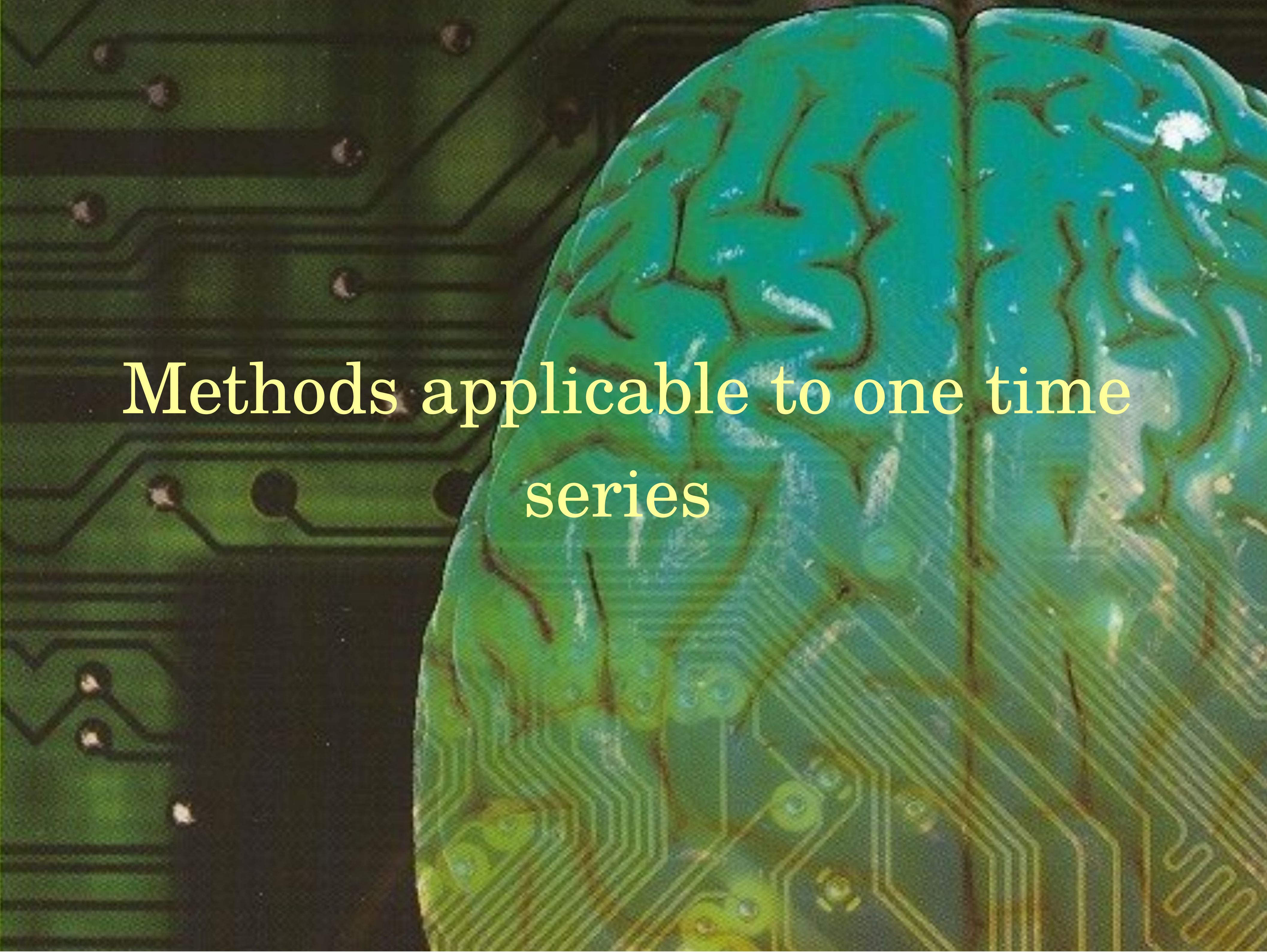
Data analysis methods in the neuroscience

Zoltán Somogyvári

Wigner Research Centre for Physics of the Hungarian Academy of Sciences

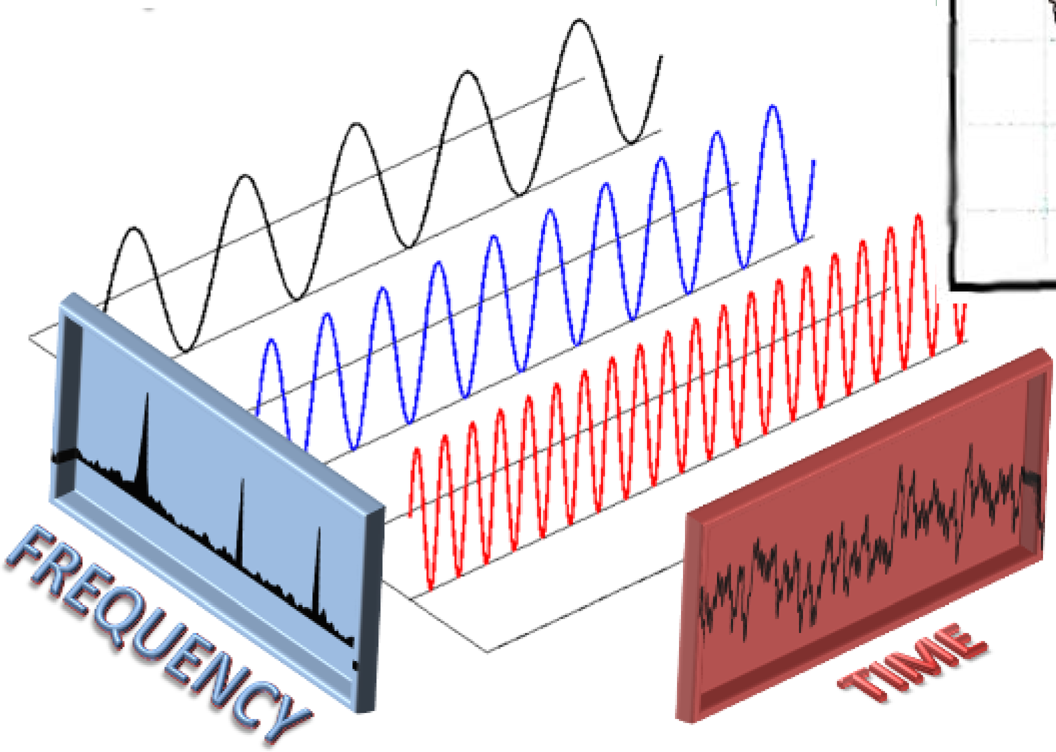


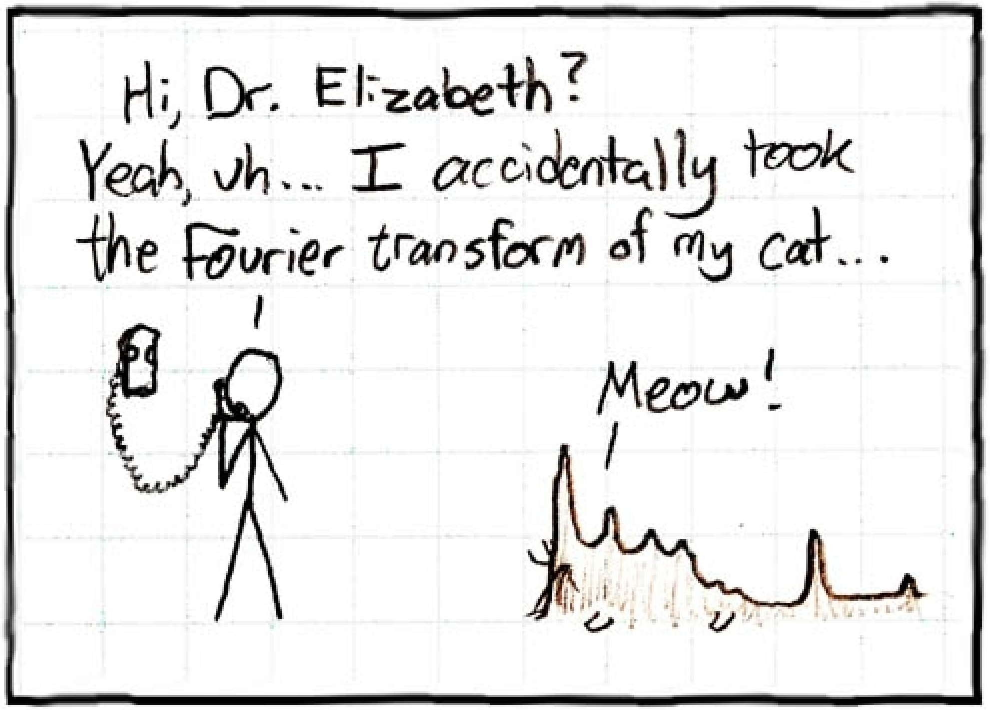
Methods applicable to one time series

The Fourier transformation

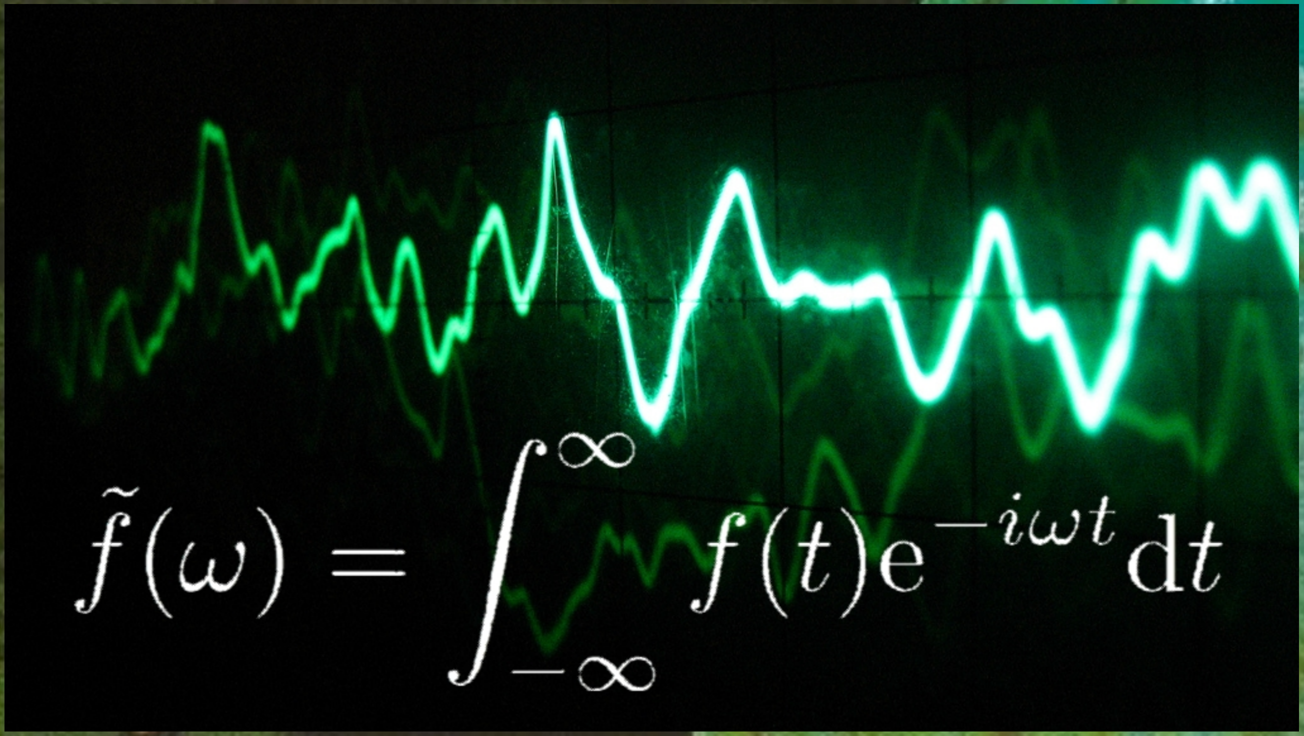
$$g(t) = a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{2\pi mt}{T}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nt}{T}\right)$$

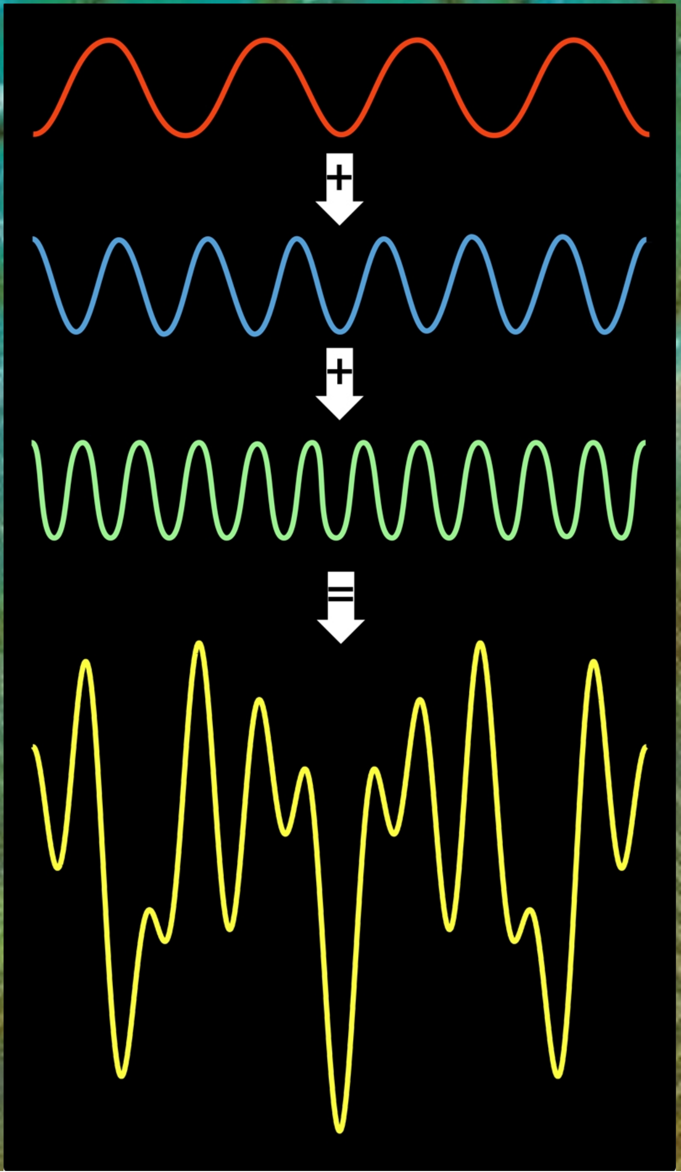
$$= \sum_{m=0}^{\infty} a_m \cos\left(\frac{2\pi mt}{T}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nt}{T}\right)$$





The Fourier transformation





The Fourier transformation

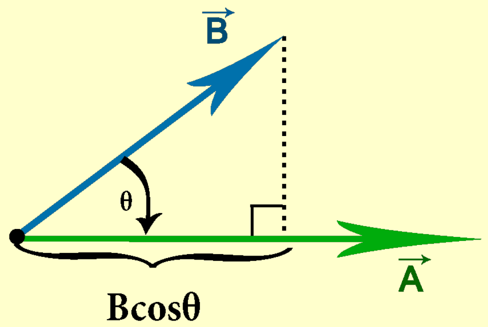
$$\bar{A}\cdot\bar{B} = AB\cos\theta = (A\cos\theta)B = A(B\cos\theta)$$

$$i\cdot i = j\cdot j = k\cdot k = 1; \quad i\cdot j = j\cdot k = k\cdot i = 0;$$

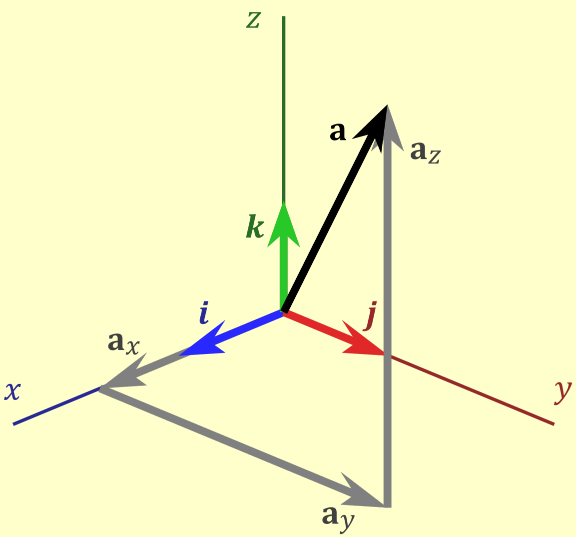
$$\bar{A}\cdot\bar{B} = (A_x i + A_y j + A_z k)\cdot(B_x i + B_y j + B_z k)$$

$$= A_x B_x + A_x B_x + A_x B_x$$

$$\bar{A}\cdot\bar{B} = AB(\hat{e}_A\cdot\hat{e}_B) = AB(1)(1)\cos\theta = AB\cos\theta$$



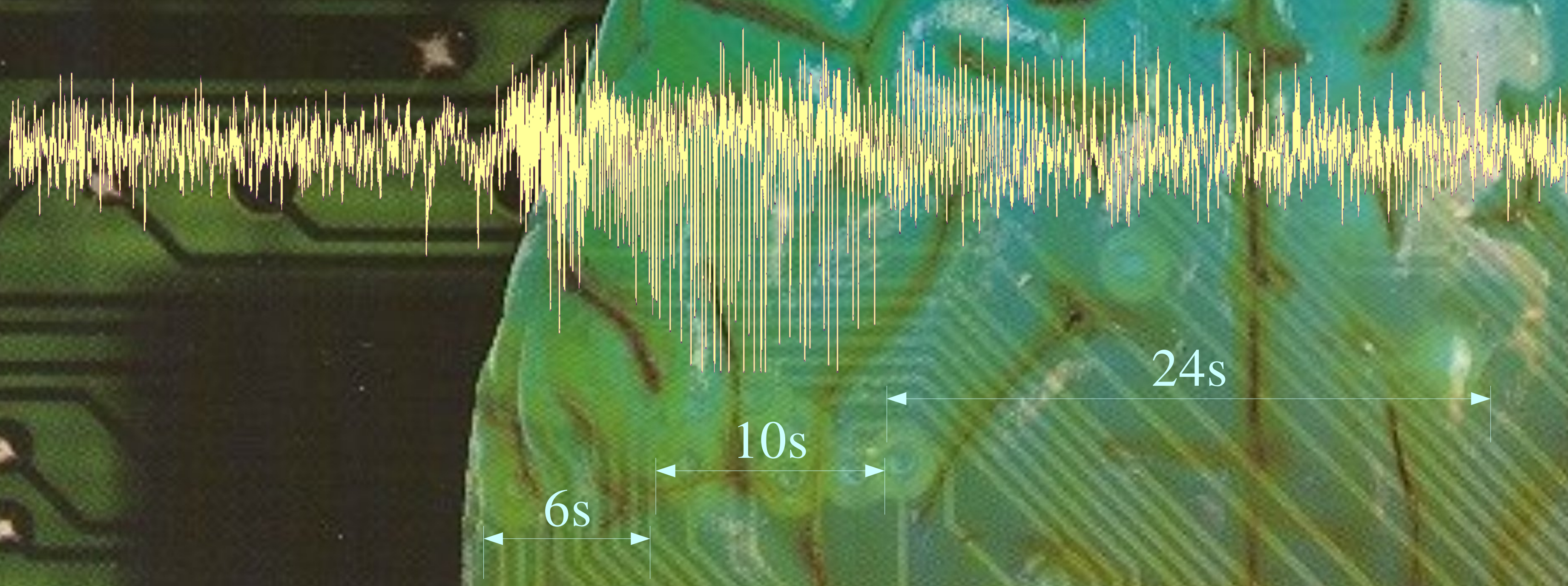
$\vec{A}\,.\,\vec{B} = |\vec{A}|.|\vec{B}|.\cos\theta$



Coordinates: projection (dot product) onto the orthogonal unit vectors (base) of the coordinate system

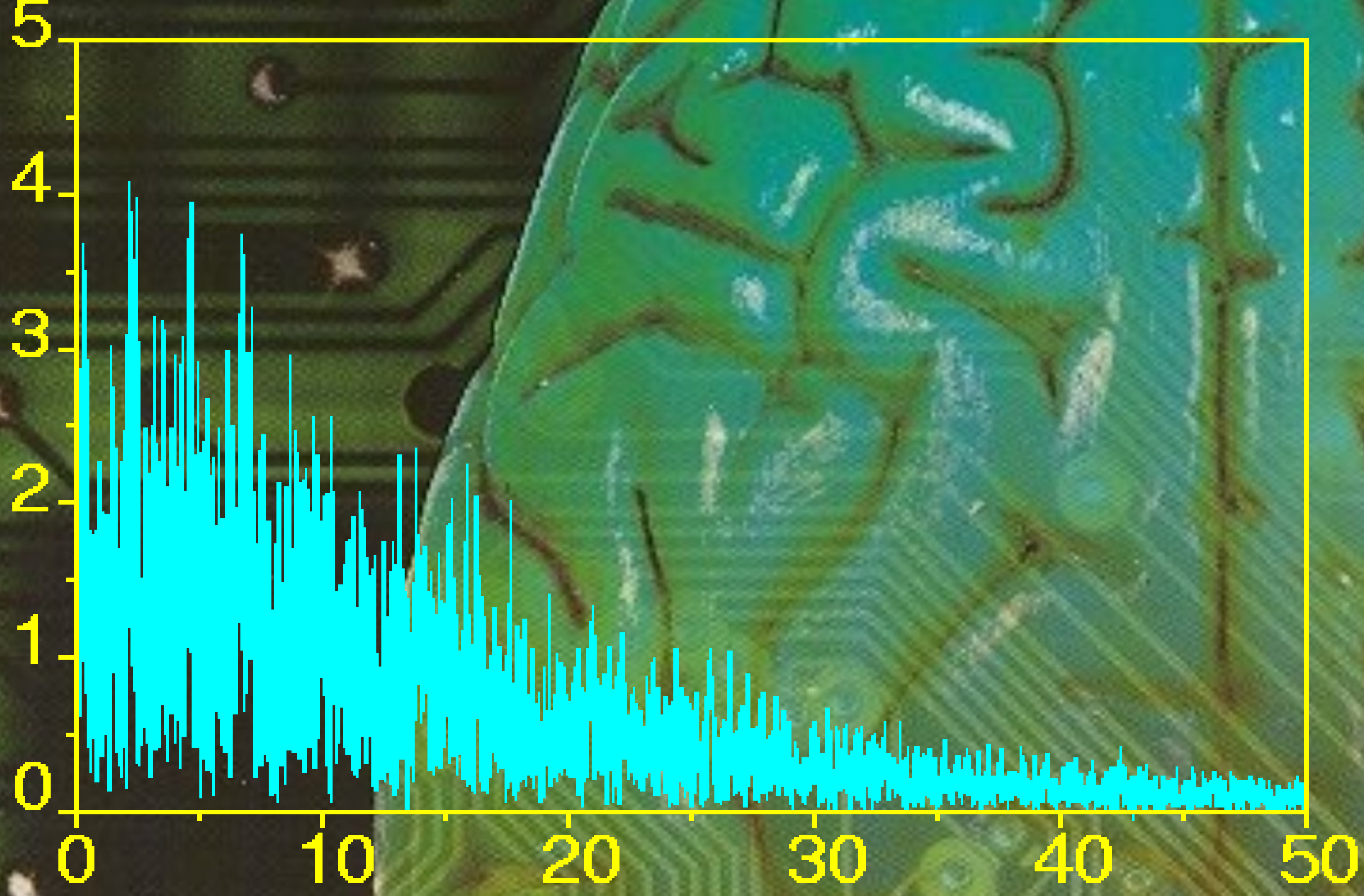
Example: Slow dynamics of the epileptic seizure

An experimental epilepsy model: Generalized epilepsy evoked by local application of 4-Aminopyridin, ECoG:



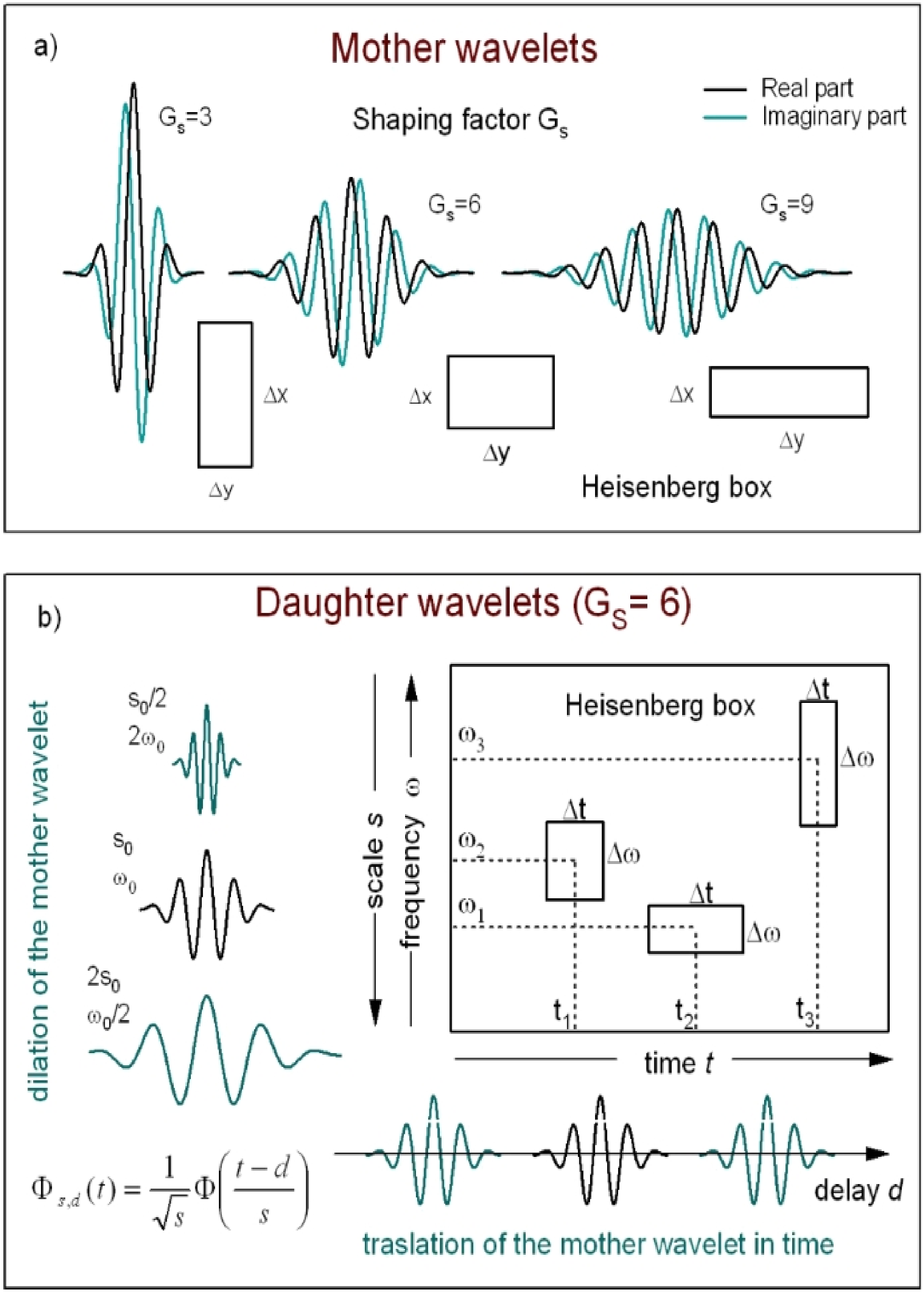
Three phases of the seizure can be distinguished, based on amplitudes, frequencies and waveforms.

The Fourier spectrum



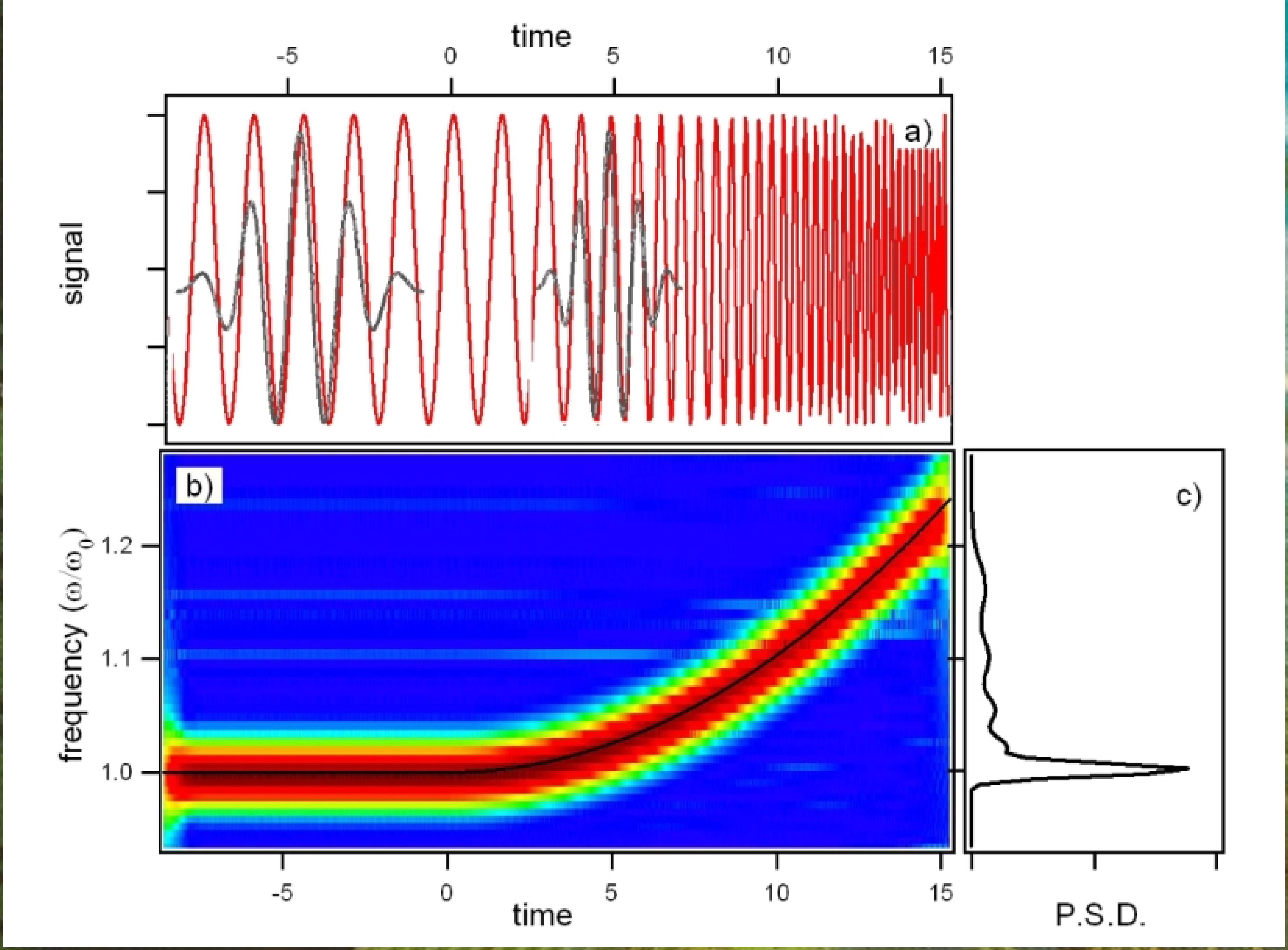
Frequency (Hz)

Wavelet-transformation

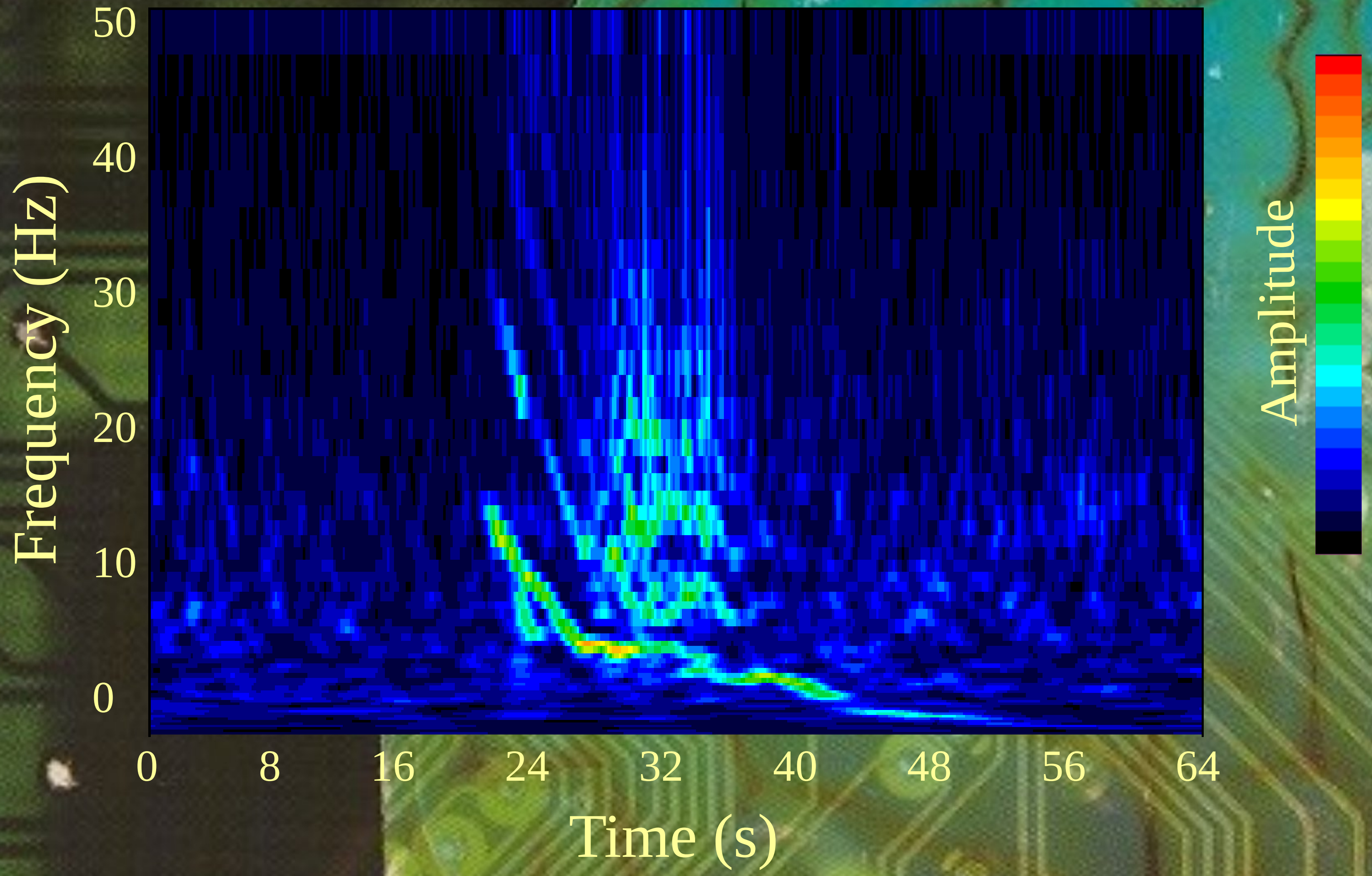


$$\Phi_{s,d}(t) = \frac{1}{\sqrt{s}}\Phi\left(\frac{t-d}{s}\right)$$

Wavelet-transformation



Wavelet-transformation of the ECoG



Information theoretical measures

Entropy:

$$H(X) = -\sum_x p_x \log(p_x)$$

Entropy is a measure of disorder and informaition content

P_x is the probability of state x

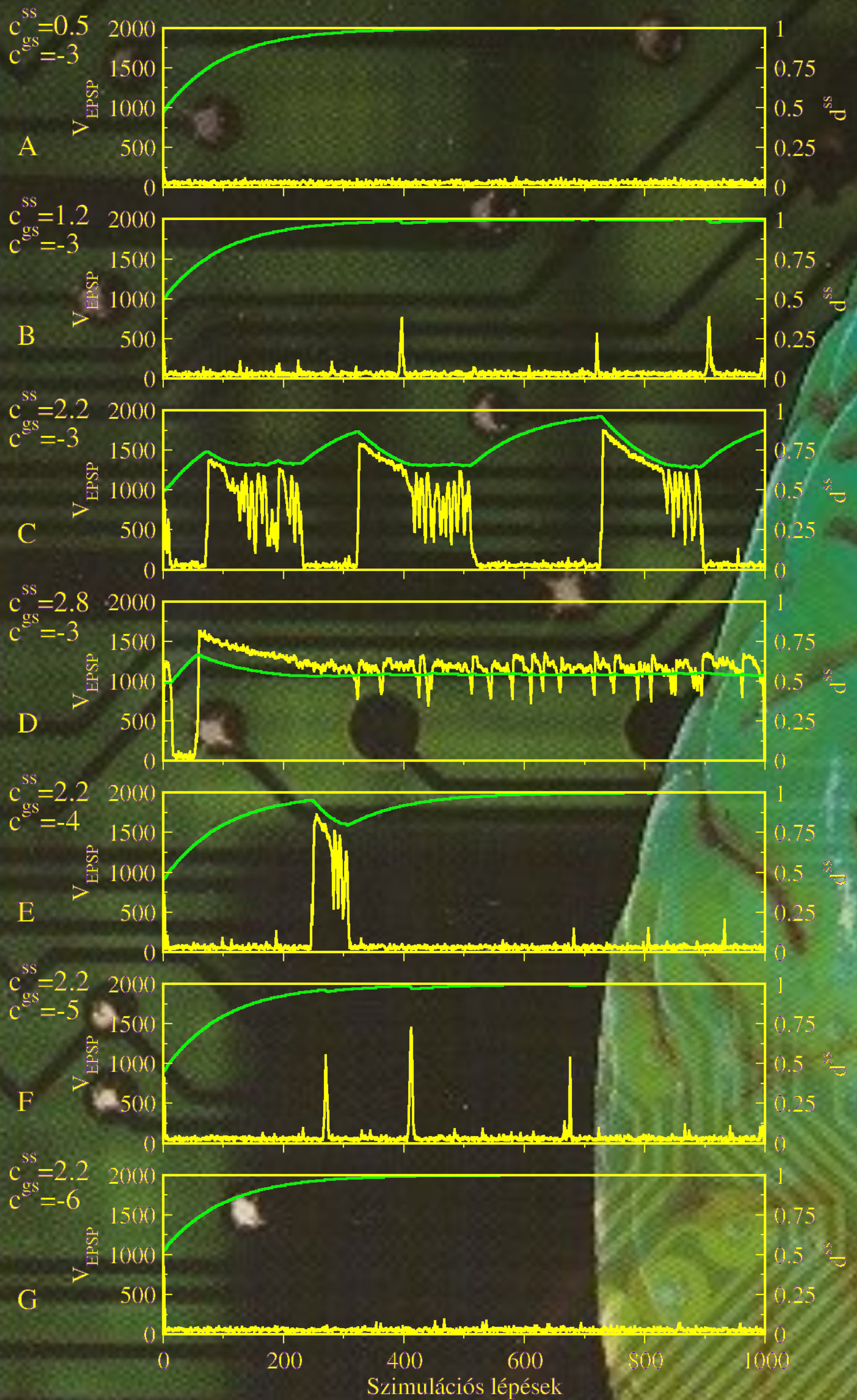
Depending on the state space, there are different entropies

Spectral entropy, approximate entropy...

Phase-space reconstruction

The reconstructed pseudo-attractor in the state space, constructed from the data and its derivatives ($a(t)$, $a^1(t)$, $a^2(t)$ ...) is topologically equivalent to the systems real attractor in its original state space, according to the Whitney theorem.

Derivation increases noise, so the ($a(t)$, $a(t+dt)$, $a(t+2dt)$ ... delayed coordinates, return maps are used in stead.

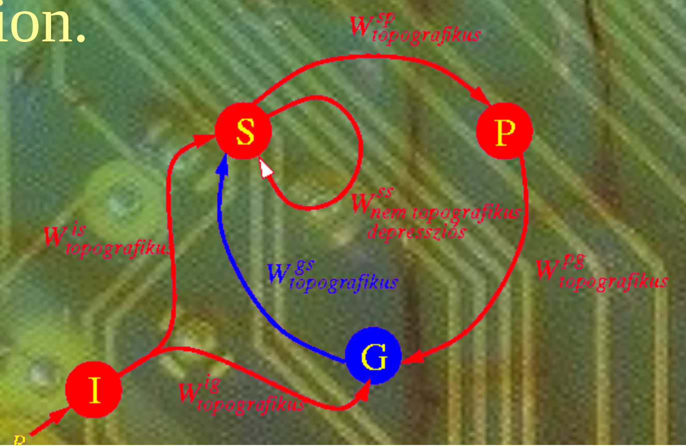


A simple epilepsy model

The change in the relative strength of the recurrent excitation and in inhibition results in:

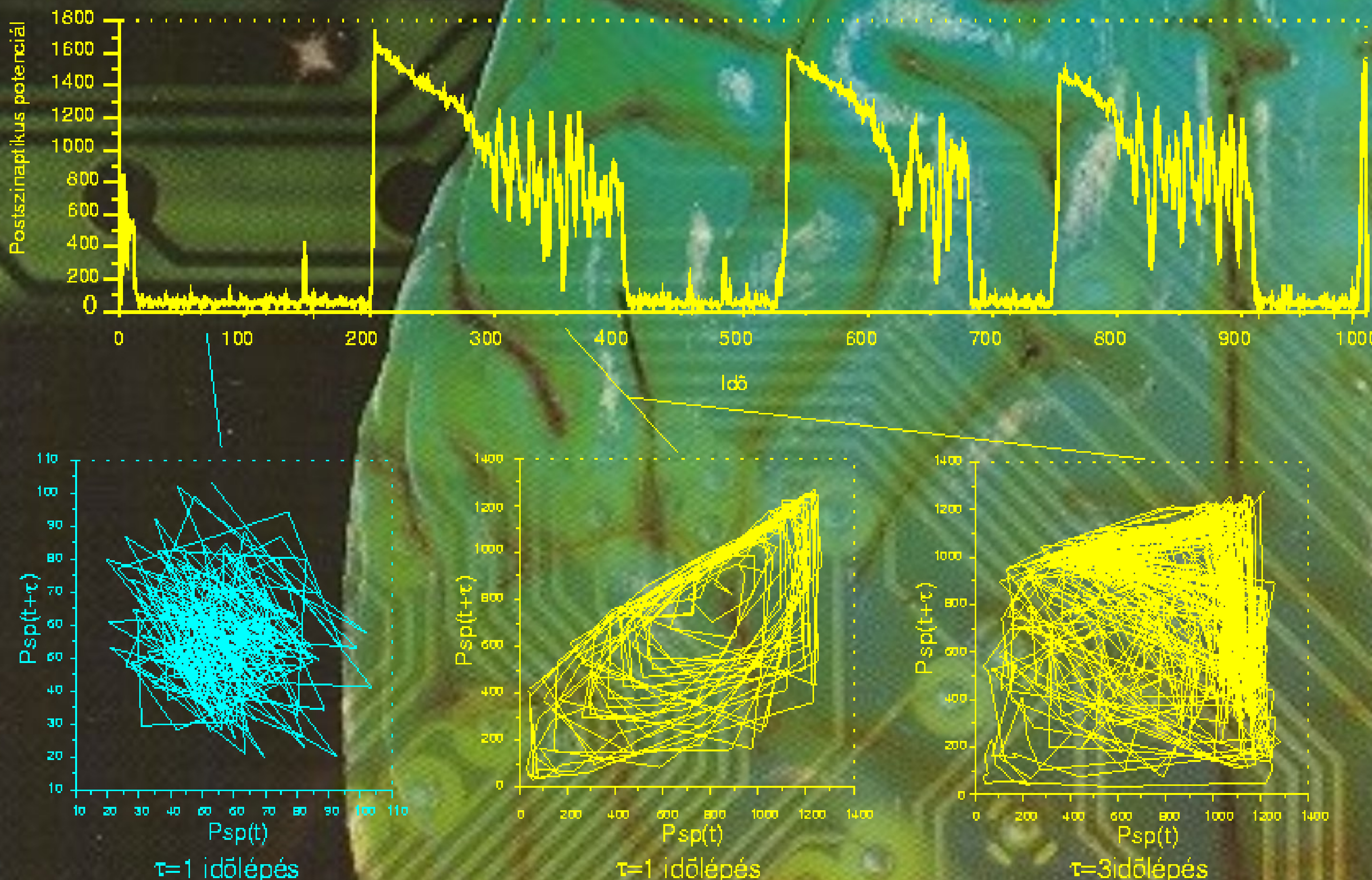
- spikes
- seizures with complex dynamics
- status epilepticus

The seizures can be eliminated by increasing the strength of the inhibition.

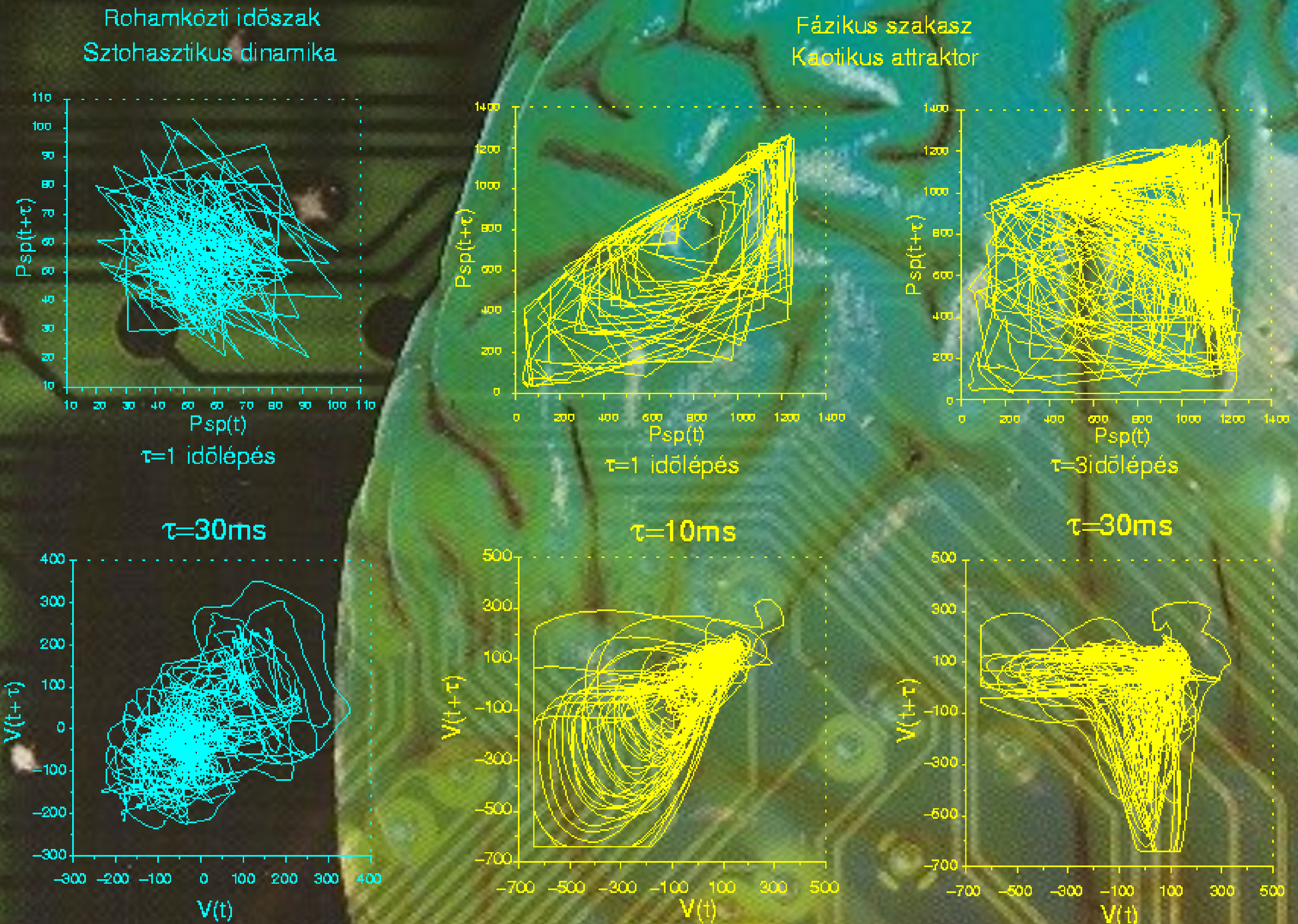


Reconstructed attractors from the simulated time series and their changes

The synaptic depression decreases the activation and drives the system into the regime of the irregular (chaotic) oscillation



Comparison of the reconstructed attractors from the simulation and the epileptic ECoG



Phase space reconstruction

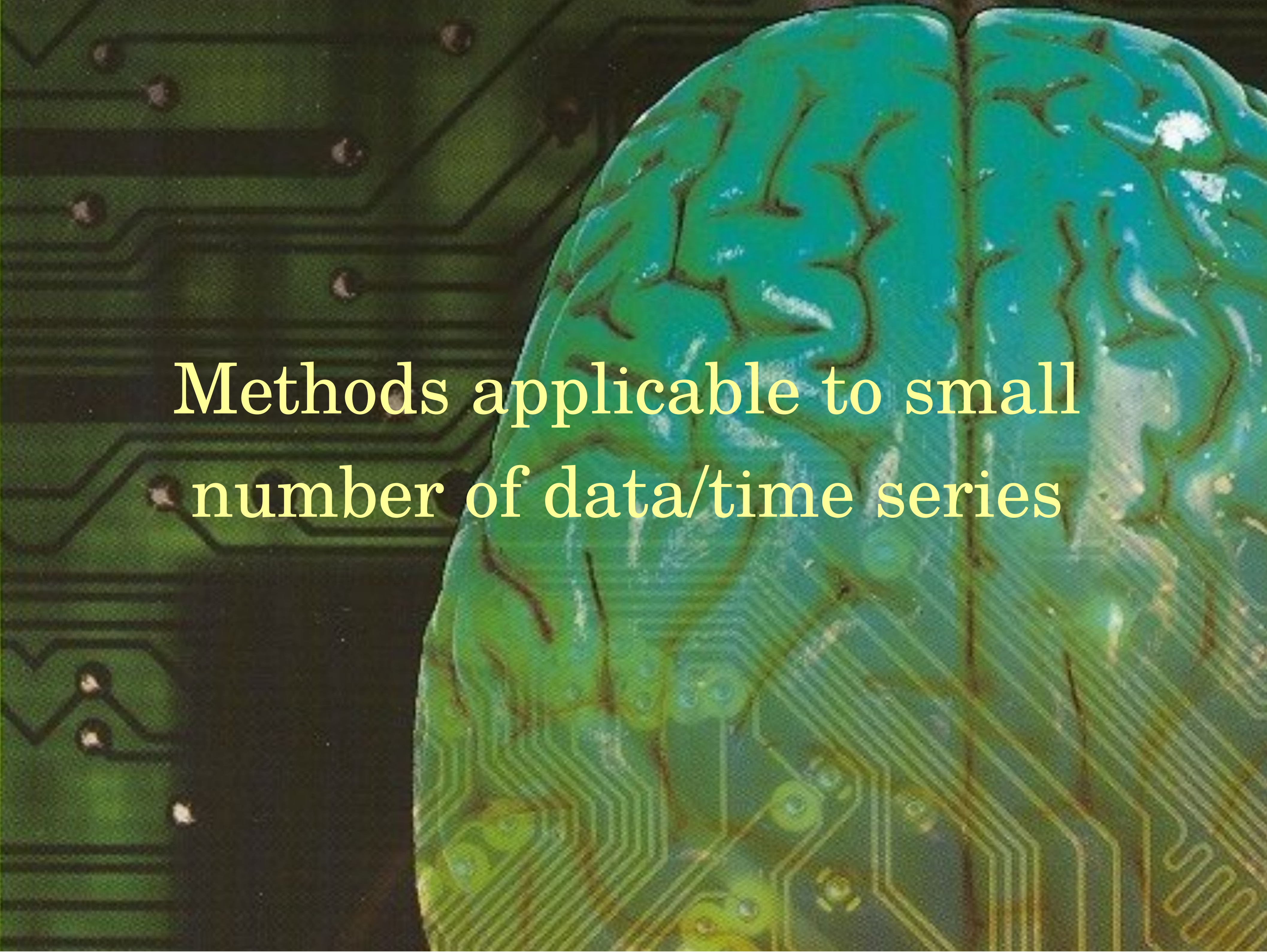
What to do with the reconstructed attractors?

It is not easy to determine the type (topology) of the attractor, based on the noisy measurements.

It is possible to measure its dimension, for example: L2-dimension. $N=L^d$ where N is the number points in a sphere with radius L.

It is possible to measure the average Ljapunov-exponent, meaning the average instability of the paths.

What else?


Methods applicable to small
number of data/time series

Correlation vs. Coherence

The linear correlation coefficient

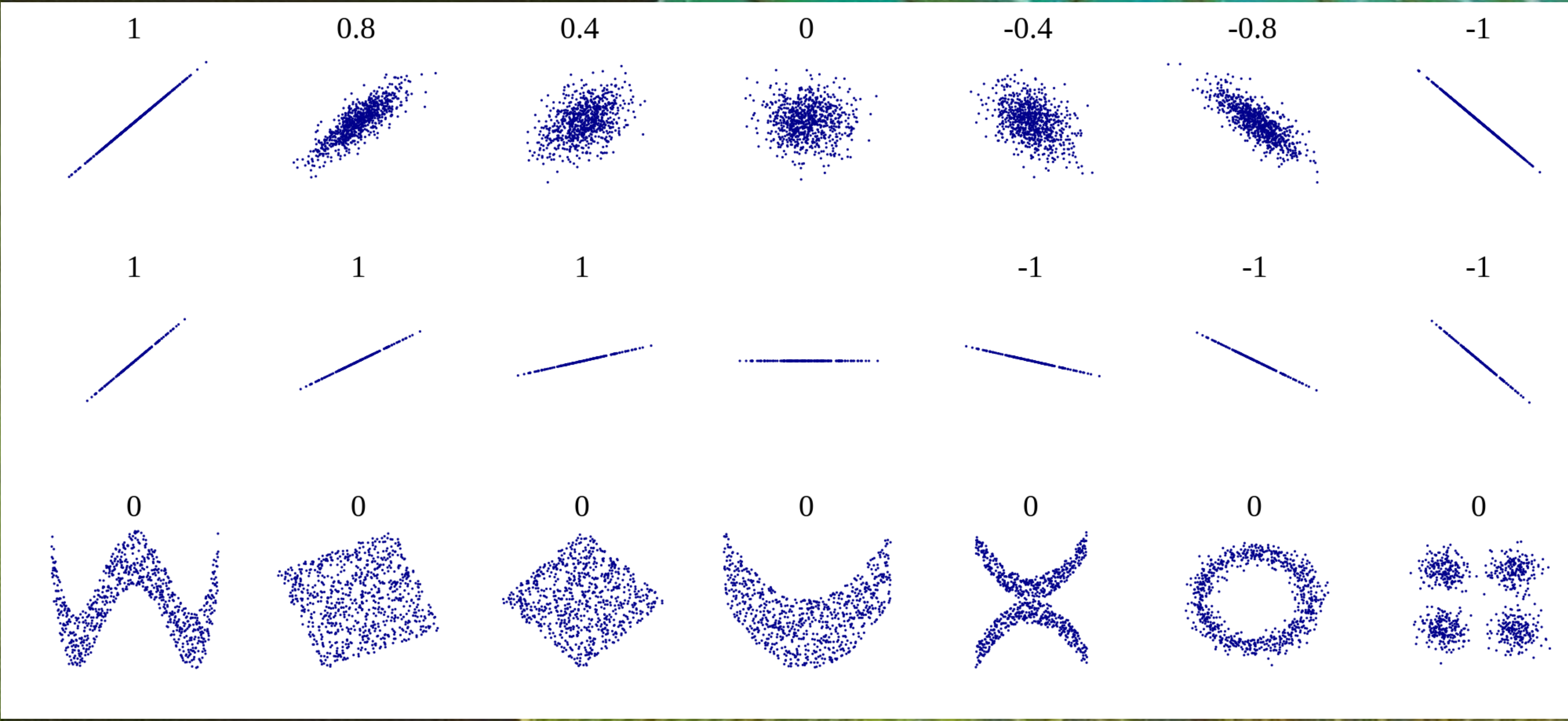
$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Coherence spectrum

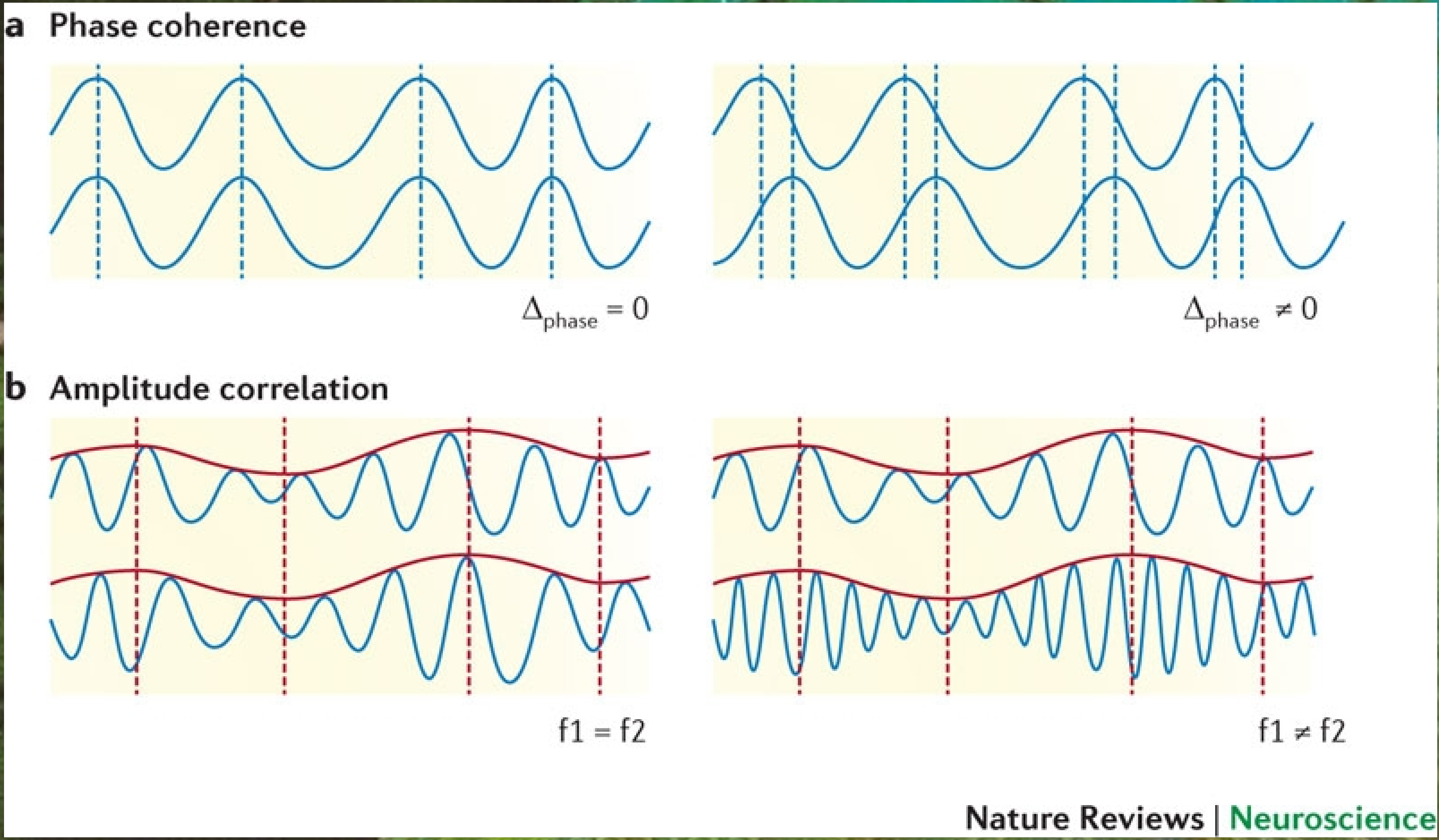
$$Coh(f) = \frac{\left|\sum_{i=1}^{N} F_1(f) \cdot F_2^*(f)\right|^2}{\sum_{i=1}^{N}\left|F_1(f)\right|^2 \cdot \sum_{i=1}^{N}\left|F_1(f)\right|^2}$$

Correlation vs. Coherence

Meaning of the linear correlation coefficient



Correlation vs. Coherence



ARMA and ARIMA model fitting

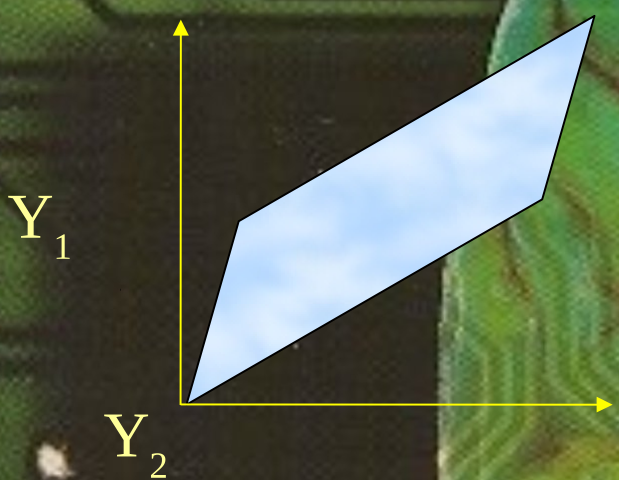
$$X(t) = \sum A_i X(t-i)$$

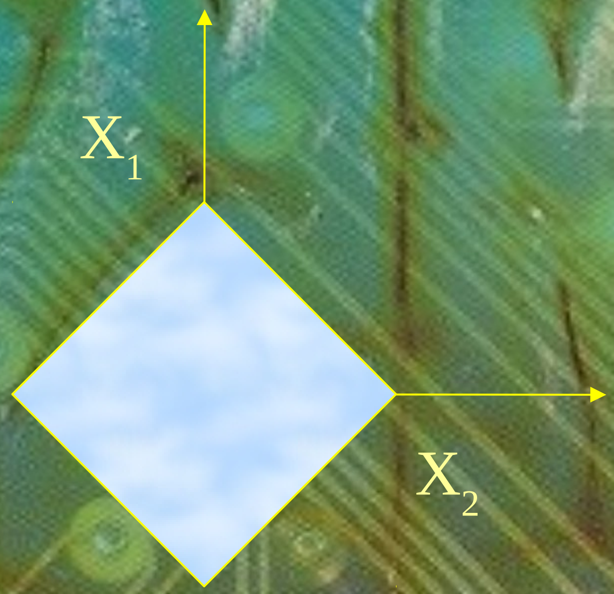
$$X(t) = \sum A_i X(t-i) + \sum B_j X'(t-j)$$

The cocktail-party problem and the principal component analysis (PCA)

$$Y_i(t) = \sum W_{ij} X_j(t)$$

Let's search for the directions correspond to maximal variance





The cocktail-party problem and the independent component analysis (ICA)

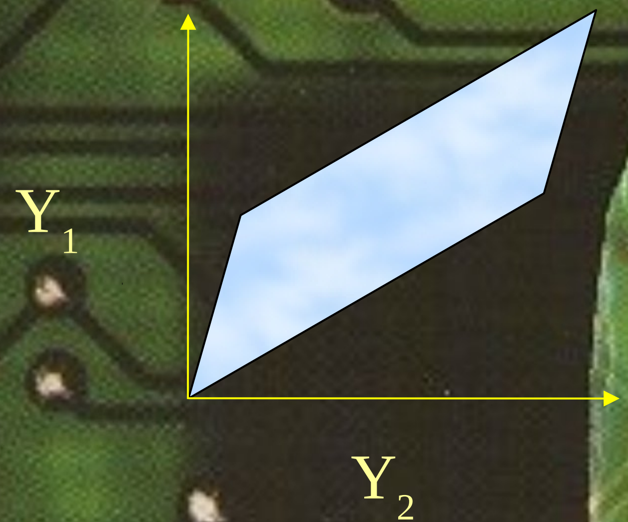
$$Y_i(t) = \sum W_{ij} X_j(t)$$

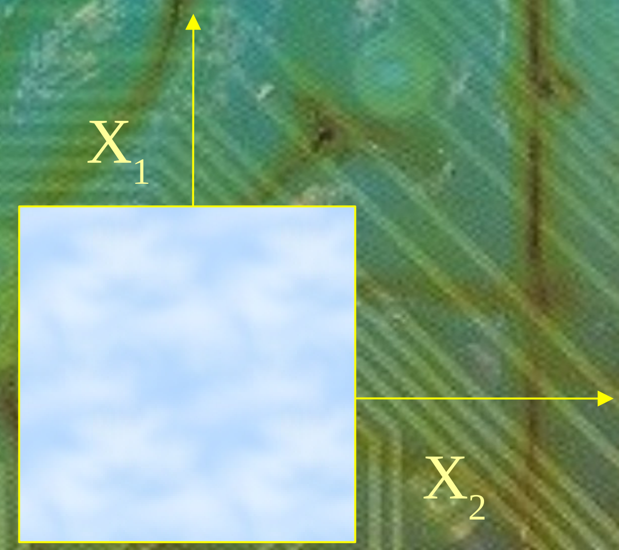
Let's search for the most independent directions!
The basic idea is the central limit theorem: Linear combination of two independent variables is closer to the Gaussian distribution than the original. Thus, let's search for the least Gaussian sources. How to measure the "non-Gaussianity"? Eq: Skewness, entropy...

The cocktail-party problem and the independent component analysis (ICA)

$$Y_i(t) = \sum W_{ij} X_j(t)$$

The most independent directions:

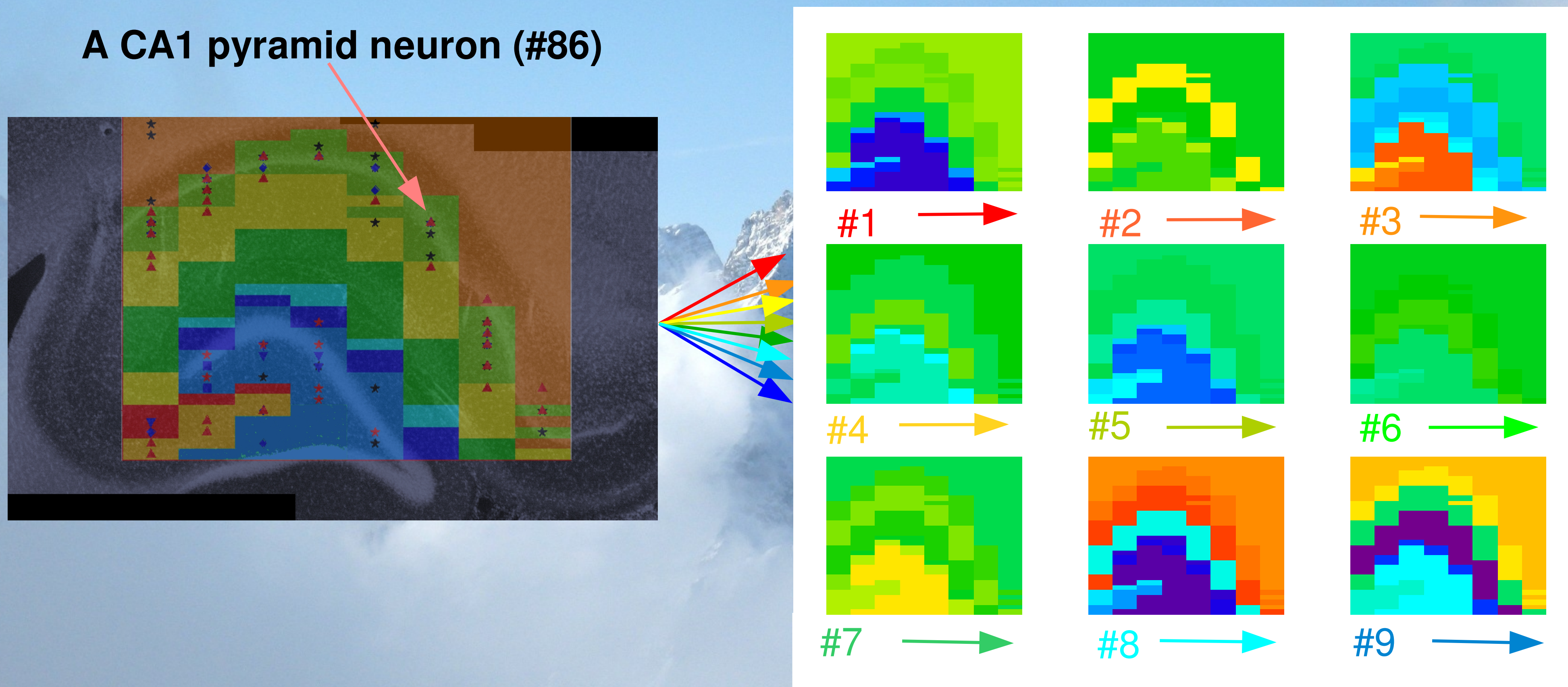




Micro-electro imaging

Inputs of a neurons from different layers

A CA1 pyramid neuron (#86)

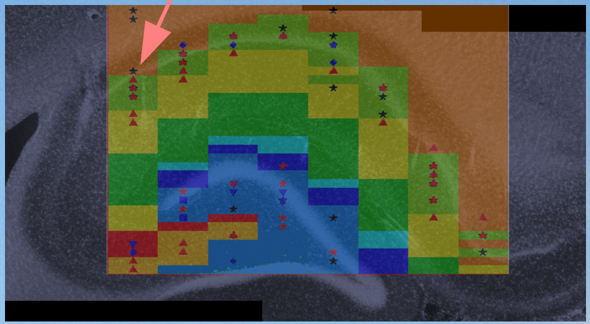


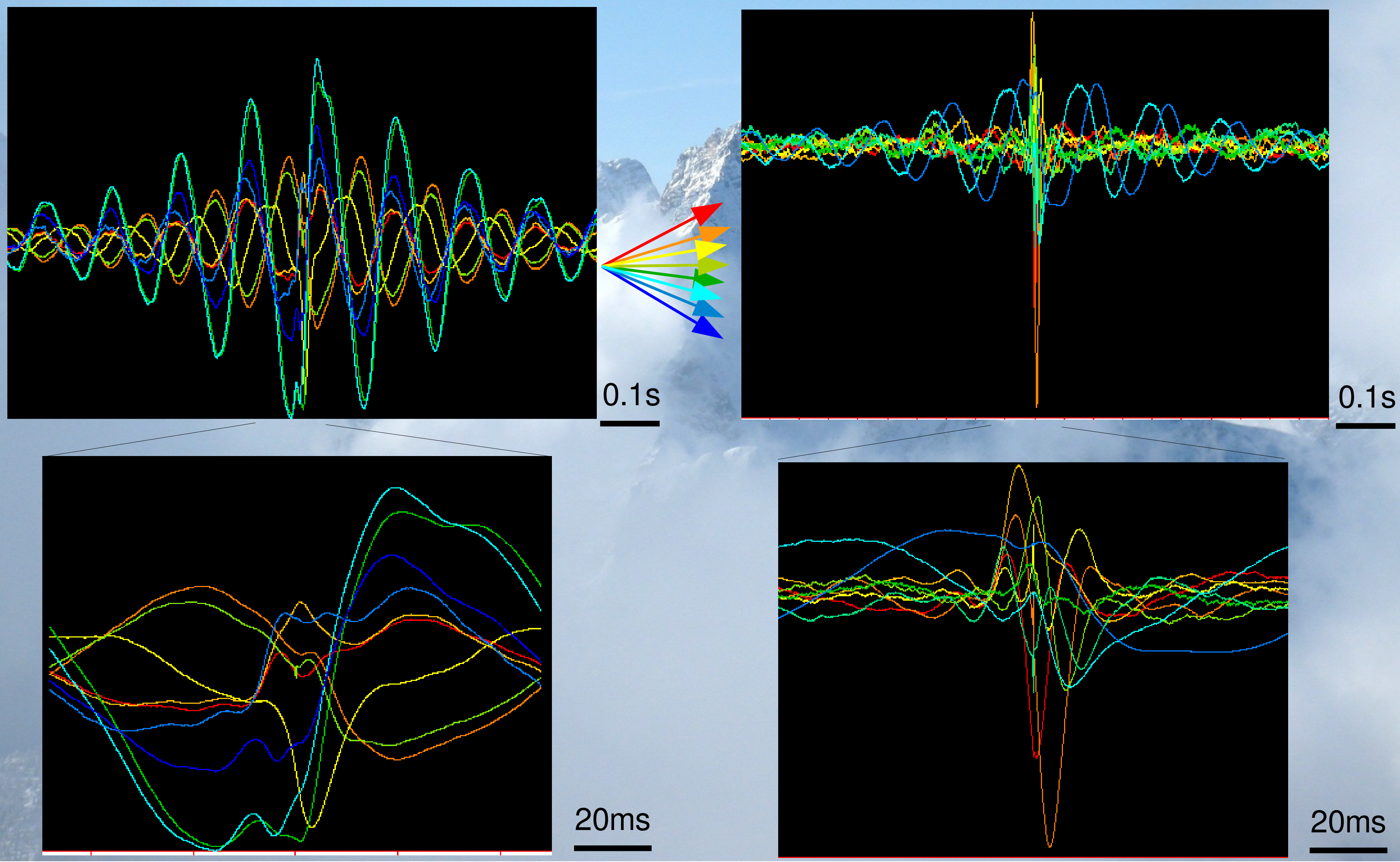
The spike triggered average EC potential patterns have been decomposed into 9 different independent components by ICA. Some of them clearly corresponds to the signals of specific pathways and mechanisms: component #2 corresponds to Schaffer collateral, #8 and #9 together correspond to the Theta.

Micro-electro imaging

Inputs of a neurons from different pathways: ICA

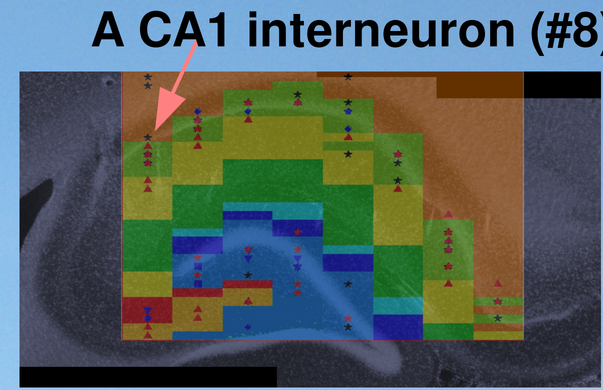
A CA1 interneuron (#8)

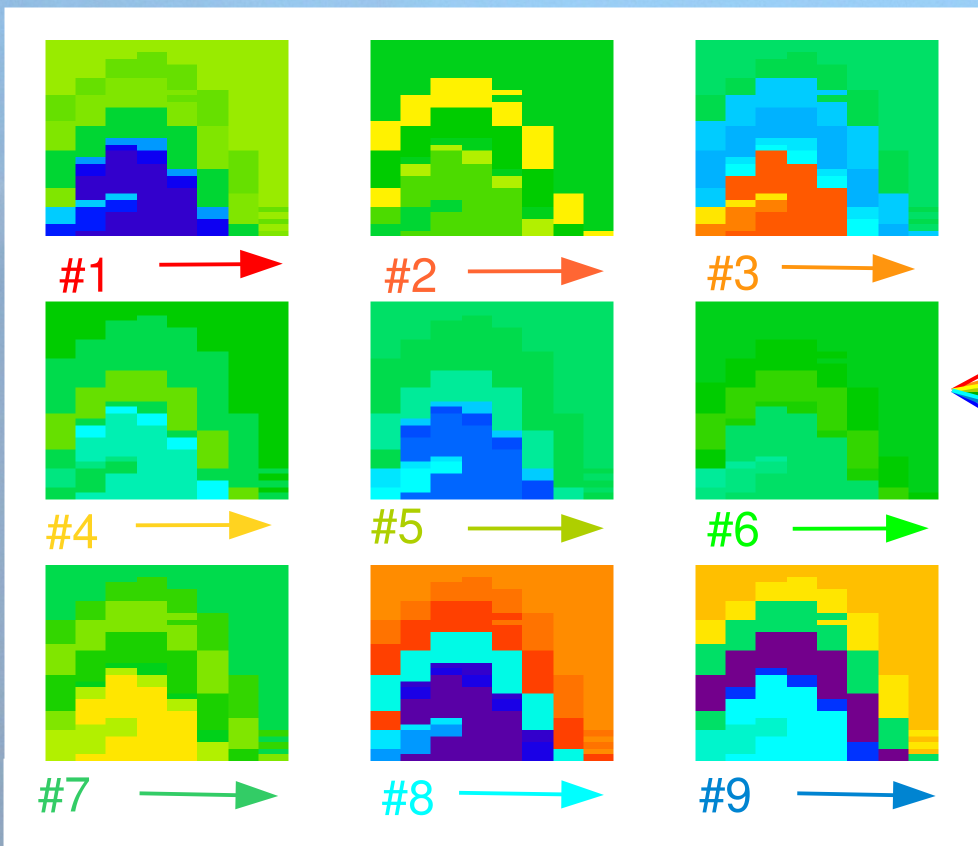


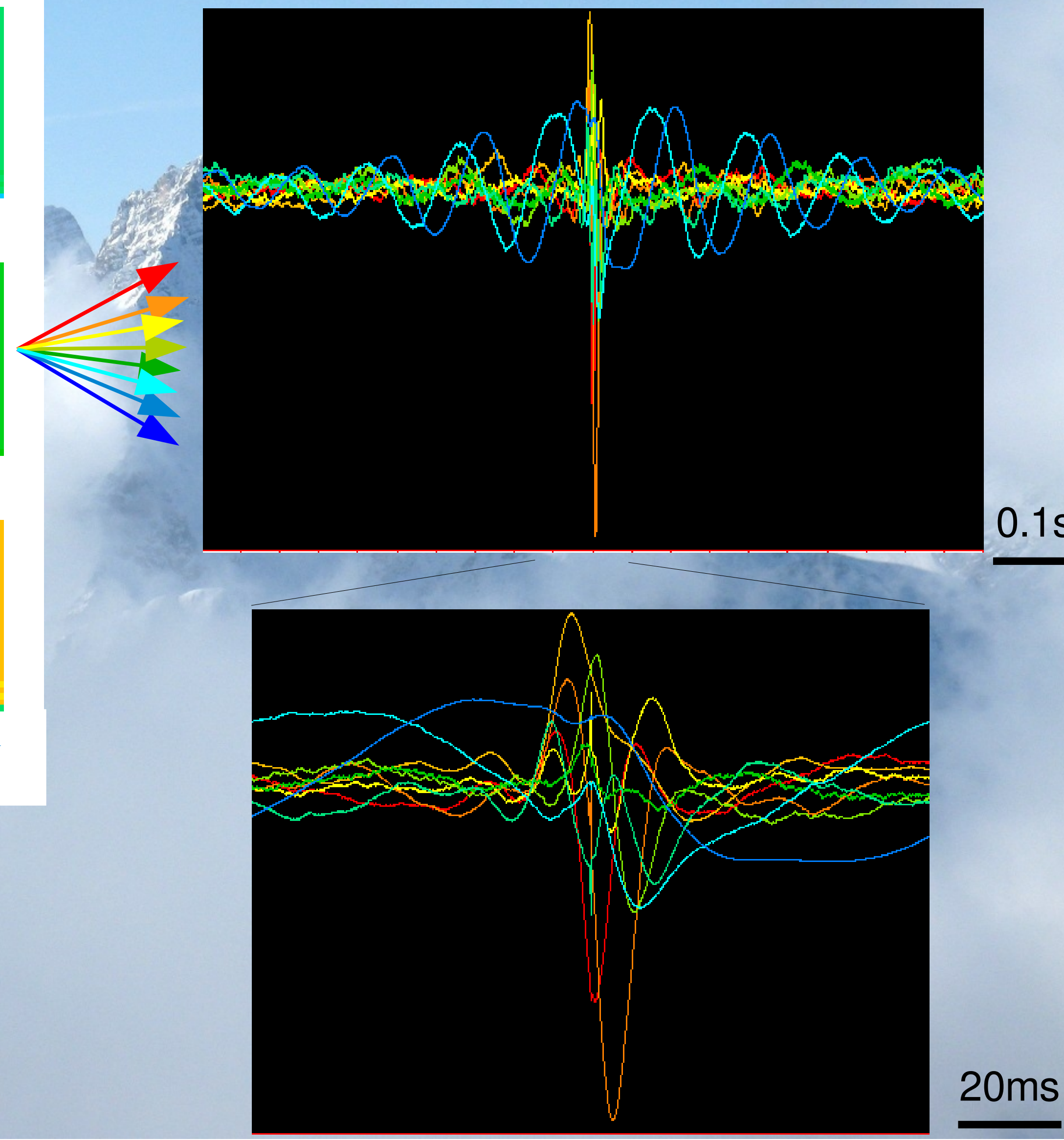


Micro-electro imaging

Inputs of a neurons from different pathways: ICA

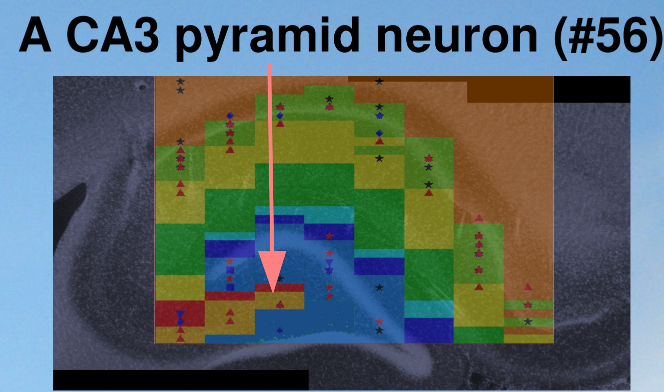


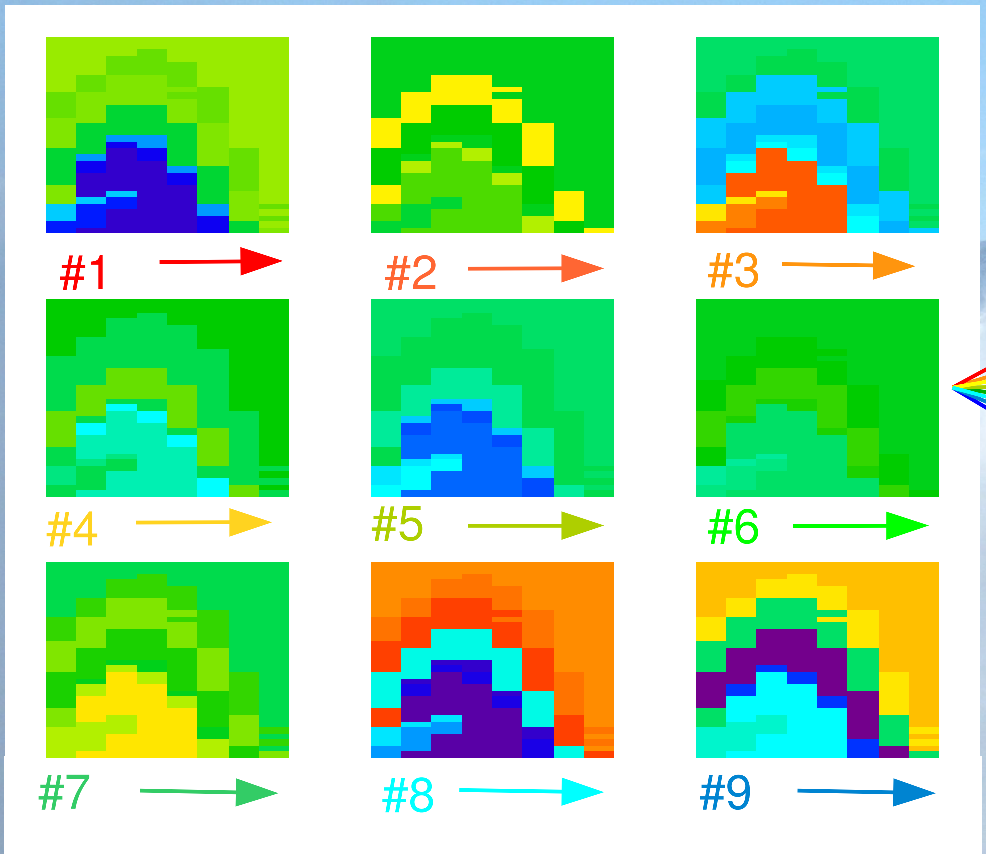


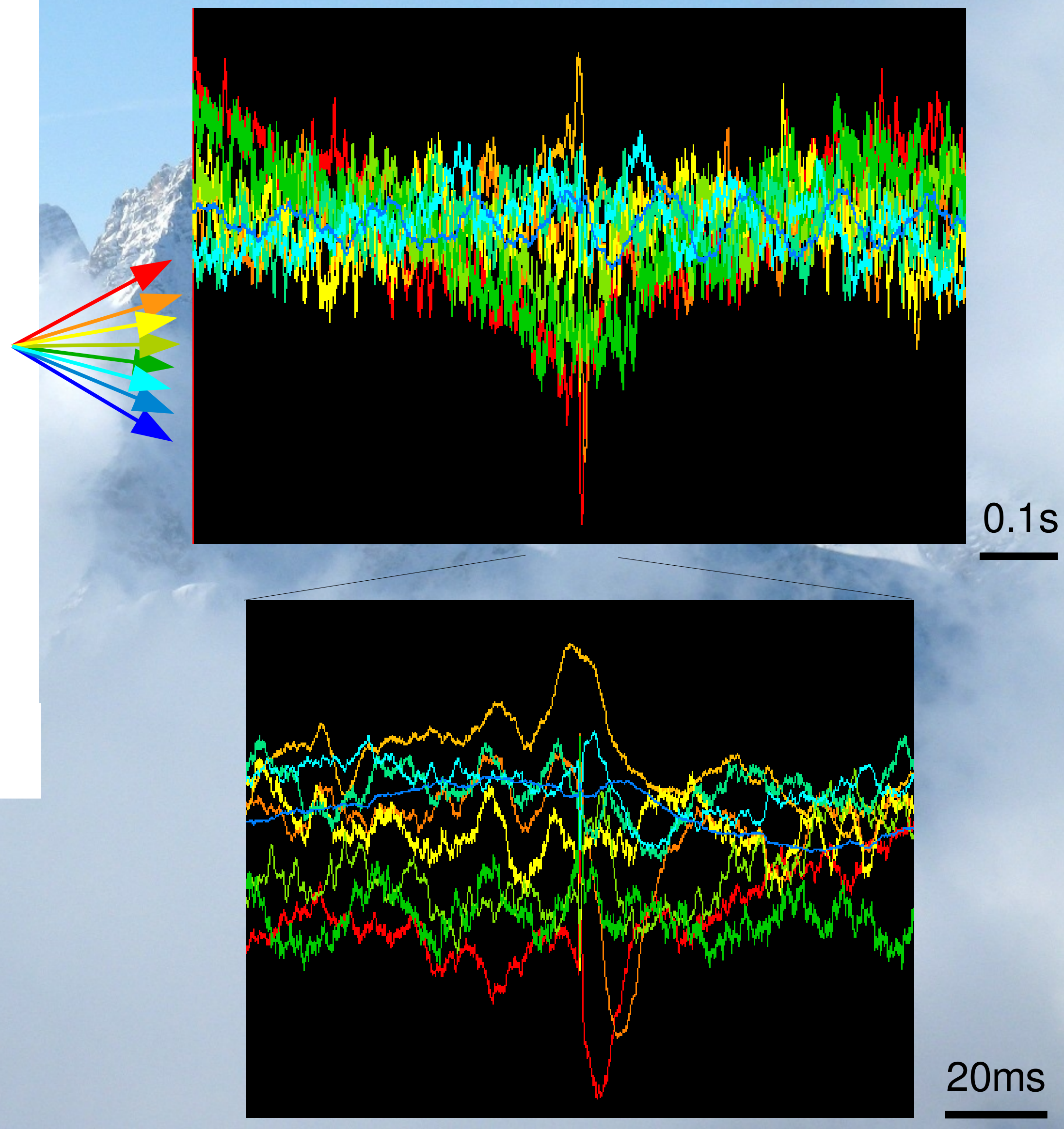


Micro-electro imaging

Inputs of a neurons from different pathways: ICA

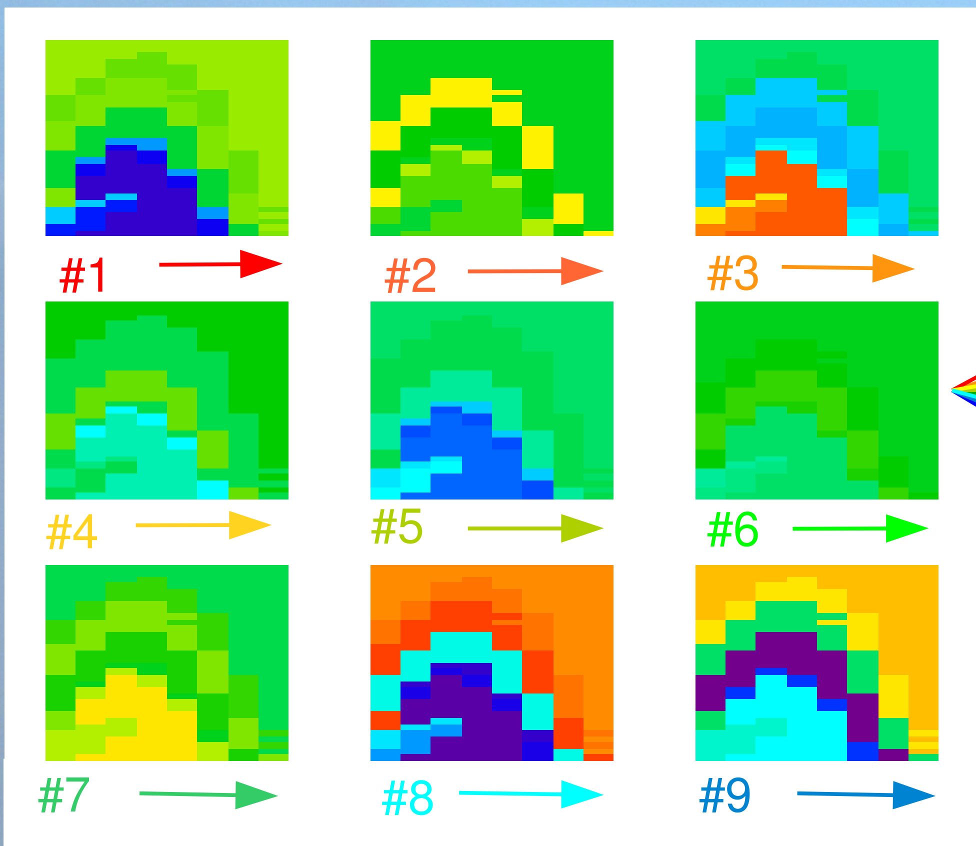




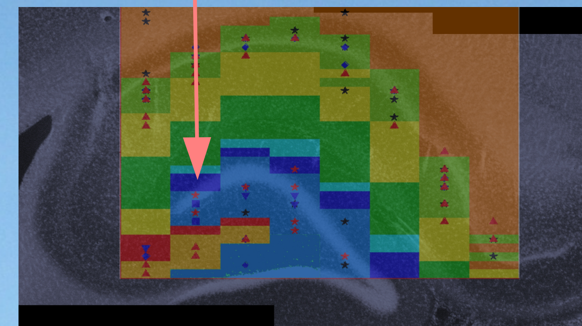


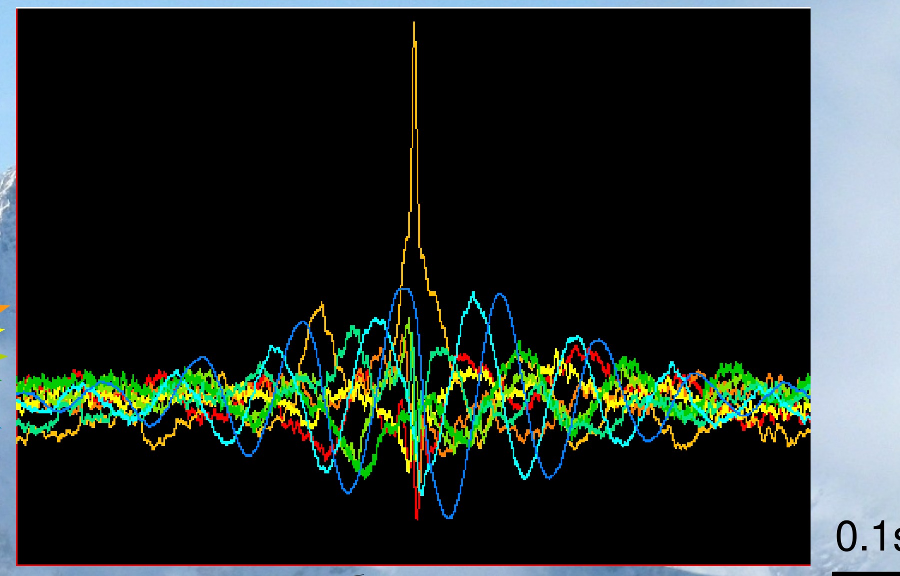
Micro-electro imaging

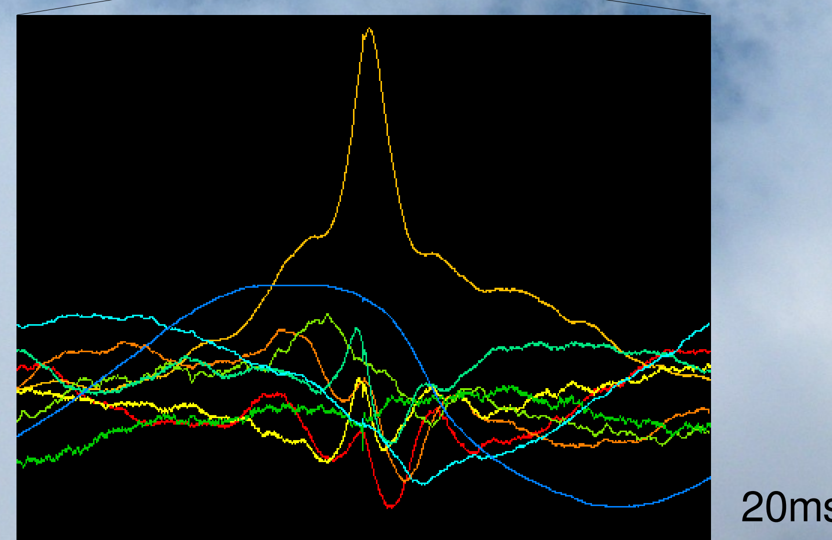
Inputs of a neurons from different pathways: ICA



A DG neuron (#36)

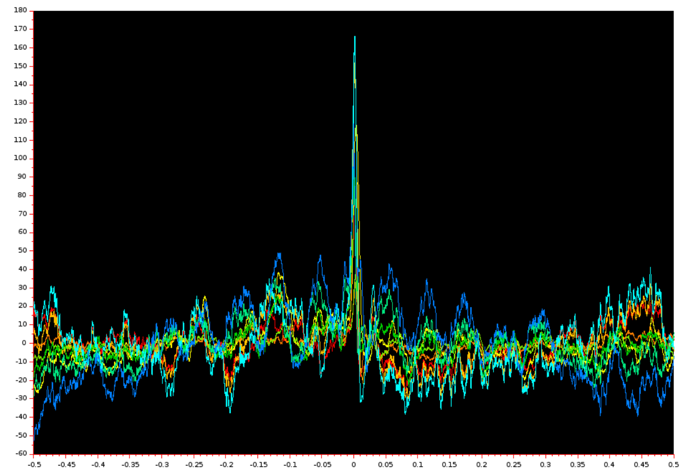




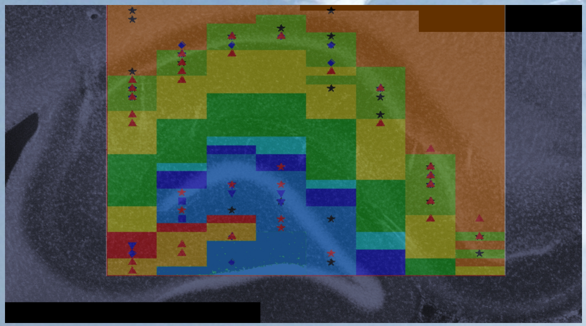


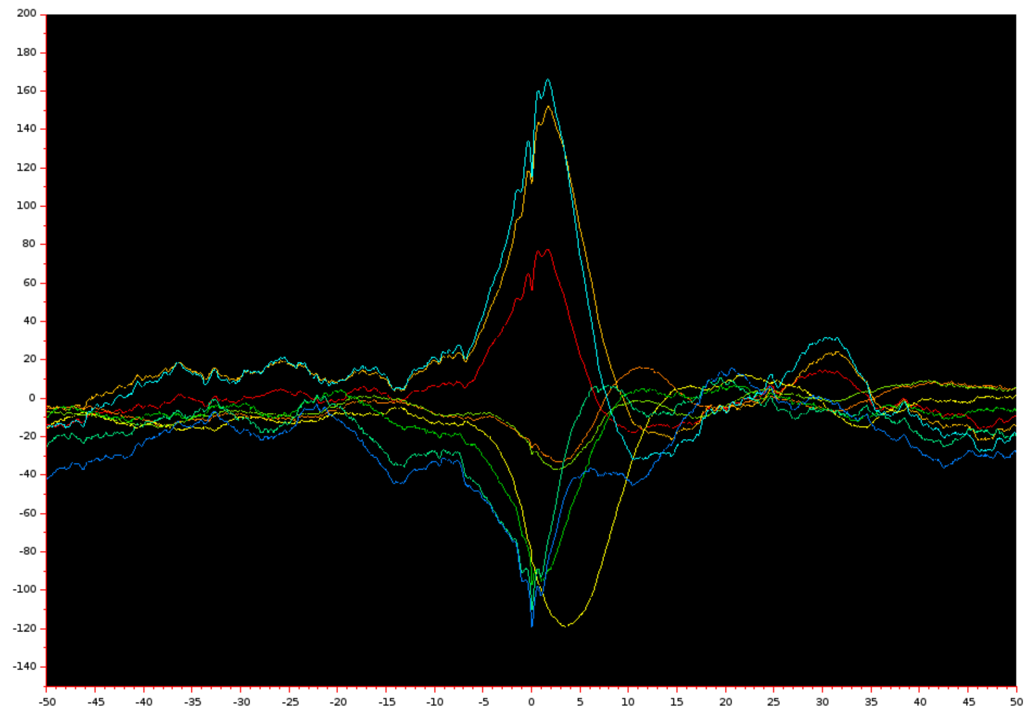
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



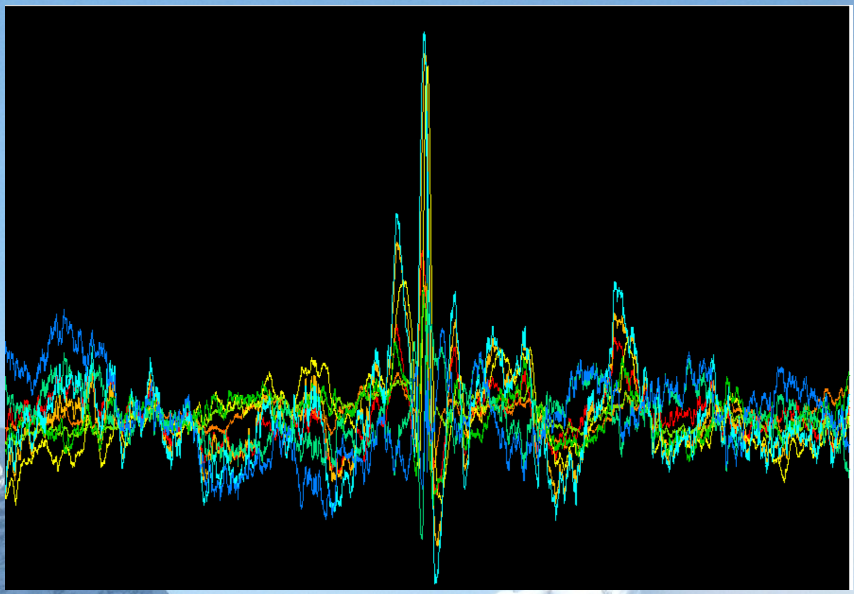
DG granular neurons (n=8)



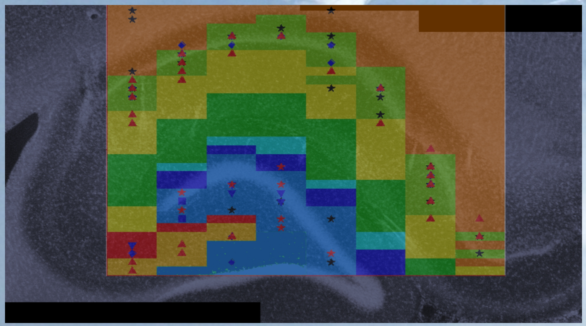


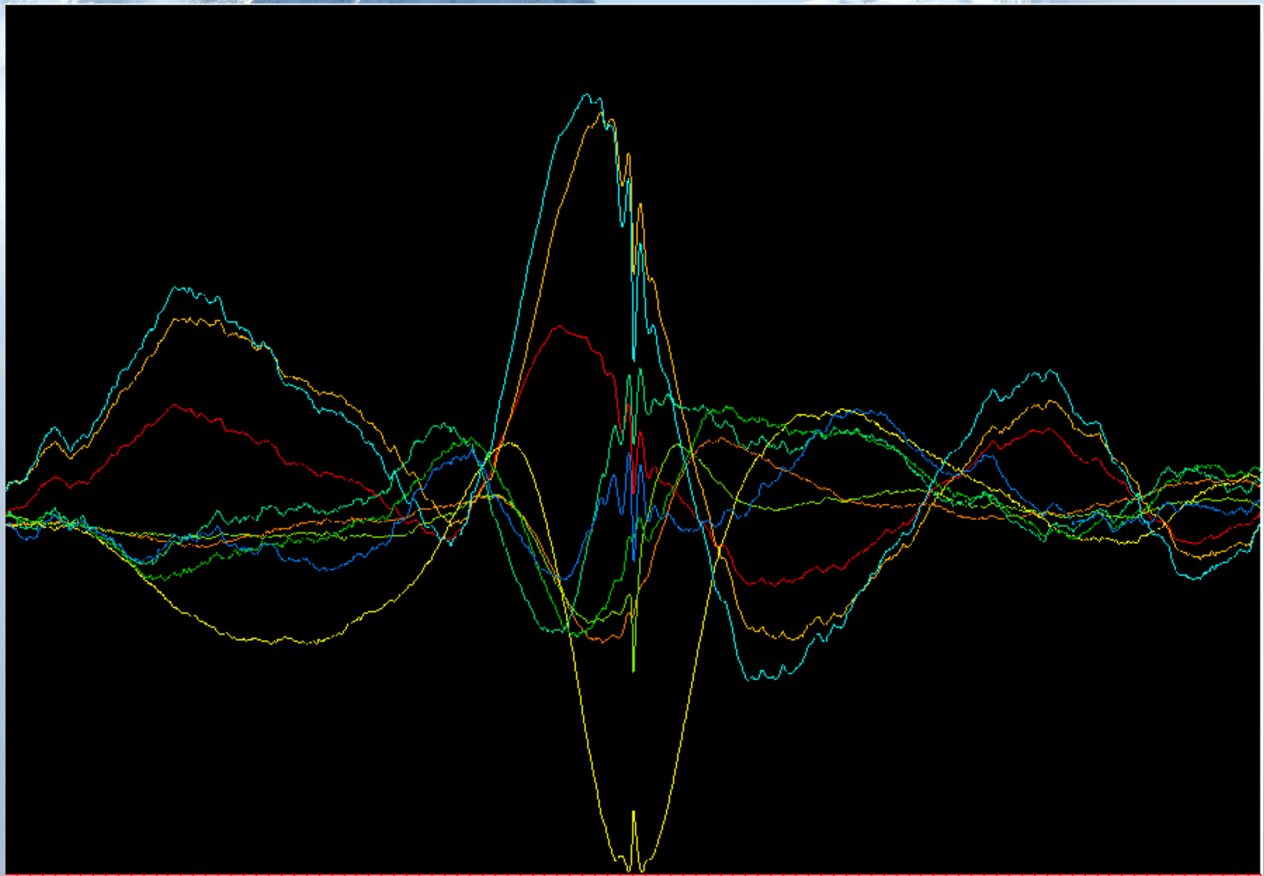
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



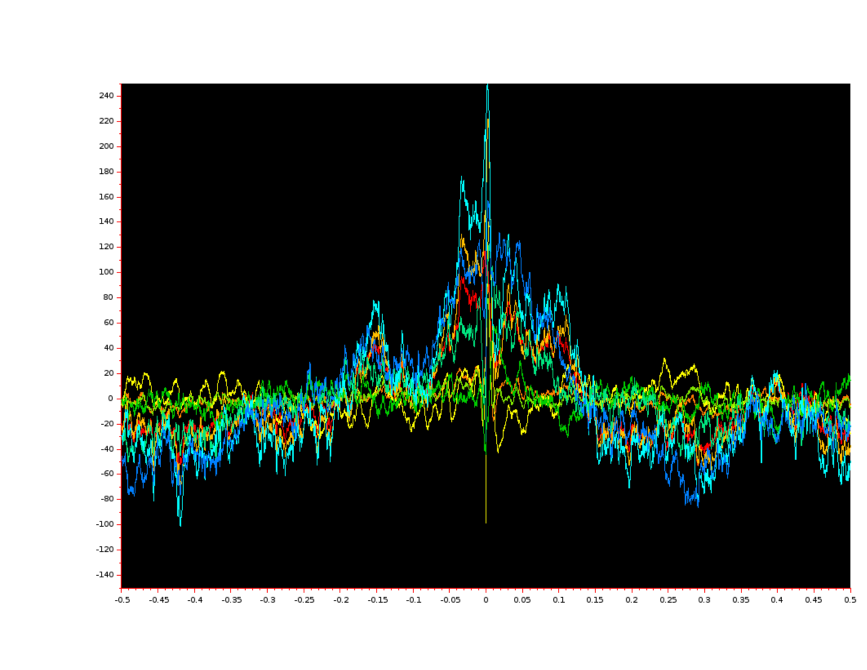
CA1 pyramidal neurons (n=29)



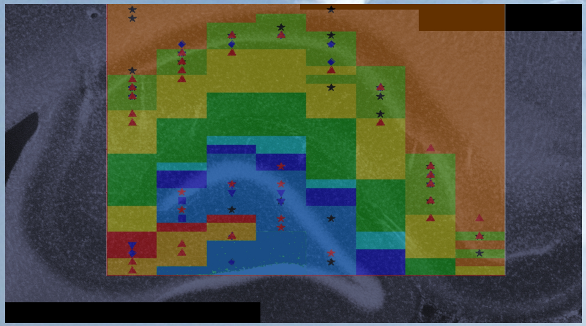


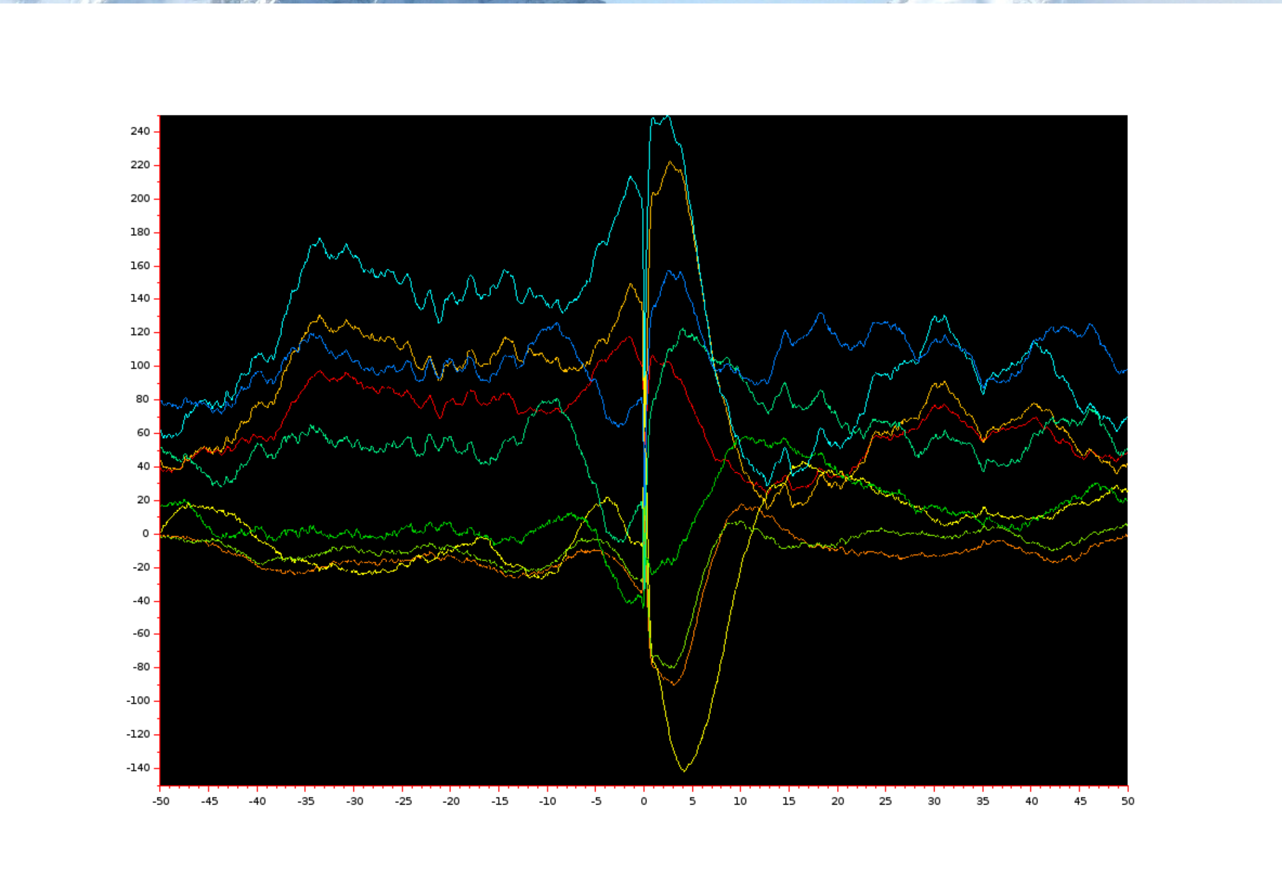
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



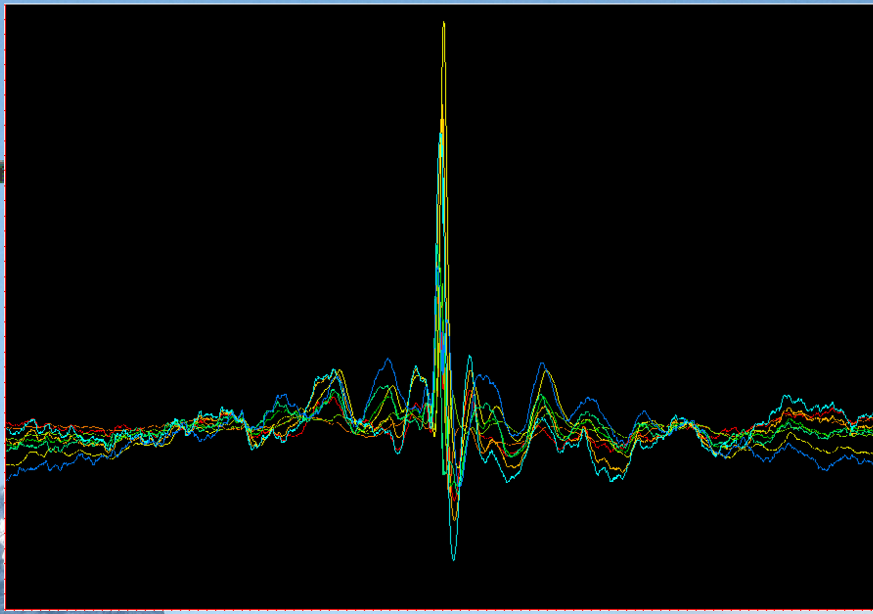
CA3 pyramidal neurons (n=8)



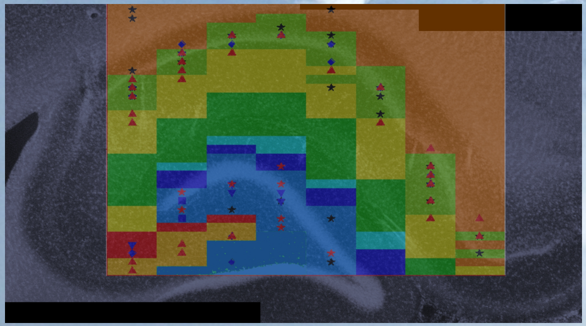


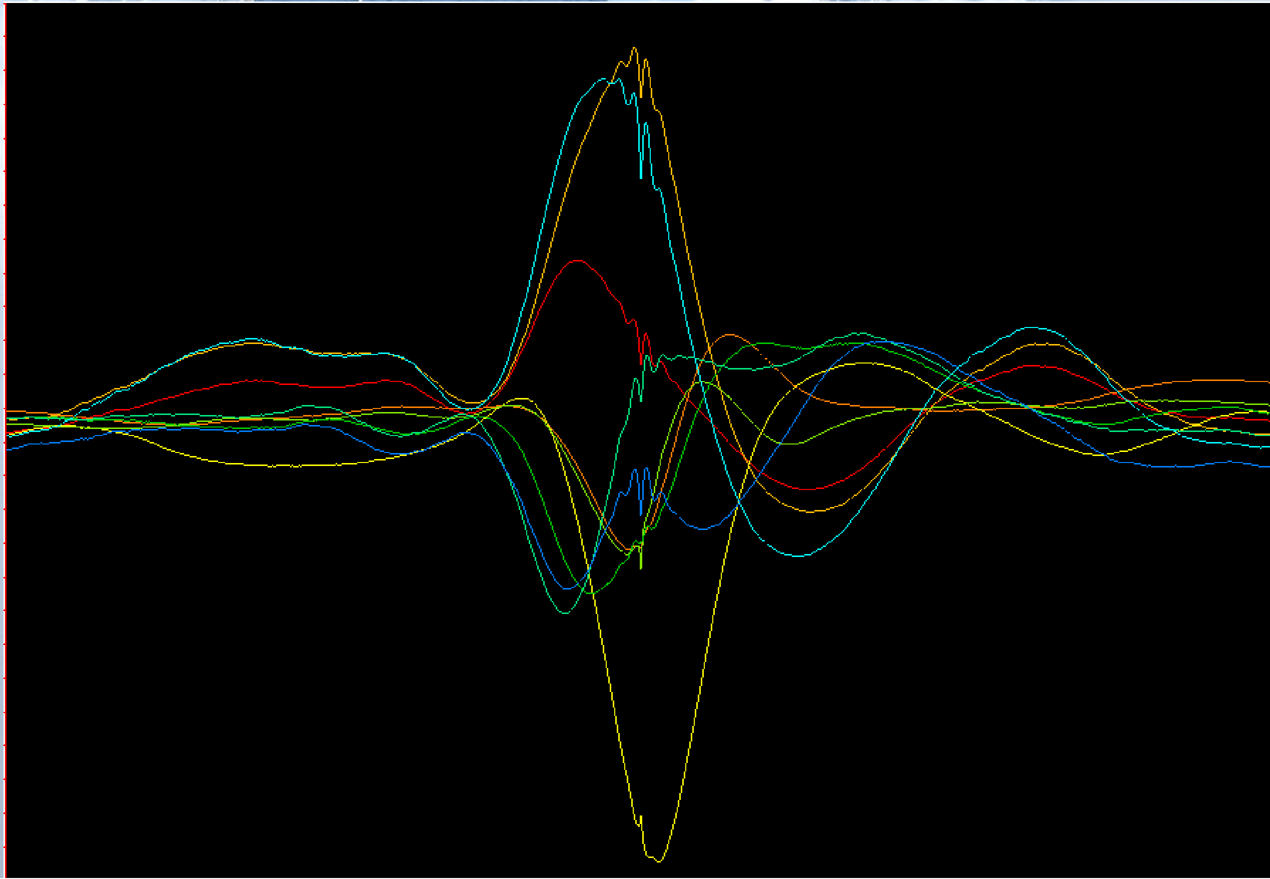
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



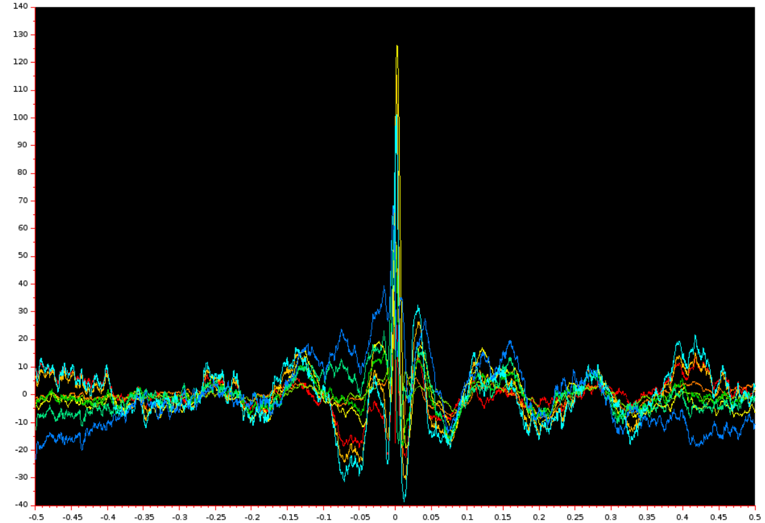
CA1 PV neurons (n=16)



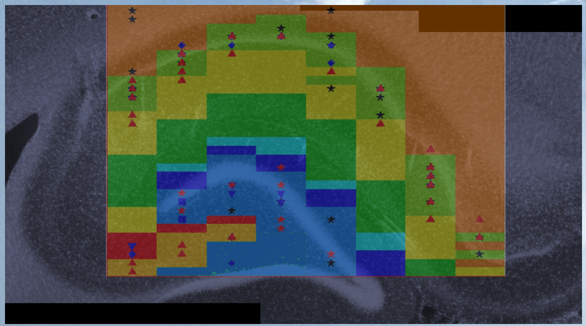


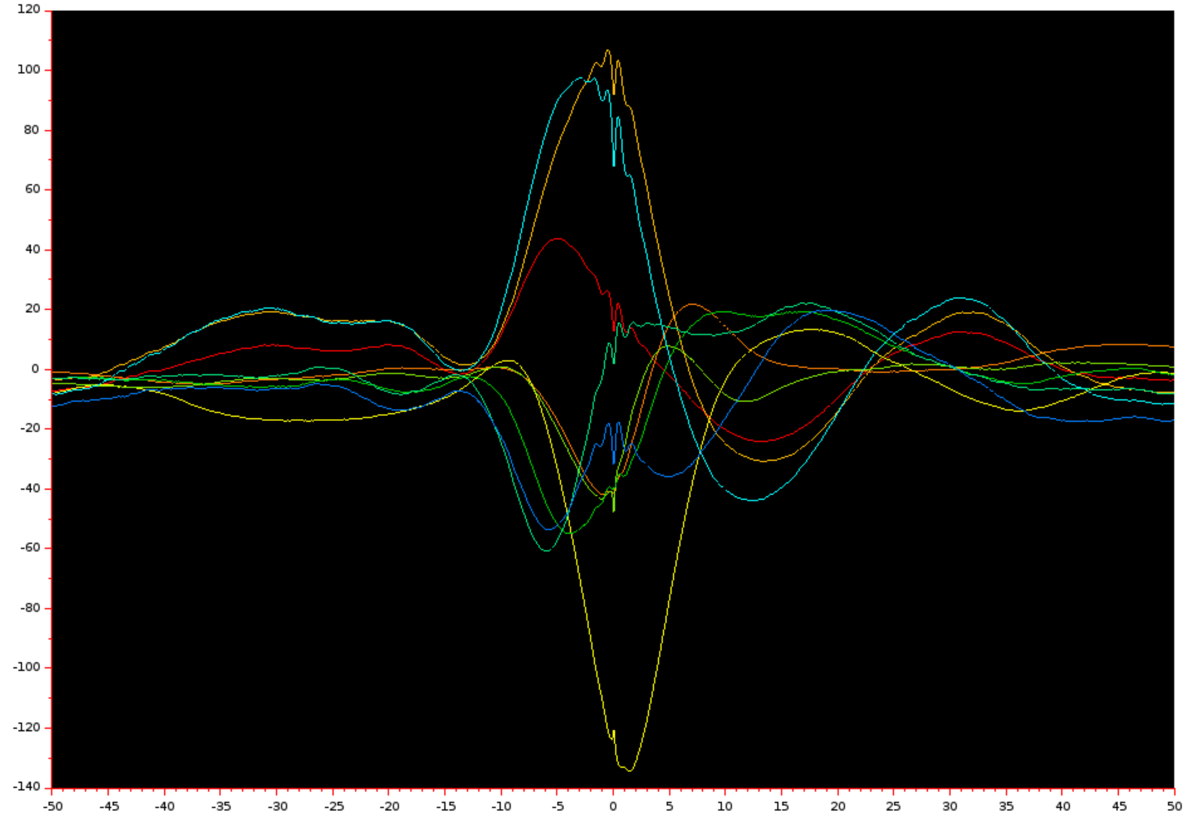
Micro-electro imaging

Cell type specific potentials
Reconstructed without theta



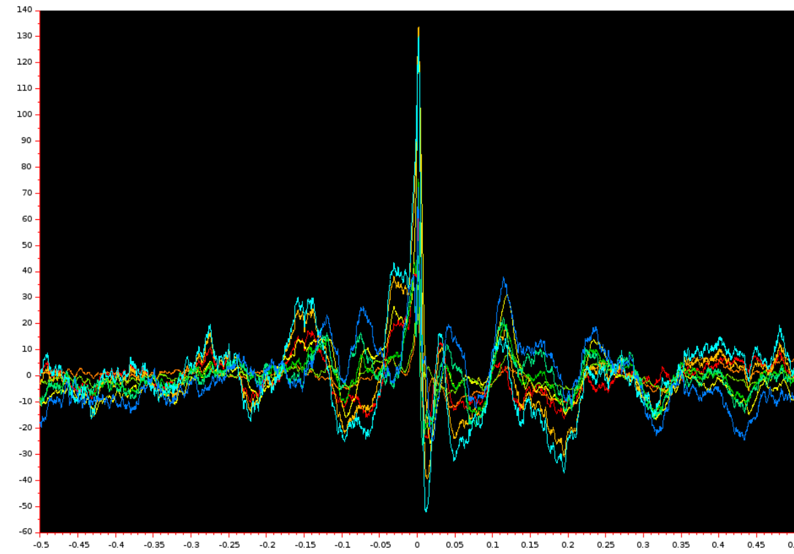
CA3 PV neurons (n=2)



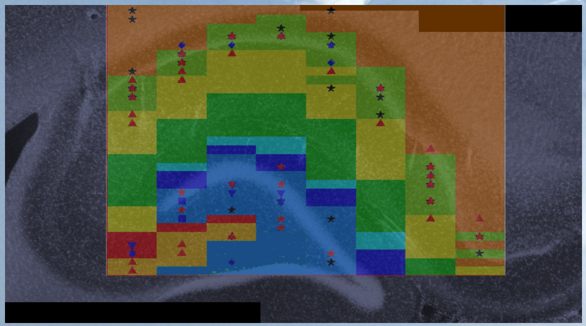


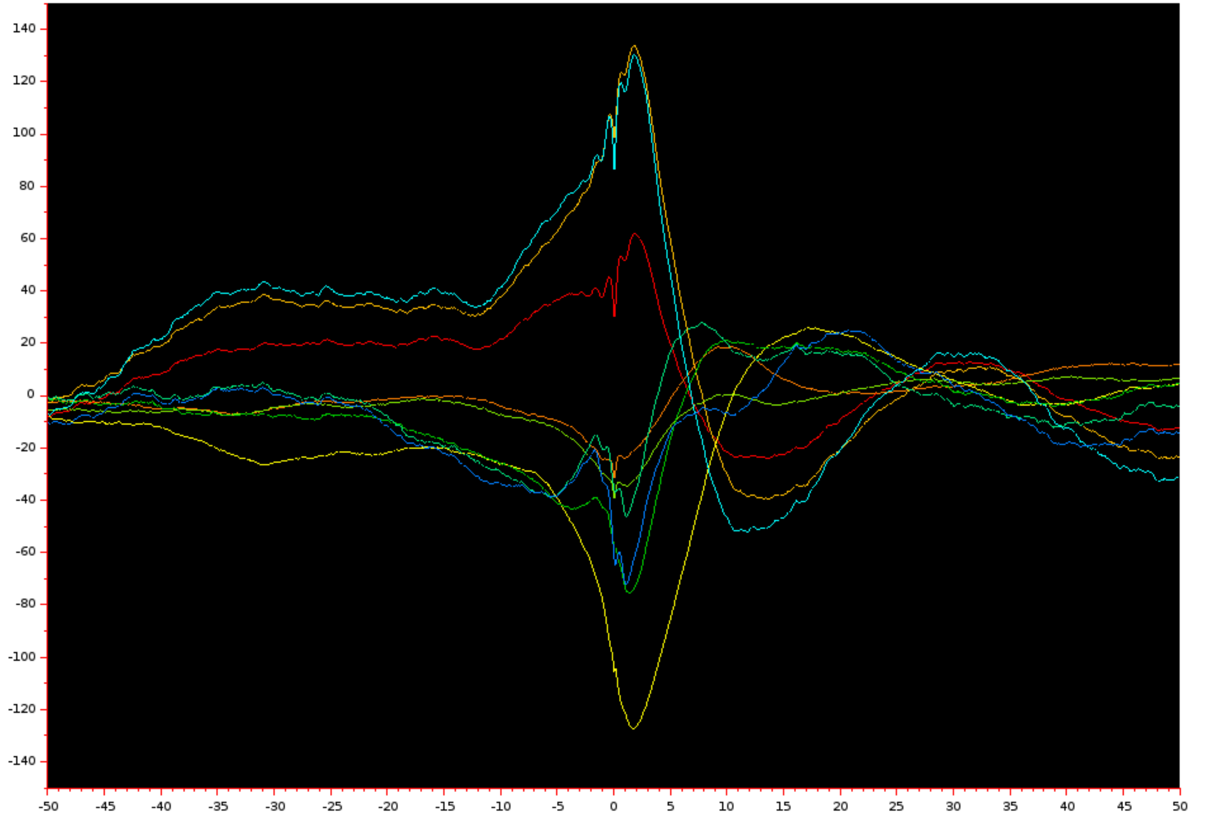
Micro-electro imaging

Cell type specific potentials
Reconstructed without theta



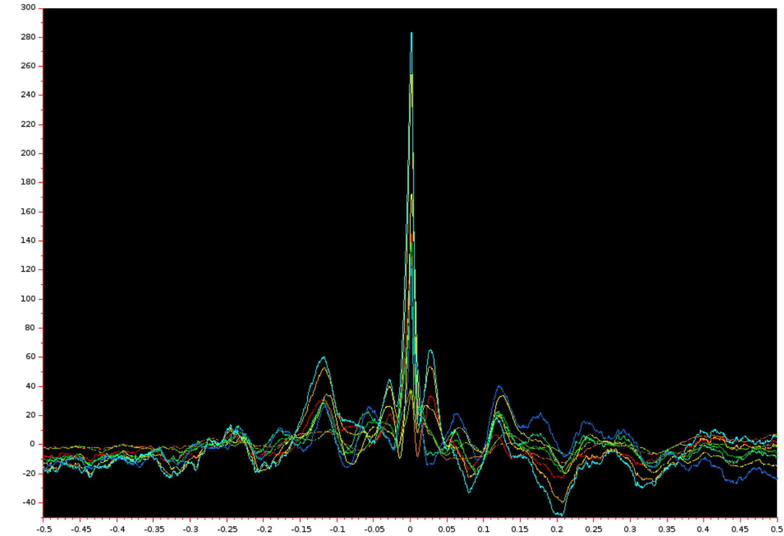
DG (CA3?) PV neurons (n=2)



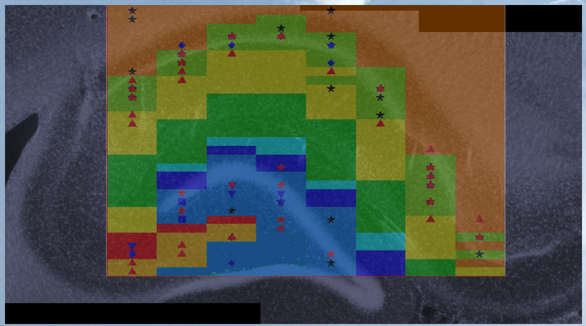


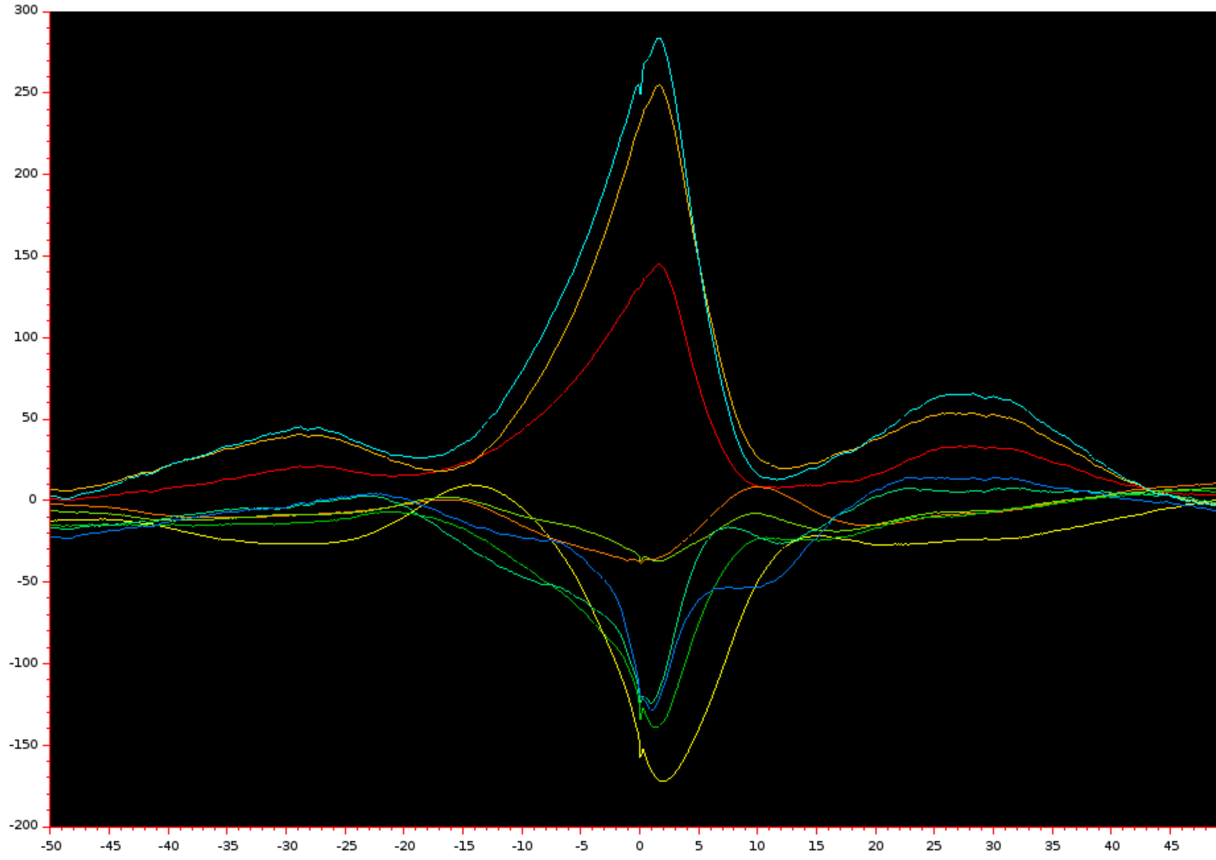
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



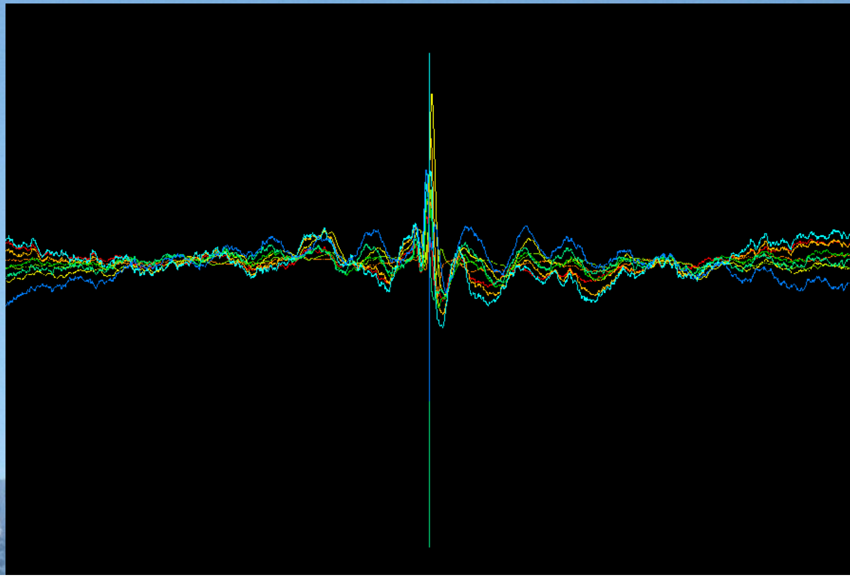
DG AxoAx neurons (n=4)



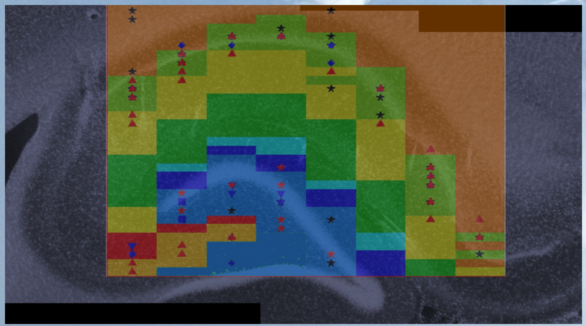


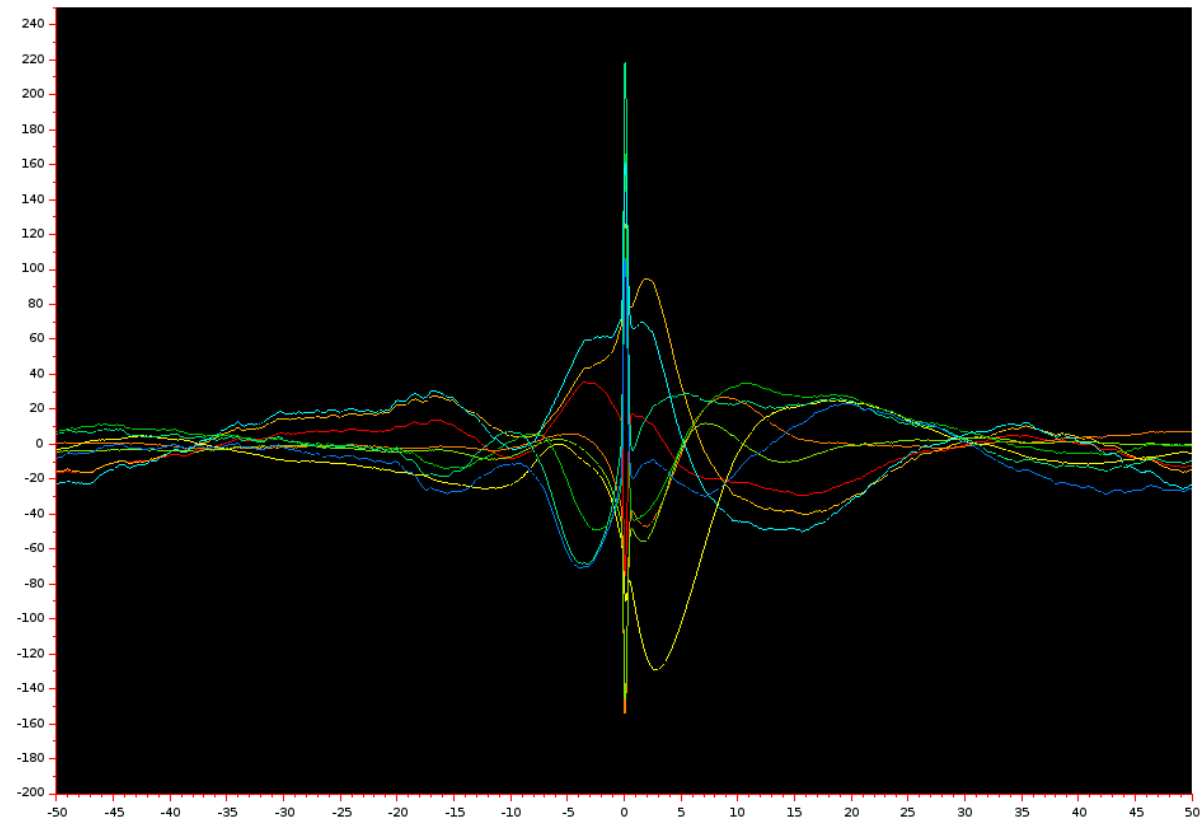
Micro-electro imaging

Cell type specific potentials Reconstructed without theta



CA3 AxoAx neurons (n=1)



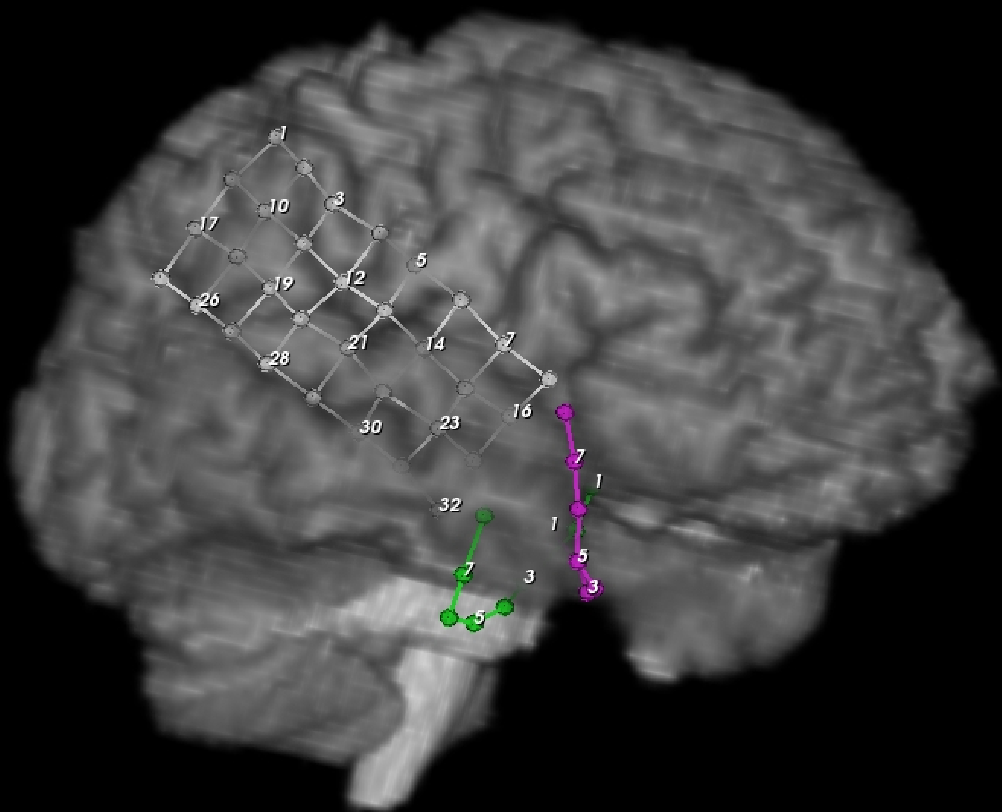


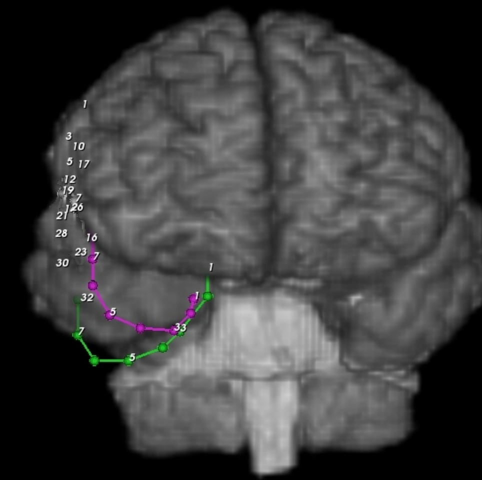
Information theoretical measures

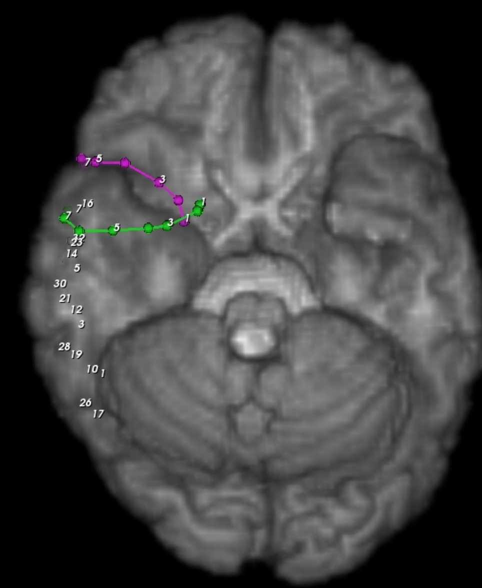
Entropy:

$$H(X) = -\sum_x p_x \log(p_x)$$

MRI with implanted subdural grid electrodes



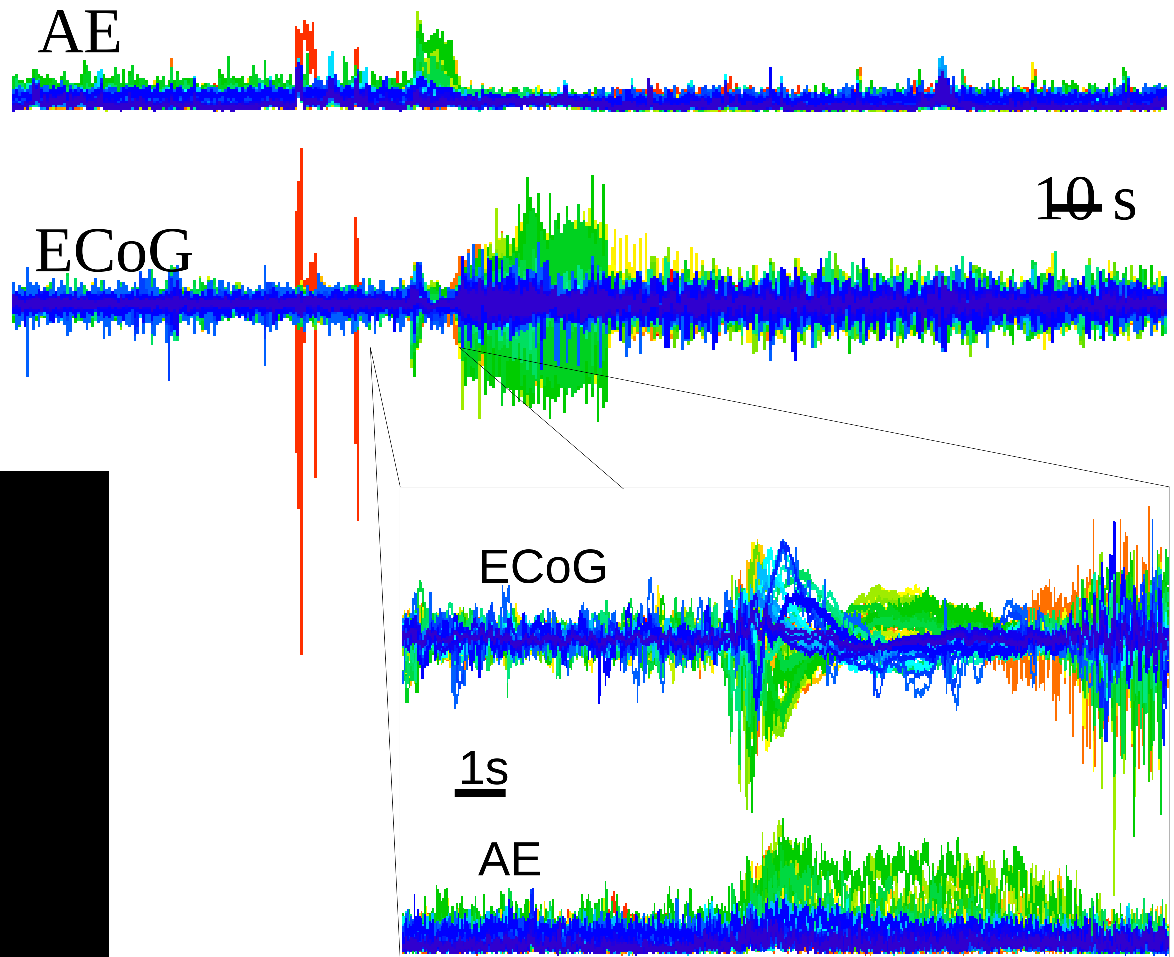


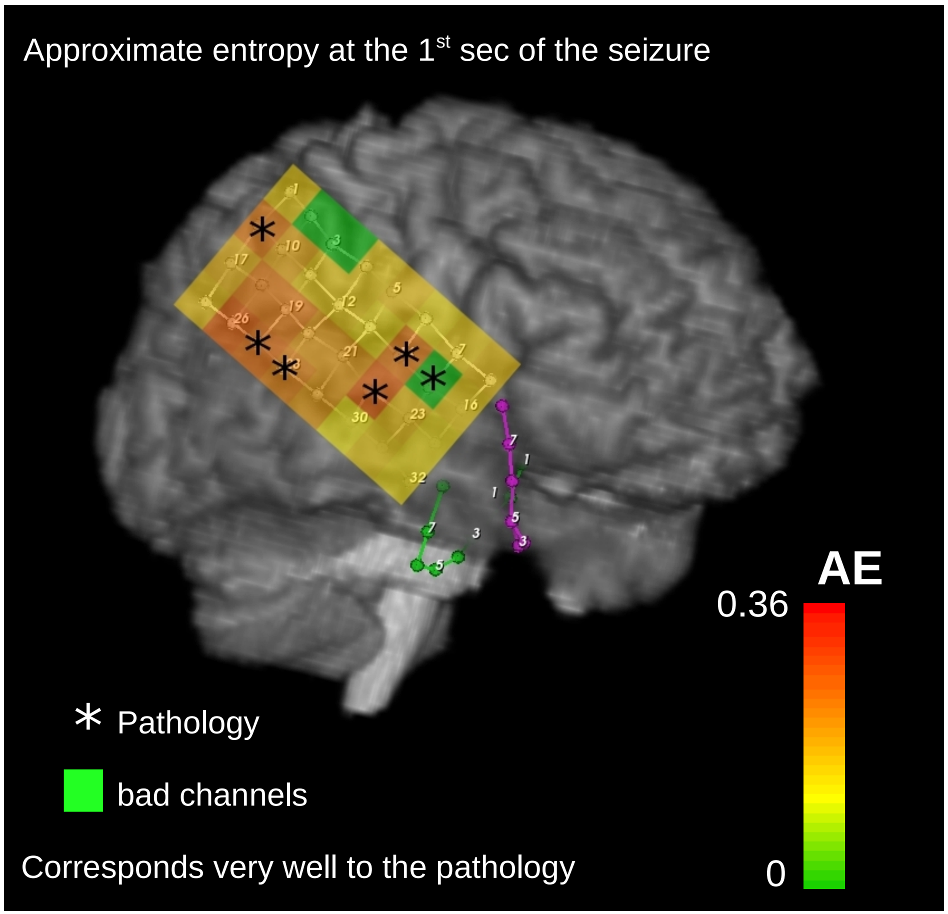


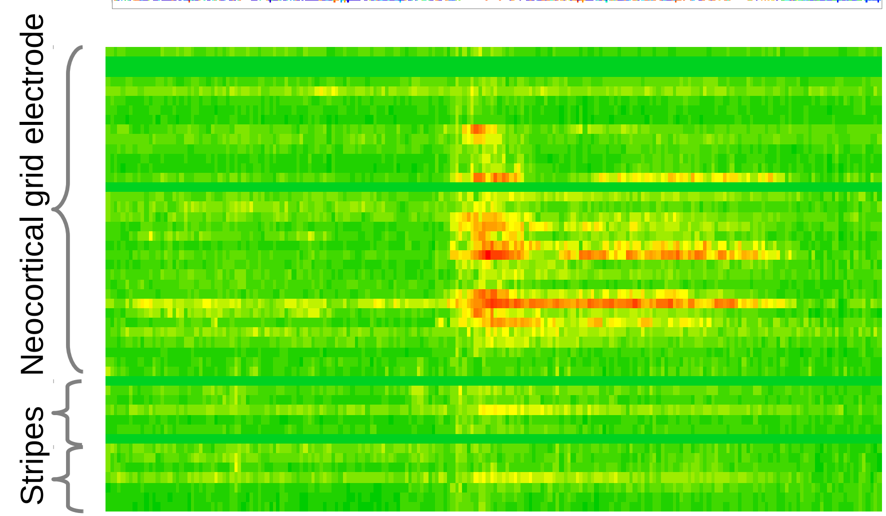
4*8 channels in the grid plus 2*8 channels
In two strip electrodes, 1024 Hz sampling

Entropy of the ECoG during seizure initialization

The Approximate Entropy (AE) is significantly increased solely during the initial, low amplitude phase of the seizure, then AE is decreased below the baseline during the high amplitude phase of the seizure. The positions of the increased AE values during the first sec of the seizure corresponds very well to the seizure onset zone,







Publisher on two conference posters: Hungarian Neuroscience Meeting 2015 and the Hungarian Neurosurgery Conference 2014

Information theoretical measures

Mutual information

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$H(X) = -\sum_x p_x \log(p_x)$$

Causality measures

Granger-causality

$$X(t) = \sum_{i}^{p} a_1(j)\, X(t-j) + \epsilon_1(t)$$

$$Y(t) = \sum_{i}^{p} d_1(j)\, X(t-j) + \eta_1(t)$$

$$X(t) = \sum_{j}^{p} a_2(j)\, X(t-j) + \sum_{j}^{p} b_2(j)\, Y(t-j) + \epsilon_2(t)$$

$$Y(t) = \sum_{j}^{p} c_2(j)\, X(t-j) + \sum_{j}^{p} d_2(j)\, Y(t-j) + \eta_2(t)$$

Causality measures

Granger-causality

$$\Sigma_1 = Var(\epsilon_1(t)) \qquad \Sigma_2 = Var(\epsilon_2(t))$$

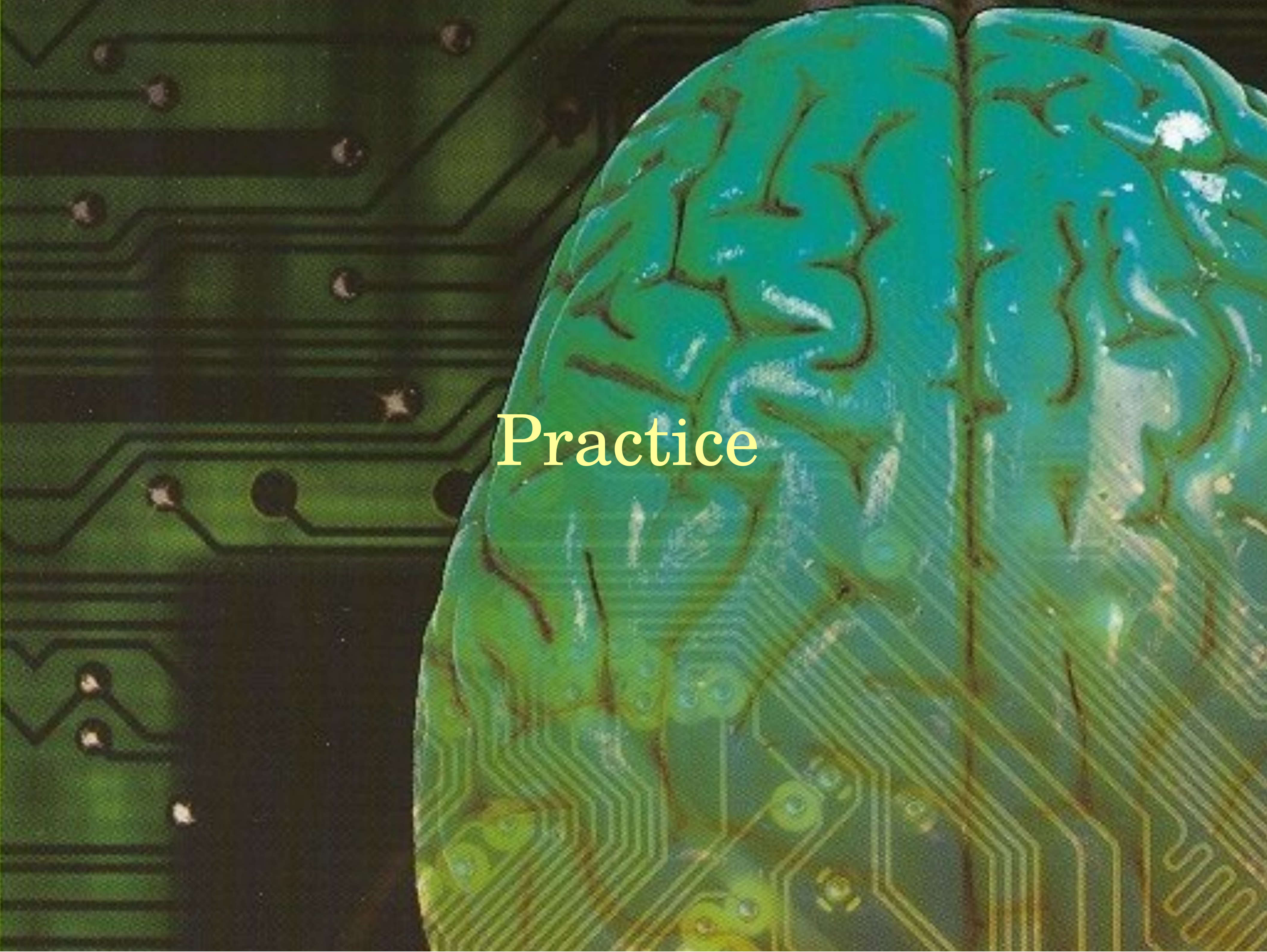
$$\Gamma_1 = Var(\eta_1(t)) \qquad \Gamma_2 = Var(\eta_2(t))$$

$$F_{X \to Y} = \log(\Sigma_1) - \log(\Sigma_2)$$

$$F_{Y \to X} = \log(\Gamma_1) - \log(\Gamma_2)$$

$$F_{YX} = \log(\Sigma_2 \Gamma_2) - \log(\Sigma_2^2 - cov^2(\epsilon_2(t)\eta_2(t)))$$

Practice



Practice

```
stacksize(2e8)
getd ~/TANIT/SummerSchool15/PRACTICE
loadmatfile('~/TANIT/SummerSchool15/PRACTICE/Se
izure1.mat');
st=1e3;
chn=43;
cm1=CorrFor(adat,1,5e3);
stn=floor(size(adat,1)/st);
scm=zeros(stn,chn);
cmm=zeros(stn*chn,chn);
for k=1:stn
l1=(k-1)*st+1;
l2=k*st;
[cm]=CorrFor(adat,l1,l2);
cmm((k-1)*chn+1:k*chn,:)=cm;
scm(k,:)=mean(cm,'r');
end

socol(24);
tplot(scm);
```