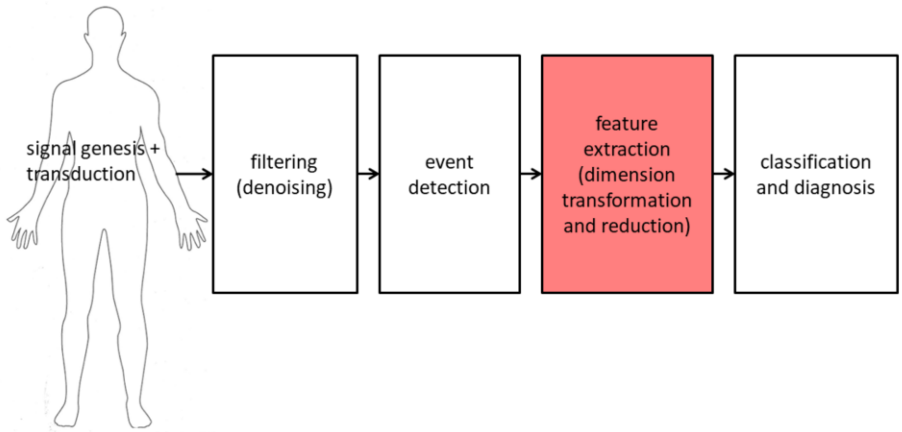
Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

Biomedical Signal Processing
2018-2019 Autumn

Source Separation

Lecturer: Janka Hatvani
Responsible lecturer: dr. Miklós Gyöngy

The BSP Flow Chart



Today's goal

- Examples on the separation problem – what does it mean? How is it related to the dimension reduction problem?
- The PCA/ICA algorithms

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

Motivation

- **Biomedical signals often arise from several additive sources/components**
  - EEG: summation of local field potentials
  - EMG: summation of motor unit action potentials (MUAPs)
- **Can we separate these components?**
  - separation in different signal domains (e.g. t/f/t-f)
  - separation using several channels/trials
    - several microphones/PCG or ECG/EEG/EMG leads
    - several recordings of heart beat/pulse or ERP

4

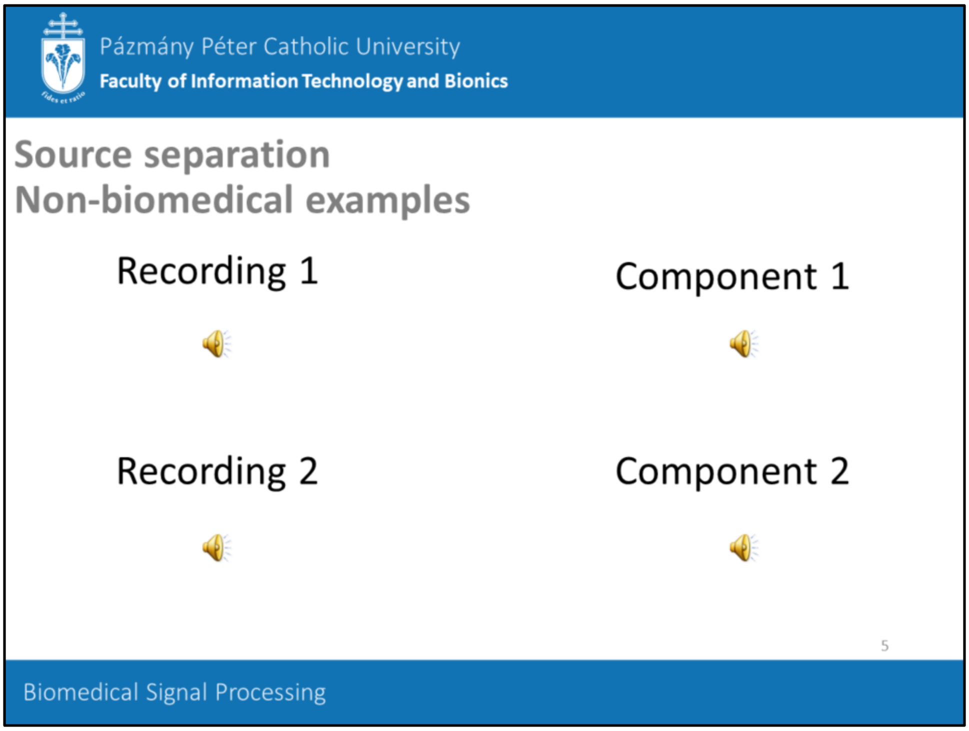
Biomedical Signal Processing

separation in different **signal domains** (e.g. t/f/t-f):

- eg. the frequency of the maternal and fetal heart rate is different, we can try to separate them in the frequency domain. Smilar can be, if the 2 sources are separate in time, and with some windowing we can separate them.

separation using **several channels/trials**

today we are going to be talking about latter. However, clearly we can make several channels by running a set of filters through a signal and then trying to run these algorithms on it



The problem:

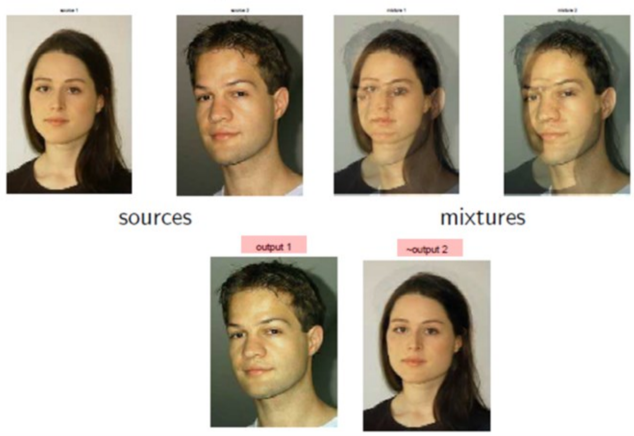
We have two microphones in a room. Two people are talking, both of their voices are recorded on the microphones. Using the two measurements, try to separate the voice of the two speker.

What if we had 3 microphones, and the ‚third' speaker' would be a noise, like the buzzing of some equipments in the room? In this case We will separate to 3 sources, but one of them (with the smallest amplitude) will be the noise. By neglecting this component, we implemented dimension reduction.

6. Non-biomedical examples: http://cnl.salk.edu/~tewon/Blind/blind_audio.html

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

Source separation
Non-biomedical examples



6

Biomedical Signal Processing

The problem:
We have two images. In some process these images are getting mixed. With the source separation algorithm try to separate the two images from each other.

7. Non-biomedical examples: http://www.biologie.uni-regensburg.de/Biophysik/Theis/research/tutorialOnBSS.pdf

Model

- N components X mapped to N channels of recordings Y using $N \times N$ mixing matrix A
- I samples in time, frequency, or *any other suitable basis*

$$\begin{aligned} Y &= AX \\ \text{find } \tilde{A} \text{ s.t. } \tilde{X} &= \tilde{A}^{-1} Y \end{aligned}$$

The N components in matrix X are the N speaker in the room. The N channels in matrix Y are the N microphones. The mixing matrix can contain the different weights of the components – eg. depending on their distance from the microphone.

If the model is not linear (the 2 sources are not additive), but multiplicative mixing: take the logarithm of the model to come to a linear model.

Haven't we found this problem of source mixing before in our earlier lectures? (breathing+ECG, EEG+EMG, etc.)
However, in those tasks we had some information about the sources to be separated (eg. We used the different frequency of the breathing and ECG). When the smixing matrix A is unknown, it needs to be estimated: blind source separation

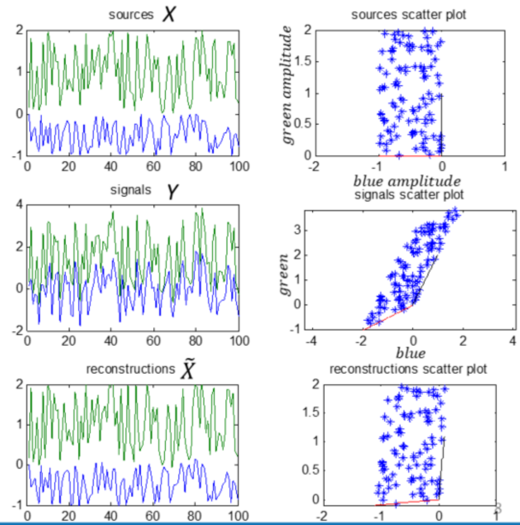
- Not all components may be sources: rest considered noise → This is dimension reduction.
- model generally cannot cope with more sources than available signal channels; in this case we need to have sparsity of the data in order to perform separation
- Choose your basis of representation well (e.g. may help with sparsity)

Pázmány Péter Catholic University

Faculty of Information Technology and Bionics

2-D visualization

$$
\begin{aligned}
Y &= AX \\
\tilde{X} &= BY \\
A &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\
A^{-1} &= \begin{bmatrix} 0.7 & -0.3 \\ -0.3 & 0.7 \end{bmatrix}
\end{aligned}
$$



Biomedical Signal Processing

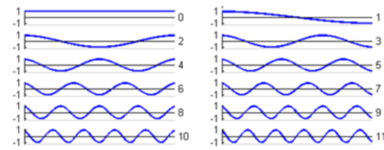
Left: On the 2 electrodes X we are measuring 2 signals. The ground truth, which should be reconstructed is in Y. X~ is the estimated version using source separation.

Right: scatter plot of the plot on the left. The axis are the green/blue amplitues, one asterisk is one timepoint.

From Fourier(/Wavelet) transforms to Component Analysis

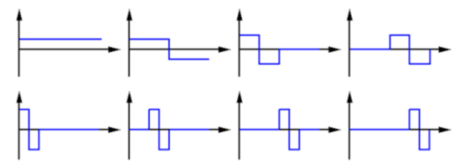
Fourier Basis

- Sinusoids at different frequencies
- Any signal can be composed from it



Wavelet Basis

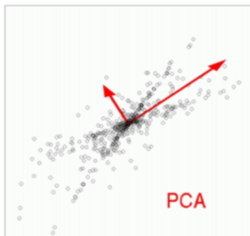
- Motherwavelets at different scales and shifts
- Might be overcomplete

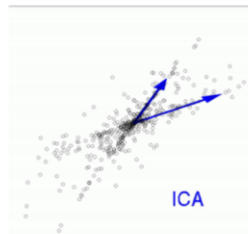


Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

From Fourier(/Wavelet) transforms to Component Analysis

Let's design a data-dependent basis!





- PCA – uncorrelated $\langle x_1 x_2 \rangle = \langle x_1 \rangle \langle x_2 \rangle$
- ICA – uncorrelated $P(x_1 = a, x_2 = b) = P(x_1 = a)P(x_2 = b)$

Biomedical Signal Processing

PCA: Finds principal components of the dataset. Each succeeding step finds direction that explains most variance.
It removes correlations, but not higher order dependence. Some components are more important than others, vectors are orthogonal.

ICA: the pdf of x is equal to the multiplication of each marginal pdf of x.
separation by statistical independence – higher order statistics are analysed compared to the variance (2nd order statictics). This statistic can be Minimization of mutual information, Maximization of non-Gaussianity
Removes correlations and higher order dependence. All components are equally important, vectors are non-orthogonal.

http://compneurosci.com/wiki/images/4/42/Intro_to_PCA_and_ICA.pdf
10. Separation by independence: Hyvärinen and Oja: Independent Component Analysis: Algorithms and Applications
Niedermeyer and Silva (2005): Electroencephalography: Basic Principles, &c

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

Principal Component Analysis (PCA)

$Y = AX$

Squared: $Y \otimes Y^* = AX \otimes X^* A^*$
$Y \otimes Y^* = R_Y$, *the covariance of the data*
$R_Y = A R_X A^*$

R_X is diagonal (as the components are supposed to be uncorrelated). So it is actually an SVD!

SVD of $R_Y = UDU^*$

$$A \leftarrow U; R_X \leftarrow D$$

11

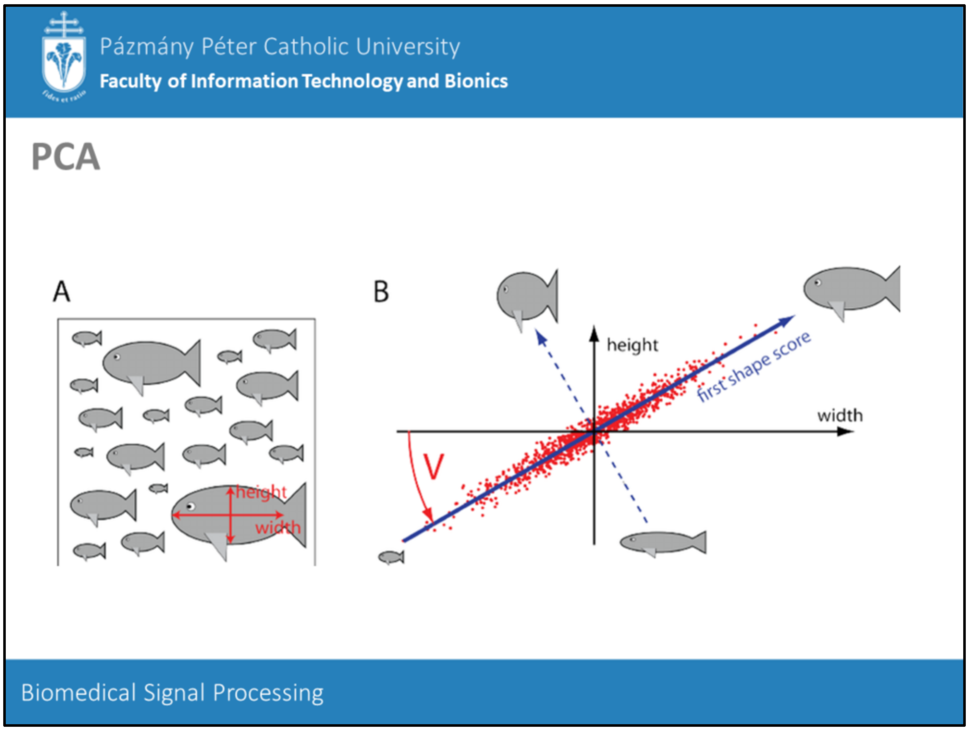
Biomedical Signal Processing

$\otimes$ is the symbol for outer product

SVD: singular value decomposition. U is the matrix of eigenvectors, D is the diagonal matrix of eigenvalues.
We calculate the singular value decomposition of the correlation matrix R_Y.
With the substitution $A \leftarrow U; R_X \leftarrow D$ we know the eigenvectors transforming us to the new basis. The eigenvalues in D correspond to their importance: the heighest eigenvalue is the most important dimension.

Fine to assume that sources are uncorrelated. But is it not a problem to assume that transformation is also orthogonal?
We have example in 2D where even the second signal is recovered reasonably well.
What happens when we have more than 2 signals?

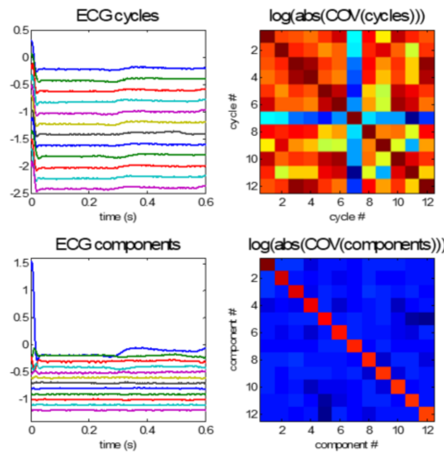
11. Principal Component Analysis (PCA):
https://en.wikipedia.org/wiki/Principal_component_analysis



http://setosa.io/ev/principal-component-analysis/
extracting relevant information from confusing data sets
PCA is a way of identifying patterns in
data, and expressing the data in such a way as to highlight their similarities and differences.

https://www.google.fr/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwiPzJbG-ereAhUih6YKHS1pBL8Qjhx6BAgBEAM&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FIllustration-of-principal-component-analysis-A-As-a-minimal-example-we-consider-a_fig1_263968032&psig=AOvVaw195Gm_0MHweAyFak99rbFX&ust=1543077435477411

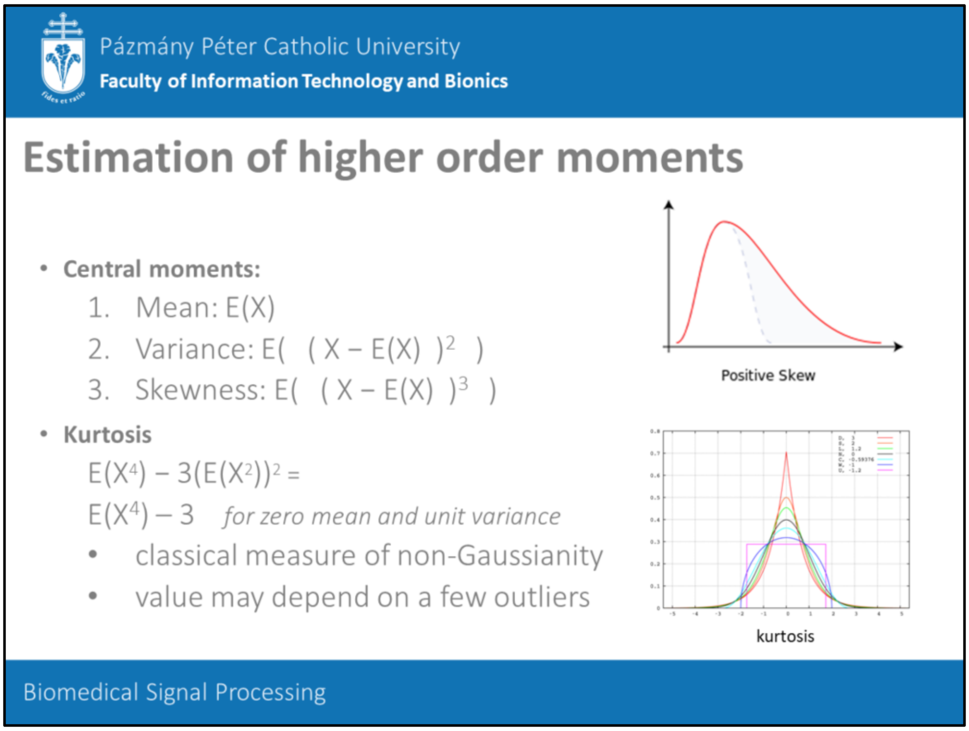
PCA Results – ECG decomposition



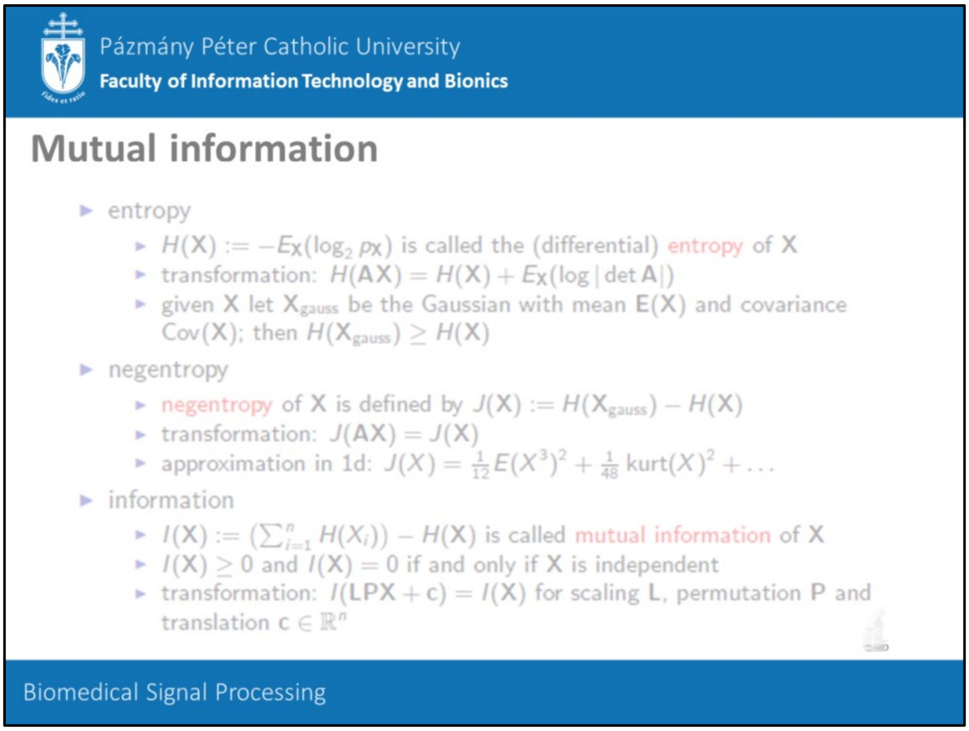
This is example of having several trials rather than several channels. We performs eigendecomposition of covariance matrix and find principal components.
Note that here, components will be uncorrelated with each other.

Independent Component Analysis (ICA)

- **Measure of independence:**
  - non-Gaussianity is maximised (cf. central limit theorem); using higher than second-order statistics (e.g. kurtosis)
  - mutual information is minimised using higher order statistics
- **ICA cannot separate Gaussian distributed data from each other!**



23. Estimation of higher order moments: Hyvärinen and Oja: Independent Component Analysis



from wikipedia

mutual information closely related to entropy.

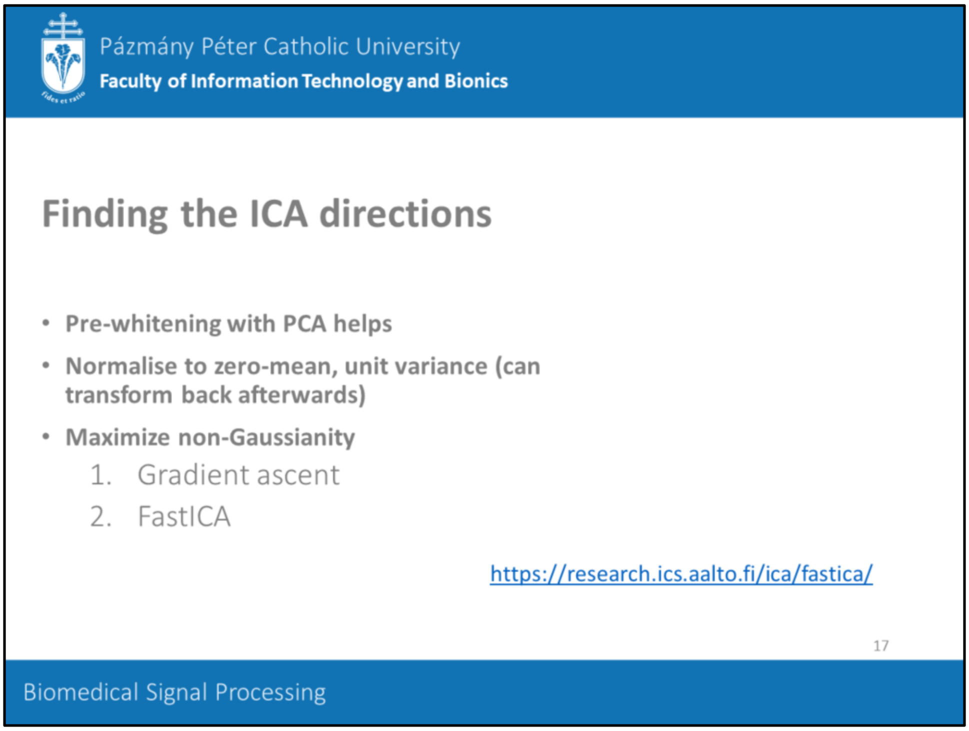
evaluation of a function of the joint and marginal probabilities.

Point to note: mutual information is always equal or greater than 0, and only exactly 0 when there is independence.

Remember indepedence? We could also evaluate it as p(x,y)-p(x)p(y).

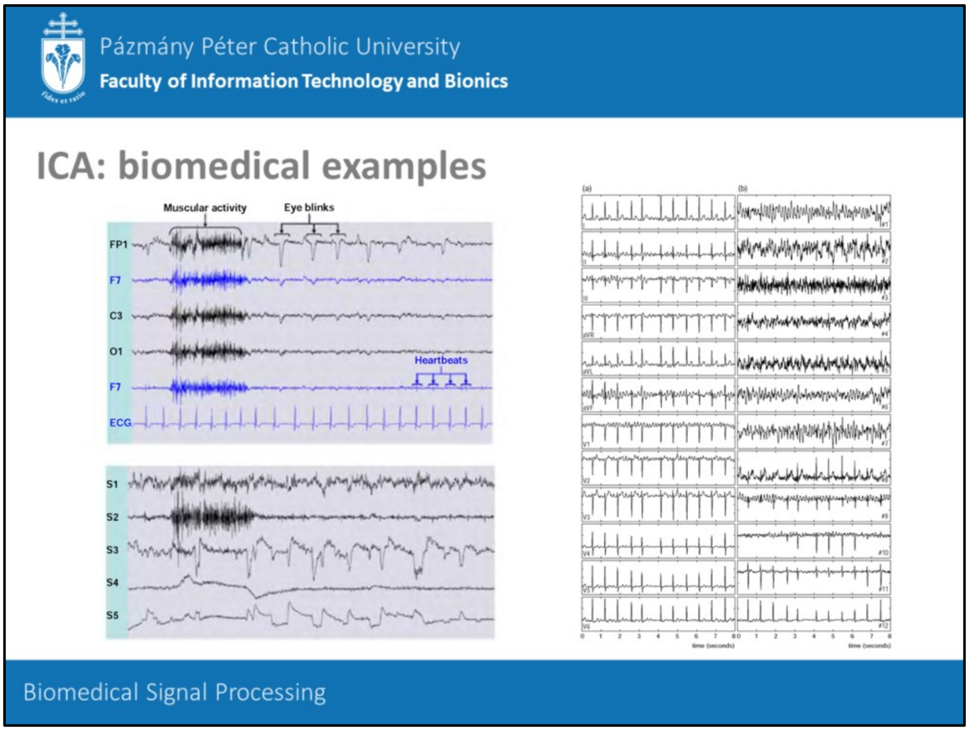
But MI can be nicely defined in terms of individual higher moments of variables.

24. Mutual information.



Optimization procedure to maximize non-Gaussianity, however it is defined.

25. Finding the ICA directions: http://www.biologie.uni-regensburg.de/Biophysik/Theis/research/tutorialOnBSS.pdf



Left: ICA on EEG + ECG signals (p. 743, 772)

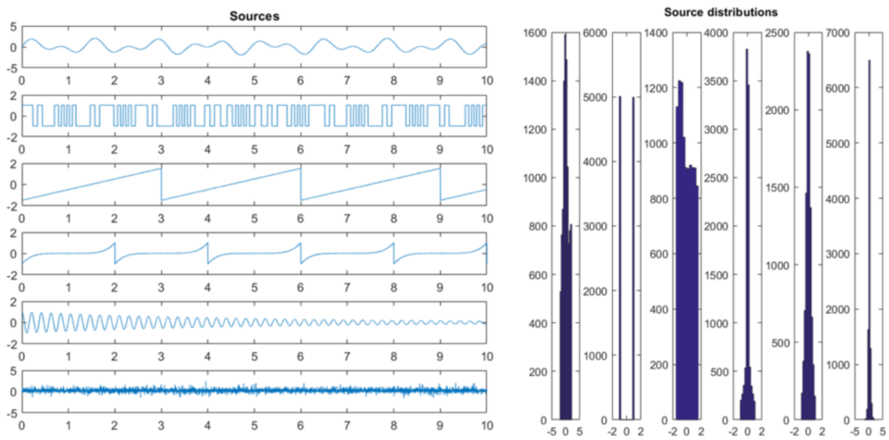
Right: ICA on 12-lead ECG (p. 755). Right, going down: increasing levels of kurtosis

Comon (2010): Handbook of Blind Source Separation

26. ICA: biomedical examples: Comon (2010): Handbook of Blind Source Separation (pp. 743,755,772)

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

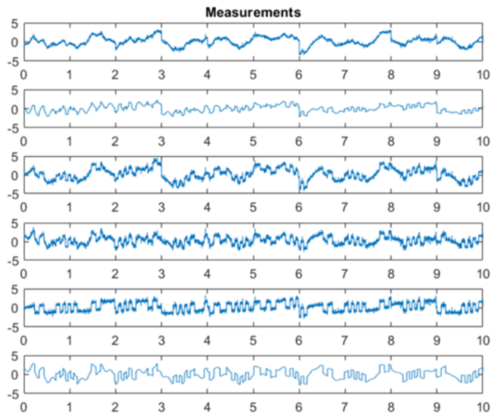
PCA vs ICA



Biomedical Signal Processing

https://pdfs.semanticscholar.org/16fc/90913c1074f568ac9ca81cf0a10fcc00c379.pdf

PCA vs. ICA

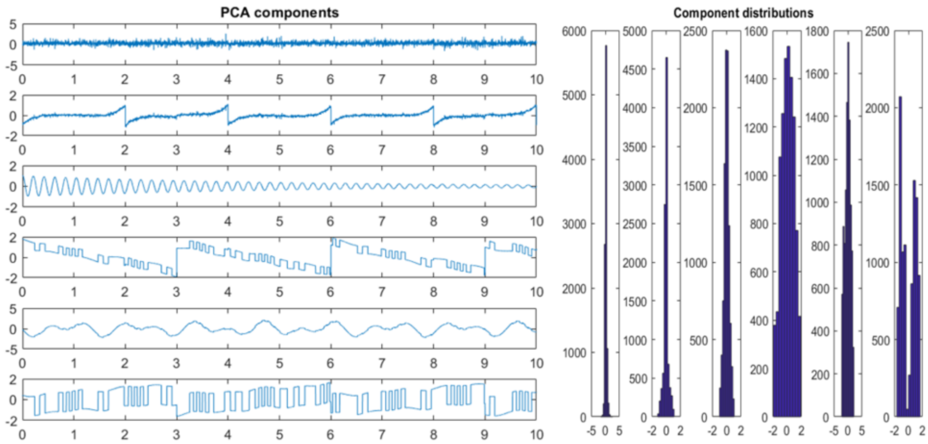


Mixing matrix A

0.88	0.15	0.59	0.96	0.52	0.54
0.21	0.44	0.71	0.23	0.99	0.14
0.97	0.67	0.98	0.00	0.29	0.75
0.60	0.79	0.02	0.45	0.82	0.86
0.08	0.99	0.48	0.80	0.06	0.76
0.71	0.98	0.14	0.40	0.81	0.05

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics

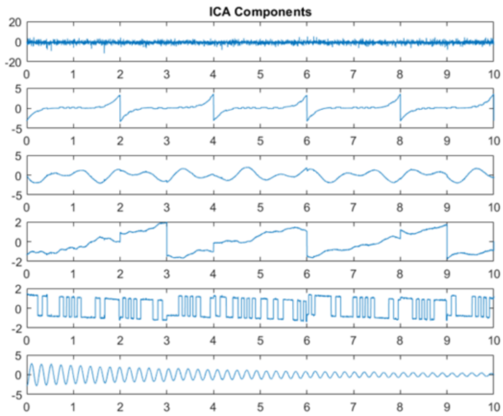
PCA vs. ICA

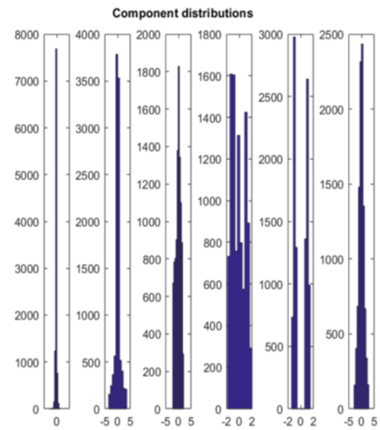


Biomedical Signal Processing

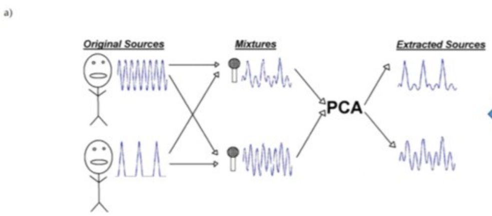
```
Rx = cov(Measurements);
[V, D] = eig(Rx);
C_pca = Measurements*V;
```

PCA vs. ICA

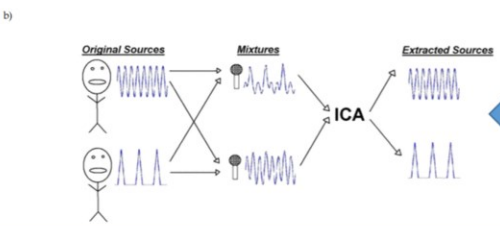




Is a soundwave Gaussian?

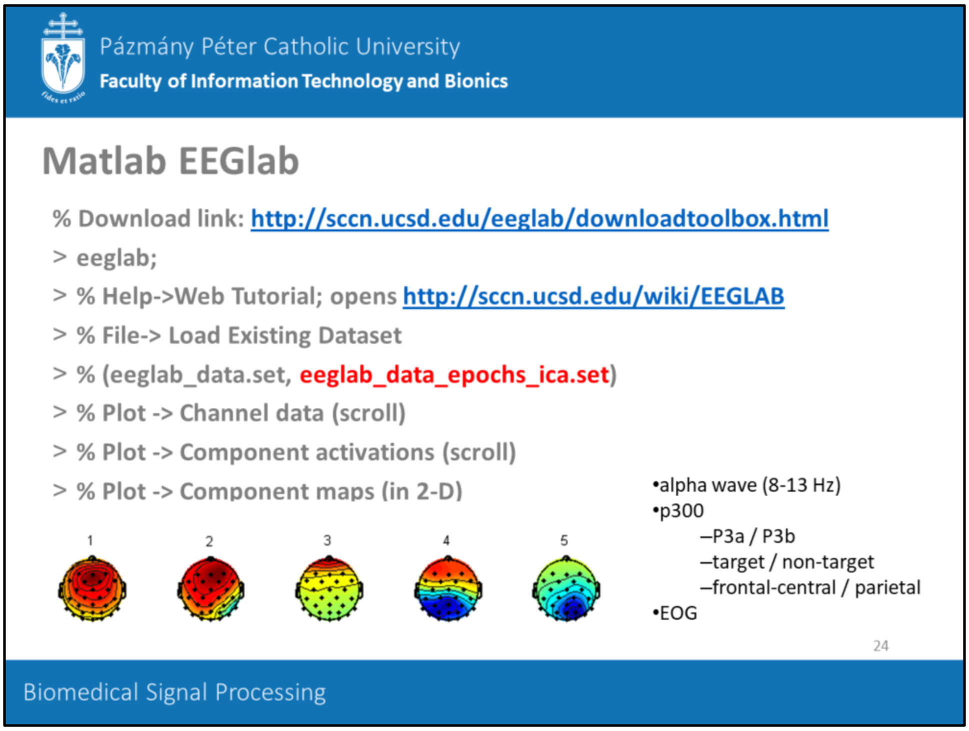


Assumption:
- Gaussianity and uncorrelatedness



Assumptions:
- statistical independence
- Equally important components

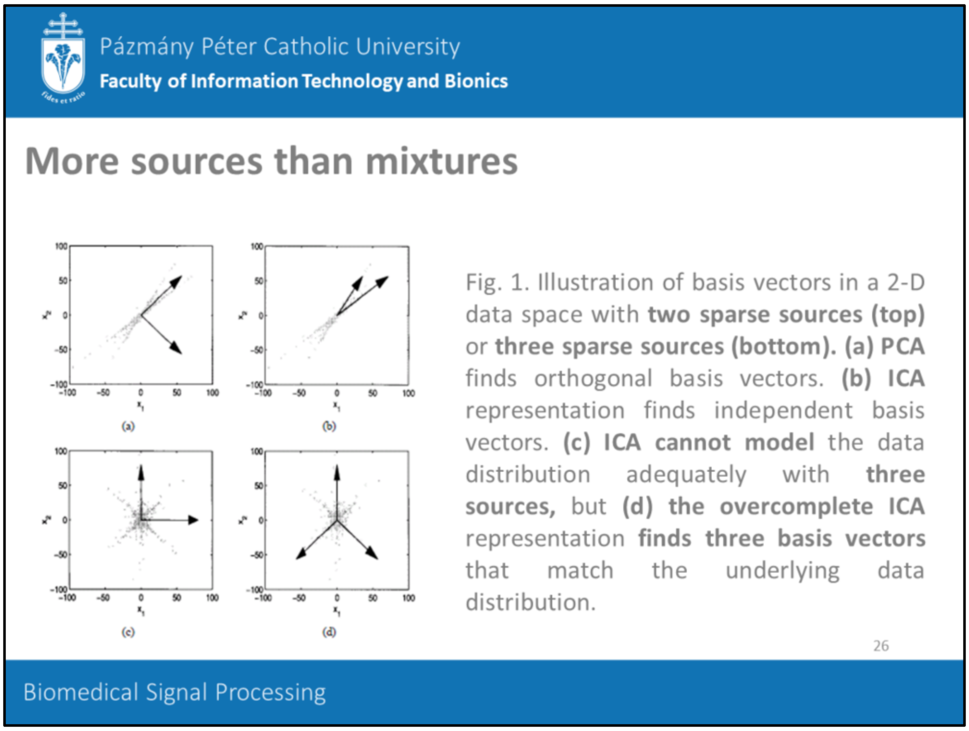
http://mail.tqmp.org/RegularArticles/vol06-1/p031/p031.pdf



ICA components happen to be nicely concentrated in different parts. Does not automatically follow but it's nice.

27. MATLAB EEGlab S. Makeig et al. (2002), Dynamic Brain Sources of Visual Evoked Responses, Science 295, pp. 690–694

Extra material



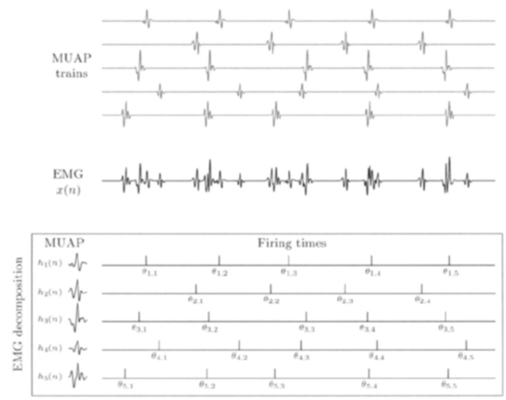
Sparsity of signals. Think of speech earlier. One, two, three

28. More sources than mixtures: Lee et al. (1999): Blind source separation of more sources than mixtures using overcomplete representations
Fig. 1. Illustration of basis vectors in a 2-D data space with two sparse sources (top) or three sparse sources (bottom). (a) PCA finds orthogonal basis vectors. (b) ICA representation finds independent basis vectors. (c) ICA cannot model the data distribution adequately with three sources, but (d) the overcomplete ICA representation finds three basis vectors that match the underlying data distribution

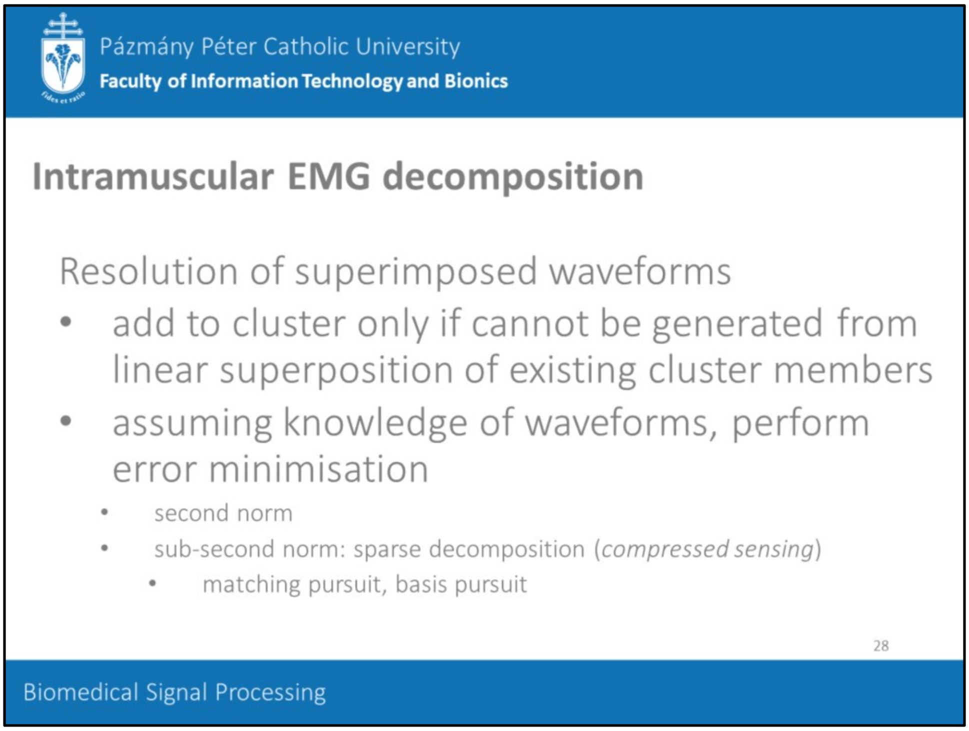
Intramuscular EMG decomposition

1. **Filtering (bandpass)**
2. **Event detection**
3. **Feature extraction**
4. **Clustering**
5. **Resolution (of superimposed)**

- *firing time model?*



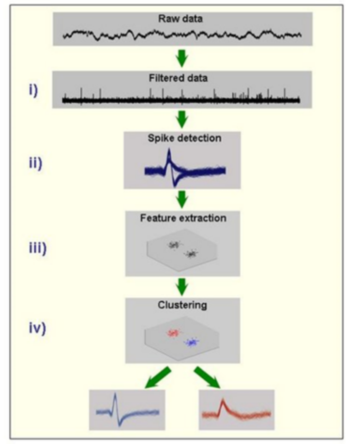
30, 31. Intramuscular EMG decomposition: Sörnmo p. 392-395



30, 31. Intramuscular EMG decomposition: Sörnmo p. 392-395

Spike sorting

- **Setting the threshold:**
  - manual
  - multiple of standard deviation
  - 5 median{|x|/0.6745}
- **Feature extraction:**
  - PCA, wavelet coefficients, other shape descriptors
- **Superimposed spikes**
  - linear superposition of existing waveforms
  - *firing time model?*



Inverse source problem

Take a modification of our simple model:

$$Y = A(R)X(R),$$

where we know $A(R)$ as a function of source locations R

Possibilities:

- the source locations $\mathbf{r}$ need to be estimated
- more (putative) sources than channels

Solutions:

- (robust) beamforming
- regularized inversion
- subspace-based methods

The forward matrix *A*(*R*)

Suppose $A_{i,j}$ describes propagation of source signal from source at $\mathbf{r}_i$ to sensor at $\mathbf{r}_j$

Possibilities:

- pressure field

$$\alpha = 0; \quad \left[\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right] p = s(\mathbf{r}, t)$$

$$p(\mathbf{r}) = \int \frac{s(\mathbf{r}', |\mathbf{r} - \mathbf{r}'|/c)}{4\pi|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}'^3$$

$$\text{LF free-field: } A_{i,j} = e^{-\alpha|\mathbf{r}_i - \mathbf{r}_j|}/|\mathbf{r}_i - \mathbf{r}_j|$$

$$\text{HF: } A_{i,j} = e^{(-\alpha + j2f)|\mathbf{r}_i - \mathbf{r}_j|}/|\mathbf{r}_i - \mathbf{r}_j|$$

- electromagnetic field

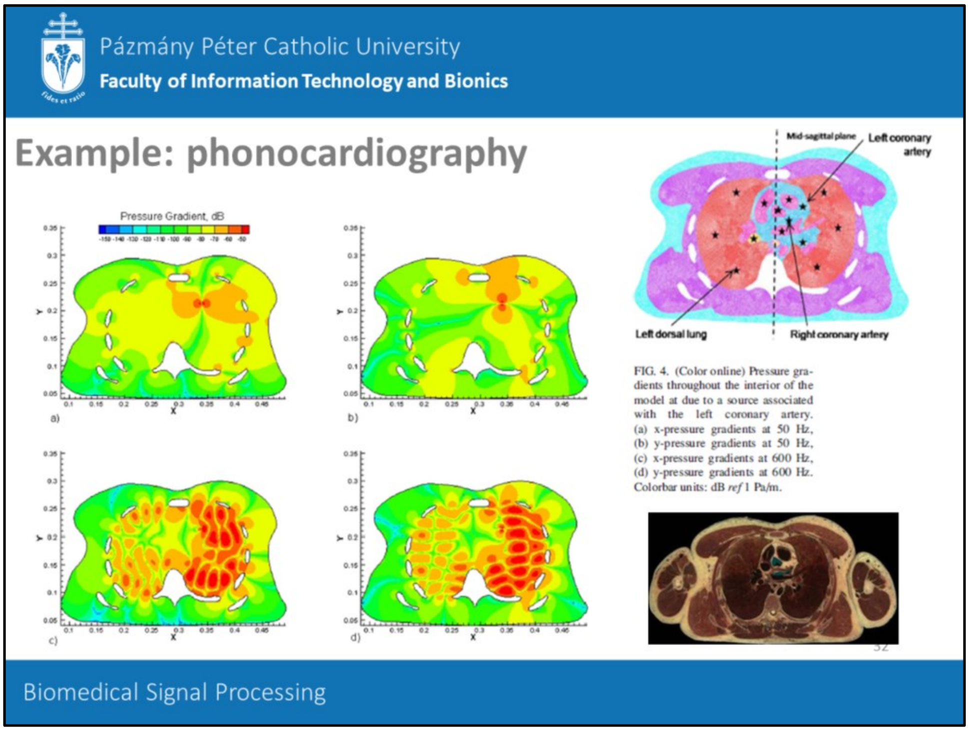
$$\nabla^2 V = \nabla \cdot J/\sigma$$

$$\mathbf{k}_{\alpha\beta} = \mathbf{k}(\mathbf{s}_\alpha, \mathbf{v}_\beta) = \frac{1}{4\pi\sigma}\frac{(\mathbf{s}_\alpha - \mathbf{v}_\beta)}{\|\mathbf{s}_\alpha - \mathbf{v}_\beta\|^3} - \frac{1}{4\pi\sigma}\frac{(\mathbf{s}_R - \mathbf{v}_\beta)}{\|\mathbf{s}_R - \mathbf{v}_\beta\|^3}$$

$$V = \frac{1}{4\pi\sigma}\int_{\mathcal{V}} \frac{\nabla \cdot J}{R} \mathrm{d}\mathbf{r}'^3; \quad V(\mathbf{r}) = \int_{\mathcal{S}} \frac{p(\mathbf{r}')\cos\theta}{4\pi\sigma|\mathbf{r} - \mathbf{r}'|^2} \mathrm{d}S$$

Ponsey p. 180

35. The forward matrix A(R): Niedermeyer and Silva (2005): Electroencephalography: Basic Principles, &c, p. 119
http://www.uzh.ch/keyinst/NewLORETA/TechnicalDetails/TechnicalDetails.htm



36: Example: phonocardiography. Cooper et al. (2011): Acoustic source separation for the detection of coronary artery sounds

Example: EEG

- LORETA: low-resolution brain electromagnetic tomography (standardized and exact: sLORETA, eLORETA): regularized inversion, *maximum smoothness*
- FOCUS (focal underdetermined system solver): sparse solution
- Laplacian: indicator of current sources, highlights underlying sources better than potential maps

$$\nabla^2 V = \nabla \cdot J/\sigma$$

37. Example: EEG. Niedermeyer and Silva (2005): Electroencephalography: Basic Principles, &c, p. 829-834, http://www.uzh.ch/keyinst/loreta.htm
http://www.uzh.ch/keyinst/NewLORETA/TechnicalDetails/TechnicalDetails.htm
Gorodnitsky and Rao Sparse signal reconstruction from limited data using FOCUSS: &c
http://virt.uni-pannon.hu/index.php/in-english/research/1090-professor-laboratory-for-bioelectric-neuroimaging