Possible Projects

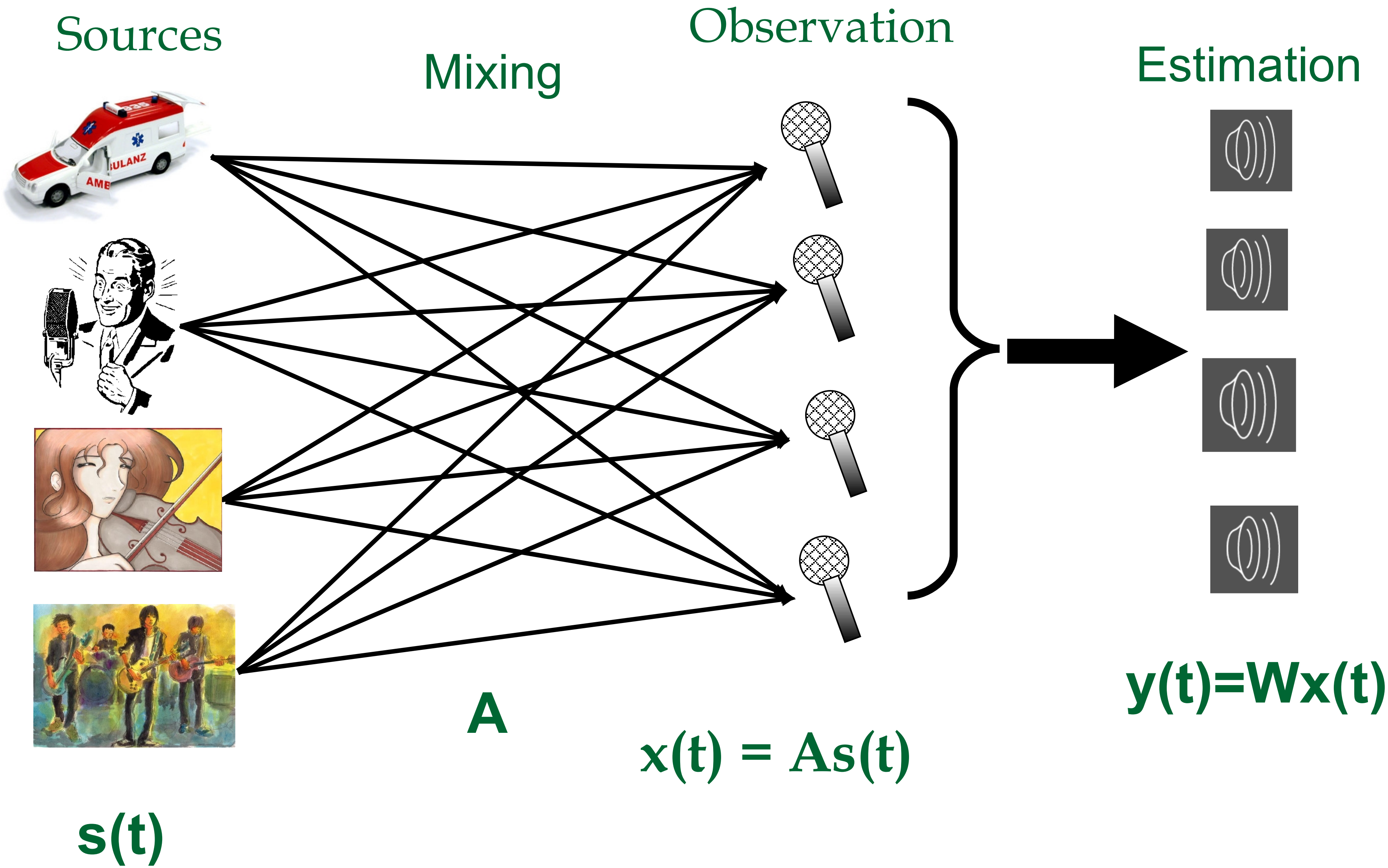
Barnabás Póczos

UofA, Edmonton

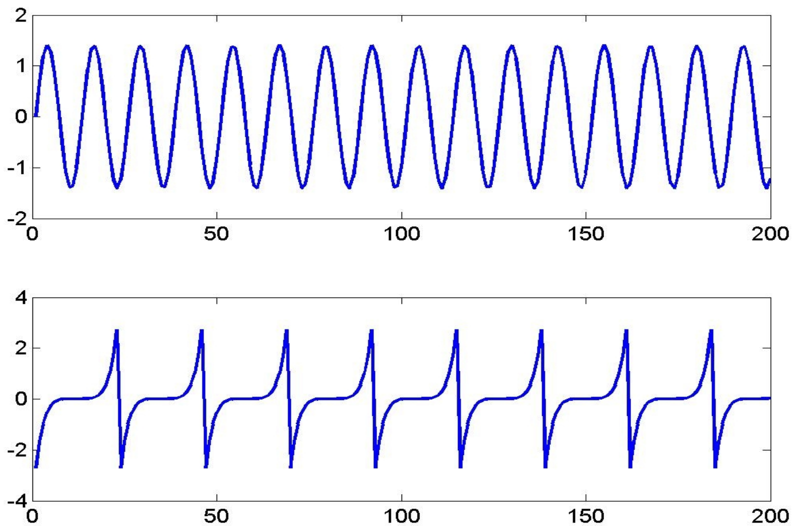
Sept 21, 2009

Blind Subspace Deconvolution

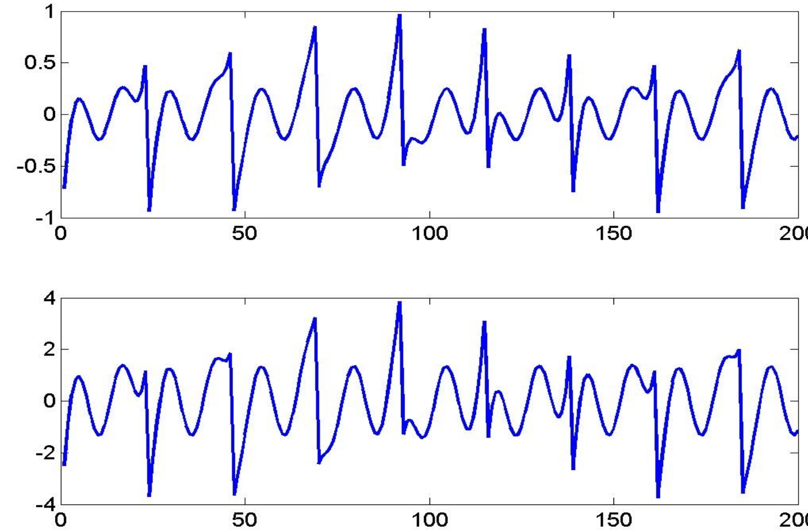
Independent Component Analysis



Independent Component Analysis



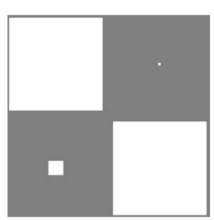
Independent signals

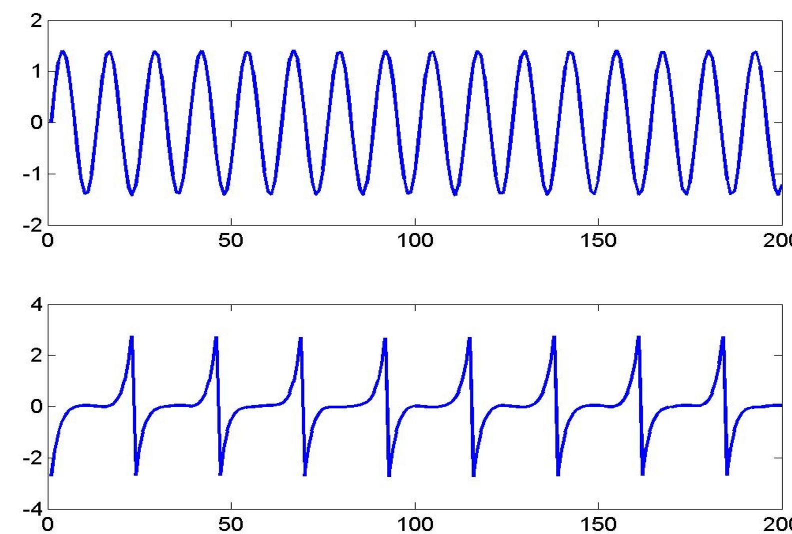


Mixtures

$$
\begin{aligned}
x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) \\
x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t)
\end{aligned}
$$

WA=



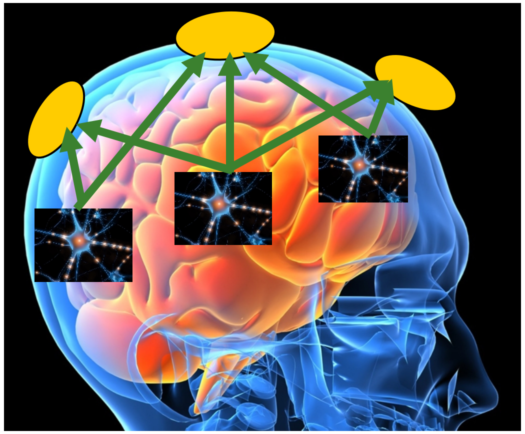


ICA estimation

Some ICA Applications

- Medical signal processing – fMRI, ECG, EEG
- Brain Computer Interfaces
- Image denoising
- Modeling of the hippocampus, place cells
- Modeling of the visual cortex
- Microarray data processing
- Decomposing the spectra of galaxies
- Blind deconvolution
- Feature extraction
- Face recognition
- Time series analysis
- Financial applications
- Clustering
- Classification

ICA Applications, Removing Artifacts from EEG

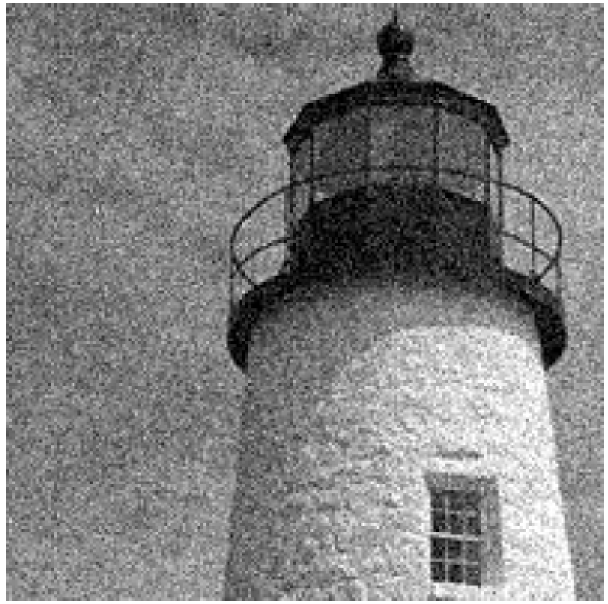


- EEG ~ *Neural cocktail party*
- Severe ***contamination*** of EEG activity by
  - eye movements,
  - blinks,
  - muscle,
  - heart, ECG artifact
  - vessel pulse
  - electrode noise
  - line noise, alternating current (60 Hz)

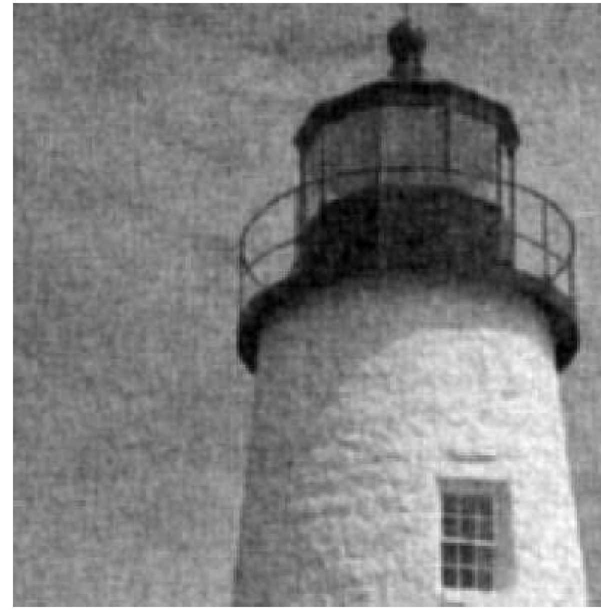
- ICA can effectively ***detect, separate and remove*** activity in EEG records from a wide variety of artifactual sources.
  (Jung, Makeig, Bell, and Sejnowski)

ICA for Image Denoising (Hoyer, Hyvarinen)

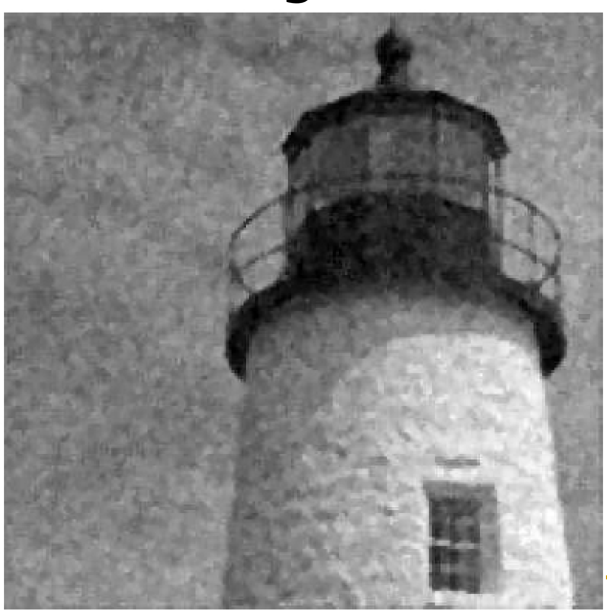
original



noisy



Wiener filtered

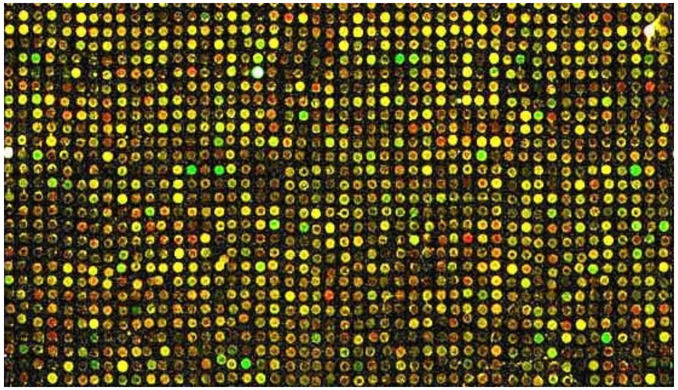


median filtered

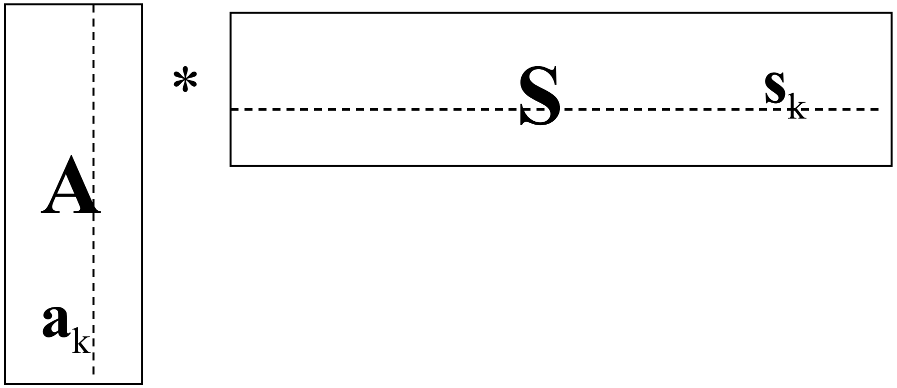
ICA denoised

ICA for Microarray data processing

$\mathbf{X}^T =$



$=$



$\mathbf{X}^T \in R^{M \times N}$

M = number of experiments

N = number of genes

Assumption:

- each experiment is a mixture of **independent expression modes** ($\mathbf{s}_1, \ldots \mathbf{s}_K$).
- some of these modes (e.g. $\mathbf{s}_k$) can be related to the difference between the classes.
- $\rightarrow$ $\mathbf{a}_k$ correlates with the class labels

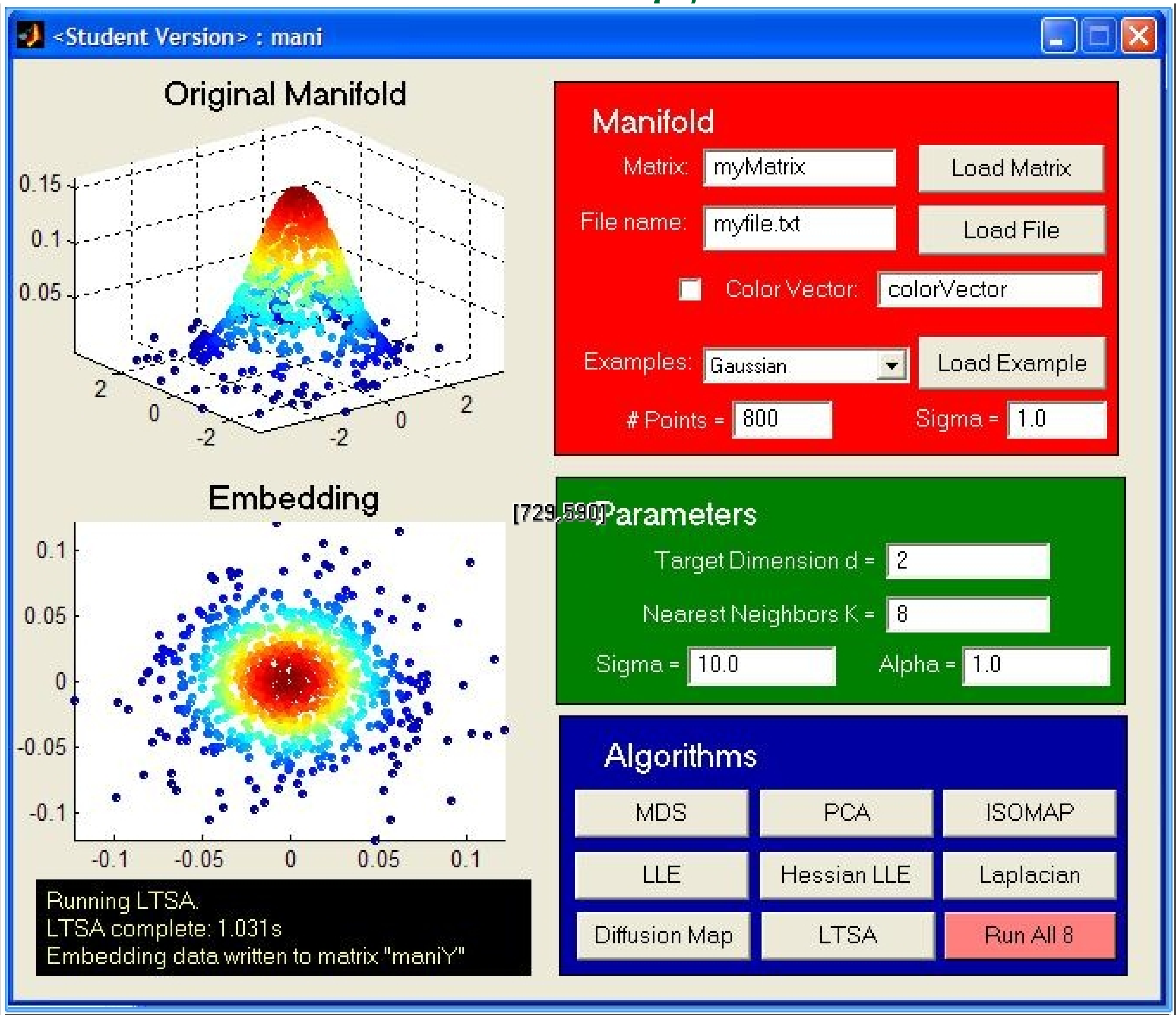
Generalizations

- **Blind Deconvolution**
  - $x(t)=A_1 s(t-1)+A_2 s(t-2)+...+A_k s(t-k)$
  - After FFT we have to do ICA in the frequency space
- **Independent Subspace Analysis**
- ***Blind Subspace Deconvolution using FFT???***

Manifold learning using time processes

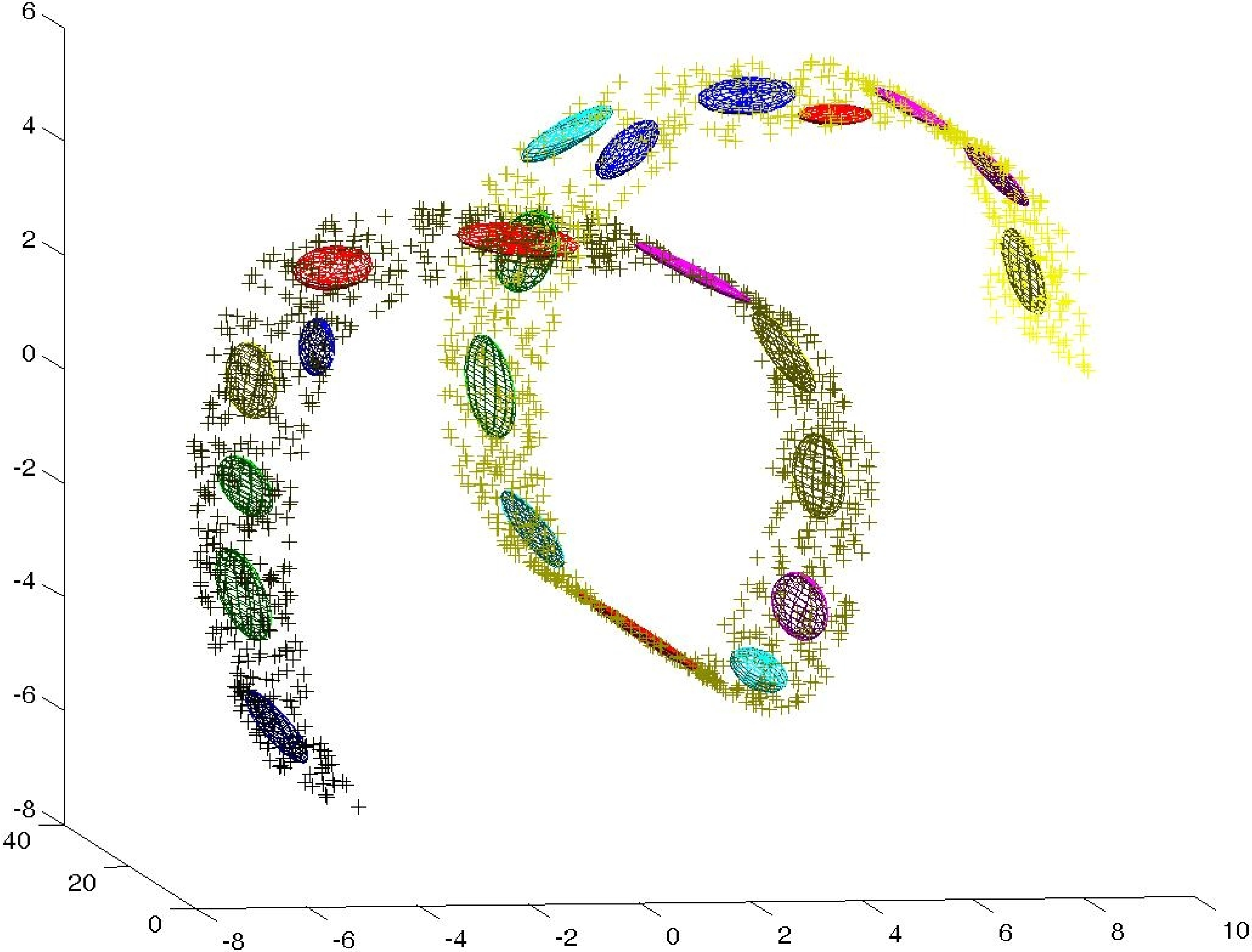
Manifold learning

http://www.math.umn.edu/~wittman/mani/

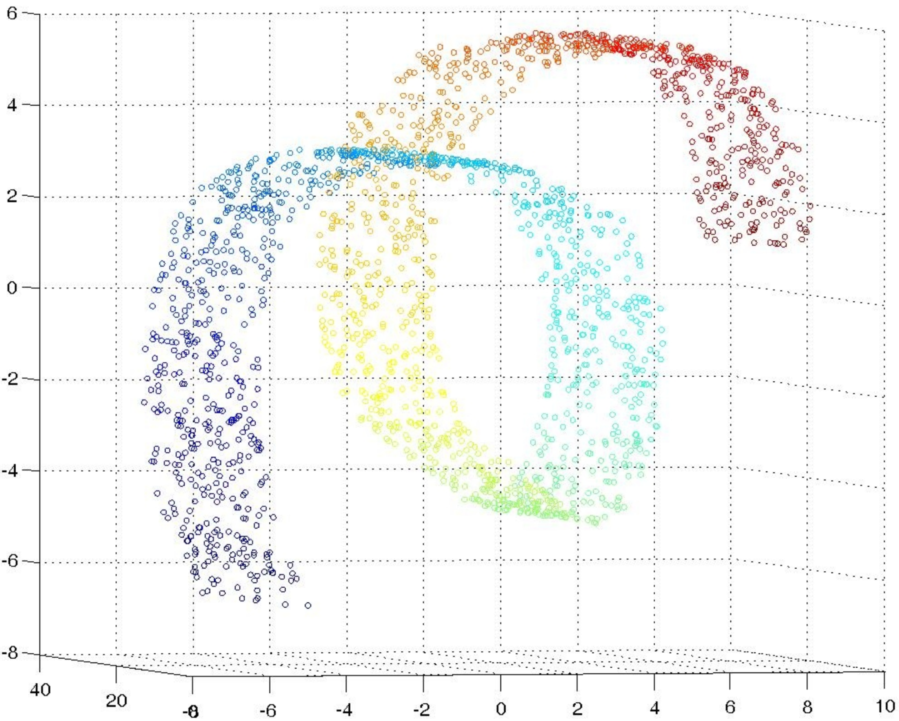


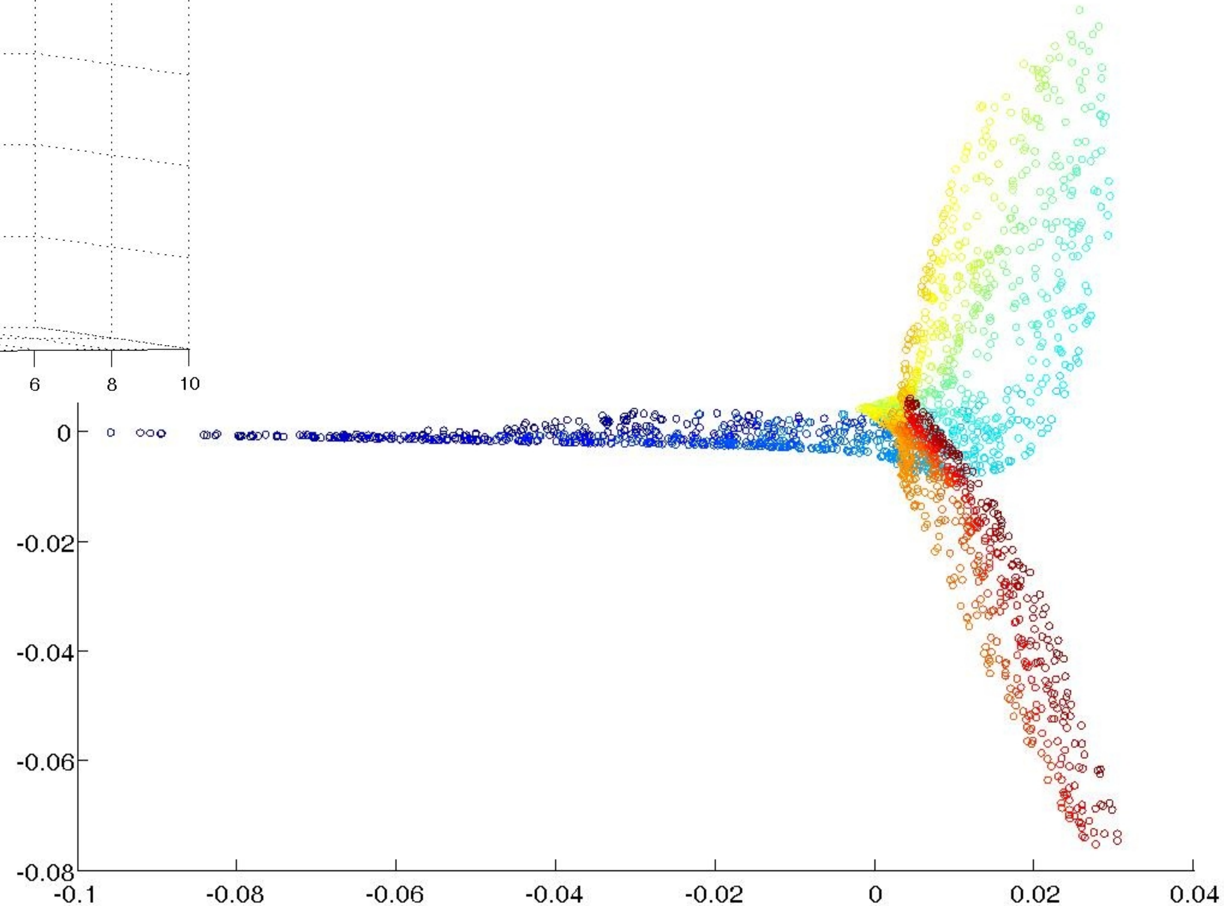
Manifold learning

- The algorithms usually assume iid sample.
- ***Can we exploit time correlations???***
  - E.g. using Hidden Markov Models in the Variational Bayesian Mixture Factor Analyzer model (vbMFA, Beal & Ghahramani)...


6
4
2
0
-2
-4
-6
-8
40
20
0
-8
-6
-4
-2
0
2
4
6
8
10

LTSA Failure





Thanks for the attention!

ICA decomposition

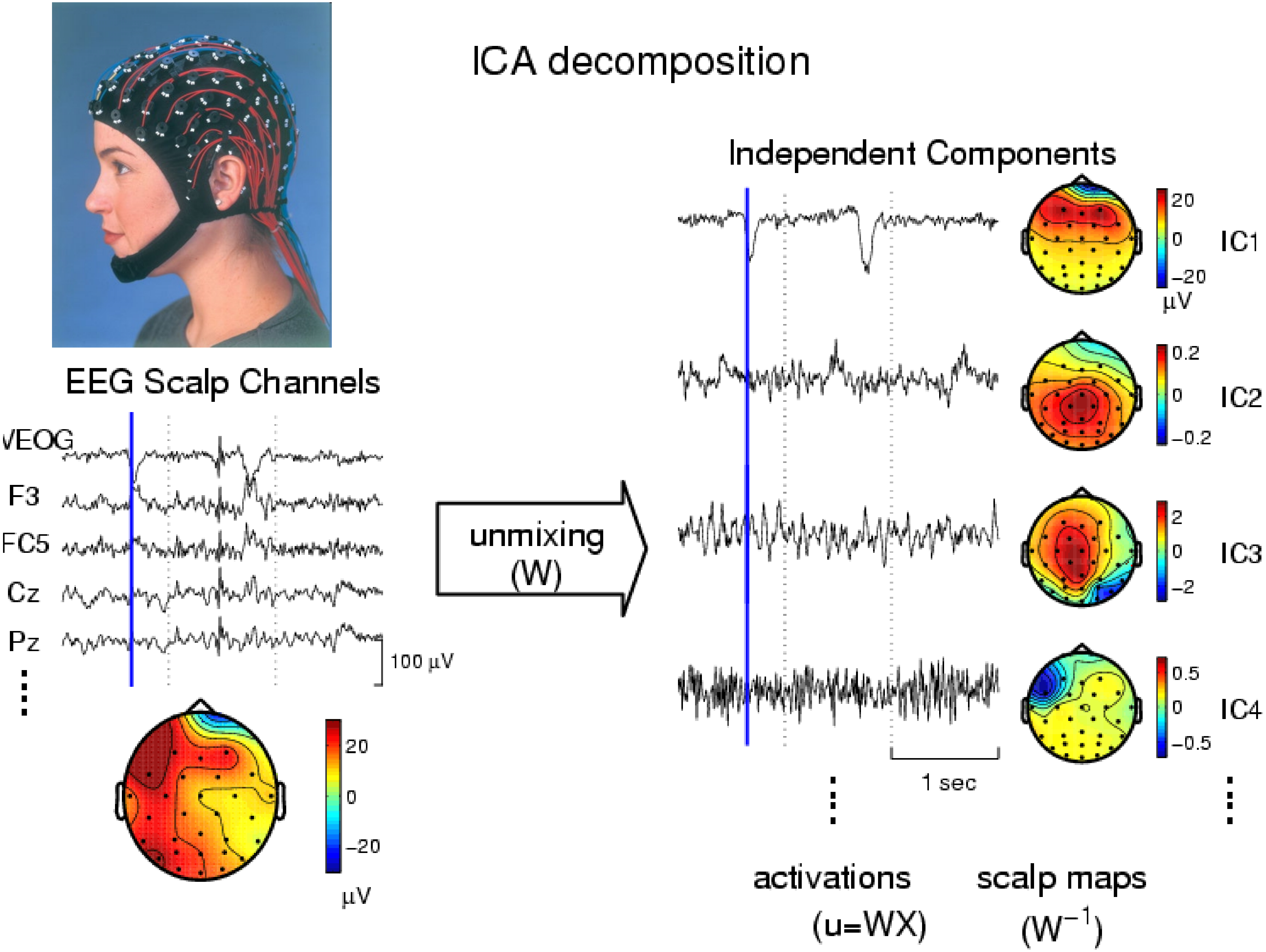


Fig from Jung

Summed Projection of Selected Components

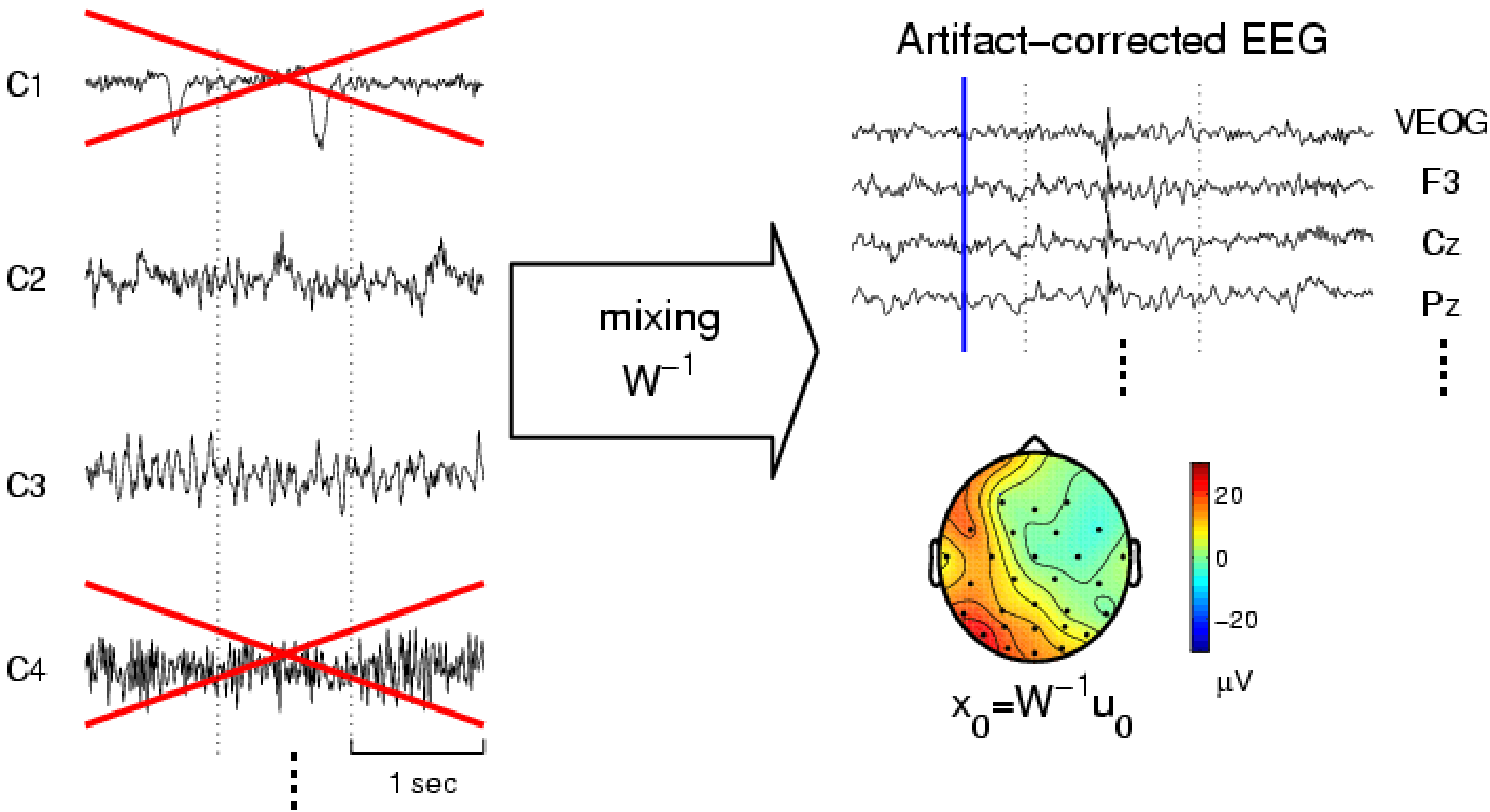


Fig from Jung

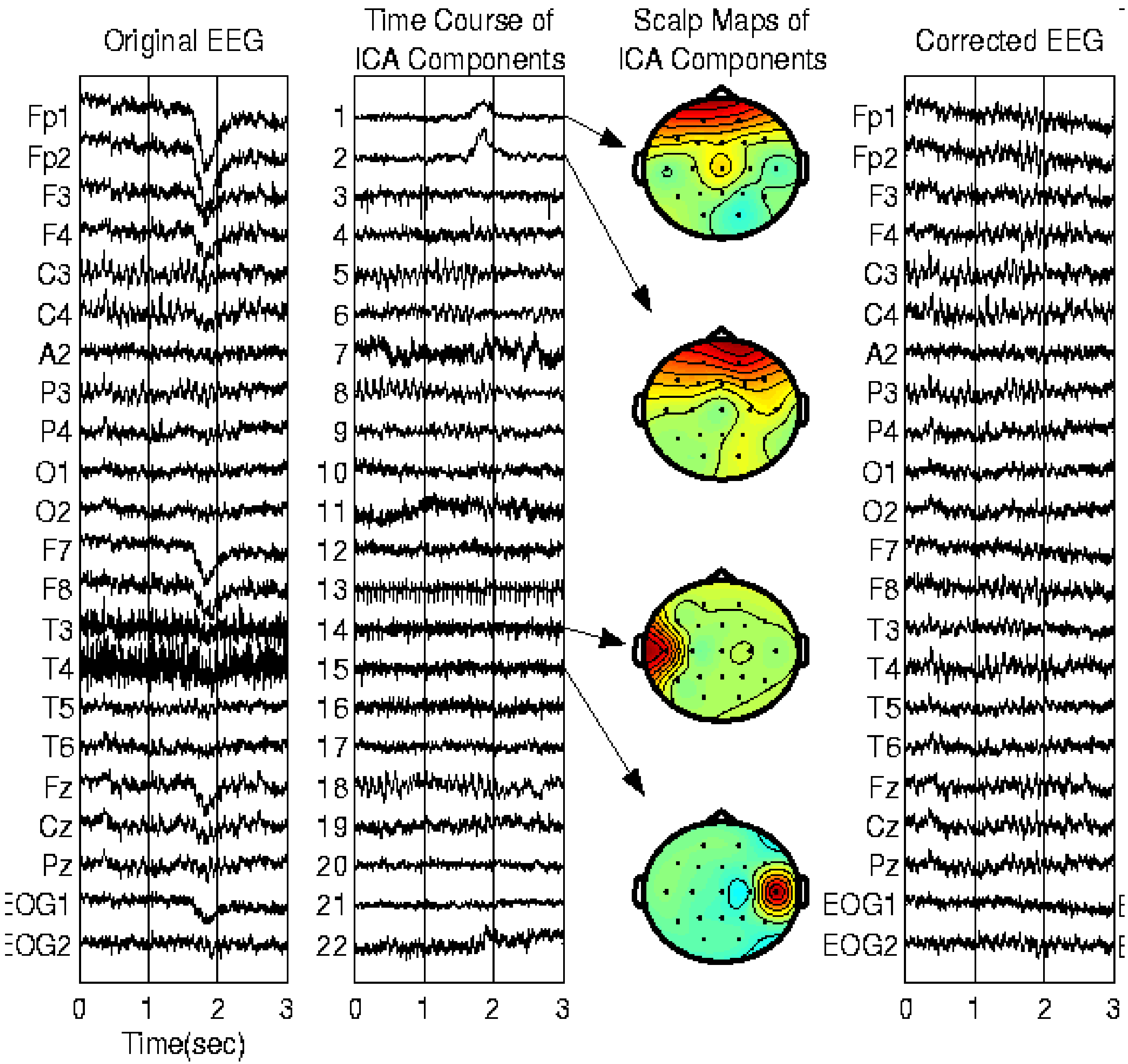


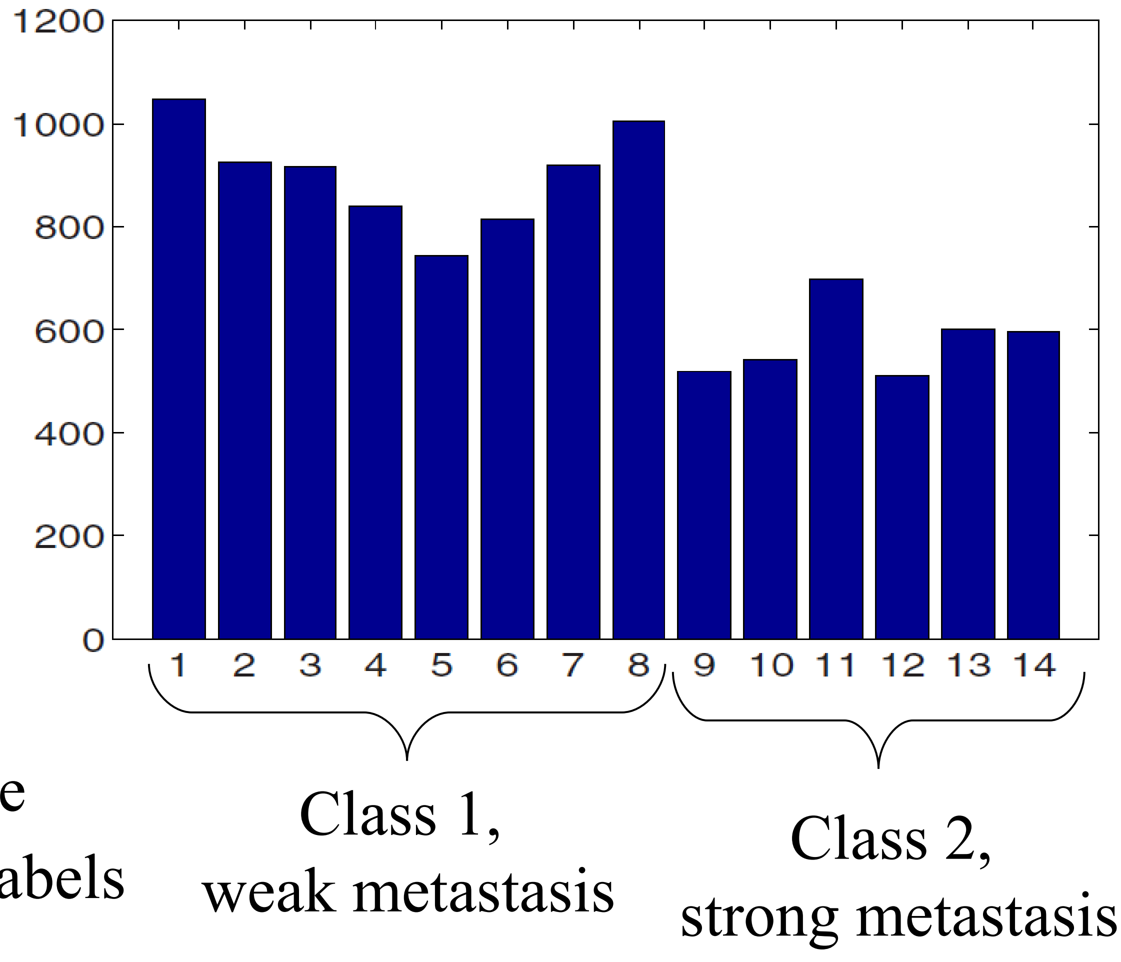
Fig from Jung

ICA for Microarray data processing (Schachtner at al, ICA07)

Brest Cancer Data set

M=14 Experiments
N=22283 genes
2 classes

9^{th} column of **A**:



$|\mathrm{Corr}(\mathbf{a}_9, \mathbf{d})|=0.89$, where
d is the vector of class labels

ICA for Microarray data processing (Schachtner at al, ICA07)

Leukemia Data set

M=38 Experiments

N=5000 genes

3 classes: ALL-B, ALL-T, AML