Matlab project

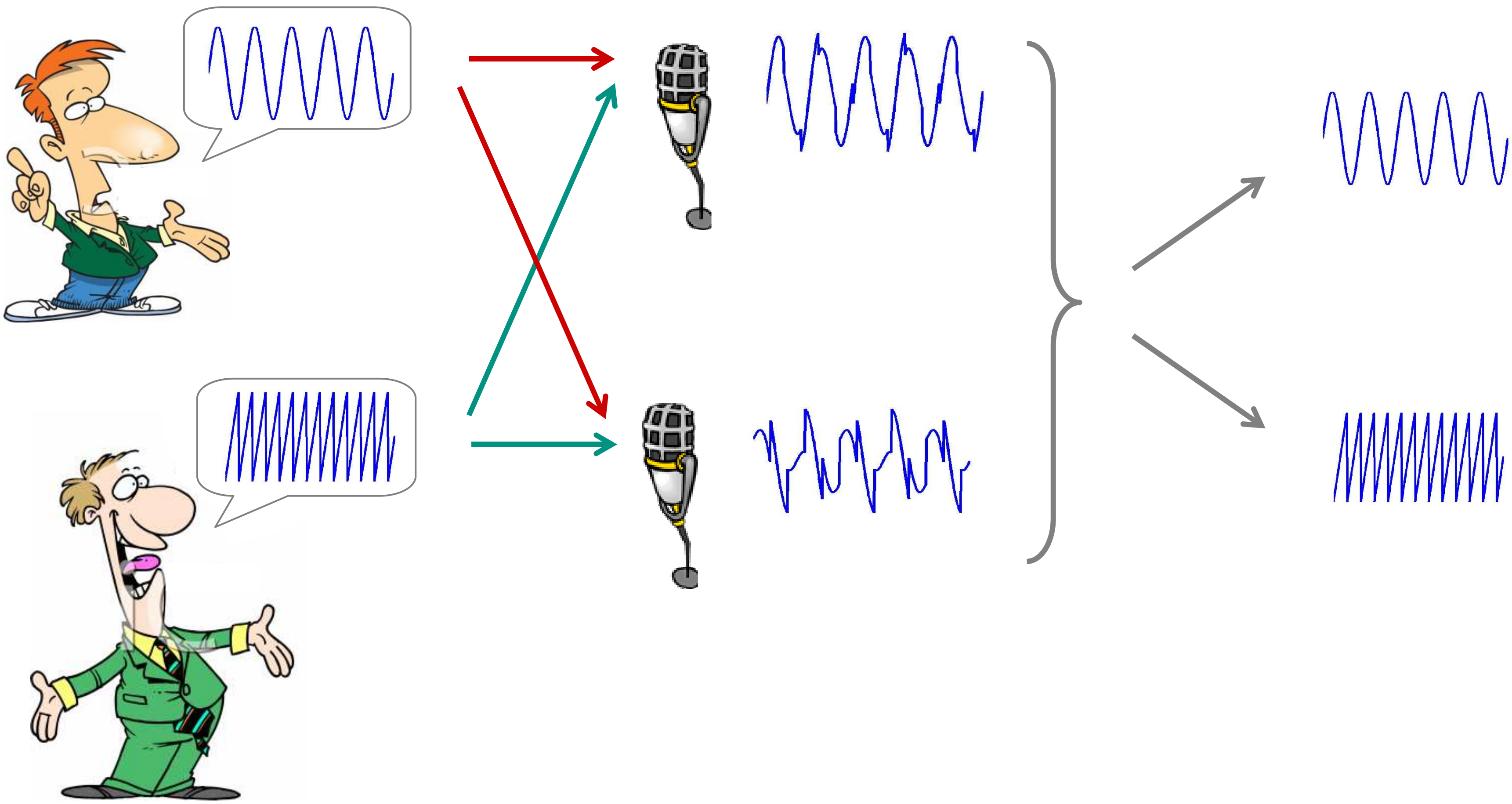
Independent component analysis

Michel Journée
Dept. of Electrical Engineering and Computer Science
University of Liège, Belgium
m.journee@ulg.ac.be

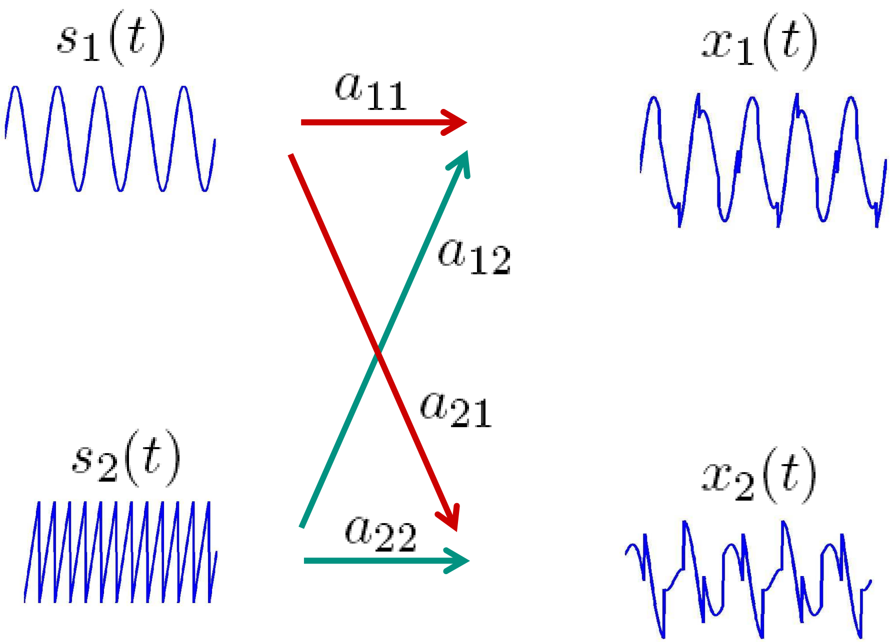
September 2008

What is Independent Component Analysis?

The cocktail party problem



ICA performs
a linear projection into independent components



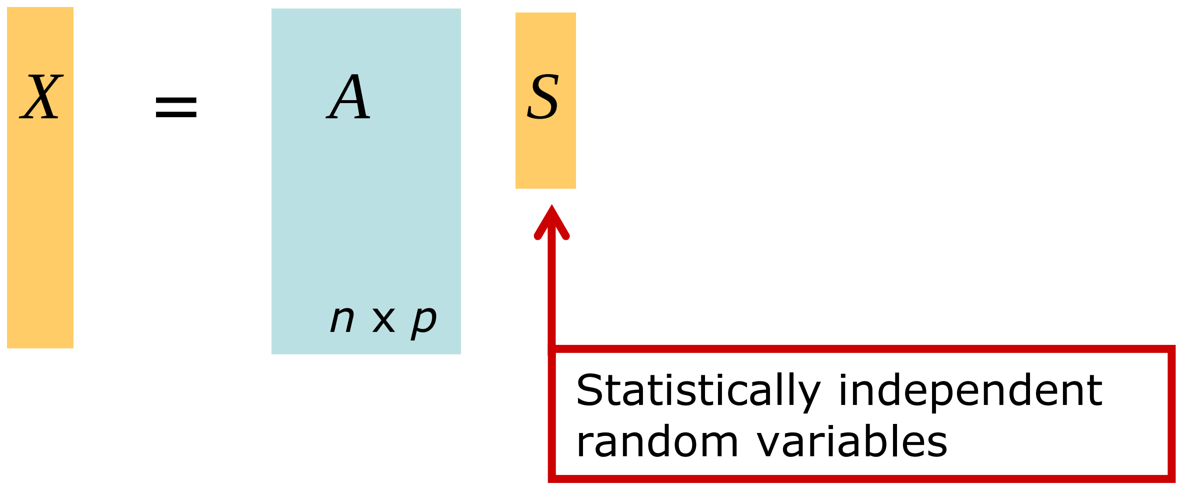
Assumptions

linearity
no delay
statistically independent sources

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

ICA performs
a linear projection into independent components

$$X = AS$$



p number of components

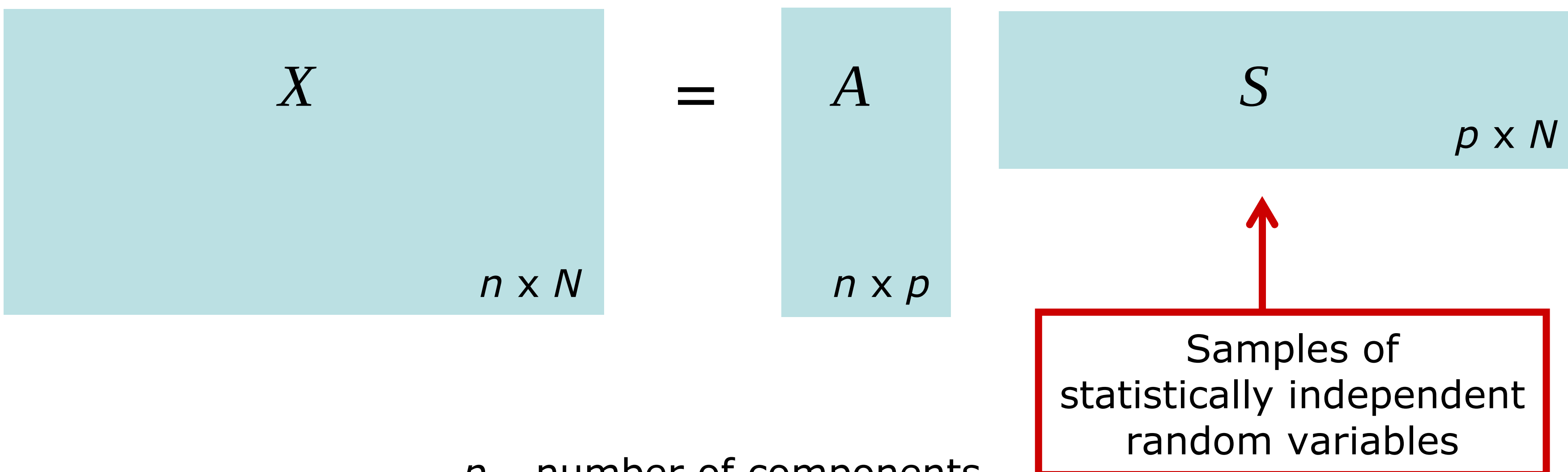
n number of variables

Real matrix

Random vector

ICA performs
a linear projection into independent components

$$X = AS$$

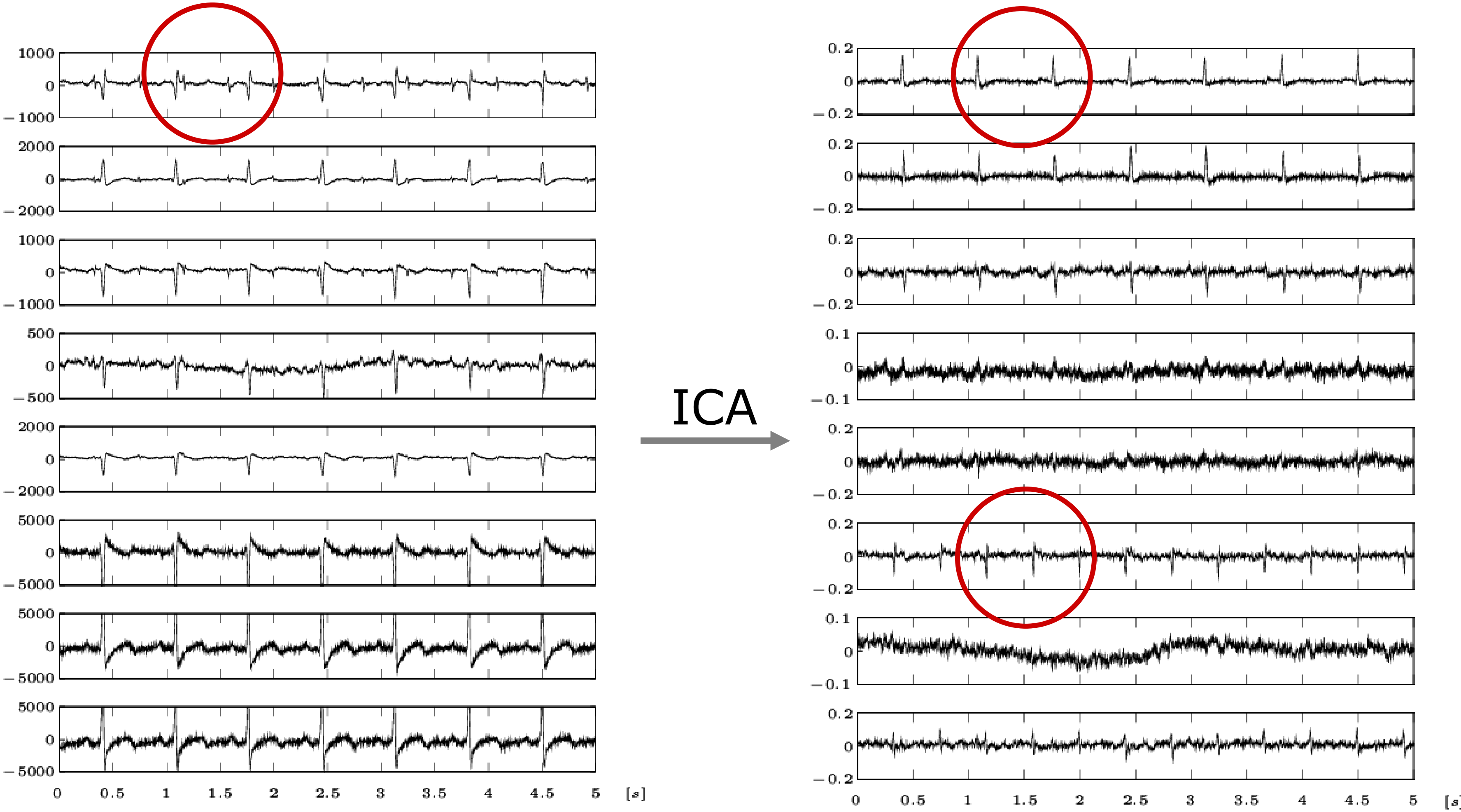


p number of components

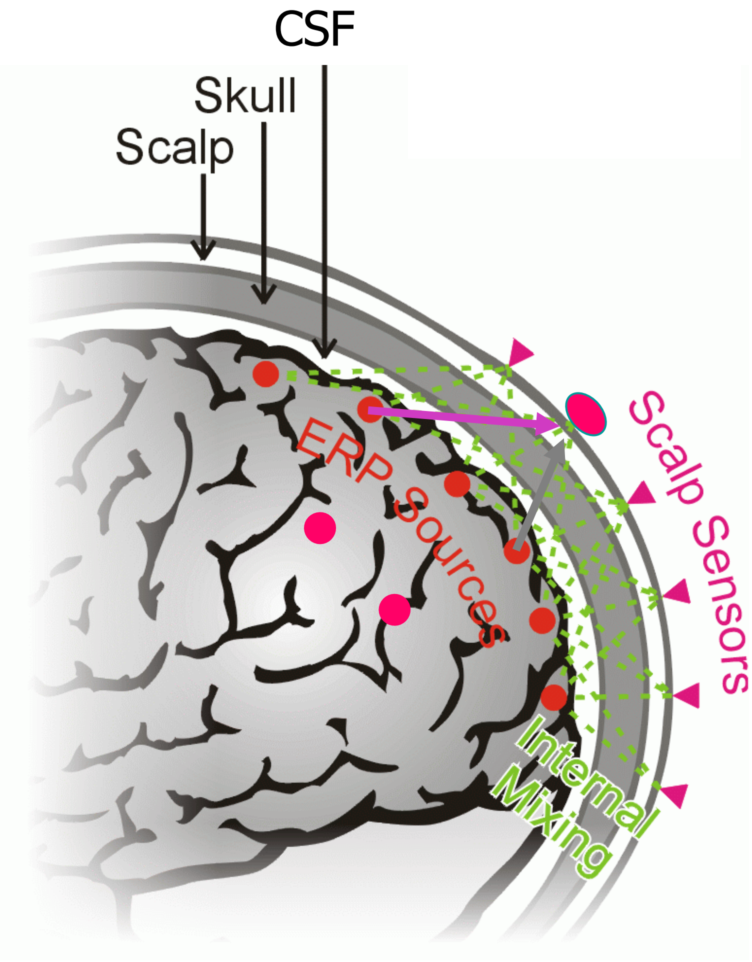
n number of variables

N number of samples

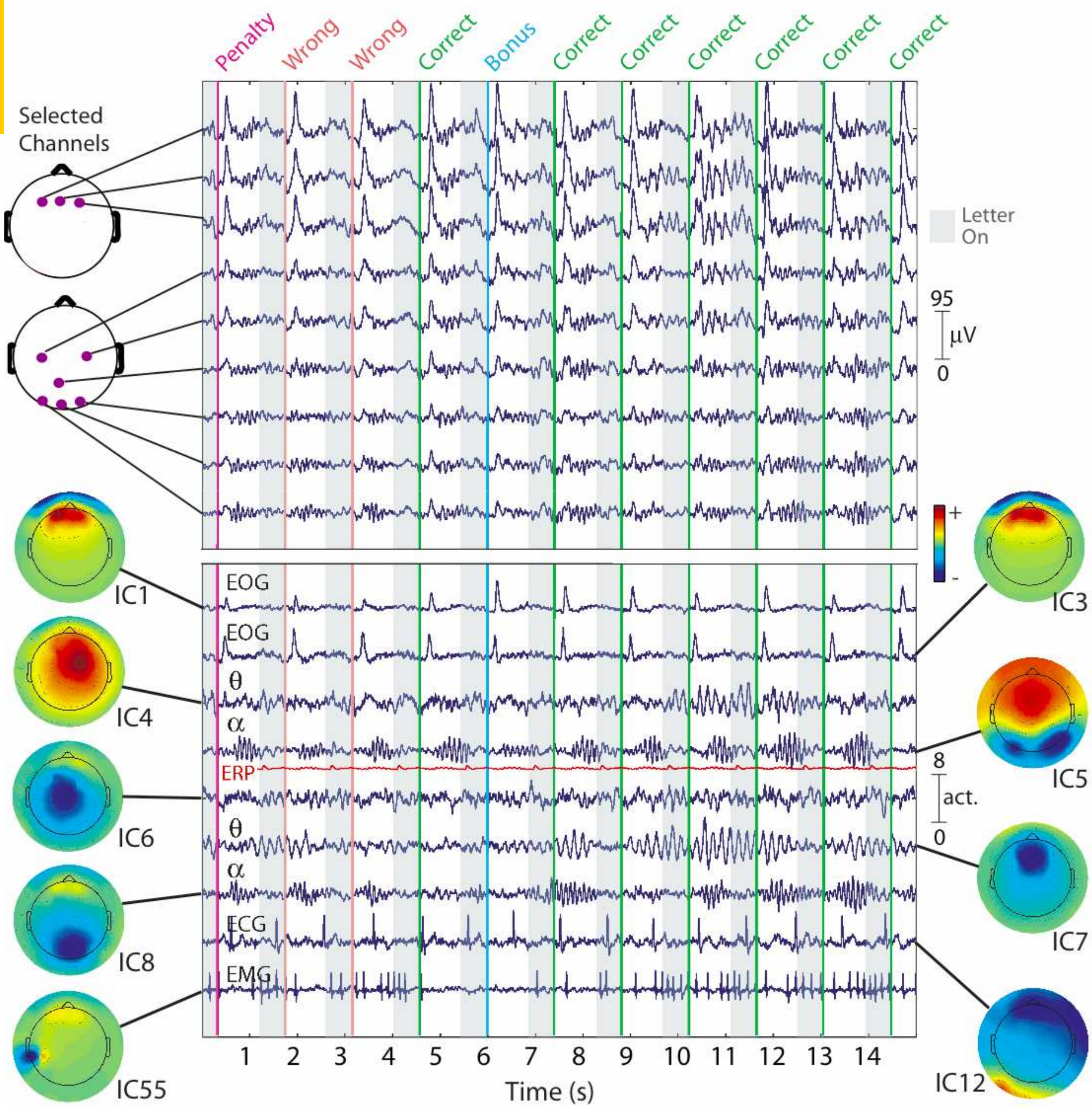
ICA for blind source separation: fECG extraction



ICA for blind source separation: Analysis of EEG

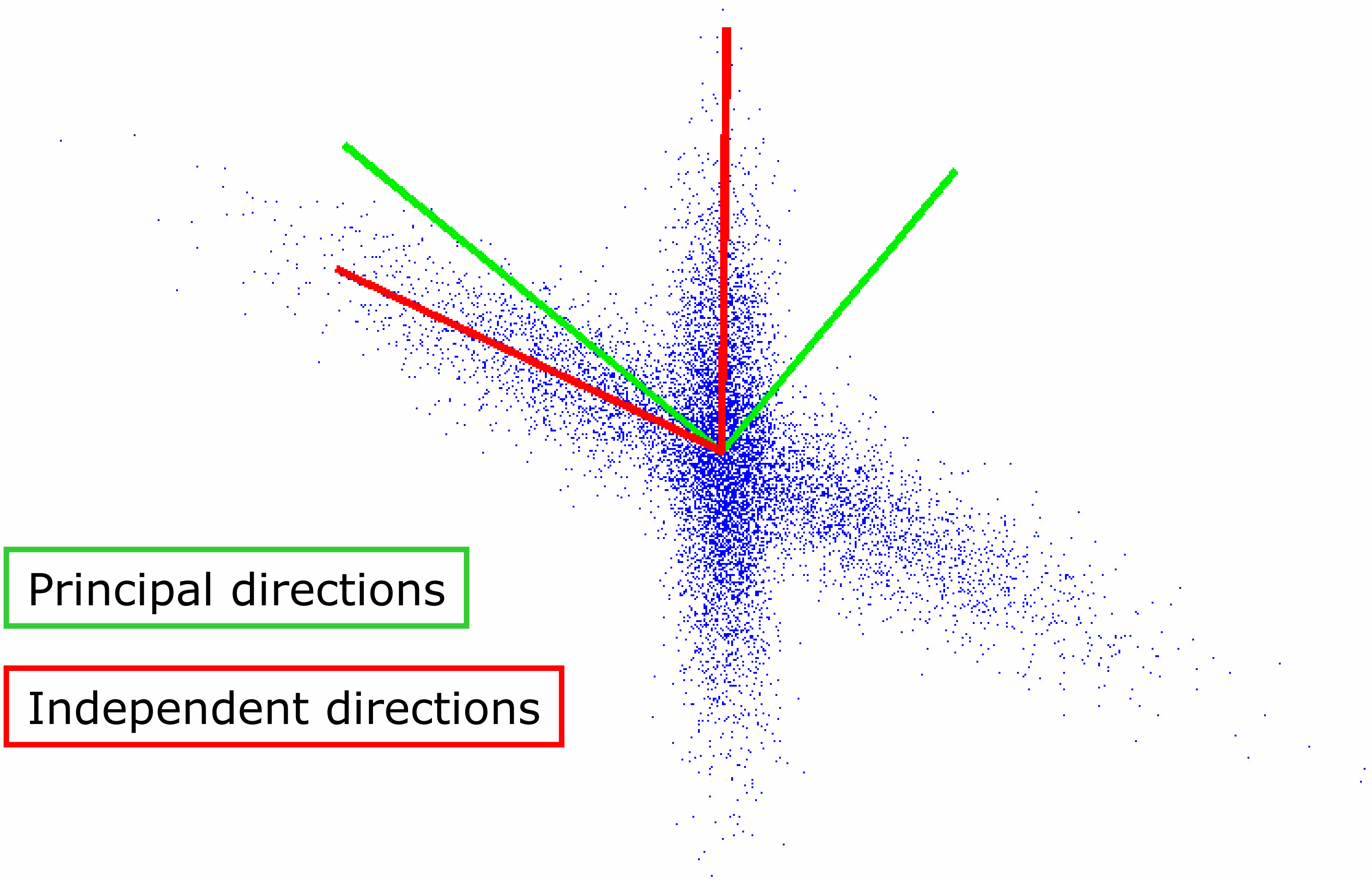


ICA for EEG



Penalty
Wrong
Wrong
Correct
Bonus
Correct
Correct
Correct
Correct
Correct
Correct
Selected Channels
Letter On
95
μV
0
EOG
EOG
θ
α
ERP
θ
α
ECG
EMG
8
act.
0
+
-
IC1
IC4
IC6
IC8
IC55
IC3
IC5
IC7
IC12
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Time (s)

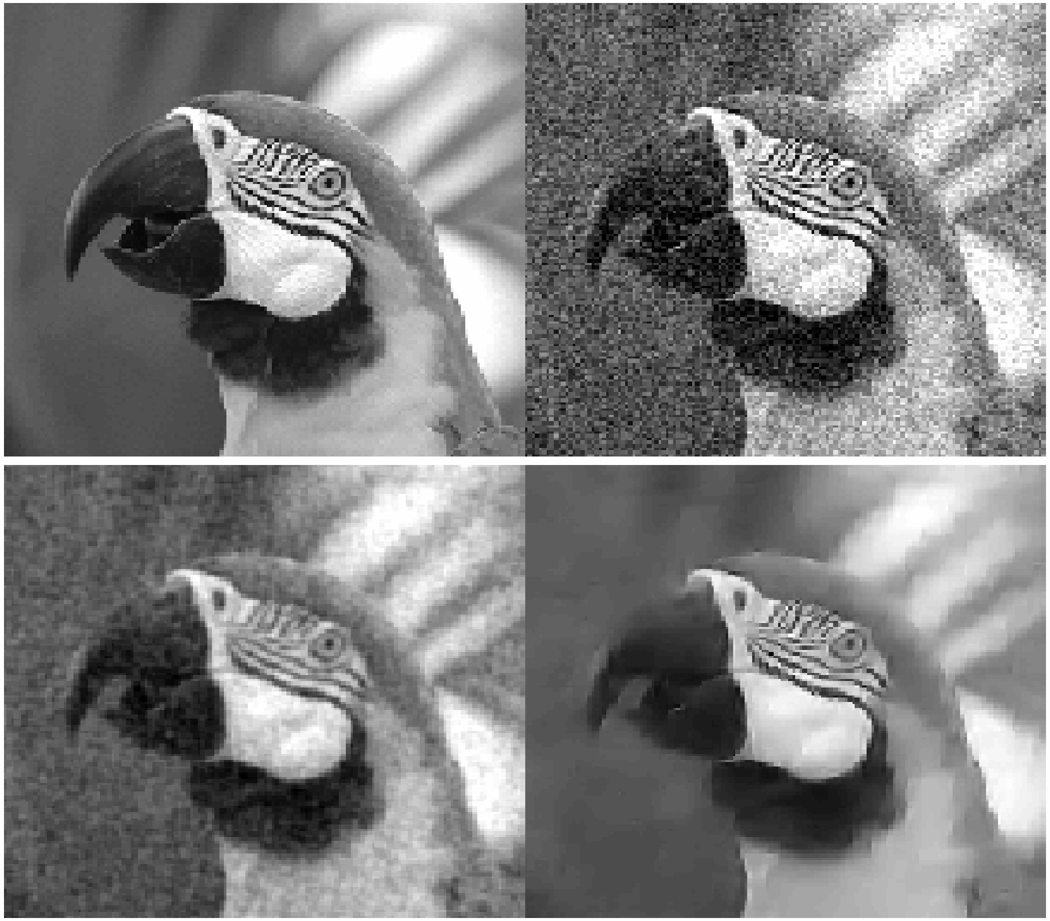
ICA for data analysis



ICA for denoising

Original image

Noisy image



Wiener filtering

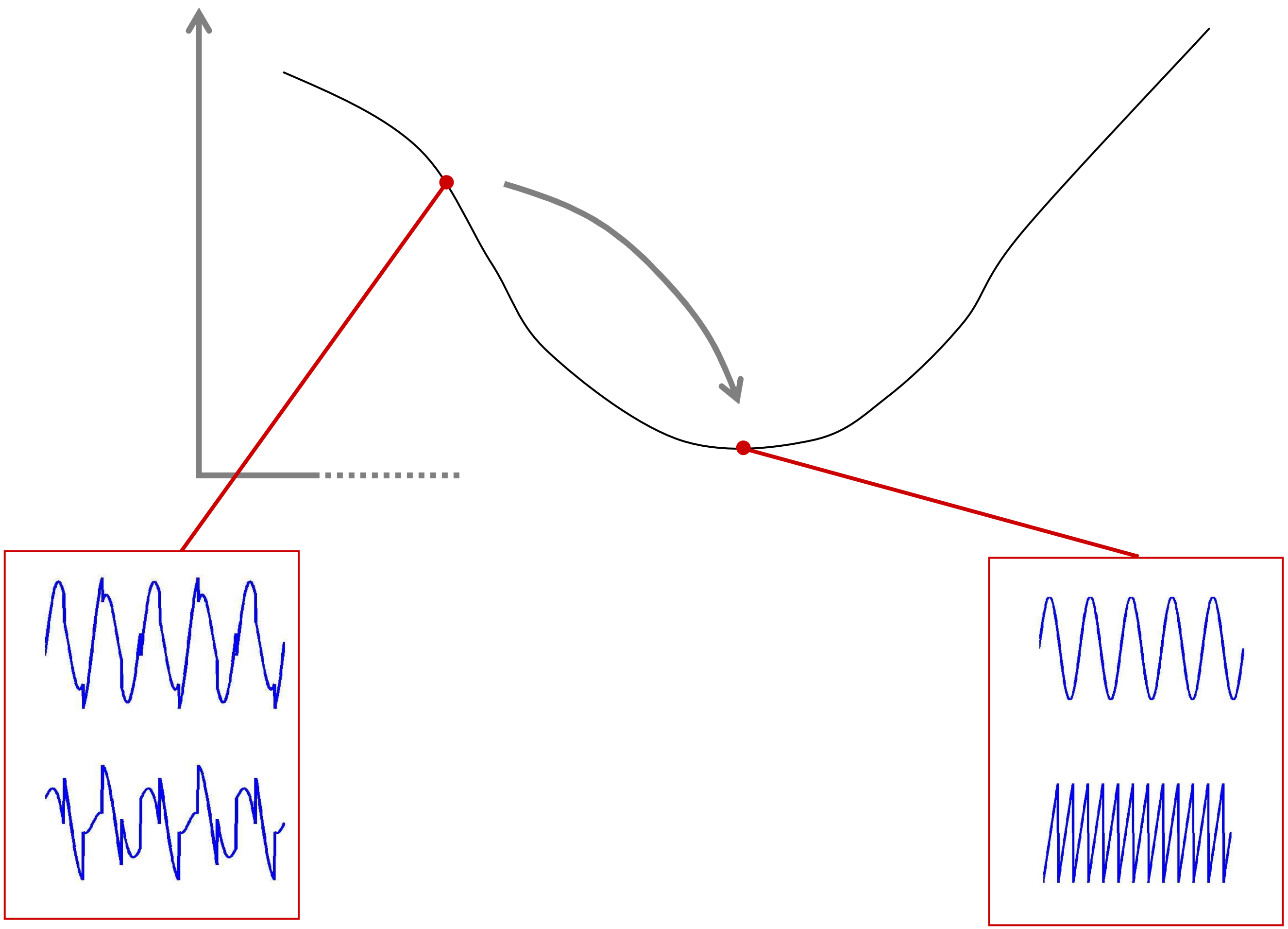
ICA filtering

ICA has applications in many areas

- Blind source separation
  (e.g., biomedical signal processing, radar and mobile communication)
- Data analysis
- Noise reduction
- Feature extraction (image, audio, video representation)
- …

ICA is an optimization problem

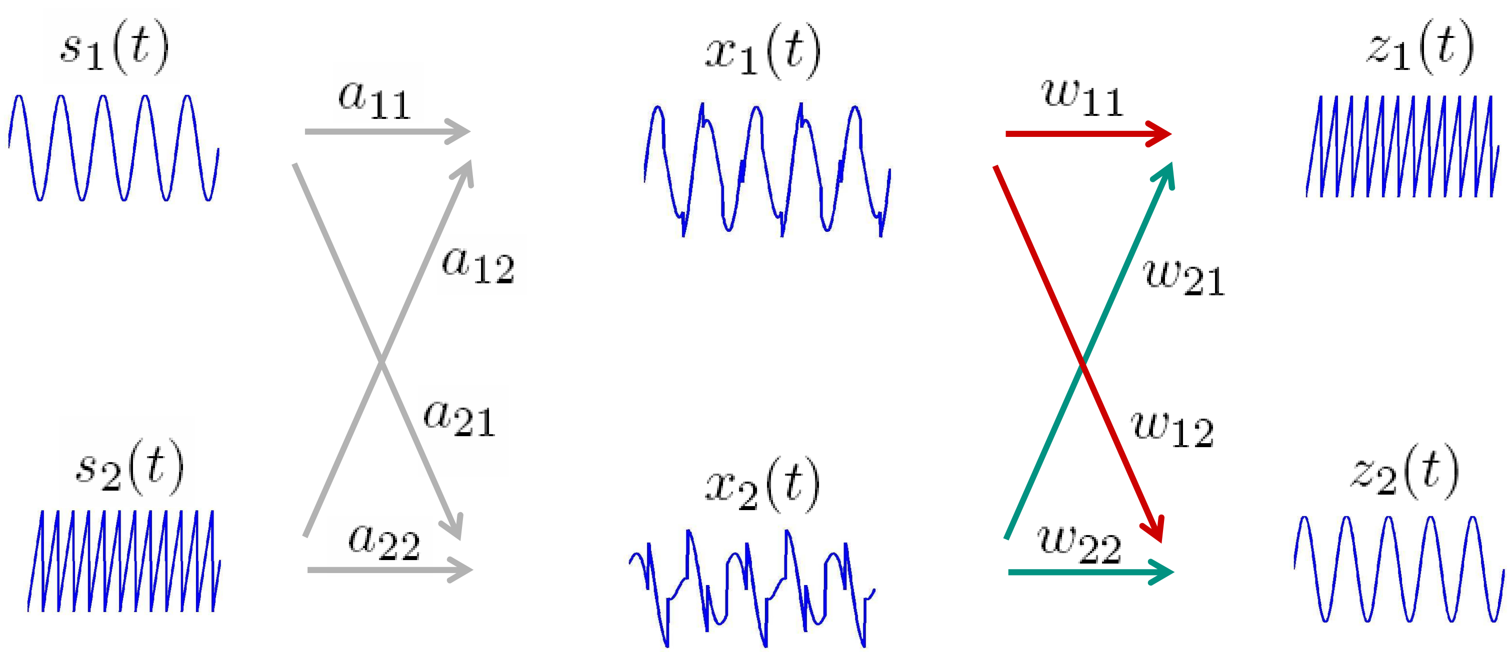
estimator of
statistical independence



ICA algorithms compute the unmixing model

Mixing model

Unmixing model



$$X = AS$$

$$Z = W^T X$$

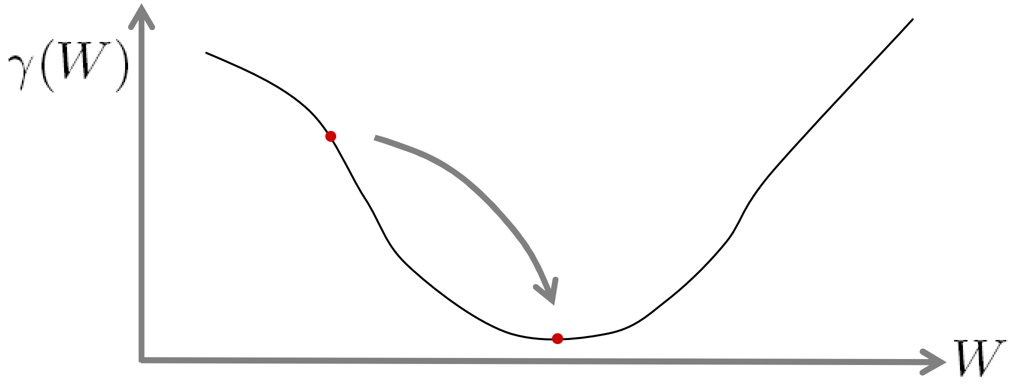
ICA is an optimization problem

1. Estimation of the statistical independence of the *z*'s:

$$\gamma(\cdot) : \mathbb{R}^{n \times p} \longrightarrow \mathbb{R}$$

2. Minimization of the contrast:

$$\min_{W \in \mathbb{R}^{n \times p}} \gamma(W)$$



The contrast presents two inherent symmetries

If $\left\{ \begin{array}{c} z_1 \\ z_2 \end{array} \right\}$ are independent,

then $\left\{ \begin{array}{c} \alpha_1 z_1 \\ \alpha_2 z_2 \end{array} \right\}$ are independent,

and $\left\{ \begin{array}{c} z_2 \\ z_1 \end{array} \right\}$ as well.

The contrast presents two inherent symmetries

If $\left\{ \begin{array}{c} z_1 \\ z_2 \end{array} \right\}$ are independent,

then $\left\{ \begin{array}{c} \alpha_1 z_1 \\ \alpha_2 z_2 \end{array} \right\}$ are independent,

and $\left\{ \begin{array}{c} z_2 \\ z_1 \end{array} \right\}$ as well.

$$\Rightarrow \quad \gamma(W) = \gamma(WP\Lambda)$$

invertible diagonal matrix

permutation matrix

Furthermore, most ICA methods use prewhitening

For any matrix $W \in R^{n \times p}$:

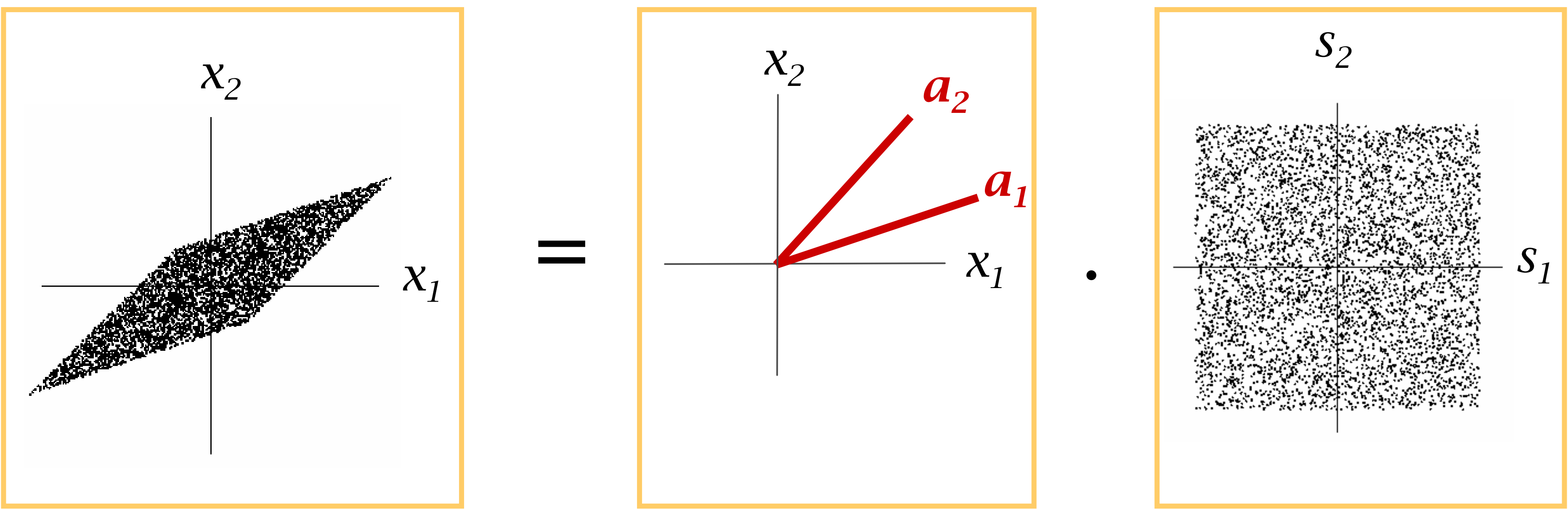
$$W = \boxed{U\,S}\,\boxed{V^T}$$

Determined by
Orthogonal ICA

Determined by SVD of ***X***
Principal Component Analysis

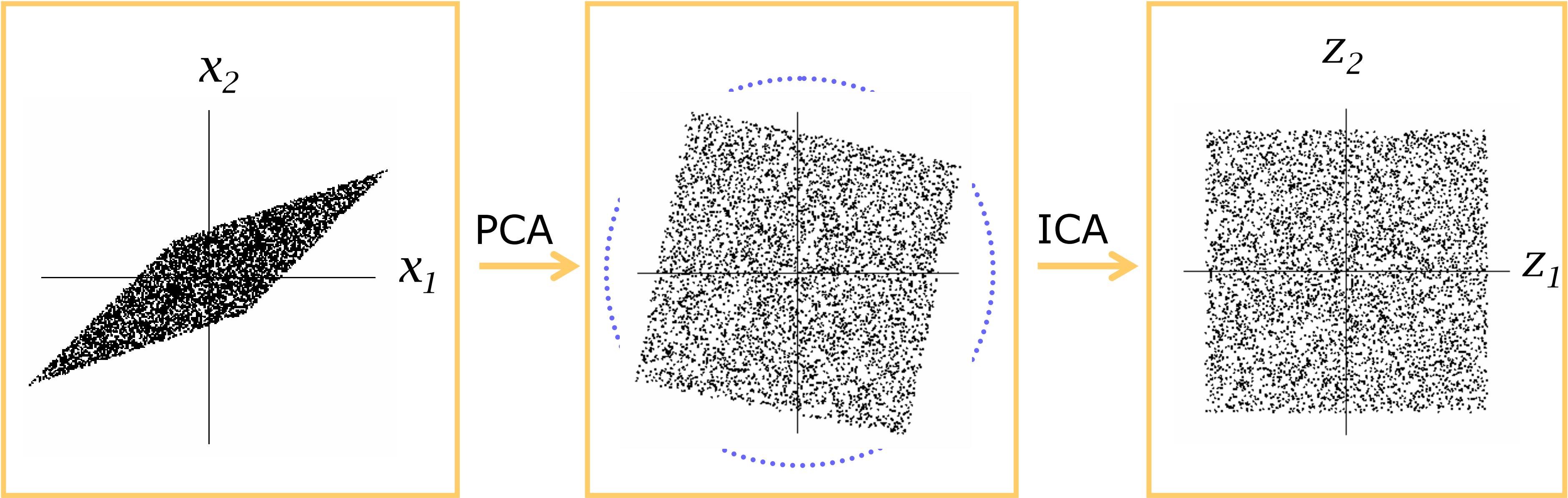
In dimension 2…

$$X = AS$$



In dimension 2…

$$Z = W^T X$$



ICA as an optimization on the orthogonal group

Orthogonal ICA (also called prewhitening-based ICA):

$$\min_{W \in \mathcal{O}_p} \gamma(W) \qquad \text{with} \qquad \mathcal{O}_p = \{Y \in \mathbb{R}^{p \times p} : Y^T Y = I_p\}$$

The orthogonal group automatically gets rid of the scaling indeterminacy.

A whole bunch of ICA algorithms…

Contrast

- Estimation of the mutual information
- Joint diagonalization of cumulant matrices
- Diagonalization of cumulant tensors
- Non-gaussianity
- Constrained covariance

Manifold

- Orthogonal group
- Stiefel manifold
- Oblique manifold
- Flag manifold (independent subspace analysis)

Optimization method

- Jacobi rotations
- Gradient descent
- Second-order approaches

Outline of the project

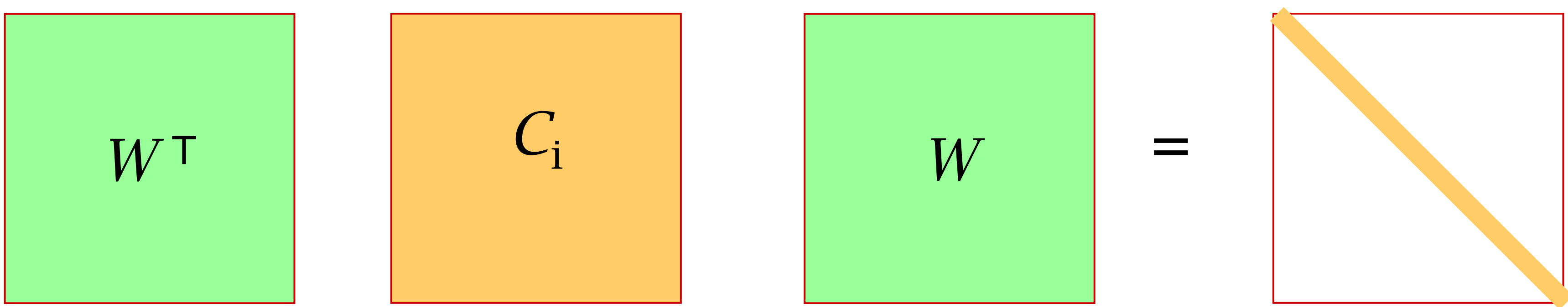
- **Manifold** : Orthogonal group
- **Contrast**: Joint diagonalization of cumulant matrices

Joint diagonalization of a set of matrices

Given m cumulant matrices C_i, minimize

$$\gamma(W) = \sum_{i=1}^{m} \|\mathrm{off}(W^T C_i W)\|_F^2$$

Diagonalization of one matrix:

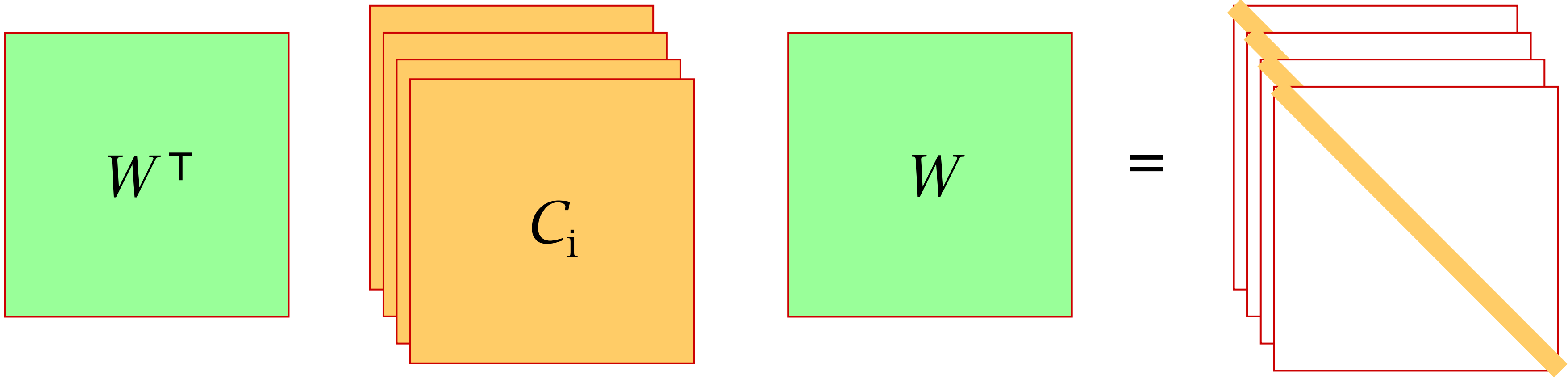


Joint diagonalization of a set of matrices

Given m cumulant matrices C_i, minimize

$$\gamma(W) = \sum_{i=1}^{m} \|\mathrm{off}(W^T C_i W)\|_F^2$$

Joint diagonalization of m matrices:



Outline of the project

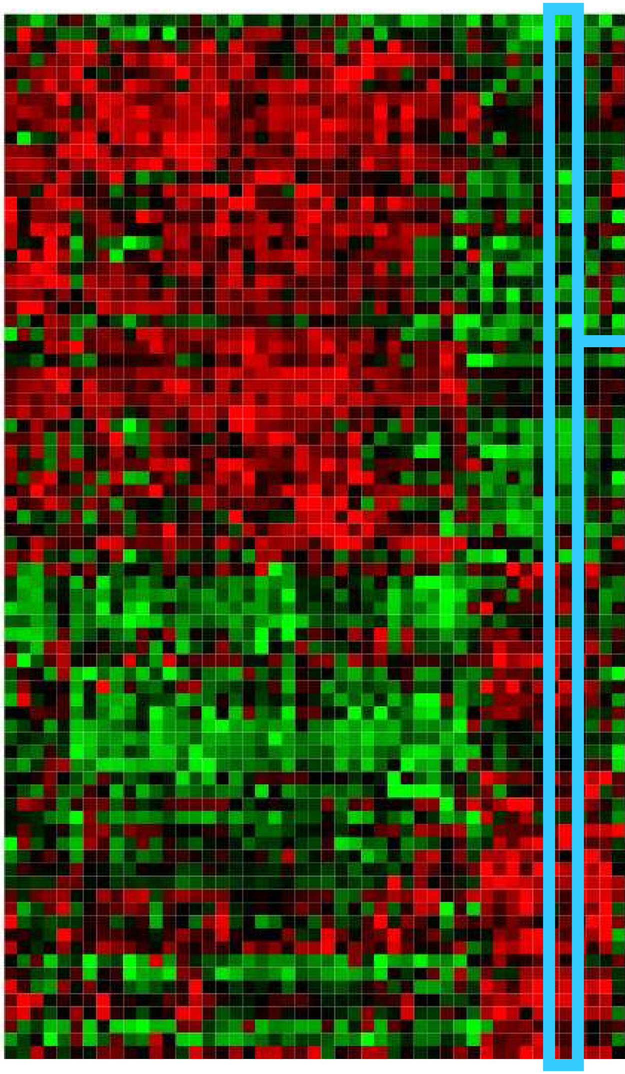
- **Manifold** : Orthogonal group
- **Contrast**: Joint diagonalization of cumulant matrices
- **Optimization method**: conjugate gradient

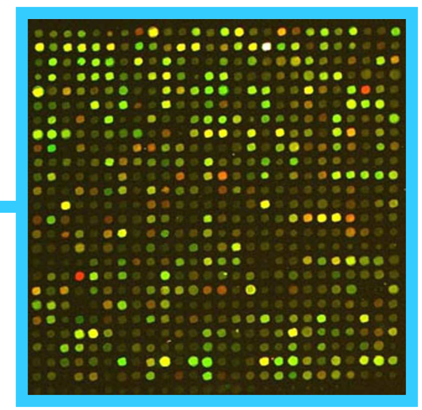
- **Applications**: blind source separation of images, bioinformatics

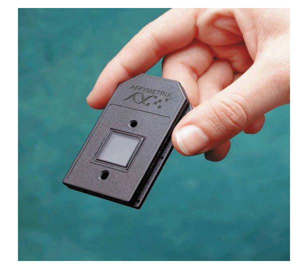
Separation of images

ICA

Analysis of gene expression data







Microarray

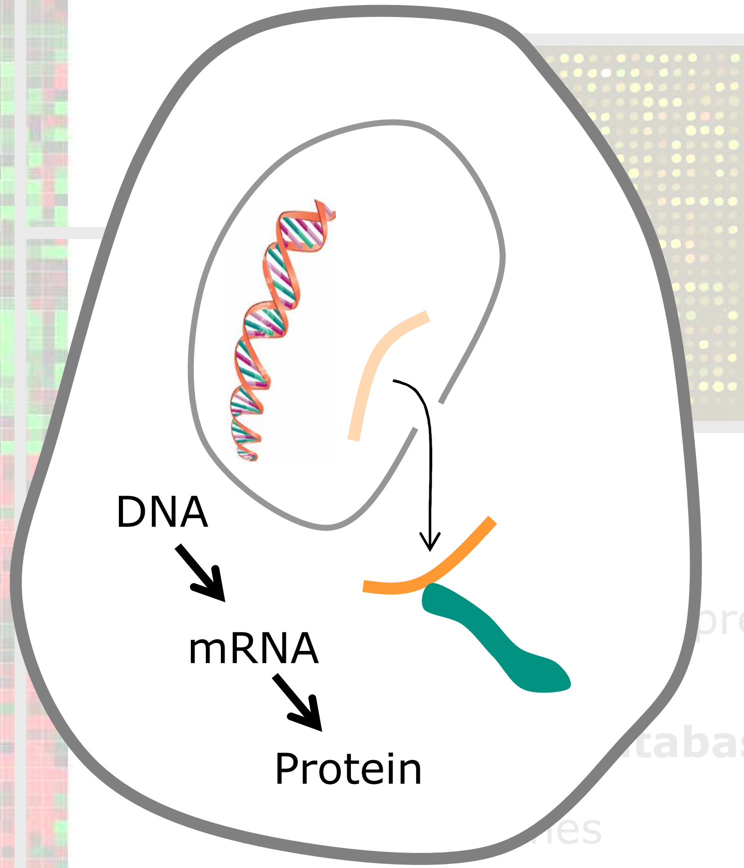
Each spot reflects the expression of a gene

Gene expression database

Rows $\leftrightarrow$ genes ($\sim 10^4$)

Columns $\leftrightarrow$ experiments ($\sim 10^2$)

Analysis of gene expression data



pression of a gene

tabase

les

Columns ↔ experiments

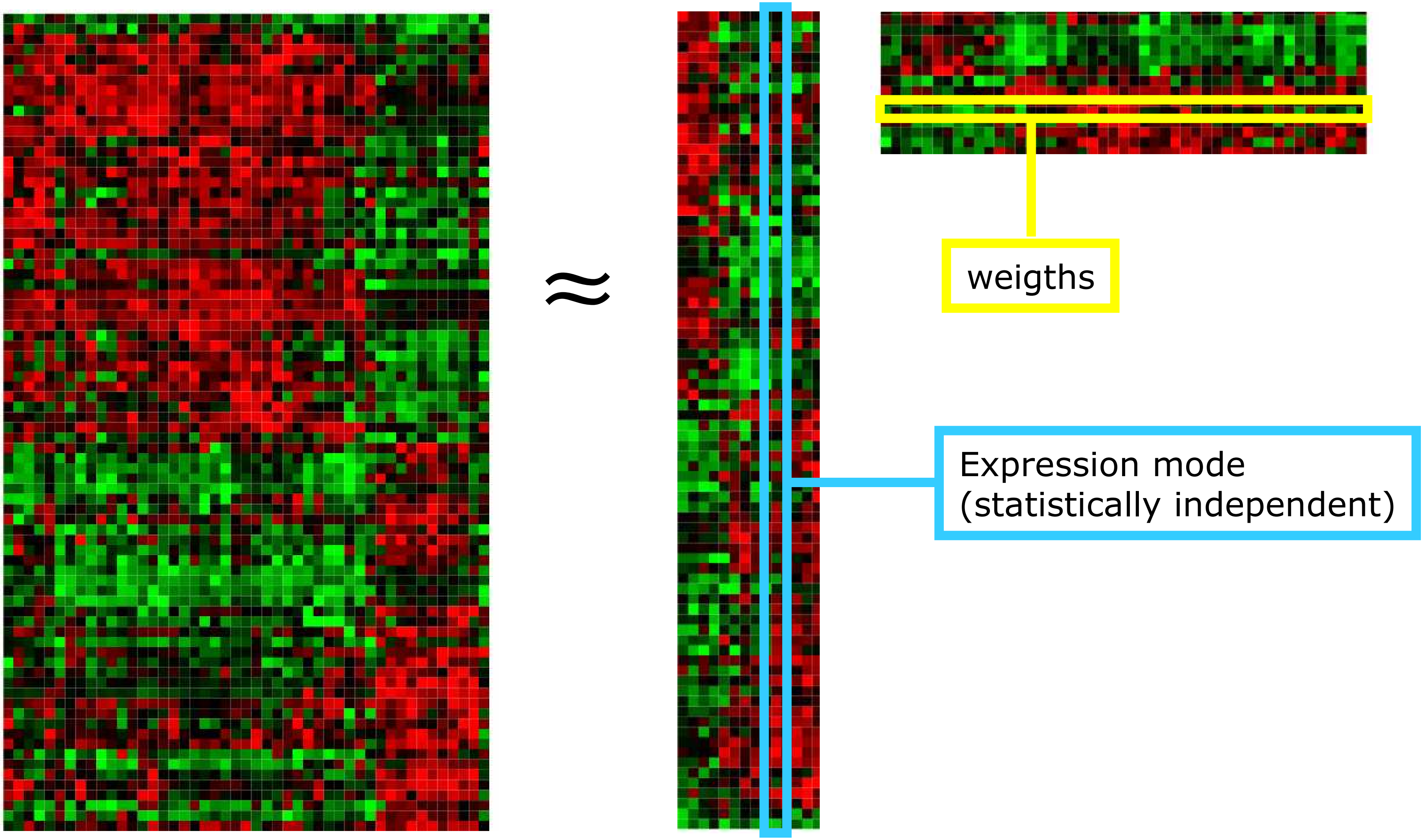
Such a database is a goldmine for new knowledge about the cellular machinery

- **Global** picture of the transcriptome under several conditions
- Genes that are **coexpressed across similar conditions** are very informative
- Identification of interesting structures in the genome

Some interesting questions:

- What does this gene do?
- Which genes are responsible of a phenotype?
- How do the genes act on a phenotype?

ICA in case of gene expression data

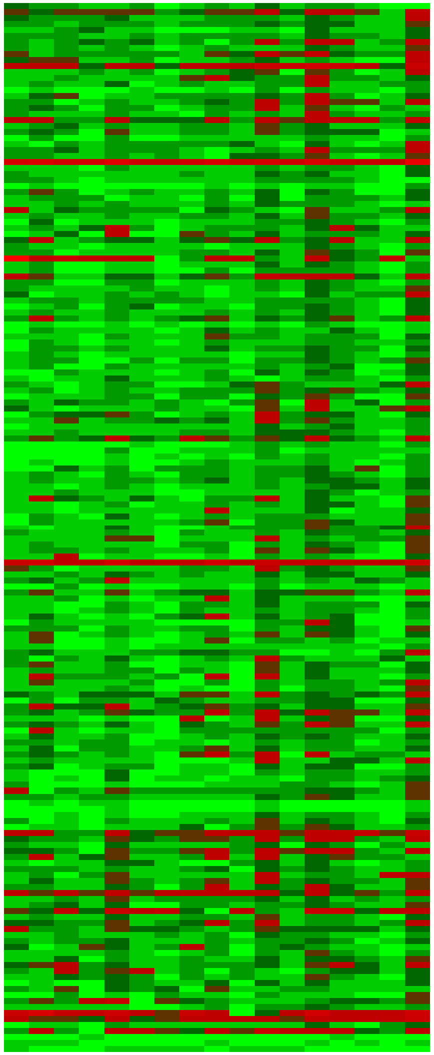


Analysis of an ovarian cancer database

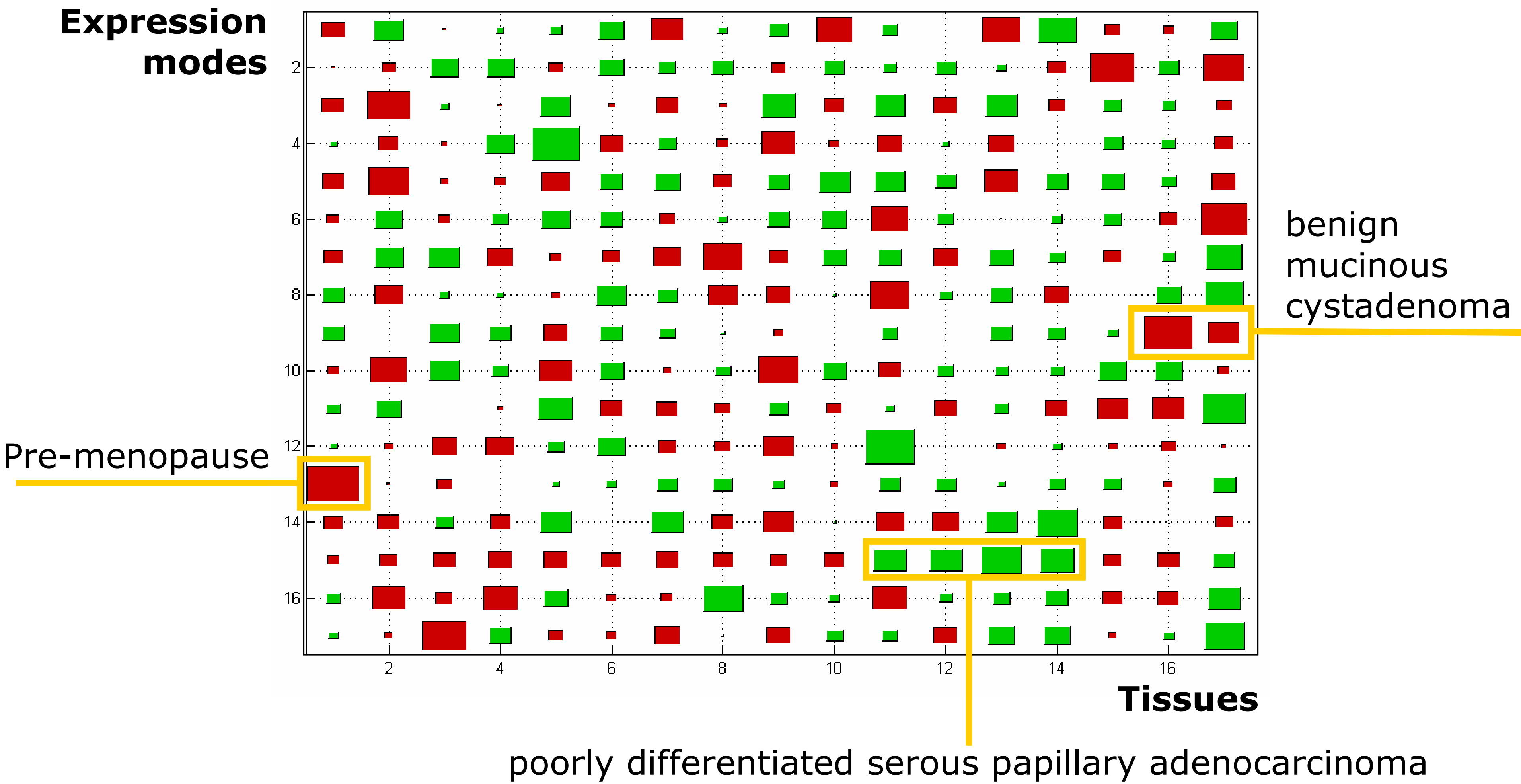
175 genes

17 tissues

+ some clinical data

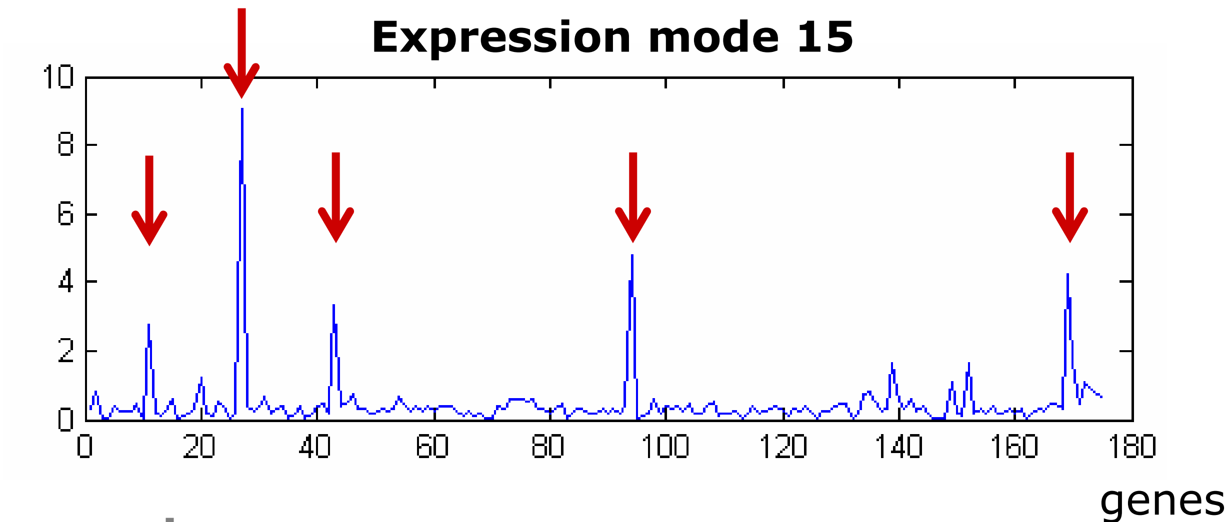


ICA expression modes are highly correlated with the observed phenotypes



ICA identifies genes likely to be coexpressed for an observed phenotype

E.g. poorly differentiated serous papillary adenocarcinoma (pd-spa)



HLA CLASS I
MEMBRANE GLYCOPROTEIN GP130
PLACENTAL-CADHERIN
COFILIN
TIE1

References

- P.-A. Absil and K. A. Gallivan, *Joint diagonalization on the oblique manifold for independent component analysis*, ICASSP 2006, 2006.
- P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, 2008.
- F. R. Bach and M. I. Jordan, Kernel independent component analysis, Journal of Machine Learning Research, 3,1-48, 2003.
- J.-F. Cardoso, *High-order contrasts for independent component analysis*, Neural Computation 11, no. 1, 157–192, 1999.
- P. Comon, *Independent Component Analysis, a new concept ?*, Signal Processing, Elsevier 36, no. 3, 287–314, Special issue on Higher-Order Statistics, 1994.
- A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*, John Wiley & Sons, 2001.
- E.G. Learned-Miller and J.W.Fisher III, *ICA using spacings estimates of entropy*, Journal of Machine Learning Research, 4, 1271-1295,2003.

References

- W. Liebermeister, *Linear modes of gene expression determined by independent component analysis*, Bioinformatics 18, 51–60, 2002.

- A.-M. Martoglio, J. W. Miskin, S. K. Smith, and D. J. C. MacKay, *A decomposition model to track gene expression signatures: preview on observer-independent classification of ovarian cancer*, Bioinformatics 18, no. 12, 1617–1624, 2002.

- A. E. Teschendorff, M Journée, P.-A. Absil, R. Sepulchre, andC. Caldas, *Elucidating the altered transcriptional programs inbreast cancer using independent component analysis*, PLoS Computational Biology 3, Number 8, page 1539-1554, 2007.

Schedule

4 Matlab session:

- Wednesday 11:30 – 12:30
- Wednesday 16:30 – 17:30
- Thursday 16:30 – 17:30
- Friday 11:30 – 12:30

Good work!