CEITEC
Central European Institute of Technology
BRNO | CZECH REPUBLIC

fMRI advanced:

Independent Component Analysis

Ing. Tomáš Slavíček

Brno, November 14th 2016

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007–13
OP Research and
Development for Innovation

Contents

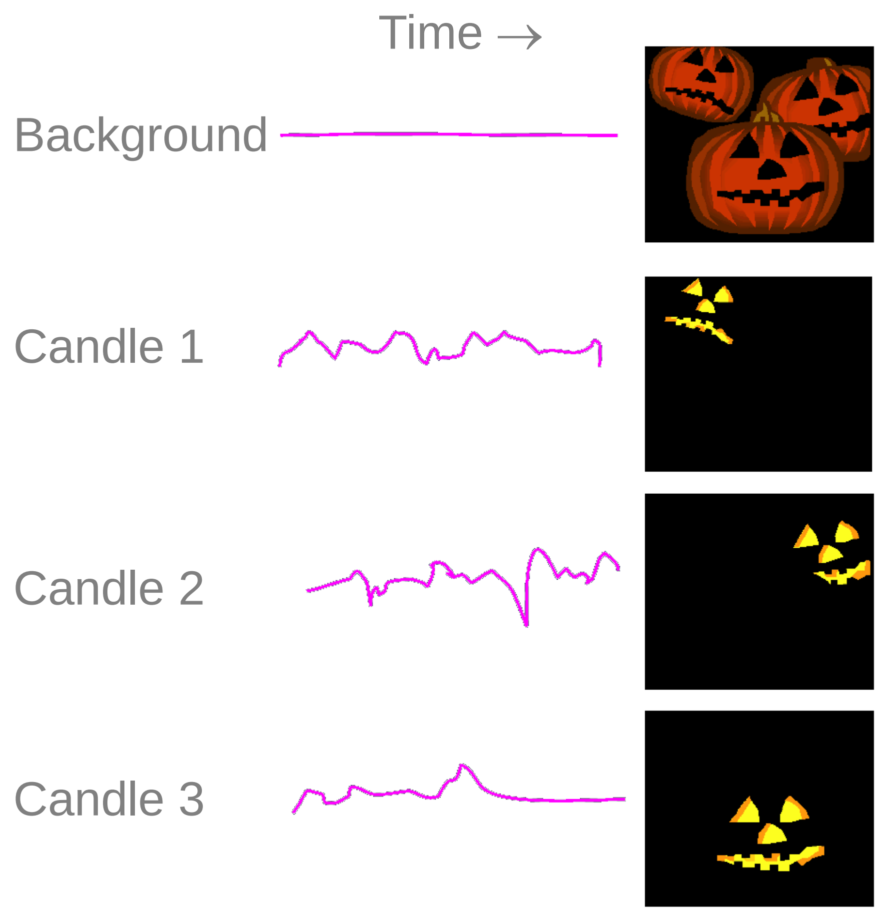
1. Introduction
2. Theory of data driven methods
   - PCA, ICA
3. ICA application to fMRI data
   - spatial ICA
   - data preprocessing
   - ICASSO
4. Results, postprocessing
   - Interpretation, IC classification, ICASSO
5. Useful applications (denoising, functional connectivity)

Introduction

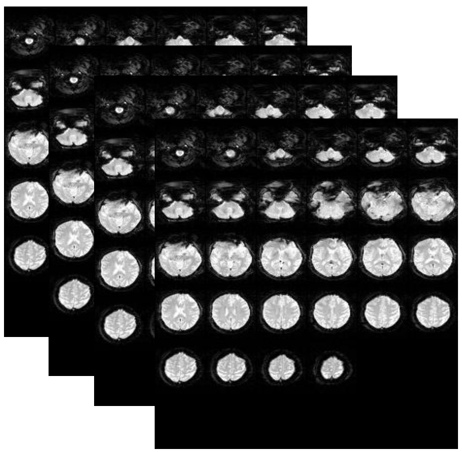
- Independent Component Analysis (ICA)
  - A multidimensional statistical method
  - Allows extraction unknown sources (Blind Source Separ.)
- No need for apriori information about data
  - Useful on Resting-State data
  - Avoid low sensitivity due inaccurate model
  - Avoid multiple statistical testing (as for single dim. methods)
- Limitations
  - Need to estimate # of sources
  - Sources must not have a Gaussian distribution (with exception of one) and be statistically independent (spatial and temporal domain)
  - Direct interpretation problematic

ICA illustration example

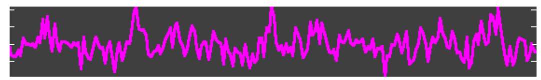
(Calhoun V.D.: Group ICA of FMRI: Introduction and Review of Current Work)

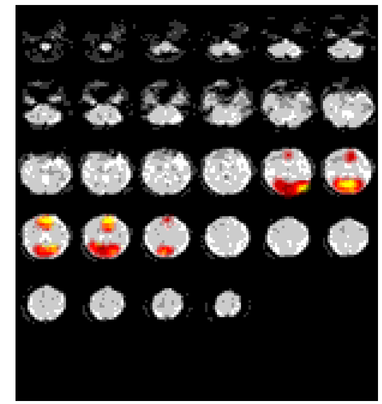


ICA on fMRI data

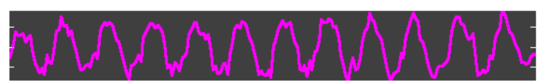


Component 1
default mode network

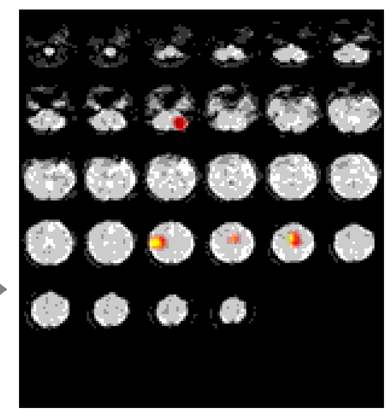




Component 2
Activity in motor cortex

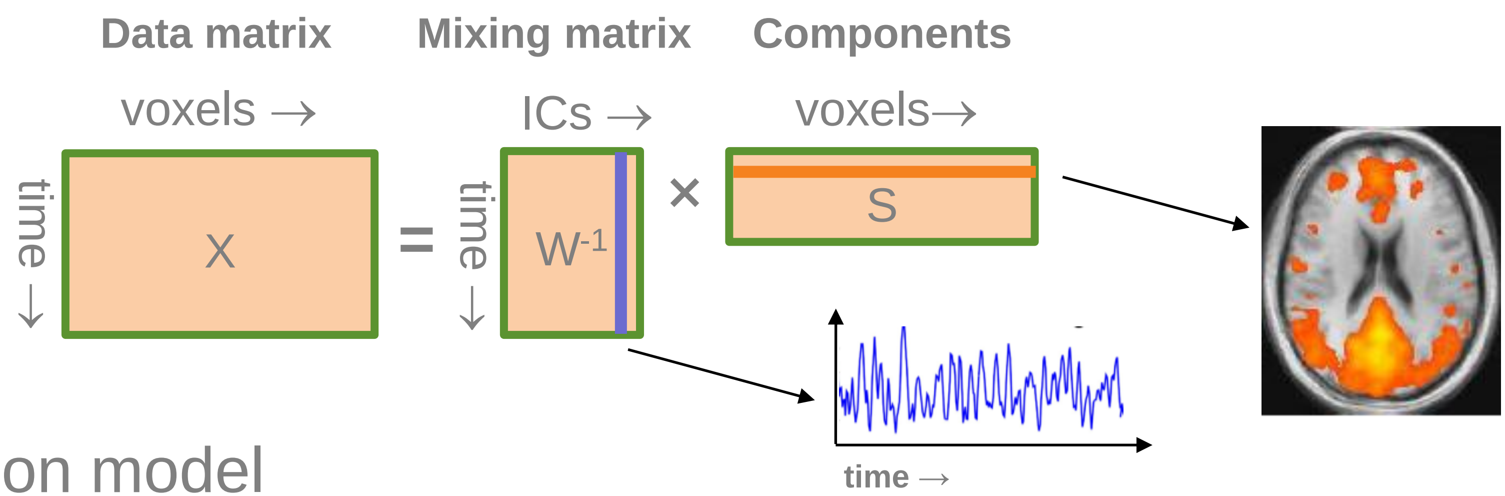


time →



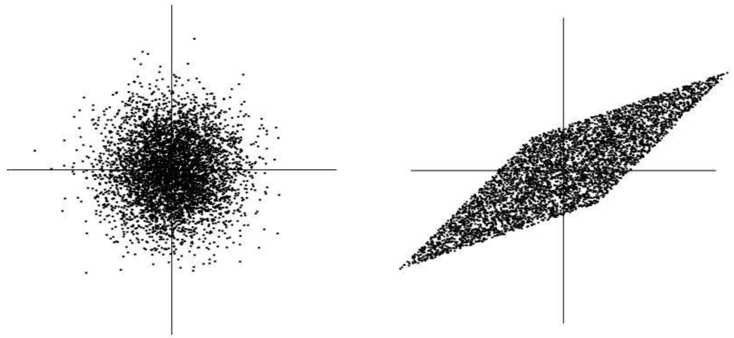
ICA principle

- ICA model - spatial ICA (sICA)
  - **$X = A \cdot S = W^{-1} \cdot S$**
  - Data matrix (X) is decomposed into linear combination of spatially independent sources (S) of variability
  - <u>Assumption that brain activity sources are not overlapping</u>



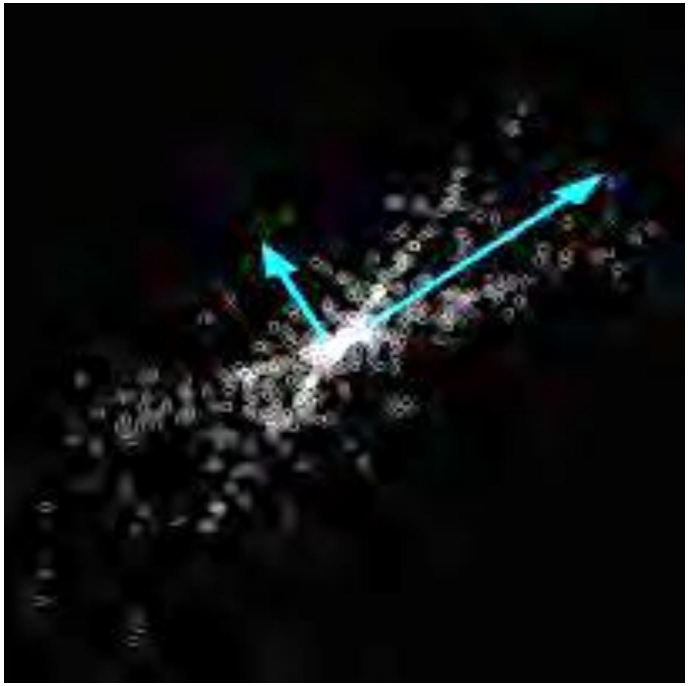
- Inversion model
  - $S = W \cdot X$
  - Estimation of demixing matrix (W) in order to maximize statistical independence of sources (rows of matrix S)

ICA principle

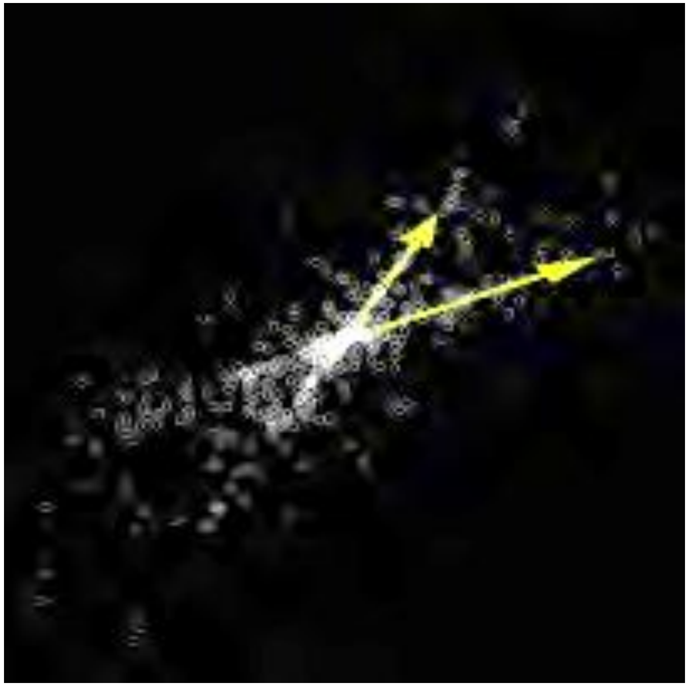


- ICA model
  - $\mathbf{X = A \cdot S = W^{-1} \cdot S}$
- Due ambiguity of right side of the equation
  - We cannot determine order of sources (contrast to PCA)
  - We cannot determine sign
  - Output component's variance is set to 1
- Noise-free model
  - All sources of variability in data are separated into ICs, i.e. brain networks, task-related activity, and also noise, such as movement artifacts, etc.
- Mixing matrix limitations
  - (maximum rank, square)
  - # of components has to be <= # of observations (= timepoints in fMRI)

ICA vs. PCA



PCA finds directions
of maximal variance
(using second order statistics)



ICA finds directions
which maximize independence
(using higher order statistics)

ICA model calculation

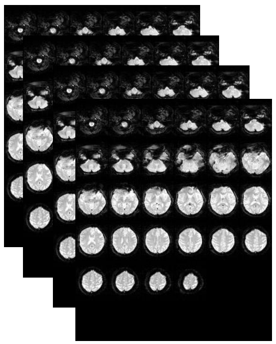
- Estimation of demixing matrix (W)
  - Iterative process

- 1.step – initialization of W (random)
- 2.step – maximization of statistical independence between components (rows of S) according to selected criterion
  - Minimization of Mutual Information
  - Maximization of non-Gaussianity (negentropy)
- Repeat until convergence condition met

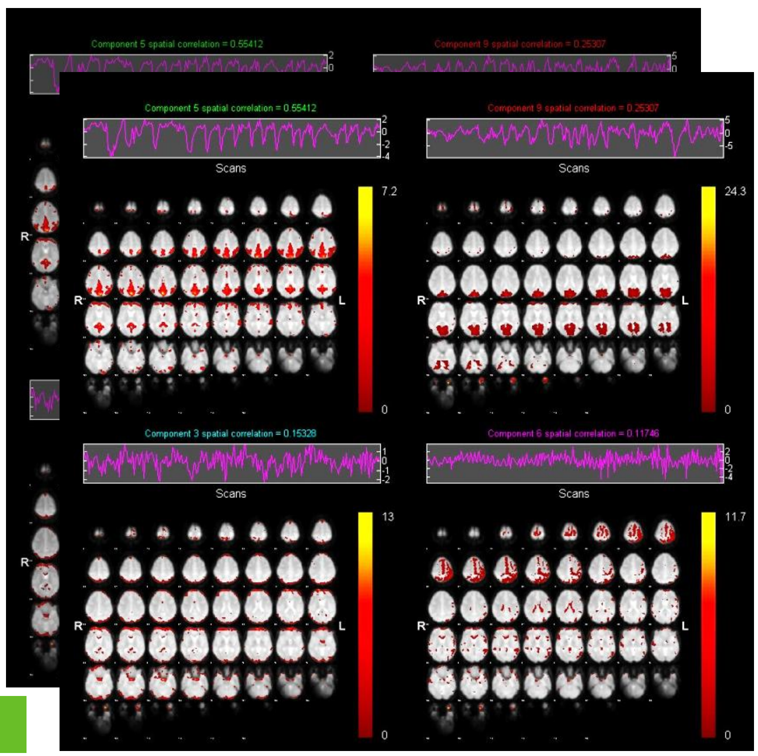
- Many algorithms available
  - Infomax, FastICA, JADE, Amuse, etc.

ICA processing pipeline

Measured data (realigned, normalized and smoothed)



→ Estimating # of components → PCA reduction → ICA calculation



Postprocessing:
- Sorting components
- Selection of interesting components
- Univariate statistical testing
- Interpretation

Estimating # of components/sources

- Each process should be represented by one component

a) Possible to set threshold based on contained variability (PCA)
b) Analytical tool, such as Minimum Description Length (MDL)
    - Information theoretic criterion based on data compression
c) Arbitrary number, usually 15 to 100 ICs for fMRI data

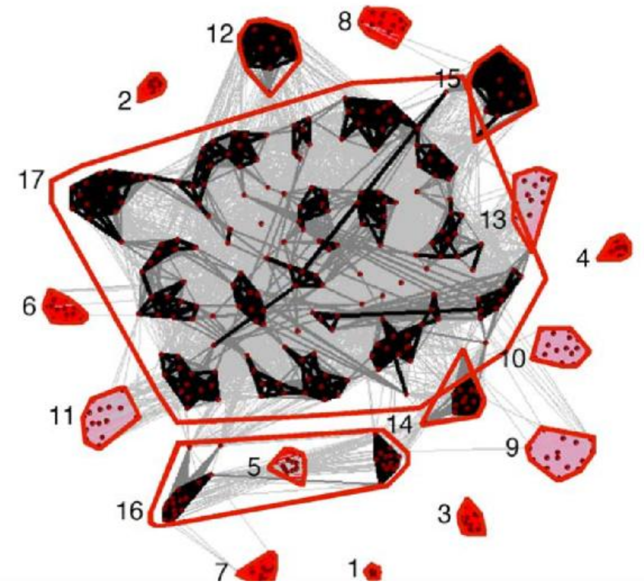
- Too many components
    - Meaningful ICs (such as brain networks) split up,
    - Increasing computational workload and time
- Too few components
    - Merging different sources (overlapping activity)

ICASSO
(J.Himberg, 2003)

- Repeating ICA calculation using
  - Bootstrapping
  - Different initial conditions (mixing matrix)
  - (Both of above)

- Projecting results to multidimensional space
  - Resulting ICs as cluster centroids
  - Cluster size corresponds to stability and reliability of ICA estimates
  - Noise ICs are usually unstable

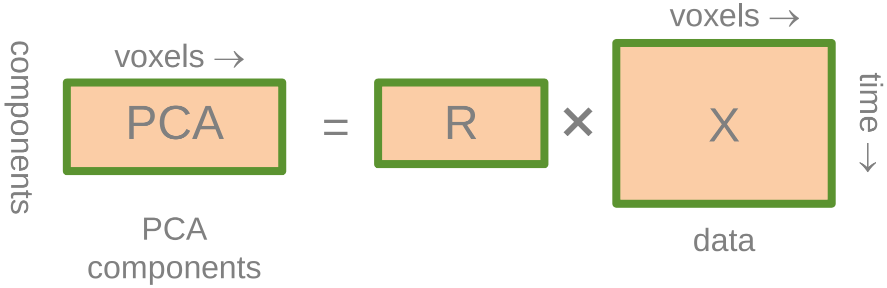
- When increasing # of components, the stability decreases quickly



CEITEC

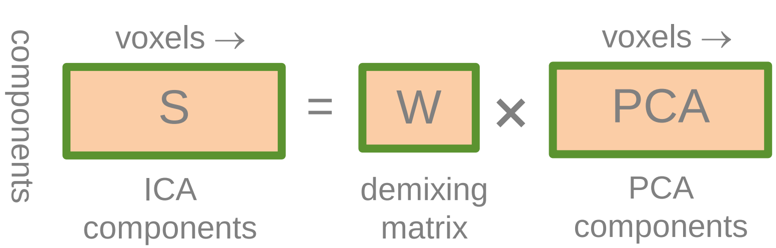
Data reduction using PCA

- Principal Component Analysis (PCA)
  - Reducing dimensionality while preserving maximum variability
  - Whitening (decorrelation) facilitates ICA calculation
  - Useful for group ICA



PCA reduction

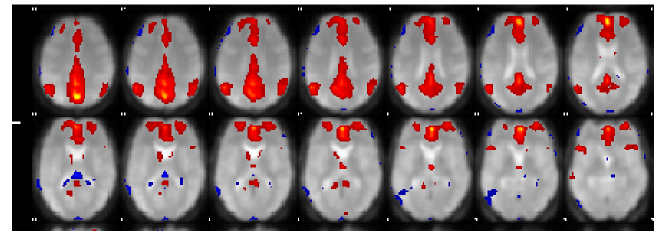
- Matrix R can be found analytically using eigenvectors and eigenvalues of matrix $X*X^T$

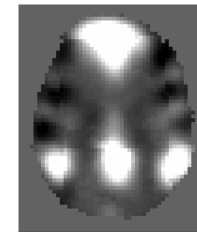


ICA inversion model

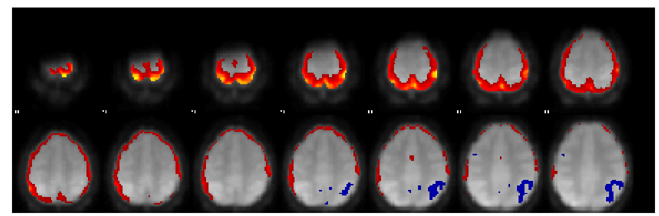
Postprocessing ICs

- Sorting based on spatial distribution
  - Similarity to well-known mask (Default Mode Network, etc.)

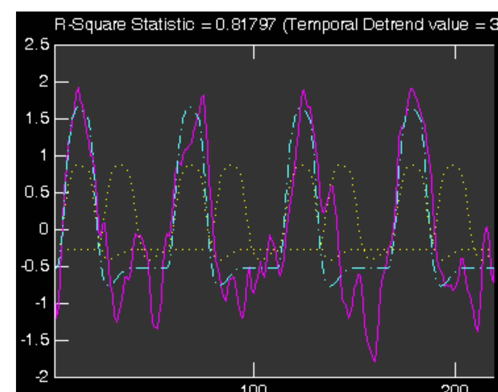




  - Expert evaluation: for example movement artifacts represented as corona shapes on the edges

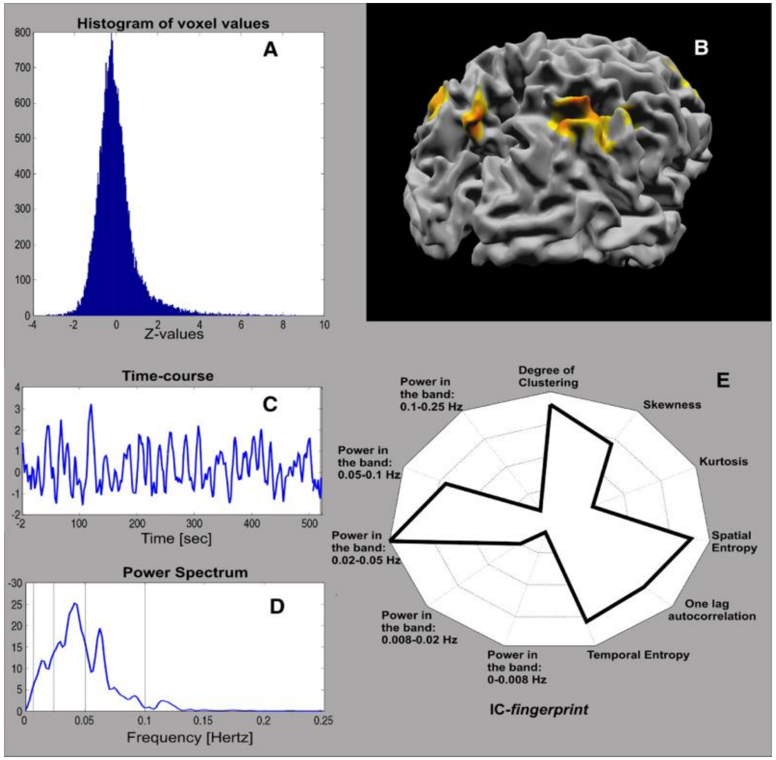


- Sorting based on timecourse parameters (in freq.domain too)
  - Low frequency components ≈ Resting State Networks
  - High freq. components ≈ noise, large vessels artifacts
  - Task-related components based on correlation with stim.timecourse



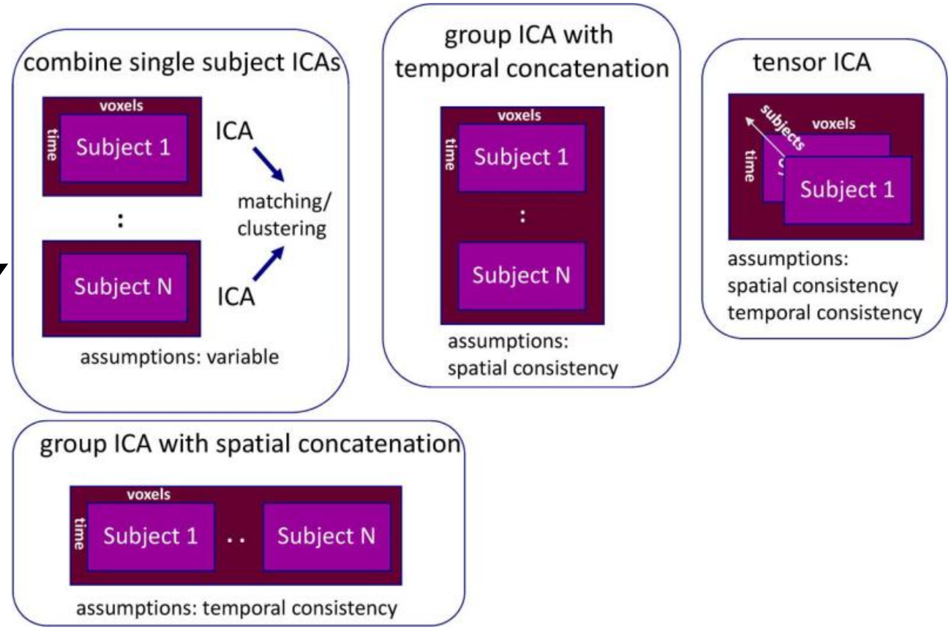
Postprocessing ICs

- Semi-automated IC-fingerprint classificator (F.DeMartino, 2007)
  - Description using 11 measures based both on spatial and temporal parameters (kurtosis,skewness, entropy, one-lag autocorrelation, etc.)
  - Components representing same process have similar „fingerprint“ in parametric space
  - SVM-classificator
  - 6 classes
    - BOLD (both task and non-task)
    - Motion artifacts
    - EPI susceptibility artifacts
    - Vessels and other noise artifacts
  - Need to expert-labelled data
  - Robust results (multiple datasets)
    - About 90% sensitivity for BOLD

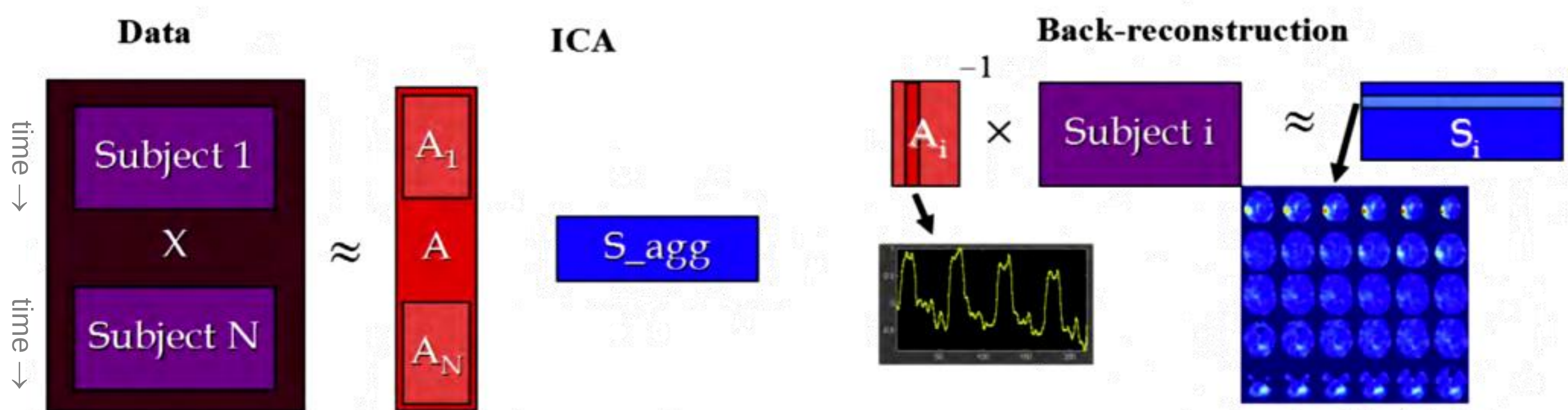


Group ICA

- Group level inference
- Many possible approaches



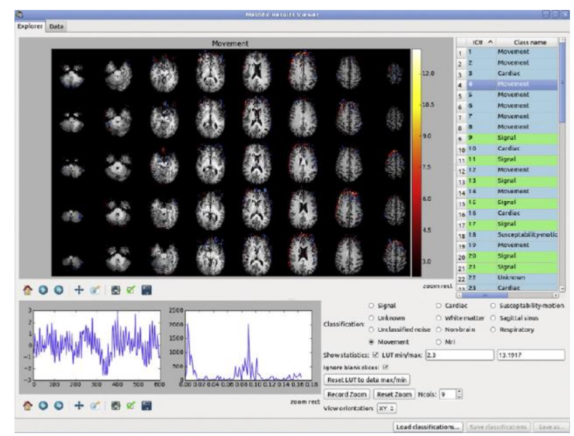
- Temporal concatenation
  - $\mathbf{Y = A \cdot S_{agg}}$ (aggregated comp.)
  - Back reconstruction of individual ICs for all subjects by separation of mixing matrix or dual regression
  - Resulting ICs can be tested for significance (among group) using voxel-by-voxel based T-test, calculation of mean spatial map and time course



Brief overview of software tools

- fMRI data toolboxes that allows ICA
  - GIFT or ICATB (http://mialab.mrn.org/software/gift)
  - MELODIC (part of FSL) (www.fmrib.ox.ac.uk/fsl)
  - BrainVoyager
  - etc.

- fMRI data denoising
  - FIX (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX)
    - uses 180 features (both spatial and temporal)
    - automated classificator
    - very high (95%) overall accuracy



  - Principle: zeroing noise IC’s and reconstruction of data

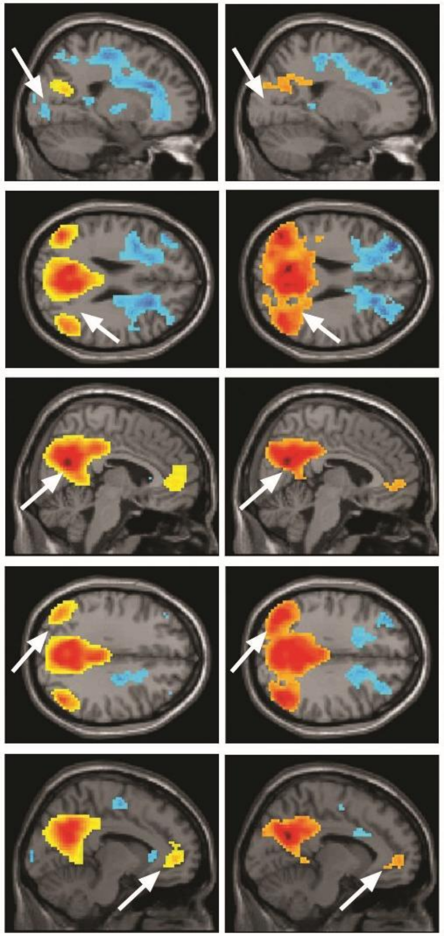
Practical example – prof.Rektorova's study

- 300scan/15min resting state
  - Healthy control group
  - Patients with Alzheimer Disease (AD)

- ICA performed for both groups
- DMN component for each subject selected using comparison with mask
- Statistical testing between groups

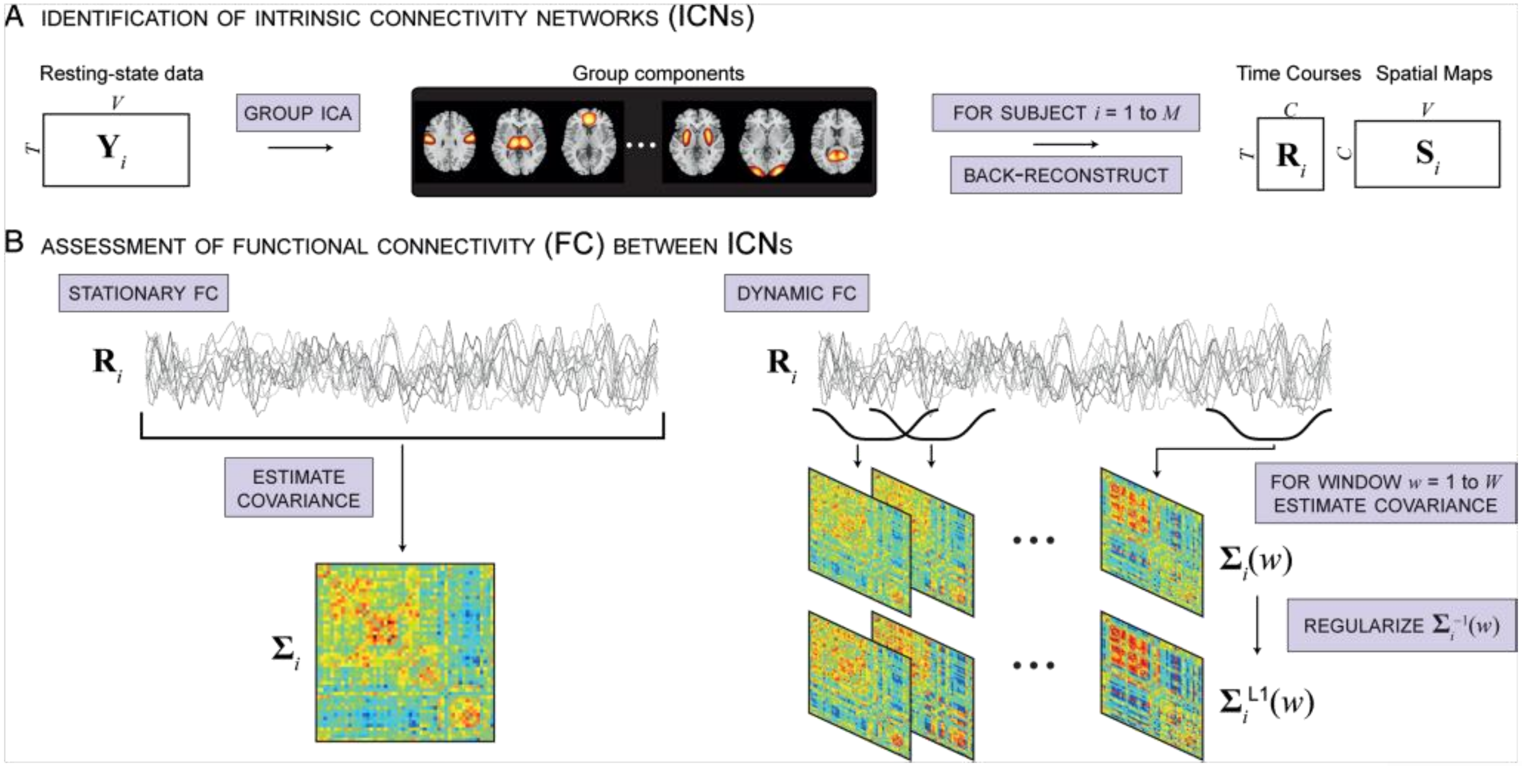
- (arrows points out significant differences)

Healthy controls | AD Patients



ICA in context of functional connectivity
(E.A.Allen, 2012)

- ICA components can be seen as functional networks
- Brain regions represented in one IC shares similar BOLD activity pattern => similar metabolism, neural activity => they are functionally connected
- Investigating connectivity or causality based on comparing IC's timecourses



CEITEC

Thank you for your attention

CEITEC

Central European Institute of Technology
c/o Masaryk University
Žerotínovo nám. 9
601 77 Brno, Czech Republic

www.ceitec.eu | info@ceitec.cz

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007–13
OP Research and
Development for Innovation