KU LEUVEN

iMinds | Medical Information Technologies Department

The power of low rank MATRIX and TENSOR Approximations in *Smart Diagnostics*

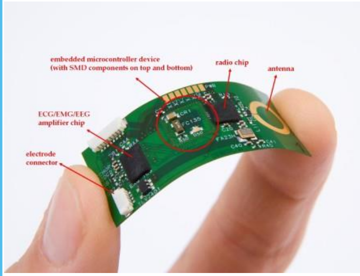
HCM workshop: Low-rank Optimization and Applications
Bonn, Germany
June 8-12, 2015

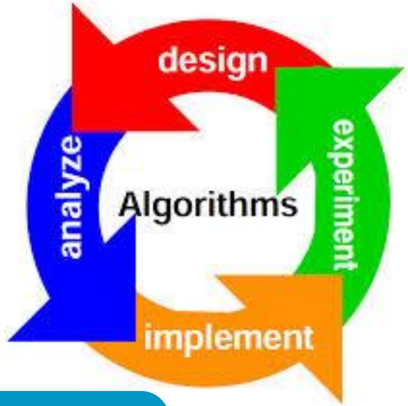
Prof. Sabine Van Huffel

European Research Council
Established by the European Commission

Contents Overview

- Introduction
  - Smart Patient Monitoring
  - Blind Source Separation
  - Tensor Decompositions
- Examples in Neonatal Brain Monitoring
- Examples in MR-based Brain Tumor Diagnosis
- Conclusions and new directions





Brain monitoring for neurological

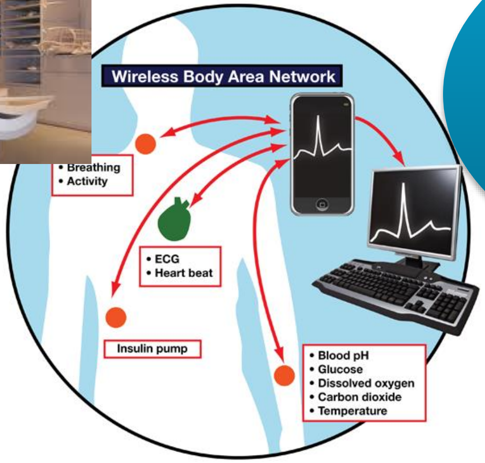
Vital signs monitoring: sleep, stress, cardio risk stratification

Algorithms (Technology)

Sensors (Carriers)

Pathologies (Applications)

Smart Diagnostics



Oncology: cancer diagnosis and prognosis

Chronic disease management & telemonitoring application

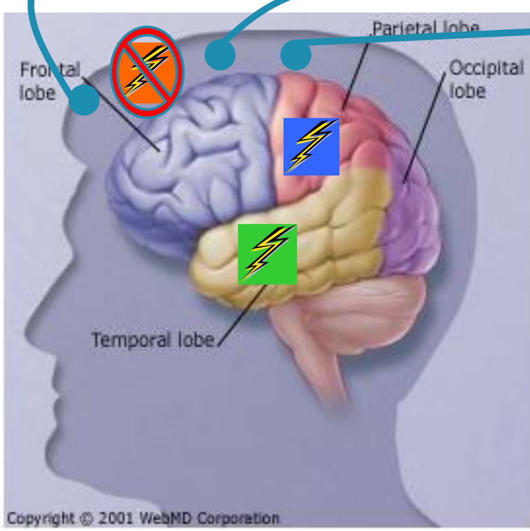
Blind source separation

Signal analysis difficult because of artefacts → REMOVE

Matrix based Blind Source Separation (BSS)

Non-unique → *Constraints are needed!*

- sources orthogonal (PCA),
- sources statistically independent (ICA)
- **sources uncorrelated and of different autocorrelation (CCA)**

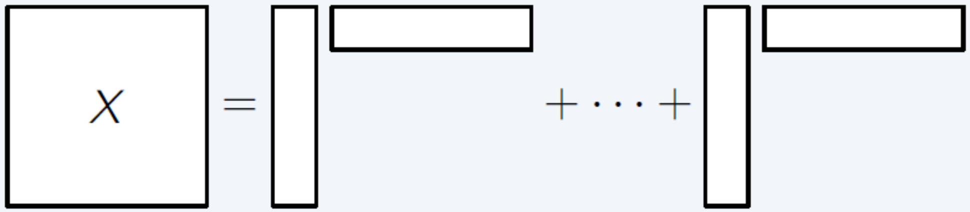


$$EEG_1 = a_{11}\mathbf{s_1} + a_{12}\mathbf{s_2} + a_{13}\mathbf{s_3}$$
$$EEG_2 = a_{21}\mathbf{s_1} + a_{22}\mathbf{s_2} + a_{23}\mathbf{s_3}$$
$$EEG_3 = a_{31}\mathbf{s_1} + a_{32}\mathbf{s_2} + a_{33}\mathbf{s_3}$$

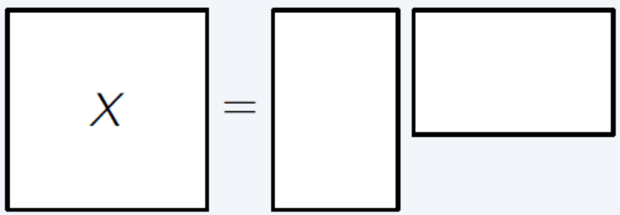
$$EEG = A?\; S^T?$$

From Matrix to Tensor rank

At its core, a matrix decomposition is



or



with some constraints

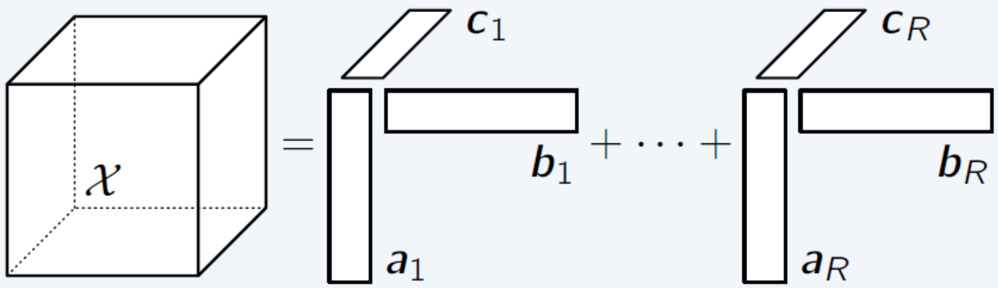
Two equivalent ways to define matrix rank

- Minimal number of rank-one matrices that sum to X
- Dimension of column (or row) space of X

For tensors, these are two different concepts!

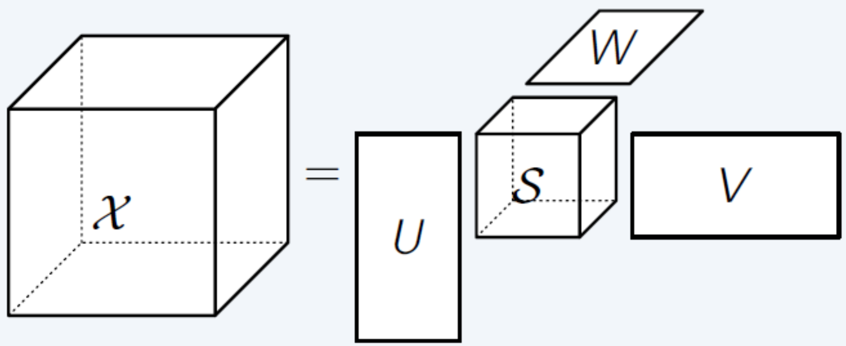
Tensor Decompositions

The canonical polyadic decomposition (CPD) decomposes a tensor into a minimal number of rank-one tensors R



The tensor's rank is defined as R

A low multilinear rank approximation (LMLRA) decomposes a tensor into a core tensor $\mathcal{S}$ and matrices U, V and W



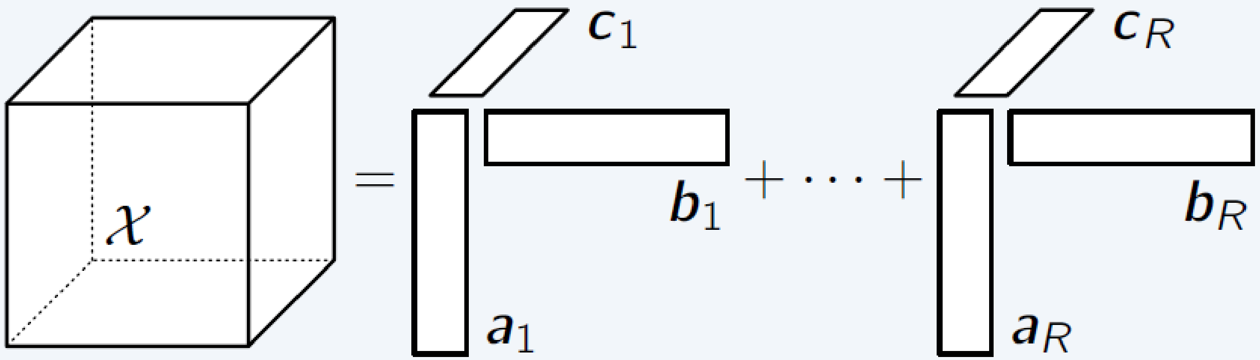
The tensor's multilinear rank is defined as the triplet $(\mathrm{rank}(U), \mathrm{rank}(V), \mathrm{rank}(W))$

Uniqueness means Interpretable

Without constraints, matrix decompositions are not unique

$$X = A \cdot B = (A \cdot M) \cdot (M^{-1} \cdot B) = \hat{A} \cdot \hat{B}$$

Tensor decompositions can be unique under mild conditions!
For example, the vectors $\boldsymbol{a}_r$, $\boldsymbol{b}_r$ and $\boldsymbol{c}_r$ in the CPD



are generically unique when $k_A + k_B + k_C$ >= 2*R+2

$A=[a_1,\ldots,a_R]$
$B=[b_1,\ldots,b_R]$
$C=[c_1,\ldots,c_R]$

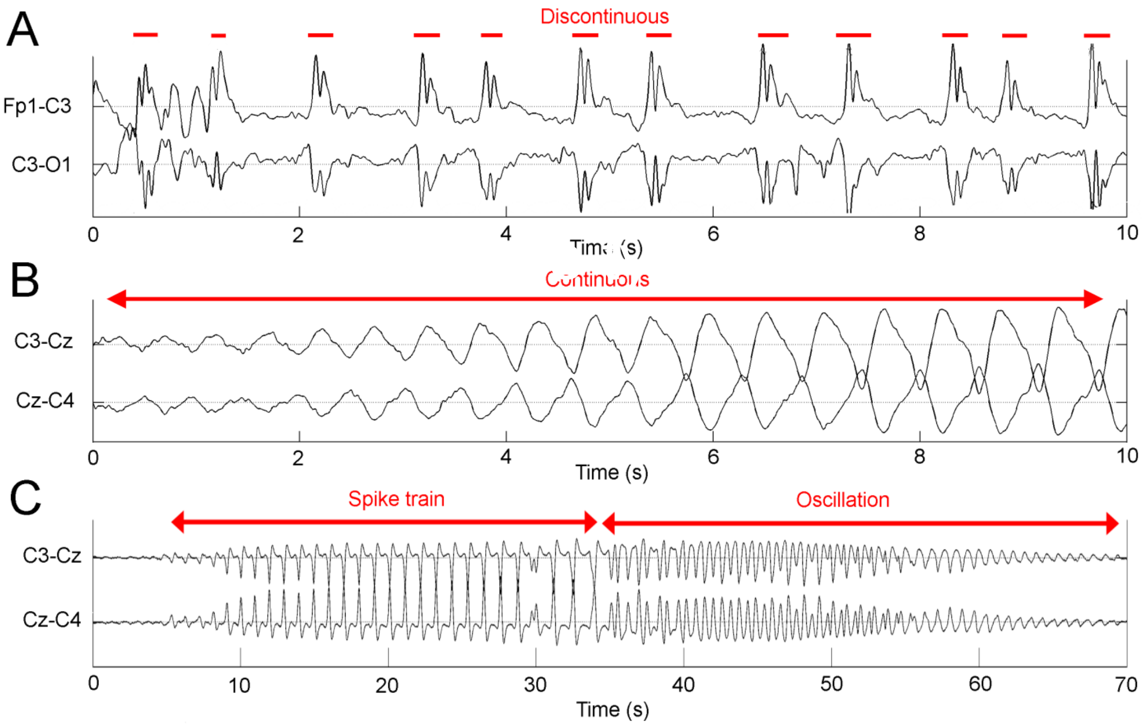
Contributors (nonexhaustive list):
L. De Lathauwer, P. Comon, T. Kolda, B. Bader, L-H Lim, C. Van Loan, E. Acar, A. Cichocki, O. Alter, R. Bro, M. Morup, N. Sidiropoulos, I. Domanov, M. Sorensen, L. Sorber, M. Ishteva, L. Albera, M. Haardt, …. and collaborators

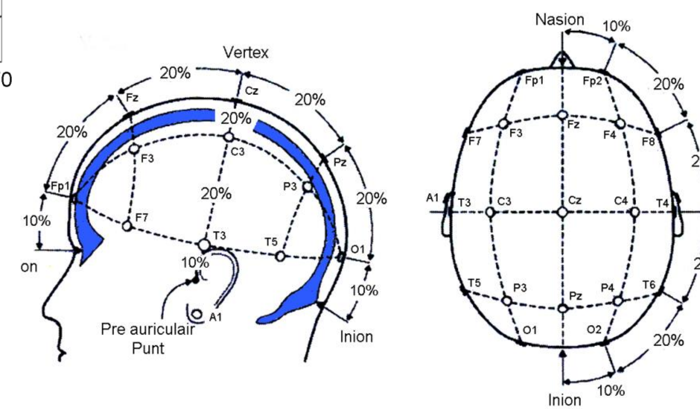
Contents Overview

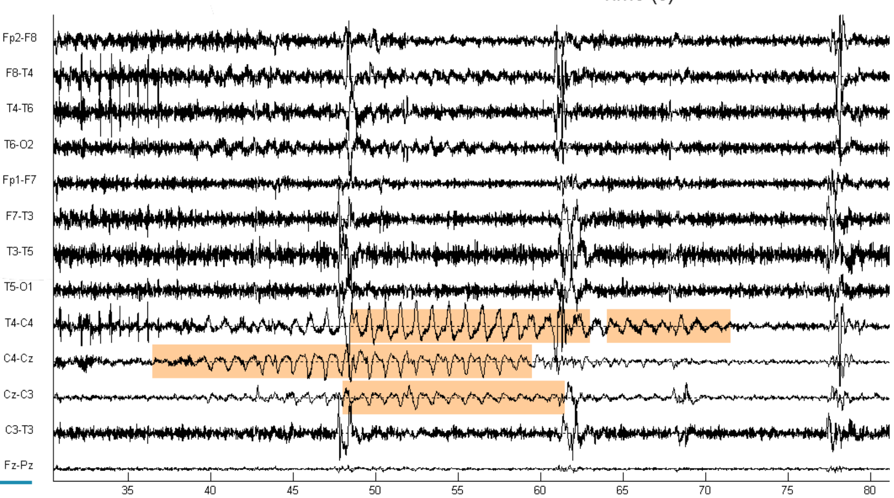
- Introduction
- Examples in Neonatal Brain Monitoring
  - Seizure detection
  - Seizure onset localization
  - Background EEG grading
- Examples in MR-based Brain Tumor Diagnosis
- Conclusions and new directions

Neonatal Brain Monitoring:

Seizure detection







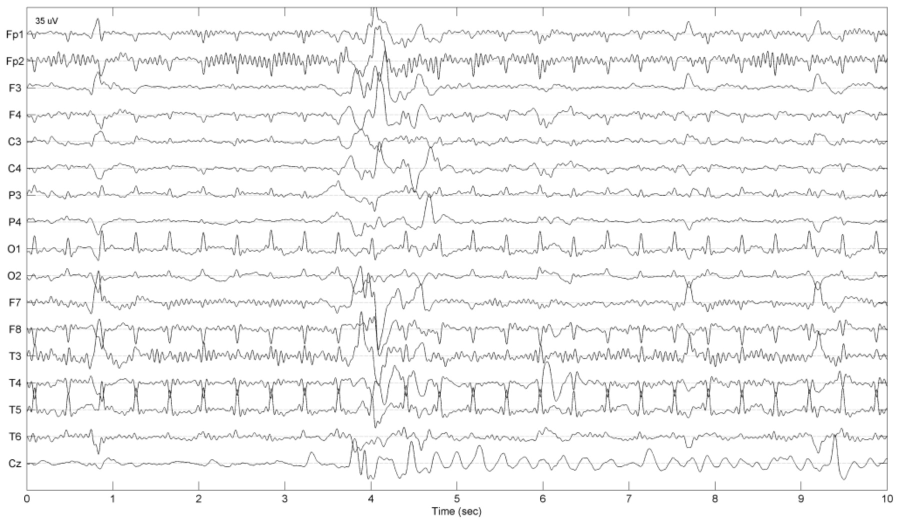
Deburchgraeve, PhD thesis, 2011; P.J. Cherian 2011; W. Deburchgraeve et al, Clin. Neurophys. 2008 & 2009
Alternatives using classification: See pubs of Temko, G. Boylan, etc.

Artefact removal: ECG, respiration, pulsation

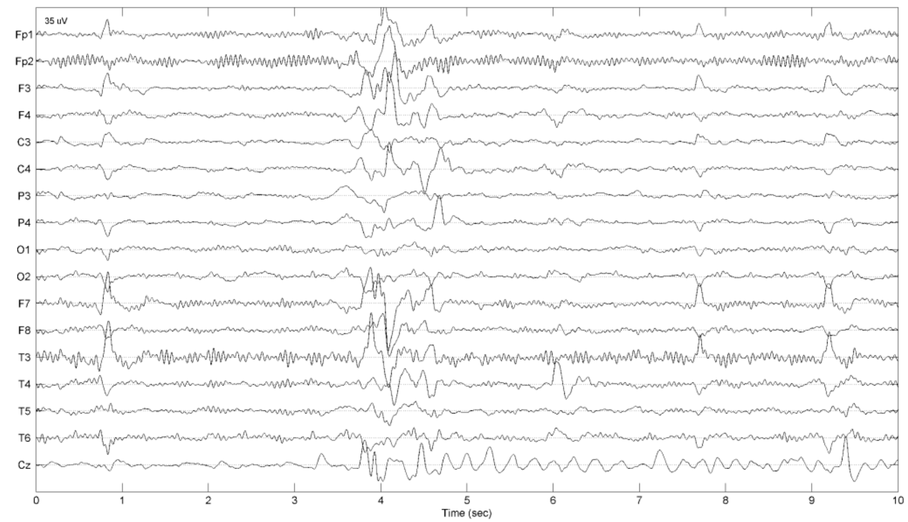
BSS based algorithm removes these 3 artifacts **before** seizure detection starts

ECG artefact removal: *with RobustICA* (spiky)
Respiration & Pulsation removal: *with SOBI* (oscillatory, autocorrelated)
Number of sources: *estimated with PCA and variance threshold*
IC source recognition: *correlation with reference after transformation*

ECG spike artefact



Cleaned EEG



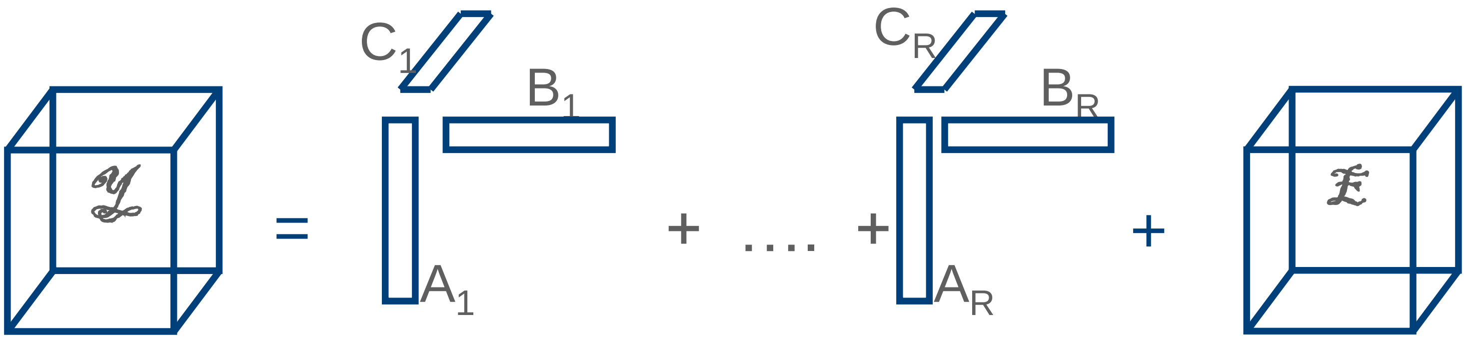
VALIDATION study artefact-removal: *M. De Vos et al., Clinical Neurophys., Dec. 2011*

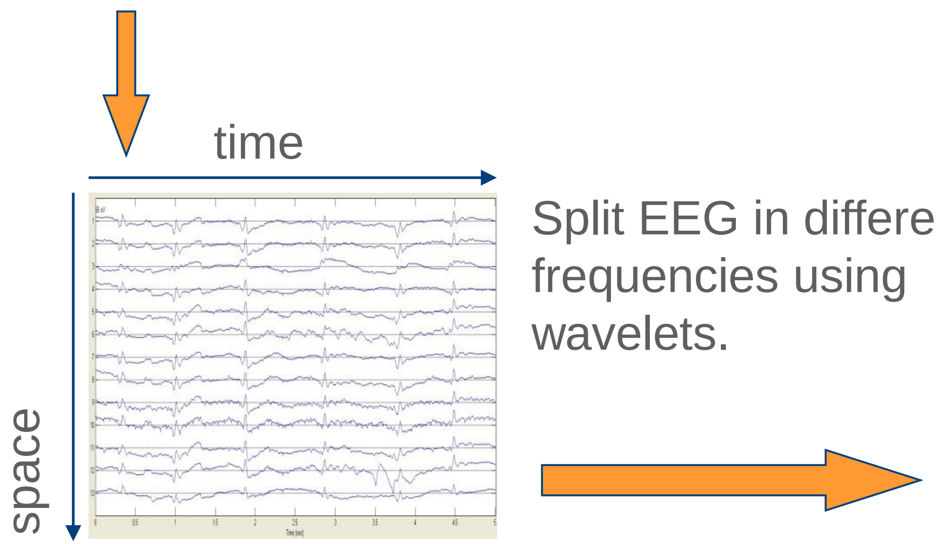
•13 patients: asphyxia (9 with artefacts, 4 artefact-free), 8h EEG per patient

	Fp/h	Seizure detection Rate (%)
No-AR:	0.38 (0-5)	87.1 (20-100)
With-AR:	0.00 (0-0.875)	100 (20-100)

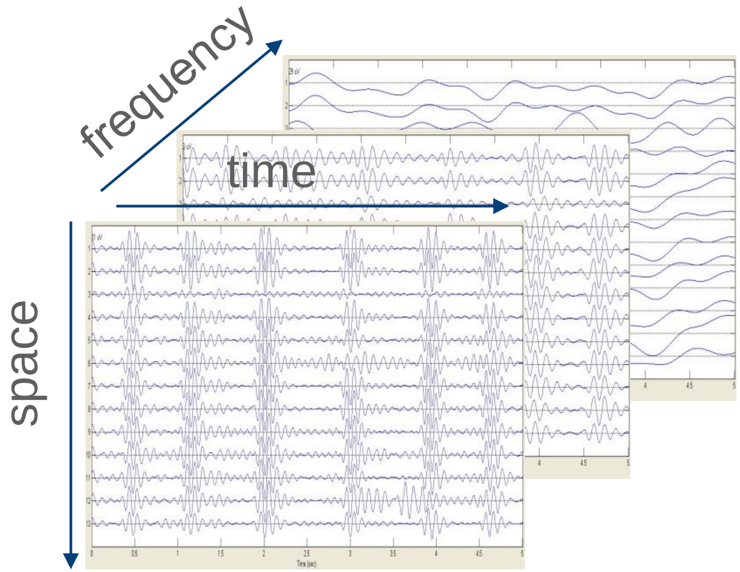
Seizure onset localization:

CPD





Split EEG in different frequencies using wavelets.

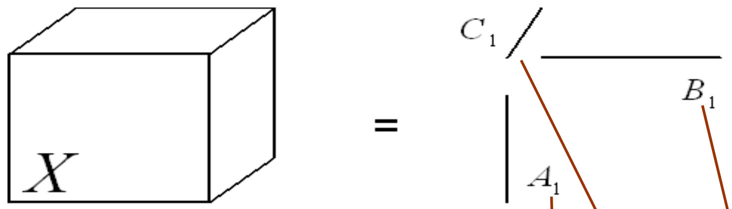


=> Analysis in 3 dimensions instead of just 2

KU LEUVEN

Interpretation of a trilinear component

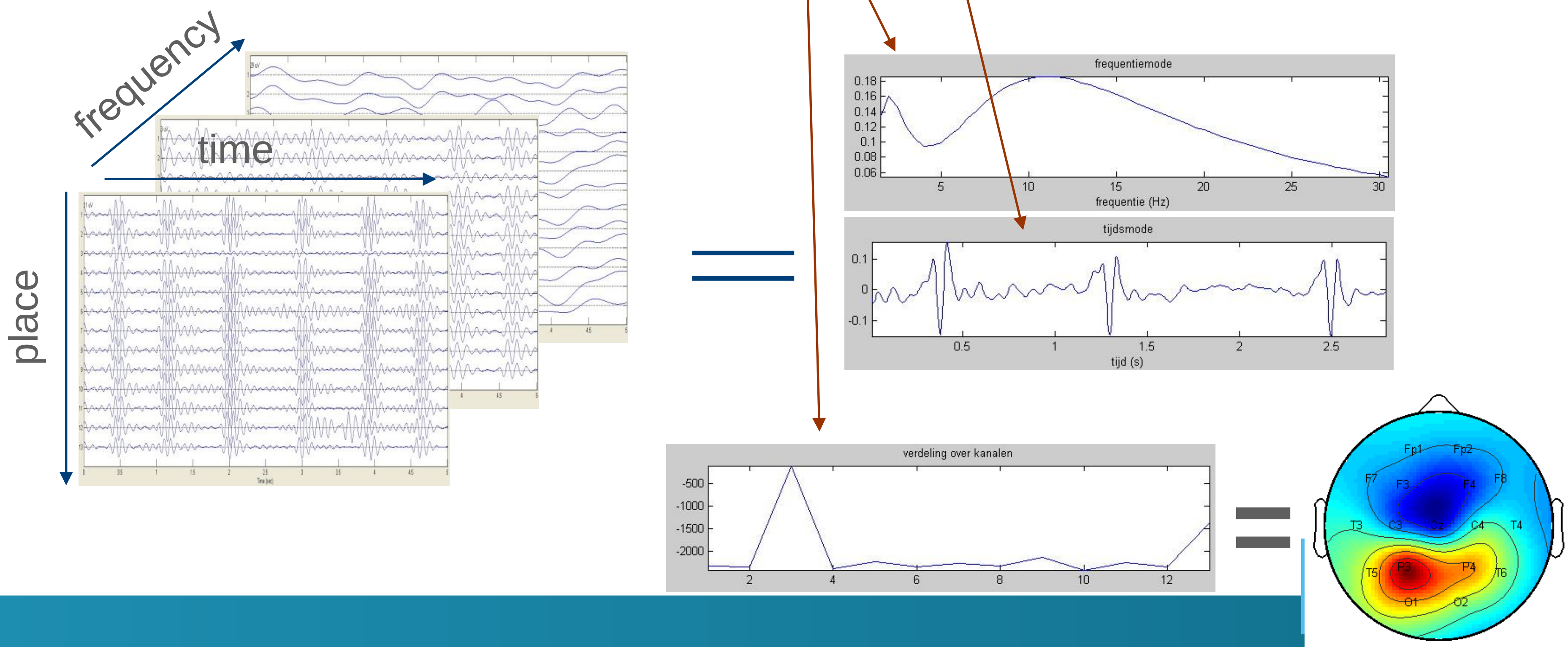
CPD: Example extracting 1 component



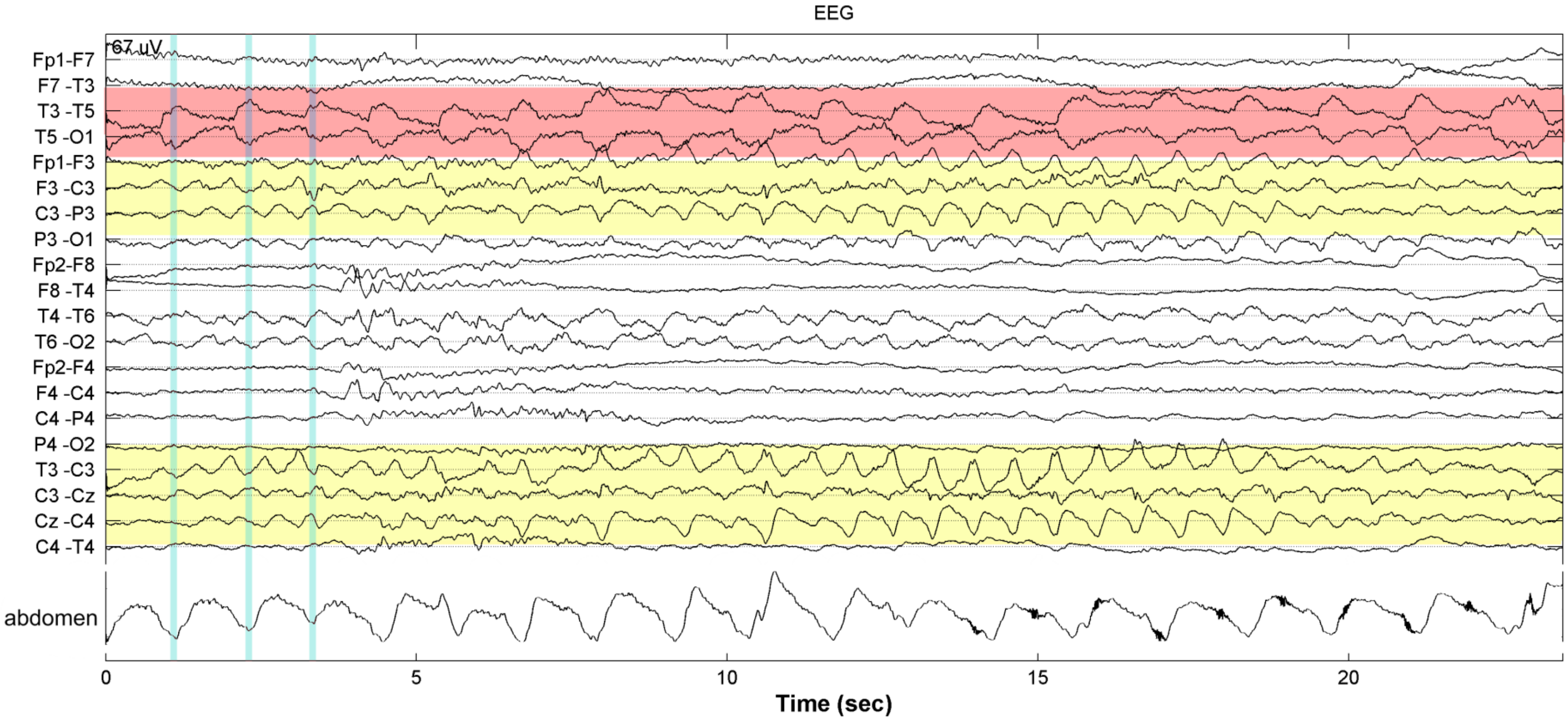
B_1: time course

A_1: distribution over channels

C_1: frequency content (distribution across scales).



separation seizure from background activity



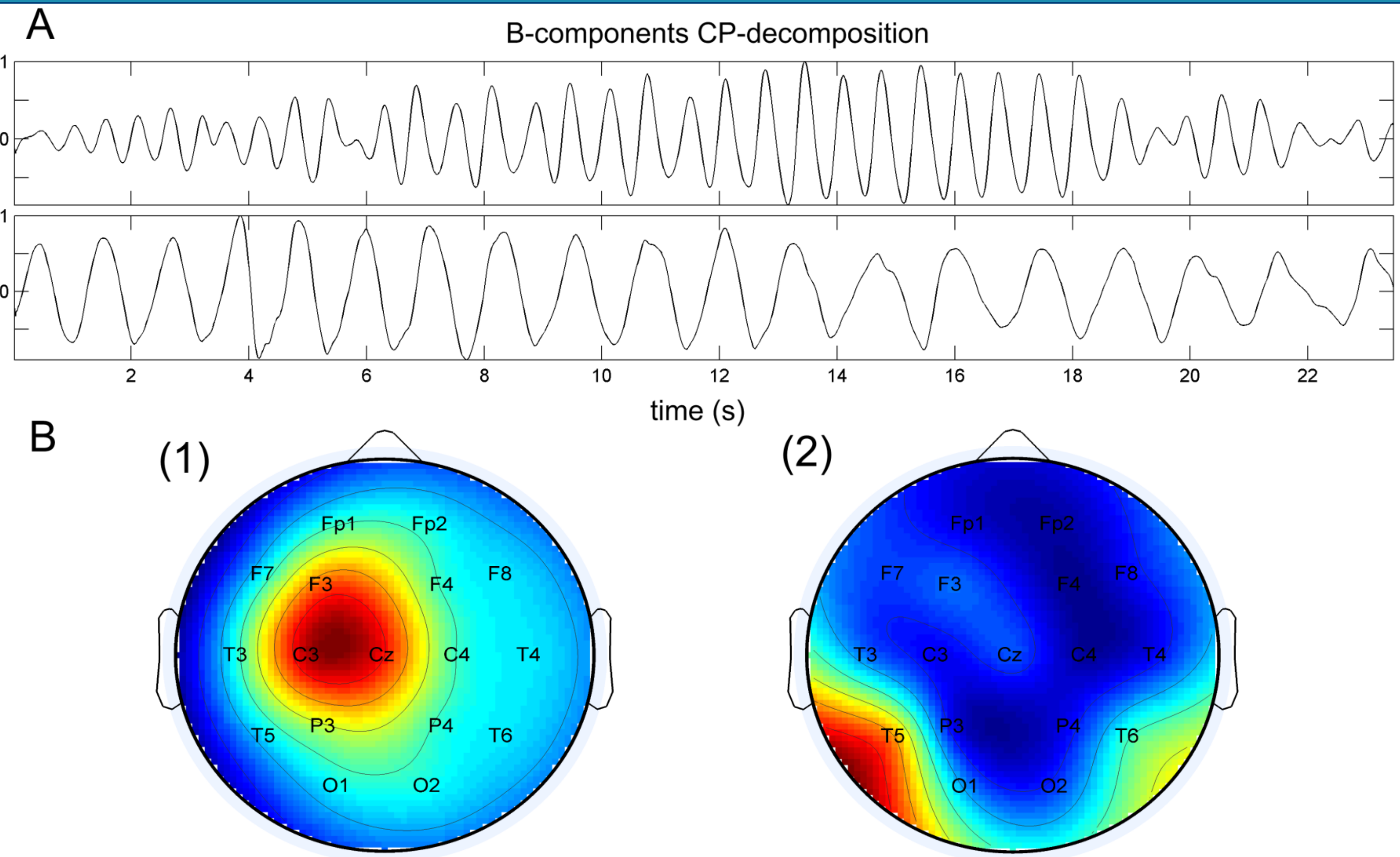
⇒ Extract spatial distribution of the seizure without distortion of the artifact

KU LEUVEN

CPD: spatial source components

(1) Seizure:

(2) Artefact:

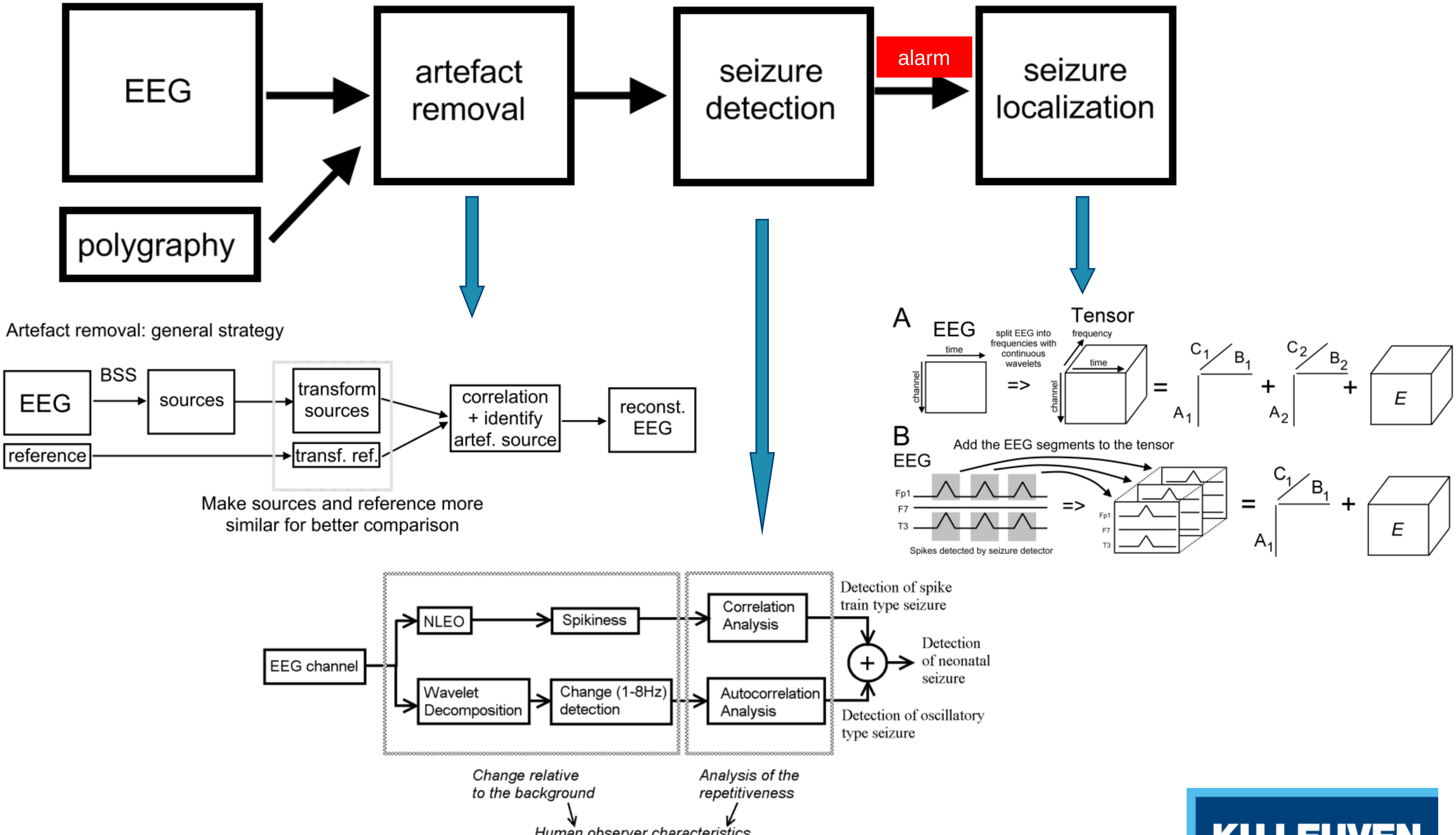


VALIDATION Study: *W. Deburchgraeve et al., Clinical Neurophys., 2009*

good qualitative correspondence between visual analysis by expert and algorithm.
Together with seizure detector useful for bedside brain monitoring

KU LEUVEN

Automated seizure monitoring



NeoGuard : decision support

Brain injury estimate

- Detection of neonatal epileptic seizures
- **Seizure onset localization**
- Inter-burst intervals

Clinician's expertise

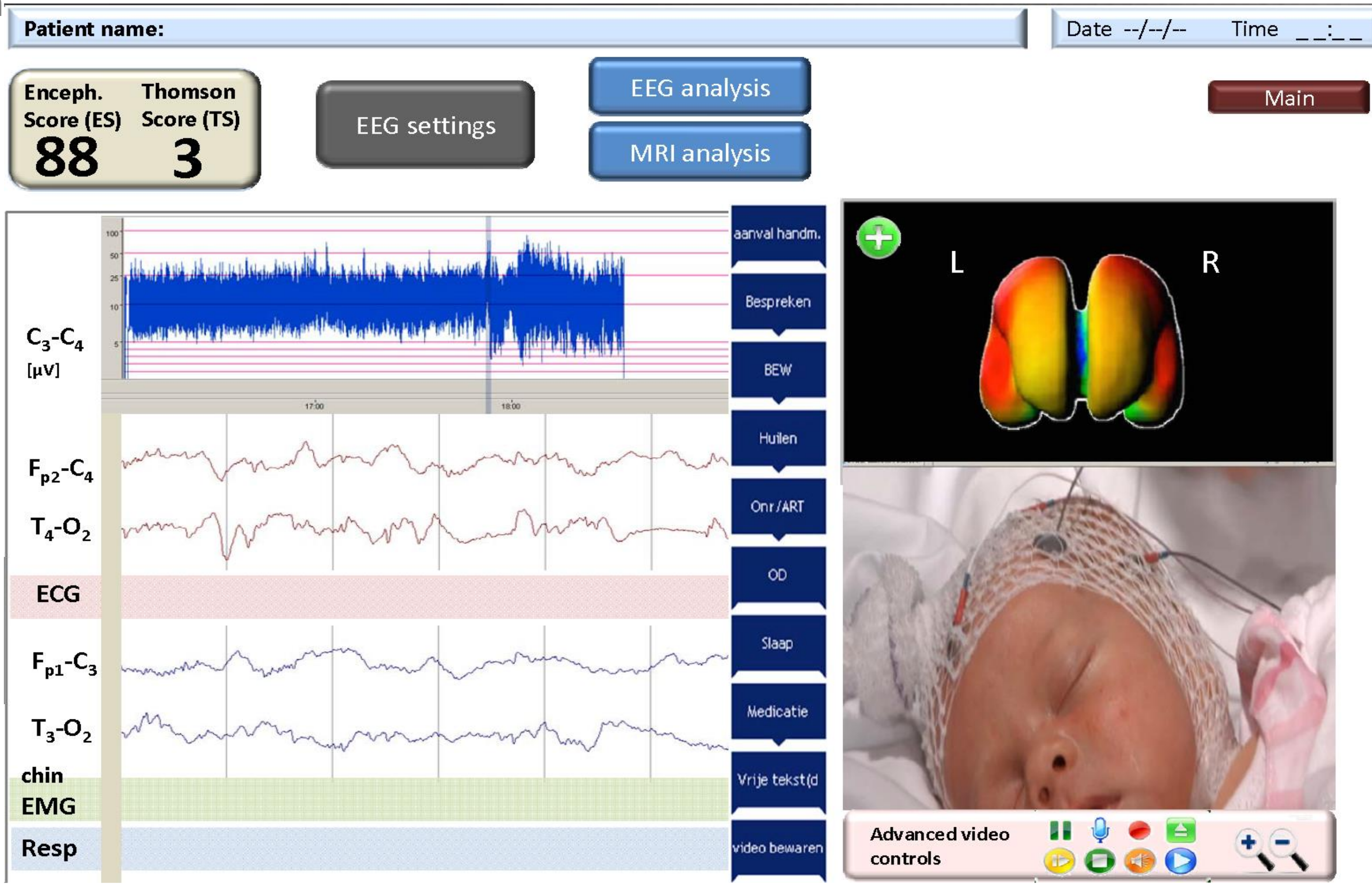
- Neurophysiological knowledge included in algorithms

Brain Monitoring

- **Recovery after damage**
- Maturation in preterms

Outcome prediction

- Good
- Poor

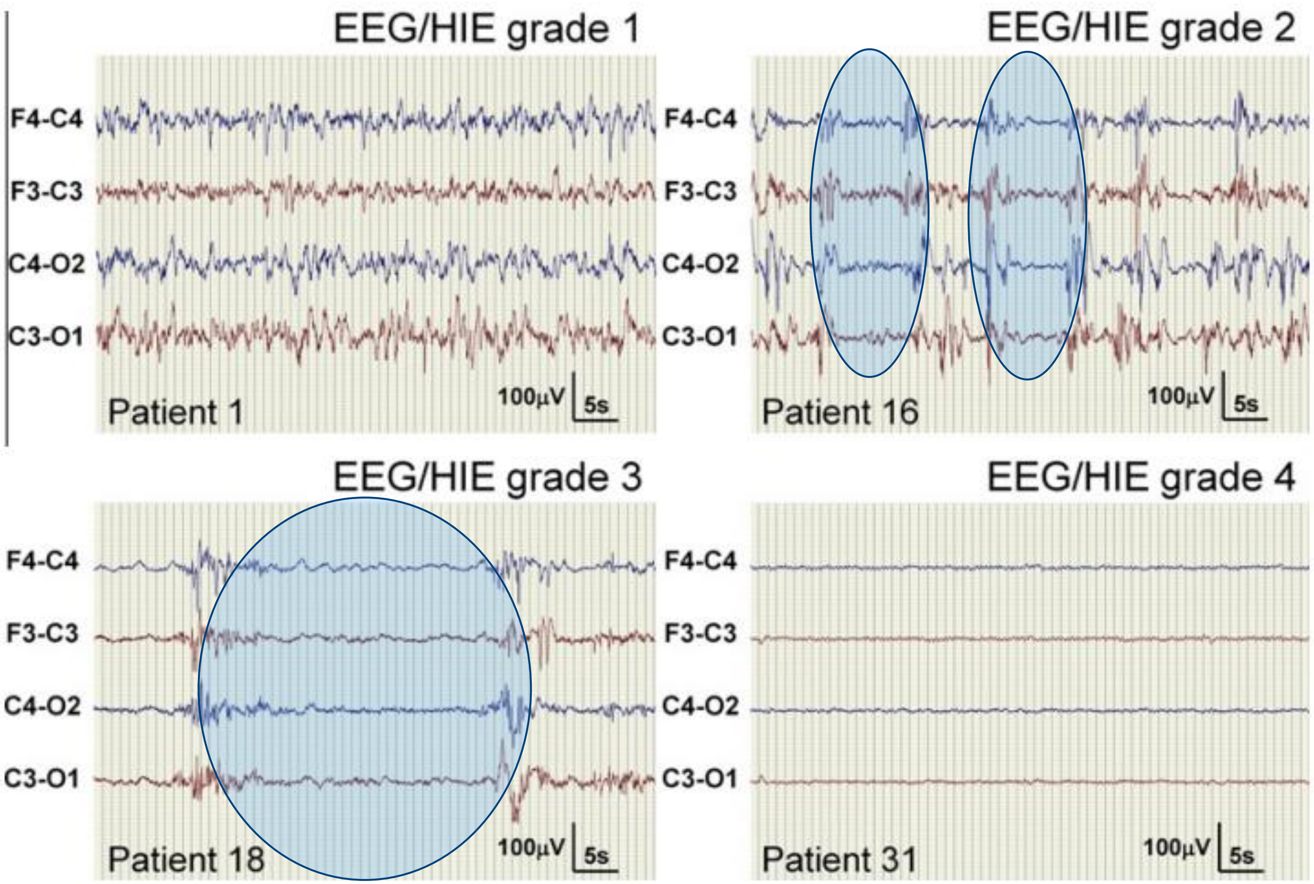


KU LEUVEN

How Assessing Brain Recovery?
→ monitor abnormality of Background EEG

mildly abnormal

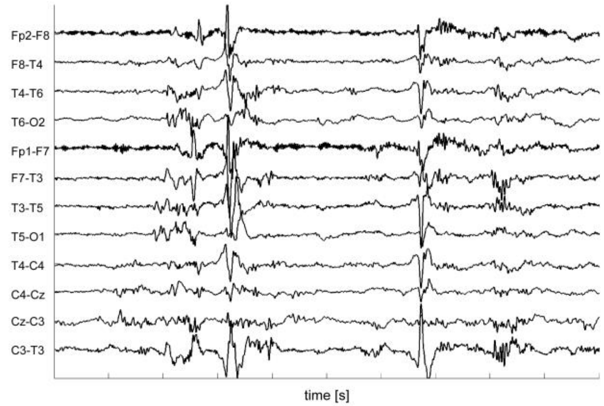
moderately abnormal

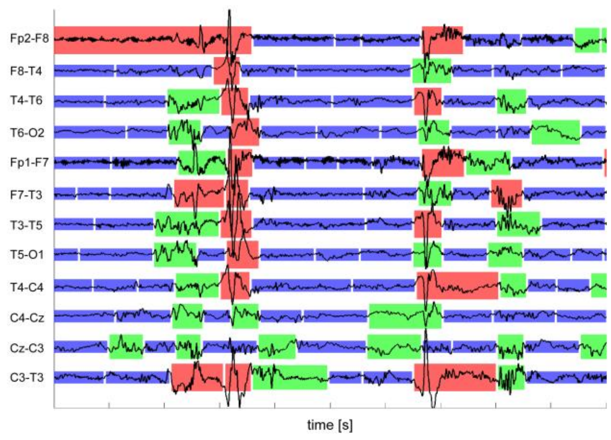


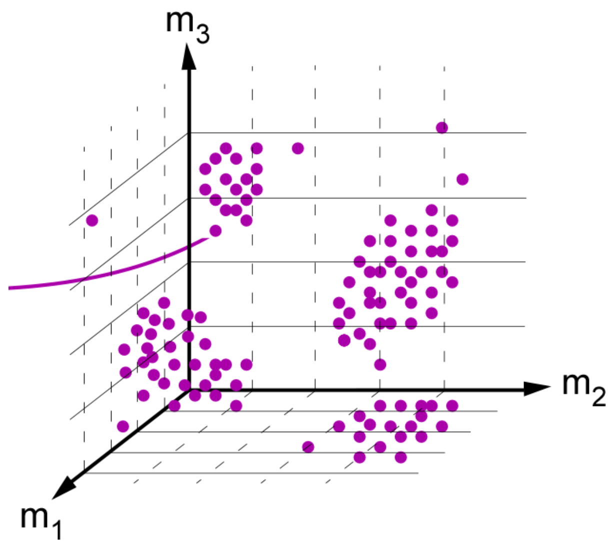
severely abnormal

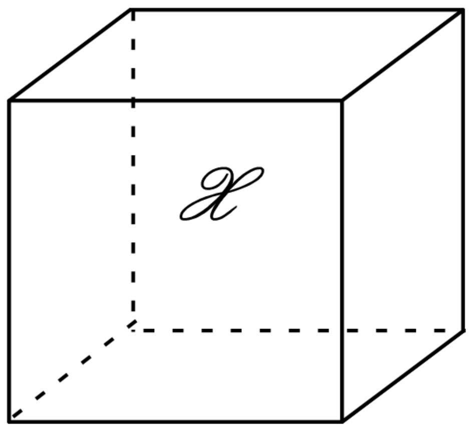
Ideal examples, taken from [Korotchikova et al., 2011]

Monitoring neonatal background EEG:
The power of structuring data









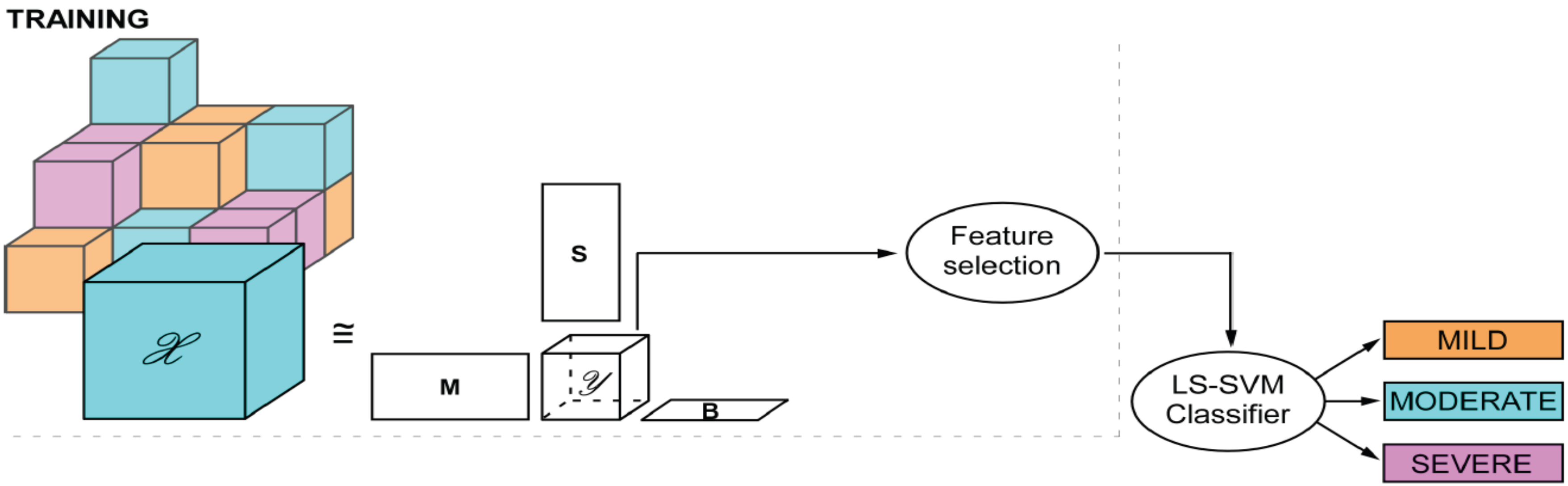
V. Matic et al., J. Neural Engineering, Oct. 2014

Higher Order Discriminant Analysis

-> compute simultaneous LMLRA

-> factors M, S, B common and orthogonal

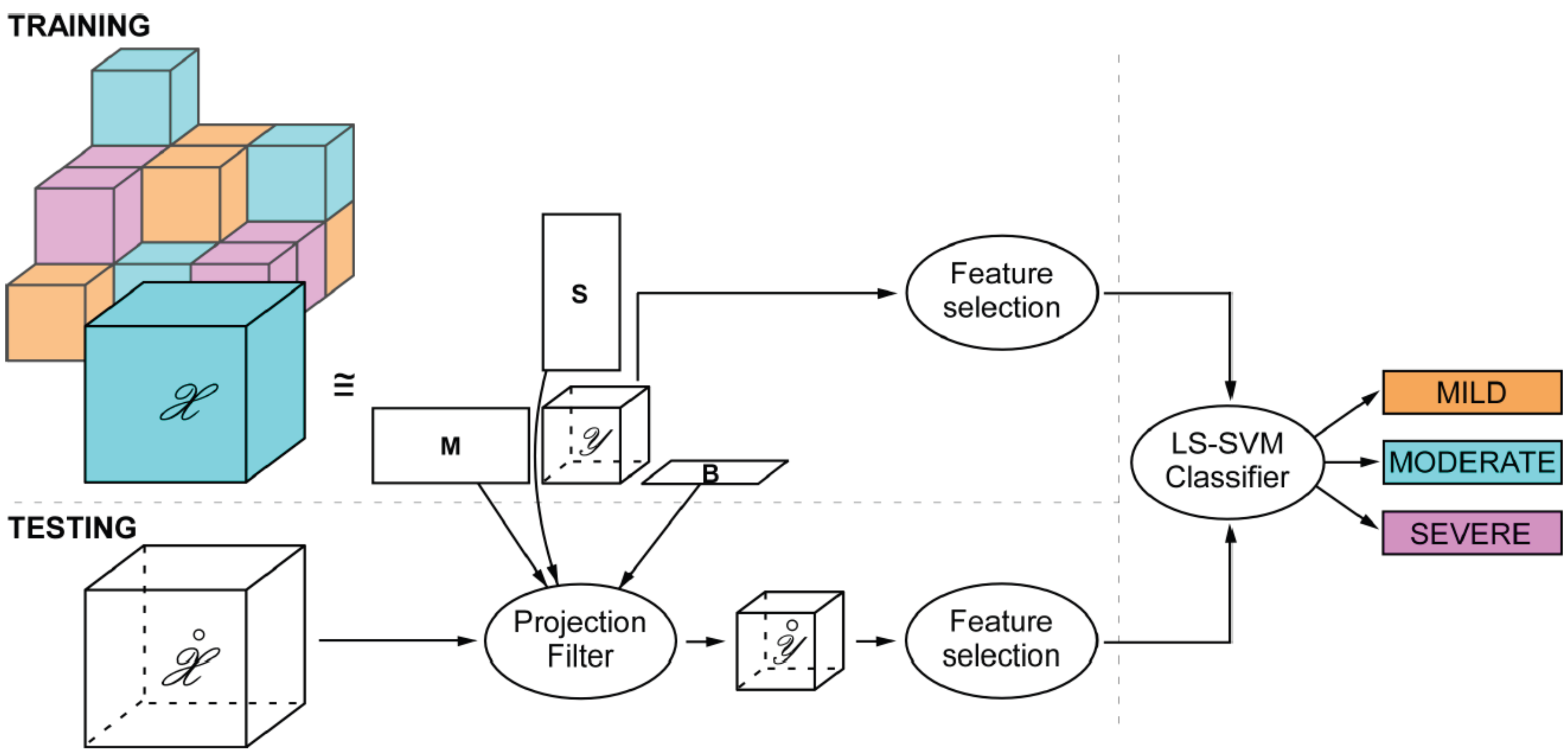
-> maximizing the Fisher ratio between core tensors



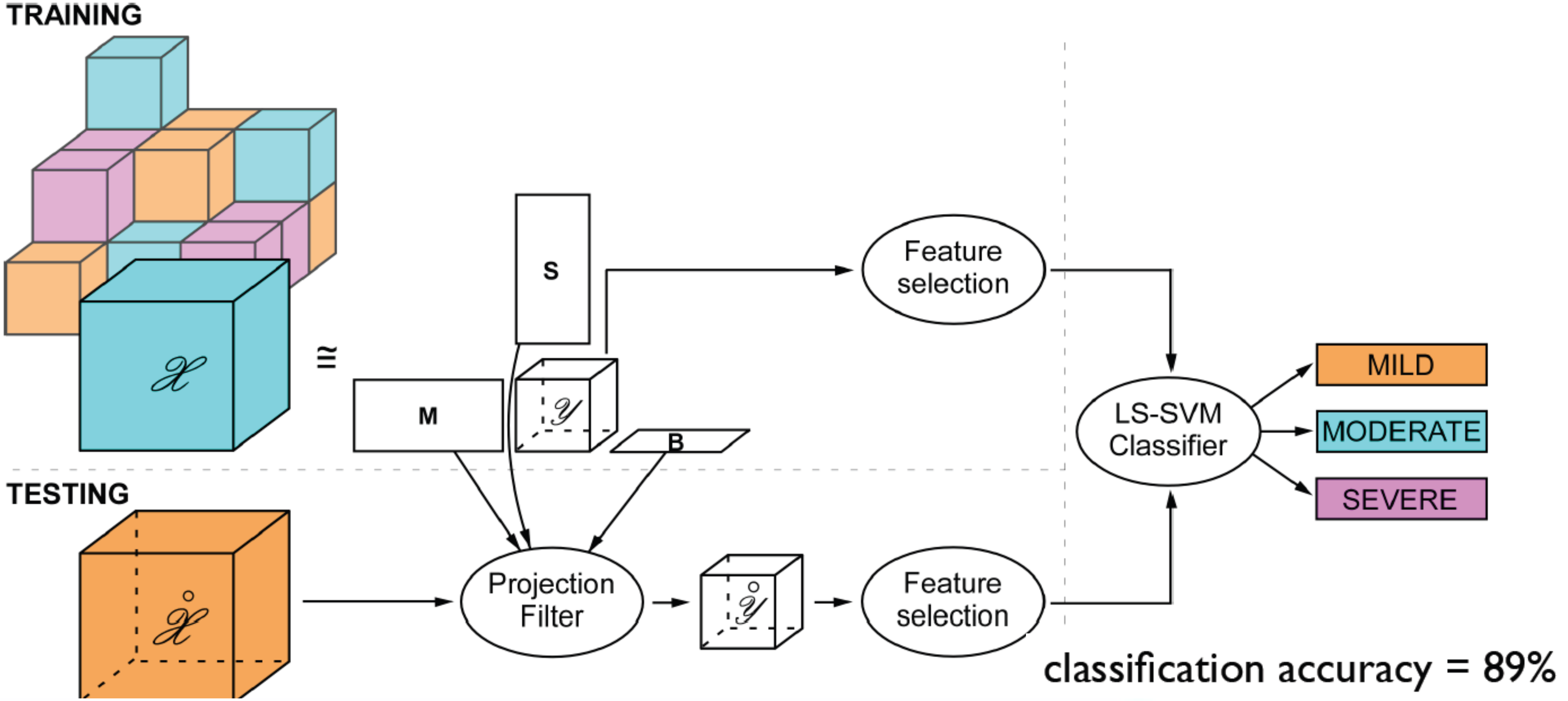
Phan A and Cichocki A, Nonlinear Theory Appl., IEICE, 2010
Phan A, 2011, Matlab Software Toolbox
(www.bsp.brain.riken.jp/˜phan/nfea/nfea.html)

KU LEUVEN

Higher Order Discriminant Analysis



Higher Order Discriminant Analysis



classification accuracy = 89%

Automated \ Expert EEG reader	MILD	MODERATE	SEVERE
MILD	73 (91%)	6	1
MODERATE	7	44 (76%)	7
SEVERE	0	8	126 (94%)
Achieved accuracy	91%	76%	(94%)

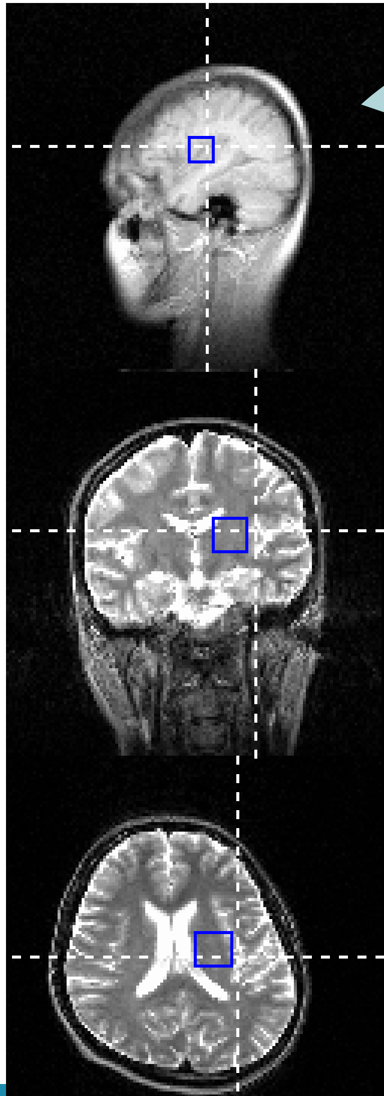
(V. Matic et al, J. Neural Eng. 11, 2014)

KU LEUVEN

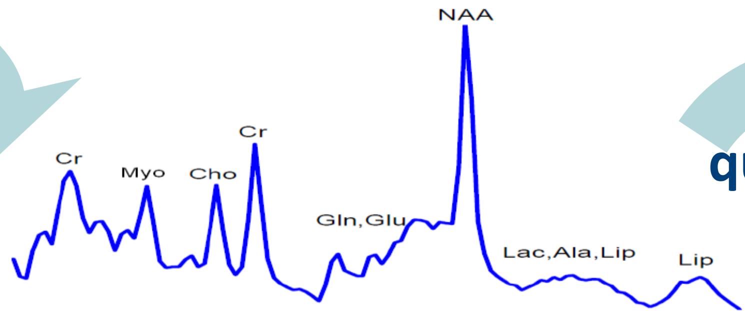
Contents Overview

- Introduction
- Examples in Neonatal brain monitoring
- Examples in MR-based Brain tumor diagnosis
  - NMF
  - Hierarchical(h) NMF
  - Multimodal hNMF
  - Non-negative Tensor Factorization
- Conclusions and new directions

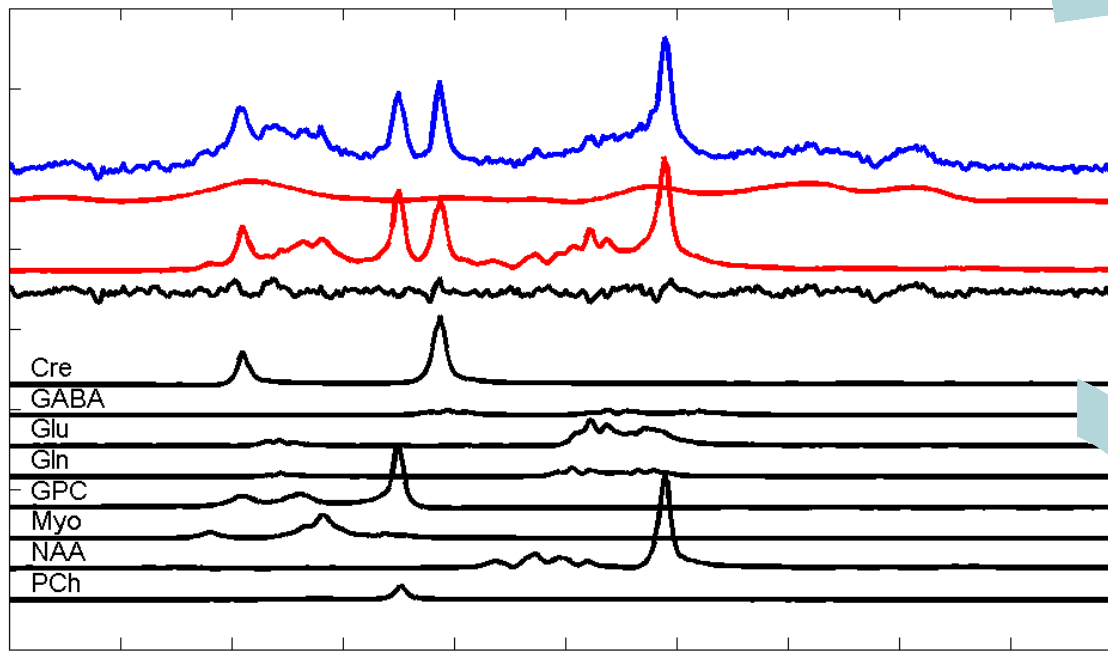
Metabolite quantification for MR Spectroscopy (MRS)



Single-voxel MRS

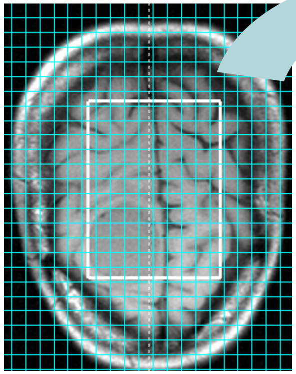


MRS quantification

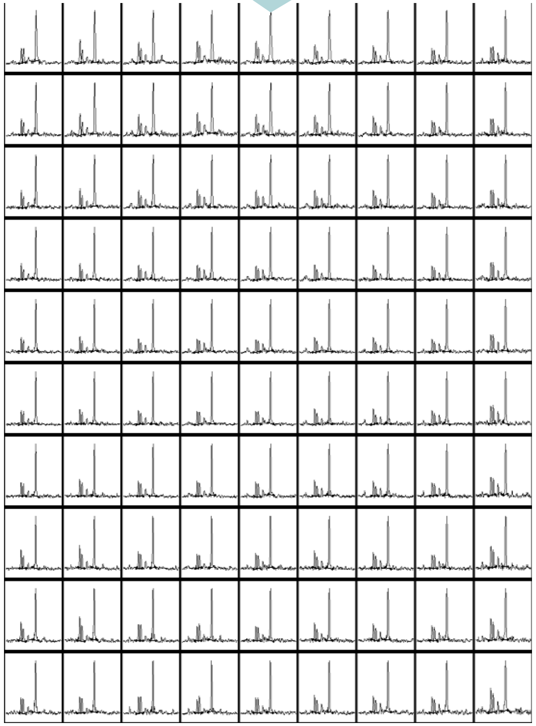


Metabolite conce

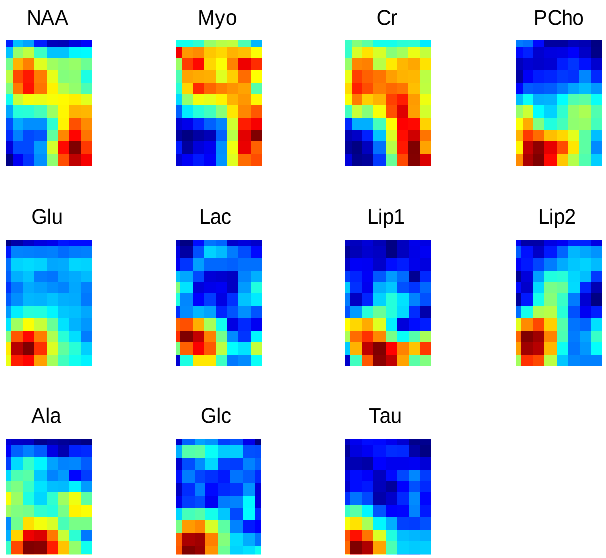
Metabolite quantification for MRS Imaging (MRSI)



Multi-voxel MRS



MRS quantification using spatial information

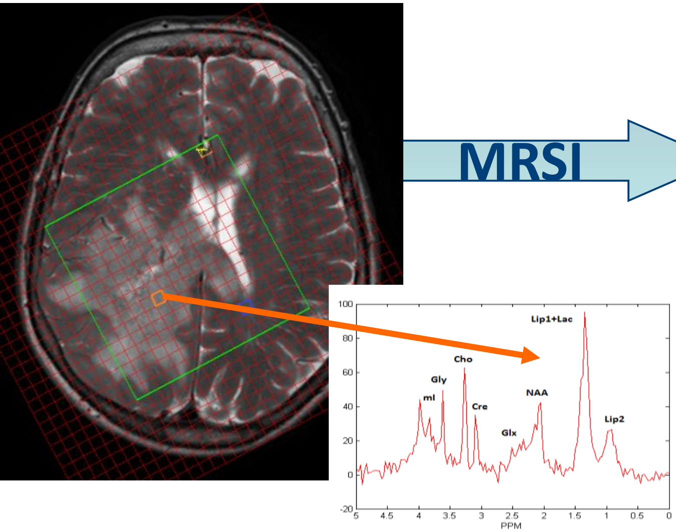


Metabolite maps

Metabolite concentrations = biomarkers of disease

KU LEUVEN

Unsupervised Brain Tumor Diagnosis using NMF



MRSI

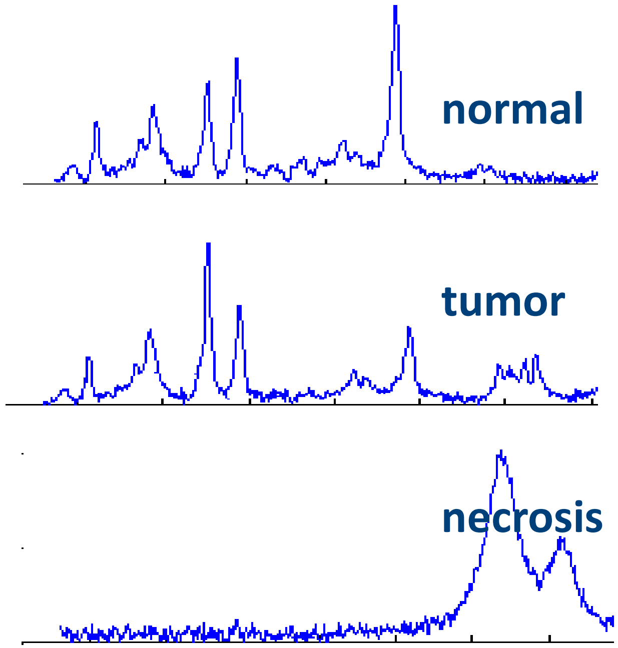
$\boldsymbol{Y}$ = matrix of spectra, $\boldsymbol{Y} \approx \boldsymbol{W}\,\boldsymbol{H}$

min || $\boldsymbol{Y}$ - $\boldsymbol{W}\,\boldsymbol{H}$ ||

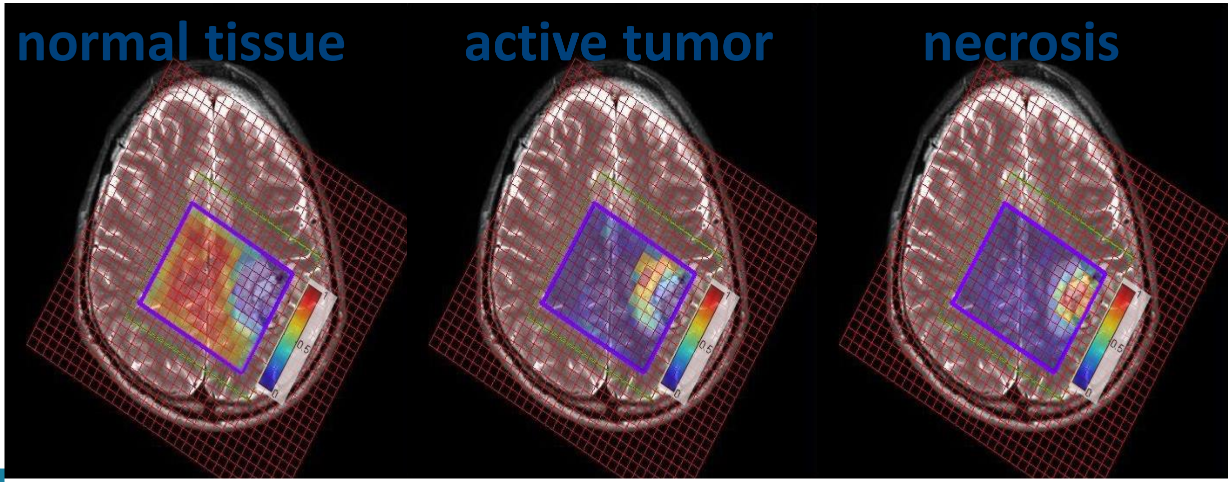
such that $\boldsymbol{W} \geq \boldsymbol{0}$, $\boldsymbol{H} \geq \boldsymbol{0}$

non-negative matrix factorization

W = tissue-specific spectral patterns:

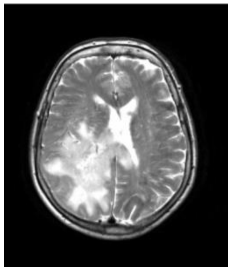


H = spatial distribution of tissue types:

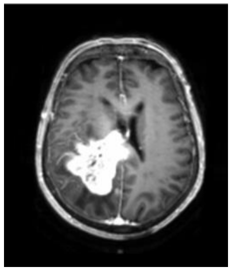


glioblastoma multiforme patient

Multi-Parametric (MP) NMF

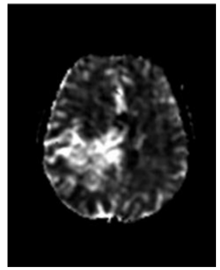


T2



T1

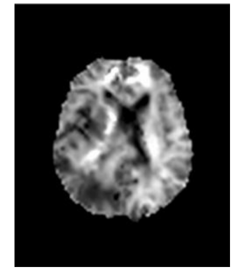
Anatomical MRI



CBV
Cerebral Blood Volume

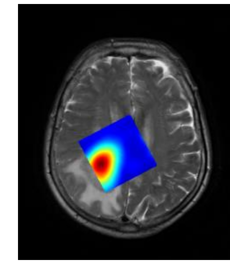
Perfusion MRI

MD
Mean Diffusion



MK
Mean Kurtosis

Diffusion MRI



NAA, Cre, tCho, Lip1,...

Spectroscopic MRI

<u>Research study:</u>

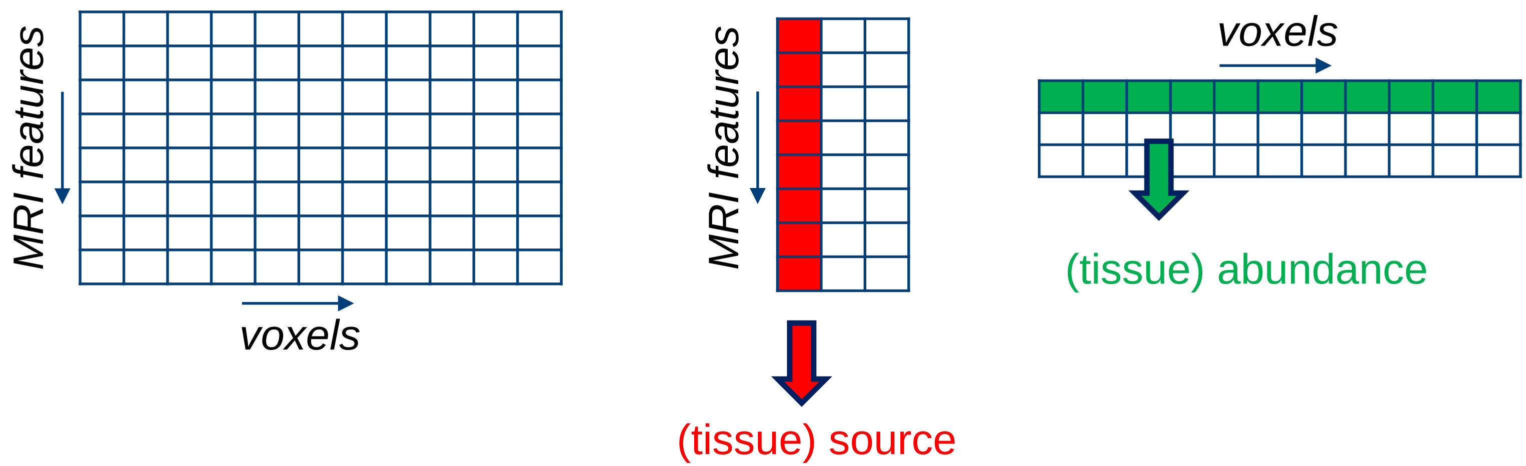
- 14 high-grade glioma (HGG) patients with complete MP-MRI dataset (3 grade III, 11 grade IV, UZ Leuven) and 10 low-grade gliomas

Research questions:

- *Do we get valid tissue differentiation with NMF on MP-MRI?*
- *What is added value of individual MRI modalities?*
- *How to improve NMF→ hierarchical NMF ? Tensor Factorisation?*

Non-negative matrix factorization (NMF)

$$\boldsymbol{Y} \approx \boldsymbol{W} \times \boldsymbol{H}$$

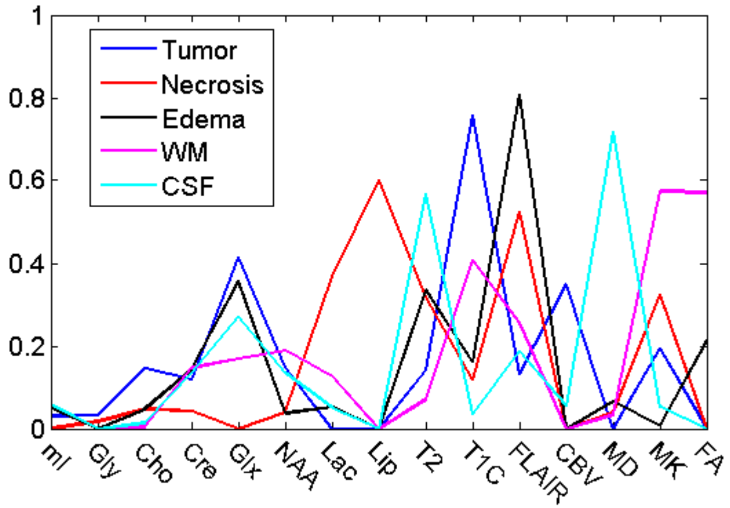


- Non-negativity constraint: $Y_{i,j}, W_{i,j}, H_{i,j} \geq 0, \forall i,j$
- Unsupervised: → applicable patient-by-patient,
  → tissue classes not a priori known
- MP-NMF:
  1. integrate ALL features of each modality into one vector
  2. use NMF

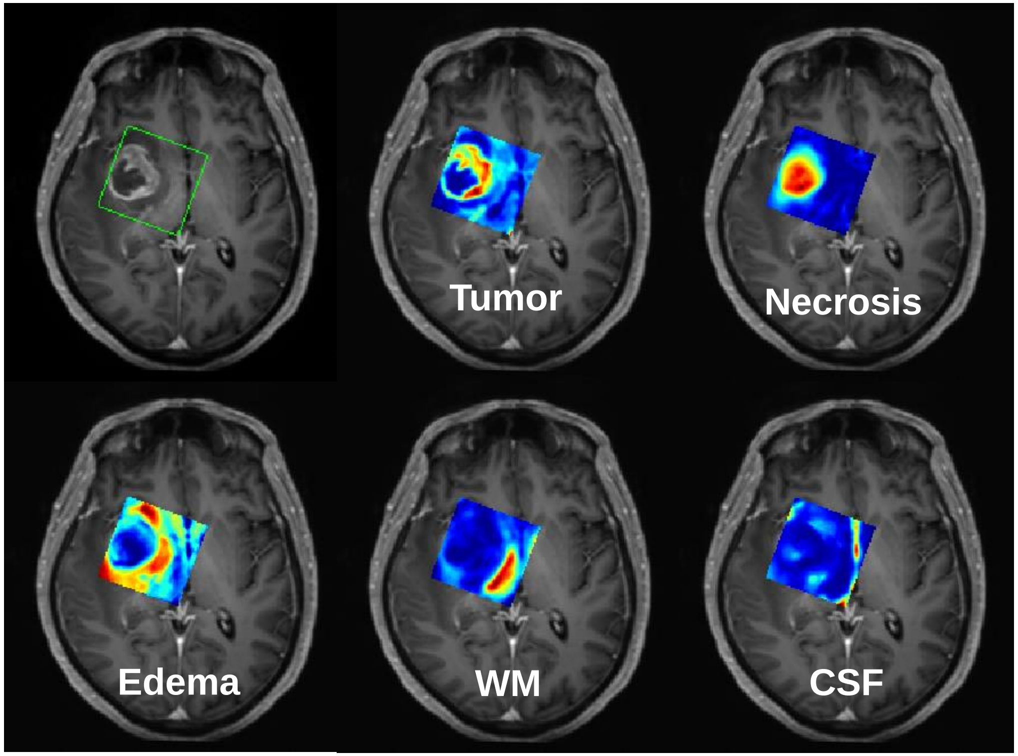
Case study: single stage NMF

5 tissue types within region of interest: active tumor, necrosis, edema, white matter, CSF

Tissue sources (W)



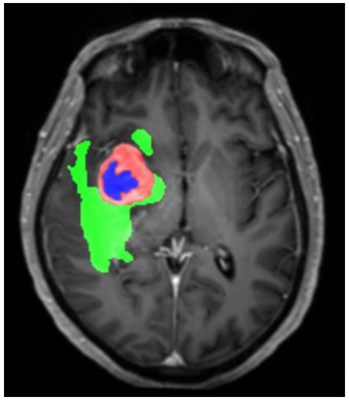
Tissue abundances (H)



Is this good/bad result?

KU LEUVEN

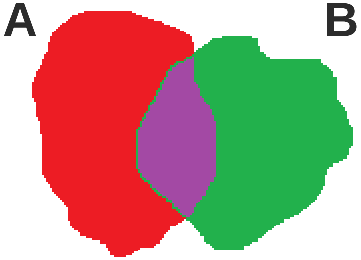
Validation



Based on manual segmentation by radiologist (only pathological tissue types)

1) **Dice-scores** (based on H)

$$Dice = 2x \frac{area(A \cap B)}{area(A) \cup area(B)}$$



2) **Correlation coefficients** (based on W)

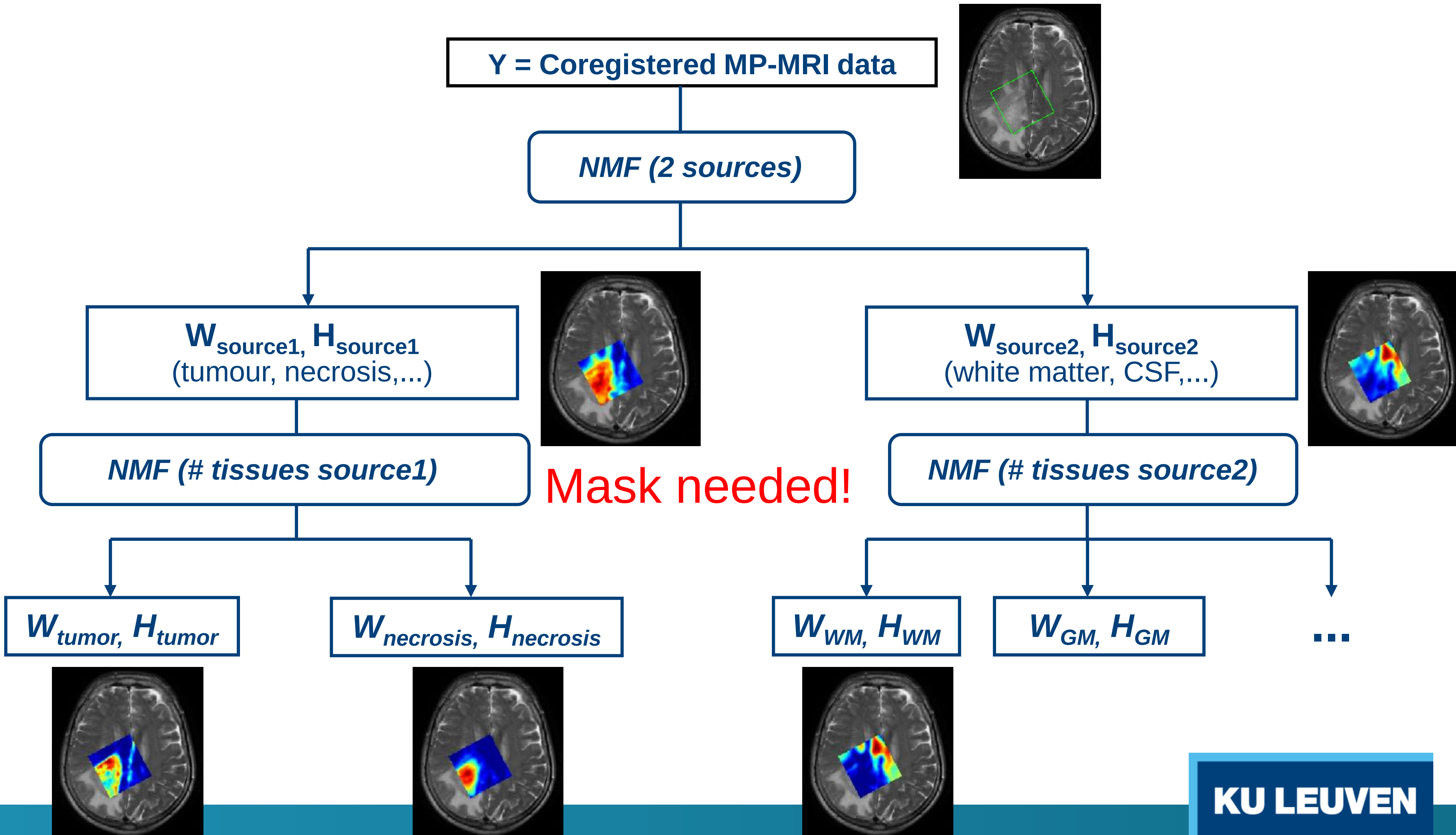
$$Corr = \frac{\vec{a}.\vec{b}}{\|\vec{a}\|.\|\vec{b}\|}$$

$\vec{a}$: tissue source vector

$\vec{b}$: average feature vector over corresponding tissue region

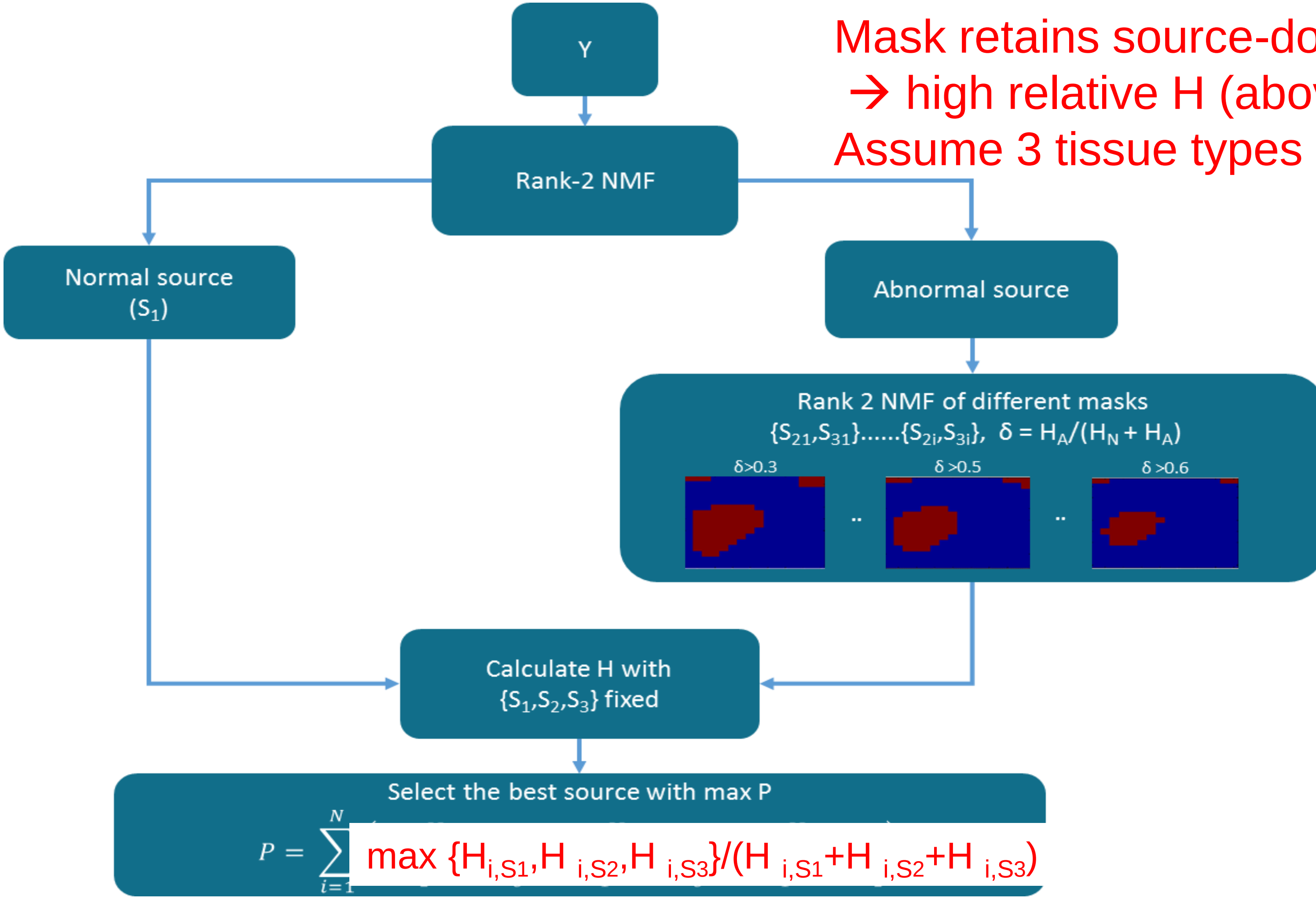
Hierarchical NMF (hNMF)

- Improved results on MRSI data only (Li et al., NMR in BioMed. 2013)



Hierarchical NMF: Select best Mask

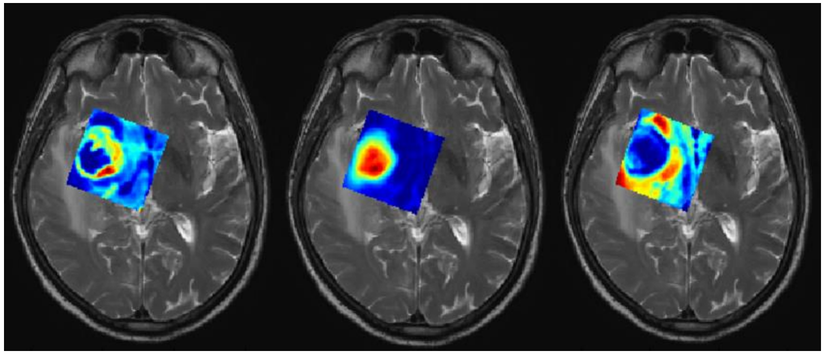
Mask retains source-dominant voxels:
→ high relative H (above threshold δ)
Assume 3 tissue types for simplicity



Sauwen N, Sima D, Van Cauter S, Veraart J, Leemans A, Maes F., Himmelreich U, Van Huffel S. Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. NMR in BioMedicine, 2015, paper in review

Case study: single stage NMF vs hNMF

Single stage NMF



$Dice_{tumor}$ = 71%

$Dice_{tumor+necrosis}$ = 83%

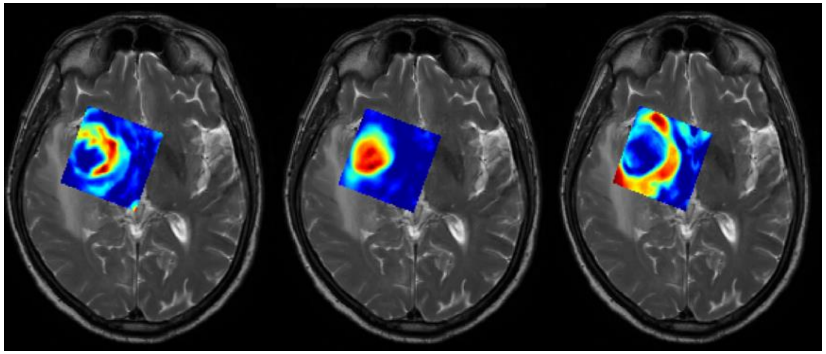
$Dice_{complete\ tumor}$ = 75%

$Corr_{tumor}$ = 0.60

$Corr_{necrosis}$ = 0.97

$Corr_{edema}$ = 0.93

hNMF



$Dice_{tumor}$ = 81%

$Dice_{tumor+necrosis}$ = 92%

$Dice_{complete\ tumor}$ = 83%

$Corr_{tumor}$ = 0.78

$Corr_{necrosis}$ = 0.98

$Corr_{edema}$ = 0.97

KU LEUVEN

Full study results: Dice-scores

HGG	Dice-score active tumor [%]						Dice-score tumor core [%] (active tumor + necrosis)						Dice-score complete tumor [%] (active tumor + necrosis + edema)					
	Full MP-MRI	no-cMRI	no-PWI	no-MRSI	no-DKI	cMRI only	**Full MP-MRI**	no-cMRI	no-PWI	no-MRSI	no-DKI	cMRI only	**Full MP-MRI**	no-cMRI	no-PWI	no-MRSI	no-DKI	cMRI only
Average	**78**	65	76	71	76	69	**85**	77	81	79	79	71	**83**	75	79	76	77	68
Std dev	**10**	14	19	16	12	18	**11**	14	18	17	14	17	**14**	15	17	18	14	18
p-value	-	0.001*	0.54	0.03*	0.18	0.004*	-	0.003*	0.28	0.01*	0.002*	0.001*	-	0.002*	0.03*	0.001*	0.003*	0.001*

** Statistical significance of lower Dice-scores compared to full MP-MRI based on one-tailed Wilcoxon signed rank test, $p<0.05$*

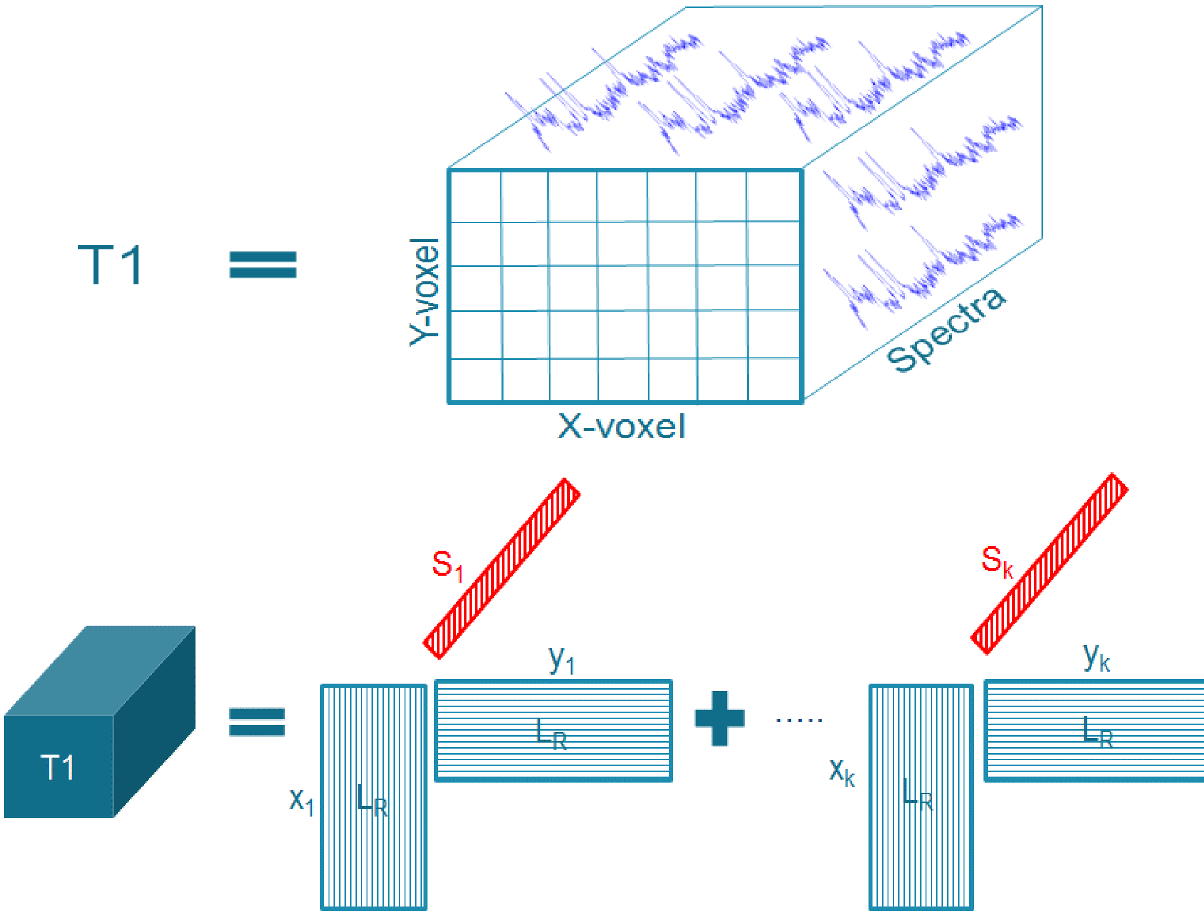
Combining 4 MRI modalities
improves brain tissue differentiation

Similar conclusions when comparing correlation coefficients

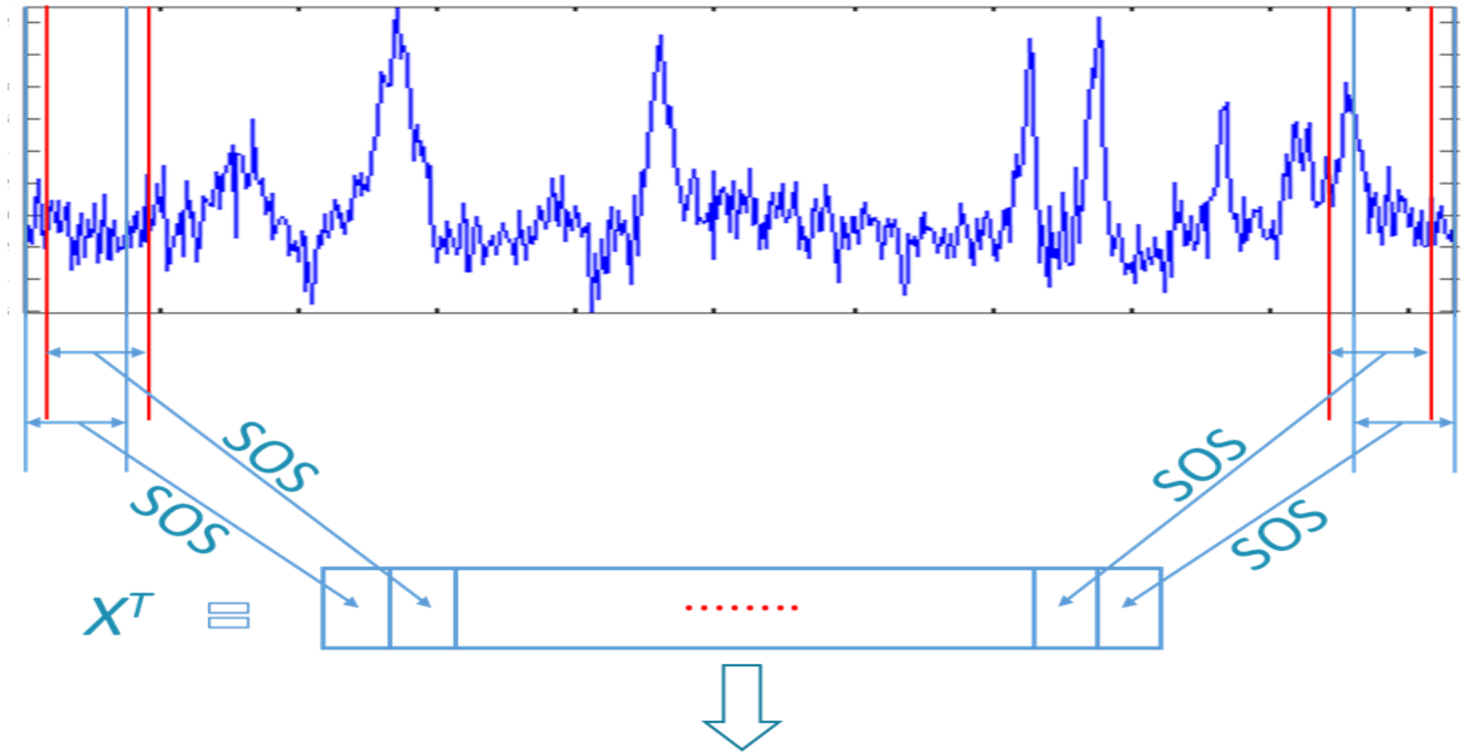
Spatial Tensor Representation (MRSI only)

- Frontal slices (XY^T) representing the spatial distribution of a tissue type does not have low rank structure.

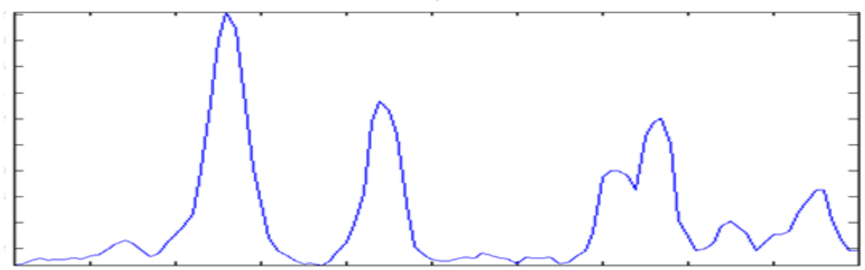
- Difficult to find the rank L_R for a particular tissue type distribution.



XX^T based Tensor Representation (MRSI only)



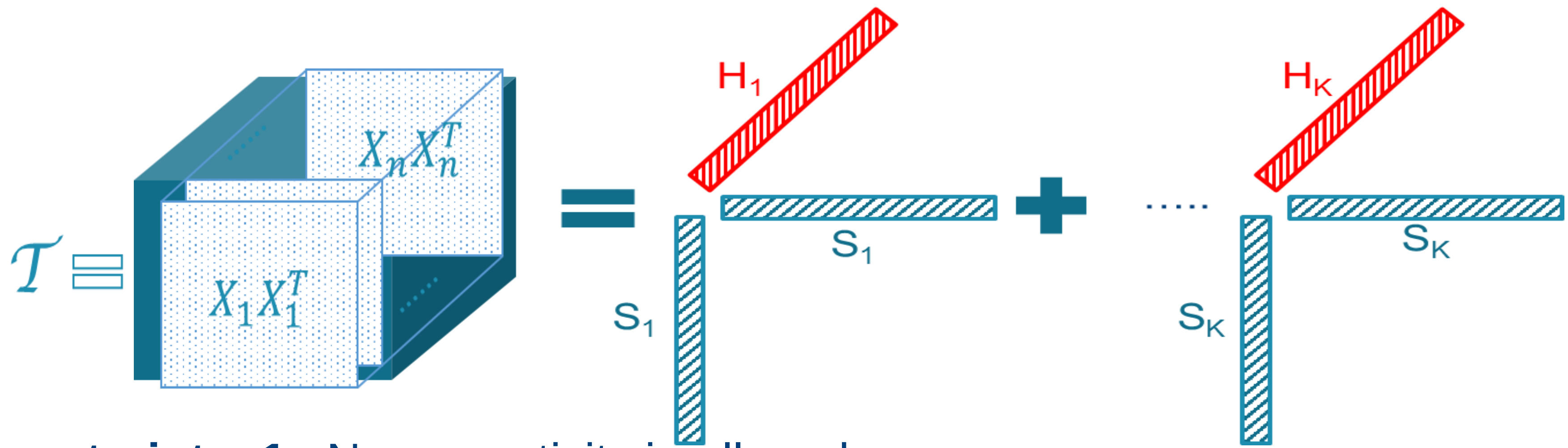
$$X(i) = \sum_{j=1}^{L} S_{ij}^2$$



- Spectra reduced in length and denoised without losing vital information,
- Peaks get higher weights,
- Peaks coupled because of XX^T in the frontal slices

HN Bharath et al, Proc. IEEE-EMB Symposium, Milan, Italy, Aug 2015, to appear

Non-negative CPD for Tumor Differentiation



- **Constraints: 1.** Non-negativity in all modes

  2. Symmetry in frontal slices

  3. Sparsity in H-factor

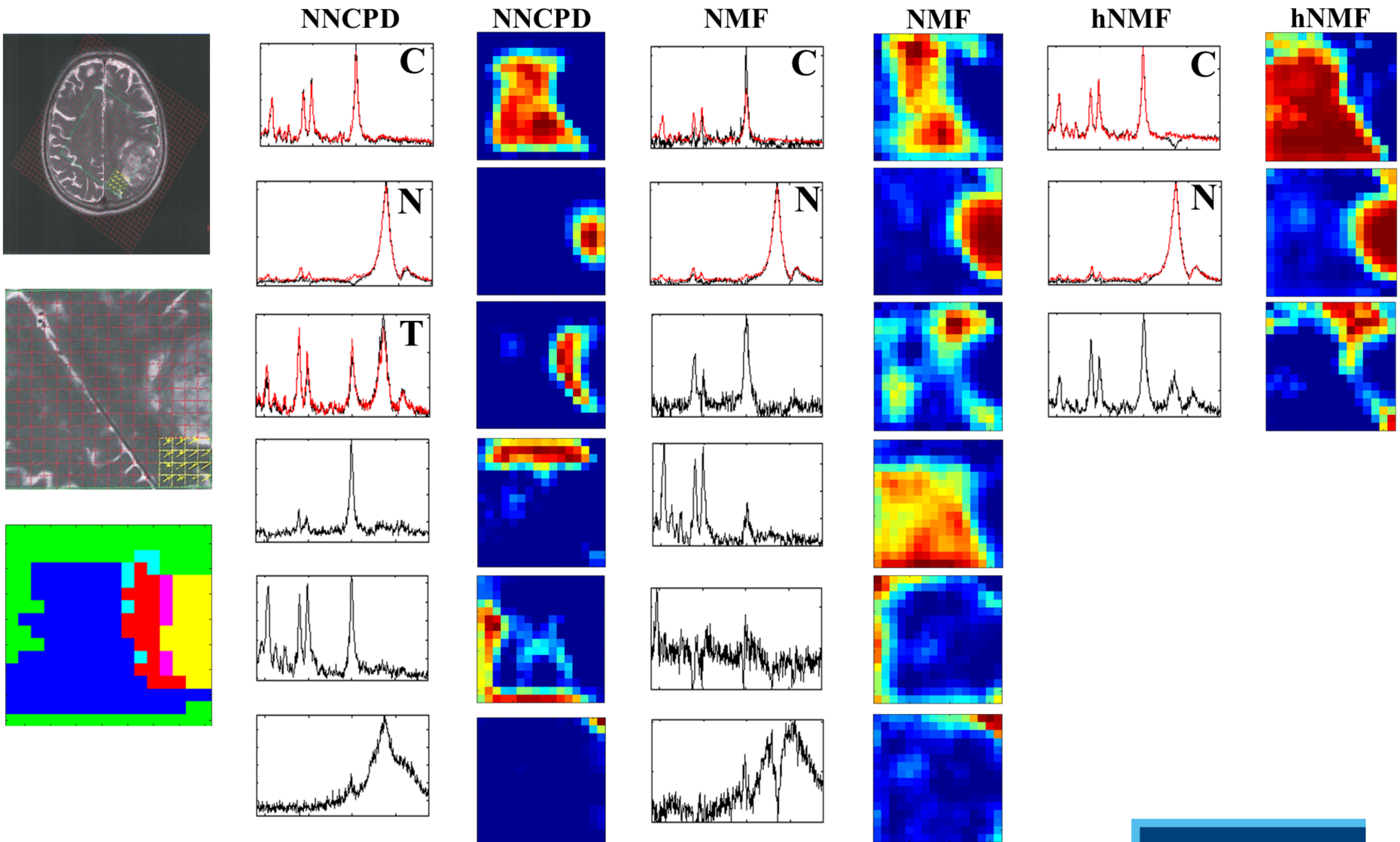
→ apply Non-negative CPD with L_1 regularization using Tensorlab[1]

$$[S^*, H^*] = \min_{S\geq 0, H\geq 0} \left\| T - \sum_{r=1}^{K} S(:,r) \circ S(:,r) \circ H(:,r), \right\|_2^2 + \lambda |Vec(H)|_1$$

- Using H*, recover tissue-specific spectra W from Y via LS
- Using W, recover tissue-type spatial distributions H from Y via NN-LS

[1] *Laurent Sorber, Marc Van Barel and Lieven De Lathauwer. Tensorlab v2.0, Available online, January 2014. URL: http://www.tensorlab.net/.*

Result: Patient-2



Source Correlation: Algorithm vs Expert labeling (MRSI only)

W		NCPD	Single stage NMF	hNMF	Grade
PATIENT-2	T	**0.99**	X	X	High
PATIENT-2	N	**0.9975**	0.9971	0.9972	High
Median/MAD	T	**0.98** /0.0112	0.67 /0.0514	0.87 /0.0093	High
Median/MAD	N	**0.9969**/0.0013	0.9945/0.0028	0.9967/0.0005	

Abundance map Correlation: Algorithm vs Expert labeling

H		NCPD	Single stage NMF	hNMF	Grade
PATIENT-2	T	**0.80**	X	X	High
PATIENT-2	N	**0.9552**	0.8949	0.8967	High
Median/MAD	T	**0.79** /0.0380	0.69 /0.0849	0.69 /0.0731	High
Median/MAD	N	**0.87**42/0.0276	0.7876/0.0662	0.8009/0.0716	

KU LEUVEN

Contents Overview

- Introduction
- Examples in Neonatal brain monitoring
- Examples in MR-based Brain tumor diagnosis
- Conclusions and new directions

Conclusions and new directions

- Many BSS problems in Smart Diagnostics are low rank
  → solve via matrix or tensor factorization plus constraints

- Successful examples shown, e.g., in neonatal brain monitoring, brain tissue typing

- Extensions to biomedical data fusion emerge, e.g. EEG-fMRI
  → solve via coupled matrix /tensor decompositions

- Other BSS applications: *bioinformatics (O. Alter, E. Acar), BCI (Cichocki, Mørup, Martinez-Montes), mobile EEG, multichannel ECG*

New directions?

- *Adaptive tensor decompositions, rank & structure estimation*
- *Applications increasing in BCI, (single-trial) ERP, ECG, MRSI*

→ *exploit full potential of Tensor toolbox*

Acknowledgment

University Hospitals Leuven Gasthuisberg
ZNA Middelheim, Queen Paola Children's hospital
EMC Rotterdam
KU Leuven, Dept. Electrical Engineering-ESAT, division STADIUS & MICAS
Ghent University, Dept. Telecommunication and Information Processing, TELIN-IPI
Eindhoven University of Technology

ERC advanced grant 339804 BIOTENSORS
in collaboration with L. De Lathauwer and group

Thank you!

TDA 2016 @ KU Leuven

http://www.esat.kuleuven.be/stadius/TDA2016/

Workshop on Tensor Decompositions and Applications
January 18 - 22, 2016, Leuven, Belgium

Local Organisers: Sabine Van Huffel and Lieven De Lathauwer

Confirmed Speakers

Orly Alter
Pierre Comon
Eva Ceulemans
Harm Derksen

Nicolas Gilllis
Daniel Kressner
Lek-Heng Lim
Ivan Markovsky

Morten Mørup
Nikos Sidiropoulos
Bart Vandereycken
Frank Verstraete

European Research Council
Established by the European Commission

siam

I L A S

KU LEUVEN