Physics 178/278 - David Kleinfeld - Winter 2024

Revised 20 February 2024

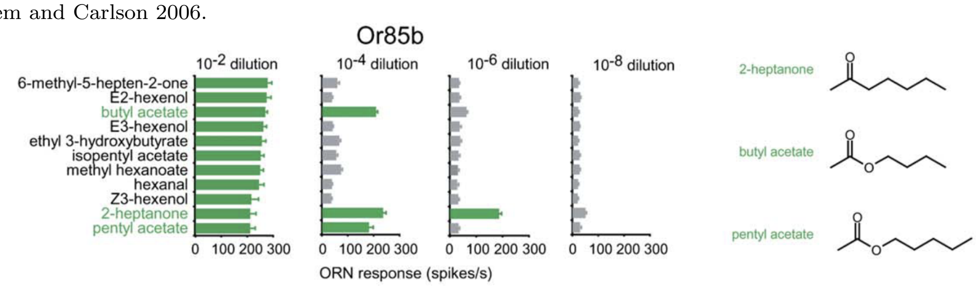
1 Information and sensory coding

Our focus has been on the pattern of spiking across neurons. We now ask if there are global constraints on the spiking. The central issue is if the overall rate is controlled by some aspect of the stimuli while the detailed pattern is controlled by another aspect. Thus the pattern of response can be solely dependent on features in the stimulus while the rate can code overall intensity. Let's see what the evidence is and consider simple models that give rise to the evidence.

Elissa Hallem and John Carlson measured the spiking output from 24 different olfactory receptor neurons in the fly in response to 100 different odorants, including natural and synthetic (Figures 1 and 2). This data was reanalyzed by Charles Stevens. He concludes that "...the probability distribution of [olfactory] sensory neuron firing rates across the population of odorant sensory neurons is an exponential for nearly all [pure] odors and odor mixtures, with the mean rate dependent on the odor concentration." Thus the aggregate population activity codes the concentration of an odorant, while the state of neuronal firing across the population, with different olfactory neuron receptor cells firing at rates that variety considerably for different odors, codes a particular odor per se (Figure 3). A similar result is found by reanalyzing Doris Tsao's data on spike rate in inferotemporal cortex in monkey for the coding of faces (Figure 4).

A exponential rate $P(r) \propto e^{-r/r_o}$, is telling. It corresponds to a simple idea from information theory where the average rate for a population is fixed and one asks for the distribution of rates, let call it $P(r)$, that satisfies this constraint.

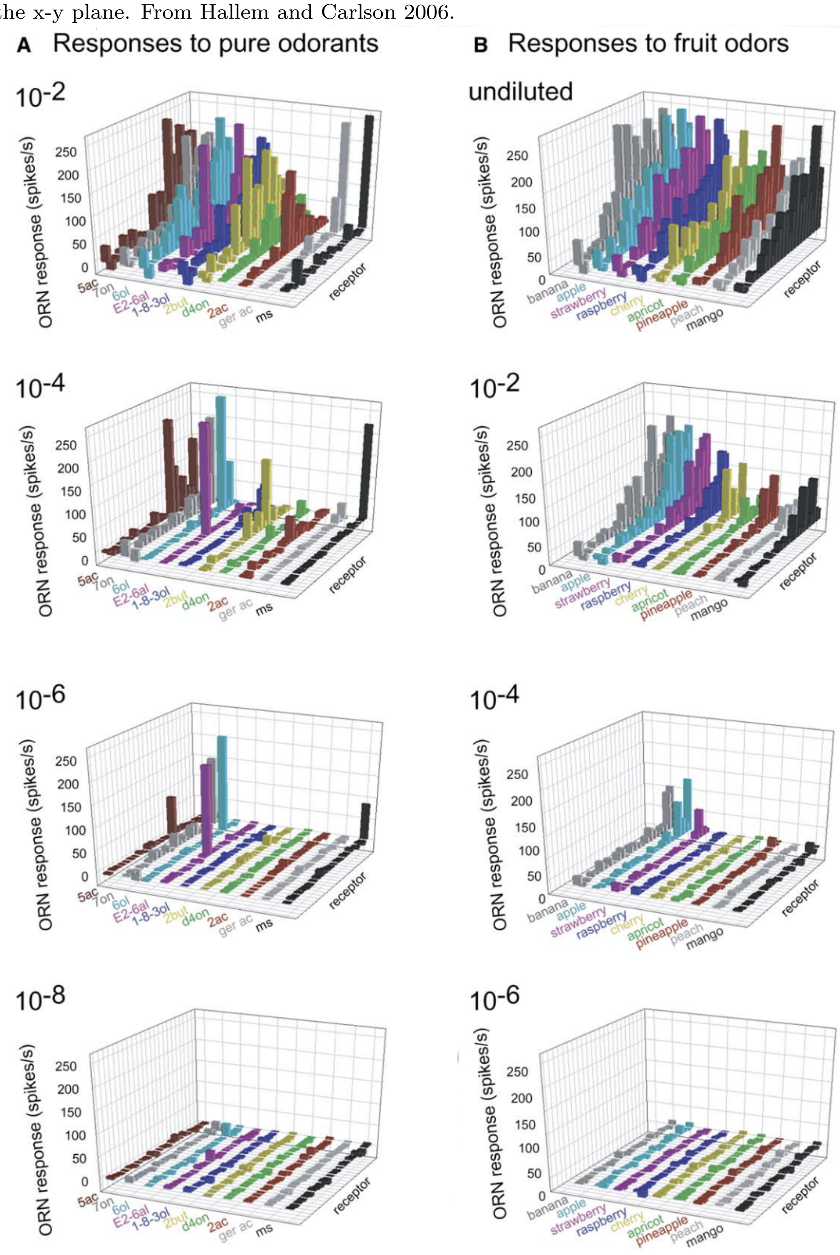
Figure 1: Specificities of broadly tuned olfactory receptor Or8rb across concentrations. The eleven odorants that elicited the strongest responses at a 10-to-2 dilution were tested at lower concentrations. Colored bars indicate strong responses (R100 spikes/s). Structures of odorants eliciting strong responses at low concentrations are indicated. Error bars = SEM. From Hallem and Carlson 2006.



Let us use information theory to get a handle on the probability distribution for spiking across a network. Let $P(n)$ re the probability of observing a response with spike count n, where this can be derived with the rate r over a time interval Δt as $n = r\Delta t$. We define a new quantity, the entropy $h[P(n)]$ as a measure of the surprise.

- We desire an increase in $h[P(n)]$ as $P(n)$ decreases. Thus $h[P(n)] \propto 1/P(n)$.

Figure 2: Odor coding across concentrations (A) Responses of receptors to pure odorants. 5ac = pentyl acetate, 7on = 2-heptanone, 6ol = 1-hexanol, E2-6al = E2-hexenal, 1-8-3ol = 1-octen-3-ol,v2but = ethyl butyrate, d4on = 2,3-butanedione, 2ac = ethyl acetate, ger ac = geranyl acetate, ms = methyl salicylate. (B) Responses to complex mixtures. Responses of each receptor to the diluent were subtracted from each odorant response. Inhibitory responses are apparent as bars extending below the x-y plane. From Hallem and Carlson 2006.



- We desire that the measure of $h[P(n)]$ that should be additive for multiple processes. Thus $h[P(n_1), P(n_2)] = h[P(n_1)] + h[P(n_2)]$.
- Together these imply Thus $h[P(n)] = -\log P(n)$. We take log and $\log_2$ unless noted.

Let's take an example of entropy (suggested by Tatyana Sharpee). In "The Good, the Bad, and the Ugly", Tuco Benedicto PacÃfico Juan MarÃa RamÃrez needs to search

Figure 3: Analysis of nine complex fruit odors at four odor concentrations each. (A) Thirty-six superimposed cumulative probability distributions of ORN firing rates for nine fruit extracts, each at four concentrations. Twenty-four ORNs were used to compile each distribution for one odor and one concentration. (B) The mean ORN firing rate (across nine odors at a single concentration each) as a function of relative concentration (dilution of fruit extract). The least-squares fitted line has a slope of 0.166. The vertical bars on the data points are the SDs for the nine odors at each odor concentration. (C) The superimposed cumulative distributions as a function of ORN rate that appear in panel A, replotted after scaling each distribution to the same mean(100 Hz). (D) The average of the 36 distribution functions from C with an exponential distribution function with a mean of 100 superimposed. From Stevens 2016.

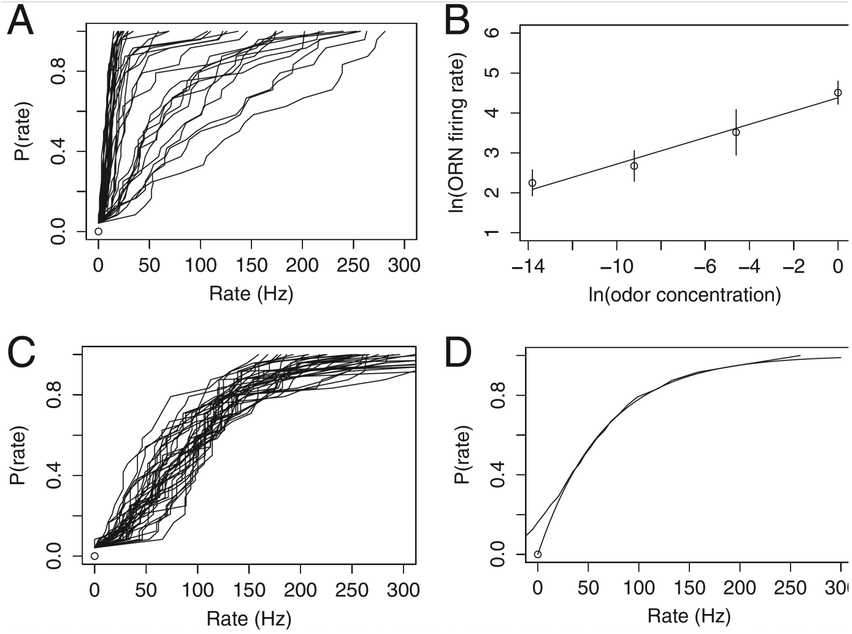
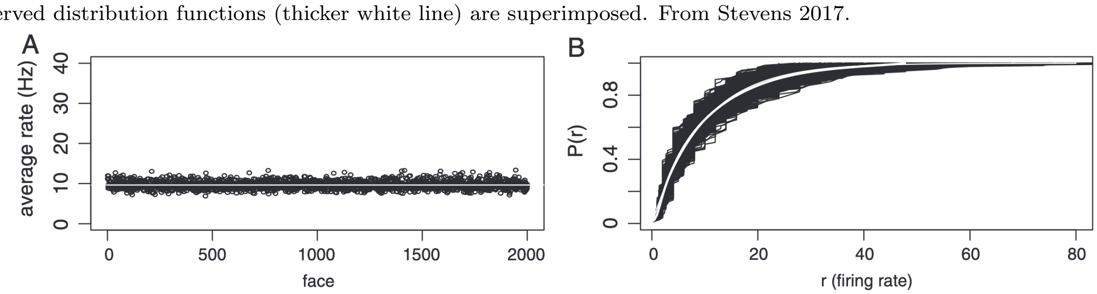


Figure 4: Constraints on the combinatorial face code. (A) The mean firing rate of active neurons as a function of face number. The white line represents the grand mean firing rate of 9.7 Hz calculated across the 2,000 individual mean rates. (B) Cumulative probability distribution of AM neuron rates. The abscissa gives the firing rate for each anterior medial (AM) neuron, and the ordinate is the cumulative probability. The distribution functions for 2,000 faces are superimposed (black band), and an exponential distribution with a mean of 9.7 Hz (thin white line) and the average across all 2,000 observed distribution functions (thicker white line) are superimposed. From Stevens 2017.



$N = 1025$ grave sites in Sad Hill Cemetery for buried treasure (Figure 5).

- Match at first site. Thus $P(n) = 1/1024$ to find treasure and $h[P(n)] = -\log(1/1024) = \log(1024) = 10$ bits.
- Miss at first site. Thus $P(n) = 1023/1024$ to not find treasure and $h[P(n)] = \log(1024/1023) = 0.0014$ bits.

Figure 5: Tuco running through Sad Hill Cemetery looking for the grave of "Arch Stanton". From Sergio Leone 1966.

- Miss at second site. Thus $P(n) = 1022/1023$ to not find treasure and $h[P(n)] = \log(1023/1022) = 0.0014$ bits.
- Match at third site. Thus $P(n) = 1/1022$ to find treasure and $h[P(n)] = \log(2022) = 9.9972$ bits. Note that $\log(1024/1023) + \log(1023/1022) + \log(2022) = \log(1024) = 10$ bits.
- In fact, the total entropy is always 10 bits. Let n be the number of empty sites. Then

$$\log\frac{N}{N-1} + \log\frac{N-1}{N-2} + \cdots + \log\frac{N-n}{N-n-1} + \log\frac{N-n}{N-n-1} = \log N$$

Let's take a second example the Shannon entropy, denoted H, which is the sum over the entropy and is a useful measure of a task.

$$\begin{aligned} H &= \int dnP(n)\ h[P(n)] \qquad (1.1) \\ &= \int dnP(n) \log(1/P(n)) \\ &= -\int dnP(n) \log P(n) \end{aligned}$$

Where the integral can be replaced $\sum_n$ for a discrete distribution. For a binary process, where the response in one of two counts (or ON versus OFF), we have $P(n) = p$ or $(1-p)$, so

$$H = -p \log p - (1-p) \log(1-p) \qquad (1.2)$$

This simple but interesting case can also be derived by considering how to arrange n objects in N sites, with one object per site (Box 1).

What is the value of p to maximize the entropy? We compute

$$\begin{aligned} 0 &= \frac{\partial H}{\partial p} \\ &= \log\left(\frac{p}{1-p}\right) \end{aligned} \tag{1.3}$$

for which $p = 1/2$.

Let's try something more interesting and maximize the entropy subject to a constraint on a function $f(n)$ using Lagrangian multipliers μ and λ.

$$F \;=\; -\int dnP(n)\log P(n) \;-\; \mu \int dn\; P(n) \;-\; \lambda\int dnP(n)f(n) \tag{1.4}$$

with

$$\int dnP(n) \;=\; 1 \tag{1.5}$$

and

$$\int dnP(n)f(n) = n_o \tag{1.6}$$

where n_o constrains We then differentiate

$$\begin{aligned} 0 &= \frac{\partial F}{\partial P} \\ &= -1 - \log P - \mu - \lambda f(n) \end{aligned} \tag{1.7}$$

so

$$P = e^{1+\mu}e^{-\lambda f(n))}. \tag{1.8}$$

From Equation 1.5

$$e^{-(1+\mu)} = \int dn\; e^{-\lambda f(n))} \tag{1.9}$$

so

$$P(n) = \frac{e^{-\lambda f(n)}}{\int dn\; e^{-\lambda f(n)}} \tag{1.10}$$

Then (See Box 2 for general; approach)

$$\begin{aligned} \int dnP(n)\; f(n) &= \frac{\int dnf(n)\; e^{-\lambda f(n)}}{\int dn\; e^{-\lambda f(n)}} \\ &= \frac{\int dn \frac{-\partial e^{-\lambda f(n)}}{\partial \lambda}}{\int dn\; e^{-\lambda f(n)}} \\ &= \frac{\frac{-\partial \int dn\; e^{-\lambda f(n)}}{\partial \lambda}}{\int dn\; e^{-\lambda f(n)}} \end{aligned} \tag{1.11}$$

Thus

$$\langle f(n)\rangle = \frac{\frac{-\partial \int dne^{-\lambda f(n)}}{\partial \lambda}}{\int dn\; e^{-\lambda f(n)}} \tag{1.12}$$

and for $\langle f(n)\rangle = n_o$ we can explicitly do the integrals over the integrals from 0 to ∞ and

$$\begin{aligned} n_o &= \frac{\frac{-\partial \frac{1}{\lambda}}{\partial \lambda}}{\frac{1}{\lambda}} \\ &= \frac{1}{\lambda} \end{aligned} \tag{1.13}$$

and

$$P(n) = \frac{e^{-n/n_o}}{n_o}. \tag{1.14}$$

We have retrieved an exponential distribution for the rate in terms of $r_o = n_o/\Delta t$.

Box 1 - Entropy from Choose

The number of ways to arrange n objects in N sites, with one object per site, is

$$\mathbf{C(N,n)} = \left(\begin{array}{c} n \\ N \end{array} \right) = \frac{N!}{(N-n)!n!}$$

We consider the limit $N \rightarrow \infty$ and $n \rightarrow \infty$ but the ratio n/N is taken as the probability of occupancy, p. Stirling's formula for $N!$ holds in the limit of large value N as $\log N! \approx N \log N - N$. The entropy is given by

$$\begin{aligned} H &= \frac{-1}{N} \log \left(\frac{N!}{(N-n)!n!} \right) \\ &= \frac{-1}{N} \left(\log N! - \log n! - \log (N-n)! \right) \\ &= \frac{-1}{N} \left(N \log N - N - n \log n + n - (N-n) \log (N-n) + (N-n) \right) \\ &= -p \log p - (1-p) \log (1-p) \end{aligned} \tag{1.15}$$

Box 2 - General approach to constraints using LaGrangian multipliers

We let

$$\frac{\partial \log f(x)}{\partial x} = \frac{1}{\log f(x)} \frac{\partial f(x)}{\partial x} \tag{1.16}$$

so

$$\begin{aligned} -\frac{\partial \log \left[\int dn\; e^{-\lambda f(n)} \right]}{\partial \lambda} &= \frac{-1}{\int dn\; e^{-\lambda f(n)}} \frac{\partial \int dn\; e^{-\lambda f(n)}}{\partial \lambda} \\ &= \frac{\int dn\; f(n)\; e^{-\lambda f(n)}}{\int dn\; e^{-\lambda f(n)}} \\ &= \int dn\; P(n)\; f(n) \\ &\simeq\; < f(n) > . \end{aligned} \tag{1.17}$$