Laplacian envelope

Subhaneil Lahiri

May 31, 2015

Abstract

We try to find the continuous time Markov process that has the maximal Laplace tranformed signal-to-noise curve.

Contents

1 Framework

`:framework`

1.1 Recognition memory

`sec:recog`

We will be trying to store patterns in a set od N synaptic weights, $\vec{w}$. Every time we try to store a pattern, these synapses are subjected to a plasticity event where each synapse is either potentiated or depressed, depending on the pattern. we will assume that these patterns are spatially and temporally independent.

At some time, suppose we wish to determine if a given pattern is one of those that we previously attempted to store. We wish to answer this question by looking at the

1

synaptic weights directly (ideal observer). For that given pattern there will be an ideal set of synaptic weights, $\vec{w}_{\text{ideal}}$, where those synapses that were supposed to be potentiated are maximised and those that were supposed to be depressed are minimised. Suppose that the given pattern was actually seen at time 0 and we are observing the synapses at time t. The actual set of synaptic weights we see, $\vec{w}(t)$, will be a vector of random variables that differs from $\vec{w}_{\text{ideal}}$ due to the stochasticity of the pattern encoding and all of the other (uncorrelated) pattern that are stored after it. As $t \to \infty$, the synaptic weights will become independent of the patter stored at $t = 0$. Thus, the vector of random variables $\vec{w}(\infty)$ also describes the synaptic weights under the null hypothesis – if that given pattern had never been stored.

We can test if the pattern had been previously stored by computing $\vec{w}_{\text{ideal}} \cdot \vec{w}$ and comparing it to some threshold. For large N, this quantity will have a Gausssian distribution. There will be a ROC curve as a function of this threshold:

$$\text{TPR} = \Phi_c \left(\frac{\Phi_c^{-1}(\text{FPR}) - \text{SNR}}{\text{NNR}} \right), \quad \text{where} \quad \Phi_c(x) = \int_x^\infty \frac{\mathrm{e}^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \, \mathrm{d}z,$$

$$\text{SNR} = \frac{\langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) \rangle - \langle \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle}{\sqrt{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))}}, \quad (1) \quad \boxed{\texttt{eq:ROC}}$$

$$\text{NNR} = \sqrt{\frac{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(t))}{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))}},$$

and TPR/FPR are the true/false positive rates. When the signal-to-noise ratio, SNR is larger than $\Phi_c^{-1}(\text{FPR})$, it is beneficial to decrease the noise-to-noise ratio, NNR. In the other case, it is beneficial to increase it. The expectations and variances are over the probability distribution of the synaptic states given the sequence of plasticity events, the probability distribution of the sequence of plasticity events themselves and the probability distribution of the pattern we are testing, $\vec{w}_{\text{ideal}}$.

The formulae above assumed that we know the time between storage and recognition. We should also compute the expectations and variances over probability distribution of the recall time, t, as well. If we only know average time, τ, a natural choice for this distribution is

$$P(t|\tau) = \frac{\mathrm{e}^{-t/\tau}}{\tau}. \quad (2) \quad \boxed{\texttt{eq:recogtime}}$$

Different parts of the brain, that store memories for different timescales, could be characterised by different values of τ.

1.2 Signal-to-noise ratio

We will model the synapses as having two possible synaptic weights. As we are looking at the information contained in the synaptic weights, rather than modelling the readout via electrical activity of the neurons, what values the synaptic weights actually take is irrelevant. It will be convenient to call these two values ± 1.

As discussed in §1.1, the signal-to-noise ratio is given by

$$\text{SNR}(t) = \frac{\langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) \rangle - \langle \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle}{\sqrt{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))}}, \qquad (3) \quad \boxed{\texttt{eq:SNRdef}}$$

Where the averages and variances are over the probability distribution of the synaptic states given the sequence of plasticity events, the probability distribution of the sequence of plasticity events themselves and the probability distribution of the pattern we are testing, $\vec{w}_{\text{ideal}}$. If we also average over the time between storage and recognition, we have

$$\overline{\text{SNR}}(\tau) = \int_0^\infty \mathrm{d}t \, P(t|\tau) \, \text{SNR}(t) = \frac{1}{\tau} \int_0^\infty \mathrm{d}t \, \mathrm{e}^{-t/\tau} \, \text{SNR}(t), \qquad (4) \quad \boxed{\texttt{eq:snrb}}$$

This is similar to the average SNR up to time τ, but with the hard cut-off replaced with a smooth exponential. This can be expressed in terms of he Laplace transform of the SNR curve in (3):

$$A(s) = \int_0^\infty \mathrm{d}t \, \mathrm{e}^{-st} \, \text{SNR}(t). \qquad \overline{\text{SNR}}(\tau) = \frac{A(1/\tau)}{\tau}, \qquad (5) \quad \boxed{\texttt{eq:laplacede}}$$

First, let's look at the variances, remembering that the states and plasticity events of each synapse are independent and identically distributed:

$$\begin{aligned}
\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(t)) &= \sum_{\alpha\beta} \left\langle \vec{w}_{\text{ideal}}^\alpha \vec{w}^\alpha(t) \vec{w}_{\text{ideal}}^\beta \vec{w}^\beta(t) \right\rangle - \left(\sum_\alpha \langle \vec{w}_{\text{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \right)^2 \\
&= \sum_\alpha \left\langle (\vec{w}_{\text{ideal}}^\alpha)^2 (\vec{w}^\alpha(t))^2 \right\rangle + \sum_{\alpha\neq\beta} \langle \vec{w}_{\text{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \langle \vec{w}^\beta_{\text{ideal}} \vec{w}^\beta(t) \rangle \\
&\qquad\qquad - \left(\sum_\alpha \langle \vec{w}_{\text{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \right)^2 \\
&= N \langle 1 \rangle + N(N-1) \left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle^2 - N^2 \left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle^2 \\
&= N(1 - \left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle^2),
\end{aligned} \qquad (6) \quad \boxed{\texttt{eq:noise}}$$

where we used $\vec{w}^\alpha = \pm 1$. We can compute $\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))$ by taking $t \to \infty$.

For the numerator, we can write

$$\langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) \rangle = \sum_\alpha \langle \vec{w}_{\text{ideal}}^\alpha \vec{w}^\alpha(t) \rangle = N \left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle, \qquad (7) \quad \boxed{\texttt{eq:overlapex}}$$

If the elements of $\vec{w}_{\text{ideal}}$ take values ± 1 with probability $f^{\text{pot/dep}}$,

$$\left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle = f^{\text{pot}} \left\langle \vec{w}^1(t) \right\rangle_{\text{pot},t=0} - f^{\text{dep}} \left\langle \vec{w}^1(t) \right\rangle_{\text{dep},t=0} \qquad (8) \quad \boxed{\texttt{eq:overlapon}}$$

To compute this quantity, we will need to discuss how we model individual synapses.

1.3 Markov models of synapses

We model a synapse as having M internal states. The synaptic weight depends deterministically on the state, given by the M-element column vector $\mathbf{w}$. We denote the probability distribution across these states by the row vector $\mathbf{p}(t)$. We denote a column vector of ones by $\mathbf{e}$, so the normalisation condition is $\mathbf{p}(t)\mathbf{e} = 1$.

When the synapse is subjected to a plasticity event the internal state will change stochastically, described by the matrices $\mathbf{M}^{\mathrm{pot/dep}}$ whose i, j elements are the transition probabilities from state i to state j. In between these plasticity events the synapse will undergo stochastic transitions described by a matrix $\mathbf{H}$ whose i, j element is the transition rate from state i to state j for $i \neq j$ and $\mathbf{H}_{ii} = -\sum_{j \neq i} \mathbf{H}_{ij}$. This allows us to describe homeostatic plasticity as well as plasticity processes that occur over a period of time – $\mathbf{M}^{\mathrm{pot/dep}}$ puts the synapse into some state where the process $\mathbf{H}$ takes over and does the rest. The matrices $\mathbf{M}^{\mathrm{pot/dep}}$ are transition matrices of discrete time Markov processes, whereas $\mathbf{H}$ is the transition matrix of a continuous time Markov process. Their elements must satisfy the following constraints

$$\mathbf{M}_{ij}^{\mathrm{pot/dep}} \geq 0, \qquad \sum_j \mathbf{M}_{ij}^{\mathrm{pot/dep}} = 1,$$

$$\mathbf{H}_{i \neq j} \geq 0, \qquad \sum_j \mathbf{H}_{ij} = 0. \tag{9}$$

The upper bounds $\mathbf{M}_{ij}^{\mathrm{pot/dep}} \leq 1$ and $\mathbf{H}_{ii} \leq 0$ follow automatically from these. There is no upper bound on $\mathbf{H}_{i \neq j}$ (or lower bound on $\mathbf{H}_{ii}$) as these are rates rather than probabilities.

If we treat the off diagonal elements as the independent degrees of freedom, with the off diagonal elements determined by the row sum constraints (second column of (9)), the constraints are

$$\mathbf{M}_{i \neq j}^{\mathrm{pot/dep}} \geq 0, \qquad \sum_{j \neq i} \mathbf{M}_{ij}^{\mathrm{pot/dep}} \leq 1, \qquad \mathbf{H}_{i \neq j} \geq 0. \tag{10}$$

The second of these will be referred to as the "diagonal" constraint.

Given a sequence of plasticity events at times $t_1, t_2, t_3, \ldots$, the probability distribution across the internal states is

$$\mathbf{p}(t_n) = \mathbf{p}(t_0) e^{\mathbf{H}(t_1 - t_0)} \mathbf{M}^{\mathrm{pot}} e^{\mathbf{H}(t_2 - t_1)} \mathbf{M}^{\mathrm{dep}} e^{\mathbf{H}(t_3 - t_4)} \mathbf{M}^{\mathrm{dep}} \ldots \mathbf{M}^{\mathrm{pot}} e^{\mathbf{H}(t_n - t_{n-1})}. \tag{11}$$

In computing the expectation in (8) the plasticity event that occurs at $t = 0$ is fixed, but all of the previous and subsequent events are marginalised over. If the plasticity events occur at Poisson rate r, with potentiation and depression occurring independently with probability $f^{\mathrm{pot/dep}}$, averaging over all previous and subsequent events gives

$$P(\text{state} = i, t \mid \text{pot/dep}, 0) = \left[\mathbf{p}^\infty \mathbf{M}^{\mathrm{pot/dep}} e^{r(f^{\mathrm{pot}} \mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}} \mathbf{M}^{\mathrm{dep}} - \mathbf{I})t + \mathbf{H}t} \right]_i, \tag{12}$$

where $\mathbf{p}^\infty$ is the steady state distribution of the continuous time transition matrix that appears in the exponential:

$$\mathbf{p}^\infty (r \mathbf{W}^{\mathrm{F}} + \mathbf{H}) = 0, \qquad \text{where} \quad \mathbf{W}^{\mathrm{F}} = f^{\mathrm{pot}} \mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}} \mathbf{M}^{\mathrm{dep}} - \mathbf{I}. \tag{13}$$

This results in

$$\left\langle \vec{w}^1_{\mathrm{ideal}} \vec{w}^1(t) \right\rangle = \mathbf{p}^\infty (f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} - f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}})\,\mathrm{e}^{(r\mathbf{W}^{\mathrm{F}}+\mathbf{H})t}\mathbf{w},$$
$$\left\langle \vec{w}^1_{\mathrm{ideal}} \vec{w}^1(\infty) \right\rangle = \mathbf{p}^\infty (f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} - f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}})\,\mathbf{e}\mathbf{p}^\infty \mathbf{w}$$
$$= \mathbf{p}^\infty (f^{\mathrm{pot}}\mathbf{e} - f^{\mathrm{dep}}\mathbf{e})\,\mathbf{p}^\infty \mathbf{w} \tag{14}$$
$$= (f^{\mathrm{pot}} - f^{\mathrm{dep}})\,\mathbf{p}^\infty \mathbf{w}$$
$$= (f^{\mathrm{pot}} - f^{\mathrm{dep}})\,\mathbf{p}^\infty \mathrm{e}^{(r\mathbf{W}^{\mathrm{F}}+\mathbf{H})t}\mathbf{w}.$$

$\boxed{\texttt{eq:overlapwt}}$

Combining these allows us to write the numerator of (3) as

$$\left\langle \vec{w}_{\mathrm{ideal}}\cdot\vec{w}(t) \right\rangle - \left\langle \vec{w}_{\mathrm{ideal}}\cdot\vec{w}(\infty) \right\rangle = N\mathbf{p}^\infty \mathbf{K}\,\mathrm{e}^{(r\mathbf{W}^{\mathrm{F}}+\mathbf{H})t}\mathbf{w},$$
$$\text{where} \qquad \mathbf{K} = f^{\mathrm{pot}}(\mathbf{M}^{\mathrm{pot}} - \mathbf{I}) - f^{\mathrm{dep}}(\mathbf{M}^{\mathrm{dep}} - \mathbf{I}) \tag{15}$$

$\boxed{\texttt{eq:signal}}$

Combining with (6) gives

$$\mathrm{SNR}(t) = \frac{\sqrt{N}\mathbf{p}^\infty \mathbf{K}\,\mathrm{e}^{(r\mathbf{W}^{\mathrm{F}}+\mathbf{H})t}\mathbf{w}}{\sqrt{1 - (f^{\mathrm{pot}} - f^{\mathrm{dep}})^2(\mathbf{p}^\infty_+ - \mathbf{p}^\infty_-)^2}}. \tag{16}$$

$\boxed{\texttt{eq:SNRcurveE}}$

We can absorb $\mathbf{H}$ into $\mathbf{Q}^{\mathrm{pot/dep}}$ by defining

$$\widetilde{\mathbf{M}}^{\mathrm{pot/dep}} = \mathbf{M}^{\mathrm{pot/dep}} + \frac{\mathbf{H}}{2rf^{\mathrm{pot/dep}}}, \qquad \widetilde{\mathbf{W}}^{\mathrm{F}} = f^{\mathrm{pot}}\widetilde{\mathbf{M}}^{\mathrm{pot}} + f^{\mathrm{dep}}\widetilde{\mathbf{M}}^{\mathrm{dep}} - \mathbf{I},$$
$$\widetilde{\mathbf{K}} = f^{\mathrm{pot}}(\widetilde{\mathbf{M}}^{\mathrm{pot}} - \mathbf{I}) - f^{\mathrm{dep}}(\widetilde{\mathbf{M}}^{\mathrm{dep}} - \mathbf{I}). \tag{17}$$

$\boxed{\texttt{eq:absorbhom}}$

Then the SNR curve is

$$\mathrm{SNR}(t) = \frac{\sqrt{N}\mathbf{p}^\infty \widetilde{\mathbf{K}}\,\mathrm{e}^{rt\widetilde{\mathbf{W}}^{\mathrm{F}}}\mathbf{w}}{\sqrt{1 - (f^{\mathrm{pot}} - f^{\mathrm{dep}})^2(\mathbf{p}^\infty_+ - \mathbf{p}^\infty_-)^2}}. \tag{18}$$

$\boxed{\texttt{eq:SNRcurveh}}$

The denominator will not play any role in what follows, as the models that maximize the various measures of memory performance all have some sort of balance between potentiation and depression, either with $f^{\mathrm{pot}} = f^{\mathrm{dep}}$ or $\mathbf{p}^\infty_+ = \mathbf{p}^\infty_-$. We can perform the maximisation in two steps. First maximise the numerator at fixed $\mathbf{p}^\infty \mathbf{w}$. Then maximise the ratio wrt. $\mathbf{p}^\infty \mathbf{w}$. This will allow us to ignore the denominator until the very end.

This formula takes the same form as the one in our NIPS paper [1], where the process $\mathbf{H}$ was not included. However, the process $\mathbf{H}$ still has an impact due to the constraints.

The constraints on $\widetilde{\mathbf{M}}^{\mathrm{pot/dep}}$ are less stringent than the constraints on $\mathbf{M}^{\mathrm{pot/dep}}$, (9). Clearly

$$\widetilde{\mathbf{M}}^{\mathrm{pot/dep}}_{i\neq j} \geq 0, \qquad \sum_j \widetilde{\mathbf{M}}^{\mathrm{pot/dep}}_{ij} = 1. \tag{19}$$

$\boxed{\texttt{eq:constrhof}}$

The "diagonal" constraints are more complicated. There can be many $\mathbf{M}^{\mathrm{pot/dep}}, \mathbf{H}$ that result in the same $\widetilde{\mathbf{M}}^{\mathrm{pot/dep}}$. However, the fact that $\mathbf{M}^{\mathrm{pot/dep}}_{ij} \in [0,1]$ and $\mathbf{H}_{i\neq j} \geq 0$ imply that $\forall (i \neq j)$:

$$2r\max\left\{f^{\mathrm{pot}}(\widetilde{\mathbf{M}}^{\mathrm{pot}}_{ij} - 1), f^{\mathrm{dep}}(\widetilde{\mathbf{M}}^{\mathrm{dep}}_{ij} - 1)\right\} \leq \mathbf{H}_{ij} \leq 2r\min\left\{f^{\mathrm{pot}}\widetilde{\mathbf{M}}^{\mathrm{pot}}_{ij}, f^{\mathrm{dep}}\widetilde{\mathbf{M}}^{\mathrm{dep}}_{ij}\right\} \tag{20}$$

$\boxed{\texttt{eq:homrange}}$

5

These two bounds are consistent when

$$-f^{\mathrm{dep}} \leq f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} - f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} \leq f^{\mathrm{pot}} \qquad \forall i \neq j. \tag{21}$$

`eq:homconsis`

To ensure that $\mathbf{M}_{ii}^{\mathrm{pot/dep}} \geq 0$, we also need to ensure that $\sum_{j \neq i} \mathbf{M}_{ij}^{\mathrm{pot/dep}} \leq 1$.

$$
\begin{aligned}
\sum_{j \neq i} \mathbf{M}_{ij}^{\mathrm{pot}} &= \sum_{j \neq i} \left(\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} - \frac{\mathbf{H}_{ij}}{2r\,f^{\mathrm{pot}}} \right) \\
&\geq \sum_{j \neq i} \left(\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} - \frac{1}{f^{\mathrm{pot}}} \min \left\{ f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}, f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} \right\} \right) \\
&= \frac{1}{f^{\mathrm{pot}}} \sum_{j \neq i} \left[f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} - f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} \right]_+ , \\
\sum_{j \neq i} \mathbf{M}_{ij}^{\mathrm{dep}} &= \frac{1}{f^{\mathrm{dep}}} \sum_{j \neq i} \left[f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} - f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} \right]_+ ,
\end{aligned}
\tag{22}
$$

`eq:homdiag`

where rectification function is $[x]_\pm = x\Theta(\pm x)$, in terms of the Heaviside step function. Demanding that these quantities are at most 1 is sufficient to ensure that (21) is satisfied.

Therefore we find the set of constraints

$$
\begin{aligned}
\widetilde{\mathbf{M}}_{i \neq j}^{\mathrm{pot/dep}} &\geq 0, & \sum_{j \neq i} \left[f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} - f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} \right]_+ &\leq f^{\mathrm{pot}}, \\
\sum_j \widetilde{\mathbf{M}}_{ij}^{\mathrm{pot/dep}} &= 1, & \sum_{j \neq i} \left[f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} - f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}} \right]_+ &\leq f^{\mathrm{dep}}.
\end{aligned}
\tag{23}
$$

`eq:constrh`

The second column of this will be referred to as the "diagonal" constraints.

We showed that these are necessary conditions for the existence of $\mathbf{M}^{\mathrm{pot/dep}}, \mathbf{H}$ satisfying (9). We can see that they are sufficient by choosing $\mathbf{H}_{ij} = 2r \min \left\{ f^{\mathrm{pot}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}, f^{\mathrm{dep}}\widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}} \right\}$ for $i \neq j$ and $\mathbf{H}_{ii} = -\sum_{j \neq i} \mathbf{H}_{ij}$.

We can then treat the off-diagonal elements of $\widetilde{\mathbf{M}}^{\mathrm{pot/dep}}$ as the independent degrees of freedom, subject to these constraints, when maximising the SNR. The diagonal elements are determined by setting the row sums to 1. We then have

$$\frac{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}}{\partial \widetilde{\mathbf{M}}_{mn}^{\mathrm{pot}}} = \frac{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}}}{\partial \widetilde{\mathbf{M}}_{mn}^{\mathrm{dep}}} = \delta_{im}(\delta_{jn} - \delta_{jm}), \qquad \frac{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}}{\partial \widetilde{\mathbf{M}}_{mn}^{\mathrm{dep}}} = \frac{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}}}{\partial \widetilde{\mathbf{M}}_{mn}^{\mathrm{pot}}} = 0. \tag{24}$$

`eq:derivpd`

It will also be convenient to define the differential operators

$$\frac{\partial}{\partial \widetilde{\mathbf{W}}_{ij}^{\mathrm{F}}} = \frac{1}{2f^{\mathrm{pot}}} \frac{\partial}{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}} + \frac{1}{2f^{\mathrm{dep}}} \frac{\partial}{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}}}, \qquad \frac{\partial}{\partial \widetilde{\mathbf{K}}_{ij}} = \frac{1}{2f^{\mathrm{pot}}} \frac{\partial}{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{pot}}} - \frac{1}{2f^{\mathrm{dep}}} \frac{\partial}{\partial \widetilde{\mathbf{M}}_{ij}^{\mathrm{dep}}}, \tag{25}$$

`eq:pertfe`

these satisfy

$$\frac{\partial \widetilde{\mathbf{W}}_{ij}^{\mathrm{F}}}{\partial \widetilde{\mathbf{W}}_{mn}^{\mathrm{F}}} = \frac{\partial \widetilde{\mathbf{K}}_{ij}}{\partial \widetilde{\mathbf{K}}_{mn}} = \delta_{im}(\delta_{jn} - \delta_{jm}), \qquad \frac{\partial \widetilde{\mathbf{W}}_{ij}^{\mathrm{F}}}{\partial \widetilde{\mathbf{K}}_{mn}} = \frac{\partial \widetilde{\mathbf{K}}_{ij}}{\partial \widetilde{\mathbf{W}}_{mn}^{\mathrm{F}}} = 0. \tag{26}$$

`eq:derivfe`

6

2 Laplace transform

2.1 Fundamental matrix etc.

In analogy to the generalised fundamental matrix of a Markov chain [2], define

$$\mathbf{Z}(s) = (s\mathbf{I} + \mathbf{e}\boldsymbol{\pi} - \mathbf{Q})^{-1}, \tag{27}$$

where $\boldsymbol{\pi}$ is an arbitrary row vector satisfying $\boldsymbol{\pi}\mathbf{e} = \frac{1}{\tau_0} \neq 0$. This reduces to the fundamental matrix at $s = 0$. We can see that $\mathbf{e}$ is an eigenvector:

$$\mathbf{Z}(s)\mathbf{e} = \frac{\tau_0}{1 + s\tau_0}\mathbf{e}. \tag{28}$$

Suppose we have some row vector $\mathbf{a}$ such that $\mathbf{a}\mathbf{e} = 0$. Then $\mathbf{a}\mathbf{Z}(s)$ is independent of $\boldsymbol{\pi}$:

$$\frac{\partial(\mathbf{a}\mathbf{Z}(s))_i}{\partial\boldsymbol{\pi}_k} = (\mathbf{a}\mathbf{Z}(s)\mathbf{e})\mathbf{Z}_{ki}(s) = \frac{\tau_0}{1 + s\tau_0}(\mathbf{a}\mathbf{e})\mathbf{Z}_{ki}(s) = 0. \tag{29}$$

Then we also define

$$\overline{\mathbf{T}}(s) = (\mathbf{E}\mathbf{Z}^{\mathrm{dg}}(s) - \mathbf{Z}(s))\mathbf{D}, \qquad \overline{\mathbf{T}}_{ij}(s) = \frac{\mathbf{Z}_{jj}(s) - \mathbf{Z}_{ij}(s)}{\mathbf{p}_j^\infty}. \tag{30}$$

This reduces to the first passage time matrix at $s = 0$, but at other s there is (probably) no relation.

Using the exact same proof as in the case $s = 0$, we see that the quantity

$$\eta(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty \tag{31}$$

is independent of the starting state i, but now $\eta(s) = \operatorname{tr}\mathbf{Z}(s) - \frac{\tau}{1+s\tau}$. We can the define

$$\eta_i^\pm(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty\left(\frac{1 \pm \mathbf{w}_j}{2}\right) = \sum_{j\in\pm} \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty. \tag{32}$$

We can arrange the states in order of decreasing $\eta_i^+(s)$ or increasing $\eta_i^-(s)$. It will be more convenient to use the quantities

$$\eta_i^w(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty\mathbf{w} = \eta_i^+(s) - \eta_i^-(s) = 2\eta_i^+(s) - \eta(s) = \eta(s) - 2\eta_i^-(s). \tag{33}$$

Arranging the states in order of decreasing $\eta_i^w(s)$ is the same as the order of decreasing $\eta_i^+(s)$ or increasing $\eta_i^-(s)$. These quantities can still be used when $\mathbf{w}$ takes more than two values.

2.2 Laplace transform of SNR curve

Consider the Laplace transform of the evolution operator:

$$\mathbf{G}(s) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{(\mathbf{Q}-s)t}. \tag{34}$$

For $\mathfrak{Re}\, s > 0$, we have

$$(s - \mathbf{Q})\mathbf{G}(s) = \int_0^\infty \mathrm{d}t\, (s - \mathbf{Q})\mathrm{e}^{(\mathbf{Q}-s)t} = \left[-\mathrm{e}^{(\mathbf{Q}-s)t}\right]_0^\infty = \mathbf{I}. \tag{35}$$

As $(s - \mathbf{Q})$ is invertible for $\mathfrak{Re}\, s > 0$, because the real part of all eigenvalues of $\mathbf{Q}$ are nonpositive, we have

$$\mathbf{G}(s) = (s - \mathbf{Q})^{-1}. \tag{36}$$

For $s = 0$, we can avoid problems by replacing $\mathbf{G}(s) \to \mathbf{Z}(s)$.

Now consider the Laplace transform of the SNR curve (18) (ignoring the denominator)

$$
\begin{aligned}
A(s) &= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st}\, \mathrm{SNR}(t) \\
&= \int_0^\infty \mathrm{d}t\, \sqrt{N}\mathbf{p}^\infty \widetilde{\mathbf{K}} \mathrm{e}^{(r\widetilde{\mathbf{W}}^{\mathrm{F}}-s)t}\mathbf{w} \\
&= \sqrt{N}\mathbf{p}^\infty \widetilde{\mathbf{K}}\left(s - r\widetilde{\mathbf{W}}^{\mathrm{F}}\right)^{-1}\mathbf{w}.
\end{aligned}
\tag{37}
$$

This expression is ill-behaved at $s = 0$. Thanks to (29), we can solve this by the replacement $\mathbf{G}(s) \to \mathbf{Z}(s)$, as $\widetilde{\mathbf{K}}\mathbf{e} = 0$.

$$
\begin{aligned}
\hat{A}(s) = \frac{A(s)}{\sqrt{N}} &= \mathbf{p}^\infty \widetilde{\mathbf{K}}\mathbf{Z}(s)\mathbf{w} \\
&= \sum_{ijk} \mathbf{p}_i^\infty \widetilde{\mathbf{K}}_{ij}\left(\overline{\mathbf{T}}_{ik}(s) - \overline{\mathbf{T}}_{jk}(s)\right)\mathbf{p}_k^\infty \mathbf{w}_k \\
&= \sum_{ij} \mathbf{p}_i^\infty \widetilde{\mathbf{K}}_{ij}(\eta_i^w(s) - \eta_j^w(s)).
\end{aligned}
\tag{38}
$$

Note that $A(0) = A$, the area, and $\lim_{s\to\infty}\{sA(s)\} = \mathrm{SNR}(0)$, the initial SNR.

In analogy with the case $s = 0$, we define

$$
\begin{aligned}
c_k(s) &= \sum_{ij} \mathbf{p}_i^\infty \widetilde{\mathbf{K}}_{ij}\left(\overline{\mathbf{T}}_{ik}(s) - \overline{\mathbf{T}}_{jk}(s)\right) \\
a_i(s) &= \sum_j \mathbf{p}_i^\infty \widetilde{\mathbf{K}}_{ij}(\eta_i^w(s) - \eta_j^w(s)), \\
\implies \hat{A}(s) &= \sum_k c_k(s)\mathbf{p}_k^\infty \mathbf{w}_k = \sum_i a_i(s).
\end{aligned}
\tag{39}
$$

8

2.3 Derivatives

Using the same approach as the area bound in [1], we find that the derivatives are

$$\frac{\partial \hat{A}(s)}{\partial \widetilde{\mathbf{M}}_{mn}^{\text{pot/dep}}} = r f^{\text{pot/dep}} \mathbf{p}_m^\infty \left[\sum_i (\overline{\mathbf{T}}_{mi}(0) - \overline{\mathbf{T}}_{ni}(0)) a_i(s) + c_m(s)(\eta_m^w(s) - \eta_n^w(s)) \right] \tag{40}$$
$$\pm f^{\text{pot/dep}} \mathbf{p}_m^\infty (\eta_m^w(s) - \eta_n^w(s)).$$

3 Upper bounds

3.1 Area bound

$$\int [_0^\infty] dt \, \text{SNR}(t) = \lim_{\tau \to \infty} \tau \overline{\text{SNR}}(\tau) = A(0). \tag{41}$$

The only effect of $\mathbf{H}$ is in the diagonal constraints (second column of (23) vs. second inequality in (10)). However, the area is invariant under the scaling mode:

$$\widetilde{\mathbf{M}}^{\text{pot/dep}} - \mathbf{I} \to \lambda(\widetilde{\mathbf{M}}^{\text{pot/dep}} - \mathbf{I}). \tag{42}$$

This means we can always take a set of matrices that violate either version of the diagonal constraints and use (42) to create a set of matrices that satisfy them without changing the area. Therefore, including $\mathbf{H}$ has no effect on the area bound and the result from [1] applies here as well.

This means that

$$A(0) \leq \frac{\sqrt{N}(M-1)}{r}. \tag{43}$$

The model that saturates this bound is one of serial topology with $\mathbf{p}^\infty$ concentrated symmtrically at the two end states.

3.2 Initial SNR

$$\text{SNR}(0) = \overline{\text{SNR}}(0) = \lim_{s \to \infty} s A(s). \tag{44}$$

$$\text{SNR}(0) \leq \sqrt{N}, \tag{45}$$

at least when $f^{\text{pot}} = f^{\text{dep}} = \frac{1}{2}$.

3.3 Envelope

Now we introduce an eigenvector decomposition for $-\widetilde{\mathbf{W}}^{\text{F}}$:

$$\widetilde{\mathbf{W}}^{\text{F}} = -\sum_a q_a \mathbf{u}^a \boldsymbol{\eta}^a, \quad \boldsymbol{\eta}^a \mathbf{u}^b = \delta_{ab}, \quad \widetilde{\mathbf{W}}^{\text{F}} \mathbf{u}^a = -q_a \mathbf{u}^a, \quad \boldsymbol{\eta}^a \widetilde{\mathbf{W}}^{\text{F}} = -q_a \boldsymbol{\eta}^a. \tag{46}$$

Then we can write

$$\text{SNR}(t) = \sqrt{N} \sum_a \mathcal{I}_a e^{-rt/\tau_a}, \qquad \overline{\text{SNR}}(\tau) = \sqrt{N} \sum_a \frac{\mathcal{I}_a}{1 + r\tau/\tau_a}, \qquad (47) \quad \boxed{\texttt{eqsnreig}}$$

where

$$\mathcal{I}_a = (\mathbf{p}^\infty \widetilde{\mathbf{K}} \mathbf{u}^a)(\boldsymbol{\eta}^a \mathbf{w}), \qquad \tau_a = \frac{1}{q_a}. \qquad (48) \quad \boxed{\texttt{eq:snrcoeffs}}$$

From (43) and (45), we see that they are subject to the constraints

$$\sum_a \mathcal{I}_a \tau_a \leq M - 1, \qquad \sum_a \mathcal{I}_a \leq 1. \qquad (49) \quad \boxed{\texttt{eq:coeffcons}}$$

No doubt there are many other constraints that these quantities must satisfy, but we can see what these ones imply for the $\overline{\text{SNR}}(\tau)$ by maximising it subject to these constraints.

Consider the Lagrangian

$$\mathcal{L} = \sum_a \frac{\mathcal{I}_a}{1 + r\tau/\tau_a} + \mu_\mathcal{I} \left(1 - \sum_a \mathcal{I}_a \right) + \mu_\mathcal{A} \left((M-1) - \sum_a \mathcal{I}_a \tau_a \right). \qquad (50) \quad \boxed{\texttt{eq:envlagran}}$$

The Karush-Kuhn-Tucker necessary conditions for a maximum are

$$\frac{\partial \mathcal{L}}{\partial \mathcal{I}_a} = 0, \qquad \frac{\partial \mathcal{L}}{\partial \tau_a} = 0, \qquad \mu_\alpha \geq 0, \qquad \frac{\partial \mathcal{L}}{\partial \mu_\alpha} \geq 0, \qquad \mu_\alpha \frac{\partial \mathcal{L}}{\partial \mu_\alpha} = 0. \qquad (51) \quad \boxed{\texttt{eq:envKTcond}}$$

These are solved by

$$\mathcal{I}_1 = 1, \qquad \mathcal{I}_{a>1} = 0, \qquad \tau_1 = M - 1,$$
$$\mu_\mathcal{I} = \frac{(M-1)^2}{(r\tau + (M-1))^2}, \qquad \mu_\mathcal{A} = \frac{r\tau}{(r\tau + (M-1))^2}. \qquad (52) \quad \boxed{\texttt{eq:envsol}}$$

This leads to the envelope

$$\overline{\text{SNR}}(\tau) \leq \frac{\sqrt{N}(M-1)}{r\tau + (M-1)}. \qquad (53) \quad \boxed{\texttt{eq:env}}$$

This is the envelope ("memory frontier") that we can prove. Finding more constraints would lower it. Clearly there have to be more constraints, as (52) indicates that this would be achieved by the same model at all times. This model would have to saturate both the initial SNR bound (45) and the area bound (43), whereas we saw that they were achieved by different models, which perform well at short/long timescales respectively.

References

$\boxed{\texttt{2013synapse}}$ [1] S. Lahiri and S. Ganguli, "A memory frontier for complex synapses," in *Advances in Neural Information Processing Systems 26*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, eds., pp. 1034–1042. NIPS, 2013.

$\boxed{\texttt{ny1981fund}}$ [2] J. G. Kemeny, "Generalization of a fundamental matrix," *Linear Algebra and its Applications* **38** (1981) no. 0, 193 – 206.