Hierarchical Selective Recruitment in Linear-Threshold Brain Networks Part I: Single-Layer Dynamics and Selective Inhibition

Erfan Nozari Jorge Cortés

***Abstract*—Goal-driven selective attention (GDSA) refers to the brain's function of prioritizing the activity of a task-relevant subset of its overall network to efficiently process relevant information while inhibiting the effects of distractions. Despite decades of research in neuroscience, a comprehensive understanding of GDSA is still lacking. We propose a novel framework using concepts and tools from control theory as well as insights and structures from neuroscience. Central to this framework is an information-processing hierarchy with two main components: selective inhibition of task-irrelevant activity and top-down recruitment of task-relevant activity. We analyze the internal dynamics of each layer of the hierarchy described as a network with linear-threshold dynamics and derive conditions on its structure to guarantee existence and uniqueness of equilibria, asymptotic stability, and boundedness of trajectories. We also provide mechanisms that enforce selective inhibition using the biologically-inspired schemes of feedforward and feedback inhibition. Despite their differences, both lead to the same conclusion: the intrinsic dynamical properties of the (not-inhibited) task-relevant subnetworks are the sole determiner of the dynamical properties that are achievable under selective inhibition.**

I. Introduction

The human brain is constantly under the influx of sensory inputs and is responsible for integrating and interpreting them to generate appropriate decisions and actions. This influx contains not only the pieces of information relevant to the present task(s), but also a myriad of distractions. Goal-driven selective attention (GDSA) refers to the active selective processing of a subset of information influx while suppressing the effects of others, and is vital for the proper function of the brain.[1] Examples range from selective audition in a crowded place to selective vision in cluttered environments to selective taste/smell in food. As a result, a long standing question in neuroscience involves understanding the brain's complex mechanisms underlying selective attention [2]–[7].

A central element in addressing this question is the role played by the hierarchical organization of the brain [8]. Broadly, this organization places primary sensory and motor areas at the bottom and integrative association areas (prefrontal cortex in particular) at the top. Accordingly, sensory information is processed while flowing up the hierarchy, where decisions are eventually made and transmitted back down the hierarchy to generate motor actions.[2] The top-down direction is also responsible for GDSA, where the higher-order areas differentially "modulate" the activity of the lower-level areas such that only relevant information is further processed. This phenomenon constitutes the basis for GDSA and has been the subject of extensive experimental research in neuroscience, see e.g., [4], [9]–[18]. However, a complete understanding of how, when (how quick), or where (within the hierarchy) it occurs is still lacking. In particular, the relationship between GDSA and the dynamics of the involved neuronal networks is poorly understood. Our goal is to address this gap from a model-based perspective, resorting to control-theoretic tools to explain various aspects of GDSA in terms of the synaptic network structure and the dynamics that emerge from it.

In this work, we propose a theoretical framework, termed *hierarchical selective recruitment (HSR)*, to explain the network dynamics underlying GDSA. This framework consists of a novel hierarchical model of brain organization (though composed of well-established sub-models at each layer), a set of analytical results regarding the multi-timescale dynamics of this model, and a careful translation between the properties of these dynamics and well known experimental observations about GDSA. The starting point in the development of HSR is the observation that different stimuli, in particular the task-relevant and task-irrelevant ones, are processed by different populations of neurons (see, e.g., [4], [5], [7], [11]–[14], [18]). With each neuronal population represented by a node in the overall neuronal network of networks and based on extensive experimental research (see below), HSR primarily relies on the *selective inhibition* of the task-irrelevant nodes and the *top-down recruitment* of the task-relevant nodes of each layer by the layer immediately above. This paper analyzes the dynamics of individual layers as well as the mechanisms for selective inhibition in a bilayer network. These results set the basis for the study of the mechanisms for top-down recruitment in multilayer networks in our accompanying work [19].

Literature Review: In this work we use dynamical networks with linear-threshold nonlinearities (the unbounded version also called rectified linear units, ReLU, in machine learning) to model the activity of neuronal populations. Linear-threshold models allow for a unique combination between the tractability of linear systems and the dynamical versatility of nonlinear systems, and thus have been widely used in computational neuroscience. They were first proposed as a model for the lateral eye of the horseshoe crab in [20] and their dynamical behavior has been studied at least as early as [21]. A detailed stability analysis of symmetric (undirected) linear-threshold networks has been carried out in continuous [22] and discrete [23] time: however, this has limited relevance for biological neuronal networks, which are fundamentally asymmetric (due to the presence of excitatory and inhibitory neurons). Regarding asymmetric networks, it was claimed

A preliminary version appeared as [1] at the American Control Conference.

Erfan Nozari is with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, enozari@seas.upenn.edu. Jorge Cortés is with the Department of Mechanical and Aerospace Engineering, UC San Diego, La Jolla, CA 92093, cortes@ucsd.edu.

[1]Note the distinction with stimulus-driven selective attention (the reactive shift of focus based on saliency of stimuli) which is not the focus here.

[2]Note that the role of memory (being distributed across the brain) is implicit in this simplified stimulus-response description. Indeed, many sensory inputs only form memories (without motor response) or many motor actions result chiefly from memory (without sensory stimulation). The hierarchical aspect is nevertheless present.

(without proper proof) in [24] that the identity minus the matrix of synaptic connectivities being a P-matrix is necessary and sufficient for the existence and uniqueness of equilibria (EUE), the negative of this matrix being totally Hurwitz is necessary and sufficient for local asymptotic stability, and the matrix of synaptic connectivities being absolutely Schur stable is sufficient for global asymptotic stability. In addition to lacking proper proof, these results were limited to fully-inhibitory networks. The latter assertion was later proved rigorously in [25] for arbitrary networks, while we prove the first two (except on certain sets of measure zero) here. Around the same time, [26] considered the more general class of monotonically non-decreasing activation functions and proved the sufficiency of identity minus the matrix of synaptic connectivities being a P-matrix for the uniqueness of equilibria (being only one of the four implications we prove here) and the sufficiency of the same matrix being Lyapunov diagonally stable for global asymptotic stability (which we relax here by allowing for arbitrary quadratic Lyapunov functions). This work was later generalized to discontinuous neural networks (though not applicable to our model here) in [27]. Also related is the work [28] showing the necessity and sufficiency of identity minus the matrix of synaptic connectivities being a P_0-matrix for EUE of similar systems but with *strictly* monotonically increasing activation functions. The work [29] provides a comprehensive review of stability analysis of a range of continuous-time recurrent neural networks, including the linear-threshold model.

Lyapunov-based methods have also been used in a number of later studies for discrete-time linear-threshold networks [30]–[32], but the extension of these results to continuous-time dynamics is unclear. In fact, the use of Lyapunov-based techniques in continuous-time networks has remained limited to planar dynamics [33] and restrictive conditions for boundedness of trajectories [33], [34]. Recently, [35] presents interesting properties of competitive (i.e., fully inhibitory) linear-threshold networks, particularly regarding the emergence of oscillations. However, the majority of neurons in biological neuronal networks are excitatory, making the implications of these results limited. Moreover, all the preceding works are limited to networks with constant exogenous inputs whereas time-varying inputs are essential for modeling inter-layer connections in HSR.

A critical property of linear-threshold networks is that their nonlinearity, while enriching their behavior beyond that of linear systems, is piecewise linear. Accordingly, almost all the theoretical analysis of these networks builds upon the formulation of them as switched affine systems. There exists a vast literature on the analysis of general switched linear/affine systems, see, e.g., [36]–[38]. Nevertheless, we have found that the conditions obtained by applying these results to linear-threshold dynamics are more conservative than the ones we obtain using direct analysis of the system dynamics. This is mainly due to the fact that such results, by the essence of their generality, are oblivious to the particular structure of linear-threshold dynamics that can be leveraged in direct analysis.

Selective inhibition has been the subject of extensive research in neuroscience. A number of early studies [4], [11], [12] provided evidence for a mechanism of selective visual attention based on a biased competition between the subnetwork of task-relevant nodes and the subnetwork of task-irrelevant ones. In this model, nodes belonging to these subnetworks compete at each layer by mutually suppressing the activity of each other, and this competition is biased towards task-relevant nodes by the layer immediately above. Later studies [13], [14] further supported this theory using functional magnetic resonance imaging (fMRI) and showed [39], in particular, the suppression of activity of task-irrelevant nodes as a result of GDSA. This suppression of activity is further shown to occur in multiple layers along the hierarchy [40], grow with increasing attention [41], [42], and be inversely related to the power of the task-irrelevant nodes' state trajectories in the alpha frequency band ($\sim$ 8-14^{Hz}) [16].

Statement of Contributions: The contributions are twofold. First, we analyze the internal dynamics of a single-layer linear-threshold network as a basis for our study of hierarchical structures. Our results here provide a comprehensive characterization of the dynamical properties of linear-threshold networks. Specifically, we show that existence and uniqueness of equilibria, asymptotic stability, and boundedness of trajectories can be characterized using simple algebraic conditions on the network structure in terms of the class of P-matrices (matrices with positive principal minors), totally-Hurwitz matrices (those with Hurwitz principal submatrices, shown to be a sub-class of P-matrices), and Schur-stable matrices, respectively. In addition to forming the basis of HSR, these results solve some long-standing open problems in the characterization of linear-threshold networks [21], [24], [25], [33]–[35] and are of independent interest. Our analysis covers both the class of unbounded (a.k.a. ReLU) as well as bounded linear-threshold networks, where the latter is a piecewise-affine approximation of sigmoidal neural networks, for which limited analytical results are available. Our second contribution pertains the problem of selective inhibition in a bilayer network. Motivated by the mechanisms of inhibition in the brain, we study feedforward and feedback mechanisms. We provide necessary and sufficient conditions on the network structure that guarantee selective inhibition of task-irrelevant nodes at the lower-level while simultaneously guaranteeing various dynamical properties of the resulting (partly inhibited, partly active) subnetwork, including existence and uniqueness of equilibria and asymptotic stability. Interestingly, under both mechanisms, these conditions require that the (not-inhibited) task-relevant part of the lower-level subnetwork intrinsically satisfies the same desired dynamical properties. This is particularly important for selective inhibition as asymptotic stability underlies it. The results unveil the important role of task-relevant nodes in constraining the dynamical properties achievable under selective inhibition and have implications for the number and centrality of nodes that need to be inhibited for an unstable-in-isolation subnetwork to gain stability through selective inhibition. For subnetworks that are not stable as a whole, these results provide conditions on the task-relevant/irrelevant partitioning of the nodes that allow for stabilization using inhibitory control.

II. PRELIMINARIES

We introduce notational conventions and basic concepts on matrix analysis and modeling of biological neuronal networks.

Notation

Throughout the paper, we employ the following notation. We use $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{\leq 0}$ to denote the set of reals, nonnegative reals, and nonpositive reals, respectively. We use boldfaced letters for vectors and matrices. $\mathbf{1}_n$, $\mathbf{0}_n$, $\boldsymbol{\ell}_n$, $\mathbf{0}_{p\times n}$, and $\mathbf{I}_n$ stand for the n-vector of all ones, the n-vector of all zeros, the n-vector of all ℓ's, the p-by-n zero matrix, and the identity n-by-n matrix (we omit the subscripts when clear from the context). Given a vector $\mathbf{x} \in \mathbb{R}^n$, x_i and $(\mathbf{x})_i$ refer to its ith component. Given $\mathbf{A} \in \mathbb{R}^{p\times n}$, a_{ij} refers to the (i,j)th entry. For block-partitioned $\mathbf{x}$ and $\mathbf{A}$, $\mathbf{x}_i$ and $\mathbf{A}_{ij}$ refer to the ith block of $\mathbf{x}$ and (i,j)th block of $\mathbf{A}$, respectively. In block representation of matrices, $\star$ denotes arbitrary blocks whose value is immaterial to the discussion. For $\mathbf{A} \in \mathbb{R}^{p\times n}$, range$(\mathbf{A})$ denotes the subspace of $\mathbb{R}^p$ spanned by the columns of $\mathbf{A}$. If $\mathbf{x}$ and $\mathbf{y}$ are vectors, $\mathbf{x} \leq \mathbf{y}$ denotes $x_i \leq y_i$ for all i. For symmetric $\mathbf{P} \in \mathbb{R}^{n\times n}$, $\mathbf{P} > \mathbf{0}$ ($\mathbf{P} < \mathbf{0}$) denotes that $\mathbf{P}$ is positive (negative) definite. Given $\mathbf{A} \in \mathbb{R}^{n\times n}$, its element-wise absolute value, determinant, spectral radius, and induced 2-norm are denoted by $|\mathbf{A}|$, $\det(\mathbf{A})$, $\rho(\mathbf{A})$, and $\|\mathbf{A}\|$, respectively. Similarly, for $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|$ is its 2-norm. Likewise, for two matrices $\mathbf{A}$ and $\mathbf{B}$, diag$(\mathbf{A},\mathbf{B})$ denotes the block-diagonal matrix with $\mathbf{A}$ and $\mathbf{B}$ on its diagonal. Given a subspace W of $\mathbb{R}^n$, $W^\perp$ denotes the orthogonal complement of W in $\mathbb{R}^n$. For $x \in \mathbb{R}$ and $m \in \mathbb{R}_{>0} \cup \{\infty\}$, $[x]_0^m = \min\{\max\{x,0\}, m\}$, which is the projection of x onto $[0,m]$. When $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{m} \in \mathbb{R}^n_{>0} \cup \{\infty\}^n$, we similarly define $[\mathbf{x}]_{\mathbf{0}}^{\mathbf{m}} = [[x_1]_0^{m_1} \quad \cdots \quad [x_n]_0^{m_n}]^T$. All measure-theoretic statements are meant in the Lebesgue sense.

Matrix Analysis

We here define and characterize several matrix classes of interest that play a key role in the forthcoming discussion.

Definition II.1. ***(Matrix classes).*** *A matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is*

(i) absolutely Schur stable *if $\rho(|\mathbf{A}|) < 1$;*
(ii) totally $\mathcal{L}$-stable*, denoted $\mathbf{A} \in \mathcal{L}$, if there exists $\mathbf{P} = \mathbf{P}^T > \mathbf{0}$ such that $(-\mathbf{I} + \mathbf{A}^T\boldsymbol{\Sigma})\mathbf{P} + \mathbf{P}(-\mathbf{I} + \boldsymbol{\Sigma}\mathbf{A}) < \mathbf{0}$ for $\boldsymbol{\Sigma} = diag(\boldsymbol{\sigma})$ and all $\boldsymbol{\sigma} \in \{0,1\}^n$;*
(iii) totally Hurwitz*, denoted $\mathbf{A} \in \mathcal{H}$, if all the principal submatrices of $\mathbf{A}$ are Hurwitz;*
(iv) a P-matrix*, denoted $\mathbf{A} \in \mathcal{P}$, if all the principal minors of $\mathbf{A}$ are positive.*

In working with P-matrices, the principal pivot transform of a matrix plays an important role. Given

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

with nonsingular $\mathbf{A}_{22}$, its principal pivot transform is the matrix

$$\pi(\mathbf{A}) \triangleq \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{22}^{-1} \end{bmatrix}.$$

Note that $\pi(\pi(\mathbf{A})) = \mathbf{A}$. The next result formalizes several equivalent characterizations of P-matrices.

Lemma II.2. ***(Properties of P-matrices [43], [44]).*** *$\mathbf{A} \in \mathbb{R}^{n\times n}$ is a P-matrix if and only if any of the following holds:*

(i) *$\mathbf{A}^{-1}$ is a P-matrix;*
(ii) *all real eigenvalues of all the principal submatrices of $\mathbf{A}$ are positive;*
(iii) *for any $\mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ there is k such that $x_k(\mathbf{A}\mathbf{x})_k > 0$;*
(iv) *the principal pivot transform of $\mathbf{A}$ is a P-matrix.*

The matrix classes in Definition II.1 have important inclusion relationships, as shown next.

Lemma II.3. ***(Inclusions among matrix classes).*** *For $\mathbf{A}, \mathbf{W} \in \mathbb{R}^{n\times n}$, we have*

(i) $\rho(|\mathbf{W}|) < 1 \Rightarrow -\mathbf{I} + \mathbf{W} \in \mathcal{H}$;
(ii) $\|\mathbf{W}\| < 1 \Rightarrow \mathbf{W} \in \mathcal{L}$;
(iii) $\mathbf{W} \in \mathcal{L} \Rightarrow -\mathbf{I} + \mathbf{W} \in \mathcal{H}$;
(iv) $\mathbf{A} \in \mathcal{H} \Rightarrow -\mathbf{A} \in \mathcal{P}$.

Proof: (i). From [45, Fact 4.11.19], we have that $\rho(|\mathbf{W}_{\boldsymbol{\sigma}}|) < 1$ for any principal submatrix $\mathbf{W}_{\boldsymbol{\sigma}}$ of $\mathbf{W}$, which implies $\rho(\mathbf{W}_{\boldsymbol{\sigma}}) < 1$ by [45, Fact 4.11.17], implying the result.

(ii) It is straightforward to check that $\mathbf{P} = \mathbf{I}_n$ satisfies $(-\mathbf{I} + \mathbf{W}^T\boldsymbol{\Sigma})\mathbf{P} + \mathbf{P}(-\mathbf{I} + \boldsymbol{\Sigma}\mathbf{W}) < \mathbf{0}$ for all $\boldsymbol{\sigma} \in \{0,1\}^n$.

(iii) Pick an arbitrary $\boldsymbol{\sigma} \in \{0,1\}^n$ and let the permutation $\boldsymbol{\Pi} \in \mathbb{R}^{n\times n}$ be such that $\boldsymbol{\Pi}\boldsymbol{\Sigma}\mathbf{W}\boldsymbol{\Pi}^T = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{W}}_{21} & \hat{\mathbf{W}}_{22} \end{bmatrix}$, where $\hat{\mathbf{W}}_{22}$ is the principal submatrix of $\mathbf{W}$ corresponding to $\boldsymbol{\sigma}$. Then

$$\begin{aligned} \mathbf{P}(-\mathbf{I} + \boldsymbol{\Sigma}\mathbf{W}) &= \mathbf{P}\boldsymbol{\Pi}^T \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \hat{\mathbf{W}}_{21} & -\mathbf{I} + \hat{\mathbf{W}}_{22} \end{bmatrix} \boldsymbol{\Pi} \\ &= \boldsymbol{\Pi}^T \Big(\underbrace{\boldsymbol{\Pi}\mathbf{P}\boldsymbol{\Pi}^T}_{\hat{\mathbf{P}}} \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \hat{\mathbf{W}}_{21} & -\mathbf{I} + \hat{\mathbf{W}}_{22} \end{bmatrix} \Big) \boldsymbol{\Pi} \\ &= \boldsymbol{\Pi}^T \begin{bmatrix} \star & \star \\ \star & \hat{\mathbf{P}}_{22}(-\mathbf{I} + \hat{\mathbf{W}}_{22}) \end{bmatrix} \boldsymbol{\Pi}, \end{aligned}$$

where $\hat{\mathbf{P}} = \begin{bmatrix} \hat{\mathbf{P}}_{11} & \hat{\mathbf{P}}_{12} \\ \hat{\mathbf{P}}_{21} & \hat{\mathbf{P}}_{22} \end{bmatrix} = \hat{\mathbf{P}}^T > \mathbf{0}$. Thus, by assumption,

$$\begin{aligned} &\boldsymbol{\Pi}^T \begin{bmatrix} \star & \star \\ \star & (-\mathbf{I} + \hat{\mathbf{W}}_{22}^T)\hat{\mathbf{P}}_{22} + \hat{\mathbf{P}}_{22}(-\mathbf{I} + \hat{\mathbf{W}}_{22}) \end{bmatrix} \boldsymbol{\Pi} < \mathbf{0} \\ &\Rightarrow \begin{bmatrix} \star & \star \\ \star & (-\mathbf{I} + \hat{\mathbf{W}}_{22}^T)\hat{\mathbf{P}}_{22} + \hat{\mathbf{P}}_{22}(-\mathbf{I} + \hat{\mathbf{W}}_{22}) \end{bmatrix} < \mathbf{0} \\ &\Rightarrow (-\mathbf{I} + \hat{\mathbf{W}}_{22}^T)\hat{\mathbf{P}}_{22} + \hat{\mathbf{P}}_{22}(-\mathbf{I} + \hat{\mathbf{W}}_{22}) < \mathbf{0}, \end{aligned}$$

proving that $-\mathbf{I} + \hat{\mathbf{W}}_{22}$ is Hurwitz. Since $\boldsymbol{\sigma}$ is arbitrary, $-\mathbf{I} + \mathbf{W}$ is totally Hurwitz.

(iv) The result follows from Lemma II.2(ii). ■

Remark II.4. ***(Counterexamples for converses of Lemma II.3).*** The converse of the implications in Lemma II.3 do not hold, as shown in the following. First, for a general matrix $\mathbf{W}$, neither of $\rho(|\mathbf{W}|)$ and $\|\mathbf{W}\|$ is bounded by the other. The former is larger for $[8,3;2,-1]$, e.g., while the latter is larger for $[0,0;1,0]$. However, if $\mathbf{W}$ satisfies the *Dale's law* (as many biological neuronal networks do), i.e., each column is either nonnegative or nonpositive, then

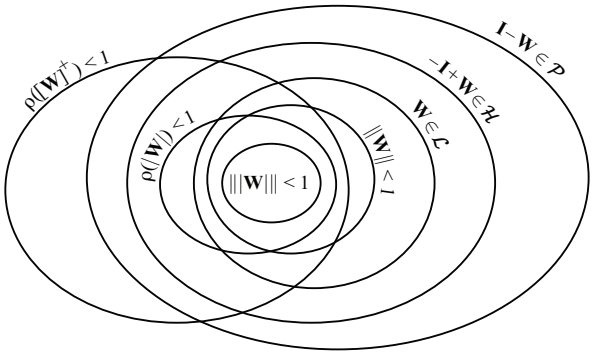


Fig. 1: Inclusion relationships between the matrix classes introduced in Definition II.1 (cf. Lemma II.3).

$\mathbf{W} = |\mathbf{W}|\mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix such that $|\mathbf{D}| = \mathbf{I}$. Then, $\|\mathbf{W}\| = \||\mathbf{W}|\| \geq \rho(|\mathbf{W}|)$, showing that, in this case, $\rho(|\mathbf{W}|) < 1$ is a less restrictive condition. Further, $-\mathbf{I} + \mathbf{W} \in \mathcal{H} \not\Rightarrow \rho(|\mathbf{W}|) < 1$ as seen, e.g., from $\mathbf{W} = -2\mathbf{I}$. The same example shows $\mathbf{W} \in \mathcal{L} \not\Rightarrow \|\mathbf{W}\| < 1$. Likewise, $-\mathbf{I} + \mathbf{W} \in \mathcal{H} \not\Rightarrow \mathbf{W} \in \mathcal{L}$, for which $[0.5, -3; 4, -1]$ serves as a counter example (note that $\mathbf{W} \in \mathcal{L}$ is an LMI feasibility problem that can be checked using standard solvers such as MATLAB `feasp` function). Finally, $\mathbf{A} = [-1, -5, 0; 0, -1, -6; -1, 0, -1]$ ensures the converse of Lemma II.3(iv) does not hold either. □

Figure 1 shows a summary of Lemma II.3 and Remark II.4.

Dynamical Rate Models of Brain Networks

Here we briefly review, following [46, §7], the construction of the linear-threshold network model used throughout the paper. In a lumped model, neurons are the smallest unit of neuronal circuits and the (directional) transmission of activity from one neuron to another takes place at a *synapse*, thus the terms *pre-synaptic* and *post-synaptic* for the two neurons, respectively. Both the input and output signals mainly consist of a sequence of spikes (action-potentials, Figure 2 top panel) which are modeled as impulse trains of the form

$$\rho(t) = \sum_k \delta(t - t_k),$$

where $\delta(\cdot)$ denotes the Dirac delta function. In many brain areas, the exact timing $\{t_k\}$ of $\rho(t)$ seems highly random while the firing rate (number of spikes per second, Figure 2 bottom panel) shows greater trial-to-trial reproducibility. Therefore, a standard approximation is to model $\rho(t)$ as the sample path of an inhomogeneous Poisson point process with rate, say, $x(t)$.

Now, consider a pair of pre- and post-synaptic neurons with rates $x_{\text{pre}}(t)$ and $x_{\text{post}}(t)$, respectively. As a result of $x_{\text{pre}}(t)$, an electrical current $I_{\text{post}}(t)$ flows in the post-synaptic neuron. Assuming fast synaptic dynamics, $I_{\text{post}}(t) \propto x_{\text{pre}}(t)$. Let $w_{\text{post,pre}}$ be the proportionality constant, so $I_{\text{post}}(t) = w_{\text{post,pre}} x_{\text{pre}}(t)$. The *pre-synaptic neuron* is called excitatory if $w_{\text{post,pre}} > 0$ and inhibitory if $w_{\text{post,pre}} < 0$. In other words, excitatory neurons increase the activity of their out-neighbors while inhibitory

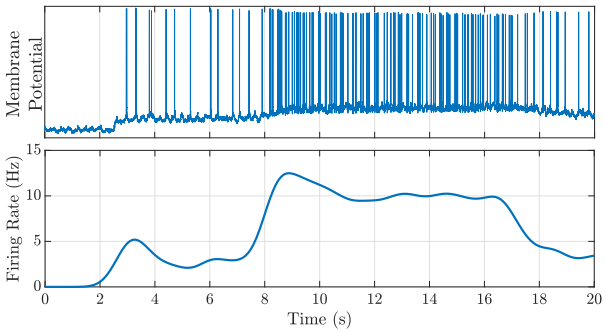


Fig. 2: A sample intracellular recording illustrating the spike train used for neuronal communication (top panel, measured intracellularly [47], [48]) and the corresponding (estimate of) firing rate (bottom panel, estimated by binning spikes in 100^{ms} bins and smoothing with Gaussian window with 500^{ms} standard deviation).

neurons decrease it.[3] If the post-synaptic neuron receives input from multiple neurons, $I_{\text{post}}(t)$ follows a superposition law,

$$I_{\text{post}}(t) = \sum_j w_{\text{post},j} x_j(t), \tag{1}$$

where the sum is taken over its in-neighbors.

If I_{post} is constant, the post-synaptic rate approximately follows $x_{\text{post}} = F(I_{\text{post}})$, where F is a nonlinear "response function". Among the two widely used response functions, sigmoidal and linear-threshold, we use the latter for its analytical tractability: $F(\cdot) = [\,\cdot\,]_0^{m_{\text{post}}}$. Finally, if $I_{\text{post}}(t)$ is time-varying, $x_{\text{post}}(t)$ "lags" $F(I_{\text{post}}(t))$ with a time constant τ, i.e.,

$$\tau \dot{x}_{\text{post}}(t) = -x_{\text{post}}(t) + [I_{\text{post}}(t)]_0^{m_{\text{post}}}. \tag{2}$$

Equations (1)-(2) are the basis for our network model described next.

III. Problem Formulation

Consider a network of neurons evolving according to (1)-(2). Since the number of neurons in a brain region is very large, it is common to consider a *population of neurons* with similar activation patterns as a single *node* with the average firing rate of its neurons. This convention also has the advantage of getting more consistent rates, as the firing pattern of individual neurons may be sparse.[4] Combining the nodal rates in a vector $\mathbf{x} \in \mathbb{R}^n$ and synaptic weights in a matrix $\mathbf{W} \in \mathbb{R}^{n \times n}$, we obtain, according to (1)-(2), the *linear-threshold network dynamics*

$$\tau \dot{\mathbf{x}}(t) = -\mathbf{x}(t) + [\mathbf{W}\mathbf{x}(t) + \mathbf{d}(t)]_{\mathbf{0}}^{\mathbf{m}}, \quad \mathbf{0} \leq \mathbf{x}(0) \leq \mathbf{m}, \quad \mathbf{m} \in \mathbb{R}^n_{>0} \cup \{\infty\}^n. \tag{3}$$

The term $\mathbf{d}(t) \in \mathbb{R}^n$ captures the *external inputs* to the network, including un-modeled background activity and possibly nonzero thresholds (i.e., if a node i becomes active when $(\mathbf{W}\mathbf{x} + \mathbf{d})_i > \vartheta_i$ for some threshold $\vartheta_i \neq 0$).

[3] While many brain networks, such as mammalian cortical networks, satisfy the Dale's law (cf. Remark II.4), all of our results in this work are applicable to arbitrary synaptic sign patterns.

[4] Our discussion is nevertheless valid irrespective of whether network nodes represent individual neurons or groups of them.

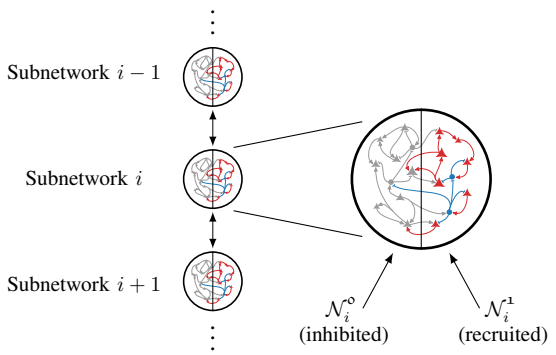


Fig. 3: The hierarchical network structure considered in this work. Each layer is only directly connected to the layers below and above it. Longer-range connections between non-successive layers do exist in thalamocortical hierarchies but are weaker than those between successive layers and are not considered in this work for simplicity.

The vector of state upper bounds $\mathbf{m}$ can be finite ($\mathbf{m} \in \mathbb{R}^n_{>0}$) or infinite ($\mathbf{m} = \infty\mathbf{1}_n$). Even though all biological neurons eventually saturate for high input values, whether finite or infinite $\mathbf{m}$ gives a more realistic/appropriate model can vary from brain region to brain region depending on whether typical (in vivo) values of $\mathbf{x}$ reach (near) saturation. Historically, the unbounded case has in fact been used and studied more extensively, both in computational neuroscience and machine learning. See, e.g., [49] and the references therein for evidence in favor of unbounded activation functions. Surprisingly, however, the analytical properties of the two cases are very similar, as we will see throughout this work. Further, note that the right-hand side of (3) is globally Lipschitz-continuous (though not smooth) and therefore a unique continuously differentiable solution exists for all $t \geq 0$ [50, Thm 3.2]. Moreover, it is straightforward to show that if $\mathbf{0} \leq \mathbf{x}(0) \leq \mathbf{m}$ then $\mathbf{0} \leq \mathbf{x}(t) \leq \mathbf{m}$ for all $t \geq 0$.

Consistent with the vision for hierarchical selective recruitment (HSR) outlined in Section I, we consider a hierarchy of linear-threshold networks of the form (3), as depicted in Figure 3. For each layer i, we use $\mathcal{N}_i$, $\mathcal{N}_i^{\mathbf{1}}$, and $\mathcal{N}_i^{\mathbf{0}}, i \in \{1, \dots, N\}$ to denote the corresponding subnetwork and its task-relevant and task-irrelevant sub-subnetworks, respectively.

Even when considered in isolation, each layer of the network exhibits rich dynamical behavior. In fact, simulations of (3) with random $\mathbf{W}$ and $\mathbf{d}$ reveal that

- locally, the dynamics may have zero, one, or many stable and/or unstable equilibrium points,
- globally, the dynamics can exhibit nonlinear phenomena such as limit cycles, multi-stability, and chaos,
- the state trajectories may grow unbounded (if $\mathbf{m} = \infty\mathbf{1}_n$) if the excitatory subnetwork $[\mathbf{W}]_{\mathbf{0}}^{\infty}$ is sufficiently strong.

This richness of behavior can only increase if layers are subject to time-varying inputs $\mathbf{d}(t)$ and, in particular, when interconnected with other layers in the hierarchy. Motivated by these observations, our ultimate goal in this work is to tackle four problems:

(i) the analysis of the relationship between structure ($\mathbf{W}$) and dynamical behavior (basic properties such as existence and uniqueness of equilibria (EUE), asymptotic stability, and boundedness of trajectories) for each subnetwork when operating in isolation from the rest of the network ($\mathbf{d}(t) \equiv \mathbf{d}$);

(ii) the analysis of the conditions on the joint structure of each two successive layers $\mathcal{N}_i$ and $\mathcal{N}_{i+1}$ that allows for selective inhibition of $\mathcal{N}_{i+1}^{\mathbf{0}}$ by its input from $\mathcal{N}_i$, being equivalent to the stabilization of $\mathcal{N}_{i+1}^{\mathbf{0}}$ to $\mathbf{0}$ (inactivity);

(iii) the analysis of the conditions on the joint structure of each two successive layers $\mathcal{N}_i$ and $\mathcal{N}_{i+1}$ that allows for top-down recruitment of $\mathcal{N}_{i+1}^{\mathbf{1}}$ by its input from $\mathcal{N}_i$, being equivalent to the stabilization of $\mathcal{N}_{i+1}^{\mathbf{1}}$ toward a desired trajectory set by $\mathcal{N}_i$ (activity);

(iv) the combination of (ii) and (iii) in a unified framework and its extension to the complete N-layer network of networks.

Problems (i) and (ii) are the focus of this paper, whereas we address problems (iii) and (iv) in the accompanying work [19].

IV. Internal Dynamics of Single-Layer Networks

In this section, we provide an in-depth study of the basic dynamical properties of the network dynamics (3) in isolation. In such case, the external input $\mathbf{d}(t)$ boils down to background activity and possibly nonzero thresholds, which are constant relative to the timescale τ. The dynamics (3) thus simplify to

$$\tau\dot{\mathbf{x}}(t) = -\mathbf{x}(t) + [\mathbf{W}\mathbf{x}(t) + \mathbf{d}]_{\mathbf{0}}^{\mathbf{m}}, \quad \mathbf{0} \leq \mathbf{x}(0) \leq \mathbf{m}, \quad \mathbf{m} \in \mathbb{R}^n_{>0} \cup \{\infty\}^n. \tag{4}$$

In the following, we derive conditions in terms of the network structure for the existence and uniqueness of equilibria (EUE), asymptotic stability, and boundedness of trajectories.

A. Dynamics as Switched Affine System

The nonlinear dynamics (4) is a switched affine system with 2^n modes if $\mathbf{m} = \infty\mathbf{1}_n$ or 3^n modes if $\mathbf{m} < \infty\mathbf{1}_n$. Each mode of this system corresponds to a switching index $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{x}) \in \{0, \ell, \mathrm{s}\}^n$, where for each $i \in \{1, \dots, n\}$, $\sigma_i = 0$ if the node is inactive (i.e., $(\mathbf{W}\mathbf{x} + \mathbf{d})_i \leq 0$), $\sigma_i = \ell$ if the node is in linear regime (i.e., $(\mathbf{W}\mathbf{x} + \mathbf{d})_i \in [0, m_i)$), and $\sigma_i = \mathrm{s}$ if the node is saturated (i.e., $(\mathbf{W}\mathbf{x}+\mathbf{d})_i \geq m_i$). Clearly, the mode of the system is state-dependent and each switching index $\boldsymbol{\sigma} \in \{0, \ell, \mathrm{s}\}^n$ corresponds to a switching region

$$\begin{aligned}\Omega_{\boldsymbol{\sigma}} = \{\mathbf{x} \mid\ & (\mathbf{W}\mathbf{x} + \mathbf{d})_i \in (-\infty, 0], \quad \forall i \ \text{ s.t. } \ \sigma_i = 0, \ \text{and} \\ & (\mathbf{W}\mathbf{x} + \mathbf{d})_i \in [0, m_i], \quad \forall i \ \text{ s.t. } \ \sigma_i = \ell, \ \text{and} \\ & (\mathbf{W}\mathbf{x} + \mathbf{d})_i \in [m_i, \infty), \quad \forall i \ \text{ s.t. } \ \sigma_i = \mathrm{s}\}.\end{aligned}$$

Within each $\Omega_{\boldsymbol{\sigma}}$, we have

$$[\mathbf{W}\mathbf{x}(t) + \mathbf{d}]_{\mathbf{0}}^{\mathbf{m}} = \boldsymbol{\Sigma}^{\ell}(\mathbf{W}\mathbf{x}(t) + \mathbf{d}) + \boldsymbol{\Sigma}^{s}\mathbf{m},$$

where $\boldsymbol{\Sigma}^{\ell} = \boldsymbol{\Sigma}^{\ell}(\boldsymbol{\sigma})$ is a diagonal matrix with $\Sigma^{\ell}_{ii} = 1$ if $\sigma_i = \ell$ and $\Sigma^{\ell}_{ii} = 0$ otherwise. $\boldsymbol{\Sigma}^{\mathrm{s}}$ is defined similarly, and we set the convention that $\boldsymbol{\Sigma}^{\mathrm{s}}\mathbf{m} = \mathbf{0}$ if $\mathbf{m} = \infty\mathbf{1}_n$. Therefore, (4) can be written in the equivalent piecewise-affine form

$$\tau\dot{\mathbf{x}} = (-\mathbf{I} + \boldsymbol{\Sigma}^{\ell}\mathbf{W})\mathbf{x} + \boldsymbol{\Sigma}^{\ell}\mathbf{d} + \boldsymbol{\Sigma}^{\mathrm{s}}\mathbf{m}, \qquad \forall \mathbf{x} \in \Omega_{\boldsymbol{\sigma}}. \tag{5}$$

This switched representation of the dynamics motivates the following assumptions on the weight matrix $\mathbf{W}$.

Assumption 1. *Assume*

(i) $det(\mathbf{W}) \neq 0$*;*

(ii) $det(\mathbf{I} - \mathbf{\Sigma}^\ell \mathbf{W}) \neq 0$ *for all the* 2^n *matrices* $\mathbf{\Sigma}^\ell(\boldsymbol{\sigma}), \boldsymbol{\sigma} \in \{0, \ell, \mathrm{s}\}^n$.

Assumption 1 is not a restriction in practice since the set of matrices for which it is not satisfied can be expressed as a finite union of measure-zero sets, and hence has measure zero. By Assumption 1(i), the system of equations $\mathbf{W}\mathbf{x}+\mathbf{d} = \mathbf{0}$ defines a non-degenerate set of n hyperplanes partitioning $\mathbb{R}^n$ into 2^n solid convex polytopic translated cones apexed at $-\mathbf{W}^{-1}\mathbf{d}$.[5]

Unlike linear systems, the existence of equilibria is not guaranteed for (5). Rather, each $\boldsymbol{\sigma} \in \{0, \ell, \mathrm{s}\}^n$ corresponds to an *equilibrium candidate*

$$\mathbf{x}^*_{\boldsymbol{\sigma}} = (\mathbf{I} - \mathbf{\Sigma}^\ell \mathbf{W})^{-1}(\mathbf{\Sigma}^\ell \mathbf{d} + \mathbf{\Sigma}^{\mathrm{s}} \mathbf{m}), \tag{6}$$

which is an equilibrium if it belongs to $\Omega_{\boldsymbol{\sigma}}$. We next identify conditions for this to be true.

B. Existence and Uniqueness of Equilibria (EUE)

The first step in analyzing any dynamical system is the characterization of its equilibria. We being our analysis of the EUE with the case of bounded activation functions ($\mathbf{m} \in \mathbb{R}^n_{>0}$).

Theorem IV.1. ***(EUE).*** *The network dynamics* (4) *has a unique equilibrium for all* $\mathbf{d} \in \mathbb{R}^n$ *if and only if* $\mathbf{I} - \mathbf{W} \in \mathcal{P}$.

Proof: Despite their similarity, different equivalences need to be established and results need to be invoked for the bounded and unbounded cases. Therefore, we prove the result separately for each case.

Case 1: $\mathbf{m} < \infty \mathbf{1}_n$. The existence of equilibria is guaranteed by the Brouwer fixed point theorem [52] for all $\mathbf{W}$ and all $\mathbf{d}$. We use results from [53] to characterize uniqueness. Following the terminology therein, the set $\mathcal{C} = \{\Omega_{\boldsymbol{\sigma}} \mid \boldsymbol{\sigma} \in \{0, \ell, \mathrm{s}\}^n\}$ is a chamber system and its branching number is 4 by Assumption 1. Let $f(\mathbf{x}; \mathbf{d}) = \mathbf{x} - [\mathbf{W}\mathbf{x} + \mathbf{d}]^{\mathbf{m}}_{\mathbf{0}}$ which, for any $\mathbf{d}$, is piecewise-affine on the chamber system $\mathcal{C}$ by (5) and is proper since $\|f(\mathbf{x}; \mathbf{d})\| \to \infty$ whenever $\|\mathbf{x}\| \to \infty$.

First, assume that $\mathbf{I}-\mathbf{W} \in \mathcal{P}$. Then, f is coherently oriented by definition and thus [53, Thm 5.3] ensures that f is bijective. In particular, there exists a unique $\mathbf{x}$ such that $f(\mathbf{x}; \mathbf{d}) = \mathbf{0}$, giving uniqueness of equilibria for any $\mathbf{d}$.

Now, assume that $\mathbf{I} - \mathbf{W} \notin \mathcal{P}$. Since the determinant of $\mathbf{I} - \mathbf{\Sigma}^\ell \mathbf{W}$ is always positive on the chamber $\Omega_{\mathbf{0}}$, f cannot be coherently oriented, thus not bijective by [53, Thm 5.3], and thus not injective by [53, Cor 5.2]. Therefore, there exists $\mathbf{x}_1, \mathbf{x}_2, \mathbf{z} \in \mathbb{R}^n$ such that $\mathbf{x}_1 \neq \mathbf{x}_2$ but

$$\mathbf{x}_1 - [\mathbf{W}\mathbf{x} + \mathbf{d}]^{\mathbf{m}}_{\mathbf{0}} = \mathbf{z} = \mathbf{x}_2 - [\mathbf{W}\mathbf{x} + \mathbf{d}]^{\mathbf{m}}_{\mathbf{0}},$$

where $\mathbf{d} \in \mathbb{R}^n$ is arbitrary. Therefore, $f(\,\cdot\,; \mathbf{W}\mathbf{z} + \mathbf{d})$ has two distinct roots, $\mathbf{x}_1 - \mathbf{z}$ and $\mathbf{x}_2 - \mathbf{z}$, proving the necessity of $\mathbf{I} - \mathbf{W} \in \mathcal{P}$ for uniqueness of equilibria.

Case 2: $\mathbf{m} = \infty \mathbf{1}_n$. In this case, we simply show the equivalence between our equilibrium equation $\mathbf{x} = [\mathbf{W}\mathbf{x} + \mathbf{d}]^\infty_{\mathbf{0}}$ and the well-studied linear complementarity problem (LCP). By Assumption 1,

$$\mathbf{x} = [\mathbf{W}\mathbf{x} + \mathbf{d}]^\infty_{\mathbf{0}} \Leftrightarrow \mathbf{W}\mathbf{x} + \mathbf{d} = \mathbf{W}[\mathbf{W}\mathbf{x} + \mathbf{d}]^\infty_{\mathbf{0}} + \mathbf{d}. \tag{7}$$

We next perform a change of variables as follows. Let

$$\mathbf{z} = [\mathbf{W}\mathbf{x} + \mathbf{d}]^\infty_{\mathbf{0}} \qquad \text{and} \qquad \mathbf{w} = [-\mathbf{W}\mathbf{x} - \mathbf{d}]^\infty_{\mathbf{0}}.$$

These vectors have the properties that

$$\mathbf{z}, \mathbf{w} \in \mathbb{R}^n_{\geq 0}, \qquad \mathbf{z}^T\mathbf{w} = 0, \quad \text{and} \quad \mathbf{W}\mathbf{x} + \mathbf{d} = \mathbf{z} - \mathbf{w}.$$

and thus provide a unique (invertible) characterization of $\mathbf{x}$. Therefore, (7) is equivalent to

$$\mathbf{w} = (\mathbf{I} - \mathbf{W})\mathbf{z} - \mathbf{d}, \qquad \mathbf{z}, \mathbf{w} \in \mathbb{R}^n_{\geq 0}, \qquad \mathbf{z}^T\mathbf{w} = 0,$$

which is the standard LCP and has a unique solution $(\mathbf{z}, \mathbf{w})$ for all $\mathbf{d} \in \mathbb{R}^n$ if and only if $\mathbf{I}-\mathbf{W} \in \mathcal{P}$ [54]. Similar to Case 1, it can also be shown that if $\mathbf{I}-\mathbf{W} \notin \mathcal{P}$, there exists at least one $\mathbf{d}$ for which two equilibrium points exists (see, e.g., the proof of [54, Thm 4.2]). This completes the proof. ■

Even though the condition $\mathbf{I} - \mathbf{W} \in \mathcal{P}$ may seem abstract, it has a nice geometric interpretation. From [53, Lem 2.2], $\mathbf{I} - \mathbf{W} \in \mathcal{P}$ if and only if the (negative) vector field $\mathbf{x} \mapsto \mathbf{x} - [\mathbf{W}\mathbf{x} + \mathbf{d}]^{\mathbf{m}}_{\mathbf{0}}$ maps each switching region Ω_σ to another polytopic region and the images of adjacent switching regions remains adjacent. In other words, the vector field has a *coherent orientation* when mapping the state space.[6]

Remark IV.2. ***(Computational complexity of verifying*** $\mathbf{I} - \mathbf{W} \in \mathcal{P}$***).*** Although the problem of determining whether a matrix is in $\mathcal{P}$ is straightforward for small n, it is known to be co-NP-complete [56], and thus expensive for large networks. Indeed, [57] shows that all the 2^n principal minors of $\mathbf{A}$ have to be checked to prove $\mathbf{A} \in \mathcal{P}$ (though disproving $\mathbf{A} \in \mathcal{P}$ is usually much easier). In these cases, one may need to rely on more conservative sufficient conditions such as $\rho(|\mathbf{W}|) < 1$ or $\|\mathbf{W}\| < 1$ (cf. Lemma II.3) to establish $\mathbf{I} - \mathbf{W} \in \mathcal{P}$. These conditions, moreover, have the added benefit of providing intuitive connections between the distribution of synaptic weights, network size, and stability. We elaborate more on this point in Section V-C. □

Example IV.3. ***(Wilson-Cowan model).*** Consider a network of n nodes in which $\alpha n, \alpha \in (0, 1)$ are excitatory, $(1-\alpha)n$ are inhibitory. Under some regularity assumptions, given next, this network can be (further) reduced to a simple, two-dimensional network commonly known as the Wilson-Cowan model [58] and widely used in computational neuroscience [59], [60]. Assume that the synaptic weight between any pair of nodes, the external input to them, and their maximal firing rate (if finite) only depends on their type: the synaptic weight of any inhibitory-to-excitatory connection is $w_{ei} < 0$, similarly for $w_{ee} > 0, w_{ie} > 0, w_{ii} < 0$, and all excitatory (inhibitory) nodes receive $d_e \in \mathbb{R}$ ($d_i \in \mathbb{R}$) and have maximal rate $m_e \in \mathbb{R}_{>0} \cup \{\infty\}$ ($m_i \in \mathbb{R}_{>0} \cup \{\infty\}$). Let $x_e(t)$ and $x_i(t)$

[5] Recall that a set of n hyperplanes is *non-degenerate* [51] if their intersection is a point or, equivalently, the matrix composed of their normal vectors is nonsingular. A set $S \subseteq \mathbb{R}^n$ is called a *polytope* if it has the form $S = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$; a *cone* if $c\mathbf{x} \in S$ for any $\mathbf{x} \in S, c \in \mathbb{R}_{\geq 0}$; a *translated cone apexed at* $\mathbf{y}$ if $\{\mathbf{x} | \mathbf{x}+\mathbf{y} \in S\}$ is a cone; *convex* if $(1-\theta)\mathbf{x}+\theta\mathbf{y} \in S$ for any $\mathbf{x}, \mathbf{y} \in S, \theta \in [0, 1]$; and *solid* if it has a non-empty interior.

[6] A closely-related class of matrices is that of M-matrices [45] with established relationships with the stability of nonlinear systems, see, e.g., [55].

be the average firing rates of excitatory and inhibitory nodes, respectively. Then, (4) simplifies to

$$\tau \begin{bmatrix} \dot{x}_e \\ \dot{x}_i \end{bmatrix} = -\begin{bmatrix} x_e \\ x_i \end{bmatrix} + \left[\begin{bmatrix} \alpha n w_{ee} & (1-\alpha) n w_{ei} \\ \alpha n w_{ie} & (1-\alpha) n w_{ii} \end{bmatrix} \begin{bmatrix} x_e \\ x_i \end{bmatrix} + \begin{bmatrix} d_e \\ d_i \end{bmatrix} \right]_{\mathbf{0}}^{\mathbf{m}}.$$

Let $\mathbf{W}_{EI} \in \mathbb{R}^{2\times 2}$ be the corresponding weight matrix above. One can check that

$$\mathbf{I} - \mathbf{W}_{EI} \in \mathcal{P} \Leftrightarrow \alpha n w_{ee} < 1, \tag{8}$$

and

$$\rho(|\mathbf{W}_{EI}|) < 1 \Leftrightarrow \alpha n w_{ee} < 1,\ (1-\alpha) n |w_{ii}| < 1, \text{ and}$$
$$\alpha(1-\alpha) n^2 w_{ie} |w_{ei}| < (1 - \alpha n w_{ee})(1 - (1-\alpha) n |w_{ii}|).$$

Thus, according to Theorem IV.1, EUE only requires the excitatory dynamics to be stable (note that w_{ee} has to become smaller as n grows), while the more conservative condition $\rho(|\mathbf{W}_{EI}|) < 1$ also requires (relatively) weak inhibitory-inhibitory synapses and a weak interconnection between excitatory and inhibitory subnetworks. □

When $\mathbf{I} - \mathbf{W} \in \mathcal{P}$, Theorem IV.1 ensures EUE for all $\mathbf{d} \in \mathbb{R}^n$. When $\mathbf{I} - \mathbf{W} \notin \mathcal{P}$, however, a more involved question is to find the values of $\mathbf{d}$ that give rise to non-unique equilibrium points. To answer this, we next perform a more direct analysis of the equilibria. For simplicity, we focus in the remainder of this section on unbounded dynamics ($\mathbf{m} = \infty \mathbf{1}_n$).

Recall the definition of an equilibrium candidate in (6). Using Assumption 1, and after some manipulations, we have

$$\begin{aligned} \mathbf{W}\mathbf{x}^*_{\boldsymbol{\sigma}} + \mathbf{d} &= \mathbf{W}(\mathbf{I} - \boldsymbol{\Sigma}^\ell \mathbf{W})^{-1} \boldsymbol{\Sigma}^\ell \mathbf{d} + \mathbf{d} \\ &= (\mathbf{W}^{-1} - \boldsymbol{\Sigma}^\ell)^{-1} \boldsymbol{\Sigma}^\ell \mathbf{d} + \mathbf{d} \\ &= (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}^\ell)^{-1} \mathbf{W} \boldsymbol{\Sigma}^\ell \mathbf{d} + \mathbf{d} \\ &= \left[(\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}^\ell)^{-1} \mathbf{W} \boldsymbol{\Sigma}^\ell + \mathbf{I} \right] \mathbf{d} = (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}^\ell)^{-1} \mathbf{d}, \end{aligned} \tag{9}$$

thus,

$$\mathbf{x}^*_{\boldsymbol{\sigma}} \in \Omega_{\boldsymbol{\sigma}} \Leftrightarrow \underbrace{(2\boldsymbol{\Sigma}^\ell - \mathbf{I})(\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}^\ell)^{-1}}_{\triangleq \mathbf{M}_{\boldsymbol{\sigma}}} \mathbf{d} \geq \mathbf{0}. \tag{10}$$

Therefore, if $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d} \geq \mathbf{0}$ for exactly one $\boldsymbol{\sigma} \in \{0,\ell\}^n$, then a unique equilibrium exists. However, when $\mathbf{M}_{\boldsymbol{\sigma}_\ell}\mathbf{d} \geq \mathbf{0}$ for multiple $\boldsymbol{\sigma}_\ell \in \{0,\ell\}^n, \ell \in \{1,\dots,\bar{\ell}\}$, the network may have either multiple equilibria or a unique one $\mathbf{x}^*_{\boldsymbol{\sigma}_1} = \cdots = \mathbf{x}^*_{\boldsymbol{\sigma}_{\bar{\ell}}}$ lying on the boundary between $\{\Omega_{\sigma_\ell}\}_{\ell=1}^{\bar{\ell}}$. The next result shows that the quantities $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d}$ can be used to distinguish between these two latter cases.

Lemma IV.4. ***(Existence of multiple equilibria).*** *Assume* $\mathbf{W}$ *satisfies Assumption 1,* $\mathbf{d} \in \mathbb{R}^n$ *is arbitrary, and* $\mathbf{M}_{\boldsymbol{\sigma}}$ *is defined as in* (10) *for* $\boldsymbol{\sigma} \in \{0,\ell\}^n$. *If there exist* $\boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2$ *such that* $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d} \geq \mathbf{0}$ *and* $\mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d} \geq \mathbf{0}$*, then* $\mathbf{x}^*_{\boldsymbol{\sigma}_1} = \mathbf{x}^*_{\boldsymbol{\sigma}_2}$ *if and only if* $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d} = \mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d}$.

Proof: Clearly,

$$\begin{aligned} \mathbf{x}^*_{\boldsymbol{\sigma}_1} = \mathbf{x}^*_{\boldsymbol{\sigma}_2} &\Leftrightarrow \mathbf{W}\mathbf{x}^*_{\boldsymbol{\sigma}_1} + \mathbf{d} = \mathbf{W}\mathbf{x}^*_{\boldsymbol{\sigma}_2} + \mathbf{d} \\ &\Leftrightarrow (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}_1)^{-1}\mathbf{d} = (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}_2)^{-1}\mathbf{d}, \end{aligned} \tag{11}$$

where we have used (9). Since both $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d}$ and $\mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d}$ are nonnegative, (11) holds if and only if $\left((\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}_1)^{-1}\mathbf{d}\right)_i = \left((\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}_2)^{-1}\mathbf{d}\right)_i = 0$ for any i such that $\sigma_{1,i} \neq \sigma_{2,i}$, which is equivalent to $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d} = \mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d}$. ■

This property of $\mathbf{M}_{\boldsymbol{\sigma}}$ can be used to derive a computationally more involved but input-dependent characterization of EUE, as follows.

Proposition IV.5. ***(Optimization-based condition for EUE).*** *Let* $\mathbf{W}$ *satisfy Assumption 1 and* $\mathbf{M}_{\boldsymbol{\sigma}}$ *be as defined in* (10) *for* $\boldsymbol{\sigma} \in \{0,\ell\}^n$. *For* $\mathbf{d} \in \mathbb{R}^n$*, define* $\mu_1(\mathbf{d})$ *and* $\mu_2(\mathbf{d})$ *to be the largest and second largest elements of the set*

$$\left\{ \min_{i=1,\dots,n} (\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d})_i \mid \boldsymbol{\sigma} \in \{0,\ell\}^n \right\},$$

respectively. Then, (4) *has a unique equilibrium for each* $\mathbf{d} \in \mathbb{R}^n$ *if and only if*

$$\max_{\|\mathbf{d}\|=1} \mu_1(\mathbf{d})\mu_2(\mathbf{d}) < 0. \tag{12}$$

Proof: First, note that $\mathbf{d} = \mathbf{0}$ is a degenerate case where the origin is the unique equilibrium belonging to all $\Omega_{\boldsymbol{\sigma}}$. For any $\mathbf{d} \neq \mathbf{0}$ and $\boldsymbol{\sigma} \in \{0,\ell\}^n$, $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d} \geq \mathbf{0}$ if and only if $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d}/\|\mathbf{d}\| \geq \mathbf{0}$. Thus, EUE for all $\mathbf{d} \in \mathbb{R}^n$ and all $\|\mathbf{d}\| = 1$ are equivalent. Then, for any $\mathbf{d}$,

$$\begin{aligned} \mu_1(\mathbf{d})\mu_2(\mathbf{d}) < 0 &\Leftrightarrow \mu_1(\mathbf{d}) > 0 \text{ and } \mu_2(\mathbf{d}) < 0 \\ &\Leftrightarrow \exists \text{ unique } \mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d} \geq \mathbf{0}, \quad \boldsymbol{\sigma} \in \{0,\ell\}^n. \end{aligned} \tag{13}$$

Note that the latter allows for the possibility of the existence of multiple $\boldsymbol{\sigma}$ with $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d} \geq \mathbf{0}$ provided that they have the same value of $\mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d}$. By Lemma IV.4, (13) is then equivalent to EUE, completing the proof. ■

In our experience, the optimization involved in (12) is usually highly non-convex but since the search space $\{\|\mathbf{d}\| = 1\}$ is compact, global search methods can be used to verify (12) numerically if n is not too large. However, note that our main interest in (12) (being equivalent to $\mathbf{I} - \mathbf{W} \in \mathcal{P}$) is when it does *not* hold. If so, then any $\mathbf{d}$ for which (12) fails gives a ray $\{\alpha\mathbf{d} \mid \alpha > 0\}$ of input values that give rise to non-unique equilibria. Combined with stability analysis of Section IV-C, e.g., this can be a basis for the characterization of multistability in linear-threshold dynamics which is itself beyond the scope of this work.

The proof of Theorem IV.1 (for the unbounded case) is based on the LCP, which makes the relationship between $\mathbf{I} - \mathbf{W} \in \mathcal{P}$ and EUE opaque, even when taking into account the proof of the LCP. The equilibrium characterization in (10), however, can be used to explain this relationship more transparently. For any given $\mathbf{d}$, the non-uniqueness of equilibria is equivalent to asking whether

$$\begin{aligned} \exists \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \in \{0,\ell\}^n \quad \text{s.t.} \quad & \mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d} \geq \mathbf{0} \ \text{ and } \ \mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d} \geq \mathbf{0} \\ & \mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{d} \neq \mathbf{M}_{\boldsymbol{\sigma}_2}\mathbf{d}, \end{aligned}$$

or, whether there exist $\mathbf{q}_{\boldsymbol{\sigma}_1} \neq \mathbf{q}_{\boldsymbol{\sigma}_2} \in \mathbb{R}^n_{\geq 0}$ such that $\mathbf{M}_{\boldsymbol{\sigma}_1}^{-1}\mathbf{q}_{\boldsymbol{\sigma}_1} = \mathbf{M}_{\boldsymbol{\sigma}_2}^{-1}\mathbf{q}_{\boldsymbol{\sigma}_2} = \mathbf{d}$. A more general question, which turns out to be relevant to EUE, is whether

$$\exists \mathbf{q}_{\boldsymbol{\sigma}_1} \neq \mathbf{q}_{\boldsymbol{\sigma}_2} \in \mathcal{O}_n \quad \text{s.t.} \quad \mathbf{q}_{\boldsymbol{\sigma}_1} = \mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1}\mathbf{q}_{\boldsymbol{\sigma}_2}, \tag{14}$$

for *any* orthant $\mathcal{O}_n$ of $\mathbb{R}^n$ (including $\mathcal{O}_n = \mathbb{R}^n_{\geq 0}$ as a special case). This depends on whether the matrix $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1}$ can map any nonzero vector to the same orthant which, by

Lemma II.2(iii), happens if and only if $-\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} \notin \mathcal{P}$. The following result, whose proof is in Appendix A, gives a necessary and sufficient condition for this to not happen.

Theorem IV.6. ***(Coherently oriented vector fields and the validity of equilibrium candidates).*** *Let* $\mathbf{W}$ *satisfy Assumption 1 and* $\mathbf{M}_\sigma$ *be defined as in* (10)*. Then,* $-\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} \in \mathcal{P}$ *for all (distinct)* $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \in \{0,\ell\}^n$ *if and only if* $\mathbf{I}-\mathbf{W} \in \mathcal{P}$.

Theorem IV.6 provides a more transparent account of the relationship between $\mathbf{I}-\mathbf{W} \in \mathcal{P}$ and EUE. If $\mathbf{I}-\mathbf{W} \in \mathcal{P}$, then Theorem IV.6 and Lemma II.2(iii) ensure that none of $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1}$ can map a vector to the same orthant. Thus, no two $\mathbf{q}_{\boldsymbol{\sigma}} = \mathbf{M}_{\boldsymbol{\sigma}}\mathbf{d}$ belong to the same orthant. Therefore, there exists a one-to-one correspondence between $\{\mathbf{q}_{\boldsymbol{\sigma}}\}$ and orthants in $\mathbb{R}^n$, ensuring that exactly one $\mathbf{q}_{\boldsymbol{\sigma}}$ belongs to $\mathbb{R}^n_{\geq 0}$, i.e., EUE.[7]

We end this subsection with a result that bounds the number and location if equilibria for the case when $\mathbf{I}-\mathbf{W} \notin \mathcal{P}$. For $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{\sigma} \in \{0,\ell\}^n$, let $\mathbf{A}_{(\boldsymbol{\sigma})}$ be the principal submatrix of $\mathbf{A}$ containing the rows and columns for which $\sigma_i = \ell$. Further, for $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \in \{0,\ell\}^n$, we say $\boldsymbol{\sigma}_1 \leq \boldsymbol{\sigma}_2$ if $\sigma_{1,i} = \ell \Rightarrow \sigma_{2,i} = \ell$ for all $i \in \{1,\dots,n\}$.

Corollary IV.7. ***(Partial EUE).*** *Consider the dynamics* (4) *and assume that Assumption 1 holds. If* $\mathbf{I}-\mathbf{W}_{(\bar{\boldsymbol{\sigma}})} \in \mathcal{P}$ *for any* $\bar{\boldsymbol{\sigma}} \in \{0,\ell\}^n$*, then* $\bigcup_{\boldsymbol{\sigma}\leq\bar{\boldsymbol{\sigma}}} \Omega_{\boldsymbol{\sigma}}$ *contains at most one equilibrium point.*

Proof: The proof follows directly from the proof of Theorem IV.6 and the fact that, using the definitions therein, $-\boldsymbol{\Gamma} \in \mathcal{P}$ only requires $\mathbf{I}-\mathbf{W}_{([\boldsymbol{\ell}_{n_1+n_2+n_3}\ \mathbf{0}_{n_4}]^T)} \in \mathcal{P}$. ∎

Even in the simplest case when $\mathbf{I}-\mathbf{W} \in \mathcal{P}$, the resulting unique equilibrium may or may not be stable, as studied next.

C. Asymptotic Stability

The EUE is an *opportunity* to shape the network activity at steady state, provided that the equilibrium corresponds to a desired state (a memory, the encoding of a spatial location, eye position, etc. [61]–[65]) *and* it attracts network trajectories. Here we investigate the latter.

Theorem IV.8. ***(Asymptotic Stability).*** *Consider the network dynamics* (4) *and assume* $\mathbf{W}$ *satisfies Assumption 1.*

(i) [Sufficient condition] If $\mathbf{W} \in \mathcal{L}$*, then for all* $\mathbf{d} \in \mathbb{R}^n$*, the network is globally exponentially stable (GES) relative to a unique equilibrium* $\mathbf{x}^*$*;*

(ii) [Necessary condition] If for all $\mathbf{d} \in \mathbb{R}^n$ *the network is locally asymptotically stable relative to a unique equilibrium* $\mathbf{x}^*$*, then* $-\mathbf{I}+\mathbf{W} \in \mathcal{H}$.

Proof: (i) The EUE follows from Lemma II.3(iii)&(iv) and Theorem IV.1. GES can be deduced from [66, Thm 1], but a simpler and direct proof can also be found in a preliminary version of this work [67] which is omitted here space reasons.

(ii) Assume, by contradiction, that $-\mathbf{I}+\mathbf{W} \notin \mathcal{H}$, which means that there exists $\boldsymbol{\sigma} \in \{0,\ell\}^n$ such that $-\mathbf{I}+\boldsymbol{\Sigma}^\ell\mathbf{W}$ is not Hurwitz. Let $\boldsymbol{\sigma}_{01} \in \{0,1\}^n$ have the same zeros as $\boldsymbol{\sigma}$, and consider the choice

$$\mathbf{d} = (2\mathbf{I}-\mathbf{W})\boldsymbol{\sigma}_{01} - \mathbf{1}_n.$$

It is straightforward to show that $\mathbf{x}^* = \boldsymbol{\sigma}_{01}$ is an equilibrium point for (4) lying in the interior of $\Omega_{\boldsymbol{\sigma}}$. By assumption, $\mathbf{x}^*$ is (unique and) locally asymptotically stable, which contradicts $-\mathbf{I}+\boldsymbol{\Sigma}^\ell\mathbf{W}$ not being Hurwitz. This completes the proof. ∎

Similar to the problem of verifying whether a matrix is a P-matrix, cf. Remark IV.2, the computational complexity of verifying total-Hurwitzness grows exponentially with n. The same applies to the verification of total $\mathcal{L}$-stability, see, e.g., [68] and the references therein. The next result gives a usually more conservative but computationally inexpensive alternative.

Proposition IV.9. ***(Computationally feasible sufficient conditions for GES).*** *Consider the network dynamics* (4) *and assume the weight matrix* $\mathbf{W}$ *satisfies Assumption 1. If* $\rho(|\mathbf{W}|) < 1$ *or* $\|\mathbf{W}\| < 1$*, then for all* $\mathbf{d} \in \mathbb{R}^n$*, the network has a unique equilibrium* $\mathbf{x}^*$ *and it is GES relative to* $\mathbf{x}^*$.

Proof: If $\|\mathbf{W}\| < 1$, the result follows from Lemma II.3(ii) and Theorem IV.8. For the case $\rho(|\mathbf{W}|) < 1$, the same proof technique as in [25, Prop. 3] can be used to prove GES, as shown in a preliminary version of this work [67], but is omitted here for space reasons. ∎

From Lemma II.3(iii), the conditions of Theorem IV.8 and Proposition IV.9 are not conclusive when $\mathbf{W}$ satisfies $-\mathbf{I}+\mathbf{W} \in \mathcal{H}$ but neither $\mathbf{W} \in \mathcal{L}$ nor $\rho(|\mathbf{W}|) < 1$. However,

(i) If a unique equilibrium $\mathbf{x}^*$ lies in the interior of a switching region (a condition that can be shown to hold for almost all $\mathbf{d}$), then $\mathbf{x}^*$ is at least locally exponentially stable (since a sufficiently small region of attraction of $\mathbf{x}^*$ is contained in that switching region).

(ii) In our extensive simulations with random $(\mathbf{W},\mathbf{d})$, any system satisfying $-\mathbf{I}+\mathbf{W} \in \mathcal{H}$ was GES for all $\mathbf{d}$.

These observations lead us to the following conjecture, whose analytic characterization remains an open problem.

Conjecture IV.10. ***(Sufficiency of total-Hurwitzness for GES).*** *Consider the dynamics* (4) *and assume* $\mathbf{W}$ *satisfies Assumption 1. The network has a unique GES equilibrium for all* $\mathbf{d} \in \mathbb{R}^n$ *if and only if* $-\mathbf{I}+\mathbf{W} \in \mathcal{H}$.

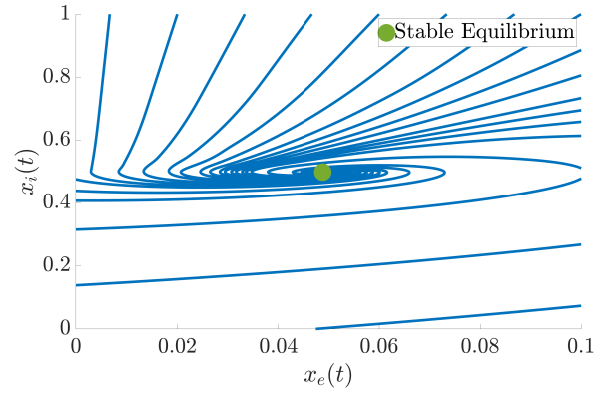
We next study the GES of the uniform excitatory-inhibitory networks of Example IV.3.

Example IV.11. ***(Wilson-Cowan model, cont'd).*** Consider the Wilson-Cowan model of Example IV.3. One can verify

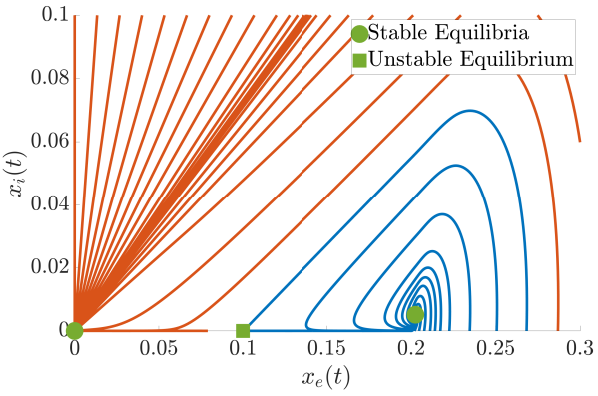
$$-\mathbf{I}+\mathbf{W}_{EI} \in \mathcal{H} \Leftrightarrow \alpha n w_{ee} < 1, \tag{15}$$

thus being equivalent (in this two-dimensional case) to $\mathbf{I}-\mathbf{W}_{EI} \in \mathcal{P}$ and, interestingly, only restricting w_{ee} while w_{ei}, w_{ie}, and w_{ii} are completely free. Figure 4 shows sample phase portraits for the cases $\alpha n w_{ee} < 1$ and $\alpha n w_{ee} > 1$, matching our expectations from Theorems IV.1 and IV.8. While our focus here is on the existence, uniqueness, and stability of equilibria, it is instructive to highlight the role of equilibrium analysis and, in particular, *lack of stable equilibria* in the generation of oscillations in the same linear-threshold Wilson-Cowan model [69]. In this case, both the linear-threshold Wilson-Cowan model and the popular Kuramoto model [70]–[72] of neural oscillations provide parallel simplifications to the (more biologically accurate) Wilson-Cowan model with smooth sigmoidal nonlinearities, cf. [73]. □

[7] With a careful resolution of the ties, this still holds in the measure-zero event that multiple $\mathbf{q}_{\boldsymbol{\sigma}}$ are equal and belong to the boundary between orthants.



(a)



(b)

Fig. 4: Network trajectories for the excitatory-inhibitory network of Example IV.11. **(a)** When $\mathbf{W}_{EI} = [0.9, -2; 5, -1.5]$, $\mathbf{d}_{EI} = [1; 1]$, the network has a unique GES equilibrium. **(b)** However, for $\mathbf{W}_{EI} = [1.1, -2; 5, -1.5]$, $\mathbf{d}_{EI} = [-0.01; -1]$, the network exhibits bi-stable behavior. $\mathbf{m} = \infty\mathbf{1}_2$ in both cases. The trajectory colors correspond to the equilibria to which they converge. Although $\alpha n w_{ee} > 1$, the network is GES for most values of $\mathbf{d}_{EI}$, so we used Proposition IV.5 for finding a $\mathbf{d}_{EI}$ that leads to multi-stability.

D. Boundedness of Solutions

Here we study the boundedness of solutions under the dynamics (3). While our discussion so far has been about (4) (with constant $\mathbf{d}$), we switch for the remainder of this section to (3) for the sake of generality, as the same results are applicable without major modifications. Since network trajectories are trivially bounded if $\mathbf{m} < \infty\mathbf{1}_n$, we limit our discussion here to the unbounded case. Note that in reality, the firing rate of any neuron is bounded by a maximum rate dictated by its refractory period (the minimum inter-spike duration). Unboundedness of solutions in this model corresponds in practice to the so-called "run-away" excitations where the firing of neurons grow beyond sustainable rates for prolonged periods of time, which is neither desirable nor safe [74].

Since GES implies boundedness of solutions, any condition that is sufficient for GES is also sufficient for boundedness. However, boundedness of solutions can be guaranteed under less restrictive conditions. The next result shows that inhibition, overall, preserves boundedness.

Lemma IV.12. ***(Inhibition preserves boundedness).*** *Let $t \mapsto \mathbf{x}(t)$ be the solution of* (3) *starting from initial state $\mathbf{x}(0) = \mathbf{x}_0$. Consider the system*

$$\tau\dot{\bar{\mathbf{x}}}(t) = -\bar{\mathbf{x}}(t) + [[\mathbf{W}]_{\mathbf{0}}^{\infty}\bar{\mathbf{x}}(t) + \mathbf{d}(t)]_{\mathbf{0}}^{\infty}, \quad \bar{\mathbf{x}}(0) = \mathbf{x}_0. \tag{16}$$

Then, $\mathbf{x}(t) \leq \bar{\mathbf{x}}(t)$ for all $t \geq 0$.

Proof: Since $\mathbf{x}(t) \geq \mathbf{0}$ for all t, we can write (3) as

$$\tau\dot{\mathbf{x}}(t) = -\mathbf{x}(t) + [[\mathbf{W}]_{\mathbf{0}}^{\infty}\mathbf{x}(t) + \mathbf{d}(t) + \boldsymbol{\delta}(t)]_{\mathbf{0}}^{\infty}, \tag{17}$$

where $\boldsymbol{\delta}(t) \triangleq (\mathbf{W} - [\mathbf{W}]_{\mathbf{0}}^{\infty})\mathbf{x}(t) \leq \mathbf{0}$. Since the vector field $(\mathbf{x}, t) \mapsto -\mathbf{x} + [[\mathbf{W}]_{\mathbf{0}}^{\infty}\mathbf{x} + \mathbf{d}(t)]_{\mathbf{0}}^{\infty}$ is quasi-monotone nondecreasing[8], the result follows by using the monotonicity of $[\,\cdot\,]_{\mathbf{0}}^{\infty}$ and applying the vector-valued extension of the Comparison Principle given in [75, Lemma 3.4] to (16) and (17). ■

While the result about preservation of boundedness under inhibition in Lemma IV.12 is intuitive, one must interpret it carefully: it is *not* in general true that adding inhibition to any dynamics (3) can only decrease $\mathbf{x}(t)$. This is only true if the network vector field is quasi-monotone nondecreasing, as is the case with the excitatory-only dynamics (16). Intuitively, this is because, if the network has inhibitory nodes, adding inhibition to their input can in turn "disinhibit" and increase the activity of the rest of the network. The next result identifies a condition on the excitatory part of the dynamics to determine if trajectories are bounded.

Theorem IV.13. ***(Boundedness).*** *Consider the network dynamics* (3)*. If the corresponding excitatory-only dynamics* (16) *has bounded trajectories, the trajectories of* (3) *are also bounded by the same bound as those of* (16)*.*

The proof of this result follows from Lemma IV.12 and is therefore omitted. The following result, similar to Proposition IV.9, provides a more conservative but computationally feasible test for boundedness.

Corollary IV.14. ***(Boundedness).*** *Consider the network dynamics* (3) *and assume that $\mathbf{d}(t)$ is bounded, i.e., there exists $\bar{\mathbf{d}} \in \mathbb{R}^n_{>0}$ such that $\mathbf{d}(t) \leq \bar{\mathbf{d}}, t \geq 0$. If $\rho([\mathbf{W}]_{\mathbf{0}}^{\infty}) < 1$, then the network trajectories remain bounded for all $t \geq 0$.*

Proof: If $\mathbf{d}(t)$ is constant, the result follows from Theorem IV.13 and Proposition IV.9. If $\mathbf{d}(t)$ is not constant, the same argument proves boundedness of trajectories for

$$\tau\dot{\bar{\mathbf{x}}}(t) = -\bar{\mathbf{x}}(t) + [[\mathbf{W}]_{\mathbf{0}}^{\infty}\bar{\mathbf{x}}(t) + \bar{\mathbf{d}}]_{\mathbf{0}}^{\infty}, \quad \bar{\mathbf{x}}(0) = \mathbf{x}_0. \tag{18}$$

The result then follows from the quasi-monotonicity of $(\mathbf{x}, t) \mapsto -\mathbf{x} + [\mathbf{W}^{+}\mathbf{x} + \bar{\mathbf{d}}]_{\mathbf{0}}^{\infty}$, similar to Lemma IV.12. ■

Example IV.15. ***(Uniform excitatory-inhibitory networks, cont'd).*** Let us revisit the excitatory-inhibitory network of Example IV.3, here with $\mathbf{m} = \infty\mathbf{1}_2$. Clearly, the excitatory-only dynamics have bounded trajectories if and only if

$$\rho([\mathbf{W}_{EI}]_{\mathbf{0}}^{\infty}) < 1 \Leftrightarrow \alpha n w_{ee} < 1, \tag{19}$$

which is the same condition as in (15) and (8). However, an exhaustive inspection of the switching regions $\{\Omega_{\boldsymbol{\sigma}}\}_{\boldsymbol{\sigma}}$ and the

[8] A vector field $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is *quasi-monotone nondecreasing* [75, Def 2.3] if for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any $i \in \{1, \ldots, n\}$,

$$\big(x_i = y_i \quad \text{and} \quad x_j \leq y_j \text{ for all } j \neq i\big) \Rightarrow f(\mathbf{x}, t) \leq f(\mathbf{y}, t).$$

eigenvalues of $\{-\mathbf{I}+\mathbf{\Sigma}^\ell\mathbf{W}\}_{\boldsymbol{\sigma}}$ reveals that the boundedness of trajectories can also be guaranteed with the weaker condition

$$-\mathbf{I}+\mathbf{W} \text{ be Hurwitz}$$
$$\Leftrightarrow \begin{cases}(1-\alpha n w_{ee})+(1-(1-\alpha)n w_{ii})>0, \text{ and}\\ (1-\alpha n w_{ee})(1-(1-\alpha)n w_{ii})>\alpha(1-\alpha)n^2 w_{ie}w_{ei},\end{cases}$$

showing that there is room for sharpening Theorem IV.13. □

Remark IV.16. ***(Comparison with the literature).*** In this section, we have provided a comprehensive list of conditions that both extend and simplify the state of the art on stability of dynamically isolated linear-threshold networks. To the best of our knowledge, all the results are novel for the bounded case with the exception of the sufficiency of $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ for the uniqueness of equilibria (one of the four implications in Theorem IV.1) shown in [26] and Theorem IV.8(i), for which we present a simpler proof. Regarding unbounded networks, for equilibria we have extended [53, Thm 5.3] (implying only the sufficiency of $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ for EUE) to show both necessity and sufficiency of $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ for both existence and uniqueness (Theorem IV.1) and provided several results that partially characterize equilibria when $\mathbf{I}-\mathbf{W}\notin\mathcal{P}$. On exponential stability of the unbounded case, Theorem IV.8 gives a simpler proof than [66, Thm 1] for the sufficiency of $\mathbf{W}\in\mathcal{L}$ and a novel proof for the necessity of $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$. Finally, our result on boundedness of trajectories (Theorem IV.13) extends Corollary IV.14 (also available in [34, Thm 1]) to a much wider class of networks by exploiting the quasi-monotonicity of excitatory-only dynamics. □

V. Selective Inhibition in Bilayer Networks

Here, we study selective inhibition in bilayer networks as a building block towards the understanding of hierarchical selective recruitment in multilayer networks. With respect to the model described in Section III, we consider two layers ($N=2$), where the dynamics of the lower layer $\mathcal{N}_2$ is described by (3) and the dynamics of the upper layer $\mathcal{N}_1$ is arbitrary (this is for generality, we consider linear-threshold dynamics for $\mathcal{N}_1$ too in a multilayer framework in our accompanying work [19]). Our goal is to study the selective inhibition of $\mathcal{N}_2^{\mathtt{o}}$ via the input that it receives from $\mathcal{N}_1$.

As pointed out in Section III, when a group of neurons are inhibited, their activity is substantially decreased, ideally such that their net input (their respective component of $\mathbf{W}\mathbf{x}(t)+\mathbf{d}(t)$) becomes negative and their firing rate decays exponentially to zero. Therefore, the problem of selective inhibition is equivalent to the exponential stabilization of the nodes $\mathcal{N}_2^{\mathtt{o}}$ to the origin. To this end, we decompose $\mathbf{d}(t)$ as

$$\mathbf{d}(t)=\mathbf{B}\mathbf{u}(t)+\tilde{\mathbf{d}}. \tag{20}$$

The role of $\mathbf{u}(t)\in\mathbb{R}^p_{\geq 0}$ is to stabilize $\mathcal{N}_2^{\mathtt{o}}$ to the origin while the role of $\tilde{\mathbf{d}}\in\mathbb{R}^n$ is to shape the activity of $\mathcal{N}_2^{\mathtt{1}}$ by determining its equilibrium. For the purpose of this section, we assume $\tilde{\mathbf{d}}$ is given and constant.

Let $r\leq n$ be the size of $\mathcal{N}_2^{\mathtt{o}}$. We partition $\mathbf{x}$, $\mathbf{W}$, $\mathbf{B}$, and $\tilde{\mathbf{d}}$ accordingly,

$$\mathbf{x}=\begin{bmatrix}\mathbf{x}^{\mathtt{o}}\\ \mathbf{x}^{\mathtt{1}}\end{bmatrix},\ \mathbf{W}=\begin{bmatrix}\mathbf{W}^{\mathtt{oo}} & \mathbf{W}^{\mathtt{o1}}\\ \mathbf{W}^{\mathtt{1o}} & \mathbf{W}^{\mathtt{11}}\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}\mathbf{B}^{\mathtt{o}}\\ \mathbf{0}\end{bmatrix},\ \tilde{\mathbf{d}}=\begin{bmatrix}\mathbf{0}\\ \tilde{\mathbf{d}}^{\mathtt{1}}\end{bmatrix}, \tag{21}$$

where $\mathbf{W}^{\mathtt{oo}}\in\mathbb{R}^{r\times r}, \mathbf{B}^{\mathtt{o}}\in\mathbb{R}^{r\times p}_{\leq 0}$ is nonpositive to deliver inhibition, and $\tilde{\mathbf{d}}^{\mathtt{1}}\in\mathbb{R}^{n-r}$. $\mathbf{m}$ is decomposed similarly. The first r rows of $\mathbf{B}$ are nonzero to allow for the inhibition of $\mathcal{N}_2^{\mathtt{o}}$ while the remaining $n-r$ rows are zero to make this inhibition selective to $\mathcal{N}_2^{\mathtt{o}}$.[9] The sparsity of the entries of $\tilde{\mathbf{d}}$ is opposite to the rows of $\mathbf{B}$ due to the complementary roles of $\mathbf{B}\mathbf{u}(t)$ and $\tilde{\mathbf{d}}$.

The mechanisms of inhibition in the brain are broadly divided [76] into two categories, feedforward and feedback, based on how the signal $\mathbf{u}(t)$ is determined. In the following, we separately study each scenario, analyzing the interplay between the corresponding mechanism and network structure. We will later combine both mechanisms when we discuss the complete HSR framework in [19], as natural selective inhibition is not purely feedback or feedforward.

A. Feedforward Selective Inhibition

Feedforward inhibition [76] refers to the scenario where $\mathcal{N}_1$ provides an inhibition based on its own "desired" activity/inactivity pattern for $\mathcal{N}_2$ and irrespective of the current state of $\mathcal{N}_2$. This is indeed possible if the inhibition is sufficiently strong, as excessive inhibition has no effect on nodal dynamics due to the (negative) thresholding in $[\,\cdot\,]_0^m$. However, this independence from the activity level of $\mathcal{N}_2$ requires some form of guaranteed boundedness, as defined next.

Definition V.1. ***(Monotone boundedness).*** *The dynamics* (3) *is* monotonically bounded *if for any $\bar{\mathbf{d}}\in\mathbb{R}^n$ there exists $\boldsymbol{\nu}(\bar{\mathbf{d}})\in\mathbb{R}^n_{\geq 0}$ such that $\mathbf{x}(t)\leq\boldsymbol{\nu}(\bar{\mathbf{d}}), t\geq 0$ for any $\mathbf{d}(t)\leq\bar{\mathbf{d}}, t\geq 0$.*

From Lemma IV.12 and Proposition IV.9, (3) is monotonically bounded if $\rho([\mathbf{W}]_{\mathbf{0}}^{\infty})<1$ and the initial condition $\mathbf{x}_0$ is restricted to a bounded domain. Also in reality, the state of any biological neuronal network is uniformly bounded due to the refractory period of its neurons, implying monotone boundedness. We next show that the GES of $\mathcal{N}_2^{\mathtt{1}}$ is both necessary and sufficient for feedforward selective inhibition.

Theorem V.2. ***(Feedforward selective inhibition).*** *Consider the dynamics* (3)*, where the external input is given by* (20)-(21) *with a constant feedforward control*

$$\mathbf{u}(t)\equiv\mathbf{u}\geq\mathbf{0}.$$

Assume that (3) *is monotonically bounded and*

$$\textit{range}([\mathbf{W}^{\mathtt{oo}}\ \mathbf{W}^{\mathtt{o1}}])\subseteq\textit{range}(\mathbf{B}^{\mathtt{o}}). \tag{22}$$

Then, for any $\tilde{\mathbf{d}}^{\mathtt{1}}\in\mathbb{R}^{n-r}$, there exists $\bar{\mathbf{u}}\in\mathbb{R}^p_{\geq 0}$ such that for all $\mathbf{u}\geq\bar{\mathbf{u}}$, $\mathcal{N}_2$ is GES relative to a unique equilibrium of the form $\mathbf{x}_*=[\mathbf{0}_r^T\ (\mathbf{x}_*^{\mathtt{1}})^T]^T$ if and only if $\mathbf{W}^{\mathtt{11}}$ is such that the internal $\mathcal{N}_2^{\mathtt{1}}$ dynamics

$$\tau\dot{\mathbf{x}}^{\mathtt{1}}=-\mathbf{x}^{\mathtt{1}}+[\mathbf{W}^{\mathtt{11}}\mathbf{x}^{\mathtt{1}}+\tilde{\mathbf{d}}^{\mathtt{1}}]_{\mathbf{0}}^{\mathbf{m}^{\mathtt{1}}}, \tag{23}$$

is GES relative to a unique equilibrium.

Proof: ($\Leftarrow$) Define $\mathbf{u}_s$ to be a solution of

$$\mathbf{B}^{\mathtt{o}}\mathbf{u}_s=-[[\mathbf{W}^{\mathtt{oo}}\ \mathbf{W}^{\mathtt{o1}}]]_{\mathbf{0}}^{\infty}\boldsymbol{\nu}(\tilde{\mathbf{d}}). \tag{24}$$

[9] This sparsity pattern can always be achieved by (re-)labeling the r directly controlled nodes as $1,\dots,r$, so that the $n-r$ last entries of $\mathbf{B}$ are 0.

This solution exists by assumption (22). Let $\bar{\mathbf{u}} = [\mathbf{u}_s]_{\mathbf{0}}^{\infty}$ and note that $\mathbf{B}^{\circ}\mathbf{u} \leq \mathbf{B}^{\circ}\bar{\mathbf{u}} \leq \mathbf{B}^{\circ}\mathbf{u}_s$. By construction, (3), (20), (21), (24) simplify to

$$\begin{aligned}\tau\dot{\mathbf{x}}^{\circ} &= -\mathbf{x}^{\circ},\\ \tau\dot{\mathbf{x}}^{1} &= -\mathbf{x}^{1} + [\mathbf{W}^{1\circ}\mathbf{x}^{\circ} + \mathbf{W}^{11}\mathbf{x}^{1} + \tilde{\mathbf{d}}^{1}]_{\mathbf{0}}^{\mathbf{m}^1},\end{aligned} \tag{25}$$

whose GES follows from Lemma A.1.

($\Rightarrow$) By monotone boundedness and nonpositivity of $\mathbf{B}^{\circ}$, $\mathbf{x}(t) \leq \nu(\tilde{\mathbf{d}})$ for all $t \geq 0$ and any $\mathbf{u} \geq \bar{\mathbf{u}}$. Let $\mathbf{u} = \bar{\mathbf{u}} + [\mathbf{u}_s]_{\mathbf{0}}^{\infty}$ where $\mathbf{u}_s$ is a solution to (24). Similar to above, this simplifies (3), (20), (21), (24) to (25), which is GES by assumption. However, for any initial condition of the form $\mathbf{x}(0) = [\mathbf{0}_r^T \ \mathbf{x}^{1}(0)^T]^T$, the trajectories of (25) are the same as (23), and the result follows. ∎

The next section shows that the condition (22) on the ability to influence the dynamics of the task-irrelevant nodes through control also plays a key role in feedback selective inhibition. We defer the discussion about the interpretation of this condition to Section V-C below.

B. Feedback Selective Inhibition

The core idea of feedback inhibition [76], as found throughout the brain, is the dependence of the amount of inhibition on the activity level of the nodes that are to be inhibited. This dependence is in particular relevant to GDSA, as the stronger and more salient a source of distraction, the harder one must try to suppress its effects on perception. The next result provides a novel characterization of several equivalences between the dynamical properties of $\mathcal{N}_2$ under linear full-state feedback inhibition and those of $\mathcal{N}_2^{1}$.

Theorem V.3. ***(Feedback selective inhibition).*** *Consider the dynamics* (3)*, where the external input is given by* (20)-(21) *with a linear state feedback* $\mathbf{u}$

$$\mathbf{u}(t) = \mathbf{K}\mathbf{x}(t), \tag{26}$$

and $\mathbf{K} \in \mathbb{R}^{p\times n}$ *is a constant control gain. Assume that* (22) *holds. Then, there almost always exists* $\mathbf{K} \in \mathbb{R}^{p\times n}$ *such that*

(i) $\mathbf{I} - (\mathbf{W} + \mathbf{B}\mathbf{K}) \in \mathcal{P}$ *if and only if* $\mathbf{I} - \mathbf{W}^{11} \in \mathcal{P}$*;*
(ii) $-\mathbf{I} + (\mathbf{W} + \mathbf{B}\mathbf{K}) \in \mathcal{H}$ *if and only if* $-\mathbf{I} + \mathbf{W}^{11} \in \mathcal{H}$*;*
(iii) $\mathbf{W} + \mathbf{B}\mathbf{K} \in \mathcal{L}$ *if and only if* $\mathbf{W}^{11} \in \mathcal{L}$*;*
(iv) $\rho(|\mathbf{W} + \mathbf{B}\mathbf{K}|) < 1$ *if and only if* $\rho(|\mathbf{W}^{11}|) < 1$*;*
(v) $\|\mathbf{W} + \mathbf{B}\mathbf{K}\| < 1$ *if and only if* $\|[\mathbf{W}^{1\circ}\ \mathbf{W}^{11}]\| < 1$.

Proof: (i) $\Rightarrow$) For any $\mathbf{K} = [\mathbf{K}^{\circ}\ \mathbf{K}^{1}] \in \mathbb{R}^{p\times n}$,

$$\mathbf{W} + \mathbf{B}\mathbf{K} = \begin{bmatrix} \mathbf{W}^{\circ\circ} + \mathbf{B}^{\circ}\mathbf{K}^{\circ} & \mathbf{W}^{\circ 1} + \mathbf{B}^{\circ}\mathbf{K}^{1} \\ \mathbf{W}^{1\circ} & \mathbf{W}^{11} \end{bmatrix}. \tag{27}$$

Thus, since any principal submatrix of a P-matrix is a P-matrix, $\mathbf{I} - \mathbf{W}^{11} \in \mathcal{P}$.
$\Leftarrow$) By (22) there exists $\bar{\mathbf{K}} \in \mathbb{R}^{p\times n}$ such that

$$-\begin{bmatrix}\mathbf{W}^{\circ\circ} & \mathbf{W}^{\circ 1}\end{bmatrix} = \mathbf{B}^{\circ}\bar{\mathbf{K}}. \tag{28}$$

Using the fact that the determinant of any block-triangular matrix is the product of the determinants of the blocks on its diagonal [45, Prop 2.8.1], it follows that $\mathbf{I} - (\mathbf{W} + \mathbf{B}\bar{\mathbf{K}}) \in \mathcal{P}$.

(ii) $\Rightarrow$) This follows from (27) and the fact that a principal submatrix of a totally-Hurwitz matrix is totally-Hurwitz.

$\Leftarrow$) Using the matrix $\bar{\mathbf{K}}$ in (28), the result follows from the fact that the eigenvalues of a block-triangular matrix are the eigenvalues of its diagonal blocks.

(iii) $\Rightarrow$) Let $\mathbf{P} = \mathbf{P}^T > \mathbf{0}$ be such that

$$(-\mathbf{I} + (\mathbf{W}+\mathbf{B}\mathbf{K})^T\mathbf{\Sigma}^{\ell})\mathbf{P} + \mathbf{P}(-\mathbf{I} + \mathbf{\Sigma}^{\ell}(\mathbf{W}+\mathbf{B}\mathbf{K})) < \mathbf{0} \tag{29}$$

for all $\boldsymbol{\sigma} \in \{0,\ell\}^n$. Consider, in particular, $\boldsymbol{\sigma} = [\mathbf{0}_r^T \ (\boldsymbol{\sigma}^{1})^T]^T$ where $\boldsymbol{\sigma}^{1} \in \{0,\ell\}^{n-r}$ is arbitrary. Let $\mathbf{\Sigma}^{\ell 1} \in \mathbb{R}^{(n-r)\times(n-r)}$ be the bottom-right block of Σ^{ℓ} and partition $\mathbf{P}$ in 2-by-2 block form similarly to $\mathbf{W}$. Since

$$\begin{aligned}&(-\mathbf{I} + (\mathbf{W} + \mathbf{B}\mathbf{K})^T\mathbf{\Sigma}^{\ell})\mathbf{P} + \mathbf{P}(-\mathbf{I} + \mathbf{\Sigma}^{\ell}(\mathbf{W} + \mathbf{B}\mathbf{K}))\\ &\quad= \begin{bmatrix} \star & \star \\ \star & (-\mathbf{I} + \mathbf{\Sigma}^{\ell 1}\mathbf{W}^{11})^T\mathbf{P}^{11} + \mathbf{P}^{11}(-\mathbf{I} + \mathbf{\Sigma}^{\ell 1}\mathbf{W}^{11}) \end{bmatrix},\end{aligned}$$

and any principal submatrix of a negative definite matrix is negative definite, we deduce $\mathbf{W}^{11} \in \mathcal{L}$.
$\Leftarrow$) Let $\mathbf{P}^{11} \in \mathbb{R}^{(n-r)\times(n-r)}$ be such that

$$(-\mathbf{I} + (\mathbf{W}^{11})^T\mathbf{\Sigma}^{\ell 1})\mathbf{P}^{11} + \mathbf{P}^{11}(-\mathbf{I} + \mathbf{\Sigma}^{\ell 1}\mathbf{W}^{11}) < \mathbf{0},$$

for all $\boldsymbol{\sigma}^{1} \in \{0,1\}^{n-r}$ and $\bar{\mathbf{K}}$ be as in (28). For any $\boldsymbol{\sigma} = [(\boldsymbol{\sigma}^{\circ})^T \ (\boldsymbol{\sigma}^{1})^T]^T$, (28) gives

$$-\mathbf{I} + \mathbf{\Sigma}^{\ell}(\mathbf{W} + \mathbf{B}\bar{\mathbf{K}}) = \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \star & -\mathbf{I} + \mathbf{\Sigma}^{\ell 1}\mathbf{W}^{11} \end{bmatrix}.$$

Thus, the dynamics $\tau\dot{\mathbf{x}} = \big(-\mathbf{I} + \mathbf{\Sigma}^{\ell}(\mathbf{W} + \mathbf{B}\bar{\mathbf{K}})\big)\mathbf{x}$ is a cascade of $\tau\dot{\mathbf{x}}^{\circ} = -\mathbf{x}^{\circ}$ and $\tau\dot{\mathbf{x}}^{1} = (-\mathbf{I} + \mathbf{\Sigma}^{\ell 1}\mathbf{W}^{11})\mathbf{x}^{1} + \star\cdot\mathbf{x}^{\circ}$, where the latter has the ISS[10]-Lyapunov function $V^{1}(\mathbf{x}^{1}) = (\mathbf{x}^{1})^T\mathbf{P}^{11}\mathbf{x}^{1}$. Using [77, Thm 3], (29) holds for $\mathbf{K} = \bar{\mathbf{K}}$, $\mathbf{P} = \mathrm{diag}(\mathbf{I}, \mathbf{P}^{11})$, and any $\boldsymbol{\sigma} \in \{0,\ell\}^n$, giving $\mathbf{W} + \mathbf{B}\bar{\mathbf{K}} \in \mathcal{L}$.

(iv) $\Rightarrow$) This follows from (27) and [45, Fact 4.11.19].
$\Leftarrow$) Consider the matrix $\bar{\mathbf{K}}$ in (28). The result then follows from the fact that the eigenvalues of a block-triangular matrix are the eigenvalues of its diagonal blocks.

(v) $\Rightarrow$) Note that for any $\mathbf{K} \in \mathbb{R}^{p\times n}$,

$$\begin{aligned}\|\mathbf{W} + \mathbf{B}\mathbf{K}\|^2 &= \rho\left(\begin{bmatrix} \star & \star \\ \star & \mathbf{W}^{1\circ}(\mathbf{W}^{1\circ})^T + \mathbf{W}^{11}(\mathbf{W}^{11})^T \end{bmatrix}\right)\\ &\geq \rho(\mathbf{W}^{1\circ}(\mathbf{W}^{1\circ})^T + \mathbf{W}^{11}(\mathbf{W}^{11})^T) = \left\|\begin{bmatrix}\mathbf{W}^{1\circ} & \mathbf{W}^{11}\end{bmatrix}\right\|^2,\end{aligned}$$

where the inequality follows from the well-known interlacing property of eigenvalues of principal submatrices (cf. [78]).
$\Leftarrow$) Consider the matrix $\bar{\mathbf{K}}$ in (28) and note that

$$\begin{aligned}\|\mathbf{W} + \mathbf{B}\bar{\mathbf{K}}\|^2 &= \rho\left(\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}^{1\circ}(\mathbf{W}^{1\circ})^T + \mathbf{W}^{11}(\mathbf{W}^{11})^T \end{bmatrix}\right)\\ &= \rho(\mathbf{W}^{1\circ}(\mathbf{W}^{1\circ})^T + \mathbf{W}^{11}(\mathbf{W}^{11})^T) = \left\|\begin{bmatrix}\mathbf{W}^{1\circ} & \mathbf{W}^{11}\end{bmatrix}\right\|^2 < 1,\end{aligned}$$

completing the proof. ∎

Remark V.4. ***(Feedback inhibition with nonnegative*** $\mathbf{u}(t)$***).*** Even though Theorem V.3 is motivated by feedback inhibition in the brain, the result illustrates some fundamental properties of linear-threshold dynamics and the corresponding matrix classes that is of independent interest, which motivates the generality of its formulation. The particular application to brain networks requires nonnegative inputs, which we discuss next. The core principle of Theorem V.3 is the cancellation of

[10] Input-to-state stability

local input $[\mathbf{W}^{\circ\circ}\ \mathbf{W}^{\circ 1}]\mathbf{x}$ to $\mathcal{N}_2^\circ$ with the top-down feedback input $\mathbf{B}^\circ\bar{\mathbf{K}}\mathbf{x}$, simplifying the dynamics of $\mathcal{N}_2^\circ$ to $\tau\dot{\mathbf{x}}^\circ = -\mathbf{x}^\circ$ that guarantee its inhibition. However, the resulting input signal $\mathbf{u} = \bar{\mathbf{K}}\mathbf{x}$ (being the firing rate of some neuronal population in $\mathcal{N}_1$) may not remain nonnegative at all times. This can be easily addressed as follows. Similar to the proof of Theorem V.2, we let

$$\mathbf{u}(t) = [\bar{\mathbf{K}}\mathbf{x}(t)]_{\mathbf{0}}^{\infty}.$$

This makes $\mathbf{u}(t)$ nonnegative without affecting the selective inhibition of $\mathcal{N}_2^\circ$ in (3) as $\mathbf{B}^\circ \leq \mathbf{0}$. In principle, a similar concern can exist as to whether $\bar{\mathbf{K}}\mathbf{x}$ becomes larger than the maximum firing rate of the corresponding populations in $\mathcal{N}_1$. However, this only relates to the magnitude of the entries in $\mathbf{B}^\circ$ (via (28), as opposed to the sign of $\bar{\mathbf{K}}\mathbf{x}$, which relates to the sign of the entries in $\mathbf{B}^\circ$), which can always be increased via synaptic long term potentiation (LTP) [79], in turn decreasing the magnitude of the entries in $\bar{\mathbf{K}}$. □

Remark V.5. ***(State vs. output feedback).*** The assumption of state feedback is a simplifying one and its generalization merits further research. However, we note that $\mathbf{x}^\circ$ is most likely always available for feedback (as feedback inhibition is highly reciprocal at the neuronal level [80] and even more so at the population level) while the availability of $\mathbf{x}^1$ for feedback remains case-specific. If the latter is not available, one of two scenarios may happen: either the local interaction of $\mathbf{x}^\circ$ and $\mathbf{x}^1$ is competitive and $\mathbf{W}^{\circ 1} \leq \mathbf{0}$ (which is not unlikely due to the prevalence of lateral inhibition in the cortex [81]), in which case the feedback of $\mathbf{x}^1$ is not even needed (similar to Remark V.4) or, at worst, the unobserved $\mathbf{x}^1$ actively excites $\mathbf{x}^\circ$, in which case a combination of feedback and feedforward inhibition can be used, similar to our full model in Part II [19, Thm 4.3]. □

C. Network Size, Weight Distribution, and Stabilization

Underlying the discussion above is the requirement that $\mathcal{N}_2$ can be asymptotically stabilized towards an equilibrium which has some components equal to zero and the remaining components determined by $\tilde{\mathbf{d}}$. Here, it is important to distinguish between the stability of $\mathcal{N}_2$ in the absence and presence of selective inhibition. In reality, the large size of biological neuronal networks often leads to highly unstable dynamics if all the nodes in a layer, say $\mathcal{N}_2$, are active. Therefore, the selective inhibition of $\mathcal{N}_2^\circ$ is not only responsible for the suppression of the task-irrelevant activity of $\mathcal{N}_2^\circ$, but also for the overall stabilization of $\mathcal{N}_2$ that allows for top-down recruitment of $\mathcal{N}_2^1$. This poses limitations on the size and structure of the subnetworks $\mathcal{N}_2^\circ$ and $\mathcal{N}_2^1$. It is in this context that one can analyze the condition (22) assumed in both Theorems V.2 and V.3. This condition requires, essentially, that there are sufficiently many "independent" external controls $\mathbf{u}$ to enforce inhibition on $\mathcal{N}_2^\circ$. The following result formalizes this statement.

Lemma V.6. ***(Equivalent characterization of*** (22)***).*** *Let the matrices $\mathbf{W}^\circ$ and $\mathbf{B}^\circ$ have dimensions $r \times n$ and $r \times p$, respectively. Then, range$(\mathbf{W}^\circ) \subseteq$ range$(\mathbf{B}^\circ)$ for almost all $(\mathbf{W}^\circ, \mathbf{B}^\circ) \in \mathbb{R}^{r\times n} \times \mathbb{R}^{r\times m}$ if and only if $p \geq r$.*

Proof: $\Rightarrow$) Assume, by contradiction, that $p < r$, so range$(\mathbf{B}^\circ) \subsetneq \mathbb{R}^r$ for any $\mathbf{B}^\circ$. Let $\mathbf{Q} = \mathbf{Q}(\mathbf{B}^\circ)$ be a matrix whose columns form a basis for range$(\mathbf{B}^\circ)^\perp$. Then, range$(\mathbf{W}^\circ) \subseteq$ range$(\mathbf{B}^\circ)$ if and only if $\mathbf{Q}(\mathbf{B}^\circ)^T\mathbf{W}^\circ = \mathbf{0}$. By Fubini's theorem [82, Ch. 20],

$$\begin{aligned}
&\int_{\mathbb{R}^{r\times n}\times\mathbb{R}^{r\times p}} \mathbb{I}_{\{\mathbf{Q}(\mathbf{B}^\circ)^T\mathbf{W}^\circ=\mathbf{0}\}}(\mathbf{W}^\circ,\mathbf{B}^\circ)d(\mathbf{W}^\circ,\mathbf{B}^\circ)\\
&= \int_{\mathbb{R}^{r\times p}} d\mathbf{B}^\circ \int_{\mathbb{R}^{r\times n}} \mathbb{I}_{\{\mathbf{Q}(\mathbf{B}^\circ)^T\mathbf{W}^\circ=\mathbf{0}\}}(\mathbf{W}^\circ,\mathbf{B}^\circ)d\mathbf{W}^\circ\\
&= \int_{\mathbb{R}^{r\times p}} 0\ d\mathbf{B}^\circ = 0,
\end{aligned}$$

where $\mathbb{I}$ denotes the indicator function. This contradiction proves $p \geq r$.

$\Leftarrow$) Let $\mathbf{B}^\circ = [\mathbf{B}_1^\circ\ \mathbf{B}_2^\circ]$ where $\mathbf{B}_1^\circ \in \mathbb{R}^{r\times r}$. It is straightforward to show that

$$\{(\mathbf{W}^\circ, \mathbf{B}^\circ) \mid \text{range}(\mathbf{W}^\circ) \nsubseteq \text{range}(\mathbf{B}^\circ)\} \subseteq \mathbb{R}^{r\times n} \times A,$$

where $A = \{\mathbf{B}^\circ \mid \det(\mathbf{B}_1^\circ) = 0\}$. Since A has measure zero, the result follows from a similar argument as above invoking Fubini's theorem. ■

Based on intuitions from linear systems theory, it may be tempting to seek a relaxation of (22) for the case where $p < r$. This is due to the fact that for a *linear* system $\tau\dot{\mathbf{x}} = \mathbf{W}\mathbf{x}+\mathbf{B}\mathbf{u}$, it is known [83, eq (4.5) and Thm 3.5] that the set of all reachable states from the origin is given by

$$\text{range}\left(\begin{bmatrix}\mathbf{B}^\circ & \mathbf{W}^{\circ\circ}\mathbf{B}^\circ & \cdots & (\mathbf{W}^{n-1})^{\circ\circ}\mathbf{B}^\circ\\ \mathbf{0} & \mathbf{W}^{1\circ}\mathbf{B}^\circ & \cdots & (\mathbf{W}^{n-1})^{1\circ}\mathbf{B}^\circ\end{bmatrix}\right),$$

which is usually much larger than range$(\mathbf{B})$. Therefore, it is reasonable to expect that (22) could be relaxed to

$$\begin{aligned}
&\text{range}([\mathbf{W}^{\circ\circ}\ \mathbf{W}^{\circ 1}])\\
&\qquad \subseteq \text{range}([\mathbf{B}^\circ\ \mathbf{W}^{\circ\circ}\mathbf{B}^\circ\ \cdots\ (\mathbf{W}^{n-1})^{\circ\circ}\mathbf{B}^\circ]). \quad (30)
\end{aligned}$$

However, it turns out that this relaxation is not possible, the reason being the (apparently simple, yet intricate) nonlinearity in (3). We show this by means of an example.

Example V.7. ***(Tightness of*** (22)***).*** Consider the feedback dynamics (3), (20), (26), where $n = 3$, $p = 1$, $r = 2$, and

$$\mathbf{W} = \left[\begin{array}{cc|c} 2\alpha & 0 & 0\\ 0 & 3\alpha & 0\\ \hline 0 & 0 & \alpha\end{array}\right], \quad \mathbf{B} = \left[\begin{array}{c} 1\\ 1\\ \hline 0\end{array}\right], \quad \alpha \in (0.5, 1).$$

Clearly, (22) does not hold (so Theorem V.3(iv) does not apply), but range$([\mathbf{W}^{\circ\circ}\ \mathbf{W}^{\circ 1}]) \subseteq$ range$([\mathbf{B}^\circ\ \mathbf{W}^{\circ\circ}\mathbf{B}^\circ])$. One can show that for all $\mathbf{K} \in \mathbb{R}^{1\times 3}$,

$$\rho(|\mathbf{W} + \mathbf{B}\mathbf{K}|) \geq 2\alpha > 1,$$

while $\rho(\mathbf{W}^{11}) = \alpha < 1$, verifying that (22) is necessary and cannot be relaxed to (30). □

Theorems V.2 and V.3 use completely different mechanisms for inhibition of $\mathcal{N}_2^\circ$, yet they are strikingly similar in one conclusion: that the dynamical properties achievable under selective inhibition are precisely those satisfied by $\mathcal{N}_2^1$. This has important implications for the size and structure of the part $\mathcal{N}_2^1$ that can be active at any instance of time without resulting in instability. The next remark elaborates on this implication.

Remark V.8. ***(Implications for the size and connection strength of $\mathcal{N}_2^1$).*** Existing experimental evidence suggest that the synaptic weights $\mathbf{W}$ in cortical networks are sparse, approximately follow a log-normal distribution, and have a pairwise connection probability that is independent of physical distance between neurons within short distances [84]. Based on simulations of matrices with such statistics, Figure 5(a, b) show how quickly the network (representing $\mathcal{N}_2$ here) moves towards instability when its size grows. On the other hand, it is well-known that increasing n (and thus the number of synaptic weights) increases network expressivity (i.e., capacity to reproduce complex trajectories). While determining the optimal size of a network that leads to the best tradeoff between stability and expressivity is beyond the scope of this work, our results suggest a critical role for selective inhibition in keeping only a limited number of nodes in $\mathcal{N}_2$ active at any given time while inhibiting others. In other words, while the overall size of subnetworks in a brain network (corresponding to, e.g., the number of neuronal populations with distinct preferred stimuli in a brain region) is inevitably large, selective inhibition offers a plausible explanation for the mechanism by which the brain keeps the number of *active populations* bounded ($O(1)$) at any given time.

Similarly, Figure 5(c, d) show the transition of networks towards instability as their synaptic connections become stronger. While excitatory synapses, as expected, have a larger impact on stability, the same trend is also observed while varying inhibitory synaptic strengths. Interestingly, several works in the neuroscience literature have shown that neuronal networks maintain stability by re-scaling their synaptic weights that change during learning, a process commonly referred to as *homeostatic synaptic plasticity* [85]. Our results thus open the way to provide rigorous and quantifiable measures of the optimal size and weight distribution of subnetworks that may be active at any given time and the homeostatic mechanisms that maintain any desired level of stability and expressivity.□

VI. Conclusions

We adopt a control-theoretic framework, termed hierarchical selective recruitment (HSR), as a mechanism to explain goal-driven selective attention. Motivated by the organization of the brain, HSR employs a hierarchical model which consists of an arbitrary number of neuronal subnetworks that operate at different layers of a hierarchy. While HSR is not confined to any family of models, we here use the well-studied linear-threshold rate models to describe the dynamics at each layer of the hierarchy. We provide a thorough analysis of the internal dynamics of each layer. Leveraging the switched-affine nature of linear-threshold dynamics, we derive several necessary and sufficient conditions for the existence and uniqueness of equilibria (corresponding to P-matrices), local and global asymptotic stability (corresponding to totally-Hurwitz matrices), and boundedness of trajectories (corresponding to stability of excitatory-only dynamics). These results set the basis for analyzing the problem of selective inhibition. We show that using either feedforward or feedback inhibition, the dynamical properties of each layer after inhibition are precisely determined by the task-relevant part that remains active. We have also provided constructive control designs that guarantee selective inhibition under both schemes. Among the directions for future research, we highlight the study of output-feedback selective inhibition and the analysis of the conditions on (single-layer) linear-threshold networks that lead to the emergence of limit cycles and their stability.

Acknowledgments

We would like to thank Dr. Erik J. Peterson for piquing our interest with his questions on dimensionality control in brain networks and for introducing us to linear-threshold modeling in neuroscience. We are further indebted to the anonymous reviewer for suggesting the proof of the necessity of $\mathbf{I}-\mathbf{W} \in \mathcal{P}$ for the EUE of unbounded networks. This work was supported by NSF Award CMMI-1826065 (EN and JC) and ARO Award W911NF-18-1-0213 (JC).

References

[1] E. Nozari and J. Cortés, "Stability analysis of complex networks with linear-threshold rate dynamics," in *American Control Conference*, Milwaukee, WI, May 2018, pp. 191–196.
[2] J. Sully, "The psycho-physical process in attention," *Brain*, vol. 13, no. 2, pp. 145–164, 1890.
[3] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *The Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975–979, 1953.
[4] R. Desimone and J. Duncan, "Neural mechanisms of selective visual attention," *Annual Review of Neuroscience*, vol. 18, no. 1, pp. 193–222, 1995.
[5] L. Itti and C. Koch, "Computational modelling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 3, p. 194, 2001.
[6] N. Lavie, A. Hirst, J. W. DeFockert, and E. Viding, "Load theory of selective attention and cognitive control." *Journal of Experimental Psychology: General*, vol. 133, no. 3, p. 339, 2004.
[7] A. Gazzaley and A. C. Nobre, "Top-down modulation: bridging selective attention and working memory," *Trends in Cognitive Sciences*, vol. 16, no. 2, pp. 129–135, 2012.
[8] J. T. Serences and S. Kastner, "A multi-level account of selective attention," *The Oxford Handbook of Attention*, p. 76, 2014.
[9] D. E. Broadbent, Ed., *Perception and communication*. Pergamon, 1958.
[10] A. M. Treisman, "Strategies and models of selective attention." *Psychological review*, vol. 76, no. 3, p. 282, 1969.
[11] J. Moran and R. Desimone, "Selective attention gates visual processing in the extrastriate cortex," *Science*, vol. 229, no. 4715, pp. 782–784, 1985.
[12] B. C. Motter, "Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli," *Journal of Neurophysiology*, vol. 70, no. 3, pp. 909–919, 1993.
[13] S. Kastner, P. DeWeerd, R. Desimone, and L. G. Ungerleider, "Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI," *Science*, vol. 282, no. 5386, pp. 108–111, 1998.
[14] M. A. Pinsk, G. M. Doniger, and S. Kastner, "Push-pull mechanism of selective attention in human extrastriate cortex," *Journal of Neurophysiology*, vol. 92, no. 1, pp. 622–629, 2004.
[15] N. Lavie, "Distracted and confused?: Selective attention under load," *Trends in Cognitive Sciences*, vol. 9, no. 2, pp. 75–82, 2005.
[16] J. J. Foxe and A. C. Snyder, "The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention," *Frontiers in Psychology*, vol. 2, p. 154, 2011.
[17] H. Pashler, *Attention*. Psychology Press, 2016.
[18] M. Gomez-Ramirez, K. Hysaj, and E. Niebur, "Neural mechanisms of selective attention in the somatosensory system," *Journal of neurophysiology*, vol. 116, no. 3, pp. 1218–1231, 2016.
[19] E. Nozari and J. Cortés, "Hierarchical selective recruitment in linear-threshold brain networks. Part II: Inter-layer dynamics and top-down recruitment," *IEEE Transactions on Automatic Control*, vol. 66, 2021, to appear. Available at https://arxiv.org/abs/1809.02493.
[20] F. Ratliff and H. K. Hartline, *Studies on Excitation and Inhibition in the Retina*. Rockefeller University Press, 1974.
[21] K. P. Hadeler, "On the theory of lateral inhibition," *Kybernetik*, vol. 14, no. 3, pp. 161–165, 1973.

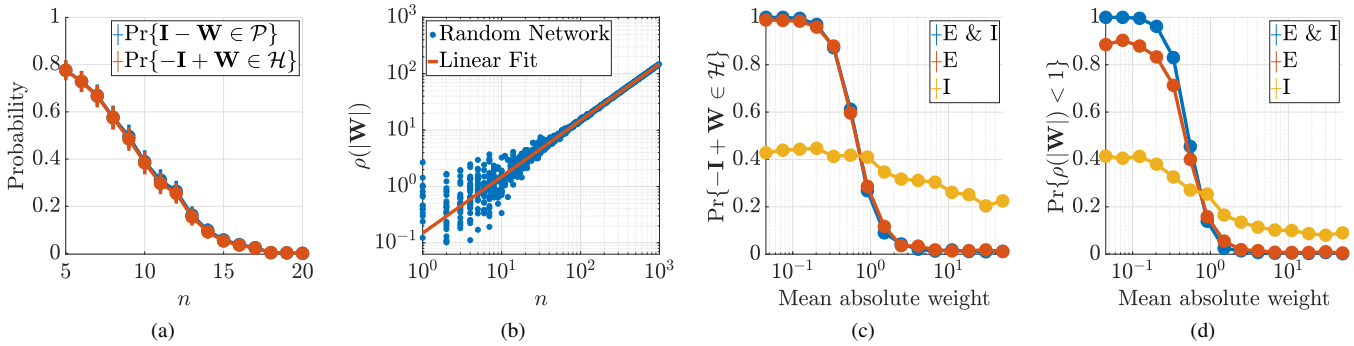


Fig. 5: The effects of network size and weight distribution on its stability. In all panels, unless otherwise stated, $\mathbf{W}$ matrices are generated randomly with log-normally distributed entries with parameters $\mu = -0.7$ and $\sigma = 0.9$ as given in [84], 20% sparsity, and 80% excitatory nodes. Probabilities are estimated empirically with 10^3 samples and error bars represent standard error of the mean (s.e.m). **(a)** Probability of $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ (as related to EUE, cf. Theorem IV.1) and $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$ (as related to asymptotic stability, cf. Theorem IV.8) showing a rapid decay with n. **(b)** For large n, while checking $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ and $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$ can become prohibitive, their sufficient condition $\rho(|\mathbf{W}|)<1$ can be checked efficiently. The line illustrates a fit of the form $\log\rho(|\mathbf{W}|)=\alpha\log n+\beta$ with $\alpha=1$ and $\beta=-1.2$, showing a linear growth of $\rho(|\mathbf{W}|)$ with n. **(c, d)** The probabilities of $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$ and $\rho(|\mathbf{W}|)<1$ while varying the μ parameter of the log-normal weight distribution over $[-3.5, 3.5]$ for both excitatory and inhibitory synapses (blue), only for excitatory ones (red), or only for inhibitory ones (yellow). The plots are very similar for $\mathbf{I}-\mathbf{W}\in\mathcal{P}$ and thus not shown. $n=10$ is fixed in these panels.

[22] R. H. R. Hahnloser, H. S. Seung, and J. J. Slotine, "Permitted and forbidden sets in symmetric threshold-linear networks," *Neural Computation*, vol. 15, no. 3, pp. 621–638, 2003.

[23] Z. Yi, L. Zhang, J. Yu, and K. K. Tan, "Permitted and forbidden sets in discrete-time linear threshold recurrent neural networks," *IEEE Transactions on Neural Networks*, vol. 20, no. 6, pp. 952–963, 2009.

[24] K. P. Hadeler and D. Kuhn, "Stationary states of the Hartline-Ratliff model," *Biological Cybernetics*, vol. 56, no. 5-6, pp. 411–417, 1987.

[25] J. Feng and K. P. Hadeler, "Qualitative behaviour of some simple networks," *Journal of Physics A: Mathematical and General*, vol. 29, no. 16, pp. 5019–5033, 1996.

[26] M. Forti and A. Tesi, "New conditions for global stability of neural networks with application to linear and quadratic programming problems," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 42, no. 7, pp. 354–366, 1995.

[27] M. Forti and P. Nistri, "Global convergence of neural networks with discontinuous neuron activations," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 11, pp. 1421–1435, 2003.

[28] I. W. Sandberg and A. N. Willson Jr., "Some theorems on properties of DC equations of nonlinear networks," *Bell System Technical Journal*, vol. 48, no. 1, pp. 1–34, 1969.

[29] H. Zhang, Z. Wang, and D. Liu, "A comprehensive review of stability analysis of continuous-time recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 7, pp. 1229–1262, 2014.

[30] Z. Yi and K. K. Tan, "Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions," *IEEE Transactions on Neural Networks*, vol. 15, no. 2, pp. 329–336, 2004.

[31] W. Zhou and J. M. Zurada, "A new stability condition for discrete time linear threshold recurrent neural networks," in *Fifth Int. Conf. on Intellig. Control and Inf. Proc.*, Aug 2014, pp. 96–99.

[32] T. Shen and I. R. Petersen, "Linear threshold discrete-time recurrent neural networks: Stability and globally attractive sets," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2650–2656, 2016.

[33] K. C. Tan, H. Tang, and W. Zhang, "Qualitative analysis for recurrent neural networks with linear threshold transfer functions," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 5, pp. 1003–1012, 2005.

[34] H. Wersing, W. Beyn, and H. Ritter, "Dynamical stability conditions for recurrent neural networks with unsaturating piecewise linear transfer functions," *Neural Computation*, vol. 13, no. 8, pp. 1811–1825, 2001.

[35] K. Morrison, A. Degeratu, V. Itskov, and C. Curto, "Diversity of emergent dynamics in competitive threshold-linear networks: a preliminary report," *arXiv preprint arXiv:1605.04463*, 2016.

[36] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: A survey of recent results," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 308–322, 2009.

[37] D. Liberzon, *Switching in Systems and Control*, ser. Systems & Control: Foundations & Applications. Birkhäuser, 2003.

[38] M. K. J. Johansson, *Piecewise Linear Control Systems: A Computational Approach*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2003.

[39] K. N. Seidl, M. V. Peelen, and S. Kastner, "Neural evidence for distracter suppression during visual search in real-world scenes," *Journal of Neuroscience*, vol. 32, no. 34, pp. 11 812–11 819, 2012.

[40] S. Kastner, P. DeWeerd, M. A. Pinsk, M. I. Elizondo, R. Desimone, and L. G. Ungerleider, "Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex," *Journal of Neurophysiology*, vol. 86, no. 3, pp. 1398–1411, 2001.

[41] G. Rees, C. D. Frith, and N. Lavie, "Modulating irrelevant motion perception by varying attentional load in an unrelated task," *Science*, vol. 278, no. 5343, pp. 1616–1619, 1997.

[42] D. H. O'Connor, M. M. Fukui, M. A. Pinsk, and S. Kastner, "Attention modulates responses in the human lateral geniculate nucleus," *Nature Neuroscience*, vol. 5, no. 11, p. 1203, 2002.

[43] M. Fiedler and V. Ptak, "On matrices with non-positive off-diagonal elements and positive principal minors," *Czechoslovak Mathematical Journal*, vol. 12, no. 3, pp. 382–400, 1962.

[44] O. Slyusareva and M. Tsatsomeros, "Mapping and preserver properties of the principal pivot transform," *Linear and Multilinear Algebra*, vol. 56, no. 3, pp. 279–292, 2008.

[45] D. S. Bernstein, *Matrix Mathematics*, 2nd ed. Princeton University Press, 2009.

[46] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, ser. Computational Neuroscience. Cambridge, MA: MIT Press, 2001.

[47] D. A. Henze, Z. Borhegyi, J. Csicsvari, M. A. Mamiya, K. D. Harris, and G. Buzsaki, "Intracellular features predicted by extracellular recordings in the hippocampus in vivo," *Journal of Neurophysiology*, vol. 84, no. 1, pp. 390–400, 2000.

[48] D. A. Henze, K. D. Harris, Z. Borhegyi, J. Csicsvari, A. Mamiya, H. Hirase, A. Sirota, and G. Buzsaki, "Simultaneous intracellular and extracellular recordings from hippocampus region ca1 of anesthetized rats," CRCNS.org, 2009. [Online]. Available: http://dx.doi.org/10.6080/K02Z13FP

[49] 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.

[50] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

[51] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, ser. Com-

prehensive Studies in Mathematics. New York: Springer, 1998, vol. 317.
[52] L. E. J. Brouwer, "Über abbildung von mannigfaltigkeiten," *Mathematische Annalen*, vol. 71, no. 1, pp. 97–115, 1911.
[53] D. Kuhn and R. Löwen, "Piecewise affine bijections of $\mathbb{R}^n$, and the equation $Sx^+ - Tx^- = y$," *Linear Algebra and its Applications*, vol. 96, pp. 109–129, 1987.
[54] K. G. Murty, "On the number of solutions to the complementarity problem and spanning properties of complementary cones," *Linear Algebra and its Applications*, vol. 5, no. 1, pp. 65–108, 1972.
[55] D. M. Stipanovic and D. D. Siljak, "Stability of polytopic systems via convex M-matrices and parameter-dependent Lyapunov functions," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 40, no. 1, pp. 589–609, 2000.
[56] G. E. Coxson, "The P-matrix problem is co-NP-complete," *Mathematical Programming*, vol. 64, no. 1, pp. 173–178, 1994.
[57] S. M. Rump, "On P-matrices," *Linear Algebra and its Applications*, vol. 363, pp. 237–250, 2003.
[58] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1–24, 1972.
[59] A. C. E. Onslow, M. W. Jones, and R. Bogacz, "A canonical circuit for generating phase-amplitude coupling," *PLOS One*, vol. 9, no. 8, p. e102591, 2014.
[60] M. P. Jadi and T. J. Sejnowski, "Regulating cortical oscillations in an inhibition-stabilized network," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 830–842, 2014.
[61] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.
[62] H. S. Seung, D. D. Lee, B. Y. Reis, and D. W. Tank, "Stability of the memory of eye position in a recurrent network of conductance-based model neurons," *Neuron*, vol. 26, no. 1, pp. 259–271, 2000.
[63] D. Durstewitz, J. K. Seamans, and T. J. Sejnowski, "Neurocomputational models of working memory," *Nature Neuroscience*, vol. 3, no. 11s, p. 1184, 2000.
[64] R. Cossart, D. Aronov, and R. Yuste, "Attractor dynamics of network up states in the neocortex," *Nature*, vol. 423, no. 6937, pp. 283–288, 2003.
[65] J. J. Knierim and K. Zhang, "Attractor dynamics of spatially correlated neural activity in the limbic system," *Annual Review of Neuroscience*, vol. 35, no. 1, pp. 267–285, 2012.
[66] A. Pavlov, N. van de Wouw, and H. Nijmeijer, "Convergent piecewise affine systems: analysis and design Part I: continuous case," in *IEEE Conf. on Decision and Control*, Dec 2005, pp. 5391–5396.
[67] E. Nozari and J. Cortés, "Hierarchical selective recruitment in linear-threshold brain networks. Part I: Intra-layer dynamics and selective inhibition," *Preliminary version, available online at https://arxiv.org/abs/1809.01674v2*, 2019.
[68] D. Liberzon and R. Tempo, "Common Lyapunov functions and gradient algorithms," *IEEE Transactions on Automatic Control*, vol. 49, no. 6, pp. 990–994, 2004.
[69] E. Nozari and J. Cortés, "Oscillations and coupling in interconnections of two-dimensional brain networks," in *American Control Conference*, Philadelphia, PA, July 2019, pp. 193–198.
[70] Y. Qin, Y. Kawano, and M. Cao, "Partial phase cohesiveness in networks of communitinized Kuramoto oscillators," in *European Control Conference*, 2018, pp. 2028–2033.
[71] M. Jafarian, X. Yi, M. Pirani, H. Sandberg, and K. Henrik Johansson, "Synchronization of kuramoto oscillators in a bidirectional frequency-dependent tree network," in *IEEE Conf. on Decision and Control*, 2018, pp. 4505–4510.
[72] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti, "Stability conditions for cluster synchronization in networks of heterogeneous Kuramoto oscillators," *IEEE Transactions on Control of Network Systems*, 2019, in press.
[73] H. G. Schuster and P. Wagner, "A model for neuronal oscillations in the visual cortex. 1. mean-field theory and derivation of the phase equations," *Biological Cybernetics*, vol. 64, no. 1, pp. 77–82, 1990.
[74] P. Jiruska, J. Csicsvari, A. D. Powell, J. E. Fox, W. Chang, M. Vreugdenhil, X. Li, M. Palus, A. F. Bujan, R. W. Dearden, and J. G. R. Jefferys, "High-frequency network activity, global increase in neuronal activity, and synchrony expansion precede epileptic seizures in vitro," *Journal of Neuroscience*, vol. 30, no. 16, pp. 5690–5701, 2010.
[75] S. K. Y. Nikravesh, *Nonlinear Systems Stability Analysis: Lyapunov-Based Approach*. CRC Press, 2013.
[76] J. S. Isaacson and M. Scanziani, "How inhibition shapes cortical activity," *Neuron*, vol. 72, no. 2, pp. 231–243, 2011.
[77] E. D. Sontag, "On the input-to-state stability property," *European Journal of Control*, vol. 1, pp. 24–36, 1995.
[78] C. R. Johnson and H. A. Robinson, "Eigenvalue inequalities for principal submatrices," *Linear Algebra and its Applications*, vol. 37, pp. 11–22, 1981.
[79] R. A. Nicoll, "A brief history of long-term potentiation," *Neuron*, vol. 93, no. 2, pp. 281–290, 2017.
[80] 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.
[81] P. Somogyi, G. Tamasab, R. Lujan, and E. H. Buhl, "Salient features of synaptic organisation in the cerebral cortex," *Brain Research Reviews*, vol. 26, no. 2, pp. 113–135, 1998.
[82] H. L. Royden and P. Fitzpatrick, *Real Analysis*. Prentice Hall, 2010.
[83] C. T. Chen, *Linear System Theory and Design*, 3rd ed. New York, NY, USA: Oxford University Press, Inc., 1998.
[84] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLOS Biology*, vol. 3, no. 3, p. e68, 2005.
[85] G. Turrigiano, "Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function," *Cold Spring Harbor Perspectives in Biology*, vol. 4, p. a005736, 2012.

APPENDIX A. ADDITIONAL LEMMAS AND PROOFS

Proof of Theorem IV.6: The necessity is trivial since $\mathbf{M_0}^{-1}\mathbf{M}_{\boldsymbol{\ell}} = -(\mathbf{I}-\mathbf{W})$. To prove sufficiency, note that for any $\boldsymbol{\sigma} \in \{0,\ell\}^n$,

$$\mathbf{M}_{\boldsymbol{\sigma}}^{-1} = (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma}^{\ell})(2\boldsymbol{\Sigma}^{\ell} - \mathbf{I}) = (2\boldsymbol{\Sigma}^{\ell} - \mathbf{I}) - \mathbf{W}\boldsymbol{\Sigma}^{\ell}. \quad (31)$$

Since nodes can be relabeled arbitrarily, we can assume without loss of generality that $\boldsymbol{\sigma}_1 = [\boldsymbol{\ell}_{n_1}^T \;\; \boldsymbol{\ell}_{n_2}^T \;\; \mathbf{0}_{n_3}^T \;\; \mathbf{0}_{n_4}^T]^T$ and $\boldsymbol{\sigma}_2 = [\boldsymbol{\ell}_{n_1}^T \;\; \mathbf{0}_{n_2}^T \;\; \boldsymbol{\ell}_{n_3}^T \;\; \mathbf{0}_{n_4}^T]^T$ where $n_1,\dots,n_4 \geq 0, \sum_{i=1}^4 n_i = n$. Then, it follows from (31) that

$$\mathbf{M}_{\boldsymbol{\sigma}_1}^{-1} = \left[\begin{array}{cc:cc} \mathbf{I}_{n_1}-\mathbf{W}_{11} & -\mathbf{W}_{12} & \mathbf{0} & \mathbf{0} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2}-\mathbf{W}_{22} & \mathbf{0} & \mathbf{0} \\ \hdashline -\mathbf{W}_{31} & -\mathbf{W}_{32} & -\mathbf{I}_{n_3} & \mathbf{0} \\ -\mathbf{W}_{41} & -\mathbf{W}_{42} & \mathbf{0} & -\mathbf{I}_{n_4} \end{array}\right],$$

$$\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} = \begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & \mathbf{0} & -\mathbf{W}_{13} & \mathbf{0} \\ -\mathbf{W}_{21} & -\mathbf{I}_{n_2} & -\mathbf{W}_{23} & \mathbf{0} \\ -\mathbf{W}_{31} & \mathbf{0} & \mathbf{I}_{n_3}-\mathbf{W}_{33} & \mathbf{0} \\ -\mathbf{W}_{41} & \mathbf{0} & -\mathbf{W}_{43} & -\mathbf{I}_{n_4} \end{bmatrix},$$

where $\mathbf{W}_{ij}$'s are submatrices of $\mathbf{W}$ with appropriate dimensions. Taking the inverse of $\mathbf{M}_{\boldsymbol{\sigma}_1}^{-1}$ as a 2-by-2 block-triangular matrix [45, Prop 2.8.7] (with the indicated blocks), we get

$$\mathbf{M}_{\boldsymbol{\sigma}_1} = \begin{bmatrix} \begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & -\mathbf{W}_{12} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2}-\mathbf{W}_{22} \end{bmatrix}^{-1} & \mathbf{0} \\ -\begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \\ \mathbf{W}_{41} & \mathbf{W}_{42} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & -\mathbf{W}_{12} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2}-\mathbf{W}_{22} \end{bmatrix}^{-1} & -\mathbf{I}_{n_3+n_4} \end{bmatrix},$$

so direct multiplication gives $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{B}_4 \end{bmatrix}$, with

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & -\mathbf{W}_{12} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2}-\mathbf{W}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & \mathbf{0} \\ -\mathbf{W}_{21} & -\mathbf{I}_{n_2} \end{bmatrix},$$

$$\mathbf{B}_2 = -\begin{bmatrix} \mathbf{I}_{n_1}-\mathbf{W}_{11} & -\mathbf{W}_{12} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2}-\mathbf{W}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{W}_{13} & \mathbf{0} \\ \mathbf{W}_{23} & \mathbf{0} \end{bmatrix},$$

$$\mathbf{B}_3 = -\begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \\ \mathbf{W}_{41} & \mathbf{W}_{42} \end{bmatrix} \mathbf{B}_1 + \begin{bmatrix} \mathbf{W}_{31} & \mathbf{0} \\ \mathbf{W}_{41} & \mathbf{0} \end{bmatrix},$$

$$\mathbf{B}_4 = -\begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \\ \mathbf{W}_{41} & \mathbf{W}_{42} \end{bmatrix} \mathbf{B}_2 - \begin{bmatrix} \mathbf{I}_{n_3}-\mathbf{W}_{33} & \mathbf{0} \\ -\mathbf{W}_{43} & -\mathbf{I}_{n_4} \end{bmatrix}.$$

With this, after some computations one can show that

$$\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} = \begin{bmatrix} \mathbf{I}_{n_1} & \star & \star & \mathbf{0} \\ \mathbf{0} & \begin{bmatrix} \star & \star \\ \star & \star \end{bmatrix} & & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0} \\ \mathbf{0} & \star & \star & \mathbf{I}_{n_4} \end{bmatrix}. \tag{32}$$

Let $\mathbf{\Gamma} \in \mathbb{R}^{(n_2+n_3)\times(n_2+n_3)}$ be the bracketed block in $\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1}$ and define

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{I}_{n_1} - \mathbf{W}_{11} & -\mathbf{W}_{12} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2} - \mathbf{W}_{22} \end{bmatrix}^{-1},$$
$$\mathbf{R} \triangleq \mathbf{I}_{n_3} - \mathbf{W}_{33} - \begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \end{bmatrix} \mathbf{Q} \begin{bmatrix} \mathbf{W}_{13} \\ \mathbf{W}_{23} \end{bmatrix}.$$

It can be shown that

$$\mathbf{\Gamma} = -\begin{bmatrix} \mathbf{Q}_{22} & \begin{bmatrix} \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{W}_{13} \\ \mathbf{W}_{23} \end{bmatrix} \\ -\begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{12} \\ \mathbf{Q}_{22} \end{bmatrix} & \mathbf{R} \end{bmatrix}.$$

Inverting the left-hand-side matrix (below) as a 2-by-2 block matrix [45, Prop 2.8.7] (the first block is $\mathbf{Q}^{-1}$) and applying the matrix inversion lemma [45, Cor 2.8.8] to the first block of the result, we obtain

$$\left[\begin{array}{cc:c} \mathbf{I}_{n_1} - \mathbf{W}_{11} & -\mathbf{W}_{12} & -\mathbf{W}_{13} \\ -\mathbf{W}_{21} & \mathbf{I}_{n_2} - \mathbf{W}_{22} & -\mathbf{W}_{23} \\ \hdashline -\mathbf{W}_{31} & -\mathbf{W}_{32} & \mathbf{I}_{n_3} - \mathbf{W}_{33} \end{array}\right]^{-1} = \begin{bmatrix} \star & \star & \star \\ \star & \hat{\mathbf{B}}_1 & \hat{\mathbf{B}}_2 \\ \star & \hat{\mathbf{B}}_3 & \hat{\mathbf{B}}_4 \end{bmatrix}$$

where

$$\hat{\mathbf{B}}_1 = \mathbf{Q}_{22} + \begin{bmatrix} \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{W}_{13} \\ \mathbf{W}_{23} \end{bmatrix} \mathbf{R}^{-1} \begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{12} \\ \mathbf{Q}_{22} \end{bmatrix},$$
$$\hat{\mathbf{B}}_2 = \begin{bmatrix} \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{W}_{13} \\ \mathbf{W}_{23} \end{bmatrix} \mathbf{R}^{-1},\ \hat{\mathbf{B}}_3 = \mathbf{R}^{-1} \begin{bmatrix} \mathbf{W}_{31} & \mathbf{W}_{32} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{12} \\ \mathbf{Q}_{22} \end{bmatrix},$$

and $\hat{\mathbf{B}}_4 = \mathbf{R}^{-1}$. Therefore, $-\mathbf{\Gamma}$ is the principal pivot transform of $\begin{bmatrix} \hat{\mathbf{B}}_1 & \hat{\mathbf{B}}_2 \\ \hat{\mathbf{B}}_3 & \hat{\mathbf{B}}_4 \end{bmatrix}$ so if $\mathbf{I} - \mathbf{W} \in \mathcal{P}$, Lemma II.2(v) and the block structure of (32) guarantee that $-\mathbf{M}_{\boldsymbol{\sigma}_1}\mathbf{M}_{\boldsymbol{\sigma}_2}^{-1} \in \mathcal{P}$. ■

The following result is used in the proof of Theorem V.2.

Lemma A.1. ***(GES of cascaded interconnections).*** *Consider the cascaded dynamics*

$$\begin{aligned} \tau\dot{\mathbf{x}}^{\mathtt{o}} &= -\mathbf{x}^{\mathtt{o}}, \\ \tau\dot{\mathbf{x}}^{\mathtt{1}} &= -\mathbf{x}^{\mathtt{1}} + [\mathbf{W}^{\mathtt{1o}}\mathbf{x}^{\mathtt{o}} + \mathbf{W}^{\mathtt{11}}\mathbf{x}^{\mathtt{1}} + \tilde{\mathbf{d}}^{\mathtt{1}}]_{\mathbf{0}}^{\mathbf{m}^{\mathtt{1}}}, \end{aligned} \tag{33}$$

where $\mathbf{x}^{\mathtt{o}} \in \mathbb{R}^r$ *and* $\mathbf{x}^{\mathtt{1}} \in \mathbb{R}^{n-r}$. *If* $\mathbf{W}^{\mathtt{11}}$ *is such that*

$$\tau\dot{\mathbf{x}}^{\mathtt{1}} = -\mathbf{x}^{\mathtt{1}} + [\mathbf{W}^{\mathtt{11}}\mathbf{x}^{\mathtt{1}} + \tilde{\mathbf{d}}^{\mathtt{1}}]_{\mathbf{0}}^{\mathbf{m}^{\mathtt{1}}}, \tag{34}$$

is GES for any constant $\tilde{\mathbf{d}}^{\mathtt{1}} \in \mathbb{R}^{n-r}$, *then the whole dynamics* (33) *is also GES for any constant* $\tilde{\mathbf{d}}^{\mathtt{1}}$.

Proof: We only prove the result for $\tilde{\mathbf{d}}^{\mathtt{1}} = \mathbf{0}$. This is without loss of generality, since for $\tilde{\mathbf{d}}^{\mathtt{1}} \neq \mathbf{0}$, we can apply the change of variables $\boldsymbol{\xi} = \mathbf{x} - \mathbf{x}^*$, where $\mathbf{x}^*$ is the equilibrium corresponding to input $[\mathbf{0}^T\ (\tilde{\mathbf{d}}^{\mathtt{1}})^T]^T$, and shift the equilibrium to the origin. Since (34) is GES, [19, Thm A.1] guarantees that there exists $\mathbf{x}^{\mathtt{1}} \mapsto V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}})$ such that

$$c_1\|\mathbf{x}^{\mathtt{1}}\|^2 \leq V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}}) \leq c_2\|\mathbf{x}^{\mathtt{1}}\|^2, \tag{35a}$$

$$\left\|\frac{\partial V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}})}{\partial \mathbf{x}^{\mathtt{1}}}\right\| \leq c_3\|\mathbf{x}^{\mathtt{1}}\|, \tag{35b}$$

for some $c_1, c_2, c_3 > 0$, and, if $\mathbf{x}^{\mathtt{1}}(t)$ is the solution of (34),

$$\tau\frac{d}{dt}V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}}(t)) \leq -c_4\|\mathbf{x}^{\mathtt{1}}\|^2, \tag{35c}$$

for some $c_4 > 0$. Since $[\cdot]_{\mathbf{0}}^{\mathbf{m}^{\mathtt{1}}}$ is Lipschitz continuous, it follows from (35b) and (35c) that if $\mathbf{x}^{\mathtt{1}}(t)$ is the solution of (33),

$$\begin{aligned} \tau\frac{d}{dt}V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}}(t)) &\leq -c_4\|\mathbf{x}^{\mathtt{1}}\|^2 + c_3\|\mathbf{x}^{\mathtt{1}}\|\|\mathbf{W}^{\mathtt{1o}}\mathbf{x}^{\mathtt{o}}\| \\ &\leq -\frac{c_4}{2}\|\mathbf{x}^{\mathtt{1}}\|^2 + \frac{c_3^2\|\mathbf{W}^{\mathtt{1o}}\|^2}{2c_4}\|\mathbf{x}^{\mathtt{o}}\|^2, \end{aligned}$$

where the second inequality follows from Young's inequality [50, p. 466]. Now, let $V(\mathbf{x}) = (c_3^2\|\mathbf{W}^{\mathtt{1o}}\|^2/2c_4)\|\mathbf{x}^{\mathtt{o}}\|^2 + V^{\mathtt{1}}(\mathbf{x}^{\mathtt{1}})$. It is straightforward to verify that V satisfies all the assumptions of [50, Thm 4.10] with $a = 2$. ■

Erfan Nozari received his B.Sc. degree in Electrical Engineering-Control in 2013 from Isfahan University of Technology, Iran and Ph.D. in Mechanical Engineering and Cognitive Science in 2019 from University of California San Diego. He is currently a postdoctoral researcher at the University of Pennsylvania Department of Electrical and Systems Engineering. He has been the (co)recipient of the 2019 IEEE Transactions on Control of Network Systems Outstanding Paper Award, the Best Student Paper Award from the 57th IEEE Conference on Decision and Control, the Best Student Paper Award from the 2018 American Control Conference, and the Mechanical and Aerospace Engineering Distinguished Fellowship Award from the University of California San Diego. His research interests include dynamical systems and control theory and its applications in computational and theoretical neuroscience and complex network systems.

Jorge Cortés (M'02-SM'06-F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE and SIAM. His current research interests include distributed control and optimization, network science, opportunistic state-triggered control, reasoning and decision making under uncertainty, and distributed coordination in power networks, robotics, and transportation.