 <https://doi.org/10.1371/journal.pcbi.1005056> | PubMed
- Reimann Michael W, Santander Daniela Egas, Ecker Andr s, Muller Eilif B (2023) Specific Inhibition and Disinhibition in the Higher-Order Structure of a Cortical Connectome. *bioRxiv* <https://doi.org/10.1101/2023.12.22.573036>
- Riquelme Juan Luis, Hemberger Mike, Laurent Gilles, Gjorgjieva Julijana (2023) Single Spikes Drive Sequential Propagation and Routing of Activity in a Cortical Network. *eLife* **12**:e79928 <https://doi.org/10.7554/eLife.79928> | PubMed
- Rubin Ran, Abbott LF, Sompolinsky Haim (2017) Balanced Excitation and Inhibition Are Required for High-Capacity, Noise-Robust Neuronal Selectivity. *Proceedings of the National Academy of Sciences* **201705841** <https://doi.org/10.1073/pnas.1705841114> | PubMed
- Sadeh Sadra, Clopath Claudia (2020) Theory of Neuronal Perturbome in Cortical Networks. *Proceedings of the National Academy of Sciences* **117**:26966-26976 <https://doi.org/10.1073/pnas.2004568117> | PubMed
- Sadeh Sadra, Clopath Claudia (2021) Excitatory-Inhibitory Balance Modulates the Formation and Dynamics of Neuronal Assemblies in Cortical Networks. *Science Advances* **7**:eabg8411 <https://doi.org/10.1126/sciadv.abg8411> | PubMed
- Schulz Auguste, Miehl Christoph, Michael J Berry II, Gjorgjieva Julijana (2021) The Generation of Cortical Novelty Responses through Inhibitory Plasticity. *eLife* **10**:e65309 <https://doi.org/10.7554/eLife.65309> | PubMed
- Shao Shuai, Riquelme Juan Luis, Gjorgjieva Julijana (2024) Spontaneous Emergence and Drifting of Sequential Neural Activity in Recurrent Networks. *bioRxiv* 2024.09.27.615499 <https://doi.org/10.1101/2024.09.27.615499>
- Shulz DE, Feldman DE (2013) Spike Timing-Dependent Plasticity. In: Rubenstein JLR, Rakic P (Eds). *Neural Circuit Development and Function in the Brain* Elsevier. pp. 155-181 <https://doi.org/10.1016/B978-0-12-397267-5.00029-7>
- Sillito AM, Jones HE (1996) Context-Dependent Interactions and Visual Processing in V1. *Journal of Physiology-Paris* **90**:205-209 [https://doi.org/10.1016/S0928-4257\(97\)81424-6](https://doi.org/10.1016/S0928-4257(97)81424-6) | PubMed

- Smith Gordon B**, Hein Bettina, Whitney David E, Fitzpatrick David, Kaschube Matthias (2018) Distributed Network Interactions and Their Emergence in Developing Neocortex. *Nature Neuroscience* **21**:1600-1608 <https://doi.org/10.1038/s41593-018-0247-5> | PubMed
- Song Sen**, Miller Kenneth D, Abbott LF (2000) Competitive Hebbian Learning through Spike-Timing-Dependent Synaptic Plasticity. *Nature Neuroscience* **3**:919-926 <https://doi.org/10.1038/78829> | PubMed
- Sreenivasan Varun**, et al. (2022) Input-Specific Control of Interneuron Numbers in Nascent Striatal Networks. *Proceedings of the National Academy of Sciences* **119**:e2118430119 <https://doi.org/10.1073/pnas.2118430119> | PubMed
- Stimberg Marcel**, Brette Romain, Goodman Dan FM (2019) Brian 2, an Intuitive and Efficient Neural Simulator. *eLife* **8**:e47314 <https://doi.org/10.7554/eLife.47314> | PubMed
- Trägenap Sigrid**, Whitney David E, Fitzpatrick David, Kaschube Matthias (2025) The Developmental Emergence of Reliable Cortical Representations. *Nature Neuroscience* **28**:394-405 <https://doi.org/10.1038/s41593-024-01857-3> | PubMed
- Tremblay Robin**, Lee Soohyun, Rudy Bernardo (2016) GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron* **91**:260-292 <https://doi.org/10.1016/j.neuron.2016.06.033> | PubMed
- Triplett Marcus A**, Avitan Lilach, Goodhill Geoffrey J (2018) Emergence of Spontaneous Assembly Activity in Developing Neural Networks without Afferent Input. *PLOS Computational Biology* **14**:e1006421 <https://doi.org/10.1371/journal.pcbi.1006421> | PubMed
- Tsodyks MV**, Sejnowski T (1995) Rapid State Switching in Balanced Cortical Network Models. *Network: Computation in Neural Systems* **6**:111 <https://doi.org/10.1088/0954-898X/6/2/001>
- Udvary Daniel**, et al. (2022) The Impact of Neuron Morphology on Cortical Network Architecture. *Cell Reports* **39**:110677 <https://doi.org/10.1016/j.celrep.2022.110677> | PubMed
- Gaëtan Vignoud**, Robert Philippe (2022) Spontaneous Dynamics of Synaptic Weights in Stochastic Models with Pair-Based Spike-Timing-Dependent Plasticity. *Physical Review E* **105**:054405 <https://doi.org/10.1103/PhysRevE.105.054405>
- Vogels Tim P**, 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 <https://doi.org/10.1126/science.1211095> | PubMed
- Van Vreeswijk C**, Sompolinsky H (1998) Chaotic Balanced State in a Model of Cortical Circuits. *Neural Computation* **10**:1321-1371 <https://doi.org/10.1162/089976698300017214>
- Wallenstein Gene V**, Hasselmo Michael E (1997) GABAergic Modulation of Hippocampal Population Activity: Sequence Learning, Place Field Development, and the Phase Precession Effect. *Journal of Neurophysiology* **78**:393-408 <https://doi.org/10.1152/jn.1997.78.1.393> | PubMed
- Wu Sherry Jingjing**, et al. (2023) Cortical Somatostatin Interneuron Subtypes Form Cell-Type-Specific Circuits. *Neuron* **111**:2675-2692.e9 <https://doi.org/10.1016/j.neuron.2023.05.032> | PubMed
- Xie Xiaohui**, Hahnloser Richard HR, Seung H Sebastian (2002) Double-Ring Network Model of the Head-Direction System. *Physical Review E* **66**:041902 <https://doi.org/10.1103/PhysRevE.66.041902> | PubMed
- Znamenskiy Petr**, et al. (2024) Functional Specificity of Recurrent Inhibition in Visual Cortex. *Neuron* **112**:991-1000.e8 <https://doi.org/10.1016/j.neuron.2023.12.013> | PubMed
- Hawkes Alan G** (1971) Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika* **58**:83-90 <https://doi.org/10.2307/2334319>
- Jovanović Stojan**, Hertz John, Rotter Stefan (2015) Cumulants of Hawkes Point Processes. *Physical Review E* **91**:042802 <https://doi.org/10.1103/PhysRevE.91.042802> | PubMed
- Kempter Richard**, Gerstner Wulfram, Leo van Hemmen J (1999) Hebbian Learning and Spiking Neurons. *Physical Review E* **59**:4498-4514 <https://doi.org/10.1103/PhysRevE.59.4498>

- Mastrogiuseppe Francesca, Ostojic Srdjan (2017) Intrinsically-Generated Fluctuating Activity in Excitatory-Inhibitory Networks. *PLOS Computational Biology* **13**:e1005498 <https://doi.org/10.1371/journal.pcbi.1005498> | PubMed
- Pernice Volker, Staude Benjamin, Cardanobile Stefano, Rotter Stefan (2011) How Structure Determines Correlations in Neuronal Networks. *PLOS Computational Biology* **7**:e1002059 <https://doi.org/10.1371/journal.pcbi.1002059> | PubMed
- Trousdale James, Hu Yu, Shea-Brown Eric, Josić Krešimir (2012) Impact of Network Structure and Cellular Response on Spike Time Correlations. *PLoS Comput Biol* **8**:e1002408 <https://doi.org/10.1371/journal.pcbi.1002408> | PubMed
- Van Vreeswijk C, Sompolinsky H (1998) Chaotic Balanced State in a Model of Cortical Circuits. *Neural Computation* **10**:1321-1371 <https://doi.org/10.1162/089976698300017214>

Peer reviews

Reviewer #1 (Public review):

Summary:

Festa et al. provide a detailed analysis of the outcome of spike-timing-dependent plasticity acting on inhibitory synapses for distinct shapes of the kernel that governs how pre- and postsynaptic spike times induce synaptic changes. The authors investigate symmetric and asymmetric kernels, providing a theoretical description of the ingredients that give rise to rate- or covariance-dominated plasticity based on a simplified two-neuron circuit. These analyses are confirmed via simulations of large recurrent networks with random excitatory connectivity. For excitatory connections arranged in a one-dimensional ring, the authors show that two distinct classes of inhibitory neurons (distinguished by their plasticity rules) form an effective Mexican-hat weight profile. Furthermore, the authors show that external inhibition of one of the inhibitory neuron types gives rise to the phenomenon of surround modulation.

Strengths:

The analytical description of the two-neuron circuit is robust and accurately captures the qualitative evolution of inhibitory weights in the recurrent network with random excitatory connectivity. The emergence of the Mexican hat from the combination of distinct inhibitory synaptic plasticity rules acting on different neuron types is an important result that reveals how such connectivity can be learned in biologically plausible networks. All the analyses are well done, and the simulation results are convincing, which supports a robust interpretation of the findings.

Weaknesses:

The two-neuron circuit model is a good choice for the analytics, but it may have hidden a covariance effect of the "rate-dominated" symmetric spike-based kernel that would appear when several inhibitory neurons, each sharing a different spike correlation with the postsynaptic neuron, converge onto it. The rate homeostasis achieved by the rate-dominated model arises from adjusting inhibitory weights according to their initial correlation with the output neuron, so that after learning, the weights are distributed such that these correlations are cancelled out (Vogels et al., 2011). In other words, even the rate-dominated rule is covariance-driven under the hood: with a single inhibitory input, the two-neuron circuit cannot expose this, but with several differently correlated inputs, the covariance dependence should reappear.

It is unclear whether the distribution of inhibitory weights has stabilised after 25 minutes of simulation time (Figure 3C), given that a considerable proportion of (mutual) weights reach

the maximum allowed weight while (unidirectional) weights appear to vanish. Without a maximum-weight bound, and given sufficiently long simulations, the weights might diverge to infinity or decay to zero, so the apparent stationarity may be imposed by the bound rather than reflecting a true steady state. This could also be a finite-size effect, given the small number of excitatory connections per neuron.

The connections from excitatory neurons to the two inhibitory populations are different in the ring model (exc to PV is wider than exc to SST according to Table 3), and it is not clear whether this width difference, rather than the plasticity rules themselves, is responsible for the emergence of the Mexican hat.

<https://doi.org/10.7554/eLife.111666.1.sa1>

Reviewer #2 (Public review):

Summary:

This study investigates how inhibitory synaptic plasticity can stabilize recurrent neural circuits while also shaping their functional connectivity. The authors analyze inhibitory spike-timing-dependent plasticity rules and show that different temporal kernels promote distinct E/I motifs, including reciprocal E/I connectivity and lateral inhibition. Using reduced circuit analyses and larger spiking network simulations, they demonstrate that inhibitory plasticity can generate structured effective connectivity, including Mexican-hat-like interactions in ring networks, while maintaining stable activity. The work therefore extends the view of inhibitory plasticity from a primarily homeostatic mechanism to one that may contribute to computationally useful circuit organization.

Strengths:

A major strength of the study is that it identifies a concrete mechanism by which the temporal shape of iSTDP rules determines the structure of learned inhibitory connectivity. The comparison between rules favoring reciprocal E/I motifs and those favoring "lateral" inhibition is shown across both reduced circuit models and larger spiking networks. The ring-network simulations further connect these learned motifs to circuit-level outcomes, including Mexican-hat-like effective connectivity, surround-suppression, and modular spontaneous activity.

Weaknesses:

The main limitations concern the extent to which the learned motifs are fully self-organized and how broadly the results generalize. In particular, the ring-network results rely on a pre-specified ring-like excitatory architecture and on two inhibitory populations with distinct plasticity rules, making it important to clarify which aspects of the Mexican-hat effective connectivity emerge from iSTDP itself. The conclusions would also be strengthened by intermediate plasticity rules. Finally, the ring-network simulations provide an interpretable proof of principle, but the authors should clarify whether the PV/SST effects depend on this specific architecture or would also arise in a more generic recurrent or cortex-like connectivity motif.

The authors largely achieve their aim of showing that inhibitory synaptic plasticity can provide structured stabilization of recurrent circuits. The results support this claim within the model framework by demonstrating that different temporal forms of iSTDP lead to distinct learned E/I motifs and can shape effective connectivity and cortical-like response patterns. However, the broader biological interpretation remains more suggestive because some results depend on specific assumptions for the network architecture and plasticity rules.

The work is likely to be valuable for researchers studying inhibitory plasticity, E/I balance, cortical circuit development, and biologically plausible learning because it provides a clear theoretical link between local inhibitory learning rules and circuit-level organization. The combination of analytically tractable motifs, spiking network simulations, and publicly available code makes the framework useful for future research.

The significance of the work lies not in showing that inhibitory plasticity can have functions beyond homeostatic stabilization, which has been established by previous theoretical and experimental studies, but in formalizing how the temporal form of iSTDP rules can bias the emergence of distinct E/I motifs. At present, the work identifies rules that are sufficient to generate these motifs in model networks, while the mapping of these rules onto specific interneuron types remains for future experimental testing.

<https://doi.org/10.7554/eLife.111666.1.sa0>

Author response:

We thank the editors for sending our work for review and the reviewers for their thorough and constructive evaluations. In response to their feedback, we will submit a revised version of the manuscript soon. Below, we provide clarification on several of the concerns raised and outline the changes planned for the revised manuscript.

Reviewer 1, weaknesses

(1) The two-neuron circuit model is a good choice for the analytics, but it may have hidden a covariance effect of the "rate-dominated" symmetric spike-based kernel that would appear when several inhibitory neurons, each sharing a different spike correlation with the postsynaptic neuron, converge onto it. The rate homeostasis achieved by the rate-dominated model arises from adjusting inhibitory weights according to their initial correlation with the output neuron, so that after learning, the weights are distributed such that these correlations are cancelled out (Vogels et al., 2011). In other words, even the rate-dominated rule is covariance-driven under the hood: with a single inhibitory input, the two-neuron circuit cannot expose this, but with several differently correlated inputs, the covariance dependence should reappear.

Reviewer 1 highlights that the rate-homeostatic rule by Vogels et al. (2011) also includes a covariance-dependent component. Thus, in a network with multiple excitatory units, differences in pre-postsynaptic correlations can drive a redistribution of weights, resulting in stronger inhibitory weights for higher correlations. Our two-neuron circuit, which contains only one plastic inhibitory input, cannot reveal this competitive effect. We note, however, that although in the Vogels rule the covariance-dependent term is non-zero, it is typically much smaller than the rate-dependent term (see Methods, Section 4.3). Consequently, covariance-dependent organization may emerge on a slower timescale (a similar effect was also shown by Lagzi and Fairhall 2024; <https://doi.org/10.1126/sciadv.adi4350>). In the revised manuscript, we will clarify this point and analyze small motifs with multiple excitatory units and heterogeneous correlations, focusing on both the timescale of weight redistribution and the resulting steady-state weights. We will also revise our interpretation of Supplementary Figure S9: within the simulated time window, the rate-dominated rule produces uniform inhibitory connectivity, but this does not exclude slower covariance-dependent reorganization.

(2) It is unclear whether the distribution of inhibitory weights has stabilised after 25 minutes of simulation time (Figure 3C), given that a considerable proportion of (mutual) weights reach the maximum allowed weight while (unidirectional) weights appear to vanish. Without a maximum-weight bound, and given sufficiently long simulations, the weights might diverge to infinity or decay to zero, so the apparent stationarity may be

imposed by the bound rather than reflecting a true steady state. This could also be a finite-size effect, given the small number of excitatory connections per neuron.

The reviewer raises the possibility that the apparent stationarity in Figure 3C is influenced by the imposed weight bounds. In the revised manuscript, we will discuss this point and present extended versions of the simulations in Figure 3 where we remove the weight bounds. We will also test whether the observed behavior depends on network size or excitatory connection density. We note that, under different external input regimes, the weights stabilize without reaching the hard bound (Figure S11), suggesting that saturation is not a necessary outcome.

(3) The connections from excitatory neurons to the two inhibitory populations are different in the ring model (exc to PV is wider than exc to SST according to Table 3), and it is not clear whether this width difference, rather than the plasticity rules themselves, is responsible for the emergence of the Mexican hat.

We recognize that we did not fully justify the parametrization used in the ring model. The pre-existing ring architecture determines the spatial correlations available to iSTDP and therefore contributes to the learned connectivity. In the revised manuscript, we will add control simulations in which the excitatory inputs to the PV and SST populations have either identical or markedly different widths, while the plasticity rules are kept fixed. These controls will allow us to assess the relative contributions of these factors to the formation of a Mexican-hat effective-connectivity profile.

Reviewer 2, weaknesses

(1) The main limitations concern the extent to which the learned motifs are fully self-organized and how broadly the results generalize. In particular, the ring-network results rely on a pre-specified ring-like excitatory architecture and on two inhibitory populations with distinct plasticity rules, making it important to clarify which aspects of the Mexican-hat effective connectivity emerge from iSTDP itself. The conclusions would also be strengthened by intermediate plasticity rules. Finally, the ring-network simulations provide an interpretable proof of principle, but the authors should clarify whether the PV/SST effects depend on this specific architecture or would also arise in a more generic recurrent or cortex-like connectivity motif.

This comment raises an important distinction between the components that are specified and those that emerge through plasticity. In our simulations, the excitatory architecture is fixed, whereas the initially weak and unstructured inhibitory-to-excitatory connections are learned through iSTDP. Thus, in the ring network, the spatial organization of excitation is prescribed, but the inhibitory connectivity and resulting Mexican hat-like effective connectivity emerge from the interaction of this architecture with the two iSTDP rules. Importantly, the central result that symmetric and antisymmetric rules promote distinct reciprocal and lateral E/I motifs is not restricted to the ring network, but is also observed in sparse randomly connected spiking networks and in networks with intrinsically generated irregular activity. The ring network is therefore used to demonstrate how these general motif-forming mechanisms can support specific circuit computations.

Our framework parametrizes a broader family of pairwise iSTDP rules rather than relying exclusively on isolated, preselected rules, allowing the contribution of rule shape and rate-dependent terms to be understood analytically. Intermediate or “mixed” plasticity rules, including those considered by Yang and Doiron (2026) <https://doi.org/10.1103/9nv2-y63v>, as well as additional network architectures, are valuable directions for extending the framework. We will revise the Discussion to clarify which components are prescribed, which emerge through plasticity, and which conclusions apply beyond the ring-network implementation.

(2) The authors largely achieve their aim [...] by demonstrating that different temporal forms of iSTDP lead to distinct learned E/I motifs and can shape effective connectivity and cortical-like response patterns. However, the broader biological interpretation remains more suggestive because some results depend on specific assumptions for the network architecture and plasticity rules.

This comment points to an important distinction between biological generality and mechanistic insight. Our models are deliberately simplified to isolate how the temporal structure of iSTDP interacts with internally generated correlations to select distinct E/I connectivity motifs, and to make this relationship analytically tractable. The resulting predictions are then reproduced in large conductance-based spiking networks with different connectivity structures and sources of irregular activity. Thus, although we do not claim that the specific biological implementations considered here capture the full diversity of cortical circuits, the conclusions are not restricted to a single minimal model or network architecture. Adding further biological detail would introduce additional parameters and architecture-specific assumptions, but would not by itself establish greater generality or provide the same mechanistic understanding. In the revised Discussion, we will clarify the distinction between the general mechanistic principles established by our framework and the more specific biological interpretations that remain to be tested.

(3) [...] At present, the work identifies rules that are sufficient to generate these motifs in model networks, while the mapping of these rules onto specific interneuron types remains for future experimental testing.

Our use of the PV and SST labels is intended as a biologically motivated implementation rather than as a universal assignment of plasticity rules to these interneuron classes. The symmetric and antisymmetric kernels were motivated by in vitro measurements from PV and SST interneurons in mouse orbitofrontal cortex, respectively (Lagzi et al., 2021; <https://doi.org/10.1101/2021.09.06.459211> ). Because inhibitory plasticity can vary across brain regions, developmental stages, and experimental conditions (Feldman, 2012; <https://doi.org/10.1016/j.neuron.2012.08.001> ), the general conclusion of our work concerns the mapping from the temporal structure of an iSTDP rule to the E/I motif it promotes, rather than a fixed correspondence between a particular rule and an interneuron identity.

Our models therefore establish more than the sufficiency of two isolated rules: the analytical framework explains how features of the plasticity kernel and rate-dependent terms determine whether reciprocal, lateral, or blanket inhibitory connectivity emerges. The specific association of these mechanisms with PV and SST interneurons in other circuits remains an experimentally testable prediction. We will clarify this distinction in the revised manuscript and emphasize that, once cell-type-specific plasticity rules are measured in a given circuit, the framework can predict the E/I connectivity motifs that those rules are expected to promote.

<https://doi.org/10.7554/eLife.111666.1.sa3>