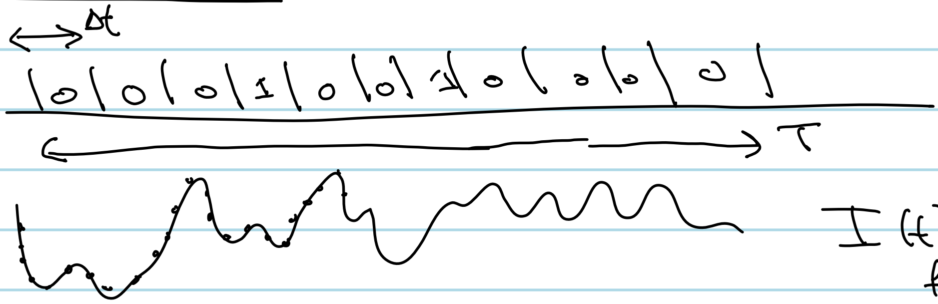
SPIKE TIME CODES AND INFORMATION IN SPIKE TRAINS:



$= \vec{W}$, "Spike word", length $T/\Delta t = N_W$

$I(t)$, or, with some fixed discretization, $\vec{I}$, vector of length $N_{\vec{I}}$.

Say stimuli come from some distribution $P(\vec{I})$

words " " " " $P(\vec{w} \mid \vec{I})$

Q1): How much information does a sample $\vec{w}$ tell us about $\vec{I}$?

Q2:) How does this information depend on timescale Δt over which we discretize spike train... ie how "temporal" is the code?

Define entropy, measure of "uncertainty" / width of a proba. distribution

$$S(\vec{I}) = -\sum_{\vec{I}} P(\vec{I}) \log_2 P(\vec{I})$$

eg: if $I \sim gsn(0, \sigma^2)$, $\quad S(I) = \frac{1}{2}\log_2(2\pi e \sigma^2)$

$= c + \log_2 \sigma$,

so entropy measures "width" of proba. distⁿ.

- **FACT** (easy to check): if I_1, I_2 are indep. Define $\vec{I} = (I_1, I_2)$

ie, $p(\vec{I}) = p(I_1, I_2)$. then $S(\vec{I}) = S(I_1) + S(I_2)$

- Likewise, if $\vec{I}$ is length-N vector w/ indep. components,

$$S(\vec{I}) = \sum_j^N S(I_j)$$

- How much is this entropy reduced by observing $\vec{W}$?

DEFINE: $S(\vec{I}|\vec{w}) = \sum_{\vec{I}} p(\vec{I}|\vec{w}) \log_2 p(\vec{I}|\vec{w})$

MUTUAL INFORMATION is average drop in uncertainty

$$MI(\vec{I}, \vec{w}) = S(\vec{I}) - \sum_{\vec{w}} p(\vec{w}) \, S(\vec{I}|\vec{w})$$

- FACT: $MI \geq 0$.
- FACT: Mutual info is symmetric, ie

$$MI(\vec{I}, \vec{w}) = MI(\vec{w}, \vec{I}) =$$

$$S(\vec{w}) - \sum_{\vec{I}} p(\vec{I}) \, S(\vec{w}|\vec{I})$$

Convenient! • $w_j \sim 0, 1$, or at least discrete ... $p(\vec{w})$ easier to sample.

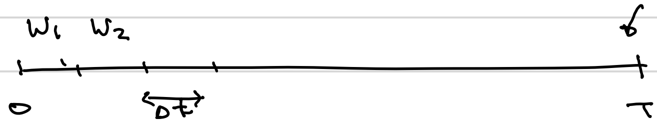
• Don't need to know statistics of $\vec{I}$ ... just when we are holding it fixed

$\longrightarrow S(\vec{w}|I)$, via $p(\vec{w}|I)$

or <u>not</u> $\longrightarrow$ $S(\vec{w})$, via $p(\vec{w}) = \int p(\vec{I})\, p(\vec{w}|\vec{I})\, d\vec{I}$.

- Strong, Koberle, van Steveninck, Bialek PRL 1998 $\quad$ $\vec{w}$ is length $T/\Delta t = N_w$

How does this depend on Δt?



- Via independence property above, <u>if</u> w_j indep from w_i ... (and stationary)

expect $S(\vec{w}) = k_1 N_w = c_1 T$

likewise $S(\vec{w}|\vec{I}) = k_1^I N_w = c_1^I T$

- If w_j is indep. from w_i for $|j-i| > K$, and that is true for "most" w_j given a w_i ... again expect

$$S(\vec{w}) \approx c_2 + k_i N_w = c_2 + c_1 T$$

for large N_w $\quad \approx c_1 T$ for large T

$$S(\vec{w}|\vec{I}) \sim c_2^I + c_1^I T$$

Does this verify ... so, plot $S(\vec{w})$ vs. T

$S(\vec{w}|\vec{I})$ vs. T ... for different (fixed) Δt

and look for linear <u>trends</u>!

POINT. Thus, expect entropy rate ~ constant.

$$\frac{S(\vec{w})}{T} = c_1 \quad , \quad \frac{S(\vec{w} \mid \vec{I})}{T} = c_1^{I}$$

Fig. 3 Strong et al... estimate that constant from data...

Key... Use measurements at small T, where can sample to estimate P(w), to extrapolate to large T estimate.

• SKIP THIS – DO WE NEED IT? •

POINT 2: Expect Entropy ... and MI ... to grow with N_w (ie, with $\frac{1}{\Delta t}$ or T)

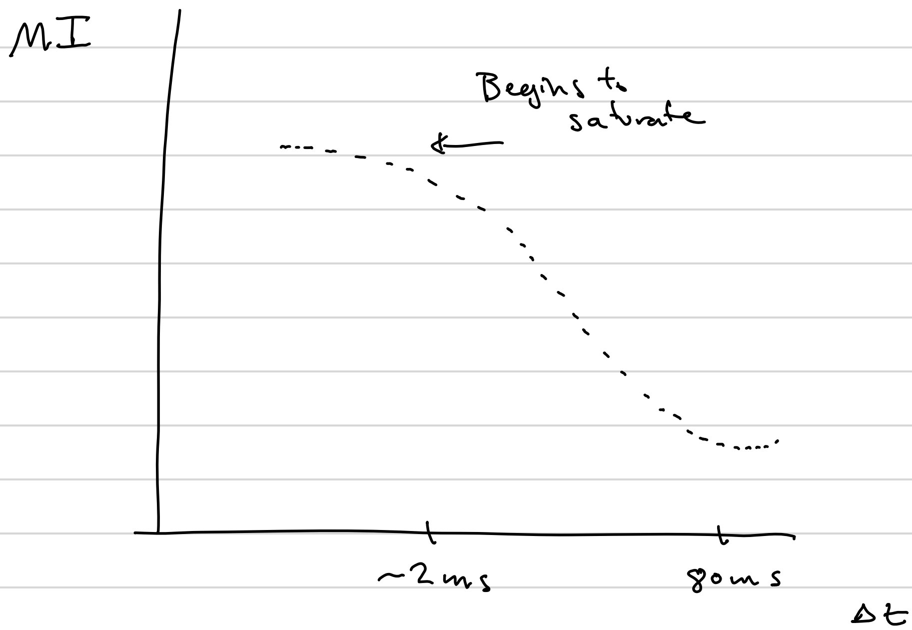
How can we compare MI at different Δt? Define...

$$\text{Efficiency } \varepsilon(\Delta t) = \frac{MI}{S(\vec{w})} = \frac{S(\vec{w}) - \sum_{I} p(I)\, S(w \mid I)}{S(\vec{w})}$$

$$\in [0,1]$$

= FRACTIONAL REDUCTION IN UNCERTAINTY ABOUT $\vec{w}$, FOR GIVEN TIMESCALE Δt.

• Thus equipped, plot...



Q: When does this begin to saturate?

Say... timing information increase info / decreases uncertainty about stimulus, at least down to that timescale.

See... Fig. 4, needs some unpacking.
(STRONG ET AL...)

For... fly H1 neurons w/ visual motion stim. $\vec{I}$.