Inhibition stabilization and paradoxical effects in recurrent neural networks with short-term plasticity

Yue Kris Wu[1, 2, *] and Julijana Gjorgjieva[1, 2]

[1] *School of Life Sciences, Technical University of Munich, Freising, Germany*
[2] *Max Planck Institute for Brain Research, Frankfurt, Germany*

(Dated: December 19, 2022)

Inhibition stabilization is considered a ubiquitous property of cortical networks, whereby inhibition controls network activity in the presence of strong recurrent excitation. In networks with fixed connectivity, an identifying characteristic of inhibition stabilization is that increasing (decreasing) excitatory input to the inhibitory population leads to a decrease (increase) in inhibitory firing, known as the paradoxical effect. However, population responses to stimulation are highly nonlinear, and drastic changes in synaptic strengths induced by short-term plasticity (STP) can occur on the timescale of perception. How neuronal nonlinearities and STP affect inhibition stabilization and the paradoxical effect is unclear. Using analytical calculations, we demonstrate that in networks with STP the paradoxical effect implies inhibition stabilization, but inhibition stabilization does not imply the paradoxical effect. Interestingly, networks with neuronal nonlinearities and STP can transition non-monotonically between inhibition-stabilization and non-inhibition-stabilization, and between paradoxically- and non-paradoxically-responding regimes with increasing excitatory activity. In summary, our work reveals the relationship between inhibition stabilization and the paradoxical effect in the presence of neuronal nonlinearity and STP, yielding several testable predictions.

I. INTRODUCTION

Cortical networks are typically characterized by inhibition stabilization, where inhibition is needed to keep network activity levels in biologically realistic ranges despite the presence of strong recurrent excitation [1]. Networks operating in the inhibition-stabilized regime are capable of performing various computations, including input amplification, response normalization, and network multistability [2–6]. In networks with fixed connectivity, a hallmark of inhibition stabilization is the paradoxical effect: an increase or a decrease of excitatory input to the inhibitory population respectively decreases or increases the inhibitory firing. Over the past decade, much effort has been made to identify the operating regime of cortical networks based on the paradoxical effect [1, 7, 8].

Yet, various aspects ranging from the network to the synaptic level can considerably affect network dynamics and the operating regime. First, if individual neurons in the network receive large excitatory and inhibitory currents which precisely cancel each other, the network operates in a balanced state characterized by a linear population response [9–11]. Recent work has argued that neuronal input-output functions are better characterized by supralinear functions, and networks with this type of nonlinearity can exhibit various nonlinear phenomena as observed in biology [12–14]. Second, synapses in the brain are highly dynamic as a result of different short-term plasticity (STP) mechanisms, operating on a timescale of milliseconds to seconds [15, 16]. Upon presynaptic stimulation, postsynaptic responses can either get depressed subject to short-term depression (STD) or facilitated subject toshort-term facilitation (STF). While short-term synaptic dynamics are widely observed in biological circuits, it is unclear how they interact with the neuronal nonlinearity to jointly determine the network operating regime. Here we ask how the neuronal nonlinearity and STP affect inhibition stabilization and the paradoxical effect.

To address this question, we determine the conditions for inhibition stabilization and the paradoxical effect in networks of excitatory and inhibitory neurons in the presence of STP with linear and supralinear population response functions. We find that irrespective of the neuronal nonlinearity, in networks with excitatory-to-excitatory (E-to-E) STD, inhibition stabilization does not necessarily imply the paradoxical effect, but the paradoxical effect implies inhibition stabilization. In contrast, in networks with static connectivity or networks with other STP mechanisms, inhibition stabilization and the paradoxical effect imply each other. Interestingly, neuronal nonlinearities and STP endow the network with unconventional behaviors. More specifically, in the presence of the neuronal nonlinearity and E-to-E STD, monotonically increasing excitatory activity can lead to non-monotonic transitions between non-inhibition-stabilization and inhibition-stabilization, as well as between non-paradoxically-responding and paradoxically-responding. In con-

clusion, our work reveals the impact of the neuronal nonlinearity and STP on inhibition stabilization and the paradoxical effect, and makes several predictions for future experiments.

II. RESULTS

To understand the relationship between inhibition stabilization and the paradoxical effect in recurrent neural networks with STP, we studied rate-based population models consisting of an excitatory (E) and an inhibitory (I) population. The dynamics of the network are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \left[J_{EE} r_E - J_{EI} r_I + g_E\right]_+^{\alpha_E} \quad (1)$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \left[J_{IE} r_E - J_{II} r_I + g_I\right]_+^{\alpha_I} \quad (2)$$

where r_E and r_I denote the firing rates of the excitatory and inhibitory population, τ_E and τ_I are the corresponding time constants, J_{AB} represents the synaptic strength from population B to population A, where A, $B \in \{E, I\}$, g_E and g_I are the external inputs to the respective populations. Finally, α_E and α_I are the exponents of the respective input-output functions. To investigate the impact of neuronal nonlinearities on inhibition stabilization and the paradoxical effect, we considered both threshold-linear networks ($\alpha_E = \alpha_I = 1$) as well as supralinear networks ($\alpha_E = \alpha_I > 1$). In the regime of positive r_E and r_I, threshold-linear networks behave as linear networks. In the following, we thus call them linear networks. Furthermore, while we keep our analysis for supralinear networks in a general form, we use $\alpha_E = \alpha_I = 2$ for the numerical simulations. Note that the neuronal nonlinearity in our study refers to the nonlinearity of population-averaged responses to input when no STP mechanisms are taken into account, which is fully determined by α_E and α_I.

We implemented short-term plasticity mechanisms based on the Tsodyks and Markram model (15; Supplemental Methods). For the sake of analytical tractability, we included one STP mechanism at a time. To investigate how inhibition stabilization is affected by the neuronal nonlinearity and STP, we computed the real part of the leading eigenvalue of the Jacobian matrix of the excitatory-to-excitatory subnetwork incorporating STP, and refer to it as the 'Inhibition Stabilization index' (IS index) (Supplemental Methods; Supplemental Table 1). A positive (negative) IS index implies that the network is in the IS (non-IS) regime. To reveal how inhibition stabilization changes with network activity and network connectivity, we investigated how the IS index changes with the excitatory activity r_E and the excitatory to excitatory connection strength J_{EE}. These two parameters, r_E and J_{EE}, are directly involved in the definition of the IS index (Supplemental Table 1).

A. INHIBITION STABILIZATION IN RECURRENT NEURAL NETWORKS WITH SHORT-TERM DEPRESSION AT E-TO-E SYNAPSES

We first examined inhibition stabilization for networks with E-to-E STD, evaluated at the fixed point of the system (Fig. 1a). The distinction between non-IS and IS is reflected in network responses to perturbations induced by injecting additional excitatory currents into excitatory neurons while inhibition is fixed. Networks initially in the non-IS regime return back to their initial activity level after a small transient perturbation to the excitatory activity when inhibition is fixed, whereas networks initially in the IS regime deviate from their initial activity (Fig. S1). For linear networks with E-to-E STD, if J_{EE} is less than 1, the network is always in the non-IS regime regardless of r_E (Fig. 1b). If J_{EE} is greater than 1, the network transitions from IS to non-IS with increasing r_E (Fig. 1b). In contrast, supralinear networks with E-to-E STD manifest different behaviors. When J_{EE} is large, the network first transitions from non-IS to IS, and then back to non-IS with increasing r_E (Fig. 1c, Fig. S1). When J_{EE} is small, the supralinear network stays in the non-IS regime for all values of r_E (Fig. 1c). To better understand the transition between non-IS and IS in the presence of neuronal nonlinearities and E-to-E STD, we investigated how the boundary between non-IS and IS, defined as 'IS boundary', changes with r_E (Supplemental Methods; Fig. S2). Mathematically, the IS boundary is determined by the recurrent excitatory-to-excitatory connection strength for different r_E at which the IS index is 0, denoted by J_{EE}^{IS}. By computing the derivative of J_{EE}^{IS} with respect to r_E (Supplemental Methods), we found that the derivative is always positive for linear networks with E-to-E STD, suggesting that the IS boundary increases with increasing r_E (Fig. 1d). Therefore, for networks with large J_{EE}, as r_E increases, only the transition from IS to non-IS is possible (Fig. 1b, d). In contrast, for supralinear networks with E-to-E STD,

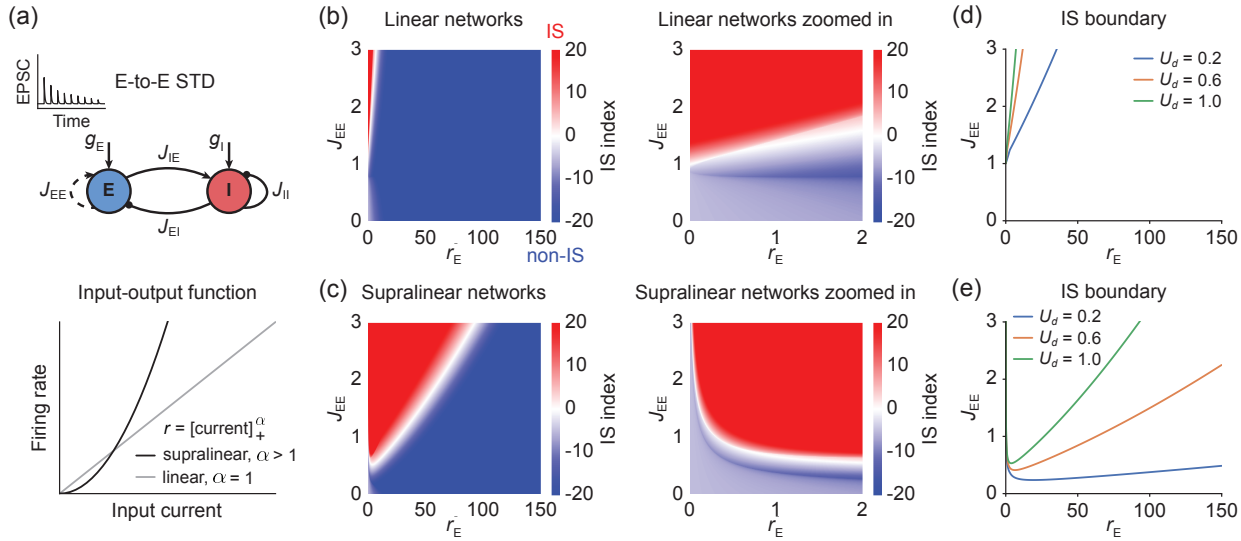

FIG. 1. Inhibition stabilization in recurrent neural networks with E-to-E short-term depression. (a) Top: Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STD. Bottom: Different nonlinearities of population-averaged responses to input. Linear (gray) and supralinear (black) input-output function given by a rectified power law with exponent $\alpha = 1$ and $\alpha = 2$ (cf. Eq. (1)-(2)), respectively. (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for linear networks with E-to-E STD. The non-IS and IS regime are marked in blue and red, respectively. Right: zoomed-in version of (b) left. Here, the depression rate U_d is 1.0 (Supplemental Methods), the firing rate r_E is plotted in a biologically realistic range from 0 to 150 Hz. (c) Same as (b) but for supralinear networks with E-to-E STD. Here, the depression rate U_d is 1.0 (Supplemental Methods). (d) IS boundary for linear networks with E-to-E STD, defined as the corresponding excitatory-to-excitatory connection strength J_{EE}^{IS} for different r_E at which the IS index is 0. Different colors represent the IS boundary for different values of STD variable U_d. (e) Same as (d) but for supralinear networks with E-to-E STD.

the derivative changes from negative to positive with increasing r_E (Supplemental Methods), implying that the IS boundary first decreases and then increases as r_E increases (Fig. 1e). As a result, networks can undergo non-monotonic transitions between non-IS and IS with increasing r_E. More specifically, networks can switch from non-IS to IS, and then back to non-IS with increasing r_E (Fig. 1c, e; Fig. S1). We found that these non-monotonic transitions cannot be observed in networks with static connectivity (Fig. S3; Supplemental Methods). Taken together, our results suggest that E-to-E STD can nontrivially affect the inhibition stabilization property, especially in the presence of neuronal nonlinearities.

B. INHIBITION STABILIZATION IN RECURRENT NEURAL NETWORKS WITH SHORT-TERM FACILITATION AT E-TO-E SYNAPSES

To determine if the observed effects are specific to the type of STP at E-to-E synapses, we next examined networks with E-to-E STF (Fig. 2a). Unlike the scenario with STD, for both linear networks or supralinear networks, only a monotonic transition from non-IS to IS is possible with increasing r_E in the presence of E-to-E STF (Fig. 2b, c). In contrast to supralinear networks, linear networks with J_{EE} larger than 1 are always in the IS regime regardless of r_E. In both cases, the parameter regime of J_{EE} and r_E which supports IS is much larger than in the corresponding network with STD (Fig. 2). Furthermore, independent of the neuronal nonlinearity, the derivative of J_{EE}^{IS} with respect to r_E is always negative (Supplemental Methods), indicating that the IS boundary decreases as r_E increases (Fig. 2d, e). These results indicate that E-to-E STF exerts a more intuitive influence on the inhibition stabilization property than E-to-E STD even in the presence of neuronal nonlinearities.

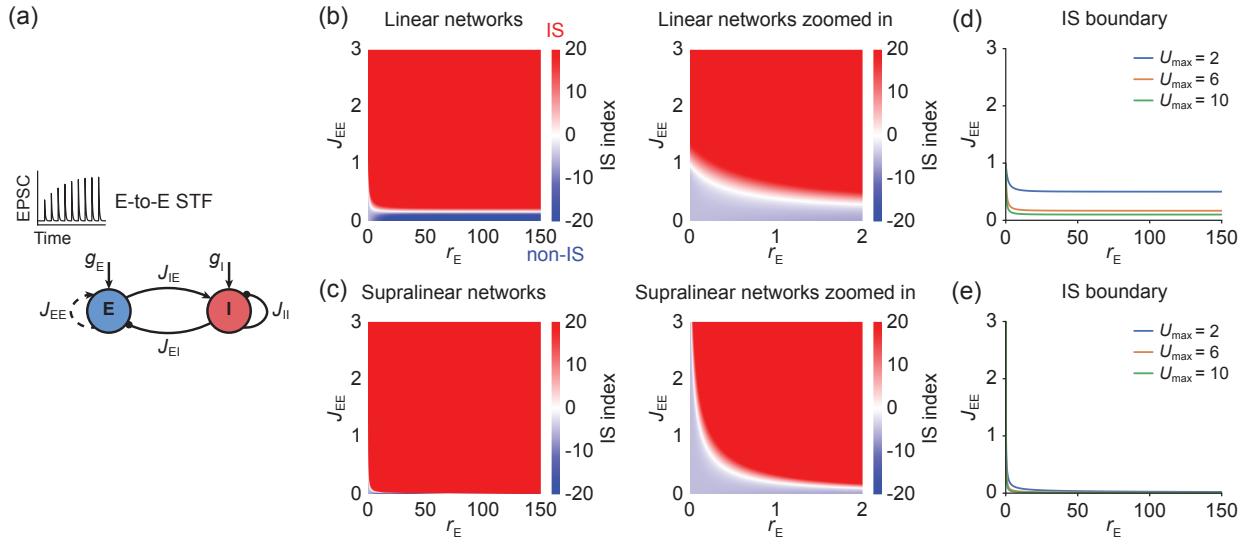

FIG. 2. Inhibition stabilization in recurrent neural networks with E-to-E short-term facilitation. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STF. (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for linear networks with E-to-E STF. The non-IS and IS regime are marked in blue and red, respectively. Right: zoomed-in version of (b) left. Here, the facilitation rate U_f is 1.0, and the maximal facilitation value U_{max} is 6.0 (Supplemental Methods). (c) Same as (b) but for supralinear networks with E-to-E STF. Here, the facilitation rate U_f is 1.0, and U_{max} is 6.0 (Supplemental Methods). (d) IS boundary for linear networks with E-to-E STF, defined as the corresponding excitatory-to-excitatory connection strength J_{EE}^{IS} for different r_E at which the IS index is 0. Different colors represent the IS boundary for different values of U_{max}. (e) Same as (d) but for supralinear networks with E-to-E STF.

C. INHIBITION STABILIZATION IN RECURRENT NEURAL NETWORKS WITH SHORT-TERM PLASTICITY AT OTHER SYNAPSES

Finally, we performed the same analyses for networks with different types of STP at all synapses other than E-to-E, including E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF, respectively (Fig. 3a; Supplemental Methods). Including these STP mechanisms does not change the IS condition relative to networks with fixed connectivity (Supplemental Methods). For networks with a linear input-output function, the IS boundary does not change with r_E (Fig. 3b, d; Supplemental Methods), and J_{EE} completely determines whether the network is non-IS or IS. In contrast, for networks with a supralinear input-output function, the derivative of J_{EE}^{IS} with respect to r_E is always negative, suggesting that the IS boundary decreases with increasing r_E (Fig. 3c, e; Supplemental Methods). Therefore, the transition from non-IS to IS with increasing r_E in static supralinear networks or supralinear networks with STP at all synapses other than E-to-E can only happen for large J_{EE} (Fig. 3c, e; Fig. S3). No transition between non-IS and IS can occur in the biological realistic firing regime from 0 to 150 Hz for small J_{EE} (Fig. 3c, e).

In summary, by considering all possible STP mechanisms, our results demonstrate a nontrivial influence of the neuronal nonlinearity and STP on inhibition stabilization. Specifically, we revealed how inhibition stabilization changes with network activity and network connectivity when considering neuronal nonlinearities and STP.

D. PARADOXICAL EFFECTS IN RECURRENT NEURAL NETWORKS WITH SHORT-TERM PLASTICITY

Previous theoretical studies have suggested that in excitatory and inhibitory networks with static connectivity one identifying characteristic of inhibition stabilization is that injecting excitatory (inhibitory) current into inhibitory neurons decreases (increases) inhibitory firing rates, known as the paradoxical effect [4, 17]. Here, we sought to identify the conditions under which a paradoxical effect can arise in recurrent neural networks with STP. We assumed that the system is stable locally around

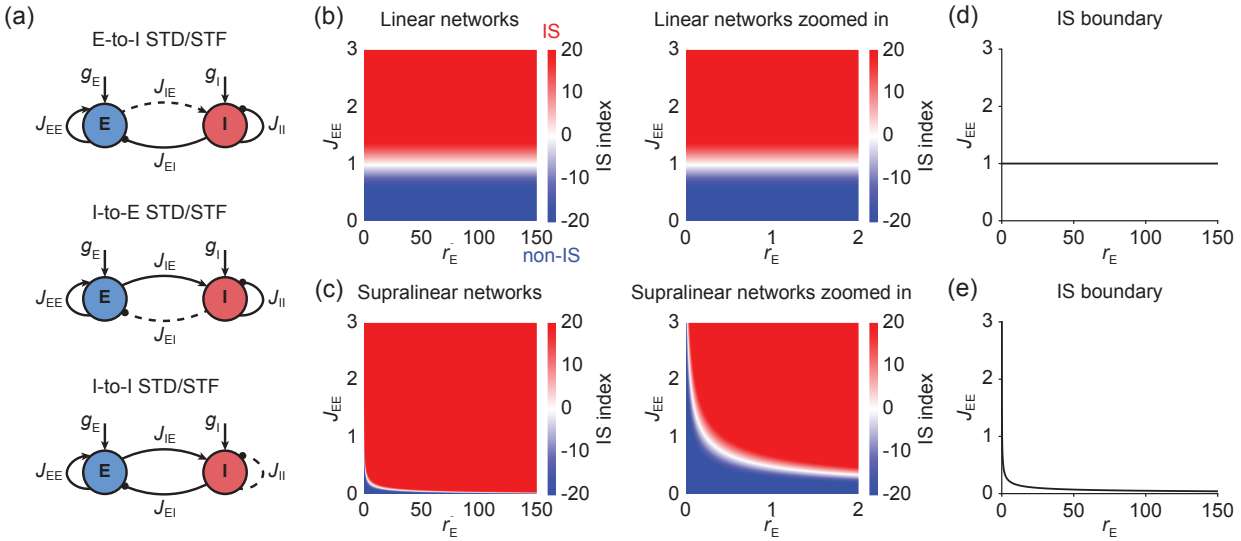

FIG. 3. Inhibition stabilization in recurrent neural networks with E-to-I short-term depression/facilitation, I-to-E short-term depression/facilitation, or I-to-I short-term depression/facilitation. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-I STD/STF (top), with I-to-E STD/STF (middle), or with I-to-I STD/STF (bottom). (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for linear networks with E-to-I STD/STF, with I-to-E STD/STF, or with I-to-I STD/STF. The non-IS and IS regime are marked in blue and red, respectively. Right: zoomed-in version of (b) left. (c) Same as (b) but for supralinear networks with E-to-I STD/STF, with I-to-E STD/STF, or with I-to-I STD/STF. (d) IS boundary for linear networks with E-to-I STD/STF, with I-to-E STD/STF, or with I-to-I STD/STF, defined as the corresponding excitatory-to-excitatory connection strength J_{EE}^{IS} for different r_E at which the IS index is 0. (e) Same as (d) but for supralinear networks with E-to-I STD/STF, with I-to-E STD/STF, or with I-to-I STD/STF.

the fixed point, in other words, a small transient perturbation to the system will not lead to deviation from the initial fixed point over time. Furthermore, the perturbation used to probe the paradoxical effect (e.g. the excitatory current injected to the inhibitory population) is small enough that it will not lead to a transition between non-IS and IS. To determine the conditions for the presence of the paradoxical effect under these assumptions, we considered the phase plane of the excitatory (x-axis) and inhibitory (y-axis) firing rate dynamics. The first condition involves a positive slope of the excitatory nullcline around the fixed point in the phase plane, while the second condition involves a larger slope of the inhibitory nullcline than of the excitatory nullcline locally around the fixed point (17; Fig. S4; Supplemental Methods). We compared the above two conditions for the presence of the paradoxical effect with the conditions to be in the IS regime. We found that irrespective of the shape of the neuronal nonlinearity, in networks with E-to-E STD, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not imply the paradoxical effect (Fig. 4a; Supplemental Methods). In contrast, for networks with E-to-E STF, E-to-I STD/STF, I-to-I STD/STF, and I-to-I STD/STF, inhibition stabilization and the paradoxical effect imply each other (Fig. 4a; Supplemental Methods). To highlight the difference between inhibition stabilization and the paradoxical effect in networks with E-to-E STD (Fig. 4b), we plotted the paradoxical effect boundary that separates the paradoxically-responding and the non-paradoxically-responding regime together with the IS boundary for both linear networks and supralinear networks with E-to-E STD (Fig. 4c-f). In the two dimensional r_E-J_{EE} plane, the parameter regime for the paradoxical effect is much narrower than the IS regime, suggesting that there is a large parameter space, in which inhibition-stabilized networks do not exhibit the paradoxical effect (Fig. 4c-f).

Furthermore, by analyzing how the paradoxical effect boundary changes with r_E, we found that it exhibits a similar trend to the IS boundary (Supplemental Methods). In particular, the paradoxical effect boundary of supralinear networks with E-to-E STD is also a non-monotonic

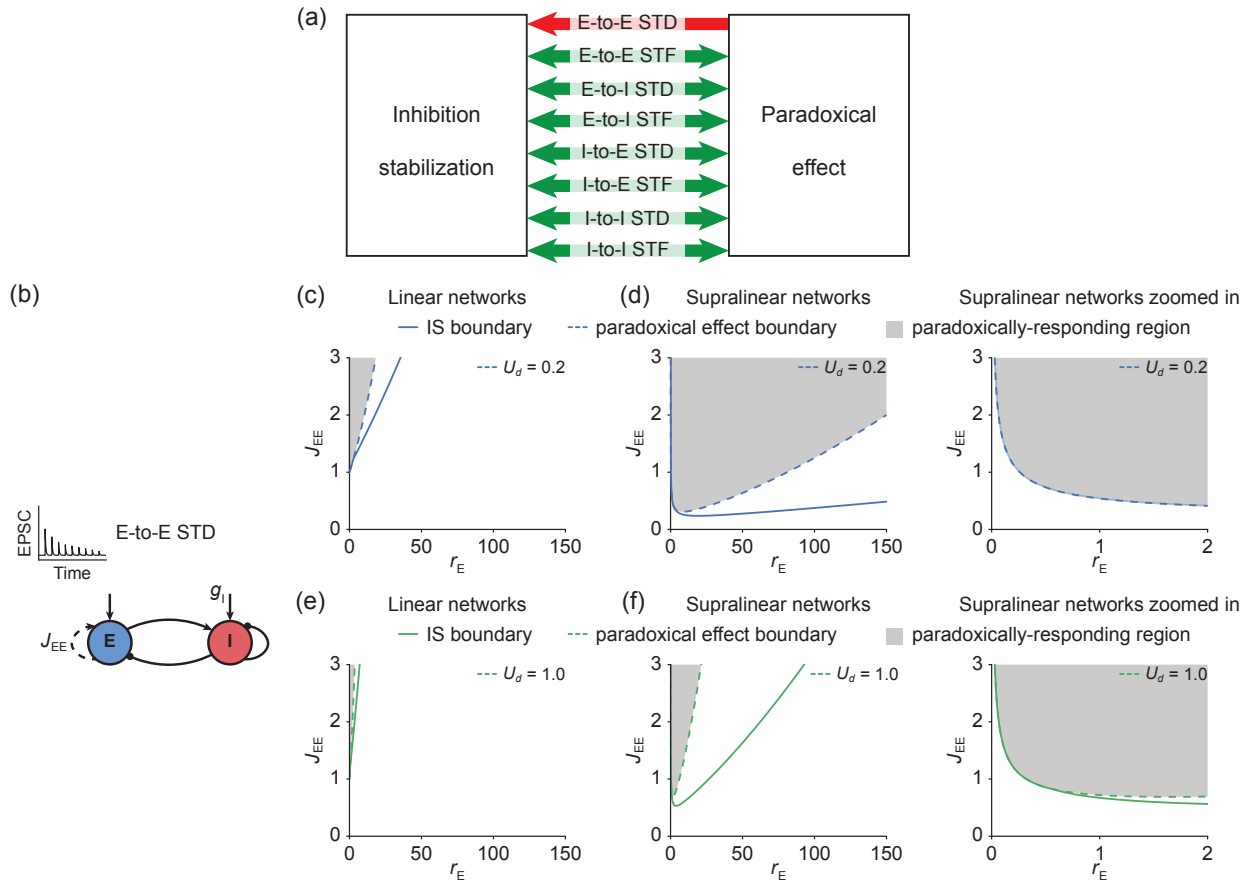

FIG. 4. Relationship between inhibition stabilization and the paradoxical effect in recurrent neural networks with short-term plasticity. (a) In networks with E-to-E STD, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not imply the paradoxical effect indicated by the unidirectional red arrow. In contrast, for networks with E-to-E STF, E-to-I STD/STF, I-to-I STD/STF, and I-to-I STD/STF, inhibition stabilization and the paradoxical effect imply each other indicated by the bidirectional green arrows. (b) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STD. (c) An example of the paradoxical effect boundary (dashed line) and the inhibition stabilization boundary (solid line) as a function of excitatory firing rate r_E for linear networks with E-to-E STD. The paradoxically-responding region is marked in gray. Here, the depression rate U_d is 0.2 (Supplemental Methods). (d) Same as (c) but for supralinear networks with E-to-E STD. Right: zoomed-in version of (d) left. (e-f) Same as (c-d) but the depression rate U_d 1.0.

function of r_E. Therefore, in this case, networks can undergo non-monotonic transitions between the paradoxically-responding regime and non-paradoxically-responding regime with monotonically changing excitatory activity r_E. Taken together, our results indicate that the relationship between inhibition stabilization and paradoxical effect in networks with STP becomes non-trivial in the presence of short-term plasticity.

III. DISCUSSION

In this study, we combined analytical and numerical methods to reveal the effects of neuronal nonlinearities and STP on inhibition stabilization, the paradoxical effect, and the relationship between them. Including STD at E-to-E synapses, in contrast to other types of STP and other synapse types, generates the most surprising results. Under these conditions, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not imply the paradoxical effect. For networks with other STP mechanisms and networks with static

connectivity, inhibition stabilization and the paradoxical effect imply each other. Additionally, in the presence of a neuronal nonlinearity and E-to-E STD, a non-monotonic transition between different regimes can occur when excitatory activity changes monotonically.

Recent studies have examined inhibition stabilization and the paradoxical effect in threshold-linear networks with E-to-E STP to demonstrate that inhibition stabilization can be probed by the paradoxical effect [1]. Recent work has conducted similar analysis for supralinear networks with short-term plasticity on specific synapses [18]. Here, we systematically analyzed networks with all forms of STP mechanisms, for both linear networks and supralinear networks. By mathematically defining the IS boundary and the paradoxical effect boundary, we further demonstrated how network activity and connectivity affect the inhibition stabilization property and the paradoxical effect.

In this work, we assumed that the network activity has reached a fixed point, and we did not consider scenarios like multistability or oscillations that could arise from neuronal nonlinearities or STP [14, 19, 20]. While multistability and oscillations have been observed in the brain [21, 22], the single stable fixed point assumed in our study is believed to be a reasonable approximation of awake sensory cortex [23]. Furthermore, to retain analytical tractability, here we only considered individual STP mechanisms one at a time. While classical STP models incorporate both STD and STF at the same synapses [15], depending on how the effective synaptic strength changes with activity, either the STD or the STF component dominates for a given activity level. Therefore, in principle, our theory approximates these scenarios when the dominant mechanism is known.

Beyond the single inhibitory population considered here, inhibitory neurons in the cortex are highly diverse in terms of anatomy, electrophysiology and function [24–26]. Recent studies have investigated inhibition stabilization and the paradoxical effect in networks with multiple interneuron subtypes in the absence of STP [27–31]. Yet, synapses between different interneuron subtypes exhibit considerable short-term dynamics [32]. It will be interesting for future work to consider STP in networks with different subclasses of interneurons, especially in the presence of neuronal nonlinearities.

Our model makes several predictions that can be tested in optogenetic experiments. Across cortical layers and across brain regions, synaptic strengths can differ by an order of magnitude [33]. Furthermore, the degrees of balance between excitation and inhibition may also vary [34, 35], resulting different neuronal nonlinearities [10, 12, 35]. Therefore, different behaviors predicted by our analysis may be observable in different neural circuits. For example, in the presence of E-to-E STD, our model shows that networks with weak excitatory to excitatory connection strength J_{EE} are always non-IS in the biologically realistic activity regime and therefore exhibit no paradoxical effects. In contrast, with E-to-E STD and strong J_{EE}, network models with a linear population-averaged response function can undergo the transition from IS to non-IS with increasing excitatory activity r_E. Different from linear networks, our model predicts that non-monotonic transitions between non-IS and IS can be found in supralinear networks. More specifically, supralinear networks can switch from non-IS to IS, and then from IS to non-IS with increasing r_E. Although inhibition stabilization does not imply the paradoxical effect in the presence of E-to-E STD, the transition between paradoxically-responding and non-paradoxically-responding regime is also non-monotonic with increasing r_E in supralinear networks, whereas in linear networks, only transitions from the paradoxically-responding to the non-paradoxically-responding regime with increasing r_E is possible. Therefore, depending on the excitation-inhibition balance, and the specific short term plasticity mechanisms operating in different brain regions, our work proposes that the circuits will exhibit different properties when interrogated with common experimental techniques.

Second, in the presence of STF on E-to-E synapses or STP on other synapses, our results demonstrate that inhibition stabilization and the paradoxical effect imply each other. Linear network models with J_{EE} larger than 1 that have E-to-E STF are always IS and thus exhibit the paradoxical effect. In linear network models with STP on other synapses, activity does not affect inhibition stabilization and the paradoxical effect. In contrast, regardless of the strength of J_{EE}, supralinear networks with E-to-E STF or STP on other synapses can switch from non-IS to IS with increasing r_E. This transition from non-IS and IS can be directly tested experimentally by probing the paradoxical effect, because of the equivalence of inhibition stabilization and the paradoxical effect found in network models with E-to-E STF or STP on other synapses.

Last, our analysis shows that in most cases substantially altering either J_{EE} or r_E can switch the network operating regime. Multiple factors can modify J_{EE} and r_E experimentally. On a short

timescale, the strength of sensory stimulation, and behavioral states like arousal [36], attention [37] and locomotion [38] can dramatically change activity levels r_E. Regime switching may therefore be experimentally observable across different stimulation conditions and different behavioral states. On a long timescale, J_{EE} or r_E can be modified by long term plasticity mechanisms [39, 40]. In this case, regime switching could be experimentally detectable across different developmental stages.

Taken together, our theoretical framework provides a systematic analysis of how short-term synaptic plasticity and response nonlinearities interact to determine the network operating regime, revealing unexpected relationships and their signatures as a guide for future experimental studies.

ACKNOWLEDGMENTS

We thank Elizabeth Herbert, Leonidas M. A. Richter, and Dylan Festa for comments on the manuscript. This work was supported by the Max Planck Society and has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 804824 to J.G.) and from the Deutsche Forschungsgemeinschaft in the Collaborative Research Centre 1080 (to J.G.). Y.K. W. is supported by the Add-on Fellowship of the Joachim Herz Foundation.

DATA AVAILABILTY

The code used for model simulations is available at GitHub `https://github.com/comp-neural-circuits/inhibition-stabilization-paradoxical-effect-STP`. All simulation parameters are listed in Supplemental Tables.

* kris.wu@tum.de

[1] A. Sanzeni, B. Akitake, H. C. Goldbach, C. E. Leedy, N. Brunel, and M. H. Histed, "Inhibition stabilization is a widespread property of cortical networks," *eLife*, vol. 9, pp. 1–39, 2020.

[2] D. J. Amit and N. Brunel, "Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex," *Cerebral Cortex*, vol. 7, no. 3, pp. 237–252, 1997.

[3] B. K. Murphy and K. D. Miller, "Balanced Amplification : A New Mechanism of Selective Amplification of Neural Activity Patterns," *Neuron*, vol. 61, no. 4, pp. 635–648, 2009.

[4] H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, and D. Ferster, "Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression," *Neuron*, vol. 62, no. 4, pp. 578–592, 2009.

[5] D. B. Rubin, S. D. VanHooser, and K. D. Miller, "The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex," *Neuron*, vol. 85, no. 2, pp. 402–417, 2015.

[6] S. Sadeh and C. Clopath, "Inhibitory stabilization and cortical computation," *Nature Reviews Neuroscience*, vol. 22, no. 1, pp. 21–37, 2021.

[7] S. Sadeh, R. A. Silver, T. D. Mrsic-Flogel, and D. R. Muir, "Assessing the role of inhibition in stabilizing neocortical networks requires large-scale perturbation of the inhibitory population," *Journal of Neuroscience*, vol. 37, no. 49, pp. 12050–12067, 2017.

[8] N. Li, S. Chen, Z. V. Guo, H. Chen, Y. Huo, H. K. Inagaki, G. Chen, C. Davis, D. Hansel, C. Guo, and K. Svoboda, "Spatiotemporal constraints on optogenetic inactivation in cortical circuits," *eLife*, vol. 8, pp. 1–31, 2019.

[9] C. van Vreeswijk and H. Sompolinsky, "Chaos in neuronal networks with balanced excitatory and inhibitory activity," *Science*, vol. 274, no. 5293, pp. 1724–1726, 1996.

[10] C. van Vreeswijk and H. Sompolinsky, "Chaotic Balanced State in a Model of Cortical Circuits," *Neural Computation*, vol. 10, no. 6, pp. 1321–1371, 1998.

[11] A. Renart, J. D. Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, "The Asynchronous State in Cortical Circuits," *Science*, vol. 327, no. January, pp. 587–590, 2010.

[12] Y. Ahmadian, D. B. Rubin, and K. D. Miller, "Analysis of the stabilized supralinear network," *Neural Computation*, vol. 25, no. 8, pp. 1994–2037, 2013.

[13] G. Hennequin, Y. Ahmadian, D. B. Rubin, M. Lengyel, and K. D. Miller, "The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability," *Neuron*, vol. 98, no. 4, pp. 846–860.e5, 2018.

[14] N. Kraynyukova and T. Tchumatchenko, "Stabilized supralinear network can give rise to bistable, oscillatory, and persistent activity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 13, pp. 3464–3469, 2018.

[15] M. V. Tsodyks and H. Markram, "The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 2, pp. 719–723, 1997.

[16] H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ail-

amaki, L. Alonso-Nanclares, N. Antille, S. Arsever, G. A. A. Kahou, T. K. Berger, A. Bilgili, N. Buncic, A. Chalimourda, G. Chindemi, J. D. Courcol, F. Delalondre, V. Delattre, S. Druckmann, R. Dumusc, J. Dynes, S. Eilemann, E. Gal, M. E. Gevaert, J. P. Ghobril, A. Gidon, J. W. Graham, A. Gupta, V. Haenel, E. Hay, T. Heinis, J. B. Hernando, M. Hines, L. Kanari, D. Keller, J. Kenyon, G. Khazen, Y. Kim, J. G. King, Z. Kisvarday, P. Kumbhar, S. Lasserre, J. V. Le Bé, B. R. Magalhães, A. Merchán-Pérez, J. Meystre, B. R. Morrice, J. Muller, A. Muñoz-Céspedes, S. Muralidhar, K. Muthurasa, D. Nachbaur, T. H. Newton, M. Nolte, A. Ovcharenko, J. Palacios, L. Pastor, R. Perin, R. Ranjan, I. Riachi, J. R. Rodríguez, J. L. Riquelme, C. Rössert, K. Sfyrakis, Y. Shi, J. C. Shillcock, G. Silberberg, R. Silva, F. Tauheed, M. Telefont, M. Toledo-Rodriguez, T. Tränkler, W. Van Geit, J. V. Díaz, R. Walker, Y. Wang, S. M. Zaninetta, J. Defelipe, S. L. Hill, I. Segev, and F. Schürmann, "Reconstruction and Simulation of Neocortical Microcircuitry," *Cell*, vol. 163, no. 2, pp. 456–492, 2015.

[17] M. V. Tsodyks, W. E. Skaggs, T. J. Sejnowski, and B. L. McNaughton, "Paradoxical effects of external modulation of inhibitory interneurons," *Journal of Neuroscience*, vol. 17, no. 11, pp. 4382–4388, 1997.

[18] Y. K. Wu and F. Zenke, "Nonlinear transient amplification in recurrent neural networks with short-term plasticity," *eLife*, vol. 10, pp. 1–43, 2021.

[19] G. Mongillo, O. Barak, and M. Tsodyks, "Synaptic theory of working memory," *Science*, vol. 319, no. 5869, pp. 1543–1546, 2008.

[20] G. Mongillo, D. Hansel, and C. van Vreeswijk, "Bistability and Spatiotemporal Irregularity in Neuronal Networks with Nonlinear Synaptic Transmission," *Physical Review Letters*, vol. 108, p. 158101, Apr. 2012.

[21] X.-J. Wang, "Synaptic reverberation underlying mnemonic persistent activity," *Trends in Neurosciences*, vol. 24, no. 8, pp. 455–463, 2001.

[22] B. Gyorgy and D. Andreas, "Neuronal Oscillations in Cortical Networks," *Science*, vol. 304, no. June, pp. 1926–1929, 2004.

[23] K. D. Miller, "Canonical computations of cerebral cortex," *Current Opinion in Neurobiology*, vol. 37, pp. 75–84, 2016. Neurobiology of cognitive behavior.

[24] C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, and M. Scanziani, "Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons," *Nature Neuroscience*, vol. 16, no. 8, pp. 1068–1076, 2013.

[25] X. Jiang, S. Shen, C. R. Cadwell, P. Berens, F. Sinz, A. S. Ecker, S. Patel, and A. S. Tolias, "Principles of connectivity among morphologically defined cell types in adult neocortex," *Science*, vol. 350, no. 6264, 2015.

[26] R. Tremblay, S. Lee, and B. Rudy, "GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits," *Neuron*, vol. 91, no. 2, pp. 260–292, 2016.

[27] A. Litwin-Kumar, R. Rosenbaum, and B. Doiron, "Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes," *Journal of Neurophysiology*, vol. 115, no. 3, pp. 1399–1409, 2016.

[28] A. Mahrach, G. Chen, N. Li, C. van Vreeswijk, and D. Hansel, "Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation," *bioRxiv*, pp. 1–37, 2019.

[29] K. D. Miller and A. Palmigiano, "Generalized paradoxical effects in excitatory/inhibitory networks," *bioRxiv*, p. 2020.10.13.336727, 2020.

[30] A. Palmigiano, F. Fumarola, D. P. Mossing, N. Kraynyukova, H. Adesnik, and K. D. Miller, "Structure and variability of optogenetic responses identify the operating regime of cortex," *bioRxiv*, pp. 1–49, 2020.

[31] Richter, M. A. Leonidas, G, and J. Gjorgjieva, "Interneuron subtypes enable independent modulation of excitatory and inhibitory firing rates after sensory deprivation," *bioRxiv*, p. 2021.05.25.445562, 2021.

[32] L. Campagnola, S. C. Seeman, T. Chartrand, L. Kim, A. Hoggarth, C. Gamlin, S. Ito, J. Trinh, P. Davoudian, C. Radaelli, *et al.*, "Local connectivity and synaptic dynamics in mouse and human neocortex," *Science*, vol. 375, no. 6585, p. eabj5861, 2022.

[33] Allen Institute for Brain Science, "Synaptic physiology coarse matrix dataset," 2019.

[34] J. Barral and A. D'Reyes, "Synaptic scaling rule preserves excitatory-inhibitory balance and salient neuronal network dynamics," *Nature Neuroscience*, vol. 19, no. 12, pp. 1690–1696, 2016.

[35] Y. Ahmadian and K. D. Miller, "What is the dynamical regime of cerebral cortex?," *Neuron*, pp. 1–19, 2021.

[36] M. Vinck, R. Batista-Brito, U. Knoblich, and J. A. Cardin, "Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding," *Neuron*, vol. 86, no. 3, pp. 740–754, 2015.

[37] J. H. Reynolds and L. Chelazzi, "Attentional modulation of visual processing," *Annual Review of Neuroscience*, vol. 27, pp. 611–647, 2004.

[38] Y. Fu, J. M. Tucciarone, J. S. Espinosa, N. Sheng, D. P. Darcy, R. A. Nicoll, Z. J. Huang, and M. P. Stryker, "A cortical circuit for gain control by behavioral state," *Cell*, vol. 156, no. 6, pp. 1139–1152, 2014.

[39] N. L. Rochefort, O. Garaschuk, R. I. Milos, M. Narushima, N. Marandi, B. Pichler, Y. Kovalchuk, and A. Konnerth, "Sparsification of neuronal activity in the visual cortex at eye-opening," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 35, pp. 15049–15054, 2009.

[40] H. Ko, L. Cossell, C. Baragli, J. Antolik, C. Clopath, S. B. Hofer, and T. D. Mrsic-Flogel, "The emergence of functional microcircuits in vi-

sual cortex," *Nature*, vol. 496, no. 7443, pp. 96–100, 2013.

Supplemental Information for 'Inhibition stabilization and paradoxical effects in recurrent neural networks with short-term plasticity'

Yue Kris Wu and Julijana Gjorgjieva

School of Life Sciences, Technical University of Munich, Freising, Germany

Max Planck Institute for Brain Research, Frankfurt, Germany

Supplemental Tables

Table 1: IS index		
E-to-E	STD	$\mathrm{Re}\left[\frac{1}{2}(\tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)-\tau_x^{-1}-U_d r_E)+\sqrt{\frac{1}{4}(\tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)+\tau_x^{-1}+U_d r_E)^2-\tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}U_d x}\right]$
	STF	$\mathrm{Re}\left[\frac{1}{2}(\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)-\tau_u^{-1}-U_f r_E)+\sqrt{\frac{1}{4}(\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)+\tau_u^{-1}+U_f r_E)^2+\tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}U_f(U_{max}-u)}\right]$
E-to-I	STD	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$
	STF	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$
I-to-E	STD	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$
	STF	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$
I-to-I	STD	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$
	STF	$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}-1)$

Table 2: Parameters for Figure 1			
Symbol	Value	Unit	Description
τ_E	20	ms	Time constant of excitatory firing dynamics
τ_x	200	ms	Time constant of STD
U_d	1.0	-	Depression rate
Note that these values are also used elsewhere unless mentioned otherwise.			

Table 3: Parameters for Figure 2			
Symbol	Value	Unit	Description
τ_u	200	ms	Time constant of STF
U_f	1.0	-	Facilitation rate
U_{max}	6.0	-	Maximal facilitation value
Note that these values are also used elsewhere unless mentioned otherwise.			

Supplementary Figures

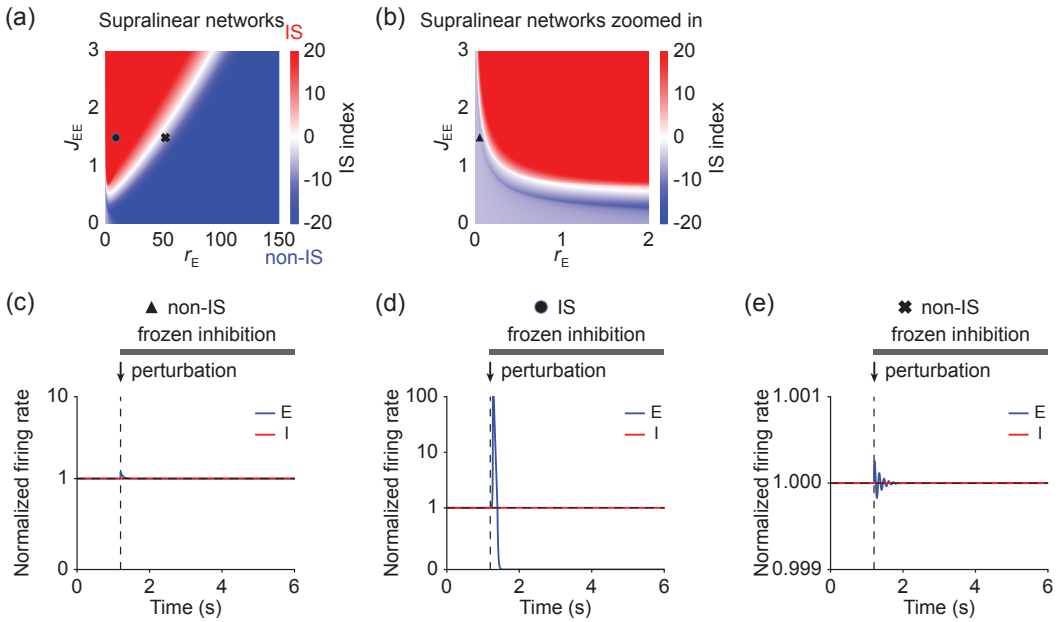

Figure S1. Examples of network activity in response to transient small perturbations to the excitatory activity while freezing inhibition at different initial states. (a) IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for supralinear networks with E-to-E STD. The non-IS and IS regimes are marked in blue and red, respectively. (b) Zoomed-in version of (a). (c) Firing rates of the excitatory (blue) and inhibitory population (red) for a network with an initially non-IS state corresponding to the triangle in (b). A small transient perturbation to excitatory population activity is introduced marked with arrows while freezing inhibition. The periods in which inhibition is frozen are marked with the black bar. (d-e) Same as (c) but for networks with initially IS and non-IS states corresponding to the circle and the cross in (a), respectively. Initial states are shifted by changing g_E. g_E is 6.65 in (c), 20 in (d), and 100 in (e).

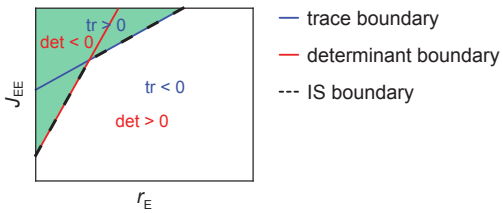

Figure S2. Schematic of IS boundary. Trace boundary (blue line) and determinant boundary (red line) correspond to the J_{EE} values at which the trace and the determinant of the Jacobian of the system with fixed inhibition are 0 for different r_E, respectively. IS boundary (black dashed line) is defined as the minimal J_{EE} values at which either the trace or the determinant of the Jacobian of the system with fixed inhibition is 0 for different r_E. The green and white regions represent the IS regime and non-IS regime, respectively.

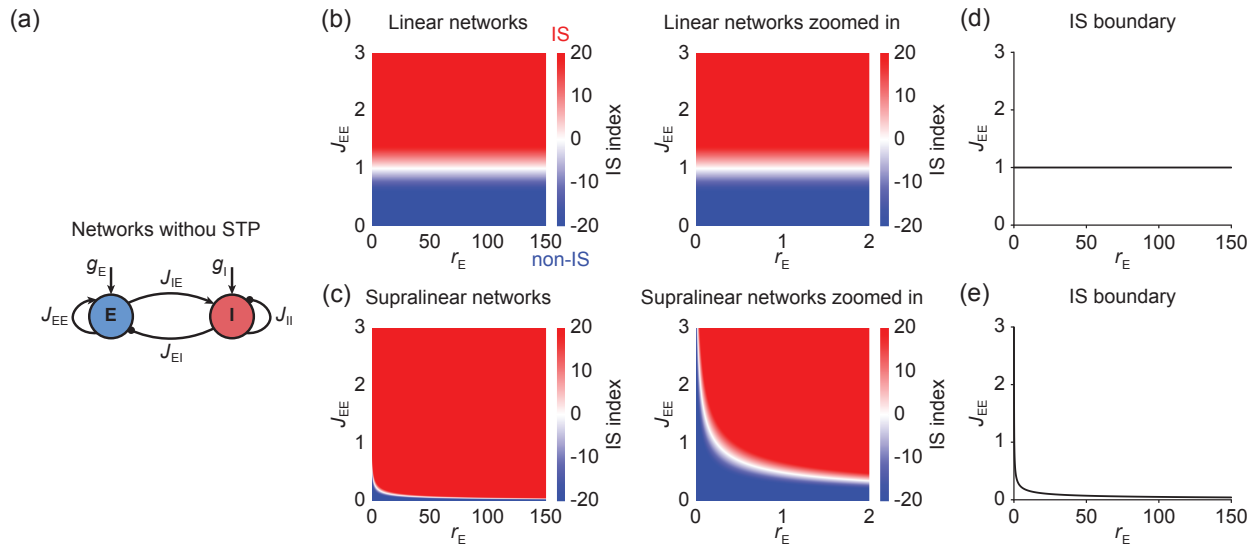

Figure S3. Inhibition stabilization in recurrent neural networks with static connectivity. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with static connectivity. (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for linear networks with static connectivity. The non-IS and IS regime are marked in blue and red, respectively. Right: zoomed-in version of (b) left. (c) Same as (b) but for supralinear networks with static connectivity. (d) IS boundary for linear networks with static connectivity, defined as the corresponding excitatory-to-excitatory connection strength J_{EE}^{IS} for different r_E at which the IS index is 0. (e) Same as (d) but for supralinear networks with static connectivity.

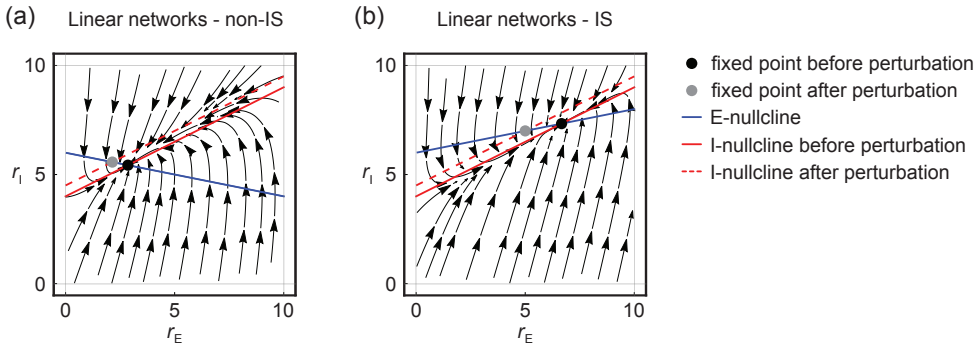

Figure S4. Conditions for the presence of the paradoxical effect from the angle of phase portraits. (a) Phase portrait of a linear non-IS network. The slope of the E-nullcline (blue) around the fixed point (black dot) is negative. The slope of the I-nucllcline (red) is larger than that of the E-nullcline. Perturbating the inhibitory population via injecting additional excitatory current shifts the I-nullcline vertically (from the red solid line to the red dashed line). The inhibitory activity at the new fixed point (gray dot) after perturbation is higher than at the original fixed point. The non-IS network therefore does not exhibit the paradoxical response. (b) Same as (a) but for a linear IS network. The slope of the E-nullcline (blue) around the fixed point is positive. The inhibitory activity at the new fixed point (gray dot) after perturbation is lower than at the original fixed point. The IS network therefore exhibits the paradoxical response.

Supplemental Methods

To incorporate short-term plasticity mechanisms, we followed the Tsodyks and Markram model [15] and modeled STD as follows:

$$\frac{dx_{AB}}{dt} = \frac{1 - x_{AB}}{\tau_x} - U_d x_{AB} r_B \tag{S0.1}$$

where x_{AB} is a depression variable that is limited to the interval $(0, 1]$ for the synaptic connection from population B to population A, τ_x is the time constant of STD, and U_d is the depression rate.

Furthermore, we modeled STF as follows:

$$\frac{du_{AB}}{dt} = \frac{1 - u_{AB}}{\tau_u} + U_f(U_{max} - u_{AB}) r_B \tag{S0.2}$$

where u_{AB} is a facilitation variable that is constrained to the interval $[1, U_{max})$ for the synaptic connection from population B to population A, τ_u is the time constant of STF, U_f is the facilitation rate, and U_{max} is the maximal facilitation value. Note that the short-term plasticity (STP) variables x_{AB} and u_{AB} multiply the corresponding synaptic connection strength J_{AB}, and therefore affect the effective synaptic strength.

Conditions for IS in networks with E-to-E STD

The dynamics of networks with E-to-E STD are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[x J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.3}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.4}$$

$$\frac{dx}{dt} = \frac{1 - x}{\tau_x} - U_d x r_E \tag{S0.5}$$

where x is the depression variable, which is limited to the interval $(0, 1]$, τ_x is the depression time constant, and U_d is the depression rate.

The Jacobian $\mathbf{M^{EE}_{STD}}$ of the system with E-to-E STD is given by

$$\mathbf{M^{EE}_{STD}} = \begin{bmatrix} \tau_E^{-1}(x J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} & \tau_E^{-1} J_{EE} \alpha_E r_E^{\frac{2\alpha_E - 1}{\alpha_E}} \\ \tau_I^{-1} J_{IE} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}}) & 0 \\ -U_d x & 0 & -\tau_x^{-1} - U_d r_E \end{bmatrix} \tag{S0.6}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{EE}_{STD}} = \begin{bmatrix} \tau_E^{-1}(x J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) & \tau_E^{-1} J_{EE} \alpha_E r_E^{\frac{2\alpha_E - 1}{\alpha_E}} \\ -U_d x & -\tau_x^{-1} - U_d r_E \end{bmatrix} \tag{S0.7}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{EE}_{STD}}) = \tau_E^{-1}(x J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E < 0 \tag{S0.8}$$

and

$$\det(\mathbf{N^{EE}_{STD}}) = \tau_E^{-1}(x J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_E) - \tau_E^{-1} J_{EE} \alpha_E r_E^{\frac{2\alpha_E - 1}{\alpha_E}} (-U_d x) > 0 \tag{S0.9}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$x > \min\Big(\sqrt{\frac{1}{J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}}, \frac{\tau_x + \tau_E + \tau_E \tau_x U_d r_E}{\tau_x J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}\Big) \tag{S0.10}$$

We further define the IS index for the system with E-to-E STD, which is the real part of the leading eigenvalue of $\mathbf{N}^{\mathbf{EE}}_{\mathbf{STD}}$, as follows:

$$\text{IS index} = \text{Re}\Bigg[\frac{\tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E}{2} + \sqrt{\frac{1}{4}[\tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) + \tau_x^{-1} + U_d r_E]^2 - \tau_E^{-1} J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_d x}\Bigg] \tag{S0.11}$$

IS boundary for networks with E-to-E STD

We next investigated how the boundary between non-IS and IS, which we called 'IS boundary', changes as a function of r_E. Mathematically, the IS boundary is determined by the corresponding recurrent excitatory-to-excitatory connection strength denoted by J_{EE}^{IS} for different r_E at which the IS index is 0. Therefore, we have

$$\mathrm{tr}(\mathbf{N_{STD}^{EE}}) = \tau_E^{-1}(xJ_{EE}^{\mathrm{tr}}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E = 0 \tag{S0.12}$$

$$J_{EE}^{\mathrm{tr}} = \frac{\tau_E\tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E\tau_x U_d^2 r_E^2 + U_d\tau_x r_E}{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}} \tag{S0.13}$$

$$\det(\mathbf{N_{STD}^{EE}}) = -\tau_E^{-1}\left(\frac{J_{EE}^{\det}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{\tau_x} - \frac{1+\tau_x U_d r_E}{\tau_x} - \frac{J_{EE}^{\det}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_d}{1+U_d r_E\tau_x}\right) = 0 \tag{S0.14}$$

$$J_{EE}^{\det} = \frac{(1+\tau_x U_d r_E)^2}{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}} \tag{S0.15}$$

The IS boundary is the minimum of J_{EE}^{tr} and $J_{EE}^{\det}$ (Fig. S2), and thus can be written as follows:

$$J_{EE}^{IS} = \min(J_{EE}^{\mathrm{tr}}, J_{EE}^{\det}) \tag{S0.16}$$

For $\alpha_E = 1$, we have

$$\frac{dJ_{EE}^{\mathrm{tr}}}{dr_E} = 2\tau_E U_d + 2\tau_E\tau_x U_d^2 r_E + U_d\tau_x \tag{S0.17}$$

$$\frac{dJ_{EE}^{\det}}{dr_E} = 2\tau_x U_d + 2\tau_x^2 U_d^2 r_E \tag{S0.18}$$

As a result, for linear networks with E-to-E STD, J_{EE}^{tr} and $J_{EE}^{\det}$ are increasing with r_E. Therefore, J_{EE}^{IS} also increases with r_E. Furthermore, the increase of J_{EE}^{IS} accelerates for larger r_E.

For $\alpha_E > 1$, we have

$$\frac{dJ_{EE}^{\mathrm{tr}}}{dr_E} = \frac{-(\tau_E\tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E\tau_x U_d^2 r_E^2 + U_d\tau_x r_E)(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (2\tau_E U_d + 2\tau_E\tau_x U_d^2 r_E + U_d\tau_x)\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}})^2} \tag{S0.19}$$

Further,

$$\begin{aligned} &-(\tau_E\tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E\tau_x U_d^2 r_E^2 + U_d\tau_x r_E)(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (2\tau_E U_d + 2\tau_E\tau_x U_d^2 r_E + U_d\tau_x)\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ &\quad \propto -(\tau_E\tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E\tau_x U_d^2 r_E^2 + U_d\tau_x r_E)(\alpha_E - 1) + (2\tau_E U_d + 2\tau_E\tau_x U_d^2 r_E + U_d\tau_x)\alpha_E r_E \end{aligned} \tag{S0.20}$$

Similarly, we have

$$\frac{dJ_{EE}^{\det}}{dr_E} = \frac{-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}})^2} \tag{S0.21}$$

Further,

$$\begin{aligned} &-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ &\quad \propto -(r_E^{-1} + \tau_x U_d)(\alpha_E - 1) + 2\tau_x U_d\alpha_E \\ &\quad = -r_E^{-1}\alpha_E + r_E^{-1} + \tau_x U_d + \tau_x U_d\alpha_E \end{aligned} \tag{S0.22}$$

Clearly, the terms $-(\tau_E\tau_x^{-1}+1+2\tau_E U_d r_E+\tau_E\tau_x U_d^2 r_E^2+U_d\tau_x r_E)(\alpha_E-1)+(2\tau_E U_d+2\tau_E\tau_x U_d^2 r_E+U_d\tau_x)\alpha_E r_E$ and $-r_E^{-1}\alpha_E+r_E^{-1}+\tau_x U_d+\tau_x U_d\alpha_E$ switches from negative to positive as r_E grows. Therefore, for supralinear networks with E-to-E STD, the IS boundary first shifts downwards and then shifts upwards as r_E increases.

Conditions for paradoxical response in networks with E-to-E STD

Next, we derived the condition of having the paradoxical effect in networks with E-to-E STD. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the depression variable to the value at its fixed point in terms of r_E. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[\frac{1}{1+\tau_x U_d r_E} J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.23}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{\frac{1}{1+\tau_x U_d r_E} J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.24}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STD}^E = \frac{1}{J_{EI}}\big[-\frac{J_{EE}}{(1+\tau_x U_d r_E)^2}\tau_x U_d r_E + \frac{J_{EE}}{1+\tau_x U_d r_E} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big] \tag{S0.25}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, the short-term depression variable x at the fixed point has to satisfy the following condition

$$x > \sqrt{\frac{1}{J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}} \tag{S0.26}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.27}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.28}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.29}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\begin{aligned} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}(\tau_x^{-1}+U_d r_E) > (1+J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})\big[-\frac{J_{EE}U_d r_E}{1+\tau_x U_d r_E}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ +\frac{J_{EE}}{1+\tau_x U_d r_E}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}(\tau_x^{-1}+U_d r_E) - (\tau_x^{-1}+U_d r_E)\big] \end{aligned} \tag{S0.30}$$

Note that to ensure that the system with E-to-E STD is stable, $\det(\mathbf{M_{STD}^{EE}})$ has to be negative. Therefore, we have

$$\begin{aligned} \det(\mathbf{N_{STD}^{EE}}) = \tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1+J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1}+U_d r_E) \\ -\tau_E^{-1}J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}\tau_I^{-1}J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}(\tau_x^{-1}+U_d r_E) + \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}\tau_I^{-1}(1+J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})U_d x < 0 \end{aligned} \tag{S0.31}$$

The condition shown in Eq. S0.30 is the same as the stability condition of the determinant of the Jacobian of the system with E-to-E STD, namely, as $\det(\mathbf{M_{STD}^{EE}}) < 0$. Thus, the condition is always satisfied when the system with E-to-E STD is stable.

Based on the condition of being IS shown in Eq. S0.10 and the condition of having paradoxical effect shown in Eq. S0.26, we therefore can conclude that in networks with E-to-E STD, the paradoxical effect implies inhibitory stabilization, whereas inhibitory stabilization does not necessarily imply paradoxical responses.

Paradoxical boundary for networks with E-to-E STD

We next investigated how the boundary between non-paradoxical and paradoxical effects, which we called 'paradoxical boundary', changes as a function of r_E. Mathematically, the paradoxical boundary is determined by the corresponding recurrent excitatory-to-excitatory connection strength denoted by J_{EE}^{PD} for different r_E at which the slope of the excitatory nullcline is 0. Therefore, we have

$$k_{STD}^{E} = \frac{1}{J_{EI}}\Big(- \frac{J_{EE}^{PD}}{(1+\tau_x U_d r_E)^2}\tau_x U_d r_E + \frac{J_{EE}^{PD}}{1+\tau_x U_d r_E} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big) = 0 \tag{S0.32}$$

$$J_{EE}^{PD} = \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}(1+\tau_x U_d r_E)^2 \tag{S0.33}$$

For $\alpha_E = 1$, we have

$$\frac{dJ_{EE}^{PD}}{dr_E} = 2\tau_x U_d + 2\tau_x^2 U_d^2 r_E \tag{S0.34}$$

Obviously, the derivative is always positive. Therefore, for linear networks with E-to-E STD, J_{EE}^{PD} is increasing with r_E.

For $\alpha_E > 1$, we have

$$\begin{aligned}\frac{dJ_{EE}^{PD}}{dr_E} =& \frac{1}{\alpha_E}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{\frac{1-2\alpha_E}{\alpha_E}} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{\frac{1-\alpha_E}{\alpha_E}} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E} r_E^{\frac{1}{\alpha_E}}\Big) \\ =& \frac{1}{\alpha_E} r_E^{-\frac{1}{\alpha_E}}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{-2} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{-1} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E}\Big)\end{aligned} \tag{S0.35}$$

Clearly, the derivative switches from negative to positive as r_E grows. Therefore, for supralinear networks with E-to-E STD, the paradoxical effect boundary first shifts downwards and then shifts upwards as r_E increases.

Conditions for IS in networks with E-to-E STF

The dynamics of networks with E-to-E STF are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[uJ_{EE}r_E - J_{EI}r_I + g_E\Big]_+^{\alpha_E} \tag{S0.36}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE}r_E - J_{II}r_I + g_I\Big]_+^{\alpha_I} \tag{S0.37}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f(U_{max}-u)r_E \tag{S0.38}$$

where u is the facilitation variable constrained to the interval $[1, U_{max})$, U_{max} is the maximal facilitation value, τ_u is the time constant of STF, and U_f is the facilitation rate.

The Jacobian $\mathbf{M^{EE}_{STF}}$ of the system with E-to-E STF is given by

$$\mathbf{M^{EE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1}J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1}J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ U_f(U_{max}-u) & 0 & -\tau_u^{-1} - U_f r_E \end{bmatrix} \tag{S0.39}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{EE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ U_f(U_{max}-u) & -\tau_u^{-1} - U_f r_E \end{bmatrix} \tag{S0.40}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{EE}_{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E < 0 \tag{S0.41}$$

and

$$\det(\mathbf{N^{EE}_{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) - \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max}-u) > 0 \tag{S0.42}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\frac{(1 + 2U_f U_{max} r_E \tau_u + U_f r_E \tau_u U_f U_{max} r_E \tau_u) J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{(1+\tau_u U_f r_E)^2} - 1 > 0 \tag{S0.43}$$

We further define the IS index for the system with E-to-E STF as follows:

$$\begin{aligned} \text{IS index} = \mathrm{Re}\Bigg[& \frac{\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E}{2} \\ & + \sqrt{\frac{1}{4}[\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) + \tau_u^{-1} + U_f r_E]^2 + \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max}-u)} \Bigg] \end{aligned} \tag{S0.44}$$

IS boundary for networks with E-to-E STF

We next investigated how the boundary between non-IS and IS, which we called 'IS boundary', changes as a function of r_E. Mathematically, the IS boundary is determined by the corresponding recurrent excitatory-to-excitatory connection strength denoted by J_{EE}^{IS} for different r_E at which the IS index is 0. Therefore, we have

$$\mathrm{tr}(\mathbf{N_{STF}^{EE}}) = \tau_E^{-1}(uJ_{EE}^{\mathrm{tr}}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E = 0 \tag{S0.45}$$

$$J_{EE}^{\mathrm{tr}} = \frac{(\tau_E\tau_u^{-1} + \tau_E U_f r_E + 1)(1 + U_f r_E \tau_u)}{(1 + U_f U_{max} r_E \tau_u)\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}} \tag{S0.46}$$

$$\det(\mathbf{N_{STF}^{EE}}) = \tau_E^{-1}(-\tau_u^{-1}\frac{1 + U_f U_{max} r_E \tau_u}{1 + U_f r_E \tau_u} J_{EE}^{\det}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u^{-1} + U_f r_E - J_{EE}^{\det}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}) = 0 \tag{S0.47}$$

$$J_{EE}^{\det} = \frac{1 + \tau_u U_f r_E}{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}} \tag{S0.48}$$

Clearly, J_{EE}^{tr} is greater than $J_{EE}^{\det}$ for any r_E. Therefore, we have

$$J_{EE}^{IS} = \min(J_{EE}^{\mathrm{tr}}, J_{EE}^{\det}) = J_{EE}^{\det} \tag{S0.49}$$

For $\alpha_E = 1$, we have

$$\frac{dJ_{EE}^{\det}}{dr_E} = \frac{dJ_{EE}^{IS}}{dr_E} = \frac{2U_f\tau_u(1 - U_{max})}{(1 + U_f r_E \tau_u)(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} + \tau_u r_E U_f U_{max})^2} \tag{S0.50}$$

Therefore, for linear networks with E-to-E STF, J_{EE}^{IS} is decreasing with r_E. Furthermore, the decrease of J_{EE}^{IS} slows down for larger r_E.

For $\alpha_E > 1$, we have

$$\begin{aligned}\frac{dJ_{EE}^{\det}}{dr_E} = \frac{dJ_{EE}^{IS}}{dr_E} = &-\frac{1 + \tau_u U_f r_E}{(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})^2}(-\frac{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u + U_f U_{max} r_E \tau_u \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u}{(1 + U_f r_E \tau_u)^2} \\ &+ \frac{(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + 2U_f U_{max}\tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u}{1 + U_f r_E \tau_u}) \\ &+ \frac{\tau_u U_f(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})}{(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})^2}\end{aligned} \tag{S0.51}$$

Further,

$$\begin{aligned}&-(1 + \tau_u U_f r_E)(-\frac{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u + U_f U_{max} r_E \tau_u \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u}{(1 + U_f r_E \tau_u)^2} \\ &+ \frac{(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + 2U_f U_{max}\tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u}{1 + U_f r_E \tau_u}) \\ &+ \tau_u U_f(\frac{1 + U_f U_{max} r_E \tau_u}{1 + U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}) \\ &= (-(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (-2U_{max}\alpha_E + 2U_{max} - U_{max}\alpha_E + \alpha_E - U_{max}\alpha_E + 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u \\ &-2U_f U_{max}\tau_u(\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f r_E \tau_u - 2\tau_u(\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u(1 + U_f r_E \tau_u))/(1 + U_f r_E \tau_u)\end{aligned} \tag{S0.52}$$

Therefore, for supralinear networks with E-to-E STF, J_{EE}^{IS} is decreasing with r_E.

Conditions for paradoxical response in networks with E-to-E STF

Next, we derived the condition of having the paradoxical effect in networks with E-to-E STF. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the facilitation variable to the value at its fixed point in terms of r_E. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.53}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.54}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STF}^E = \frac{1}{J_{EI}}\Big(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \Big(\frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2}\Big) J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big) \tag{S0.55}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$\frac{1}{J_{EI}}\Big(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big) > 0 \tag{S0.56}$$

The above condition can be simplified as follows:

$$\frac{1+2U_f U_{max} r_E \tau_u + U_f U_{max} r_E \tau_u U_f r_E \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0 \tag{S0.57}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.58}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.59}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.60}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big] \tag{S0.61}$$

Note that to ensure that the system with E-to-E STF is stable, $\det(\mathbf{M}^{\mathbf{EE}}_{\mathbf{STF}})$ has to be negative. Therefore, we have

$$\det(\mathbf{N}^{\mathbf{EE}}_{\mathbf{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) - \tau_E^{-1} J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max} - u) < 0 \tag{S0.62}$$

The condition shown in Eq. S0.61 is the same as the stability condition of the determinant of the Jacobian of the system with E-to-E STF, namely, as $\det(\mathbf{M}^{\mathbf{EE}}_{\mathbf{STF}}) < 0$. Thus, the condition is always satisfied when the system with E-to-E STF is stable.

The condition of being IS shown in Eq. S0.43 is identical to the condition of having paradoxical effect shown in Eq. S0.57, we therefore can conclude that in networks with E-to-E STF, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with E-to-I STD

The dynamics of networks with E-to-I STD are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.63}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[x J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.64}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_E \tag{S0.65}$$

where x is the depression variable, which is limited to the interval $(0, 1]$, τ_x is the depression time constant, and U_d is the depression rate.

The Jacobian $\mathbf{M}^{\mathbf{IE}}_{\mathbf{STD}}$ of the system with E-to-I STD is given by

$$\mathbf{M}^{\mathbf{IE}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} x J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} \\ -U_d x & 0 & -\tau_x^{-1} - U_d r_E \end{bmatrix} \tag{S0.66}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ -U_d x & -\tau_x^{-1} - U_d r_E \end{bmatrix} \tag{S0.67}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E < 0 \tag{S0.68}$$

and

$$\det(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_E) > 0 \tag{S0.69}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0 \tag{S0.70}$$

We further define the IS index for the system with E-to-I STD as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) \tag{S0.71}$$

Conditions for paradoxical response in networks with E-to-I STD

Next, we derived the condition of having the paradoxical effect in networks with E-to-I STD. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the depression variable to the value at its fixed point in terms of r_E. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.72}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.73}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STD}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.74}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0 \tag{S0.75}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[\frac{1}{1+U_d r_E \tau_x} J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.76}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{\frac{1}{1+U_d r_E \tau_x} J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.77}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{\frac{1}{(1+U_d r_E \tau_x)^2} J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.78}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{\frac{1}{(1+U_d r_E \tau_x)^2} J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.79}$$

Note that to ensure that the system with E-to-I STD is stable, $\det(\mathbf{M_{STD}^{IE}})$ has to be negative. Therefore, we have

$$\begin{aligned}\det(\mathbf{M_{STD}^{IE}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1} + U_d r_E) \\ +\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}[(\tau_I^{-1} x J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(-\tau_x^{-1} - U_d r_E) + \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_d x] < 0\end{aligned} \tag{S0.80}$$

The condition shown in Eq. S0.79 is fulfilled if $\det(\mathbf{M_{STD}^{IE}}) < 0$. Thus, the condition is always satisfied when the system with E-to-I STD is stable.

The condition of being IS shown in Eq. S0.70 is identical to the condition of having paradoxical effect shown in Eq. S0.75, we therefore can conclude that in networks with E-to-I STD, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with E-to-I STF

The dynamics of networks with E-to-I STF are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.81}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[u J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.82}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_E \tag{S0.83}$$

where u is the facilitation variable constrained to the interval $[1, U_{max})$, U_{max} is the maximal facilitation value, τ_u is the time constant of STF, and U_f is the facilitation rate.

The Jacobian $\mathbf{M^{IE}_{STF}}$ of the system with E-to-I STF is given by

$$\mathbf{M^{IE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} & 0 \\ \tau_I^{-1} u J_{IE}\alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}}) & \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} \\ U_f(U_{max} - u) & 0 & -\tau_u^{-1} - U_f r_E \end{bmatrix} \tag{S0.84}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{IE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) & 0 \\ U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_E \end{bmatrix} \tag{S0.85}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{IE}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E < 0 \tag{S0.86}$$

and

$$\det(\mathbf{N^{IE}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) > 0 \tag{S0.87}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) > 0 \tag{S0.88}$$

We further define the IS index for the system with E-to-I STF as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1) \tag{S0.89}$$

Conditions for paradoxical response in networks with E-to-I STF

Next, we derived the condition of having the paradoxical effect in networks with E-to-I STF. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the facilitation variable to the value at its fixed point in terms of r_E. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.90}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.91}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STF}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.92}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0 \tag{S0.93}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.94}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.95}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{IE} r_E}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.96}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{IE} r_E}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.97}$$

Note that to ensure that the system with E-to-I STF is stable, $\det(\mathbf{M_{STF}^{IE}})$ has to be negative. Therefore, we have

$$\begin{aligned} \det(\mathbf{M_{STF}^{IE}}) &= \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_E) \\ &+ \tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} [(\tau_I^{-1} u J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(-\tau_u^{-1} - U_f r_E) - \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_f(U_{max} - u)] < 0 \end{aligned} \tag{S0.98}$$

The condition shown in Eq. S0.97 is the same as the stability condition of the determinant of the Jacobian of the system with E-to-I STF, namely, as $\det(\mathbf{M_{STF}^{IE}}) < 0$. Thus, the condition is always satisfied when the system with E-to-I STF is stable.

The condition of being IS shown in Eq. S0.88 is identical to the condition of having paradoxical effect shown in Eq. S0.93, we therefore can conclude that in networks with E-to-I STF, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with I-to-E STD

The dynamics of networks with I-to-E STD are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - x J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.99}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.100}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_I \tag{S0.101}$$

where x is the depression variable, which is limited to the interval $(0, 1]$, τ_x is the depression time constant, and U_d is the depression rate.

The Jacobian $\mathbf{M_{STD}^{EI}}$ of the system with I-to-E STD is given by

$$\mathbf{M_{STD}^{EI}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} x J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ 0 & -U_d x & -\tau_x^{-1} - U_d r_I \end{bmatrix} \tag{S0.102}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N_{STD}^{EI}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ 0 & -\tau_x^{-1} - U_d r_I \end{bmatrix} \tag{S0.103}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N_{STD}^{EI}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_I < 0 \tag{S0.104}$$

and

$$\det(\mathbf{N_{STD}^{EI}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_I) > 0 \tag{S0.105}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0 \tag{S0.106}$$

We further define the IS index for the system with I-to-E STD as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) \tag{S0.107}$$

Conditions for paradoxical response in networks with I-to-E STD

Next, we derived the condition of having the paradoxical effect in networks with I-to-E STD. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the depression variable to the value at its fixed point in terms of r_I. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \left[J_{EE} r_E - \frac{1}{1 + U_d r_I \tau_x} J_{EI} r_I + g_E \right]_+^{\alpha_E} = 0 \tag{S0.108}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{\frac{1}{1+U_d r_I \tau_x} J_{EI}} \tag{S0.109}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STF}^E = \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{EI} r_I + \frac{1}{1+U_d r_I \tau_x} J_{EI}} \tag{S0.110}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0 \tag{S0.111}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \left[J_{IE} r_E - J_{II} r_I + g_I \right]_+^{\alpha_I} = 0 \tag{S0.112}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.113}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.114}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{EI} r_I + \frac{1}{1+U_d r_I \tau_x} J_{EI}} \tag{S0.115}$$

Note that to ensure that the system with I-to-E STD is stable, $\det(\mathbf{M_{STD}^{EI}})$ has to be negative. Therefore, we have

$$\begin{aligned} \det(\mathbf{M_{STD}^{EI}}) = (J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1)(1 + J_{II} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}})(1 + \tau_x U_d r_I) \\ -x J_{EI} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} J_{IE} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} (1 + \tau_x U_d r_I) + J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} J_{IE} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} \tau_x U_d x < 0 \end{aligned} \tag{S0.116}$$

The condition shown in Eq. S0.115 is the same as the stability condition of the determinant of the Jacobian of the system with I-to-E STD, namely, as $\det(\mathbf{M}^{\mathbf{EI}}_{\mathbf{STD}}) < 0$. Thus, the condition is always satisfied when the system with I-to-E STD is stable.

The condition of being IS shown in Eq. S0.106 is identical to the condition of having paradoxical effect shown in Eq. S0.111, we therefore can conclude that in networks with I-to-E STD, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with I-to-E STF

The dynamics of networks with I-to-E STF are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - u J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.117}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.118}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_I \tag{S0.119}$$

where u is the facilitation variable constrained to the interval $[1, U_{max})$, U_{max} is the maximal facilitation value, τ_u is the time constant of STF, and U_f is the facilitation rate.

The Jacobian $\mathbf{M^{EI}_{STF}}$ of the system with I-to-E STF is given by

$$\mathbf{M^{EI}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} u J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ 0 & U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_I \end{bmatrix} \tag{S0.120}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{EI}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ 0 & -\tau_u^{-1} - U_f r_I \end{bmatrix} \tag{S0.121}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{EI}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_I < 0 \tag{S0.122}$$

and

$$\det(\mathbf{N^{EI}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_I) > 0 \tag{S0.123}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0 \tag{S0.124}$$

We further define the IS index for the system with I-to-E STF as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) \tag{S0.125}$$

Conditions for paradoxical response in networks with I-to-E STF

Next, we derived the condition of having the paradoxical effect in networks with I-to-E STF. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the facilitation variable to the value at its fixed point in terms of r_I. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \left[J_{EE} r_E - \frac{1 + U_f U_{max} r_I \tau_u}{1 + U_f r_I \tau_u} J_{EI} r_I + g_E \right]_+^{\alpha_E} = 0 \tag{S0.126}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI}} \tag{S0.127}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STF}^E = \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{EI} r_I} \tag{S0.128}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0 \tag{S0.129}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \left[J_{IE} r_E - J_{II} r_I + g_I \right]_+^{\alpha_I} = 0 \tag{S0.130}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}} \tag{S0.131}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.132}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{EI} r_I} \tag{S0.133}$$

Note that to ensure that the system with I-to-E STF is stable, $\det(\mathbf{M_{STF}^{EI}})$ has to be negative. Therefore, we have

$$\begin{aligned} \det(\mathbf{M_{STF}^{EI}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_I) \\ + \tau_E^{-1} u J_{EI} \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE} \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (-\tau_u^{-1} - U_f r_I) - \tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE} \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_f (U_{max} - u) < 0 \end{aligned} \tag{S0.134}$$

The condition shown in Eq. S0.133 is the same as the stability condition of the determinant of the Jacobian of the system with I-to-E STF, namely, as $\det(\mathbf{M}^{\mathbf{EI}}_{\mathbf{STF}}) < 0$. Thus, the condition is always satisfied when the system with I-to-E STF is stable.

The condition of being IS shown in Eq. S0.124 is identical to the condition of having paradoxical effect shown in Eq. S0.129, we therefore can conclude that in networks with I-to-E STF, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with I-to-I STD

The dynamics of networks with I-to-I STD are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.135}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - x J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.136}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_I \tag{S0.137}$$

where x is the depression variable, which is limited to the interval $(0, 1]$, τ_x is the depression time constant, and U_d is the depression rate.

The Jacobian $\mathbf{M^{II}_{STD}}$ of the system with I-to-I STD is given by

$$\mathbf{M^{II}_{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + x J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & -\tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} \\ 0 & -U_d x & -\tau_x^{-1} - U_d r_I \end{bmatrix} \tag{S0.138}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{II}_{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ 0 & -\tau_x^{-1} - U_d r_I \end{bmatrix} \tag{S0.139}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_I < 0 \tag{S0.140}$$

and

$$\det(\mathbf{N^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_I) > 0 \tag{S0.141}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0 \tag{S0.142}$$

We further define the IS index for the system with I-to-I STD as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) \tag{S0.143}$$

Conditions for paradoxical response in networks with I-to-I STD

Next, we derived the condition of having the paradoxical effect in networks with I-to-I STD. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the depression variable to the value at its fixed point in terms of r_I. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.144}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.145}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k^E_{STD} = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.146}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0 \tag{S0.147}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} \quad = \quad -r_I + \Big[J_{IE}r_E - \frac{1}{1+U_d r_I \tau_x} J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.148}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I \quad = \quad \frac{J_{IE}r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{\frac{1}{1+U_d r_I \tau_x} J_{II}} \tag{S0.149}$$

The slope of the inhibitory nullcline can be written as follows

$$k^I_{STD} = \frac{J_{IE}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{II} r_I + \frac{1}{1+U_d r_I \tau_x} J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.150}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{J_{IE}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{II} r_I + \frac{1}{1+U_d r_I \tau_x} J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.151}$$

Note that to ensure that the system with I-to-I STD is stable, $\det(\mathbf{M^{II}_{STD}})$ has to be negative. Therefore, we have

$$\begin{aligned}\det(\mathbf{M^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)[\tau_I^{-1}(1 + xJ_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1} + U_d r_I) - \tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} U_d x] \\ + \tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (-\tau_x^{-1} - U_d r_I) < 0\end{aligned} \tag{S0.152}$$

The condition shown in Eq. S0.151 is the same as the stability condition of the determinant of the Jacobian of the system with I-to-I STD, namely, as $\det(\mathbf{M^{II}_{STD}}) < 0$. Thus, the condition is always satisfied when the system with I-to-I STD is stable.

The condition of being IS shown in Eq. S0.142 is identical to the condition of having paradoxical effect shown in Eq. S0.147, we therefore can conclude that in networks with I-to-I STD, inhibitory stabilization and the paradoxical effect imply each other.

Conditions for IS in networks with I-to-I STF

The dynamics of networks with I-to-I STF are given by

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.153}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - u J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.154}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_I \tag{S0.155}$$

where u is the facilitation variable constrained to the interval $[1, U_{max})$, U_{max} is the maximal facilitation value, τ_u is the time constant of STF, and U_f is the facilitation rate.

The Jacobian $\mathbf{M^{II}_{STF}}$ of the system with I-to-I STF is given by

$$\mathbf{M^{II}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + u J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & -\tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} \\ 0 & U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_I \end{bmatrix} \tag{S0.156}$$

If inhibition is frozen, in other words, if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{II}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ 0 & -\tau_u^{-1} - U_f r_I \end{bmatrix} \tag{S0.157}$$

For the system with frozen inhibition, the dynamics are stable if

$$\mathrm{tr}(\mathbf{N^{II}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_I < 0 \tag{S0.158}$$

and

$$\det(\mathbf{N^{II}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_I) > 0 \tag{S0.159}$$

Therefore, if the network is an IS at the fixed point, the following condition has to be satisfied:

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0 \tag{S0.160}$$

We further define the IS index for the system with I-to-I STF as follows:

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) \tag{S0.161}$$

Conditions for paradoxical response in networks with I-to-I STF

Next, we derived the condition of having the paradoxical effect in networks with I-to-I STF. Under the assumption that sufficient small perturbations would not lead to unstable network dynamics and regime transition between non-IS and IS, the conditions of having paradoxical effects are a positive slope of the excitatory nullcline and a larger slope of the inhibitory nullcline than that of the excitatory nullcline locally around the fixed point [17]. We set the facilitation variable to the value at its fixed point in terms of r_I. Then we can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0 \tag{S0.162}$$

For $r_{E,I} > 0$, we have

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}} \tag{S0.163}$$

The slope of the excitatory nullcline in the r_E/r_I plane where x axis is r_E and y axis is r_I can be written as follows

$$k_{STF}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.164}$$

To have paradoxical effect, the slope of the excitatory nullcline at the fixed point of the system has to be positive. Therefore, we have

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0 \tag{S0.165}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - \frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0 \tag{S0.166}$$

In the region of rates $r_{E,I} > 0$, we have

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II}} \tag{S0.167}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} + \frac{U_f U_{max} \tau_u (1+U_f r_I \tau_u) - (1+U_f U_{max} r_I \tau_u) U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{II} r_I + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} \tag{S0.168}$$

In addition to the positive slope of the excitatory nullcline, the slope of the inhibitory nullcline at the fixed point of the system has to be larger than the slope of the excitatory nullcline. We therefore have

$$\frac{J_{IE}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} + \frac{U_f U_{max} \tau_u (1+U_f r_I \tau_u) - (1+U_f U_{max} r_I \tau_u) U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{II} r_I + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big) \tag{S0.169}$$

Note that to ensure that the system with I-to-I STF is stable, $\det(\mathbf{M_{STF}^{II}})$ has to be negative. Therefore, we have

$$\begin{aligned}\det(\mathbf{M_{STF}^{II}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)[\tau_I^{-1}(1 + uJ_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_I) + \tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} U_f (U_{max} - u)] \\ -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (\tau_u^{-1} + U_f r_I) < 0\end{aligned} \tag{S0.170}$$

The condition shown in Eq. S0.169 is the same as the stability condition of the determinant of the Jacobian of the system with I-to-I STF, namely, as $\det(\mathbf{M}^{\mathbf{II}}_{\mathbf{STF}}) < 0$. Thus, the condition is always satisfied when the system with I-to-I STF is stable.

The condition of being IS shown in Eq. S0.160 is identical to the condition of having paradoxical effect shown in Eq. S0.165, we therefore can conclude that in networks with I-to-I STF, inhibitory stabilization and the paradoxical effect imply each other.

IS boundary for networks with E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF

We next investigated how the boundary between non-IS and IS, which we called 'IS boundary', changes as a function of r_E. Mathematically, the IS boundary is determined by the corresponding recurrent excitatory-to-excitatory connection strength denoted by J_{EE}^{IS} for different r_E at which the IS index is 0. Therefore, we have

$$\begin{aligned}\det(\mathbf{N_{STD}^{IE}}) = \det(\mathbf{N_{STF}^{IE}}) = \det(\mathbf{N_{STD}^{EI}}) = \det(\mathbf{N_{STF}^{EI}}) = \det(\mathbf{N_{STD}^{II}}) = \det(\mathbf{N_{STF}^{II}}) \\ = \tau_E^{-1}(J_{EE}^{IS}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_E) = 0\end{aligned} \tag{S0.171}$$

$$J_{EE}^{IS} = \frac{1}{\alpha_E} r_E^{-\frac{\alpha_E-1}{\alpha_E}} \tag{S0.172}$$

For $\alpha = 1$, we have

$$\frac{dJ_{EE}^{IS}}{dr_E} = 0 \tag{S0.173}$$

Therefore, for linear networks with E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF, the IS boundary does not change with r_E.

For $\alpha > 1$, we have

$$\frac{dJ_{EE}^{IS}}{dr_E} = -\frac{\alpha_E - 1}{\alpha_E^2} r_E^{-\frac{2\alpha_E-1}{\alpha_E}} \tag{S0.174}$$

As a result, for supralinear networks with E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF, $\frac{dJ_{EE}^{IS}}{dr_E}$ is always negative, suggesting that the boundary shifts downwards as r_E increases. Furthermore, the decrease of J_{EE}^{IS} slows down for larger r_E.

Paradoxical boundary for networks with E-to-E STF, E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF

As inhibition stabilization and paradoxical effect imply each other in networks with E-to-E STF, E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF, the change of paradoxical boundary is identical to the change of IS boundary as we demonstrated before.

Note that conditions for IS and the paradoxical effect in networks with E-to-E STD and networks with E-to-I STF are shown in recent studies [18]. For the sake of completeness, we also included them in the Methods section.