A memory frontier for complex synapses

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Blah blah blah.

1 Introduction

sec:intro

2 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 performance in memory storage. 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 [1–3]). In this approach, potentiating and depressing plasticity events occur at random times, with all information about the neural activity and learning rules responsible for them absorbed into their rates. We assume that there are no spatial or temporal correlations in the pattern of potentiating and depressing events, and these events cause Markovian transitions between the internal states of the synapses. As a result of these assumptions, the states of different synapses will be independent, and the system can be fully described by the probability distribution across these states, which we will indicate with the row-vector $\mathbf{p}(t)$.

We also employ an "ideal observer" approach to the memory readout, where the synaptic weights are read directly. This provides an upper bound to the quality of any readout using neural activity. These synaptic weights will be restricted to two values, which we can shift and scale to ± 1.

For any single memory, there will be an ideal pattern of synaptic weights, the N-element vector $\vec{w}_{\text{ideal}}$, that is $+1$ at all synapses that were potentiated and -1 at all synapses that were depressed. The actual pattern of synaptic weights at some later time, t, will be $\vec{w}(t)$. 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 in question. 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):

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

eq:SNRdef

The noise 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.

All of the preceding plasticity events will put the system in its steady-state distribution, $\mathbf{p}^{\infty}$. The memory we are tracking at $t = 0$ will change this 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, where $\mathbf{M}^{\text{pot/dep}}$ are $M \times M$ matrices of transition probabilities for the Markov processes describing potentiation and depression. As the potentiating/depressing nature of the subsequent memories is independent of $\vec{w}_{\text{ideal}}$, we can average over all

1

sequences, resulting in 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, r is the rate of the Poisson process describing the timing of plasticity events, $f^{\mathrm{pot/dep}}$ is the fraction of events that are potentiating/depressing and $\mathbf{I}$ is the identity matrix.

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.

The parameters must satisfy the following constraints:

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

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

The question is: what do these constraints imply for the memory curve above? In practice, to make any statements about finite times, we need to got to 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`

Where q_a are the negative of the eigenvalues, $\mathbf{u}^a$ are the right (column) eigenvectors and $\mathbf{v}^a$ are the left (row) eigenvectors. This allow us to write

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

We can ask then ask the question: what are the constraints on these quantities implied by the constraints (4)? We will find some of these constraints in the next section.

3 Upper bounds

`sec:bounds`

In this section, we will find some upper bounds on certain properties of the memory curve discussed in the previous section, namely its initial value and the area under it. We will use these bounds to determine an upper bound on the SNR at finite times in the next section.

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 $\mathcal{B}$:

$$\mathbf{\Phi}_{\mathcal{A}\mathcal{B}} = \sum_{i \in \mathcal{A}}\sum_{j \in \mathcal{B}} r\mathbf{p}^{\infty}_i \mathbf{W}^{\mathrm{F}}_{ij}. \tag{8}$$ `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`
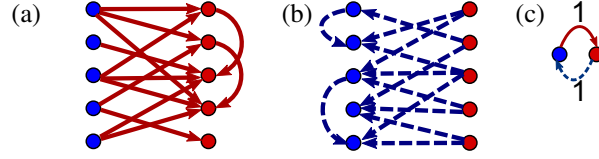
2

Figure 1: 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.

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 maximise 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 1(a,b). These models have a property known as "lumpability" (see [4, §6.3] for the discrete time version and [5, 6] 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 1(c).

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

$$\text{SNR}(0) \leq \sqrt{N}. \tag{10}$$

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.

3.2 Area

Now consider the area under the memory curve:

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

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 $\text{SNR}(\tau) = \epsilon$:

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

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, $\mathbf{M}_{ii}^{\text{pot/dep}} > 0$. 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

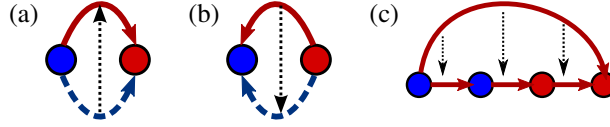$$\eta \equiv \sum_j \overline{\mathbf{T}}_{ij} \mathbf{p}_j^\infty \tag{13}$$

3

Figure 2: 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.

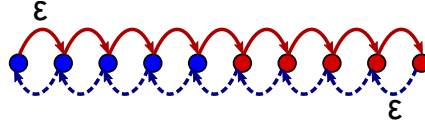Figure 3: Model that maximizes the area. Unlabeled transitions have probability 1. Labeled transitions have probability $\epsilon \to 0$.

is independent of the starting state i. It is known as Kemeny's constant (see [4, §4.4]). We can define analogous quantities

$$\eta_i^+ = \sum_{j \in +} \overline{\mathbf{T}}_{ij} \mathbf{p}_j^\infty, \qquad \eta_i^- = \sum_{j \in -} \overline{\mathbf{T}}_{ij} \mathbf{p}_j^\infty. \tag{14}$$

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 2. 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 in Figure 3, with potentiation moving one step to the right and depression moving one step to the left. Any other topology would allow some perturbation that could 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

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

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 3. This makes it very difficult to get out of the end states once they have been entered. The resulting area is

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

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

4 Memory curve envelope

5 Discussion

References

[1] S. Fusi, P. J. Drew, and L. F. Abbott, "Cascade models of synaptically stored memories," *Neuron* **45** (Feb, 2005) 599–611, `PubMed:15721245`.

[2] S. Fusi and L. F. Abbott, "Limits on the memory storage capacity of bounded synapses," *Nat. Neurosci.* **10** (Apr, 2007) 485–493, `PubMed:17351638`.

[3] A. B. Barrett and M. C. van Rossum, "Optimal learning rules for discrete synapses," *PLoS Comput. Biol.* **4** (Nov, 2008) e1000230, `PubMed:19043540`.

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

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

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