Click this field to add text

Automatic artifact reduction in fMRI using ICA and global decision trees

Jussi Tohka
Department of Signal Processing
Tampere University of Technology
Finland

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 39 (2008) 1227–1245

Automatic independent component labeling for artifact removal in fMRI

Jussi Tohka,[a,*] Karin Foerde,[b] Adam R. Aron,[f] Sabrina M. Tom,[b] Arthur W. Toga,[e] and Russell A. Poldrack[b,c,d]

[a]*Institute of Signal Processing, Tampere University of Technology, Tampere, Finland*
[b]*Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA*
[c]*Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, USA*
[d]*Brain Research Institute, University of California Los Angeles, Los Angeles, CA, USA*
[e]*Laboratory of Neuro Imaging, Department of Neurology, University of California Los Angeles Medical School, Los Angeles, CA, USA*
[f]*Department of Psychology, University of California, San Diego, CA, USA*

Outline

- Structured noise in fMRI
- Component analysis based structured noise/artifact reduction
- Automatic artifactual component identification
- Effects to group level GLM/GRF based analysis

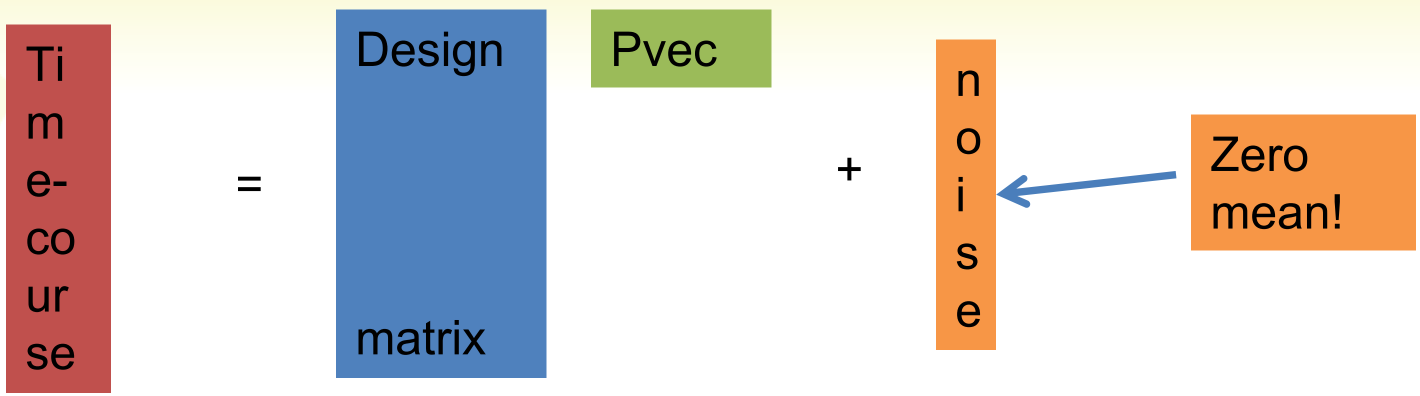
Structured noise in BOLD fMRI

- Various types of noise in BOLD fMRI
  - Gross head motion. Induces also spin history artifacts (Friston MRM 96) and motion by susceptibility interactions (Wu MRI 97)
  - Even after motion correction some motion artifacts remain, both due to inability of motion correction algorithms to correct within scan or nonlinear motion effects and interpolation artifacts (Groontoonk NI 00)
  - Eye and jar movement related artifacts (Beauchamp MRM 03)
  - Physiological noise due to respiration, heart beat, blood flow, etc. (Kruger and Glover MRM 01). Often aliased due to slow TR and aliased noise not restricted to a certain frequency band (Lund NI 07)

- **Important:** Noise/artifacts have complex spatial, temporal, and frequency domain patterns.

Structured noise and GLM

The basic GLM:



We need: 1) the estimate of the parameter vector (OLS) AND
2) the estimate of noise covariance

The structured noise effects both

Noise reduction

- Filtering (spatial, temporal), movement correction, various specialized techniques
- If one knows (or can estimate) the noise/artifact time course, one can add a regressor into design matrix or subtract the artifact time-course from the data
  - RETROKCOR (Hu 95)
  - RETROICOR (Glover MRM 00)
  - Friston MRM 1996
  - Lund NI 2007
- Usually external measurements are needed (E.G. photoplethysmograph and pneumatic belt in RETROICOR)

Noise reduction

- Noise reduction based on component analyses
  - The focus of this presentation

  The 'algorithm':

  1. Decompose the timeseries into linear latent variables models
  2. Detect the those variables/components that are clearly artifact related
  3. Re-Reconstruct the timeseries without clearly artifact related components
- Can be purely image-based
- Can be used in conjunction of GLM or without GLM

Independent Component Analysis in fMRI

- Divide the 4-D timeseries into components consisting of 1-D time series and a 3-D component map so that random variables representing component maps are maximally independent

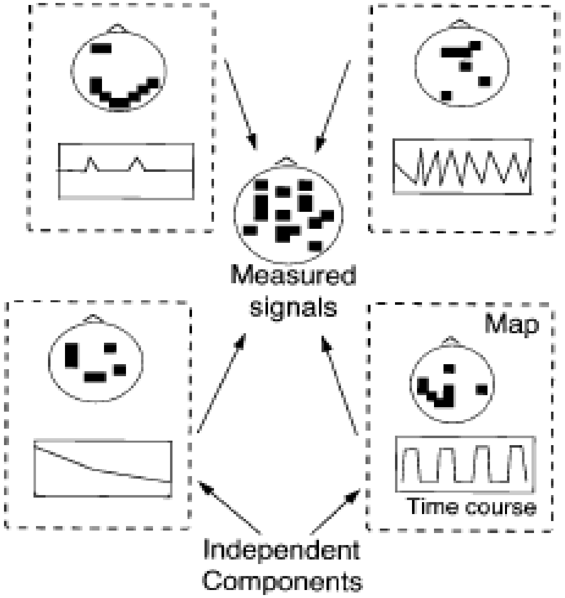
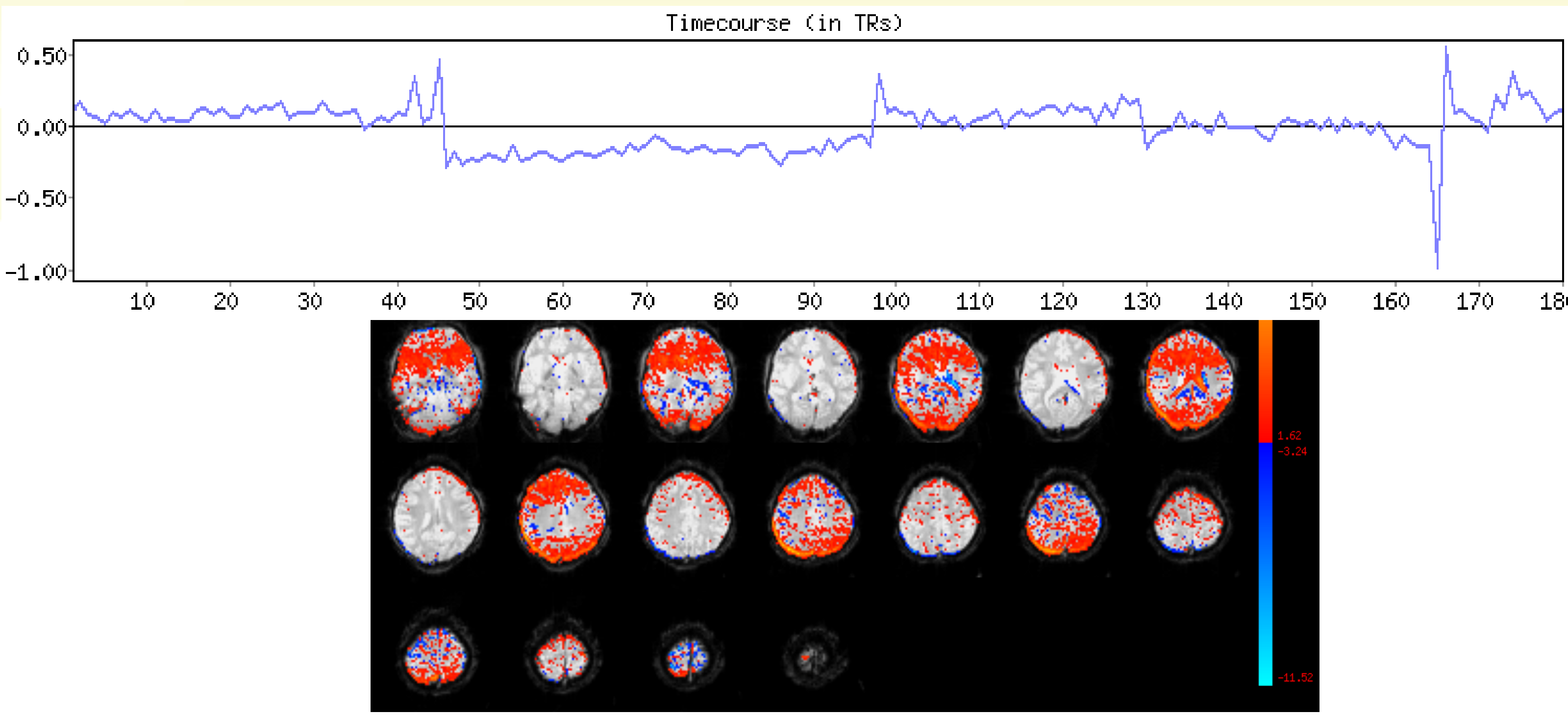


Image from Mckeown et al HBM 1998

ICA based denoising

- Some components are clearly artifact/noise related



ICA based denoising

- $\mathbf{y}_i$ is the ith voxel time series
- C is the matrix of spatial component maps; c_{ij} is the value corresponding voxel i and jth component
- $\mathbf{t}_j$ is the jth component time course, T is the mixing matrix

$$\mathbf{y}_i = \sum\nolimits_{\text{non-artifacts}} c_{ij}\mathbf{t}_j + \sum\nolimits_{\text{artifacts}} c_{ij}\mathbf{t}_j + e$$

$$\mathbf{y}_i^{\text{denoised}} = \sum\nolimits_{\text{non-artifacts}} c_{ij}\mathbf{t}_j + e = \mathbf{y}_i - \sum\nolimits_{\text{artifacts}} c_{ij}\mathbf{t}_j$$

- In practise: $\mathbf{y}_i^{\text{denoised}} = (I - TPT^{+})\,\mathbf{y}_i$,

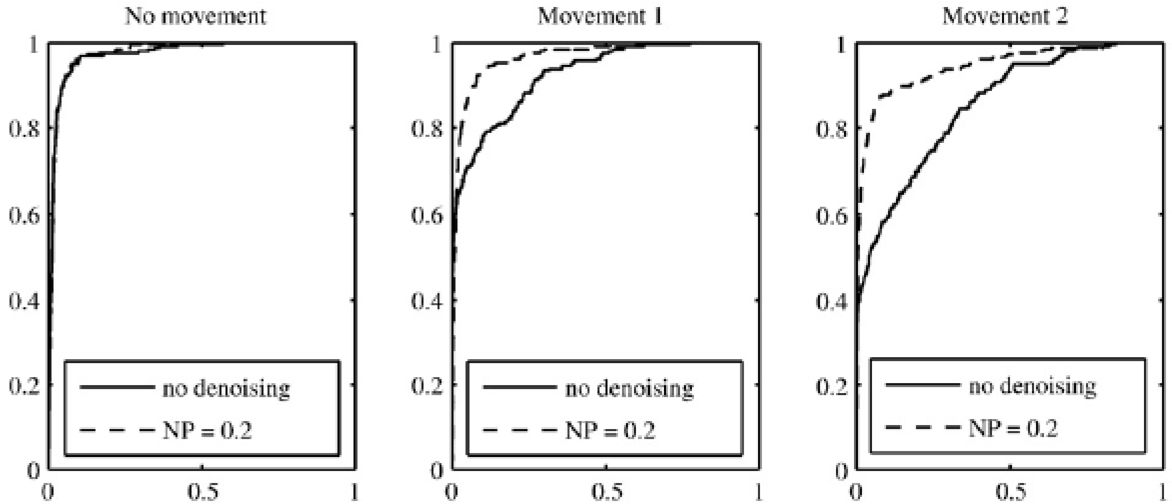
where P is the artifact selection matrix and T^{+} is the Moore-Penrose pseudoinverse of T (FSL's regfilt)

Demonstration

Simulations

Resting state data from FBIRN traveling subjects database; A simulated activation pattern overlaid to the resting state data and finally added some simulated motion artifacts (interleaved acquisition)

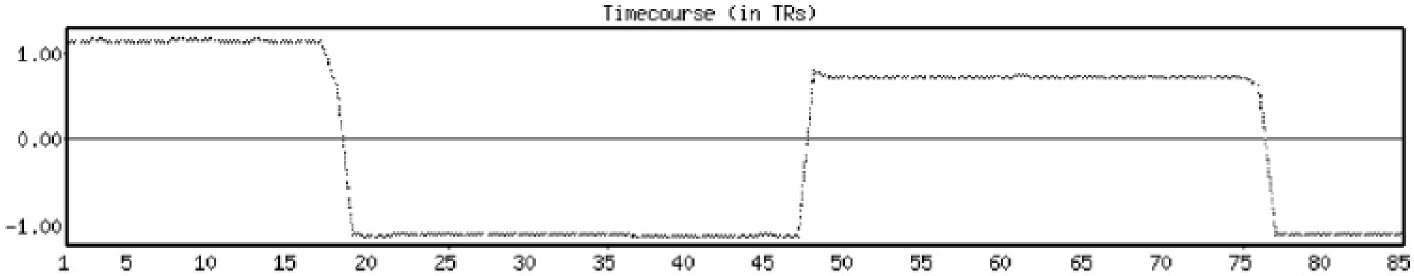
Sensitivity



1 - specificity

Demonstration

ICA (FSL's Melodic, Beckmann IEEE-TMI 04) picked nicely components corresponding to motions

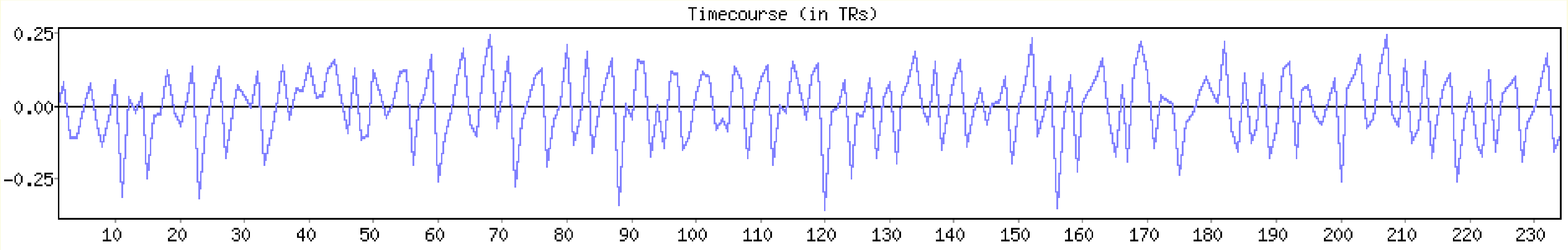


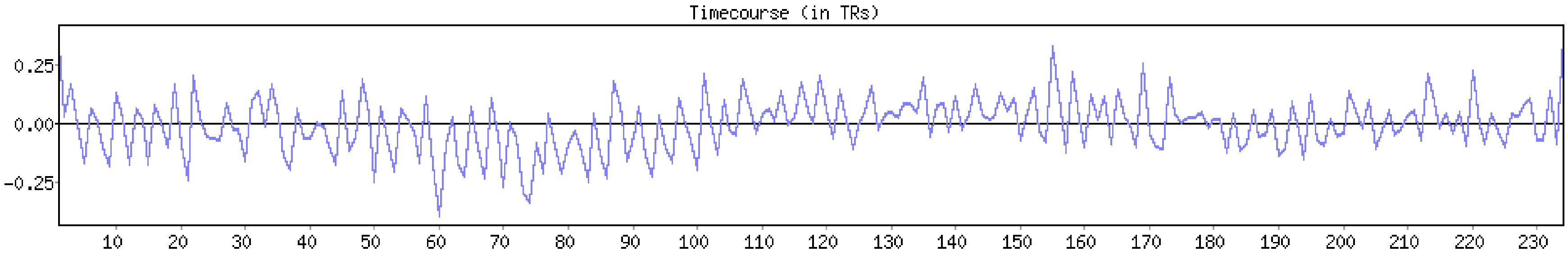
Note: the estimated number of ICA components dropped from 44 to 7 when the simulated motion was added (4 of those 7 could be attributed to motion)

-> bad artifacts mix up melodic's component number estimation approach, and probably others as well

Some components are borderline components

Component time course #2 (artifact)





Component time course #42 (not artifact)

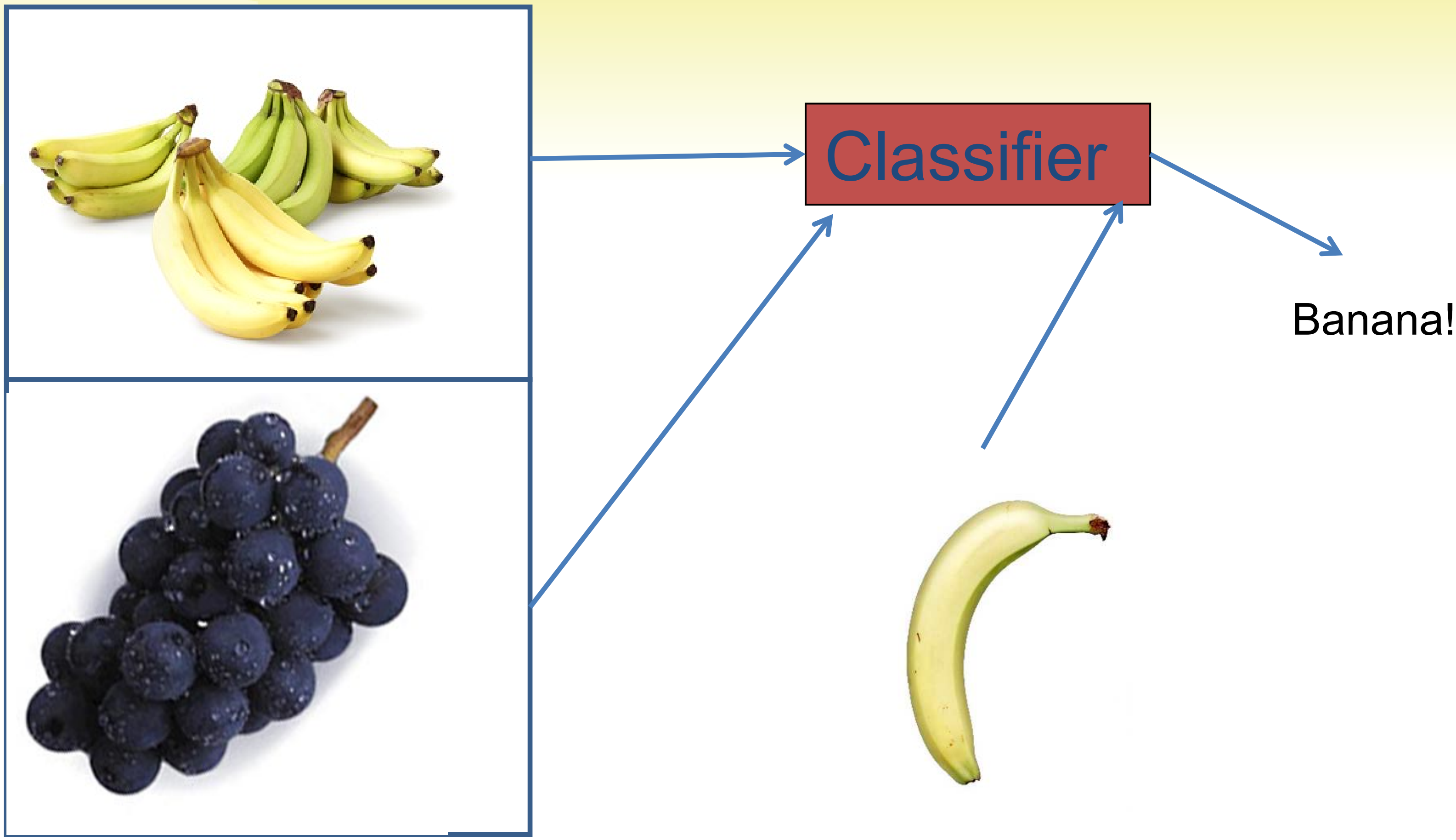
Component selection: Other works

- It is possible to manually pick the components that are clearly artifact related (e.g. Foerde PNAS06, FEAT user guide)
  - Time consuming
  - Subjective
- Automatic approaches:
  - McKeown NI 2000: Hybrid component analysis (adding regressors to GLM)
  - Thomas NI 2002: Short TR cardiac noise and respiration noise removal (visual checkerboard stimulus)
  - Kochiyama NI2005: Task related motion removal (motor task)
  - Perlbarg MRI 2007: Global and respiration related movements and local cardiac fluctuations (motor task)
  - Beall NI 2007: RETROICOR + ICA (resting-state)
  - Sui NI 2009

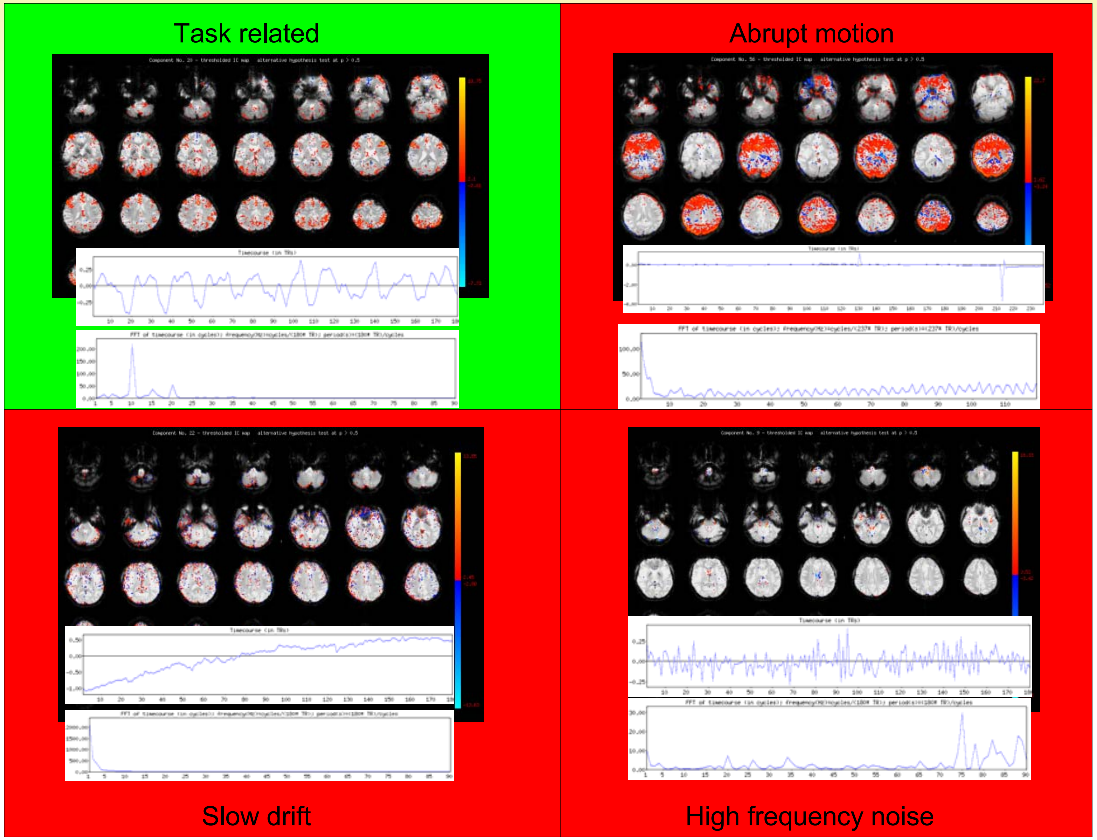
ICA based denoising: Our approach

- Perform ICA on motion corrected and filtered data
- Train a supervised classifier based on manually classified exemplars to detect obvious artifact/noise related component
- Thereafter, the classification can be performed automatically
- Our aim was in improving the accuracy and reliability of GLM-based group level analyses.
- However, the method applies no knowledge of the behavioral paradigm -> could be useful in resting state

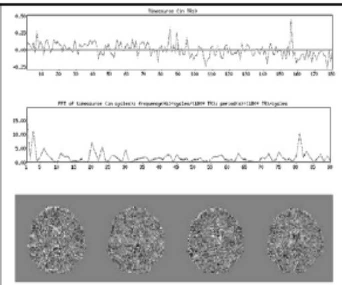
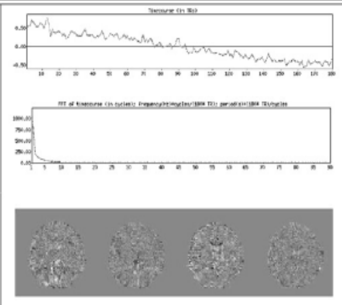
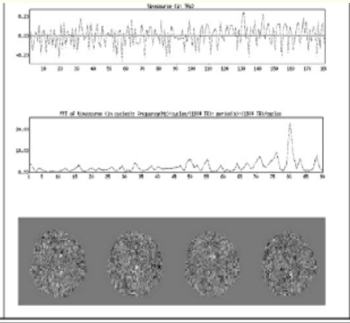
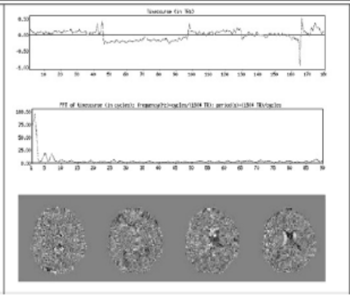
Supervised classification



ICA based denoising: Noise classes



Manual classification procedure

1		Component timecourse varies due to large range of frequencies Little/no peak at task frequencies for the power spectrum with peaks across the entire spectrum range ; Component maps are noisy;
2		Component timecourse drifts slowly, sometimes with an abrupt spike in motion; Small peak at the task frequencies swamped by large peaks at very low frequency on the power spectrum; Component maps show ringing around the brain or alternating patterns of activity for every other slice (in data acquired using interleaved acquisition);
3		Component timecourse shows high frequency activity that does not look task related. Little/no peak at the task frequencies for the power spectrum with peaks across most of the range, but centered mostly in the high frequency range Component maps are noisy;
4		Timecourse shows abrupt spike which swamps other variation several times. May have peak(s) at the task frequency, present with other peaks across frequency range. Component maps look artifactual, specifically they will most likely show alternating patterns of activity for every other slice.

Pattern Classification Systems

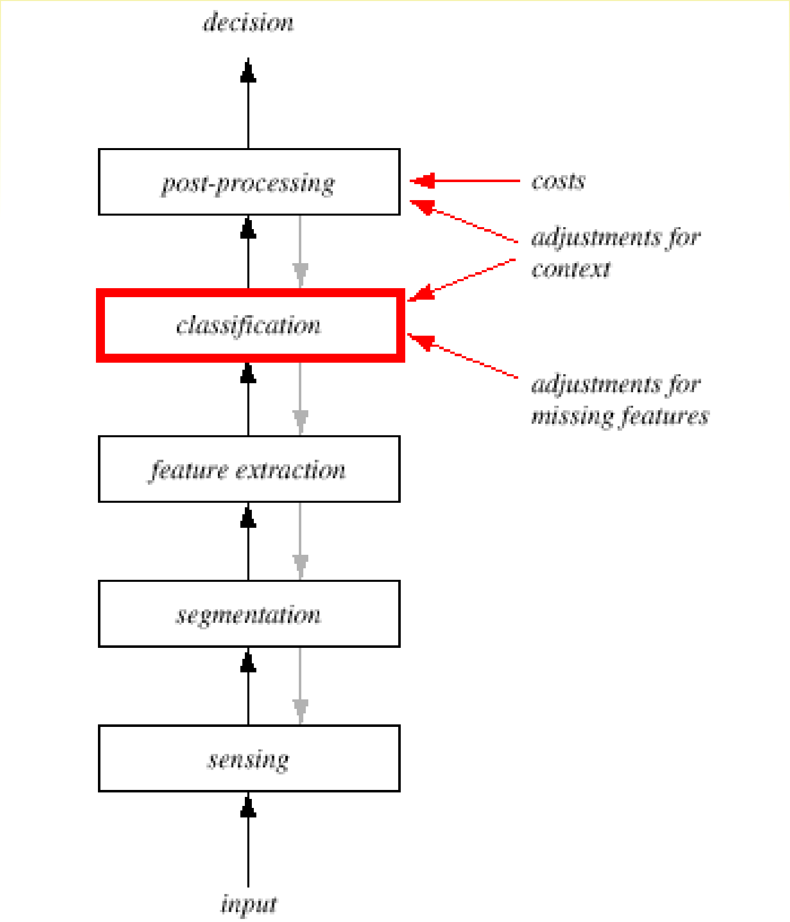


Image from Duda, Hart, Stork: Pattern Classification, 2nd edition

Classifier training

- For each component, compute six features characterising the admissibility of the component
- 2 features based on PSD of the component time course, 2 features based on component time course itself, 2 features based on component map
- Train a Global decision tree classifier based on manually classified examples
  - Simple decision regions, good generalization rate, imitates manual procedure
- Training is done in Neyman Pearson setting
  - Implications: False alarm rate can be controlled, Divide and conquer/exhaustive search algorithm is ok.

Features

$$f_1 = \frac{\sum_{\omega_j \in \Omega_{\text{target}}} p[\omega_j]}{\sum_{\omega_j \in \Omega_{\text{target}}} p[\omega_j] + \sum_{\omega_j \in \Omega_{\text{low}}} p[\omega_j]}.$$

$$f_4 = -\frac{|\sum_i \mathrm{Var}[c^{2i-1}] - \sum_i \mathrm{Var}[c^{2i}]|}{\sum_i \mathrm{Var}[c^i]}$$

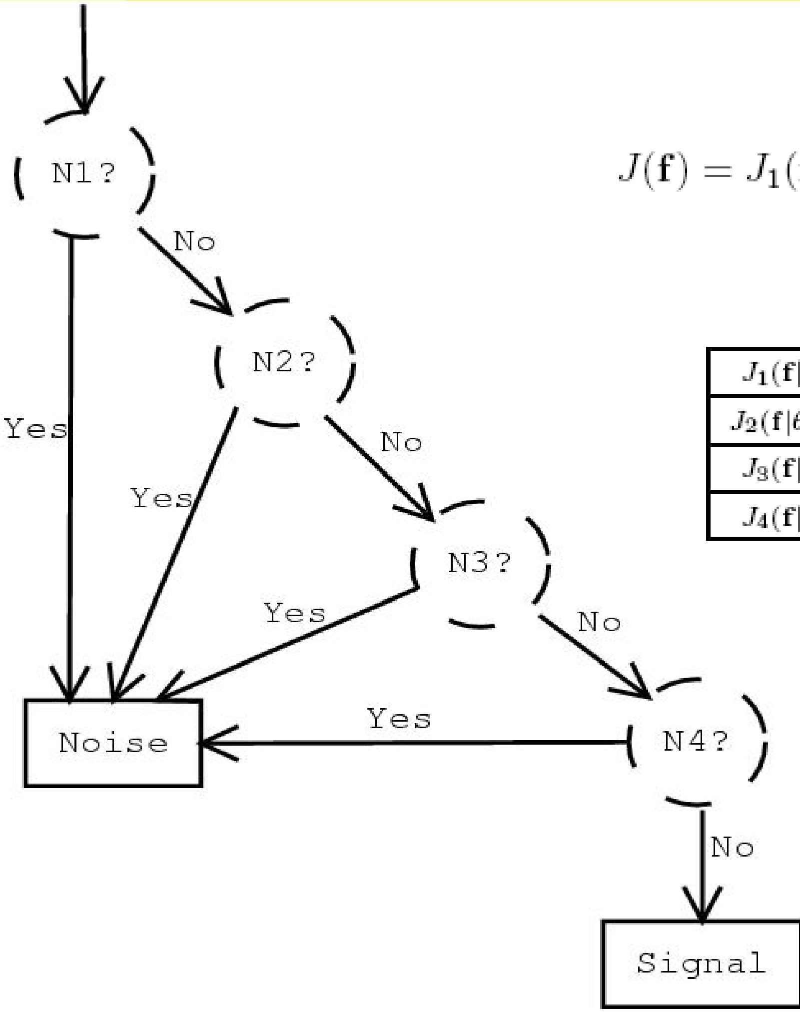
$$f_2 = \frac{\sum_{\omega_j \in \Omega_{\text{target}}} p[\omega_j]}{\sum_{\omega_j \in \Omega} p[\omega_j]}.$$

$$f_5 = -\frac{\max_j |t[j] - t[j-1]|}{\sum_{j=2}^{T} |t[j] - t[j-1]| - \sum_{j=\text{amax}-2}^{\text{amax}+2} |t[j] - t[j-1]|},$$

$$f_3 = \frac{\mathrm{Var}\{c[i] : i \in B\} - \mathrm{Var}\{c[i] : i \in \partial B\}}{\mathrm{Var}\{c[i] : i \in B\}},$$

$$f_6 = \frac{T \sum_{j=2}^{T} t^*[j]t^*[j-1]}{(T-1)\sum_{j=1}^{T} t^*[j]t^*[j]}.$$

ICA based denoising: Classifier



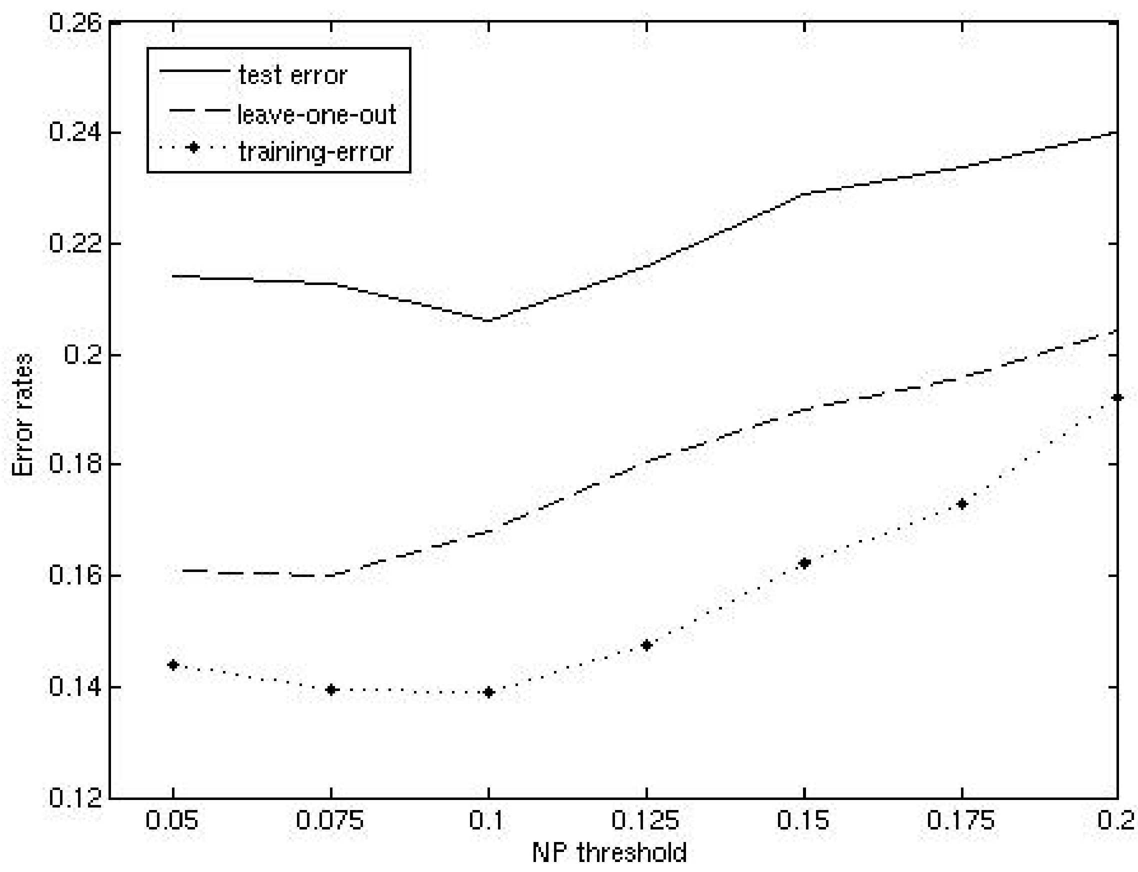
$$J(\mathbf{f}) = J_1(\mathbf{f}|\theta_1) \vee J_2(\mathbf{f}|\theta_2) \vee J_3(\mathbf{f}|\theta_3) \vee J_4(\mathbf{f}|\theta_4),$$

$J_1(\mathbf{f}\|\theta_1) = (f_2 < \theta_{12}) \wedge (f_4 < \theta_{14}) \wedge (f_6 < \theta_{16})$
$J_2(\mathbf{f}\|\theta_2) = (f_1 < \theta_{21}) \wedge ((f_4 < \theta_{24}) \vee (f_3 < \theta_{23}))$
$J_3(\mathbf{f}\|\theta_3) = (f_2 < \theta_{32}) \wedge (f_4 < \theta_{34}) \wedge (f_6 < \theta_{36})$
$J_4(\mathbf{f}\|\theta_4) = (f_4 < \theta_{42}) \wedge (f_4 < \theta_{44}) \wedge (f_5 < \theta_{45})$

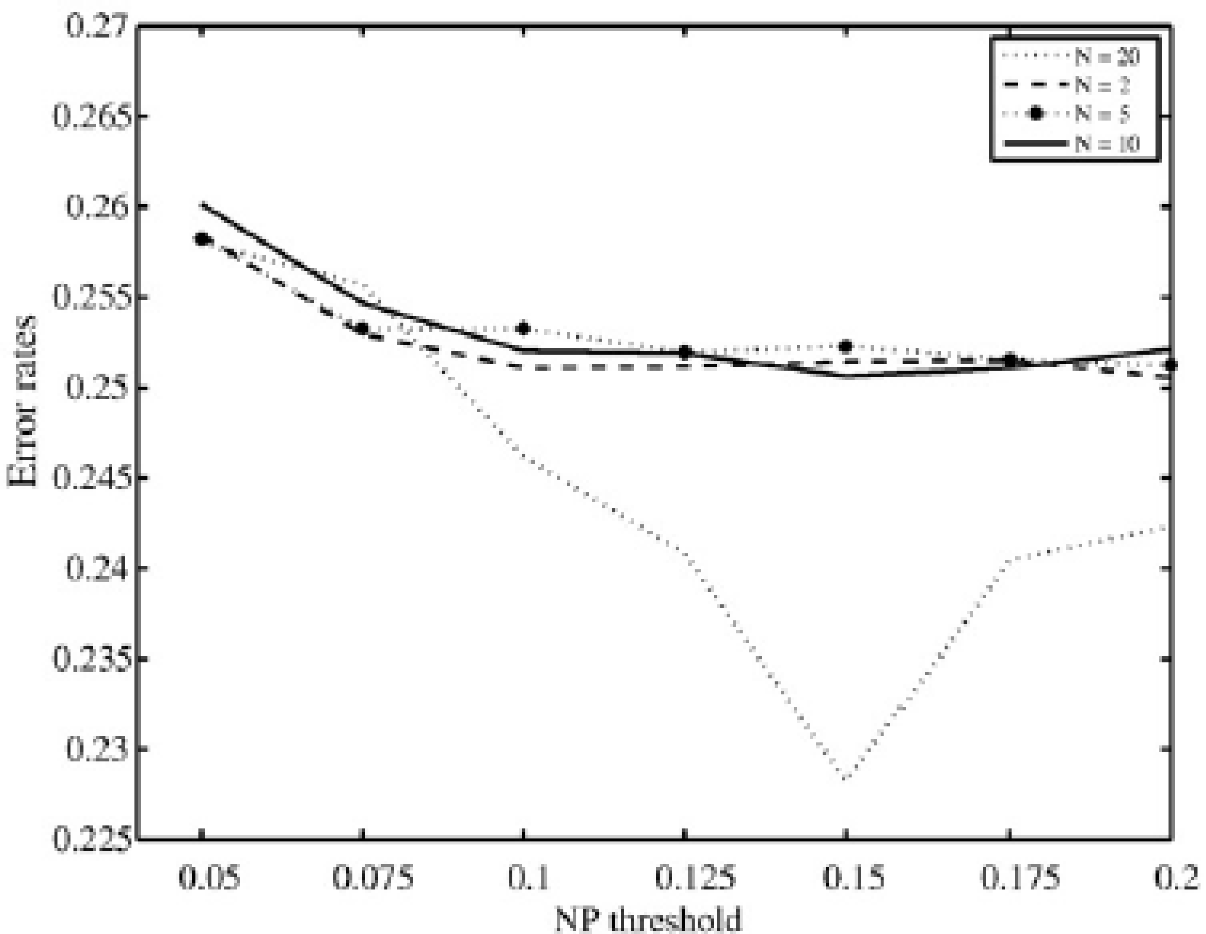
Experiments: Material

- Data from category learning tasks, both block design and event related design
  - 3T Siemens Allegra head only scanner, interleaved acquisition, gradient each echo planar pulse sequence (TR 2000 ms, TE 30ms, 64 x 64 x 25/30 image size, 3mm x 3mm x 4mm voxel size)
- Training Set
  - Used for classifier training, 20 subjects, 4 blocked design and 2 event related design runs per subject, total 5321 ICA components hand classified to generate training data
- Test Set
  - NOT used for classifier training. 12 subjects. 5 blocked design and 1 event related design probe run per subject. ICA components hand-classified for evaluation purposes
- Simulations
  - Resting state data from FBIRN traveling subjects database; A simulated activation pattern overlaid to the resting state data and finally added some simulated motion artifacts (interleaved acquisition)

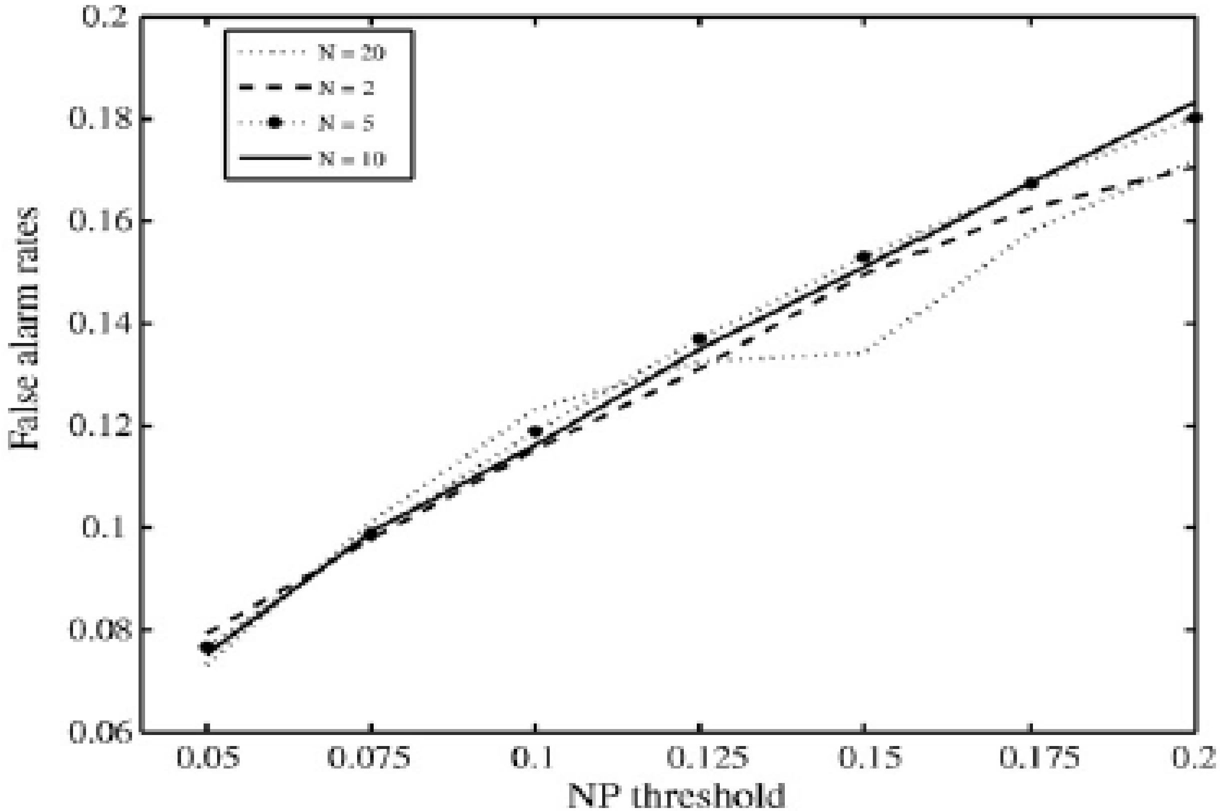
Results: training vs. test errors (event related design)



Training set size with test data (blocked design)

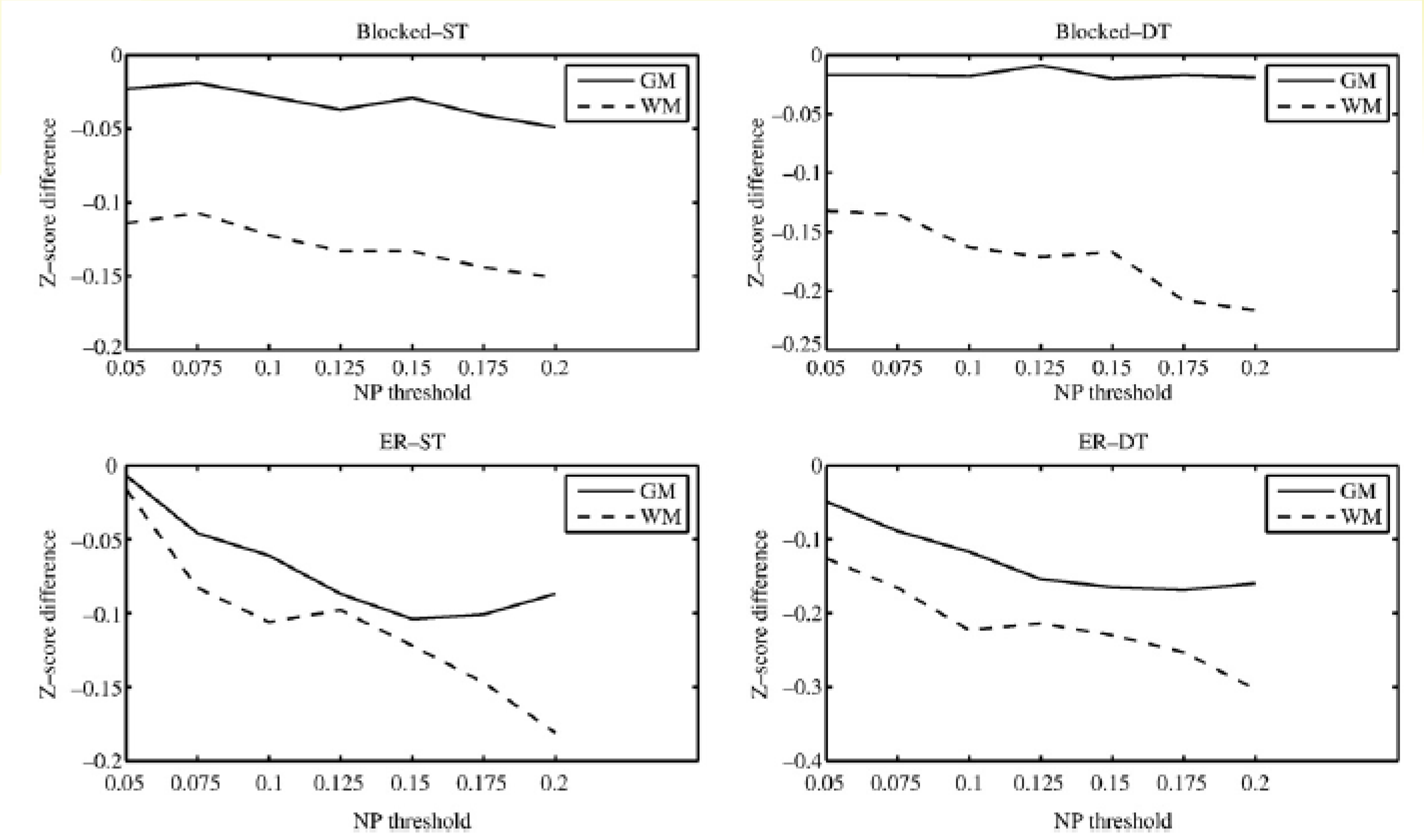


NP threshold with test data (blocked design)

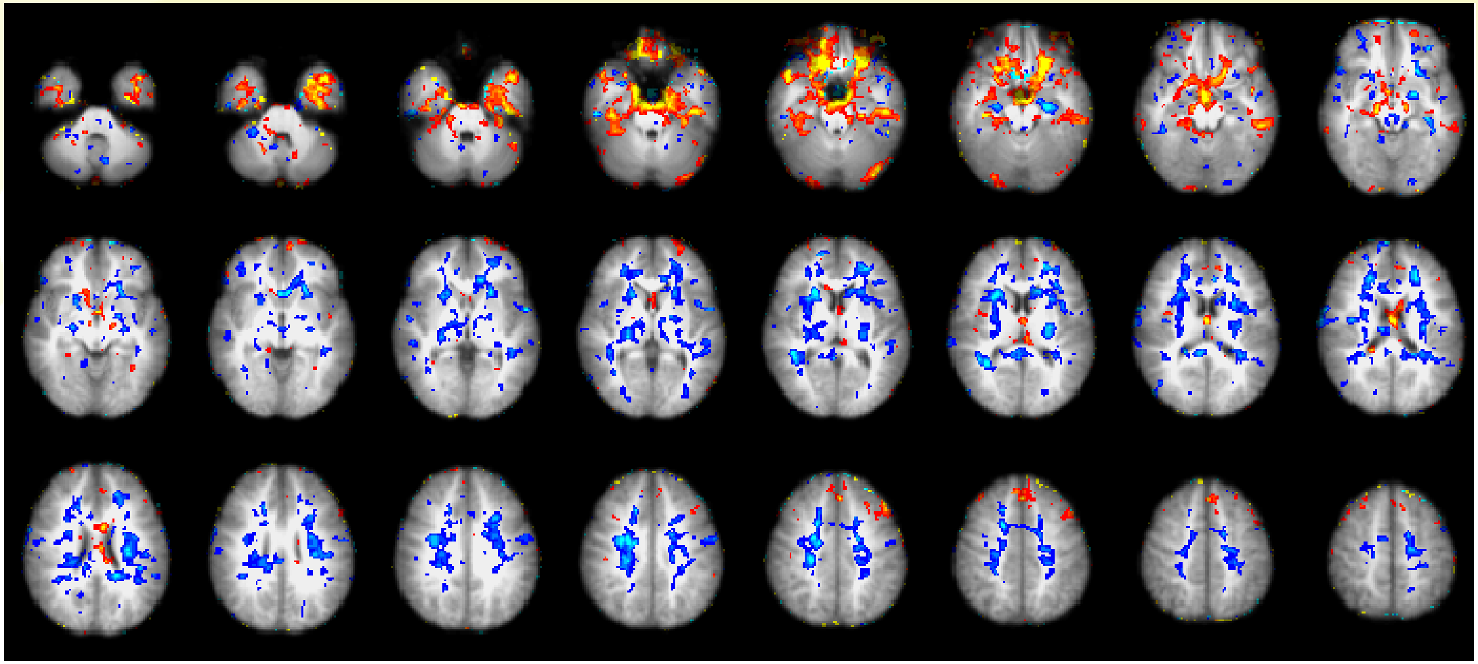


Average False Alarm rate

Effects to group level analyses

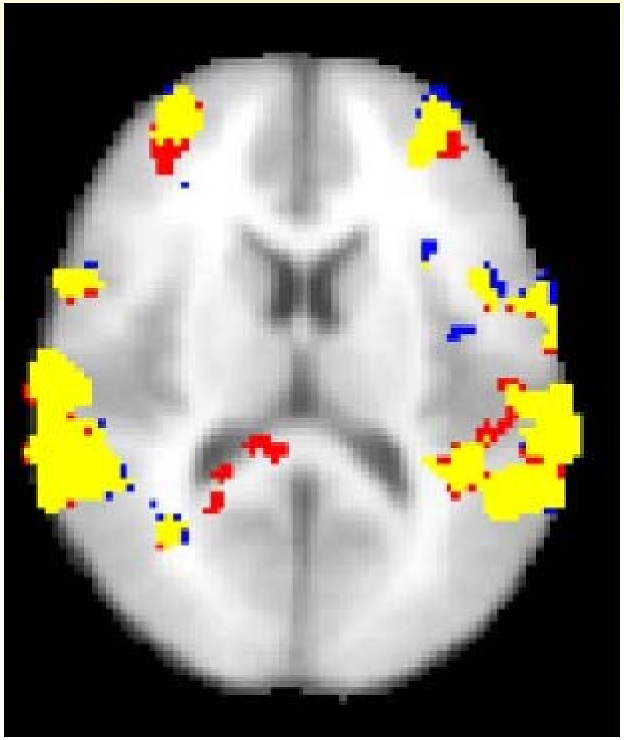


Effects to group level analyses (test data, blocked design)



Hot colors: Group level Z score increased more than 0.33
Cold colors: Group level Z score decreased more than 0.33

Effects to group level analyses



- Greater dual task activity than single task activity in the group analysis (P = 0.05, corrected for whole brain using GRF theory)
- Yellow: The voxel activated for original and denoised datasets
- Blue: The voxel is activated only for denoised dataset
- Red: The voxel is activated only for original dataset

Results: Number of Components: ER design

NP	0.05	0.1	0.15	0.2
Mean M	51.3	51.3	51.3	51.3
Mean R	4.58	9.83	12.3	14.42
Mean Mdn	47.1	42.8	40.5	38.7
Mean |Mdn – (M - R)|	0.67	1.50	1.50	1.75

- NP ~ NP threshold
- M ~ Number of estimated components
- R ~ Number of rejected components
- Mdn ~ Number of estimated components from denoised data

Conclusions

- A method for automatic identification of artifact/structured noise related independent components has been described
- The method was based on supervised classification using Global Neyman Pearson decision trees
  - We have tested also other classification schemes. Good generalization rates were hard to achieve. However, this data is preliminary.
- The Matlab-code & classifiers can be obtained from http://www.cs.tut.fi/~jupeto/software.html

Thank you!