Multimodal Imaging and BCI

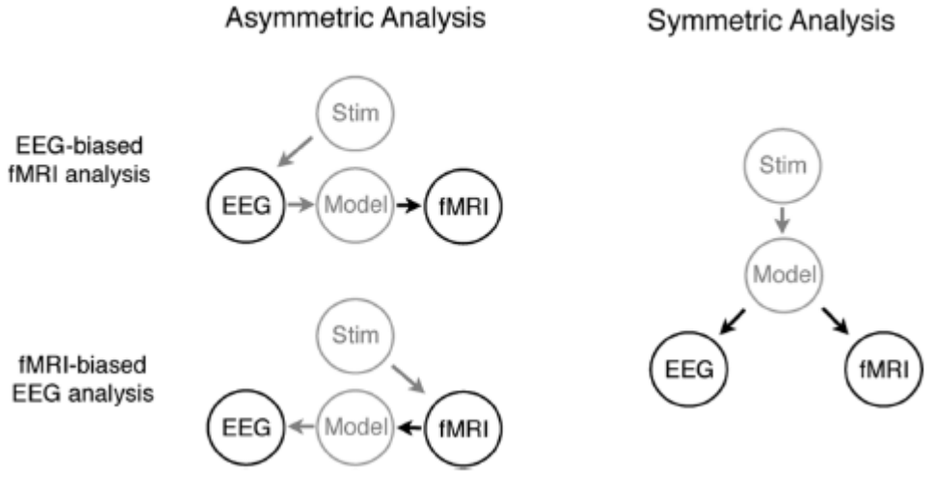
berlin
brain computer
interface

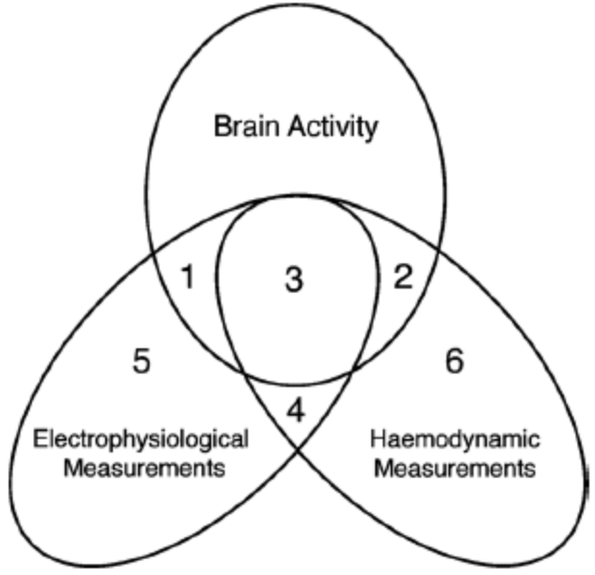
Klaus-Robert Müller, Siamac Fazli, Jan Mehnert, Stefan Haufe, Frank Meinecke, Wojciech Samek, Felix Biessmann, Sven Dähne, Gabriel Curio, Benjamin Blankertz et al.

Analysis of Multimodal Neuroimaging Data

Felix Bießmann, Sergey Plis, Frank C. Meinecke, Tom Eichele, and Klaus-Robert Müller

al Review





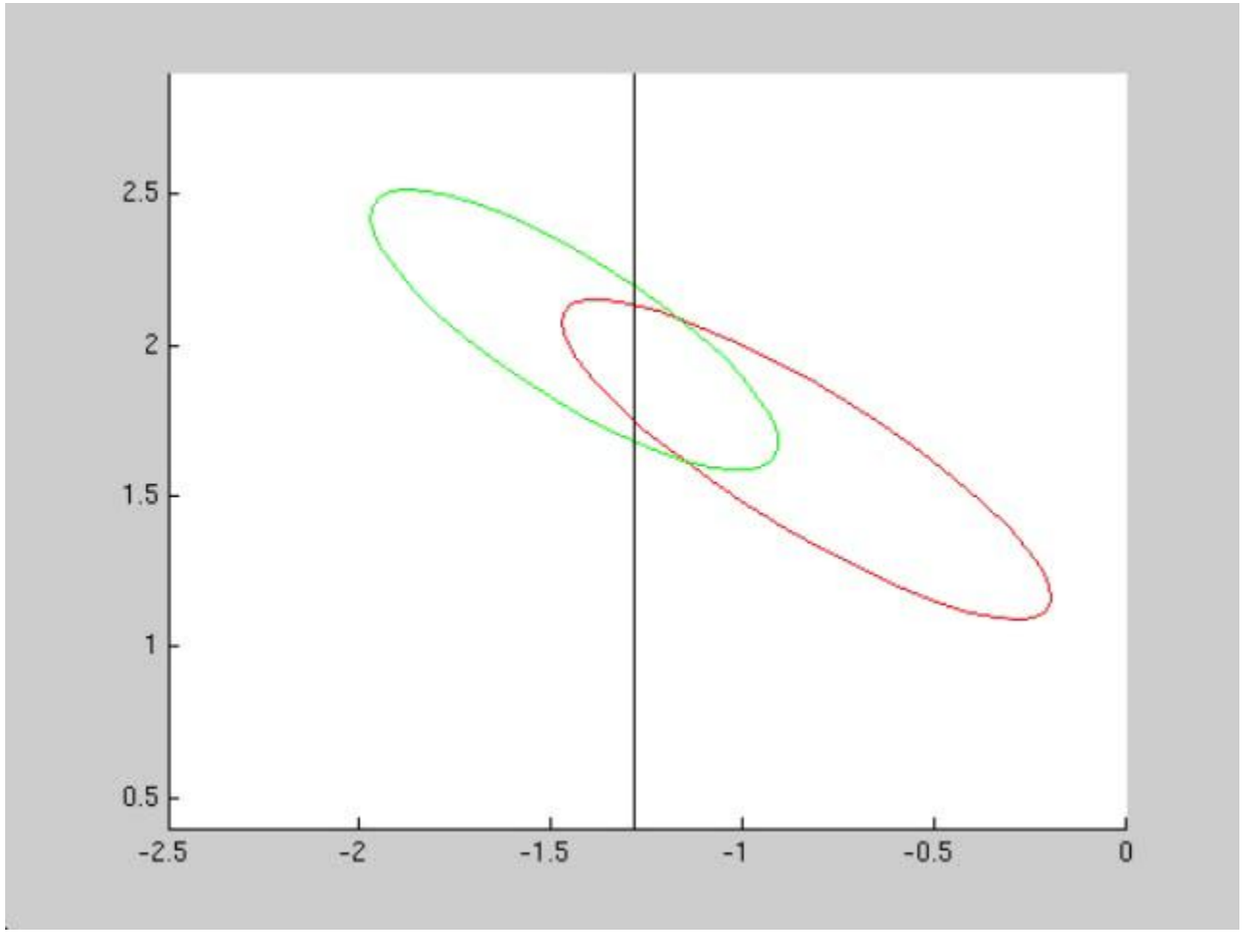
Multimodal Data

- Combining **multiple** physiological **features** (oscillations, SCPs, maps)
  (cf. Dornhege et al 2006)
- Combining **multiple subjects** data (cf. nonstationarity)
  (cf. Fazli et al 2009 and 2011, Samek, Meinecke & Müller in Press)

- **Correlating apples & oranges** i.e. computing correlations between multiple measuring modalities (EEG & EMG, EEG & NIRS, EEG & fMRI, LFP & fMRI)
  (cf. tkCCA Biessmann et al 2010, 2011, 2012)
- Combining **multiple measuring modalities** (EEG & EMG, EEG & NIRS, EEG & fMRI)
  (see Fazli et al 2012, Biessmann et al 2011; Pfurtscheller, Müller-Putz, Calhoun, Adali, Ritter et al, Cohen, Villringer, Eichele,)
- **Nonlinear correlations between modalities** NIRS & EEG
  (see Dähne et al to appear)

Multimodal $\longleftrightarrow$ Nonstationary

Motivation: Shifting distributions within experiment



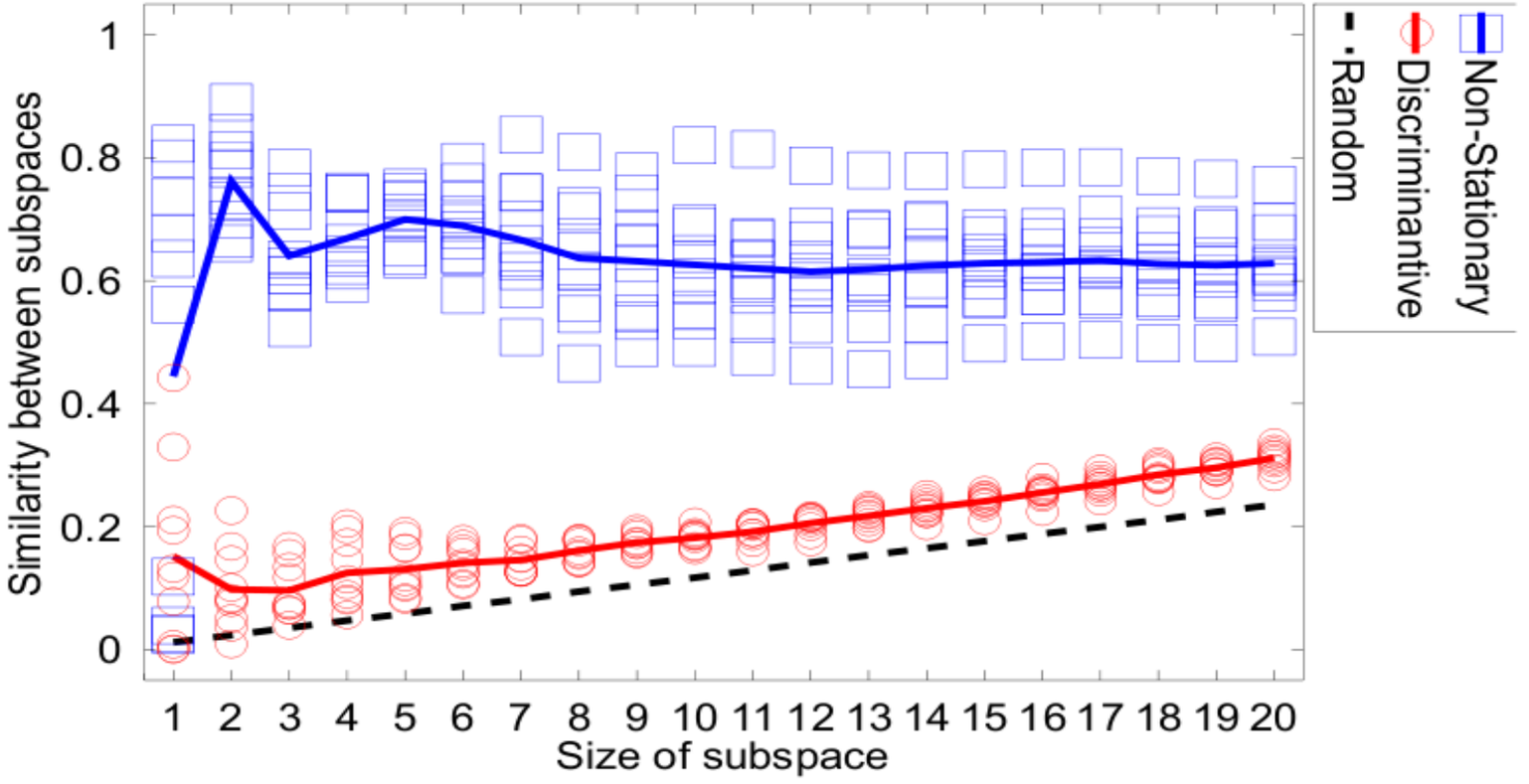
But: Is the nonstationarity different between subjects, i.e. could we learn it from other subjects?

Changes are similar !

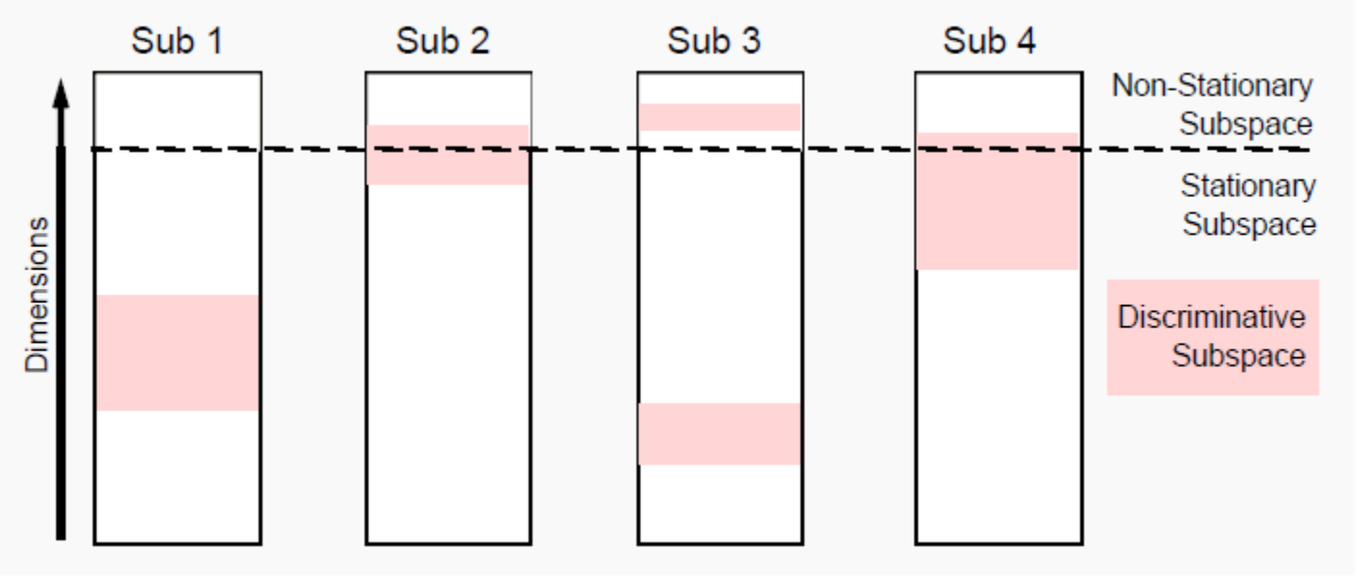
Modalities = Other Subjects

Changes between training and test data are similar between users.

Other multi-subject methods, e.g. cov matrix shrinkage, may improve estimation quality but do not reduce non-stationarities.



Cartoon: learn from adverse nonstationary subspace across subjects



Usually discriminative information is transfered between subjects.

Algorithm

(1) For each subject $i = 1 \ldots n,\ i \neq i^*$ compute the eigenvectors $\mathbf{v}_i^{(1)} \ldots \mathbf{v}_i^{(d)}$ of $\mathbf{\Sigma}_i^{train} - \mathbf{\Sigma}_i^{test}$.

(2) For each subject i select the l eigenvectors with largest absolute eigenvalues.

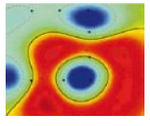
(3) Aggregate the vectors into a matrix P.

(4) Apply PCA to reduce the dimensionality of the non-stationary subspace $\mathcal{S}_P = \mathrm{span}(P)$ to ν.

(5) Compute the projection matrix $P^{\perp}$ to the orthogonal complement of $\mathcal{S}_P$.

(6) Make i^*s data invariant to the changes by projecting out non-stationarities $\tilde{\mathbf{X}} = (P^{\perp})^T P^{\perp} \mathbf{X}$.

(7) Compute spatial filters from $\tilde{\mathbf{X}}$ using CSP.



Results

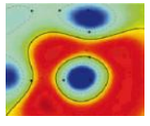
Two data sets with different stimulus cues in training and test

1. visual cue in training & auditory cue in test
2. letters in training & moving objects in test

The size of the non-stationary subspace is determined by CV in a leave-one-subject-out manner on the other users.

	Audio-Visual Data Set					BCI Competition III					Overall		
Subject	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Mean	Median	Std
CSP	79.5	80.0	65.8	59.2	94.2	66.1	96.4	58.2	88.8	81.0	76.9	79.8	14.0
ssCSP	87.1	80.8	67.5	65.8	93.3	67.0	94.6	58.2	89.3	85.7	78.9	83.3	13.1

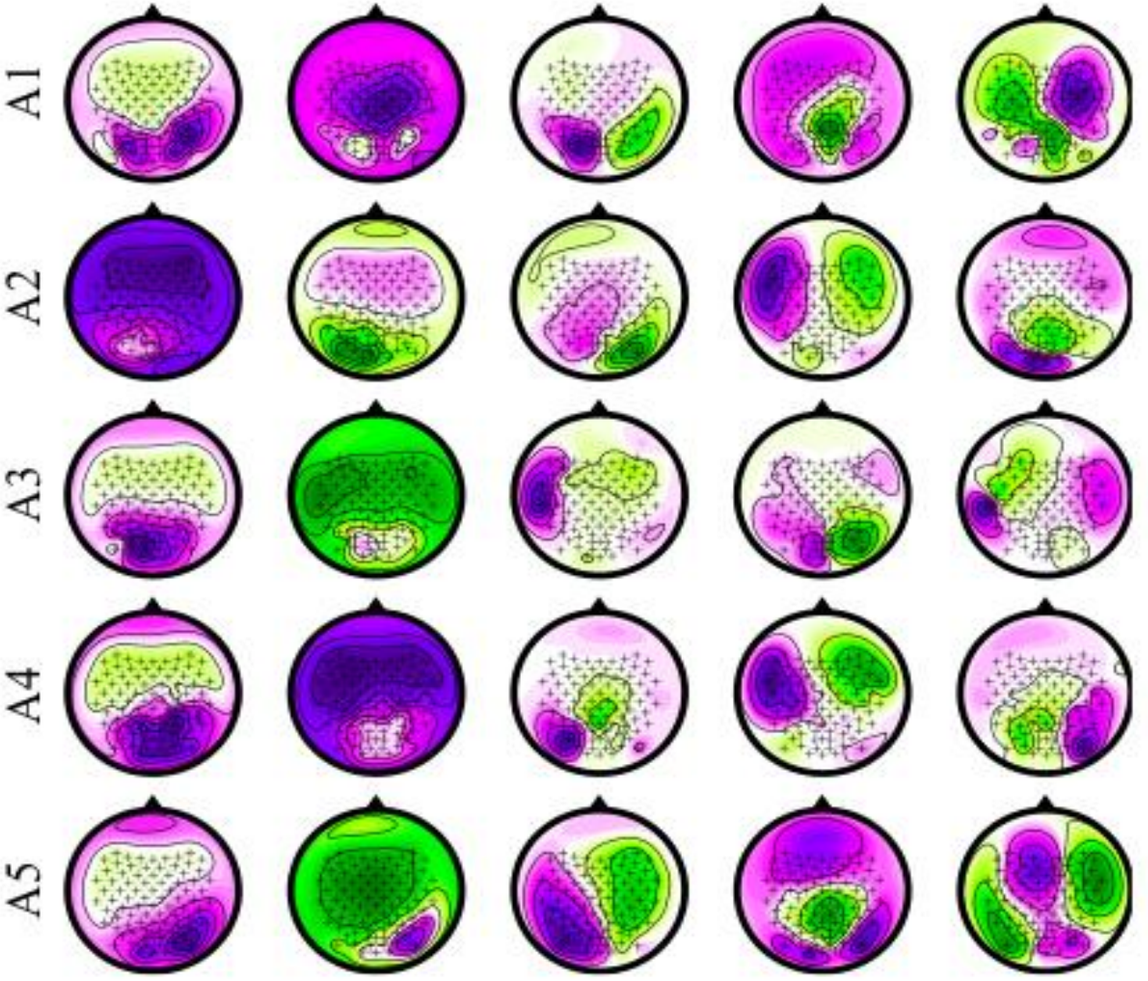
ssCSP: stationary subspace CSP



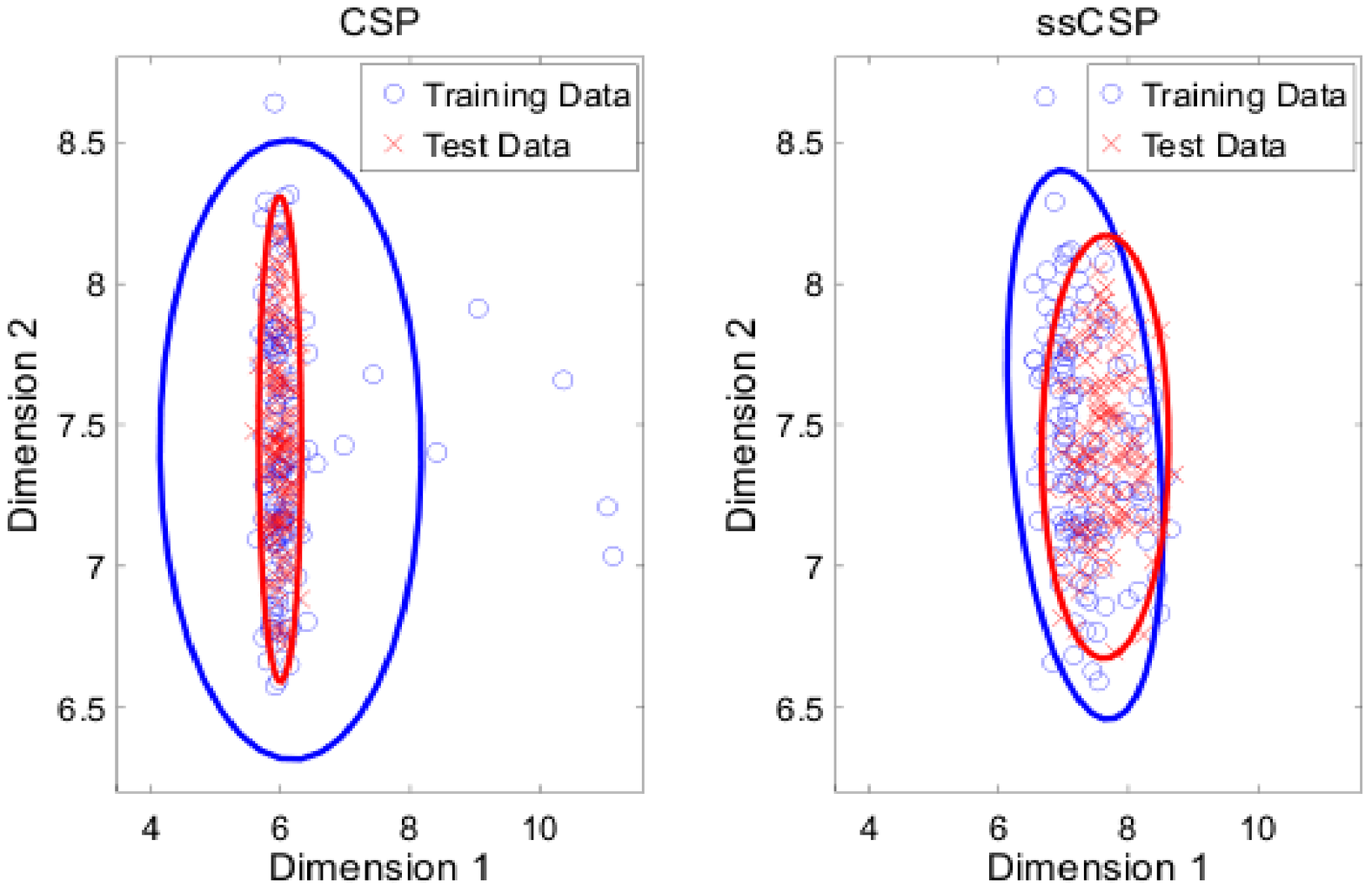
Interpretation

The most non-stationary directions are very similar between users.

Activity in occipital and temporal areas is penalized as these regions are mainly responsible for visual and auditory processing.

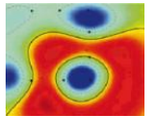


Feature distribution becomes stationary



Summary Part I

- Novel “multi-modal“ approach to reduce non-stationarities in data
- In contrast to other multi-subject methods it does NOT transfer discriminative information, thus is more robust if subject similarity is low.
- Non-stationary information appears physiologically interpretable and meaningful.
- The idea of transfering stationary subspaces between subjects can be applied to many other problems.



Samek, Meinecke & Müller IEEE TBME in Press

NIRS-EEG Brain Computer Interfaces

[Fazli et al. Neuroimage 2012]

Experimental Setup and Paradigm

EEG: 37 electrodes

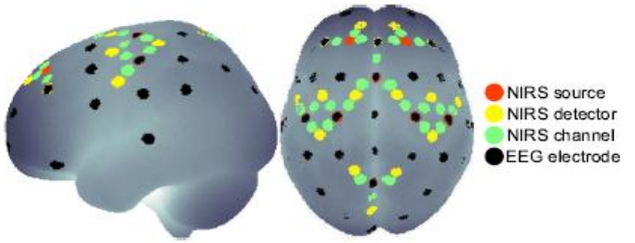
NIRS 26 channels (frontal, parietal, occipital)

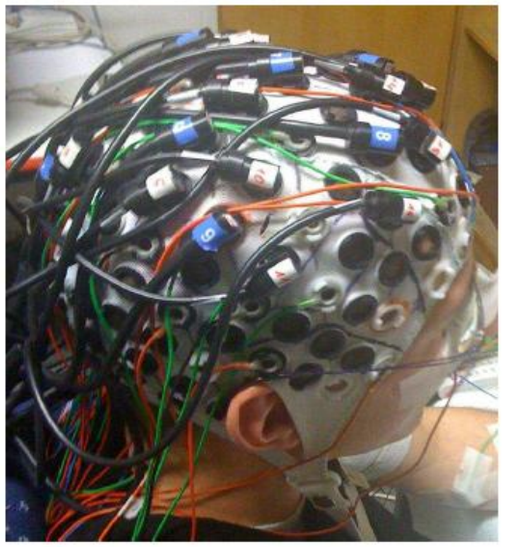
EEG-based cursor feedback (ISI = 15 s)

Executed movement vs imagery movements

Imagery movements: EEG-feedback for left and right motor imagery

Number of subjects: 14





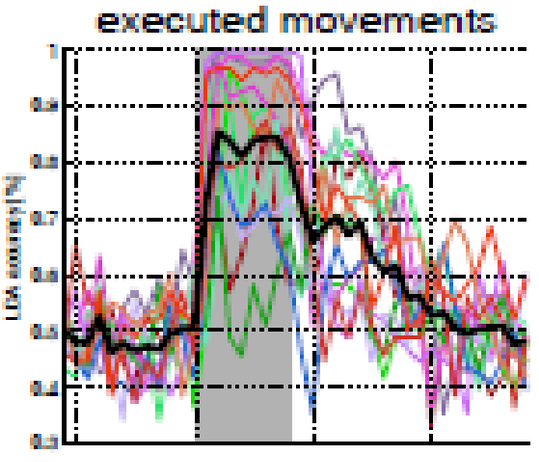
Can a simultaneous measurement of NIRS and EEG during Brain Computer Interfacing enhance the classification accuracy?

Are the results physiologically reliable?

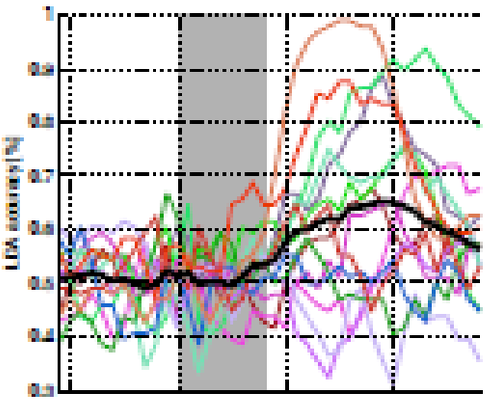
Fazli et al. 2012

Temporal Dependency of Classification in Executed Movements

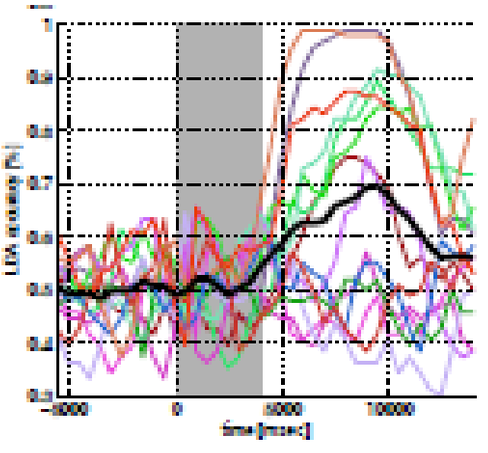
EEG



HbO



HbR

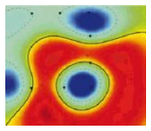


Fazli et al. 2012

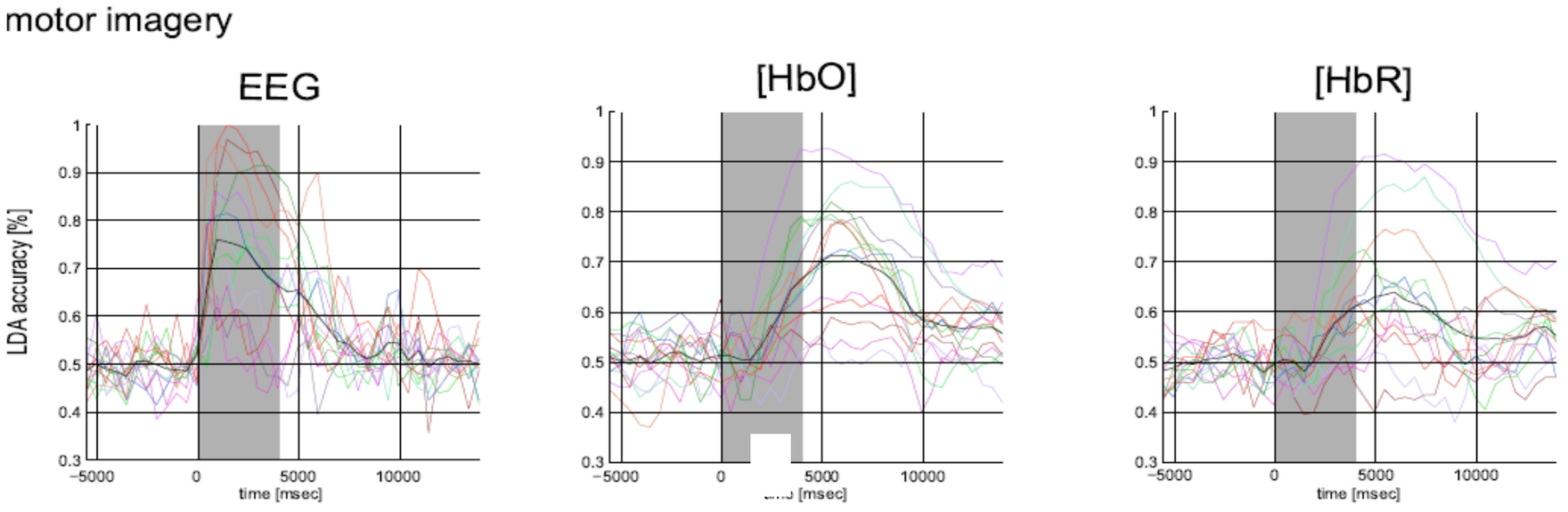
EEG peaks earlier as compared to HbO and HbR

Physiological reliability: HRF shaped classification accuracies over time

Classification accuracy higher for EEG



Temporal Dependency of Classification in Motor Imagery

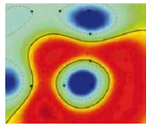


EEG peaks earlier as compared to HbO and HbR

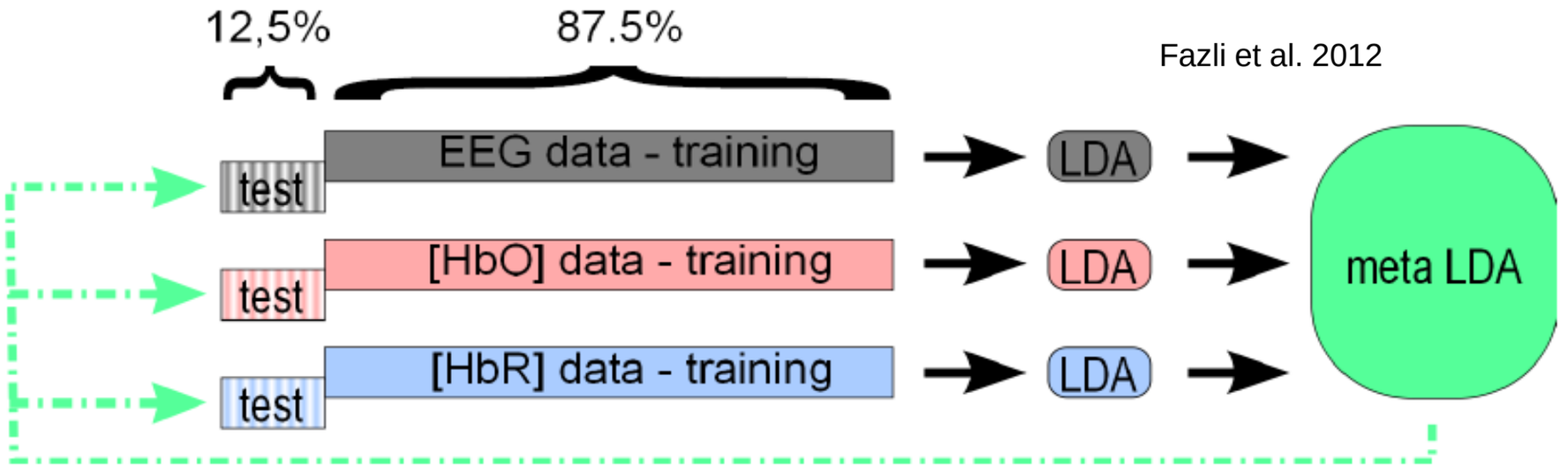
Physiological reliability: HRF shaped classification accuracies over time

Classification accuracy higher for EEG

Classification accuracy lower than in executed movements



Combination of EEG and NIRS



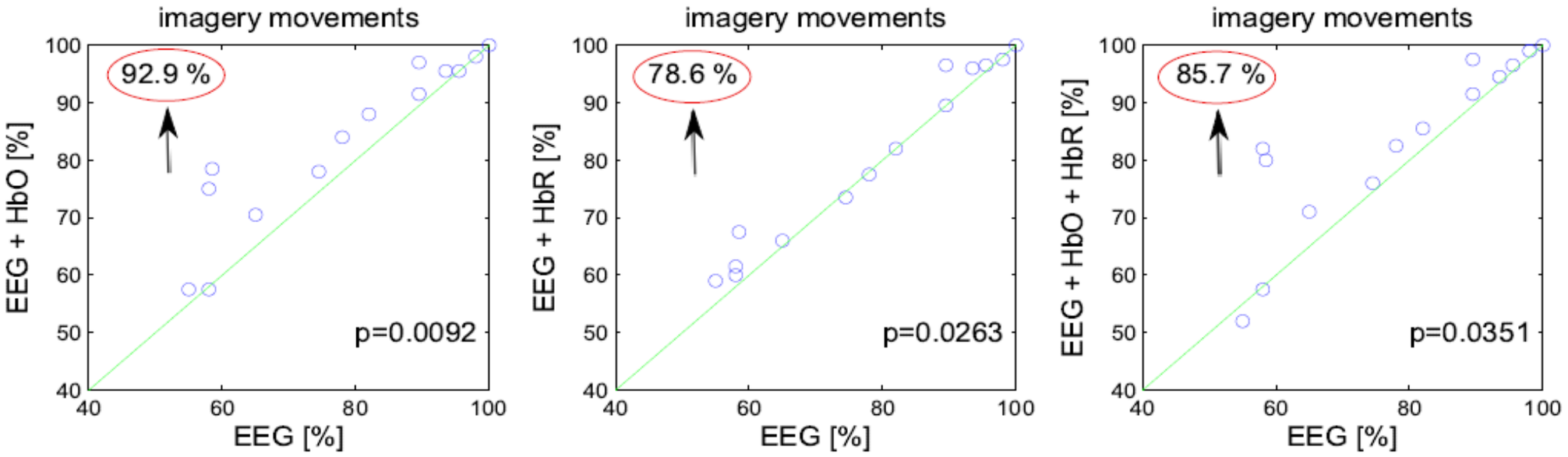
LDA classifier estimated for EEG, HbO and HbR (individually)

Meta-classifier estimated for combination in each subject

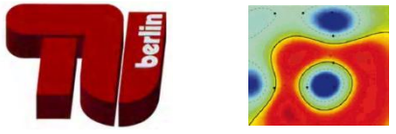
All within cross-validation (8 chronological splits)

Feature Combination

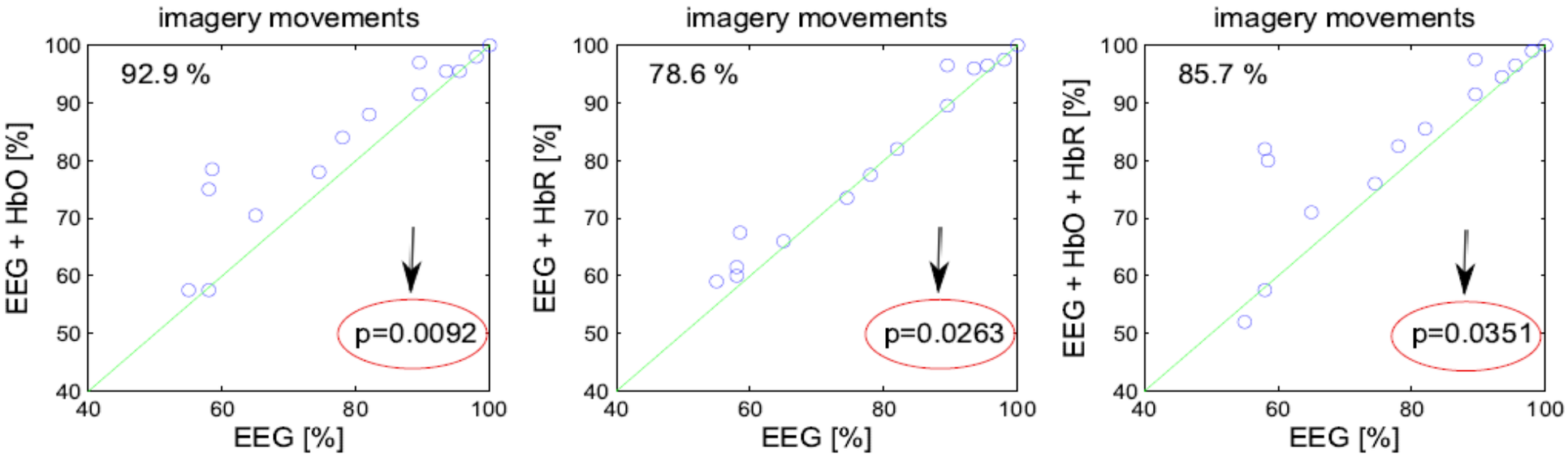
Fazli et al. 2012



NIRS-EEG combinations have higher classification accuracies for vast majority of subjects



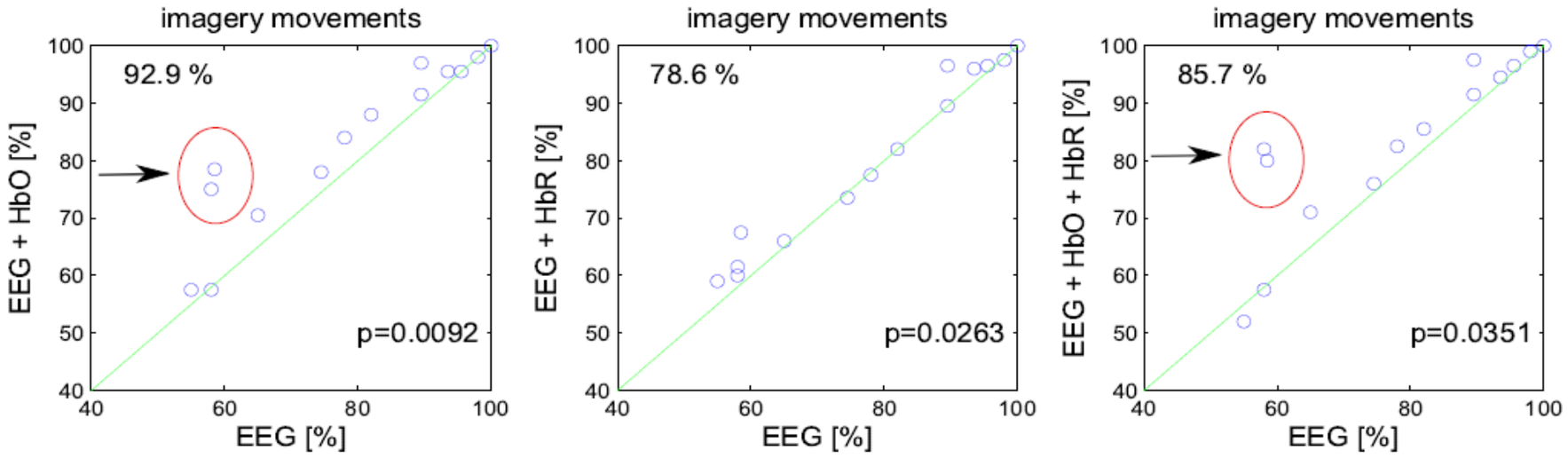
Feature Combination



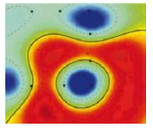
Fazli et al. 2012

t-tests reveal a significant increase of classification accuracy for combination

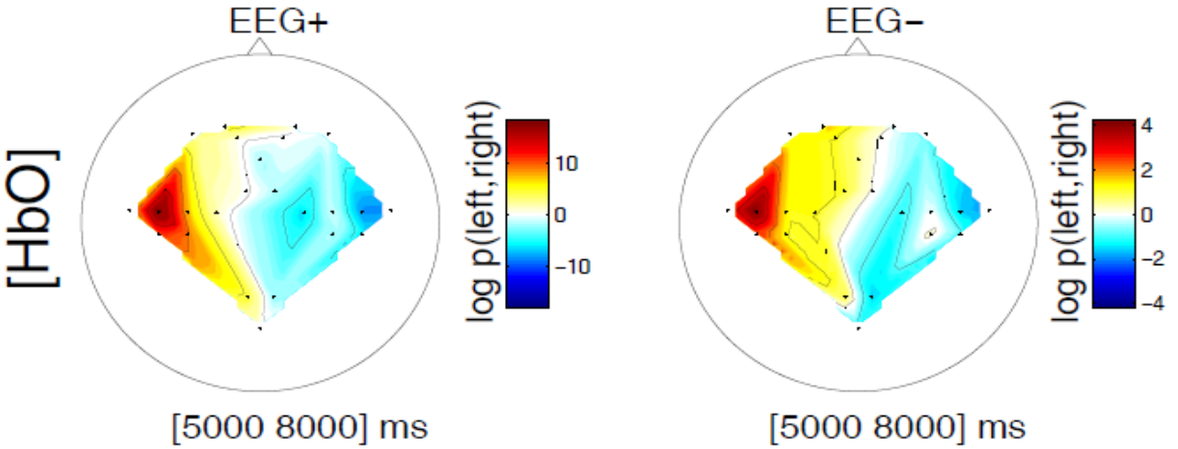
Feature Combination



Some subjects, which were not classifiable with EEG become classifiable by a meta-classifier in combination with NIRS



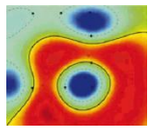
Mutual Information



NIRS features for all correct EEG trials (EEG+) and incorrect EEG trials (EEG-)

Pattern is similar although the significance drops

NIRS can complement the EEG with physiological meaningful information



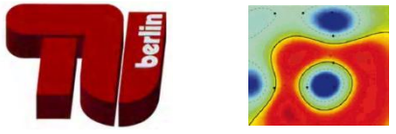
Discussion Part II

Problems

- Different temporal properties of the measurement devices (e.g. EEG: 1000 Hz, NIRS: max. 10 Hz)
- Temporal lag between parameters
- Different signal qualities

Ideas to Overcome the Temporal Lag

- NIRS as a measure of subjects' attention to predict EEG-based performance
- NIRS as a localizer of the source of EEG signals
- NIRS as a 'stop', e.g. to discard a EEG-based classified trial when not confirmed by NIRS



Correlating apples and oranges

[Biessmann et al. Neuroimage 2012, Machine Learning 2010]

CCA: correlating apples and oranges

Given two (or more) multivariate variables

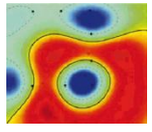
$$X \in \mathbb{R}^M, Y \in \mathbb{R}^N$$

CCA finds projections

$$w_x \in \mathbb{R}^M, w_y \in \mathbb{R}^N$$

that maximise the covariance between the variables

$$\begin{bmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} = \alpha \begin{bmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix}$$



kCCA: solving CCA on data kernels

Intuition behind the Kernel Trick:

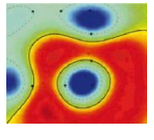
The solution of CCA in kernel space is obtained by solving the generalised eigenvalue problem

$$\begin{bmatrix} 0 & K_x K_y \\ K_y K_x & 0 \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} = \rho \begin{bmatrix} K_x^2 & 0 \\ 0 & K_y^2 \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix}$$

The solutions in the input space can be recovered by

$$w_x = X\alpha_x$$
$$w_y = Y\alpha_y$$

No need to compute big covariance matrices!

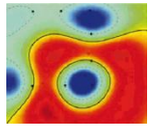


tkCCA: correlating apples and oranges over time

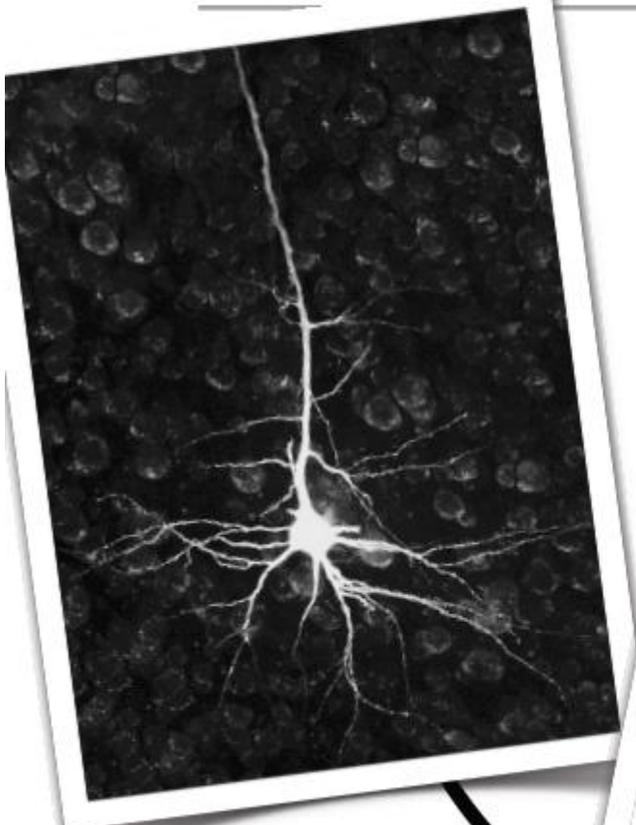
$$\underset{w_x(\tau), w_y}{\operatorname{argmax}} \operatorname{Corr} \left(\sum_{\tau} w_x(\tau)^\top x(t-\tau),\ w_y^\top y(t) \right)$$

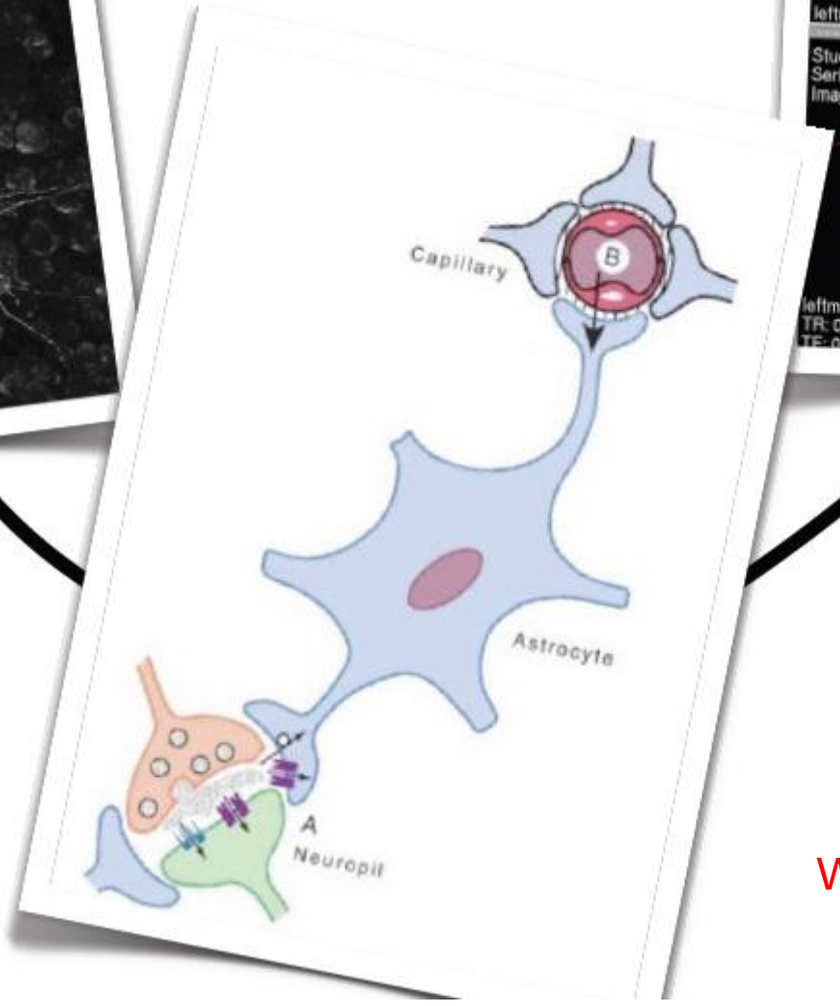
$$\tilde{X} = \begin{bmatrix} X_{\tau_1} \\ X_{\tau_2} \\ \vdots \\ X_{\tau_T} \end{bmatrix} \qquad \tilde{w}_x = \begin{bmatrix} w_x(\tau_1) \\ w_x(\tau_2) \\ \vdots \\ w_x(\tau_T) \end{bmatrix}$$

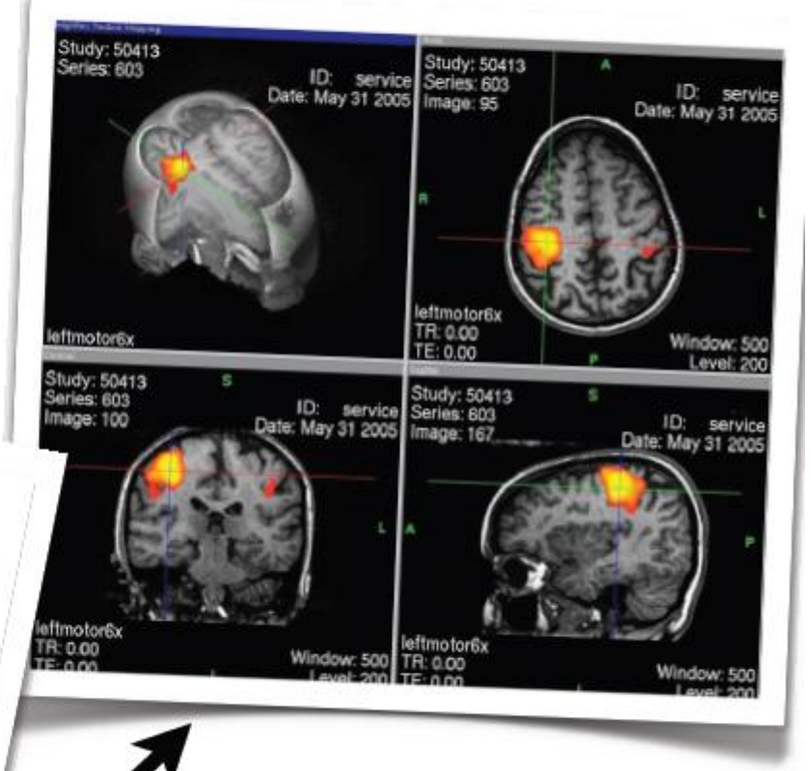
$$\underset{w_{\tilde{x}}, w_y}{\operatorname{argmax}} \operatorname{Corr} \left(\tilde{w}_x^\top \tilde{X}, w_y^\top Y \right)$$



Application: Neuro-Vascular Coupling



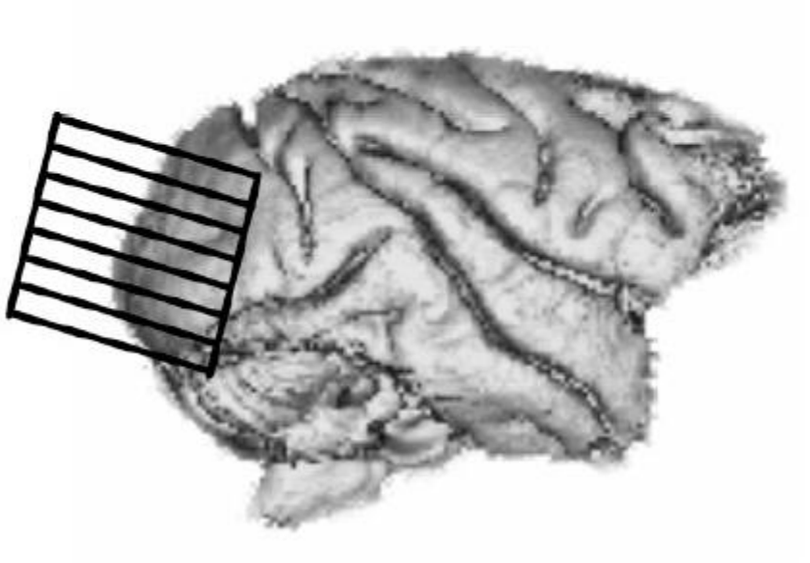


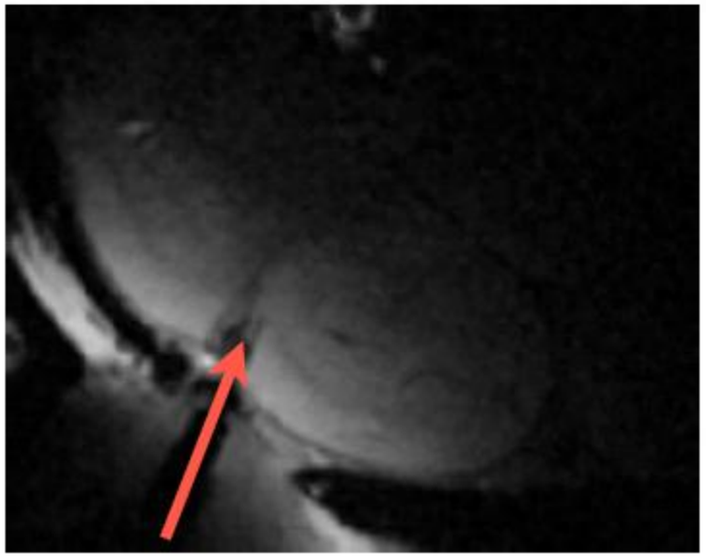


With Logothetis

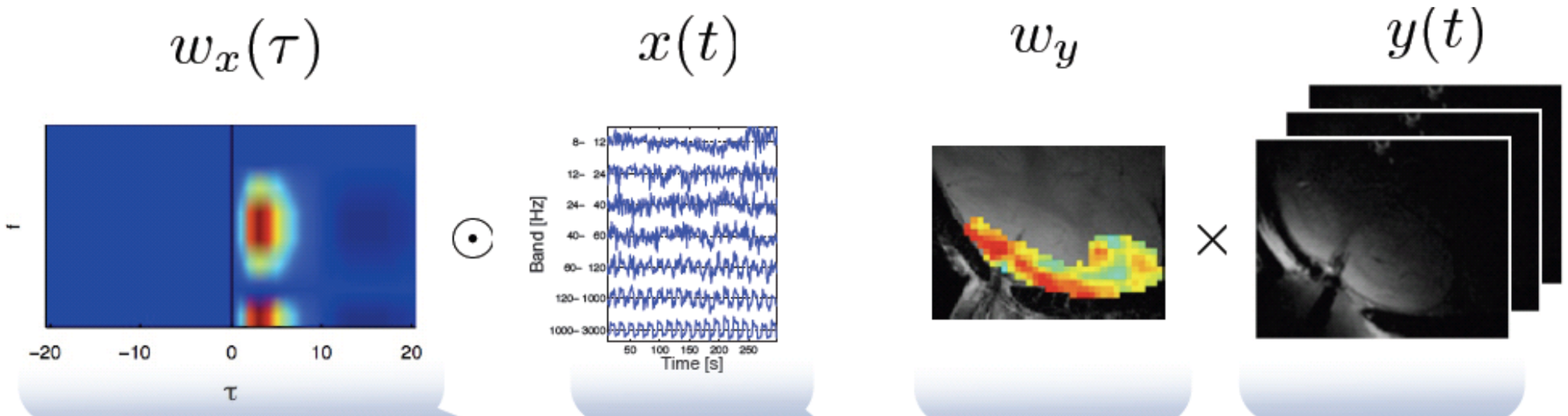
Experimental Setup

- **Simultaneous measurements of**
  - fMRI/ BOLD signal
  - Intracortical neural activity





Temporal Kernel CCA

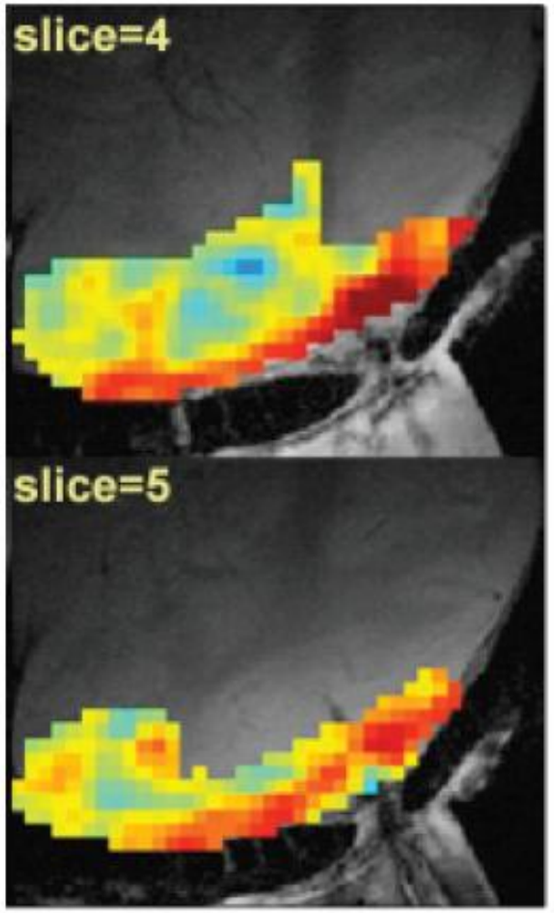


$$\underset{w_x(\tau), w_y}{\operatorname{argmax}} \operatorname{Corr}\left(\sum_{\tau} w_x(\tau)^\top x(t-\tau),\ w_y^\top y(t)\right)$$

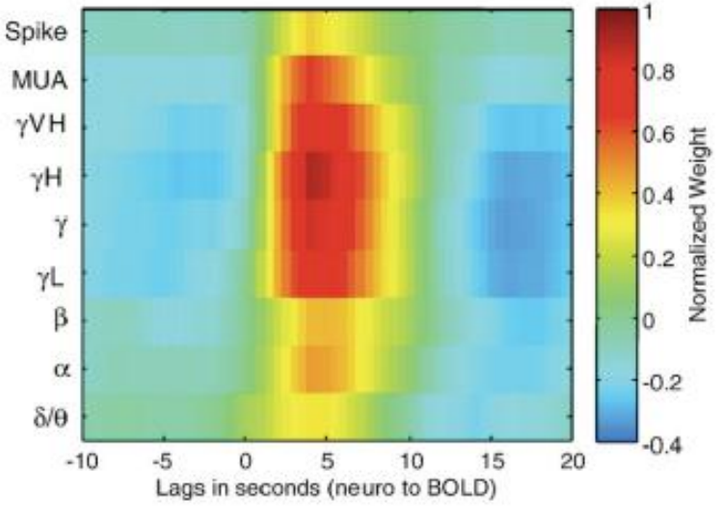
multivariate convolution of the neurophysiological signal with frequency dependent HRF

spatial weighting of voxels with activation pattern.

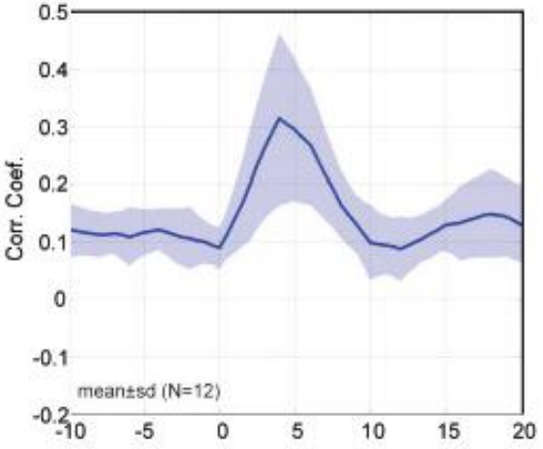
Results tkCCA: spatial dependencies and HRF



Spatial Dependencies



Haemodynamic Response Function



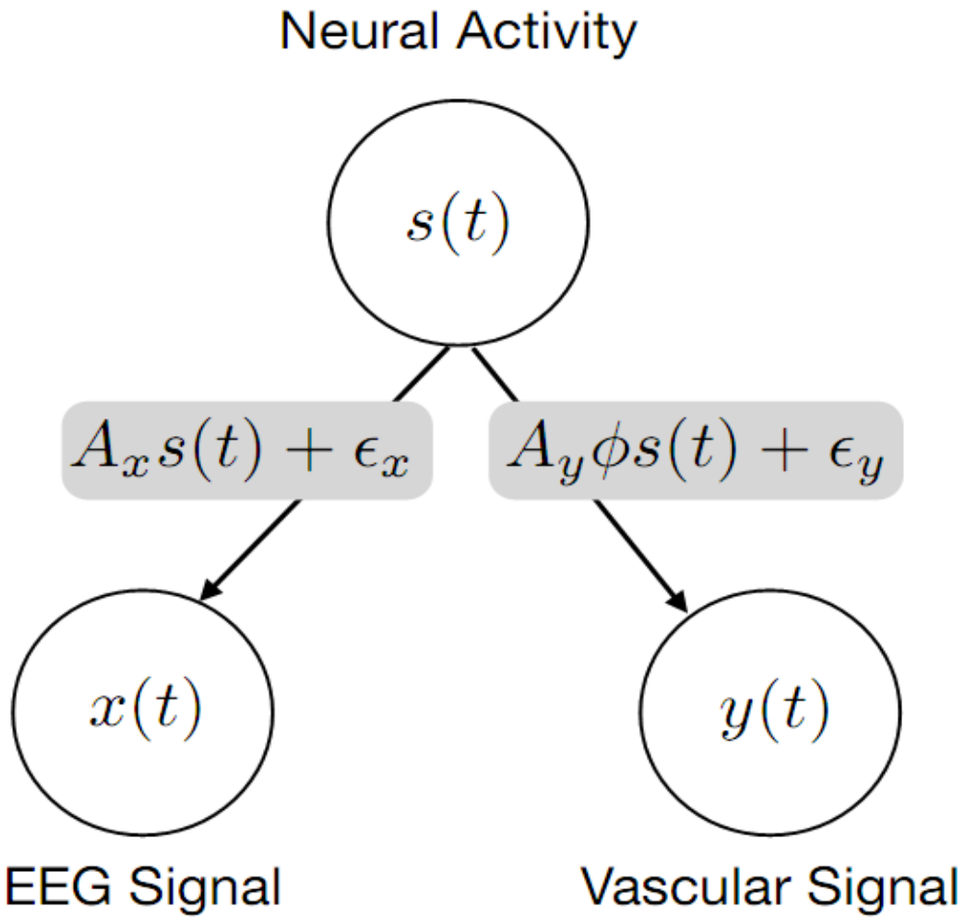
Canonical Correlogram

Murayama et al., "*Relationship between neural and haemodynamic signals during spontaneous activity studied with temporal kernel CCA*", Magnetic Resonance Imaging, 2010

Finding nonlinear correlations between NIRS &EEG

[Dähne, Biessmann et al. In Press]

Generative Model



Mapping to sensor space:

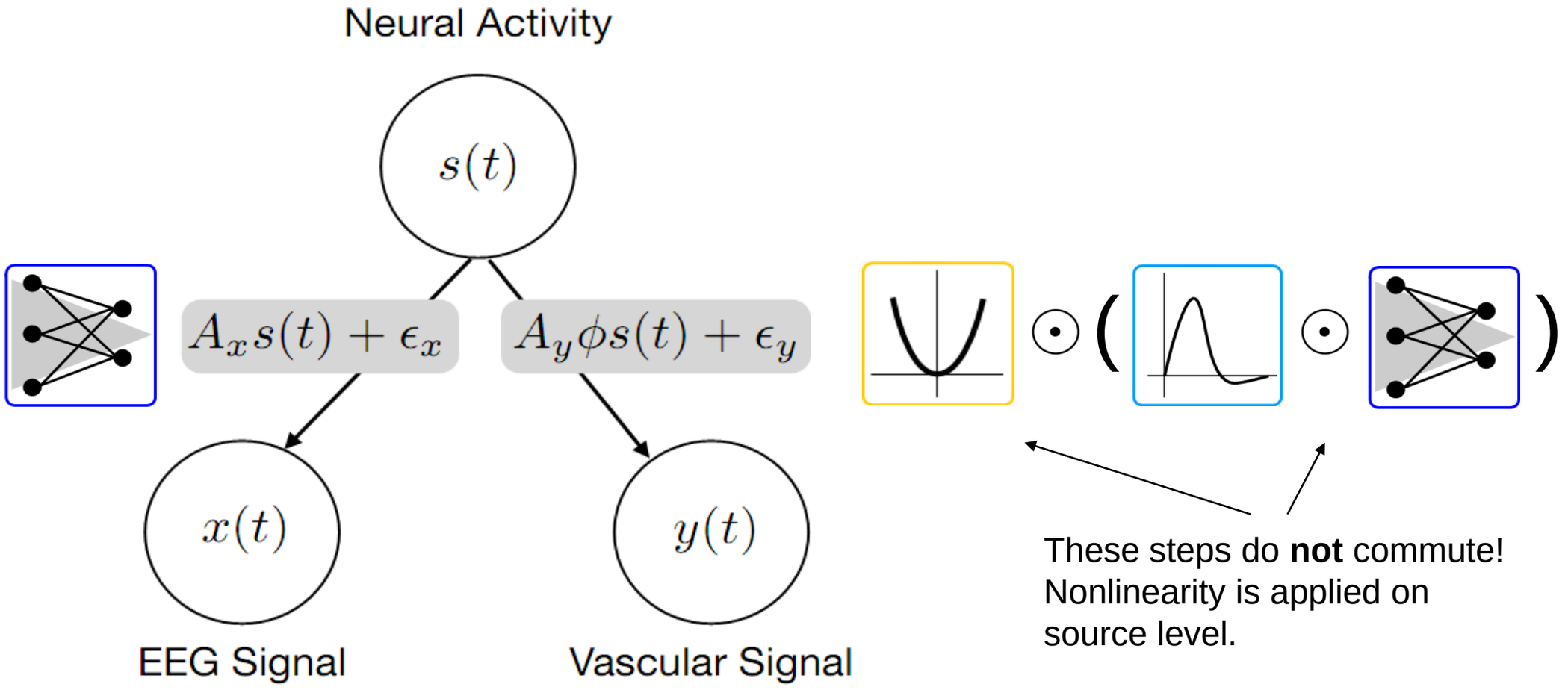
$$\mathbf{x}(t) = \mathbf{A}_x \mathbf{s}(t) + \epsilon_x(t)$$

$$\mathbf{y}(t) = \mathbf{A}_y \Phi\left(\mathbf{s}(t)\right) + \epsilon_y(t)$$

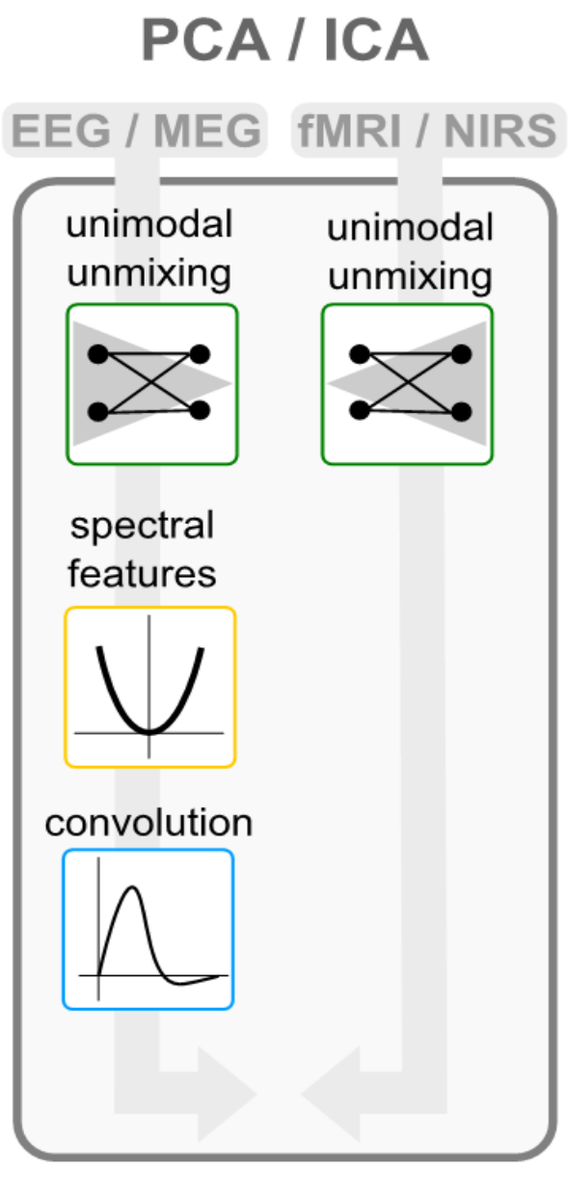
$$\phi\left(s_i(t)\right) := (h \star p_{s_i})(t) = \sum_{\tau>0} h(\tau)\, p_{s_i}(t-\tau)$$

"Convolution of *source* band power
with hemodynamic response function (HRF)"

Generative Model



Approaches to multimodal data analysis

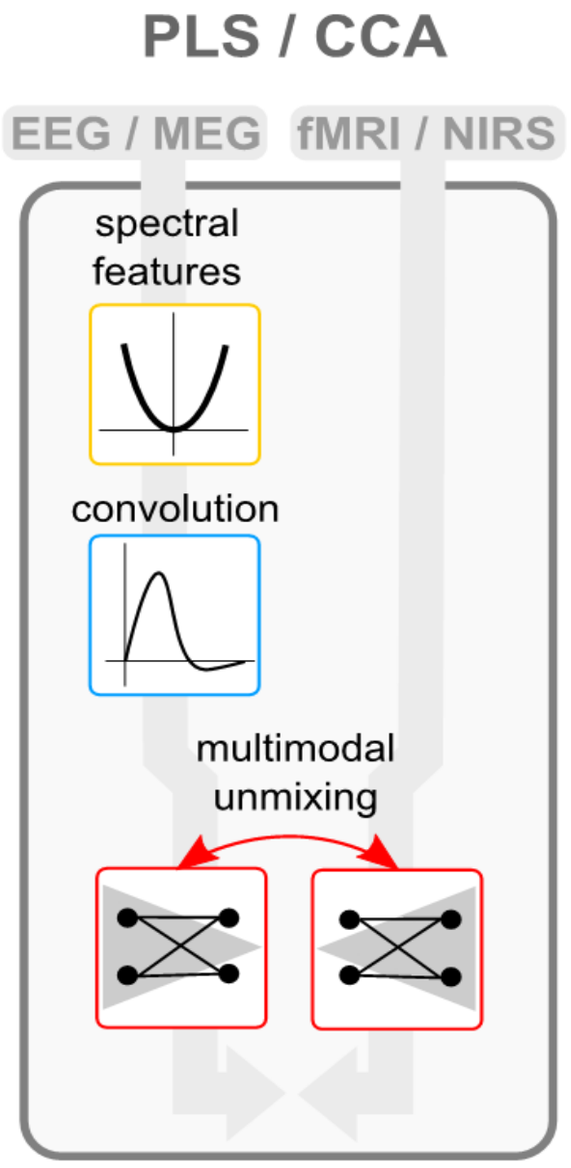


- order of processing steps in line with generative model, i.e. first transformation into source space, then computation of spectral power

- modality-specific unmixing → does not take information from other modality into account to guide the unmixing

- post-hoc matching of components, thus not truly multimodal

Approaches to multimodal data analysis

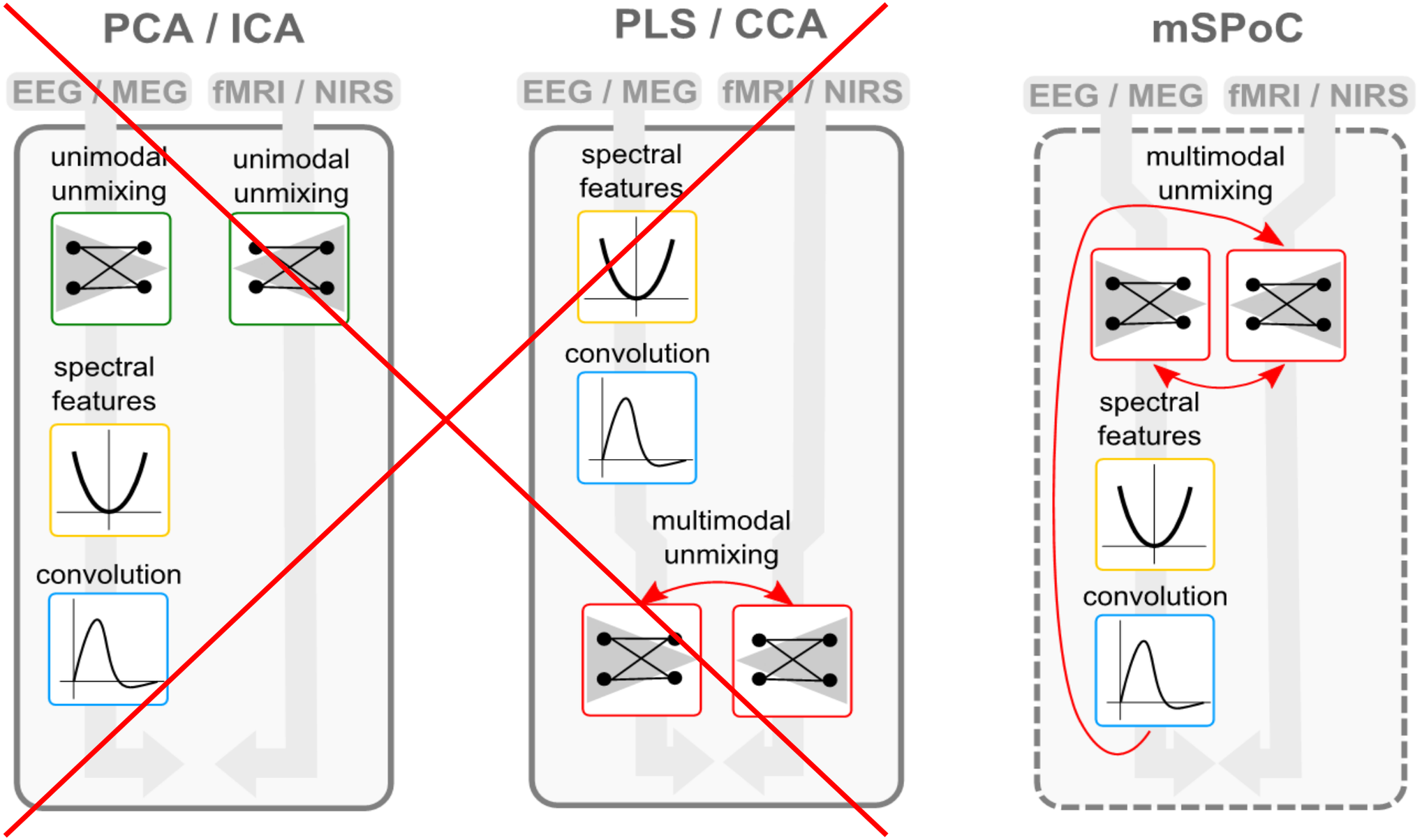


- multi-modal unmixing → optimizes the coupling between components

- order of processing steps **not** in line with generative model, i.e. nonlinearity is applied in sensor space instead of source space, !**WRONG!**

- resulting "EEG/MEG power patterns" cannot be subjected to standard source localization techniques, because these methods are designed to localize time-domain patterns, not spectral-domain patterns

Approaches to multimodal data analysis



Multimodal source power correlation analysis (mSPoC)

mSPoC objective function:

$$f_{\text{obj}}(\mathbf{w_x}, \mathbf{w_y}, \mathbf{w}_\tau) := \text{Cov}\left(\hat{h}(\hat{p}_{s_x}), \hat{s}_y\right)$$

norm constraints:

$$\begin{aligned} \|\mathbf{w_x}\|_{\mathbf{C_{xx}}} &:= \mathbf{w_x}^\top \mathbf{C_{xx}} \mathbf{w_x} = 1 \\ \|\mathbf{w_y}\|_{\mathbf{C_{yy}}} &:= \mathbf{w_y}^\top \mathbf{C_{yy}} \mathbf{w_y} = 1 \\ \|\mathbf{w}_\tau\|_{\mathbf{B}} &:= \mathbf{w}_\tau^\top \mathbf{B} \mathbf{w}_\tau = 1 \end{aligned}$$

$\mathbf{C_{xx}}$ $\mathbf{C_{yy}}$: modality specific covariance matrices

$\mathbf{B}$: auto-correlation matrix of $\hat{p}_{s_x}$

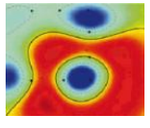
mSPoC model:

$$\begin{aligned} \mathbf{w_x}^\top \mathbf{x} &= \hat{s}_x \\ \mathbf{w_y}^\top \mathbf{y} &= \hat{s}_y \end{aligned}$$

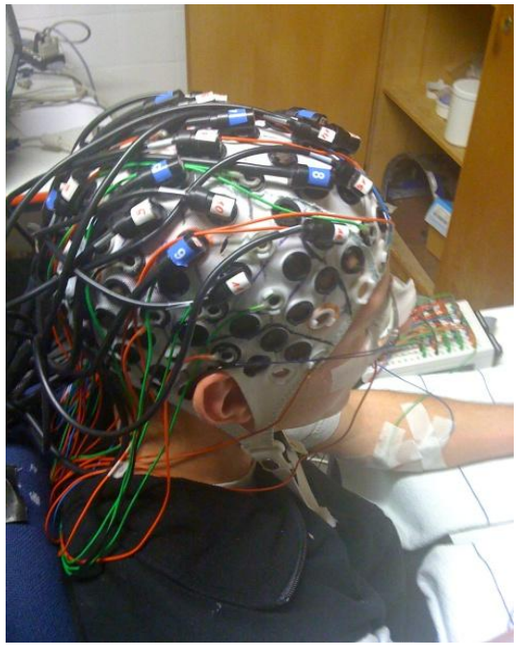
$$\begin{aligned} \hat{p}_{s_x}(e) &= \left\langle \left(\mathbf{w_x}^\top \mathbf{x}(t)\right)^2 \right\rangle_{t \in T_e} \\ &= \mathbf{w_x}^\top \mathbf{C}_{xx}(e) \mathbf{w_x}. \end{aligned}$$

$$\hat{h}(\hat{p}_{s_x})(e) = \sum_n^{N_\tau} \mathbf{w}_{\tau n} \hat{p}_{s_x}(e-n)$$

$$\begin{aligned} \mathbf{C_{xx}}^{-1} \frac{(\mathbf{C_{xxy}}_\tau \bar{\times}_{(3)} \mathbf{w_y} \bar{\times}_{(4)} \mathbf{w}_\tau)}{\|\mathbf{w_y}\|_{\mathbf{C_{yy}}} \|\mathbf{w}_\tau\|_{\mathbf{B}}} \mathbf{w_x} &= \lambda \mathbf{w_x} \\ \mathbf{C_{yy}}^{-1} \frac{(\mathbf{C_{xxy}}_\tau \bar{\times}_{(1)} \mathbf{w_x} \bar{\times}_{(2)} \mathbf{w_x} \bar{\times}_{(4)} \mathbf{w}_\tau)}{\|\mathbf{w_x}\|^2_{\mathbf{C_{xx}}} \|\mathbf{w}_\tau\|_{\mathbf{B}}} &= \lambda \mathbf{w_y} \\ \mathbf{B}^{-1} \frac{(\mathbf{C_{xxy}}_\tau \bar{\times}_{(1)} \mathbf{w_x} \bar{\times}_{(2)} \mathbf{w_x} \bar{\times}_{(3)} \mathbf{w_y})}{\|\mathbf{w_x}\|^2_{\mathbf{C_{xx}}} \|\mathbf{w_y}\|_{\mathbf{C_{yy}}}} &= \lambda \mathbf{w}_\tau \end{aligned}$$



Multimodal analysis of simultaneously recorded EEG and NIRS



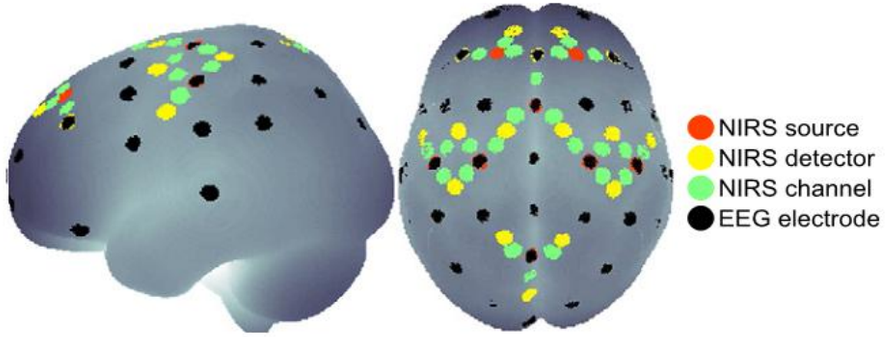
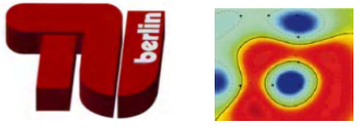


Fig. 1. Locations of EEG electrodes; sources, detectors and actual measurement channels of NIRS. Note that electrodes and optodes might share a location.

Data from Fazli et al. 2012 → 96 trials of (left/right) hand gripping.
Comparison of **mSPoC** to convolutive **CCA**.



mSPoC vs CCA

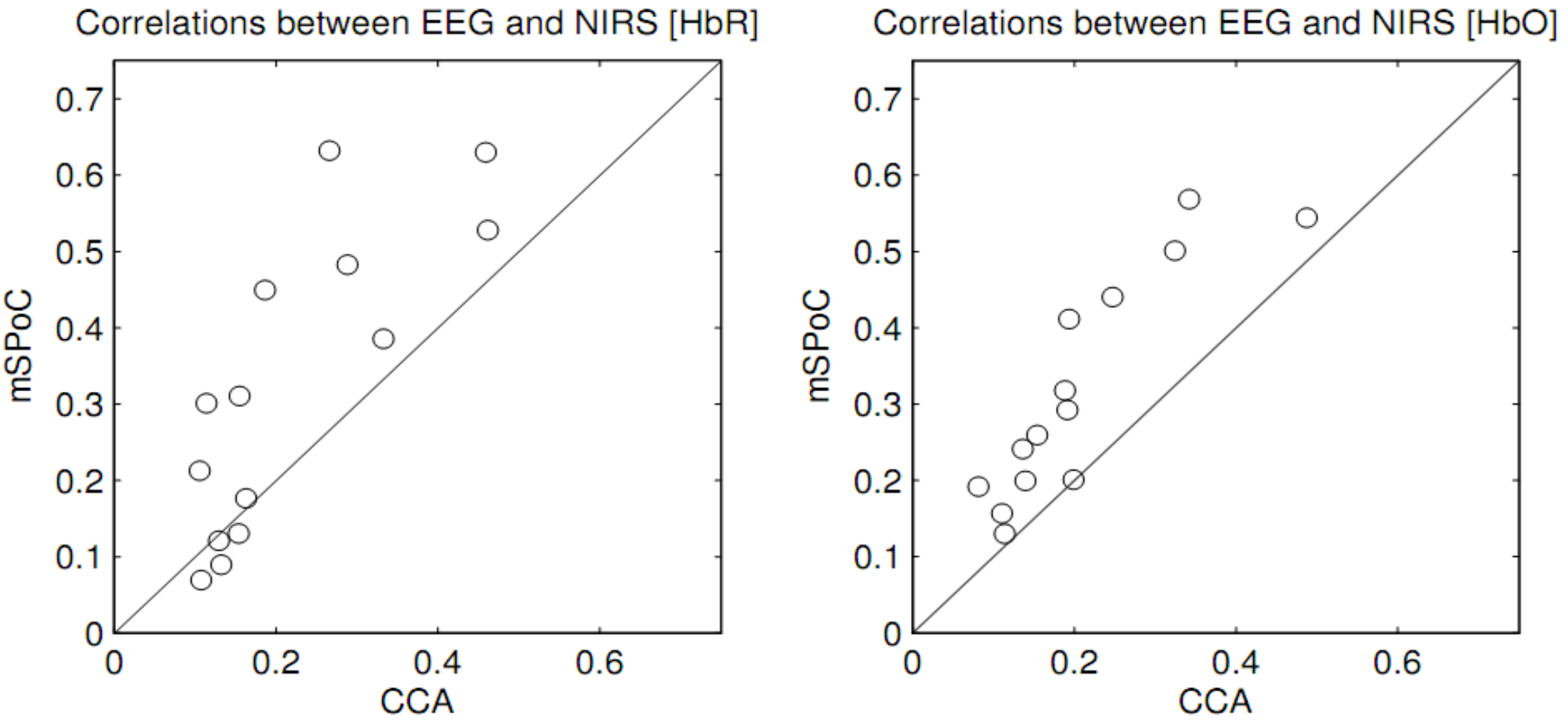


Fig. 5. Cross-validated correlations between EEG and NIRS (left HbR, right HbO) in the motor execution task for each subject. Results of mSPoC and CCA are compared. Each point corresponds to the correlations obtained for the first set of $\mathbf{w}_x$, $\mathbf{w}_y$, and $\mathbf{w}_\tau$ from a single subject by CCA (x-axis) and mSPoC (y-axis).

mSPoC vs CCA

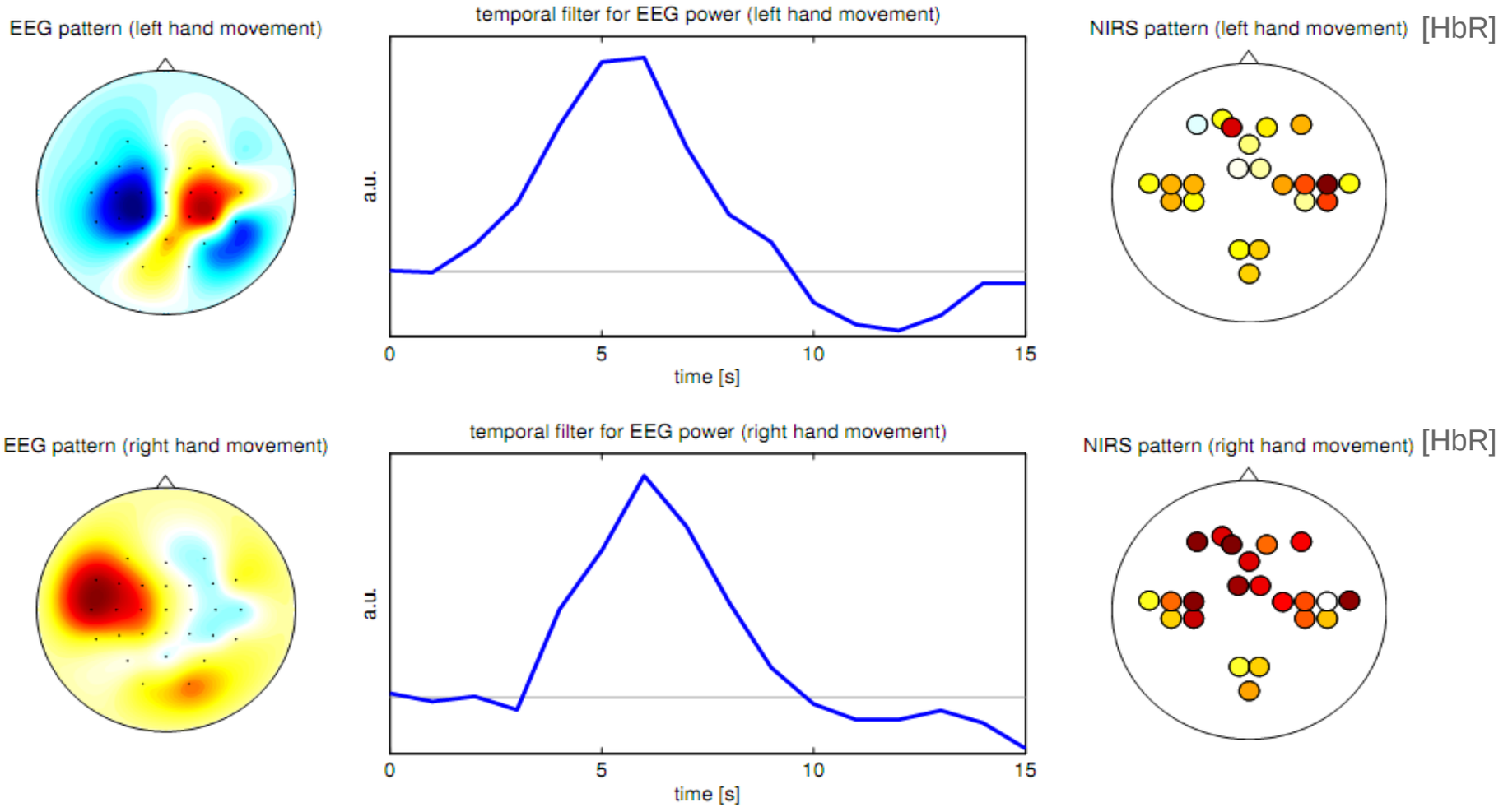
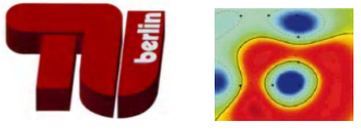


Fig. 7. Exemplary results for one subject (VPean) as derived by mSPoC. The scalp-plots on the left side show the EEG pattern that corresponds to the obtained filter $\mathbf{w}_x$. In the middle plot we show the temporal filter for the EEG power of the component shown left. The rightmost scalp-plots depict the spatial pattern that corresponds to the filter $\mathbf{w}_y$, i.e. the NIRS patterns. The top row shows the results for applying mSPoC to left hand movement trials, while in the bottom row results for right hand movement trials are shown.



Conclusion

- Information from multimodal measurements increases the understanding of physiology in neuroscience
- Multimodal Imaging is of interest for numerous research questions and clinical application
- The specific fusion of data depends on the research question and the used instruments
- Numerous algorithms have been developed to merge the data

FOR INFORMATION SEE:
www.bbci.de

Thanks to BBCI core team:

Gabriel Curio
Florian Losch
Volker Kunzmann
Frederike Holefeld
Jan Mehnert
Vadim Nikulin@Charite

Florin Popescu
Motoaki Kawanabe
Guido Nolte
Thorsten Dickhaus

Benjamin Blankertz
Michael Tangermann
Claudia Sannelli
Carmen Vidaurre
Siamac Fazli
Martijn Schreuter
Stefan Haufe
Andreas Ziehe
Laura Acqualagna
Frank Meinecke
Wojciech Samek
Sven Dähne
Paul von Bünau
Felix Biessmann@TUB

Matthias Krauledat
Guido Dornhege
Steven Lemm
Yakob Badower
Marton Danozci
Roman Krepki@industry

Collaboration with: U Tübingen, Bremen, Albany, KU, TU Graz, EPFL, Daimler, Siemens, MES, MPIs, U Tokyo, TIT, RIKEN, Bernstein Center for Computational Neuroscience Berlin, Columbia, CUNY
Funding by: EU, BMBF and DFG