Calculating information in spike trains

Two methods:

- Information in spike patterns
- Information in single spikes

Calculating mutual information

Mutual information is the difference between
the total response entropy and the mean noise entropy:

$$I(S;R) = H[R] - \Sigma_s P(s)\, H[R|s)] \,.$$

Grandma's famous mutual information recipe

Take one stimulus *s* and repeat many times to obtain $P(R|s)$.

Compute variability due to noise: *noise entropy* $H[R|s]$

Repeat for all s and average: $\Sigma_s P(s)\, H[R|s)]$.

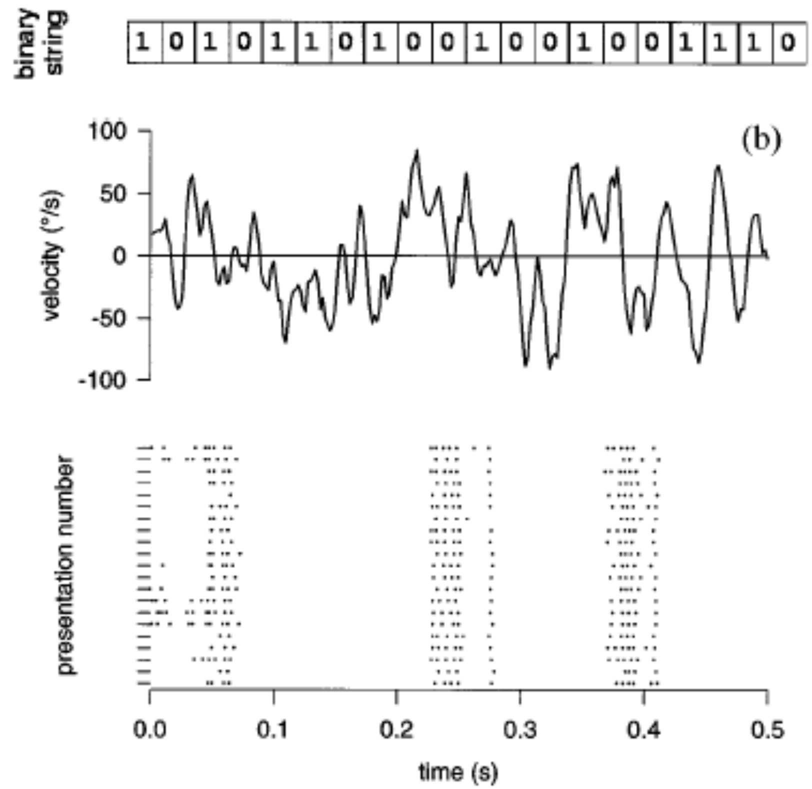
Compute $P(R) = \Sigma_s P(s)\, P(R|s)$ **and the total entropy** $H[R]$

Calculating information in spike patterns

So far only dealt with single spikes, or firing rates.

What information is carried by patterns of spikes?

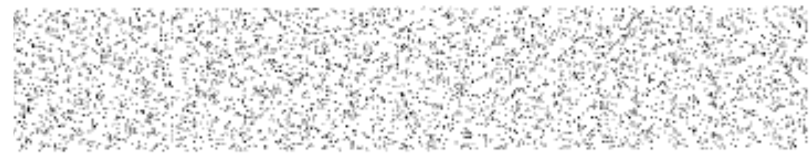
Analyze patterns of the code: how informative are they?



Strong et al., 1997

Calculating information in spike trains

Entropy:



- Binary words *w* with letter size Δt, length T.

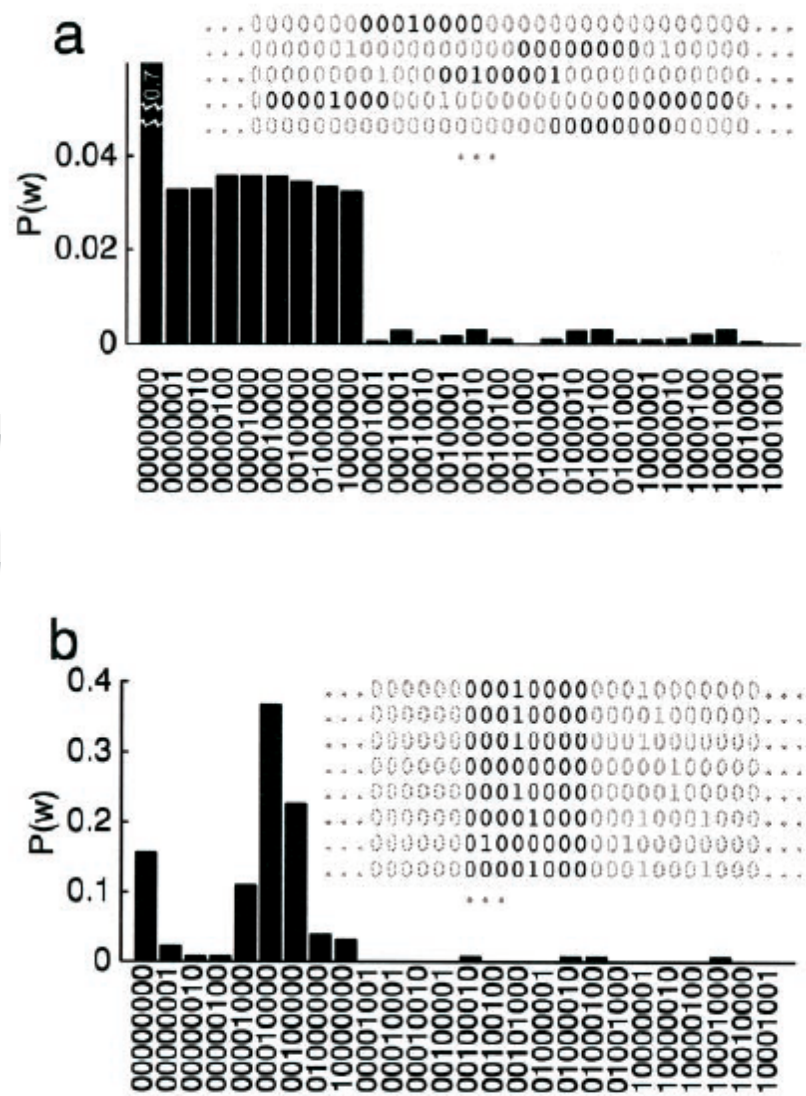
- Compute p(w_i)

$$H[w] = -\sum p(w_i)\log_2 p(w_i)$$

Strong et al., 1997; Reinagel and Reid, 2000

Calculating information in spike trains

Information :
difference between the total variability driven by stimuli and that due to noise, averaged over stimuli.



Reinagel and Reid, ‘00

Apply grandma's recipe

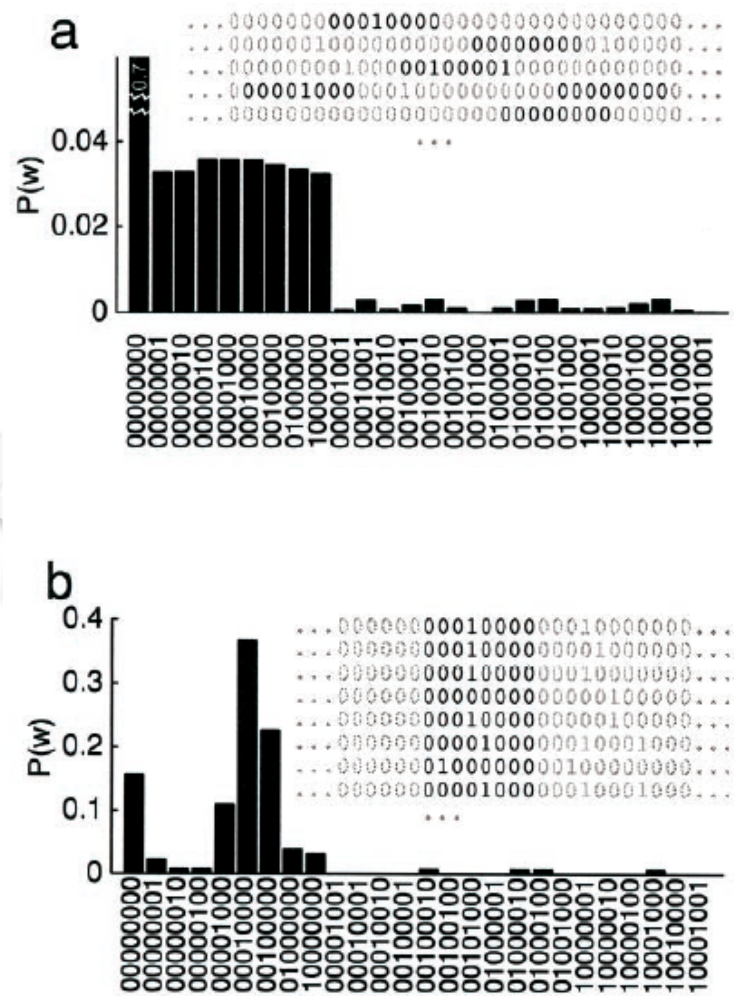
Take a stimulus sequence and repeat many times.

How to sample P(s)?
Average over s → average over time:

For each time in the repeated stimulus, get a set of words P(w|s(t)).

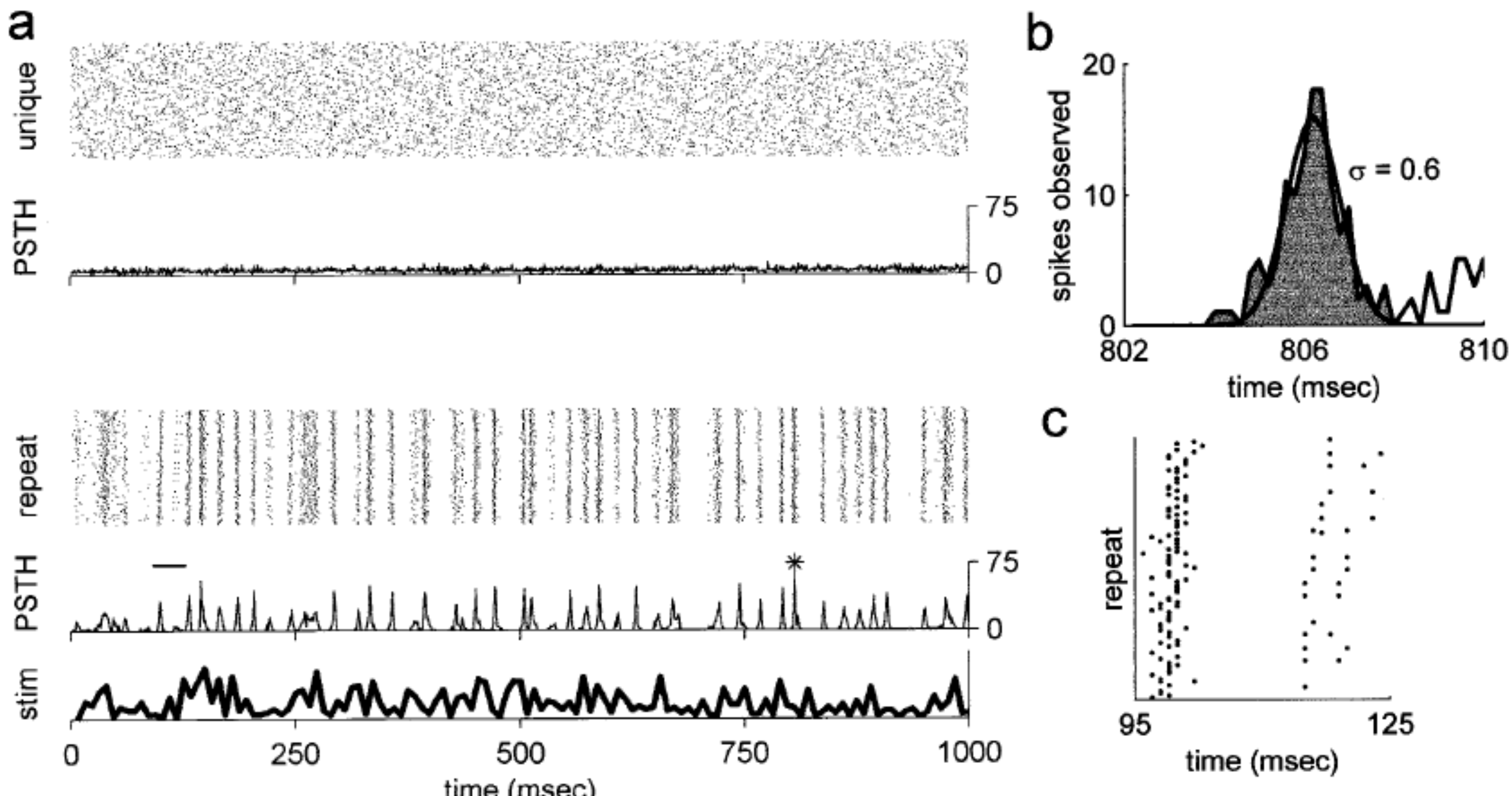
$$H_{noise} = \langle H[P(w|s_i)] \rangle_i.$$

Choose length of repeated sequence long enough to sample the noise entropy adequately.



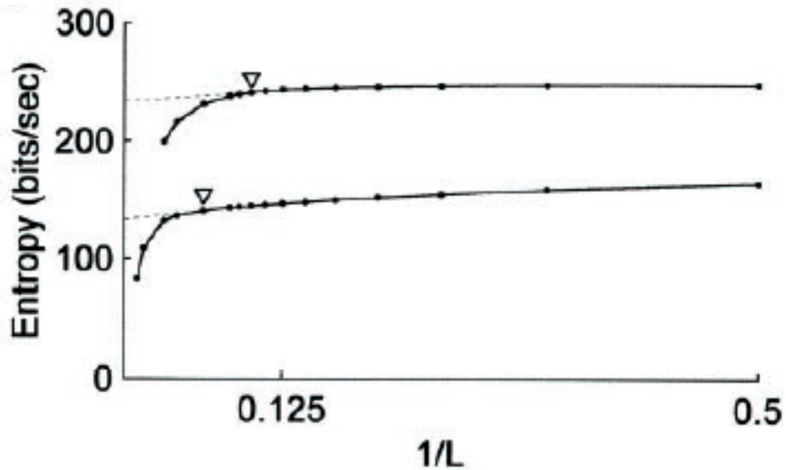
Reinagel and Reid (2000)

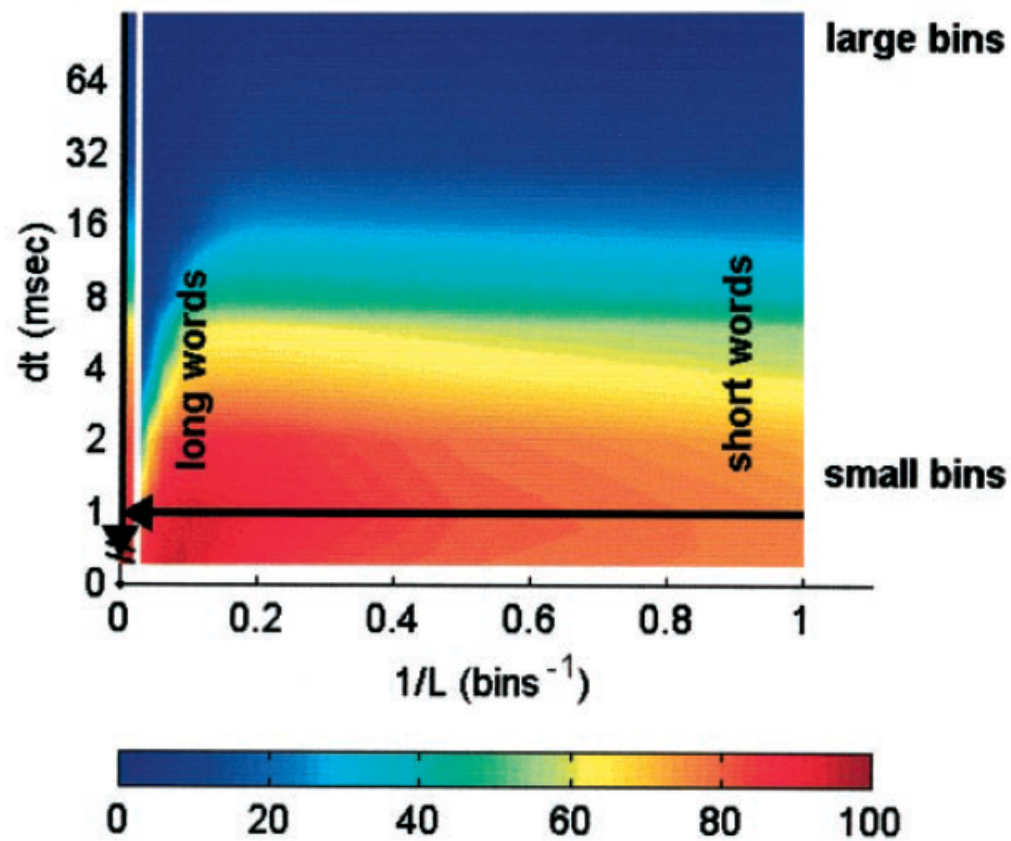
Calculating information in the LGN



Reinagel and Reid (2000)

Learning about the LGN's code





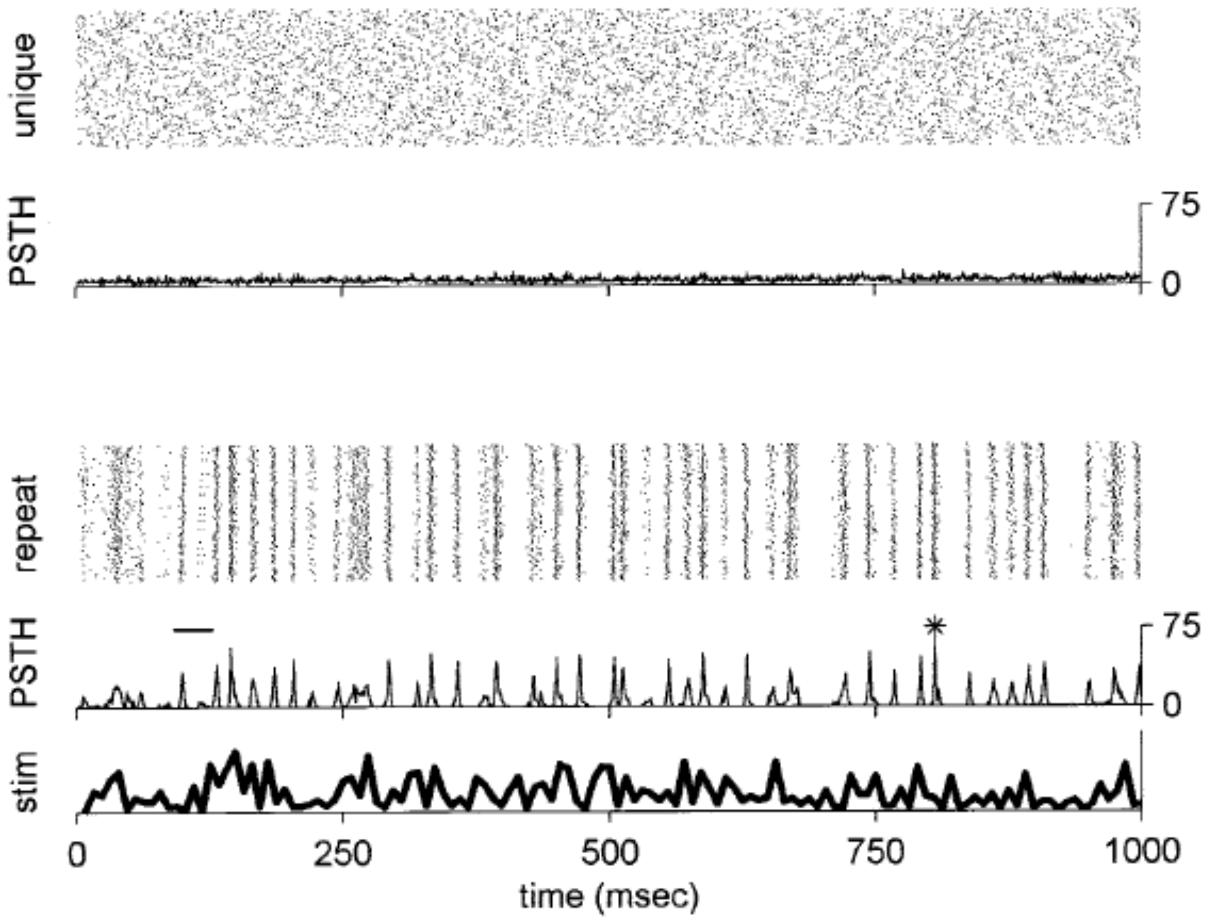
Reinagel and Reid (2000)

Sampling and bias

- Never enough data!
- Corrections for finite sample size
- Panzeri, Nemenman, ..

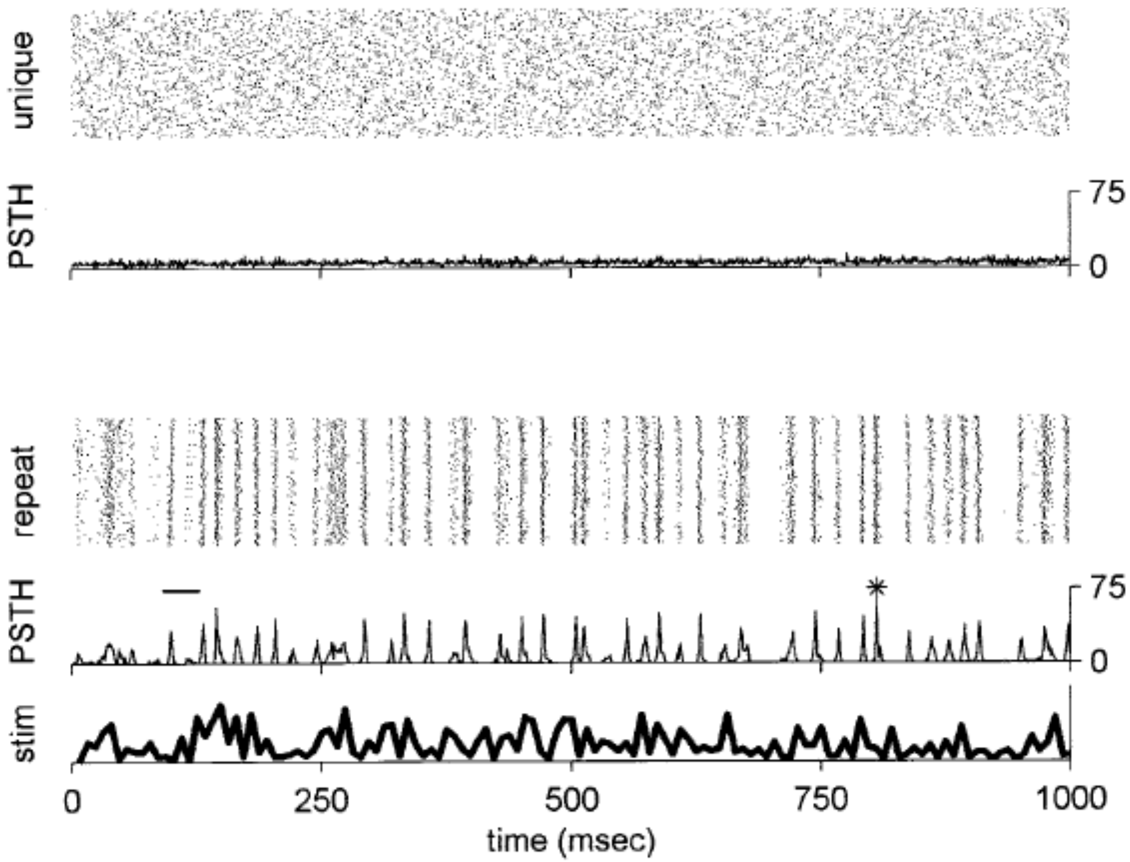
Information in single spikes

By how much does knowing that a particular stimulus occurred reduce the entropy of the response?



Brenner et al. (2000), data Reinagel and Reid (2000)

Information in single spikes



$$\begin{aligned} P(r=1) &= \bar{r}\Delta t, \\ P(r=0) &= 1-\bar{r}\Delta t, \end{aligned}$$

$$\begin{aligned} P(r=1|s) &= r(t)\Delta t, \\ P(r=0|s) &= 1-r(t)\Delta t. \end{aligned}$$

Information in single spikes

Now compute the entropy difference: $p = \bar{r}\Delta t \quad p(t) = r(t)\Delta t$

$$I(r,s) = -p\log p - (1-p)\log(1-p) + \qquad \leftarrow \text{Total}$$

$$+\frac{1}{T}\int_0^T dt\, \left[p(t)\log p(t) + (1-p(t))\log(1-p(t))\right]. \qquad \leftarrow \text{Noise}$$

Every time *t* stands in for a sample of *s*

A time average is equivalent to averaging over the *s* ensemble.

Ergodicity

Information in single spikes

$$I(r,s) = -p\log p - (1-p)\log(1-p) +$$ ← Total

$$+\frac{1}{T}\int_0^T dt\,\left[p(t)\log p(t) + (1-p(t))\log(1-p(t))\right].$$ ← Noise

Assuming $p \ll 1$, $\log(1-p) \sim p$ and using $\frac{1}{T}\int_0^T dt\, p(t) \to p$

$$I(r,s) = \frac{1}{T}\int_0^T dt\,\Delta t\, r(t)\log\frac{r(t)}{\bar{r}} + Var(p(t))/2ln2 + O(p^3).$$

To get *information per spike*, divide by $\bar{r}\Delta t$:

$$I(r,s) = \frac{1}{T}\int_0^T dt\,\frac{r(t)}{\bar{r}}\log\frac{r(t)}{\bar{r}}$$

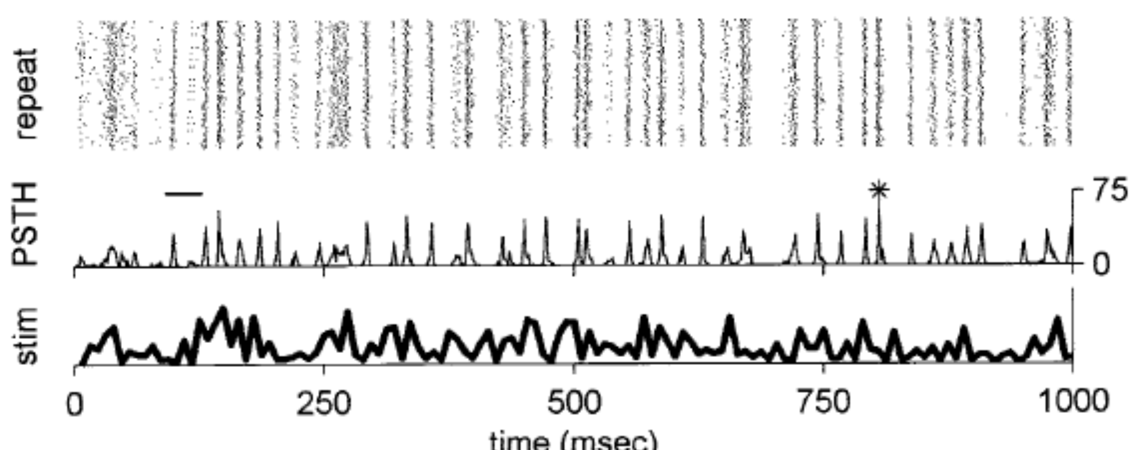
Information in single spikes

Information per spike: $$I(r,s) = \frac{1}{T}\int_0^T dt\, \frac{r(t)}{\bar{r}} \log \frac{r(t)}{\bar{r}}$$

- No explicit stimulus dependence (no coding/decoding model)
- The rate *r* does not have to mean rate of spikes; rate of any event.

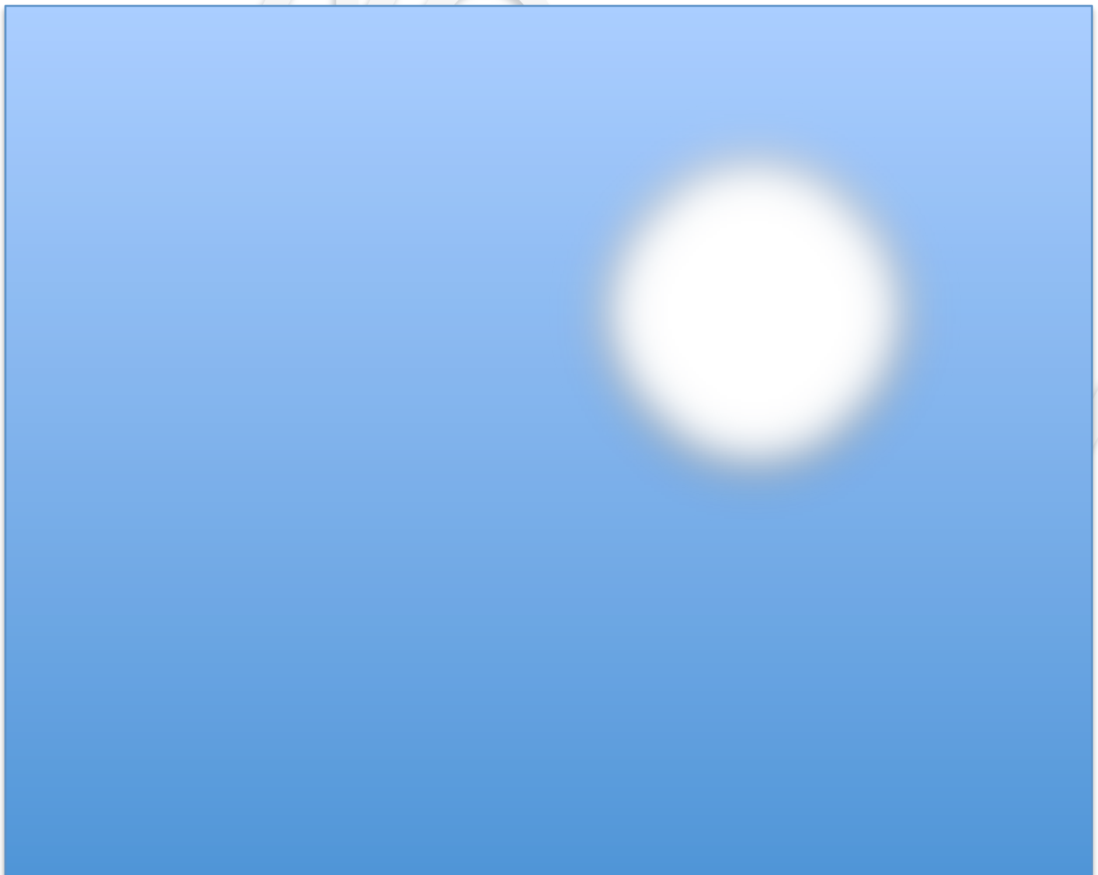
What limits information?

- spike precision, which blurs r(t)
- the mean spike rate.



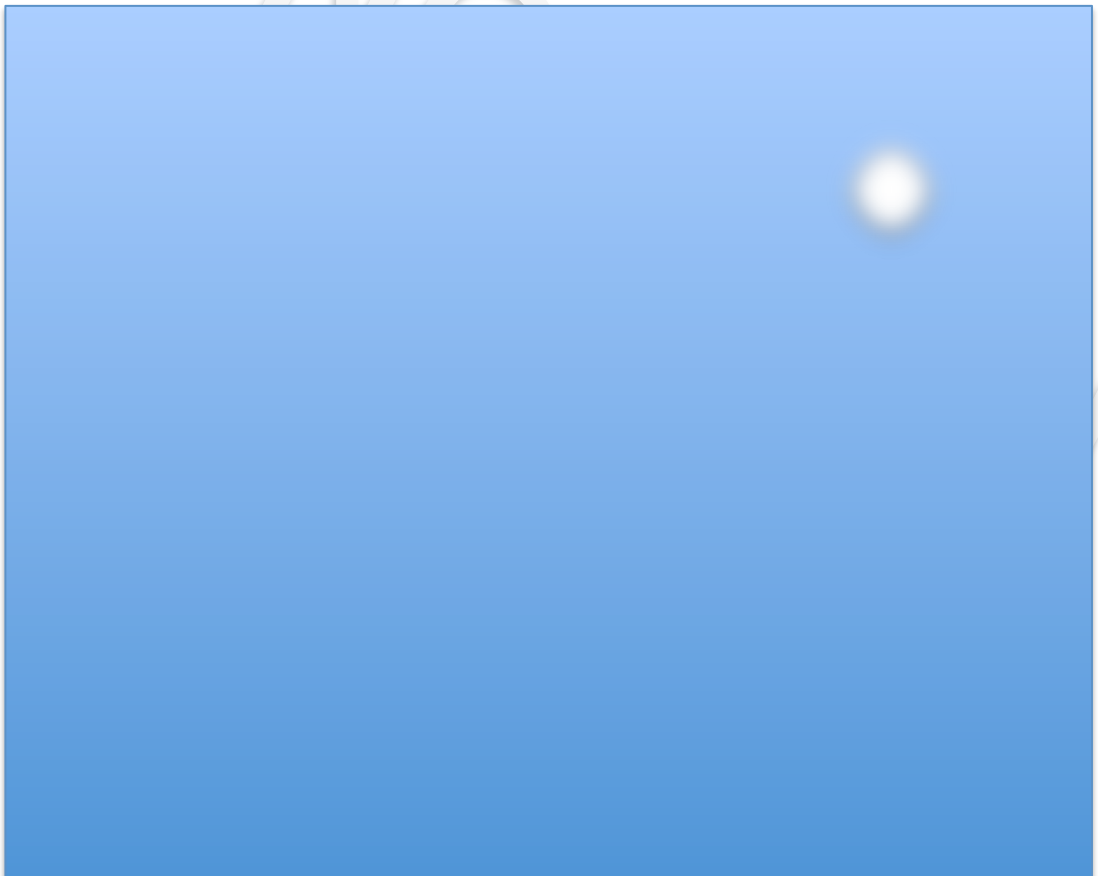
Information in single spikes

Information per spike: $$I(r,s) = \frac{1}{T}\int_0^T dt\, \frac{r(t)}{\bar{r}} \log \frac{r(t)}{\bar{r}}$$



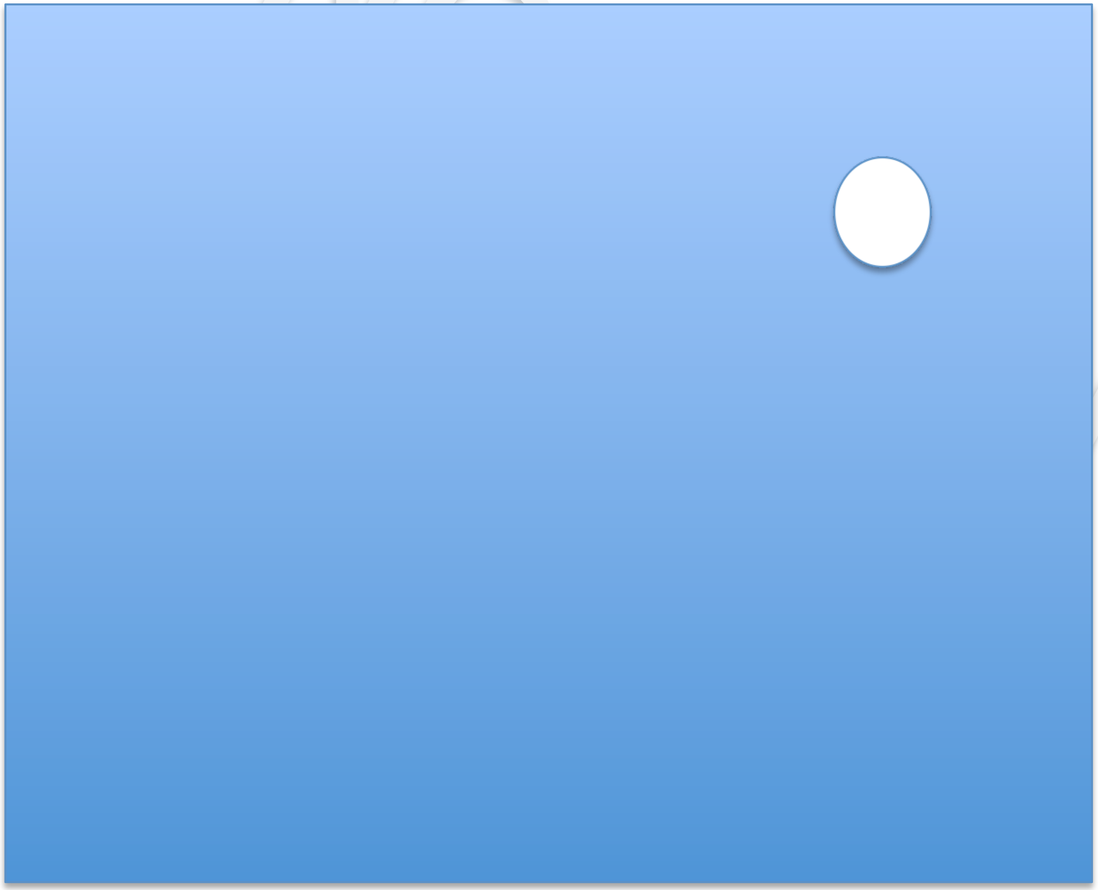
Information in single spikes

Information per spike: $$I(r,s) = \frac{1}{T}\int_0^T dt\, \frac{r(t)}{\bar{r}} \log \frac{r(t)}{\bar{r}}$$



Information in single spikes

Information per spike: $$I(r,s) = \frac{1}{T}\int_0^T dt\, \frac{r(t)}{\bar{r}} \log \frac{r(t)}{\bar{r}}$$



Next up: information and coding efficiency

- What are the challenges posed by natural stimuli?
- What do information theoretic concepts suggest that neural systems should do?
- What principles seem to be at work in shaping the neural code?