

Sequence generation in inhibition-dominated neural networks

Caitlyn Parmelee, Juliana Londono Alvarez, Carina Curto*, Katherine Morrison*

* equal contribution

Sequences of neural activity arise in many brain areas, including cortex [1, 2, 3], hippocampus [4, 5, 6], and central pattern generator circuits that underlie rhythmic behaviors like locomotion [7, 8]. Such sequences are examples of emergent or *internally-generated activity*: neural activity that is shaped primarily by the structure of a recurrent network rather than inherited from a changing external input. A fundamental question is to understand how a network’s connectivity shapes neural activity. In particular, what types of network architectures give rise to sequential activity?

While the architectures underlying sequence generation vary considerably, a common feature is an abundance of inhibition. Inhibition-dominated networks exhibit emergent sequences even in the absence of an obvious chain-like architecture, such as a synfire chain [9, 10]. Here, we provide an overview of results from [11] on network architectures that produce sequential dynamics in a special family of inhibition-dominated networks known as combinatorial threshold-linear networks (CTLNs).

Emergent sequences from CTLNs

Combinatorial threshold-linear networks (CTLNs) are a special family of threshold-linear networks (TLNs), whose dynamics are prescribed by the equations:

$$\frac{dx_i}{dt} = -x_i + \left[\sum_{j=1}^n W_{ij}x_j + \theta_i \right]_+, \quad i = 1, \dots, n. \quad (1)$$

Here, $x_1(t), \dots, x_n(t)$ represent the firing rate of n recurrently-connected neurons, $[\cdot]_+ = \max\{0, \cdot\}$ is the threshold nonlinearity, and the connection strength matrix $W = W(G, \varepsilon, \delta)$ is determined by a simple¹ directed graph G as follows:

$$W_{ij} = \begin{cases} 0 & \text{if } i = j, \\ -1 + \varepsilon & \text{if } j \rightarrow i \text{ in } G, \\ -1 - \delta & \text{if } j \not\rightarrow i \text{ in } G. \end{cases} \quad (2)$$

We require the parameters to satisfy $\delta > 0$, and $0 < \varepsilon < \frac{\delta}{\delta+1}$, and typically we take the external input to be constant, $\theta_i = \theta > 0$, to ensure the dynamics are internally generated. Note that the upper bound on ε implies $\varepsilon < 1$, and so the W matrix is always effectively inhibitory. We think of this as modeling the activity of excitatory neurons in a sea of inhibition: an edge indicates an excitatory connection in the presence of background inhibition, while the absence of an edge indicates no excitatory connection. The graph G thus captures the effective pattern of weak and strong inhibition in the network.

The simplest architectures that give rise to sequences in CTLNs are graphs that are cycles. These networks produce limit cycles in which the neurons reach their peak activation in cyclic order. For example, in Figure 1A-C, we see such sequential limit cycles when the graph G is a 3-cycle, 4-cycle, or a 5-cycle. Interestingly, these simple cycles can also be embedded in a larger network, as in Figure 1D, and still produce sequences that match those of the subnetwork in isolation. Each of the sequences from panels A-C may be transiently activated in the network in panel D by a time-dependent external drive $\theta_i(t)$ consisting of pulses that are concentrated on different subnetworks. A single simulation is shown, with localized pulses activating the 5-cycle, the 3-cycle, and finally the 4-cycle. Although these cycles overlap, each pulse activates a sequence involving only the neurons in the stimulated subnetwork. Depending on the duration of the pulse, the sequence may play only once or repeat two or more times.

What other architectures, beyond simple cycles, produce sequential dynamics? Is there a key feature that we can generalize from cycles that will allow us to build more interesting architectures supporting sequential activity? To better understand how graph connectivity shapes dynamics, we focus our attention on the relationship between connectivity and fixed points of the dynamics.

¹A graph is *simple* if it does not have self-loops or multiple edges (in the same direction) between a pair of nodes.

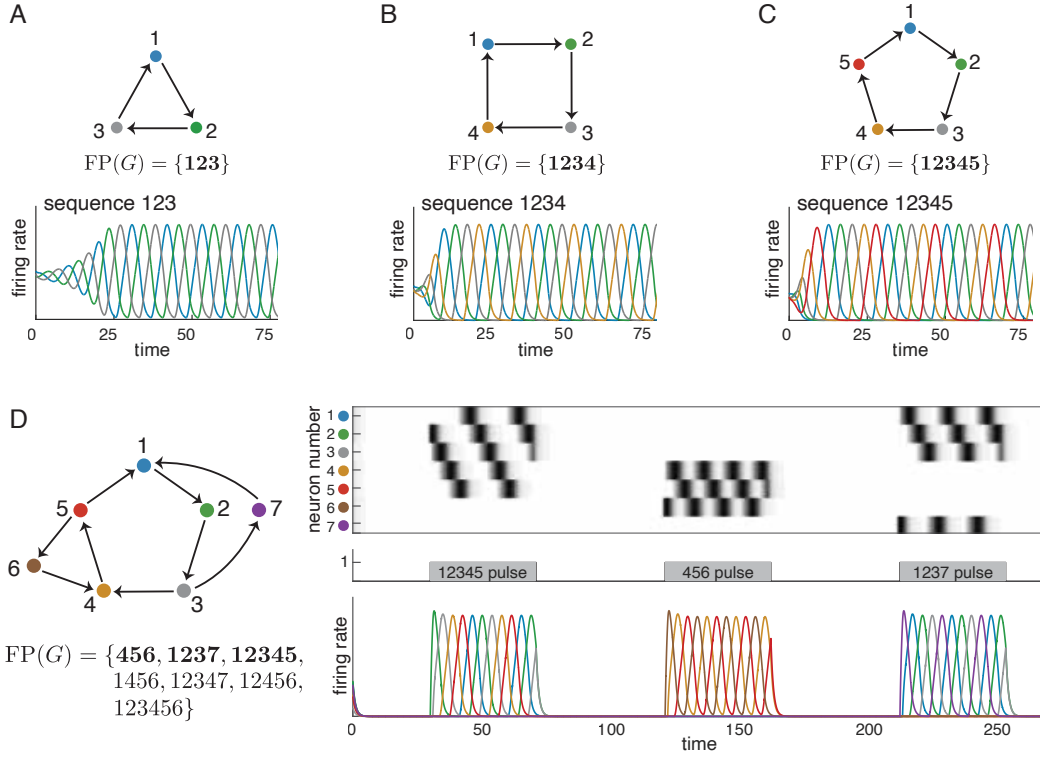


Figure 1: **Sequential attractors from simple cycles.** (A-C) CTLNs corresponding to a 3-cycle, a 4-cycle, and a 5-cycle each produce a limit cycle where the neurons reach their peak activations in the expected sequence. Colored curves correspond to solutions $x_i(t)$ for matching node i in the graph. (D) Attractors corresponding to the embedded 3-cycle, 4-cycle, and 5-cycle of the network are transiently activated to produce sequences matching those of the isolated cycle networks in A-C. For each network in A-D, $FP(G)$ is shown, with the minimal fixed points in bold. Note that we say a fixed point is *minimal* if its support is minimal under inclusion within $FP(G)$. To simplify notation for $FP(G)$, we denote a subset $\{i_1, \dots, i_k\}$ by $i_1 \dots i_k$. For example, 12345 denotes the set $\{1, 2, 3, 4, 5\}$. Unless otherwise noted, all simulations have CTLN parameters $\varepsilon = 0.25$, $\delta = 0.5$, and $\theta = 1$.

Fixed points of CTLNs

One of the most striking features of CTLNs is the strong connection between dynamic attractors and *unstable* fixed points [12, 13]. (Of course, stable fixed points are also attractors of the network, but these are static.) Recall that a fixed point x^* of a CTLN is a solution that satisfies $dx_i/dt|_{x=x^*} = 0$ for each $i \in [n]$. The *support* of a fixed point is the subset of active neurons, $\text{supp } x \stackrel{\text{def}}{=} \{i \mid x_i^* > 0\}$. For a given nondegenerate TLN, there can be at most one fixed point per support, and it is straightforward to recover the actual fixed point values from knowledge of the support. Thus, it suffices to label all the fixed points of a network by their support, $\sigma = \text{supp } x^* \subseteq [n]$, where $[n] \stackrel{\text{def}}{=} \{1, \dots, n\}$. We denote the collection of fixed point supports by

$$FP(G) = FP(G, \varepsilon, \delta) \stackrel{\text{def}}{=} \{\sigma \subseteq [n] \mid \sigma \text{ is a fixed point support of } W(G, \varepsilon, \delta)\}.$$

For example, in Figure 1A, 123 is the only fixed point support of the 3-cycle. In contrast, the network in Figure 1D supports several fixed points. Can these fixed point supports give insight into the dynamics?

It turns out that the *minimal* fixed points of the network in Figure 1D reflect the subsets of neurons that are active in each of the attractors. This phenomenon has been consistently observed in simulations; in [13] it was shown that there is a close correspondence between certain minimal fixed points in $FP(G)$, known as *core motifs*, and the attractors of a network. Additionally, prior work proved a series of *graph rules* that can be used to determine fixed points of a CTLN by analyzing the structure of the graph G [14, 15]. These rules directly connect the network connectivity to $FP(G)$, and are all independent of the choice of parameters ε , δ , and θ . Here, we investigate families of network architectures that give strong constraints on $FP(G)$, and for which this fixed point structure is predictive of sequential attractors of the networks.

What other architectures can support sequential attractors?

Simple cycles were the most obvious candidate to produce sequential attractors. But what other architectures can support sequential attractors? Here, we investigate four architectures that generalize different features of simple cycles: cyclic unions, directional cycles, simply-embedded partitions, and simply-embedded directional cycles. A common feature of all these architectures is that the neurons of the network are partitioned into components $\tau_1, \dots, \tau_N$, organized in a cyclic manner, whose disjoint union equals the full set of neurons $[n] \stackrel{\text{def}}{=} \{1, \dots, n\}$. The induced subgraphs $G|_{\tau_i}$ are called component subgraphs. In [11] we prove a series of theorems about these architectures, connecting the fixed points of a graph G to the fixed points of the component subgraphs $G|_{\tau_i}$. In the following, we give an overview of these architectures and summarize the key results proven about them in [11].

Cyclic unions. The most straightforward generalization of a cycle is the *cyclic union*, an architecture first introduced in [14]. Given a set of component subgraphs $G|_{\tau_1}, \dots, G|_{\tau_N}$, on subsets of nodes $\tau_1, \dots, \tau_N$, the *cyclic union* is constructed by connecting these subgraphs in a cyclic fashion so that there are edges forward from every node in τ_i to every node in τ_{i+1} (cyclically identifying τ_N with τ_0), and there are no other edges between components (see Figure 2A).

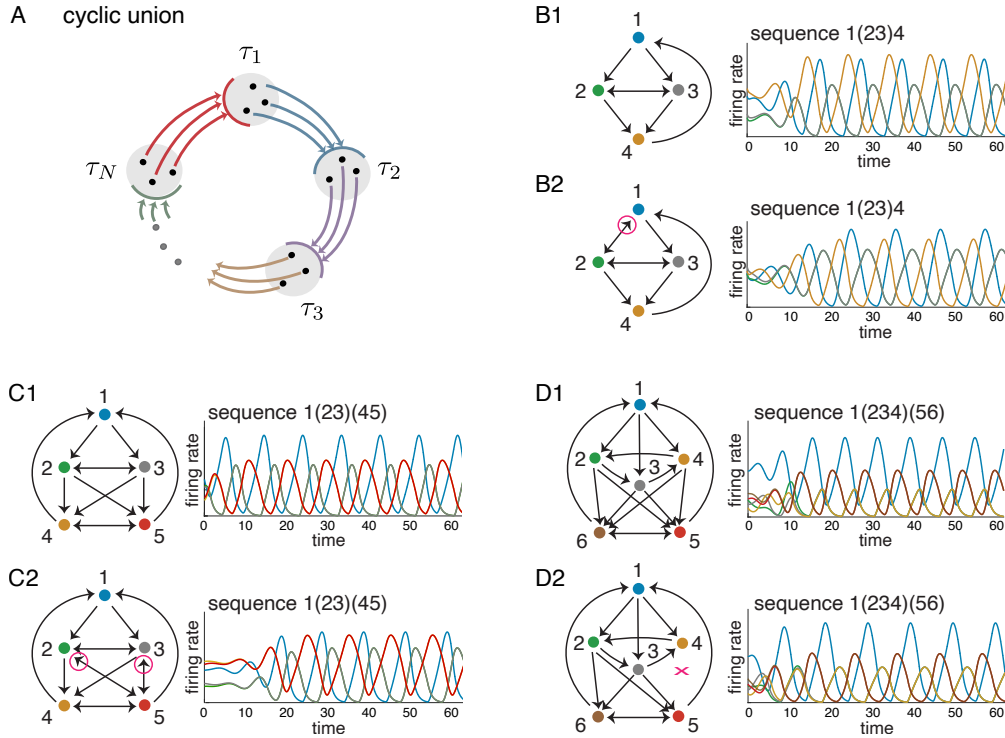


Figure 2: **Cyclic unions and related variations.** (A) A cyclic union has component subgraphs with subsets of nodes $\tau_1, \dots, \tau_N$, organized in a cyclic manner. (B1,C1,D1) Three cyclic unions with firing rate plots showing solutions to a corresponding CTLN. Above each solution the associated sequence of firing rate peaks is given, with synchronously firing neurons denoted by parentheses. (B2,C2,D2) These graphs are all variations on the cyclic unions above them, with some edges added or dropped (highlighted in magenta). Solutions of the corresponding CLTNs qualitatively match the solutions of the corresponding cyclic unions. In each case, the sequence is identical.

The top graphs in Figure 2B-D are examples of cyclic unions with three components. All the nodes at a given height comprise a τ_i component. Next to each graph is a solution to a corresponding CTLN, which is a global attractor of the network. Notice that the activity traverses the components in cyclic order. Cyclic unions are particularly well-behaved architectures where the fixed point supports can be fully characterized in terms of those of the components. Specifically, the fixed points of a cyclic union G are precisely the unions of supports of the component subgraphs, exactly one per component [14], that is,

$$\sigma \in \text{FP}(G) \iff \sigma \cap \tau_i \in \text{FP}(G|_{\tau_i}) \text{ for all } i \in [N].$$

The bottom graphs in Figure 2B-D have very similar dynamics to the ones above them, but do not have the perfect cyclic union structure. Despite deviations from the cyclic union architecture, these graphs produce sequential dynamics that similarly traverse the components in cyclic order. In fact, they are examples of a more general class of architecture: directional cycles.

Directional cycles. In a cyclic union, if we restrict to the subnetwork consisting of a pair of consecutive components, $G|_{\tau_i \cup \tau_{i+1}}$, we find that activity initialized on τ_i flows forward and ends up concentrated on τ_{i+1} . Moreover, the fixed points of $G|_{\tau_i \cup \tau_{i+1}}$ are all confined to live in τ_{i+1} , and so the concentration of neural activity coincides with the subnetwork supporting the fixed points. We say that a graph is *directional* whenever its fixed points are confined to a proper subnetwork. In simulations, we have seen that directional graphs have the desired directionality of neural activity, so that activity flows to the subnetwork supporting the fixed points.

We can chain together directional graphs in a cyclic manner to produce *directional cycles*. We predict that these graphs will have a cyclic flow to their dynamics, hitting each τ_i component in cyclic order. While we have not been able to explicitly prove this property of the dynamics, we can prove that all the fixed point supports have cyclic structure (Theorem 1.2 in [11]). However, unlike with cyclic unions, for directional cycles we do not necessarily have the property that fixed points σ of the full network restrict to fixed points $\sigma_i \stackrel{\text{def}}{=} \sigma \cap \tau_i$ of the component subnetworks $G|_{\tau_i}$. The key structural feature of cyclic unions that guarantees this property of the fixed points is what we call a *simply-embedded partition*.

Simply-embedded partitions. Given a graph G , a partition of its nodes $\{\tau_1 | \dots | \tau_N\}$ is called *simply-embedded* if all nodes in a component τ_i are treated identically by each node outside that component. For example, the pair of graphs in Figure 2C have $\{1 | 2, 3 | 4, 5\}$ as a simply-embedded partition: in each graph, nodes 2 and 3 receive identical inputs from 1, 4 and 5 and nodes 4 and 5 receive identical inputs from 1, 2 and 3. It turns out that this simply-embedded partition structure is sufficient to guarantee that all the fixed points of G restrict to fixed points of the component subgraphs. This means that the fixed points of the components provide a kind of “menu” from which the fixed points of G are made: each fixed point of the full network has support that is the union of component fixed point supports, at most one per component (Theorem 1.4 in [11]).

Simply-embedded directional cycles. While the simply-embedded partition generalizes one key property of $\text{FP}(G)$ from cyclic unions, it does not guarantee that every fixed point intersects every component, nor does it guarantee the cyclic flow of the dynamics through the components. But combining the previous two constructions, we immediately see that a *simply-embedded directional cycle* will have the desired fixed point properties while maintaining cyclic dynamics. In particular, every fixed point support of G is a union of (nonempty) component fixed point supports, exactly one per component (Theorem 1.5 in [11]).

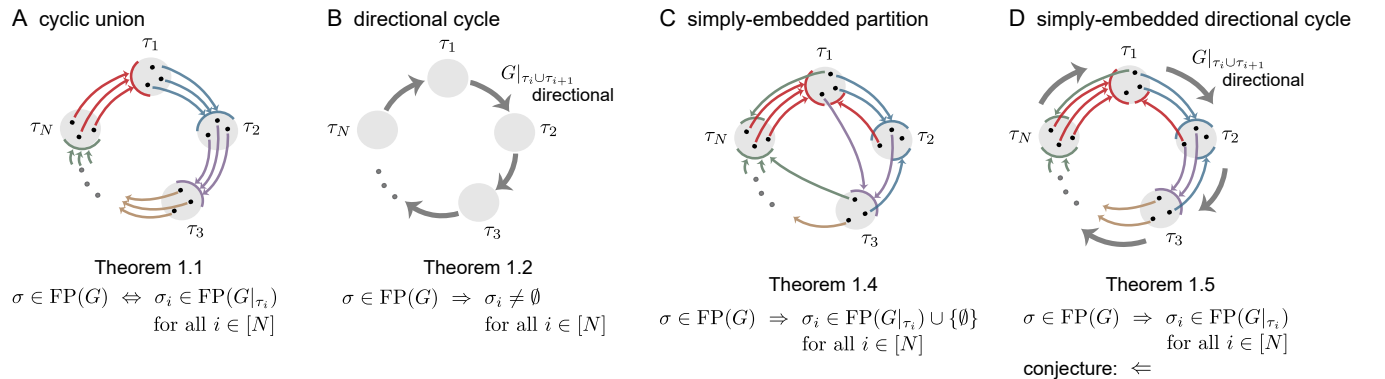


Figure 3: Summary of main results. In each graph, colored edges from a node to a component indicate that the node projects edges out to all the nodes in the receiving component, as needed for a simply-embedded partition. Thick gray edges indicate directionality of the subgraph $G|_{\tau_i \cup \tau_{i+1}}$. For $\sigma \subseteq [n]$, $\sigma_i \stackrel{\text{def}}{=} \sigma \cap \tau_i$. Theorem numbers match those in [11].

Can a single network support multiple types of sequential attractors?

The answer is yes! These new architectures can be embedded as subnetworks in a larger network, akin to the example in Figure 1D, with each subnetwork yielding a different sequential attractor. Figure 4A shows a network of eleven neurons made up of three subnetworks on overlapping nodes: $G|_{\{1,\dots,6\}}$ is the cyclic union from Figure 2 panel D2, $G|_{\{5,6,7,8\}}$ is the directional cycle from Figure 2 panel B2, and $G|_{\{4,9,10,11\}}$ is a 4-cycle (Figure 1B). The corresponding attractors can be activated by either of two types of external drive. In Figure 4B, we chose $\theta = 0$ as a baseline and step function pulses of $\theta_i = 1$, so that only a given subnetwork is receiving positive external drive at a time, and thus only that subnetwork can be active. The pulses activate sequences involving only the neurons in the stimulated subnetwork, which match precisely those in the isolated graphs. In Figure 4C, we see that it is also possible for the network to converge to each of the different attractors of these subnetworks even when *all* the neurons receive positive external drive. Additionally, the network can transition between these attractors simply by providing a short pulse to the relevant subnetwork. In both types of simulations, we see that the network has multiple *coexisting* sequential attractors. Moreover, the minimal fixed points of the network again reflect the subsets of neurons that are high firing in the attractors.

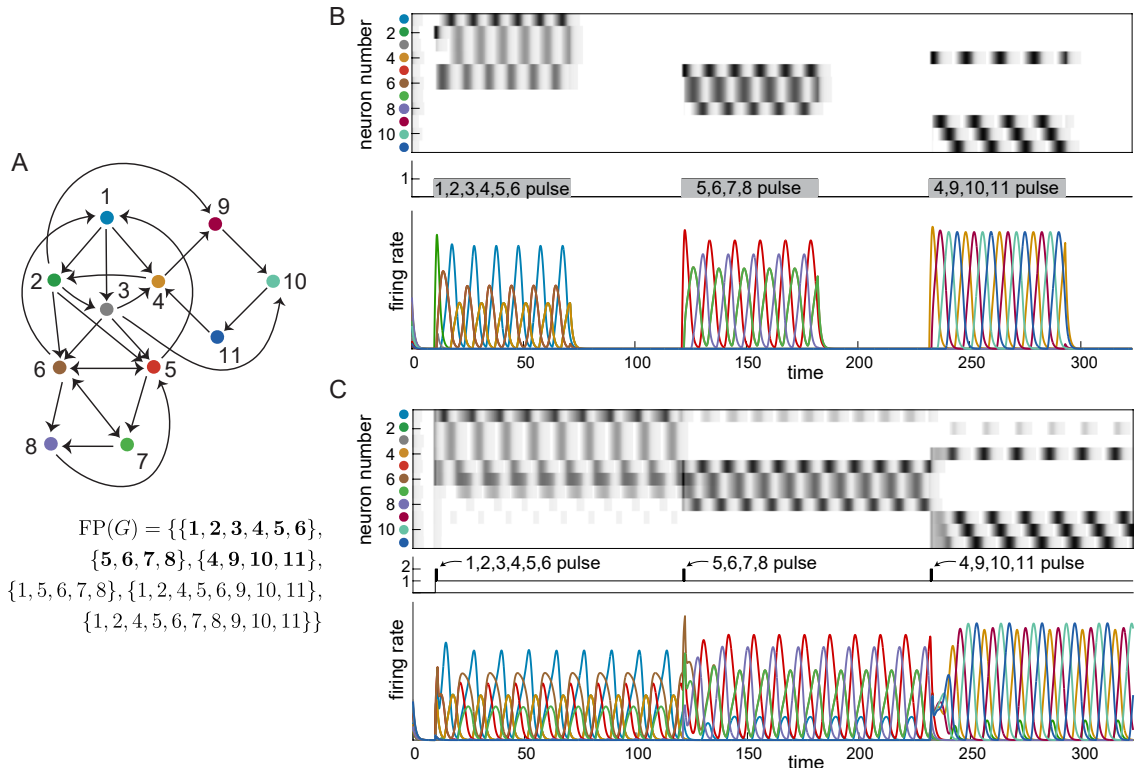


Figure 4: Coexistence of three sequential attractors in larger network. Three embedded attractors are transiently activated to produce sequences matching those of the isolated networks in Figures 1 and 2. (A) Three different example graphs are embedded in a single network. $FP(G)$ is shown, with minimal supports in bold. Notice that the minimal supports correspond exactly to the three subgraphs $G|_{\{1,\dots,6\}}$, $G|_{\{5,6,7,8\}}$, $G|_{\{4,9,10,11\}}$. (B) Long pulses with baseline $\theta = 0$ and pulse values $\theta_i = 1$ activate sequences exactly matching those of the isolated graphs. (C) Short pulses with baseline $\theta = 1$ and pulse values $\theta_i = 2$ activate sequences closely matching those of the isolated graphs.

Curious to know more?

The architectures shown in Figure 3 all give rise to sequential attractors and have been shown to have well-behaved fixed point structure. In [11], we further explore these architectures as well as other related network structures, prove theorems, and investigate related questions in depth. We explore sequential attractors of a variety of minimal subnetworks, known as *core motifs*, to find that the directional cycle structure is indeed predictive of the sequential activity, particularly when the directional cycle partition is also simply-embedded. From this analysis, we see that these architectures give insight into the sequences of neural activity that emerge, going beyond the structure of fixed point supports.

References

- [1] A. Luczak, P. Barthó, S.L. Marguet, G. Buzsáki, and K.D. Harris. Sequential structure of neocortical spontaneous activity *in vivo*. *Proc. Natl. Acad. Sci.*, 104(1):347–352, 2007.
- [2] L. Carillo-Reid, J.K. Miller, J.P. Hamm, J. Jackson, and R. Yuste. Endogenous sequential cortical activity evoked by visual stimuli. *J. Neurosci.*, 35(23):8813–8828, 2015.
- [3] R. Yuste, J.N. MacLean, J. Smith, and A. Lansner. The cortex as a central pattern generator. *Nat. Rev. Neurosci.*, 6:477–483, 2005.
- [4] E. Stark, L. Roux, R. Eichler, and G. Buzsáki. Local generation of multineuronal spike sequences in the hippocampal CA1 region. *Proc. Natl. Acad. Sci.*, 112(33):10521–10526, 2015.
- [5] E. Pastalkova, V. Itskov, A. Amarasingham, and G. Buzsáki. Internally generated cell assembly sequences in the rat hippocampus. *Science*, 321(5894):1322–1327, 2008.
- [6] V. Itskov, C. Curto, E. Pastalkova, and G. Buzsáki. Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus. *J. Neurosci.*, 31(8):2828–2834, 2011.
- [7] E. Marder and D. Bucher. Central pattern generators and the control of rhythmic movements. *Curr. Bio.*, 11(23):R986–996, 2001.
- [8] S. Grillner and P. Wallén. Cellular bases of a vertebrate locomotor system – steering, intersegmental and segmental co-ordination and sensory control. *Brain Res. Rev.*, 40:92–106, 2002.
- [9] M.A. Whittington, R.D. Traub, N. Kopell, B. Ermentrout, and E.H. Buhl. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol.*, 38(3):315–336, 2000.
- [10] P. Malerba, G.P. Krishnan, J-M. Fellous, and M. Bazhenov. Hippocampal CA1 ripples as inhibitory transients. *PLoS Comput Biol*, 12(4), 2016.
- [11] C. Parmelee, J. Londono Alvarez, C. Curto, and K. Morrison. Sequential attractors in combinatorial threshold-linear networks. *SIAM Journal on Applied Dynamical Systems*, 21(2):1597–1630, 2022.
- [12] K. Morrison and C. Curto. *Predicting neural network dynamics via graphical analysis*. Book chapter in Algebraic and Combinatorial Computational Biology, edited by R. Robeva and M. Macaulay. Elsevier, 2018.
- [13] C. Parmelee, S. Moore, K. Morrison, and C. Curto. Core motifs predict dynamic attractors in combinatorial threshold-linear networks. *PLOS ONE*, 17, 2022.
- [14] C. Curto, J. Geneson, and K. Morrison. Fixed points of competitive threshold-linear networks. *Neural Comput.*, 31(1):94–155, 2019.
- [15] C. Curto, J. Geneson, and K. Morrison. Stable fixed points of combinatorial threshold-linear networks. Available at <https://arxiv.org/abs/1909.02947>