A memory frontier for complex synapses

Anonymous Author(s)
Affiliation
Address
email

Abstract

Blah blah blah.

1 Introduction

sec:intro

It is widely thought that our very ability to remember the past over long time scales depends crucially on our ability to modify synapses in our brain in an experience dependent manner. Classical models of synaptic plasticity model synaptic efficacy as an analog scalar value, denoting the size of a post-synaptic potential injected into one neuron from another. Theoretical work has shown that such models have a reasonable, extensive memory capacity, in which the number of long term associations that can be stored by a neuron is proportional its number of afferent synapses [1–3]. However, recent experimental work has shown that many synapses are more digital than analog; they cannot robustly assume an infinite continuum of analog values, but rather can only take on a finite number of distinguishable strengths, a number than can be as small as two [4–6] (though see [7]). This one simple modification leads to a catastrophe in memory capacity: classical models with digital synapses, when operating in a palimpset mode in which the ongoing storage of new memories can overwrite previous memories, have a memory capacity proportional to the logarithm of the number of synapses [8, 9]. Intuitively, when synapses are digital, the storage of a new memory can flip a population of synaptic switches, thereby rapidly erasing previous memories. This result indicates that the dominant theoretical basis for the storage of long term memories in modifiable synaptic switches is flawed.

Recent work [10–12] has suggested that a way out of this logarithmic catastrophe is to expand our theoretical conception of a synapse from a single scalar value to an entire stochastic dynamical system in its own right. This conceptual expansion is further necessitated by the experimental reality that synapses contain within them immensely complex molecular signaling pathways, with many internal molecular functional states (e.g. see [4, 13, 14]). While externally, synaptic efficacy could be digital, candidate patterns of electrical activity leading to potentiation or depression could yield transitions between these internal molecular states without necessarily inducing an associated change in synaptic efficacy. This form of synaptic change, known as metaplasticity [15, 16], can allow the probability of synaptic potentiation or depression to acquire a rich dependence on the history of prior changes in efficacy, thereby potentially improving memory capacity.

Theoretical studies of complex, metaplastic synapses, have focused on analyzing the memory performance of a limited number of very specific, molecular dynamical systems, characterized by a number of internal states in which potentiation and depression each induce a specific set of allowable transitions between states (e.g. see Figure 1 below). While these models can vastly outperform simple binary synaptic switches, these analyses leave open several deep and important questions. For example, how does the structure of a synaptic dynamical system determine its memory performance? What are the fundamental limits of memory performance over the space of all possible synaptic dynamical systems? What is the structural organization of synaptic dynamical systems that achieve these limits? Moreover, from an experimental perspective, it is unlikely to be the case that there is a single canonical synaptic model; just like the case of neurons, there is an incredible diver-

sity of molecular networks underlying synapses both across species and across brain regions within a single organism [17]. In order to elucidate the functional contribution of this diverse molecular complexity to learning and memory, it is essential to move beyond the analysis of specific models and instead develop a general theory of learning and memory for complex synapses. Moreover, such a general theory of complex synapses could aid in development of novel artificial memory storage devices.

Here we initiate such a general theory by proving upper bounds on the memory curve associated with any synaptic dynamical system, within the well established ideal observer framework of [10, 11, 18]. Along the way we develop principles based on first passage time theory to order the structure of synaptic dynamical systems and relate this structure to memory performance. We summarize our main results in the discussion section.

2 Overall framework: synaptic models and their memory curves

sec:setup

In this section, we will describe the class of models of synaptic plasticity that we are studying and how we quantify their memory performance. In the subsequent sections, we will find upper bounds on this performance.

We use a well established formalism for the study of learning and memory with complex synapses (see [10, 11, 18]). In this approach, electrical patterns of activity corresponding to candidate potentiating and depressing plasticity events occur randomly and independently at all synapses at a Poisson rate r. These events reflect possible synaptic changes due to either spontaneous network activity, or the storage of new memories. We let f^{pot} and f^{dep} denote the fraction of these events that are candidate potentiating or depressing events respectively. Furthermore, we assume our synaptic model has M internal molecular functional states, and that a candidate potentiating (depotentiating) event induces a stochastic transition in the internal state described by an $M \times M$ discrete time Markov transition matrix $\mathbf{M}^{\text{pot}}$ ($\mathbf{M}^{\text{dep}}$). In this framework, the states of different synapses will be independent, and the entire synaptic population can be fully described by the probability distribution across these states, which we will indicate with the row-vector $\mathbf{p}(t)$. Thus the i'th component of $\mathbf{p}(t)$ denotes the fraction of the synaptic population in state i. Furthermore, each state i has its own synaptic weight, which we take, in the worst case scenario, to be restricted to two values. After shifting and scaling these two values, we can assume they are ± 1, without loss of generality.

We also employ an "ideal observer" approach to the memory readout, where the synaptic weights are read directly. This provides an upper bound on the quality of any readout using neural activity.

For any single memory, stored at time $t = 0$, we assume there will be an ideal pattern of synaptic weights across a population of N synapses, the N-element vector $\vec{w}_{\text{ideal}}$, that is $+1$ at all synapses that experience a candidate potentiation event, and -1 at all synapses that experienced a candidate depression event at the time of memory storage. The actual pattern of synaptic weights at some later time, t, will change to $\vec{w}(t)$ due to further modifications from subsequent memories. We can use the overlap between these, $\vec{w}_{\text{ideal}} \cdot \vec{w}(t)$, as a measure of the quality of the memory. As $t \to \infty$, the system will return to its steady state distribution which will be uncorrelated with the memory stored at $t = 0$. The probability distribution of the quantity $\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty)$ can be used as a "null model" for comparison.

The extent to which the memory has been stored is described by a signal-to-noise ratio (SNR) [10, 11, 18]:

$$\text{SNR}(t) = \frac{\langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) - \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle}{\sqrt{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))}}. \tag{1}$$

eq:SNRdef

The noise in the denominator is essentially $\sqrt{N}$. There is a correction when potentiation and depression are imbalanced, but this will not affect the upper bounds that we will discuss below and will be ignored in the subsequent formulae.

A simple mean memory curve can be derived as follows. All of the preceding plasticity events, prior to $t = 0$, will put the population of synapses in its steady-state distribution, $\mathbf{p}^{\infty}$. The memory we are tracking at $t = 0$ will change the internal state distribution to $\mathbf{p}^{\infty}\mathbf{M}^{\text{pot}}$ in those synapses that are potentiated and $\mathbf{p}^{\infty}\mathbf{M}^{\text{dep}}$ in those synapses that are depressed. As the potentiating/depressing nature of the subsequent memories is independent of $\vec{w}_{\text{ideal}}$, we can average over all sequences, resulting in

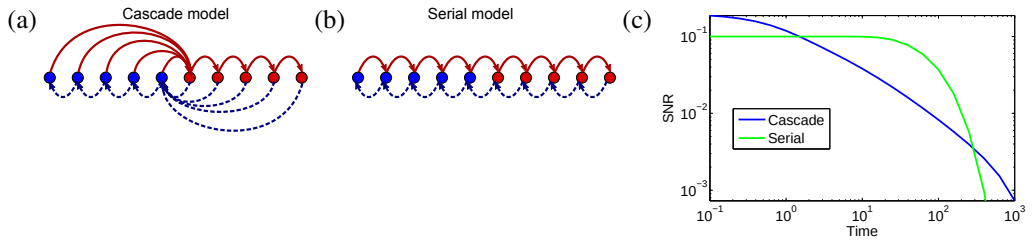

fig:cascade fig:cascade

Figure 1: Models of complex synapses. (a) The cascade model of [10], showing transitions between states of high/low synaptic weight (red/blue circles) due to potentiation/depression (solid red/dashed blue arrows). (b) The serial model of [12]. (c) The memory curves of these two models, showing the decay of the signal-to-noise ratio (to be defined in §2) as subsequent memories are stored.

fig:models

the evolution of the probability distribution:

$$\frac{\mathrm{d}\mathbf{p}(t)}{\mathrm{d}t} = r\mathbf{p}(t)\mathbf{W}^{\mathrm{F}}, \qquad \text{where} \quad \mathbf{W}^{\mathrm{F}} = f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}} - \mathbf{I}. \tag{2}$$

eq:evol

Here $\mathbf{W}^{\mathrm{F}}$ is a continuous time transition matrix that models the process of forgetting the memory stored at time $t = 0$ due to random candidate potentiation/depression events occurring at each synapse due to the storage of subsequent memories.

This results in the following SNR

$$\mathrm{SNR}(t) = \sqrt{N}\left(2f^{\mathrm{pot}}f^{\mathrm{dep}}\right)\mathbf{p}^{\infty}\left(\mathbf{M}^{\mathrm{pot}} - \mathbf{M}^{\mathrm{dep}}\right)\mathrm{e}^{rt\mathbf{W}^{\mathrm{F}}}\mathbf{w}, \tag{3}$$

eq:SNRcalc

where the weight of the synapse when it is in its i'th state is given by the corresponding element of the column vector $\mathbf{w}$. We will frequently refer to this function as the memory curve. It can be thought of as the fraction of synapses that maintain the ideal synaptic strength at time t, dictated by the stored memory at time $t = 0$.

Much of the previous work on these types of complex synaptic models has focused on understanding the memory curves of specific models, or choices of $\mathbf{M}^{\mathrm{pot/dep}}$. Two examples of these models are shown in Figure 1. We see that they have different memory properties. The serial model performs relatively well at one particular timescale, but it performs poorly at other times. The cascade model does not perform quite as well at that time, but it maintains it performance over a wider range of timescales.

In this work, rather than analyzing specific models, we take a different approach, in order to obtain a more general theory. We consider the *entire* space of these models and find upper bounds on the memory capacity of *any* of them. The space of models with a fixed number of internal states M is parameterized by the pair of $M \times M$ discrete time stochastic transition matrices $\mathbf{M}^{\mathrm{pot}}$ and $\mathbf{M}^{\mathrm{dep}}$, in addition to $f^{\mathrm{pot/dep}}$. The parameters must satisfy the following constraints:

$$\begin{aligned} &\mathbf{M}_{ij}^{\mathrm{pot/dep}} \in [0,1], \qquad f^{\mathrm{pot/dep}} \in [0,1], \qquad \mathbf{p}^{\infty}\mathbf{W}^{\mathrm{F}} = 0, \qquad \mathbf{w}_i = \pm 1, \\ &\sum_j \mathbf{M}_{ij}^{\mathrm{pot/dep}} = 1, \qquad f^{\mathrm{pot}} + f^{\mathrm{dep}} = 1, \qquad \sum_i \mathbf{p}_i^{\infty} = 1. \end{aligned} \tag{4}$$

eq:constr

The upper bounds on $\mathbf{M}_{ij}^{\mathrm{pot/dep}}$ and $f^{\mathrm{pot/dep}}$ follow automatically from the other constraints.

The critical question is: what do these constraints imply about the space of achievable memory curves in (3)? To answer this question, especially for limits on achievable memory at finite times, it will be useful to employ the eigenmode description:

$$\mathbf{W}^{\mathrm{F}} = \sum_a -q_a \mathbf{u}^a \mathbf{v}^a, \quad \mathbf{v}^a\mathbf{u}^b = \delta_{ab}, \quad \mathbf{W}^{\mathrm{F}}\mathbf{u}^a = -q_a\mathbf{u}^a, \quad \mathbf{v}^a\mathbf{W}^{\mathrm{F}} = -q_a\mathbf{v}^a. \tag{5}$$

eq:eigendecomp

Here q_a are the negative of the eigenvalues of the forgetting process $\mathbf{W}^{\mathrm{F}}$, $\mathbf{u}^a$ are the right (column) eigenvectors and $\mathbf{v}^a$ are the left (row) eigenvectors. This decomposition allows us to write the

memory curve as a sum of exponentials,

$$\mathrm{SNR}(t) = \sqrt{N}\sum_a \mathcal{I}_a \mathrm{e}^{-rt/\tau_a}, \qquad \text{where} \quad \mathcal{I}_a = \left(2f^{\mathrm{pot}}f^{\mathrm{dep}}\right)\mathbf{p}^\infty(\mathbf{M}^{\mathrm{pot}}-\mathbf{M}^{\mathrm{dep}})\mathbf{u}^a\mathbf{v}^a\mathbf{w}, \quad \text{and} \quad \tau_a = \frac{1}{q_a}. \tag{6}$$

eq:SNReigen

We can then ask the question: what are the constraints on these quantities, namely initial SNR's $\mathcal{I}_a$ and time constants τ_a for the modes, implied by the constraints in (4)? We will derive some of these constraints in the next section.

3 Upper bounds on achievable memory capacity

sec:bounds

In the previous section, in (3) we have described an analytic expression for a memory curve as a function of the structure of a synaptic dynamical system, described by the pair of stochastic transition matrices $\mathbf{M}^{\mathrm{pot/dep}}$. Since the performance measure for memory is an entire memory curve, and not just a single number, there is no universal scalar notion of optimal memory in the space of synaptic dynamical systems. Instead there are tradeoffs between storing proximal and distal memories; often attempts to increase memory at late (early) times by changing $\mathbf{M}^{\mathrm{pot/dep}}$, incurs a performance loss in memory at early (late) times in specific models considered so far [10–12]. Thus our goal, achieved in §4, is to derive an envelope memory curve in the SNR-time plane, or a curve that forms an upper-bound on the entire memory curve of *any* model. In order to achieve this goal, in this section, we must first derive upper bounds, over the space of all possible synaptic models, on two different scalar functions of the memory curve: its initial SNR, and the area under the memory curve. As a corollary of these results, we also derive an upper-bound on the memory lifetime, or the time at which the SNR drops to a given detection threshold.

3.1 Initial SNR

sec:initial

Now we will discuss the SNR at $t = 0$:

$$\mathrm{SNR}(0) = \sqrt{N}\left(2f^{\mathrm{pot}}f^{\mathrm{dep}}\right)\mathbf{p}^\infty\left(\mathbf{M}^{\mathrm{pot}}-\mathbf{M}^{\mathrm{dep}}\right)\mathbf{w}. \tag{7}$$

eq:init

We will find an upper bound on this quantity for *all* possible models and also find the model that saturates this bound. As maximizing initial SNR does not place any value on later parts of the memory curve, this will produce models with fast decay that may not be good for memory storage. Nevertheless, the upper bound will be useful in §4.

A useful quantity is the equilibrium probability flux between two disjoint sets of states, $\mathcal{A}$ and $\mathcal{B}$:

$$\mathbf{\Phi}_{\mathcal{A}\mathcal{B}} = \sum_{i\in\mathcal{A}}\sum_{j\in\mathcal{B}} r\mathbf{p}_i^\infty \mathbf{W}_{ij}^{\mathrm{F}}. \tag{8}$$

eq:flux

The initial SNR is closely related to the flux between the states with $\mathbf{w}_i = -1$ and those with $\mathbf{w}_j = +1$:

$$\mathrm{SNR}(0) \le \frac{4\sqrt{N}\mathbf{\Phi}_{-+}}{r}. \tag{9}$$

eq:initflux

This inequality becomes an equality if potentiation never decreases the synaptic weight and depression never increases it, which should be a property of any sensible model.

To maximize this flux, potentiation from a weak state must be guaranteed to end in a strong state, and depression must do the reverse. An example of such a model is shown in Figure 2(a,b). These models have a property known as "lumpability" (see [19, §6.3] for the discrete time version and [20, 21] for continuous time). They are completely equivalent (i.e. have the same memory curve) as a two state model with transition probabilities equal to 1, as shown in Figure 2(c).

This two state model has the equilibrium distribution $\mathbf{p}^\infty = (f^{\mathrm{dep}}, f^{\mathrm{pot}})$ and its flux is given by $\mathbf{\Phi}_{-+} = rf^{\mathrm{pot}}f^{\mathrm{dep}}$. This is maximized when $f^{\mathrm{pot}} = f^{\mathrm{dep}} = \frac{1}{2}$, leading to the upper bound:

$$\mathrm{SNR}(0) \le \sqrt{N}. \tag{10}$$

eq:max_init

Whilst this model has high initial SNR, it also has very fast decay – with a timescale $\tau \sim \frac{1}{r}$. As the synapse is very plastic, the initial memory is encoded very easily, but the subsequent memories also overwrite it easily. This is not a good design for a synapse, but the resulting bound on the initial SNR will prove useful in §4.

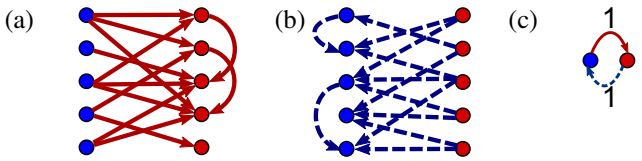

fig:max_init_dep

Figure 2: Synaptic models that maximize initial SNR. (a) For potentiation, all transitions starting from a weak state lead to a strong state, and the probabilities for all transitions leaving a given weak state sum to 1. (a) Depression is similar to potentiation, but with strong and weak interchanged. (c) The equivalent two state model, with transition probabilities under potentiation and depression equal to one.

fig:max_init

3.2 Area

sec:area

Now consider the area under the memory curve:

$$A = \int_0^\infty \mathrm{d}t\ \mathrm{SNR}(t). \tag{11}$$

eq:area

We will find an upper bound on this quantity as well as the model that saturates it. The area is useful as it also provides an upper bound on the lifetime of a memory, $\tau(\epsilon)$, defined as the time at which $\mathrm{SNR}(\tau) = \epsilon$ (we often take $\epsilon = 1$):

$$A = \int_0^\infty \mathrm{d}t\ \mathrm{SNR}(t) > \int_0^{\tau(\epsilon)} \mathrm{d}t\ \mathrm{SNR}(t) > \int_0^{\tau(\epsilon)} \mathrm{d}t\, \epsilon = \epsilon\tau(\epsilon). \tag{12}$$

eq:arealife

However, the area still receives a contribution from the tail of the memory curve, after the SNR has decayed below any useful value. This means that maximizing SNR will favor models with long tails that will not necessarily be good for storing memories.

One important property of the area is that it has a null "scaling" degree of freedom. If all off-diagonal elements of $\mathbf{M}^{\text{pot/dep}}$ are multiplied by a constant, with diagonal elements adjusted to keep row-sums fixed, this will have two effects. First, it will scale up the initial signal creation. Second, it will scale down time. Combining these two effects, it will leave the area invariant.

This feature allows us to ignore the lower bound on the diagonal elements, $\mathbf{M}^{\text{pot/dep}}_{ii} > 0$. If we maximize the area ignoring this constraint, the "scaling" mode can be used to satisfy it without changing the area.

We also develop an important organizing principle using the theory of first passage times. The mean first passage time matrix, $\overline{\mathbf{T}}_{ij}$, is defined as the average time it takes to reach state j for the first time, starting from state j. The diagonal elements are defined to be zero. Then the quantity

$$\eta \equiv \sum_j \overline{\mathbf{T}}_{ij}\mathbf{p}^\infty_j \tag{13}$$

eq:kemenyconst

is independent of the starting state i. It is known as Kemeny's constant (see [19, §4.4]). We can define analogous quantities

$$\eta^+_i = \sum_{j\in+} \overline{\mathbf{T}}_{ij}\mathbf{p}^\infty_j, \qquad \eta^-_i = \sum_{j\in-} \overline{\mathbf{T}}_{ij}\mathbf{p}^\infty_j. \tag{14}$$

eq:kemenypm

These can be thought of as the average time it takes to reach the strong/weak states respectively. We can put the states in order of decreasing η^+_i or increasing η^-_i. Because $\eta^+_i + \eta^-_i = \eta$ is independent of i, the two orderings are the same.

With the states in this order, we can find perturbations of $\mathbf{M}^{\text{pot/dep}}$ that will always increase the area, whilst leaving the equilibrium distribution, $\mathbf{p}^\infty$, unchanged. Some of these perturbations are shown in Figure 3. The only thing that can prevent these perturbations from increasing the area is when they require the decrease of a matrix element that has already been set to 0. This determines the topology (non-zero transition probabilities) of the model with maximal area. It is of the form shown

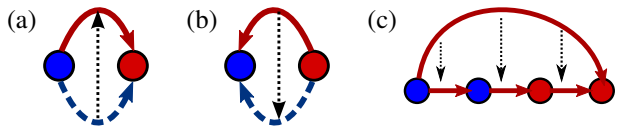

fig:shortcut

Figure 3: Perturbations that increase the area. (a) Perturbations that increase elements of $\mathbf{M}^{\text{pot}}$ above the diagonal and decrease the corresponding elements of $\mathbf{M}^{\text{dep}}$. It can no longer be used when $\mathbf{M}^{\text{dep}}$ is lower triangular. (b) Perturbations that decrease elements of $\mathbf{M}^{\text{pot}}$ below the diagonal and increase the corresponding elements of $\mathbf{M}^{\text{dep}}$. It can no longer be used when $\mathbf{M}^{\text{pot}}$ is upper triangular. (c) Perturbation that decreases "shortcut" transitions and increases the bypassed "direct" transitions. It can no longer be used when there are only nearest-neighbor "direct" transitions.

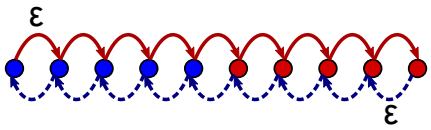

fig:perts

Figure 4: Model that maximizes the area. Unlabeled transitions have probability 1. Labeled transitions have probability $\varepsilon \to 0$.

fig:max_area

in Figure 4, with potentiation moving one step to the right and depression moving one step to the left. Any other topology would allow some perturbation to further increase the area.

As these perturbations do not change the equilibrium distribution, this means that the area of *any* model is bounded by that of a linear chain with the same equilibrium distribution. This area is given by

$$A \le \frac{2\sqrt{N}}{r} \sum_k \left[k - \sum_j j \mathbf{p}_j^\infty \right] \mathbf{p}_k^\infty \mathbf{w}_k = \frac{2\sqrt{N}}{r} \sum_k \left| k - \sum_j j \mathbf{p}_j^\infty \right| \mathbf{p}_k^\infty, \tag{15}$$

eq:multistatearea

where we chose $\mathbf{w}_k = \operatorname{sgn}\left[k - \sum_j j \mathbf{p}_j^\infty\right]$. We can then maximize this by pushing all of the equilibrium distribution symmetrically to the two end states. This can be done by reducing the transition probabilities out of these states, as in Figure 4. This makes it very difficult to get out of the end states once they have been entered. The resulting area is

$$A \le \frac{\sqrt{N}(M-1)}{r}. \tag{16}$$

eq:max_area

This result is similar to that found numerically by Barrett and van Rossum [18] in a slightly different situation.

The "sticky" end states result in very slow decay, but they also make it difficult to encode the memory in the first place. The memory curve for this model starts off very low, but hovers above zero for a very long time. This is also not a good design for a synapse, but the resulting bound on the area will also prove useful in §4.

4 Memory curve envelope

sec:env

Now we will look at the implications of the upper bounds found in the previous section for the SNR at finite times. As argued in (6), the memory curve can be written in the form

$$\mathrm{SNR}(t) = \sqrt{N} \sum_a \mathcal{I}_a \mathrm{e}^{-rt/\tau_a}. \tag{17}$$

eq:SNRmodes

The upper bounds on the initial SNR, (10), and the area, (16), imply the following constraints on the parameters $\{\mathcal{I}_a, \tau_a\}$:

$$\sum_a \mathcal{I}_a \le 1, \qquad \sum_a \mathcal{I}_a \tau_a \le M - 1. \tag{18}$$

eq:modeconstr

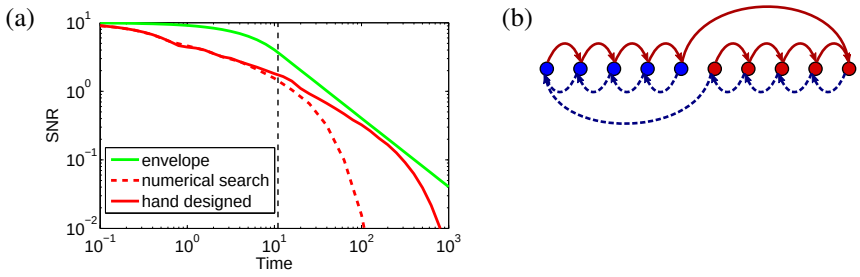

Figure 5: The memory curve envelope for $N = 100$, $M = 12$. (a) An upper bound on the SNR at any time is shown in green. The red dashed curve shows the result of numerical optimization of synaptic models with random initialization. The solid red curve shows the highest SNR we have found with hand designed models. At early times these models are of the form shown in (b) with different numbers of states. At late times they are of the form shown in Figure 4 with different values of ε.

We are not claiming that these are a complete set of constraints: not every set $\{\mathcal{I}_a, \tau_a\}$ that satisfies these inequalities will actually be achievable by a synaptic model. However, any set that violates either inequality will definitely not be achievable.

Now we can pick some fixed time, t_0, and maximize the SNR at that time wrt. the parameters $\{\mathcal{I}_a, \tau_a\}$, subject to the constraints above. This always results in a single nonzero $\mathcal{I}_a$:

$$\begin{aligned} t_0 \le \frac{M-1}{r} \quad &\Longrightarrow \mathrm{SNR}(t) = \sqrt{N}\mathrm{e}^{-rt/(M-1)} \quad &&\Longrightarrow \mathrm{SNR}(t_0) = \sqrt{N}\mathrm{e}^{-rt_0/(M-1)}, \\ t_0 \ge \frac{M-1}{r} \quad &\Longrightarrow \mathrm{SNR}(t) = \frac{\sqrt{N}(M-1)\mathrm{e}^{-t/t_0}}{rt_0} \quad &&\Longrightarrow \mathrm{SNR}(t_0) = \frac{\sqrt{N}(M-1)}{\mathrm{e}rt_0}. \end{aligned} \tag{19}$$

eq:env

Both the initial SNR bound and the area bound are saturated at early times. At late times, only the area bound is saturated. The function $\mathrm{SNR}(t_0)$, the green curve in Figure 5(a), above forms a memory curve envelope. No synaptic model can have an SNR that is greater than this at any time. We can use this to find an upper bound on the memory lifetime, $\tau(\epsilon)$, by finding the point at which the envelope crosses ϵ:

$$\tau(\epsilon) \le \frac{\sqrt{N}(M-1)}{\epsilon \mathrm{e}r}, \tag{20}$$

eq:lifebound

where we assume $N > (\epsilon \mathrm{e})^2$.

This leaves the question of whether this bound is achievable. At any time, can we find a model whose memory curve touched the envelope? The red curves in Figure 5(a) show the closest we have come to the envelope with actual models, by repeated numerical optimization with random initialization and by hand designed models.

We see that here is a gap between the upper bound that we can prove and what we can achieve with actual models. There may be other models we haven't found that could beat the ones we have, but it is probably too much to expect them to reach this envelope. That would require saturating the area bound, and we saw that the only model that does that has very low initial SNR. It seems that we need to find additional constraints to add to those we have (18), thereby bringing the envelope down. We do not necessarily need to find *every* constraint. We have already seen that not every constraints is saturated for certain time ranges, so it may be that a small subset of them will be sufficient to find an envelope that is tight.

5 Discussion

sec:disc

References

[1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. U.S.A.* **79** (1982) no. 8, 2554–2558.

Hopfield1982model

Amit1985hopfield

[2] D. J. Amit, H. Gutfreund, and H. Sompolinsky, "Spin-glass models of neural networks," *Phys. Rev. A* **32** (Aug, 1985) 1007–1018. http://link.aps.org/doi/10.1103/PhysRevA.32.1007.

dner1988perceptron

[3] E. Gardner, "The space of interactions in neural network models," *Journal of Physics A: Mathematical and General* **21** (1988) no. 1, 257.

Bliss1993LTP

[4] T. V. P. Bliss and G. L. Collingridge, "A synaptic model of memory: long-term potentiation in the hippocampus," *Nature* **361** (Jan, 1993) 31–39.

ersen1998allornone

[5] C. C. H. Petersen, R. C. Malenka, R. A. Nicoll, and J. J. Hopfield, "All-or-none potentiation at CA3-CA1 synapses," *Proc. Natl. Acad. Sci. U.S.A.* **95** (1998) no. 8, 4732–4737.

O'Connor2005switch

[6] D. H. O'Connor, G. M. Wittenberg, and S. S.-H. Wang, "Graded bidirectional synaptic plasticity is composed of switch-like unitary events," *Proc. Natl. Acad. Sci. U.S.A.* **102** (2005) no. 27, 9679–9684.

Enoki2009graded

[7] R. Enoki, Y. ling Hu, D. Hamilton, and A. Fine, "Expression of Long-Term Plasticity at Individual Synapses in Hippocampus Is Graded, Bidirectional, and Mainly Presynaptic: Optical Quantal Analysis," *Neuron* **62** (2009) no. 2, 242 – 253.

mit1992constraints

[8] D. J. Amit and S. Fusi, "Constraints on learning in dynamic synapses," *Network: Computation in Neural Systems* **3** (1992) no. 4, 443–464.

amit1994learning

[9] D. J. Amit and S. Fusi, "Learning in neural networks with material synapses," *Neural Computation* **6** (1994) no. 5, 957–982.

Fusi2005cascade

[10] S. Fusi, P. J. Drew, and L. F. Abbott, "Cascade models of synaptically stored memories," *Neuron* **45** (Feb, 2005) 599–611.

Fusi2007multistate

[11] S. Fusi and L. F. Abbott, "Limits on the memory storage capacity of bounded synapses," *Nat. Neurosci.* **10** (Apr, 2007) 485–493.

Leibold2008serial

[12] C. Leibold and R. Kempter, "Sparseness Constrains the Prolongation of Memory Lifetime via Synaptic Metaplasticity," *Cerebral Cortex* **18** (2008) no. 1, 67–77.

Bredt2003AMPA

[13] D. S. Bredt and R. A. Nicoll, "AMPA Receptor Trafficking at Excitatory Synapses," *Neuron* **40** (2003) no. 2, 361 – 379.

009phosphorylation

[14] M. P. Coba, A. J. Pocklington, M. O. Collins, M. V. Kopanitsa, R. T. Uren, S. Swamy, M. D. Croning, J. S. Choudhary, and S. G. Grant, "Neurotransmitters drive combinatorial multistate postsynaptic density networks," *Sci Signal* **2** (2009) no. 68, ra19.

1996metaplasticity

[15] W. C. Abraham and M. F. Bear, "Metaplasticity: the plasticity of synaptic plasticity," *Trends in Neurosciences* **19** (1996) no. 4, 126 – 130.

Montgomery2002765

[16] J. M. Montgomery and D. V. Madison, "State-Dependent Heterogeneity in Synaptic Depression between Pyramidal Cell Pairs," *Neuron* **33** (2002) no. 5, 765 – 777.

Emes2012synapserev

[17] R. D. Emes and S. G. Grant, "Evolution of Synapse Complexity and Diversity," *Annual Review of Neuroscience* **35** (2012) no. 1, 111–131.

arrett2008discrete

[18] A. B. Barrett and M. C. van Rossum, "Optimal learning rules for discrete synapses," *PLoS Comput. Biol.* **4** (Nov, 2008) e1000230.

kemeny1960finite

[19] J. Kemeny and J. Snell, *Finite markov chains*. Springer, 1960.

ourke1958markovian

[20] C. Burke and M. Rosenblatt, "A Markovian function of a Markov chain," *The Annals of Mathematical Statistics* **29** (1958) no. 4, 1112–1122.

all1993Lumpability

[21] F. Ball and G. F. Yeo, "Lumpability and Marginalisability for Continuous-Time Markov Chains," *Journal of Applied Probability* **30** (1993) no. 3, 518–528.