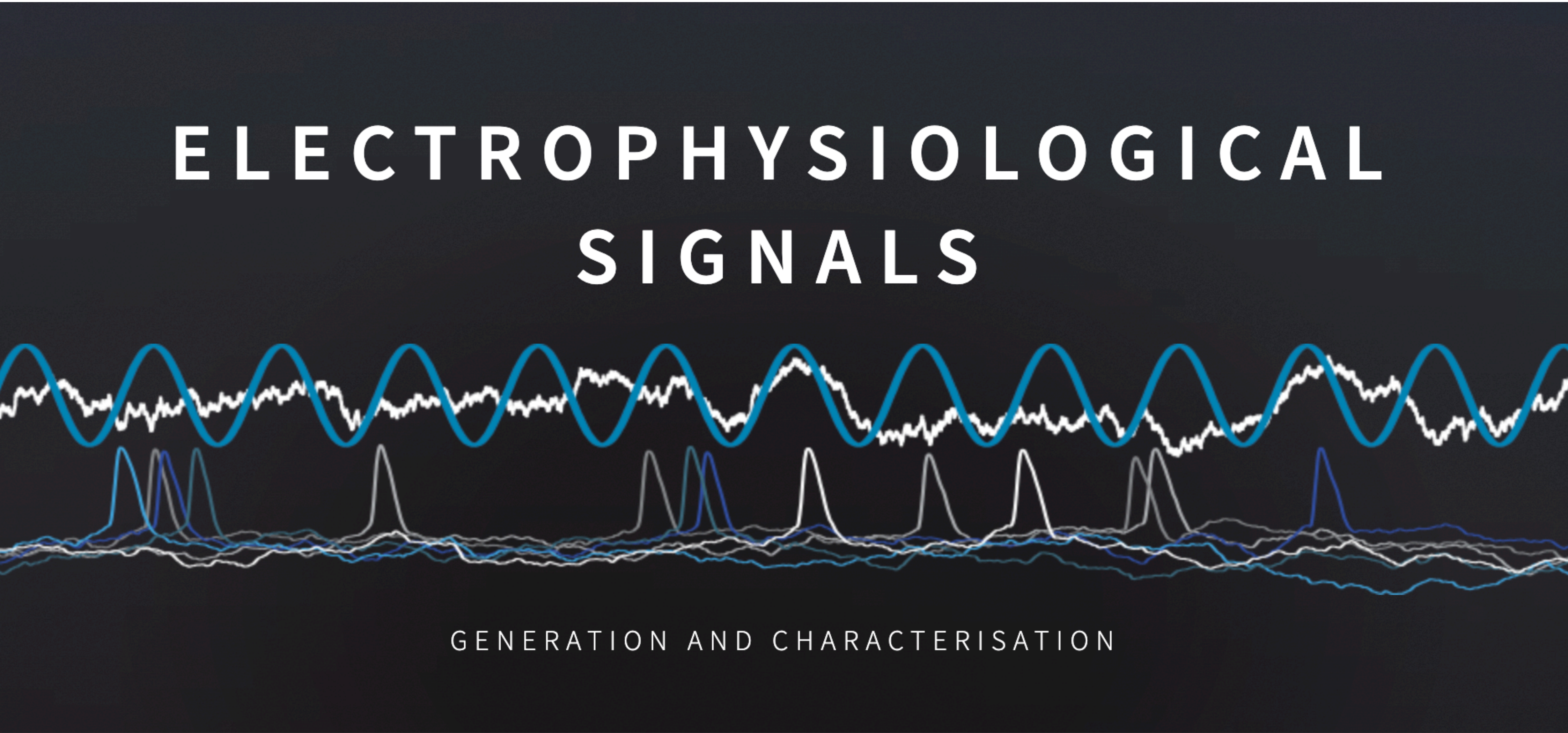
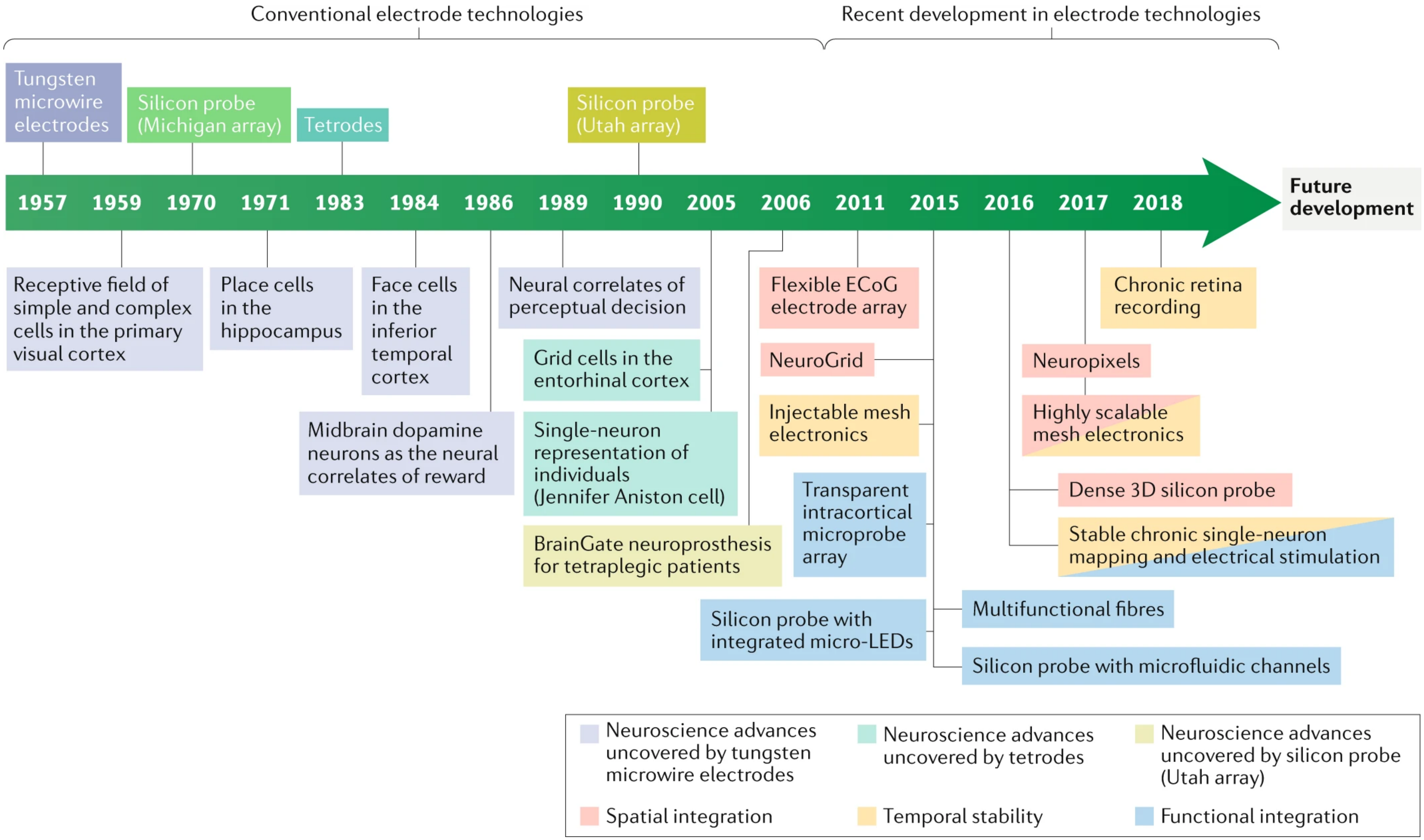
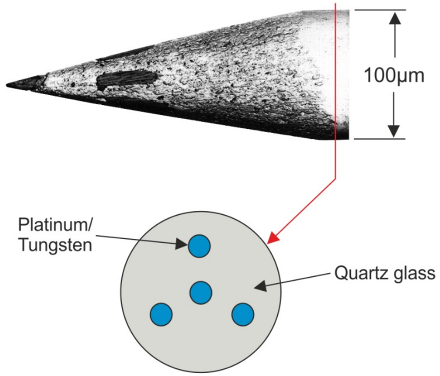


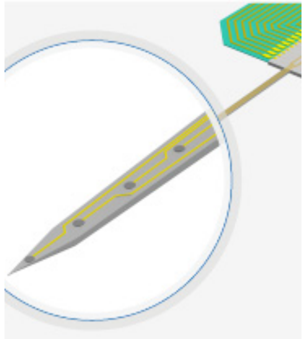
Michele GIUGLIANO

Analysis of Electrophys. Signals

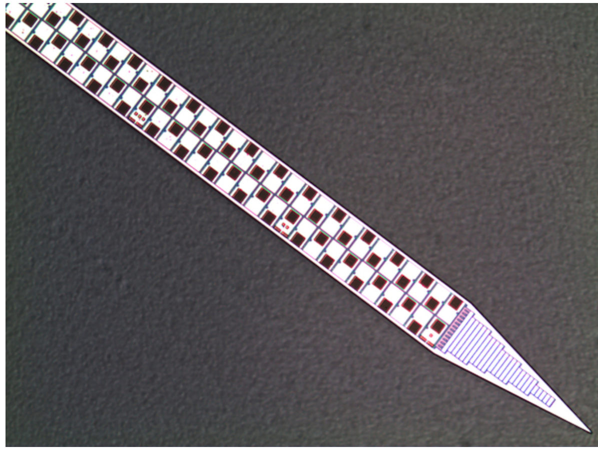


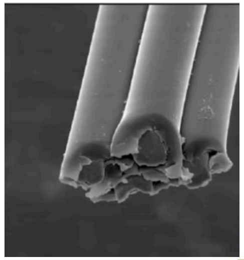
Hong & Lieber, *2019*

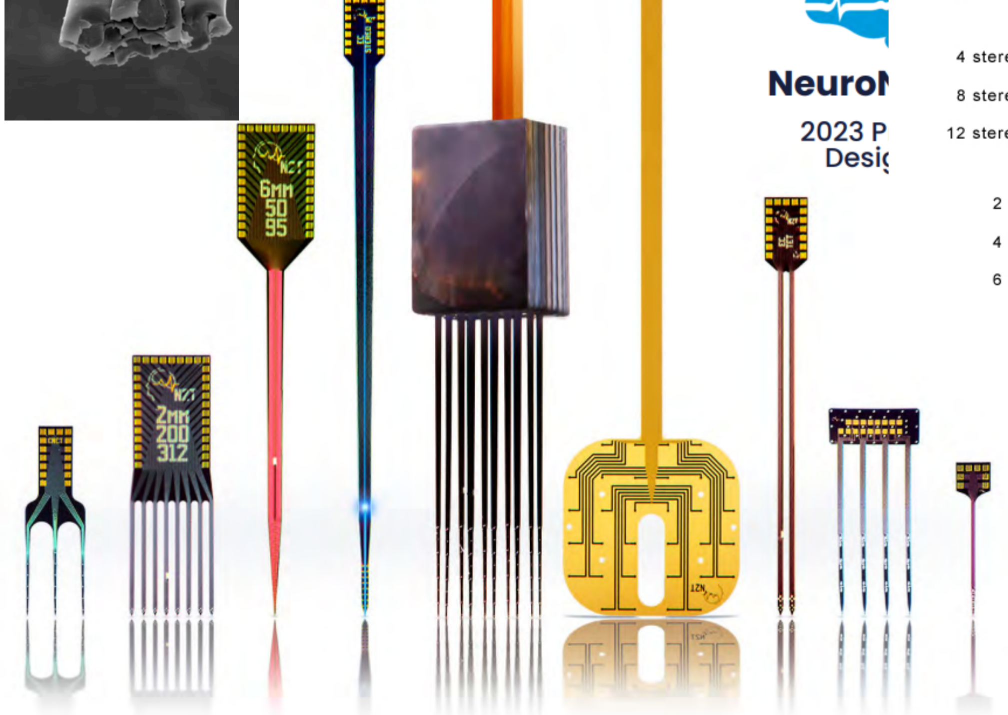


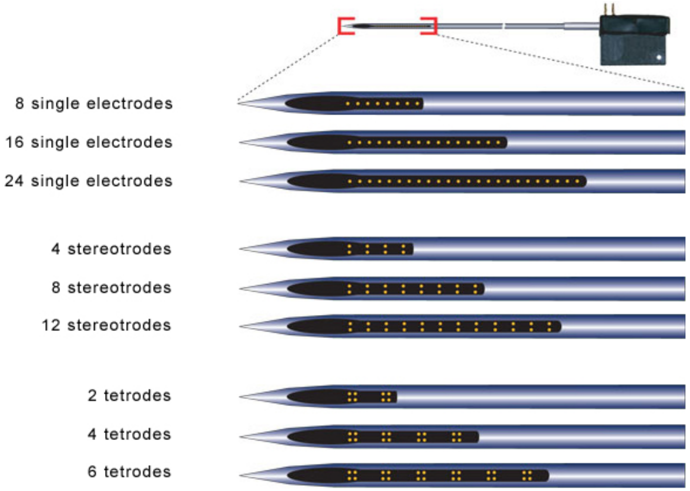


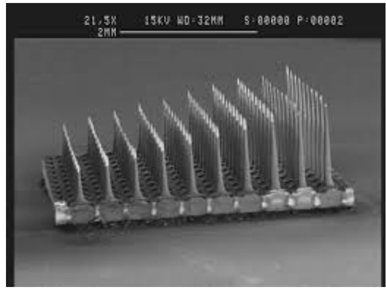
Tetrodes, Silicon Probes, Neuropixels, Utah Arrays

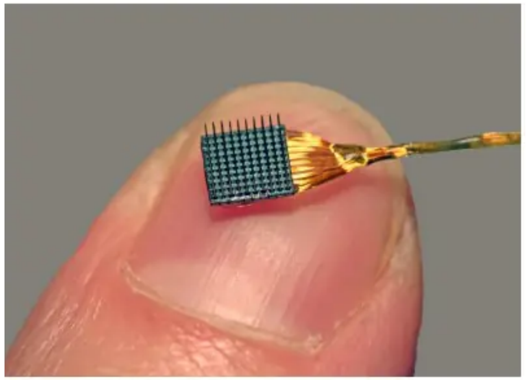


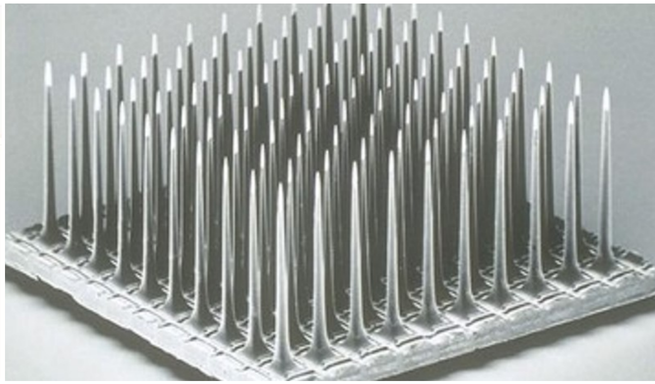




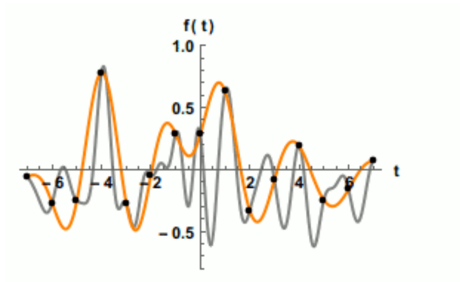


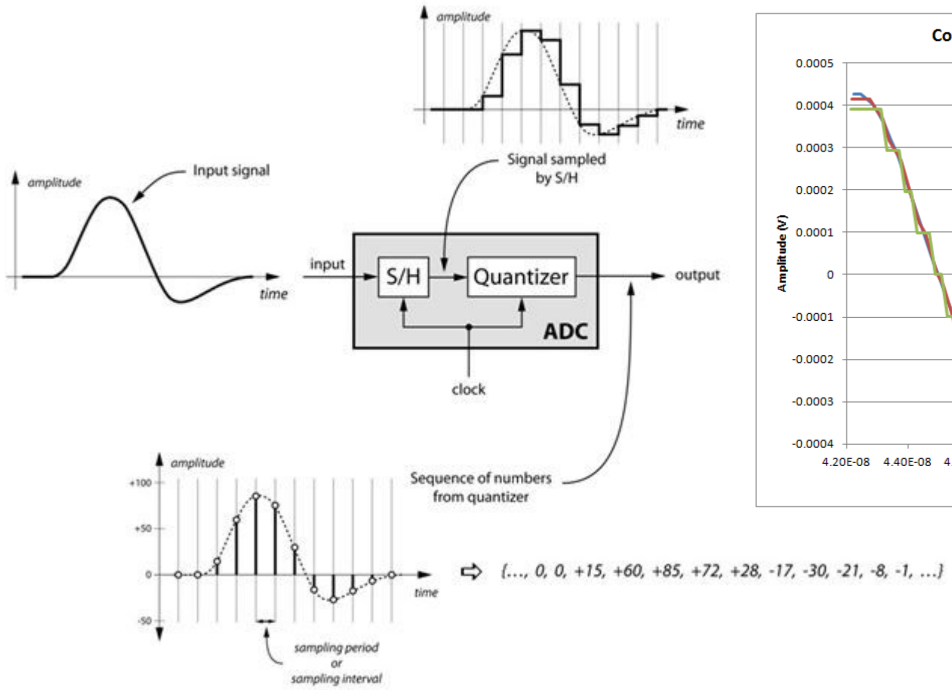


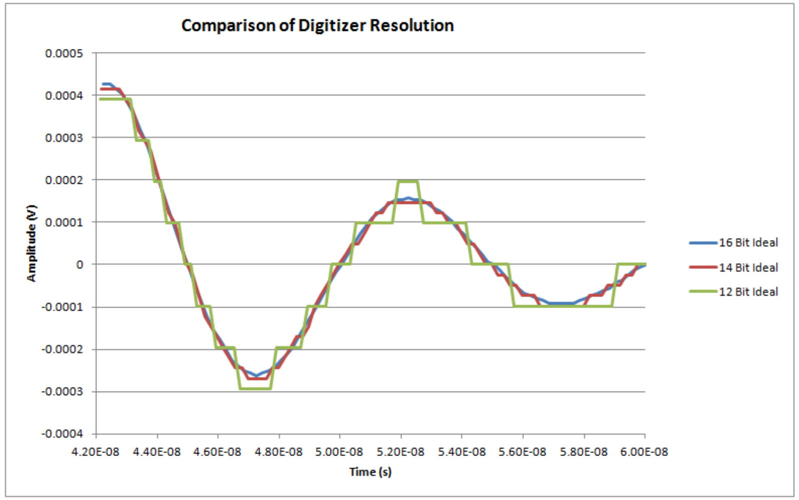




Sampling and Analog-to-Digital (A/D) Conversion (after amplification! e.g. x100)

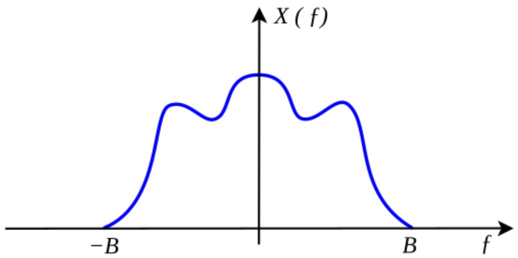






Nyquist-Shannon Theorem

$$f_s > 2\ B$$



$$100\mu V \times 100 = 10mV$$

Range of A/D and Resolution of A/D

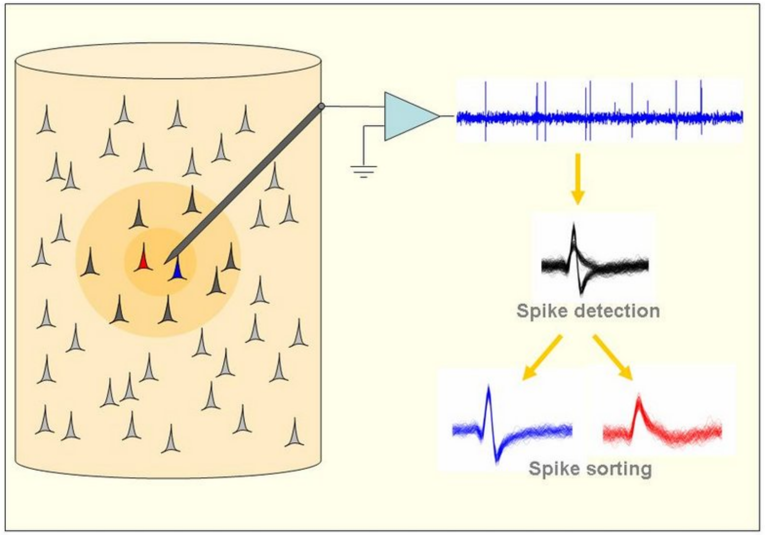
$\pm 5V$ $\pm 10V$

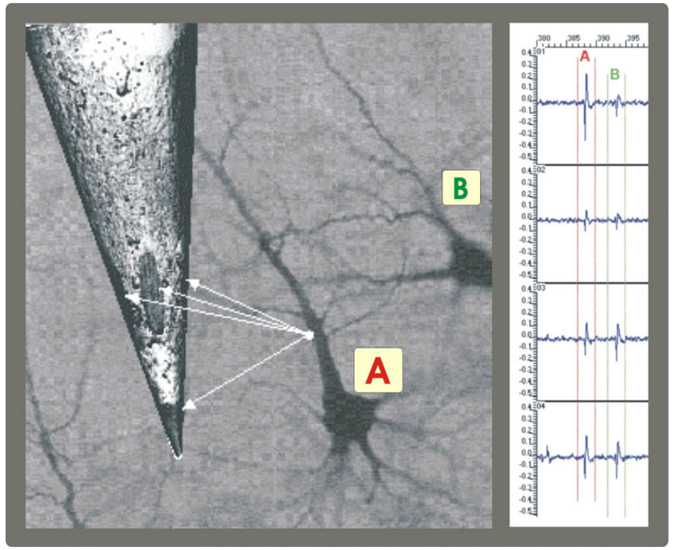
2^N

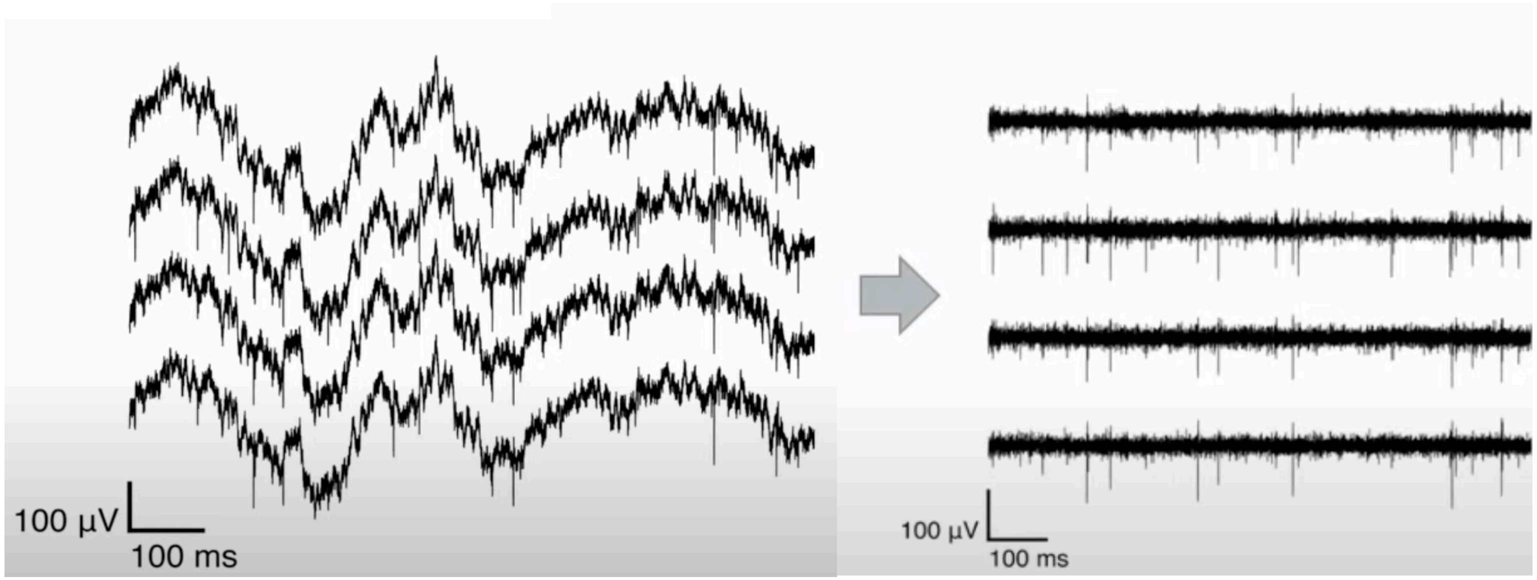
$2^{12} = 4096$ $2^{14} = 16384$ $2^{16} = 65536$

$2.4mV$ $0.6mV$ $0.15mV$

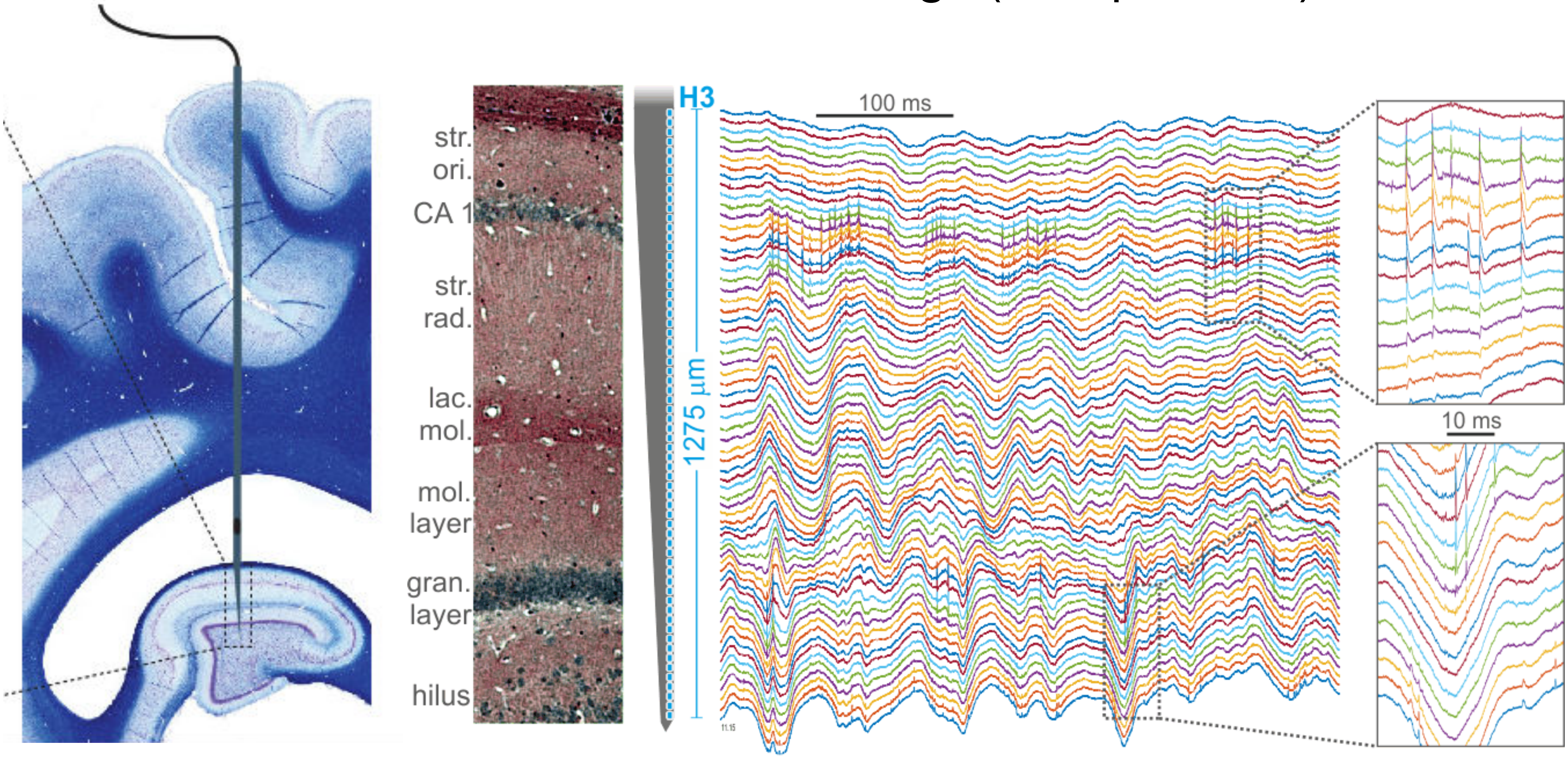
Spike detection and spike sorting





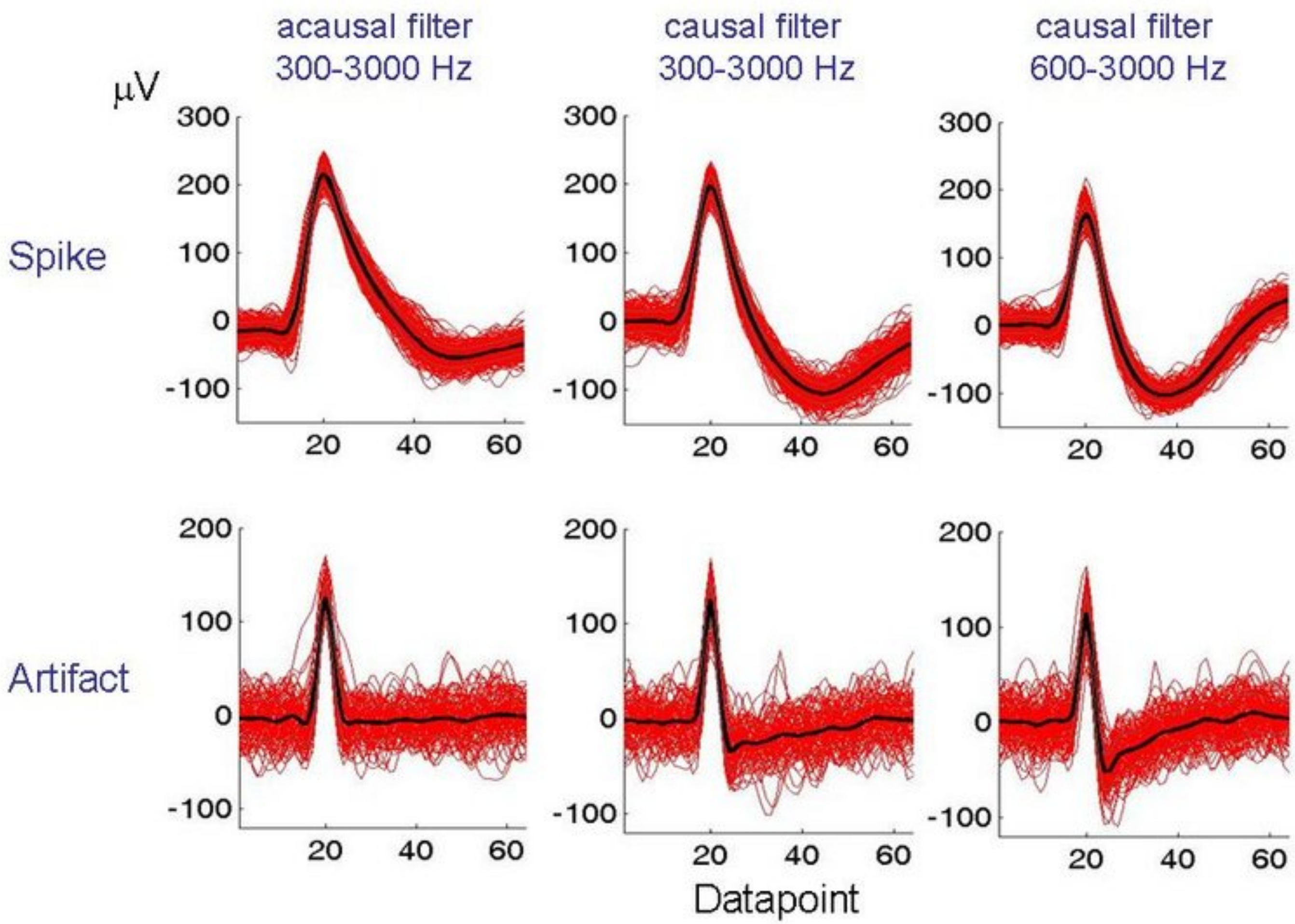


Raw extracellular recordings (multiple sites)



Wolf's lab, UPenn

Beware of filtering: a-causal is preferred



Peak-detection / Threshold crossings: estimates of the “baseline” noise level (from data itself!)

- **median:** the value separating the higher half from the lower half of a data sample… $\tilde{X}$

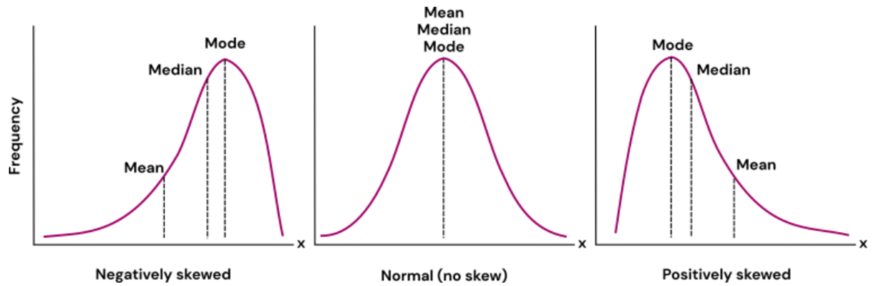
1, 3, 3, **6**, 7, 8, 9

Median = <u>6</u>

1, 2, 3, **4**, **5**, 6, 8, 9

Median = (4 + 5) ÷ 2

= <u>4.5</u>



- **MAD (median absolute deviation):** the *median* of the absolute deviations from the data's median…

$$MAD \ = \ median(|X_i - \tilde{X}|)$$

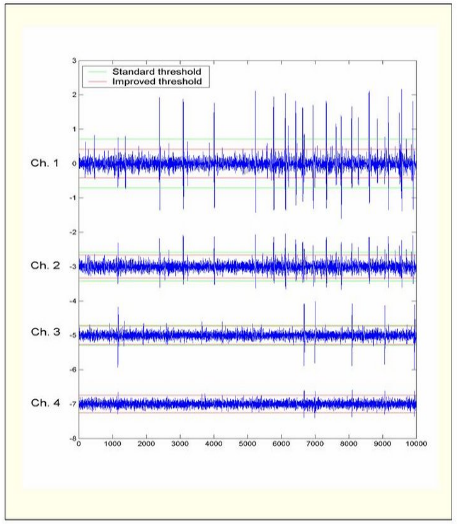
$$\hat{\sigma} = k \ MAD$$

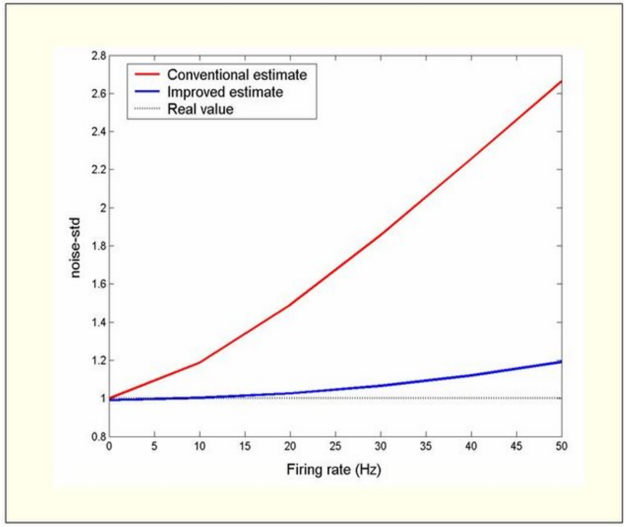
It is related to an *unbiased estimator* of the standard deviation (i.e. converge in probability to the true value)

For normal distribution, the scaling factor is **1/0.6645**.

$$\hat{\sigma} = \frac{MAD}{0.66449}$$

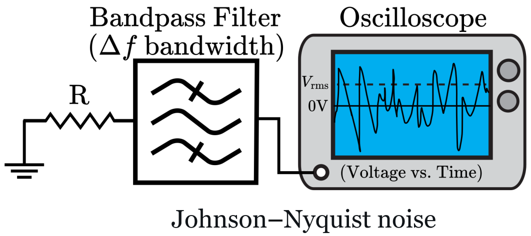
Peak-detection / Threshold crossings: estimates of the “baseline” noise level (from data itself!)





$$Thr = 5\,\sigma_n$$

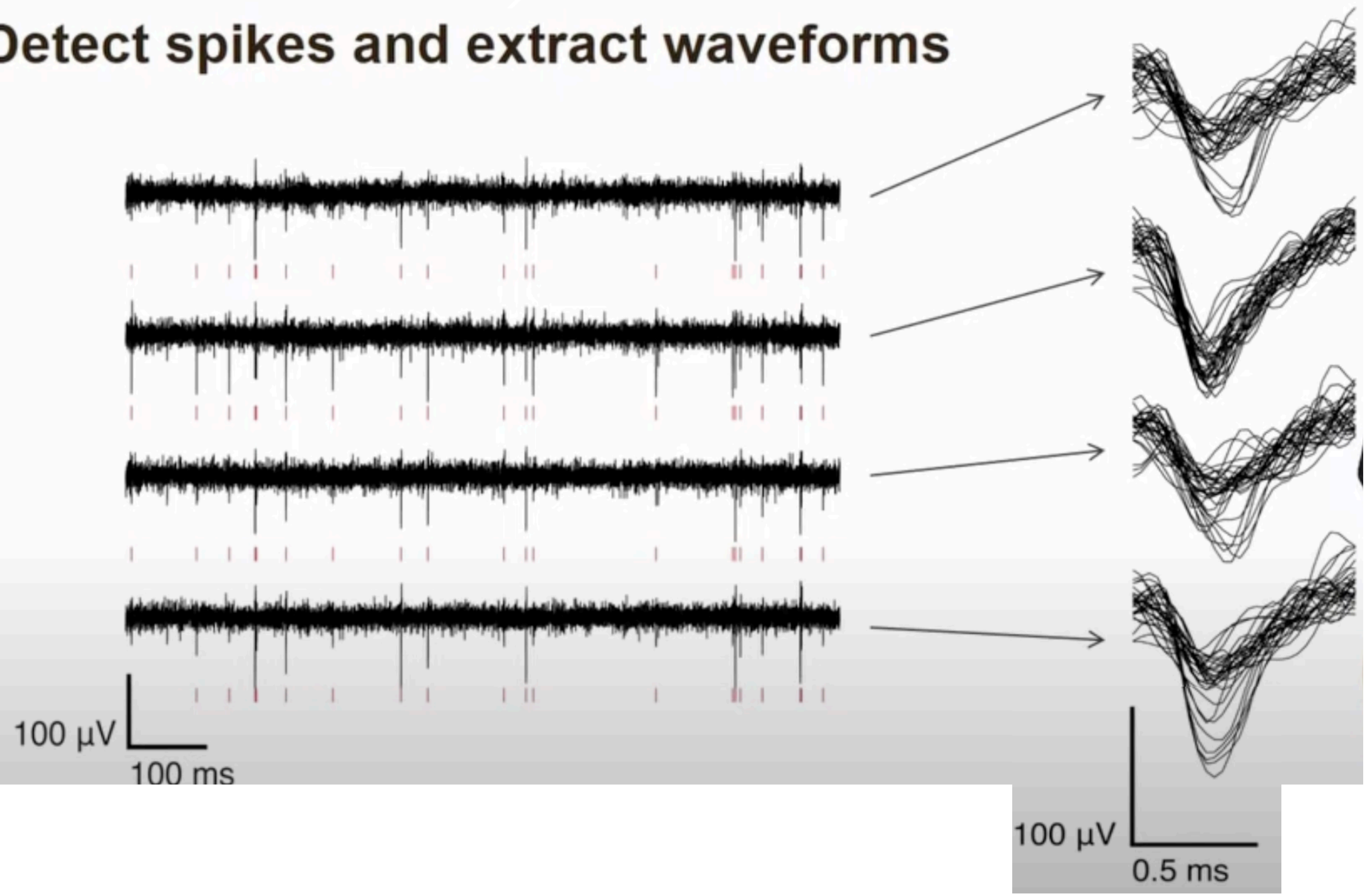
$$\sigma_n = \frac{median\left(|x|\right)}{0.6745}$$



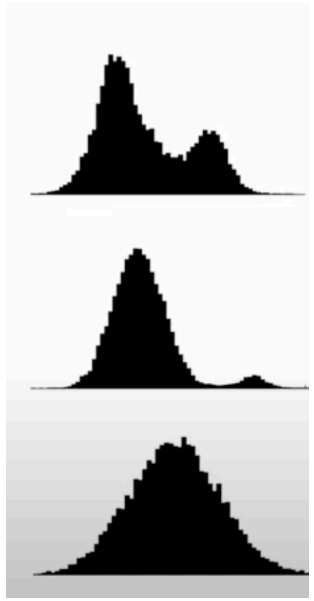
$$\sigma = \sqrt{4 k_B \mathrm{T}\,\mathrm{R}\,\Delta f}$$

- <u>adaptive</u> *versus* fixed threshold for the detection
- based on the hypothesis that *data = signal + noise* (noise has <u>unknown</u> statistical properties!)
- amount of noise <u>estimated from the data itself</u>, hopefully not altered by the signal itself (e.g. lot of spikes)
- <u>robust estimators</u> of the standard deviation: e.g. based on median of the band-pass filtered data.

Detect spikes and extract waveforms

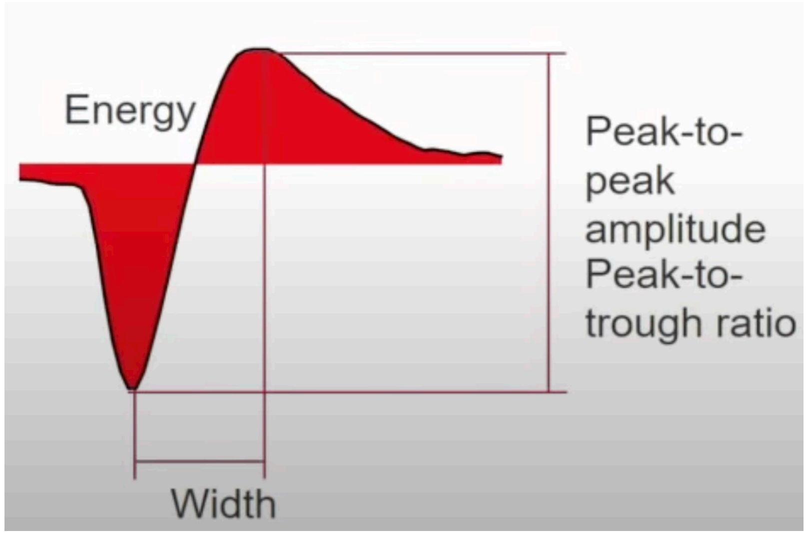


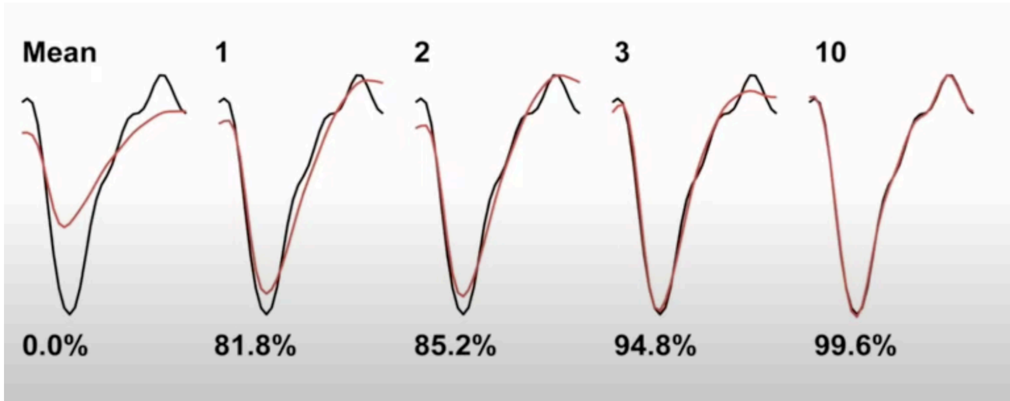
Feature extraction



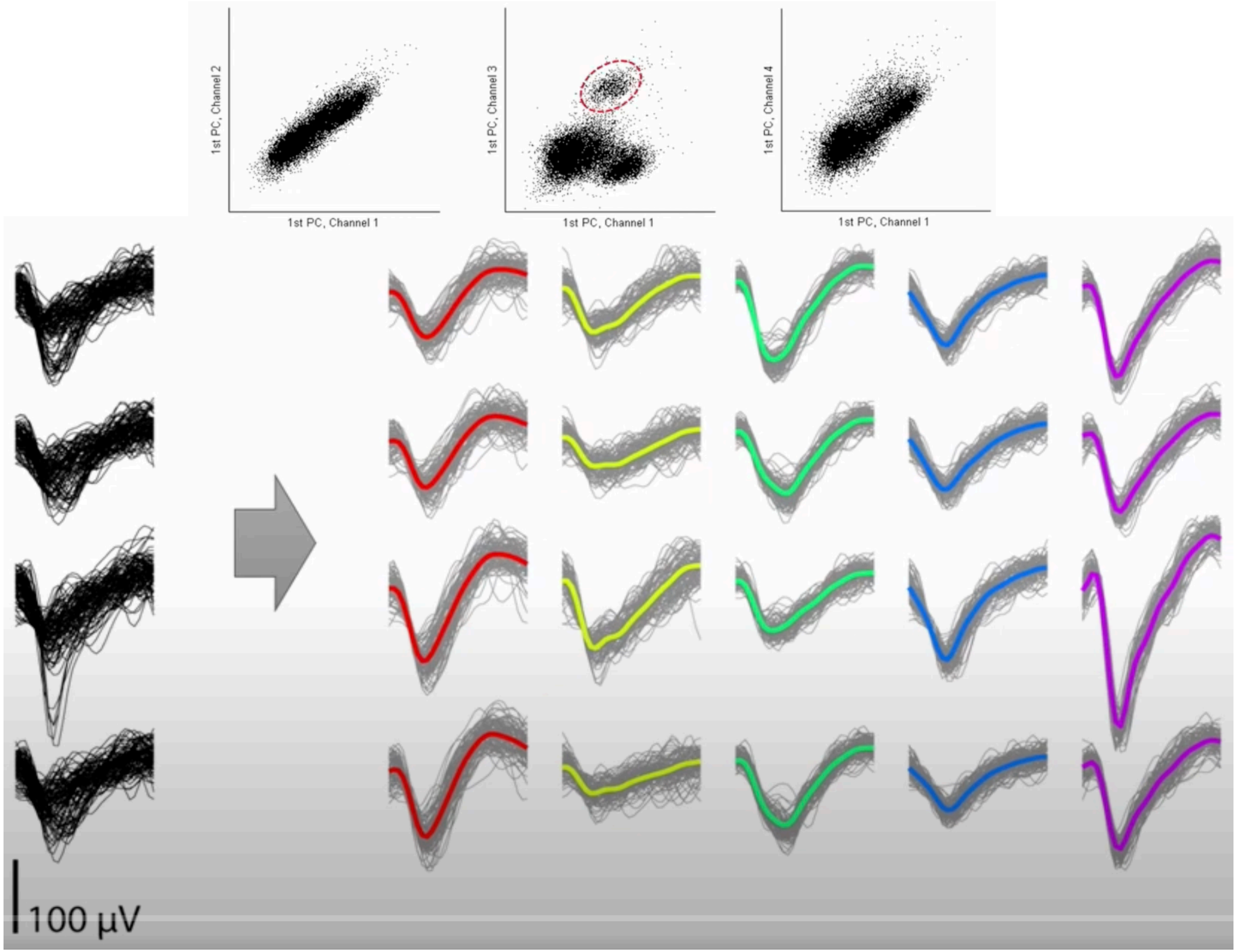
- smallest set of features, as possible
- best set of features to discriminate spike waveforms, robust against noise
- might depend on the choice of the classification algorithm

PCA - Principal Component Analysis

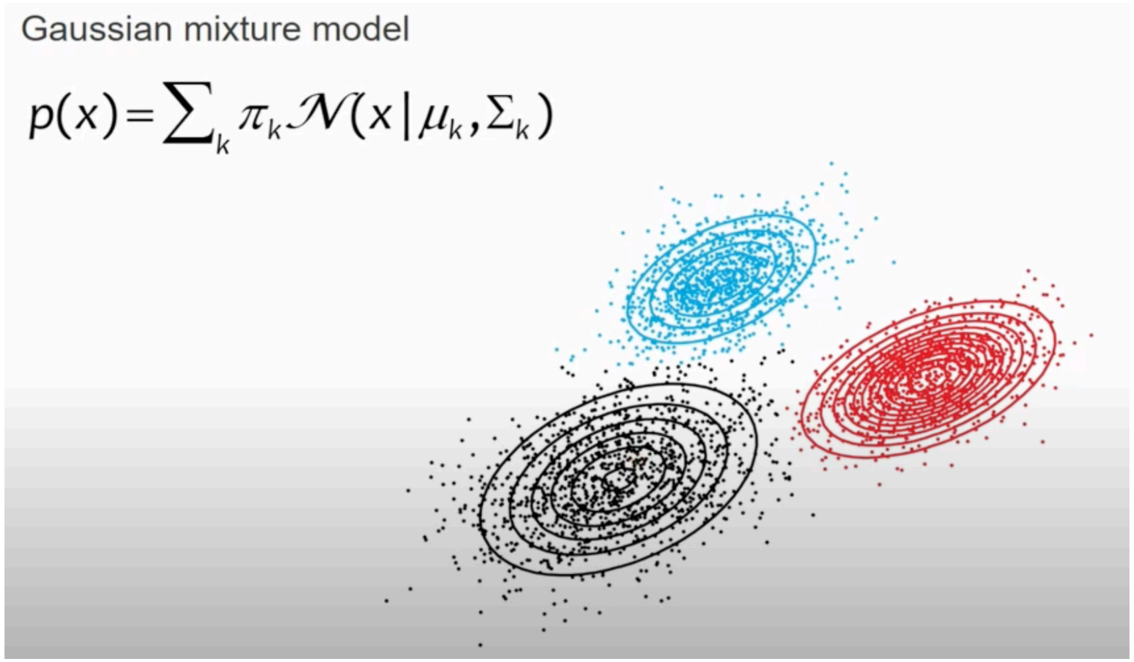


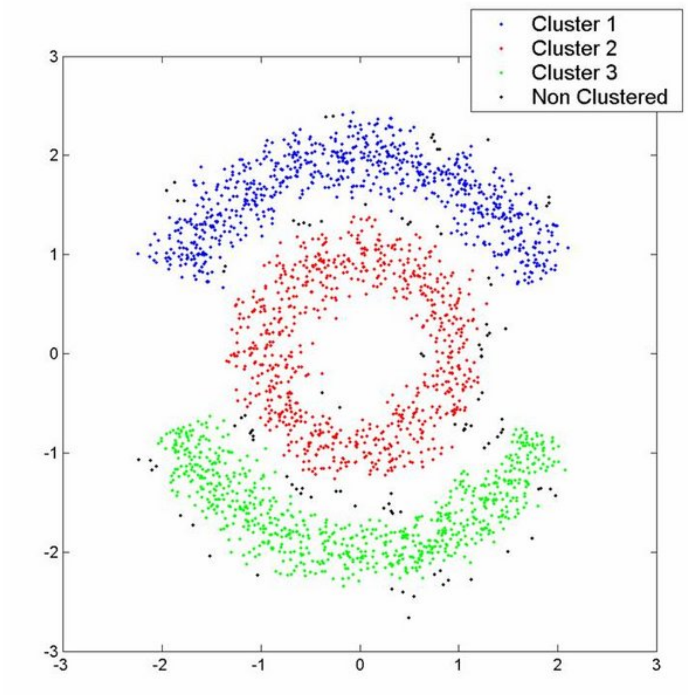


- Finds an orthonormal basis for the data
- First PC = direction of the largest variance
- very few components describe well the data
- (alternatives: non-linear PCA, autoencoder networks,


1st PC, Channel 2
1st PC, Channel 1
1st PC, Channel 3
1st PC, Channel 1
1st PC, Channel 4
1st PC, Channel 1
100 µV

Clustering: manual *vs* unsupervised

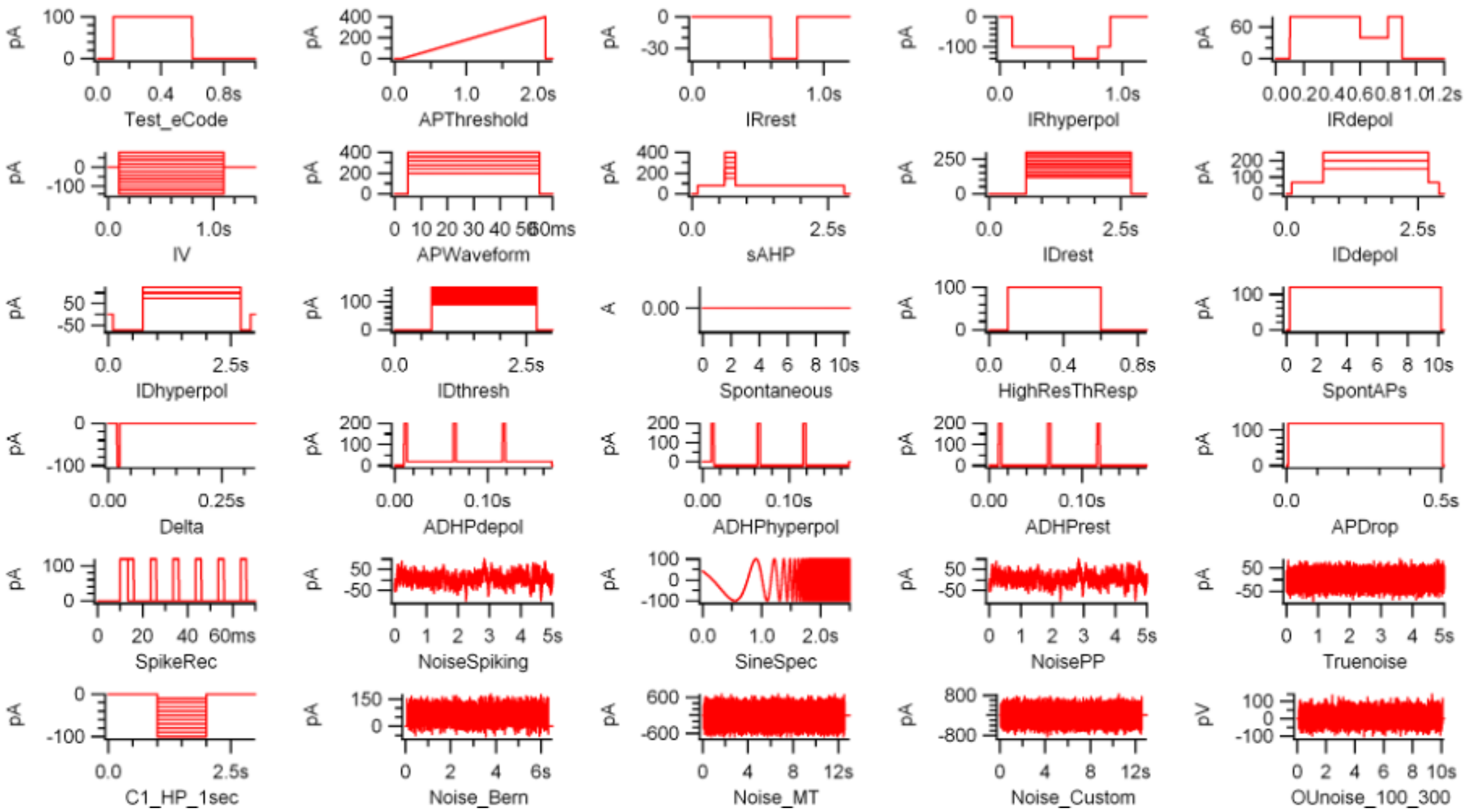




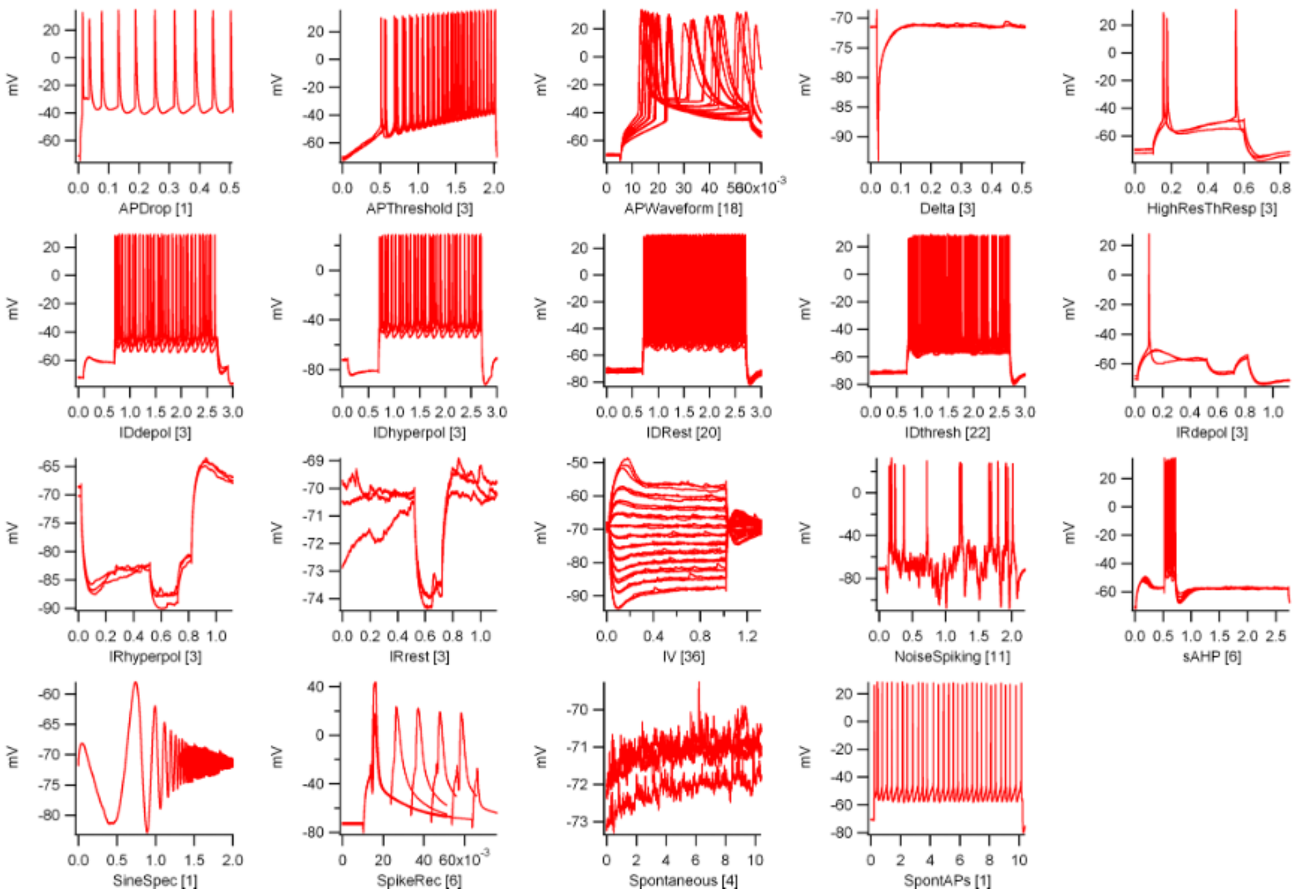
- manual (subjective, high error rate, not reproducible, time consuming)
- unsupervised/automatic clustering: k-Means, Mixture of Gaussian, super paramagnetic clustering, etc.

- tetrodes? bursting neurons? overlapping spikes?

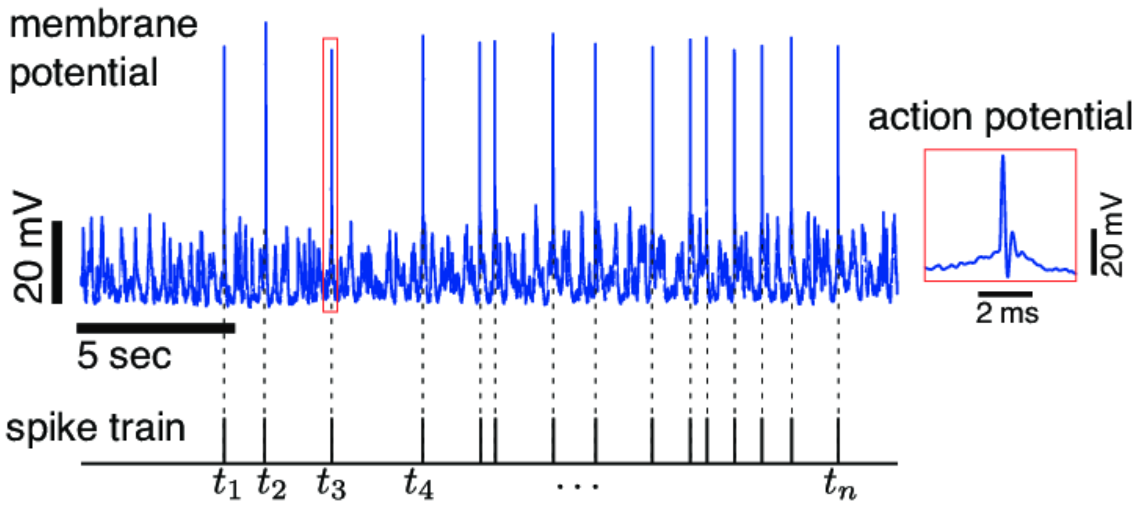
Intracellular recordings - stimulus/response for extracting the *e-Code* of a cell



Intracellular recordings - stimulus/response for extracting the *e-Code* of a cell



Analysis of *spike trains*



- Neurons fire irregularly
- Upon (sensory/electrical) stimulation, large trial-to-trial variability (both for timing and number of events)
- Analyzing *spike trains* benefits from “idealized”/“reference” situations, known as ***spike train models***.
- They are (probabilistic) models - without biophysical realism! - hopefully capturing some feature or statistics of real data, effects of *past history dependence*, variability, regularity, etc.
- Stochastic models of spike trains are the most widely used.

Demo

(Handling) Spike Trains

$$s(t) \;=\; \sum_{i=0}^{N} \delta(t - t_i)$$

$$f_t(t_1, t_2, t_3, ..., t_N)$$

- **joint PDF (probability density function)**

$t_1, t_2, t_3, ...$

$$x \;\; R.V.$$

$$F_x(X) \;=\; Prob\{x \leq X\}$$

- **Cumulative probability distribution**

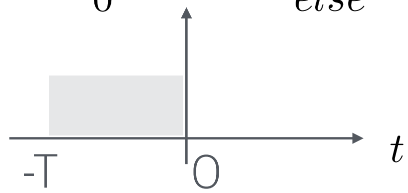
$$f_x(X) \;=\; \frac{d}{dX} F_x(X)$$

- **Probability density function**

$$Prob\{X < x \leq X + \Delta X\} \;=\; Prob\{X \leq X + \Delta X\} - Prob\{X \leq X\} \;=\; F_x(X + \Delta X) - F_x(X) \;\rightarrow f_x(X)\, \Delta X$$

(Handling) Spike Counts

$$w_T(t) = \begin{cases} 1 & -T < t < 0 \\ 0 & else \end{cases}$$



$$n(t) \;=\; \int_t^{t+T} s(x)\,dx \;=\; s(t) \;*\; w_T(t)$$

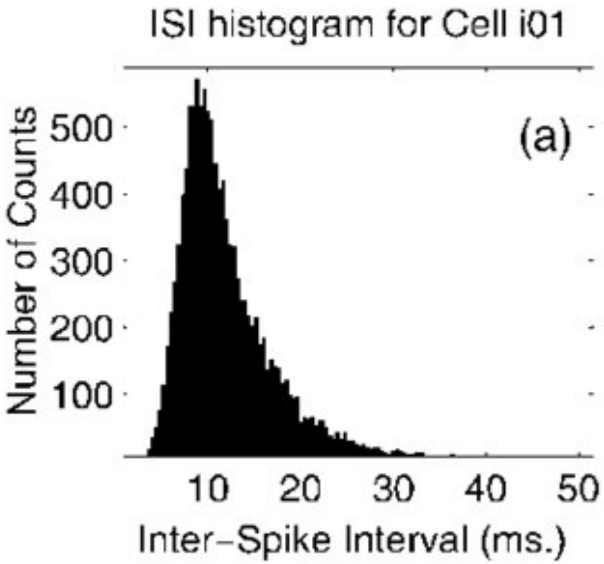
$$\int_{-\infty}^{+\infty} s(t-x)\,w_T(x)\,dx$$

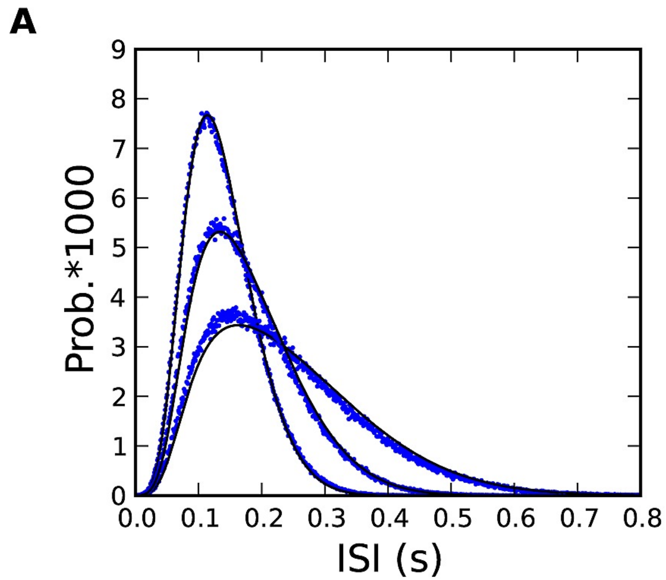
The inter-spike intervals distribution

$\{t_k\}$ $\quad t_1, t_2, t_3, ...$ $\quad \Delta_1, \Delta_2, \Delta_3, ...$

$$\Delta_k = t_k - t_{k-1}$$

- **Notation:** spike times, inter-spike intervals.





Poisson point process (simplest *renewal* process)

$$f_t(t)\ \delta t \qquad\qquad Prob\{spike\ in\ [t\ ;\ t+\delta t)\}\ =\ R\ \delta t$$

- **R** is called **firing rate** (for simplicity chosen here to be constant in time, as in *homogeneous processes*)
- The probability of spike occurrence does **not** depend on the past history (no refractory period captured!)

$\{t_k\}$ $\qquad t_1, t_2, t_3, \ldots$ $\qquad \Delta_1, \Delta_2, \Delta_3, \ldots$

$$\Delta_k\ =\ t_k\ -\ t_{k-1}$$

- **Notation:** spike times, inter-spike intervals.
- Renewal property = *the inter-event intervals are independent and drawn from the same probability density.*

Survival Probability for a Poisson point process

$$S(T) \; = \; Prob\{next \; spike \; in \; (t_0 \; + T \; ; \; +\infty)/spike \; at \; t_0\}$$

- **Survival Probability (or survivor function) *S(T)*:** prob of occurrence **after** *t + T* of the next spike, if previous at *t.*

$$S(0) \; = 1 \qquad\qquad S(T + \delta t) \; = S(T) \; (1 - R\delta t)$$

- The *Survival probability* can be made explicit, by solving …the usual boring ordinary differential equation!

$$\frac{dS(T)}{dT} \; = \; -R \; S(T) \qquad\qquad S(T) \; = \; e^{-R \; T}$$

- The *Survival probability* is the complementary of the cumulative distribution of ISI. Then, we can calculate the probability density of ISI…

$$S(T) \; = \; Prob\{ISI > T\} \qquad Prob\{ISI \leq T\} \; = \; 1 - S(T) \qquad f_{ISI}(T) \; = \; \frac{d}{dT}(1 - S(T))$$

- ISI are drawn from an *exponential* distribution $f_{ISI}(T) \; = \; R \; e^{-RT}$ **(renewal)**

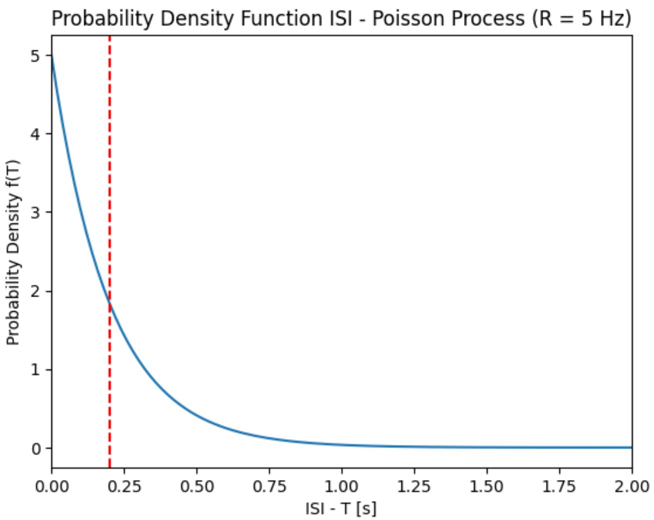
Poisson point process (simplest *renewal* process)

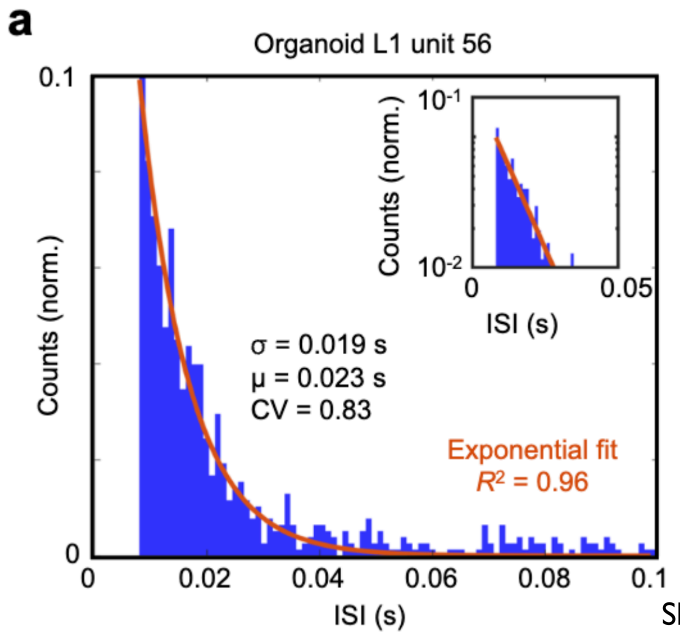
$$\int_0^{+\infty} f_{ISI}(T)\, dT \;=\; 1 \qquad\qquad \bar{T} \;=\; E\{T\} \;=\; \int_0^{+\infty} T\, f_{ISI}(T)\, dT \;=\; \frac{1}{R}$$

$$Var\{T\} \;=\; E\{T^2\} \;-(E\{T\})^2 \;=\; \int_0^{+\infty} T^2\, f_{ISI}(T)\, dT \;-\; \frac{1}{R^2} \;=\; \frac{1}{R^2}$$

- **Mean and Standard Deviation are the same**
- The Coefficient of Variation is therefore **1**

$$CV_{ISI} \;=\; \frac{stdDev\{T\}}{\bar{T}} \;=\; 1$$





Sharf et al., 2022

Integration by part (I never remember it)

$$(a(x)\ b(x))' \ = \ a(x)'\ b(x) \ + \ a(x)\ b(x)'$$

$$\int (a(x)\ b(x))' \ = \ \int a(x)'\ b(x) \ + \ \int a(x)\ b(x)'$$

$$a(x)\ b(x) \ = \ \int a(x)'\ b(x) \ + \ \int a(x)\ b(x)'$$

$$\int a(x)'\ b(x) \ = a(x)\ b(x) \ - \ \int a(x)\ b(x)'$$

$$\int_0^{+\infty} T\ R\ e^{-RT}\ dT \ = \ -T\ e^{-RT}|_0^{+\infty} \ - \ \int_0^{+\infty} -e^{-RT}\ dT \ = \ 0 \ + \ \frac{1}{R}$$

More general renewal processes

$$f_t(t)\ \delta t$$

$$Prob\{spike\ in\ [t\ ;\ t+\delta t)/last\ spike\ at\ t_0\}\ =\ H(t-t_0)\delta t$$

- *H* is called **stochastic *intensity*** (or ***hazard rate***)

$$S(T\ +\ \delta t)\ =\ S(T)\ [1\ -\ H(T)\delta t]$$

$$S(0)\ = 1$$

$$\frac{dS(T)}{dT}\ = -H(T)S(T)$$

$$f_{ISI}(\Delta)\ = \frac{d}{d\Delta}[1-S(\Delta)]\ =\ -\frac{d}{d\Delta}S(\Delta)\ =\ H(\Delta)S(\Delta)$$

$$\int_0^\Delta \frac{dS(T)}{S(T)}\ = \int_0^\Delta -H(T)\ dT \qquad ln(S(\Delta))\ -\ 0\ = \int_0^\Delta -H(T)\ dT \qquad S(\Delta)\ =\ e^{\int_0^\Delta -H(T)\ dT}$$

$$f_{ISI}(\Delta)\ = H(\Delta)e^{\int_0^\Delta -H(T)\ dT}$$

More general renewal processes

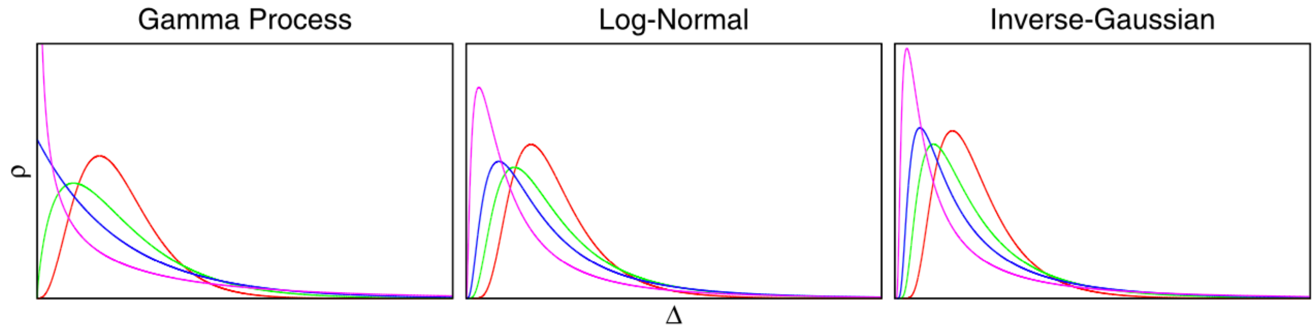
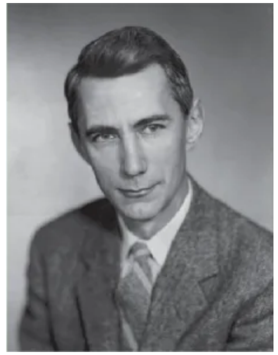


Fig. 1.1 ISI distributions for Gamma, Log-Normal, and inverse Gaussian processes. For each class, the distribution is shown for $C_V = 0.5$ (*red*), $C_V = 0.75$ (*green*), $C_V = 1$ (*blue*), and $C_V = 1.5$ (*purple*)

Information Theory (for spike-train analysis)



Claude Shannon

(1916 - 2001)

Used to understand how much information is contained in a spike-train, about (e.g.) a stimulus.
(e.g. neuroprosthetics inferring desired motor commands, from a tetraplegic patient)

Source/Sender (stimulus)

Communication channel (noisy)

Observer/Receiver (observed spike train)

X - symbol (discrete)

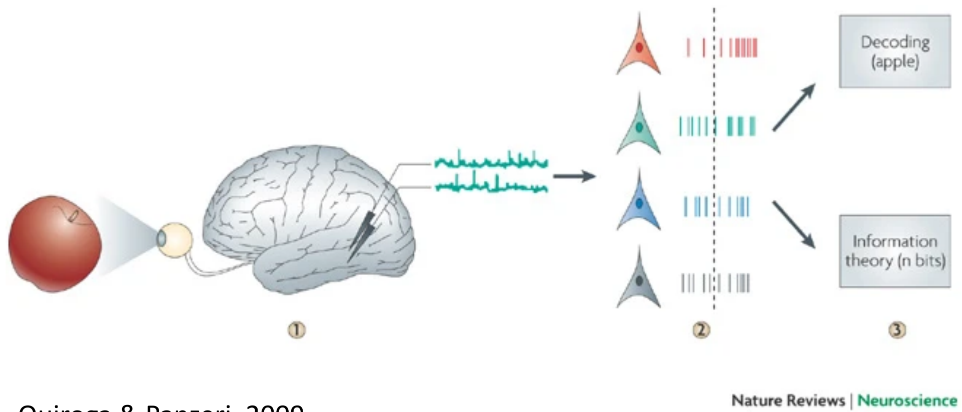
$p_x(X)$

Y - spike-train

$p_y(Y)$

1 0 1 0 0 1 0 0 1 1

- binning time (e.g. 10 bins)
- 2^10 =1024 possible spike-trains
- repeated observations (yet finite)



Quiroga & Panzeri, 2009

Information and Entropy of a "source"

Information? something about an outcome/observation that **reduces my uncertainty** (measured in <u>*bit*</u>, if log2)

- something that is non-negative
- something that decreases for increasing *p*
- something that is *0* for *p = 1*
- continuous in *p*
- something that is additive for independent events observed

$$-\log_2 p$$

(discrete) Entropy? The average *information*, related to an experiment (e.g. to a *source*)

$$0 \leq H[X] \;=\; \sum_k \; p_k(-\log_2 p_k) \;\leq \sum_k \; \frac{1}{K}(-\log_2 \frac{1}{K}) \;=\; \log_2 K$$

maximum for equally probable outcomes

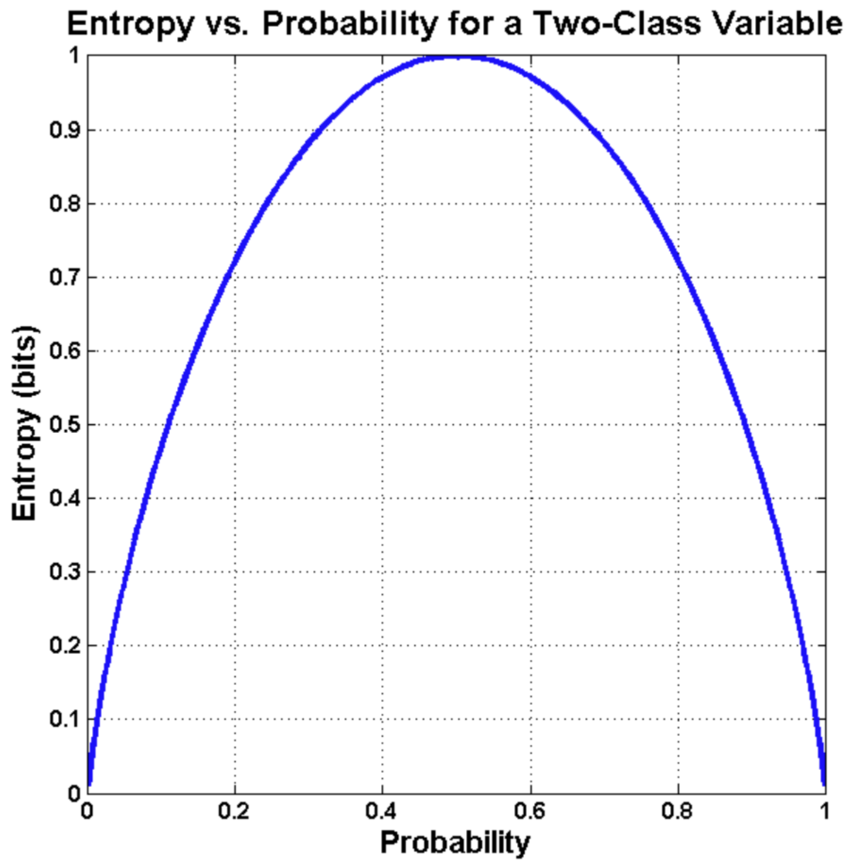
(on average) the number of yes/no questions required to know which symbol was transmitted

Information and Entropy of a “source”

Entropy of a coin (i.e. Bernoulli experiment)

$$H[X] \ = \ p(-\log_2 p) + (1-p)(-\log_2(1-p))$$

It is maximal at *p = 0.5* (equally probable outcomes)



Conditional Entropy: “given we observed *y*”

Conditional Entropy?

Source/Sender
(stimulus)

Communication channel (noisy)

Observer/Receiver
(observed spike train)

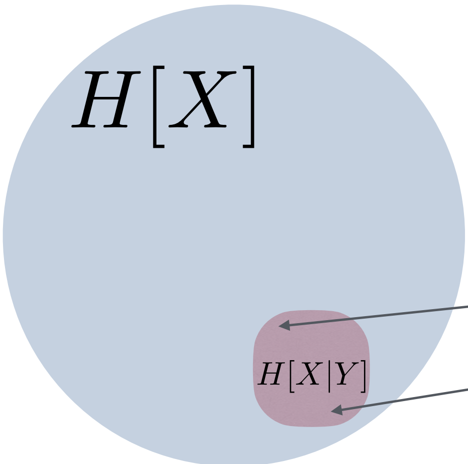
X - symbol (discrete)

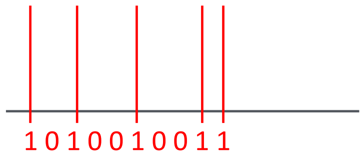
$p_x(X)$

Y - spike-train

$p_y(Y)$

$$H[X|Y] = \sum_x p(x,y)(-\log_2 p(x|y))$$





The channel is noisy!

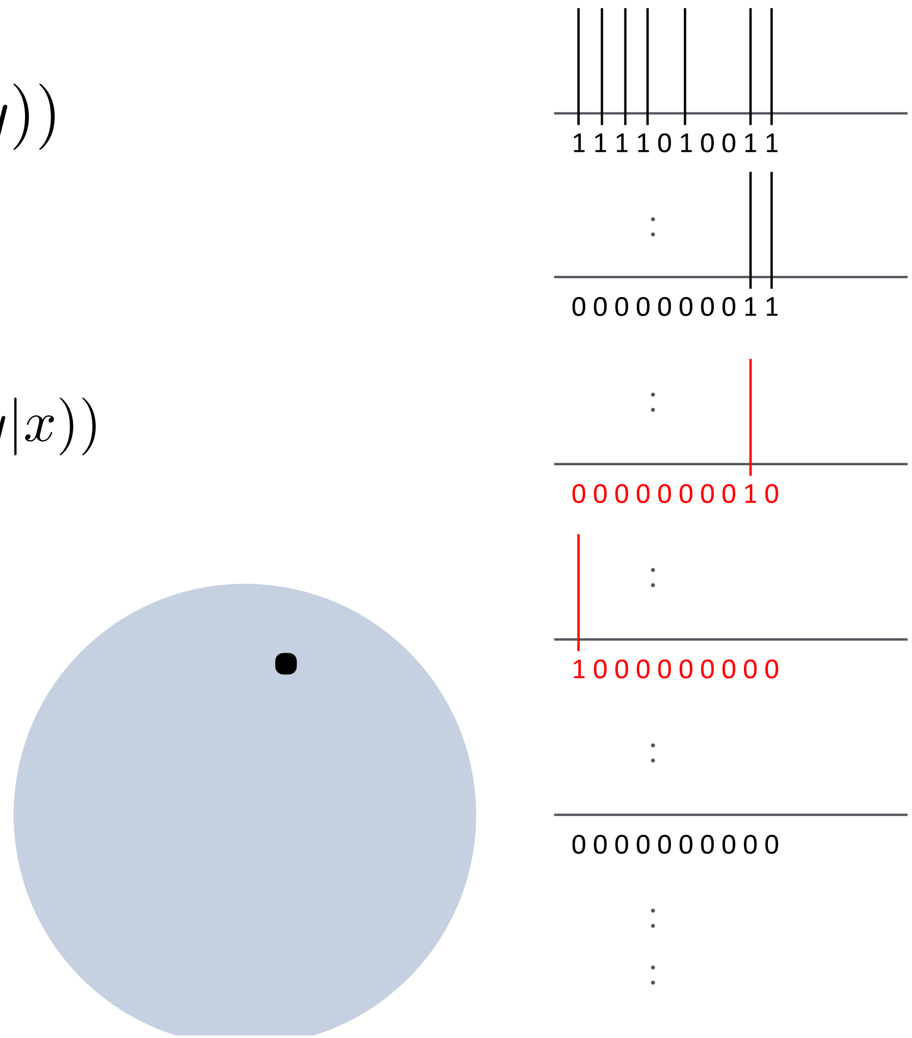
More than one stimulus

consistent with the observation.

Entropy of a spike train and its cond. entropy (given a stim.)

$$H[Y] = \sum_y p(y)(-\log_2 p(y))$$

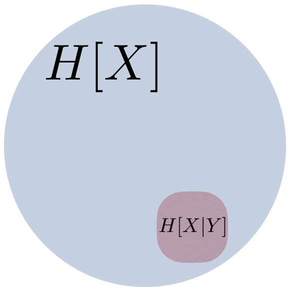
$$H[Y|X] = \sum_y p(x,y)(-\log_2 p(y|x))$$

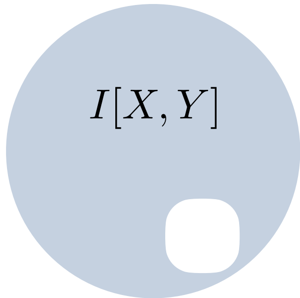


Mutual Information

$$I[X,Y] = H[X] - H[X|Y]$$

Reduction of uncertainty on the stimulus X, after observing a spike train Y.





$$I[X,Y] = H[Y] - H[Y|X]$$

Reduction of uncertainty on which spike train Y will be observed, knowing the stimulus X used.

Data Processing Inequality

$$I[X,Y] \;\geq\; I[X,Z] \qquad\qquad Z = f(Y)$$

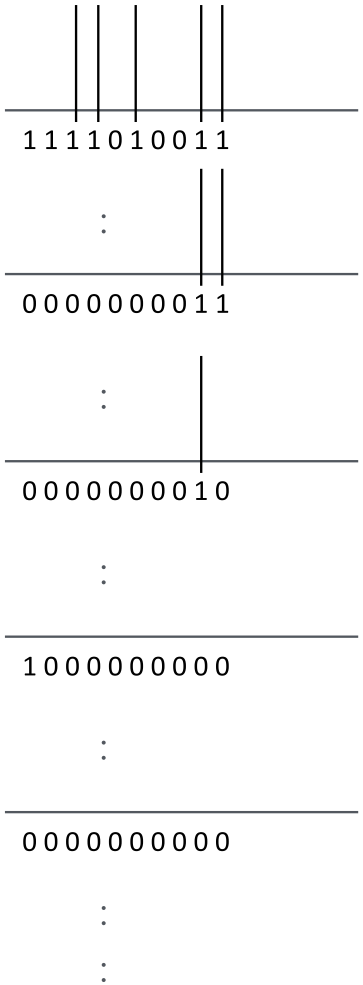
Whatever analysis/transformation/quantification Z from the spike train Y we may do,

the mutual information will NOT be over-estimated.

Criticism of Information Theory in Neuroscience

- it may not be meaningful for the brain/behavior itself
- information is difficult to estimate from (limited) data/observations
- existence of a *bias* due to our *finite* capability to observe outcomes of an experiment.

Estimating Entropy of a spike-train: biased!



M = 1024 possible spike trains (2^10 - in this *toy*-example)

$$p_1, p_2, \ldots, p_{1024}$$

$$p_i = \frac{1}{M} \qquad H[X] = \log_2 1024 = 10 \; bit$$

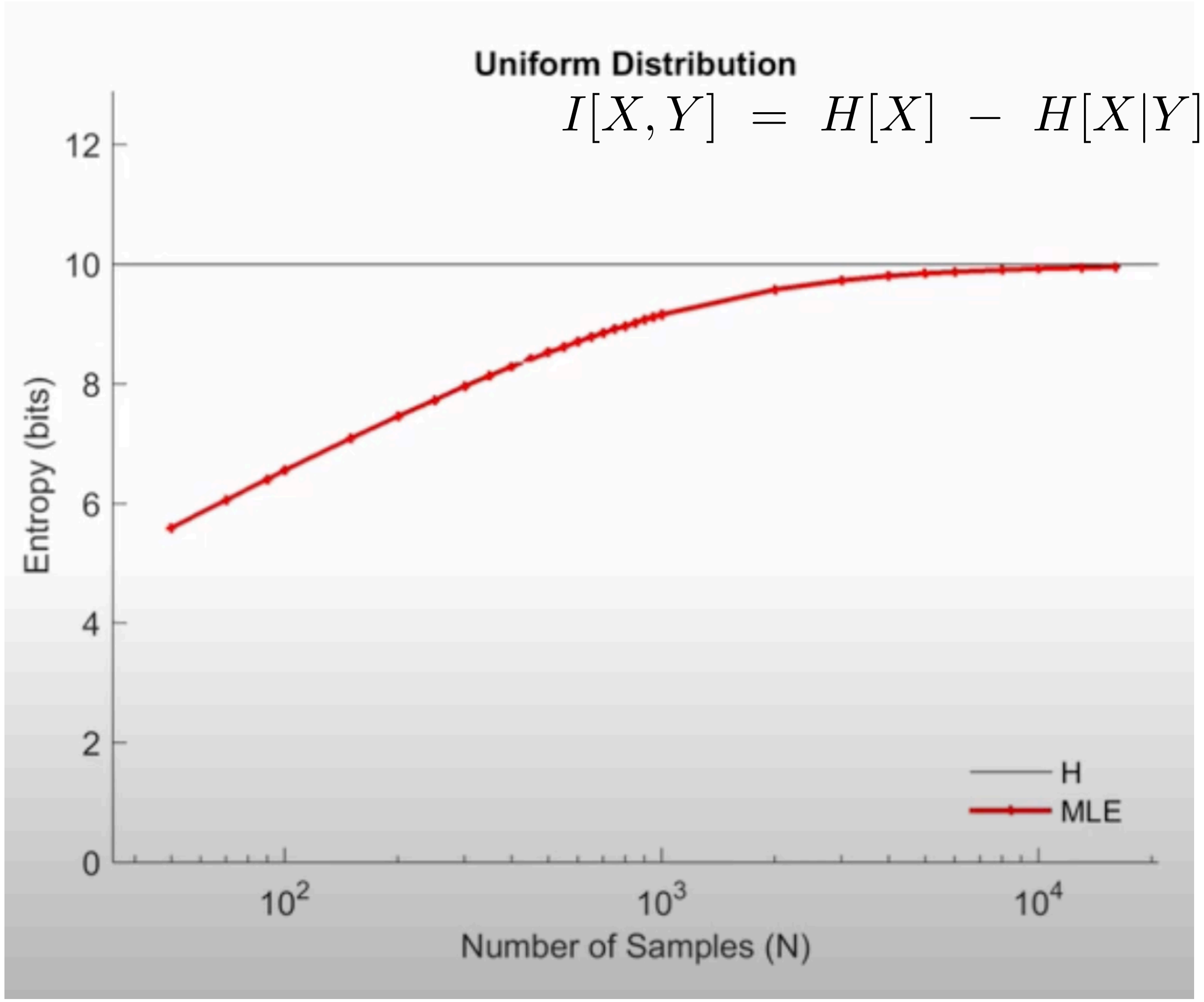
Estimating probability of a certain observation Y, requires (e.g.) computing the relative frequency *f* of occurrence (i.e. times it occurred / all trials done).

$$f_1/N, f_2/N, \ldots, f_{1024}/N$$

Experiments are finite in time! It may be difficult to estimate the *p* associated with each spike-train, or even… observe all 1024 spike-train occurring at least once!!!

$$\hat{H}[X] = \sum_k \hat{p}_k(-\log_2 \hat{p}_k)$$ max likelihood estimator

Estimating Entropy of a spike-train: biased!



$$I[X,Y] = H[X] - H[X|Y]$$

Severely underestimated (unless N>>M)

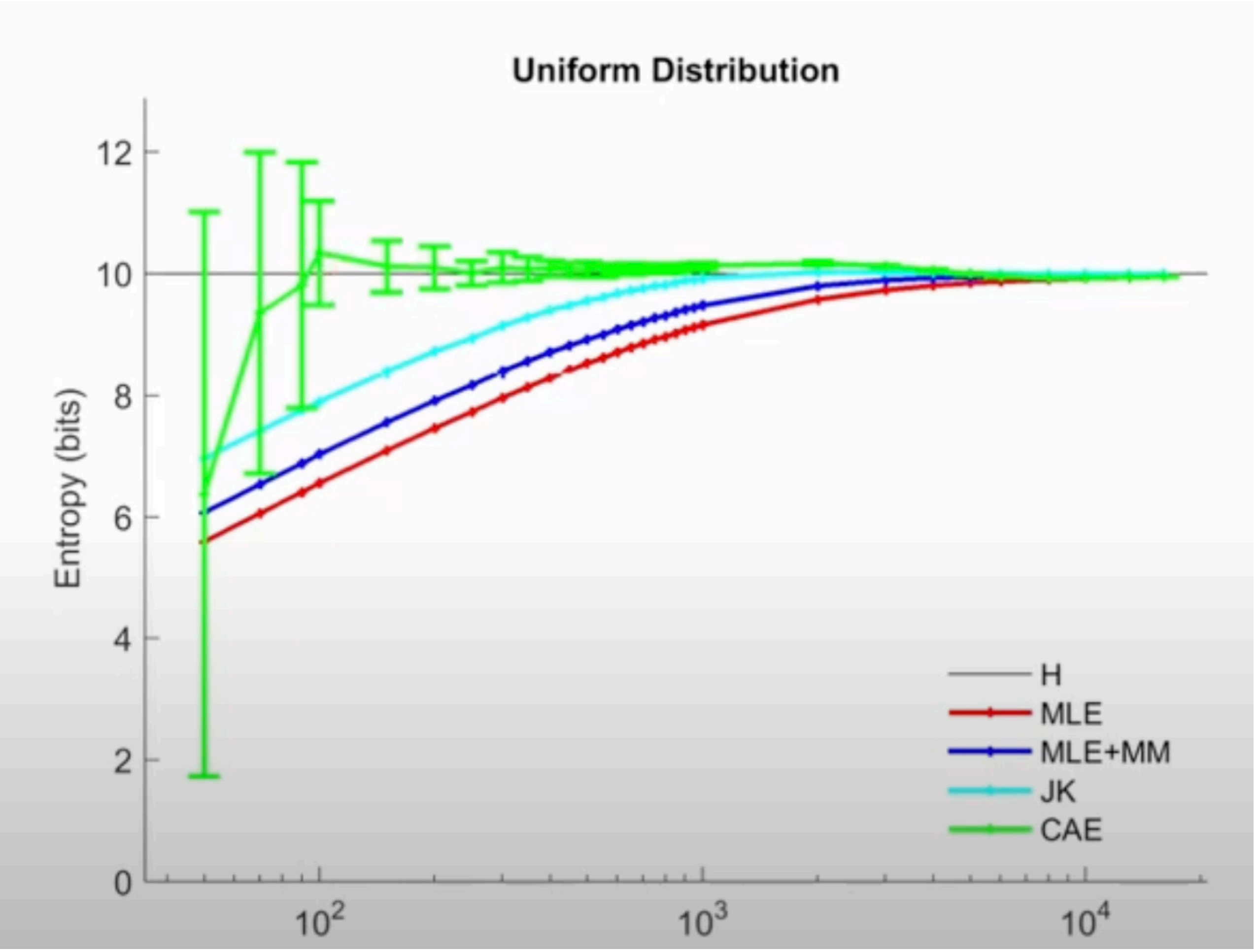
$$I[X,Y] = H[Y] - H[Y|X]$$

across **all** stim conditions
so all spike-trains observations
can be pooled together.
—> small bias, because large(r) number.

across a (small) subset of observations! Large bias!

<u>So mutual observation is over-estimated!!!</u>

Several “correction” methods exist



Miller Maddow (MM) correction
(subtracts from the entropy, one first-order Taylor estimate of the bias)

Jack-Knife (JK) Montecarlo resampling
(subtracts from the entropy the entropy calculated with all observations Y leaving one out)

Coverage-adjusted estimator (CAE)
(shrinks *p* estimates for observed Y and inflate *p* for rare Y)