Laplacian envelope

Subhaneil Lahiri

June 15, 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 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:

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

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

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

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

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

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

$$\mathrm{SNR}(t) = \frac{\langle \vec{w}_{\mathrm{ideal}} \cdot \vec{w}(t) \rangle - \langle \vec{w}_{\mathrm{ideal}} \cdot \vec{w}(\infty) \rangle}{\sqrt{\mathrm{Var}(\vec{w}_{\mathrm{ideal}} \cdot \vec{w}(\infty))}}, \tag{3}$$

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, $\vec{w}_{\mathrm{ideal}}$. If we also average over the time between storage and recognition, we have

$$\overline{\mathrm{SNR}}(\tau) = \int_0^\infty \mathrm{d}t\, P(t|\tau)\, \mathrm{SNR}(t) = \frac{1}{\tau} \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t/\tau}\, \mathrm{SNR}(t), \tag{4}$$

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):

$$A(s) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st}\, \mathrm{SNR}(t). \qquad \overline{\mathrm{SNR}}(\tau) = \frac{A(1/\tau)}{\tau}, \tag{5}$$

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

$$
\begin{aligned}
\mathrm{Var}(\vec{w}_{\mathrm{ideal}} \cdot \vec{w}(t)) &= \sum_{\alpha\beta} \left\langle \vec{w}_{\mathrm{ideal}}^\alpha \vec{w}^\alpha(t) \vec{w}_{\mathrm{ideal}}^\beta \vec{w}^\beta(t) \right\rangle - \left(\sum_\alpha \langle \vec{w}_{\mathrm{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \right)^2 \\
&= \sum_\alpha \left\langle (\vec{w}_{\mathrm{ideal}}^\alpha)^2 (\vec{w}^\alpha(t))^2 \right\rangle + \sum_{\alpha \neq \beta} \langle \vec{w}_{\mathrm{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \langle \vec{w}_{\mathrm{ideal}}^\beta \vec{w}^\beta(t) \rangle \\
&\qquad\qquad - \left(\sum_\alpha \langle \vec{w}_{\mathrm{ideal}}^\alpha \vec{w}^\alpha(t) \rangle \right)^2 \\
&= N \langle 1 \rangle + N(N-1) \left\langle \vec{w}_{\mathrm{ideal}}^1 \vec{w}^1(t) \right\rangle^2 - N^2 \left\langle \vec{w}_{\mathrm{ideal}}^1 \vec{w}^1(t) \right\rangle^2 \\
&= N(1 - \left\langle \vec{w}_{\mathrm{ideal}}^1 \vec{w}^1(t) \right\rangle^2),
\end{aligned} \tag{6}
$$

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

For the numerator, we can write

$$\langle \vec{w}_{\mathrm{ideal}} \cdot \vec{w}(t) \rangle = \sum_\alpha \langle \vec{w}_{\mathrm{ideal}}^\alpha \vec{w}^\alpha(t) \rangle = N \left\langle \vec{w}_{\mathrm{ideal}}^1 \vec{w}^1(t) \right\rangle, \tag{7}$$

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

$$\left\langle \vec{w}_{\mathrm{ideal}}^1 \vec{w}^1(t) \right\rangle = f^{\mathrm{pot}} \left\langle \vec{w}^1(t) \right\rangle_{\mathrm{pot},t=0} - f^{\mathrm{dep}} \left\langle \vec{w}^1(t) \right\rangle_{\mathrm{dep},t=0} \tag{8}$$

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

3

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. The matrices $\mathbf{M}^{\mathrm{pot/dep}}$ are transition matrices of discrete time Markov processes. Their elements must satisfy the following constraints

$$\mathbf{M}_{ij}^{\mathrm{pot/dep}} \geq 0, \qquad \sum_j \mathbf{M}_{ij}^{\mathrm{pot/dep}} = 1. \tag{9}$$

The upper bounds $\mathbf{M}_{ij}^{\mathrm{pot/dep}} \leq 1$ follow automatically from these.

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

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

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 at time t_n is

$$\mathbf{p}(t_n) = \mathbf{p}(t_0)\mathbf{M}^{\mathrm{pot}}\mathbf{M}^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}}\mathbf{M}^{\mathrm{pot}}\mathbf{M}^{\mathrm{dep}}\mathbf{M}^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}}\mathbf{M}^{\mathrm{dep}} \ldots \mathbf{M}^{\mathrm{pot}}. \tag{11}$$

If the probability of a given plasticity event being poetitiating/depressing os $f^{\mathrm{pot/dep}}$, the probability of the sequence above (conditioned on the number of events) is given by $f^{\mathrm{pot}} f^{\mathrm{dep}} f^{\mathrm{dep}} f^{\mathrm{pot}} f^{\mathrm{dep}} f^{\mathrm{pot}} f^{\mathrm{pot}} f^{\mathrm{dep}} \ldots f^{\mathrm{pot}}$.

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 we marginalise over the types (potentiating/depressing) of these other events (conditioned on the number of events) we find

$$\mathbf{p}(t_n) = \mathbf{p}(t_0)(f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}}) \ldots (f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}})$$
$$\mathbf{M}^{\mathrm{pot/dep}}(f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}}) \ldots (f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}}), \tag{12}$$

where the factor of $\mathbf{M}^{\mathrm{pot/dep}}$ is the event at $t = 0$. If we expand out all of the parenthetical factors, we will get a sum of terms of the form (11) weighted by their probabilities.

If $t_0 \ll 0$, the effect of the large number of earlier events will be to put the synapse in the steady state distribution, $\mathbf{p}^{\infty}$, of the stochastic process $(f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}})$:

$$\mathbf{p}(t_n) = \mathbf{p}^{\infty}\mathbf{M}^{\mathrm{pot/dep}}(f^{\mathrm{pot}}\mathbf{M}^{\mathrm{pot}} + f^{\mathrm{dep}}\mathbf{M}^{\mathrm{dep}})^n, \tag{13}$$

where n is the number of plasticity events that took place.

4

If the plasticity events occur at Poisson rate r, the probability of n events occurring in time t is given by a Poisson distribution with mean rt. If we marginalise over n as well

$$
\begin{aligned}
P(\text{state} = i, t \mid \text{pot/dep}, 0) &= \sum_n \frac{(rt)^n}{n!} \mathrm{e}^{-rt} \mathbf{p}^\infty \mathbf{M}^{\text{pot/dep}} (f^{\text{pot}} \mathbf{M}^{\text{pot}} + f^{\text{dep}} \mathbf{M}^{\text{dep}})^n \\
&= \left[\mathbf{p}^\infty \mathbf{M}^{\text{pot/dep}} \, \mathrm{e}^{rt(f^{\text{pot}} \mathbf{M}^{\text{pot}} + f^{\text{dep}} \mathbf{M}^{\text{dep}} - \mathbf{I})} \right]_i,
\end{aligned}
\tag{14}
$$
`eq:plastseqa`

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

$$
\mathbf{p}^\infty \mathbf{W}^{\text{F}} = 0, \qquad \text{where} \quad \mathbf{W}^{\text{F}} = f^{\text{pot}} \mathbf{M}^{\text{pot}} + f^{\text{dep}} \mathbf{M}^{\text{dep}} - \mathbf{I}.
\tag{15}
$$
`eq:eq`

This results in

$$
\begin{aligned}
\left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(t) \right\rangle &= \mathbf{p}^\infty (f^{\text{pot}} \mathbf{M}^{\text{pot}} - f^{\text{dep}} \mathbf{M}^{\text{dep}}) \, \mathrm{e}^{rt\mathbf{W}^{\text{F}}} \mathbf{w}, \\
\left\langle \vec{w}_{\text{ideal}}^1 \vec{w}^1(\infty) \right\rangle &= \mathbf{p}^\infty (f^{\text{pot}} \mathbf{M}^{\text{pot}} - f^{\text{dep}} \mathbf{M}^{\text{dep}}) \, \mathbf{e} \mathbf{p}^\infty \mathbf{w} \\
&= \mathbf{p}^\infty (f^{\text{pot}} \mathbf{e} - f^{\text{dep}} \mathbf{e}) \, \mathbf{p}^\infty \mathbf{w} \\
&= (f^{\text{pot}} - f^{\text{dep}}) \, \mathbf{p}^\infty \mathbf{w} \\
&= (f^{\text{pot}} - f^{\text{dep}}) \, \mathbf{p}^\infty \mathrm{e}^{rt\mathbf{W}^{\text{F}}} \mathbf{w}.
\end{aligned}
\tag{16}
$$
`eq:overlapwt`

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

$$
\begin{aligned}
\langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) \rangle - \langle \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle &= N \mathbf{p}^\infty \mathbf{K} \, \mathrm{e}^{(r\mathbf{W}^{\text{F}} + \mathbf{H})t} \mathbf{w}, \\
\text{where} \qquad \mathbf{K} &= f^{\text{pot}} (\mathbf{M}^{\text{pot}} - \mathbf{I}) - f^{\text{dep}} (\mathbf{M}^{\text{dep}} - \mathbf{I})
\end{aligned}
\tag{17}
$$
`eq:signal`

Combining with (6) gives

$$
\text{SNR}(t) = \frac{\sqrt{N} \mathbf{p}^\infty \mathbf{K} \, \mathrm{e}^{(r\mathbf{W}^{\text{F}} + \mathbf{H})t} \mathbf{w}}{\sqrt{1 - (f^{\text{pot}} - f^{\text{dep}})^2 (\mathbf{p}^\infty \mathbf{w})^2}}.
\tag{18}
$$
`eq:SNRcurveE`

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

$$
\widetilde{\mathbf{M}}^{\text{pot/dep}} = \mathbf{M}^{\text{pot/dep}} + \frac{\mathbf{H}}{2r f^{\text{pot/dep}}}, \qquad \widetilde{\mathbf{W}}^{\text{F}} = f^{\text{pot}} \widetilde{\mathbf{M}}^{\text{pot}} + f^{\text{dep}} \widetilde{\mathbf{M}}^{\text{dep}} - \mathbf{I},
$$
$$
\widetilde{\mathbf{K}} = f^{\text{pot}} (\widetilde{\mathbf{M}}^{\text{pot}} - \mathbf{I}) - f^{\text{dep}} (\widetilde{\mathbf{M}}^{\text{dep}} - \mathbf{I}).
\tag{19}
$$
`eq:absorbhom`

Then the SNR curve is

$$
\text{SNR}(t) = \frac{\sqrt{N} \mathbf{p}^\infty \widetilde{\mathbf{K}} \, \mathrm{e}^{rt\widetilde{\mathbf{W}}^{\text{F}}} \mathbf{w}}{\sqrt{1 - (f^{\text{pot}} - f^{\text{dep}})^2 (\mathbf{p}^\infty \mathbf{w})^2}}.
\tag{20}
$$
`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 $f^{\text{pot}} = f^{\text{dep}}$ or $\mathbf{p}_+^\infty = \mathbf{p}_-^\infty$. We can perform the

5

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.

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

$$\frac{\partial \mathbf{M}^{\mathrm{pot}}_{ij}}{\partial \mathbf{M}^{\mathrm{pot}}_{mn}} = \frac{\partial \mathbf{M}^{\mathrm{dep}}_{ij}}{\partial \mathbf{M}^{\mathrm{dep}}_{mn}} = \delta_{im}(\delta_{jn} - \delta_{jm}), \qquad \frac{\partial \mathbf{M}^{\mathrm{pot}}_{ij}}{\partial \mathbf{M}^{\mathrm{dep}}_{mn}} = \frac{\partial \mathbf{M}^{\mathrm{dep}}_{ij}}{\partial \mathbf{M}^{\mathrm{pot}}_{mn}} = 0. \qquad (21) \quad \boxed{\texttt{eq:derivpd}}$$

It will also be convenient to define the differential operators

$$\frac{\partial}{\partial \mathbf{W}^{\mathrm{F}}_{ij}} = \frac{1}{2f^{\mathrm{pot}}}\frac{\partial}{\partial \mathbf{M}^{\mathrm{pot}}_{ij}} + \frac{1}{2f^{\mathrm{dep}}}\frac{\partial}{\partial \mathbf{M}^{\mathrm{dep}}_{ij}}, \qquad \frac{\partial}{\partial \mathbf{K}_{ij}} = \frac{1}{2f^{\mathrm{pot}}}\frac{\partial}{\partial \mathbf{M}^{\mathrm{pot}}_{ij}} - \frac{1}{2f^{\mathrm{dep}}}\frac{\partial}{\partial \mathbf{M}^{\mathrm{dep}}_{ij}}, \qquad (22) \quad \boxed{\texttt{eq:pertfe}}$$

these satisfy

$$\frac{\partial \mathbf{W}^{\mathrm{F}}_{ij}}{\partial \mathbf{W}^{\mathrm{F}}_{mn}} = \frac{\partial \mathbf{K}_{ij}}{\partial \mathbf{K}_{mn}} = \delta_{im}(\delta_{jn} - \delta_{jm}), \qquad \frac{\partial \mathbf{W}^{\mathrm{F}}_{ij}}{\partial \mathbf{K}_{mn}} = \frac{\partial \mathbf{K}_{ij}}{\partial \mathbf{W}^{\mathrm{F}}_{mn}} = 0. \qquad (23) \quad \boxed{\texttt{eq:derivfe}}$$

If we wish to use the off-diagonal elements of $\mathbf{W}^{\mathrm{F}}$ and $\mathbf{K}$ as the degrees of freedom, the constraints are

$$\mathbf{W}^{\mathrm{F}}_{i\neq j} \geq |\mathbf{K}_{i\neq j}|, \qquad \sum_{j\neq i} \mathbf{W}^{\mathrm{F}}_{ij} \pm |\mathbf{K}_{ij}| \leq f^{\mathrm{pot/dep}},$$
$$\sum_j \mathbf{W}^{\mathrm{F}}_{ij} = 0, \qquad \sum_j \mathbf{K}_{ij} = 0. \qquad (24) \quad \boxed{\texttt{eq:constrfe}}$$

2 Laplace transform

2.1 Fundamental matrix etc.

In analogy to the generalised fundamental matrix of a Markov chain [1], define

$$\mathbf{Z}(s) = (s\mathbf{I} + \mathbf{e}\boldsymbol{\pi} - \mathbf{Q})^{-1}, \qquad (25) \quad \boxed{\texttt{eq:lfund}}$$

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:

$$\mathbf{Z}(s)\mathbf{e} = \frac{\tau_0}{1 + s\tau_0}\mathbf{e}. \qquad (26) \quad \boxed{\texttt{eq:lfundrows}}$$

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

$$\frac{\partial (\mathbf{a}\mathbf{Z}(s))_i}{\partial \boldsymbol{\pi}_k} = -(\mathbf{a}\mathbf{Z}(s)\mathbf{e})\mathbf{Z}_{ki}(s) = -\frac{\tau_0}{1 + s\tau_0}(\mathbf{a}\mathbf{e})\mathbf{Z}_{ki}(s) = 0. \qquad (27) \quad \boxed{\texttt{eq:alfund}}$$

6

Then we also define

$$\overline{\mathbf{T}}(s) = (\mathbf{E}\mathbf{Z}^{\mathrm{dg}}(s) - \mathbf{Z}(s))\mathbf{D}, \qquad \overline{\mathbf{T}}_{ij}(s) = \frac{\mathbf{Z}_{jj}(s) - \mathbf{Z}_{ij}(s)}{\mathbf{p}_j^\infty}. \qquad (28) \quad \boxed{\texttt{eq:lfptb}}$$

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

$$\eta(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty \qquad (29) \quad \boxed{\texttt{eq:lkemeny}}$$

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

$$\eta_i^\pm(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty \left(\frac{1 \pm \mathbf{w}_j}{2} \right) = \sum_{j \in \pm} \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty. \qquad (30) \quad \boxed{\texttt{eq:lpmixtime}}$$

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

$$\eta_i^w(s) = \sum_j \overline{\mathbf{T}}_{ij}(s)\mathbf{p}_j^\infty \mathbf{w} = \eta_i^+(s) - \eta_i^-(s) = 2\eta_i^+(s) - \eta(s) = \eta(s) - 2\eta_i^-(s). \qquad (31) \quad \boxed{\texttt{eq:lwpmixtim}}$$

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:

$$\mathbf{G}(s) = \int_0^\infty \mathrm{d}t \, \mathrm{e}^{(\mathbf{Q}-s)t}. \qquad (32) \quad \boxed{\texttt{eq:levol}}$$

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

$$(s - \mathbf{Q})\mathbf{G}(s) = \int_0^\infty \mathrm{d}t \, (s - \mathbf{Q})\mathrm{e}^{(\mathbf{Q}-s)t} = \left[-\mathrm{e}^{(\mathbf{Q}-s)t} \right]_0^\infty = \mathbf{I}. \qquad (33) \quad \boxed{\texttt{eq:levolcalc}}$$

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

$$\mathbf{G}(s) = (s - \mathbf{Q})^{-1}. \qquad (34) \quad \boxed{\texttt{eq:levolres}}$$

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 (20) (ignoring the denominator)

$$
\begin{aligned}
A(s) &= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st}\,\mathrm{SNR}(t) \\
&= \int_0^\infty \mathrm{d}t\, \frac{\sqrt{N}\mathbf{p}^\infty \mathbf{K} \mathrm{e}^{(r\mathbf{W}^{\mathrm{F}}-s)t}\mathbf{w}}{\sqrt{1-(f^{\mathrm{pot}}-f^{\mathrm{dep}})^2(\mathbf{p}^\infty\mathbf{w})^2}} \qquad (35) \quad \boxed{\texttt{eq:lsnr}} \\
&= \frac{\sqrt{N}\mathbf{p}^\infty \mathbf{K} \left(s-r\mathbf{W}^{\mathrm{F}}\right)^{-1}\mathbf{w}}{\sqrt{1-(f^{\mathrm{pot}}-f^{\mathrm{dep}})^2(\mathbf{p}^\infty\mathbf{w})^2}}.
\end{aligned}
$$

This expression is ill-behaved at $s=0$. Thanks to (27), we can solve this by the replacement $\mathbf{G}(s)\to\mathbf{Z}(s)$, as $\mathbf{Ke}=0$.

$$
\begin{aligned}
\hat{A}(s) = A(s)\sqrt{\frac{1-(f^{\mathrm{pot}}-f^{\mathrm{dep}})^2(\mathbf{p}^\infty\mathbf{w})^2}{N}} &= \mathbf{p}^\infty\mathbf{K}\mathbf{Z}(s)\mathbf{w} \\
&= \sum_{ijk}\mathbf{p}_i^\infty\mathbf{K}_{ij}\left(\overline{\mathbf{T}}_{ik}(s)-\overline{\mathbf{T}}_{jk}(s)\right)\mathbf{p}_k^\infty\mathbf{w}_k \qquad (36) \quad \boxed{\texttt{eq:larea}} \\
&= \sum_{ij}\mathbf{p}_i^\infty\mathbf{K}_{ij}(\eta_i^w(s)-\eta_j^w(s)).
\end{aligned}
$$

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

$$
\begin{aligned}
c_k(s) &= \sum_{ij}\mathbf{p}_i^\infty\mathbf{K}_{ij}\left(\overline{\mathbf{T}}_{ik}(s)-\overline{\mathbf{T}}_{jk}(s)\right) \\
a_i(s) &= \sum_j \mathbf{p}_i^\infty\mathbf{K}_{ij}(\eta_i^w(s)-\eta_j^w(s)), \qquad (37) \quad \boxed{\texttt{eq:lareacoef}} \\
\implies \hat{A}(s) &= \sum_k c_k(s)\mathbf{p}_k^\infty\mathbf{w}_k = \sum_i a_i(s).
\end{aligned}
$$

2.3   Derivatives

Using the same approach as the area bound in [2], and defining $\Theta_i(s)=\sum_j \overline{\mathbf{T}}_{ij}(0)a_j(s)$, we find that the derivatives are

$$
\frac{\partial\hat{A}(s)}{\partial\widetilde{\mathbf{M}}_{mn}^{\mathrm{pot/dep}}} = f^{\mathrm{pot/dep}}\mathbf{p}_m^\infty\left(r\left[\Theta_m(s)-\Theta_n(s)\right]+(rc_m(s)\pm 1)\left[\eta_m^w(s)-\eta_n^w(s)\right]\right), \qquad (38) \quad \boxed{\texttt{eq:lAdiff}}
$$

with the upper/lower sign in $\pm$ corresponding to derivatives wrt. $\mathbf{M}^{\mathrm{pot/dep}}$. The term with $\Theta(s)$ comes from differentiating $\mathbf{p}^\infty$, the term with $c(s)$ comes from differentiating $\eta^w(s)$ and the term with ± 1 comes from differentiating $\mathbf{K}$ in (36).

3 Upper bounds

3.1 Initial SNR

The initial SNR is:

$$\text{SNR}(0) = \overline{\text{SNR}}(0) = \lim_{s \to \infty} sA(s). \tag{39}$$

As shown in [2], this satisfies the inequality

$$\text{SNR}(0) \leq \sqrt{N}, \tag{40}$$

The model that saturates this bound is (equivalent to) a two-state model with deterministic transitions and $f^{\text{pot/dep}} = \frac{1}{2}$.

3.2 Area bound

The area under the SNR curve is:

$$\int_0^\infty \mathrm{d}t \, \text{SNR}(t) = \lim_{\tau \to \infty} \tau \overline{\text{SNR}}(\tau) = A(0). \tag{41}$$

As shown in [2], this satisfies the inequality

$$A(0) \leq \frac{\sqrt{N}(M-1)}{r}. \tag{42}$$

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

3.3 Envelope

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

$$\mathbf{W}^{\text{F}} = -\sum_a q_a \mathbf{u}^a \boldsymbol{\eta}^a, \quad \boldsymbol{\eta}^a \mathbf{u}^b = \delta_{ab}, \quad \mathbf{W}^{\text{F}} \mathbf{u}^a = -q_a \mathbf{u}^a, \quad \boldsymbol{\eta}^a \mathbf{W}^{\text{F}} = -q_a \boldsymbol{\eta}^a. \tag{43}$$

Then we can write

$$\text{SNR}(t) = \sqrt{N} \sum_a \mathcal{I}_a \mathrm{e}^{-rt/\tau_a}, \qquad \overline{\text{SNR}}(\tau) = \sqrt{N} \sum_a \frac{\mathcal{I}_a}{1 + r\tau/\tau_a}, \tag{44}$$

where

$$\mathcal{I}_a = (\mathbf{p}^\infty \widetilde{\mathbf{K}} \mathbf{u}^a)(\boldsymbol{\eta}^a \mathbf{w}), \qquad \tau_a = \frac{1}{q_a}. \tag{45}$$

From (42) and (40), we see that they are subject to the constraints

$$\sum_a \mathcal{I}_a \tau_a \leq M - 1, \qquad \sum_a \mathcal{I}_a \leq 1. \tag{46}$$

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

$$\mathcal{L} = \sum_a \frac{\mathcal{I}_a}{1 + r\tau/\tau_a} + \mu_{\mathcal{I}}\left(1 - \sum_a \mathcal{I}_a\right) + \mu_{\mathcal{A}}\left((M-1) - \sum_a \mathcal{I}_a\tau_a\right). \qquad (47) \quad \boxed{\texttt{eq:envlagran}}$$

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

$$\frac{\partial\mathcal{L}}{\partial\mathcal{I}_a} = 0, \qquad \frac{\partial\mathcal{L}}{\partial\tau_a} = 0, \qquad \mu_\alpha \geq 0, \qquad \frac{\partial\mathcal{L}}{\partial\mu_\alpha} \geq 0, \qquad \mu_\alpha\frac{\partial\mathcal{L}}{\partial\mu_\alpha} = 0. \qquad (48) \quad \boxed{\texttt{eq:envKTcond}}$$

These are solved by

$$\mathcal{I}_1 = 1, \qquad \mathcal{I}_{a>1} = 0, \qquad \tau_1 = M - 1,$$
$$\mu_{\mathcal{I}} = \frac{(M-1)^2}{(r\tau + (M-1))^2}, \qquad \mu_{\mathcal{A}} = \frac{r\tau}{(r\tau + (M-1))^2}. \qquad (49) \quad \boxed{\texttt{eq:envsol}}$$

This leads to the envelope

$$\overline{\text{SNR}}(\tau) \leq \frac{\sqrt{N}(M-1)}{r\tau + (M-1)}. \qquad (50) \quad \boxed{\texttt{eq:env}}$$

This is the envelope ("memory frontier") that we can prove. Finding more constraints would lower it. Clearly there have to be more constraints, as (49) indicates that this would be achieved by the same model at all times. This model would have to saturate both the initial SNR bound (40) and the area bound (42), whereas we saw that they were achieved by different models, which perform well at short/long timescales respectively.

4 Finite time

$\boxed{\texttt{sec:finite}}$

Now we'll maximise $\overline{\text{SNR}}(\tau)$ for some fixed τ, subject to the constraints (9). First, we will maximise the numerator of (35) holding $\mathbf{p}^\infty\mathbf{w}$ fixed. We can maximise the ration wrt. $\mathbf{p}^\infty\mathbf{w}$ later. Consider the Lagrangian

$$\mathcal{L} = \sum_{ij} \mathbf{p}_i^\infty\mathbf{K}_{ij}(\eta_i^w(s) - \eta_j^w(s)) + \sum_{ij}\left(f^{\text{pot}}\mu_{ij}^{\text{pot}}\mathbf{M}_{ij}^{\text{pot}} + f^{\text{dep}}\mu_{ij}^{\text{dep}}\mathbf{M}_{ij}^{\text{dep}}\right)$$
$$+ \lambda\left(\Delta\mathbf{p} - \sum_i \mathbf{p}_i^\infty\mathbf{w}_i\right), \qquad (51) \quad \boxed{\texttt{eq:finitelag}}$$

where $\Delta\mathbf{p}$ is the constant value we are holding $\mathbf{p}^\infty\mathbf{w}$ at. After maximising The Karush-Kuhn-Tucker necessary conditions for a maximum are

$$\frac{\partial\mathcal{L}}{\partial\mathbf{M}_{ij}^{\text{pot/dep}}} = 0, \quad \mu_{ij}^{\text{pot/dep}} \geq 0, \quad \frac{\partial\mathcal{L}}{\partial\mu_{ij}^{\text{pot/dep}}} \geq 0, \quad \mu_{ij}^{\text{pot/dep}}\frac{\partial\mathcal{L}}{\partial\mu_{ij}^{\text{pot/dep}}} = 0, \quad \frac{\partial\mathcal{L}}{\partial\lambda} = 0. \quad (52) \quad \boxed{\texttt{eq:KTfinite}}$$

The derivatives, for $m \neq n$ are

$$\frac{1}{f^{\mathrm{pot/dep}}} \frac{\partial \mathcal{L}}{\partial \widetilde{\mathbf{M}}_{mn}^{\mathrm{pot/dep}}} = \mathbf{p}_m^\infty \left(r \left[\Theta_m(s) - \Theta_n(s) \right] + (rc_m(s) \pm 1) \left[\eta_m^w(s) - \eta_n^w(s) \right] \right)$$
$$+ \mu_{mn}^{\mathrm{pot/dep}} - \mu_{mm}^{\mathrm{pot/dep}} + \lambda \mathbf{p}_m^\infty \sum_i (\overline{\mathbf{T}}_{mi}(0) - \overline{\mathbf{T}}_{ni}(0)) \mathbf{p}_i^\infty \mathbf{w}_i. \tag{53}$$

eq:finiteder

Below, we will consider various combinations of these derivatives that preserve $\mathbf{p}^\infty$.

4.1 Increasing c_k

eacoeffincr

Consider the following combinations of derivatives:

$$\Delta_{mn} \equiv \frac{1}{\mathbf{p}_m^\infty} \left(\frac{\partial}{\partial \widetilde{\mathbf{W}}_{mn}^{\mathrm{F}}} \right) + \frac{1}{\mathbf{p}_n^\infty} \left(\frac{\partial}{\partial \widetilde{\mathbf{W}}_{nm}^{\mathrm{F}}} \right), \tag{54}$$

eq:areacoeff

One can show that the equilibrium probabilities, $\mathbf{p}^\infty$, are invariant under these operators:

$$\Delta_{mn} \mathbf{p}_i^\infty = 0, \tag{55}$$

eq:sareacoef

which makes it possible to integrate the perturbation:

$$\widetilde{\mathbf{M}}^{\mathrm{pot/dep}} \to \widetilde{\mathbf{M}}^{\mathrm{pot/dep}} + \frac{\mathbf{D}\boldsymbol{\epsilon}}{2 f^{\mathrm{pot/dep}}}, \qquad \begin{aligned} \boldsymbol{\epsilon} &= \boldsymbol{\epsilon}^{\mathrm{T}}, \\ \boldsymbol{\epsilon}\mathbf{e} &= 0. \end{aligned} \tag{56}$$

eq:areacoeff

But more interestingly:

$$\Delta_{mn} \mathcal{L} = \frac{\mu_{mn}^{\mathrm{pot}} + \mu_{mn}^{\mathrm{dep}}}{2\mathbf{p}_m^\infty} + \frac{\mu_{nm}^{\mathrm{pot}} + \mu_{nm}^{\mathrm{dep}}}{2\mathbf{p}_n^\infty} + r \left(c_m(s) - c_n(s) \right) \left(\eta_m^w(s) - \eta_n^w(s) \right), \tag{57}$$

eq:areacoeff

$$\tag{58}$$

where c_k were defined in (37). As this derivative is zero at a maximum, and the $\mu_{ij}^{\mathrm{pot/dep}}$ are non negative, the last term must be nonpositive. Therefore, if we arrange the states in order of decreasing $\eta_i^w(s)$, the $c_k(s)$ must be nondecreasing.

4.1.1 Shortcuts

:shortcuts

Now consider the following combinations of derivatives for $n > 1$:

$$\widetilde{\Delta}_{m,n} \equiv \left[\sum_{k=0}^{n-1} \frac{1}{\mathbf{p}_{m+k}^\infty} \left(\frac{\partial}{\partial \widetilde{\mathbf{W}}_{m+k,m+k+1}^{\mathrm{F}}} \right) \right] - \frac{1}{\mathbf{p}_m^\infty} \left(\frac{\partial}{\partial \widetilde{\mathbf{W}}_{m,m+n}^{\mathrm{F}}} \right). \tag{59}$$

eq:shortcutd

One can show that the equilibrium probabilities, $\mathbf{p}^\infty$, are invariant under these operators:

$$\widetilde{\Delta}_{m,n} \mathbf{p}_i^\infty = 0, \tag{60}$$

eq:shortcutp

which makes it possible to integrate the perturbation:

$$
\widetilde{\mathbf{M}}^{\text{pot/dep}} \to \widetilde{\mathbf{M}}^{\text{pot/dep}} + \frac{\mathbf{D}\boldsymbol{\epsilon}^{(m,n)}}{2f^{\text{pot/dep}}}, \qquad
\begin{aligned}
\left(\boldsymbol{\epsilon}^{(m,n)}\right)_{m,m+n} &= -\epsilon, \\
\left(\boldsymbol{\epsilon}^{(m,n)}\right)_{m+k,m+k+1} &= \epsilon && \forall\, k \in [0, n-1], \\
\left(\boldsymbol{\epsilon}^{(m,n)}\right)_{m+k,m+k} &= -\epsilon && \forall\, k \in [1, n-1].
\end{aligned}
\qquad (61)
$$

`eq:shortcutf`

But more interestingly for our purposes:

$$
\begin{aligned}
\widetilde{\Delta}_{m,n}\mathcal{L} = {} & \left[\sum_{k=0}^{n-1} \frac{\mu^{\text{pot}}_{m+k,m+k+1} + \mu^{\text{dep}}_{m+k,m+k+1}}{\mathbf{p}^{\infty}_{m+k}} - \frac{\mu^{\text{pot}}_{m,m+n} + \mu^{\text{dep}}_{m,m+n}}{\mathbf{p}^{\infty}_{m}} \right] \\
& + r \sum_{k=0}^{n-1} \left(\eta^{w}_{m+k}(s) - \eta^{w}_{m+k+1}(s) \right) \left(c_{m+k}(s) - c_{m}(s) \right),
\end{aligned}
\qquad (62)
$$

`eq:shortcuta`

If we put the states in order of decreasing $\eta^{w}_{k}(s)$, the results of the §4.1 tell us that the $c_{k}(s)$ are non-decreasing. This implies that the last term of the final expression in (62) is non-negative. If it is non-zero (there would need to be a lot of degeneracy for it to be zero), this would imply that $\mu^{\text{pot/dep}}_{m,m+n} > 0$, which in turn implies that $\widetilde{\mathbf{M}}^{\text{pot/dep}}_{m,m+n} = 0$ for all $n > 1$. NO! ONLY ONE OF THEM.

A similar technique also works for $n < 1$. This implies that the only nonzero off-diagonal transitions are nearest neighbour.

References

`eny1981fund` [1] J. G. Kemeny, "Generalization of a fundamental matrix," *Linear Algebra and its Applications* **38** (1981) no. 0, 193 – 206.

`2013synapse` [2] 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.