

Stability Analysis of Complex Networks with Linear-Threshold Rate Dynamics

Erfan Nozari Jorge Cortés

Abstract—Network models with linear-threshold rate dynamics have been widely used to explain the behavior of biological neural networks and replicate it using artificial neural networks. A full characterization of the stability properties of these networks, nevertheless, has remained elusive. This paper addresses the study of the existence and uniqueness of equilibria and asymptotic stability, leading to a thorough understanding of the conditions on the network structure that determine these properties. Given the stringency of these conditions for large-scale complex networks, we then study the stabilizability of linear-threshold network dynamics and show that, using either feedback or feedforward control, stabilization of the entire network is solely determined by the subnetwork of nodes that are not *directly* controlled. Illustrative examples demonstrate our results.

I. INTRODUCTION

Theoretical and computational approaches to neuroscience have received a surge of attention in recent years, with two main contributing factors: the exponential growth in the computational power of digital computers, allowing for large-scale simulations of neural networks and the associated data analyses; and the remarkable advancements in neurotechnologies of brain recording and stimulation. Unlike decades ago when the scarcity of data was the main challenge ahead of neuroscience, new technologies have allowed the collection of large datasets the full analysis of which relies on novel theoretical and computational approaches.

A fundamental question in neuroscience is how the brain dynamically employs its fixed neuronal circuits to perform different tasks at different times over short timescales. Unlike the classical modular model of cognitive behavior, the more recent network model of cognition attributes distinct functions of a brain region to the emergent behavior of the same recurrent network under different external conditions [1]. In this view, the *dynamic dimensionality* of a network, namely, the number of neurons that are active during every period of time, is arguably the most important network variable that determines its flexibility and capacity for generating complex behaviors [2]–[5]. Here, we propose dynamical stability as a key constraint for dynamic dimensionality control in a (biological) neuronal network. Using linear-threshold models, we seek to find mechanisms for feedforward and/or feedback interconnections of local networks that can achieve dynamic dimensionality control by inhibiting different subsets of nodes during different time intervals.

Literature review

First proposed as a model for the lateral eye of the horseshoe crab by Hartline and Ratliff [6], the dynamical behavior of linear-threshold (not necessarily symmetric) network models have been studied at least as early as [7]. The work [8] provides an initial summary of their properties, including existence and uniqueness of equilibria and asymptotic stability, with limited rigorous justification provided later in [9]. These results were further expanded using Lyapunov-based analysis in [10] for boundedness of solutions and [11]–[14] for asymptotic stability. Nevertheless, the use of a different Lyapunov function in each work has impeded a unified Lyapunov-based stability analysis of linear-threshold networks. The recent work [15] presents several interesting properties of competitive (i.e., fully inhibitory) linear threshold networks. These properties, however, turn out to be sensitive to the assumption that all neurons receive equal background excitation. Overall, the existing literature only considers constant input signals, corresponding to background modulatory activity. Here, after extending these results on constant input signals, we study the role of time-varying input controls used for stabilization and response-shaping. Finally, our work has connections with classical tools from nonlinear systems [16] and switched linear and affine systems [17]–[19].

Statement of contributions

Our contributions are threefold. First, we formulate and prove several results on the existence and uniqueness of equilibria and local and global exponential stability for dynamically-isolated networks. Our results subsume the ones available in the literature and provide a comprehensive characterization of the properties of the network structure that determine the stability of linear-threshold networks. Our second contribution is the analysis of stabilization of linear-threshold networks using state feedback. Here, we show that stabilizability by means of linear state feedback is precisely determined by the stability of the subset of nodes that are not directly controlled, a relatively counter-intuitive result that stems from the (apparently simple, yet intricate) nonlinearity in the network dynamics. Finally we show, as our third contribution, that the exact same critical role is played by the not-directly controlled subnetwork when we consider stabilization through feedforward inhibition. These results give intuition on both the number and the centrality of nodes that need to be inhibited in order for an unstable-in-isolation network to gain stability. Due to space constraints, proofs are omitted here and will be available elsewhere.

E. Nozari and J. Cortés are with the Department of Mechanical and Aerospace Engineering, UC San Diego, {enozari,cortes}@ucsd.edu.

II. PRELIMINARIES

Here, we introduce notational conventions and review basic concepts on matrix analysis and modeling of biological neural networks.

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, resp. We use bold-faced letters for vectors and matrices. $\mathbf{1}_n$, $\mathbf{0}_n$, $\mathbf{0}_{m \times n}$, and \mathbf{I}_n stand for the n -vector of all ones, the n -vector of all zeros, the m -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 i th component. Given $\mathbf{A} \in \mathbb{R}^{n \times m}$, a_{ij} refers to the (i, j) th entry. For block-partitioned \mathbf{x} and \mathbf{A} , \mathbf{x}_i , \mathbf{A}_i , and \mathbf{A}_{ij} refer to the i th block of \mathbf{x} , i th block (e.g., row) of \mathbf{A} , and (i, j) th block of \mathbf{A} , resp. For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\text{range}(\mathbf{A})$ denotes the subspace of \mathbb{R}^m 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 2-norm are denoted by $|\mathbf{A}|$, $\det(\mathbf{A})$, $\rho(\mathbf{A})$, and $\|\mathbf{A}\|$, resp. Similarly, for $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|$ is its 2-norm. For $\boldsymbol{\sigma} \in \{0, 1\}^n$, we make the convention that $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma}) \in \mathbb{R}^{n \times n}$ denotes the diagonal matrix with the elements of $\boldsymbol{\sigma}$ on its diagonal. For a set S , $|S|$ denotes its cardinality. For $x \in \mathbb{R}$, $[x]^+ = \max\{0, x\}$, which is extended entry-wise to vectors and matrices. $\delta(\cdot)$ denotes Dirac delta function. In block representation of matrices, \star denotes arbitrary blocks whose value is immaterial.

Matrix Analysis: Here, we provide a brief description of the matrix classes of interest and their inclusion relationships.

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 all $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})$ and $\boldsymbol{\sigma} \in \{0, 1\}^n$;
- (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. \square

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 [20], [21]). $\mathbf{A} \in \mathbb{R}^{n \times n}$ 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 $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 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. \square

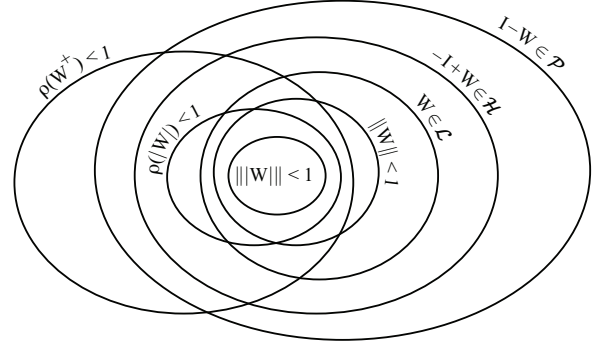


Fig. 1: Inclusion relationships (cf. Lemma II.3) between the matrix classes introduced in Definition II.1.

The next result states inclusion relationships among the matrix classes in Definition II.1 that will be used in our ensuing discussion.

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) $\mathbf{A} \in \mathcal{H} \Rightarrow -\mathbf{A} \in \mathcal{P}$. \square

If \mathbf{W} satisfies the *Dale's law* (as biological neural networks do), i.e., each column is either nonnegative or nonpositive, then one also has $\|\mathbf{W}\| = \|\mathbf{W}\| \geq \rho(|\mathbf{W}|)$. Figure 1 depicts a Venn diagram of the various matrix classes of interest to help visualize their relationships.

Dynamical Rate Models of Cortical Circuits: Here we briefly review, following [22, §7], the fundamental concepts and assumptions that underlie the linear-threshold network model used throughout the paper.

In a lumped model, neural circuits are composed of neurons communicating through spike trains of the form $\rho(t) = \sum_k \delta(t - t_k)$. In many cortical areas, however, the exact timing $\{t_k\}$ of $\rho(t)$ seems essentially random, with the information mainly encoded in its firing rate (i.e., number of spikes per second). Thus, $\rho(t)$ is modeled as 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 $x_{\text{post}}(t)$, resp. As a result of $x_{\text{pre}}(t)$, an electrical current $I_{\text{post}}(t)$ forms in the post-synaptic neuron's body. Assuming fast synaptic dynamics, $I_{\text{post}}(t) \propto x_{\text{pre}}(t)$. Let $w_{\text{post,pre}}$ be the proportionality constant, so $I_{\text{post}}(t) = w_{\text{post,pre}}x_{\text{pre}}(t)$. If the post-synaptic neuron receives input from multiple neurons, $I_{\text{post}}(t)$ follows a superposition law,

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

where $w_{\text{post},j} > 0$ (< 0) if neuron j is excitatory (inhibitory).

If I_{post} is constant, the post-synaptic rate follows $x_{\text{post}} = F(I_{\text{post}})$, where F is a nonlinear "response function". Here, among the two widely used response functions, namely, sigmoidal and linear-threshold, we use the latter: $F(\cdot) = [\cdot]^+$.

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.,

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

(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 groups of neurons with similar activation patterns as a single *node*. The “firing rate” of such a node is then defined as the average of the individual firing rates. This convention also has the advantage of getting more consistent rates, as the firing pattern of individual neurons may be at times sparse.¹ 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*

$$\tau \dot{\mathbf{x}}(t) = -\mathbf{x}(t) + [\mathbf{W}\mathbf{x}(t) + \mathbf{d}(t)]^+, \quad t \geq 0. \quad (3)$$

The extra term $\mathbf{d}(t) \in \mathbb{R}^n$ is the *external input* to the network that accounts for un-modeled background activity, possibly nonzero thresholds, and (more importantly) *top-down* controls (signals coming to the network from other brain regions). Note that (3) is a switched linear system with 2^n modes. In each mode, corresponding to a $\boldsymbol{\sigma} \in \{0, 1\}^n$,

$$[\mathbf{W}\mathbf{x}(t) + \mathbf{d}(t)]^+ = \boldsymbol{\Sigma}(\mathbf{W}\mathbf{x}(t) + \mathbf{d}(t)),$$

where $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})$. This observation motivates the following assumptions on the weight matrix \mathbf{W} .

Assumption 1. *Assume*

- (i) $\det(\mathbf{W}) \neq 0$;
- (ii) $\det(\mathbf{I} - \boldsymbol{\Sigma}\mathbf{W}) \neq 0$ for all $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})$, $\boldsymbol{\sigma} \in \{0, 1\}^n$. \square

This assumption is not a restriction in practical applications. In fact, the set of matrices for which Assumption 1 is not satisfied can be expressed as a finite union of measure-zero sets, and hence has (Lebesgue) measure zero.

Simulations of the dynamics (3) for a few random cases of network matrix \mathbf{W} and constant \mathbf{d} reveal that

(i) locally, the network may have zero, one, or many equilibrium points, where each equilibrium may be stable or unstable independent of others,

(ii) globally, the network is capable of exhibiting different nonlinear phenomena such as limit cycles, multi-stability, and chaos. Furthermore, the state trajectories grow unbounded (in reality until saturation) if the excitatory sub-network $[\mathbf{W}]^+$ is sufficiently strong.

This richness of behavior only increases if the network is subject to a time-varying $\mathbf{d}(t)$. Motivated by these observations, our goal is to characterize the stability properties of the network and the extent to which $\mathbf{d}(t)$ can stabilize an unstable network. We begin the discussion in Section IV

with the stability analysis of the simplified case of dynamical isolation, i.e., constant \mathbf{d} . Building on this analysis, we then tackle in Section V the more realistic case of time-varying $\mathbf{d}(t)$ and see how it can be used to stabilize an unstable network, utilizing both feedforward and feedback structures.

IV. STABILITY ANALYSIS

In this section, we provide an in-depth study of the stability properties of the network dynamics (3) for constant $\mathbf{d}(t) = \mathbf{d} \in \mathbb{R}^n$,

$$\tau \dot{\mathbf{x}}(t) = -\mathbf{x}(t) + [\mathbf{W}\mathbf{x}(t) + \mathbf{d}]^+, \quad t \geq 0. \quad (4)$$

Specifically, we derive conditions for existence and uniqueness of equilibria and local/global asymptotic stability.

By Assumption 1(i), $\mathbf{W}\mathbf{x} + \mathbf{d} = \mathbf{0}$ defines a non-degenerate set of n hyperplanes partitioning \mathbb{R}^n into 2^n solid convex polytopic translated cones apexed at $-\mathbf{W}^{-1}\mathbf{d}$. For each $\boldsymbol{\sigma} \in \{0, 1\}^n$, let $\Omega_{\boldsymbol{\sigma}}$ be the associated *switching region*,

$$\Omega_{\boldsymbol{\sigma}} = \{\mathbf{x} \in \mathbb{R}_{\geq 0}^n \mid (2\boldsymbol{\Sigma} - \mathbf{I})(\mathbf{W}\mathbf{x} + \mathbf{d}) \geq \mathbf{0}\}.$$

Clearly, the piecewise-affine dynamics (4) can be written in the equivalent form

$$\tau \dot{\mathbf{x}} = (-\mathbf{I} + \boldsymbol{\Sigma}\mathbf{W})\mathbf{x} + \boldsymbol{\Sigma}\mathbf{d}, \quad \mathbf{x} \in \Omega_{\boldsymbol{\sigma}}, \quad \boldsymbol{\sigma} \in \{0, 1\}^n. \quad (5)$$

In other words, $\Omega_{\boldsymbol{\sigma}}$ is the region where any neuron i for which $\sigma_i = 1$ is active (it receives a net positive input) while any neuron i for which $\sigma_i = 0$ is inactive (its firing rate decays exponentially as $\tau \dot{x}_i = -x_i$). Interestingly, the existence of equilibria is not guaranteed for this system. In fact, for each $\boldsymbol{\sigma} \in \{0, 1\}^n$, according to (5), the point

$$\mathbf{x}_{\boldsymbol{\sigma}}^* = \mathbf{x}_{\boldsymbol{\sigma}}^*(\mathbf{d}) = (\mathbf{I} - \boldsymbol{\Sigma}\mathbf{W})^{-1}\boldsymbol{\Sigma}\mathbf{d}, \quad (6)$$

is the corresponding *equilibrium candidate*. We refer to it as candidate because in general it might not belong to the switching region $\Omega_{\boldsymbol{\sigma}}$ where the description (5) is valid.

A. Existence and Uniqueness of Equilibria

Here we characterize the existence and uniqueness of equilibria for the dynamics (4). Given $\mathbf{W} \in \mathbb{R}^{n \times n}$, define the *equilibria set-valued map* $h : \mathbb{R}^n \rightrightarrows \mathbb{R}_{\geq 0}^n$ by

$$h(\mathbf{d}) \triangleq \{\mathbf{x} \in \mathbb{R}_{\geq 0}^n \mid \mathbf{x} = [\mathbf{W}\mathbf{x} + \mathbf{d}]_+\}. \quad (7)$$

Existence and uniqueness of equilibria precisely corresponds to h being single-valued on \mathbb{R}^n . If so, with slight abuse of notation, we take $h : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}^n$ to be an ordinary function.

From the definition (6) of equilibrium candidate, note that $\mathbf{x}_{\boldsymbol{\sigma}}^* \in h(\mathbf{d})$ if and only if $\mathbf{x}_{\boldsymbol{\sigma}}^* \in \Omega_{\boldsymbol{\sigma}}$. Then, using Assumption 1, and after some manipulations, we have

$$\begin{aligned} \mathbf{W}\mathbf{x}_{\boldsymbol{\sigma}}^* + \mathbf{d} &= \mathbf{W}(\mathbf{I} - \boldsymbol{\Sigma}\mathbf{W})^{-1}\boldsymbol{\Sigma}\mathbf{d} + \mathbf{d} \\ &= [(\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma})^{-1}\mathbf{W}\boldsymbol{\Sigma} + \mathbf{I}]\mathbf{d} = (\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma})^{-1}\mathbf{d}. \end{aligned} \quad (8)$$

Therefore,

$$\mathbf{x}_{\boldsymbol{\sigma}}^* \in h(\mathbf{d}) \Leftrightarrow \underbrace{(2\boldsymbol{\Sigma} - \mathbf{I})(\mathbf{I} - \mathbf{W}\boldsymbol{\Sigma})^{-1}\mathbf{d}}_{\triangleq \mathbf{M}_{\boldsymbol{\sigma}}} \geq \mathbf{0}. \quad (9)$$

¹Our subsequent discussion is nevertheless valid irrespective of whether network nodes represent individual neurons or groups of them.

Accordingly, for $\sigma \in \{0, 1\}^n$, let $\Delta_\sigma \triangleq \{\mathbf{d} \in \mathbb{R}^n \mid \mathbf{M}_\sigma \mathbf{d} \geq \mathbf{0}\}$ be the set of external inputs \mathbf{d} such that (4) has an equilibrium in Ω_σ , which is a closed convex polytopic cone.

Note that if $\mathbf{M}_\sigma \mathbf{d} \geq \mathbf{0}$ for exactly one $\sigma \in \{0, 1\}^n$, then a unique equilibrium exists according to (9). However, when $\mathbf{M}_{\sigma_\ell} \mathbf{d} \geq \mathbf{0}$ for multiple $\sigma_\ell \in \{0, 1\}^n, \ell \in \{1, \dots, \bar{\ell}\}$, the network may have either multiple equilibria or a unique one $\mathbf{x}_{\sigma_1}^* = \dots = \mathbf{x}_{\sigma_{\bar{\ell}}}^*$ lying on the boundary between $\{\Omega_{\sigma_\ell}\}_{\ell=1}^{\bar{\ell}}$. The next result shows that the quantities $\mathbf{M}_\sigma \mathbf{d}$ can be used to distinguish between these two latter cases.

Lemma IV.1. (Existence of multiple equilibria). *Assume \mathbf{W} satisfies Assumption 1, $\mathbf{d} \in \mathbb{R}^n$ is arbitrary, and \mathbf{M}_σ is defined as in (9). If there exist $\sigma_1 \neq \sigma_2$ such that $\mathbf{d} \in \Delta_{\sigma_1} \cap \Delta_{\sigma_2}$, then $\mathbf{x}_{\sigma_1}^* = \mathbf{x}_{\sigma_2}^*$ if and only if $\mathbf{M}_{\sigma_1} \mathbf{d} = \mathbf{M}_{\sigma_2} \mathbf{d}$. \square*

Our next result provides an optimization-based condition for existence and uniqueness of equilibria that is both necessary and sufficient.

Proposition IV.2. (Optimization-based condition for existence and uniqueness of equilibria). *Let \mathbf{W} satisfy Assumption 1 and \mathbf{M}_σ be as defined in (9). For $\mathbf{d} \in \mathbb{R}^n$, define $\mu_1(\mathbf{d})$ and $\mu_2(\mathbf{d})$ to be the largest and second largest element of $\{\min_{i=1, \dots, n} (\mathbf{M}_\sigma \mathbf{d})_i \mid \sigma \in \{0, 1\}^n\}$, resp. Then, (4) has a unique equilibrium for each $\mathbf{d} \in \mathbb{R}^n$ if and only if*

$$\max_{\|\mathbf{d}\|=1} \mu_1(\mathbf{d})\mu_2(\mathbf{d}) < 0. \quad (10)$$

\square

The optimization involved in (10) is usually highly non-convex. However, since the search space $\|\mathbf{d}\| = 1$ is compact, global search methods can be used to verify (10) numerically if n is small. Next, we give our main result regarding the existence and uniqueness of equilibria that not only is analytically verifiable for large n but also provides significant intuition into the class of matrices \mathbf{W} that satisfy existence and uniqueness of equilibria.

Theorem IV.3. (Existence and uniqueness of equilibria). *Consider the network dynamics (4) and assume the weight matrix \mathbf{W} satisfies Assumption 1. Then, (4) has a unique equilibrium for each $\mathbf{d} \in \mathbb{R}^n$ if $\mathbf{I} - \mathbf{W} \in \mathcal{P}$. \square*

The fact that the condition in Lemma II.2(iv) is an equivalent characterization of P-matrices suggests that the sufficient condition of Theorem IV.3 is tight. Indeed, extensive simulations with random matrices did not reveal any instance of \mathbf{W} that is not a P-matrix but for which (4) has a unique equilibrium for all $\mathbf{d} \in \mathbb{R}^n$ (where we checked the latter using the characterization of Proposition IV.2). This leads us to the following conjecture, which was also made in [8] as a claim without proof.

Conjecture IV.4. (Necessity of $\mathbf{I} - \mathbf{W} \in \mathcal{P}$). *Assume the weight matrix \mathbf{W} satisfies Assumption 1. Then, (4) has a unique equilibrium for all $\mathbf{d} \in \mathbb{R}^n$ if and only if $\mathbf{I} - \mathbf{W} \in \mathcal{P}$. \square*

The next example analyzes the existence and uniqueness of equilibria for one of the simplest and most widely-used models in computational neuroscience.

Example IV.5. (Uniform E-I networks). Consider a network of n nodes in which $\alpha n, \alpha \in (0, 1)$ are excitatory (E), $(1 - \alpha)n$ are inhibitory (I), and the synaptic weight between any pair of nodes only depends on their type (the synaptic weight of any I-to-E connection is $w_{ei} < 0$, and similarly for $w_{ee} > 0, w_{ie} > 0, w_{ii} < 0$). Also, assume common external inputs $d_e, d_i \in \mathbb{R}$ for all E and I nodes, resp. Let $x_e(t)$ and $x_i(t)$ be the average firing rates of E and I nodes, resp. Then,²

$$\tau \begin{bmatrix} \dot{x}_e \\ \dot{x}_i \end{bmatrix} = - \begin{bmatrix} x_e \\ x_i \end{bmatrix} + \begin{bmatrix} \alpha n w_{ee} & (1 - \alpha) n w_{ei} \\ \alpha n w_{ie} & (1 - \alpha) n w_{ii} \end{bmatrix} \begin{bmatrix} x_e \\ x_i \end{bmatrix} + \begin{bmatrix} d_e \\ d_i \end{bmatrix}.$$

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

$$\begin{aligned} \mathbf{I} - \mathbf{W}_{EI} \in \mathcal{P} &\Leftrightarrow \alpha n w_{ee} < 1, \\ \rho(|\mathbf{W}_{EI}|) < 1 &\Leftrightarrow \alpha n w_{ee} < 1, (1 - \alpha) n |w_{ii}| < 1, \\ \text{and } \alpha(1 - \alpha) n^2 w_{ie} |w_{ei}| &< (1 - \alpha n w_{ee})(1 - (1 - \alpha) n |w_{ii}|). \end{aligned}$$

Thus, according to Theorem IV.3, existence and uniqueness of equilibria only requires the E dynamics to be stable (note that w_{ee} has to be smaller as n grows), while the more conservative $\rho(|\mathbf{W}_{EI}|) < 1$ requires two extra conditions: the stability of I dynamics and a weak interconnection between E and I subnetworks (a small-gain type of condition). \square

B. Asymptotic Stability

The existence and uniqueness of an equilibrium, as discussed above, is an opportunity to shape the network state, provided the equilibrium corresponds to a desired state (e.g., a memory, location, or eye position) and it attracts network trajectories [28]–[30]. Here we investigate when the latter holds, i.e., the network equilibrium is asymptotically stable. Our main result on asymptotic stability is the following.

Theorem IV.6. (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 $-\mathbf{I} + \mathbf{W} \in \mathcal{H}$. \square*

With regards to [31, Thm 1], our approach in Theorem IV.6 provides a simpler proof for the sufficiency of $\mathbf{W} \in \mathcal{L}$ and a novel proof for the necessity of $-\mathbf{I} + \mathbf{W} \in \mathcal{H}$. From Lemma II.3(iii), the conditions of Theorem IV.6 are not conclusive when \mathbf{W} satisfies $-\mathbf{I} + \mathbf{W} \in \mathcal{H}$ but does not satisfy $\mathbf{W} \in \mathcal{L}$. However,

- (i) If $\rho(|\mathbf{W}|) < 1$, then the network is GES relative to a unique equilibrium for all $\mathbf{d} \in \mathbb{R}^n$. This can be shown similar to [9, Prop. 3].

²This simplification of n -dimensional networks to planar dynamics is commonly known as the Wilson-Cowan model [25], see, e.g. [26], [27].

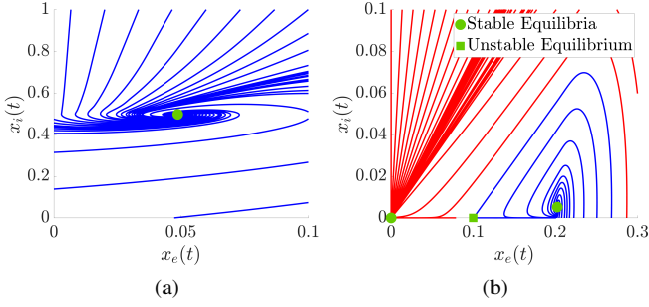


Fig. 2: Network trajectories for the E-I network of Example IV.8. a) When $\mathbf{W}_{EI} = [0.9, -2; 5, -1.5]$, $\mathbf{d}_{EI} = [1; 1]$, network has a unique GES equilibrium. b) However, for $\mathbf{W}_{EI} = [1.1, -2; 5, -1.5]$, $\mathbf{d}_{EI} = [-0.01; -1]$, the network exhibits bistable behavior. The color of the trajectories corresponds to the attractor to which they converge. Note that although $\alpha n w_{ee} > 1$, the network is GES for most values of \mathbf{d}_{EI} , so we used Proposition IV.2 for finding a \mathbf{d}_{EI} that leads to multi-stability.

(ii) If a unique equilibrium \mathbf{x}^* lies in the interior of an Ω_σ (a condition that can be shown to hold for Lebesgue-almost all \mathbf{d}), then \mathbf{x}^* is at least locally exponentially stable.

(iii) 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.7. (Sufficiency of total-Hurwitzness for GES). Consider the network 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 $-\mathbf{I} + \mathbf{W} \in \mathcal{H}$. \square

We next study the GES of the uniform E-I networks of Example IV.5.

Example IV.8. (Uniform E-I networks, cont'd). Consider again the E-I network of Example IV.5. One can verify that

$$-\mathbf{I} + \mathbf{W}_{EI} \in \mathcal{H} \Leftrightarrow \alpha n w_{ee} < 1. \quad (11)$$

Thus, the (sufficient) condition for existence and uniqueness of equilibria and (necessary) condition for GES coincide in this case, and they interestingly only restrict w_{ee} . Figure 2 shows sample phase portraits for $\alpha n w_{ee} < 1$ and $\alpha n w_{ee} > 1$. \square

V. STABILIZATION BY EXOGENOUS CONTROLS

Here we consider the case when the network matrix \mathbf{W} does not satisfy the conditions identified in Section IV ensuring the existence and uniqueness of equilibria or exponential stability. Specifically, we are interested in determining to what extent, in the general dynamics (3), it is possible to use the external input $\mathbf{d}(t)$ to enforce existence and uniqueness of equilibria/stability. To this end, we decompose $\mathbf{d}(t)$ as

$$\mathbf{d}(t) = \mathbf{B}\mathbf{u}(t) + \tilde{\mathbf{d}}. \quad (12)$$

Here, The role of $\tilde{\mathbf{d}} \in \mathbb{R}^n$ is to *shape* the network state $\mathbf{x}(t)$ by determining the network equilibrium (the same role that

\mathbf{d} was originally playing in Section IV when \mathbf{W} had the desired stability properties), while the role of $\mathbf{u}(t) \in \mathbb{R}^m$ is to make such equilibrium asymptotically stable. Let $r \leq n$ be the number of nodes that (directly) receive the input $\mathbf{u}(t)$, and partition \mathbf{W} and \mathbf{B} accordingly, i.e.,

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{0} \end{bmatrix}, \quad (13)$$

where $\mathbf{W}_{11} \in \mathbb{R}^{r \times r}$, $\mathbf{B}_1 \in \mathbb{R}^{r \times m}$ (this 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).

Since feedback (also termed re-entry) plays a fundamental role in network stability in the nervous system, we first consider the case where the input $\mathbf{u}(t)$ is in the state feedback form. The next result gives several necessary and sufficient conditions for network stabilization using state feedback.

Theorem V.1. (Role of the uncontrolled subnetwork in feedback stabilization). Consider the network dynamics (3), (12), where \mathbf{u} has the state feedback form

$$\mathbf{u}(t) = \mathbf{K}\mathbf{x}(t), \quad (14)$$

and $\mathbf{K} \in \mathbb{R}^{m \times n}$ is a constant control gain. Assume that

$$\text{range}([\mathbf{W}_{11} \ \mathbf{W}_{12}]) \subseteq \text{range}(\mathbf{B}_1). \quad (15)$$

Then, there exists $\mathbf{K} \in \mathbb{R}^{m \times n}$ such that

- (i) $\mathbf{I} - (\mathbf{W} + \mathbf{B}\mathbf{K}) \in \mathcal{P}$ if and only if $\mathbf{I} - \mathbf{W}_{22} \in \mathcal{P}$;
- (ii) $-\mathbf{I} + (\mathbf{W} + \mathbf{B}\mathbf{K}) \in \mathcal{H}$ if and only if $-\mathbf{I} + \mathbf{W}_{22} \in \mathcal{H}$;
- (iii) $\mathbf{W} + \mathbf{B}\mathbf{K} \in \mathcal{L}$ if and only if $\mathbf{W}_{22} \in \mathcal{L}$;
- (iv) $\rho(|\mathbf{W} + \mathbf{B}\mathbf{K}|) < 1$ if and only if $\rho(|\mathbf{W}_{22}|) < 1$;
- (v) $\|\mathbf{W} + \mathbf{B}\mathbf{K}\| < 1$ if and only if $\|[\mathbf{W}_{21} \ \mathbf{W}_{22}]\| < 1$. \square

The main conclusion of Theorem V.1 is that the possibility of stabilization by state feedback is determined by the subnetwork \mathbf{W}_{22} of nodes that are not *directly* controlled. The condition (15) requires, essentially, that there are sufficiently many “independent” external controls.

Besides feedback/re-entry, *feedforward inhibition* is also known to play a key role in network stability in the nervous system by *modulating* the baseline activity of neural populations and decreasing their excitability. The next result arrives at the same conclusion as Theorem V.1 (namely, that \mathbf{W}_{22} is the sole determiner of network stabilizability) from the completely independent perspective of feedforward inhibition. First, we need a definition.

Definition V.2. (Monotone boundedness). The dynamics (3) is monotonically bounded if for any $\bar{\mathbf{d}} \in \mathbb{R}^n$ there exists $\nu(\bar{\mathbf{d}})$ such that $\mathbf{x}(t) \leq \nu(\bar{\mathbf{d}})$, $t \geq 0$ for any $\mathbf{d}(t) \leq \bar{\mathbf{d}}$, $t \geq 0$. \square

Note that in reality, the state of any biological neural network is uniformly bounded due to the refractory period of its neurons, implying monotone boundedness.

Theorem V.3. (Role of the uncontrolled subnetwork in feedforward stabilization). Consider the network dynamics (3), (12) with the decomposition (13), where

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

and $\mathbf{B}_1 \in \mathbb{R}_{\leq 0}^{r \times m}$ is purely inhibitory. Assume that,

(i) $m \geq r$;

(ii) the network is monotonically bounded.

Then, for all $\tilde{\mathbf{d}} \in \mathbb{R}^n$ there exists $\mathbf{u} \geq \mathbf{0}$ such that the network is GES relative to a unique equilibrium if and only if for all $\tilde{\mathbf{d}}_2 \in \mathbb{R}^{n-r}$ the uncontrolled subnetwork

$$\tau \dot{\mathbf{x}}_2 = -\mathbf{x}_2 + [\mathbf{W}_{22}\mathbf{x}_2 + \tilde{\mathbf{d}}_2]^+, \quad (16)$$

is GES relative to a unique equilibrium. \square

Theorems V.1 and V.3 have illustrative similarities and differences. In Theorem V.1, the stabilizing control $\mathbf{u}(t)$ is determined based on $\mathbf{x}(t)$ according to a *feedback* law, while in Theorem V.3, \mathbf{u} is constant and independent of $\mathbf{x}(t)$ (thus *feedforward*). However, in both cases, $\mathbf{u}(t)$ can enforce any stability property *if and only if* the not-directly controlled \mathbf{x}_2 -dynamics already enjoys that property. Also, note that the role of $\mathbf{u}(t)$ in Theorem V.3 is purely inhibitory as $\mathbf{B}\mathbf{u} \leq \mathbf{0}$.

VI. CONCLUSIONS

We have studied the stability and stabilizability for networks with linear-threshold dynamics. Regarding stability, we have shown that the classes of P-matrices and totally-Hurwitz matrices characterize network structures that yield existence and uniqueness of equilibria and global exponential stability, resp. Regarding stabilization, we have shown that any stability property is achievable through external controls (either in feedback or feedforward form) if and only if it holds for the subnetwork of nodes that are not directly controlled. These results give rise to several opportunities for future work. Combined with the (log-normal) distribution of synaptic weights and scaling laws of network eigenvalues with its size, these results may play an important role in explaining the inter-regional connectivity structures observed in biological neural networks. Furthermore, the joint analysis of stabilization and time-varying external controls which themselves are outputs of dynamical networks can help explain the inter-regional information flow in cortical networks.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Eric Perterson for multiple conversations on neuroscience and dimensional control. This research was supported by NSF Award ECCS-1307176.

REFERENCES

- [1] S. Haykin and J. M. Fuster, "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 608–628, 2014.
- [2] M. Bianchini and F. Scarselli, "On the complexity of neural network classifiers: A comparison between shallow and deep architectures," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 8, pp. 1553–1565, 2014.
- [3] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein, "On the expressive power of deep neural networks," *arXiv preprint arXiv:1606.05336*, 2016.
- [4] I. Ispolatov, V. Madhok, S. Allende, and M. Doebeli, "Chaos in high-dimensional dissipative dynamical systems," *Scientific reports*, vol. 5, 2015.
- [5] N. Bertschinger and T. Natschläger, "Real-time computation at the edge of chaos in recurrent neural networks," *Neural computation*, vol. 16, no. 7, pp. 1413–1436, 2004.

- [6] F. Ratliff and H. K. Hartline, *Studies on Excitation and Inhibition in the Retina*. Rockefeller University Press, 1974.
- [7] K. P. Hadel, "On the theory of lateral inhibition," *Kybernetik*, vol. 14, no. 3, pp. 161–165, 1973.
- [8] K. P. Hadel and D. Kuhn, "Stationary states of the Hartline-Ratliff model," *Biological Cybernetics*, vol. 56, no. 5-6, pp. 411–417, 1987.
- [9] J. Feng and K. P. Hadel, "Qualitative behaviour of some simple networks," *Journal of Physics A: Mathematical and General*, vol. 29, no. 16, pp. 5019–5033, 1996.
- [10] H. Wersing, W. Beyn, and H. Ritter, "Dynamical stability conditions for recurrent neural networks with unsaturating piecewise linear transfer functions," *Neural Comp.*, vol. 13, no. 8, pp. 1811–1825, 2001.
- [11] Z. Yi and K. K. Tan, "Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions," *IEEE Trans. Neural Networks*, vol. 15, no. 2, pp. 329–336, 2004.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.
- [17] 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.
- [18] D. Liberzon, *Switching in Systems and Control*, ser. Systems & Control: Foundations & Applications. Birkhäuser, 2003.
- [19] M. K. J. Johansson, *Piecewise Linear Control Systems: A Computational Approach*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 2003.
- [20] 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.
- [21] 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.
- [22] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*, ser. Computational Neuroscience, T. J. Sejnowski and T. Poggio, Eds. MIT Press, 2001.
- [23] 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.
- [24] D. S. Bernstein, *Matrix Mathematics*, 2nd ed. Princeton University Press, 2009.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] E. D. Sontag, "On the input-to-state stability property," *European Journal of Control*, vol. 1, pp. 24–36, 1995.