

On the Transformational Complexity of Acyclic Networks of Neurons

UF CISE Technical Report REP-2011-517¹

April 14, 2011

Venkatakrishnan Ramaswamy and Arunava Banerjee

{vr1, arunava}@cise.ufl.edu

Computer and Information Science and Engineering,

University of Florida, Gainesville, FL 32611, USA.

Abstract

Given an acyclic network of neurons, we ask how the network's architecture and specific properties of its neurons constrain the computations it might be able to perform. We begin by setting up a framework that allows us to view computation in acyclic networks as transformations that map input spike-trains to output spike-trains. We then ask what transformations *all* acyclic networks of specific architectures cannot accomplish, that networks of other architectures can. Our neurons are abstract mathematical objects that satisfy a small number of axioms, which correspond to basic properties shared by a vast majority of biological neurons. We show some classes of architectures, in which these axioms constrain computation enough so that they cannot effect certain types of transformations, that other networks can. We also prove that for certain classes of architectures, more assumptions need to be made about the neurons, before results of this form are manifested. This type of theory might find application in extracting meaning from connectome data, where only network structure is available whereas information on dynamical properties of the neurons involved is not.

¹supercedes REP-2009-481.

1 Introduction

Neurons and their networks, fundamentally, are machines that transform spike-trains into spike-trains. It is these transformations that form the basis of information processing, indeed even cognition, in the brain.

This work is broadly motivated by the following question: What constraints, if any, do local properties of neurons impose on the computational abilities of various networks? At a coarse level, we would like to say something about what specific networks *cannot* do, by virtue of their architecture. Now the answer may depend on the specific neuron model in use. To mitigate this possibility, we instead assume that the neuron is an abstract object that satisfies certain axioms that correspond to well-known properties satisfied by a vast majority of biological neurons. This admits a variety of models as long as they satisfy the axioms (including ones that are more powerful than what is biologically possible). We then ask, if (in spite of this additional power), there exist specific transformations that cannot be effected by networks with certain architectures. In order to then rule out the prospect that the transformation in question is so hard that no network can do it, we also ask for a network (of a different architecture) comprising simple neurons that can in fact effect this transformation.

Since the functional role of single neurons and small networks in the brain is not yet well understood, we do not make any assumptions about particular “tasks” that the network is trying to perform; we are just interested in physical spike-train to spike-train transformations. Furthermore, since the kinds of neural code employed are still unclear, we make no overarching assumptions about the neural code either. We study precise spike times since there is widespread evidence (Strehler & Lestienne, 1986; Rieke et al., 1997) that precise spike time play a role in information processing in the brain. Studying spike times also subsumes cases where spiking rate may be the relevant parameter and therefore there is no loss of generality in making this assumption.

In this paper, as a first step, we restrict our study to acyclic² networks of neurons, i.e. networks that do not have a directed cycle. While even single neurons are func-

²while the term *feedforward* network is widely used in the literature to refer to this type of network, we prefer to call these acyclic networks to emphasize that these networks are not feedforward in the system-theoretic sense.

tionally recurrent³, acyclic networks quickly settle down to quiescence upon receiving no input. On the other hand, recurrent networks have been known (Banerjee, 2006) to have complex dynamics, in general, even on receiving no input for unbounded periods of time. Also, as a simplifying assumption, we do not treat synaptic plasticity effects in this paper.

Several researchers have studied related questions. In (Poirazi et al., 2003), Poirazi et al., model a compartmental model of a pyramidal neuron using a two layer neural network, assuming rate codes. Bohte et al., (Bohte et al., 2002) derive a supervised learning rule for a network of spiking neurons, where the output is restricted to a single spike in a given period of observation. Gutig and Sompolinsky (Gutig & Sompolinsky, 2006) describe a model to learn spike time decisions. They also have a task with two outcomes, which are mapped to notions of presence or absence of spikes. Maass (Maass, 1996) investigates the computational power of networks of neurons relating them to well-known models of machine computation such as Turing Machines and Random Access Machines. Finally, Bartlett and Maass (Bartlett & Maass, 2003) summarize some work that analyzes the discriminative capacity of a pulse-coded neuron from a statistical learning theory perspective.

In contrast to all the above approaches, we seek to investigate the relative complexity of the physical spike-train to spike-train transformations that are instantiated by systems of spiking neurons, without making overarching assumptions about the underlying computational dynamics of the system.

While we have done our best to keep the axioms of the single neuron as general and as biologically faithful as we could, we must remark that there are properties assumed that all biological neurons do not satisfy and known phenomena⁴ which we do not model. Also, as a first step, we have only treated the restricted case of acyclic networks here⁵. We therefore see this work as an attempt to ask if the type of approach espoused here is viable and can bear non-trivial fruit.

Roadmap. In Section 2 we introduce notation used in the remainder of the paper. In

³owing to the membrane potential also depending on past output spikes to account for effects during the absolute and relative refractory periods.

⁴Synaptic plasticity and gap junctions, for example.

⁵We are in the process of exploring ways to ask similar questions of recurrent networks.

Section 3, we describe our abstract mathematical model of the neuron and motivate its biological underpinnings. In Section 4, we find that even a single neuron cannot be consistently viewed as a spike-train to spike-train transformation, in general (in a sense that we will make precise). Notwithstanding this, in Section 5, we prove that, under conditions consistent with spiking regimes observed in-vivo, the aforementioned notions of transformations are indeed well-defined and correspond to ones mapping finite-length input spike-trains to finite-length output spike-trains. However, these conditions are also mathematically unwieldy when we try to ask what transformations networks with specific architectures cannot do. This motivates the development of a more tractable criterion, in Section 6; notably, in Section 7, we prove that this more tractable criterion can be used to prove every complexity result that is accessible to our more biologically-realistic criterion and therefore there is no loss of generality in using it. Armed with this framework, in Section 7, we then proceed to set up definitions that allow us to ask what transformations networks of specific network architectures *cannot* effect. Further, to avert the possibility that the transformation in question is so hard that no network can effect it, we stipulate that there must exist a network of a different architecture that can in fact perform this transformation. In Section 8, some results of this form are established; First, we show a transformation that a single neuron cannot effect but a network consisting of two neurons can. Next, we prove a result which shows that a class of architectures that share a certain abstract graph-theoretic property also share in their inability in effecting a particular class of transformations. Also, while this class of architectures may have arbitrarily many neurons, we show a class of networks with just two neurons which can effect this class of transformations. While attempting to ask how increase in depth of the network constrains the transformations it can effect, we discovered that the current abstract model of the neuron does not adequately constrain networks in this respect. That is, with the current model, we prove in Section 9 that every acyclic network, having arbitrary depth, has an equivalent acyclic network of depth equal to two that effects *exactly* the same transformation. The implication of this result is that we need to add more axioms to the present abstract model in order for such results to be manifested. Finally, we close in Section 10 with some concluding remarks and directions for future work.

2 Notation and Preliminaries

An *action potential* or *spike* is a stereotypical event characterized by the time instant at which it is initiated⁶ in the neuron, which is referred to as its *spike time*. Spike times are represented relative to the present by real numbers, with positive values denoting past spike times and negative values denoting future spike times⁷. A *spike-train* $\vec{x} = \langle x^1, x^2, \dots, x^k, \dots \rangle$ is an increasing sequence of spike times, with every pair of spike times being more than α apart, where $\alpha > 0$ is the absolute refractory period⁸ and x^i is the spike time of spike i . An *empty spike-train*, denoted by $\vec{\phi}$, is one which has no spikes. Let \mathcal{S} denote the set of all spike trains. A *time-bounded spike-train* (with *bound* $[a, b]$) is one where all spike times lie in the bounded interval $[a, b]$, for $a, b \in \mathbb{R}$. Note that, owing to the absolute refractory period, a time-bounded spike-train is also a finite-length sequence. Let $\bar{\mathcal{S}}_{[a,b]}$ denote the set of all time-bounded spike-trains with bound $[a, b]$. A spike-train is said to have a *gap* in the interval $[c, d]$, if it has no spikes in that time interval. Furthermore, this gap is said to be of *length* $d - c$. Two spike-trains are said to be *identical* in the interval $[a, b]$, if they have exactly the same spike times in that time interval.

We use the term *spike-train ensemble* to denote a collection of spike-trains. Thus, formally, a *spike-train ensemble* $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ is a tuple of spike-trains. The *order* of a spike-train ensemble is the number of spike-trains in it. For example, $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ is a spike-train ensemble of order m . A *time-bounded spike-train ensemble* (with *bound* $[a, b]$) is one in which each of its spike-trains is time-bounded (with *bound* $[a, b]$). Two spike-train ensembles are said to be *identical* in the interval $[a, b]$, if they have the same order and their corresponding spike-trains are identical in the same time interval. A spike-train ensemble χ is said to have a *gap* in the interval $[c, d]$, if each of its spike trains has a gap in the interval $[c, d]$.

Next, we define some operators to time-shift spike-trains/spike-train ensembles, create bounded-time spike-trains/spike-train ensembles from spike-trains/spike-train en-

⁶The exact location of spike initiation is not important as long as the spike time is measured as the time at which the spike passes through a certain fixed point in the axon, upon initiation.

⁷Note that this convention implies that larger a spike's spike time, the more ancient the spike is.

⁸We assume a single fixed absolute refractory period for all neurons, for convenience, although our results would be no different if different neurons had different absolute refractory periods.

sembles and to assemble/disassemble spike trains into/from spike-train ensembles. Let $\vec{x} = \langle x^1, x^2, \dots, x^k, \dots \rangle$ be a spike-train and $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ be a spike-train ensemble. The *shift operator for spike-trains* is defined as $\sigma_t(\vec{x}) = \langle x^1 - t, x^2 - t, \dots, x^k - t, \dots \rangle$. The *shift operator for spike-train ensembles* is defined as $\sigma_t(\chi) = \langle \sigma_t(\vec{x}_1), \dots, \sigma_t(\vec{x}_m) \rangle$. The *truncation operator for spike-trains* is defined as follows: $\Xi_{[a,b]}(\vec{x})$ is the time-bounded spike-train with bound $[a, b]$ that is identical to \vec{x} in the interval $[a, b]$. $\Xi_{(a,b]}(\vec{x})$ and $\Xi_{[a,b)}(\vec{x})$ are defined likewise. In the same vein, $\Xi_{[a,\infty)}(\vec{x})$ is the spike-train that is identical to \vec{x} in the interval $[a, \infty)$ and has no spikes in the interval $(-\infty, a)$. Similarly, $\Xi_{(-\infty,b]}(\vec{x})$ is the spike-train that is identical to \vec{x} in the interval $(-\infty, b]$ and has no spikes in the interval (b, ∞) . The *truncation operator for spike-train ensembles* is defined as $\Xi_{[a,b]}(\chi) = \langle \Xi_{[a,b]}(\vec{x}_1), \dots, \Xi_{[a,b]}(\vec{x}_m) \rangle$; $\Xi_{(a,b]}(\chi)$, $\Xi_{[a,b)}(\chi)$, $\Xi_{[a,\infty)}(\chi)$ and $\Xi_{(-\infty,b]}(\chi)$ are defined likewise. Furthermore, $\Xi_t(\cdot)$ is shorthand for $\Xi_{[t,t]}(\cdot)$. The *projection operator for spike-train ensembles* is defined as $\Pi_i(\chi) = \vec{x}_i$, where $1 \leq i \leq m$. Let $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n$ be spike-trains. The *join operator for spike-trains* is defined as $\vec{y}_1 \sqcup \vec{y}_2 \sqcup \dots \sqcup \vec{y}_n = \bigsqcup_{i=1}^n \vec{y}_i = \langle \vec{y}_1, \vec{y}_2, \dots, \vec{y}_n \rangle$.

3 The Abstract Model

3.1 Assumptions

Our neurons are abstract mathematical objects that are constrained to satisfy a small set of axioms. First, we informally describe the assumptions that underlie the axioms. Notable cases where the assumptions do not hold are also pointed out. This is followed by formal definitions. The approach taken here closely follows the one in (Banerjee, 2001).

The following are our assumptions:

1. We assume that the neuron is a device that receives input from other neurons exclusively by spikes which are received via chemical synapses.⁹
2. The neuron is a finite-precision device with fading memory. Hence, the underlying potential function can be determined¹⁰ from a bounded past. That is, we

⁹In this work, we do not treat electrical synapses or ephaptic interactions (Shepherd, 2004).

¹⁰We do not treat noise, in this paper.

assume that, for each neuron, there exist real numbers Υ and ρ , so that the current membrane potential of the neuron can be determined as a function of the input spikes received in the past Υ seconds and the spikes produced by the neuron in the past ρ seconds.¹¹

3. We assume that the membrane potential of the neuron can be written down as a real-valued, everywhere bounded function of the form $P(\chi; \vec{x}_0)$, where \vec{x}_0 is a time-bounded spike-train, with bound $[0, \rho]$ and $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ is a time-bounded spike-train ensemble with bound $[0, \Upsilon]$. Informally, \vec{x}_i , for $1 \leq i \leq m$, is the sequence of spikes afferent in synapse i in the past Υ seconds and \vec{x}_0 is the sequence of spikes efferent from the current neuron in the past ρ seconds. The function $P(\cdot)$ characterizes the entire spatiotemporal response of the neuron to spikes including synaptic strengths, their location on dendrites, and their modulation of each other's effects at the soma, spike-propagation delays, and the postspike hyperpolarization.
4. Without loss of generality, we assume the resting membrane potential to be 0. Let $\tau > 0$ be the threshold membrane potential. Observe that the model allows for variable thresholds, as long as the threshold itself is a function of spikes afferent in the past Υ seconds and spikes efferent from the present neuron in the past ρ seconds.
5. The neuron outputs a spike whenever $P(\cdot) = \tau$. Additionally, when a new output spike is produced, we assume that the membrane potential immediately goes below threshold. That is, the membrane potential function in the abstract model takes values that are at most that of the threshold. This assumption is made without loss of generality and helps ease the formulation of the abstract model.
6. Owing to the absolute refractory period $\alpha > 0$, no two input or output spikes can occur closer than α . That is, suppose $\vec{x}_0 = \langle x_0^1, x_0^2, \dots, x_0^k \rangle$, where $x_0^1 < \alpha$. Then $P(\chi; \vec{x}_0) < \tau$, for all "legal" χ .
7. Past output spikes have an inhibitory effect, in the following sense¹²:

¹¹ ρ corresponds to the notion of *relative refractory period*.

¹²This is violated, notably, in neurons that have a post-inhibitory rebound.

$$P(\chi; \vec{x}_0) \leq P(\chi; \vec{\phi}), \text{ for all "legal" } \chi \text{ and } \vec{x}_0.$$

8. Finally, on receiving no input spikes in the past Υ seconds and no output spikes in the past ρ seconds, the neuron settles to its resting potential. That is,

$$P(\langle \vec{\phi}, \vec{\phi}, \dots, \vec{\phi} \rangle; \vec{\phi}) = 0.$$

An *acyclic network of neurons*, informally, is a Directed Acyclic Graph where each vertex corresponds to an instantiation of the neuron model, with some vertices designated input vertices (which are placeholders for input spike-trains), and one neuron designated the output neuron. The *depth* of an acyclic network is the length of the longest path from an input vertex to the output vertex.

3.2 Formal Definitions

Next, we formalize the above notions into a rigorous definition of a neuron as an abstract mathematical object which is followed by some more definitions. The reader may skim these on the first reading and revisit them if a specific technical point needs clarification later on.

Definition 1 (Neuron). A *neuron* N is a 7-tuple $\langle \alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau] \rangle$, where $\alpha, \Upsilon, \rho, \tau \in \mathbb{R}^+$ with $\rho \geq \alpha$, $\lambda \in \mathbb{R}^-$ and $m \in \mathbb{Z}^+$. Furthermore,

1. If $\vec{x}_0 = \langle x_0^1, x_0^2, \dots, x_0^k \rangle$ with $x_0^1 < \alpha$, then $P(\chi; \vec{x}_0) < \tau$, for all $\chi \in \bar{\mathcal{S}}_{[0, \Upsilon]}^m$ and for all $\vec{x}_0 \in \bar{\mathcal{S}}_{[0, \rho]}$.
2. $P(\chi; \vec{x}_0) \leq P(\chi; \vec{\phi})$, for all $\chi \in \bar{\mathcal{S}}_{[0, \Upsilon]}^m$ and for all $\vec{x}_0 \in \bar{\mathcal{S}}_{[0, \rho]}$.
3. $P(\langle \vec{\phi}, \vec{\phi}, \dots, \vec{\phi} \rangle; \vec{\phi}) = 0$.

A neuron is said to *generate a spike* whenever $P(\cdot) = \tau$.

Definition 2 (Network Architecture). A *Network Architecture* \mathcal{A} is a 5-tuple $\langle G \langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ where

$G\langle V, E \rangle$ is a directed graph¹³ with input vertices $I \subset V$, an¹⁴ output vertex $o \in V$, a bijection $\# : \{1, 2, \dots, |I|\} \rightarrow I$ that numbers the input vertices and a bijection $\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V$, for each vertex $v \in (V - I)$, that numbers the vertices incident on the in-edges of v . Furthermore,

1. For each $v \in I$, $\text{indegree}(v) = 0$.
2. $\text{outdegree}(o) = 0$.

Definition 3 (Acyclic Network Architecture). A Network Architecture $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ is called an *Acyclic Network Architecture* if $G\langle V, E \rangle$ is an acyclic graph.

Definition 4 (Network). A *Network* \mathcal{N} is a 3-tuple $\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ where $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ is a network architecture, N is a set of neurons and $L : (V - I) \rightarrow N$ is a bijection that maps non-input vertices of G to neurons in N such that the following is true: For every $v \in (V - I)$, that has $L(v) = \mathbf{N}\langle \alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau] \rangle$, we have $\text{indegree}(v) = m$. Further, $L(o)$ is called the *output neuron*.

Definition 5 (Acyclic Network). An *Acyclic Network* $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ is one whose network architecture \mathcal{A} is acyclic.

Definition 6. In an acyclic network $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$, the *depth* of a neuron $\mathbf{N} \in N$ is the length of the longest path from an input vertex to $L^{-1}(\mathbf{N})$.

Definition 7. The *depth* of an acyclic network $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ is the depth of its output neuron.

4 Acyclic Networks as Input-Output transformations

As mentioned earlier, we wish to look at acyclic networks of neurons as transformations that map input spike-trains to output spike-trains. Therefore, we first need to clarify in

¹³A directed edge represents the fact that there is *at least* one synapse from the neuron represented by the outgoing vertex to the neuron represented by the incoming vertex. The $P(\cdot)$ function of the efferent neuron gets only one spike-train from the afferent neuron and encodes the response of all the synapses from the afferent neuron.

¹⁴In this work, we only study networks with a single output neuron.

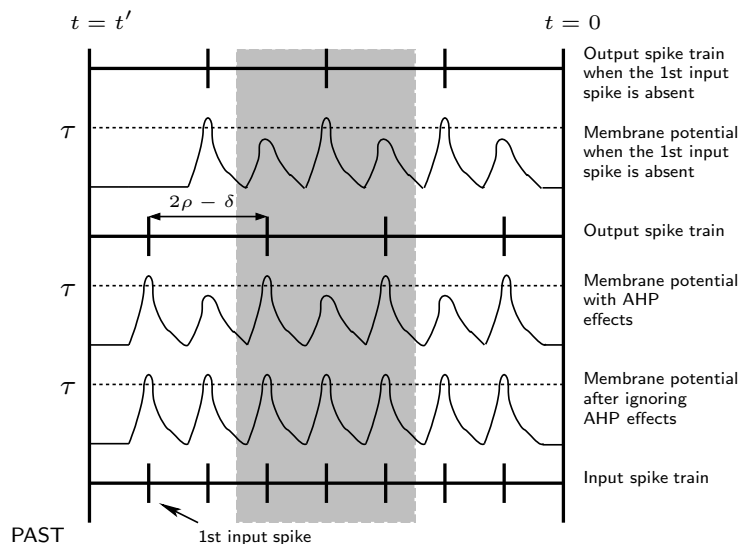


Figure 1: This counterexample describes a single neuron which has just one afferent synapse. Until time t' in the past, it received no input. After this time, its input was spikes that arrived every $\rho - \delta/2$ seconds, where $\delta > 0$. An input spike alone (if there were no output spikes in the past ρ seconds) can cause this neuron to produce an output spike. However, if there were an output spike within the past ρ seconds, the afterhyperpolarization (AHP) due to that spike is sufficient to bring the potential below threshold, so that the neuron does not spike currently. We therefore observe that if the first spike of the input spike-train is absent, then the output spike-train changes drastically. Note that this change occurs no matter how often the shaded segment in the middle is repeated, i.e. it does not depend on how long ago the first spike occurred.

what sense, if at all, these networks constitute the said transformations.

Let us first consider the simplest acyclic network, namely the single neuron. Given that our abstract neuron model does not explicitly prescribe an output spike-train for a given input spike-train, we need to ask what it means for a neuron to *produce* an output spike train, when supplied with an input spike-train. Recall, from the previous section, that the membrane potential of the neuron depends not only on the input spikes received in the past Υ seconds, it also depends on the output spikes produced by it in the past ρ seconds. Therefore, knowledge of just input spikes in the past Υ seconds does not uniquely determine the current membrane potential (and therefore the output spike-train produced from it). It might be tempting to then somehow use the fact that

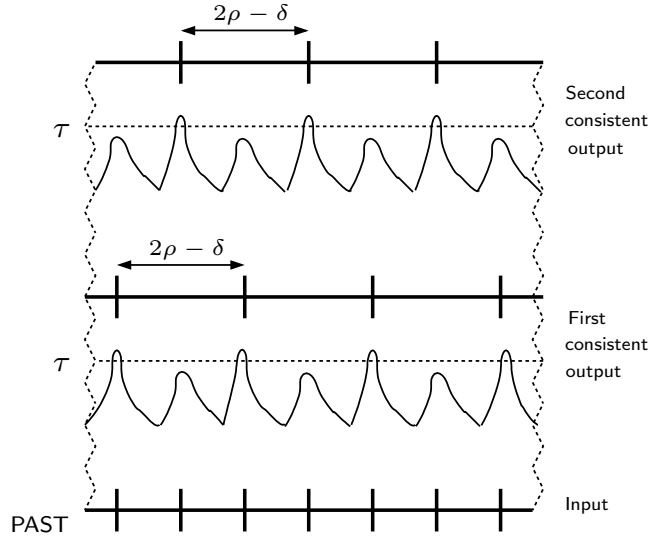


Figure 2: The counterexample here is very similar to the one in Figure 1, except that, instead of there being no input spikes before t' , we have an unbounded input spike-train ensemble, with the same periodic input spikes occurring since the infinite past. The neuron here has the exact same response as the one in Figure 1. Observe that both output spike-trains are consistent with this input, for each $t \in \mathbb{R}$.

the past output spikes are themselves a function of input and output received in the more distant past, and attempt to make the membrane potential a function of a bounded albeit larger “window” of input spikes alone. The counterexample described in Figure 1 shows that this does not work. In particular, the current membrane potential of the neuron may depend on the position of an input spike that has occurred arbitrarily long time ago in the past (if we attempt to characterize the membrane potential as a function of input spikes alone). One is then forced to ask if given the infinite history of input spikes received by the neuron, the membrane potential is then uniquely determined. Before we can answer this question, we need to rigorously define when we can consistently associate an (unbounded) output spike-train with an (unbounded) input spike-train ensemble, for a single neuron.

Definition 8. An output spike-train \vec{x}_o is said to be *consistent* with an input spike-train ensemble χ , with respect to a neuron $N\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau]\rangle$, if the following holds. For every $t \in \mathbb{R}$, $t \in \vec{x}_o$ if and only if

$$P(\Xi_{[0,\tau]}(\sigma_t(\chi)), \Xi_{[0,\rho]}(\sigma_t(\vec{x}_0))) = \tau.$$

The question we posed now becomes equivalent to the following : For every (unbounded) input spike-train ensemble χ , does there exist exactly one (unbounded) output spike train \vec{x}_o , so that \vec{x}_o is consistent with χ with respect to a given neuron N? Interestingly, we find that the answer is still in the negative. The counterexample in Figure 2 describes a neuron and an infinitely long input spike-train, which has two consistent output spike-trains.

Now, it could be argued that the input spike-train cannot possibly be infinitely long, since every neuron begins existence at a certain point in time. However, this begs the question of if the neuron was at the resting potential when the first input spikes arrived¹⁵. An assumption to this effect would be significant, particularly if the current membrane potential depended on it. It is easy to construct an example along the lines of the example described in Figure 1, where the current membrane potential is different depending on whether this assumption is made or not. Assuming infinitely long input spike-train ensembles, on the other hand, obviates the need to make any such assumption. We retain this viewpoint for the rest of the paper with the understanding that the alternative viewpoint discussed at the beginning of this paragraph can also be expounded along similar lines.

Nevertheless, the underlying difficulty in defining even single neurons as spike-train to spike-train transformations, with both viewpoints discussed above is dependence, in general, of current membrane potential on “initial state”. However, this still leaves open the possibility of considering just a subset of input/output spike trains, which have the property of the current membrane potential being independent of the input spike-train beyond a certain time in the past. Such a subset would exclude the examples discussed in this section. The caveat, of course, is that even if such a subset exists, unless it is also biologically well-motivated, claims of subsequent results having biological relevance are questionable.

In what follows, we come up with a biologically-realistic condition that implies independence as alluded to above; roughly speaking, the condition is that if a neuron has had a recent gap in its output spike-train equal to atleast *twice* its relative refractory

¹⁵Note that our axiomatic definition of a neuron does not address this question.

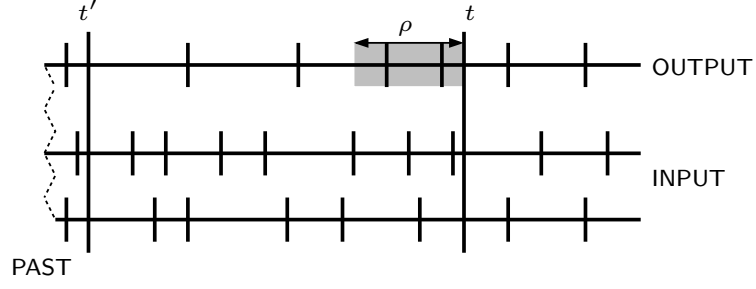
period, then its current membrane potential is independent of the input beyond the relatively recent past. We show that this leads to the notion of spike-train to spike-train transformations to be well-defined for acyclic networks.

5 The Gap Lemma and Criteria

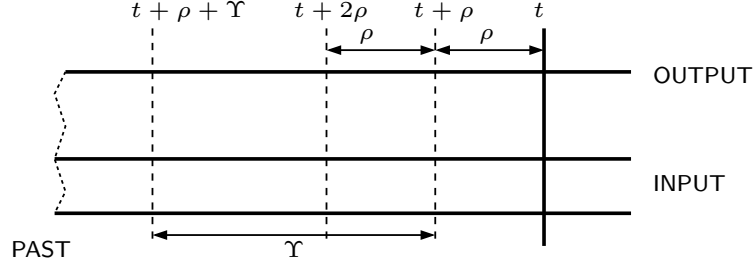
In this section, we devise a biologically well-motivated condition that guarantees independence of current membrane potential from input spikes beyond the recent past. This condition is used in constructing a criterion for single neurons which when satisfied, guarantees a unique consistent (unbounded) output spike-train. After this, similar criteria are defined for acyclic networks.

For a neuron, the way input spikes that happened sufficiently earlier affect current membrane potential is via a causal sequence of output spikes, causal in the sense that each of them had an effect on the membrane potential while the subsequent one in the sequence was being produced. The condition in the Gap Lemma basically seeks to break this causal chain.

To see the main idea that leads to the condition, suppose there exists a neuron, with Υ and ρ being the lengths of input and output windows respectively, that “effects” the transformation shown in Figure 3(a). Let $(t' - t) \geq \Upsilon$. Suppose, the spikes in the shaded region (which is an interval of length ρ) occurred at the exact same position, for all input spike-train ensembles that are identical in the range $[t, t']$, but have spikes occurring at arbitrary positions older than time instant t' . Then, the membrane potential of that neuron at t is identical in all those cases. Furthermore, the spikes in the shaded region are a function of exactly the input spikes in the interval $[t, t']$; in particular, they are independent of input spikes occurring before t' . The question is what condition will guarantee such a situation. Let us suppose we choose the shaded region to have no spikes, i.e. to be a *gap* of length ρ . Now, just requiring this does not imply that this gap is preserved when input spikes sufficiently old are changed. This is clear from the counterexample that was described in Figure 1, where an input spike arbitrarily old when removed, causes a spike, where a gap of length atleast ρ had existed previously. We need some additional constraints. How about if the gap were longer? Curiously, we find that a gap of 2ρ instead suffices, as the next lemma shows. That is, if the neuron



(a) The idea behind the Gap Lemma.



(b) Illustration accompanying the Gap Lemma.

Figure 3:

effects a transformation with a 2ρ gap, say ending at t , present in the output, then there exists a t' , such that no matter how input spikes older than t' are changed, the latter half of the 2ρ gap is guaranteed to have no spikes in each case. Therefore, membrane potential starting at t , is the same in all such cases. 2ρ is also the smallest gap for which this works. The details are in the following lemma. Figure 3(b) accompanies the lemma.

Lemma 1 (Gap Lemma). Consider a neuron $\mathbf{N}(\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau])$, a spike-train ensemble χ^* of order m and a spike-train \vec{x}_0^* which has a gap in the interval $[t, t + 2\rho]$, so that \vec{x}_0^* is consistent with χ^* , with respect to \mathbf{N} . Let χ be an arbitrary spike-train ensemble that is identical to χ^* in the interval $[t, t + \Upsilon + \rho]$.

Then, every output spike-train consistent with χ , with respect to \mathbf{N} , has a gap in the interval $[t, t + \rho]$. Furthermore, 2ρ is the smallest gap length in \vec{x}_0^* , for which this is true.

Proof. Since, in each \vec{x}_o consistent with χ , the interval $[t + 2\rho, t + 3\rho]$ of \vec{x}_o and the $[t + \Upsilon + \rho, t + \Upsilon + 2\rho]$ of χ are arbitrary, the sequence of spikes present in the interval $[t + \rho, t + 2\rho]$ of \vec{x}_o could be arbitrary. However, χ^* and χ are identical in

$[t, t + \rho + \Upsilon]$. Thus, it follows from Axiom 2 in the formal definition of a neuron that for every $t' \in [t, t + \rho]$, $P(\Xi_{[0, \Upsilon]}(\sigma_{t'}(\chi)), \Xi_{[0, \rho]}(\sigma_{t'}(\vec{x}_0)))$ is at most the value of $P(\Xi_{[0, \Upsilon]}(\sigma_{t'}(\chi^*)), \Xi_{[0, \rho]}(\sigma_{t'}(\vec{x}_0^*)))$, because $\Xi_{[0, \rho]}(\sigma_{t'}(\vec{x}_0^*))$ is $\vec{\phi}$, i.e. empty. Since $P(\Xi_{[0, \Upsilon]}(\sigma_{t'}(\chi^*)), \Xi_{[0, \rho]}(\sigma_{t'}(\vec{x}_0^*)))$ is less than τ for every $t' \in [t, t + \rho]$, $P(\Xi_{[0, \Upsilon]}(\sigma_{t'}(\chi)), \Xi_{[0, \rho]}(\sigma_{t'}(\vec{x}_0)))$ is less than τ in the same interval, as well. Therefore, \vec{x}_o has no output spikes in $[t, t + \rho]$.

That 2ρ is the smallest possible gap length in \vec{x}_o^* for this to hold, follows from the example in Figure 1, where the present conclusion did not hold, when \vec{x}_o^* had gaps of length $2\rho - \delta$, for arbitrary $\delta > 0$.

□

Corollary 1. Consider a neuron $N\langle \alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau] \rangle$, a spike-train ensemble χ^* of order m and a spike-train \vec{x}_0^* which has a gap in the interval $[t, t + 2\rho]$ so that \vec{x}_0^* is consistent with χ^* , with respect to N . Then

1. Every \vec{x}_0 consistent with χ^* , with respect to N , has a gap in the interval $[t, t + \rho]$.
2. Every \vec{x}_0 consistent with χ^* , with respect to N , is identical to \vec{x}_0^* in the interval $(-\infty, t + \rho]$, i.e. into the future starting from $t + \rho$.
3. For every t' more recent than $(t + \rho)$, the membrane potential at t' , is precisely a function of spikes in $\Xi_{[t', t + \Upsilon + \rho]}(\chi^*)$.

Proof sketch. (1) follows immediately from the Gap Lemma, when we set $\chi = \chi^*$. To see that (2) and (3) hold, note that for each time instant beyond t , the spikes in the input and output window are identical and therefore $P(\cdot)$ yields the same membrane potential in each such time instant. Formally,¹⁶ (2) can be proved by strong induction on the number of spikes since t and (3) follows from (2) and the Gap Lemma.

□

The upshot of the Gap Lemma and its corollary is that whenever a neuron goes through a period of time equal to twice its relative refractory period where it has produced no output spikes, its membrane potential from then on becomes independent of input spikes that are older than $\Upsilon + \rho$ seconds before the end of the gap.

¹⁶A more formal proof is available in the Appendix.

Large gaps in the output spike-trains of neurons seem to be extensively prevalent in the human brain. In parts of the brain where the neurons spike persistently, such as in the frontal cortex, the spike rate is very low (0.1Hz-10Hz) (Shepherd, 2004). In contrast, the typical spike rate of retinal ganglion cells can be very high but the activity is generally interspersed with large gaps during which no spikes are emitted (Nirenberg et al., 2001).

These observations motivate our definition of a criterion for input spike-train ensembles afferent on single neurons. The criterion dictates that there be intermittent gaps of length atleast twice the relative refractory period in an output spike-train consistent with the spike train ensemble.

Definition 9 (Gap Criterion for a single neuron). For $T \in \mathbb{R}^+$, a spike-train ensemble χ is said to satisfy a T -Gap Criterion for a neuron $N\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau]\rangle$ if the following is true: χ is of order m and there exists a spike-train \vec{x}_0 with at least one gap of length 2ρ in every interval of time of length $T - \Upsilon + 2\rho$, so that \vec{x}_0 is consistent with χ with respect to N .

Such input spike-train ensembles also have exactly one consistent output spike-train, as shown next.

Proposition 1. *Let χ be a spike-train ensemble that satisfies a T -Gap criterion for a neuron $N\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau]\rangle$. Then, there is exactly one spike-train \vec{x}_0 , such that \vec{x}_0 is consistent with χ , with respect to N .*

Proof. Since χ satisfies a T -Gap criterion, there exists a spike-train \vec{x}_0 with at least one gap of length 2ρ in every interval of time of length $T - \Upsilon + 2\rho$, so that \vec{x}_0 is consistent with χ with respect to N . For the sake of contradiction, assume that there exists another spike-train \vec{x}'_0 , not identical to \vec{x}_0 , which is consistent with χ , with respect to N . Let t' be the time at which one spike-train has a spike but another doesn't. Let $t > t'$ be such that \vec{x}_0 has a gap in the interval $[t, t + 2\rho]$. By Corollary 1 to the Gap Lemma, it follows that \vec{x}'_0 is identical to \vec{x}_0 starting from time instant $t + \rho$. This contradicts the hypothesis that \vec{x}'_0 is different from \vec{x}_0 at t' . □

For an input spike-train ensemble χ that satisfies a T -Gap criterion for a neuron, the membrane potential at any point in time is dependent on atmost T seconds of input

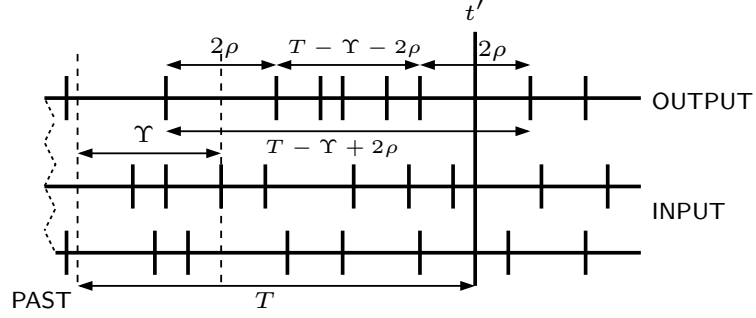


Figure 4: Illustration demonstrating that for an input spike configuration χ that satisfies a T -Gap criterion, the membrane potential at any point in time is dependent on at most T seconds of input spikes in χ before it.

spikes in χ before it. This can be seen from Figure 4, which illustrates a section of the input spike train ensemble and the output spike-train. Because of the T -Gap criterion the distance between any two gaps of length 2ρ on the output spike-train is at most $T - \Upsilon - 2\rho$. Up to the earlier half of a 2ρ gap (whose latest point is denoted by t') is dependent on input corresponding to the previous 2ρ gap. The membrane potential at t' depends on input spikes in the interval of length T , as depicted.

With inputs that satisfy the T -Gap Criterion, here is what we need to do to physically determine the current membrane potential, even if the neuron has been receiving input since the infinite past: Start off the neuron from an arbitrary state, and drive it with input that the neuron received in the past T seconds. The Gap Lemma guarantees that the membrane potential we see now will be identical to the actual membrane potential.

The Gap Criterion we have defined for single neurons can be naturally extended to acyclic networks. The criterion is simply that the input spike-train ensemble to the network is such that every neuron's input obeys a scaled Gap criterion for single neurons.

Definition 10 (Gap Criterion for an acyclic network). Consider an acyclic network $\mathcal{N} \langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$. Let d be the depth of \mathcal{N} . Let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i . For $T \in \mathbb{R}^+$, a spike-train ensemble χ is said to satisfy a T -Gap Criterion for the acyclic network \mathcal{N} if χ is of order $|I|$ and the following are true:

1. For each $N \in N_1$, the input spike-train ensemble received by N , namely

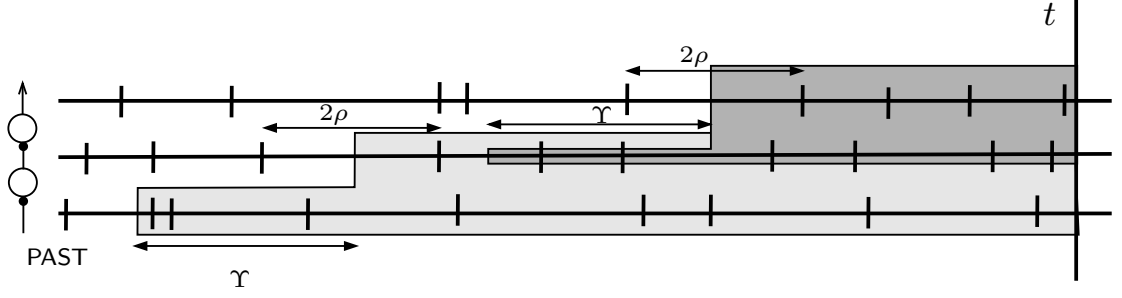


Figure 5: Schematic diagram illustrating how the Gap criterion works for the simple two-neuron network on the left. The membrane potential of the output neuron at t depends on input received from the “intermediate” neuron, as depicted in the darkly-shaded region, owing to the Gap Lemma. The output of the intermediate neuron in the darkly-shaded region, in turn, depends on input it received in the lightly-shaded region. Thus, transitively, membrane potential of the output neuron at t is dependent atmost on input received by the network in the lightly-shaded region.

$$\bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi), \text{ satisfies a } \left(\frac{T}{d}\right)\text{-Gap Criterion for } N, \text{ where } v = L^{-1}(N).$$

2. For $i = 2, \dots, d$ respectively, for each $N \in N_i$, the input spike-train ensemble received by N satisfies a $\left(\frac{T}{d}\right)$ -Gap Criterion for N .

As with the criterion for the single neuron, the membrane potential of the output neuron at any point is dependent on atmost T seconds of past input, if the input spike-train ensemble to the acyclic network satisfies a T -Gap criterion. The situation is illustrated in Figure 5. Additionally, the output spike-train is unique.

Lemma 2. Consider an acyclic network $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$. Let χ satisfy a T -Gap criterion for \mathcal{N} . Then \mathcal{N} produces a unique output spike-train when it receives χ as input. Furthermore, the membrane potential of the output neuron at any time instant depends on atmost the past T seconds of input in χ .

Proof sketch. That the output spike-train is unique can be proved by a strong induction argument, upon noting that the output spike train of each neuron is unique, since it satisfies a Gap criterion for a neuron.

To prove that membrane potential at any point is dependent on at most T seconds of past input, we essentially formalize the observations made in Figure 5. The full proof is in the Appendix. \square

We are thus at a juncture where questions we initially asked can even be posed in a rigorous manner. So looking back at the big picture, we had initially wished to see acyclic networks as transformations that mapped finite-length input spike-trains to finite-length output spike trains. However, we found that this notion was not always well-defined. We then showed that if we restrict the set of input spike-trains so they satisfied certain criteria, one can consistently speak of output spike-trains that such inputs are mapped to, by the acyclic network in question. We also noted that this restricted set of spike-trains seems to encompass biologically-relevant spiking regimes. Thus, acyclic networks can be seen as transformations that map this restricted set of input spike-trains to output spike trains. Indeed, this will be the sense in which acyclic networks are treated as transformations henceforth. Next, we formalize these observations.

Given an acyclic network \mathcal{N} , let $\mathcal{G}_{\mathcal{N}}^T$ be the set of all input spike-train ensembles that satisfy a T -Gap Criterion for \mathcal{N} . Let $\mathcal{G}_{\mathcal{N}} = \bigcup_{T \in \mathbb{R}^+} \mathcal{G}_{\mathcal{N}}^T$. Therefore, every acyclic network \mathcal{N} induces a transformation $\mathcal{T}_{\mathcal{N}} : \mathcal{G}_{\mathcal{N}} \rightarrow \mathcal{S}$ that maps each spike-train ensemble in $\mathcal{G}_{\mathcal{N}}$ to a unique output spike train in \mathcal{S} . Suppose $\mathcal{G}' \subseteq \mathcal{G}_{\mathcal{N}}$. Then, let $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}'} : \mathcal{G}' \rightarrow \mathcal{S}$ be the map defined as $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}'}(\chi) = \mathcal{T}_{\mathcal{N}}(\chi)$, for all $\chi \in \mathcal{G}'$.

The Gap Criteria are very general and biologically well-motivated. However, given a neuron or an acyclic network, there does not appear to be an easy way to characterize all the input spike-train ensembles that satisfy a certain Gap Criterion for it. So, for a neuron, whether an input spike-train ensemble satisfies a Gap Criterion for it seems to depend intimately on the exact form of its membrane potential function. That is, a spike-train ensemble that satisfies a Gap criterion for one neuron may not satisfy any Gap Criterion for another neuron. For an acyclic network, the problem becomes even more difficult, since intermediate neurons must satisfy Gap Criteria, and also produce output spike-trains that satisfy Gap Criteria for neurons further downstream. Furthermore, in order to compare transformations effected by two different networks, we need to study inputs that satisfy some Gap criterion for both of them, for otherwise the notion of a transformation may no longer hold. Now, we sought to ask what transformations *all* acyclic networks with a certain architecture could not do. For this, we need to

characterize inputs that satisfy a Gap Criterion for all the networks involved, which seems to be an especially formidable, even intractable, problem.

This brings up the question of the existence of another criterion according to which the set of spike-train ensembles is easier to characterize and is *common* across different networks. Next, we propose one such criterion and show that it consists of spike-train ensembles which are a subset of those induced by the Gap criteria for all acyclic networks. Loosely speaking, these are input spike-train ensembles which, before a certain time instant in the past, have had no spikes. The spike-train ensembles satisfying the said criterion, which we call the Flush criterion, allow us to sidestep the difficult issues just discussed. While this is a purely theoretical construct with no claim of biological relevance, in Section 7, we show that there is no loss by restricting ourselves to the Flush criterion. That is, not only is a complexity result proved using the Flush criterion applicable with the Gap criterion, every complexity result true with the Gap criterion can be proved by using the Flush criterion exclusively.

6 Flush Criterion

The idea of the Flush Criterion is to force the neuron to produce no output spikes for sufficiently long so as to guarantee that a Gap criterion is being satisfied. This is done by having a semi-infinitely long interval with no input spikes. This “flushes” the neuron by bringing it to the resting potential and keeps it there for an appropriately long time, during which it produces no output spikes. In an acyclic network, the flush is propagated so that all neurons have had a sufficiently long gap in their output spike-trains. Note that the Flush Criterion is not defined with reference to any acyclic network. We formalize this notion below.

Definition 11 (Flush Criterion). A spike-train ensemble χ is said to satisfy a *T-Flush Criterion*, if all its spikes lie in the interval $[0, T]$, i.e. it has no spikes before time instant T and after time instant 0.

First, we show that an input spike-train ensemble to a neuron that satisfies a Flush criterion also satisfies a Gap criterion. Figure 6 accompanies the following lemma.

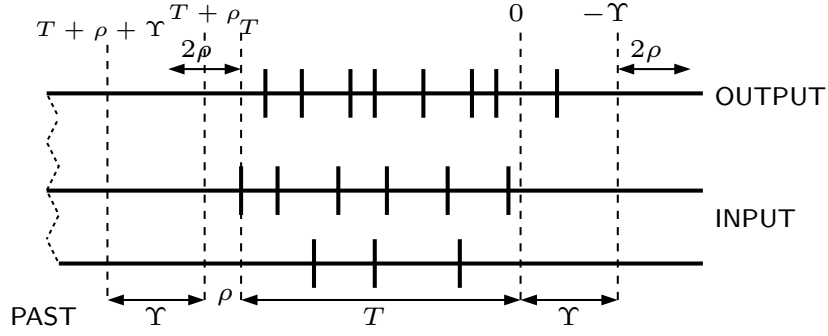


Figure 6: Illustration showing that an input spike configuration satisfying a Flush Criterion also satisfies a Gap Criterion.

Lemma 3. *An input spike-train ensemble χ for a neuron that satisfies a T -Flush Criterion also satisfies a $(T + 2\Upsilon + 2\rho)$ -Gap Criterion for that neuron.*

Proof. The neuron on being driven by χ cannot have output spikes outside the interval $[-\Upsilon, T]$. This easily follows from Axiom 2 and 3 of the neuron because the neuron does not have input spikes before time instant T and in the interval $[-\Upsilon, 0]$ and onwards. Now to see that χ satisfies a $(T + 2\Upsilon + 2\rho)$ -Gap Criterion, recall that with a T' -Gap Criterion, distance between any two gaps of length 2ρ on the output spike-train is at most $T' - \Upsilon - 2\rho$. With χ , we observe that the distance between any two 2ρ gaps on the output spike-train is at most $T + \Upsilon$. Thus, $T' - \Upsilon - 2\rho = T + \Upsilon$, which gives us $T' = T + 2\Upsilon + 2\rho$. The result follows. \square

Next, it is shown that an input spike-train ensemble to an acyclic network satisfying a Flush criterion also satisfies a Gap criterion for that network.

Lemma 4. *An input spike-train ensemble χ for an acyclic network that satisfies a T -Flush Criterion also satisfies a $(dT + d(d+1)\Upsilon + 2d\rho)$ -Gap Criterion for that network, where Υ, ρ are upper bounds on the same parameters taken over all the neurons in the network and d is the depth of the network.*

Proof. Following the proof of the previous lemma, we know that neurons that receive all their inputs from χ have no output spikes outside the interval $[-\Upsilon, T]$. Similarly, neurons that have depth 2 with respect to the input vertices of the network have no

output spikes outside $[-2\Upsilon, T]$. Likewise, the output neuron, which has depth d , has no output spikes outside $[-d\Upsilon, T]$. It follows that the output neuron obeys a $(T + (d + 1)\Upsilon + 2\rho)$ -Gap Criterion. Also, every other neuron follows this criterion because the distance between the 2ρ output gaps for every neuron is at most that of the output neuron, since their depth is bounded from above by the depth of the output neuron. Thus, from the definition of the Gap criterion for acyclic networks, we have that χ satisfies a $(dT + d(d + 1)\Upsilon + 2d\rho)$ -Gap Criterion for the current network. \square

The Flush criterion is a construct made for mathematical expedience and *prima facie* does not have any biological relevance. It is a network-independent criterion which enables us to circumvent difficulties that working with the Gap criterion entailed. It will soon become clear why it is a useful construction, when we show that it is equivalent to the Gap criterion insofar as the questions we seek to ask are concerned.

7 Transformational Complexity

In this section, we define notions of relative complexity of sets of acyclic networks of neurons, with respect to transformations effected by them. For brevity, we refer to these notions as *Transformational Complexity*. What we would like to capture with the definition is the following: Given two classes of networks with the second class encompassing the first, we wish to ask if there are networks in the second class whose transformation cannot be performed by any network in the first class. That is, do the extra networks in the second class make it richer in terms of transformational power? The classes could correspond to ones that contain all networks with specific network architectures, although for the purpose of the definition, there is no reason to require this to be the case. While comparing a set of networks, we restrict ourselves to inputs for which all the networks satisfy a certain Gap Criterion (though, not necessarily for the same T), so that the notion of a transformation is well-defined on the input set, for all networks under consideration.

Definition 12 (Transformational Complexity). Let Σ_1 and Σ_2 be two sets of acyclic networks, each network being of order m , with $\Sigma_1 \subseteq \Sigma_2$. Define $\mathcal{G}_{12} = \bigcap_{\mathcal{N} \in \Sigma_1 \cup \Sigma_2} \mathcal{G}_{\mathcal{N}}$.

Σ_2 is said to be *more complex than* Σ_1 , if $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1$, $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$.

Note that \mathcal{G}_{12} is always nonempty, because \mathcal{G}_{12} contains all inputs satisfying the Flush criterion. Henceforth, any result that establishes a relationship of the form defined above is called a *complexity result*.

Before we proceed, we introduce some useful notation. Let the set of spike-train ensembles containing exactly m spike-trains that satisfy the T-Flush criterion be \mathcal{F}_m^T . Let $\mathcal{F}_m = \bigcup_{T \in \mathbb{R}^+} \mathcal{F}_m^T$. What we have established in the previous section is that $\mathcal{F}_m \subseteq \mathcal{G}_{\mathcal{N}}$, for every acyclic network \mathcal{N} that has exactly m input vertices.

Next, is the main lemma of this section. We show that if one class of networks is more complex than another, then inputs that satisfy the Flush Criterion are sufficient to prove this. That is, to prove this type of complexity result, one can work exclusively with Flush inputs without losing any generality. This is not obvious because Flush inputs form a subset of the more biologically-realistic Gap inputs. We provide an informal proof sketch below; the full proof is available in the Appendix.

Lemma 5 (Equivalence of Flush and Gap Criteria with respect to Transformational Complexity). *Let Σ_1 and Σ_2 be two sets of acyclic networks, each network being of order m , with $\Sigma_1 \subseteq \Sigma_2$. Then, Σ_2 is more complex than Σ_1 if and only if $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1$, $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$.*

Proof sketch. One direction is easy. By using Flush inputs if we can show a complexity result, that holds for Gap inputs too, because $\mathcal{F}_m \subseteq \mathcal{G}_{12}$. To show the other direction, we assume a complexity result proved using Gap inputs and construct Flush inputs such that the result can be shown using those Flush inputs alone. Now suppose $\mathcal{N}' \in \Sigma_2$ be the network such that no network in Σ_1 effects the same transformation as \mathcal{N}' , when restricted to the set \mathcal{G}_{12} . Now, consider arbitrary $\mathcal{N} \in \Sigma_1$. There must exist a $\chi \in \mathcal{G}_{12}$ such that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m}(\chi) \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}(\chi)$. By definition, this χ satisfies a T_1 -Gap Criterion for \mathcal{N} and a T_2 -Gap Criterion for \mathcal{N}' . Let $T = \max(T_1, T_2)$. The claim is that if χ is cut up into “chunks” of length $2T$, where each “chunk” satisfies a $2T$ -Flush criterion, then \mathcal{N} and \mathcal{N}' will map at least one of those chunks to different output spike trains, since the output in the latter half of the chunk is identical to that in the corresponding segment of χ . This process when “completed” for each $\mathcal{N} \in \Sigma_1$ yields a subset of Flush inputs,

using which the complexity result can be established. \square

Assured by this theoretical guarantee that there is no loss of generality by doing so, we will henceforth only work with inputs satisfying the Flush Criterion, while faced with the task of proving complexity results. This buys us a great deal of mathematical expedience at no cost. From now on, unless qualified otherwise, when we say a *transformation*, we mean a map of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$.

8 Complexity results for the abstract model

In this section, we prove some complexity results for networks whose neurons obey the abstract model of the neuron described in Section 3. Before we proceed, a couple of remarks are in order.

For two sets of acyclic networks, Σ_1 and Σ_2 with $\Sigma_1 \subseteq \Sigma_2$, in order to prove that Σ_2 is more complex than Σ_1 , it is sufficient to show a transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ that no network present in Σ_1 can perform, while demonstrating a network in Σ_2 that can effect it. This involves constructing such a transformation, i.e. prescribing an output spike train for every element in \mathcal{F}_m . Recall that \mathcal{F}_m consists of spike-train ensembles of order m , with the property that for each such ensemble there exists a positive real number T , so that the ensemble satisfies a T -Flush criterion. In practice, however, it usually suffices to prescribe output spike trains for a small subset¹⁷ of elements of \mathcal{F}_m , and prove that no network in Σ_1 can map the input spike trains in that subset to their prescribed outputs. The second step would involve demonstrating a network in Σ_2 that maps this subset of \mathcal{F}_m to the prescribed output, while mapping the rest of \mathcal{F}_m to arbitrary output spike trains. Strictly speaking then, the transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ we are prescribing comprises the mapping from \mathcal{F}_m to output spike trains, as effected by *this* network in Σ_2 . For convenience however, we will refer to the mapping prescribed for some small subset of \mathcal{F}_m and call it the prescribed transformation.

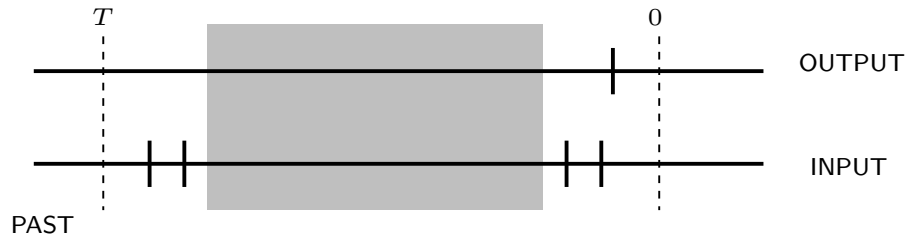
Now, the perceptive reader might wonder if allowing inputs satisfying the T -Flush Criterion, for arbitrarily large T , is contrived, particularly since physiologically it is

¹⁷albeit typically one that contains, for each positive real number T , at least one spike-train ensemble satisfying a T -Flush Criterion.

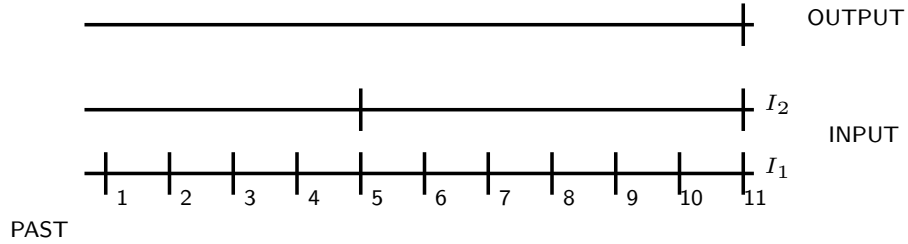
hard for a neuron to sustain high firing rates for arbitrarily large periods of time. Moreover, is this unrealistic assumption being used to push through results, thereby making them only of theoretical interest? The alternative would be to instead put a global bound on T . This however would allow any transformation to be effected even by a single neuron, by making its Υ suitably (and unrealistically) high. To mitigate this, one might propose putting a global bound on Υ , which brings us to the difficult question of having to quantify the relationship between the two bounds, with complexity results depending on the answer to this question. We sidestep these issues with our current formulation, but whenever we present a transformation that cannot be effected by a network architecture the following can be done. Given an upper bound on the values of Υ and ρ of the neurons, it is straightforward to write down a T as a function of these bounds, so that for all $T' \geq T$, inputs in the prescribed transformation satisfying a T' -Flush criterion cannot (provably) be mapped to the prescribed output by any network in Σ_1 which also obeys the same bounds on Υ and ρ .

The second remark concerns our abstract model of the neuron. The model admits a wide variety of membrane potential functions, which are meant to subsume the class of biologically-realistic potential functions. Therefore, when one proves a negative result about the abstract model, the negative result applies to all of the subsumed potential functions as well. However, when we wish to prove a positive result, some caution is required. So, suppose we are trying to come up with a network that *can* effect a transformation, coming up with a network whose neurons merely satisfy the abstract model, leaves open the possibility that the potential functions of the neurons in question are biologically unrealistic. Therefore, when we are proving such positive results we must restrict ourselves to as simple a neuron model as possible. In such cases, we describe the neurons in the construction, so that they can certainly be effected, for example, using the Spike Response Model SRM_0 (Gerstner & Kistler, 2002).

Let us now look at some complexity results. First, we point out that it is straightforward to construct a transformation that cannot be effected by any acyclic network. One of its input spike-train ensembles with the prescribed output is shown in Figure 7(a). For larger T , the shaded region is simply replicated over and over again. Informally, the reason this transformation cannot be effected by any network is that, for any network, beyond a certain value of T , the shaded region tends to act as a flush, erasing



(a) Example of a transformation that no acyclic network can effect. The shaded region is replicated over, to obtain mappings for larger and larger values of T



(b) A transformation that no single neuron can effect, that a network with two neurons can.

Figure 7:

“memory” of the first spike. When the network receives another input spike, it is in the exact same “state” it was when it received the first spike, and therefore produced no output spike.

Next, we prove that the set of networks with atmost two neurons is more complex than the set of single neurons. The proof is by prescribing a transformation which cannot be done by any single neuron. We then construct a network with two neurons that can indeed effect this transformation.

Theorem 1. *For $m \geq 2$, the set of acyclic networks with atmost two neurons which have exactly m input vertices is more complex than the set of single neurons with exactly m afferent synapses.*

Proof. We first prove the result for $m = 2$ and later indicate how it can be trivially extended for larger values of m . Let the two input spike-trains in each input spike-train ensemble, which satisfies a Flush Criterion be I_1 and I_2 . Figure 7(b) illustrates the transformation. Informally, I_1 has regularly-spaced spikes starting at time instant T until 0. I_2 has two spikes, with the first one, loosely speaking, in the “middle” of

$[0, T]$ and the second one at the end, i.e. at time instant 0. An output spike is always prescribed when the second spike in I_2 occurs, and not elsewhere. For larger T , the number of spikes on I_1 increases so as to maintain the same regular spacing; I_2 , in contrast, still has just two spikes, the first one roughly in the middle and the second in the end. For the sake of exposition, we call the distance between consecutive spikes on I_1 , one time unit and we number the spikes of I_1 with the first spike being the oldest one.

Formally, the transformation is prescribed for a subset of \mathcal{F}_m , whose elements are indexed by $i = 1, 2, \dots$. Figure 7(b) illustrates the transformation, for $i = 2$. The i th input spike-train ensemble in this subset satisfies a T -Flush criterion, where $T = 4i + 3$ time units. In the i th spike-train ensemble, I_2 has spikes at time instants at which spike numbers $2i + 1$ and $4i + 3$ occur in I_1 . Finally, the output spike-train corresponding to the the i th input spike-train ensemble has exactly one spike at the time instant at which I_1 has spike number $4i + 3$.

Next, we prove that the transformation prescribed above cannot be effected by any single neuron. For the sake of contradiction, suppose it can, by a neuron with associated Υ and ρ . Let $\max(\Upsilon, \rho)$ be bounded from above by k time units. We show that for $i \geq \lceil \frac{k}{2} \rceil$, the i th input spike-train ensemble cannot be mapped by this neuron to the prescribed output spike train. For $i = \lceil \frac{k}{2} \rceil$, consider the output of the neuron at the time instants corresponding to the $(k + 1)$ th spike number and $(2k + 3)$ rd spike number of I_1 . At each of these time instants, the input received in the past k time units and the output produced by the neuron in the past k time units are the same. Therefore, the neuron's membrane potential must be identical. However, the transformation prescribes no spike in the first time instant and a spike in the second, which is a contradiction. It follows that no single neuron can effect the prescribed transformation.

We now construct a two-neuron network which can carry out the prescribed transformation. The network is shown in Figure 8. I_1 and I_2 arrive instantaneously at N_2 . I_1 arrives instantaneously at N_1 but I_2 arrives at N_1 after a delay of 1 time unit. Spikes output by N_1 take one time unit to arrive at N_2 , which is the output neuron of the network. The functioning of this network for $i = 2$ is described in Figure 8. The generalization for larger i is straightforward. All inputs are excitatory. N_1 is akin to the neuron described in Figure 1, in that while the depolarization due to a spike in I_1 causes po-

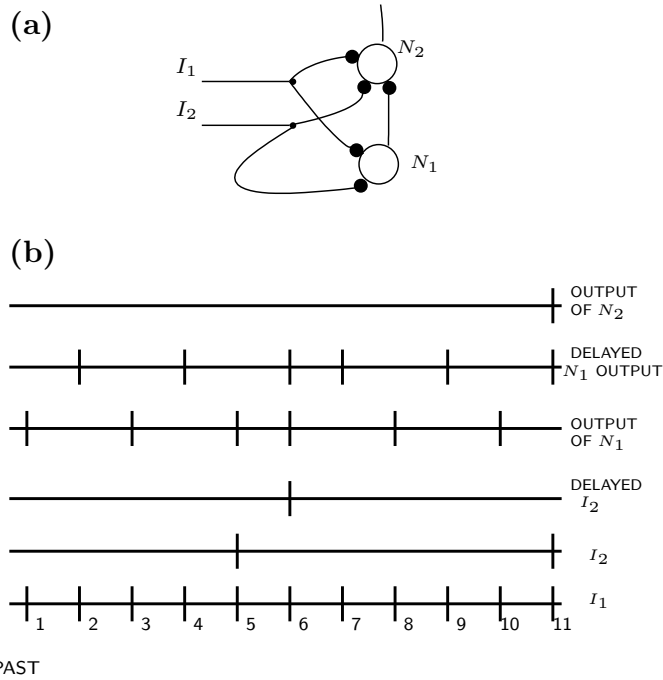


Figure 8: (a) The network that can effect the transformation described in Figure 7(b). (b) Figure describing the operation of this network.

tential to cross threshold, if, additionally, the previous output spike happened one time unit ago, the associated hyperpolarization is sufficient to keep the membrane potential below threshold now. However, if there is a spike from I_2 also at the same time, the depolarization is sufficient to cause an output spike, irrespective of if there was an output spike one time unit ago. The Υ corresponding to N_2 is shorter than 1 time unit. Further, N_2 produces a spike if and only if all three of its afferent synapses receive spikes at the same time. In the figure, I_1 spikes at times 1, 3, 5. It spikes at 6 because it received spikes both from I_1 and I_2 at that time instant. Subsequently, it spikes at 8 and 10. The only time wherein N_2 received spikes at all three synapses at the same time is at 11, which is the prescribed time for the output spike. The generalization for larger i is straightforward.

For larger m , one can just have no input on the extra input spike trains and the same proof generalizes trivially. \square

The above proof also points to a large class of transformations that cannot be done by a single neuron. Informally, these are transformations for which there is no fixed

bound, so that one can always determine whether there is an output spike or not, just by looking a window of past input and past output, where the window has length being atmost this bound.

The previous result might suggest that the more the number of neurons the larger the variety of transformations possible. The next complexity result demonstrates, on the contrary, that the structure of the network architecture is crucial. That is, we can construct network architectures with arbitrarily large number of neurons which cannot perform transformations that a simple two-neuron network can.

First, we define the abstract graph-theoretic property that characterizes this class of architectures.

Definition 13 (Path-disjoint Architecture). An acyclic network architecture $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$, where $|I| = m$ is called *path-disjoint* if for every set of m paths, where the i th path starts at input vertex $\#(i)$ and ends at the output vertex, the intersection of the m paths is exactly the output vertex.

Theorem 2. For $m \geq 3$, let Σ_1 be the set of all acyclic networks whose architecture is path-disjoint. Let Σ_2 be the union of Σ_1 with the set of all networks with atmost 2 neurons which have m input vertices. Then Σ_2 is more complex than Σ_1 .

Proof. We prove the theorem for $m = 3$; the generalization for larger m is straightforward. The following transformation is prescribed for $m = 3$. Let the three input spike-trains in each input spike train ensemble, which satisfies a Flush Criterion be I_1 , I_2 and I_3 . As before, we will use regularly spaced spikes; we call the distance between two such consecutive spikes one time unit and number these spike time instants with the oldest being numbered 1; we call this numbering the spike index. Again, the transformation is prescribed for a subset of \mathcal{F}_m , whose elements are indexed by $i = 1, 2, \dots$. Figure 9 illustrates the transformation for $i = 2$. The i th input spike-train ensemble in the subset satisfies a T -Flush Criterion for $T = 4im$ time units. The first $2i$ time units have spikes on I_2 spaced one time unit apart, the next $2i$ on I_3 and so forth. In addition, at spike index $2im$, I_m has a single spike. The input spike pattern from the beginning is repeated once again for the latter $2im$ time units. The output spike-train has exactly one spike at spike index $4im$.

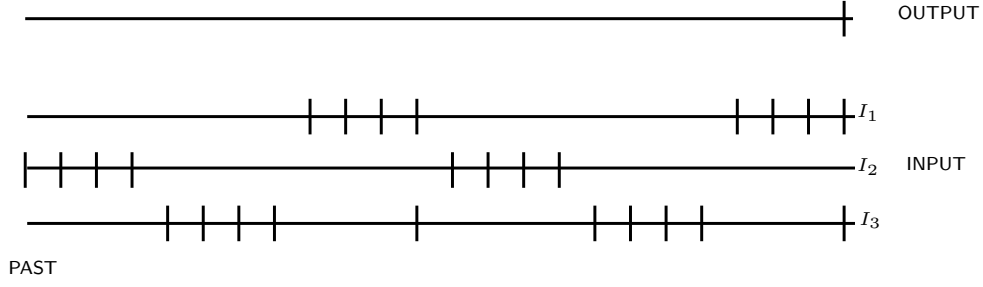


Figure 9: A transformation that no network with a path-disjoint architecture can effect.

Next we prove that the transformation prescribed above cannot be effected by any network in Σ_1 . For the sake of contradiction, assume that $\mathcal{N} \in \Sigma_1$ can effect the transformation. Let Υ and ρ be upper bounds on the same parameters over all of the neurons in \mathcal{N} and let d be the depth of \mathcal{N} . By construction of Σ_1 , every neuron in \mathcal{N} that is afferent on the output neuron receives input from atmost $m - 1$ of the input spike-trains; for, otherwise there would exist a set of m paths, one from each input vertex to the output neuron, whose intersection would contain the neuron in question. The claim, now, is that for $i > \frac{\Upsilon d}{2} + \rho$, the output neuron of \mathcal{N} has the same membrane potential at spike index $2im$ and $4im$, and therefore either has to spike at both those instants or not. Intuitively, this is so because each neuron afferent on the output neuron receives a “flush” at some point after $2im$, so that the output produced by it Υ seconds before time index $2im$ and Υ seconds before time index $4im$ are the same. This is straightforward to verify.

We now construct a two-neuron network that can effect this transformation. The construction is similar to the one used in Theorem 1. For $m = 3$, the network is shown in Figure 10. I_1, I_2 and I_3 arrive instantaneously at N_1 and N_2 . Spikes output by N_1 take two time units to arrive at N_2 , which is the output neuron of the network. The functioning of this network for $i = 2$ is described in Figure 10. The generalization for larger i is straightforward. All inputs are excitatory. N_1 is akin to the the N_1 used in the network in Theorem 1 except that that periodic input may arrive from any one of I_1, I_2 or I_3 . As before, if two input spikes arrive at the same time, as in spike index $2im$, the depolarization is sufficient to cause an output spike in N_1 , irrespective of if there was an output spike one time unit ago. Again, the Υ corresponding to N_2 is shorter than 1 time unit and N_2 produces a spike if and only if three of its afferent synapses receive

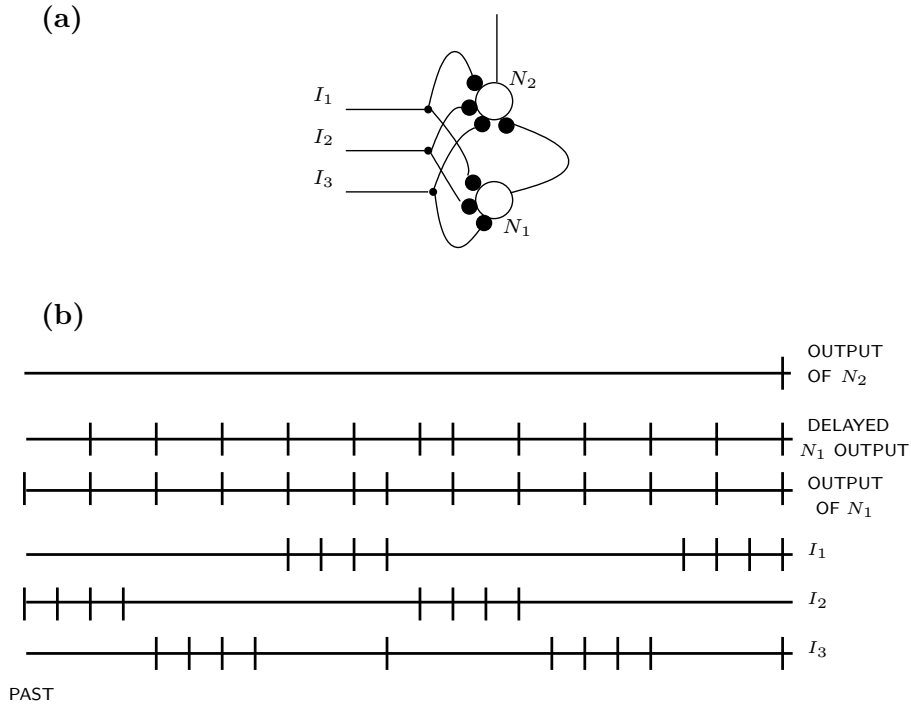


Figure 10: (a) Network that can effect the transformation described in Figure 9. (b) Figure describing the operation of this network.

spikes at the same time. As before, the idea is that at time $2im$, N_2 , receives two spikes, but not a spike from N_1 , since it is “out of sync”. However, at time $4im$, additionally, there is a spike from N_1 arriving at N_2 , which causes N_2 to spike. \square

In the next section, we investigate complexity results with increase in depth of the acyclic network. What we prove is that the present abstract model of the neuron does not cause such complexity results to exist beyond depth two, in general, and therefore more axioms need to be added to the abstract model of the neuron, in order for such results to be manifested.

9 Complexity results with increasing network depth:

A barrier at depth two

While, in the previous section, we have proved a number of complexity results, the perceptive reader will observe that we did not explicitly address the question of complexity

results pertaining to depth, in general. In other words, the question is, does increase in depth of the network, in general, buy us a larger variety of transformations effectable by it? More formally, is the set of *all* acyclic networks more complex than the set of acyclic networks of depth k , for $k \geq 2$? Much to our surprise¹⁸, it turns out that the answer is *No*, even for $k = 2$, and we can prove it. That is, given an arbitrary acyclic network (consisting of neurons obeying the abstract model described in Section 3), there exists a network of depth two (equipped with neurons obeying the same abstract model), so that the latter network induces *exactly* the same transformation as the former. The implication of this result is that one needs to add more axioms to the abstract model of Section 3, in order to break this barrier to the manifestation of complexity results.¹⁹

The difficulty in proving that every acyclic network, having arbitrary depth, has an equivalent network of depth two, appears to be in devising a way of “collapsing” the depth of the former network, while keeping the effected transformation the same. Our proof actually does not demonstrate this head-on, but instead proves it to be the case indirectly. The broad attack consists of starting off with a certain subset of the set of all possible transformations and showing that the transformation effected by every acyclic network certainly lies within this subset. Thereafter, we prove, by providing a construction, that every transformation in this subset can in fact be effected by an acyclic network of depth two. Together, this implies that, for every transformation that can be effected by an acyclic network, there exists an acyclic network of depth two that can effect exactly that transformation.

Technical structure of the proof

The main theorem of this section is the following.

Theorem. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then it can be effected by an acyclic network of depth two.

¹⁸In hindsight, of course, something like this should have been expected at some point, given that the abstract model assumed so little about single neurons.

¹⁹Note that the complexity results proved thus far still hold, if the “new” abstract model also has all the axioms of the one described in Section 3.

This theorem follows from the following two lemmas which are proved in the two subsections that follow:

Lemma. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then $\mathcal{T}(\cdot)$ is causal, time-invariant and resettable.

Lemma. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.

9.1 Causal, Time-Invariant and Resettable Transformations

In this section, we first define notions of causal, time-invariant and resettable transformations²⁰. Transformations that are causal, time-invariant and resettable form a strict subset of the set of all transformations. Notice that these notions exist independent of the existence of neurons or their networks. We then show that transformations effected by acyclic networks always lie within this subset. This is the relatively easy part of the proof. The next subsection proves the harder part, namely that every transformation in this subset can indeed be effected by an acyclic network of depth equal to two.

As in systems theory, informally, a *causal transformation* is one whose current output depends only on its current input and past input (and not future input). Abstractly, it is convenient to define a causal transformation as one that given two different inputs that are identical up to a certain point in time, also have their outputs, according to the transformation, be identical up to (atleast) the same point.

Definition 14 (Causal Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be *causal* if, for every $\chi_1, \chi_2 \in \mathcal{F}_m$, with $\Xi_{[t,\infty)}\chi_1 = \Xi_{[t,\infty)}\chi_2$, for some $t \in \mathbb{R}$, we have $\Xi_{[t,\infty)}\mathcal{T}(\chi_1) = \Xi_{[t,\infty)}\mathcal{T}(\chi_2)$.

Again, as in systems theory, a *time-invariant transformation* is one which always transforms the time-shifted version of an input, to a time-shifted version of its corresponding output. To keep the definition sound, we also need to ensure that the time-shifted input, in fact, also satisfies the Flush criterion.

²⁰Recall that when we say transformation without further qualification, we mean one, of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$.

Definition 15 (Time-Invariant Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be *time-invariant* if, for every $\chi \in \mathcal{F}_m$ and every $t \in \mathbb{R}$ with $\sigma_t(\chi) \in \mathcal{F}_m$, we have $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$.

A *resettable transformation* is one for which there exists a positive real number W , so that an input gap of the form $(t, t + W]$ “resets” it, i.e. output beyond t is independent of input received before it. Again, abstractly, it becomes convenient to say that the output in this case is identical to that produced by an input which has no spikes before t , but is identical to the present input thereafter.

Definition 16 (W -Resettable Transformation). For $W \in \mathbb{R}^+$, a transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be *W -resettable* if, for every $\chi \in \mathcal{F}_m$ which has a gap in the interval $(t, t + W]$, for some $t \in \mathbb{R}$, we have $\Xi_{(-\infty, t]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]} \chi)$.

Definition 17 (Resettable Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be *resettable* if, there exists a $W \in \mathbb{R}^+$, so that it is W -resettable.

Next, we prove that every transformation that can be effected by an acyclic network is causal, time-invariant and resettable. This implies that every transformation that is not causal, time-invariant or resettable cannot be effected by any acyclic network.

Lemma 6. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then $\mathcal{T}(\cdot)$ is causal, time-invariant and resettable.*

Proof sketch. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by a single neuron it is relatively straightforward to verify that $\mathcal{T}(\cdot)$ is causal, time-invariant and resettable. That it is causal and time-invariant follows from the fact that the $P(\cdot)$ function of the neuron only “looks” at the recent past and not the future to determine membrane potential. That $\mathcal{T}(\cdot)$ is resettable follows from Axiom (3) of the neuron. For an acyclic network, the proof proceeds by mathematical induction on the depth of the network. A full proof is provided in the Appendix. □

9.2 Construction of a depth two acyclic network for every causal, time-invariant and resettable transformation

In this subsection, we prove the following lemma.

Lemma. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.

Before we dive into the proofs, we offer some intuition.

Suppose we had a transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ which is causal, time-invariant and resettable. For the moment, pretend it satisfies the following property: There exist constant-sized input and output “windows” so that, for every input spike-train ensemble satisfying a flush criterion, just given knowledge of spikes in those windows of past input and output, one can unambiguously determine, at any point in time, if the transformation prescribes an output spike or not. Intuitively, it seems reasonable that such a transformation can be effected by a single neuron²¹ by setting the Υ and ρ of the neuron to the sizes of the input and output windows mentioned above.

Of course, one easily sees that not every transformation that is causal, time-invariant and resettable satisfies the aforementioned property. That is, there could exist two different input instances, which have past inputs and outputs be identical in the aforementioned windows at some points in time; yet in one instance, the transformation prescribes an output spike, whereas it prescribes none in the other. Indeed, the two input instances do differ at some point in the past, for otherwise the transformation would not be causal. Therefore, in such a situation, it is natural to ask if a single “intermediate” neuron can “break the tie”. That is, if two input instances differ at some point in the past, the output of the intermediate neuron since then, in any interval of time of length U , must be different, where U is a fixed constant. This is so that a neuron receiving input from the intermediate neuron can disambiguate the two, were an output spike demanded for one instance but not the other. Unfortunately, this exact property cannot be achieved by any single neuron because the transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is resettable, and so is every transformation induced by the intermediate neuron. In other words, the problem is that, suppose two input instances differ at a certain point in time; however, since then, both have had an arbitrarily large input gap. The input gap serves to “erase memory” in any network that received it and therefore it cannot disambiguate the two inputs beyond this gap. Now, fortunately, it does not have to, since this gap also causes a “reset” in the transformation (which is resettable). That is, if such an arbitrarily large

²¹Presuming that no axioms of the neuron are violated.

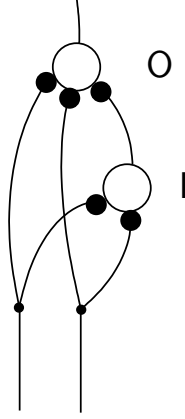


Figure 11: Network architecture for $m = 2$.

gap were present in the input, the transformation would not afterward demand an output spike in one case and no output spike in another. This is because it is W -resettable and therefore cannot make such demands, for input gaps²² larger than W . Thus, we can make do with a slightly weaker condition; that the intermediate neuron is only guaranteed to break the tie, when it is required to do so. That is, suppose there are two input instances, whose outputs according to $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ are different at certain points in time. Then, the corresponding inputs must be different too at some point in the past with no reset gaps in the intervening time and therefore the intermediate neuron should break the tie. Additionally, for technical reasons that will become clear later, we say that the outputs of the intermediate neuron in the preceding U seconds are guaranteed to be different, only if the inputs themselves in the past U seconds are not different.

The network we have in mind is illustrated in Figure 11, for $m = 2$. In the following proposition, we prove that if the intermediate neuron satisfies the condition alluded to above, then there exists an output neuron, so that the network effects the transformation in question. By way of notation, recall that $\Xi_t(\cdot)$ is shorthand for $\Xi_{[t,t]}(\cdot)$

Proposition 2. *Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and resettable. Let \mathfrak{l} be a neuron with $\mathcal{T}_{\mathfrak{l}} : \mathcal{F}_m \rightarrow \mathcal{S}$, so that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_{\mathfrak{l}}(\chi)$ is consistent with χ with respect to \mathfrak{l} . Further, suppose there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_{\mathfrak{l}}(\chi_1) \sqcup \chi_1)) \neq$*

²²which we call a “reset gap” from now on, for the sake of exposition.

$\Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$.

Then there exists a neuron O , so that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_I(\chi) \sqcup \chi$ with respect to O .

Proof sketch. The straightforward way for the neuron O to effect $\mathcal{T}(\cdot)$ is to determine the points of time wherein an output spike is prescribed and set its membrane potential function to hit threshold at those instances. Since the neuron I essentially “disambiguates” the input, this assignment can be done without conflict. However, we also need to show that doing this does not violate any of three axioms of our abstract model, for the neuron O . Axiom (1) follows easily from the fact that the co-domain of $\mathcal{T}(\cdot)$ is \mathcal{S} . Axiom (3) is also relatively straightforward to show and uses the fact that $\mathcal{T}(\cdot)$ is causal, time-invariant and resettable. Axiom (2), on the other hand, presents some subtleties. Now, in addition to setting membrane potential to threshold at the aforementioned points, in order to satisfy Axiom (2), we would also need to set it to hit threshold, when the input window has the same pattern and the output window is empty instead. However, with this assignment, we need to then show that no spurious spikes are generated. This takes a little work and again uses the “disambiguation” property of the intermediate neuron I . The full proof is available in the Appendix.

□

The next proposition shows that one can always construct an intermediate neuron with the said “disambiguation” property.

Proposition 3. *Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and resettable. Then there exists a neuron I and $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_I(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_I : \mathcal{F}_m \rightarrow \mathcal{S}$ is such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_I(\chi)$ is consistent with χ with respect to I .*

Proof idea. The basic idea is to “encode”, in the time difference of two successive output spikes, the positions of all the input spikes that occurred since the last output gap of the form $(t, t + W]$, where $\mathcal{T}(\cdot)$ is W -resettable. Such pairs of output spikes occur once every p seconds, with the time difference within each pair being a function of the time difference within the previous pair and the input spikes encountered since. Intuitively,

it is convenient to think of this encoding as one from which we can “reconstruct” the entire past input spike-train since the last reset. We first describe the encoding function for the case of a single input spike-train after which we remark on how it can be generalized.

So, suppose the time difference of the successive spikes lies in the interval $[0, 1)$. Define the encoding function as $\varepsilon_0 : [0, 1) \times \bar{\mathcal{S}}_{(0,p]} \rightarrow [0, 1)$. p is chosen to be such that there are at most 8 spikes in any interval of the form $(t, t + p]$. We now describe how $\varepsilon_0(e, \vec{x})$ is computed, given $e \in [0, 1)$ and $\vec{x} = \langle x^1, x^2, \dots, x^k \rangle$, such that each spike time lies in the interval $(0, p]$. Let e have a decimal expansion²³, so that $e = 0.c_1s_1c_2s_2c_3s_3 \dots$. Accordingly, let $c = 0.c_1c_2c_3 \dots$ and $s = 0.s_1s_2s_3 \dots$. c is a real number that encodes the number of spikes in each interval of length p encountered, since the last reset. Since each interval of length p has between 0 and 8 spikes, the digit 9 is used as a “termination symbol”. So, for example, suppose there have been 4 intervals of length p , since the last reset with 5, 0, 8 and 2 spikes apiece respectively, then $c = 0.8059$ and $c' = 0.28059$, where c' is the “updated” value of c . Likewise, s is a real number that stores the positions of all input spikes encountered since the last reset. Let each spike time be of the form $x^i = 0.x_1^i x_2^i x_3^i \dots \times 10^q$, for appropriate q . Then the updated value of s is $s' = 0.x_1^1 x_1^2 \dots x_1^k s_1 x_2^1 x_2^2 \dots x_2^k s_2 \dots$. Suppose the c' and s' obtained above were of the form $c' = 0.c'_1 c'_2 c'_3 \dots$ and $s' = 0.s'_1 s'_2 s'_3 \dots$, then $\varepsilon_0(e, \vec{x}) = 0.c'_1 s'_1 c'_2 s'_2 \dots$. Observe that the decimal expansion constructed by $\varepsilon_0(e, \vec{x})$ cannot have infinitely many successive 9s, for c' has only a finite number of non-zero digits. Suppose the input were a spike-train ensemble of order m , then for each spike-train an encoding would be computed as above and in the final step, the m real numbers obtained would be interleaved together, so as to produce the encoding.

Given knowledge of the encoding function, we now describe how l exactly works. Figure 12 provides an illustration. Suppose $\chi \in \mathcal{F}_m$ is an input spike-train. Let its oldest spike be T seconds ago. Then l produces a spike at time $T - p^{24}$ and at every $T - kp$, for $k \in \mathbb{Z}^+$, unless in the previous p seconds to when it is to spike, there is

²³Whenever we say decimal expansion, we forbid decimal expansions with an infinite number of successive 9s. With this restriction, each real number has a unique decimal expansion.

²⁴i.e. p seconds after time instant T .

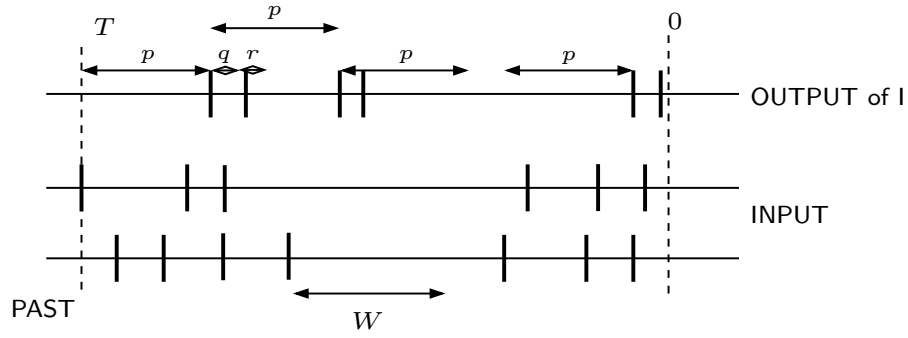


Figure 12: Example illustrating the operation of the intermediate neuron I.

a gap²⁵ of the form $(t, t + W]$. For the sake of exposition, let's call these the “clock” spikes. Now, suppose there is a gap of the form $(t, t + W]$ in the input and there is an input spike at time t , then the neuron spikes at time $t - p$ and every p seconds thereafter subject to the same “rules” as above. These clock spikes are followed by “encoding” spikes, which occur atleast q seconds after the clock spike, but atmost $q + r$ seconds after, where q is greater than the absolute refractory period α . As expected, the position of the current encoding spike is a function of the time difference between the previous encoding and clock spike²⁶ and the positions of the input spikes in the past p seconds. The output of the encoding function is appropriately scaled to “fit” in this interval of length r ; the details are available in the proof.

The claim then is that if two input spike-trains are different at some point with no intervening “reset” gaps, then the output of I in the past U seconds, where $U = p + q + r$ will be different. Intuitively, this is because the difference between the latest encoding and clock spike in each case would be different, as they encode different “histories” of input spikes.

Finally, we remark that the above is just an informal description that glosses over several technical details contained in the full proof, which is available in the Appendix.

□

The preceding two propositions thus imply the following lemma.

²⁵We set $W > p$ to force a spike at $T - p$.

²⁶unless the present clock spike is the first after a reset gap in the input.

Lemma 7. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.*

Lemma 6 and 7 imply the following theorem. From this it follows that, with the abstract model currently under consideration, the set of all acyclic networks is *not* more complex than the set of acyclic networks of depth two.

Theorem 3. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then it can be effected by an acyclic network of depth two.*

Corollary 2. *The set of all acyclic networks is not more complex than the set of acyclic networks of depth equal to two.*

Incidentally, Lemma 6 and 7 also lead to a full characterization of the class of transformations effected by all acyclic networks equipped with neurons obeying the abstract model of Section 3. This is formalized in the next theorem.

Theorem 4. *A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network if and only if it is causal, time-invariant and resettable.*

Directions for further constraining the abstract model

The meticulous reader would have observed that the proof above exploited a number of phenomena that the abstract model allowed for, but are not particularly biologically well-founded. These are good candidates for new axioms that we hope will overcome the “depth barrier” that this section describes. One is that spike-times in the abstract model are real numbers, i.e. numbers with infinite precision. When noise is taken into account, this assumption is no longer true. The other easy observation is that we did not assume much about the membrane potential function here, which in nature, changes smoothly with time. And, finally, an assumption consistent with Dale’s principle, that each neuron has either an excitatory effect at all its postsynaptic neurons or an inhibitory effect could also help in this direction.

10 Discussion

As one spans the gamut from neuron models that exactly describe the operation of neurons to the abstract model of the type espoused here, the type of computations that

networks equipped with such neurons can undertake, increases in variety. While the former types of neurons have been the subject of several investigations, the latter type seems to have received much less attention. These latter type of models are attractive due to their simplicity and their ability to subsume large classes of neurons. Also, this type of first-principles approach forces us to examine which network-level computational capabilities follow from which properties of neurons. The flip side to them is that, being less constrained, they may not embody the entire variety of phenomena that networks use for computation.

In this work, we were motivated by the goal of starting with as emaciated an abstract model as we could and see how much structure we could extract from it. This was because, the more one assumes about neurons, the smaller the class of biological neurons one tends to cover. With this model, we showed that notions of acyclic networks as spike-train to spike-train transformations, while not well-defined in general, could be rigorously defined for biologically-relevant spiking regimes. We also showed the existence of some classes of architectures, in which these axioms constrain computation enough so they cannot effect certain types of transformations that other networks can. However, we also showed that the present axioms do not adequately constrain networks with respect to depth. That is, these axioms make acyclic networks of depth two as powerful, in terms of computational ability, as the class of all acyclic networks. Thus the theory itself is able to tell us when it cannot reveal more about the computational capabilities of certain classes of networks, unless we first say something more about their constituent neurons.

The present work raises a number of questions. What new axioms will manifest complexity results with increasing depth? Can additional insight be gained by studying acyclic networks with more than one output neuron²⁷? Can similar questions be asked of recurrent networks? How does the presence of noise affect these results? How does “learning” fit into the picture? Also, this work does not address specific computational tasks that networks are trying to perform and link them to the ability of networks to accomplish them. That is, loosely speaking, suppose we can experimentally characterize

²⁷While such a network can be viewed as a set of several “independent” networks with one output neuron apiece, that view does not exploit how the transformation effected by one such “independent” network constrains the transformation effected by another.

the class of transformations that specific tasks, say face recognition, require. Then, conceivably, the present type of theory might be useful in eliminating classes of networks that cannot do the task in question. This would go a long way in relating structure of networks to their function.

In recent times, there have been a number of efforts to map the connectome of various organisms including the just-launched Human Connectome Project. These efforts are focused on ascertaining the wiring diagram of neurons in the brain; regrettably, also determining the dynamical properties of all the neurons involved is out of reach of current experimental technology. It seems that a theory such as the one initiated here, that tries to assume as little as possible about neurons but tries to elucidate how network structure constrains computation, might prove invaluable in using data from such projects to rule out what computations these networks cannot perform. Larry Abbott has asked (Abbott, 2008) what we might be able to learn from connectome data alone. We believe this paper gives a sense of what one answer to that question might be like. An important open question here is about what set of axioms one needs to assume for a single neuron, in order for the theory to yield non-trivial results about large classes of network architectures, particularly those that one might encounter from these projects.

In short, challenges abound and promise us interesting times ahead.

Acknowledgments

This work was supported, in part, by a National Science Foundation grant (NSF IIS-0902230) to A.B.

Appendix

Proofs from Section 5

Proof of Corollary 1. (1) is immediate from the Gap Lemma, when we set $\chi = \chi^*$.

For (2), the proof is by strong induction on the number of spikes since t . Let \vec{x}_0 be an arbitrary spike-train that is consistent with χ^* , with respect to N . Notice that from (1) we have that every \vec{x}_0 is identical to \vec{x}_0^* in $[t, t + \rho]$. The base case is to show that both \vec{x}_0^* and \vec{x}_0 have their first spike after t at the same time. Assume, without loss

of generality, that the first spike of \vec{x}_0 at $t_1 < t$, is no later than the first spike of \vec{x}_0^* . We have $P(\Xi_{[0,\Upsilon]}(\sigma_{t_1}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0^*))) = P(\Xi_{[0,\Upsilon]}(\sigma_{t_1}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0)))$ since $\Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0^*)) = \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0)) = \vec{\phi}$. Therefore \vec{x}_0^* also has its first spike at t_1 . Let the induction hypothesis be that both \vec{x}_0^* and \vec{x}_0 have their first k spikes since t at the same times. We show that this implies that the $(k + 1)^{th}$ spike in each spike-train is also at the same time instant. Assume, without loss of generality, that the $(k + 1)^{th}$ spike since t of \vec{x}_0 at t_{k+1} , is no later than the $(k + 1)^{th}$ spike since t of \vec{x}_0^* . Now, $\Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0^*))$ is identical to $\Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0))$ from the induction hypothesis since $(t + \rho) - t_{k+1} \geq \rho$. Thus, $P(\Xi_{[0,\Upsilon]}(\sigma_{t_{k+1}}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0^*))) = P(\Xi_{[0,\Upsilon]}(\sigma_{t_{k+1}}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0)))$ and therefore \vec{x}_0^* also has its $(k + 1)^{th}$ spike at t_{k+1} . This completes the proof of (2).

(3) follows from the Gap Lemma and (2). □

Proof of Lemma 2. We prove that the output of the network is unique by strong induction on depth. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i . Each neuron $N \in N_1$ receives all inputs from spike-trains in χ . Since, N satisfies a Gap criterion with those input spike-trains, its output is unique. The induction hypothesis then is that for all $i \leq k < d$, each neuron $N \in N_i$ produces a unique output when \mathcal{N} is driven by χ . Consider arbitrary $N' \in N_{k+1}$. It is clear that all inputs to N' are from spike-trains from χ or neurons in $\bigcup_{i=1}^k N_i$, for otherwise the depth of N' would be greater than $k + 1$. Since, all its inputs are unique by the induction hypothesis and they satisfy a Gap criterion for N' , its output is also unique.

Next, we show that the membrane potential of the output neuron at any time instant depends on atmost the past T seconds of input in χ . Since the output neuron satisfies a $(\frac{T}{d})$ -Gap Criterion, its membrane potential at any point depends on atmost the past $(\frac{T}{d})$ seconds of the inputs it receives (some of which are output spike-trains of other neurons). Consider one such ‘‘penultimate layer’’ neuron. Again, its output membrane potential at any time instant, likewise, depends on its inputs in the past $(\frac{T}{d})$ seconds. Therefore, the current potential of the output neuron is dependent on the input received by the penultimate layer neuron in atmost the past $(\frac{2T}{d})$ seconds. Similar arguments can be put forth until, for each path, one reaches a neuron, all of whose inputs do not come from other neurons. Since the longest such path is of length d , it is straightforward to verify that the membrane potential of the output neuron depends on atmost T seconds

of past input.

□

Proofs from Section 7

Proof of Lemma 5. We prove the easy direction first. If $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$, then it follows that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$ because $\mathcal{F}_m \subseteq \mathcal{G}_{\mathcal{N}}$.

For the other direction, let $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$. We construct $\mathcal{F}' \subseteq \mathcal{F}_m$, so that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}'} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}'}$. This immediately implies $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$. Consider arbitrary $\mathcal{N} \in \Sigma_1$. From the hypothesis, we have $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$. Therefore $\exists \chi \in \mathcal{G}_{12}$ such that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}}(\chi) \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}(\chi)$. Additionally, there exist $T_1, T_2 \in \mathbb{R}^+$, so that χ satisfies a T_1 -Gap Criterion for \mathcal{N} and a T_2 -Gap Criterion for \mathcal{N}' . Let $T = \max(T_1, T_2)$. Let $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}}(\chi) = \vec{x}'_0$ and $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}(\chi) = \vec{x}_0$. Let $\tilde{\mathcal{F}} = \bigcup_{t \in \mathbb{R}} \Xi_{[0, 2T]}(\sigma_t(\chi))$. Note that each element of $\tilde{\mathcal{F}}$ satisfies a $2T$ -Flush Criterion. The claim, then, is that $\mathcal{T}_{\mathcal{N}'}|_{\tilde{\mathcal{F}}} \neq \mathcal{T}_{\mathcal{N}}|_{\tilde{\mathcal{F}}}$. We have $\Xi_{[0, T]}(\mathcal{T}_{\mathcal{N}'}(\Xi_{[0, 2T]}(\sigma_t(\chi)))) = \Xi_{[0, T]}(\sigma_t(\vec{x}'_0))$ and $\Xi_{[0, T]}(\mathcal{T}_{\mathcal{N}}(\Xi_{[0, 2T]}(\sigma_t(\chi)))) = \Xi_{[0, T]}(\sigma_t(\vec{x}_0))$. This follows from the fact that χ satisfies the T -Gap Criterion with both \mathcal{N} and \mathcal{N}' and therefore when \mathcal{N} and \mathcal{N}' are driven by any segment of χ of length $2T$, the output produced in the last T seconds of that interval agrees with \vec{x}_0 and \vec{x}'_0 respectively. Therefore, if $\vec{x}_0 \neq \vec{x}'_0$, it is clear that there exists a t , so that $\mathcal{T}_{\mathcal{N}'}(\Xi_{[0, 2T]}(\sigma_t(\chi))) \neq \mathcal{T}_{\mathcal{N}}(\Xi_{[0, 2T]}(\sigma_t(\chi)))$. \mathcal{F}' is obtained by taking the union of such $\tilde{\mathcal{F}}$ for every $\mathcal{N} \in \Sigma_1$. The result follows.

□

Proofs from Section 9

Proof of Lemma 6. Let $\mathcal{N} \langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ be a network that effects $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$.

$\mathcal{T}(\cdot)$ is causal. Consider arbitrary $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_{[t, \infty)}\chi_1 = \Xi_{[t, \infty)}\chi_2$, for some $t \in \mathbb{R}$. We wish to show that $\Xi_{[t, \infty)}\mathcal{T}(\chi_1) = \Xi_{[t, \infty)}\mathcal{T}(\chi_2)$. Let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $N \in N_1$ receives all its inputs from spike-trains in χ . When the network receives χ_1 and χ_2 as input, suppose N receives χ'_1 and χ'_2 respectively as input. Formally, $\chi'_1 = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi_1)$ and $\chi'_2 = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi_2)$, where $v = L^{-1}(N)$.

Also, clearly $\Xi_{[t,\infty)}\chi'_1 = \Xi_{[t,\infty)}\chi'_2$. Let \vec{x}'_1 and \vec{x}'_2 be the output produced by N on receiving χ'_1 and χ'_2 respectively. Since $\chi'_1, \chi'_2 \in \mathcal{F}_m$, there exists a $T \in \mathbb{R}^+$, so that $\Xi_{[T,\infty)}\chi'_1 = \Xi_{[T,\infty)}\chi'_2 = \vec{\phi}^{m'}$, where m' is the number of inputs to N. Therefore, by Axiom (3) of the neuron, we have $\Xi_{[T,\infty)}\vec{x}'_1 = \Xi_{[T,\infty)}\vec{x}'_2 = \vec{\phi}$. Now, for all $t' \in \mathbb{R}$, $t' \in \vec{x}'_j$ if and only if $P_N(\Xi_{[0,\Upsilon_N]}(\sigma_{t'}(\chi'_j)), \Xi_{[0,\rho_N]}(\sigma_{t'}(\vec{x}'_j))) = \tau_N$, for $j = 1, 2$. It is immediate that for $t' \leq t$, we have $\Xi_{[0,\Upsilon_N]}(\sigma_{t'}(\chi'_1)) = \Xi_{[0,\Upsilon_N]}(\sigma_{t'}(\chi'_2))$. Now, by an induction argument on the spike number since T , it is straightforward to show that for all $t' \leq t$, $\Xi_{[0,\rho_N]}(\sigma_{t'}(\vec{x}'_1)) = \Xi_{[0,\rho_N]}(\sigma_{t'}(\vec{x}'_2))$. Thus, we have $\Xi_{[t,\infty)}\vec{x}'_1 = \Xi_{[t,\infty)}\vec{x}'_2$. Similarly, using a straightforward induction argument on depth, one can show that for every neuron in the network, its output until time instant t is identical in either case. We therefore have $\Xi_{[t,\infty)}\mathcal{T}(\chi_1) = \Xi_{[t,\infty)}\mathcal{T}(\chi_2)$.

$\mathcal{T}(\cdot)$ is time-invariant. Consider arbitrary $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ with $\sigma_t(\chi) \in \mathcal{F}_m$. We wish to show that $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $N \in N_1$ receives all its inputs from spike-trains in χ . When the network receives χ and $\sigma_t(\chi)$ as input, suppose N receives χ' and $\sigma_t(\chi')$ respectively as input. Formally, $\chi' = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi)$. Let \vec{x}'_1 and \vec{x}'_2 be the output produced by N on receiving χ' and $\sigma_t(\chi')$ as input respectively. We wish to show that $\vec{x}'_2 = \sigma_t(\vec{x}'_1)$. Since $\chi' \in \mathcal{F}_m$, there exists a $T \in \mathbb{R}^+$, so that $\Xi_{[T,\infty)}\chi' = \Xi_{[T-t,\infty)}\sigma_t(\chi') = \vec{\phi}^{m'}$, where m' is the number of inputs to N. Therefore, by Axiom (3) of the neuron, we have $\Xi_{[T,\infty)}\vec{x}'_1 = \Xi_{[T-t,\infty)}\vec{x}'_2 = \vec{\phi}$. Now, for all $t' \in \mathbb{R}$, $t' \in \vec{x}'_1$ if and only if $P_N(\Xi_{[0,\Upsilon_N]}(\sigma_{t'}(\chi')), \Xi_{[0,\rho_N]}(\sigma_{t'}(\vec{x}'_1))) = \tau_N$. It is therefore straightforward to make an induction argument on the spike number, starting from the oldest spike in \vec{x}'_1 to show that \vec{x}'_1 has a spike at some t' iff \vec{x}'_2 has a spike at $t' - t$ and therefore we have $\vec{x}'_2 = \sigma_t(\vec{x}'_1)$. Similarly, using a straightforward induction argument on depth, one can show that for every neuron in the network, its output in the second case is a time-shifted version of the one in the first case. We therefore have $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$.

$\mathcal{T}(\cdot)$ is resettable. Let Υ and ρ be upper bounds on those parameters over all the neurons in \mathcal{N} . If $\Upsilon < \rho$, then set $\Upsilon = \rho$. The claim is that for $W = d(\Upsilon + \rho) + \rho + \delta$, for arbitrary $\delta > 0$, $\mathcal{T}(\cdot)$ is W -resettable. Consider arbitrary $\chi \in \mathcal{F}_m$ so that χ has a gap in the interval $(t, t + \delta + d(\Upsilon + \rho) + \rho]$, for some $t \in \mathbb{R}^+$. Thus χ has a gap in the interval $[t + \delta, t + \delta + d(\Upsilon + \rho) + \rho]$. We will show that $\Xi_{(-\infty, t + \delta]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t + \delta]} \chi)$

which implies the required result $\Xi_{(-\infty, t]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]} \chi)$. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $N \in N_1$ receives all its inputs from spike-trains in χ . Therefore by Axiom (3) of the neuron, it is straightforward to see that the output of N has a gap in the interval $(t, t + \delta + (d-1)(\Upsilon + \rho) + 2\rho]$. By similar arguments, we have that output of each neuron $N \in N_i$, for $1 \leq i \leq d$ has a gap in the interval $(t, t + \delta + (d-i)(\Upsilon + \rho) + (i+1)\rho]$. Thus in particular the output neuron has a gap in the interval $[t + \delta, t + \delta + (d+1)\rho]$. Since $d \geq 1$, the Gap Lemma applies, and at time instant $t + \delta$ the output of the output neuron depends on spikes in the interval $[t + \delta, t + \delta + (\Upsilon + \rho)]$ of its inputs. All inputs to the output neuron have a gap in the interval $[t + \delta, t + \delta + (\Upsilon + \rho) + d\rho]$, since they have depth atmost $(d-1)$. Since those inputs have a gap in the interval $[t + \delta + (\Upsilon + \rho), t + \delta + (\Upsilon + \rho) + d\rho]$, for $d \geq 2$, the Gap Lemma applies and the output neuron's output at time instant $t + \delta$ depends on outputs of the "penultimate" layer in the interval $[t + \delta, t + \delta + 2(\Upsilon + \rho)]$. Therefore by similar arguments, the output of the output neuron at time instant $t + \delta$ atmost depends on inputs from χ in the interval $[t + \delta, t + \delta + d(\Upsilon + \rho)]$. That is to say that $\mathcal{T}(\chi')$, for every χ' identical to χ in the interval $(-\infty, t + \delta]$, has the same output as $\mathcal{T}(\chi)$ in that same interval, following the corollary to the Gap Lemma. In particular, $\Xi_{(-\infty, t + \delta]} \chi$ is one such χ' . We therefore have $\Xi_{(-\infty, t + \delta]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t + \delta]} \chi)$ which implies $\Xi_{(-\infty, t]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]} \chi)$. Thus, $\mathcal{T}(\cdot)$ is resettable. □

Proof of Proposition 2. Assume that the hypothesis in the proposition is true. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be W -Resettable for some $W \in \mathbb{R}^+$.

We first show a construction for the neuron O , prove that it obeys all the axioms and then show that it has the property that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_I(\chi) \sqcup \chi$ with respect to O .

We first construct the neuron $O(\alpha_O, \Upsilon_O, \rho_O, \tau_O, \lambda_O, m_O, P_O : \bar{\mathcal{S}}_{[0, \Upsilon_O]}^{m_O} \times \bar{\mathcal{S}}_{[0, \rho_O]} \rightarrow [\lambda_O, \tau_O])$. Set $\alpha_O = \alpha$ and $\rho_O, \tau_O \in \mathbb{R}^+$, $\lambda_O \in \mathbb{R}^-$ arbitrarily with $\rho_O \geq \alpha_O$. Set $\Upsilon_O = \max\{U, W\}$ and $m_O = m + 1$. The function $P_O : \bar{\mathcal{S}}_{[0, \Upsilon_O]}^{m_O} \times \bar{\mathcal{S}}_{[0, \rho_O]} \rightarrow [\lambda_O, \tau_O]$ is constructed as follows.

For $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_O]}^{m_O}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_O]}$, set $P_O(\chi', \vec{x}'_0) = \tau_O$ and $P_O(\chi', \vec{\phi}) = \tau_O$ if and only if there exists $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ so that $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ and $\chi' = \Xi_{[0, \Upsilon_O]}(\sigma_t(\mathcal{T}_I(\chi) \sqcup \chi))$

and $\vec{x}'_0 = \Xi_{[0, \rho_0]}(\sigma_t(\mathcal{T}(\chi)))$. Everywhere else, the value of this function is set to zero.

Next, we show it obeys all of the axioms of the single neuron.

We prove that \mathcal{O} satisfies Axiom (1) by showing that its contrapositive is true. Let $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_0]}^{m_0}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_0]}$ be arbitrary so that $P_{\mathcal{O}}(\chi', \vec{x}'_0) = \tau_{\mathcal{O}}$. If $\vec{x}'_0 = \vec{\phi}$, Axiom (1) is immediately satisfied. Thus consider the case when $\vec{x}'_0 = \langle x_0^{1'}, x_0^{2'}, \dots, x_0^{k'} \rangle$. Then $x_0^{1'} \geq \alpha$, otherwise, from the construction of $P_{\mathcal{O}}(\cdot)$, it is immediate that there exists a $\chi \in \mathcal{F}_m$ with $\mathcal{T}(\chi) \notin \mathcal{S}$.

Next, we prove that \mathcal{O} satisfies Axiom (2). Let $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_0]}^{m_0}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_0]}$ be arbitrary. If $P_{\mathcal{O}}(\chi', \vec{x}'_0) = \tau_{\mathcal{O}}$, then it is immediate from the construction that $P_{\mathcal{O}}(\chi', \vec{\phi}) = \tau_{\mathcal{O}}$. On the contrary, if $P_{\mathcal{O}}(\chi', \vec{x}'_0) \neq \tau_{\mathcal{O}}$, from the construction of \mathcal{O} , we have $P_{\mathcal{O}}(\chi', \vec{x}'_0) = 0$. Then the condition in the hypothesis implies that $P_{\mathcal{O}}(\chi', \vec{\phi}) \neq \tau_{\mathcal{O}}$. Therefore, $P_{\mathcal{O}}(\chi', \vec{\phi}) = 0$. Thus, Axiom (2) is satisfied either way.

With Axiom (3), we wish to show $P_{\mathcal{O}}(\vec{\phi}^{m+1}, \vec{\phi}) = 0$. Here, we will show that $P_{\mathcal{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = 0$, for all $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_0]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_0]}$ which implies the required result. Assume, for the sake of contradiction, that there exists a $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_0]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_0]}$, so that $P_{\mathcal{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = \tau_{\mathcal{O}}$. From the construction of \mathcal{O} , this implies that there exists $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ so that $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ and $\Xi_{[0, \Upsilon_0]}(\sigma_t(\chi)) = \vec{\phi}^m$. That is, χ has a gap in the interval $[t, t + W]$, since $\Upsilon_{\mathcal{O}} \geq W$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and W -resettable, by Corollary 3, we have $\Xi_t \mathcal{T}(\chi) = \vec{\phi}$, which is a contradiction. Therefore, we have $P_{\mathcal{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) \neq \tau_{\mathcal{O}}$ and by construction of \mathcal{O} , $P_{\mathcal{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = 0$, for all $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_0]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_0]}$. This implies $P_{\mathcal{O}}(\vec{\phi}^{m+1}, \vec{\phi}) = 0$.

Finally, we wish to show that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_1(\chi) \sqcup \chi$ with respect to \mathcal{O} . That is, we wish to show that for every $\chi \in \mathcal{F}_m$ and for every $t \in \mathbb{R}$, $\Xi_0 \sigma_t(\mathcal{T}(\chi)) = \langle 0 \rangle$ if and only if $P_{\mathcal{O}}(\Xi_{[0, \Upsilon_0]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho_0]}(\sigma_t(\mathcal{T}(\chi)))) = \tau_{\mathcal{O}}$. Consider arbitrary $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$. If $\Xi_0 \sigma_t(\mathcal{T}(\chi)) = \langle 0 \rangle$, then it is immediate from the construction of \mathcal{O} that $P_{\mathcal{O}}(\Xi_{[0, \Upsilon_0]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho_0]}(\sigma_t(\mathcal{T}(\chi)))) = \tau_{\mathcal{O}}$. To prove the converse, suppose $\Xi_0 \sigma_t(\mathcal{T}(\chi)) \neq \langle 0 \rangle$. Then, from the condition in the hypothesis, it follows that for all $\tilde{\chi} \in \mathcal{F}_m$ and for all $\tilde{t} \in \mathbb{R}$ with $\Xi_{[0, \Upsilon_0]}(\sigma_{\tilde{t}}(\mathcal{T}_1(\tilde{\chi}) \sqcup \tilde{\chi})) = \Xi_{[0, \Upsilon_0]}(\sigma_{\tilde{t}}(\mathcal{T}_1(\chi) \sqcup \chi))$, we have $\Xi_0 \sigma_{\tilde{t}}(\mathcal{T}(\tilde{\chi})) = \Xi_0 \sigma_{\tilde{t}}(\mathcal{T}(\chi)) \neq \langle 0 \rangle$. Therefore, from the construction, we have $P_{\mathcal{O}}(\Xi_{[0, \Upsilon_0]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho_0]}(\sigma_t(\mathcal{T}(\chi)))) \neq \tau_{\mathcal{O}}$.

□

Proof of Proposition 3. Assume that the hypothesis in the proposition is true. Let $\mathcal{T} :$

$\mathcal{F}_m \rightarrow \mathcal{S}$ be W' -Resetable for some $W' \in \mathbb{R}^+$. Set $W = \max\{W', 12\alpha\}$. One readily verifies that $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is also W -resettable.

We first show a construction for the neuron l , prove that it obeys all the axioms and then show that it has the property that there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_l(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_l(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_l : \mathcal{F}_m \rightarrow \mathcal{S}$ is such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_l(\chi)$ is consistent with χ with respect to l .

We first construct the neuron $l \langle \alpha_l, \Upsilon_l, \rho_l, \tau_l, \lambda_l, m_l, P_l : \bar{\mathcal{S}}_{[0,\Upsilon_l]}^{m_l} \times \bar{\mathcal{S}}_{[0,\rho_l]} \rightarrow [\lambda_l, \tau_l] \rangle$. Set $\alpha_l = \alpha$. Let $p, q, r \in \mathbb{R}^+$, with²⁸ $p = 8\alpha, q = 2\alpha$ and $r = \alpha$. Set $\Upsilon_l = p + q + r + W$, $\rho_l = p + q + r$ and $m_l = m$. Let $\tau_l \in \mathbb{R}^+, \lambda_l \in \mathbb{R}^-$ be chosen arbitrarily. The function $P_l : \bar{\mathcal{S}}_{[0,\Upsilon_l]}^{m_l} \times \bar{\mathcal{S}}_{[0,\rho_l]} \rightarrow [\lambda_l, \tau_l]$ is constructed as follows.

For $\chi \in \bar{\mathcal{S}}_{[0,\Upsilon_l]}^{m_l}$ and $\vec{x}_0 \in \bar{\mathcal{S}}_{[0,\rho_l]}$, set $P_l(\chi, \vec{x}_0) = \tau_l$ and $P_l(\chi, \vec{\phi}) = \tau_l$ if and only if one of the following is true; everywhere else, the function is set to zero.

1. $\Xi_{(p,p+W]}\chi = \vec{\phi}^{m_l}, \Xi_p\chi \neq \vec{\phi}^{m_l}$ and $\Xi_{[0,p]}\vec{x}_0 = \vec{\phi}$.
2. $\Xi_{[0,p+q]}\vec{x}_0 = \langle t \rangle$, where $q \leq t < (q+r)$ and $(t-q) = \varepsilon(0, \Xi_{(0,p]}\sigma_t(\chi))$. Moreover, $\Xi_{(t+p,t+p+W]}\chi = \vec{\phi}^{m_l}$ and $\Xi_{(p+t)}\chi \neq \vec{\phi}^{m_l}$.
3. $\Xi_{[0,2p-(q+r)]}\vec{x}_0 = \langle t_x, t_y \rangle$ with $(p - (q + r)) \leq t_x \leq (p - q) \leq t_y = p$. Also, for all $t' \in [0, p]$, $\Xi_{(t',t'+W]}\chi \neq \vec{\phi}^{m_l}$.
4. $\Xi_{[0,2p-r]}\vec{x}_0 = \langle t, t_x, t_y \rangle$ with $q \leq t \leq (q + r) < (p - r) \leq t_x \leq p \leq t_y = p + t$ and $(t - q) = \varepsilon((t_y - t_x - q), \Xi_{(0,p]}\sigma_t(\chi))$. Furthermore, for all $t' \in [0, p + t]$, $\Xi_{(t',t'+W]}\chi \neq \vec{\phi}^{m_l}$.

where $\varepsilon : [0, r] \times \bar{\mathcal{S}}_{(0,p]}^{m_l} \rightarrow [0, r]$ is as defined below.

For convenience, we define an operator $\iota_j^k : [0, 1) \rightarrow [0, 1)$, for $j, k \in \mathbb{Z}^+$, that constructs a new number obtained by concatenating every i^{th} digit of a given number, where $i \equiv j \pmod k$. More formally, for $x \in [0, 1)$, $\iota_j^k(x) = \sum_{i=1}^{\infty} ((\lfloor x \times 10^{j+(i-1)k} \rfloor - 10 \lfloor x \times 10^{j+(i-1)k-1} \rfloor) \times 10^{-i})$.

Also, we define another operator $\zeta_k : [0, 1)^k \rightarrow [0, 1)$, for $k \in \mathbb{Z}^+$ which ‘‘interleaves’’ the digits of k given numbers in order to produce a new number. More formally,

²⁸The choice of values for p, q, r and W was made so as to satisfy the following inequalities, which we will need in the proof: $p < W, p > 2(q + r)$ and $q > \alpha$.

for $x_0, x_1, \dots, x_{k-1} \in [0, 1)$, $\zeta_k(x_0, x_1, \dots, x_{k-1}) = \sum_{i=0}^{\infty} ((\lfloor x_{k(\frac{i}{k} - \lfloor \frac{i}{k} \rfloor)} \rfloor \times 10^{1+\lfloor \frac{i}{k} \rfloor}) - 10 \lfloor x_{k(\frac{i}{k} - \lfloor \frac{i}{k} \rfloor)} \rfloor \times 10^{\lfloor \frac{i}{k} \rfloor}) \times 10^{-(i+1)}$.

Let d be the largest integer so that, for all $x' \in [0, r]$, we have $x' \times 10^d < 1$. For $x' \in [0, r]$, let $x = x' \times 10^d$. For $\chi \in \bar{\mathcal{S}}_{(0,p)}^{m_1}$, define²⁹ $\varepsilon(x', \chi) = 10^{-d} \times \zeta_{m_1}(\varepsilon_0(\iota_1^{m_1}(x), \Pi_1(\chi)), \varepsilon_0(\iota_2^{m_1}(x), \Pi_2(\chi)), \dots, \varepsilon_0(\iota_{m_1}^{m_1}(x), \Pi_{m_1}(\chi)))$, where $\varepsilon_0 : [0, 1) \times \bar{\mathcal{S}}_{(0,p]} \rightarrow [0, 1)$ is as defined below.

Let $n \in [0, 1)$ and $\vec{x} \in \bar{\mathcal{S}}_{(0,p]}$. Furthermore, let $c = \iota_1^2(n)$ and $s = \iota_2^2(n)$. Let $\vec{x} = \langle x^1, x^2, \dots, x^k \rangle$. We have $0 \leq k \leq 8$, because $p = 8\alpha$. Also, since $p = 8r$, we have $x^i \times 10^{d-1} < 1$, for $1 \leq i \leq k$. Let $s' = \zeta_{k+1}(x^1 \times 10^{d-1}, x^2 \times 10^{d-1}, \dots, x^k \times 10^{d-1}, s)$. If $c = 0$, then let $c' = \frac{k}{10} + 0.09$ else let $c' = \frac{k}{10} + \frac{c}{10}$. Finally, define $\varepsilon_0(n, \vec{x}) = \zeta_2(c', s')$.

Next, we show that l satisfies all the axioms of the neuron.

It is immediate that l satisfies Axiom (1), since all output spikes in the above construction are atleast q seconds apart, and $q = 2\alpha$.

We now prove that l satisfies Axiom (2). Let $\chi' \in \bar{\mathcal{S}}_{[0, \gamma_1]}^{m_1}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_1]}$ be arbitrary. If $P_1(\chi', \vec{x}'_0) = \tau_1$, then it is immediate from the construction that $P_1(\chi', \vec{\phi}) = \tau_1$ which satisfies Axiom (2). On the contrary, if $P_1(\chi', \vec{x}'_0) \neq \tau_1$, from the construction of l , we have $P_1(\chi', \vec{x}'_0) = 0$. Also, from the construction we have either $P_1(\chi', \vec{\phi}) = 0$ or $P_1(\chi', \vec{\phi}) = \tau_1$. Axiom (2) is satisfied in either case.

Also, l satisfies Axiom (3), since it is clear that $\chi = \vec{\phi}^{m_1}$ does not satisfy any of the conditions enumerated above. We therefore have $P_1(\vec{\phi}^{m_1}, \vec{\phi}) = 0$.

Finally, we show that there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_I(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_I : \mathcal{F}_m \rightarrow \mathcal{S}$ such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_I(\chi)$ is consistent with χ with respect to l . From Proposition 5, it follows that there exist $V_1, V_2 \in \mathbb{R}^+$ so that $\Xi_{[0,V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0,V_2]}(\sigma_{t_2}(\chi_2))$. Let $U = p + q + r$. If $\Xi_{[0,U]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\chi_2))$, it is immediate that $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_I(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$. It therefore suffices to prove that if $\Xi_{[U,V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[U,V_2]}(\sigma_{t_2}(\chi_2))$ then $\Xi_{[0,U]}(\sigma_{t_1} \mathcal{T}_I(\chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2} \mathcal{T}_I(\chi_2))$. Proposition 5 implies that $\Xi_{(V_1, V_1+W]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$ and $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$. Therefore, by Case (1) of the construction, $\Xi_{(V_1-p)} \sigma_{t_1} \mathcal{T}_I(\chi_1) = \langle V_1-p \rangle$. Moreover, since Proposition 5 implies that for all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1, t'_1+W]}(\sigma_{t_1}(\chi_1)) \neq$

²⁹Recall that the *projection operator for spike-train ensembles* is defined as $\Pi_i(\chi) = \vec{x}_i$, for $1 \leq i \leq m$, where $\chi = \langle \vec{x}_1, \vec{x}_2, \dots, \vec{x}_m \rangle$.

$\vec{\phi}^m$, from Case (3) of the construction, we have that for every $k \in \mathbb{Z}^+$ with $V_1 - kp > 0$, $\Xi_{(V_1 - kp)\sigma_{t_1}} \mathcal{T}_1(\chi_1) = \langle V_1 - kp \rangle$. Let k_1 be³⁰ the smallest positive integer, so that $V_1 - k_1 p \leq U$. From the previous arguments, we have $\Xi_{(V_1 - k_1 p)\sigma_{t_1}} \mathcal{T}_1(\chi_1) = \langle V_1 - k_1 p \rangle$. Also, it is easy to see that $V_1 - k_1 p \geq (q + r)$. Let k_2 be similarly defined with respect to χ_2 so that $\Xi_{(V_2 - k_2 p)\sigma_{t_2}} \mathcal{T}_1(\chi_2) = \langle V_2 - k_2 p \rangle$ and $V_2 - k_2 p \leq U$. Now, there are two cases:

1. If $V_1 - k_1 p \neq V_2 - k_2 p$, we now show that $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$, which is the required result. Assume, without loss of generality, that $V_1 - k_1 p < V_2 - k_2 p$. If these two quantities are less than $p - r$ apart, we have $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$, because by Case (4) of the construction $\mathcal{T}_1(\chi_1)$ has a spike in the interval $[V_1 - k_1 p - (q + r), V_1 - k_1 p - q]$ and by Case (3) of the construction, $\mathcal{T}_1(\chi_2)$ has no spike in the interval $(V_2 - k_2 p, V_2 - k_2 p + p - (q + r)]$. In other words, the spike following the one at $V_1 - k_1 p$ in $\mathcal{T}_1(\chi_1)$ has no counterpart in $\mathcal{T}_1(\chi_2)$. On the other hand, if they are less than p apart but at most $p - r$ apart, by similar arguments, it is easy to show that the spike at $V_2 - k_2 p$ in $\mathcal{T}_1(\chi_2)$ has no counterpart in $\mathcal{T}_1(\chi_1)$. Finally, if they are at least p apart, then k_2 does not satisfy the property that it is the smallest positive integer, so that $V_2 - k_2 p \leq U$, which is a contradiction.
2. On the contrary, consider the case when $V_1 - k_1 p = V_2 - k_2 p$. We have two cases:
 - (a) Suppose $k_1 \neq k_2$. Let t'_1 be the largest positive integer so that $\Xi_{t'_1 \sigma_{t_1}} \mathcal{T}_1(\chi_1) = \langle t'_1 \rangle$ and $t'_1 < V_1 - k_1 p$. From Case (4) of the construction, we have that $q \leq (V_1 - k_1 p) - t'_1 \leq q + r$. Let t'_2 be defined likewise, with respect to χ_2 . Further, let $n'_1 = (V_1 - k_1 p) - t'_1 - q$ and $n'_2 = (V_2 - k_2 p) - t'_2 - q$ and $n_1 = n'_1 \times 10^d$ and $n_2 = n'_2 \times 10^d$. Since $k_1 \neq k_2$, it is straightforward to verify that for all j with $1 \leq j \leq m_1$, $\iota_1^2(\iota_j^{m_1}(n_1)) \neq \iota_1^2(\iota_j^{m_1}(n_2))$, for the former number has 9 in the $(k_1 + 1)^{th}$ decimal place, while the latter number does in the $(k_2 + 1)^{th}$ decimal place and not in the $(k_1 + 1)^{th}$ decimal place since $k_1 \neq k_2$. Therefore, $n_1 \neq n_2$ and consequently $t'_1 \neq t'_2$ which gives us $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$, which is the required result.

³⁰ k_1 exists because $U > p$.

(b) On the other hand, suppose $k_1 = k_2$. Again, we have two cases:

- i. Suppose, there exists a j with $1 \leq j \leq m_1$ and a $k' \leq k_1$, so that $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ has a different number of spikes when compared to $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$. Let n_1, n_2 be defined as before. It is straightforward to verify that $\iota_1^2(\iota_j^{m_1}(n_1)) \neq \iota_1^2(\iota_j^{m_1}(n_2))$, because they differ in the $(k_1 - k' + 1)^{th}$ decimal place³¹. Therefore, $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$.
- ii. Now consider the case where for all j with $1 \leq j \leq m_1$ and $k' \leq k_1$, we have $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ have the same number of spikes when compared to $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$. Now, by hypothesis, we have $\Xi_{[U, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[U, V_2]}(\sigma_{t_2}(\chi_2))$. Therefore there must exist a $1 \leq j \leq m_1$ and $k' \leq k_1$, so that there is a point in time where one of the spike-trains $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ and $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$ has a spike, while the other does not. Let t' be the latest time instant at which this is so. Also, assume without loss of generality that $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1)) = \langle x^1, \dots, x^q \rangle$ has a spike at time instant t' while $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$ does not. Let p be the number so that $t' = x^p$. Let n_1, n_2 be defined as before. Also, for each h with $1 \leq h \leq k_1$, let r_h be the number of spikes in $\Xi_{(V_1-hp, V_1-(h-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$. Each r_h can be determined from n_1 . Then, it is straightforward to verify³² that $\iota_p^{r_{k'}} \iota_{r_{k'-1}}^{r_{k'-1}} \dots \iota_{r_1}^{r_1} \iota_2^2 \iota_j^{m_1} n_1 \neq \iota_p^{r_{k'}} \iota_{r_{k'-1}}^{r_{k'-1}} \dots \iota_{r_1}^{r_1} \iota_2^2 \iota_j^{m_1} n_2$. Therefore, $n_1 \neq n_2$ and it follows that $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$.

□

Some auxiliary propositions used in the proofs of Propositions 2 and 3

Proposition 4. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant, then $\mathcal{T}(\vec{\phi}^m) = \vec{\phi}$.*

³¹Which in n_1 and n_2 encodes the number of spikes in the interval $(V_2 - k'p, V_2 - (k' - 1)p]$ on the j^{th} spike-train of χ_1 and χ_2 respectively.

³²The expression on either side of the inequality is a real number that encodes for the p^{th} spike time in the spike-trains $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ and $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$ respectively.

Proof. For the sake of contradiction, suppose $\mathcal{T}(\vec{\phi}^m) = \vec{x}_0$, where $\vec{x}_0 \neq \vec{\phi}$. That is, there exists a $t \in \mathbb{R}$ with $\Xi_t \vec{x}_0 = \langle t \rangle$. Let $\delta < \alpha$. Clearly, $\sigma_\delta(\vec{\phi}^m) = \vec{\phi}^m \in \mathcal{F}_m$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant, $\mathcal{T}(\sigma_\delta(\vec{\phi}^m)) = \sigma_\delta(\mathcal{T}(\vec{\phi}^m)) = \sigma_\delta(\vec{x}_0)$. Now, $\sigma_\delta(\vec{x}_0) \neq \vec{x}_0$ since $\Xi_{(t-\delta)} \sigma_\delta(\vec{x}_0) = \langle t - \delta \rangle$ whereas $\Xi_{(t-\delta)} \vec{x}_0 = \vec{\phi}$, for otherwise $\vec{x}_0 \notin \mathcal{S}$. This is a contradiction. Therefore, $\mathcal{T}(\vec{\phi}^m) = \vec{\phi}$. \square

Corollary 3. *Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and W -resettable, for some $W \in \mathbb{R}^+$. If $\chi \in \mathcal{F}_m$ has a gap in the interval $[t, t + W]$, then $\Xi_t \mathcal{T}(\chi) = \vec{\phi}$.*

Proof. Assume the hypothesis of the above statement. One readily sees that $\Xi_t \mathcal{T}(\chi) = \Xi_{[t, \infty)} \Xi_{(-\infty, t]} \mathcal{T}(\chi)$. Now, since χ has a gap in the interval $(t, t + W]$ and $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is W -resettable, we have $\Xi_{[t, \infty)} \Xi_{(-\infty, t]} \mathcal{T}(\chi) = \Xi_{[t, \infty)} \mathcal{T}(\Xi_{(-\infty, t]} \chi)$. Further, since $\Xi_t \chi = \vec{\phi}^m$, we have $\Xi_{[t, \infty)} \Xi_{(-\infty, t]} \chi = \Xi_{[t, \infty)} \vec{\phi}^m$. Therefore, since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, it follows that $\Xi_{[t, \infty)} \mathcal{T}(\Xi_{(-\infty, t]} \chi) = \Xi_{[t, \infty)} \mathcal{T}(\vec{\phi}^m) = \vec{\phi}$, with the last equality following from the previous proposition. \square

Proposition 5. *Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and W' -resettable, for some $W' \in \mathbb{R}^+$. Then for all $W \in \mathbb{R}^+$ with $W \geq W'$, $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, there exist $V_1, V_2 \in \mathbb{R}^+$ so that the following are true.*

1. $\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2))$
2. $\Xi_{(V_1, V_1 + W]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$, $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and $\Xi_{(V_2, V_2 + W]}(\sigma_{t_2}(\chi_2)) = \vec{\phi}^m$, $\Xi_{V_2}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$
3. For all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1, t'_1 + W]}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and for all $t'_2 \in [0, V_2)$, $\Xi_{(t'_2, t'_2 + W]}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$.

Proof. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, we have $\Xi_{[t_1, \infty)} \mathcal{T}(\chi_1) = \Xi_{[t_1, \infty)} \mathcal{T}(\Xi_{[t_1, \infty)} \chi_1)$. This implies $\sigma_{t_1}(\Xi_{[t_1, \infty)} \mathcal{T}(\chi_1)) = \sigma_{t_1}(\Xi_{[t_1, \infty)} \mathcal{T}(\Xi_{[t_1, \infty)} \chi_1))$ which gives us $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\Xi_{[t_1, \infty)} \chi_1))$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant and $\sigma_{t_1}(\Xi_{[t_1, \infty)} \chi_1) = \Xi_{[0, \infty)} \sigma_{t_1}(\chi_1) \in \mathcal{F}_m$, we have $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\Xi_{[t_1, \infty)} \chi_1)) = \Xi_{[0, \infty)} \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_1}(\chi_1))$. In short, $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_{[0, \infty)} \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_1}(\chi_1))$ which implies

$\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_0 \mathcal{T}(\Xi_{[0,\infty)} \sigma_{t_1}(\chi_1))$. Similarly, $\Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2)) = \Xi_0 \mathcal{T}(\Xi_{[0,\infty)} \sigma_{t_2}(\chi_2))$. Therefore, it follows from the hypothesis that $\Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))) \neq \Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_2}(\chi_2)))$.

Let $V_1, V_2 \in \mathbb{R}^+$ be the smallest positive real numbers so that $\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))$ and $\Xi_{[0,\infty)}(\sigma_{t_2}(\chi_2))$ have gaps in the intervals $(V_1, V_1 + W]$ and $(V_2, V_2 + W]$ respectively. That such V_1, V_2 exist follows from the fact that $\chi_1, \chi_2 \in \mathcal{F}_m$. Since, $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is W' -resettable, it is also W -resettable for $W \geq W'$. It therefore follows that $\Xi_{(-\infty, V_1]} \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))) = \mathcal{T}(\Xi_{(-\infty, V_1]} \Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1)))$ which equals $\mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)))$. This implies that $\Xi_0 \Xi_{(-\infty, V_1]} \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))) = \Xi_0 \mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)))$ due to which we have $\Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))) = \Xi_0 \mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)))$. Likewise, $\Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_2}(\chi_2))) = \Xi_0 \mathcal{T}(\Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2)))$. We therefore have $\Xi_0 \mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1))) \neq \Xi_0 \mathcal{T}(\Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2)))$. This readily implies $\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2))$ and, from the construction, it follows that $\Xi_{(V_1, V_1+W]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$, $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and $\Xi_{(V_2, V_2+W]}(\sigma_{t_2}(\chi_2)) = \vec{\phi}^m$, $\Xi_{V_2}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$, for otherwise V_1 or V_2 would not be the smallest choice of numbers with the said property. Furthermore, for the same reasons, for all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1, t'_1+W]}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and for all $t'_2 \in [0, V_2)$, $\Xi_{(t'_2, t'_2+W]}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$.

□

References

- Abbott, L. (2008). Theoretical neuroscience rising. *Neuron*, 60(3), 489–495.
- Banerjee, A. (2001). On the phase-space dynamics of systems of spiking neurons. I: Model and experiments. *Neural Comp.*, 13(1), 161–193.
- Banerjee, A. (2006). On the sensitive dependence on initial conditions of the dynamics of networks of spiking neurons. *J. Comput. Neurosci.*, 20(3), 321–348.
- Bartlett, P., & Maass, W. (2003). Vapnik-Chervonenkis dimension of neural nets. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed., pp. 1188–1192). Cambridge, MA: MIT Press.
- Bohte, S., Kok, J., & La Poutre, H. (2002). Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing*, 48(1-4), 17–37.

- Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. New York, NY: Cambridge University Press.
- Gutig, R., & Sompolinsky, H. (2006). The tempotron: a neuron that learns spike timing–based decisions. *Nat. Neurosci.*, 9(3), 420–428.
- Maass, W. (1996). Lower bounds for the computational power of networks of spiking neurons. *Neural Comp.*, 8(1), 1–40.
- Nirenberg, S., Carcieri, S., Jacobs, A., & Latham, P. (2001). Retinal ganglion cells act largely as independent encoders. *Nature*, 411(6838), 698–701.
- Poirazi, P., Brannon, T., & Mel, B. (2003). Pyramidal neuron as two-layer neural network. *Neuron*, 37(6), 989–999.
- Rieke, F., Warland, D., Steveninck, R. van, & Bialek, W. (1997). *Spikes: exploring the neural code*. Cambridge, MA: MIT Press.
- Shepherd, G. (2004). *The synaptic organization of the brain*. New York, NY: Oxford University Press.
- Strehler, B., & Lestienne, R. (1986). Evidence on precise time-coded symbols and memory of patterns in monkey cortical neuronal spike trains. *Proc. Nat. Acad. Sci. USA*, 83(24), 9812.