Neural Signaling III – Neuronal Coding

Two views of neuronal coding

There are two opposing views about how information is encoded in the spike trains, with many intermediate standpoint's. One view holds that it is only the *average rate*, the number of action potentials over some suitable interval (a few hundred msecs to 1 or more seconds), is relevant for information processing. The opposing view, known as *temporal coding*, argues that the seemingly random patterns of the spikes, even the exact temporal arrangement of interspike intervals both for the single cell as well as between multiple cells, encodes information.

The broad idea of temporal coding is supported by evidence from a variety of sensory systems pointing toward the role of spike timing, in particular across ensembles of cells, in encoding various aspects of the stimulus.

A more balanced view would be somewhere in between of the two extremes. On the one hand we realize temporal patterns of the spike trains do encode information, on the other hand, we also acknowledge the random nature of the physiology of neuronal firing, such as the fluctuation of the injected current I (caused by, for example, the number and magnitudes of the synaptic inputs) and the threshold potential V_{th}.

Review of linear, time invariant system (LTI)

Neurons can be treated as linear encoding systems. Before further discussion, we first review the theory of linear, time invariant system (LTI).

- **Response of a LTI system**

  If a LTI system receives a signal $x(t)$ as the input (stimulus), its output (response) $y(t)$ can be found by this *convolution* operation:

  $$y(t) = h(t) * x(t) = \int_{t=-\infty}^{\infty} h(t-\tau)x(\tau)d\tau$$

  where $h(t)$ is the *impulse response function* of the LTI system (also called the point spread function, Green's function). The Fourier transform of $h(t)$ is called the *transfer function* of the neuron:

  $$H(\omega) = \mathcal{F}h(t) = \int_{t=-\infty}^{\infty} h(t)e^{-j\omega t}dt$$

This LTI system can also be equivalently described in frequency domain as

$$Y(\omega) = H(\omega)\, X(\omega)$$

where $X(\omega)$, $Y(\omega)$, and $H(\omega)$ are the Fourier transforms of $x(t)$, $y(t)$ and $h(t)$, respectively.

- **Correlation and power spectrum**

  The *auto-correlation* of a signal $x(t)$ is defined as

  $$r_{xx}(t) = \int_{\tau=-\infty}^{\infty} x(\tau - t)x(\tau)d\tau$$

  Note that $r_{xx}(t)$ is even symmetric

  $$r_{xx}(t) = r_{xx}(-t)$$

  The *cross-correlation* between two signals $x(t)$ and $y(t)$ is

  $$r_{xy}(t) = \int_{\tau=-\infty}^{\infty} x(\tau - t)y(\tau)d\tau$$

  The auto-correlation of a signal $x(t)$ can be Fourier transformed to get its *power spectrum*

  $$S_{xx}(\omega) = \mathcal{F}r_{xx}(t) = \int_{t=-\infty}^{\infty} r_{xx}(t)e^{-j\omega t}dt$$

  and similarly for the cross-correlation of $x(t)$ and $y(t)$, their power spectrum is

  $$S_{xy}(\omega) = \mathcal{F}r_{xy}(t) = \int_{t=-\infty}^{\infty} r_{xy}(t)e^{-j\omega t}dt$$

- **Obtain transfer function from power spectra**

  The cross-correlation of the input $x(t)$ and output $y(t)$ of a LTI system can be obtained from the auto-correlation of the input and the impulse response $h(t)$ of the system:

  $$r_{xy}(t) \quad = \quad \int_{\tau=-\infty}^{\infty} x(\tau - t)y(\tau)d\tau$$

$$
\begin{aligned}
&= \int_{\tau=-\infty}^{\infty} x(\tau - t) \int_{\tau=-\infty}^{\infty} x(\tau - t')h(t')dt'd\tau \\
&= \int_{\tau=-\infty}^{\infty} h(t') \int_{\tau=-\infty}^{\infty} x(\tau - t)x(\tau - t')h(t')d\tau dt' \\
&= \int_{\tau=-\infty}^{\infty} h(t')r_{xx}(t - t')dt' = h(t) * r_{xx}(t)
\end{aligned}
$$

And similarly, the auto-correlation of the output can be obtained by

$$
\begin{aligned}
r_{yy}(t) &= \int_{\tau=-\infty}^{\infty} y(\tau - t)y(\tau)d\tau \\
&= \int_{\tau=-\infty}^{\infty} y(\tau - t) \int_{\tau=-\infty}^{\infty} x(\tau - t')h(t')dt'd\tau \\
&= \int_{\tau=-\infty}^{\infty} h(t') \int_{\tau=-\infty}^{\infty} y(\tau - t)x(\tau - t')h(t')d\tau dt' \\
&= \int_{\tau=-\infty}^{\infty} h(t')r_{xy}(t - t')dt' \\
&= h(t) * r_{xy}(t) = h(t) * h(t) * r_{xx}(t)
\end{aligned}
$$

These relations can be expressed in frequency domain as

$$S_{xy}(\omega) = H(\omega)S_{xx}(\omega)$$

and

$$S_{yy}(\omega) = H(\omega)S_{xy}(\omega)H(\omega)S_{xx}(\omega) = |H(\omega)|^2 S_{xx}(\omega)$$

Neuronal response as a LTI system

A neuronal system can be approximated by a linear, time-invariant system (LTI) with the relevant stimulus $s(t)$ as the input and the observed spike train $x(t)$ as the output:

$$x(t) = h(t) * s(t) = \int_{t=-\infty}^{\infty} x(t - \tau)s(\tau)d\tau$$

Here the neuron responds to the stimulus $s(t)$ in a particular way specified by the impulse response function $h(t)$.

$h(t)$ can be obtained experimentally by observing the spike train $x(t)$ as the response to the stimulus $s(t)$ by

$$H(\omega) = \frac{S_{sx}(\omega)}{S_{ss}(\omega)}$$

Specially if the stimulus is *white* or uncorrelated

$$r_{ss}(t) = \frac{1}{\sigma^2}\delta(t)$$

its power spectrum becomes a constant

$$S_{ss}(\omega) = \frac{1}{\sigma^2}$$

and the response spike train has N spikes in time window T

$$x(t) = \sum_{k=1}^{N} \delta(t - t_k)$$

then the transfer function $H(\omega)$ can be obtained as

$$H(\omega) = fracS_{sx}(\omega)S_{ss}(\omega)$$

or in time domain

$$\begin{aligned} h(t) &= \frac{1}{\sigma^2} r_{sx}(t) = \frac{1}{\sigma^2} \int_{-\infty}^{\infty} s(\tau - t)x(\tau)d\tau \\ &= \frac{1}{\sigma^2} \int_{-\infty}^{\infty} s(\tau - t) \sum_{k=1}^{N} \delta(\tau - t_k)d\tau = \frac{1}{\sigma^2} \sum_{k=1}^{N} s(t_k - t) \end{aligned}$$

This is the *reverse-correlation formula* of h which states that the linear impulse response function of a neuron can be recovered by a simple spike-triggered average of the stimulus preceding the spikes.

Neural coding of information

- Conditional probabilities

  Let A and B be two random events and $P(A)$ and $P(B)$ be their probabilities. The probability of the joint event of both A and B, represented by $P(A, B)$, can be obtained as

  $$P(A, B) = P(A)P(B/A) = P(B)P(A/B)$$

  where $P(A/B)$ is the conditional probability of event A given the condition that event B, and $P(B/A)$ is defined similarly.

If the two events are independent of each other, i.e., how likely event A will occur is independent of whether event B occurs, and vice versa, then

$$P(A/B) = P(A), \qquad P(B/A) = P(B)$$

and the joint probability becomes

$$P(A, B) = P(A)P(B)$$

- Stimuli and responses as random variables

  The response property of a neuron can be characterized by the *tuning curve* which represents how the neuron responds differently to stimulus with a varying parameter. When there are two varying parameter, the *tuning surface* is used. The shape of one dimensional tuning curves usually fall into one of two categories, sigmoidal shaped, as the tuning for contrasts or near and far stereo cells for distances, and bell shaped (Gaussian), as for orientation or direction of motion.

  The a neuron is treated as a communication channel which receives the stimuli $s(t)$ as the input and generates the spike trains $x(t) = \sum_k \delta(t - t_k)$ as the output. Here both $s(t)$ and $x(t)$ are treated as random variables and the observing $x(t)$ as the response to $s(t)$ is considered as a process in which certain information (e.g., about the external world from the sensory stimuli) is gained. To describe the relationship between them, that is, how the stimulus affects the response and how the response reflects the stimulus, we define

  - $p(s(t))$, the probability distribution of $s(t)$;
  - $p(x(t))$, the probability distribution of $x(t)$;
  - $p[x(t)/s(t)]$, conditional probability of observing $x(t)$ given the condition of $s(t)$. This describes how the information in the stimulus is coded in the response of the neuron. The tuning curve is the result of measuring this probability.
  - $p[s(t)/x(t)]$, conditional probability of stimulus $s(t)$ given the observed $x(t)$ as the condition. This is related to the inverse problem of how to decode the response to get information about the stimulus.

These probabilities are all related to the joint probability of both $s(t)$ and $x(t)$

$$p[s(t),\ x(t)] = p(s(t))p(x(t)/s(t)) = p(x(t))p(s(t)/x(t))$$

- Neuronal response as communication channel

  $p(s)$ is also called the *a priori* probability of input s, as it represents the likelihood of s before the neuron responding to it; and $p(s/x)$ is called the *post priori* probability of s, as it represents the likelihood of s given that the neuron has responded to it by of x. As observing x will always more or less gain some information about s, we have

  $$p(s/x) \geq p(s)$$

  where the equal sign is for the worst case where x is completely irrelevant to s therefore nothing about s can be learned from x.

  The total information gained from this process is quantitatively given by

  $$I = log\frac{p(s/x)}{p(s)} \geq 0$$

  if the base of log is 2, then the unit of information is bit.

- **Information I – ideal case**

  In the ideal case without noise, complete information about the stimuli can be obtained in the sense that any given stimulus s is responded to with a specific x, i.e,, $p(s/x) = 1$,

  $$I = log\frac{1}{p(s)} = -log\ p(s)$$

  Note that the less likely an event (small $p(s)$), the more information is gained by learning this event occurs. (e.g., snow vs. fine whether in Los Angeles.) In particular, if $p(s) = 1$, no information is gained ($I = 0$), due to the log operation.

  If the input is not a binary random event (either happens or not with probability P), but a random variable of varying values s with probability distribution $p(s)$, then the information gained in an ideal communication channel is the weighted average of information for all possible values of s:

$$I = -\sum_i p(s_i) log\ p(s_i)$$

if s is discrete, or

$$I = -\int p(s) log\ p(s) ds$$

if s is continuous.

- **Information II – with noise**

  If the system is noisy, the stimulus-response relationship is no longer definite, as the same s may be responded to with different x due to noise, i.e., $p(s/x) < 1$. Here we first find the information about s gained when observing a particular x:

$$I(x) = \int p(s/x) log \frac{p(s/x)}{p(s)} ds$$

  Then we find the total information about s gained when observing all possible x:

$$\begin{aligned} I &= \int p(x) I(x) dx \\ &= \int [\int p(s/x) log \frac{p(s/x)}{p(s)} ds] dx \\ &= \int p(x) [\int p(s/x) logp(s/x) ds] dx - \int p(x) [\int p(s/x) logp(s) ds] dx \\ &= \int p(x) [\int p(s/x) logp(s/x) ds] dx - \int [\int p(x) p(s/x) dx] logp(s) ds \\ &= H(s) - H(s/x) \end{aligned}$$

  where

$$H(s) \triangleq -\int p(s) logp(s)$$

  and

$$H(s/x) \triangleq \int p(x) [\int p(s/x) logp(s/x) ds] dx$$

- **Entropy**

  We define $H(s)$ as the *entropy* of s. Note that as $H(s)$ is the same as the complete information gained in the ideal case, it represents the maximum possible information available about s that can be gained. Since information gaining can be considered the same process as uncertainty reduction, $H(s)$ also represents the uncertainty about s before observing x. And $H(s/x)$ is the *conditional entropy*, representing the uncertainty of s after observing x. If the base of logarithm is 2, the unit of entropy is bit.

The information gained $I = H(s) - H(s/x)$ is thus called *mutual information* and represents the uncertainty reduction from $H(s)$ before observing x to $H(s/x)$ after observing x. In the ideal case $p(s/x) = 1$, after observing s the uncertainty becomes zero $H(s/x) = 0$, and the complete information about s $I = H(s) = -\int p(s) log\, p(s) ds$ is gained.

As an example, consider an experiment with n equally likely possible outcomes s_i $(i = 1, \cdots, n)$. The probability for a particular s_i to occur is $p(s_i) = 1/n$. Now the total uncertainty, the entropy, of this experiment is

$$H(s) = -\sum_{i=1}^{n} p(s_i) log\, p(s_i) = log\, n$$

We see that under the condition of equal likelihood, the entropy is simply the logarithm of the total number of possible outcomes. If $n = 8$, then the entropy, i.e, the maximum possible information available about this experiment, is $log_2 8 = 3$ bits.

- **Entropy under given conditions**

  Here we find the probability distribution $p(x)$ to maximize the entropy $H(x) = \int p(x) log\, p(x) dx$ subject to some given conditions, as well as

  $$\int_{x=-\infty}^{\infty} p(x) dx = 1$$

  or

  $$\sum_i p(x_i) = 1$$

  This kind of constrained optimization problem can be solved by Lagrange multiplier method.

– Given the range of a random variable ($0 \le x \le a$), the uniform distribution

$$p(x) = \begin{cases} 1/a & 0 \le x \le a \\ 0 & else \end{cases}$$

has the maximum entropy

$$H(x) = loga$$

– Given the range of a random variable ($0 \le x < \infty$) and its mean μ, the exponential decay distribution

$$p(x) = \frac{1}{\mu} e^{-x/\mu}$$

has the maximum entropy

$$H(x) = log(e\mu)$$

– Given the variance σ^2 of a random variable x, we find the normal (Gaussian) distribution

$$p(x) = N(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

has the maximum entropy

$$H(x) = log\sqrt{2\pi e\sigma^2}$$

The maximum entropy is proportional to $log\sigma$ (independent of μ). Also as the variance σ^2 can be considered as the dynamic energy contained in x, we see that among all possible distributions with the same uncertainty, Gaussian distribution requires the minimum energy.

If the only knowledge about a random experiment is the mean μ and variance σ^2 (which can be estimated by carrying the experiment a large number of times), then its unknown probability distribution can be best estimated by a normal distribution $N(x, \mu, \sigma^2)$, as it imposes minimum artificial constraint by allowing the maximum uncertainty.

- *Entropy of spike trains*

  We now know that the amount of information available about the external world is limited by the entropy of the input sensory signals. And, on the other hand, how much information about the sensory input the spike trains can provide is limited by the entropy of these spike trains, as considered below.

  – Entropy of spike count

    Here we ignore the temporal pattern of the spike trains and characterize each spike train only by the total number of spikes n in a time window T. The range of n is $0 \leq n < \infty$. The mean of n is

    $$\mu = \sum_{n=0}^{\infty} np(n) = rT$$

    where r is the firing rate representing number of spikes per unit time (e.g., msec).

    It can be shown that the distribution with maximum entropy is

    $$p(n) = (1 - e^{-\lambda})e^{-\lambda n} = \frac{\mu^n}{(\mu + 1)^{n+1}}$$

    where

    $$\lambda = ln(1 + \frac{1}{\mu})$$

    and the maximum entropy is

    $$H(n) = log(1 + \mu) + \mu log(1 + \frac{1}{\mu})$$

    Here the plot of $p(n)$ is very similar to that of the $p(x) = exp(-x/\mu)/\mu$ in the continuous case.

    To find the entropy per unit time, we divide $H(n)$ by T to obtain the *entropy rate* (bit per second) representing amount of information gained per unit time. Or, to find the information carried by each spike, we divide $H(n)$ by μ.

    Shown in the plot is the plot of entropy per spike as a function of μ.

– Entropy of spike train

To compute the entropy without ignoring the temporal pattern of the spike train, we divide the time window T into $N = T/\tau$ bins of width τ where τ is short to allow no more than one spike to fit in (but long enough so that a spike never needs to take two bins time). Now each spike train can be represented by a N-bit string of 0's and 1's. The number of 1's in the string is $N_1 = rT$, where r is the firing rate, number of spikes per unit time). And the probability for any bit to be 1 is $p = N_1/N$.

The total number of different equally likely patterns of the string is (choosing N_1 of N bins to be 1):

$$\frac{N!}{N_1!(N-N_1)!}$$

As we know the entropy of a random experiment is the logarithm of the total number of equally likely outcomes, the entropy of the string is

$$H = log\frac{N!}{N_1!(N-N_1)!} = logN! - logN_1! - log(N-N_1)!$$

To deal with the factorials, use Stirling's formula

$$n! \approx \sqrt{2\pi n}e^{-n}n^n \qquad \text{when } n \to \infty$$

or

$$logn! \approx n(logn-1) + log(2\pi n)/2 \approx n(logn-1)$$

With this approximation, the entropy can be simplified to

$$\begin{aligned} H &= -N[logp + (1-p)log(1-p)] \\ &= -\frac{T}{\tau}[(r\tau)log(r\tau) + (1-r\tau)log(1-r\tau)] \end{aligned}$$

Again dividing H by T, we get the entropy rate, and dividing H by the total number of spikes $N_1 = rT$, we get the information carried by each spike. The plot shows this information per spike as a function of τ (with $r = 50/msec$).

Comparing this plot with that of the spike count entropy, we see that each spike now carry more information.

Gaussian channel

Assume the stimulus is draw from a Gaussian distribution with variance σ_s^2 and the response is a linear function of the stimulus:

$$x = gs + n$$

where g is the gain and n is additive noise also normally distributed with variance σ_n^2, then the mutual information gained by observing x is

$$I = \frac{1}{2} log(1 + \frac{\sigma_s^2}{\sigma_n^2 g^2}) = \frac{1}{2} log(1 + SNR)$$

where SNR is the signal-to-noise ratio defined as the signal variance to the effective noise variance. If the stimulus is not Gaussian (but of the same variance), according to the maximum entropy property of Gaussian, the mutual information is less

$$I < \frac{1}{2} log(1 + SNR)$$