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: May 31, 2023)

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. Furthermore, we generalize our results to more complex scenarios including networks with multiple interneuron subtypes and any monotonically increasing neuronal nonlinearities. 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 [7]. Over the past decade, much effort has been made to identify the operating regime of cortical networks based on the paradoxical effect [1, 8, 9].

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 [10–12]. 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 [13–15]. 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 [16, 17]. Upon presynaptic stimulation, postsynaptic responses can either get depressed subject to short-term depression (STD) or facilitated subject to short-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 a neuronal nonlinearity and E-to-E STD, monotonically increasing excita-

tory activity can lead to non-monotonic transitions between non-inhibition-stabilization and inhibition-stabilization, as well as between non-paradoxically-responding and paradoxically-responding regimes. Furthermore, we generalize our results to more complex scenarios including networks with multiple interneuron subtypes and any monotonically increasing neuronal nonlinearities. In conclusion, our work reveals the impact of neuronal nonlinearities 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:

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[p_{EE}J_{EE}r_E - p_{EI}J_{EI}r_I + g_E\Big]_+^{\alpha_E}, \tag{1}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[p_{IE}J_{IE}r_E - p_{II}J_{II}r_I + g_I\Big]_+^{\alpha_I}, \tag{2}$$

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, α_E and α_I are the exponents of the respective input-output functions. Finally, p_{AB} represents the short-term plasticity variable from population B to population A. We implemented short-term plasticity mechanisms based on the Tsodyks and Markram model [16]. For STD, we replaced p_{AB} by x_{AB} and described the STD dynamics as follows:

$$\frac{dx_{AB}}{dt} = \frac{1 - x_{AB}}{\tau_x} - U_d x_{AB} r_B, \tag{3}$$

where x_{AB} is a short-term depression variable that is limited to the interval (0,1] for the synaptic connection from population B to population A. Biophysically, the short-term depression variable x represents the fraction of vesicles available for release, τ_x is the time constant of STD, and U_d is the depression factor controlling the degree of depression induced by the presynaptic activity.

For STF, we replaced p_{AB} by u_{AB} and expressed the STF dynamics as follows:

$$\frac{du_{AB}}{dt} = \frac{1 - u_{AB}}{\tau_u} + U_f(U_{max} - u_{AB})r_B, \tag{4}$$

where u_{AB} is a short-term facilitation variable that is constrained to the interval [1, U_{max}) for the synaptic connection from population B to population A. Biophysically, the short-term facilitation variable u represents the ability of releasing neurotransmitter, τ_u is the time constant of STF, U_f is the facilitation factor controlling the degree of facilitation induced by the presynaptic activity, and U_{max} is the maximal facilitation value.

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.

In addition, 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 quantities, 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

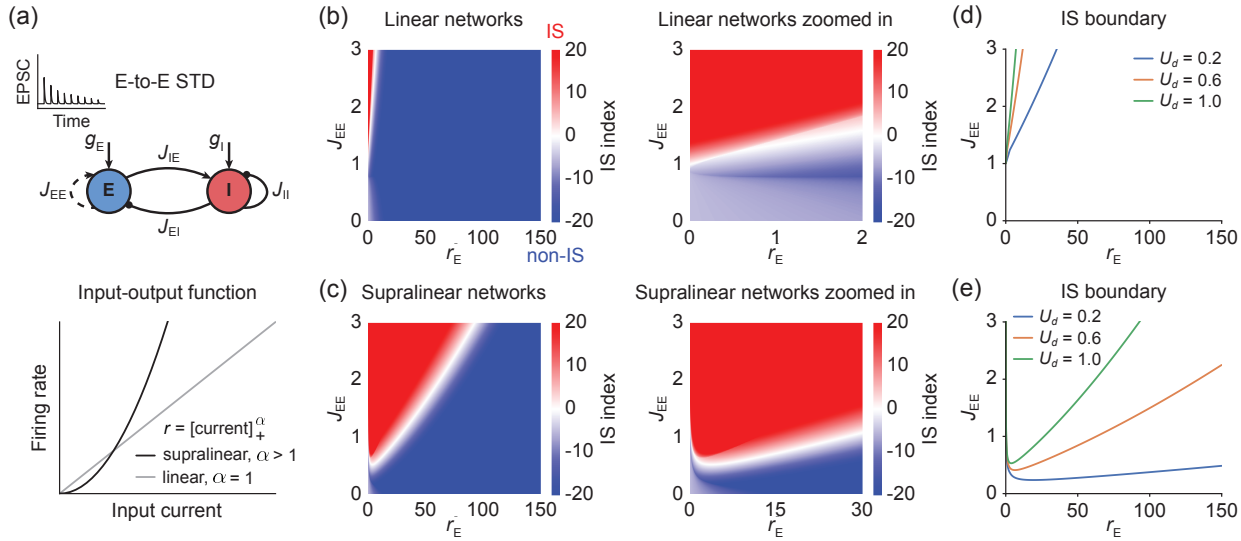

FIG. 1. Inhibition stabilization in recurrent neural networks with E-to-E short-term depression. (a) Top: Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STD. Bottom: Different nonlinearities of population-averaged responses to 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 factor 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 factor 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 depression factor U_d. (e) Same as (d) but for supralinear networks with E-to-E STD. Here, τ_x is 200 ms in (b-e).

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, 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). The non-monotonic transitions arise from the com-

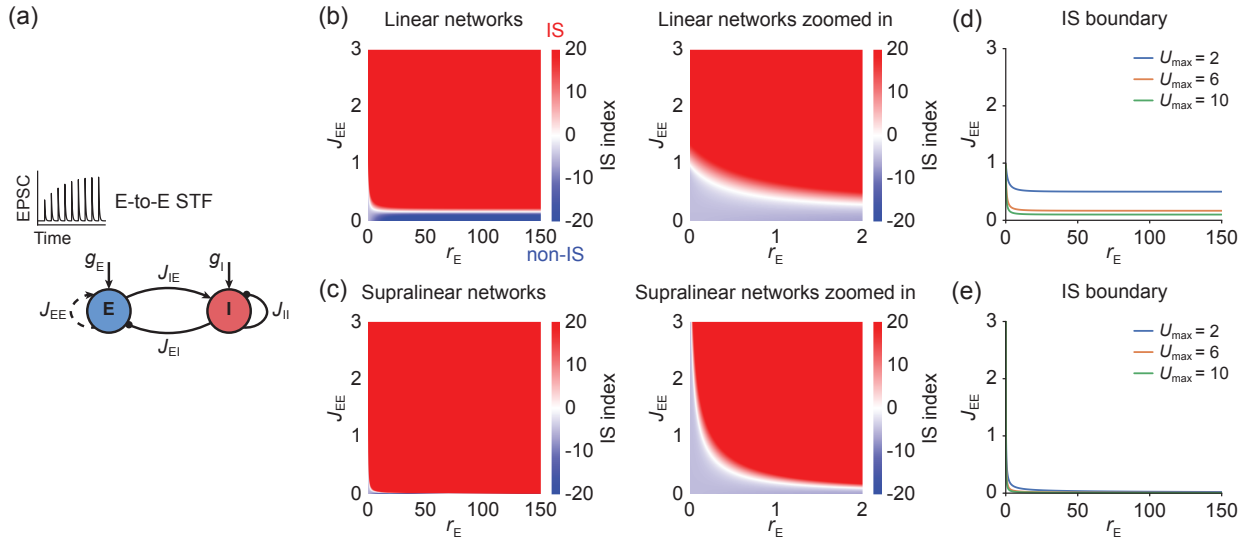

FIG. 2. Inhibition stabilization in recurrent neural networks with E-to-E short-term facilitation. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STF. (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory 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 factor 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 factor 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. Here, τ_u is 200 ms in (b-e).

petition between the increasing neuronal gain due to the supralinear neuronal input-output function and the decreasing synaptic strength due to STD. Despite high firing rates in the large r_E limit, E-to-E synaptic strengths are significantly depressed and STD effectively decouples excitatory neurons, rendering the network non-inhibition stabilized. Furthermore, the turning point of the IS boundary appears when $U_d x_{EE} r_E$ is of order 1 (Supplemental Methods). Increasing the depression factor U_d or the STD time constant τ_x shifts the turning point to the upper left in the (r_E, J_{EE}) coordinate system (Fig. S3; Supplemental Methods). We also found that these non-monotonic transitions cannot be observed in networks with static connectivity (Fig. S4; 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.

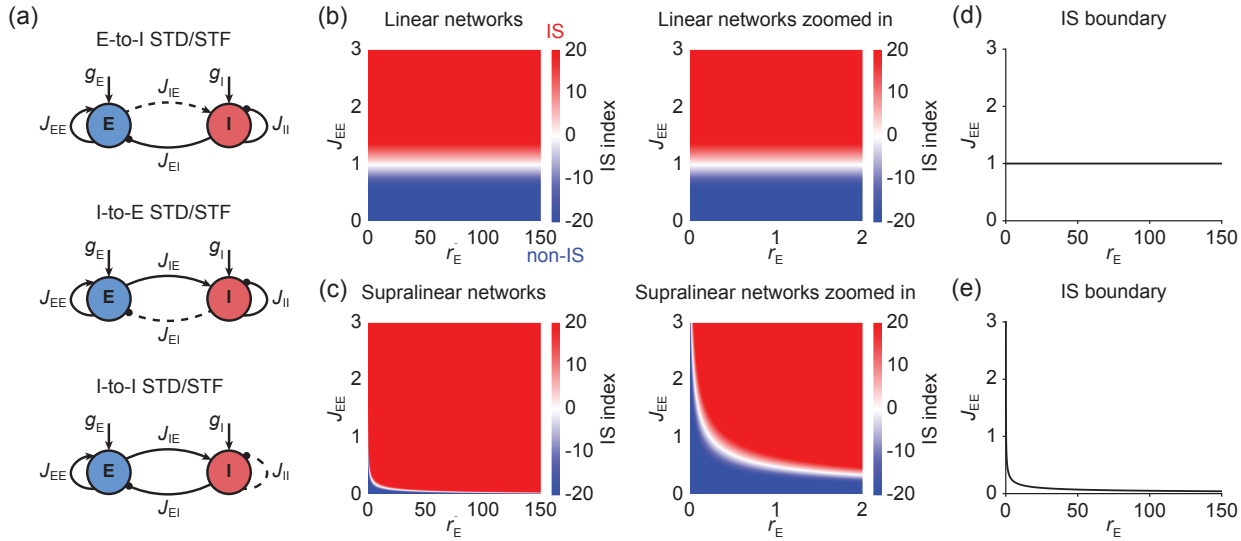

FIG. 3. Inhibition stabilization in recurrent neural networks with E-to-I short-term depression/facilitation, I-to-E short-term depression/facilitation, or I-to-I short-term depression/facilitation. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (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.

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. S4). 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 excitatory 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, 7]. Here, we sought to identify the conditions under which a paradoxical effect can

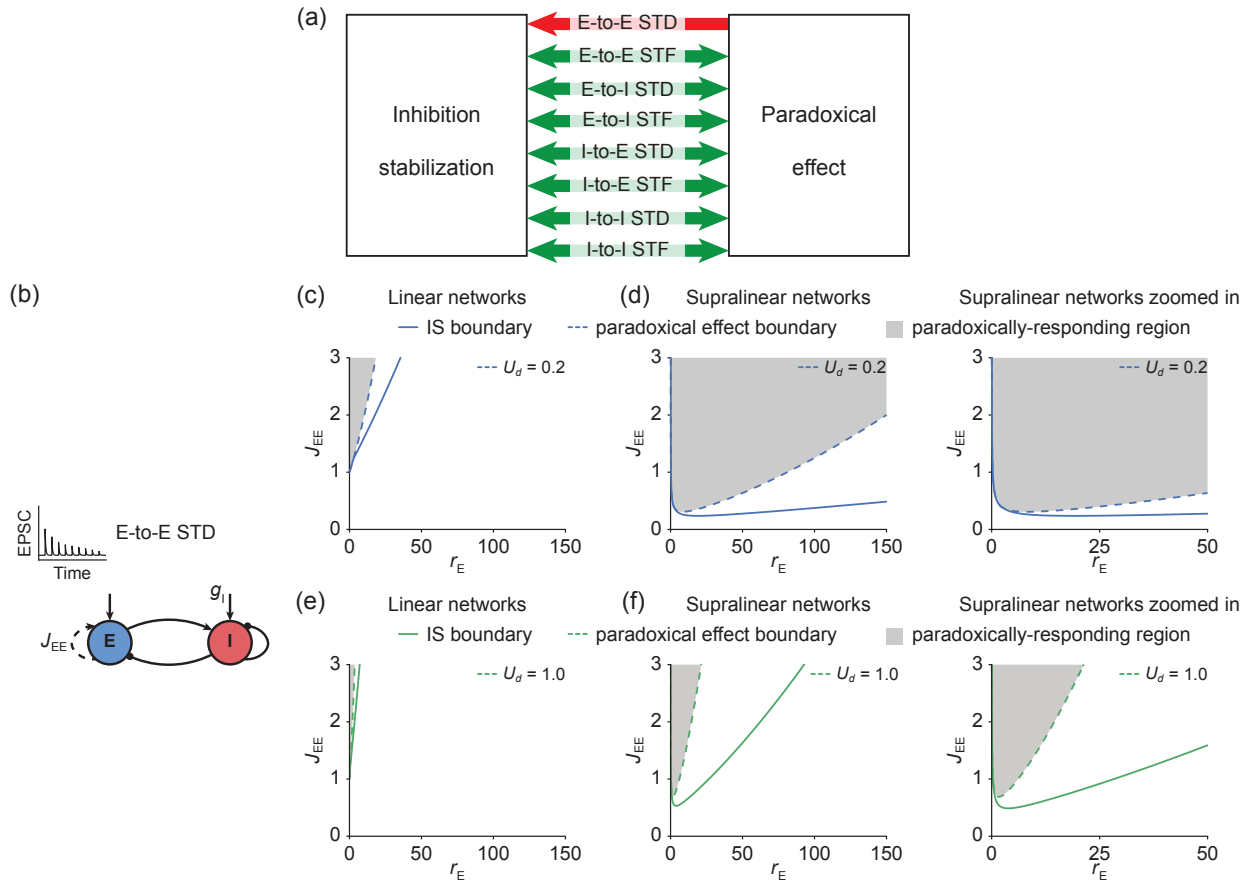

FIG. 4. Relationship between inhibition stabilization and the paradoxical effect in recurrent neural networks with short-term plasticity. (a) In networks with E-to-E STD, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not imply the paradoxical effect indicated 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 factor 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 factor U_d 1.0. Here, τ_x is 200 ms in (c-f).

arise in recurrent neural networks with STP. We assumed that the system is stable locally around 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 (abscissa) and inhibitory (ordinate) firing rate dynamics. The first condition involves a larger slope of the inhibitory nullcline than of the excitatory nullcline locally around the fixed point in the phase plane, while the second condition involves a positive slope of the excitatory nullcline around the fixed point (7; Fig. S5; 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). It is noteworthy that in the presence of E-to-E STD, networks in which dynamic STD regulation is required to ensure stability, as studied in [18, 19], are a subset of inhibition stabilized networks which do not exhibit paradoxical effects (Supplemental Methods).

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 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.

E. GENERALIZATION TO NETWORKS WITH A MIXTURE OF STP MECHANISMS AND MULTIPLE INTERNEURON SUBTYPES

So far, we only considered individual STP mechanisms one at a time. Here, we generalized our findings to four more complex and biologically realistic scenarios. First, our results can be generalized to networks with a mixture of STP mechanisms at other types of synapses except E-to-E synapses where we assume only STD or STF (Supplemental Methods). More specifically, the conditions and results derived for networks with either E-to-E STD or STF alone are the same for networks with either E-to-E STD orSTF and a mixture of STP mechanisms at other types of synapses.

Second, by incorporating both STD and STF at E-to-E synapses, we found that the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not necessarily imply the paradoxical effect (Supplemental Methods).

Third, inhibitory neurons in the cortex are highly diverse in terms of anatomy, electrophysiology and function [20–22]. Recent studies have investigated the relationship between inhibition stabilization and the paradoxical effect in networks with multiple interneuron subtypes in the absence of STP [23–27]. Yet, synapses between different interneuron subtypes exhibit considerable short-term dynamics [28]. We then extended the analysis to networks with multiple interneuron subtypes. Theoretical studies have shown that in networks with static connectivity, if the excitatory subnetwork is non-IS (IS), then the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are opposite (the same) [23]. We found that in the presence of E-to-E STD, if the network is IS, the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are the same (Supplemental Methods). However, the same sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population does not imply that the network is IS (Supplemental Methods). In the presence of E-to-E STF and STP at other types of synapses, inhibition stabilization and the same sign change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population imply each other (Supplemental Methods).

Last, the derived relationship between inhibition stabilization and the paradoxical effect is independent from the power of the power-law input-output function. Therefore, the derived relationship holds for any monotonically increasing neuronal nonlinearities including sublinear (Fig. S6, Supplemental Methods), despite the fact that sublinear networks have different IS transitions from supralinear networks (Fig. S7, Supplemental Methods). 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.

Network models applied to neural circuit development have previously investigated inhibition stabilization in the presence of STP [29]. 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 [30]. 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. Importantly, we generalized our results to several more complex scenarios including networks with a mixture of STP mechanisms, networks with both E-to-E STD and STF, networks with multiple interneuron subtypes, and any monotonically increasing neuronal nonlinearities.

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 [15, 31, 32]. While multistability and oscillations have been observed in the brain [33, 34], the single stable fixed point assumed in our study is believed to be a reasonable approximation of awake sensory cortex [35].

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 [36]. Furthermore, the degrees of balance between excitation and inhibition may also vary [37, 38], resulting different neuronal nonlinearities [11, 13, 38]. 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 [39], attention [40] and locomotion [41] 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 [42, 43]. 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, Felix Waitzmann, 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] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever, et al., "Reconstruction and Simulation of Neocortical Microcircuitry," Cell, vol. 163, no. 2, pp. 456–492, 2015.

[18] A. Loebel and M. Tsodyks, "Computation by Ensemble Synchronization in Recurrent Networks with Synaptic Depression," J Comput Neurosci, vol. 13, pp. 111–124, Sept. 2002.

[19] A. Loebel, I. Nelken, and M. Tsodyks, "Processing of sounds by population spikes in a model of primary auditory cortex," Frontiers in Neuroscience, vol. 1, p. 15, 2007.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] A. Mahrach, G. Chen, N. Li, C. van Vreeswijk, and D. Hansel, "Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation," eLife, vol. 9, pp. 1–37, 2020.

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

[26] 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.

[27] L. M. Richter and J. Gjorgjieva, "A circuit mechanism for independent modulation of excitatory and inhibitory firing rates after sensory deprivation," Proceedings of the National Academy of Sciences of the United States of America, vol. 119, no. 32, pp. 1–12, 2022.

[28] 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.

[29] V. Rahmati, K. Kirmse, K. Holthoff, L. Schwabe, and S. J. Kiebel, "Developmental Emergence of Sparse Coding: A Dynamic Systems Approach," Scientific Reports, vol. 7, no. 1, pp. 1–13, 2017.

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

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

[32] 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.

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

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

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

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

[37] 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.

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

[39] 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.

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

[41] 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.

[42] 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.

[43] H. Ko, L. Cossell, C. Baragli, J. Antolik, C. Clopath, S. B. Hofer, and T. D. Mrsic-Flogel, "The emergence of functional microcircuits in visual 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 factor
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 factor
U_{max}	6.0	-	Maximal facilitation value
Note that these values are also used elsewhere unless mentioned otherwise.			

Supplementary Figures

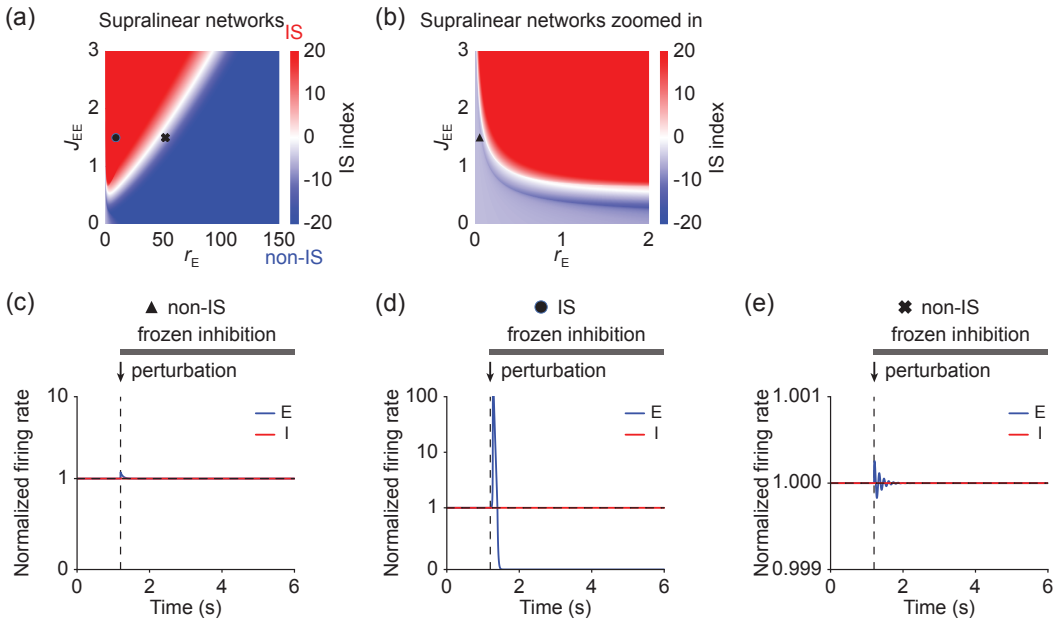

Figure S1. Examples of network activity in response to transient small perturbations to the excitatory activity while freezing inhibition at different initial states. (a) IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for supralinear networks 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).

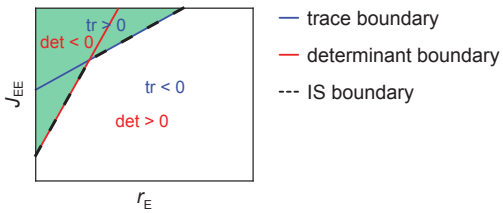

Figure S2. Schematic of IS boundary. Trace boundary (blue line) and determinant boundary (red line) correspond to the J_{EE} values at which the trace and the determinant of the Jacobian of the system with fixed inhibition are 0 for different r_E, respectively. IS boundary (black dashed line) is defined 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.

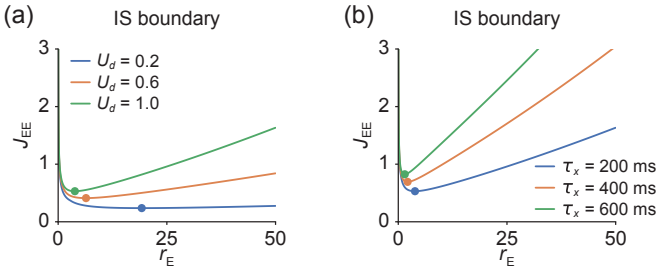

Figure S3. Turning points of the IS boundary in supralinear networks with E-to-E STD. (a) Turning points (dots) of the IS boundary for different values of depression factor U_d represented by different colors. Here, τ_x is 200 ms. (b) Same as (a) but for different values of STD time constant τ_x. Here, U_d is 1.

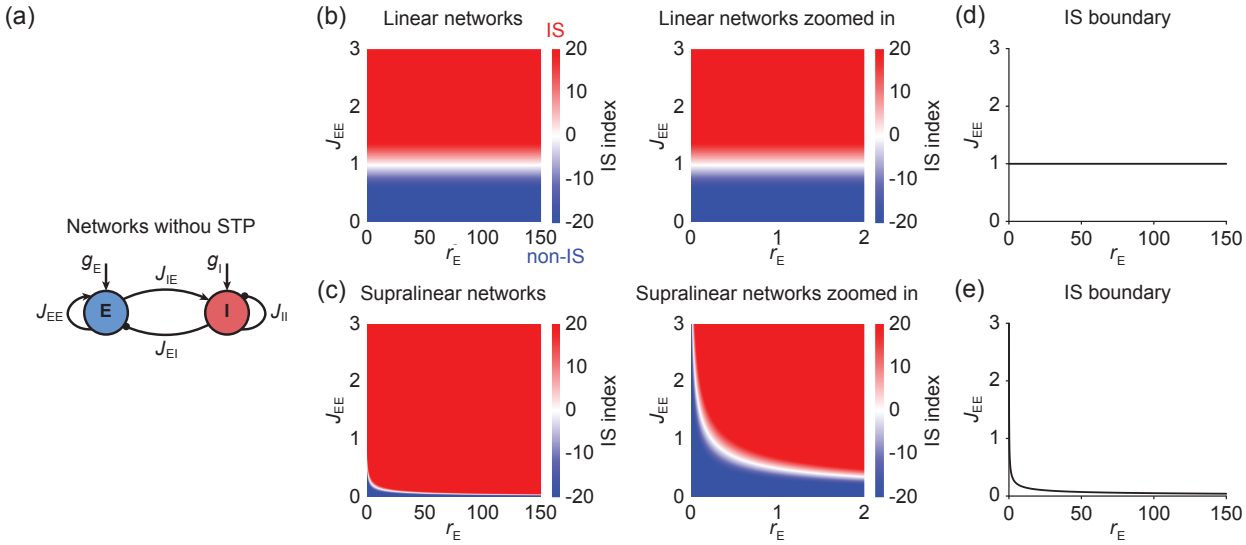

Figure S4. Inhibition stabilization in recurrent neural networks with static connectivity. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with static connectivity. (b) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory 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.

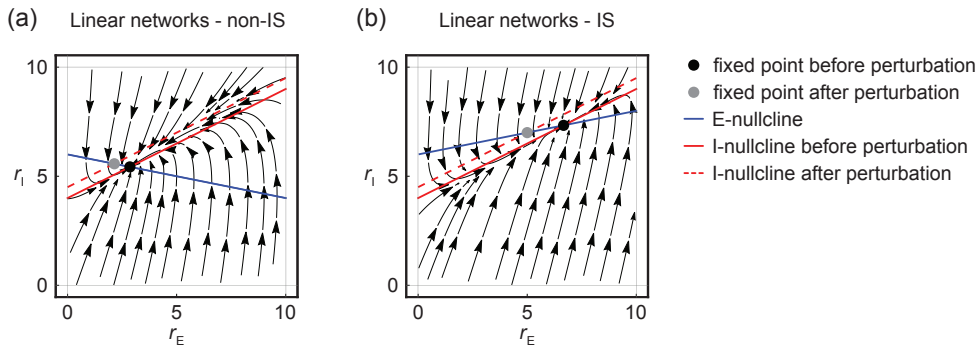

Figure S5. Conditions for the presence of the paradoxical effect from the angle of phase portraits. (a) Phase portrait of a linear non-IS network. The slope of the E-nullcline (blue) around the fixed point (black dot) is negative. The slope of the I-nucllcline (red) is larger than that of the E-nullcline. 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.

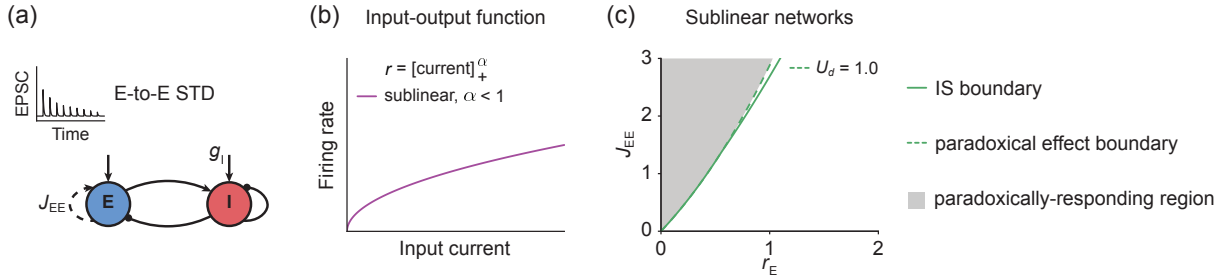

Figure S6. Relationship between inhibition stabilization and the paradoxical effect in sublinear networks with E-to-E STD. (a) Schematic of the recurrent network model consisting of an excitatory (blue) and an inhibitory population (red) with E-to-E STD. (b) Sublinear (purple) input-output function given by a rectified power law with exponent $\alpha = 0.5$. (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 sublinear networks with E-to-E STD. The paradoxically-responding region is marked in gray. Here, τ_x is 200 ms.

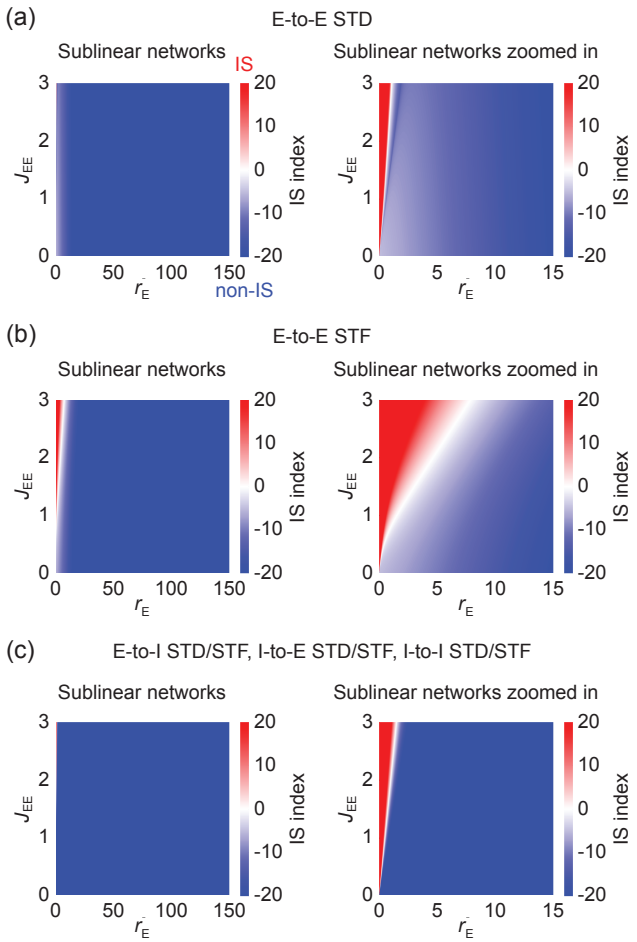

Figure S7. Inhibition stabilization in sublinear networks with short-term plasticity. (a) Left: IS index as a function of excitatory firing rate r_E and excitatory-to-excitatory connection strength J_{EE} for sublinear networks with E-to-E STD. The non-IS and IS regime are marked in blue and red, respectively. Right: zoomed-in version of (a) left. Here, the depression factor U_d is 1.0, τ_x is 200 ms, the firing rate r_E is plotted in a biologically realistic range from 0 to 150 Hz. (b) Same as (a) but for sublinear networks with E-to-E STF. Here, the facilitation factor U_f is 1.0, U_{max} is 6.0, and τ_u is 200 ms. (c) Same as (a) but for sublinear networks with E-to-I STD/STF, I-to-E STD/STF, or I-to-I STD/STF.

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 factor.

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 factor, 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 factor.

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

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

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

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

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

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

and

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

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

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

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

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

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

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

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

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

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

In the small r_E limit, we have:

$$\begin{aligned}-(\alpha_E-1)\tau_E - (\alpha_E-1)\tau_x + \tau_E\alpha_E\tau_x^2 U_d^2 r_E^2 + \tau_E\tau_x^2 U_d^2 r_E^2 + 2\tau_E\tau_x U_d r_E + \tau_x^2 U_d r_E \\ \approx -(\alpha_E-1)\tau_E - (\alpha_E-1)\tau_x < 0.\end{aligned} \tag{S0.20}$$

In the large r_E limit, the increasing neural gain-related term begins to dominate, and we have:

$$\begin{aligned}-(\alpha_E-1)\tau_E - (\alpha_E-1)\tau_x + \tau_E\alpha_E\tau_x^2 U_d^2 r_E^2 + \tau_E\tau_x^2 U_d^2 r_E^2 + 2\tau_E\tau_x U_d r_E + \tau_x^2 U_d r_E \\ \approx \tau_E\alpha_E\tau_x^2 U_d^2 r_E^2 + \tau_E\tau_x^2 U_d^2 r_E^2 + 2\tau_E\tau_x U_d r_E + \tau_x^2 U_d r_E > 0.\end{aligned} \tag{S0.21}$$

Therefore, J_{EE}^{tr} first decreases and then increases with r_E.

Similarly, we have:

$$\frac{dJ_{EE}^{\det}}{dr_E} = \frac{-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}})^2} \tag{S0.22}$$

and

$$-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} = \frac{1}{x^2} r_E^{-\frac{1}{\alpha_E}}(-(\alpha_E - 1) + \frac{2\tau_x U_d \alpha_E r_E}{1+\tau_x U_d r_E}). \tag{S0.23}$$

In the small r_E limit, the short-term depression variable $\frac{1}{1+\tau_x U_d r_E}$ is close to 1, whereas the neural gain-related term $2\tau_x U_d \alpha_E r_E$ is close to 0, so we have:

$$\frac{1}{x^2} r_E^{-\frac{1}{\alpha_E}}(-(\alpha_E - 1) + \frac{2\tau_x U_d \alpha_E r_E}{1+\tau_x U_d r_E}) \approx \frac{1}{x^2} r_E^{-\frac{1}{\alpha_E}}(-(\alpha_E - 1)) < 0. \tag{S0.24}$$

As r_E increases, the short-term depression variable $\frac{1}{1+\tau_x U_d r_E}$ tends to decrease the expression $-(\alpha_E - 1) + \frac{2\tau_x U_d \alpha_E r_E}{1+\tau_x U_d r_E}$, whereas the increasing neural gain-related term $2\tau_x U_d \alpha_E r_E$ tends to increase the expression $-(\alpha_E - 1) + \frac{2\tau_x U_d \alpha_E r_E}{1+\tau_x U_d r_E}$. In the large r_E limit, we have:

$$\frac{1}{x^2} r_E^{-\frac{1}{\alpha_E}}(-(\alpha_E - 1) + \frac{2\tau_x U_d \alpha_E r_E}{1+\tau_x U_d r_E}) \approx \frac{1}{x^2} r_E^{-\frac{1}{\alpha_E}}(-(\alpha_E - 1) + 2\alpha_E) > 0. \tag{S0.25}$$

Therefore, $J_{EE}^{\det}$ first decreases and then increases with r_E.

In sum, for supralinear networks with E-to-E STD, the IS boundary changes non-monotonically with r_E. More specifically, it first shifts downwards and then shifts upwards as r_E increases.

For $\alpha_E < 1$, we have

$$\begin{aligned}\frac{dJ_{EE}^{\mathrm{tr}}}{dr_E} &= \frac{-(\tau_E \tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E \tau_x U_d^2 r_E^2 + U_d \tau_x r_E)(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (2\tau_E U_d + 2\tau_E \tau_x U_d^2 r_E + U_d \tau_x)\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}})^2} \\ &= \frac{-(\alpha_E - 1)\tau_E - (\alpha_E - 1)\tau_x + \tau_E \alpha_E \tau_x^2 U_d^2 r_E^2 + \tau_E \tau_x^2 U_d^2 r_E^2 + 2\tau_E \tau_x U_d r_E + \tau_x^2 U_d r_E}{\tau_x (\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}})^2} r_E^{\frac{\alpha_E - 1}{\alpha_E}} > 0\end{aligned} \tag{S0.26}$$

and

$$\frac{dJ_{EE}^{\det}}{dr_E} = \frac{-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}})^2} > 0. \tag{S0.27}$$

Therefore, for sublinear networks with E-to-E STD, the IS boundary changes monotonically with r_E. More specifically, it always shifts upwards as r_E increases.

Turning point of the IS boundary for networks with E-to-E

Here, we sought to investigate how the turning point of the IS boundary changes with the short-term depression factor U_d and the short-term depression time constant τ_x. Assuming that around the turning point, J_{EE}^{IS} is some smooth function of r_E, the turning point r_E^* of the IS boundary can then be defined as follows:

$$\frac{dJ_{EE}^{IS}}{dr_E}\Big|_{r_E^*} = \frac{d\min(J_{EE}^{\text{tr}}, J_{EE}^{\text{det}})}{dr_E}\Big|_{r_E^*} = 0. \tag{S0.28}$$

Furthermore, the turning point r_E^* of the IS boundary can be described as the minimum of the turning point $r_{E_1}^*$ of J_{EE}^{tr} and the turning point $r_{E_2}^*$ of J_{EE}^{det}, i.e.,

$$r_E^* = \min(r_{E_1}^*, r_{E_2}^*), \tag{S0.29}$$

with

$$\frac{dJ_{EE}^{\text{tr}}}{dr_E}\Big|_{r_{E_1}^*} = 0 \tag{S0.30}$$

$$\frac{dJ_{EE}^{\text{det}}}{dr_E}\Big|_{r_{E_2}^*} = 0, \tag{S0.31}$$

$$\frac{dJ_{EE}^{\text{tr}}}{dr_E} = \frac{-(\tau_E\tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E\tau_x U_d^2 r_E^2 + U_d\tau_x r_E)(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (2\tau_E U_d + 2\tau_E\tau_x U_d^2 r_E + U_d\tau_x)\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}})^2}. \tag{S0.32}$$

Therefore, we have

$$r_{E_1}^* = \frac{-(2\tau_E + \tau_x) + \sqrt{4\alpha_E^2\tau_E^2 + 4\alpha_E^2\tau_E\tau_x + \tau_x^2}}{2(\alpha_E + 1)\tau_E\tau_x U_d}. \tag{S0.33}$$

The turning point $r_{E_1}^*$ of J_{EE}^{tr} can therefore be expressed as a function of U_d and τ_x. To identify how U_d and τ_x affect $r_{E_1}^*$, we computed the following derivatives:

$$\frac{dr_{E_1}^*}{dU_d} = \frac{(2\tau_E + \tau_x) - \sqrt{4\alpha_E^2\tau_E^2 + 4\alpha_E^2\tau_E\tau_x + \tau_x^2}}{2(\alpha_E + 1)\tau_E\tau_x U_d^2}, \tag{S0.34}$$

$$\frac{dr_{E_1}^*}{d\tau_x} = \frac{\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2} - \alpha_E^2(2\tau_E + \tau_x)}{(\alpha_E + 1)\tau_x^2 U_d\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2}}. \tag{S0.35}$$

For $\alpha > 1$, we have:

$$\frac{dr_{E_1}^*}{dU_d} = \frac{(2\tau_E + \tau_x) - \sqrt{4\alpha_E^2\tau_E^2 + 4\alpha_E^2\tau_E\tau_x + \tau_x^2}}{2(\alpha_E + 1)\tau_E\tau_x U_d^2} < \frac{(2\tau_E + \tau_x) - \sqrt{4\tau_E^2 + 4\tau_E\tau_x + \tau_x^2}}{2(\alpha_E + 1)\tau_E\tau_x U_d^2} = 0 \tag{S0.36}$$

$$\begin{aligned}\frac{dr_{E_1}^*}{d\tau_x} &= \frac{\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2} - \alpha_E^2(2\tau_E + \tau_x)}{(\alpha_E + 1)\tau_x^2 U_d\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2}} \\ &= \frac{4\tau_E^2\alpha_E^2(1 - \alpha_E^2) + 4\tau_E\tau_x(1 - \alpha_E^2) + \tau_x^2(1 - \alpha_E^4)}{((\alpha_E + 1)\tau_x^2 U_d\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2})(\sqrt{4\tau_E\tau_x\alpha_E^2 + 4\tau_E^2\alpha_E^2 + \tau_x^2} + \alpha_E^2(2\tau_E + \tau_x))} < 0.\end{aligned} \tag{S0.37}$$

Therefore, for $\alpha > 1$, increasing either U_d or τ_x shifts the turning point $r_{E_1}^*$ of J_{EE}^{tr} to the left in the coordinate system in which the abscissa is r_E and the ordinate is J_{EE}.
Furthermore,

$$U_d\tau_x r_{E_1}^* = \frac{-(2\tau_E + \tau_x) + \sqrt{4\alpha_E^2\tau_E^2 + 4\alpha_E^2\tau_E\tau_x + \tau_x^2}}{2(\alpha_E + 1)\tau_E}. \tag{S0.38}$$

When α_E is of order 1, then $U_d\tau_x r_{E_1}^*$ is approximately of the same order as $\frac{-(2\tau_E+\tau_x)+\sqrt{4\alpha_E^2\tau_E^2+4\alpha_E\tau_E\tau_x+\tau_x^2}}{2(\alpha_E+1)\tau_E}$, which is of order 1. Therefore, the turning point $r_{E_1}^*$ of J_{EE}^{tr} appears when $U_d\tau_x r_{E_1}^*$ is of order 1.

Similarly, we have:

$$\frac{dJ_{EE}^{\text{det}}}{dr_E} = \frac{-(1+\tau_x U_d r_E)^2(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (1+\tau_x U_d r_E)2\tau_x U_d \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}{(\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}})^2}, \tag{S0.39}$$

$$r_{E_2}^* = \frac{\alpha_E - 1}{\tau_x U_d (1+\alpha_E)}. \tag{S0.40}$$

Therefore, for $\alpha > 1$, increasing either U_d or τ_x shifts the turning point $r_{E_2}^*$ of J_{EE}^{det} to the left in the coordinate system in which the abscissa is r_E and the ordinate is J_{EE}.
Moreover,

$$U_d \tau_x r_{E_2}^* = \frac{\alpha_E - 1}{\alpha_E + 1}. \tag{S0.41}$$

Therefore, the turning point $r_{E_2}^*$ of J_{EE}^{det} appears when $U_d \tau_x r_{E_2}^*$ is of order 1.

As the turning point $r_{E_1}^*$ of J_{EE}^{tr} and the turning point $r_{E_2}^*$ of J_{EE}^{det} are affected by U_d and τ_x in the same way, increasing either U_d or τ_x also shifts the turning point r_E^* of the IS boundary to the left in the same coordinate system. Furthermore, both $r_{E_1}^*$ and $r_{E_2}^*$ appear when $U_d \tau_x r_{E_1}^*$ and $U_d \tau_x r_{E_2}^*$ are of order 1, respectively, the turning point r_E^* of the IS boundary therefore appears when $U_d \tau_x r_E^*$ is of order 1.

In addition, since

$$J_{EE}^{\text{tr}} = \frac{\tau_E \tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E \tau_x U_d^2 r_E^2 + U_d \tau_x r_E}{\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}, \tag{S0.42}$$

$$J_{EE}^{\text{det}} = \frac{(1+\tau_x U_d r_E)^2}{\alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}}}, \tag{S0.43}$$

it is obvious that increasing U_d shifts the turning point $r_{E_1}^*$ up in the (r_E, J_{EE}) coordinate system, and increasing U_d or τ_x shifts the turning point $r_{E_2}^*$ up. For fixed r_E, we have:

$$\frac{d}{d\tau_x}(\tau_E \tau_x^{-1} + 1 + 2\tau_E U_d r_E + \tau_E \tau_x U_d^2 r_E^2 + U_d \tau_x r_E) = \tau_x^{-1}(-\tau_E \tau_x^{-1} + \tau_E \tau_x U_d^2 r_E^2 + \tau_x U_d r_E). \tag{S0.44}$$

The rate dynamics is typically an order of magnitude larger than the dynamics of the short-term plasticity mechanisms. As we derived before, at the turning point, $\tau_x U_d r_E$ is of order 1. So the above expression is positive in biologically realistic regimes, indicating that increasing τ_x shifts the turning point $r_{E_1}^*$ up in the (r_E, J_{EE}) coordinate system. Therefore, increasing either U_d or τ_x also shifts the turning point r_E^* of the IS boundary up in that coordinate system.

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[\frac{1}{1+\tau_x U_d r_E} J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.45}$$

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

$$r_I = \frac{\frac{1}{1+\tau_x U_d r_E} J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.46}$$

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

$$k_{STD}^E = \frac{1}{J_{EI}}\Big[-\frac{J_{EE}}{(1+\tau_x U_d r_E)^2}\tau_x U_d r_E + \frac{J_{EE}}{1+\tau_x U_d r_E} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big]. \tag{S0.47}$$

Note that as the short-term depression variable x_{EE} depends on r_E and changes dynamically, the slope of the excitatory nullcline k_{STD}^E also contains the term $-\frac{J_{EE}}{(1+\tau_x U_d r_E)^2}\tau_x U_d r_E$, which is the product of $J_{EE} r_E$ and the derivative of short-term depression variable x_{EE} with respect to r_E.

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

$$x > \sqrt{\frac{1}{J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}}. \tag{S0.48}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.49}$$

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

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.50}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.51}$$

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

$$\begin{aligned} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}(\tau_x^{-1} + U_d r_E) > (1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})\Big[-\frac{J_{EE}U_d r_E}{1+\tau_x U_d r_E}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ + \frac{J_{EE}}{1+\tau_x U_d r_E}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}(\tau_x^{-1} + U_d r_E) - (\tau_x^{-1} + U_d r_E)\Big]. \end{aligned} \tag{S0.52}$$

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

$$\begin{aligned}\det(\mathbf{N_{STD}^{EE}}) = \tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1} + U_d r_E) \\ -\tau_E^{-1}J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}\tau_I^{-1}J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}(\tau_x^{-1} + U_d r_E) + \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})U_d x < 0.\end{aligned} \tag{S0.53}$$

The condition shown in Eq. S0.52 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. S0.48, 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.

Furthermore, for networks with E-to-E STD exhibiting no paradoxical effects, $\det(\mathbf{N_{STD}^{EE}})$ is larger than 0, indicating that the excitatory nullcline has a negative slope around the fixed point in the r_E/r_I plane where x axis is r_E and y axis is r_I. However, the networks can be inhibition stabilized when

$$\mathrm{tr}(\mathbf{N_{STD}^{EE}}) = \tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E > 0. \tag{S0.54}$$

Then, $\tau_E^{-1}(xJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)$ is also positive, suggesting that the excitatory subnetwork is unstable if STD variables and inhibition are fixed. Therefore, in the presence of E-to-E STD, networks in which dynamic STD regulation is required to ensure stability as studied in [18, 19] are a subset of inhibition stabilized networks which do not exhibit paradoxical effects.

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

$$k_{STD}^{E} = \frac{1}{J_{EI}}\Big(- \frac{J_{EE}^{PD}}{(1+\tau_x U_d r_E)^2}\tau_x U_d r_E + \frac{J_{EE}^{PD}}{1+\tau_x U_d r_E} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big) = 0, \tag{S0.55}$$

$$J_{EE}^{PD} = \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}(1+\tau_x U_d r_E)^2. \tag{S0.56}$$

For $\alpha_E = 1$, we have

$$\frac{dJ_{EE}^{PD}}{dr_E} = 2\tau_x U_d + 2\tau_x^2 U_d^2 r_E. \tag{S0.57}$$

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

$$\begin{aligned}\frac{dJ_{EE}^{PD}}{dr_E} =& \frac{1}{\alpha_E}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{\frac{1-2\alpha_E}{\alpha_E}} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{\frac{1-\alpha_E}{\alpha_E}} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E} r_E^{\frac{1}{\alpha_E}}\Big) \\ =& \frac{1}{\alpha_E} r_E^{-\frac{1}{\alpha_E}}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{-2} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{-1} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E}\Big).\end{aligned} \tag{S0.58}$$

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.

For $\alpha_E < 1$, we have

$$\begin{aligned}\frac{dJ_{EE}^{PD}}{dr_E} =& \frac{1}{\alpha_E}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{\frac{1-2\alpha_E}{\alpha_E}} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{\frac{1-\alpha_E}{\alpha_E}} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E} r_E^{\frac{1}{\alpha_E}}\Big) \\ =& \frac{1}{\alpha_E} r_E^{-\frac{1}{\alpha_E}}\Big(\frac{1-\alpha_E}{\alpha_E} r_E^{-2} + 2\tau_x U_d \frac{1}{\alpha_E} r_E^{-1} + \tau_x^2 U_d^2 \frac{1+\alpha_E}{\alpha_E}\Big) > 0.\end{aligned} \tag{S0.59}$$

Therefore, for sublinear networks with E-to-E STD, the paradoxical effect boundary always 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[uJ_{EE}r_E - J_{EI}r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.60}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE}r_E - J_{II}r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.61}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f(U_{max}-u)r_E, \tag{S0.62}$$

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 factor.

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

$$\mathbf{M^{EE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1}J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1}J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ U_f(U_{max}-u) & 0 & -\tau_u^{-1} - U_f r_E \end{bmatrix}. \tag{S0.63}$$

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

$$\mathbf{N^{EE}_{STF}} = \begin{bmatrix} \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ U_f(U_{max}-u) & -\tau_u^{-1} - U_f r_E \end{bmatrix}. \tag{S0.64}$$

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

$$\mathrm{tr}(\mathbf{N^{EE}_{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E < 0 \tag{S0.65}$$

and

$$\det(\mathbf{N^{EE}_{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) - \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max}-u) > 0. \tag{S0.66}$$

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

$$\frac{(1 + 2U_f U_{max} r_E \tau_u + U_f r_E \tau_u U_f U_{max} r_E \tau_u) J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}{(1+\tau_u U_f r_E)^2} - 1 > 0. \tag{S0.67}$$

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

$$\begin{aligned} \text{IS index} = \mathrm{Re}\Bigg[&\frac{\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E}{2} \\ &+ \sqrt{\frac{1}{4}[\tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) + \tau_u^{-1} + U_f r_E]^2 + \tau_E^{-1}J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max}-u)}\Bigg]. \end{aligned} \tag{S0.68}$$

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

$$\mathrm{tr}(\mathbf{N}_{\mathbf{STF}}^{\mathbf{EE}}) = \tau_E^{-1}(uJ_{EE}^{\mathrm{tr}}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E = 0, \tag{S0.69}$$

$$J_{EE}^{\mathrm{tr}} = \frac{(\tau_E\tau_u^{-1} + \tau_E U_f r_E + 1)(1 + U_f r_E \tau_u)}{(1 + U_f U_{max} r_E \tau_u)\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}}, \tag{S0.70}$$

$$\det(\mathbf{N}_{\mathbf{STF}}^{\mathbf{EE}}) = \tau_E^{-1}(-\tau_u^{-1}\frac{1 + U_f U_{max} r_E \tau_u}{1 + U_f r_E \tau_u} J_{EE}^{\det}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u^{-1} + U_f r_E - J_{EE}^{\det}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}) = 0, \tag{S0.71}$$

$$J_{EE}^{\det} = \frac{1 + \tau_u U_f r_E}{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}}. \tag{S0.72}$$

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

$$J_{EE}^{IS} = \min(J_{EE}^{\mathrm{tr}}, J_{EE}^{\det}) = J_{EE}^{\det}. \tag{S0.73}$$

For $\alpha_E = 1$, we have

$$\frac{dJ_{EE}^{\det}}{dr_E} = \frac{dJ_{EE}^{IS}}{dr_E} = \frac{2U_f\tau_u(1 - U_{max})}{(1 + U_f r_E \tau_u)(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} + \tau_u r_E U_f U_{max})^2}. \tag{S0.74}$$

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

$$\begin{aligned}\frac{dJ_{EE}^{\det}}{dr_E} = \frac{dJ_{EE}^{IS}}{dr_E} = &-\frac{1 + \tau_u U_f r_E}{(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})^2}(-\frac{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f\tau_u + U_f U_{max} r_E \tau_u \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u}{(1 + U_f r_E \tau_u)^2} \\ &+ \frac{(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + 2U_f U_{max}\tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u}{1 + U_f r_E \tau_u}) \\ &+ \frac{\tau_u U_f(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})}{(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max})^2}.\end{aligned} \tag{S0.75}$$

Further,

$$\begin{aligned}&-(1 + \tau_u U_f r_E)(-\frac{\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f\tau_u + U_f U_{max} r_E \tau_u \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f \tau_u}{(1 + U_f r_E \tau_u)^2} \\ &+ \frac{(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + 2U_f U_{max}\tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u(2\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u}{1 + U_f r_E \tau_u}) \\ &+ \tau_u U_f(\frac{1 + U_f U_{max} r_E \tau_u}{1 + U_f r_E \tau_u}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} + \tau_u \alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f U_{max}) \\ &= (-(\alpha_E - 1)r_E^{-\frac{1}{\alpha_E}} + (-2U_{max}\alpha_E + 2U_{max} - U_{max}\alpha_E + \alpha_E - U_{max}\alpha_E + 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f\tau_u \\ &-2U_f U_{max}\tau_u(\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f r_E \tau_u - 2\tau_u(\alpha_E - 1)r_E^{\frac{\alpha_E-1}{\alpha_E}} U_f U_{max} U_f r_E \tau_u(1 + U_f r_E \tau_u))/(1 + U_f r_E \tau_u).\end{aligned} \tag{S0.76}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.77}$$

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

$$r_I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.78}$$

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

$$k_{STF}^E = \frac{1}{J_{EI}}\Big(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \Big(\frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2}\Big) J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big). \tag{S0.79}$$

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

$$\frac{1}{J_{EI}}\Big(\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big) > 0. \tag{S0.80}$$

The above condition can be simplified as follows:

$$\frac{1+2U_f U_{max} r_E \tau_u + U_f U_{max} r_E \tau_u U_f r_E \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0. \tag{S0.81}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.82}$$

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

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.83}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.84}$$

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

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{EE} + \frac{U_f U_{max} \tau_u (1+U_f r_E \tau_u) - (1+U_f U_{max} r_E \tau_u) U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{EE} r_E - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\Big]. \tag{S0.85}$$

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

$$\det(\mathbf{N}^{\mathbf{EE}}_{\mathbf{STF}}) = \tau_E^{-1}(uJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) - \tau_E^{-1} J_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} U_f(U_{max} - u) < 0. \tag{S0.86}$$

The condition shown in Eq. S0.85 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. S0.67 is identical to the condition of having paradoxical effect shown in Eq. S0.81, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.87}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[x J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.88}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_E. \tag{S0.89}$$

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 factor.

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

$$\mathbf{M}^{\mathbf{IE}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} x J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} \\ -U_d x & 0 & -\tau_x^{-1} - U_d r_E \end{bmatrix}. \tag{S0.90}$$

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

$$\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ -U_d x & -\tau_x^{-1} - U_d r_E \end{bmatrix}. \tag{S0.91}$$

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

$$\mathrm{tr}(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_E < 0 \tag{S0.92}$$

and

$$\det(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_E) > 0. \tag{S0.93}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.94}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.95}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.96}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.97}$$

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

$$k_{STD}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.98}$$

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

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0. \tag{S0.99}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[\frac{1}{1+U_d r_E \tau_x} J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.100}$$

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

$$r_I = \frac{\frac{1}{1+U_d r_E \tau_x} J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.101}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{\frac{1}{(1+U_d r_E \tau_x)^2} J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.102}$$

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

$$\frac{\frac{1}{(1+U_d r_E \tau_x)^2} J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.103}$$

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

$$\begin{aligned}\det(\mathbf{M_{STD}^{IE}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1} + U_d r_E) \\ +\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}[(\tau_I^{-1} x J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(-\tau_x^{-1} - U_d r_E) + \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_d x] < 0.\end{aligned} \tag{S0.104}$$

The condition shown in Eq. S0.103 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. S0.94 is identical to the condition of having paradoxical effect shown in Eq. S0.99, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.105}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[u J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.106}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_E, \tag{S0.107}$$

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 factor.

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

$$\mathbf{M}^{\mathbf{IE}}_{\mathbf{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} u J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} \\ U_f(U_{max} - u) & 0 & -\tau_u^{-1} - U_f r_E \end{bmatrix}. \tag{S0.108}$$

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

$$\mathbf{N}^{\mathbf{IE}}_{\mathbf{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_E \end{bmatrix}. \tag{S0.109}$$

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

$$\mathrm{tr}(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_E < 0 \tag{S0.110}$$

and

$$\det(\mathbf{N}^{\mathbf{IE}}_{\mathbf{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_E) > 0. \tag{S0.111}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.112}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.113}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.114}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.115}$$

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

$$k_{STF}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.116}$$

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

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0. \tag{S0.117}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.118}$$

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

$$r_I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.119}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STD}^I = \frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} + \frac{U_f U_{max}\tau_u - U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{IE} r_E}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.120}$$

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

$$\frac{\frac{1+U_f U_{max} r_E \tau_u}{1+U_f r_E \tau_u} J_{IE} + \frac{U_f U_{max}\tau_u - U_f \tau_u}{(1+U_f r_E \tau_u)^2} J_{IE} r_E}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.121}$$

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

$$\begin{aligned}\det(\mathbf{M_{STF}^{IE}}) &= \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_E) \\ &+\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}[(\tau_I^{-1} u J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(-\tau_u^{-1} - U_f r_E) - \tau_I^{-1} J_{IE} r_E \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_f(U_{max} - u)] < 0.\end{aligned} \tag{S0.122}$$

The condition shown in Eq. S0.121 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. S0.112 is identical to the condition of having paradoxical effect shown in Eq. S0.117, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - x J_{EI} r_I + g_E\Big]_+^{\alpha_E} \tag{S0.123}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I} \tag{S0.124}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_I, \tag{S0.125}$$

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 factor.

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

$$\mathbf{M}^{\mathbf{EI}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} x J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ 0 & -U_d x & -\tau_x^{-1} - U_d r_I \end{bmatrix}. \tag{S0.126}$$

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

$$\mathbf{N}^{\mathbf{EI}}_{\mathbf{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ 0 & -\tau_x^{-1} - U_d r_I \end{bmatrix}. \tag{S0.127}$$

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

$$\mathrm{tr}(\mathbf{N}^{\mathbf{EI}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_I < 0 \tag{S0.128}$$

and

$$\det(\mathbf{N}^{\mathbf{EI}}_{\mathbf{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_I) > 0. \tag{S0.129}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.130}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.131}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \left[J_{EE} r_E - \frac{1}{1+U_d r_I \tau_x} J_{EI} r_I + g_E \right]_+^{\alpha_E} = 0. \tag{S0.132}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{\frac{1}{1+U_d r_I \tau_x} J_{EI}}. \tag{S0.133}$$

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

$$k_{STF}^E = \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{EI} r_I + \frac{1}{1+U_d r_I \tau_x} J_{EI}}. \tag{S0.134}$$

Provided that the inhibitory current to the excitatory network is always an increasing function (at least locally around the fixed point) of the inhibitory firing rate, we have

$$-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{EI} r_I + \frac{1}{1+U_d r_I \tau_x} J_{EI} > 0. \tag{S0.135}$$

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

$$J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0. \tag{S0.136}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \left[J_{IE} r_E - J_{II} r_I + g_I \right]_+^{\alpha_I} = 0. \tag{S0.137}$$

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

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.138}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.139}$$

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

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{EI} r_I + \frac{1}{1+U_d r_I \tau_x} J_{EI}}. \tag{S0.140}$$

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

$$\begin{aligned} \det(\mathbf{M_{STD}^{EI}}) = (J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1)(1 + J_{II} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}})(1 + \tau_x U_d r_I) \\ -x J_{EI} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} J_{IE} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} (1 + \tau_x U_d r_I) + J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} J_{IE} \alpha_I r_I^{\frac{\alpha_I - 1}{\alpha_I}} \tau_x U_d x < 0. \end{aligned} \tag{S0.141}$$

The condition shown in Eq. S0.140 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. S0.130 is identical to the condition of having paradoxical effect shown in Eq. S0.136, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - u J_{EI} r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.142}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - J_{II} r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.143}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_I, \tag{S0.144}$$

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 factor.

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

$$\mathbf{M^{EI}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} u J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 \\ 0 & U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_I \end{bmatrix}. \tag{S0.145}$$

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

$$\mathbf{N^{EI}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \\ 0 & -\tau_u^{-1} - U_f r_I \end{bmatrix}. \tag{S0.146}$$

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

$$\mathrm{tr}(\mathbf{N^{EI}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_I < 0 \tag{S0.147}$$

and

$$\det(\mathbf{N^{EI}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_I) > 0. \tag{S0.148}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.149}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.150}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \left[J_{EE} r_E - \frac{1 + U_f U_{max} r_I \tau_u}{1 + U_f r_I \tau_u} J_{EI} r_I + g_E \right]_+^{\alpha_E} = 0. \tag{S0.151}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI}}. \tag{S0.152}$$

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

$$k_{STF}^E = \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{EI} r_I}. \tag{S0.153}$$

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

$$J_{EE} \alpha_E r_E^{\frac{\alpha_E - 1}{\alpha_E}} - 1 > 0. \tag{S0.154}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \left[J_{IE} r_E - J_{II} r_I + g_I \right]_+^{\alpha_I} = 0. \tag{S0.155}$$

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

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{J_{II}}. \tag{S0.156}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.157}$$

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

$$\frac{J_{IE}}{J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{EI} + \frac{U_f U_{max} \tau_u - U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{EI} r_I}. \tag{S0.158}$$

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

$$\begin{aligned} \det(\mathbf{M_{STF}^{EI}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)\tau_I^{-1}(1 + J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_I) \\ + \tau_E^{-1} u J_{EI} \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE} \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (-\tau_u^{-1} - U_f r_I) - \tau_E^{-1} J_{EI} r_I \alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE} \alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} U_f (U_{max} - u) < 0. \end{aligned} \tag{S0.159}$$

The condition shown in Eq. S0.158 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. S0.149 is identical to the condition of having paradoxical effect shown in Eq. S0.154, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.160}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - x J_{II} r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.161}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - U_d x r_I, \tag{S0.162}$$

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 factor.

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

$$\mathbf{M^{II}_{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + x J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & -\tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} \\ 0 & -U_d x & -\tau_x^{-1} - U_d r_I \end{bmatrix}. \tag{S0.163}$$

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

$$\mathbf{N^{II}_{STD}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ 0 & -\tau_x^{-1} - U_d r_I \end{bmatrix}. \tag{S0.164}$$

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

$$\mathrm{tr}(\mathbf{N^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - U_d r_I < 0 \tag{S0.165}$$

and

$$\det(\mathbf{N^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_I) > 0. \tag{S0.166}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.167}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.168}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.169}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.170}$$

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

$$k^E_{STD} = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.171}$$

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

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0. \tag{S0.172}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} \quad = \quad -r_I + \Big[J_{IE} r_E - \frac{1}{1+U_d r_I \tau_x} J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.173}$$

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

$$r_I \quad = \quad \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{\frac{1}{1+U_d r_I \tau_x} J_{II}}. \tag{S0.174}$$

The slope of the inhibitory nullcline can be written as follows

$$k^I_{STD} = \frac{J_{IE}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{II} r_I + \frac{1}{1+U_d r_I \tau_x} J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.175}$$

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

$$\frac{J_{IE}}{-\frac{1}{(1+U_d r_I \tau_x)^2} U_d \tau_x J_{II} r_I + \frac{1}{1+U_d r_I \tau_x} J_{II} + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.176}$$

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

$$\begin{aligned}\det(\mathbf{M^{II}_{STD}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)[\tau_I^{-1}(1 + xJ_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_x^{-1} + U_d r_I) - \tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} U_d x] \\ + \tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (-\tau_x^{-1} - U_d r_I) < 0.\end{aligned} \tag{S0.177}$$

The condition shown in Eq. S0.176 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. S0.167 is identical to the condition of having paradoxical effect shown in Eq. S0.172, 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

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.178}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - u J_{II} r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.179}$$

$$\frac{du}{dt} = \frac{1-u}{\tau_u} + U_f (U_{max} - u) r_I, \tag{S0.180}$$

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 factor.

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

$$\mathbf{M^{II}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & 0 \\ \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1 + u J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & -\tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} \\ 0 & U_f(U_{max} - u) & -\tau_u^{-1} - U_f r_I \end{bmatrix}. \tag{S0.181}$$

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

$$\mathbf{N^{II}_{STF}} = \begin{bmatrix} \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & 0 \\ 0 & -\tau_u^{-1} - U_f r_I \end{bmatrix}. \tag{S0.182}$$

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

$$\mathrm{tr}(\mathbf{N^{II}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_u^{-1} - U_f r_I < 0 \tag{S0.183}$$

and

$$\det(\mathbf{N^{II}_{STF}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_u^{-1} - U_f r_I) > 0. \tag{S0.184}$$

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

$$\tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) > 0. \tag{S0.185}$$

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

$$\text{IS index} = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1). \tag{S0.186}$$

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 can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[J_{EE} r_E - J_{EI} r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.187}$$

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

$$r_I = \frac{J_{EE} r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{J_{EI}}. \tag{S0.188}$$

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

$$k_{STF}^E = \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.189}$$

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

$$J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1 > 0. \tag{S0.190}$$

The inhibitory nullcline can be written as follows

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE} r_E - \frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} r_I + g_I\Big]_+^{\alpha_I} = 0. \tag{S0.191}$$

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

$$r_I = \frac{J_{IE} r_E - r_I^{\frac{1}{\alpha_I}} + g_I}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II}}. \tag{S0.192}$$

The slope of the inhibitory nullcline can be written as follows

$$k_{STF}^I = \frac{J_{IE}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} + \frac{U_f U_{max} \tau_u (1+U_f r_I \tau_u) - (1+U_f U_{max} r_I \tau_u) U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{II} r_I + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}}. \tag{S0.193}$$

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

$$\frac{J_{IE}}{\frac{1+U_f U_{max} r_I \tau_u}{1+U_f r_I \tau_u} J_{II} + \frac{U_f U_{max} \tau_u (1+U_f r_I \tau_u) - (1+U_f U_{max} r_I \tau_u) U_f \tau_u}{(1+U_f r_I \tau_u)^2} J_{II} r_I + \frac{1}{\alpha_I} r_I^{\frac{1-\alpha_I}{\alpha_I}}} > \frac{1}{J_{EI}}\big(J_{EE} - \frac{1}{\alpha_E} r_E^{\frac{1}{\alpha_E}-1}\big). \tag{S0.194}$$

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

$$\begin{aligned}\det(\mathbf{M_{STF}^{II}}) = \tau_E^{-1}(J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)[\tau_I^{-1}(1 + uJ_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}})(\tau_u^{-1} + U_f r_I) + \tau_I^{-1} J_{II}\alpha_I r_I^{\frac{2\alpha_I-1}{\alpha_I}} U_f(U_{max} - u)] \\ -\tau_E^{-1} J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} \tau_I^{-1} J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} (\tau_u^{-1} + U_f r_I) < 0.\end{aligned} \tag{S0.195}$$

The condition shown in Eq. S0.194 is the same as the stability condition of the determinant of the Jacobian of the system with I-to-I STF, namely, as $\det(\mathbf{M_{STF}^{II}}) < 0$. Thus, the condition is always satisfied when the system with I-to-I STF is stable.

The condition of being IS shown in Eq. S0.185 is identical to the condition of having paradoxical effect shown in Eq. S0.190, 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

$$\begin{aligned}\det(\mathbf{N_{STD}^{IE}}) = \det(\mathbf{N_{STF}^{IE}}) = \det(\mathbf{N_{STD}^{EI}}) = \det(\mathbf{N_{STF}^{EI}}) = \det(\mathbf{N_{STD}^{II}}) = \det(\mathbf{N_{STF}^{II}}) \\ = \tau_E^{-1}(J_{EE}^{IS}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - U_d r_E) = 0,\end{aligned} \tag{S0.196}$$

$$J_{EE}^{IS} = \frac{1}{\alpha_E} r_E^{-\frac{\alpha_E-1}{\alpha_E}}. \tag{S0.197}$$

For $\alpha = 1$, we have

$$\frac{dJ_{EE}^{IS}}{dr_E} = 0. \tag{S0.198}$$

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

$$\frac{dJ_{EE}^{IS}}{dr_E} = -\frac{\alpha_E - 1}{\alpha_E^2} r_E^{-\frac{2\alpha_E-1}{\alpha_E}}. \tag{S0.199}$$

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.

For $\alpha < 1$, we have

$$\frac{dJ_{EE}^{IS}}{dr_E} = -\frac{\alpha_E - 1}{\alpha_E^2} r_E^{-\frac{2\alpha_E-1}{\alpha_E}} > 0. \tag{S0.200}$$

Therefore, for sublinear networks with E-to-I STD/STF, I-to-E STD/STF, and I-to-I STD/STF, the IS boundary always shifts upwards as r_E increases.

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 [29]. For the sake of completeness, we also included them in the Methods section.

Generalization to networks with a mixture of short-term plasticity mechanisms

To generalize our results to networks with a mixture of short-term plasticity mechanisms, we described the dynamics of the network as follows:

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[p_{EE}J_{EE}r_E - p_{EI}J_{EI}r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.201}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[p_{IE}J_{IE}r_E - p_{II}J_{II}r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.202}$$

where p_{XY} denotes the short-term plasticity variable for the synapse from Y to X, with $X, Y \in \{E, I\}$. A p_{XY} larger than 1 suggests that the synapse is facilitating, whereas a p_{XY} less than 1 indicates that the synapse is depressing.

Here, we derived the condition of having the paradoxical effect. 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. We can write the excitatory nullcline as follows

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[p_{EE}J_{EE}r_E - p_{EI}J_{EI}r_I + g_E\Big]_+^{\alpha_E} = 0. \tag{S0.203}$$

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

$$r_I = \frac{p_{EE}J_{EE}r_E - r_E^{\frac{1}{\alpha_E}} + g_E}{p_{EI}J_{EI}}. \tag{S0.204}$$

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

$$k^E = \frac{1}{p_{EI}J_{EI} + p_{EI}^{'}J_{EI}r_I}(p_{EE}J_{EE} + p_{EE}^{'}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1}). \tag{S0.205}$$

Provided that the inhibitory current to the excitatory network is always an increasing function (at least locally around the fixed point) of the inhibitory firing rate, we have

$$p_{EI}J_{EI} + p_{EI}^{'}J_{EI}r_I > 0. \tag{S0.206}$$

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. This condition is satisfied as long as the stability of the network is granted.

Next, we sought to identify the relationship between the inhibition stabilization property of the network and the paradoxical effect.

In the presence of E-to-E STD and STP at other types of synapses, the condition of the paradoxical effect is

$$p_{EE}J_{EE} + p_{EE}^{'}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} = x_{EE}J_{EE} + x_{EE}^{'}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} > 0. \tag{S0.207}$$

The above condition is the same as the condition that $\det(\mathbf{N}_{STD}^{EE}) < 0$. In networks with E-to-E STD as well as STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is impermissible. Thus, the IS conditions for networks with E-to-E STD and STP at other types of synapses are the same as these for networks with E-to-E STD alone. As we showed before, $\det(\mathbf{N}_{STD}^{EE}) < 0$ is the sufficient condition for IS in networks with E-to-E STD, but not the necessary condition. Therefore, in the presence of E-to-E STD and STP at other types of synapses, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not necessarily imply the paradoxical effect.

In the presence of E-to-E STF and STP at other types of synapses, the condition of the paradoxical effect changes from Eq. S0.207 to

$$p_{EE}J_{EE} + p^{'}_{EE}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} = u_{EE}J_{EE} + u^{'}_{EE}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} > 0. \tag{S0.208}$$

The above condition is the same as the condition that $\det(\mathbf{N}^{EE}_{STF}) < 0$. In networks with E-to-E STF and STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is impermissible. Thus, the IS conditions for networks with E-to-E STF and STP at other types of synapses are the same as these for networks with E-to-E STF alone. As we showed before, $\det(\mathbf{N}^{EE}_{STF}) < 0$ is the necessary and sufficient condition for IS in networks with E-to-E STF. Therefore, in the presence of E-to-E STF and STP at other types of synapses, inhibition stabilization and the paradoxical effect imply each other.

Similarly, in the absence of E-to-E STP but in the presence of STP at other types of types of synapses, the condition of the paradoxical effect instead becomes

$$p_{EE}J_{EE} + p^{'}_{EE}J_{EE}r_E - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} = J_{EE} - \frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1} > 0 \tag{S0.209}$$

In networks without E-to-E STD but with STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is impermissible. The IS conditions for networks without E-to-E STP but with STP at other types of synapses are the same as these for networks with static connectivity. The above condition is the necessary and sufficient condition for IS in networks with static connectivity. Therefore, in the absence of E-to-E STP but in the presence of STP at other types of synapses, inhibition stabilization and the paradoxical effect imply each other.

Generalization to networks with both E-to-E STD and E-to-E STF

To generalize our results to networks with both STD and STF on the E-to-E synapses, we described the dynamics of the network as follows:

$$\tau_E \frac{dr_E}{dt} = -r_E + \Big[uxJ_{EE}r_E - J_{EI}r_I + g_E\Big]_+^{\alpha_E}, \tag{S0.210}$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Big[J_{IE}r_E - J_{II}r_I + g_I\Big]_+^{\alpha_I}, \tag{S0.211}$$

$$\frac{dx}{dt} = \frac{1-x}{\tau_x} - uxr_E, \tag{S0.212}$$

$$\frac{du}{dt} = -\frac{u}{\tau_u} + U(1-u)r_E. \tag{S0.213}$$

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

$$\mathbf{M^{EE}_{STP}} = \begin{bmatrix} \tau_E^{-1}(uxJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & -\tau_E^{-1}J_{EI}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} & \tau_E^{-1}uJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} & \tau_E^{-1}xJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ \tau_I^{-1}J_{IE}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}} & -\tau_I^{-1}(1+J_{II}\alpha_I r_I^{\frac{\alpha_I-1}{\alpha_I}}) & 0 & 0 \\ -ux & 0 & -\tau_x^{-1} - ur_E & -xr_E \\ U(1-u) & 0 & 0 & -\tau_u^{-1} - Ur_E \end{bmatrix} \tag{S0.214}$$

If inhibition is frozen, i.e., if feedback inhibition is absent, the Jacobian of the system becomes as follows:

$$\mathbf{N^{EE}_{STP}} = \begin{bmatrix} \tau_E^{-1}(uxJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) & \tau_E^{-1}uJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} & \tau_E^{-1}xJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}} \\ -ux & -\tau_x^{-1} - ur_E & -xr_E \\ U(1-u) & 0 & -\tau_u^{-1} - Ur_E \end{bmatrix}. \tag{S0.215}$$

As the product of three eigenvalues of $\mathbf{N^{EE}_{STP}}$ is the determinant $det(\mathbf{N^{EE}_{STP}})$, the network is inhibition stabilized if $det(\mathbf{N^{EE}_{STP}}) > 0$ so we can write:

$$\begin{aligned} det(\mathbf{N^{EE}_{STP}}) &= \tau_E^{-1}(uxJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1)(-\tau_x^{-1} - ur_E)(-\tau_u^{-1} - Ur_E) \\ &-\tau_E^{-1}uJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}[(-ux)(-\tau_u^{-1} - Ur_E) - (-xr_E)U(1-u)] - \tau_E^{-1}xJ_{EE}\alpha_E r_E^{\frac{2\alpha_E-1}{\alpha_E}}(-\tau_x^{-1} - ur_E)U(1-u) > 0. \end{aligned} \tag{S0.216}$$

Therefore, we have

$$\frac{\tau_u Ur_E(2+\tau_u Ur_E)}{(1+\tau_x\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E)^2(1+\tau_u Ur_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} > 1 \tag{S0.217}$$

As the sum of three eigenvalues of $\mathbf{N^{EE}_{STP}}$ is the trace $tr(\mathbf{N^{EE}_{STP}})$, the network is inhibition stabilized if $tr(\mathbf{N^{EE}_{STP}}) > 0$ and we can write:

$$tr(\mathbf{N^{EE}_{STP}}) = \tau_E^{-1}(uxJ_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} - 1) - \tau_x^{-1} - ur_E - \tau_u^{-1} - Ur_E > 0. \tag{S0.218}$$

Therefore, we have

$$\begin{aligned} \frac{\tau_u Ur_E(2+\tau_u Ur_E)}{(1+\tau_u Ur_E)^2(1+\tau_x\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}} &> 1 + \frac{-\tau_x\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E + \tau_E\tau_x^{-1} + \tau_E\tau_u^{-1} + \tau_E\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E + \tau_E Ur_E}{1+\tau_x\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E} \\ &+ \frac{\tau_u Ur_E}{(1+\tau_u Ur_E)^2(1+\tau_x\frac{\tau_u Ur_E}{1+\tau_u Ur_E}r_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}. \end{aligned} \tag{S0.219}$$

Note that the trace condition is independent from the determinant condition. For instance, in the limit of small τ_x and large τ_u when E-to-E STD dominates over E-to-E STF, we have:

$$1+\frac{-\tau_x\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E+\tau_E\tau_x^{-1}+\tau_E\tau_u^{-1}+\tau_E\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E+\tau_E U r_E}{1+\tau_x\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E}+\frac{\tau_u U r_E}{(1+\tau_u U r_E)^2(1+\tau_x\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}$$
$$\approx 1+\tau_E\tau_x^{-1}+\tau_E r_E+\tau_E U r_E \tag{S0.220}$$

The trace condition then becomes as follows

$$\frac{\tau_u U r_E(2+\tau_u U r_E)}{(1+\tau_u U r_E)^2(1+\tau_x\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}>1+\tau_E\tau_x^{-1}+\tau_E r_E+\tau_E U r_E>1. \tag{S0.221}$$

Therefore, when E-to-E STD dominates over E-to-E STF, the network can be inhibition stabilized by satisfying the determinant condition, but it does not have to satisfy the trace condition.

Furthermore, it is worth mentioning that as $\mathbf{N}^{\mathbf{EE}}_{\mathbf{STP}}$ is a 3 by 3 matrix, the determinant and trace conditions are sufficient conditions to identify IS but not the necessary conditions.

Next, we derived the condition of having the paradoxical effect. 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. We can write the excitatory nullcline as follows

$$\tau_E\frac{dr_E}{dt}=-r_E+\Big[uxJ_{EE}r_E-J_{EI}r_I+g_E\Big]_+^{\alpha_E}=0 \tag{S0.222}$$

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

$$r_I=\frac{uxJ_{EE}r_E-r_E^{\frac{1}{\alpha_E}}+g_E}{J_{EI}} \tag{S0.223}$$

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

$$k^E=\frac{1}{J_{EI}}(uxJ_{EE}+u^{'}xJ_{EE}r_E+ux^{'}J_{EE}r_E-\frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1}) \tag{S0.224}$$

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

$$k^E=\frac{1}{J_{EI}}(uxJ_{EE}+u^{'}xJ_{EE}r_E+ux^{'}J_{EE}r_E-\frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1})>0 \tag{S0.225}$$

$$uxJ_{EE}+u^{'}xJ_{EE}r_E+ux^{'}J_{EE}r_E-\frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1}$$
$$=uxJ_{EE}+\frac{\tau_u U}{(1+\tau_u U r_E)^2}xJ_{EE}r_E-\frac{\tau_x\tau_u r_E U(2+\tau_u r_E U)}{(1+\tau_u r_E(1+\tau_x r_E)U)^2}uJ_{EE}r_E-\frac{1}{\alpha_E}r_E^{\frac{1}{\alpha_E}-1}>0 \tag{S0.226}$$

We can then obtain the condition below:

$$\frac{\tau_u U r_E(2+\tau_u U r_E)}{(1+\tau_u U r_E)^2(1+\tau_x\frac{\tau_u U r_E}{1+\tau_u U r_E}r_E)^2}J_{EE}\alpha_E r_E^{\frac{\alpha_E-1}{\alpha_E}}>1 \tag{S0.227}$$

This condition is identical to the determinant condition for being IS.

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. This condition is satisfied as long as the stability of the network is granted.

Therefore, in the presence of E-to-E STD and STF, the paradoxical effect implies inhibition stabilization, whereas inhibition stabilization does not necessarily imply the paradoxical effect.

Generalization to networks with multiple inhibitory cell types

To generalize our results to networks with multiple inhibitory cell types, we considered a network with one excitatory population and $N-1$ inhibitory populations. The dynamics of the network can be described as follows:

$$\mathbf{T}\frac{d}{dt}\mathbf{r} = -\mathbf{r} + \mathbf{f}(\mathbf{J_{eff}}\mathbf{r} + \mathbf{g}), \tag{S0.228}$$

where $\mathbf{T}$ is a diagonal matrix containing the time constants of firing rate dynamics for different populations, $\mathbf{r}$ is a vector containing the firing rates of different populations, $\mathbf{f}(\mathbf{x})$ is a vector containing the input-output functions for different populations, $\mathbf{g}$ is a vector containing the inputs to different populations.

$$\mathbf{T} = \begin{bmatrix} \tau_E & 0 & \cdots & 0 \\ 0 & \tau_{I_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tau_{I_{N-1}} \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} r_E \\ r_{I_1} \\ \vdots \\ r_{I_{N-1}} \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_E(x_1) \\ f_{I_1}(x_2) \\ \vdots \\ f_{I_{N-1}}(x_N) \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_E \\ g_{I_1} \\ \vdots \\ g_{I_{N-1}} \end{bmatrix}. \tag{S0.229}$$

$\mathbf{J_{eff}}$ is the effective connectivity matrix defined as follows:

$$\mathbf{J_{eff}} = \begin{bmatrix} p_{EE}J_{EE} & -p_{EI_1}J_{EI_1} & \cdots & -p_{EI_{N-1}}J_{EI_{N-1}} \\ p_{I_1E}J_{I_1E} & -p_{I_1I_1}J_{I_1I_1} & \cdots & -p_{I_1I_{N-1}}J_{I_1I_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ p_{I_{N-1}E}J_{I_{N-1}E} & -p_{I_{N-1}I_1}J_{I_{N-1}I_1} & \cdots & -p_{I_{N-1}I_{N-1}}J_{I_{N-1}I_{N-1}} \end{bmatrix}, \tag{S0.230}$$

where p_{XY} denotes the short-term plasticity variable for the synapse from Y to X, with $X, Y \in \{E, I_1, \cdots, I_{N-1}\}$. A p_{XY} larger than 1 suggests that the synapse is facilitating, whereas a p_{XY} less than 1 indicates that the synapse is depressing.

The equation linearized about the fixed point can be written as follows:

$$\mathbf{T}\frac{d}{dt}\delta\mathbf{r} = -\delta\mathbf{r} + \mathbf{F}\mathbf{J_{eff}}\delta\mathbf{r} + \mathbf{F}(\mathbf{J}'_{\mathbf{eff}} \bullet \mathbf{r})\delta\mathbf{r} + \mathbf{F}\delta\mathbf{g} \tag{S0.231}$$

where $\delta\mathbf{r}$ is a vector containing the changes in firing rates of different populations, $\delta\mathbf{g}$ is a vector containing the changes in the inputs to different populations, $\mathbf{F}$ is a diagonal matrix containing the derivatives of the input-output functions at the fixed point:

$$\mathbf{F} = \begin{bmatrix} f'_E & 0 & \cdots & 0 \\ 0 & f'_{I_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f'_{I_{N-1}} \end{bmatrix}, \tag{S0.232}$$

$\mathbf{J}'_{\mathbf{eff}}$ is a tensor and given by

$$\mathbf{J}'_{\mathbf{eff}}[:,:,1] = \begin{bmatrix} p'_{EE}J_{EE} & 0 & \cdots & 0 \\ p'_{I_1E}J_{I_1E} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p'_{I_{N-1}E}J_{I_{N-1}E} & 0 & \cdots & 0 \end{bmatrix}, \tag{S0.233}$$

$$\mathbf{J}'_{\mathbf{eff}}[:,:,2] = \begin{bmatrix} 0 & -p'_{EI_1}J_{EI_1} & \cdots & 0 \\ 0 & -p'_{I_1I_1}J_{I_1I_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -p'_{I_{N-1}I_1}J_{I_{N-1}I_1} & \cdots & 0 \end{bmatrix}, \tag{S0.234}$$

$$\vdots \tag{S0.235}$$

$$\mathbf{J}_{\mathbf{eff}}^{'}[:,:,N] = \begin{bmatrix} 0 & 0 & \cdots & -p_{EI_{N-1}}^{'} J_{EI_{N-1}} \\ 0 & 0 & \cdots & -p_{I_1 I_{N-1}}^{'} J_{I_1 I_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -p_{I_{N-1} I_{N-1}}^{'} J_{I_{N-1} I_{N-1}} \end{bmatrix}. \tag{S0.236}$$

$p_{XY}^{'}$ denotes the derivative of the short-term plasticity variable p_{XY} with respect to r_Y. $\mathbf{J}_{\mathbf{eff}}^{'} \bullet \mathbf{r}$ is a 3-mode product of the tensor $\mathbf{J}_{\mathbf{eff}}^{'}$ and the vector $\mathbf{r}$:

$$\mathbf{J}_{\mathbf{eff}}^{'} \bullet \mathbf{r} = \begin{bmatrix} p_{EE}^{'} J_{EE} r_E & -p_{EI_1}^{'} J_{EI_1} r_{I_1} & \cdots & -p_{EI_{N-1}}^{'} J_{EI_{N-1}} r_{I_{N-1}} \\ p_{I_1 E}^{'} J_{I_1 E} r_E & -p_{I_1 I_1}^{'} J_{I_1 I_1} r_{I_1} & \cdots & -p_{I_1 I_{N-1}}^{'} J_{I_1 I_{N-1}} r_{I_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ p_{I_{N-1} E}^{'} J_{I_{N-1} E} r_E & -p_{I_{N-1} I_1}^{'} J_{I_{N-1} I_1} r_{I_1} & \cdots & -p_{I_{N-1} I_{N-1}}^{'} J_{I_{N-1} I_{N-1}} r_{I_{N-1}} \end{bmatrix}. \tag{S0.237}$$

If only inhibitory populations receive external perturbative inputs, i.e., δg_E is 0, we have:

$$(1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E)\delta r_E = f_E^{'} \sum_{i=1}^{N-1} (-p_{EI_i} J_{EI_i} - p_{EI_i}^{'} J_{EI_i} r_{I_i})\delta r_{I_i}. \tag{S0.238}$$

Note that $f_E^{'} \sum_{i=1}^{N-1} (-p_{EI_i} J_{EI_i} - p_{EI_i}^{'} J_{EI_i} r_{I_i})\delta r_{I_i}$ represents the change in the total inhibitory current to the excitatory population. If it is negative, the total inhibitory current to the excitatory population increases. Provided that $f_E^{'}$ is positive, namely, the input-output function is a monotonically increasing function, we have the following two cases:
1) If $1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E$ is negative, then the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are the same. This implies that if $1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E$ is negative, additional inputs to the inhibitory populations increase or decrease the firing rate of the excitatory population as well as the total inhibitory current to the excitatory population.
2) If $1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E$ is positive, then the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are opposite.

Next, we sought to identify the relationship between the sign of $1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E$ and the inhibition stabilization property of the network.

In the presence of E-to-E STD and STP at other types of synapses, we have

$$\begin{aligned} (1 - f_E^{'} p_{EE} J_{EE} - f_E^{'} p_{EE}^{'} J_{EE} r_E)\delta r_E &= (1 - f_E^{'} x_{EE} J_{EE} - f_E^{'} x_{EE}^{'} J_{EE} r_E)\delta r_E \\ &= f_E^{'} \sum_{i=1}^{N-1} (-p_{EI_i} J_{EI_i} - p_{EI_i}^{'} J_{EI_i} r_{I_i})\delta r_{I_i}. \end{aligned} \tag{S0.239}$$

In this case, the condition that $1 - f_E^{'} x_{EE} J_{EE} - f_E^{'} x_{EE}^{'} J_{EE} r_E < 0$ is the same as the condition that $\det(\mathbf{N}_{\mathbf{STD}}^{\mathbf{EE}}) < 0$. In networks with E-to-E STD and STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is not allowed (frozen inhibition). Thus, the IS conditions for networks with E-to-E STD and STP at other types of synapses are the same as these for networks with E-to-E STD alone. As we showed before, $\det(\mathbf{N}_{\mathbf{STD}}^{\mathbf{EE}}) < 0$ is a sufficient condition for IS in networks with E-to-E STD, but not a necessary condition. Therefore, in the presence of E-to-E STD, if the network is IS, the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are the same. However, the same sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population does not imply that the network is IS.

In the presence of E-to-E STF and STP at other types of synapses, we have:

$$(1-f_E^{'}p_{EE}J_{EE}-f_E^{'}p_{EE}^{'}J_{EE}r_E)\delta r_E = (1-f_E^{'}u_{EE}J_{EE}-f_E^{'}u_{EE}^{'}J_{EE}r_E)\delta r_E = f_E^{'}\sum_{i=1}^{N-1}(-p_{EI_i}J_{EI_i}-p_{EI_i}^{'}J_{EI_i}r_{I_i})\delta r_{I_i}. \tag{S0.240}$$

In this case, the condition that $1 - f_E^{'}u_{EE}J_{EE} - f_E^{'}u_{EE}^{'}J_{EE}r_E < 0$ is the same as the condition that $\det(\mathbf{N}_{\mathbf{STF}}^{\mathbf{EE}}) < 0$. In networks with E-to-E STF and STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is not allowed. Thus, the IS conditions for networks with E-to-E STF and STP at other types of synapses are the same as those for networks with E-to-E STF alone. As we showed before, $\det(\mathbf{N}_{\mathbf{STF}}^{\mathbf{EE}}) < 0$ is a necessary and a sufficient condition for IS in networks with E-to-E STF. Therefore, in the presence of E-to-E STF, if the network is non-IS, then the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are opposite. And if the network is IS, the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are the same.

In the absence of E-to-E STP but in the presence of STP at other types of synapses, we have:

$$(1 - f_E^{'}p_{EE}J_{EE} - f_E^{'}p_{EE}^{'}J_{EE}r_E)\delta r_E = (1 - f_E^{'}J_{EE})\delta r_E = f_E^{'}\sum_{i=1}^{N-1}(-p_{EI_i}J_{EI_i} - p_{EI_i}^{'}J_{EI_i}r_{I_i})\delta r_{I_i}. \tag{S0.241}$$

In networks without E-to-E STD but with STP at other types of synapses, the effect of STP at other types of synapses does not act on the excitatory subnetwork when feedback inhibition is impermissible. The IS conditions for networks without E-to-E STP but with STP at other types of synapses are the same as these for networks with static connectivity. The above condition is a necessary and a sufficient condition for IS in networks with static connectivity. Therefore, in the absence of E-to-E STP but in the presence of STP at other types of synapses, if the network is non-IS, then the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are opposite. And if the network is IS, the sign of the change in the firing rate of the excitatory population and of the change in the total inhibitory current to the excitatory population are the same.