Studies in Applied Mathematics. Vol. XLIX No. 2, June 1970.

Some Networks that can Learn, Remember, and Reproduce any Number of Complicated Space-time Patterns, II

By Stephen Grossberg

1. Introduction

This paper describes some networks $\mathcal{M}$ that can learn, simultaneously remember, and perform individually upon demand any number of spatiotemporal patterns (e.g., "motor sequences" and "internal perceptual representations") of essentially arbitrary complexity. Because these networks are *embedding fields*, they can be given a suggestive psychological, neurophysiological, and anatomical interpretation ([1]–[14]). [14] describes some of the mathematical properties of these networks using this heuristic interpretation. They include the following:

a) "Practice makes perfect".
b) Learning occurs by a mixture of operant and respondant conditioning factors, which can include different network responses to "novel" vs. "habituated" stimuli, the existence of "nonspecific arousal" and "internal drive" stimuli, of "sensory" feedback due to prior "motor" outputs, and of "paying attention" by the network to those inputs which at any time help the network achieve its "goals".
c) New patterns can be learned without at all destroying the memory of old patterns.
d) All errors can be corrected.
e) Memory either decays at an exponential rate—which can be made arbitrarily small—or is perfect until "unrewarded" recall trials occur, during which memory is "extinguished". In both cases, "spontaneous recovery" and spontaneous improvement of memory (i.e., "reminiscence") can occur.
f) A single network "nerve", with sufficiently many "axon collaterals" activated successively by "avalanche conduction" can, in principle, learn an essentially arbitrarily complicated pattern, though in a rote way.
g) A concrete "stimulus sampling" operation occurs in the networks, and concrete analogs of "stimulus sampling probabilities" exist.
h) The network is insensitive to wild "behaviorally irrelevant" oscillations of inputs and often has a monotonic response to them.
i) Network dynamics can be globally analyzed.

[12] discusses a related class of networks whose memory is essentially perfect even during recall trials.

2. Network equations

We will establish the above properties for a class of closely related network equations. These include examples chosen from the following two network types.

$$\dot{x}_i(t) = -\alpha_i x_i(t) + \sum_{m=1}^{n} [x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ p_{mi} z_{mi}(t)$$
$$- \sum_{m=1}^{n} [x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ q_{mi} + I_i(t), \tag{1}$$

and

$$\dot{z}_{jk}(t) = -u_{jk} z_{jk}(t) + v_{jk}[x_j(t - \tau_{jk}) - \Gamma_{jk}]^+ x_k(t), \tag{2}$$

or (1) taken along with

$$\dot{z}_{jk}(t) = [-u_{jk} z_{jk}(t) + v_{jk} x_k(t)][x_j(t - \tau_{jk}) - \Gamma_{jk}]^+, \tag{3}$$

for $i, j, k = 1, 2, \ldots, n$, where for any real number η,

$$[\eta]^+ = \max(\eta, 0)$$

and introduces various threshold cut-offs. The system (1) and (2) was derived in [2] and [3], and is reviewed in [1]. (3) differs from (2) only by replacing the decay rate u_{jk} of (2) by

$$u_{jk}[x_j(t - \tau_{jk}) - \Gamma_{jk}]^+, \tag{4}$$

which, it will appear, produces an essentially perfect network memory. Henceforth the system (1) and (2) will occasionally be denoted by (*), whereas (1) and (3) will be denoted by (**).

The parameters, initial data, and inputs of (1)–(3) satisfy the following constraints.

(I) *Parameters*

(1) All constant parameters are nonnegative, e.g., $\alpha_i, p_{jk}, \Gamma_{jk}, u_{jk}$.
(2) v_{jk} is positive only if p_{jk} is positive.
(3) All time lags τ_{jk} are positive.
(4) $p_{jk} q_{jk} = 0$.

(II) *Initial data*

All initial data of $x_i(v)$ and $z_{jk}(v)$ for $v \leq 0$ is nonnegative and continuous. Moreover we suppose for convenience that $z_{jk}(0) > 0$ if and only if $p_{jk} > 0$.

(III) *Inputs*

All inputs $I_i(t)$ are bounded, nonnegative, and continuous for $t \geq 0$ and vanish for $t \leq 0$.

When we say henceforth that parameters, initial data, or inputs are chosen "arbitrarily", we will always mean "arbitrarily subject to (I)–(III)".

3. Cross-correlated flows on signed networks

For every choice of parameters, initial data, and inputs, (*) or (**) describes a cross-correlated flow on a signed network $\mathscr{M}$. Since variants of this flow have been previously described ([1]–[14]), the following summary will be brief.

A finite directed network $G = (V, E)$ is determined by its vertices $V = \{v_i: i = 1, 2, \ldots, n\}$ and its directed edges $E = \{e_{jk}: j, k = 1, 2, \ldots, n\}$. e_{jk} is drawn as an arrow facing from the point v_j with its arrowhead N_{jk} touching the point v_k. Henceforth the following heuristic terminology will sometimes be used to discuss G:

$$v_i = i\text{th "cell body" cluster,}$$

$$e_{jk} = \text{cluster of "axons" from } v_j \text{ to } v_k,$$

and

$$N_{jk} = \text{cluster of "synaptic knobs" at the terminal ends of } e_{jk} \text{ axons.}$$

$x_i(t)$ is a process fluctuating at v_i, and $z_{jk}(t)$ is a process fluctuating at N_{jk}. Each of these processes has a mathematical, psychological, and neural name. The psychological and neural names are used to facilitate comparison and contrast of network dynamics with the behavior of living organisms. Thus

$$\begin{aligned} x_i(t) &= i\text{th vertex function, or} \\ &= i\text{th "stimulus trace", or} \\ &= i\text{th "average membrane potential",} \end{aligned}$$

and

$$\begin{aligned} z_{jk}(t) &= (j, k)\text{th edge (or interaction) function, or} \\ &= \text{"associational strength" from } v_j \text{ to } v_k \text{, or} \\ &= \text{"average activity of excitatory transmitter producing process" in } N_{jk}. \end{aligned}$$

By (I4) of Section 2, either $p_{mi} = 0$ or $q_{mi} = 0$ for every m and i. If $p_{mi} > 0$, then e_{mi} is called an *excitatory edge*. If $q_{mi} > 0$, then e_{mi} is called an *inhibitory edge*. Suppose that e_{mi} is excitatory, then at every time $t - \tau_{mi}$ the "average membrane potential" $x_m(t - \tau_{mi})$ at v_m creates an excitatory signal (or "spiking frequency") of size

$$[x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ p_{mi} \tag{5}$$

in e_{mi}. (5) is positive only if $x_m(t - \tau_{mi})$ exceeds the *signal threshold* Γ_{mi}. The signal (5) flows, or is transmitted, at a finite velocity along e_{mi}, and reaches the arrowhead N_{mi} at time t. It thereupon interacts with the "transmitter process" $z_{mi}(t)$, and a signal of size

$$[x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ p_{mi} z_{mi}(t) \tag{6}$$

is released from N_{mi}, reaches v_i by crossing the "synaptic cleft" between N_{mi} and v_i, and thereupon perturbs the "average postsynaptic potential" x_i. All excitatory signals from the various "presynaptic cells" v_m with $p_{mi} > 0$ combine additively at v_i, yielding the second term on the right hand side of (1).

If $q_{mi} > 0$, then an inhibitory signal of size

$$[x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ q_{mi} \tag{7}$$

leaves v_m at time $t - \tau_{mi}$ along the inhibitory edge e_{mi}, reaches v_i at time t, and thereby perturbs x_i. All inhibitory signals combine additively at v_i to yield the third term on the right hand side of (1). The minus sign before this term shows that increasing (7) decreases x_i, other things equal. Hence the term "inhibitory".

$x_i(t)$ also decays exponentially at a rate α_i, and is perturbed by the input $I_i(t)$ which is under the control of an experimentalist, other external environmental factors, or other control cells. See Figure 1.

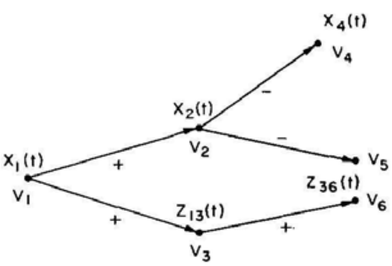

Figure 1.

In both (2) and (3), $z_{mi}(t)$ sits in N_{mi} and *cross-correlates* the signal (5) received by N_{mi} at time t with the contiguous value $x_i(t)$ of v_i; hence the condition (I2) of Section 2. Speaking heuristically, the transmitter production rate is controlled by cross-correlation of the pre- and post-synaptic potentials. This is the main learning mechanism of the networks. In (2), $z_{mi}(t)$ also delays exponentially at the rate u_{mi}, whereas in (3), $z_{mi}(t)$ decays at the rate (4), which is proportional to the spiking frequency created by v_m in e_{mi} τ_{mi} time units earlier. Processes in N_{mi} which are coupled to spiking frequency are interpreted in [3], by analogy with physiological data concerning the action potential, as being triggered by an increase in Na^+ and a decrease in K^+ concentration within N_{mi}. Given this interpretation, our mathematical results show that coupling the decay of transmitter production activity to the action potential via Na^+ and K^+ produces essentially perfect memory.

Each choice of the matrices $P = \|p_{jk}\|$ and $Q = \|q_{jk}\|$ defines a different "anatomy" for a network $\mathscr{M}$ by picking out the directed paths $v_j \xrightarrow{+} v_k$ or $v_j \xrightarrow{-} v_k$ over which excitatory or inhibitory signals, respectively, can be transmitted, and the relative strengths of these signals. Variations in (P, Q) can dramatically change the qualitative properties of learning, memory, and recall in $\mathscr{M}$ ([1]–[13]).

4. Space-time pattern learning by an alternative system

The following type of network was studied in [12].

$$\dot{x}_i(t) = -\alpha_i x_i(t) + \sum_{m=1}^{n} \beta_m [x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ y_{mi}(t)$$

$$- \sum_{m=1}^{n} [x_m(t - \tau_{mi}) - \Gamma_{mi}]^+ q_{mi} + I_i(t), \tag{8}$$

$$y_{jk}(t) = p_{jk} z_{jk}(t) \left[\sum_{m=1}^{n} p_{jm} z_{jm}(t) \right]^{-1}, \tag{9}$$

and

$$\dot{z}_{jk}(t) = -u_{jk}z_{jk}(t) + v_{jk}[x_j(t - \tau_{jk}) - \Gamma_{jk}]^+ x_k(t). \tag{10}$$

The excitatory signal reaching v_i from v_m in (8) depends on y_{mi}, rather than z_{mi}, and y_{mi} is a ratio of cross-correlators z_{mk}, as in (9). This property gives the network a perfect memory even during recall trials for suitable choices of anatomy. It creates two deficiencies, however, which are easily understood in the special case of a Γ-*outstar* $\mathscr{M}^{(1)}$ defined below.

$$\dot{x}_1(t) = -\alpha_1 x_1(t) + I_1(t), \tag{11}$$

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta[x_1(t - \tau) - \Gamma_1]^+ y_{1i}(t) + I_i(t), \tag{12}$$

$$y_{1i}(t) = z_{1i}(t)\left[\sum_{m=2}^{n} z_{1m}(t)\right]^{-1}, \tag{13}$$

and

$$\dot{z}_{1i}(t) = -uz_{1i}(t) + v[x_1(t - \tau) - \Gamma_1]^+ x_i(t), \tag{14}$$

$i = 2, \ldots, n$. Since only the excitatory edges e_{1i}, $i \neq 1$, transmit signals in (11)–(14), v_1 is called the *source vertex*, v_i a *sink vertex*, $i \neq 1$, and $B_n = \{v_i : i = 2, \ldots, n\}$ is called the *border*. B_n is thought of as a grid embedded in a region $\mathscr{R}$ upon which inputs play, as in Figure 2. $\mathscr{M}^{(1)}$ learns by "respondant conditioning". For

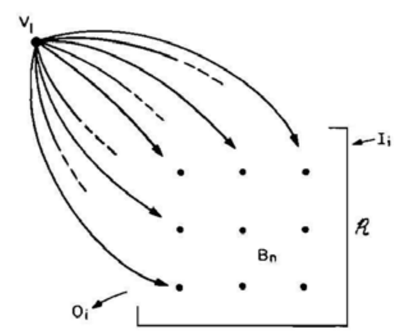

Figure 2.

example, let an intense input pulse ("conditioned stimulus") perturb v_1 and create excitatory signals in the edges e_{1i}. Suppose that these signals reach the arrowheads N_{1i} while a spatial pattern ("unconditioned stimulus")

$$I_i(t) = \theta_i I(t) \tag{15}$$

reaches B_n, where θ_i is the relative intensity of the pattern reaching v_i, and $I(t)$ is the total pattern intensity at time t. If such a pairing of input pulses to v_1 and B_n occurs sufficiently often with sufficient intensity, then a later input pulse to v_1 alone will recreate the spatial pattern with weights

$$\theta = \{\theta_i : i = 2, \ldots, n\} \text{ on } B_n.$$

There exists only one set of weights θ that $\mathscr{M}^{(1)}$ cannot learn. This is the trivial pattern $\theta = 0$. To see this, let v_1 receive an input pulse that creates a positive signal $\beta[x_1(t - \tau) - \Gamma_1]^+$ in each e_{1i}. Since, by (13), $\sum_{i=2}^{n} y_{1i}(t) = 1$, the total signal received by B_n from v_1 equals

$$\sum_{i=2}^{n} \beta[x_1(t-\tau) - \Gamma_1]^+ y_{1i}(t) = \beta[x_1(t-\tau) - \Gamma_1]^+.$$

This signal does not approach zero even if the cross-correlators $z_{1i}(t)$ approach zero in response to the zero pattern at B_n.

Ability to learn the zero pattern is important when a space-time pattern delivered to B_n is being learned. Suppose that a space-time pattern is given with total intensity $I(t) = \sum_{m=2}^{n} I_m(t)$ and relative intensities $\theta_i(t) = I_i(t)/I(t)$. If $I(t)$ is never zero, then a sequence of Γ-outstars can approximately learn the weights $\theta_i(t)$ by the mechanism depicted in Figure 3. Figure 3 describes a Γ-*outstar avalanche*

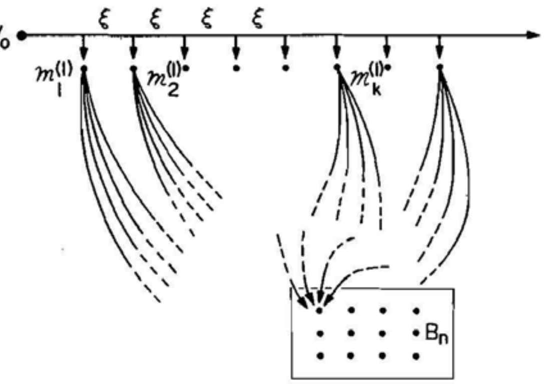

Figure 3.

[12]. The control vertex v_0 creates a signal in the long horizontal edge which perturbs a different Γ-outstar every ξ time units. Each successive Γ-outstar $\mathcal{M}_k^{(1)}$, $k = 1, 2, \ldots, K$, briefly samples the weights $\theta_i(t)$ playing on B_n during successive time intervals. In this way, the Γ-outstar avalanche learns a space-time pattern $\theta(t)$ as a sequence

$$\theta(T + k\xi), \qquad k = 1, 2, \ldots$$

of spatial patterns, and the kth Γ-outstar $\mathcal{M}_k^{(1)}$ learns the kth spatial approximation $\theta(T + k\xi)$ to the space-time pattern. That is, the "moving picture" $\theta(t)$ is learned as a sequence of "still pictures". If, however, certain spatial approximations $\theta(T + k\xi)$ with $k = m_1, m_2, \ldots, m_r$, have zero intensity, the Γ-outstars $\mathcal{M}_k^{(1)}$, $k = m_1, \ldots, m_r$, will spray B_n with substantial amounts of background noise on recall trials. The source of this difficulty is the normalization condition $\sum_{i=2}^{n} y_{1i}(t) = 1$ which keeps the total signal from v_1 to B_n large if $x_1(t-\tau)$ is large even when each z_{1i} is small. (*) and (**) eliminate this difficulty.

The second difficulty with (11)–(14) is a conceptual one. y_{1i} sits in N_{1i} and controls the flow from v_1 to v_i. By (13), y_{1i} depends on all z_{1m}, and z_{1m} sits in N_{1m}. How does the z_{1m} value instantaneously jump from N_{1m} to N_{1i} so that y_{1i} can be computed? [3] discusses the physical meaning of this jumping process in terms of "competition between response alternatives" and replaces the jumping process by "postsynaptic lateral inhibition coupled to the presynaptic transmitter production process". This process creates $\mathcal{M}^{(1)}$'s perfect memory.

The system (*), which lacks y_{1i}, forgets what it has learned at the exponential rate u_{jk}. (**), by contrast, remembers perfectly in the absence of practice and recall trials. Thus either of two mechanisms can improve the network's memory:

a) coupling of increases in transmitter production to postsynaptic lateral inhibition, or
b) coupling of decreases in transmitter production to presynaptic radial excitation.

Note the dualism in (a) and (b) between the terms

increases
postsynaptic
lateral
inhibition

and the terms

decreases
presynaptic
radial
excitation.

Both (*) and (**) will be seen to have essentially perfect memory of *relative* associational strengths y_{1i} in a Γ-outstar anatomy.

5. Γ-Outstars of type (*) and (**)

A Γ-outstar of type (*) is given by

$$\dot{x}_1(t) = -\alpha_1 x_1(t) + I_1(t), \tag{16}$$

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta[x_1(t-\tau) - \Gamma_1]^+ z_{1i}(t) + I_i(t), \tag{17}$$

and

$$\dot{z}_{1i}(t) = -uz_{1i}(t) + v[x_1(t-\tau) - \Gamma_1]^+ x_i(t), \tag{18}$$

$i = 2, \ldots, n$. A Γ-outstar of type (**) satisfies (16), (17), and

$$\dot{z}_{1i}(t) = [-uz_{1i}(t) + vx_i(t)][x_1(t-\tau) - \Gamma_1]^+, \tag{19}$$

$i = 2, \ldots, n$. Henceforth (16)–(18) will be denoted by $\mathscr{M}^{(*)}$, and (16), (17), and (19) will be denoted by $\mathscr{M}^{(**)}$. In both systems, the source vertex function x_1 merely provides a sufficiently strong signal $[x_1(t-\tau) - \Gamma_1]^+$ to drive the associational strengths $z_{1i}(t)$ and the stimulus traces $x_i(t)$ towards the values imposed on the border B_n during learning trials. The following systems will be seen to include both $\mathscr{M}^{(*)}$ and $\mathscr{M}^{(**)}$ as special cases. Let

$$\dot{x}_i(t) = -a(t)x_i(t) + b(t)z_{1i}(t) + I_i(t) \tag{20}$$

and

$$\dot{z}_{1i}(t) = -c(t)z_{1i}(t) + d(t)x_i(t), \tag{21}$$

$i = 2, \ldots, n$, where the coefficients a, b, c, and d are bounded, nonnegative, and continuous functions that satisfy the following conditions. There exist positive constants K_1, K_2, and T_1, and continuous functions $U(t)$ and $V(t)$, with $V(t)$ nonnegative and bounded, such that for all $t \geq 0$:

a) $$\int_0^\infty a(v)\,dv = \infty; \tag{22}$$

b) $$\int_0^\infty c(v)\,dv = \infty; \tag{23}$$

c) $$\dot{N}(t) \leq U(t)N(t) + V(t), \tag{24}$$

where

$$N(t) = \sqrt{(x^{(1)}(t))^2 + (z^{(1)}(t))^2},$$

with $x^{(1)} = \sum_{i=2}^{n} x_i$ and $z^{(1)} = \sum_{i=2}^{n} z_{1i}$,

$$\int_T^t \exp\left[\int_v^t U(\xi)\,d\xi\right] dv \leq K_1, \tag{25}$$

for all $T \geq 0$ and $t \geq T$, and

$$\int_0^\infty U(v)\,dv = -\infty; \tag{26}$$

d) $$\int_0^t d(v) \exp\left[-\int_v^t c(\xi)\,d\xi\right] dv \geq K_2 \tag{27}$$

for $t \geq T_1$.

All of the conditions (a)–(d) can be imposed instead for all t larger than a prescribed, but otherwise arbitrary, positive constant. Any system satisfying (20)–(27) will be denoted generically by $\tilde{\mathscr{M}}$. It will henceforth be assumed that coefficients a, b, c, and d are fixed once and for all throughout the discussion of Theorem 1 below.

We will first consider the probability distributions $X = \{X_i : i = 2, \ldots, n\}$ and $y = \{y_{1i} : i = 2, \ldots, n\}$ defined by

$$X_i = x_i\left[\sum_{m=2}^{n} x_m\right]^{-1}$$

and

$$y_{1i} = z_{1i}\left[\sum_{m=2}^{n} z_{1m}\right]^{-1},$$

which measure relative stimulus traces and associational strengths through time. We will show that these probabilities can learn the weights θ_i of any nontrivial spatial pattern θ, and that for suitable parameter and input choices the outstars $\mathscr{M}^{(*)}$ and $\mathscr{M}^{(**)}$ are of type $\tilde{\mathscr{M}}$. We therefore seek an analog of Theorem 2 in [13], which discusses learning of a nontrivial spatial pattern θ by a sequence $G^{(1)}$, $G^{(2)}, \ldots, G^{(N)}, \ldots$ of Γ-outstars. Each $G^{(N)}$ differs from $G^{(N-1)}$ only by being subjected to a longer presentation of θ on B_n (i.e., "more practice") and possibly a different input to its source vertex after the $I_i(t)$ input to B_n ceases (i.e., a different "recall trial"). Correspondingly, the following theorem will discuss a sequence

$\tilde{\mathcal{M}}^{(1)}, \ldots, \tilde{\mathcal{M}}^{(N)}, \ldots$ of systems constructed from $\tilde{\mathcal{M}}$, with identical initial data, such that $\tilde{\mathcal{M}}^{(N)}$ satisfies the equations

$$\dot{x}_i^{(N)}(t) = -a^{(N)}(t)x_i^{(N)}(t) + b^{(N)}(t)z_{1i}^{(N)}(t) + I_i^{(N)}(t) \tag{20'}$$

and

$$\dot{z}_{1i}^{(N)}(t) = -c^{(N)}(t)z_{1i}^{(N)}(t) + d^{(N)}(t)x_i^{(N)}(t), \tag{21'}$$

$i = 2, \ldots, n$, where $a^{(N)}, b^{(N)}, c^{(N)}$, and $d^{(N)}$ are constrained as follows. There exists a nonnegative, strictly increasing function $V(N)$ of $N \geq 1$ such that

1) for all $N \geq 1$ and $t \in [0, V(N)]$ (the "practice interval"), $a^{(N)}(t) = a(t)$, $b^{(N)}(t) = b(t)$, $c^{(N)}(t) = c(t)$, and $d^{(N)}(t) = d(t)$;
2) for all $N \geq 1$ and $t > V(N)$, $a^{(N)}(t)$, $b^{(N)}(t)$, $c^{(N)}(t)$, and $d^{(N)}(t)$ satisfy condition (*c*) above.

Motivated by condition (1), we call $\tilde{\mathcal{M}}^{(N)}$ an *N-truncation* of $\tilde{\mathcal{M}}$, and the sequence $\tilde{\mathcal{M}}^{(1)}, \ldots, \tilde{\mathcal{M}}^{(N)}, \ldots$ of N-truncations is said to be *derived* from $\tilde{\mathcal{M}}$. To emphasize that $\tilde{\mathcal{M}}$ is untruncated, one can write $\mathcal{M}^{(\infty)}$ instead of $\tilde{\mathcal{M}}$. The following theorem will discuss the probabilities

$$X_i^{(N)} = x_i^{(N)}\left[\sum_{m=2}^{n} x_m^{(N)}\right]^{-1}$$

and

$$y_{1i}^{(N)} = z_{1i}^{(N)}\left[\sum_{m=2}^{n} z_{1m}^{(N)}\right]^{-1}$$

of each $\tilde{\mathcal{M}}^{(N)}$. Superscripts "$(N)$" will be omitted where the untruncated system $\tilde{\mathcal{M}}$ is being discussed. Since Theorem 2 [13] shows how to derive results for truncated outstars from untruncated outstars, all estimates below will be aimed at the untruncated case. To avoid trivialities, the sums $x^{(1)}(t) = \sum_{m=2}^{n} x_m(t)$ and $z^{(1)}(t) = \sum_{m=2}^{n} z_{1m}(t)$ will be taken positive at $t = 0$.

THEOREM 1. *Let $\tilde{\mathcal{M}}^{(1)}, \ldots, \tilde{\mathcal{M}}^{(N)}, \ldots$ be any sequence of N-truncations of any $\tilde{\mathcal{M}}^{(\infty)}$ having arbitrary initial data. Let the inputs $I_i^{(N)}$ of $\tilde{\mathcal{M}}^{(N)}$ have the form*

$$I_i^{(N)}(t) = \theta_i I(t)\chi(t - U(N)), \tag{28}$$

$i = 2, \ldots, n$, where

a) *$\theta = \{\theta_i : i = 2, \ldots, n\}$ is a fixed arbitrary probability distribution,*
b) *there exist positive constants T_2 and K_3 such that for every $T \geq 0$,*

$$\int_T^{t+T} I(v)\exp\left[-\int_v^{t+T} a(\xi)\,d\xi\right]dv \geq K_3 \tag{29}$$

for $t \geq T_2$;

c) *$U(N)$ is a nonnegative and strictly increasing function of $N \geq 1$;*

and

d) $$\chi(t) = \begin{cases} 1 & t < 0 \\ 0 & t \geq 0. \end{cases}$$

Then

A) *for every $N \geq 1$, the limits*

$$Q_i^{(N)} = \lim_{t\to\infty} X_i^{(N)}(t)$$

and

$$P_{1i}^{(N)} = \lim_{t\to\infty} y_{1i}^{(N)}(t)$$

exist;

B) *for every $N \geq 1$ and $t \geq U(N)$, the functions $X_i^{(N)}(t)$ and $y_{1i}^{(N)}(t)$ are monotonic in opposite senses with $|y_{1i}^{(N)}(t) - X_i^{(N)}(t)|$ nonincreasing, and are constant on intervals for which*

$$b^{(N)}(t) = d^{(N)}(t) = 0;$$

C) $$\lim_{N\to\infty} m_i^{(N)} = \lim_{N\to\infty} M_i^{(N)} = \theta_i,$$

where

$$m_i^{(N)} = \min(X_i^{(N)}(U(N)), \quad y_{1i}^{(N)}(U(N)))$$

and

$$M_i^{(N)} = \max(X_i^{(N)}(U(N)), \quad y_{1i}^{(N)}(U(N))).$$

By (A)–(C),

$$\lim_{N\to\infty} \lim_{t\to\infty} X_i^{(N)}(t) = \lim_{N\to\infty} \lim_{t\to\infty} y_{1i}^{(N)}(t) = \theta_i. \tag{30}$$

D) *The functions $\dot{y}_{1i}^{(N)}, f_i^{(N)} = y_{1i}^{(N)} - X_i^{(N)}$, and $g_i^{(N)} = X_i^{(N)} - \theta_i$ change sign at most once and not at all if $f_i^{(N)}(0)g_i^{(N)}(0) \geq 0$. Moreover, $f_i^{(N)}(0)g_i^{(N)}(0) > 0$ implies $f_i^{(N)}(t)g_i^{(N)}(t) > 0$ for all $t \geq 0$.*

The proof will follow that of Theorem 2 [13] as closely as possible. The first step is the following change of variables.

Lemma 1. X_i and y_{1i} *satisfy the equations*

$$\dot{X}_i = A_1(y_{1i} - X_i) + B_1(\theta_i - X_i) \tag{31}$$

and

$$\dot{y}_{1i} = C_1(X_i - y_{1i}), \tag{32}$$

where

$$A_1 = \frac{bz^{(1)}}{x^{(1)}}, \tag{33}$$

$$B_1 = \frac{I}{x^{(1)}}, \tag{34}$$

and

$$C_1 = \frac{dx^{(1)}}{z^{(1)}}. \tag{35}$$

Proof: Summing (20) and (21) over $i \neq 1$ yields

$$\dot{x}^{(1)} = -ax^{(1)} + bz^{(1)} + I \tag{36}$$

and

$$\dot{z}^{(1)} = -cz^{(1)} + dx^{(1)}. \tag{37}$$

Since $X_i = x_i/x^{(1)}$,

$$\dot{X}_i = \frac{1}{x^{(1)}}\left[\dot{x}_i - x_i\frac{\dot{x}^{(1)}}{x^{(1)}}\right],$$

which by (20) and (36) yields (31). Since $y_{1i} = z_{1i}/z^{(1)}$,

$$\dot{y}_{1i} = \frac{1}{z^{(1)}}\left[\dot{z}_{1i} - z_{1i}\frac{\dot{z}^{(1)}}{z^{(1)}}\right],$$

which by (21) and (37) yields (32).

Equations (31) and (32) have the same form as equations (54) and (55) in [13] for the probabilities X_i and y_{1i} of (11)–(14). Only the coefficients A_1, B_1, and C_1 differ. To prove Theorem 1, it therefore suffices to check that the estimates on these coefficients used to prove Theorem 2 [13] hold in the present situation.

Clearly A_1, B_1, and C_1 are nonnegative and continuous. Hence the following lemma concerning memory and recall trials holds.

LEMMA. 2. *Let $I(t) = 0$ for $t \geq t_0$. Then for $t \geq t_0$,*

a) *$X_i(t)$ and $y_{1i}(t)$ are constant in intervals for which $b(t) = d(t) = 0$,*
b) *$X_i(t)$ and $y_{1i}(t)$ are monotonic in opposite senses with $|y_{1i}(t) - X_i(t)|$ monotone non-increasing, and thus*
c) *the limits $Q_i = \lim_{t\to\infty} X_i(t)$ and $P_{1u} = \lim_{t\to\infty} y_{1i}(t)$ exist and lie in $[m_i(t_0), M_i(t_0)]$, where*

$$m_i(t_0) = \min(X_i(t_0), \quad y_{1i}(t_0))$$

and

$$M_i(t_0) = \max(X_i(t_0), \quad y_{1i}(t_0)).$$

Proof: By (31),

$$\dot{X}_i = A_1(y_{1i} - X_i)$$

which along with (32) and nonnegativity of A_1 and C_1 readily yields the proof. See Theorem 2 [13].

The next lemma studies the oscillations of X_i and y_{1i} relative to θ_i.

LEMMA 3. *The functions $\dot{y}_{1i}$, $f_i = y_{1i} - X_i$, and $g_i = X_i - \theta_i$ change sign at most once for $t \geq t_0$, and not at all if $f_i(t_0)g_i(t_0) \geq 0$. Moreover, $f_i(t_0)g_i(t_0) > 0$ implies $f_i(t)g_i(t) > 0$ for $t \geq t_0$. In particular P_{1i} exists.*

Proof: By (31) and (32)

$$\dot{f}_i = -D_1 f_i + B_1 g_i \tag{38}$$

and

$$\dot{g}_i = -B_1 g_i + A_1 f_i, \tag{39}$$

where $D_1 = A_1 + C_1$. Since A_1 and B_1 are nonnegative, the lemma follows from Lemma 3 of [13].

By Lemma 3, two cases arise: either

(A) $f_i(t)g_i(t) < 0$ for all large t, or

(B) $f_i(t)g_i(t) \geq 0$ for all large t.

Suppose (A) holds. Then by (38) and (39), f_i and g_i are monotonic for all large t, hence have limits as $t \to \infty$, and thus the limits Q_i and P_{1i} exist. It remains in Case (A) to show that the existence of Q_i and P_{1i} implies $Q_i = P_{1i}$, which in turn implies $Q_i = \theta_i$. In Case (B), we must in addition show that Q_i exists. The following estimates on $x^{(1)}$ and $z^{(1)}$ are needed to establish these facts.

LEMMA 4. *Let (29) hold. Then $x^{(1)}$ and $z^{(1)}$ are bounded from above and below by positive constants.*

Proof: The existence of positive lower bounds is proved as follows. By (36),

$$\dot{x}^{(1)} \geq -ax^{(1)} + I,$$

and thus

$$x^{(1)}(t) \geq \int_0^t I(v) \exp\left[-\int_v^t a(\xi)\,d\xi\right] dv,$$

which by (29) yields

$$x^{(1)}(t) \geq K_3,\ t \geq T_2. \tag{40}$$

(40) along with the positivity of $x^{(1)}(t)$ for $t \in [0, T_2]$ proves the existence of a positive lower bound K_4 for $x^{(1)}(t)$ for $t \geq 0$.

By (37),

$$\dot{z}^{(1)} \geq -cz^{(1)} + K_4 d,$$

and thus

$$z^{(1)}(t) \geq K_4 \int_0^t d(v) \exp\left[-\int_v^t c(\xi)\,d\xi\right] dv, \tag{41}$$

which by (27) implies

$$z^{(1)}(t) \geq K_4 K_2 > 0, \quad \text{for } t \geq T_1.$$

Hence $z^{(1)}(t)$ has a positive lower bound for $t \geq 0$.

Upper bounds follow from (24)–(26). By (24), for $t \geq 0$,

$$N(t) \leq N(0) \exp\left[\int_0^t U(v)\,dv\right] + \int_0^t V(v) \exp\left[\int_v^t U(\xi)\,d\xi\right] dv.$$

By (25),

$$N(t) \leq N(0) \exp\left[\int_0^t U(v)\,dv\right] + \|V\|_\infty K_1,$$

and by (26), $\exp[\int_0^t U(v)\,dv]$ is bounded. The boundedness of $N(t)$ implies that of $x^{(1)}(t)$ and $z^{(1)}(t)$.

Remark: Any condition that guarantees upper bounds for $x^{(1)}$ and $z^{(1)}$ can be used to replace (24)–(26).

Lemma 4 can be used to show that $Q_i = P_{1i}$ in Case (A).

LEMMA 5. *Suppose Q_i and P_{1i} exist. Then $Q_i = P_{1i}$.*

Proof: Suppose not, and let $Q_i > P_{1i}$ hold for definiteness. Then there exists a T_3 such that

$$X_i(t) - y_{1i}(t) \geq \tfrac{1}{2}(Q_i - P_{1i}) \tag{42}$$

for $t \geq T_3$. By Lemma 4, there exists a positive constant K_5 such that $x^{(1)}(t)/z^{(1)}(t) \geq K_5$ for $t \geq 0$. Thus by (32) and (42)

$$\dot{y}_{1i}(t) \geq \frac{K_5(Q_i - P_{1i})}{2} d(t) \tag{43}$$

for $t \geq T_3$. Integrating (43) yields

$$1 \geq y_{1i}(t) - y_{1i}(T_3) \geq \frac{K_5(Q_i - P_{1i})}{2} \int_{T_3}^{t} d(v)\, dv$$

for every $t \geq T_3$. Thus

$$\int_0^\infty d(v)\, dv < \infty. \tag{44}$$

A contradiction will now be drawn by showing that (27) implies

$$\int_0^\infty d(v)\, dv = \infty \tag{45}$$

by using a variant of Gronwall's Lemma ([15], p. 31). For any positive $K_6 > 0$ and $t \geq T_1$, (27) implies

$$K_2 E(t) \leq K_6 + \int_0^t d(v)E(v)\, dv \tag{46}$$

where

$$E(t) = \exp\left[\int_0^t c(v)\, dv\right].$$

Thus for $t \geq T_1$,

$$K_2 \frac{d}{dt} \log\left(K_6 + \int_0^t d(v)E(v)\, dv\right) \leq d(t),$$

and

$$K_6 + \int_0^t d(v)E(v)\, dv \leq K_7 \exp\left[K_2^{-1} \int_{T_1}^{t} d(v)\, dv\right]$$

where

$$K_7 = K_6 + \int_0^{T_1} d(v)E(v)\, dv.$$

By (46), we therefore find

$$K_2\left(\log K_2K_7^{-1} + \int_0^t c(v)\,dv\right) \le \int_{T_1}^t d(v)\,dv$$

which, by (23), implies (45). A similar contradiction holds if $Q_i < P_{1i}$, thereby proving $Q_i = P_{1i}$.

The identities $Q_i = P_{1i} = \theta_i$ for Case A are now established as follows.

LEMMA 6. *Suppose Case A holds. Then* $Q_i = \theta_i$.

Proof: For definiteness, let f_i be positive for large t. It can in fact be assumed without loss of generality that f_i is positive for all t. Then by (31),

$$\dot{X}_i \ge B_1(\theta_i - X_i)$$

with X_i monotone increasing to the limit Q_i. Supposing that $\theta_i > Q_i$, we will deduce a contradiction.

Denoting the finite upper bound of $x^{(1)}(t)$ by K_8^{-1}, then

$$\dot{X}_i \ge K_8(\theta_i - Q_i)I,$$

or

$$X_i(t) \ge X_i(0) + K_8(\theta_i - Q_i)\int_0^t I(v)\,dv,$$

for all $t \ge 0$. Since $1 \ge X_i(t), \theta_i > Q_i$ implies

$$\int_0^\infty I(v)\,dv < \infty.$$

By (29), however, for any $K_9 > 0$ and $t \ge T_2$,

$$K_3F(t) \le K_9 + \int_0^t I(v)F(v)\,dv,$$

where

$$F(t) = \exp\left[\int_0^t a(v)\,dv\right].$$

Arguing as in Lemma 5, using (22), yields the contradiction

$$\int_0^\infty I(v)\,dv = \infty,$$

and thus $Q_i = \theta_i$.

It remains only to treat Case (B). It suffices to consider the subcase of Case (B) for which $y_{1i} \ge X_i \ge \theta_i$ and y_{1i} is monotone decreasing, since the case $\theta_i \le X_i \le y_{1i}$ can be similarly treated. We can assume without loss of generality that $y_{1i}(0) \ge X_i(0) \ge \theta_i$ to simplify our formulas.

LEMMA 7. *Suppose* $y_{1i}(0) \ge X_i(0) \ge \theta_i$. *Then there exists a* $\mu \in (0, 1)$ *such that for* $t \ge T_2$,

$$X_i(t) - \theta_i \le (1 - \mu)(y_{1i}(t - T_2) - \theta_i).$$

Proof: Let $X_i^{(\theta)} = X_i - \theta_i$ and $y_i^{(\theta)} = y_{1i} - \theta_i$. Then (31) becomes

$$\dot{X}_i^{(\theta)} = -(A_1 + B_1)X_i^{(\theta)} + A_1 y_i^{(\theta)},$$

which can be integrated from any $T \geq 0$ to $t \geq T$, yielding

$$X_i^{(\theta)}(t) = U_i(t, T) + V_i(t, T), \tag{47}$$

where

$$U_i(t, T) = X_i^{(\theta)}(T)Z^{-1}(t, T), \tag{48}$$

$$V_i(t, T) = \int_T^t A_1(v)y_i^{(\theta)}(v)Z^{-1}(t, v)\,dv, \tag{49}$$

$$Z(t, T) = \exp\left[\int_T^t (A_1 + B_1)\,dv\right]. \tag{50}$$

By (33), (34), and (36)

$$A_1 + B_1 = \frac{bz^{(1)} + I}{x^{(1)}}$$

$$= \frac{\dot{x}^{(1)}}{x^{(1)}} + a,$$

and thus

$$Z(t, T) = \frac{x^{(1)}(t)F(t)}{x^{(1)}(T)F(T)}, \tag{51}$$

where as above

$$F(t) = \exp\left[\int_0^t a(v)\,dv\right].$$

By (51) and the monotone decrease of $y_i^{(\theta)}$, (49) implies

$$V_i(t, T) \leq y_i^{(\theta)}(T)R(t, T), \tag{52}$$

where

$$R(t, T) = \frac{1}{x^{(1)}(t)F(t)}\int_T^t A_1(v)x^{(1)}(v)F(v)\,dv.$$

By (33) and (36),

$$A_1(v)x^{(1)}(v)F(v) = b(v)z^{(1)}(v)F(v)$$

$$= \frac{d}{dv}[x^{(1)}(v)F(v)] - I(v)F(v),$$

and thus

$$R(t, T) = 1 - Z^{-1}(t, T) - \frac{1}{x^{(1)}(t)F(t)}\int_T^t I(v)F(v)\,dv. \tag{53}$$

Since also by (48),

$$U_i(t, T) \leq y_i^{(\theta)}(T)Z^{-1}(t, T), \tag{54}$$

(47), (52), (53), and (54) imply that

$$X_i^{(\theta)}(t) \le y_i^{(\theta)}(T)P(t, T), \tag{55}$$

where

$$P(t, T) = 1 - \frac{1}{x^{(1)}(t)F(t)} \int_T^t I(v)F(v)\, dv.$$

Since $x^{(1)}(t) \le K_8^{-1}$,

$$P(t, T) \le 1 - K_8 F^{-1}(t) \int_T^t I(v)F(v)\, dv.$$

By (29), for all $t \ge T + T_2$,

$$P(t, T) \le 1 - \mu, \tag{56}$$

with $\mu = K_3 K_8$. (55) with (56) complete the proof. Lemma 7 implies the following lemma.

LEMMA 8. *If* $y_i(0) \ge X_i(0) \ge \theta_i$, *then* $P_{1i} = Q_i = \theta_i$.

Proof: P_{1i} exists in this case, since $y_{1i}(t)$ is monotone decreasing. Suppose $P_{1i} > \theta_i$. Then by Lemma 7, for t sufficiently large,

$$\begin{aligned} y_{1i}(t) - X_i(t) &= y_{1i}(t) - y_{1i}(t - T_2) + y_{1i}^{(\theta)}(t - T_2) - X_i^{(\theta)}(t) \\ &\ge \mu y_i^{(\theta)}(t - T_2) + (y_{1i}(t) - y_{1i}(t - T_2)) \\ &\ge \mu(P_{1i} - \theta_i) - \frac{\mu}{2}(P_{1i} - \theta_i) \\ &= \frac{\mu}{2}(P_{1i} - \theta_i) > 0, \end{aligned}$$

which by the argument of Lemma 5 yields a contradiction. Thus $P_{1i} = \theta_i$, and hence Q_i exists and equals θ_i.

We have hereby shown that $\theta_i = Q_i = P_{1i}$ in all cases. The application of these results to spatial patterns truncated at finite times $U(N)$ now follows just as in Theorem 2 [12]. The proof of Theorem 1 is therefore complete.

Theorem 1 shows that, given any nontrivial spatial pattern θ, and t sufficiently large, $y_{1i}(t) \cong \theta_i$ or $z_{1i}(t) \cong \theta_i z^{(1)}(t)$. Since $z^{(1)}(t)$ is bounded from above, $z_{1i}(t) \cong 0$ if $\theta_i = 0$. Since $z^{(1)}(t)$ is bounded from below by a positive constant (which increases in applications as input intensity and duration increase), the absolute size of $z_{1i}(t)$ is bounded away from zero if $\theta_i > 0$, and will therefore be able to reproduce the pattern weights θ_i on the grid, given a *bounded* signal from the source vertex v_1 with an intensity that is greater than sufficiently small amounts of noise in the grid.

Theorem 1 does not consider the case of learning the trivial spatial pattern, characterized by $I(t) = V(t) \equiv 0$ in the applications below.

PROPOSITION 1. *Suppose* $I(t) = 0$ *for* $t \ge T$ *in* $\tilde{\mathcal{M}}$. *Then given arbitrary initial data for* $t \le T$, $x_i(t)$ *and* $z_{1i}(t)$ *converge to zero as* $t \to \infty$, $i = 2, \ldots, n$, *in such a way that* X_i *and* y_{1i} *do not oscillate, and converge exponentially to zero if*

$$\int_0^t U(v)\, dv \le K_{10} - K_{11} t \tag{57}$$

with K_{11} *positive.*

Proof: By (24),

$$0 \leq N(t) \leq N(T) \exp\left[\int_T^t U(v)\,dv\right]. \tag{58}$$

By (25), $\lim_{t\to\infty} N(t) = 0$, and by nonnegativity of each x_i and z_{1i}, each x_i and z_{1i} has zero limit as $t \to \infty$. The statement concerning oscillations is proved in Lemma 2. The statement concerning exponential decay is obvious from (57) and (58).

6. Conditions which imply boundedness

The following corollaries describe special conditions that imply (24)–(26), and which will be used to study $\mathscr{M}^{(*)}$ and $\mathscr{M}^{(**)}$.

COROLLARY 1. (24) *holds if* $U(t)$ *and* $V(t)$ *are defined by*

$$U = \tfrac{1}{2}\{-a - c + \sqrt{(a + c)^2 - [4ac - (b + d)^2]}\} \tag{59}$$

and $V = I$.

Proof: By Lemma 4.2, p. 56, of [16], (24) holds with $V = I$ if also $U(t)$ is the greatest eigenvalue of the Hermitian part

$$A^H(t) = \begin{pmatrix} -a(t) & \tfrac{1}{2}(b(t) + d(t)) \\ \tfrac{1}{2}(b(t) + d(t)) & -c(t) \end{pmatrix} \tag{60}$$

of

$$A(t) = \begin{pmatrix} -a(t) & b(t) \\ d(t) & -c(t) \end{pmatrix}, \tag{61}$$

which is given by (59).

COROLLARY 2. *Theorem* 1 *holds if* (24)–(26) *are replaced by the conditions*

$$a + c \geq \varepsilon \tag{62}$$

and

$$4ac - (b + d)^2 \geq \eta \tag{63}$$

for some positive ε *and* η.

Proof: By (62) and (63), $U(t)$ in (59) is negative and bounded away from zero, say by $U_0 < 0$. (26) follows immediately. (25) follows by the simple estimates

$$\begin{aligned}\int_T^t \exp\left[\int_v^t U(\xi)\,d\xi\right] dv &= \exp\left[\int_0^t U(v)\,dv\right]\int_T^t |U(v)|^{-1}\frac{d}{dv}\exp\left[-\int_0^v U(\xi)\,d\xi\right] dv \\ &\leq |U_0|^{-1}\left(1 - \exp\left[-\int_T^t |U(v)|\,dv\right]\right) \\ &\leq |U_0|^{-1}.\end{aligned}$$

COROLLARY 3. *Theorem* 1 *holds if* (24)–(26) *are replaced by the following conditions.*

i) $4ac - (b + d)^2 \geq 0;$ (64)

ii) $b + d \to 0$ iff $4ac - (b + d)^2 \to 0;$ (65)

iii) $a - c \geq \varepsilon$, for some $\varepsilon > 0$, (66)

iv) there exist positive functions $\mu(\delta)$ and $T(\delta)$ of $\delta \geq \|a - c\|_\infty$ such that

$$\int_0^t e^{-\delta(t-v)} b(v)\, dv \geq \mu(\delta), \qquad \text{for } t \geq T(\delta). \tag{67}$$

Proof: The existence of positive constants K_{12} and K_{13} such that

$$K_{12} z^{(1)}(t) \leq x^{(1)}(t) \leq K_{13} z^{(1)}(t) \tag{68}$$

will be proved, and then used to refine Lemma 4.2, p. 56, of [16]. Let $f = x^{(1)}/z^{(1)}$. Then

$$\dot{f} = \frac{1}{z^{(1)}}\left[\dot{x}^{(1)} - x^{(1)}\frac{\dot{z}^{(1)}}{z^{(1)}}\right],$$

from which (36) and (37) imply

$$\dot{f} = (c - a)f - df^2 + b + I/z^{(1)}.$$

The proof in Lemma 4 that $z^{(1)}$ has a positive lower bound λ^{-1} does not use (24)–(26). Thus

$$-(a - c + df)f + b \leq \dot{f} \leq -(a - c)f + b + \lambda I. \tag{69}$$

By (66) and (69),

$$\dot{f} \leq -\varepsilon f + g,$$

where $g \equiv b + \lambda I$ is bounded. Thus for all $t \geq 0$,

$$f(t) \leq K_{13} \equiv f(0) + \varepsilon^{-1}\|g\|_\infty. \tag{70}$$

By (70),

$$a - c + df \leq \delta \equiv \|a - c\|_\infty + K_{13}\|d\|_\infty$$

which by (69) implies

$$\dot{f} \geq -\delta f + b,$$

and thus

$$f(t) \geq \int_0^t e^{-\delta(t-v)} b(v)\, dv.$$

Since $\delta > \|a - c\|_\infty$, (67) implies that

$$f(t) \geq \mu(\delta) > 0, \quad t \geq T(\delta).$$

$f(t)$ is also positive for $t \in [0, T(\delta)]$. Letting

$$K_{12} = \min\{\mu(\delta), f(t) : t \in [0, T(\delta)]\}$$

completes the proof of (68).

We now refine Lemma 4.2 of [16] using (68). By definition, $N(t) = \|y(t)\| \equiv \sqrt{y(t) \cdot y(t)}$, where $y(t) = \begin{pmatrix} x^{(1)}(t) \\ z^{(1)}(t) \end{pmatrix}$.

Since $\|y(t)\| > 0$

$$\dot{N}(t) = \frac{\frac{1}{2}(d/dt)(\|y(t)\|)^2}{\|y(t)\|},$$

and since $x^{(1)}$ and $z^{(1)}$ are real-valued,

$$\dot{N}(t) = \frac{\dot{y}(t) \cdot y(t)}{\|y(t)\|},$$

where

$$\dot{y}(t) = A(t)y(t) + B(t),$$

with $A(t)$ given by (61), and $B(t) = \begin{pmatrix} I(t) \\ 0 \end{pmatrix}$. Thus

$$\dot{N}(t) \le [A(t)w(t) \cdot w(t)]N(t) + I(t), \tag{71}$$

where $w(t) = y(t)/\|y(t)\|$, and in particular $\|w(t)\| = 1$. Lemma 4.2 is proved by noting that $U(t)$ as defined by (59) equals $\sup\{A(t)w \cdot w : \|w\| = 1\}$. A better upper bound for (71) can be found in the present case, since by (68), there exists an $\eta \in (0, 1)$ such that $w_i(t) \ge \eta$, $i = 1, 2$, where $w(t) = \begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix}$. Thus by (71),

$$\dot{N}(t) \le U(t)N(t) + I(t),$$

with

$$U(t) = \sup\{A(t)w \cdot w : \|w\| = 1, w_i \ge \eta, i = 1, 2\}. \tag{72}$$

We now show that this $U(t)$ is negative and bounded away from zero. The proof can then be completed as in Corollary 2.

Transform $A(t)w \cdot w$ to principal axes ([17], p. 23). Then there exists a vector $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ with $\|v\| = 1$, such that

$$A(t)w \cdot w = U_1(t)v_1^2 + U_2(t)v_2^2, \tag{73}$$

where

$$U_1 = \tfrac{1}{2}\{-a - c + \sqrt{(a + c)^2 - [4ac - (b + d)^2]}\},$$

$$U_2 = \tfrac{1}{2}\{-a - c - \sqrt{(a + c)^2 - [4ac - (b + d)^2]}\},$$

and the matrix $R = \|r_{ij}\|$ such that

$$v = Rw \tag{74}$$

is orthogonal; i.e., $\sum_{k=1}^{2} r_{ki}r_{kj} = \delta_{ij} = \sum_{k=1}^{2} r_{ik}r_{jk}$. Clearly $U_2 \le U_1$, and by (64), $U_1 \le 0$. Since $\|v\| = 1$, it suffices to show that v_1^2 and U_2 are bounded away from zero and that v_2^2 is bounded away from zero if U_1 approaches zero.

We first show that $v_1^2 \geq \eta^2 > 0$. By (74),

$$v_1^2 = r_{11}^2 w_1^2 + r_{12}^2 w_2^2 + 2r_{11}r_{12}w_1w_2. \tag{75}$$

By (73) and (74),

$$b + d = 2(U_1 r_{11} r_{12} + U_2 r_{21} r_{22}).$$

By the orthogonality of R,

$$\begin{aligned} b + d &= 2(U_1 - U_2) r_{11} r_{12} \\ &= 2 r_{11} r_{12} \sqrt{(a - c)^2 + (b + d)^2}. \end{aligned}$$

Since $b + d \geq 0$, $r_{11}r_{12} \geq 0$. Thus (75) implies

$$v_1^2 \geq \eta^2 (r_{11}^2 + r_{12}^2) = \eta^2.$$

To show that U_2 is bounded away from zero, note that (66) implies $a + c \geq \varepsilon$, and thus $U_2 \leq -\varepsilon$.

By (66), U_1 can approach zero iff $4ac - (b + d)^2$ approaches zero, and thus by (65) only if $b + d$ approaches zero. By (73), (74), and the orthogonality of R,

$$\begin{aligned} b + d &= 2(U_2 - U_1) r_{21} r_{22} \\ &= -2 r_{21} r_{22} \sqrt{(a - c)^2 + (b + d)^2}. \end{aligned}$$

By (66),

$$|b + d| \geq 2\varepsilon |r_{21} r_{22}| \geq 0. \tag{76}$$

Thus if U_1 approaches zero $r_{21}r_{22}$ approaches zero. By (74),

$$v_2^2 = r_{21}^2 w_1^2 + r_{22}^2 w_2^2 + 2 r_{21} r_{22} w_1 w_2.$$

Since $1 \geq w_i \geq \eta$, $i = 1, 2$,

$$v_2^2 \geq \eta^2 - 2|r_{21} r_{22}|.$$

Thus $v_2^2 \geq \frac{1}{2}\eta^2$ if $|r_{21}r_{22}| \leq \eta^2/4$, or by (76), if $|b + d| \leq \varepsilon\eta^2/2$. The proof is therefore complete.

7. $\mathcal{M}^{(*)}$ and $\mathcal{M}^{(**)}$ are of type $\tilde{\mathcal{M}}$.

Using Corollaries 1–3, it can be shown that $\mathcal{M}^{(*)}$ and $\mathcal{M}^{(**)}$ satisfy (20)–(27) if the parameters and inputs are suitably chosen. The following choice of inputs suffices for many applications.

DEFINITION. *An* input pulse $J(t)$ *is a nonnegative and continuous function that is positive on a finite open interval.*

Let

$$I_1(t) = \sum_{n=1}^{\infty} J_{1n}(t - t_1(n)) \tag{77}$$

and

$$I(t) = \sum_{n=1}^{\infty} J_n(t - t(n)) \tag{78}$$

where

(1) The sequences $\{t_1(n):n \geq 1\}$ and $\{t(n):n \geq 1\}$ satisfy

$$\varepsilon_1 \leq t_1(n+1) - t_1(n) \leq \varepsilon_2 \tag{79}$$

and

$$\delta_1 \leq t(n+1) - t(n) \leq \delta_2 \tag{80}$$

for some choice of positive constants $\varepsilon_1, \varepsilon_2, \delta_1$, and δ_2; and

(2)

$$M_1(t) \leq J_{1n}(t) \leq L_1(t) \tag{81}$$

and

$$M(t) \leq J_n(t) \leq L(t), \tag{82}$$

where M_1, L_1, M, and L are input pulses. The left hand endpoints of the intervals of positivity of these input pulses are chosen equal to zero, for convenience. Thus, $I_1(t)$ and $I(t)$ dominate and are dominated by a sum of infinitely many iterations of an input pulse with bounded spacing. Letting

$$N(t) = \begin{cases} 0, & t < 0 \\ \displaystyle\int_0^t M_1(v)\, e^{-\alpha_1(t-v)}\, dv, & t \geq 0 \end{cases}$$

we find the following corollaries.

COROLLARY 4. *Theorem 1 holds for* $\mathcal{M}^{(*)}$ *if* I_1 *and* I *satisfy* (77)–(82),

$$\|N\|_\infty > \Gamma_1, \tag{83}$$

and

$$\Gamma_1 + \sqrt{\frac{4\alpha u}{(\beta + v)^2} - \varepsilon_1} \geq \|x_1\|_\infty \tag{84}$$

for some $\varepsilon_1 > 0$.

Proof: In $\mathcal{M}^{(*)}$, $a(t) = \alpha$, $b(t) = \beta[x_1(t-\tau) - \Gamma_1]^+$, $c(t) = u$, and $d(t) = v[x_1(t-\tau) - \Gamma_1]^+$. Thus (22) and (23) hold trivially. To verify (24)–(26), we will use Corollary 2. Clearly $a + c \geq \varepsilon$ if $\varepsilon = \alpha + u$.

Since

$$\begin{aligned} 4ac - (b+d)^2 &= 4\alpha u - (\beta + v)^2\{[x_1(t-\tau) - \Gamma_1]^+\}^2 \\ &\geq 4\alpha u - (\beta + v)^2(\|x_1\|_\infty - \Gamma_1)^2, \end{aligned}$$

(84) implies

$$4ac - (b+d)^2 \geq \eta$$

where $\eta = \varepsilon_1(\beta + v)^2$.

Hence (62) and (63), and thus (24)–(26), hold. (27) requires that

$$\int_0^t e^{-u(t-v)}[x_1(v-\tau) - \Gamma_1]^+\, dv \geq K_{14} \tag{85}$$

for some $K_{14} > 0$. By (16) and (81)

$$x_1(v) \geq x_1(0)\, e^{-\alpha_1 v} + \sum_{k=1}^{\infty} N(v - t_1(k)),$$

where by (83),

$$\|N(\cdot - t_1(k))\|_\infty > \Gamma_1$$

for every $k = 1, 2, \ldots$. Thus (85) merely requires that an exponentially weighted sum of a positive input pulse that is iterated in time with bounded spacing has a positive lower bound. This is obvious. (See, for example, Corollary 3 of [7].) (29) also holds for the same reason.

COROLLARY 5. *Theorem* 1 *holds for* $\mathcal{M}^{(**)}$ *if* I_1 *and* I *satisfy* (77)–(82),

$$\|N\|_\infty > \Gamma_1, \tag{83}$$

and

$$\Gamma_1 + \min\left(\frac{\alpha}{u}, \frac{4\alpha u}{(\beta + v)^2}\right) - \varepsilon_2 \geq \|x_1\|_\infty, \tag{86}$$

for some $\varepsilon_2 > 0$.

Proof: In $\mathcal{M}^{(**)}$, $a(t) = \alpha$, $b(t) = \beta[x_1(t - \tau) - \Gamma_1]^+$, $c(t) = u[x_1(t - \tau) - \Gamma_1]^+$, and $d(t) = v[x_1(t - \tau) - \Gamma_1]^+$. Thus (22) holds trivially. (23) holds because

$$\int_0^\infty c(v)\, dv \geq u \int_0^\infty \left[\sum_{k=1}^{\infty} N(v - \tau - t_1(k)) - \Gamma_1\right]^+ dv,$$

where by (83), each function $N(\cdot - \tau - t_1(k)) - \Gamma_1$ becomes positive in an open interval, $k = 1, 2, \ldots$.

(24)–(26) are proved using Corollary 3. Since

$$4ac - (b + d)^2 = \{4\alpha u - (\beta + v)^2[x_1(t - \tau) - \Gamma_1]^+\}[x_1(t - \tau) - \Gamma_1]^+,$$

where (86) implies

$$4au - (\beta + v)^2[x_1(t - \tau) - \Gamma_1]^+ \geq \varepsilon_2(\beta + v)^2[x_1(t - \tau) - \Gamma_1]^+,$$

$4ac - (b + d)^2$ is nonnegative. Since

$$b + d = (\beta + v)[x_1(t - \tau) - \Gamma_1]^+,$$

$4ac - (b + d)^2 \to 0$ iff $b + d \to 0$. By (86),

$$a - c = u\left(\frac{\alpha}{u} - [x_1(t - \tau) - \Gamma_1]^+\right) \geq \varepsilon,$$

where $\varepsilon = u\varepsilon_2$. It remains only to prove (67), or that

$$\int_0^t e^{-\delta(t-v)}[x_1(v - \tau) - \Gamma_1]^+\, dv \geq \mu(\delta) > 0$$

for $t \geq T(\delta)$. This is true for every $\delta > 0$ and some $\mu(\delta) > 0$ since $[x_1(v - \tau) - \Gamma_1]^+$ dominates the iteration in time with bounded spacing of a positive input pulse. This completes the proof of (24)–(26). (27) follows just as in Corollary 4.

The term $[x_1(t - \tau) - \Gamma_1]^+$ in (17), (18), and (19) can, in principle, be replaced by any function $g(t) = f(x_1(t - \tau))$ which is continuous, nonnegative, bounded and monotone increasing in $x_1(t-\tau)$. It is also convenient, for purposes of space-time pattern learning, that $g(t) = 0$ whenever $x_1(t-\tau) < \Gamma$ for some $\Gamma > 0$. The following alternatives are, for example, possible.

$$\log(1 + [x_1(t-\tau) - \Gamma_1]^+),$$

$$[\log(1 + x_1(t-\tau)) - \Gamma_1]^+,$$

$$[(1 + x_1(t-\tau))^\omega - \Gamma_1]^+, \quad \omega > 1, \quad \Gamma_1 > 1,$$

$$[x_1(t-\tau) - \Gamma_1]^+\{1 + [x_1(t-\tau) - \Gamma_1]^+\}^\omega, \quad \omega > 1.$$

These alternatives require trivial modifications of (84) and (86) to ensure boundedness of $x^{(1)}$ and $z^{(1)}$.

Any conditions such as (84) and (86) can be guaranteed for an arbitrarily large input I_1 to v_1 by modifying (16) slightly. For example, let

$$\dot{x}_1(t) = -\alpha_1 x_1(t) + (\Omega_1 - x_1(t))I_1(t), \tag{87}$$

with $0 \leq x_1(0) \leq \Omega_1$. Then $0 \leq x_1(t) \leq \Omega_1$ for $t \geq 0$ given any nonnegative input I_1, bounded or not. The response of $x_1(t)$ to $I_1(t)$ if $x_1(t) \cong 0$ is approximately linear, and $x_1(t)$ "saturates" at Ω_1 if very large and prolonged input pulses arrive. Replacing $\|x_1\|_\infty$ by Ω_1 in (84) and (86) then fulfills the boundedness conditions for any input I_1.

8. Forgetting, extinction, spontaneous recovery, and post-tetanic potentiation

The above corollaries show that $\mathcal{M}^{(*)}$ and $\mathcal{M}^{(**)}$ are of type $\tilde{\mathcal{M}}$ in cases of practical interest. The main constraints on parameters, such as (84) and (86), aim at keeping $x^{(1)}$ and $z^{(1)}$ bounded, as physical intuition—as well as the method of proof in Theorem 1—require. The constraint on the decay rate u of associational strengths z_{1i} is of particular interest. It is important that small values of u be permissible to allow a slow decay of memory. In both (84) and (86), this can be achieved, for *any* choice of input $I_1(t)$ satisfying (77), (79), and (81) by choosing the decay rate α of the grid vertex functions x_i sufficiently large. A large choice of α is physically desirable to allow old perturbations of the grid to decay rapidly and thereby prepare the grid to receive new perturbations without bias. Speaking heuristically, Corollaries 4 and 5 show that $\mathcal{M}^{(*)}$ and $\mathcal{M}^{(**)}$ can learn any spatial pattern with an arbitrarily small rate of memory decay if the response of the grid to perturbations is sufficiently rapid.

The exponential decay rate u for associational strengths $z_{1i}(t)$ in $\mathcal{M}^{(*)}$ is reminiscent of Ebbinghaus forgetting curves ([18], p. 555). The decay rate of $z_{1i}(t)$ in $\mathcal{M}^{(**)}$ is zero whenever $x_1(t-\tau) \leq \Gamma_1$, and is positive if $x_1(t-\tau) > \Gamma_1$ and no spatial pattern reaches the grid at time t. In other words, $\mathcal{M}^{(**)}$ has a perfect memory that is "extinguished" on "unrewarded" recall trials ([19], p. 727). In both $\mathcal{M}^{(*)}$ and $\mathcal{M}^{(**)}$, a form of "reminiscence" ([18], p. 509) or spontaneous improvement of memory occurs, since if practice ceases at a time $t = T$ for which $f_i(T)g_i(T) \geq 0$, then $y_{1i}(t)$ for $t \geq T$ will continue to approach θ_i.

The interplay between decay of $z_{1i}(t)$ and approximate constancy of $y_{1i}(t)$ during memory or recall intervals helps to understand the phenomenon of "spontaneous recovery" of memory after an interval of forgetting or extinction

([18], p. 733). Since the pattern weights θ_i are recorded in the $y_{1i}(t)$, which are not forgotten, any mechanism that bolsters the absolute size of $z^{(1)}(t)$ will create spontaneous recovery of memory. A sufficiently large and prolonged "presynaptic" signal $x_1(t - \tau) - \Gamma_1$ can, for example, accomplish this. This increase in $z_{1i}(t)$ is analogous to the phenomenon of "post-tetanic potentiation" ([20], p. 98). Alternatively, any mechanism that recreates the spatial pattern θ_i on the grid when $x_1(t - \tau) > \Gamma_1$ will tend to increase $z_{1i}(t)$ as well as to drive $y_{1i}(t)$ closer to θ_i.

9. Γ-Outstar avalanches and stimulus sampling

A space-time pattern $\theta(t)$ with continuous weights $\theta_i(t) = I_i(t)[\sum_{m=1}^n I_m(t)]^{-1}$ can be approximated by a sequence $\{\theta(k\xi):k = 1, 2, \ldots\}$ of spatial patterns, where if ξ is taken sufficiently small, then the approximation becomes arbitrarily good. Since a Γ-outstar $\tilde{\mathscr{M}}_k$ can learn any one spatial pattern $\theta(k\xi)$, we will arrange a sequence $\tilde{\mathscr{M}}_1, \tilde{\mathscr{M}}_2, \ldots, \tilde{\mathscr{M}}_k, \ldots$ of Γ-outstars in such a way that the source vertex $v_{k,1}$ of the kth Γ-outstar $\tilde{\mathscr{M}}_k$ is excited briefly ξ time units after $v_{k-1,1}$ is excited [13]. $\tilde{\mathscr{M}}_k$ will learn from the grid only at times t such that $x_{k,1}(t - \tau) > \Gamma_1$, and in this sense $\tilde{\mathscr{M}}_k$ "samples" the grid at prescribed times during which $\theta(t) \cong \theta(k\xi)$. Such a sequence of Γ-outstars is called a *Γ-outstar avalanche*. We remark in passing that this "sampling" operation is a concrete analog within our networks of the abstract sampling operation of Stimulus Sampling Theory [21]. A network analog of stimulus sampling probabilities are the normalized associational strengths $y_{k,1i}(t)$ from $v_{k,1}$ to the ith grid vertex.

The following examples illustrate several ways of sequentially activating Γ-outstars, or related network components, to sample an arbitrary continuous function $\theta(t)$, bounded by 0 and 1, as a sequence $\{\theta(k\xi):k = 1, 2, \ldots\}$ of spatial patterns. In particular, a single "nerve" with sufficiently many "axon collaterals" activated by "avalanche conduction" will in principle be able to learn an essentially arbitrarily complicated space-time pattern. The only unavoidable limitation of learning accuracy will be some smoothing of $\theta(t)$ in memory and recall, which is, in fact, often desirable for producing smoothly modulated motor performance. A single "nerve" will, however, perform its space-time pattern in a wholly rote or ritualistic way. For example, it is known that perturbing a single nerve in insects can activate significant portions of their feeding, withdrawal, or running reflexes ([22], p. 8). If less ritualistic performance is desired, the learning mechanism should be sensitive to feedback created by its own prior outputs, much as notes previously played by a pianist help to determine the future notes to be played. To accomplish this, one can encode only a portion of a given space-time pattern in any one "nerve". Then these "nerves" must be arranged so that clusters of them are excited in the proper temporal sequence to reproduce the entire space-time pattern, where the next cluster to be excited is partially determined by feedback from the last few clusters to have been excited. In this situation, one must also rapidly switch off all clusters after they have been played out, and all clusters that will compete with ongoing performance of the pattern. Otherwise background noise will accumulate on the grids and interfere with accurate learning and performance. It has been suggested that, *in vivo*, the cerebellum is just such a switching-off, or inhibitory, mechanism of excitatory pattern controls [12].

Experiments have, in fact, recently shown that the output from the cerebellum is inhibitory [23].

Figure 4 schematically describes some Γ-outstar avalanches.

Type A describes a single "nerve" with a long edge, or "correlational axon", and with clusters of "axon collaterals" spaced regularly along the axon. Different collaterals in each cluster terminate at a different grid vertex. This avalanche is said to be *homogeneous* if each grid vertex receives one collateral from every cluster. The case in which different clusters sample nonidentical sets of grid vertices can also be readily studied.

Each axon collateral cluster in Type B perturbs a different grid. The grids are copies of one another, in the sense that the same input I_i is delivered, via axon collaterals, to the ith vertex $v_{k,i}$ of the kth grid G_k. Spreading out the input in this way eliminates the background noise found in Type A due to activation of prior clusters while a given cluster is sampling the grid.

Type C is like Type A with one addition. Instead of sending its own axon collaterals to the grid, the correlational axon in Type C perturbs a Γ-outstar, which in turn perturbs the grid. Type D differs from Type B in a similar fashion.

Type E can accomplish more than Types A–D can, and without as much background noise. Type E is constructed from rows of correlational axons. Each correlational axon sends out an axon collateral at regularly spaced intervals to a single vertex. The pattern to be learned by Type E depends on which vertices in the sets $V_j = \{v_{ji}: i \geq 1\}$ receive inputs from common sources; i.e., which vertices distributed perpendicular to the direction of flow in correlational axons receive inputs with the same phase relations. Speaking rigorously, let $S_j(m, R)$, $m = 1, 2, \ldots, M_j(R)$, be the maximal subsets of indices such that

(i) if $j = 0$:

$$I_{0i}(t) \cong \theta_0^{(m)} \sum_{k \in S_0(m,R)} I_{0k}(t)$$

for *all* $i \in S_0(m, R)$, *all* $t \in R$, and suitable nonnegative constants $\theta_0^{(m)}$;

(ii) if $j \geq 1$:

$$I_{ji}(t) \cong \theta_{ji}^{(m)} \sum_{k \in S_j(m,R)} I_{jk}(t)$$

for *all* $i \in S_j(m, R)$ and *all* $t \in R$, where the $\theta_{ji}^{(m)}$ are nonnegative constants such that $\sum_{k \in S_j(m,R)} \theta_{jk}^{(m)} = 1$. The pattern that will be learned for $t \in R$ depends on the structure of the sets

$$\mathscr{S}_m = \bigcap_{j=0}^{N} S_j(m(j), R)$$

parametrized by all functions m such that $1 \leq m(j) \leq M_j(R)$, $j = 1, 2, \ldots, n$, and as usual, on whether or not an input regularly arrives at the grid vertex v_{ji} as the contiguous associational strength $z_{ji}(t)$ is activated by a signal from its axon collateral in the ith avalanche.

For example, let all $S_j(m, R) = \{1, 2, \ldots n\}$ and $R = [0, \infty)$. Also suppose that V_{j+1} only receives an input pulse ξ time units after V_j receives an input pulse, $i \geq 1$, where ξ is the time needed for a signal to flow between successive axon collaterals. Then all avalanche control vertices receive the same input, each V_j receives a spatial pattern, and the time lag between arrival of successive spatial

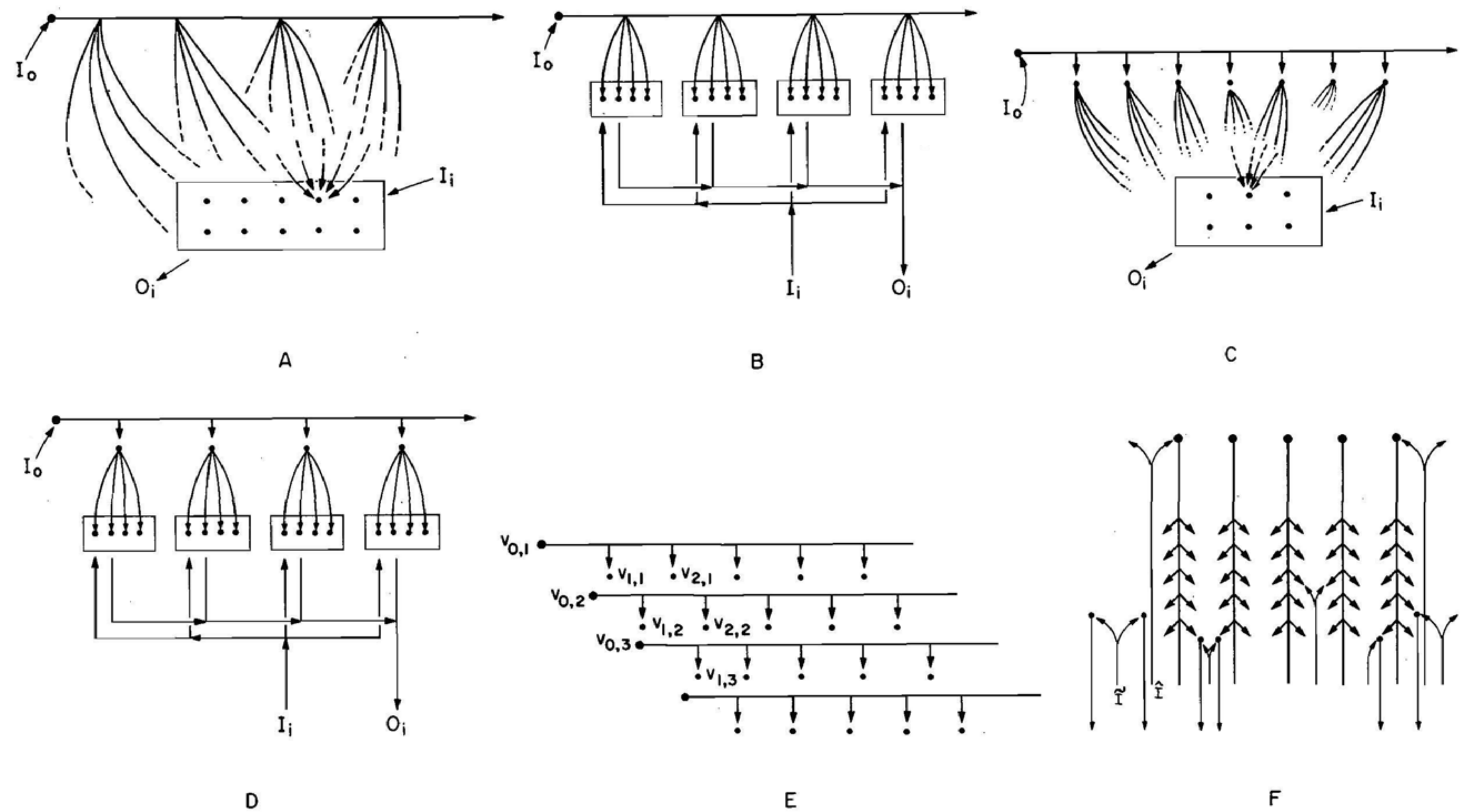

Figure 4.

patterns equals the time lag between activation of successive axon collaterals. Hence letting $y_{ji}(t) = z_{ji}(t)[\sum_{k=1}^{n} z_{jk}(t)]^{-1}$, Theorem 1 guarantees that $\lim_{t\to\infty} y_{ji}(t) = \theta_{ji}$ under rather weak conditions.

Another example is depicted in Figure 5, and illustrates the general situation. In Figure 5, V_0 receives inputs from two different sources, V_1 from two sources,

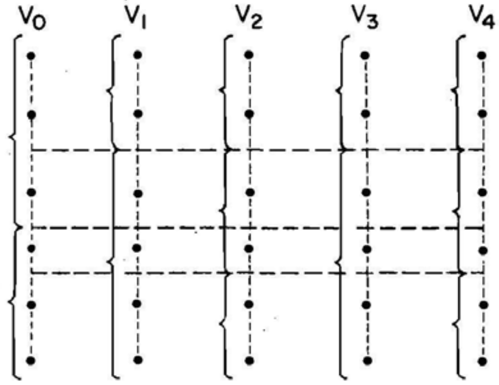

Figure 5.

V_2 from three sources, etc. Forming the intersections $\mathscr{S}_m$, the vertices v_{ji} with i in some fixed $\mathscr{S}_m$, can be divided by successive horizontal dotted lines as in the figure. Between any two dotted lines, vertices in each column receive a common input from their avalanches and a common spatial pattern from the grid. Thus the ratios $y_{ji}^{(m)}(t) = z_{ji}(t)[\sum_{k\in\mathscr{S}_m} z_{jk}(t)]^{-1}$, for all nonempty $\mathscr{S}_m$, will learn the weights $\theta_{ji}^{(m)} = \theta_{ji}[\sum_{k\in\mathscr{S}_m} \theta_{jk}]^{-1}$, in cases for which the spatial pattern arrives at the grid vertices v_{jk}, $k \in \mathscr{S}_m$, while the axon collaterals leading to these vertices sample the grid.

Type F is essentially a collection of avalanches of Type B with horizontal correlational axes rotated 90° to a vertical position, and with axon collaterals branching out in several directions. The input sources $\tilde{I}$ excite the grid elements and the input sources $\hat{I}$ excite the avalanche control vertices.

For completeness, we list the equations of types A–F below, using Γ-outstars of type $\mathscr{M}^{(*)}$ for specificity.

Type A

$$\dot{x}_0(t) = -\alpha_0 x_0(t) + I_0(t)$$

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta_0 \sum_{k=1}^{K} [x_0(t - T - \tau - k\xi) - \Gamma_0]^+ z_{k,0i}(t) + I_i(t),$$

$$\dot{z}_{k,0i}(t) = -u z_{k,0i}(t) + v_0[x_0(t - T - \tau - k\xi) - \Gamma_0]^+ x_i(t),$$

with outputs

$$O_i(t) = \delta[x_i(t) - \Gamma]^+,$$

$i = 1, 2, \ldots,$

Type B

$$\dot{x}_0(t) = -\alpha_0 x_0(t) + I_0(t),$$
$$\dot{x}_{k,i}(t) = -\alpha x_{k,i}(t) + \beta_0[x_0(t - T - \tau - k\xi) - \Gamma_0]^+ z_{k,0i}(t) + I_i(t),$$
$$\dot{z}_{k,0i}(t) = -uz_{k,0i}(t) + v_0[x_0(t - T - \tau - k\xi) - \Gamma_0]^+ x_{k,i}(t),$$

with outputs

$$O_i(t) = \delta \sum_{k=1}^{K} [x_{k,i}(t) - \Gamma]^+,$$

$i = 1, 2, \ldots, n$.

Type C

$$\dot{x}_0(t) = -\alpha_0 x_0(t) + I_0(t)$$
$$\dot{x}_{k,1}(t) = -\alpha_1 x_{k,1}(t) + \beta_0[x_0(t - T - k\xi) - \Gamma_0]^+,$$
$$\dot{x}_i(t) = -\alpha x_i(t) + \beta \sum_{k=1}^{K} [x_{k,1}(t - \tau) - \Gamma_1]^+ z_{k,1i}(t) + I_i(t),$$
$$\dot{z}_{k,1i}(t) = -uz_{k,1i}(t) + v[x_{k,1}(t - \tau) - \Gamma_1]^+ x_i(t),$$

with outputs

$$O_i(t) = \delta[x_i(t) - \Gamma]^+,$$

$i = 2, 3, \ldots, n$.

Type D

$$\dot{x}_0(t) = -\alpha_0 x_0(t) + I_0(t),$$
$$\dot{x}_{k,1}(t) = -\alpha_1 x_{k,1}(t) + \beta_0[x_0(t - T - k\xi) - \Gamma_0]^+,$$
$$\dot{x}_{k,i}(t) = -\alpha x_{k,i}(t) + \beta[x_{k,1}(t - \tau) - \Gamma_1]^+ z_{k,1i}(t) + I_i(t),$$
$$\dot{z}_{k,1i}(t) = -uz_{k,1i}(t) + v[x_{k,1}(t - \tau) - \Gamma_1]^+ x_{k,i}(t),$$

with outputs

$$O_i(t) = \delta \sum_{k=1}^{K} [x_{k,i}(t) - \Gamma]^+,$$

$i = 2, 3, \ldots, n$.

Type E

$$\dot{x}_{k,0}(t) = -\alpha_0 x_{k,0}(t) + I_{k,0}(t),$$
$$\dot{x}_{k,i}(t) = -\alpha x_{k,i}(t) + \beta_0[x_{k,0}(t - T - \tau - i\xi) - \Gamma_0]^+ z_{k,0i}(t) + I_{k,i}(t),$$
$$\dot{z}_{k,0i}(t) = -uz_{k,0i}(t) + v_0[x_{k,0}(t - T - \tau - i\xi) - \Gamma_0]^+ x_{k,i}(t)$$

$k = 1, 2, \ldots, N$; $i = 1, 2, \ldots, M$, where the output

$$O_{k,i}(t) = \delta[x_{k,i}(t) - \Gamma]^+$$

can be summed over the indices i in some nonempty $\mathscr{S}_m$.

Type F

$$\dot{x}_{k,0}(t) = -\alpha_0 x_{k,0}(t) + I_{k,0}(t),$$

$$\dot{x}_i(t) = -\alpha x_i(t) + \beta_0 \sum_{k \in K_i} [x_{k,0}(t - T - \tau - \xi_{k,i}) - \Gamma_0]^+ z_{k,0i}(t) + I_i(t),$$

$$\dot{z}_{j,0i}(t) = -u z_{j,0i}(t) + v_0[x_{j,0}(t - T - \tau - \xi_{j,i}) - \Gamma_0]^+ x_i(t),$$

$k = 1, 2, \ldots, N$; $j \in K_i$; $i = 1, 2, \ldots, M$. More general parameter choices are possible, just so long as the parameters chosen give rise to Γ-outstars when relativized to the sets $\mathscr{S}_m$. See [13], Section 10, for some more general parameter choices.

10. Some systems bounded under arbitrarily large inputs

There exist systems which learn spatial patterns in the manner of Theorem 1—and hence space-time patterns when sequentially arranged in avalanches—whose variables remain bounded under arbitrarily large inputs. Consider the following system, for example

$$\dot{x}_1(t) = (\Omega_1 - x_1(t))I_1(t) - \alpha_1 x_1(t), \tag{87}$$

$$\dot{x}_i(t) = (\Omega - x_i(t))H_i^+(t) - x_i(t)H_i^-(t), \tag{88}$$

$$H_i^+(t) = \delta(I_i(t) + \beta[x_1(t - \tau) - \Gamma_1]^+ z_{1i}(t)), \tag{89}$$

$$H_i^-(t) = \alpha + \delta \left\{ \sum_{j \neq 1,i} I_j(t) + \beta[x_1(t - \tau) - \Gamma_1]^+ \sum_{j \neq 1,i} z_{1j}(t) \right\} \tag{90}$$

and

$$\dot{z}_{1i}(t) = -u z_{1i}(t) + v[x_1(t - \tau) - \Gamma_1]^+ x_i(t), \tag{18}$$

$i = 2, \ldots, n$. We perturb this system with any inputs of the form

$$I_1(t) = \sum_{n=0}^{\infty} J_{1n}(t - t_1(n)) \tag{77}$$

and

$$I_i(t) = \theta_i \sum_{n=0}^{\infty} J_n(t - t(n)), \tag{91}$$

$i = 2, \ldots, n$, which satisfy (79)–(82), and define the probabilities $X_i(t)$ and $y_{1i}(t)$ in the usual way. Then the following corollary of Theorem 1 holds.

COROLLARY 6. *Given* (87)–(91) *and* (18) *with nonnegative and continuous initial data satisfying* $x_1(0) \leq \Omega_1$, $x_i(0) \leq \Omega$, *and*

$$\|N\|_\infty > \Gamma_1. \tag{83}$$

Then Theorem 1 *holds for the probabilities* X_i *and* y_{1i}.

Proof: $x_i(0) \leq \Omega$ implies $x_i(t) \leq \Omega$ for all $t \geq 0$ $i = 2, \ldots, n$. Since also $x_1(0) \leq \Omega$, implies $x_1(t) \leq \Omega_1$ for $t \geq 0$, $z_{1i}(t)$ is bounded. Both $x^{(1)}(t)$ and $z^{(1)}(t)$ are therefore bounded. The equations for $x^{(1)}(t)$ and $z^{(1)}(t)$ are

$$\dot{x}^{(1)}(t) = -\alpha x^{(1)} + \delta(\Omega - x^{(1)}(t))\{I(t) + \beta[x_1(t - \tau) - \Gamma_1]^+ z^{(1)}(t)\} \tag{92}$$

and

$$\dot{z}^{(1)}(t) = -uz^{(1)}(t) + \delta[x_1(t - \tau) - \Gamma_1]^+ x^{(1)}(t).$$

By (92),

$$\begin{aligned}\dot{x}^{(1)}(t) &\geq -(\alpha + \delta I(t))x^{(1)}(t) + \delta\Omega I(t)\\ &\geq -(\alpha + \delta\|I\|_\infty)x^{(1)}(t) + \delta\Omega I(t),\end{aligned}$$

and thus

$$x^{(1)}(t) \geq \delta\Omega \int_0^t I(v) \exp[-(\alpha + \delta\|I\|_\infty)(t - v)]\, dv,$$

which by (80), (82), and (91) proves the existence of a positive lower bound for $x^{(1)}$. The existence of a positive lower bound for $z^{(1)}$ readily follows.

It remains to check that X_i and y_{1i} satisfy equations of the form (31) and (32). These equations hold with

$$A_1(t) = \frac{\beta[x_1(t - \tau) - \Gamma_1]^+ z^{(1)}(t)}{x^{(1)}(t)},$$

$$B_1(t) = \frac{\delta\Omega I(t)}{x^{(1)}(t)},$$

and

$$C_1(t) = \frac{\delta[x_1(t - \tau) - \Gamma_1]^+ x^{(1)}(t)}{z^{(1)}(t)}.$$

Given the above estimates on $x^{(1)}$ and $z^{(1)}$ the proof can now be completed in the usual way.

System (87)–(91) and (18) can therefore learn a spatial pattern given an arbitrarily small memory decay rate u that is chosen independently from the value of α.

Replacing (18) by

$$\dot{z}_{1i}(t) = [-uz_{1i}(t) + vx_i(t)][x_1(t - \tau) - \Gamma_1]^+, \tag{19}$$

we again find a system whose memory is perfect in the absence of recall trials, but in which u and α can be chosen independently.

COROLLARY 7. *Let* (87)–(91) *and* (19) *be given with any nonnegative and continuous initial data which satisfy* $x_1(0) \leq \Omega_1$ *and* $x_i(0) \leq \Omega$, $i = 2, \ldots, n$, *and any inputs of the form* (77) *and* (91) *which satisfy* (79)–(82) *and* (83). *Then Theorem* 1 *holds for the probabilities* X_i *and* y_{1i}.

Proof: Since $x_i(t) \leq \Omega$ for all $t \geq 0$, $\dot{z}_{1i}(t) \leq 0$ if $z_{1i}(t) \geq v\Omega u^{-1}$. Once again $x^{(1)}$ and $z^{(1)}$ are bounded, and the remainder of the proof proceeds in the usual way.

11. Arousal and inhibition

[13] suggests that when more than one avalanche sends signals to a given grid vertex, then inhibitory signals are needed in some form to turn off some of the avalanches before massive background noise accumulates on the grid and interferes with learning and performance. Also "diffuse arousal" inputs, created by the "unconditioned stimulus", are needed in conjunction with the "conditioned

stimulus" at the avalanche control vertices to guarantee that a given Γ-outstar in an avalanche practices the same spatial approximant to a space-time pattern on successive practice trials. Analogs of "novel" stimuli, "habituation", "internal drive states", the process of "paying attention", "feedback inhibition", etc., also arise as aids for coordinating the activation and inhibition through time of avalanche clusters to produce smoothly modulated output behavior that fulfills the network's "behavioral goals". [13] should be consulted for a further discussion of these phenomena.

12. Cerebellar and cerebral analogies?

The horizontally displaced rows of avalanches of Type E are reminiscent of the lattice-like anatomy of the mammalian cerebellum ([13], and [23]). The correlational axons would then be analogs of parallel fibers, the inputs to the parallel fibers would be analogs of mossy fibers, and the inputs to the grid vertices would be analogs of climbing fibers. In the case $S_j(m, R) = \{1, 2, \ldots, n\}$ with $R = [0, \infty)$ above, we would then say that the somatotopic representations of mossy fibers and parallel fibers are mutually orthogonal. Experimental data has been collected which reports such a finding [24]. See [13] for a more thorough investigation of this analogy.

The vertically displaced rows of avalanches of Type F are reminiscent of the columnar anatomy of the mammalian cerebral cortex ([25], p. 437). The vertical correlation axons would presumably be analogs of pyramidal cells running through the cortical layers, and the existence of at least two distinct types I and $\tilde{I}$ of input sources, segregated in different cortical layers, would be suggested, one carrying unconditioned stimuli and the other conditioned stimuli. Actually a third input source, carrying diffuse arousal inputs, would also be suggested, and would presumably be distributed by the horizontal cells of Cajal ([25], p. 435) to the apical dendrites of the pyramidal cells. A forthcoming paper will discuss this cerebral analogy in greater detail.

13. Acknowledgment

The preparation of this work was supported in part by the National Science Foundation (GP9003), the Office of Naval Research (N00014-67-A-0204-0016), and the A. P. Sloan Foundation (71609).

References

1. GROSSBERG, S., "Some Physiological and Biochemical Consequences of Psychological Postulates", *Proc. Natl. Acad. Sci.*, **60**, 758–765 (1968).
2. GROSSBERG, S., "Embedding Fields: A Theory of Learning with Physiological Implications", *J. Math. Psych.* **6** (2), 209–239 (1969).
3. GROSSBERG, S., "On Learning, Information, Lateral Inhibition, and Transmitters", *Math. Biosci.* **4**, 255–310 (1969).
4. GROSSBERG, S., "On the Production and Release of Chemical Transmitters and Related Topics in Cellular Control", *J. Theoret. Biol.* **22**, 325–364 (1969).
5. GROSSBERG, S., "Nonlinear Difference-Differential Equations in Prediction and Learning Theory", *Proc. Natl. Acad. Sci.* **58**, 1329–1334 (1967).
6. GROSSBERG, S., "Some Nonlinear Networks Capable of Learning a Spatial Pattern of Arbitrary Complexity", *Proc. Natl. Acad. Sci.* **59**, 368–372 (1968).

7. Grossberg, S., "A Prediction Theory for Some Nonlinear Functional-Differential Equations (I)", *J. Math. Anal. and Applics.* **21**, 643–694 (1968).
8. Grossberg, S., "A Prediction Theory for Some Nonlinear Functional-Differential Equations (II)", *J. Math. Anal. and Applics.* **22**, 490–522 (1968).
9. Grossberg, S., "On the Serial Learning of Lists", *Math. Biosci.* **4**, 201–253 (1969).
10. Grossberg, S., "On the Global Limits and Oscillations of a System of Nonlinear Differential Equations Describing a Flow on a Probabilistic Network", *J. Diff. Eqns.* **5**(3), 531–563 (1969).
11. Grossberg, S., "On the Variational Systems of Some Nonlinear Difference-Differential Equations, *J. Diff Eqns.*, **6** (3), 544–577 (1969).
12. Grossberg, S., "Some Networks that Can Learn, Remember, and Reproduce any Number of Complicated Space-Time Patterns, I", *J. of Math. and Mechanics* **19** (1), 53–91 (1969).
13. Grossberg, S., "On Learning of Spatiotemporal Patterns by Networks with Ordered Sensory and Motor Components, I—Excitatory Components of the Cerebellum", *Studies in Applied Math.* XLVIII (2), 105–132 (1969).
14. Grossberg, S. "Embedding Fields: Underlying Philosophy, Mathematics, and Applications to Psychology, Physiology, and Anatomy". To appear in Proceedings of annual American Society of Cybernetics meeting (1969).
15. Bellman, R. E., and Cooke, K. L., "Differential-Difference Equations", New York: Academic (1963).
16. Hartman, P., "Ordinary Differential Equations", New York: Wiley (1964).
17. Courant, R., and Hilbert, D., "Methods of Mathematical Physics", Vol. 1, New York: Interscience (1953).
18. Osgood, C. E., "Method and Theory in Experimental Psychology", New York: Oxford (1953).
19. Ferster, C. B., and Skinner, B. F., "Schedules of Reinforcement", New York: Appleton-Century-Crofts (1957).
20. Eccles, J. C., "The Physiology of Synapses", New York: Academic (1964).
21. Atkinson, R. C., and Estes, W. K., "Stimulus Sampling Theory", in "Handbook of Mathematical Psychology", (ed., Luce, R. D., Bush, R. R., and Galanter, E.), Vol. II, Chapters 9–14, New York: Wiley (1963), p. 121–268.
22. Dethier, V. G., "Physiology of Insect Senses", London: Methuen and Company (1963).
23. Eccles, J. C., Ito, M., and Szentagothai, J., "The Cerebellum as a Neuronal Machine", New York: Springer-Verlag (1967).
24. Oscarsson, O., and Uddenberg, N., "Somatotopic Termination of Spino-Olivocerebellar Pathway", *Brain Res.* **3**, (1966), 204–207.
25. Crosby, E. C., Humphrey, T., and Lauer, E. W., "Correlative Anatomy of the Nervous System", New York: Macmillan (1962).

Massachusetts Institute of Technology

(Received October 20, 1969)