Information Theory and Neuroscience II

John Z. Sun and Da Wang

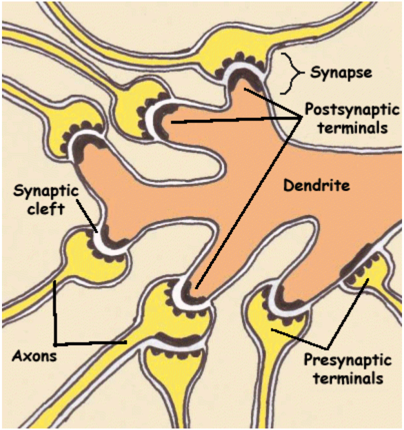
Massachusetts Institute of Technology

October 14, 2009

Outline

- System Model & Problem Formulation
- Information Rate Analysis
- Recap

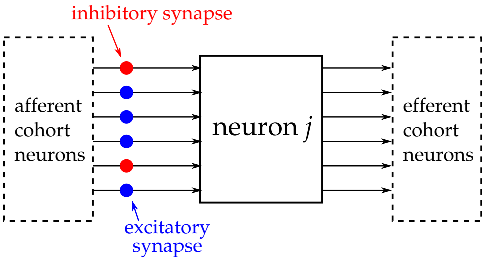
Neurons



Neuron (denoted by j)

- I/O: via synapses
- excitatory & inhibitory synapses
- afferent cohort $a(j)$, efferent cohort

Neurons



85% synapses are excitatory

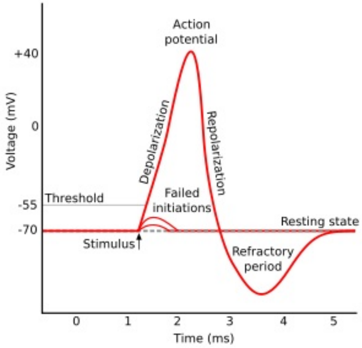
- amount $n = 8500$ typical in primary sensory cortex

$|a(j)|$ percentagewise close to n

- though some axons form more than one synapses with j

Neural Spike Train

Neural spike pulse shape:



time of arrival (TOA) conventions:

- several options
- needs to use one consistently

asynchronous operations among $a(j)$:

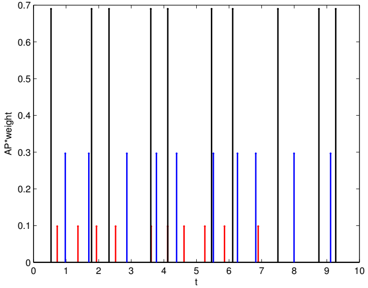
- no links among spike production time from different neurons

$\Rightarrow$ when $n > 2$, possible for two spikes to be arbitrary close

- when $n = 1$, need to wait the refractory period Δ

Neural Spike Train

Union of spikes from three neurons:



time of arrival (TOA) conventions:

- several options
- needs to use one consistently

asynchronous operations among $a(j)$:

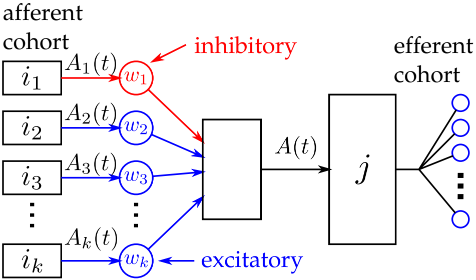
- no links among spike production time from different neurons

$\Rightarrow$ when $n > 2$, possible for two spikes to be arbitrary close

- when $n = 1$, need to wait the refractory period Δ

Afferent Spike Train Process

$$A_i(t) \triangleq \lim_{dt \downarrow 0} \frac{\mathbb{P}\left[i \text{ produces a spike in } (t, t+dt)\right]}{dt}$$



$A_+(t)$ & $A_-(t)$: random excitatory/inhibitory intensity

$$A_+(t) \triangleq \sum_{i:w_i>0} w_i A_i(t)$$

$$A_-(t) \triangleq \sum_{i:w_i<0} |w_i| A_i(t)$$

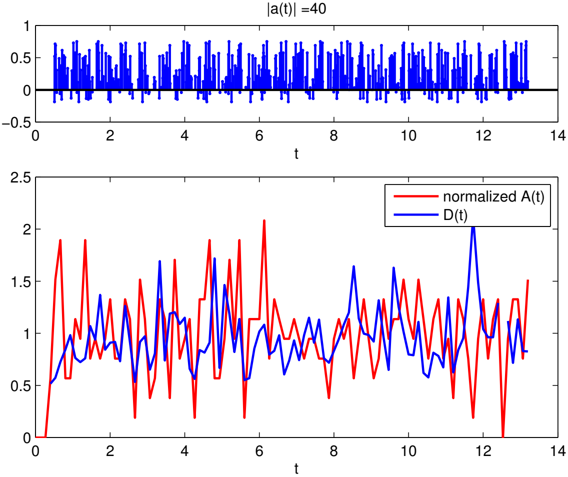
Union of spike trains:

$$A(t) = \frac{A_+(t)}{1 + C A_-(t)}$$

$A(t)$ always non-negative.

Spike train process (cont.)

- spikes $\leftrightarrow$ inter-spike interval (ISI)
- $A(t)$: instantaneous random mean spiking frequency (unit: spikes/second)
- $D(t)$: instantaneous random mean excitatory afferent ISI duration (unit: time unit)



Why $D(t)$?

Introducing $D(t)$ allows us to conceptualize the neuron as a communication channel:

- input (excitations): has unit of time
- output (firing): has unit of time. [coming soon]
- stochastically converts input time signals to output time signals

Statistics of $D(t)$

- $\rho_D(\tau)$: correlation of $D(t)$
- τ_D: "coherence time" of $D(t)$

$$\rho_D(\tau_D) = 1/2$$

- $\tau_D \gg \Delta$

- stationary assumption
- analogous to the modeling of wireless channel

Channel Model: input

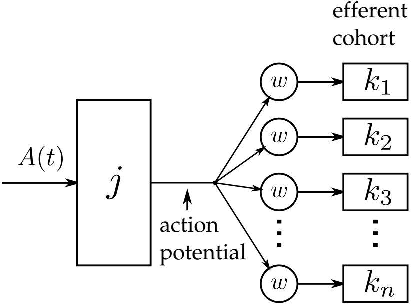
What is the output?

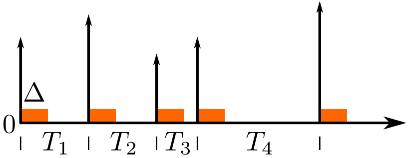
Efferent Spike Train Process

- Efferent spike: action potential to neurons in efferent cohort
- T_k: the duration for k-th ISI ($T_k \sim P_T, \forall k$)
- S_k: the time at which k-th AP is generated:

$$S_k = \sum_{i=1}^{k} T_i$$

- $T_k \geq \Delta =$ refractory period



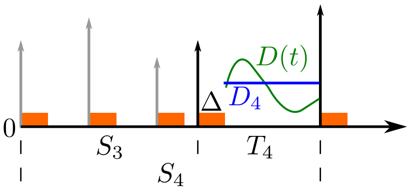


Channel Model: input & output

- both input & output have units of time
- but different indexing!

$$D_k = \frac{1}{T_k - \Delta} \int_{S_{k-1}+\Delta}^{S_k} D(u)\, du$$

$k = 4$:



Construct a piecewise constant random process $\tilde{D}(t)$, where

$\tilde{D}(t) = D_k$ for $S_{k-1}+\Delta \leq t < S_k$

Mean Value Assumption

$\left\{\tilde{D}(t)\right\}$ adequately represents $D(t)$ for the purpose of analysis.

Feedback among neurons

For neuron j, its excitations come from:

- top-down feedback from higher levels of the cortex
- horizontal feedback from the same neuron region
- bottom-up signals from lower levels of the cortex

$\Rightarrow$ causal feedback

$\Rightarrow$

$$T_n \perp\!\!\!\perp (D_1, \ldots, D_{n-1}) | D_n$$
$$T_n \perp\!\!\!\perp (T_1, \ldots, T_{n-1}) | D_n$$

$$\begin{tikzcd}
D_1 \arrow[r] & T_1 \arrow[dl] \arrow[ddl] \arrow[ddddl] \\
D_2 \arrow[r] & T_2 \arrow[dl] \arrow[dddl] \\
D_3 \arrow[r] & T_3 \arrow[ddl] \\
\vdots & \vdots \\
D_n \arrow[r] & T_n
\end{tikzcd}$$

Channel Model: integer-indexed input & output

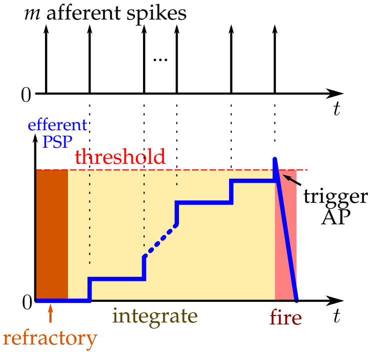
Any model for the channel?

Classical Integrate and Fire (CIF) Neuron

CIF neuron:

- excitatory synapses have the same weight w.
- unit step response to each afferent spike
- fixed threshold $\eta \in ((m-1)w, mw]$

Always need m spikes to fire



Input-output relationship

$$T_k = \Delta + mI$$
$$\mathbb{E}[I] = d_k$$
$$d_k \sim D_k$$

CIF Neuron: Energy Constraints

Energy that j expands during an ISI:

- metabolic energy:

$$e_1 = C_1 T$$

- energy to construct PSP during $(\Delta, T]$:

$$e_2 = C_2 M$$

- energy to generate AP:

$$e_3 = C_3$$

- For CIF model, we have $M = m$ always, hence

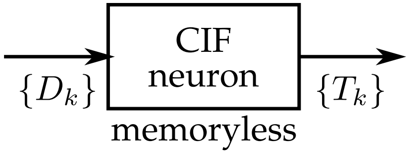
$$e_{CIF}(T) = C_0 + C_1 T$$

- T is the random duration of the ISI
- M is the random number of afferent spikes in $(\Delta, T]$
- $C_0 = C_2 m + C_3$.

Problem Formulation

Channel model:

- Input: $\{D_k\}$
- Output: $\{T_k\}$
- memoryless and time-invariant channel
- but with (lots of) causal feedback!



Central tenet

Te optimality criterion apropos of neuronal information transmission is the maximization of bits per joule (bpj).

Main objective

Determines the optimal input & output distributions $f_D(\cdot)$ and $f_T(\cdot)$ based on the above principle.

Information Rate

We start by investigating the information rate $I(\mathbf{D};\mathbf{T})$.

$$\mathbf{D} \triangleq (D_1, D_2, \cdots, D_n)$$
$$\mathbf{T} \triangleq (T_1, T_2, \cdots, T_n)$$

Memoryless $\Rightarrow$

$$f_{\mathbf{T}|\mathbf{D}}(\mathbf{t}|\mathbf{d}) = \prod_{i=1}^{n} f_{T|D}(t_i|d_i)$$

However, as $\{D_i\}$ may not be independent,

$$I(\mathbf{D};\mathbf{T}) \leq \sum_{i=1}^{n} I(D_i;T_i)$$

Equality only when $\{D_i\}$ independent. Are they?

Implications of τ_D

Recall

$$\tau_D \gg \Delta$$

When T_k slightly larger than Δ

$\Rightarrow$ D_{k+1} and D_k highly correlated

$\Rightarrow$ T_{k+1} and T_k similarly correlated

$\Rightarrow$ T_{k+1} will be similarly small

For two jointly Gaussian r.v.,

$$I_{\text{jointly Gaussian}} = -\frac{1}{2}\log(1-\rho^2)$$

where $I \to \infty$ as $\rho \to 1$.

Message: correlation between D_k and D_{k+1} results in a penalty in information rate.

Long Term Information Rate

Incremental conditional mutual information:

$$I = \lim_{n\to\infty} I_n$$

where

$$\begin{aligned}
I_n &= I(D_1,\ldots,D_n;T_n|T_1,\ldots,T_{n-1}) \\
&= I(D_n;T_n|T_1,\ldots,T_{n-1}) + \cancel{I(D_1,\ldots,D_{n-1};T_n|D_n,T_1,\ldots,T_{n-1})} \\
&= h(T_n|T_1,\ldots,T_{n-1}) - h(T_n|D_n,T_1,\ldots,T_{n-1}) \\
&= h(T_n|T_1,\ldots,T_{n-1}) - h(T_n|D_n) \\
&= I(D_n;T_n) - I(T_n;T_1,\ldots,T_{n-1})
\end{aligned}$$

Since $\{(D_k,T_k)\}$ strictly stationary:

$$I = I(D_1;T_1) - \underbrace{\left[I(T_1,T_2) + \lim_{n\to\infty} I(T_n;T_1,\ldots,T_{n-2}|T_{n-1})\right]}_{\text{information decrement}}$$

Long Term Information Rate (cont.)

Long Term Information Rate

$$I = I(D_1; T_1) - \underbrace{\left[I(T_1, T_2) + \lim_{n\to\infty} I(T_n; T_1, \ldots, T_{n-2} | T_{n-1})\right]}_{\text{information decrement}}$$

$I(T_1, T_2)$: principal information decrement

$$\lim_{n\to\infty} I(T_n; T_1, \ldots, T_{n-2} | T_{n-1})$$

negligible, as having T_{n-1} is almost as effective as having D_{n-1}.

Information Decrement $I(T_1; T_2)$

$$
\begin{aligned}
&I(T_1; T_2) \\
=&\mathbb{P}\left[T_1 > \tau_D\right] I(T_1; T_2 | T_1 > \tau_D) \\
&+ \mathbb{P}\left[T_1 \leq \tau_D\right] I(T_1; T_2 | T_1 \leq \tau_D) \\
\approx&\mathbb{P}\left[T_1 \leq \tau_D\right] I(T_1; T_2 | T_1 \leq \tau_D)
\end{aligned}
$$

$\leftarrow$ when $T_1 > \tau_D$, $T_2 \perp\!\!\!\perp T_1$ effectively.

When $T_1 < 2\Delta \ll \tau_D$,

$$
\begin{aligned}
\rho_{T_1, T_2} &\approx \rho_D(T_1) \\
I(T_1; T_2) &\approx -\kappa \mathbb{E}\left[\log T_1\right] + C
\end{aligned}
$$

$\leftarrow$ high correlation

$\leftarrow$ via analyzing the conditional variance of T_2.

where κ and C are constants based on $\rho_D(\tau)$.

Recap

- Model channel input and output as time signals
- Mean value assumption: simplify analysis
- Memoryless channel with causal feedback
- CIF neuron model
  - energy

$$e_{CIF}(T) = C_0 + C_1 T$$

- Information rate analysis
  - coherence time
  - information decrement
  - information rate

$$I = I(D_1; T_1) - I(T_1, T_2)$$

- Lots of “hand-waving” in modeling...