SIAM J. APPL. MATH.
Vol. 84, No. 6, pp. 2522–2544

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

THE COMBINATORIAL CODE AND THE GRAPH RULES OF DALE NETWORKS*

NIKOLA MILIĆEVIĆ† AND VLADIMIR ITSKOV†

Abstract. We describe the combinatorics of equilibria and steady states of neurons in threshold-linear networks that satisfy Dale's law. The combinatorial code of a Dale network is characterized in terms of two conditions: (i) a condition on the network connectivity graph, and (ii) a spectral condition on the synaptic matrix. We find that in the weak coupling regime the combinatorial code depends only on the connectivity graph, and not on the particulars of the synaptic strengths. Moreover, we prove that the combinatorial code of a weakly coupled network is a sublattice, and we provide a learning rule for encoding a sublattice in a weakly coupled excitatory network. In the strong coupling regime we prove that the combinatorial code of a generic Dale network is intersection-complete and is therefore a convex code, as is common in some sensory systems in the brain.

Key words. Dale's law, recurrent neural networks, combinatorial neural codes

MSC code. 92B20

DOI. 10.1137/23M1599069

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

1. Introduction.

We study the equilibria and steady states of threshold-linear networks that satisfy Dale's law. Dale's law, introduced by Dale in 1935 [11, 25], postulates that a single neuron utilizes the same set of chemical messengers. This usually implies that each neuron's efferent (outgoing) synapses are either all excitatory or all inhibitory. While this restriction on the signs of the synaptic weights has a number of exceptions in *some* neural systems, it is observed in most neural circuits on the brain [20].

Threshold-linear networks have been extensively studied in [15, 32, 14, 6, 7, 10, 24], and a number of results regarding the stable fixed points of these networks were obtained, especially in the case of a symmetric synaptic matrix. Networks that satisfy Dale's law have been previously investigated in the context of large random networks [28, 1, 18] (this list is very incomplete), where statistical properties such as the spectrum, the number of equilibria, and phase transitions have been investigated. In contrast, we are interested in describing the exact combinatorics of the neural code of these networks, as opposed to investigating statistical features.

To this end we developed a method for understanding the combinatorial codes of threshold-linear networks that obey Dale's law. A combinatorial code is the collection of patterns of neuronal activation at the equilibria that is possible in a given network. In particular, this discards the details of the firing rates, and only keeps track of what neurons are coactive at a fixed point. It turns out it is possible to directly translate features of the connectivity graph into the combinatorial code of a Dale network. We show that the connectivity features completely determine the combinatorial neural code in the weak coupling regime. In the strong coupling regime, the code is described

*Received by the editors September 6, 2023; accepted for publication (in revised form) August 29, 2024; published electronically December 11, 2024.
https://doi.org/10.1137/23M1599069
Funding: This work was supported by the NSF Next Generation Networks for Neuroscience Program award 2014217.

†Department of Mathematics, The Pennsylvania State University, University Park, PA 16802 USA (nqm5625@psu.edu, https://sites.google.com/view/nikolamilicevic; vladimir.itskov@psu.edu, https://vitskov.github.io/homepage/).

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

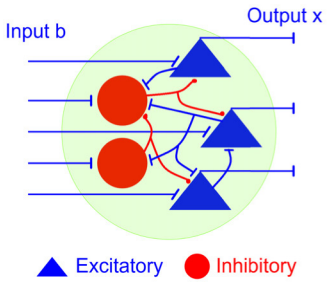


FIG. 1. *A Dale recurrent network.*

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

in terms of connectivity and certain spectral conditions on the excitatory subnetworks. We also show that these combinatorial codes are *convex,* that is, they are compatible with patterns of overlaps of convex receptive fields that are common in many sensory systems of the brain [5].

This paper is organized as follows. In section 2 we introduce the necessary background for threshold-linear networks and Dale's law and define the combinatorial and stable combinatorial codes. In section 3 we state our main results. In section 4 we discuss the findings of our manuscript. The proofs are relegated to section 5. Appendix A contains some necessary results concerning the stability of fixed points of threshold-linear networks.

2. Preliminaries: The combinatorial codes of Dale networks. We consider a standard firing rate model of a recurrent neural network of n neurons, where the dynamics of the firing rates $x_i(t) \geq 0$ is described by the differential equations

$$\dot{x}_i = -x_i + \left[\sum_{j=1}^{n} W_{ij}x_j + b_i\right]_+, \quad i = 1, \ldots, n, \tag{2.1}$$

and $[y]_+ = \max(0, y)$ is the ReLU transfer function.

We assume that this excitatory-inhibitory network respects *Dale's law* [12], where the neurons are either excitatory (denoted as $\mathcal{E}$) or inhibitory (denoted as $\mathcal{I}$), and the synaptic weights satisfy the following sign constraints:

$$\begin{aligned} &\text{excitatory synapses: } i \in \mathcal{E} \implies W_{ji} \geq 0 \quad \forall j = 1 \ldots n, \\ &\text{inhibitory synapses: } i \in \mathcal{I} \implies W_{ji} \leq 0 \quad \forall j = 1 \ldots n. \end{aligned}$$

This follows the neuroscience convention that the weight W_{ij} is the j to i synaptic strength. Furthermore, we assume that the diagonal entries of W are zeros, $W_{ii} = 0$, and denote the collection of all $n \times n$ Dale matrices by $\mathbb{D}_n$.

Following a common architecture of the neocortex, we also assume that the excitatory neurons "broadcast" the output, while the activity of the inhibitory neurons is not observable *directly* outside of the network. We thus consider the setup where the inputs to the network are excitatory, while only the excitatory output can be "read" from the network (Figure 1). We shall call such a network *a Dale network.*

We are interested in the combinatorics of the excitatory output of Dale networks. A combinatorial code is the set of possible patterns of neural activation at the fixed points (or steady states). For a firing rate vector $x \in \mathbb{R}^n_{\geq 0}$, we consider the *excitatory support,* i.e., the set of active excitatory neurons:

$$\operatorname{supp}_+ x = \{i \in \mathcal{E} \mid x_i > 0\} \subset \mathcal{E}.$$

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Recall that $x^* \in \mathbb{R}^n$ is a fixed point of a network (2.1) if $x(t) = x^*$ is a constant solution. For a given Dale network (2.1), we denote the set of excitatory supports of all the possible fixed points as

$$FP_+(W,b) = \left\{ \operatorname{supp}_+ x^* \;\middle|\; x^* \in \mathbb{R}^n_{\geq 0} \text{ is a fixed point of (2.1)} \right\}.$$

Here the plus sign highlights the difference from a somewhat different definition in [23], which considers the combinatorics of all fixed points, which were previously investigated.

The *combinatorial code* of a Dale synaptic matrix W is the collection of all possible excitatory supports of all fixed points in response to all possible inputs:

$$\mathcal{C}(W) \stackrel{\text{def}}{=} \bigcup_{b \in \mathbb{R}^n_{\geq 0}} FP_+(W,b). \tag{2.2}$$

The *stable combinatorial code* is the set of the excitatory supports of asymptotically stable fixed points:

$$\mathcal{SC}(W) \stackrel{\text{def}}{=} \bigcup_{b \in \mathbb{R}^n_{\geq 0}} \{\operatorname{supp}_+ x^* \,|\, x^* \in \mathbb{R}^n_{\geq 0} \text{ is an asymptotically stable fixed point of (2.1)}\}.$$

It turns out that the combinatorial codes of Dale networks can be completely described in terms of connectivity and spectral radius of the synaptic matrix. We also show that the combinatorial code is always intersection complete, which implies it is also always a convex combinatorial code [5].

3. The main results. Here we state the main results, while all the proofs of the theorems are provided in section 5. To simplify the mathematical analyses of the network, we make the following mild assumption.

GROUND ASSUMPTION 3.1. *The synaptic matrix W of the network* (2.1) *satisfies the condition that for every nonempty subset of neurons $\sigma \subset [n]$ the principal submatrix $(I - W)_\sigma$ is nonsingular.*

Note that for a square $n \times n$ matrix A and a subset $\sigma \subset [n]$ we denote by A_σ the appropriate principal submatrix, i.e., the submatrix of A obtained by restricting it to the rows and columns indexed by σ. In all our results Ground Assumption 3.1 is always implicitly assumed. Note that the set of matrices that do *not* have this property has measure zero, thus this assumption is generically satisfied in any network without fine tuning.

3.1. The role of excitatory-inhibitory connectivity in shaping the combinatorial code. We first observe that to understand the combinatorial code, one can streamline the excitatory-inhibitory connectivity to its "essential features" as follows.

THEOREM 3.2. *Let $W \in \mathbb{D}_n$ be a Dale matrix with a set of excitatory neurons $\mathcal{E}$ and inhibitory neurons $\mathcal{I}$. Let $m = |\mathcal{E}| + 1$ and let $W' \in \mathbb{D}_m$ be any Dale matrix such that $W'_\mathcal{E} = W_\mathcal{E}$ and for all $i \in \mathcal{E}$,*

$$W'_{im} = \begin{cases} -1, & \exists j \in \mathcal{I} \text{ with } W_{ij} \neq 0, \\ 0 & \forall j \in \mathcal{I} \text{ with } W_{ij} = 0. \end{cases}$$

Then $\mathcal{C}(W) = \mathcal{C}(W')$.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

In other words, the combinatorial code $\mathcal{C}(W)$ remains unchanged if we replace all the inhibitory neurons with a single inhibitory neuron that mimics the connectivity of the entire inhibitory population to each excitatory neuron. Note that here the numerical values of the inhibitory-excitatory weights of the synaptic matrix W do not affect the combinatorial code, even though they may determine the stability of the appropriate fixed points. Furthermore, the following result states that the connections from excitatory to inhibitory neurons have no influence over the combinatorial code $\mathcal{C}(W)$, while they still may determine the stability of the appropriate fixed points, as well as other dynamical properties.

THEOREM 3.3. *Let $W \in \mathbb{D}_n$ be a Dale matrix with a set of excitatory neurons $\mathcal{E}$ and inhibitory neurons $\mathcal{I}$, and let $W' \in \mathbb{D}_n$ be such that*

$$W'_{ij} = \begin{cases} 0 & \forall i \in \mathcal{I} \ \forall j \in \mathcal{E}, \\ W_{ij} & \text{otherwise.} \end{cases}$$

Then $\mathcal{C}(W) = \mathcal{C}(W')$.

3.2. Characterization of the combinatorial code of a Dale network. Theorems 3.2 and 3.3 imply that only *excitatory − excitatory* synapses and a particular feature of the inhibitory-excitatory connectivity play a role in shaping the combinatorial code. To describe these features we give the following definitions.

DEFINITION 3.4. *Let W be a Dale matrix, and $\mathcal{E}$ and $\mathcal{I}$ denote the set of its excitatory and inhibitory neurons, respectively.*

- *An excitatory connectivity graph of W is a directed graph $G_{\mathcal{E}}$ whose vertices are the exitatory neurons $\mathcal{E}$, and whose arcs are defined as*

$$i \to j \iff W_{ji} > 0, \quad \text{where } i \neq j \in \mathcal{E}.$$

- *The uninhibited set $\mathcal{E}_{\mathcal{U}} \subset \mathcal{E}$ is the subset of excitatory neurons that* do not *receive any inhibition, i.e.,*

$$\mathcal{E}_{\mathcal{U}} \stackrel{def}{=} \{j \in \mathcal{E} \mid \forall i \in \mathcal{I}, W_{ji} = 0\}.$$

- *The inhibited set $\mathcal{E}_{\mathcal{I}} \subset \mathcal{E}$ is the subset of excitatory neurons that receive inhibition, i.e., $\mathcal{E}_{\mathcal{I}} = \mathcal{E} \setminus \mathcal{E}_{\mathcal{U}}$.*
- *For a directed graph G on a set of vertices $\mathcal{E}$ and a subset $\mathcal{E}_{\mathcal{U}} \subset \mathcal{E}$, we define a code of a pair $(G, \mathcal{E}_{\mathcal{U}})$ as*

$$\text{code}(G, \mathcal{E}_{\mathcal{U}}) \stackrel{def}{=} \left\{\sigma \subset \mathcal{E} \mid N_G^+(\sigma) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma\right\} = \{\sigma \subset \mathcal{E} \mid N_{G_{\mathcal{E}}}^+(\sigma) \setminus \sigma \subset \mathcal{E}_{\mathcal{I}}\},$$

where

$$N_G^+(\sigma) \stackrel{def}{=} \bigcup_{i \in \sigma} \{j \in \mathcal{E} \mid i \to j\}$$

denotes the out-neighborhood of a set of vertices σ in the graph G. In other words, $\text{code}(G, \mathcal{E}_{\mathcal{U}})$ *is the set of σ such that all the outneighbors of σ are inhibited (other than possibly the nodes in σ themselves). We consider an excitatory connectivity graph of some Dale network $G = G_{\mathcal{E}}$ and refer to the set of excitatory neurons $N_{G_{\mathcal{E}}}^+(\sigma)$ as the synaptic targets of a subset σ.*

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

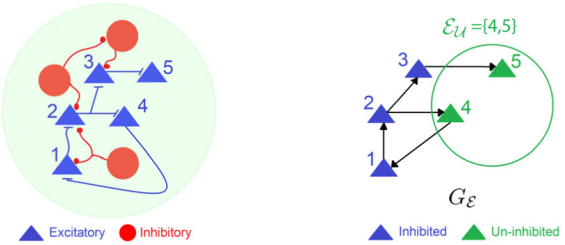


FIG. 2. *A Dale network (left) and its connectivity features (right).*

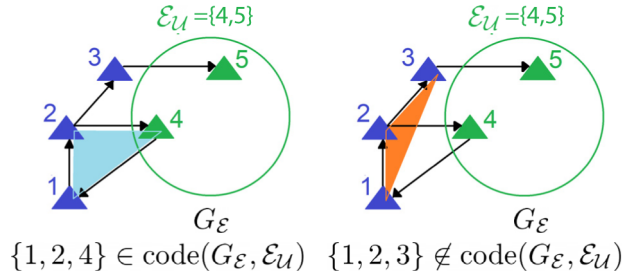


FIG. 3. *An excitatory connectivity graph of a Dale network. A codeword in* $\text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U},)$ *(left) and a codeword not in* $\text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U},)$ *(right).*

These connectivity features are illustrated on an example in Figure 2.

Intuitively, a network has a fixed point, supported at excitatory neurons $\sigma \subset \mathcal{E}$, if all the other excitatory neurons are silent. This will not occur if neurons outside of σ receive excitatory input from neurons in σ and don't have any inhibitory neurons to silence them. This is equivalent to requiring that the only unihibitted nodes that can receive excitation from σ are nodes in σ itself. This intuition is formalized in the definition of $\text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$.

The following theorem describes the combinatorial code of a Dale network.

THEOREM 3.5. *Let W be a Dale matrix, and $\sigma \subset \mathcal{E}$ be a nonempty subset of excitatory neurons. Then $\sigma \in \mathcal{C}(W)$ if and only if the following two conditions are both satisfied:*

(i) *(the spectral condition)* $\rho(W_{\mathcal{E}_\mathcal{U} \cap \sigma}) < 1$*;*

(ii) *(the graph condition)* $\sigma \in \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$*,*

where $W_{\mathcal{E}_\mathcal{U} \cap \sigma}$ denotes the synaptic weights of the excitatory subnetwork on the subset $\mathcal{E}_\mathcal{U} \cap \sigma$, and $\rho(W_{\mathcal{E}_\mathcal{U} \cap \sigma})$ denotes the spectral radius of the matrix $W_{\mathcal{E}_\mathcal{U} \cap \sigma}$.

To illustrate an application of Theorem 3.5, consider the excitatory connectivity graph of a Dale network from Figure 2. From Figure 3, we see that $N^+_{G_\mathcal{E}}(\{1,2,4\}) = \{1,2,3,4\}$ and $N^+_{G_\mathcal{E}}(\{1,2,3\}) = \{2,3,4,5\}$. Thus

$$N^+_{G_\mathcal{E}}(\{1,2,4\}) \cap \mathcal{E}_\mathcal{U} = \{4\} \subset \{1,2,4\}$$

(equivalently $N^+_{G_\mathcal{E}}(\{1,2,4\}) \setminus \{1,2,4\} = \{3\} \subset \mathcal{E}_\mathcal{I}$)), so $\{1,2,4\} \in \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$. On the other hand,

$$N^+_{G_\mathcal{E}}(\{1,2,3\}) \cap \mathcal{E}_\mathcal{U} = \{4,5\} \not\subset \{1,2,3\},$$

(equivalently, $N^+_{G_\mathcal{E}}(\{1,2,3\}) \setminus \{1,2,3\} = \{4,5\} \not\subset \mathcal{E}_\mathcal{I}$), so $\{1,2,3\} \notin \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$. Thus, since $\{1,2,4\} \in \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$, we have a *candidate* for an element in $\mathcal{C}(W)$.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Assuming one has access to the synaptic weights, the last remaining step would be checking the spectral radius condition to see if $\{1,2,4\}$ is indeed in the code $\mathcal{C}(W)$ of the network.

We say that the network (2.1) is *weakly coupled* [31, 16, 22] if the Frobenius matrix norm

$$||W||_F = \sqrt{\text{trace}(W^T W)}$$

of its synaptic matrix W is smaller than 1. It is natural to consider the weak coupling regime separately, as in this regime the spectral condition in Theorem 3.5 is always satisfied.

THEOREM 3.6. *Let W be a nonsingular Dale matrix that is weakly coupled, i.e., $||W||_F < 1$. Then every fixed point is asymptotically stable, and the combinatorial code is completely described by the graph condition*

$$\mathcal{C}(W) = \mathcal{SC}(W) = \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U}).$$

Furthermore, for all $b \in \mathbb{R}^n_{\geq 0}$ there is a unique globally exponentially stable fixed point of (2.1).

The above theorem implies that the combinatorial code of weakly coupled networks is completely determined by the connectivity features *alone,* and does not depend on the strengths of the synapses as long as the network is in the weak coupling regime. The following cautionary example illustrates two Dale matrices that are not weakly coupled. These matrices have the same connectivity, but they exhibit different eigenvalue spectra of excitatory subnetworks, resulting in different combinatorial codes. Consider Dale matrices

$$W = \begin{pmatrix} 0 & 2 & 0 & -2 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 0 & 3 & 0 & -2 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

and note that both these matrices have the same excitatory connectivity graph $G_\mathcal{E}$ and the same uninhibited excitatory neurons $\mathcal{E}_\mathcal{U} = \{2,3\}$. Let $\sigma = \{1,2,3\}$, thus $\mathcal{E}_\mathcal{U} \cap \sigma = \{2,3\}$ and observe that $\rho(W_{\mathcal{E}_\mathcal{U} \cap \sigma}) = \rho(W_{\{2,3\}}) = \sqrt{2} > 1$; thus, by Theorem 3.5, $\sigma \notin \mathcal{C}(W)$. On the other hand $\rho(U_{\mathcal{E}_\mathcal{U} \cap \sigma}) = \rho(U_{\{2,3\}}) = 0.5 < 1$. Furthermore, because $\sigma = \mathcal{E}$, the graph condition $N^+_{G_\mathcal{E}}(\sigma) \cap \mathcal{E}_\mathcal{U} \subset \sigma$ is satisfied and thus by Theorem 3.5, $\sigma \in \mathcal{C}(U)$.

3.3. Combinatorial codes of Dale recurrent networks are convex. Perhaps the most surprising implication of Theorem 3.5 is that Dale networks naturally produce convex combinatorial codes. To explain the relevant background, we first motivate and define convex combinatorial codes.

There are two complementary viewpoints on what determines neural activity in sensory systems. One viewpoint is that the brain represents information via patterns of neural activity that arise as a result of neural dynamics. A different perspective is that the neural activity in sensory neural systems is induced by external stimuli, whereby each neuron responds to a given stimulus according to its own receptive field. If these two views are both correct, then the patterns of neural activity resulting from the dynamics should be compatible with those determined by the intersection patterns of the receptive fields.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

We define a receptive field of an individual neuron as a subset $U \subset \mathbb{R}^d$ in a stimulus space $\mathbb{R}^d$, such that the firing rate $x(t)$ of a given neuron is nonzero at the times when the stimulus is in the region U and the firing rate is 0 (or below some threshold) when the stimulus is outside U. Given a collection of receptive fields $\mathcal{U} = \{U_i\}$ of a population of neurons $\mathcal{E}$, consider *the code* of $\mathcal{U}$ that describes all possible intersection patterns of the receptive fields U_i as

$$\operatorname{code}(\mathcal{U}) \stackrel{\text{def}}{=} \{\sigma \subset \mathcal{E} \,|\, R_\sigma \neq \varnothing\},$$

where

$$R_\sigma \stackrel{\text{def}}{=} \left(\bigcap_{i\in\sigma} U_i\right) \setminus \bigcup_{j\notin\sigma} U_j$$

denotes the region in the stimulus space where each of the neurons in σ is activated and no other considered neuron is active, and $R_\varnothing \stackrel{\text{def}}{=} \mathbb{R}^d \setminus \bigcup_{j\in\mathcal{E}} U_j$. The regions R_σ partition the stimulus space $\mathbb{R}^d$ (see Figure 4).

Note that $\operatorname{code}(\mathcal{U})$ is the set of all possible patterns of neuronal activation that are compatible with the receptive fields $\mathcal{U}$. The same patterns should be allowed by the neural network dynamics. We associate a combinatorial neural code $\mathcal{C}(W)$ with a Dale network (2.1), and we interpret the fixed points of the excitatory neurons, *both stable and unstable,* as our model for transiently activated stimulus representations. This is because the neural activity spends a significant amount of time not only in the neighborhood of stable fixed points, but also around saddle fixed points, while evolving along heteroclinic trajectories (see, e.g., [17, 3, 27]). Following these interpretations we posit that the combinatorial code of the network must be the same as the intersection patterns of the receptive fields:

$$\mathcal{C}(W) = \operatorname{code}(\mathcal{U}).$$

Recall that a set $U \subset \mathbb{R}^d$ is called convex if for any two points $x, y \in U$ the line segment $[x, y]$ also lies in U. A number of sensory areas in the brain possess convex receptive fields; a very incomplete list of such areas includes the hippocampus, the primary visual cortex, the primary auditory cortex, etc. Following [9, 5, 19, 8], we define convex combinatorial codes as follows.

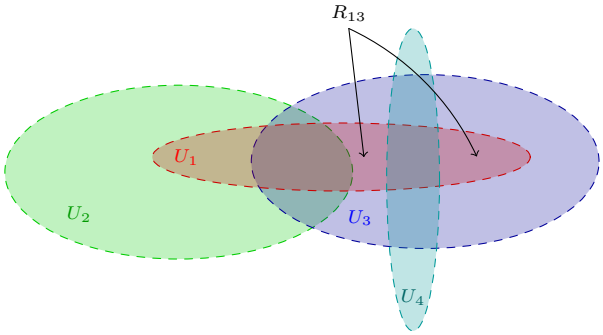


FIG. 4. *An example of a collection of receptive fields* $\mathcal{U} = \{U_1, U_2, U_3, U_4\}$ *and its code* $\mathcal{C} = \operatorname{code}(\mathcal{U}) = \{\varnothing, 2, 3, 4, 12, 13, 23, 34, 123, 134\}$. *Here we denote a codeword* $\{i_1, i_2, \ldots, i_k\} \in \mathcal{C}$ *by the string* $i_1 i_2 \ldots i_k$*; for example,* $\{1, 3, 4\}$ *is abbreviated to* 134. *Note that all* U_i *are convex (and thus connected), but the regions* R_σ *are typically nonconvex for nonmaximal subsets* σ. *For example* R_{13} *is neither convex nor connected.*

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

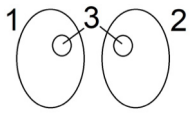


FIG. 5. *A nonconvex code* $\mathcal{C} = \{1, 2, 13, 23\}$.

DEFINITION 3.7. *A combinatorial code* $\mathcal{C} \subset 2^{[n]}$ *is* open convex *if* $\mathcal{C} = code(\mathcal{U})$ *for a collection* $\mathcal{U} = \{U_i\}$ *of open convex subsets in a Euclidean space* $\mathbb{R}^d$ *for some* $d \geq 1$.

It has been previously established in [9, 13, 5] that not every combinatorial code is open convex. The smallest example of a nonconvex code is illustrated in Figure 5. The existence of the codewords 13 and 23 without 123 guarantee that the receptive field U_3 must be disconnected and thus nonconvex. The topological properties that prevent a combinatorial code from being convex were studied in [5, 8, 21]. In fact, a randomly chosen[1] combinatorial code on a large number of neurons is nonconvex with a high probability.

Following [5], recall that a combinatorial code $\mathcal{C}$ is called *intersection-complete,* if the intersection of any two codewords in $\mathcal{C}$ also belongs to $\mathcal{C}$:

$$\sigma, \nu \in \mathcal{C} \implies \sigma \cap \nu \in \mathcal{C}.$$

If a code is intersection-complete and also has the property that a union of two codewords is a codeword, then such code is a *sublattice* of $2^{\mathcal{E}}$ [29]. It has been previously established in [5] that intersection-complete codes are open convex.

THEOREM 3.8 (see [5]). *Every intersection-complete code* $\mathcal{C} \subset 2^{[n]}$ *is open convex.*

We use this result to show that the combinatorial codes of a Dale network are convex. Consider a directed graph $G_{\mathcal{E}}$, whose vertices are the excitatory neurons $\mathcal{E}$, and whose edges are derived from the excitatory connectivity as in Definition 3.4. First we observe the following.

PROPOSITION 3.9. *The code* $\text{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$ *is a sublattice of the Boolean lattice* $2^{\mathcal{E}}$.

This observation, combined with Theorem 3.6 implies that the combinatorial code of a weakly coupled network is a sublattice. If the network is not weakly coupled, then the spectral condition in Theorem 3.5 may prevent the code of the network from being a sublattice; however, the code $\mathcal{C}(W)$ of a Dale network remains intersection-complete.

THEOREM 3.10. *For a Dale synaptic matrix* W*, its code* $\mathcal{C}(W)$ *is intersection-complete and is thus convex.*

The proof of this theorem is given in subsection 5.2. We suggest that this theorem may provide a new explanation for the prevalence of convex receptive fields of excitatory neurons in recurrent circuits of sensory systems.

3.4. Constructing a Dale network from a combinatorial code. Can one encode an arbitrary combinatorial code as the combinatorial code of a Dale network? Theorem 3.10 implies that the code needs to be intersection-complete to be realized on a generic Dale network. In addition, Proposition 3.9 tells us that if one wants to build a weakly coupled network with a prescribed code, this code needs to be a

[1]This, of course, requires a proper definition of the probability distribution on the set of all codes that we omit here.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

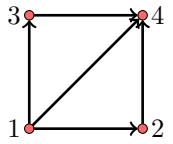


FIG. 6. *The graph* G^c *on four vertices with* $\mathrm{code}(G^c,[4]) = \{\varnothing,\{4\},\{2,4\},\{3,4\},\{2,3,4\},[4]\}$.

sublattice. It turns out that any sublattice can be encoded on a weakly coupled Dale network.

THEOREM 3.11. *Given a sublattice* $\mathcal{C} \subset 2^{\mathcal{E}}$ *with* $\varnothing, \mathcal{E} \in \mathcal{C}$*, define a map* $c : 2^{\mathcal{E}} \to \mathcal{C}$ *as*

$$c(\sigma) \stackrel{def}{=} \bigcap_{\substack{\nu \in \mathcal{C}, \\ \nu \supset \sigma}} \nu, \tag{3.1}$$

and consider a directed graph G^c *whose vertices are* $\mathcal{E}$ *and whose edges are defined via the rule*

$$i \to j \iff j \in c(\{i\}) \text{ and } i \neq j. \tag{3.2}$$

Then $\mathcal{C} = \mathrm{code}(G^c, \mathcal{E})$.

Note that $\mathrm{code}(G^c, \mathcal{E})$ from Theorem 3.11 is realized by a network with no inhibition. This theorem, combined with Theorem 3.6 translates into the following.

COROLLARY 3.12. *Given a sublattice* $\mathcal{C} \subset 2^{\mathcal{E}}$ *with* $\varnothing, \mathcal{E} \in \mathcal{C}$*, one can find a weakly coupled network* W *of excitatory neurons with* $\mathcal{C} = \mathcal{C}(W)$.

This "*learning rule*" for an excitatory synaptic matrix W amounts to assigning small nonzero excitatory synaptic weights according to the directed arcs in the graph G^c defined in (3.1), (3.2).

The following example illustrates the procedure described above to construct a graph of connectivity from a given code. Consider a sublattice $C = \{\varnothing, \{4\}, \{2,4\}, \{3,4\}, \{2,3,4\}, \{1,2,3,4\}\}$. One can directly compute that $c(\{1\}) = \{1,2,3,4\}$, $c(\{2\}) = \{2,4\}$, $c(\{3\}) = \{3,4\}$, and $c(\{4\}) = \{4\}$. It is then straightforward to verify that G^c is the graph in Figure 6 and thus $C = \mathrm{code}(G^c, \mathcal{E}_{\mathcal{U}} = \{1,2,3,4\})$.

Remark 3.13. This construction is common and known in finitely generated topologies or, equivalently, *Alexandroff topological spaces* [2]. Note that $\varnothing, \mathcal{E} \in C$ and C being a lattice mean that C is a collection of closed sets of a topology on $\mathcal{E}$. In particular, the code $\mathcal{C}$ is the closed sets of a topology on a finite set. The digraph we obtained is also known as the *specialization preorder*,[2] associated with the topology of $\mathcal{C}$.

4. Discussion. We investigated threshold-linear recurrent networks that satisfy Dale's law and established which features of the synaptic connectivity are responsible for determining their combinatorial codes. Theorem 3.5 describes the combinatorial code in terms of the connectivity graph and the spectral radii of the appropriate submatrices. In the case of weakly coupled networks, the spectral conditions are always

[2]Technically, the digraph we would obtain would have all self-loops as it is a preorder and hence a reflexive relation on the finite set. However, we ignore the loops in our construction which is not a problem in our context.

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

satisfied and the combinatorial code depends only on the features in code$(G, \mathcal{E}_\mathcal{U})$ that are derived from the connectivity graph with a given set of uninhibited neurons $\mathcal{E}_\mathcal{U}$ (Theorem 3.6). In this situation, all the fixed points are stable; moreover, the combinatorial code is a sublattice. We also proved that any sublattice can be encoded as a combinatorial code of an excitatory network (Corollary 3.12).

If the network is not weakly coupled, the spectral condition in Theorem 3.5 may prevent the combinatorial code from being a sublattice. Intuitively, this is because the union of two codewords in $\mathcal{C}(W)$ may not be in $\mathcal{C}(W)$, as the spectral radius of the appropriate larger matrix may (or may not) increase. We have proven in Theorem 3.10 that every combinatorial code is intersection-complete and therefore convex. Can any intersection-complete code be encoded on a Dale network? The answer to this question is currently unknown, as it requires a better understanding of the interplay of the spectral and the graph conditions in Theorem 3.5.

Finally, we hypothesize that the result (Theorem 3.10) that the combinatorial codes of Dale recurrent networks are always convex may provide a natural explanation for the ubiquity of convex receptive fields in many sensory systems in the mammalian brain.

5. Proofs. In this section we give proofs for our main results stated in section 3. First, we recall the conditions for having a fixed point of the dynamics of (2.1).

LEMMA 5.1 (see [6, Proposition 2.1]). *For a threshold linear network in* (2.1)*, characterized by* (W, b)*, a point* $x^* \in \mathbb{R}^n$ *is a fixed point with support* $\sigma \subset [n]$ *if and only if the following conditions are all satisfied:*

1. $(I - W)_\sigma x^*_\sigma = b_\sigma$*,;*
2. $x^*_\sigma > 0$*; and*
3. $b_{\overline{\sigma}} \leq -W_{\overline{\sigma}\sigma} x^*_\sigma$*.*

Here b_σ *is the restriction of a vector* b *to the indices in* σ*,* $\overline{\sigma}$ *denotes the complement in* $[n]$ *to a subset* σ*,* $W_{\overline{\sigma}\sigma}$ *is the rectangular submatrix of* W*, restricted to the rows in* $\overline{\sigma}$ *and columns in* σ*, and* $(I - W)_\sigma = (I - W)_{\sigma\sigma}$ *is the principal submatrix restricted to the subset* σ*. Note that if* $(I - W)_\sigma$ *is invertible, then there is at most one fixed point with support* σ *and it is given by* $x^*_\sigma = (I_\sigma - W_\sigma)^{-1} b_\sigma$*.*

The conditions for having a fixed point in Lemma 5.1 follow (with some work) from the dynamical system in (2.1). In particular, $(I - W)_\sigma x^*_\sigma = b_\sigma$ follows from needing to evaluate the ReLU nonlinearity to a positive number, and $b_{\overline{\sigma}} \leq -W_{\overline{\sigma}\sigma} x^*_\sigma$ follows from evaluating the ReLU nonlinearity to 0.

A key ingredient in our proofs is the observation that the first condition in Lemma 5.1 can be translated to the language of semipositive matrices (Definition 5.5 below). From there we make use of the theory of M-matrices (Definition 5.2 below). Note that the machinery of semipositive matrices has been previously used in a different context of recurrent networks with a Heaviside transfer function in [30].

5.1. Necessary matrix theory results. First, we recall some definitions and results from the theory of M-matrices, following [4]. An $m \times n$ matrix A is *positive* (*nonnegative*) if $A_{ij} > 0$ ($A_{ij} \geq 0$) for all i and j. If A is positive (nonnegative) we denote this by $A > 0$ ($A \geq 0$). For a vector $x \in \mathbb{R}^n$, we write $x > 0$, $x < 0$, $x \geq 0$, or $x \leq 0$ if all the entries of corresponding vector x satisfy the appropriate inequality.

DEFINITION 5.2 (see [4, Chapter 6]). *A square matrix* A *is a* Z-matrix *if* $A_{ij} \leq 0$ *for all* $i \neq j$*. A* Z*-matrix* A *is called an* M-matrix *if* $A = sI - B$*, where* I *denotes the identity matrix,* $B_{ij} \geq 0$ *for all* $i \neq j$*, and the scalar* s *is not smaller than the spectral radius of the matrix* B*,* $s \geq \rho(B)$*.*

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

The following lemma will be used later.

LEMMA 5.3 (see [4, Lemma 2.1]). *Assume that $A = I - W$, where $W \geq 0$. Then the following two statements are equivalent:*

(i) *The matrix A is a nonsingular M-matrix.*

(ii) *The spectral radius of W is smaller than one: $\rho(W) < 1$.*

We shall also make use of the following.

LEMMA 5.4 (see [4, Chapter 2, Corollary 1.6]). *Let A be a square nonnegative matrix. Suppose that B is a principal submatrix of A. Then $\rho(B) \leq \rho(A)$.*

DEFINITION 5.5. *A square matrix A is* semipositive *if $\exists x > 0$ such that $Ax > 0$.*

LEMMA 5.6. *Let A be a nonsingular $n \times n$ matrix. Then A is semipositive if and only if there exists a vector $x > 0$ such that $Ax \geq 0$.*

Proof. The forward direction is immediate. For the converse suppose that $\exists x > 0$ such that $Ax \geq 0$. Let $a = ||A^{-1}||_{op}$ be the operator norm of A^{-1}, $x_m = \min_{1 \leq i \leq n}(x_i)$, $\varepsilon = \frac{x_m}{2a\sqrt{n}}$, and $\vec{\varepsilon} \in \mathbb{R}^n$ be the column vector whose components are all equal to ε. Then $||\vec{\varepsilon}||_2 = \frac{x_m}{2a}$ by construction. Furthermore, $||A^{-1}\vec{\varepsilon}||_2 \leq a||\vec{\varepsilon}||_2 = a\frac{x_m}{2a} = \frac{x_m}{2}$. In particular, $\sum_{i=1}^{n}(A^{-1}\vec{\varepsilon})_i^2 \leq \frac{x_m^2}{4}$ for all $1 \leq i \leq n$. Therefore, $(A^{-1}\vec{\varepsilon})_i^2 \leq \frac{x_m^2}{4}$ which implies $|(A^{-1}\vec{\varepsilon})_i| \leq \frac{x_m}{2}$ for all $1 \leq i \leq n$. Thus, $x + A^{-1}\vec{\varepsilon} > 0$ because of the way x_m was defined. Finally, $A(x + A^{-1}\vec{\varepsilon}) = Ax + \vec{\varepsilon} \geq 0 + \vec{\varepsilon} > 0$ and thus A is semipositive. ∎

The following example illustrates the necessity of the hypothesis that A is a nonsingular matrix in Lemma 5.6. Consider a singular matrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $x = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$. Then note that $Ax \geq 0$. Now let $y \in \mathbb{R}^3_{>0}$. Suppose that $Ay > 0$. This implies that $y_1 > y_2$ and $y_2 > y_1$ which cannot be. Thus A is not semipositive.

There are numerous characterizations of M-matrices; one can find several dozen conditions on a matrix that are all equivalent to being an M-matrix [4]. However, our proofs rely only on the following observation.

THEOREM 5.7 (see [4, Chapter 6, Theorem 2.3]). *Let A be a Z-matrix. Then A is a nonsingular M-matrix if and only if A is semipositive.*

5.2. Proof of Theorem 3.2 and related results. In order to prove Theorem 3.2 a few preliminary results are necessary. Namely, Theorem 5.9 and Corollary 5.10. We also introduce the following notation for convenience.

DEFINITION 5.8. *For the dynamical system in* (2.1) *we denote the set of supports of all the fixed points by $FP(W, b)$ and we let $FP(W) \stackrel{def}{=} \bigcup_{b \geq 0} FP(W, b)$.*

THEOREM 5.9. *Let $W \in \mathbb{D}_n$ and $W' \in \mathbb{D}_{n+1}$ be Dale matrices such that $W'_{[n]} = W$, and the last $(n+1)$th column of W' is inhibitory and satisfies the following condition:*

$$\exists i \in \mathcal{E} \text{ with } W'_{i(n+1)} < 0 \quad \Longrightarrow \quad \exists j \in \mathcal{I} \text{ with } W_{ij} < 0. \tag{5.1}$$

Then $\mathcal{C}(W) = \mathcal{C}(W')$.

Proof. The proof needs that both matrices W and W' satisfy Ground Assumption 3.1. Since being singular is a fine-tuned condition for matrices, the assumption is generically true. Furthermore, we can explicitly construct a W' in such a way that it satisfies Ground Assumption 3.1 and the hypothesis in the statement of the theorem.

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

In particular we can choose the $(n+1)$th row of W' to be the zero vector. This will make it so that for any $\sigma \subset [n+1]$, $(I - W')_\sigma$ is nonsingular, assuming that W was already satisfying Ground Assumption 3.1. We now proceed with the proof.

Observe that by construction, as we are only adding an additional inhibitory neuron to the network, both W and W' have the same collection of excitatory neurons, $\mathcal{E}$. Furthermore, $\mathcal{E}_\mathcal{I} \subset \mathcal{E}_\mathcal{I}{}'$, where the $\mathcal{E}_\mathcal{I}{}'$ are the excitatory neurons in W' that receive inhibition. However, (5.1) in fact yields us $\mathcal{E}_\mathcal{I}{}' \subset \mathcal{E}_\mathcal{I}$. Therefore $\mathcal{E}_\mathcal{I} = \mathcal{E}_\mathcal{I}{}'$, that is to say, the collection of inhibited and uninhibited excitatory neurons in W and W' are the same. Hence, there is no confusion when we write the excitatory neurons in W' as a partition $\mathcal{E} = \mathcal{E}_\mathcal{U} \sqcup \mathcal{E}_\mathcal{I}$.

Suppose that $\sigma \in \mathcal{C}(W)$. Then, there exists $\tau \subset \mathcal{I}$ such that $\nu = \sigma \sqcup \tau \in FP(W)$, and (by Lemma 5.1) a vector $b \in \mathbb{R}^n_{\geq 0}$ such that $x_\nu = (I_\nu - W_\nu)^{-1} b_\nu > 0$ and $0 \leq b_{\overline{\nu}} \leq -W_{\overline{\nu}\nu} x_\nu$, where $\overline{\nu} = [n] \setminus \nu$. Let $\nu' = \nu \sqcup \{n+1\}$ and define $y \in \mathbb{R}^{|\nu|+1}_{>0}$ as follows. The first $|\nu|$ entries of y are those of x_ν, i.e., $y_\nu = x_\nu \in \mathbb{R}^{|\nu|}_{>0}$, and the last entry is defined as

$$y_{|\nu|+1} = \left| \sum_{k \in \nu} (I - W')_{(n+1)k} x_k \right| + 1.$$

Note that for $i \in \nu$, we have

$$((I - W')_{\nu'} y)_i = \sum_{k \in \nu'} (I - W')_{ik} y_k = \sum_{k \in \nu} (I - W)_{ik} x_k - W'_{i(|\nu|+1)} y_{|\nu|+1} \geq 0,$$

since $-W'_{i(|\nu+1)} \geq 0$, and x was such that $(I_\nu - W_\nu) x_\nu = b_\nu \geq 0$. Thus, by construction we have $(I - W')_{\nu'} y \geq 0$. Furthermore, $-W'_{\overline{\nu'}\nu'} y \geq -W_{\overline{\nu}\nu} x_\nu \geq 0$. Let $b' \in \mathbb{R}^{n+1}_{\geq 0}$ be defined by $b'_{\nu'} = (I - W')_{\nu'} y$ and $b'_{\overline{\nu'}} = 0$. Then, by Lemma 5.1 $\nu' = \sigma \sqcup \tau \cup \{n+1\} \in FP(W', b') \subset FP(W')$, and thus $\sigma \in \mathcal{C}(W')$.

Now suppose that $\sigma \in \mathcal{C}(W')$. Then there exists a $\tau' \subset \mathcal{I} \sqcup \{n+1\}$ such that $\nu = \sigma \sqcup \tau' \in FP(W')$. Since $\nu \in FP(W')$, there exists a $b' \in \mathbb{R}^{n+1}_{\geq 0}$ such that

$$x \overset{\text{def}}{=} (I_\nu - W'_\nu)^{-1} b'_\nu > 0 \text{ and } 0 \leq b'_{[n+1]\setminus\nu} \leq -W'_{([n+1]\setminus\nu)\nu} x,$$

by Lemma 5.1. We have two cases to consider: either $\tau' \subset \mathcal{I}$ or $(n+1) \in \tau'$.

In the case that $\tau' \subset \mathcal{I}$, observe that $(I_\nu - W'_\nu)^{-1} = (I_\nu - W_\nu)^{-1}$ and that $W'_{([n+1]\setminus\nu)\nu}$ is almost identical to $W_{([n]\setminus\nu)\nu}$; $W'_{([n+1]\setminus\nu)\nu}$ can be obtained from $W_{([n]\setminus\nu)\nu}$ by adding an extra row of entries induced from the $(n+1)$th row of W'. Define $b \in \mathbb{R}^n_{\geq 0}$ by $b_\nu = (I - W)_\nu x = (I - W')_\nu x = b'_\nu$ and $b_{[n]\setminus\nu} = 0$. Then by construction

$$x = (I_\nu - W_\nu)^{-1} b_\nu > 0 \text{ and } 0 \leq b_{[n]\setminus\nu} \leq -W_{([n]\setminus\nu)\nu} x$$

and, thus, $\nu = \sigma \sqcup \tau' \in FP(W, b) \subset FP(W)$ by Lemma 5.1. Therefore $\sigma \in \mathcal{C}(W)$.

In the case that $(n+1) \in \tau'$, let $\mathcal{E}_{\tau'}$ be the set of indices of nonzero rows of $W'_{\mathcal{E}\tau'}$ (the set of inhibited excitatory neurons of the Dale network $W'_{\mathcal{E}_{\tau'}}$). For each $i \in \alpha$ let $j_i \in \mathcal{I}$ be such that $W'_{ij_i} = W_{ij_i} < 0$ (we can always choose j_i to not equal $n+1$ by the hypothesis of the theorem, and thus we can assume $j_i \in \mathcal{I}$). Let $\mu = \sigma \sqcup \mathcal{I}$ and define

$$c \overset{\text{def}}{=} \max \left\{ \max_{i \in \sigma \cap \mathcal{E}_{\tau'}} -\frac{1}{W_{ij_i}} \left| \sum_{k \in \sigma} (I - W)_{ik} x_k \right|, \max_{i \in \mathcal{I}} \left| \sum_{k \in \sigma} -W_{ik} x_k \right|, \right.$$
$$\left. \max_{i \in ([n]\setminus\mu) \cap \alpha} -\frac{1}{W_{ij_i}} \left| \sum_{k \in \sigma} -W_{ik} x_k \right| \right\}.$$

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Let $y \in \mathbb{R}_{>0}^{|\mu|}$ be defined by $y_\sigma = x_\sigma$ and $y_j = 1 + c$ for $j \in \mathcal{I}$. We aim to show that $(I - W)_\mu y > 0$ and that $0 \leq -W_{([n]\setminus\mu)\mu} y$ as that will give us that $\mu \in FP(W)$ by Lemma 5.1. We prove these inequalities by considering the partitions

$$\mu = \sigma \sqcup \mathcal{I} = (\sigma \cap \mathcal{E}_{\tau'}) \sqcup (\sigma \setminus \mathcal{E}_{\tau'}) \sqcup \mathcal{I}, [n] \setminus \mu = (([n] \setminus \mu) \cap \mathcal{E}_{\tau'}) \sqcup (([n] \setminus \mu) \setminus \mathcal{E}_{\tau'}).$$

Then for $i \in \sigma \cap \mathcal{E}_{\tau'}$, we have

$$\begin{aligned}
((I - W)_\mu y)_i &= \sum_{k \in \mu} (I - W)_{ik} y_k = \sum_{k \in \sigma} (I - W)_{ik} y_k + \sum_{k \in \mathcal{I}} (I - W)_{ik} y_k \\
&= \sum_{k \in \sigma} (I - W)_{ik} x_k + (1 + c) \sum_{k \in \mathcal{I}} (I - W)_{ik} \\
&= \sum_{k \in \sigma} (I - W)_{ik} x_k + (1 + c)(I - W)_{ij_i} + (1 + c) \sum_{j_i \neq k \in \mathcal{I}} (I - W)_{ik} \\
&\geq \sum_{k \in \sigma} (I - W)_{ik} x_k + c(-W_{ij_i}) - W_{ij_i} + (1 + c)0 \\
&\geq \sum_{k \in \sigma} (I - W)_{ik} x_k + (-W_{ij_i})(-\frac{1}{W_{ij_i}}) \left| \sum_{k \in \sigma} (I - W)_{ik} x_k \right| - W_{ij_i} \\
&\geq 0 - W_{ij_i} > 0.
\end{aligned}$$

For $i \in \sigma \setminus \mathcal{E}_{\tau'}$, we have

$$\begin{aligned}
((I - W)_\mu y)_i &= \sum_{k \in \mu} (I - W)_{ik} y_k = \sum_{k \in \sigma} (I - W)_{ik} y_k + \sum_{k \in \mathcal{I}} (I - W)_{ik} y_k \\
&= \sum_{k \in \sigma} (I - W)_{ik} x_k + \sum_{k \in \mathcal{I}} (I - W)_{ik} y_k \geq \sum_{k \in \sigma} (I - W)_{ik} x_k + 0 > 0,
\end{aligned}$$

because the ith row of $W'_{\mathcal{E}_{\tau'}}$ is zero for $i \notin \mathcal{E}_{\tau'}$ by assumption and x was by hypothesis such that for $i \in \sigma \setminus \mathcal{E}_{\tau'}$, $((I - W')_\nu x)_i = \sum_{k \in \nu} (I - W')_{ik} x_k = \sum_{k \in \sigma} (I - W')_{ik} x_k + \sum_{k \in \tau'} (I - W')_{ik} x_k = \sum_{k \in \sigma} (I - W)_{ik} x_k + 0 > 0$.

Let $i \in \mathcal{I}$. Since $i \notin \sigma$, we have $\sum_{k \in \sigma} (I - W)_{ik} x_k = \sum_{k \in \sigma} -W_{ik} x_k$. Furthermore, note that $\sum_{i \neq k \in \mathcal{I}} -W_{ik} \geq 0$. From all of this we have

$$\begin{aligned}
((I - W)_\mu y)_i &= \sum_{k \in \mu} (I - W)_{ik} y_k = \sum_{k \in \sigma} (I - W)_{ik} y_k + \sum_{k \in \mathcal{I}} (I - W)_{ik} y_k \\
&= \sum_{k \in \sigma} (I - W)_{ik} x_k + (1 + c) \sum_{k \in \mathcal{I}} (I - W)_{ik} \\
&= \sum_{k \in \sigma} -W_{ik} x_k + (1 + c)((I - W)_{ii} + \sum_{i \neq k \in \mathcal{I}} -W_{ik}) \\
&\geq \sum_{k \in \sigma} -W_{ik} x_k + 1 + c \geq \sum_{k \in \sigma} -W_{ik} x_k + 1 + \left| \sum_{k \in \sigma} -W_{ik} x_k \right| > 0.
\end{aligned}$$

Together, all these inequalities yield that $(I - W)_\mu y > 0$ and $(I_\mu - W_\mu)^{-1}(I - W)_\mu y = y > 0$.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

On the other hand, for $i \in ([n] \setminus \mu) \cap \mathcal{E}_{\tau'}$, we have

$$\begin{aligned}(-W_{([n]\setminus\mu)\mu}y)_i &= (-W_{([n]\setminus\mu)\mu}y)_i = \sum_{k\in\mu} -W_{ik}y_k \\ &= \sum_{k\in\sigma} -W_{ik}y_k + \sum_{k\in\mathcal{I}} -W_{ik}y_k = \sum_{k\in\sigma} -W_{ik}x_k + \sum_{k\in\mathcal{I}} -W_{ik} + c\sum_{k\in\mathcal{I}} -W_{ik} \\ &= \sum_{k\in\sigma} -W_{ik}x_k - W_{ij_i} + \sum_{j_i\neq k\in\mathcal{I}} -W_{ik} + c(-W_{ij_i}) + c\sum_{j_i\neq k\in\mathcal{I}} -W_{ik} \\ &\geq \sum_{k\in\sigma} -W_{ik}x_k - W_{ij_i} + 0 + (-W_{ij_i})\left(-\frac{1}{W_{ij_i}}\right)\left|\sum_{k\in\sigma} -W_{ik}x_k\right| + 0 \\ &\geq -W_{ij_i} > 0.\end{aligned}$$

Furthermore, for $i \in ([n] \setminus \mu) \setminus \mathcal{E}_{\tau'}$, we have

$$\begin{aligned}(-W_{([n]\setminus\mu)\mu}y)_i &= \sum_{k\in\mu} -W_{ik}y_k \\ &= \sum_{k\in\sigma} -W_{ik}y_k + \sum_{k\in\mathcal{I}} -W_{ik}y_k \geq \sum_{k\in\sigma} -W_{ik}x_k + 0 \geq 0,\end{aligned}$$

because the ith row of $W'_{\mathcal{E}_{\tau'}}$ is zero for $i \notin \mathcal{E}_{\tau'}$ by assumption and x was by hypothesis such that for $i \in ([n] \setminus \mu) \setminus \mathcal{E}_{\tau'} \subset \mathcal{E}$,

$$(-W'_{([n+1]\setminus\nu)\nu}x)_i = \sum_{k\in\nu} -W'_{ik}x_k = \sum_{k\in\sigma} -W'_{ik}x_k + \sum_{k\in\tau'} -W'_{ik}x_k = \sum_{k\in\sigma} -W_{ik}x_k + 0 \geq 0.$$

Let $b \in \mathbb{R}^n_{\geq 0}$ be defined by $b_\mu = (I - W)_\mu y$ and $b_{[n]\setminus\mu} = 0$. Then, by construction $\mu = \sigma \sqcup \mathcal{I} \in FP(W, b) \subset FP(W)$ by Lemma 5.1. Therefore $\sigma \in \mathcal{C}(W)$. ∎

COROLLARY 5.10. *Let $W \in \mathbb{D}_n$, and $W' \in \mathbb{D}_{n+1}$ be Dale matrices such that $W'_{[n]} = W$, and the last $(n+1)$th column of W' is inhibitory and such that for all $i \in \mathcal{E}$, $W_{i(n+1)} < 0$ if and only if the ith row vector of $W_{\mathcal{E}\mathcal{I}}$ is nonzero. Let W'' be a Dale matrix obtained by deleting all the inhibitory columns and rows of W' except the $(n+1)$th row and column. Then $\mathcal{C}(W) = \mathcal{C}(W'')$.*

Proof. By Theorem 5.9 we have that $\mathcal{C}(W) = \mathcal{C}(W')$. Now we can reindex the inhibitory columns and rows so that the $(n+1)$th column of W' becomes the first inhibitory column. Then we start deleting all the other inhibitory rows and columns one at a time and at the end we obtain the matrix W''. The code will be preserved at each deletion step by Theorem 5.9 and thus $\mathcal{C}(W) = \mathcal{C}(W'')$. Note that W' can be generically chosen such that at each deletion step all relevant matrices satisfy Ground Assumption 3.1, which was necessary in the proof of Theorem 5.9. ∎

Proof of Theorem 3.2. Note that we can obtain W' from W by following the procedure from Corollary 5.10 where we obtained W''. The result thus follows. ∎

Before proceeding to prove Theorem 3.3, we make a few observations in Proposition 5.11 and Corollary 5.12 which will simplify our proof strategy.

PROPOSITION 5.11. *Let $W \in \mathbb{D}_n$, $\sigma \subset \mathcal{E}$ and $\tau \subset \mathcal{I}$. If $\sigma \sqcup \tau \in FP(W)$, then $\sigma \sqcup \tau' \in FP(W)$ for any $\tau \subset \tau' \subset \mathcal{I}$.*

Proof. Suppose that $\nu = \sigma \sqcup \tau \in FP(W)$ and let $\tau \subset \tau' \subset \mathcal{I}$, $\nu' = \sigma \sqcup \tau'$. Then, there exists a $b \in \mathbb{R}^n_{\geq 0}$ such that

$$x \stackrel{\text{def}}{=} (I_\nu - W_\nu)^{-1} b_\nu > 0 \text{ and } 0 \leq b_{\overline{\nu}} \leq -W_{\overline{\nu}\nu}x$$

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

by Lemma 5.1. Define

$$c \overset{\text{def}}{=} \max_{i \in \tau' \setminus \tau} \left| \sum_{k \in \sigma} (I - W)_{ik} x_k \right|.$$

Let $x' \in \mathbb{R}^{|\nu'|}_{>0}$ be defined by $x'_\nu = x$ and $x'_i = 1 + c$ for $i \in \nu' \setminus \nu$. Then for $i \in \nu$ we have

$$\begin{aligned}((I - W)_{\nu'} x')_i &= \sum_{k \in \nu'} (I - W)_{ik} x'_k \\ &= \sum_{k \in \nu} (I - W)_{ik} x'_k + \sum_{k \in \nu' \setminus \nu} (I - W)_{ik} x'_k \\ &\geq \sum_{k \in \nu} (I - W)_{ik} x_k + 0 = b_i \geq 0.\end{aligned}$$

For $i \in \nu' \setminus \nu$ we have

$$\begin{aligned}&((I - W)_{\nu'} x')_i = \sum_{k \in \nu'} (I - W)_{ik} x'_k \\ &= \sum_{k \in \sigma} (I - W)_{ik} x'_k + \sum_{k \in \tau} (I - W)_{ik} x'_k + \sum_{k \in \tau' \setminus \tau} (I - W)_{ik} x'_k \\ &= \sum_{k \in \sigma} (I - W)_{ik} x'_k + 0 + (I - W)_{ii} x'_i + \sum_{i \neq k \in \tau' \setminus \tau} (I - W)_{ik} x'_k \\ &\geq \sum_{k \in \sigma} (I - W)_{ik} x_k + 1 + c + 0 \geq \sum_{k \in \sigma} (I - W)_{ik} x_k + 1 + \left| \sum_{k \in \sigma} (I - W)_{ik} x_k \right| \geq 1 > 0.\end{aligned}$$

Furthermore, for $i \in \overline{\nu'} \subset \overline{\nu}$ we have

$$(-W_{\overline{\nu'} \nu'} x')_i = \sum_{k \in \nu'} -W_{ik} x'_k = \sum_{k \in \nu} -W_{ik} x'_k + \sum_{k \in \tau' \setminus \tau} -W_{ik} x'_k \geq \sum_{k \in \nu} -W_{ik} x_k + 0 \geq b_i \geq 0.$$

Define $b' \in \mathbb{R}^n_{\geq 0}$ by $b'_{\nu'} = (I - W)_{\nu'} x'$ and $b'_{\overline{\nu'}} = 0$. Then by construction and by Lemma 5.1, $\nu' = \sigma \sqcup \tau' \in FP(W, b') \subset FP(W)$. ∎

COROLLARY 5.12. *Let $W \in \mathbb{D}_n$. Then $\sigma \in \mathcal{C}(W)$ if and only if $\sigma \sqcup \mathcal{I} \in FP(W)$.*

We now prove Theorem 3.3, reprinted below for convenience.

THEOREM 3.3. *Let $W \in \mathbb{D}_n$ be a Dale matrix with a set of excitatory neurons $\mathcal{E}$ and inhibitory neurons $\mathcal{I}$, and let $W' \in \mathbb{D}_n$ be such that*

$$W'_{ij} = \begin{cases} 0 & \forall i \in \mathcal{I} \ \forall j \in \mathcal{E}, \\ W_{ij} & \textit{otherwise.} \end{cases}$$

Then $\mathcal{C}(W) = \mathcal{C}(W')$.

Proof of Theorem 3.3. The proof relies on W satisfying Ground Assumption 3.1, however, we will also need W' to satisfy Ground Assumption 3.1. Because of the way W' was constructed from W, this is the case. Indeed, observe that for $\sigma \subset [n]$, we have a block matrix decomposition $(I - W')_\sigma = \left(\begin{smallmatrix} (I-W)_{\mathcal{E} \cap \sigma} & -W_{(\mathcal{E} \cap \sigma)(\mathcal{I} \cap \sigma)} \\ 0 & (I-W)_{\mathcal{I} \cap \sigma} \end{smallmatrix}\right)$. Therefore $\det((I - W')_\sigma) = \det((I - W)_{\mathcal{E} \cap \sigma}) \det((I - W)_{\mathcal{I} \cap \sigma}) \neq 0$ by Ground Assumption 3.1 for W. Having this in mind, we proceed with the proof.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Let $\sigma \in \mathcal{C}(W)$. Then there exists a $\tau \subset \mathcal{I}$ such that $\nu = \sigma \sqcup \tau \in FP(W)$. Thus, there exists a $b \in \mathbb{R}^n_{\geq 0}$ such that

$$x \stackrel{\text{def}}{=} (I_\nu - W_\nu)^{-1} b_\nu > 0 \text{ and } -W_{\overline{\nu}\nu} x \geq b_{\overline{\nu}} \geq 0$$

by Lemma 5.1. Observe that because of the way W' was defined we have that

$$(I - W')_{ij} = \begin{cases} 0, & i \in \mathcal{I}, j \in \mathcal{E}, \\ (I-W)_{ij} & \text{otherwise.} \end{cases}$$

Therefore $y \stackrel{\text{def}}{=} (I - W')_\nu x \geq (I - W)_\nu x = b_\nu \geq 0$. Observe that $x = (I_\nu - W'_\nu)^{-1} y = (I_\nu - W_\nu)^{-1} b_\nu$. Furthermore, $-W'_{\overline{\nu}\nu} x \geq -W_{\overline{\nu}\nu} x \geq 0$. Define $b' \in \mathbb{R}^n_{\geq 0}$ by $b'_\nu = y$ and $b'_{\overline{\nu}} = 0$. Then by construction $\nu = \sigma \sqcup \tau \in FP(W', b') \subset FP(W')$ by Lemma 5.1. Therefore $\sigma \in \mathcal{C}(W')$.

Now suppose that $\sigma \in \mathcal{C}(W')$. By Corollary 5.12 it follows that $\nu = \sigma \sqcup \mathcal{I} \in FP(W')$. Thus, there exists a $b' \in \mathbb{R}^n_{\geq 0}$ such that

$$x' \stackrel{\text{def}}{=} (I_\nu - W'_\nu)^{-1} b'_\nu > 0 \text{ and } -W'_{\overline{\nu}\nu} x' \geq b'_{\overline{\nu}} \geq 0$$

by Lemma 5.1. Define

$$c \stackrel{\text{def}}{=} \max_{i \in \mathcal{I}} \left| \sum_{k \in \sigma} (I - W)_{ik} x'_k \right|.$$

Let $x \in \mathbb{R}^{|\nu|}_{>0}$ be defined by $x_\sigma = x'_\sigma$ and $x_i = x'_i + 1 + c$ for all $i \in \mathcal{I}$. Then, for $i \in \sigma$ we have

$$\begin{aligned} ((I - W)_\nu x)_i &= \sum_{k \in \nu} (I - W)_{ik} x_k \\ &= \sum_{k \in \nu} (I - W')_{ik} x'_k + (1 + c) \sum_{k \in \mathcal{I}} (I - W)_{ik} \\ &\geq \sum_{k \in \nu} (I - W')_{ik} x'_k + 0 \geq 0. \end{aligned}$$

For $i \in \mathcal{I}$ we have

$$\begin{aligned} ((I - W)_\nu x)_i &= \sum_{k \in \nu} (I - W)_{ik} x_k = \sum_{k \in \sigma} (I - W)_{ik} x_k + \sum_{k \in \mathcal{I}} (I - W)_{ik} x_k \\ &= \sum_{k \in \sigma} (I - W)_{ik} x'_k + (I - W)_{ii} x_i + \sum_{i \neq k \in \mathcal{I}} (I - W)_{ik} x_k \\ &\geq \sum_{k \in \sigma} (I - W)_{ik} x'_k + (x'_i + 1 + c) + 0 \\ &\geq \sum_{k \in \sigma} (I - W)_{ik} x'_k + 1 + \left| \sum_{k \in \sigma} (I - W)_{ik} x'_k \right| \geq 1 > 0. \end{aligned}$$

Furthermore, for $i \in \overline{\nu} = \mathcal{E} \setminus \sigma$ we have

$$(-W_{\overline{\nu}\nu} x)_i = \sum_{k \in \nu} -W_{ik} x_k = \sum_{k \in \nu} -W'_{ik} x'_k + (1 + c) \sum_{k \in \mathcal{I}} -W'_{ik} \geq \sum_{k \in \nu} -W'_{ik} x'_k + 0 \geq 0.$$

Define $b \in \mathbb{R}^n_{\geq 0}$ by $b_\nu = (I - W)_\nu x$ and $b_{\overline{\nu}} = 0$. Then by construction and by Lemma 5.1, $\nu = \sigma \sqcup \overline{\mathcal{I}} \in FP(W, b) \subset FP(W)$. Therefore, $\sigma \in \mathcal{C}(W)$. ∎

From Theorems 3.2 and 3.3 we have the following corollary.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

COROLLARY 5.13. *Let $W \in \mathbb{D}_n$ with $\mathcal{E} = \{1, 2, \ldots, m\}$, $|\mathcal{I}| = k \geq 1$, $m + k = n$. Let $W' \in \mathbb{D}_{m+1}$ be a Dale matrix obtained from W in the following way: $W'_{ij} = W_{ij}$ for $i, j \in \mathcal{E}$, and for all $i \in \mathcal{E}$, $W'_{i(m+1)} < 0$ if and only if the ith row of $W_{\mathcal{E}\mathcal{I}}$ is a nonzero row vector, and the $(m+1)$th row of W' is zero. Then $\mathcal{C}(W) = \mathcal{C}(W')$.*

Before we proceed to prove Theorem 3.5, we will need the following lemma.

LEMMA 5.14. *Let A be an $n \times m$ matrix, $n \leq m$ such that $\exists \sigma \subset [m]$ with $|\sigma| = n$ and that $A_{[n]\sigma}$ is an invertible Z-matrix, and $A_{[n]([m]\setminus\sigma)} \leq 0$. Here $A_{[n]\tau}$ is a submatrix of A obtained by deleting columns outside of τ. Then $\exists x \in \mathbb{R}^m_{>0}$ such that $Ax \geq 0$ if and only if $A_{[n]\sigma}$ is an M-matrix.*

Proof. We can assume without loss of generality that $\sigma = [n]$, otherwise we can reindex. Suppose that $\exists x \in \mathbb{R}^m_{>0}$ such that $Ax \geq 0$. For all $i \in [n]$ we then have $(Ax)_i = \sum_{k\in[m]} A_{ik}x_k = \sum_{k\in[n]} A_{ik}x_k + \sum_{k\in[m]\setminus[n]} A_{ik}x_k \geq 0$, from where it follows that $\sum_{k\in[n]} A_{ik}x_k = A_{i[n]}x_{[n]} \geq 0$ because $A_{i([m]\setminus[n])} \leq 0$ by assumption. Thus, $A_{[n]}x_{[n]} \geq 0$ and thus $A_{[n]}$ is semipositive by Lemma 5.6. Since it is also a Z-matrix (off-diagonal entries are nonpositive) by assumption, we have that it is an M-matrix by Theorem 5.7, meaning $\rho(A) < 1$.

Conversely, suppose that $A_{[n]}$ is an M-matrix. Then, since it is also invertible by assumption, from Theorem 5.7 we have that $A_{[n]}$ is semipositive. Therefore, $\exists y \in \mathbb{R}^n_{>0}$ such that $A_{[n]}y \geq 0$ by Lemma 5.6. Note that for all $i \in [n]$, $\sum_{k\in[m]\setminus[n]} A_{ik} \leq 0$. Define

$$M \stackrel{\text{def}}{=} \min_{i\in[n]} \sum_{k\in[n]} A_{ik}y_k,\ N \stackrel{\text{def}}{=} \min_{i\in[n]} \sum_{k\in[m]\setminus[n]} A_{ik}.$$

Observe that since $A_{[n]}$ is invertible by assumption and $y > 0$ it cannot be that $A_{[n]}y = 0$. Thus, $M > 0$. Suppose that $N = 0$. Since, $A_{[n]([m]\setminus[n])} \leq 0$ by assumption, this implies that $A_{[n]([m]\setminus[n])} = 0$. Therefore, we can define $x \in \mathbb{R}^m_{>0}$ by $x_{[n]} = y$ and $x_j = 1$ for all $j \in [m] \setminus [n]$. Then, by construction $Ax \geq 0$. Now suppose that $N \neq 0$, that is, $N < 0$ and define $x \in \mathbb{R}^m_{>0}$ by $x_{[n]} = y$ and $x_j = \frac{-M}{N}$ for all $j \in [m] \setminus [n]$. Then for all $i \in [n]$ we have

$$\begin{aligned}(Ax)_i &= \sum_{k\in[m]} A_{ik}x_k = \sum_{k\in[n]} A_{ik}x_k + \sum_{k\in[m]\setminus[n]} A_{ik}x_k \\ &= \sum_{k\in[n]} A_{ik}y_k - \frac{M}{N} \sum_{k\in[m]\setminus[n]} A_{ik} \geq M - \frac{M}{N}N = M - M = 0.\end{aligned}$$

Therefore $Ax \geq 0$. ∎

Finally, we prove Theorem 3.5, reprinted below for convenience.

THEOREM 3.5. *Let W be a Dale matrix, and $\sigma \subset \mathcal{E}$ be a nonempty subset of excitatory neurons. Then $\sigma \in \mathcal{C}(W)$ if and only if the following two conditions are both satisfied:*

(i) *(the spectral condition) $\rho(W_{\mathcal{E}_\mathcal{U}\cap\sigma}) < 1$;*

(ii) *(the graph condition) $\sigma \in \text{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$,*

where $W_{\mathcal{E}_\mathcal{U}\cap\sigma}$ denotes the synaptic weights of the excitatory subnetwork on the subset $\mathcal{E}_\mathcal{U} \cap \sigma$, and $\rho(W_{\mathcal{E}_\mathcal{U}\cap\sigma})$ denotes the spectral radius of the matrix $W_{\mathcal{E}_\mathcal{U}\cap\sigma}$.

Proof of Theorem 3.5. By Theorems 3.2 and 3.3, we can assume that there is only one inhibitory neuron, let us say the nth neuron, and that the nth row of W is all 0. Recall that $\mathcal{E}_\mathcal{U}$ (resp., $\mathcal{E}_\mathcal{I}$) denotes the set of uninhibited (resp., inhibited) excitatory neurons of W, i.e., $\mathcal{E} = \mathcal{E}_\mathcal{U} \sqcup \mathcal{E}_\mathcal{I}$, and that $N^+_{G_\mathcal{E}}(\sigma)$ is the out-neighborhood

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

or the synaptic targets of $\sigma \subset \mathcal{E}$ in the excitatory connectivity graph $G_{\mathcal{E}}$ (Definition 3.4).

Suppose that $\sigma \in \mathcal{C}(W)$. As observed in Corollary 5.12, $\sigma \in \mathcal{C}(W)$ if and only if $\tau = \sigma \sqcup \{n\} \in FP(W)$. By Lemma 5.1, $\tau \in FP(W)$ if and only if $\exists x \in \mathbb{R}^{|\tau|}_{>0}$ such that the following two conditions are satisfied:

$$(I - W)_\tau x \geq 0, \tag{5.2a}$$

$$0 \leq -W_{\overline{\tau}\tau} x. \tag{5.2b}$$

For any vector $y \in \mathbb{R}^{|\tau|}_{>0}$ and any $i \in \mathcal{E}_{\mathcal{U}} \cap \sigma$ we have

$$((I - W)_\tau y)_i = \sum_{k \in \tau} (I - W)_{ik} y_k = \sum_{k \in \sigma} (I - W)_{ik} y_k$$

since $(I - W)_{in} = 0$. Therefore we can make the following two observations:

(a) The condition (5.2a) ensures that $x_\sigma > 0$ satisfies $(I - W)_{(\mathcal{E}_{\mathcal{U}} \cap \sigma)\sigma} x_\sigma \geq 0$. Observe that $\sigma \setminus \mathcal{E}_{\mathcal{U}} = \mathcal{E}_{\mathcal{I}} \cap \sigma$, i.e., $\sigma = (\mathcal{E}_{\mathcal{U}} \cap \sigma) \sqcup (\mathcal{E}_{\mathcal{I}} \cap \sigma)$. Because $(I - W)_{\mathcal{E}_{\mathcal{U}} \cap \sigma}$ is a nonsingular Z-matrix (off-diagonal entries are nonpositive) and $(I - W)_{(\mathcal{E}_{\mathcal{U}} \cap \sigma)(\mathcal{E}_{\mathcal{I}} \cap \sigma)} \leq 0$, by Lemma 5.15 we have that this is equivalent to $(I - W)_{\mathcal{E}_{\mathcal{U}} \cap \sigma}$ being an M-matrix, that is, $\rho(W_{\mathcal{E}_{\mathcal{U}} \cap \sigma}) < 1$.

(b) By observing that $\overline{\tau} = \mathcal{E} \setminus \sigma$, the condition (5.2b) is equivalent to the requirement that for all $i \in \mathcal{E} \setminus \sigma$ we need to have $(-W_{(\mathcal{E} \setminus \sigma)\tau} x)_i = \sum_{k \in \tau} -W_{ik} x_k \geq 0$. For $i \in \mathcal{E} \setminus \sigma$, if $W_{in} = 0$ the only possibility is that $W_{i\sigma} = 0$ for the inequality to be true. Thus to be able to find an $x \in \mathbb{R}^n_{>0}$ for which condition (5.2b) is satisfied we need that for all $i \in \mathcal{E} \setminus \sigma$, $W_{i\sigma} = 0$ or $W_{in} < 0$. This implies that $N^+_{G_{\mathcal{E}}}(\sigma) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma$.

Now suppose the following two conditions are both satisfied:

$$\rho(W_{\mathcal{E}_{\mathcal{U}} \cap \sigma}) < 1, \tag{5.3a}$$

$$N^+_{G_{\mathcal{E}}}(\sigma) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma. \tag{5.3b}$$

Condition (5.3a) means that $(I - W)_{\mathcal{E}_{\mathcal{U}} \cap \sigma}$ is an M-matrix. Since $(I - W)_{(\mathcal{E}_{\mathcal{U}} \cap \sigma)(\mathcal{E}_{\mathcal{I}} \cap \sigma)} \leq 0$, by Lemma 5.15 this means that $\exists x \in \mathbb{R}^{|\sigma|}_{>0}$ such that $(I - W)_{(\mathcal{E}_{\mathcal{U}} \cap \sigma)\sigma} x \geq 0$. Let $\tau = \sigma \sqcup \{n\}$ and define

$$c \stackrel{\text{def}}{=} \max \left\{ \begin{array}{l} \max\limits_{i \in \mathcal{E}_{\mathcal{I}} \cap \sigma} -\dfrac{1}{W_{in}} \left| \sum\limits_{k \in \sigma} (I - W)_{ik} x_k \right|, \\ \max\limits_{i \in \mathcal{E}_{\mathcal{I}} \setminus \sigma} -\dfrac{1}{W_{in}} \left| \sum\limits_{k \in \sigma} -W_{ik} x_k \right|, \left| \sum\limits_{k \in \sigma} -W_{nk} x_k \right| \end{array} \right\}.$$

Define $y \in \mathbb{R}^{|\tau|}_{>0}$ by $y_\sigma = x$ and $y_{|\tau|} = 1 + c$. Then, for $i \in \mathcal{E}_{\mathcal{U}} \cap \sigma$ we have

$$\begin{aligned} ((I - W)_\tau y)_i &= \sum_{k \in \tau} (I - W)_{ik} y_k = \sum_{k \in \sigma} (I - W)_{ik} y_k + (I - W)_{in} y_{|\tau|} \\ &= \sum_{k \in \sigma} (I - W)_{ik} x_k - W_{in}(1 + c) \geq 0 + 0(1 + c) = 0. \end{aligned}$$

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

For $i \in \mathcal{E}_{\mathcal{I}} \cap \sigma$ we have

$$\begin{aligned}((I-W)_\tau y)_i &= \sum_{k\in\tau}(I-W)_{ik}y_k = \sum_{k\in\sigma}(I-W)_{ik}y_k + (I-W)_{in}y_{|\tau|}\\ &= \sum_{k\in\sigma}(I-W)_{ik}x_k - W_{in}(1+c) = \sum_{k\in\sigma}(I-W)_{ik}x_k - W_{in} - W_{in}c\\ &\geq \sum_{k\in\sigma}(I-W)_{ik}x_k - W_{in} + W_{in}\frac{1}{W_{in}}\left|\sum_{k\in\sigma}(I-W)_{ik}x_k\right| \geq -W_{in} > 0.\end{aligned}$$

Similarly if $i = n$, we will also have that $((I-W)_\tau y)_i \geq 0$ because of the definition of $y_{|\tau|}$. Therefore we have that $(I-W)_\tau y \geq 0$. Furthermore, $N^+_{G_\mathcal{E}}(\sigma) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma$ implies that for all $i \in \mathcal{E} \setminus \sigma$, $W_{i\sigma} = 0$ or $W_{in} < 0$. Note that $\overline{\tau} = \mathcal{E} \setminus \sigma = (\mathcal{E}_{\mathcal{U}} \setminus \sigma) \sqcup (\mathcal{E}_{\mathcal{I}} \setminus \sigma)$. Then for $i \in \mathcal{E}_{\mathcal{U}} \setminus \sigma$, and thus $W_{in} = 0$, we have that it must be that $W_{i\sigma} = 0$ and therefore

$$(-W_{\overline{\tau}\tau}y)_i = \sum_{k\in\tau} -W_{ik}y_k = \sum_{k\in\sigma} -W_{ik}y_k - W_{in}y_{|\tau|} = 0 - 0(1+c) = 0 \geq 0.$$

On the other hand for $i \in \mathcal{E}_{\mathcal{I}} \setminus \sigma$ we have that $W_{in} < 0$ and thus

$$\begin{aligned}(-W_{\overline{\tau}\tau}y)_i &= \sum_{k\in\tau} -W_{ik}y_k = \sum_{k\in\sigma} -W_{ik}y_k - W_{in}y_{|\tau|} = \sum_{k\in\sigma} -W_{ik}x_k - W_{in}(1+c)\\ &= \sum_{k\in\sigma} -W_{ik}x_k - W_{in} - W_{in}c \geq \sum_{k\in\sigma} -W_{ik}x_k - W_{in} + W_{in}\frac{1}{W_{in}}\left|\sum_{k\in\sigma} -W_{ik}x_k\right|\\ &\geq -W_{in} > 0.\end{aligned}$$

Thus, we have that $0 \leq -W_{\overline{\tau}\tau}y$. Let $b \in \mathbb{R}^n$ be defined by $b_\tau = (I-W)_\tau y$ and $b_{\overline{\tau}} = 0$. Then $\tau = \sigma \sqcup \{n\} \in FP(W,b) \subset FP(W)$ (Definition 5.8) by Lemma 5.1. Hence $\sigma \in \mathcal{C}(W)$. ∎

We now prove Theorem 3.6, mainly relying on results stated in Appendix A. We reprinted the theorem below for convenience.

THEOREM 3.6. *Let W be a nonsingular Dale matrix that is weakly coupled, i.e., $||W||_F < 1$. Then every fixed point is asymptotically stable, and the combinatorial code is completely described by the graph condition:*

$$\mathcal{C}(W) = \mathcal{SC}(W) = \mathrm{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U}).$$

Furthermore, for all $b \in \mathbb{R}^n_{\geq 0}$ there is a unique globally exponentially stable fixed point of (2.1).

Proof of Theorem 3.6. Since $||W||_F < 1$, this implies that $\rho(W_\eta) < 1$ for any principal submatrix W_η, $\eta \subset [n]$. Therefore by Theorem 3.5, $\mathcal{C}(W) = \mathrm{code}(G_\mathcal{E}, \mathcal{E}_\mathcal{U})$. Furthermore, $||W||_F < 1$ implies that $W \in \mathcal{L}$ (Definition A.1) by Lemma A.2. Since Ground Assumption 3.1 is assumed to hold, by Proposition A.3, for all $b \in \mathbb{R}^n_{\geq 0}$, (2.1) has a unique globally exponentially stable fixed point. In particular, if $\sigma \in \mathcal{C}(W)$, by definition $\sigma \sqcup \tau$ is a support of a fixed point of (2.1) for some $\tau \subset \mathcal{I}$ and some input $b \in \mathbb{R}^n_{\geq 0}$. This fixed point is unique and asymptotically (exponentially) stable as argued above. Therefore $\sigma \in \mathcal{SC}(W)$. The inclusion $\mathcal{SC}(W) \subset \mathcal{C}(W)$ holds by definition and thus $\mathcal{C}(W) = \mathcal{SC}(W)$. ∎

The proof of Proposition 3.9 is straightforward and it gives us the important Corollary 5.15. We reprinted the proposition below for convenience.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

PROPOSITION 3.9. *The code* $\operatorname{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$ *is a sublattice of the Boolean lattice* $2^{\mathcal{E}}$.

Proof of Proposition 3.9. It's easy to see that $\varnothing, \mathcal{E} \in \operatorname{code}(G, \mathcal{E}_{\mathcal{U}})$. To show that $\operatorname{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$ is respected by intersections and unions assume that $N^+_{G_{\mathcal{E}}}(\sigma_i) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma_i$ for $i = 1, 2$, $\sigma_i \subset \mathcal{E}$. Since $N^+_{G_{\mathcal{E}}}(\sigma_1 \cap \sigma_2) \subset N^+_{G_{\mathcal{E}}}(\sigma_1) \cap N^+_{G_{\mathcal{E}}}(\sigma_2)$, we obtain that

$$N^+_{G_{\mathcal{E}}}(\sigma_1 \cap \sigma_2) \cap \mathcal{E}_{\mathcal{U}} \subset N^+_{G_{\mathcal{E}}}(\sigma_1) \cap N^+_{G_{\mathcal{E}}}(\sigma_1) \cap \mathcal{E}_{\mathcal{U}} \subset \sigma_1 \cap \sigma_2.$$

Similarly,

$$\begin{aligned} N^+_{G_{\mathcal{E}}}(\sigma_1 \cup \sigma_2) \cap \mathcal{E}_{\mathcal{U}} &= (N^+_{G_{\mathcal{E}}}(\sigma_1) \cup N^+_{G_{\mathcal{E}}}(\sigma_2)) \cap \mathcal{E}_{\mathcal{U}} \\ &= (N^+_{G_{\mathcal{E}}}(\sigma_1) \cap \mathcal{E}_{\mathcal{U}}) \cup (N^+_{G_{\mathcal{E}}}(\sigma_2) \cap \mathcal{E}_{\mathcal{U}}) \subset \sigma_1 \cup \sigma_2. \end{aligned}$$ ∎

COROLLARY 5.15. *Let* $W \in \mathbb{D}_n$. *The code* $\mathcal{C}(W)$ *is closed under intersections.*

Proof. Suppose that $\sigma, \tau \in \mathcal{C}(W)$. By Theorem 3.5 we have that $\rho(W_{\mathcal{E}_{\mathcal{U}} \cap \sigma})$, $\rho(W_{\mathcal{E}_{\mathcal{U}} \tau}) < 1$ and that $\sigma, \tau \in \operatorname{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$. Note that $W_{\mathcal{E}_{\mathcal{U}} \cap (\sigma \cap \tau)}$ is a principal submatrix of $W_{\mathcal{E}_{\mathcal{U}} \cap \sigma}$ which is nonnegative. Thus, by Lemma 5.4, $\rho(W_{\mathcal{E}_{\mathcal{U}} \cap (\sigma \cap \tau)}) < 1$. Furthermore, $\operatorname{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$ is closed under intersections by Proposition 3.9 and thus $\sigma \cap \tau \in \operatorname{code}(G_{\mathcal{E}}, \mathcal{E}_{\mathcal{U}})$. Thus, by Theorem 3.5 it follows that $\sigma \cap \tau \in \mathcal{C}(W)$. ∎

By Corollary 5.15 and Theorem 3.8 we immediately get Theorem 3.10.

We now proceed to prove Theorem 3.11. Given a code $C \subset 2^{\mathcal{E}}$ that is a sublattice with $\varnothing, \mathcal{E} \in C$, recall the definition of $c : 2^{\mathcal{E}} \to C$ and the graph $(G^c, \mathcal{E})$ from Theorem 3.11. Let $N^+_{G^c}(\sigma)$ denote the targets of $\sigma \subset \mathcal{E}$ (the out-neighborhood of σ) in the graph $(G_c, \mathcal{E})$. By definition, one sees that $c(\varnothing) = \varnothing$, $c(\mathcal{E}) = \mathcal{E}$, and $\sigma \subset c(\sigma)$. We immediately make the following observations that will help us prove Theorem 3.11.

LEMMA 5.16. *Let* $C \subset 2^{\mathcal{E}}$ *be a sublattice with* $\varnothing, \mathcal{E} \in C$. *Then the following are true:*

1. *Let* $\sigma, \tau \subset \mathcal{E}$. *Then* $c(\sigma \cup \tau) = c(\sigma) \cup c(\tau)$.
2. $C = \{\sigma \subset \mathcal{E} \mid \sigma = c(\sigma)\}$.

Proof. For part 1, we first show that $c(\sigma) \cup c(\tau) \subset c(\sigma \cup \tau)$. Note that since $\sigma \subset \sigma \cup \tau$ it follows that $c(\sigma) \subset c(\sigma \cup \tau)$. Similarly $c(\tau) \subset c(\sigma \cup \tau)$. Therefore $c(\sigma) \cup c(\tau) \subset c(\sigma \cup \tau)$. Now we show the other inclusion. Observe that since C is a lattice, it follows that $c(\sigma), c(\tau)$ and $c(\sigma) \cup c(\tau) \in C$. Furthermore, $\sigma \cup \tau \subset c(\sigma) \cup c(\tau)$ and therefore by definition $c(\sigma \cup \tau) \subset c(\sigma) \cup c(\tau)$.

To prove part 2, suppose that $\sigma \in C$. By definition,

$$c(\sigma) = \bigcap_{\nu \in C, \sigma \subset \nu} \nu.$$

Since $\sigma \in C$, it follows that

$$\bigcap_{\nu \in \mathcal{C}, \sigma \subset \nu} \nu = \sigma,$$

and thus $c(\sigma) = \sigma$. Now suppose that $\sigma \subset \mathcal{E}$ is such that

$$c(\sigma) = \bigcap_{\nu \in C, \sigma \subset \nu} \nu = \sigma.$$

Since C is finite and is closed under finite intersections, it follows that $\sigma \in C$. ∎

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

LEMMA 5.17. *The arcs in G^c are a transitive relation on $\mathcal{E}$.*

Proof. Let $i,j,k \in \mathcal{E}$ be such that $i \to j$ and $j \to k$. In other words, we have $j \in c(i)$ and $k \in c(j)$. Let $\nu \in C$ be such that $i \in \nu$. Then, by definition, $j \in \nu$. Since $k \in c(j)$, by definition this implies that $k \in \nu$. Therefore, by definition $i \to k$. ∎

LEMMA 5.18. *Let G be any graph whose transitive closure is G^c. Then* $\operatorname{code}(G,\mathcal{E}) = \operatorname{code}(G^c,\mathcal{E})$.

Proof. Let $N_G^+(\sigma), N_{G^c}^+(\sigma)$ be the targets of $\sigma \subset \mathcal{E}$, in G and G^c, respectively. By definition $\operatorname{code}(G,\mathcal{E}) = \{\sigma \subset \mathcal{E} \mid N_G^+(\sigma) \subset \sigma\}$ and $\operatorname{code}(G^c,\mathcal{E}) = \{\sigma \subset \mathcal{E} \mid N_{G^c}^+(\sigma) \subset \sigma\}$. By assumption $N_G^+(\sigma) \subset N_{G^c}^+(\sigma)$ for all $\sigma \subset \mathcal{E}$. Thus for any $\sigma \subset \mathcal{E}$ if $N_{G^c}^+(\sigma) \subset \sigma$, then $N_G^+(\sigma) \subset \sigma$. Therefore $\operatorname{code}(G^c,\mathcal{E}) \subset \operatorname{code}(G,\mathcal{E})$. Furthermore, if for any $\sigma \subset \mathcal{E}$, $N_G^+(\sigma) \subset \sigma$, then $(N_{G^c}^+)^m(\sigma) \subset \sigma$, where $(N_{G^c}^+)^m$ is an m-fold application of the $N_{G^c}^+$ operator to σ, for any $m \geq 1$. Thus, $N_c^+(\sigma) \subset \sigma$. Therefore $\operatorname{code}(G,\mathcal{E}) \subset \operatorname{code}(G^c,\mathcal{E})$. ∎

Finally, we can prove Theorem 3.11.

Proof of Theorem 3.11. By definition,

$$\operatorname{code}(G^c,\mathcal{E}) = \{\sigma \subset \mathcal{E} \mid N_{G^c}^+(\sigma) \cap \mathcal{E} = N_{G^c}^+(\sigma) \subset \sigma\}.$$

Thus, we need to show that $\{\sigma \subset \mathcal{E} \mid N_{G^c}^+(\sigma) \cap \mathcal{E} = N_{G^c}^+(\sigma) \subset \sigma\} = \{\sigma \subset \mathcal{E} \mid c(\sigma) = \sigma\}$. Let G be the digraph on $\mathcal{E}$ and let it be defined by $i \to j$ if and only if $j \in c(i)$ for all $i,j \in \mathcal{E}$. Let $N_G^+(\sigma)$ be the targets of $\sigma \subset \mathcal{E}$ in G. Note that by construction, $N_G^+(\sigma) = N_{G^c}^+(\sigma) \cup \sigma$. Therefore for all $\sigma \subset \mathcal{E}$, $N_G^+(\sigma) = \sigma$ if and only if $N_{G^c}^+(\sigma) \subset \sigma$. Furthermore, by construction $N_G^+(i) = c(i)$. Hence by Lemma 5.16 for all $\sigma \subset \mathcal{E}$ we have

$$N_G^+(\sigma) = \bigcup_{i \in \sigma} N_G^+(i) = \bigcup_{i \in \sigma} c(i) = c\left(\bigcup_{i \in \sigma} i\right) = c(\sigma).$$

Therefore for all $\sigma \subset \mathcal{E}$, $c(\sigma) = \sigma$ if and only if $N_{G^c}^+(\sigma) \subset \sigma$. ∎

Appendix A. Stability of linear-threshold rate dynamics. Here we recall the necessary results on the stability of the linear threshold dynamics from [24, 26] that we used in the proofs of section 5. For a matrix A let $||A||$ be its 2-norm and let $|A|$ denote the matrix $|A|_{ij} = |A_{ij}|$. Given a $\sigma \in \{0,1\}^n$ let $\Sigma = \operatorname{diag}(\sigma)$, that is the diagonal matrix with the elements of σ on the diagonal.

DEFINITION A.1. *An $n \times n$ matrix A is totally-$\mathcal{L}$ stable, written $A \in \mathcal{L}$, if there exists $P = P^T \succ 0$ such that $(-I + A^T\Sigma)P + P(-I + \Sigma A) \prec 0$ for all $\Sigma = \operatorname{diag}(\sigma)$ and $\sigma \in \{0,1\}^n$.*

LEMMA A.2 (see [24, Lemma 2.3]).
1. $\rho(|W|) < 1 \Longrightarrow W \in \mathcal{L}$.
2. $||W|| < 1 \Longrightarrow W \in \mathcal{L}$.

In the statement of the following result the authors in [24] had a hypothesis that W is nonsingular and that for all $\sigma \in \{0,1\}^n$, $(I - \Sigma W)$ is nonsingular as well. By the standard correspondence between finite binary sequences and finite subsets of $2^{[n]}$, it is not hard to observe that for a given $\sigma \in \{0,1\}^n$, $(I - \Sigma W)$ is nonsingular if and only if $(I - W)_\sigma$ is nonsingular. Thus, the second hypothesis is equivalent to Ground Assumption 3.1. We thus continue not writing out Ground Assumption 3.1 in the

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

statement of the following theorem as is the case in the rest of the paper, but the reader should note that it is indeed necessary for the statement to be true.

PROPOSITION A.3 (see [24, Proposition 4.9]). *Consider the network dynamics in* (2.1). *If* $\rho(|W|) < 1$ *or* $||W|| < 1$, *then for all* $b \in R^n$, *the network has a unique fixed point* x^* *and it is globally exponentially stable relative to* x^*.

Note that it has been shown in [24], that a more general condition [24, Theorem 4.8] can also guarantee the above result. However, checking if the matrix W satisfies those conditions is significantly harder in our context, thus we used the statement above instead.

Acknowledgment. We thank the anonymous referees for their many helpful comments that improved the manuscript.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

REFERENCES

[1] Y. AHMADIAN, F. FUMAROLA, AND K. D. MILLER, *Properties of networks with partially structured and partially random connectivity*, Phys. Rev. E (3), 91 (2015), 012820, https://doi.org/10.1103/PhysRevE.91.012820.

[2] P. ALEXANDROFF, *Diskrete Räume*, Mat. Sb. (NS), 2 (1937), pp. 501–519, https://www.mathnet.ru/eng/sm5579.

[3] P. ASHWIN AND M. TIMME, *When instability makes sense*, Nature, 436 (2005), pp. 36–37, https://doi.org/10.1038/436036b.

[4] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Classics Appl. Sci. 9, SIAM, Philadelphia, 1994.

[5] J. CRUZ, C. GIUSTI, V. ITSKOV, AND B. KRONHOLM, *On open and closed convex codes*, Discrete Comput. Geom., 61 (2019), pp. 247–270, https://doi.org/10.1007/s00454-018-00050-1.

[6] C. CURTO, A. DEGERATU, AND V. ITSKOV, *Flexible memory networks*, Bull. Math. Biol., 74 (2012), pp. 590–614, https://doi.org/10.1007/s11538-011-9678-9.

[7] C. CURTO, A. DEGERATU, AND V. ITSKOV, *Encoding binary neural codes in networks of threshold-linear neurons*, Neural Comput., 25 (2013), pp. 2858–2903, https://doi.org/10.1162/NECO_a_00504.

[8] C. CURTO, E. GROSS, J. JEFFRIES, K. MORRISON, M. OMAR, Z. ROSEN, A. SHIU, AND N. YOUNGS, *What makes a neural code convex?*, SIAM J. Appl. Algebra Geom., 1 (2017), pp. 222–238, https://doi.org/10.1137/16M1073170.

[9] C. CURTO, V. ITSKOV, A. VELIZ-CUBA, AND N. YOUNGS, *The neural ring: An algebraic tool for analyzing the intrinsic structure of neural codes*, Bull. Math. Biol., 75 (2013), pp. 1571–1611, https://doi.org/10.1007/s11538-013-9860-3.

[10] C. CURTO AND K. MORRISON, *Pattern completion in symmetric threshold-linear networks*, Neural Comput., 28 (2015), pp. 2825–2852, https://doi.org/10.1162/NECO_a_00869.

[11] H. DALE, *Pharmacology and nerve-endings*, Proc. R. Soc. Med. (1935), pp. 319–332.

[12] J. C. ECCLES, *From electrical to chemical transmission in the central nervous system,* Notes Rec. Roy. Soc. Lond., 30 (1976), pp. 219–230, https://doi.org/10.1098/rsnr.1976.0015.

[13] C. GIUSTI AND V. ITSKOV, *A no-go theorem for one-layer feedforward networks*, Neural Comput., 26 (2014), pp. 2527–2540, https://doi.org/10.1162/NECO_a_00657.

[14] R. H. HAHNLOSER, R. SARPESHKAR, M. MAHOWALD, R. DOUGLAS, AND H. SEUNG, *Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit*, Nature, 405 (2000), pp. 947–951, https://doi.org/10.1038/35016072.

[15] R. H. HAHNLOSER, H. SEUNG, AND J. SLOTINE, *Permitted and forbidden sets in symmetric threshold-linear networks*, Neural Comput., 15 (2003), pp. 621–638, https://doi.org/10.1162/089976603321192103.

[16] F. C. HOPPENSTEADT AND E. M. IZHIKEVICH, *Weakly Connected Neural Networks*, Appl. Math. Sci. 126, Springer, New York, 1997.

[17] R. HUERTA AND M. RABINOVICH, *Reproducible sequence generation in random neural ensembles*, Phys. Rev. Lett., 93 (2004), 238104, https://doi.org/10.1103/PhysRevLett.93.238104.

[18] J. R. IPSEN AND A. D. PETERSON, *Consequences of Dale's law on the stability-complexity relationship of random neural networks*, Phys. Rev. E (3), 101 (2020), 052412, https://doi.org/10.1103/PhysRevE.101.052412.

[19] V. ITSKOV, A. KUNIN, AND Z. ROZEN, *Hyperplane neural codes and the polar complex*, Abel Symposia Proc., 15 (2020), pp. 343–369, https://doi.org/10.1007/978-3-030-43408-3_13.

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license

[20] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, 6th ed., Elsevier, New York, 1991.
[21] C. Lienkaemper, A. Shiu, and Z. Woodstock, *Obstructions to convexity in neural codes*, Adv. Appl. Math., 85 (2017), pp. 31–59, https://doi.org/10.1016/j.aam.2016.11.006.
[22] F. Mastrogiuseppe and S. Ostojic, *Linking connectivity, dynamics, and computations in low-rank recurrent neural networks*, Neuron, 99 (2018), pp. 609–623, https://doi.org/10.1016/j.neuron.2018.07.003.
[23] K. Morrison, A. Degeratu, V. Itskov, and C. Curto, *Diversity of emergent dynamics in competitive threshold-linear networks*, SIAM J. Appl. Dyn. Syst., 23 (2024), pp. 855–884, https://doi.org/10.1137/22M1541666.
[24] E. Nozari and J. Cortés, *Hierarchical selective recruitment in linear-threshold brain networks—part* I*: Single-layer dynamics and selective inhibition*, IEEE Trans. Automat. Control, 66 (2020), pp. 949–964, https://doi.org/10.1109/TAC.2020.3004801.
[25] N. N. Osborne, *Dale's Principle and Communication Between Neurones: Based on a Colloquium of the Neurochemical Group of the Biochemical Society, July* 1982, Held at Oxford University, Elsevier, Burlington, MA, 2013.
[26] A. Pavlov, N. van de Wouw, and H. Nijmeijer, *Convergent piecewise affine systems: Analysis and design part* I*: Continuous case*, in Proceedings of the 44th IEEE Conference on Decision and Control, IEEE, Piscataway, NJ, 2005, pp. 5391–5396.
[27] M. I. Rabinovich and P. Varona, *Robust transient dynamics and brain functions*, Front. Comput. Neurosci., 5 (2011), 24, https://doi.org/10.3389/fncom.2011.00024.
[28] K. Rajan and L. F. Abbott, *Eigenvalue spectra of random matrices for neural networks*, Phys. Rev. Lett., 97 (2006), 188104, https://doi.org/10.1103/PhysRevLett.97.188104.
[29] S. Roman, *Lattices and Ordered Sets*, Springer New York, 2010.
[30] M. Tournoy and B. Doiron, *A step towards uncovering the structure of multistable neural networks*, https://doi.org/10.48550/arXiv.2210.03241, 2022.
[31] C. Van Vreeswijk and H. Sompolinsky, *Chaos in neuronal networks with balanced excitatory and inhibitory activity*, Science, 274 (1996), pp. 1724–1726, https://doi.org/10.1126/science.274.5293.1724.
[32] X. Xie, R. H. Hahnloser, and H. Seung, *Selectively grouping neurons in recurrent networks of lateral inhibition*, Neural Comput., 14 (2002), pp. 2627–2646, https://doi.org/10.1162/089976602760408008.

Downloaded 12/11/24 to 165.85.38.17 . Redistribution subject to CCBY license

© 2024 SIAM. Published by SIAM under the terms of the Creative Commons 4.0 license