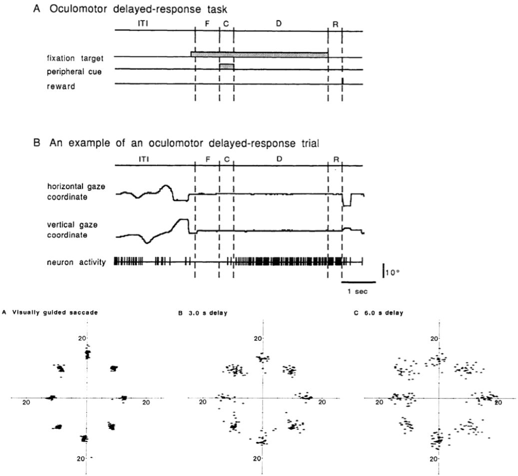
Neuroinformatics

April 26, 2012

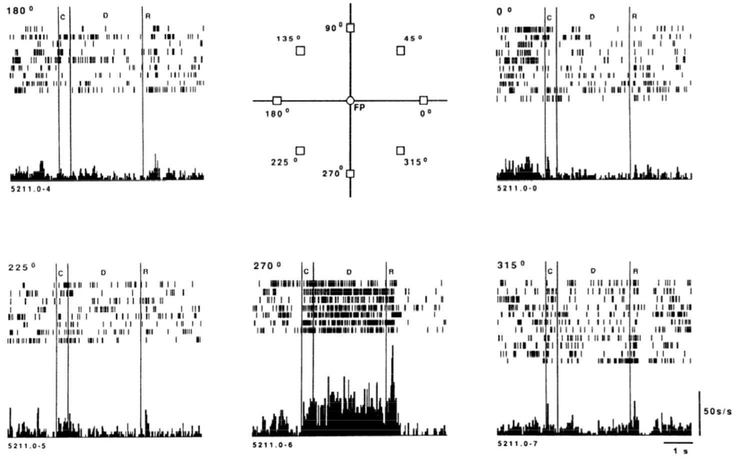
Lecture 10: Decoding and Encoding

Working memory by ongoing firing - sustained DNF buble

- F- fixation period (0.75s), C-cue period (0.5s), D - delay period (3-6 s), R - response period (0.5s) $\rightarrow$ reward



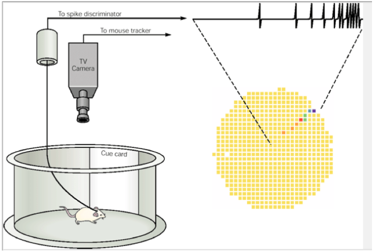
Directional delay period activity

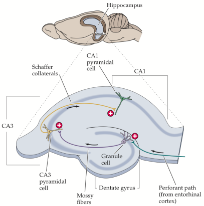


S. Funahashi, C.J. Bruce and P.S. Goldman-Rakic, Mnemonic coding of visual space in the monkeys dorsolateral prefrontal cortex, J Neurophysiol 61:331349, 1989

Place cells

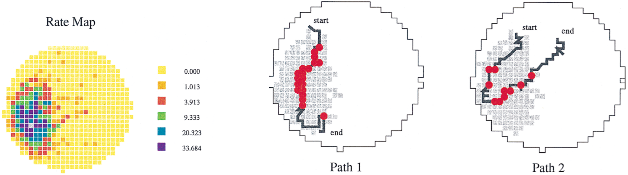
- Place cells are neurons in the hippocampus that exhibit a high rate of firing whenever an animal is in a specific location (pyramidal cells in CA1,CA4)
- On initial exposure to a new environment, place fields become established within minutes. The place fields of cells tend to be stable over repeated exposures to the same environment.
- Remapping - In a different environment, however, a cell may have a completely different place field or no place field at all





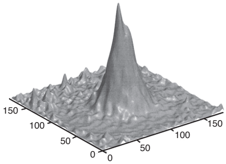
Place cells - 16 mins experiment

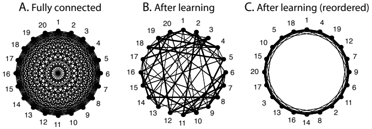
- colored circular region is an overhead view of a 76 cm diameter cylinder, each small square region (pixel) is about 2.5 cm squared, firing rate $\rightarrow$ total number of spikes fired in the pixel divided by the total time spent in the pixel.
- hungry rat ran around for 16 min chasing small food pellets, the black line indicates the rat's path and the red dots the locations at which action potentials were fired, action potentials were fired all along the second path even though the rat turned and ran out of the field in the direction opposite to its entry; this is an indication that the firing is not directionally selective.
- http://www.youtube.com/watch?v=PGHRDcPKio8



Place cells - is there any topography ?

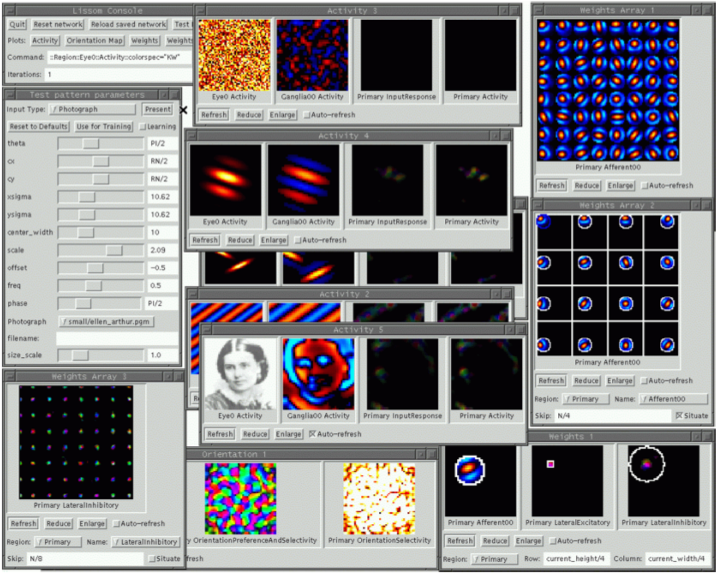
- no specific topography found with respect to neuron's maximal response to a particular place
- rearranging plot - neurons firing maximally in response to adjacent location $\rightarrow$ plot neurons adjacent to each other
- direction head cells $\rightarrow$ recurrent AAN simulation, high dimensionality - (i) before learning - equal weights for all nodes (ii) training- each node assigned (Gaussian profile) to preferred direction where fires maximally, competitive Hebb's rule (iii) strongly connected nodes adjacent to each other
- dimensionality was reduced to 1D model, networks self-organized to reflect the dimensionality of feature space





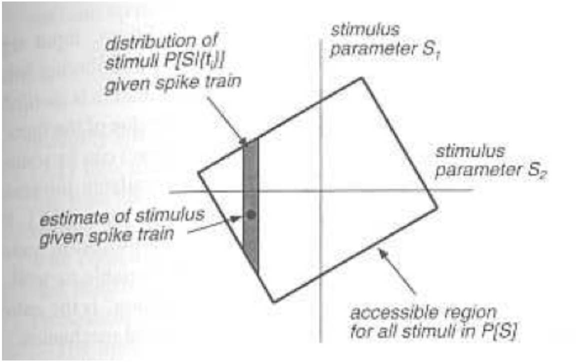
Topographica - general simulator for cortical maps

`http://topographica.org/`



Why information theory

- quantifying information that sensory neuron convey about the world
- how much information is spike train $t_i$transmitting, is this transmission large or small?
- stimulation estimate (dot) estimated by ML or Bayes



Communication channel as studied by Shanon

- (i) information depends on frequency of messages $p_i = P(y_i)$, (ii) independent information should be additive $(f(x,y) = f(x) + f(y))$: $p(x,y) = p(x)p(y)$
- logarithm has these characteristics: $I(y_i) = -log_2(p_i)$, how much we can learn relative what is known a priori.
- ENTROPY: average amount of information: $S(X) = -\sum_i p_i log_2(p_i)$
- Gaussian probability: $p(x) = \frac{1}{\sqrt{(2\pi)}\sigma} e^{\frac{(x-\mu)^2}{2\sigma^2}} \leftarrow$

  $S = \frac{1}{2}\log_2(2\pi e\sigma^2)$, depends only on variability (ENERGY)

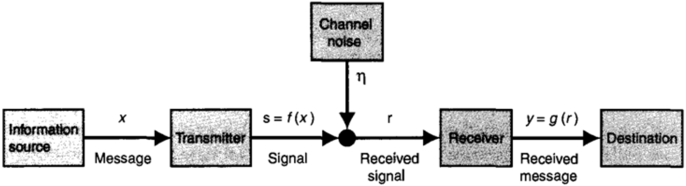
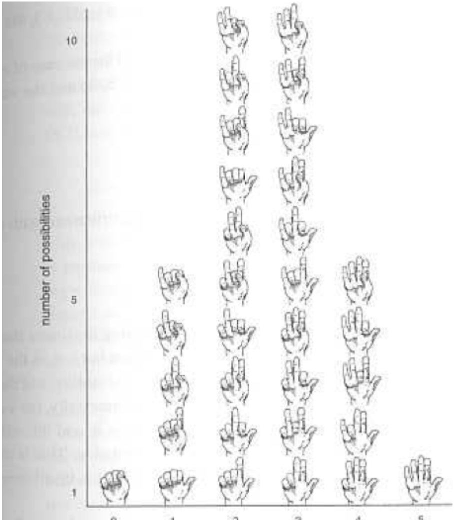


Fig. 5.6 The communication channel as studied by Shannon. A message x is converted into a signal $s = f(x)$ by a transmitter that sends this signal subsequently to the receiver. The receiver generally receives a distorted signal r consisting of the sent signal s convoluted by noise η. The received signal is then converted into a message $y = g(r)$ [adapted from C. Shannon, *The Bell System Technical Journal* 27: 379–423 (1948)].

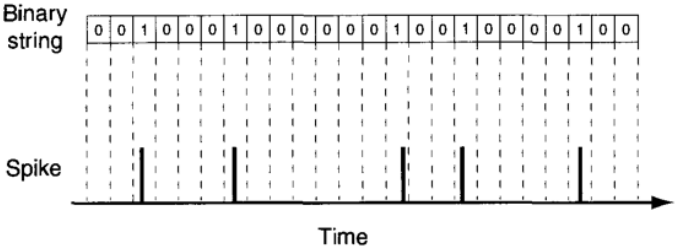
Example: ordering drink in restaurant

- rate coding: one hand can carry $log_2(6) = 2.58$bits of information, MORE ROBUST TO NOISE
- temporal coding: one hand can convey $2^5 = 32$ distinct messages, 5 bits of information



Entropy of Spike Train with temporal coding

- estimated in 1952 by MacKay and McCulloch, first application of information theory to nervous system, 4 years after Shannon
- $\bar{r}$: mean rate, $\Delta\tau$: time resolution, T : length of spike, occurrence of 1 $p = \bar{r}\Delta\tau$
- set of different strings, e.g. $1111111\ldots111111$
- counting the number of different spike trains that can be distinguished given our time resolution



Spike Train Calculation

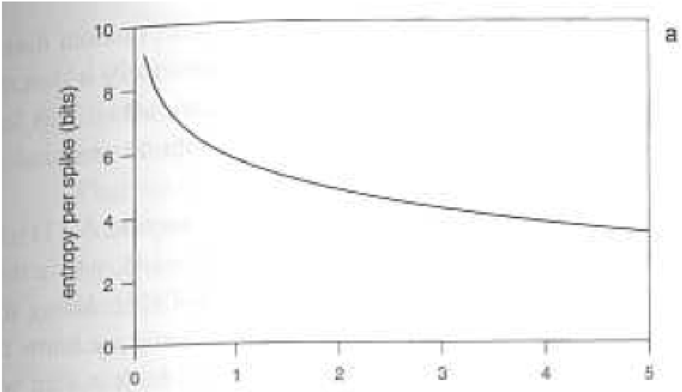
- total number of bins $N = N/\Delta\tau$, number of 1 (spikes) $N_1 = pN$, of 0 $N_0 = (1-p)N$, number of possible LARGE strings $N_{strings} = \frac{N!}{N_1!N_0!}$
- entropy $S = log_2 \frac{N!}{N_1!N_0!} = \frac{1}{ln2}(lnN! - lnN_1! - lnN_0!)$
- Stirling's approximation $lnx! = x(lnx - 1) + \ldots$, $ln_2(x) = ln(x)\ ln2$, all symbols K are equal, $S = -\sum_{i=1}^{K}(1/K)log_2(1/K) = log_2 K$

$$
\begin{aligned}
S &= \frac{1}{ln2}(lnN! - lnN_1! - lnN_0!) \\
&= \frac{1}{ln2}(NlnN - N_1 lnN_1 - N_0 lnN_0 - (N - N1 - N_0)), N = N_0 + N_1 \\
&= -\frac{1}{ln2}N(\frac{N1}{N}ln\frac{N1}{N} + \frac{N0}{N}ln\frac{N0}{N}) \\
&= \frac{N}{ln2}(plnp + (1-p)ln(1-p)) \\
&= -\frac{T}{\Delta\tau ln2}(\bar{r}\Delta\tau ln(\bar{r}\Delta\tau) + (1-\bar{r}\Delta\tau)ln(1-\bar{r}\Delta\tau)) \propto T, \bar{r}
\end{aligned}
$$

Entropy rate S/T approximation

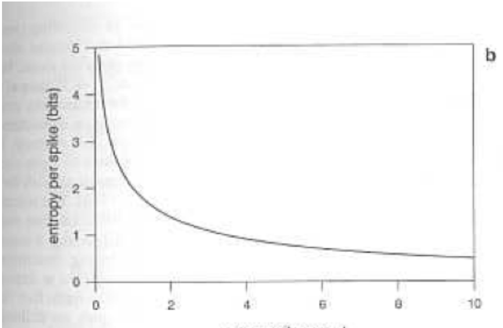
- approximating the entropy of spike trains
- limiting behaviour, $\Delta\tau$ is small, small bins, high resolution $\rightarrow$ Taylor series
- entropy is lager than 1 bits, $\bar{r} \sim 50s^{-}1$, $\Delta\tau \sim 1ms$, 5.76 bits per spike $\sim log_2(e/\bar{r}\Delta\tau)$, 288 bits/sec

$$S/T \approx \bar{r} log_2(\frac{e}{\bar{r}\Delta\tau})$$



Entropy of spike rate count

- different coding scheme - before the position was relevant! Now we want to calculate $S(spikecount) = -\sum_n p(n) log_2 p(n)$
- we are counting spikes in some large window $T \rightarrow$ measuring rate of spiking, $p(n)$ is probability of observing n spikes in window of length T
- $p(n) =?$, $\sum_n p(n) = 1$, average spike count $\langle n \rangle = \bar{r}T$, MAXIMAZING spiking count ENTROPY
- $p(n) \propto \exp(-\lambda n)$, $\lambda = ln(1 + (\bar{r}T)^-1)$, substituing
- $S(spikecount) \leq log_2(1 + \langle n \rangle) + \langle n \rangle log_2(1 + 1/\langle n \rangle)$ bits
- capacity 1 per bit, $\langle n \rangle \leq 3.4$ bits

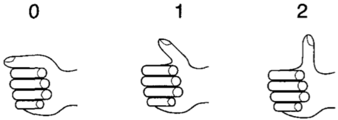


Channel capacity

- Mutual information $I_{mutual} = S(X) + S(Y) - S(X,Y)$, model of channel $y = gs + \eta$, where η is normal distribution and g is gain
- adding the noise to the signal itself and than transducing, $y = g(s + n_{eff})$, $n_{eff} = \eta/g$
- Example:resolution of our visual system, noise introduced by the motor system
- information transmission can be increased by increasing variability of the input signals. High variability of spike trains is well suited for transmission in noisy neural systems

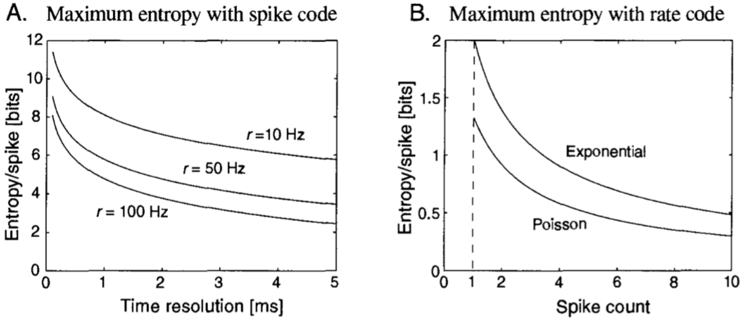
$$I = \frac{1}{2} log_2 \left(1 + \frac{\langle s^2 \rangle}{\langle \eta^2 \rangle / \langle g^2 \rangle} \right)$$

$$I = \frac{1}{2} log_2 \left(1 + \frac{\langle s^2 \rangle}{\langle n_{eff}^2 \rangle} \right) = \frac{1}{2} log_2 (1 + SNR)$$



Summary

- measuring entropy is difficult → estimating probability distributions
- small events in the entropy → large factor in entropy (log). Realiable measurements of rare events
- overestimating entropy due to potential miss of rare events with high information content



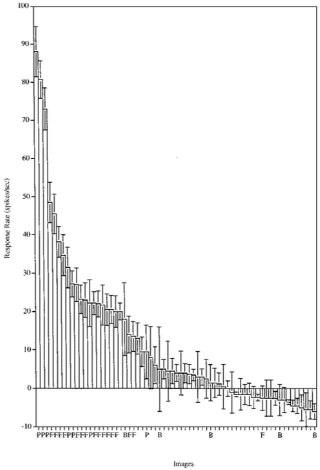
Entropy measured from single neuron

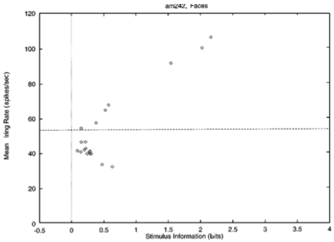
- 65 visual stimuli in macaques performing a visual task (23 monkey and human faces and 42 nonfaces images from real word), 14 face-selective neurons
- how much information is available about each stimulus in the set
- measuring firing rate in poststimulus phase (100 . . . 500ms).
- defining information between stimulus $S = \{s_i\}$ and responce $R = \{r_i\}$, I(s,R): amount of information about stimulus s, I(S,R) -average information gain

$$I(s, R) = \sum_r P(r|s) log_2 \frac{P(r|s)}{P(r)}$$

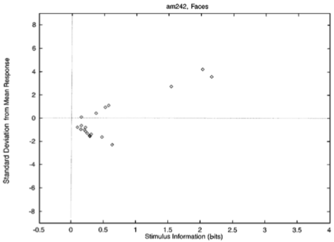
E. Rolls, Information in the Neuronal Representation of Individual Stimuli in the Primate Temporal Visual Cortex, Journal of Computational Neuroscience 4,309-333,1997

AM242 - quantitative analyses

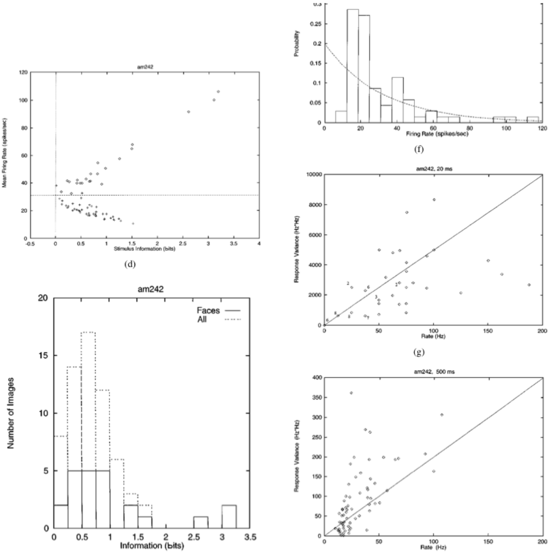




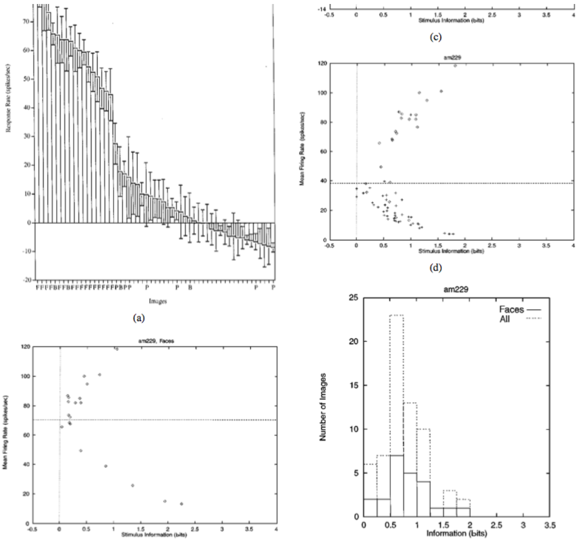
(b)



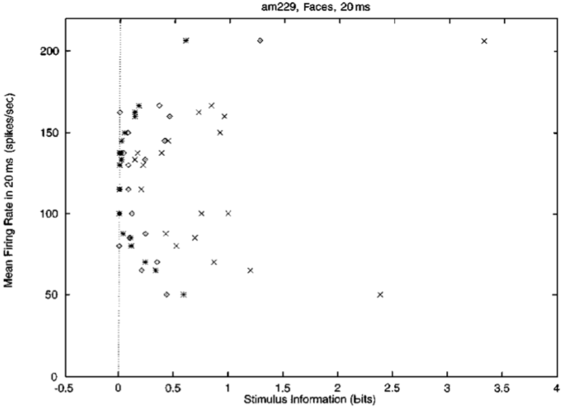
AM242 - quantitative analyses



AM242 - quantitative analyses



AM242 - any coding at the beginning?



Population coding (encoding and decoding)

Probability of neural response for a sensory input (encoding):

$$P(\mathbf{r}|s) = P(r_1^s, r_2^s, r_3^s, ...|s)$$

Decoding: $P(s|\mathbf{r}) = P(s|r_1^s, r_2^s, r_3^s, ...)$

Stimulus estimate: $\hat{s} = \arg\max_s P(s|\mathbf{r})$

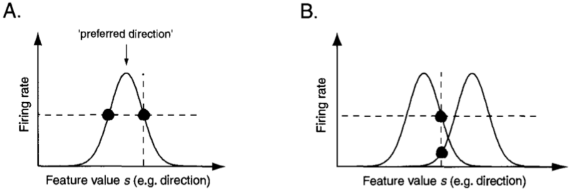
Bayes's theorem: $P(s|\mathbf{r}) = \frac{P(\mathbf{r}|s)P(s)}{P(\mathbf{r})}$

Likelihood: $P(\mathbf{r}|s), P = f(s)$

Decoding with response tuning curves

- we need at least two tuning curves $r_i = f_i(s)$to estimate the stimulus
- responses of neurons r_i are not correlated and tuning curves have Gaussian probability
- decoding using ML estimate - equivalent to least square fit

$$
\begin{aligned}
P(\mathbf{r}|s) &= \prod_i P(r_i|s) \\
P(r_i|s) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-}(r_i - f_s(s))^2/2\sigma_i^2 \\
\hat{s} &= \operatorname{argmin} \sum_i \left(\frac{r_i - f_i(s)}{\sigma_i} \right)^2
\end{aligned}
$$



Population vector decoding

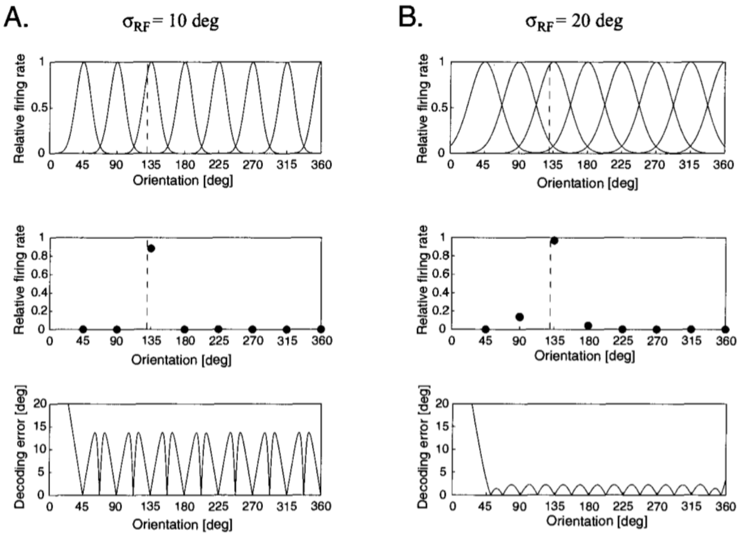
- e.g. Gaussian or cosine tuning curve: $f_i(s) = e^-(s - s_i^{pref}/2\sigma_{RF}^2)$, $\sigma_R F$ is receptibe field size
- Easy implementation in brain: dot product, normalization needed

$$\hat{s} = \sum_i r_i s_i^{pref}$$

$$\hat{r}_i = \frac{r_i - r_i^{min}}{r_i^{max}}$$

$$\hat{s}_{pop} = \sum_i \frac{\hat{r}_i}{\sum_j \hat{r}_j} s_i^{pref}$$

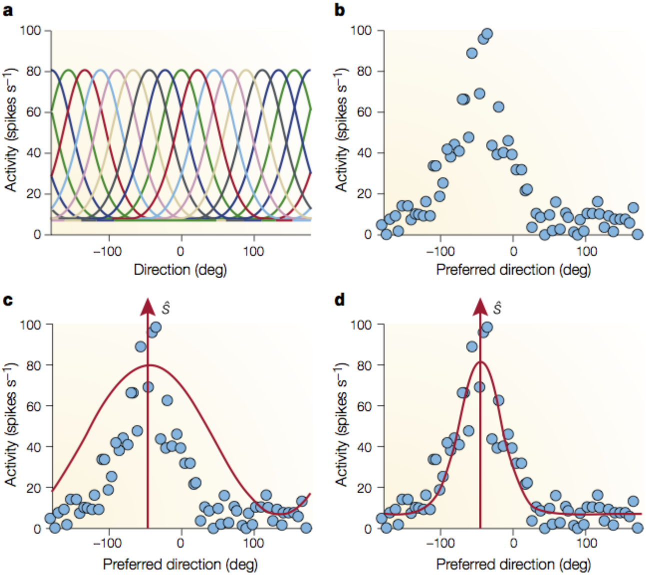
Population vector decoding - example



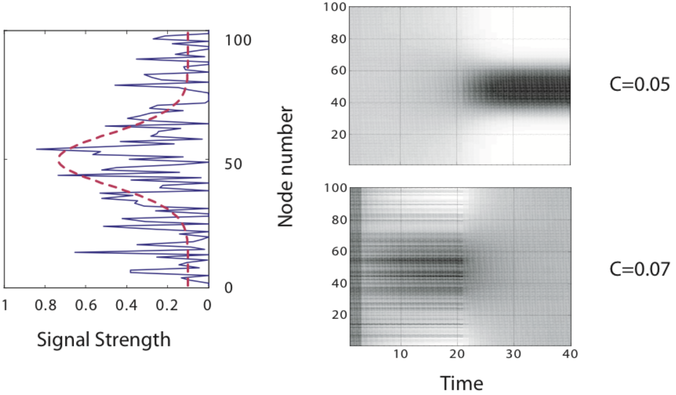
Example of coding model

- noisy model is used: $r_i = f_i(s) + \eta_i$, $f_i(s) = e^{-}(s - s_i^{pref}/2\sigma_{RF}^2)$

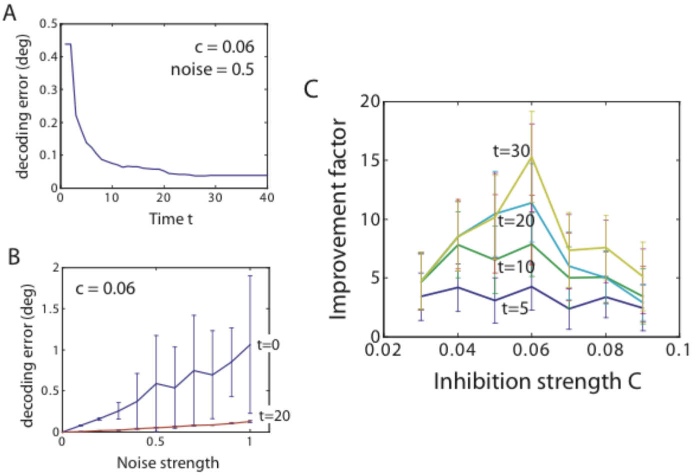
A. Pouget, Information Processing with population Codes, Nature Reviews, Neuroscience,2000



Implementations of decoding mechanisms with DNF



Quality of decoding



Further Readings

Fred Rieke (1995), **Spikes, exploring the neural code**, The MIT Press, 3st edition.

E. Rolls, Information in the Neuronal Representation of Individual Stimuli in the Primate Temporal Visual Cortex, Journal of Computational Neuroscience 4,309-333,1997

S. Funahashi, C.J. Bruce and P.S. Goldman-Rakic, Mnemonic coding of visual space in the monkeys dorsolateral prefrontal cortex, J Neurophysiol 61:331349, 1989

A. Pouget, Information Processing with population Codes, Nature Reviews, Neuroscience,2000