Neural Data Analysis

Intro:

Sensory Coding and Computational Neurosciences.

Tutorial 1: Signal and Noise and Coherence

The Brain

$$\begin{tikzcd} & \vec{i}(t) \arrow[d, leftrightarrow] & \\ \vec{s}(t) \arrow[r] & \vec{r}(t) \arrow[r] & \vec{m}(t) \end{tikzcd}$$

Stimulus

Representation
Response

Behavior/Motor Response

Highly Dimensional, Dynamic, Recurrent, Multiple time scales,

Stochastic

Dimensionality Reduction, Visualization (modeling)

Stochastic Response

$\vec{r}(t)$ is characterized by a probability distribution

$E\left[\vec{r}(t)\right] = \overline{\vec{r}(t)}$ is the expected value or the mean of this distribution

$$\vec{r}(t) \sim N(\overline{\vec{r}(t)}, n)$$

Normal distribution

$$\vec{r}(t) \sim P(\overline{\vec{r}(t)})$$

Poisson distribution

Probability Distributions (Review)
Normal Distribution – 1d

Average across trials

$$p(r(t)) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{1}{2}\frac{(r(t)-\overline{r(t)})^2}{n^2}}$$

$$n^2 = \overline{n(t)^2}$$

Average across trials and time

Noise power

$$n = \sqrt{n^2}$$

Probability Distributions (Review)
Poisson Distribution

$$p(r(t)\Delta t = k) = \frac{(\bar{r}\Delta t)^k}{k!} e^{-(\bar{r}\Delta t)}$$

Stimulus is also probabilistic…

$\vec{s}(t)$ is characterized by a probability distribution

$p(\vec{s}(t), \vec{r}(t))$ fully describes sensory coding

But i) curse of dimensionality and ii) no insight

3 Approaches

- Encoding

$$\overbrace{\vec{r}(t)}^{Est} = f(\vec{s}(t)) \quad \text{Or} \quad \vec{s}(t) \xrightarrow{f} p(\vec{r}(t))$$

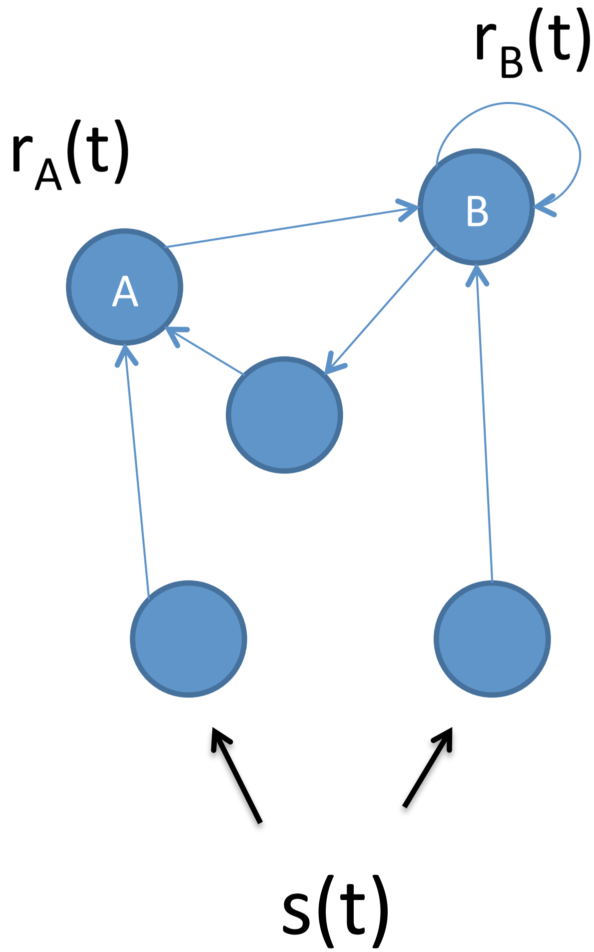
- Decoding

$$\overbrace{\vec{s}(t)}^{Est} = g(\vec{r}(t)) \quad \text{Or} \quad \vec{r}(t) \xrightarrow{g} p(\vec{s}(t))$$

- Information/Ceiling Values

$$I(S,R) = \iint p(\vec{s}(t),\vec{r}(t)) \log_2 \frac{p(\vec{s}(t),\vec{r}(t))}{p(\vec{s}(t))p(\vec{r}(t))}$$

Sensory Coding: Encoding

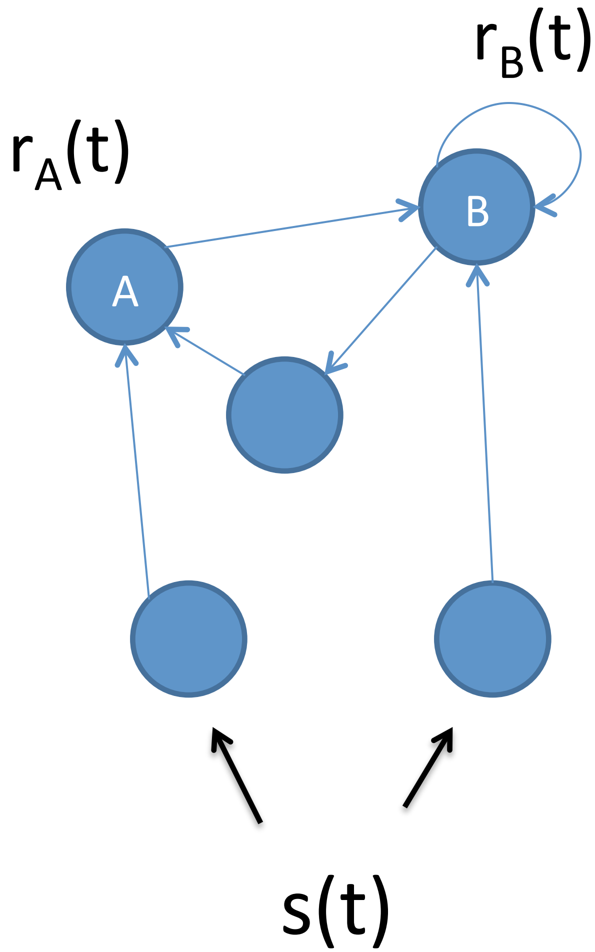


$$\vec{r}(t) = \{r_A(t); r_B(t)\}$$

$$\overbrace{\vec{r}(t)}^{Est} = f(\vec{s}(t))$$

$$\vec{s}(t)$$

Sensory Coding Decoding

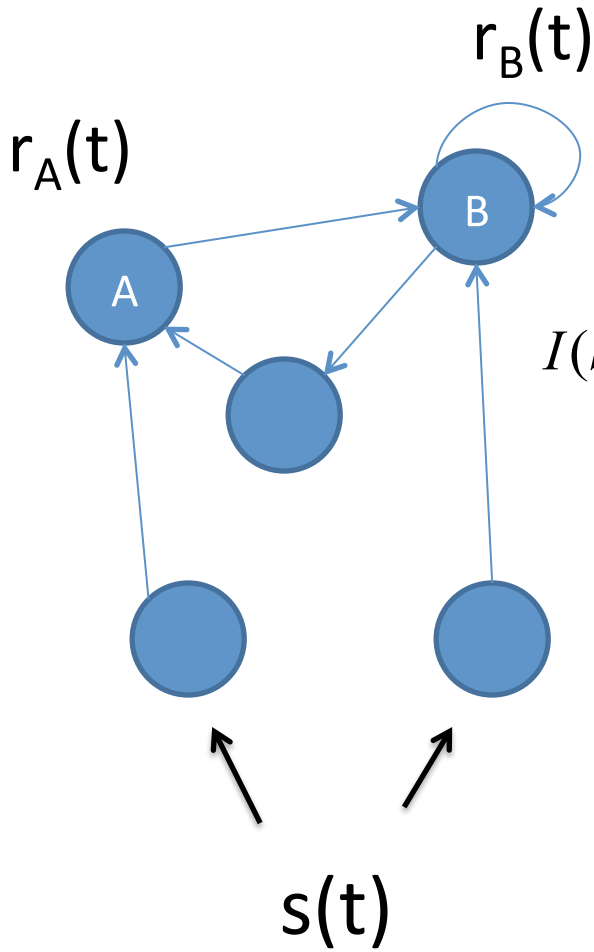


$$\vec{r}(t) = \{r_A(t); r_B(t)\}$$

$$\overbrace{\overline{\vec{s}(t)}}^{Est} = g(\vec{r}(t))$$

$$\vec{s}(t)$$

Sensory Coding: Mutual Info



$$\vec{r}(t) = \{r_A(t); r_B(t)\}$$

$$I(S,R) = \iint p(\vec{s}(t),\vec{r}(t)) \log_2 \frac{p(\vec{s}(t),\vec{r}(t))}{p(\vec{s}(t))\,p(\vec{r}(t))}$$

$$\vec{s}(t)$$

3 Approaches

- Encoding

Starting this PM

$$\overbrace{\overline{\vec{r}(t)}}^{Est} = f(\vec{s}(t))$$

Or

$$\vec{s}(t) \xrightarrow{f} p(\vec{r}(t))$$

- Decoding

$$\overbrace{\overline{\vec{s}(t)}}^{Est} = g(\vec{r}(t))$$

Or

$$\vec{r}(t) \xrightarrow{g} p(\vec{s}(t))$$

- Information/Ceiling Values

NOW!

$$I(S,R) = \iint p(\vec{s}(t),\vec{r}(t)) \log_2 \frac{p(\vec{s}(t),\vec{r}(t))}{p(\vec{s}(t))\,p(\vec{r}(t))}$$

Combining approaches…

- Encoding Model Validation

$$I\big(f(s),r\big) \le I(s,r)$$

- Neural Code

$$I\big(s,g_1(r)\big) \le I(s,r)$$

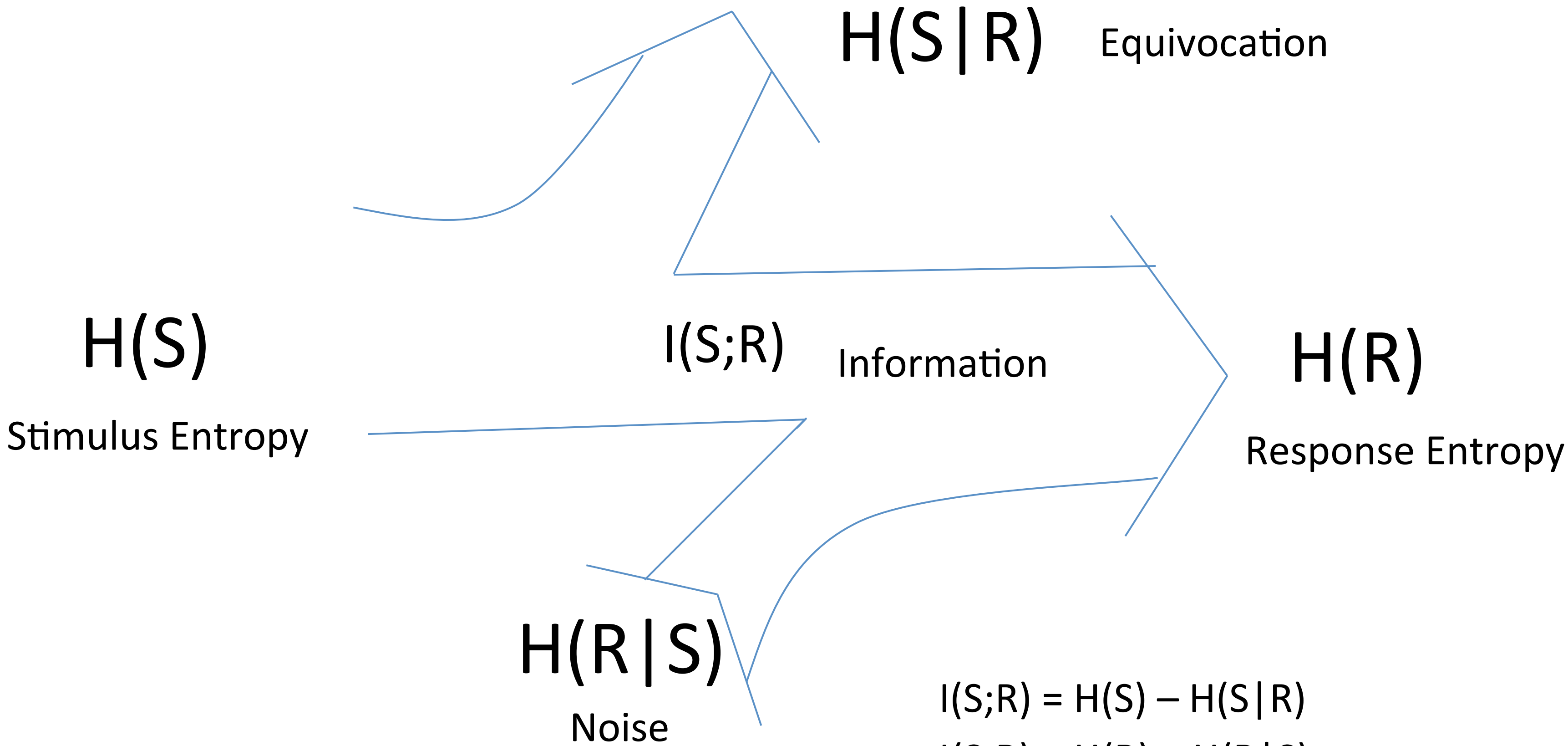
- Efficiency, Redundancy in ensemble responses

$$I(s,r) \big/ H(r)$$

$$\frac{I(S;R_1,R_2,\ldots,R_M)}{\sum_{m=1}^{M} I(S;R_m)}$$

Any signal out there??

Information and Entropies

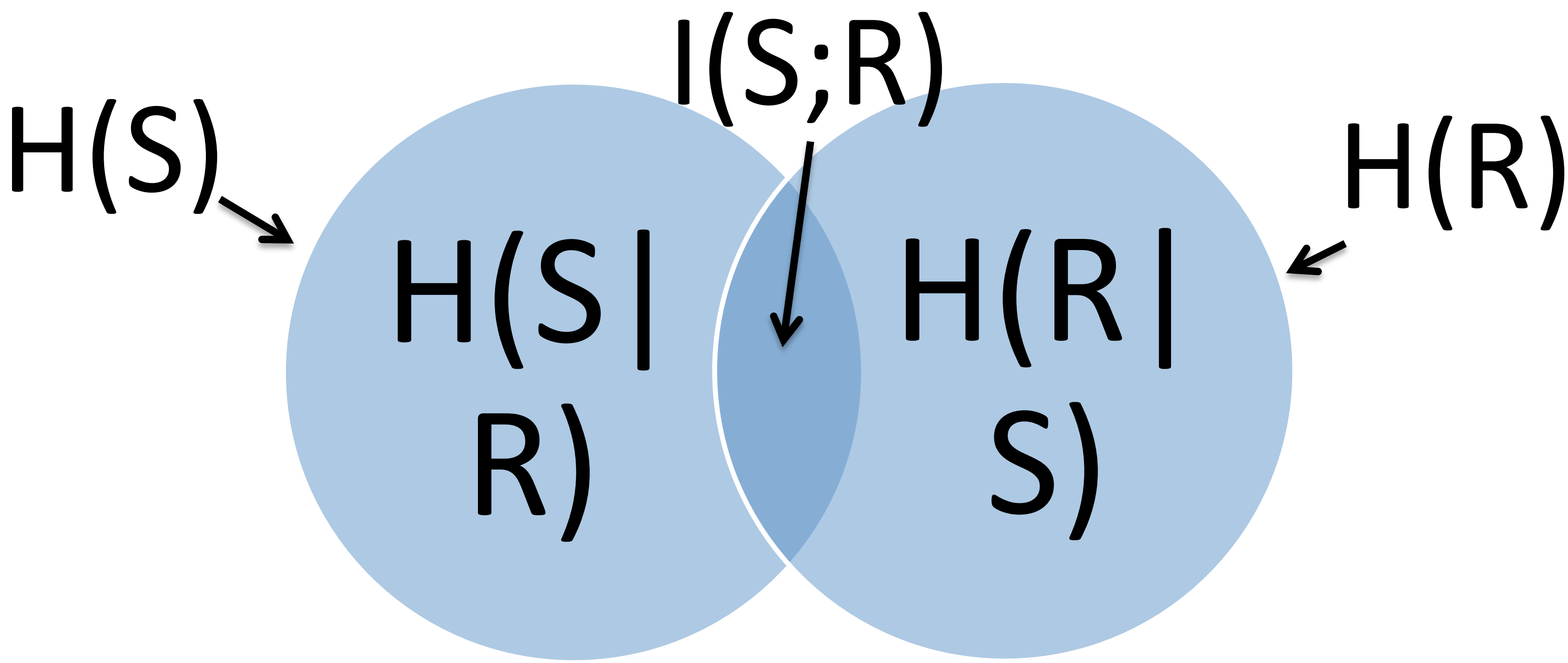


$$I(S;R) = H(S) - H(S|R)$$

$$I(S;R) = H(R) - H(R|S)$$

$$I(S;R) = H(S) + H(R) - H(R,S)$$

Information and Entropies



$I(S;R) = H(S) - H(S|R)$

$I(S;R) = H(R) - H(R|S)$

$I(S;R) = H(S) + H(R) - H(R,S)$

Entropy

- Uncertainty about source.
- Number of items of equal probability.
- Log 2 units or bits.

$$H(S) = \log_2(N)$$

$$H(S) = \sum_i p(s_i)\log_2\left(\frac{1}{p(s_i)}\right)$$

$$H(S) = \sum_i - p(s_i)\log_2(p(s_i))$$

Formulas for Information

I(S;R) = H(S) – H(S|R)

$$H(S) = \sum_i -p(s_i)\log_2\big(p(s_i)\big)$$

$$H(S\mid R) = \sum_j p(r_j)\sum_i -p(s_i\mid r_j)\log_2\big(p(s_i\mid r_j)\big)$$

$$I(S;R) = \sum_{i,j} p(s_i,r_j)\log_2\left(\frac{p(s_i\mid r_j)}{p(s_i)}\right)$$

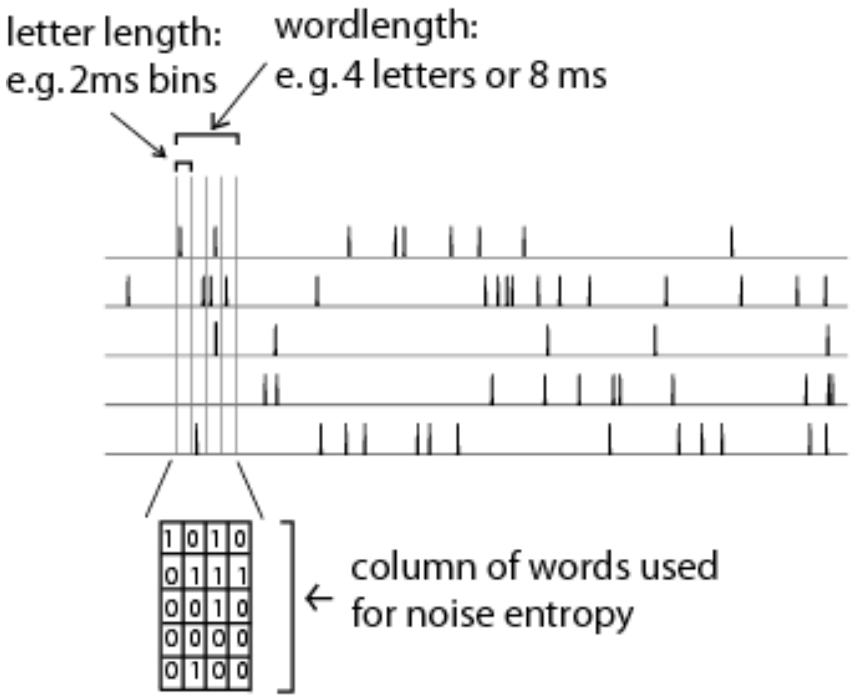
Formulas for Information

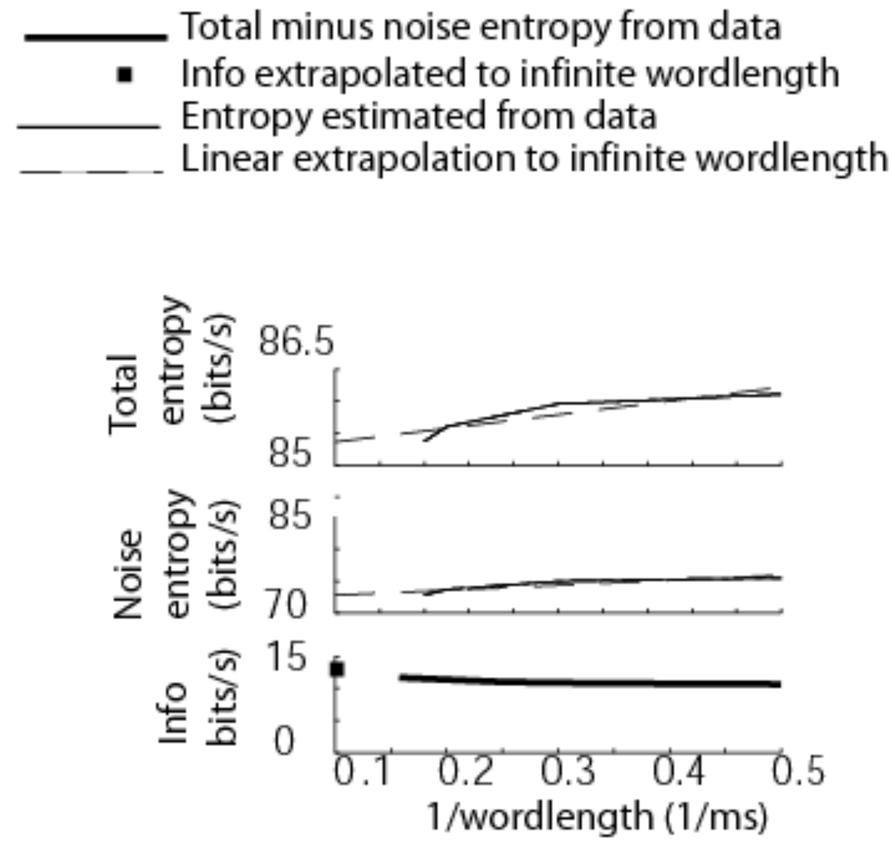
I(S;R) = H(S) – H(S|R) $\longrightarrow$ $$I(S;R) = \sum_{i,j} p(s_i, r_j)\log_2\left(\frac{p(s_i \mid r_j)}{p(s_i)}\right)$$

I(R;S) = H(R) – H(R|S) $\longrightarrow$ $$I(S;R) = \sum_{i,j} p(s_i, r_j)\log_2\left(\frac{p(r_j \mid s_i)}{p(r_j)}\right)$$

I(R;S) = H(S) + H(R) – H(R,S) $\longrightarrow$ $$I(S;R) = \sum_{i,j} p(s_i, r_j)\log_2\left(\frac{p(r_j, s_i)}{p(s_i)p(r_j)}\right)$$

Direct Method for estimating MI





H(Total)

H(Noise)

I=H(Total)-H(Noise)

$$I(s,r) = H(r) - H(r \mid s)$$

I_{LB}=64 bits/s → I=75 bits/s

Strong et al. *Phys Rev Letters* 1998

Dimensionality Reduction for Information Calculation

Today

Gaussian
Additive Noise

$$r(t) = Sig(t) + N(t) = \overline{r(t)} + n(t)$$

$$r(t) \sim N(\overline{r(t)}, n) \Leftrightarrow n(t) \sim N(0, n)$$

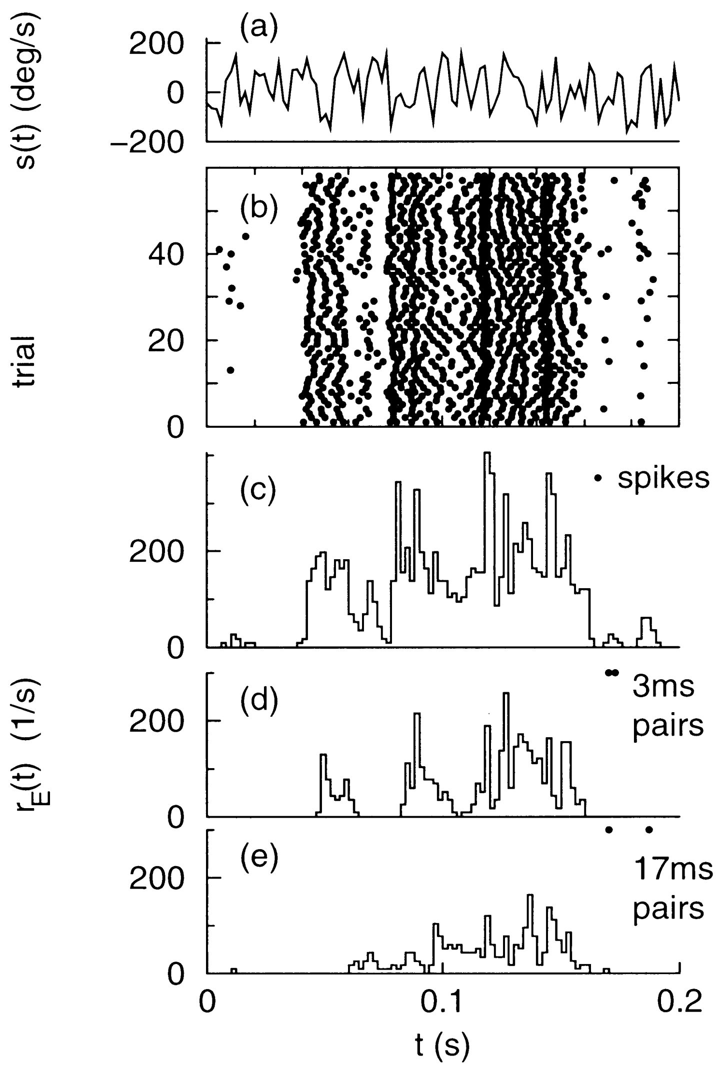
Information from SNR/Coherence

Poisson Noise

$$\vec{r}(t) \sim P(\overline{\vec{r}(t)})$$

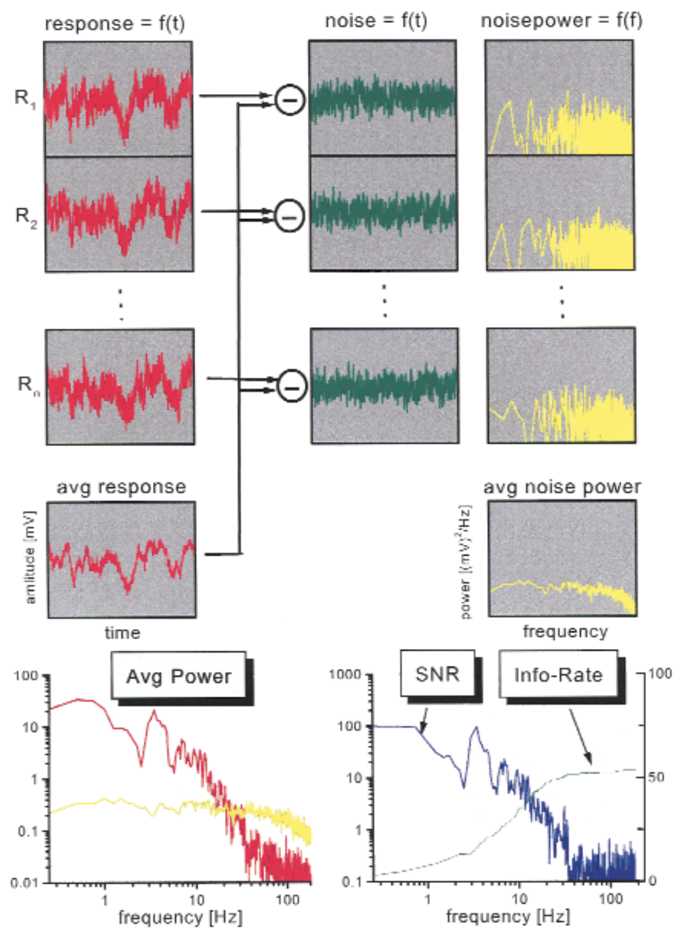
Single Spike Information

Single Event Information



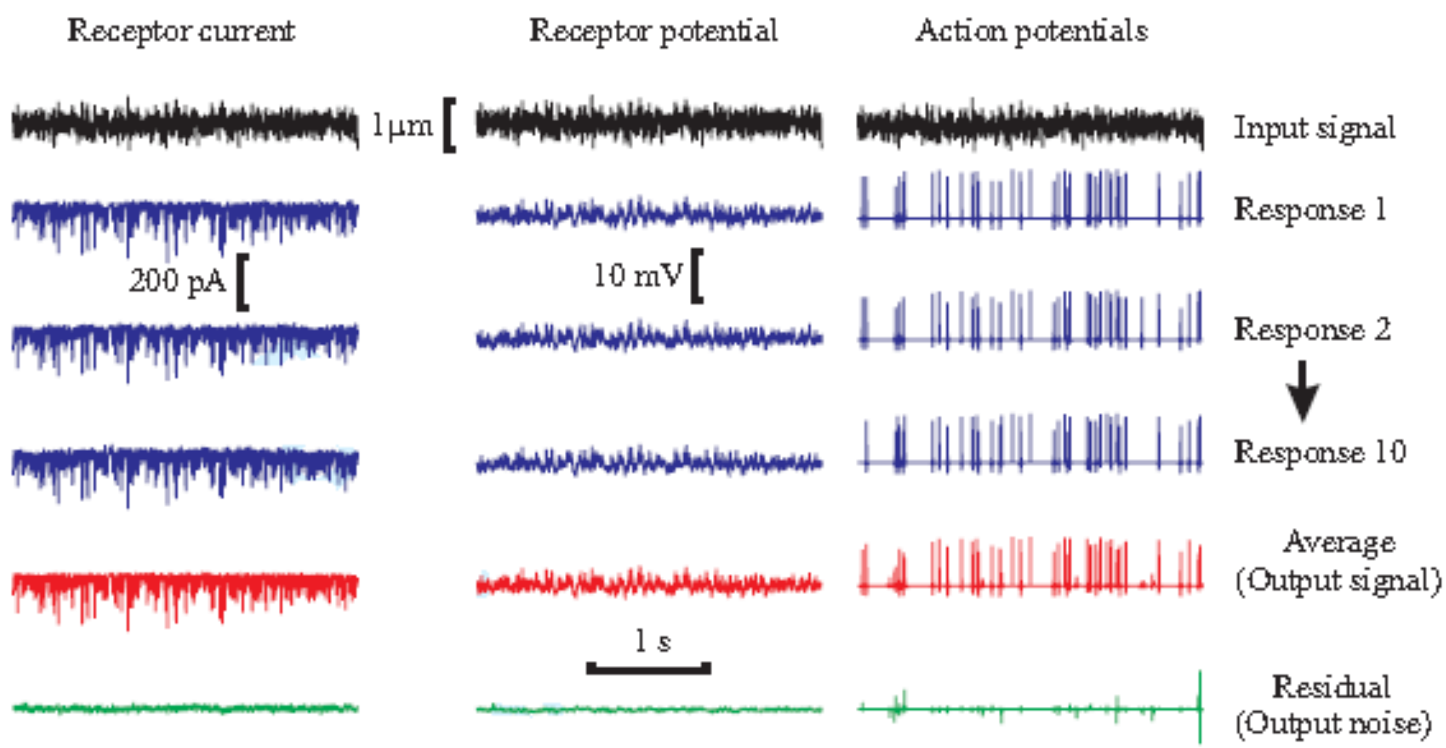
$$I(E,s) = \frac{1}{T}\int_0^T dt \left(\frac{r_E(t)}{\bar{r}_E}\right) \log_2\left(\frac{r_E(t)}{\bar{r}_E}\right)$$

Getting the "signal" by averaging…



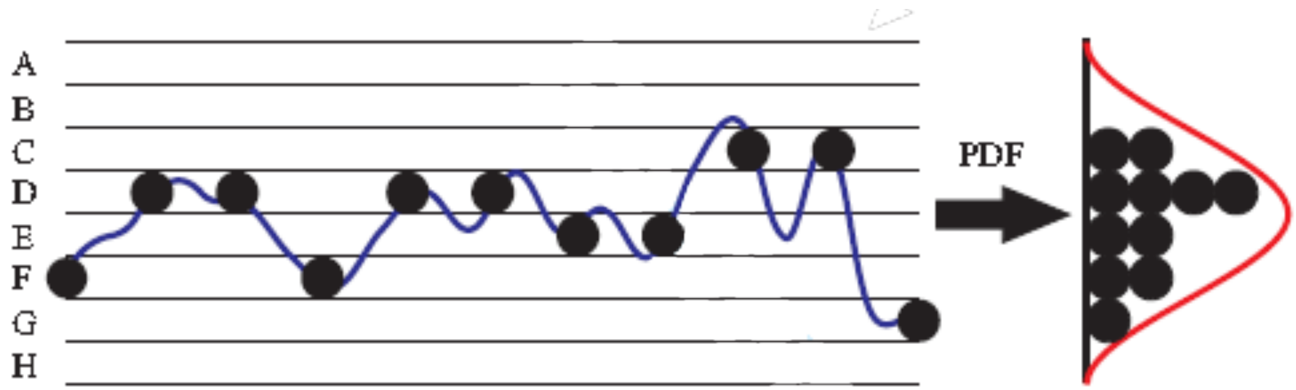
From Borst and Theunissen: Fly visual neurons

Getting the signal by “averaging”



From Andrew French: Spider Mechanoreceptors

Is the noise Gaussian?



From SNR to Info

$$I = \int \log_2 \left(1 + \frac{\left\langle Sig(f)^2 \right\rangle}{\left\langle N(f)^2 \right\rangle} \right) df$$

This is the Channel Capacity or Information Upper Bound.

It is an upper bound is Noise is Gaussian.

The coherency and the coherence

$$S_{xy}(f) = \langle \bar{x}(f) y(f) \rangle = FT\left\{ \frac{1}{T} \int_0^T \bar{x}(t) y(t+\tau) dt \right\}$$

$$\gamma(f) = \frac{S_{xy}(f)}{\left| S_{xx}(f) S_{yy}(f) \right|^{1/2}}$$

$$\gamma^2(f) = \frac{\left| S_{xy}(f) \right|^2}{\left| S_{xx}(f) S_{yy}(f) \right|}$$

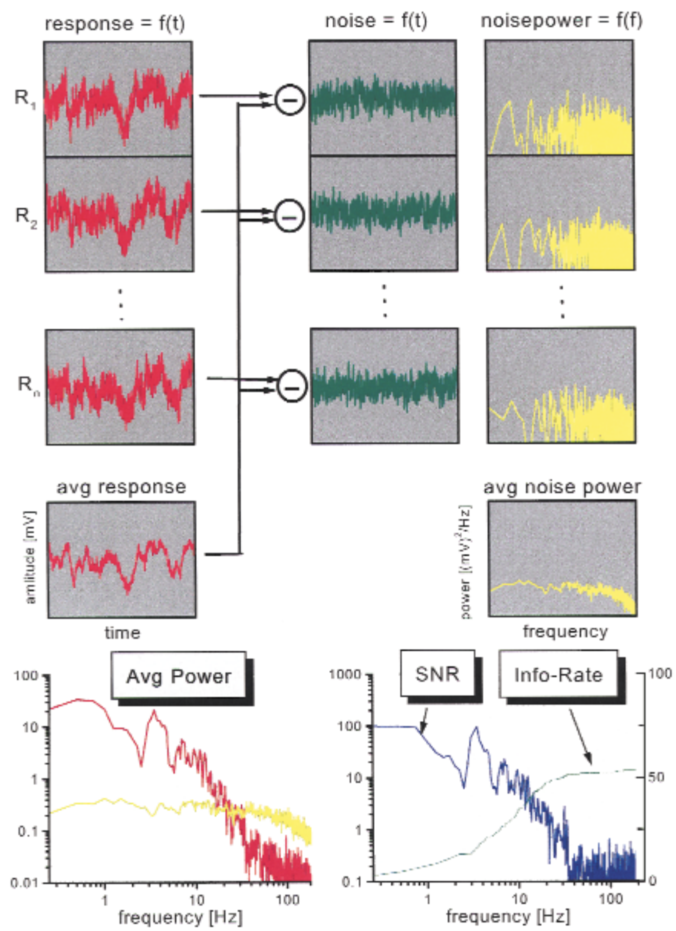
The coherence between a signal and signal+ noise

$$\gamma^2_{s,s+n} = \frac{\left\langle S^2(f) \right\rangle}{\left\langle S^2(f) \right\rangle + \left\langle N^2(f) \right\rangle}$$

$$\frac{S^2}{N^2} = \frac{\gamma^2}{1-\gamma^2}$$

$$I = \int \log_2 \left(1 + \frac{\left\langle S(f)^2 \right\rangle}{\left\langle N(f)^2 \right\rangle} \right) df = \int -\log_2 (1-\gamma^2(f)) df$$

From SNR to info capacity…



Signal Averaging

Neural Responses are noisy

$$R^2(\omega) = Sig^2(\omega) + N^2(\omega)$$

An average response obtained from M trials is still noisy

$$\overline{R}_M^2(\omega) = Sig^2(\omega) + N^2(\omega) / M$$

Actual signal (S) is unknown and we are
overestimating signal power…

Noise Estimation (note here Sig=S)

- If we estimate noise with $N_{est} = R - \overline{R}$

$$N_{est}^2 = \left|S - S\right|^2 + \left|N\left(1 - \frac{1}{M}\right)\right|^2 + \frac{M-1}{M^2}\left|N\right|^2$$

- We can recover the unbiased noise power by

$$N_{est}^2 = \frac{M-1}{M}\left|N\right|^2 \Rightarrow \left|N\right|^2 = \frac{M}{M-1}N_{est}$$

- We can recover the unbiased signal power by:

$$\left|S\right|^2 = \left|\overline{R}\right|^2 - \frac{\left|N\right|^2}{M} = \left|\overline{R}\right|^2 - \frac{N_{Est}}{M-1}$$

Validation

Solution: Compare the coherence function between a single spike train, R, and the actual rate, A, with the coherence function between a single spike train, R, and the predicted rate, B.

Coherence between R and A is estimated by resampling:

$$\frac{1}{\gamma_{AR}^2} - 1 = \frac{-M + M\sqrt{\left(\frac{1}{\gamma^2_{\overline{R}_{1,M/2}\overline{R}_{2,M/2}}}\right)}}{2}$$

Equation 8 in Hsu et al

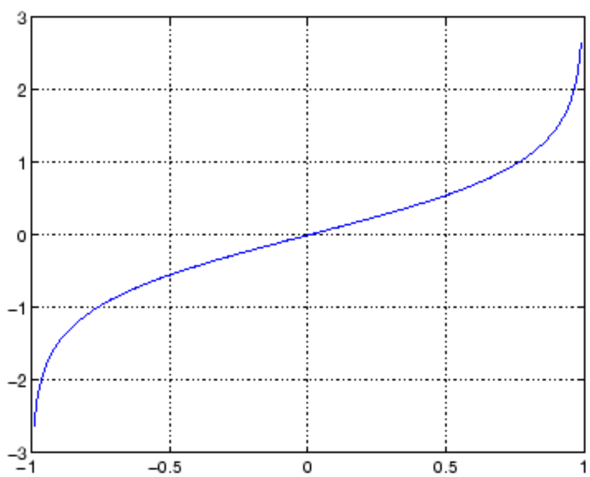
Coherence between R and B can be estimated from the coherence between the PSTH and B:

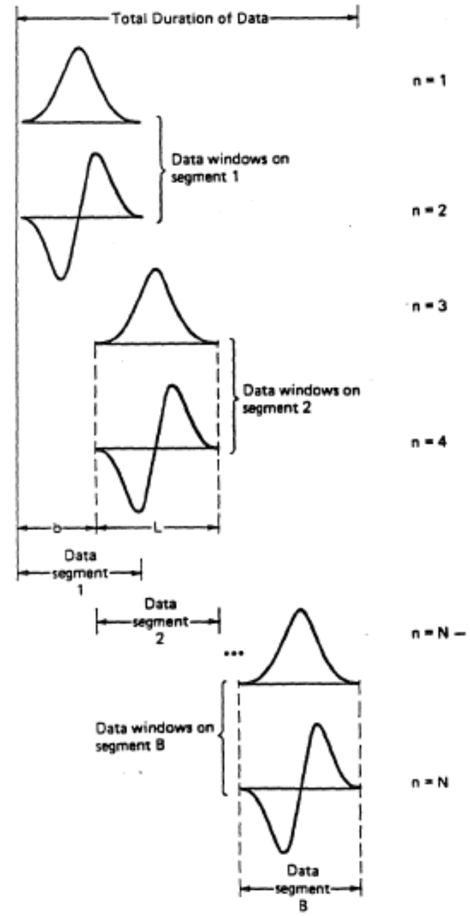
$$\frac{\gamma_{BR}^2}{\gamma_{B\overline{R}}^2} = \frac{\gamma_{AR}^2}{\gamma_{A\overline{R}}^2} = \frac{1 + \sqrt{\left(\frac{1}{\gamma^2_{\overline{R}_{1,M/2}\overline{R}_{2,M/2}}}\right)}}{-M + M\sqrt{\left(\frac{1}{\gamma^2_{\overline{R}_{1,M/2}\overline{R}_{2,M/2}}}\right)} + 2}$$

Equation 11 in Hsu et al

Jack-knifing, Multi-taper methods, Data Transformations

$$p_i = N\vartheta_{All} - (N-1)\vartheta_{\hat{i}}$$





How about methods for non-Gaussian noise?

- Spike trains can be modeled as non-homogeneous Poisson or Gamma Processes.
- Time-rescaling can be used to assess order of Gamma model.
- Information rates can be calculated numerically.
- Direct Information Theoretic Measures.
- Robert Kass' Lectures

Hsu et al, J. Neuroscience, 2004

Maximum Entropy Rate for Spikes

$$E \approx \bar{r} \log_2 \left(\frac{e}{\bar{r}\Delta t} \right)$$

$$\bar{r}\Delta t << 1$$

From Spikes (Rieke et al.) Ch 3.

Reducing the dimensions…

- Encoding Analysis

$$\overbrace{\overline{\vec{r}(t)}}^{Est} = f(\vec{s}(t)) \Rightarrow \overbrace{\overline{\vec{r}(t)}}^{Est} = f_{Static}(\vec{h} \cdot \vec{\phi}(\vec{s}(t))) \Rightarrow \overbrace{\overline{\vec{r}(t)}}^{Est} = \vec{h} \cdot \vec{s}(t)$$

- Mutual information

$$I(s,r) = H(r) - H(r \mid s)$$

Gaussian Channel

Single Spike

Extrapolation (Direct)

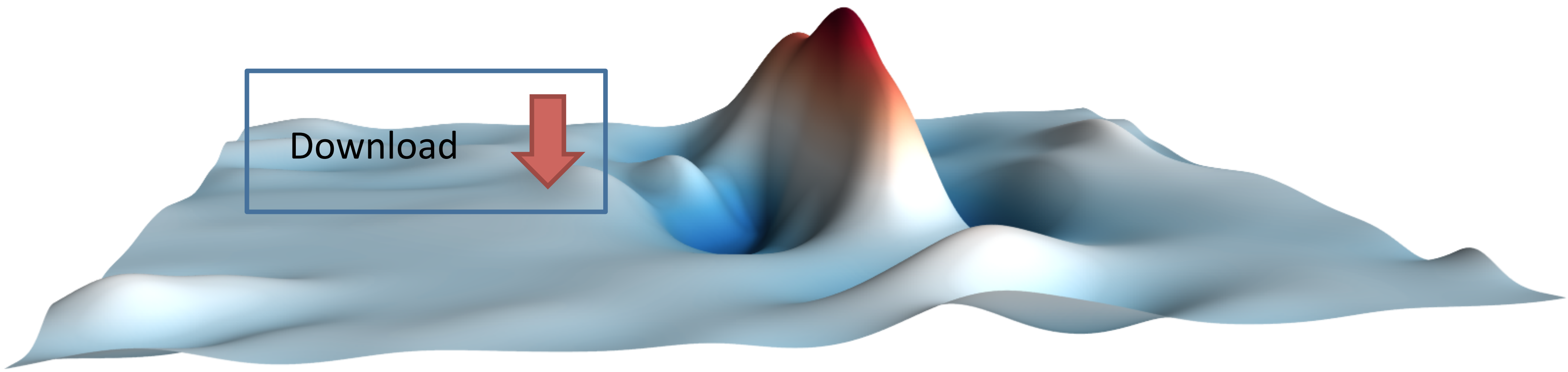
+ Bounds

+ Methods for estimating H

Tutorials inspired by STRFLab

STRFPAK 5.3

STRFLab 1.0



http://strflab.berkeley.edu

Key References

- ***Direct Estimation***: Strong, S. P., R. Koberle, R. de Ruyter van Steveninck and W. Bialek (1998). "Entropy and information in neural spike trains." Phys Rev Letters **80**(1): 197-200.
- ***Single Spike Information***: Brenner, N., S. P. Strong, R. Koberle, W. Bialek and R. R. de Ruyter van Steveninck (2000). "Synergy in a neural code." Neural Comput **12**(7): 1531-1552.
- ***Review+Gaussian Channel Information***: Borst, A. and F. E. Theunissen (1999). "Information theory and neural coding." Nat Neurosci **2**(11): 947-957.
- ***Gaussian Channel, Coherence and Model Validation***: Hsu, A., A. Borst and F. E. Theunissen (2004). "Quantifying variability in neural responses and its application for the validation of model predictions." Network **15**(2): 91-109.

Assignment

- coherence_tutorial.m