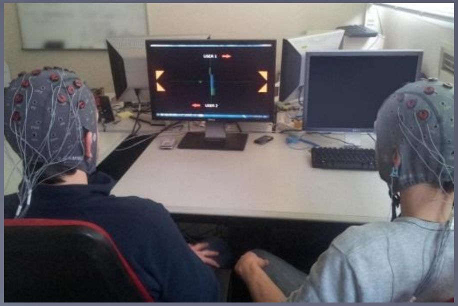
$$\frac{wC_1w^T}{wC_2w^T}$$

Oscillatory EEG-based BCI design:
signal processing and more

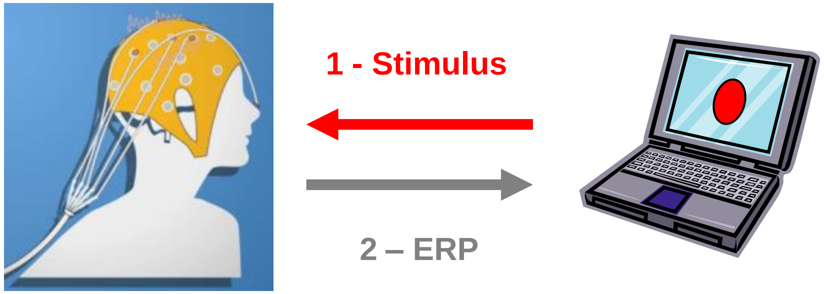
Fabien LOTTE
Inria Bordeaux Sud-Ouest, France
Potioc team



Oscillatory activity-based BCI

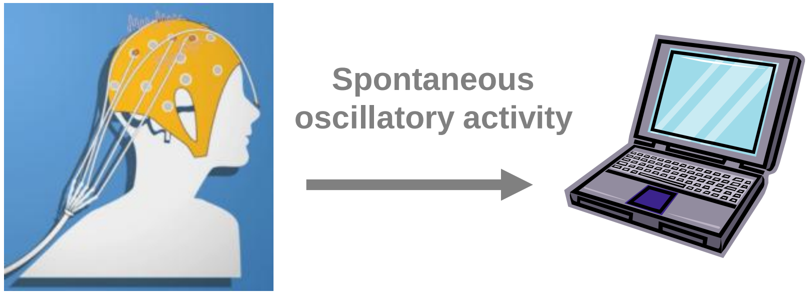
Event Related Potentials (ERP)

- Following a stimulus
- Modulated by attention



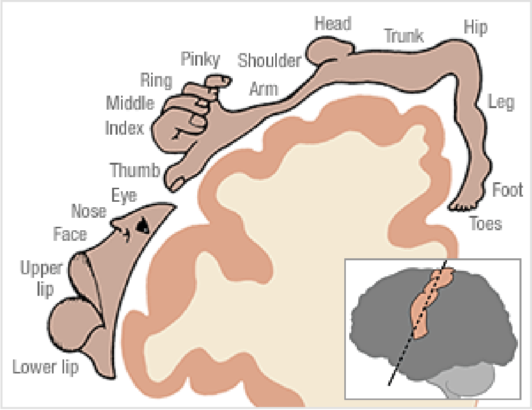
Oscillatory activity-based BCI

- Spontaneous activity (no stimulus)
- Uses mental imagery
  - Ex: limb movement imagination
- Change in EEG oscillations

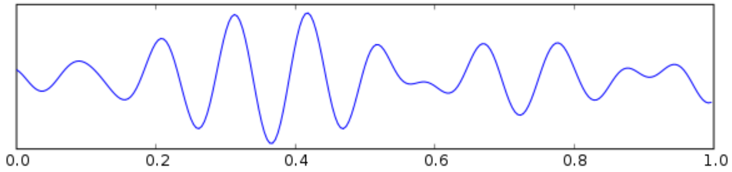


Example: Motor Imagery (MI)

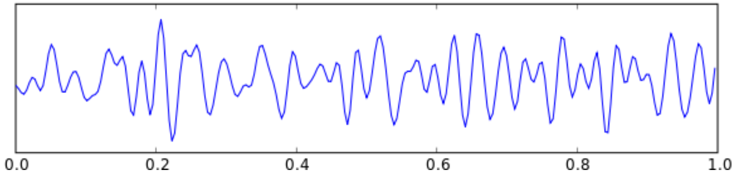
- Imagination of limb movements (e.g., left hand, right hand, feet)
- Contralateral ERD in μ (~8-12 Hz) or β (~12-30 Hz) during MI + β ERS (rebound) after MI



Penfield homonculus [Penfield54]



μ (~8-12 Hz) oscillations

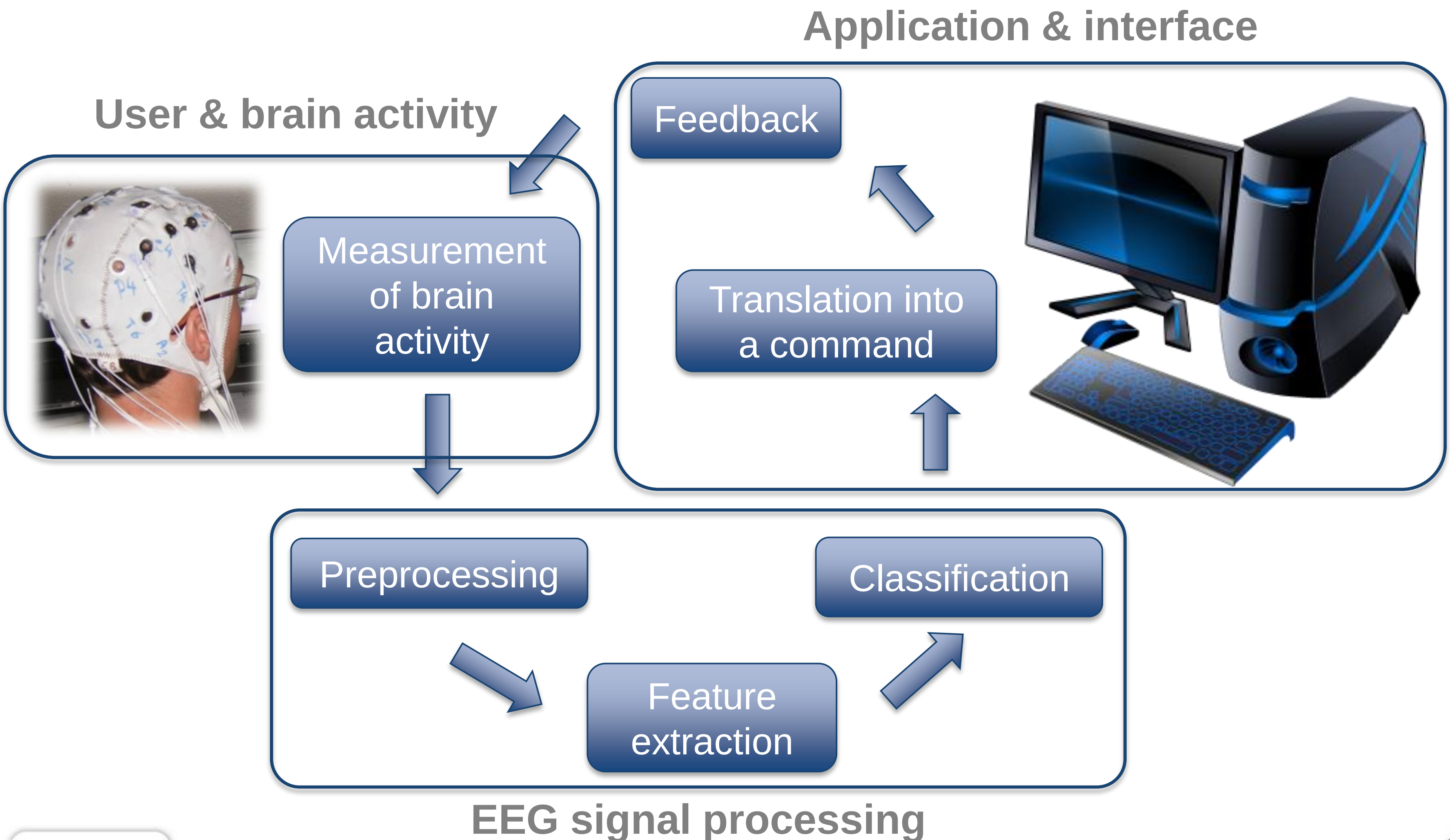


β (~12-30 Hz) oscillations

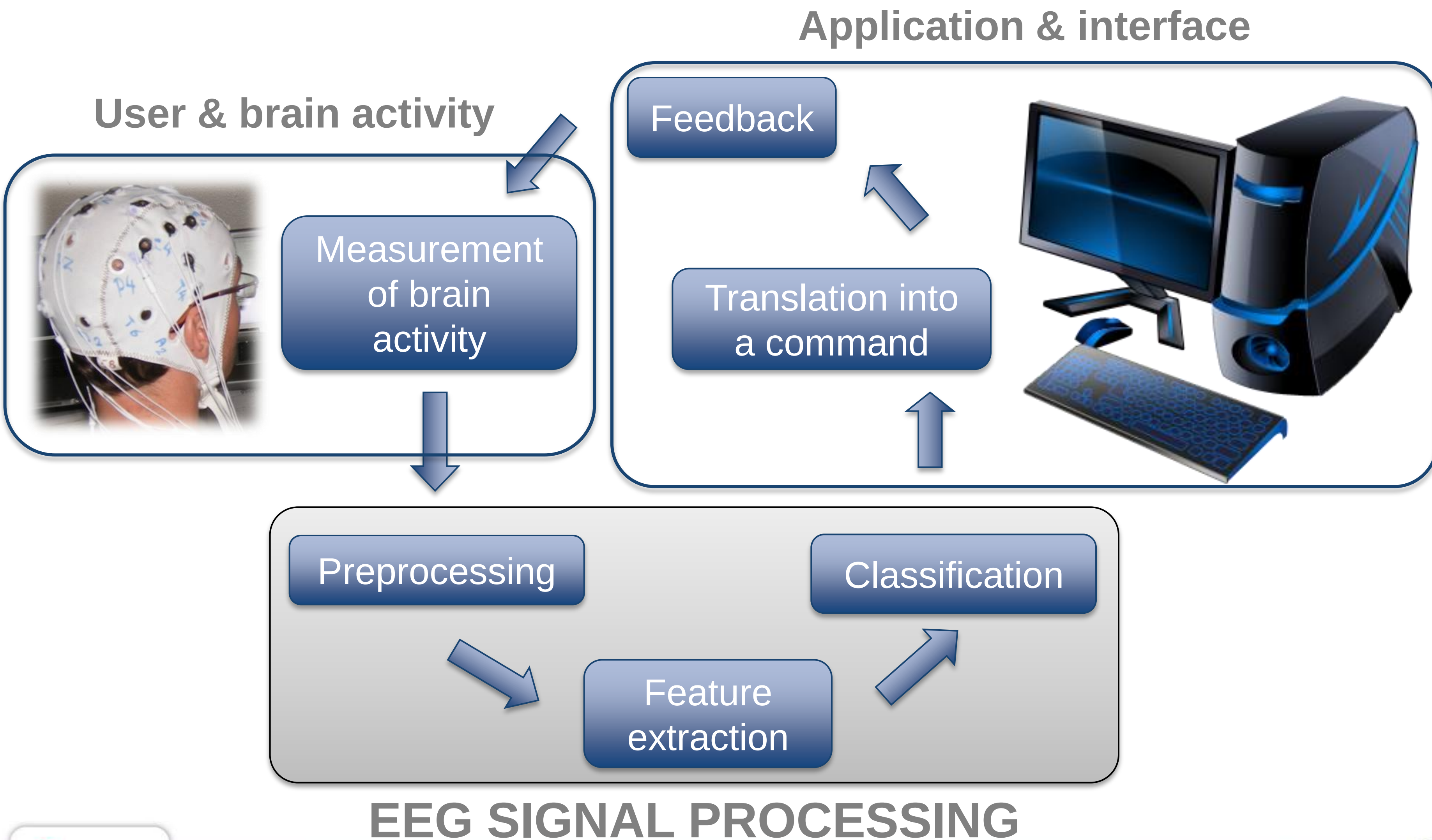
ERD/ERS = Event Related (De)Synchronization

Pfurtscheller & Neuper, "Motor imagery and direct brain-computer communication", Proceedings of the IEEE, 2001

ARCHITECTURE OF A BCI



THIS TALK MAIN FOCUS



Lecture Outline

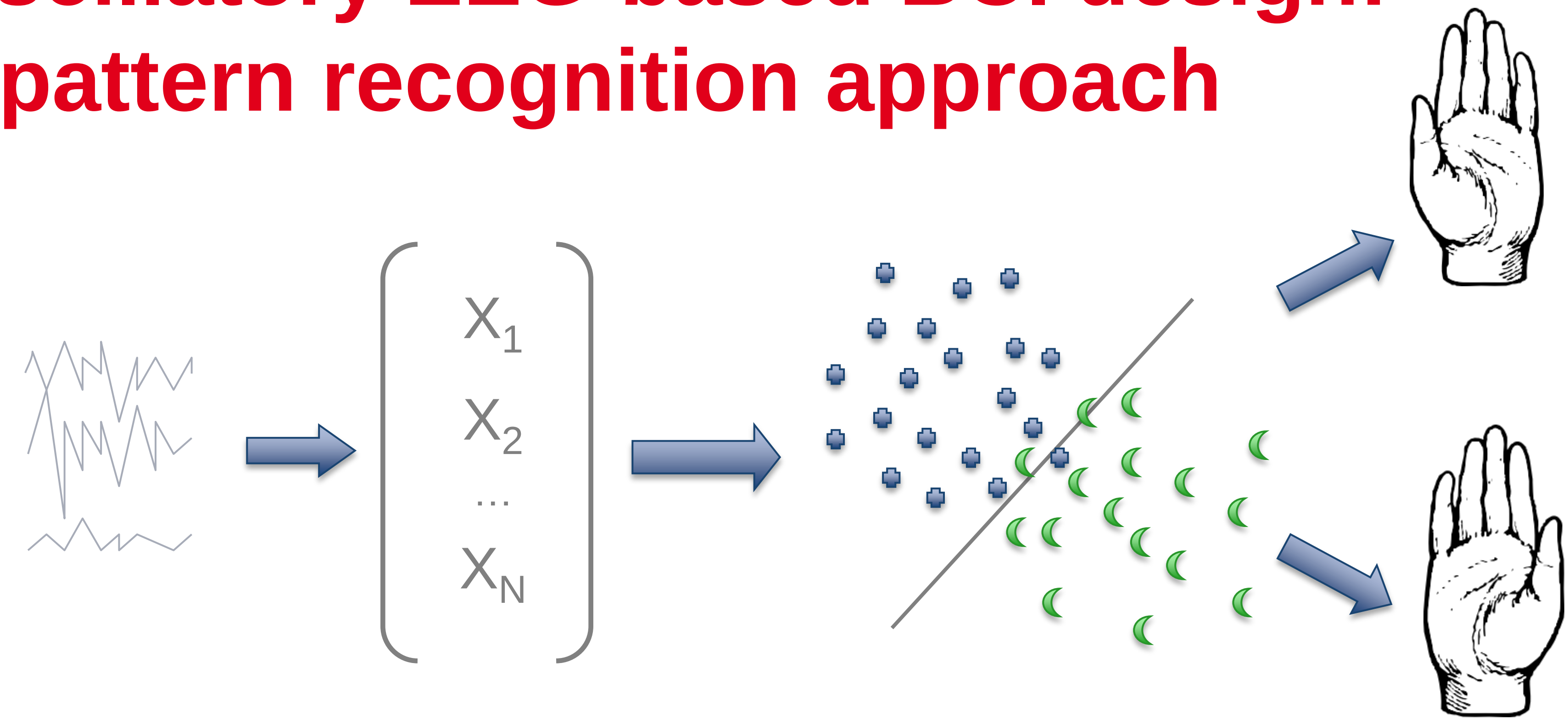
1. Basic oscillatory activity-based BCI design
2. Spatial Filters & Common Spatial Patterns
3. Common Spatial Patterns Extensions
4. Optimizing Spectral Filtering
5. Alternative Features for Oscillatory activity based BCI
6. Feedback & User training

1

Basic oscillatory activity-based BCI design

Inria

Oscillatory EEG-based BCI design: a pattern recognition approach



EEG signals

Ex: signal recorded during left or right hand motor imagery

Feature extraction

Ex: band power in the μ and β rhythms for electrodes located over the motor cortex

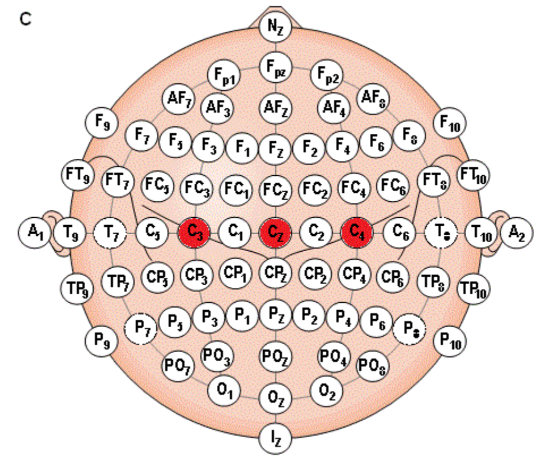
Classification

Ex: Linear Discriminant Analysis (LDA)

Estimated class

Ex: Left or Right (imagined hand movement)

Example: Features for Motor Imagery-based BCI



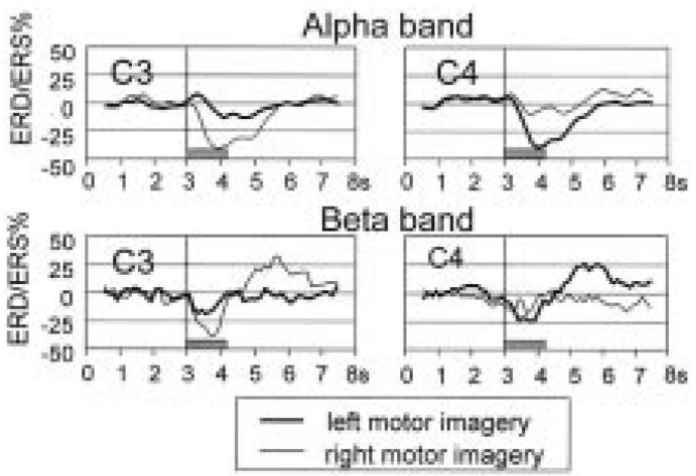
Channels C3, Cz, C4

Spatial information

- focusing on channels
  - C3: right hand MI
  - Cz: foot MI
  - C4: left hand MI

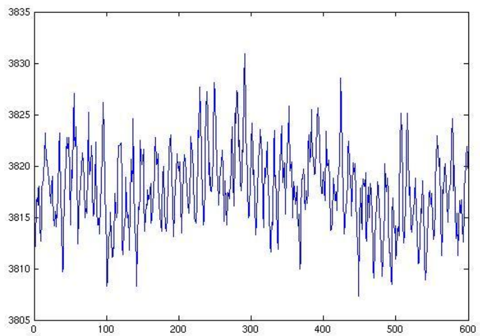
Spectral information

- Focusing on frequency bands
  - µ (mu: ~8-12 Hz)
  - β (beta: ~12-30 Hz)
- Features = power in such bands



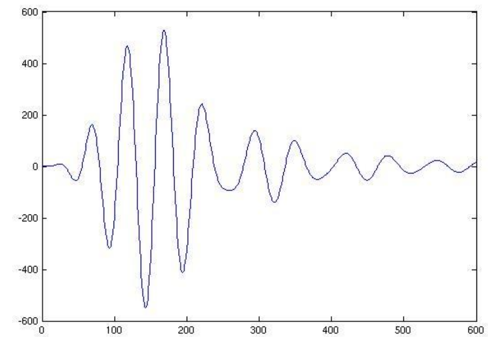
Band power features

Signal power in a given frequency band (here μ=8-12Hz)

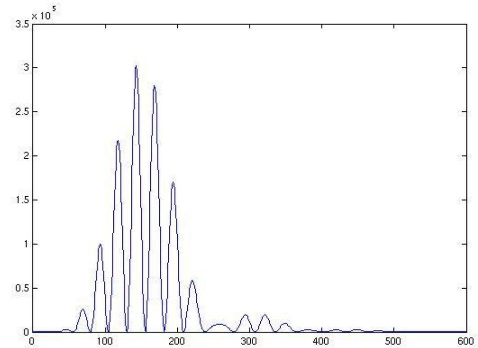


Raw EEG at C3
(left motor cortex)

Band-pass filtering in 8-12 Hz (μ)



Power estimation (squaring)



Temporal average

1 feature:
μ band power for channel C3
($P_{C3\text{-}\mu}$)

Band power features (2)

Other ways to compute them

- Periodogram (Fourier Decomposition)
- Power spectral density from AutoRegressive (AR) coefficients
- Wavelet scalogram (time-scale representation)
- Spectrogram (time-frequency decomposition then spectrums are averaged over time)
- Etc.

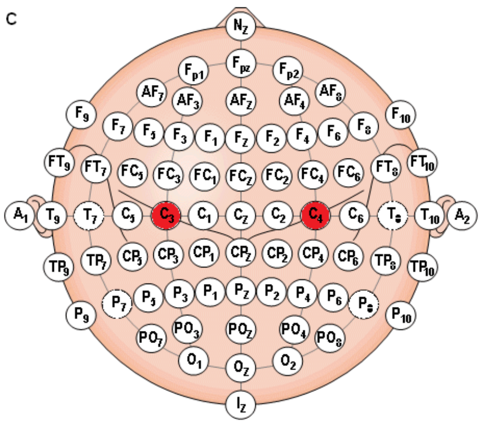
Herman, Prasad, McGinnity, Coyle, IEEE TNSRE, 2008
Brodu, Lotte, Lécuyer, IEEE SSCI, 2011

Basic design for left and right hand motor imagery-based BCI

- Computing band power features P
  - In frequency bands
    - μ (8-12 Hz) & β (12-30 Hz)
  - For channels
    - C3 & C4

- Gathering them into a feature vector
  $v = [P_{C3\text{-}\mu}, P_{C4\text{-}\mu}, P_{C3\text{-}\beta}, P_{C4\text{-}\beta}]$

- v is used as input to the classifier
  - E.g., LDA



C3 and C4 electrode location

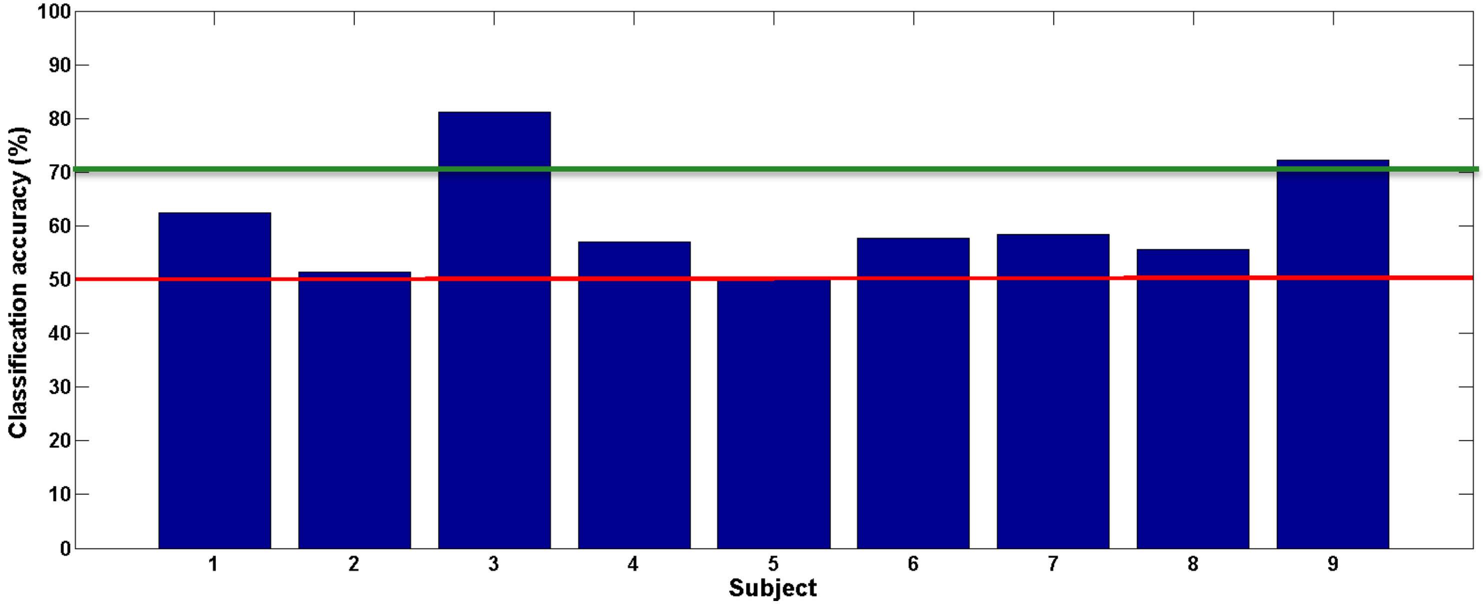
That’s all? It’s that simple?

Yes… but this basic design is far from being optimal

- Only 2 channels
  - $\Rightarrow$ Information might be missing
- Fixed channels (C3 & C4)
  - $\Rightarrow$ The optimal channels are subject-dependent
- Fixed frequency bands (8-12 Hz, 12-30 Hz)
  - $\Rightarrow$ The optimal frequency bands are subject-dependent

Basic design performance examples

- BCI competition IV (Tangermann et al, Frontiers, 2012), data set IIa
  - 9 subjects, Motor Imagery tasks
    (here: left and right hand movement imagination)
  - 72 trials per class for training and testing
  - Band Power in 8-30Hz, channels C3-C4, LDA classifier



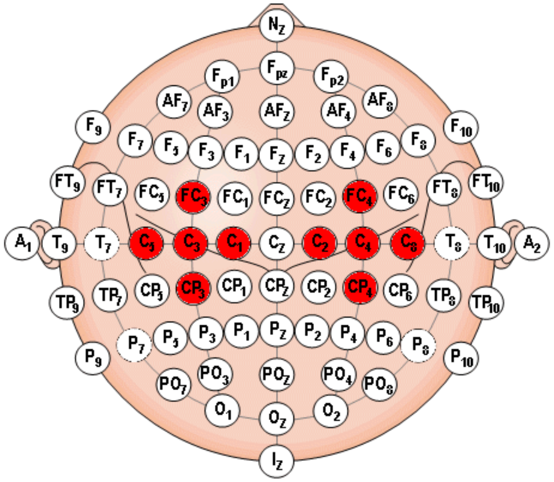
Average accuracy C3-C4: **60,7%**

Using more channels

- Extracting features from neighboring channels as well

- Problem
  - More channels
    $\Rightarrow$ More features
    $\Rightarrow$ Need for more training data (Curse-of-dimensionality)
  - Redundancies and correlations between channels

- Solution
  - Spatial filtering!



2

Spatial Filters & Common Spatial Patterns

(linear) Spatial Filtering

- Definition
  - Using a small number of new channels defined as a linear combination of the original ones

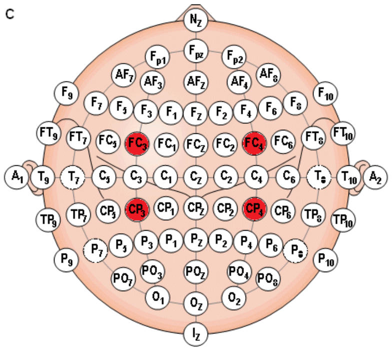
$$x' = \sum_i w_i x_i = wX$$

- Due to the smearing effect of the skull and brain, the underlying source signal is spread over several channels
  $\Rightarrow$ Spatial filtering helps in recovering this source signal

Some basic spatial filters

Bipolar filters

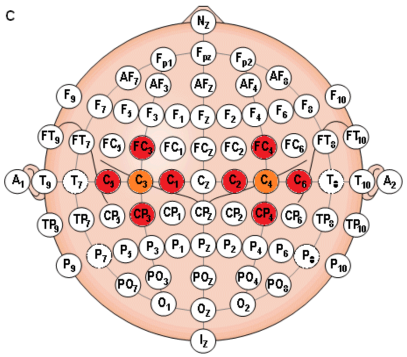
- C3’ = FC3 - CP3



Bipolar C3 & C4

Laplacian filters

- C3’= 4*C3 - FC3 - C5 - C1 - CP3



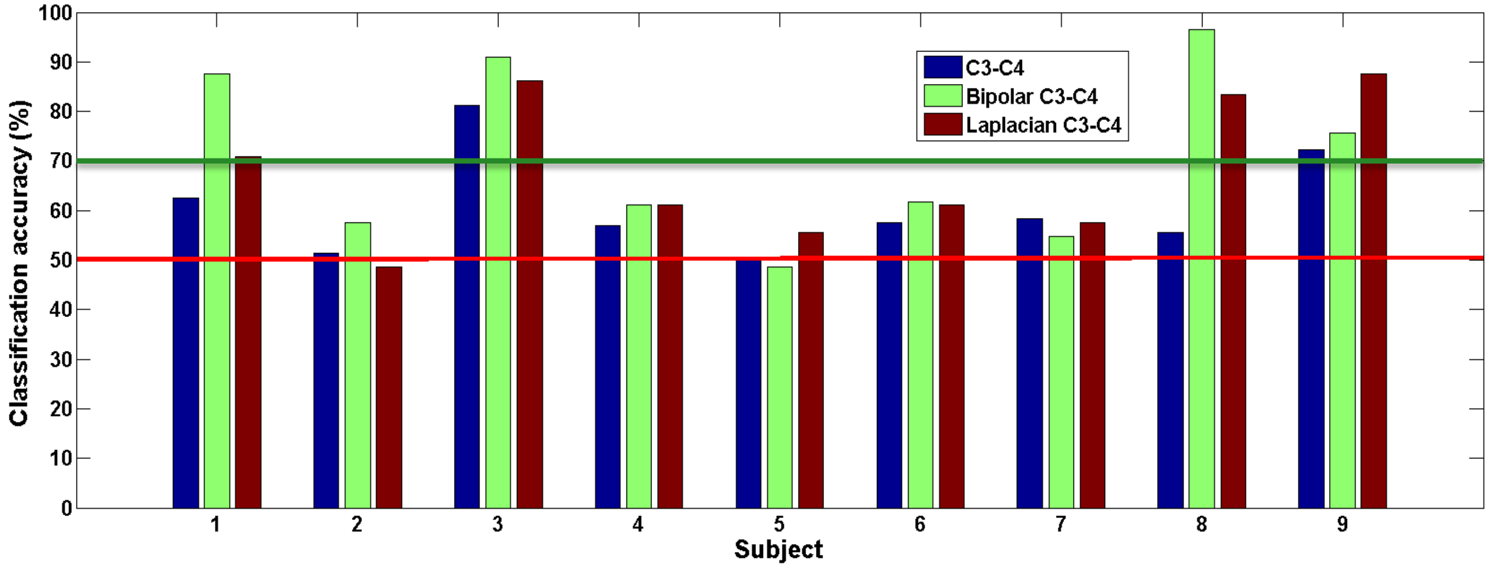
Laplacian C3 & C4

Emphasize localized activity and reduce diffuse spatial activity

[McFarland et al, EEG and Clinical Neurophysiology, 1997]

Basic spatial filters performance examples

- BCI competition IV, data set IIa
  - 9 subjects, Left vs right hand motor imagery
  - 8-30 Hz band power+LDA, different spatial filters



Average Accuracy:

C3-C4: **60,7%**

Bipolar C3-C4: **70,5%**

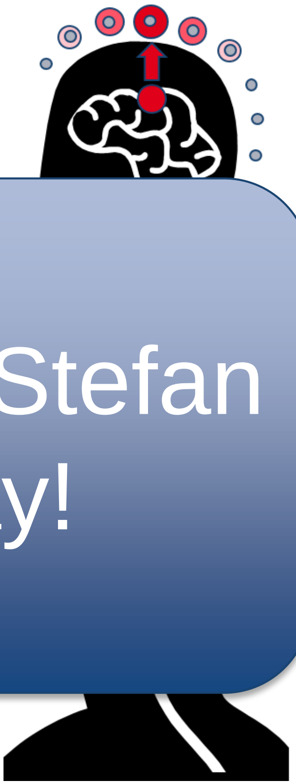
Laplacian C3-C4: **68%**

More advanced spatial filters

- **Inverse solutions**
  - Identify the spatial filter weights based on physical considerations
- **Supervised spatial filters**
  - Identify the spatial filter weights based on the EEG data and the class labels
  - Common Spatial Patterns (CSP)

Inverse Solutions

- Context
  - EEG are scalp measurements ***m*** resulting from the mixing ***A*** of several



More on Inverse solutions with Stefan Haufe's lecture on Thursday!

$\boldsymbol{s=Tm}$ ⟵ **Inverse/backward model**

  - Was shown efficient for BCI design
    Kamousi 2005, Congedo 2006, Lotte 2009, Besserve 2011

Supervised spatial filtering: Common Spatial Patterns (CSP) informally…

- Find spatial filters *w* such that the variance of the filtered signal is maximal for one class and minimal for the other class

➔ Variance of band-pass filtered signal

(we typically use 8-30 Hz by default)

= band-power of this frequency band

➔ CSP learns spatial filters that lead to optimally discriminant band-power features

Ramoser et al, "Optimal spatial filtering of single trial EEG during imagined hand movements", IEEE Trans. On Rehab. Eng., 2000

CSP formally

It consists in extremizing

$$J(w) = \frac{wX_1X_1^Tw^T}{\underbrace{\underbrace{wX_2}_{\text{Spatially filtered signal from class 2}}X_2^Tw^T}_{\text{Variance of the spatially filtered signal}}} = \frac{wC_1w^T}{wC_2w^T}$$

$\boldsymbol{C_i}$: EEG spatial covariance matrix for class i

$\boldsymbol{w}$: spatial filter to optimize

$\boldsymbol{X_i}$: multichannel EEG signals from class i

Solved by Generalized EVD (GEVD) of C_1 and C_2

- We typically use 3 CSP filter pairs hence obtained

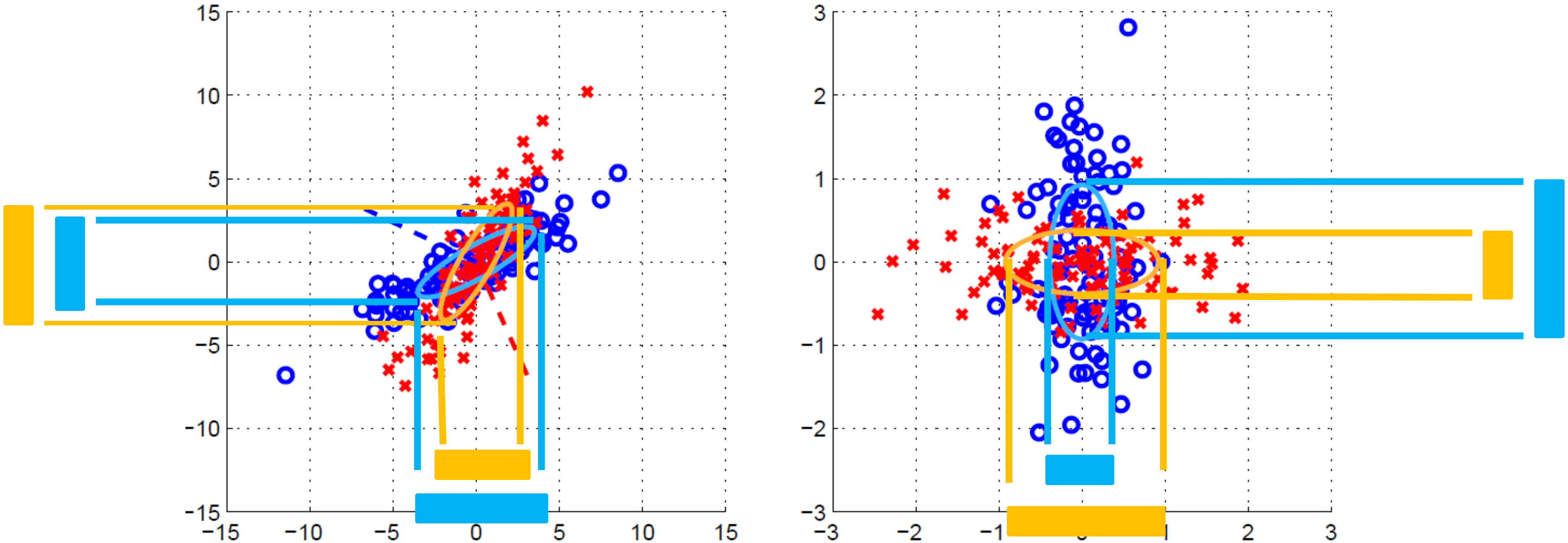
Once the filter *w* obtained, the feature *f* used is

$$f = \log(var(wX)) = \log(wCw^T)$$

CSP in action

Average feature value for class 1

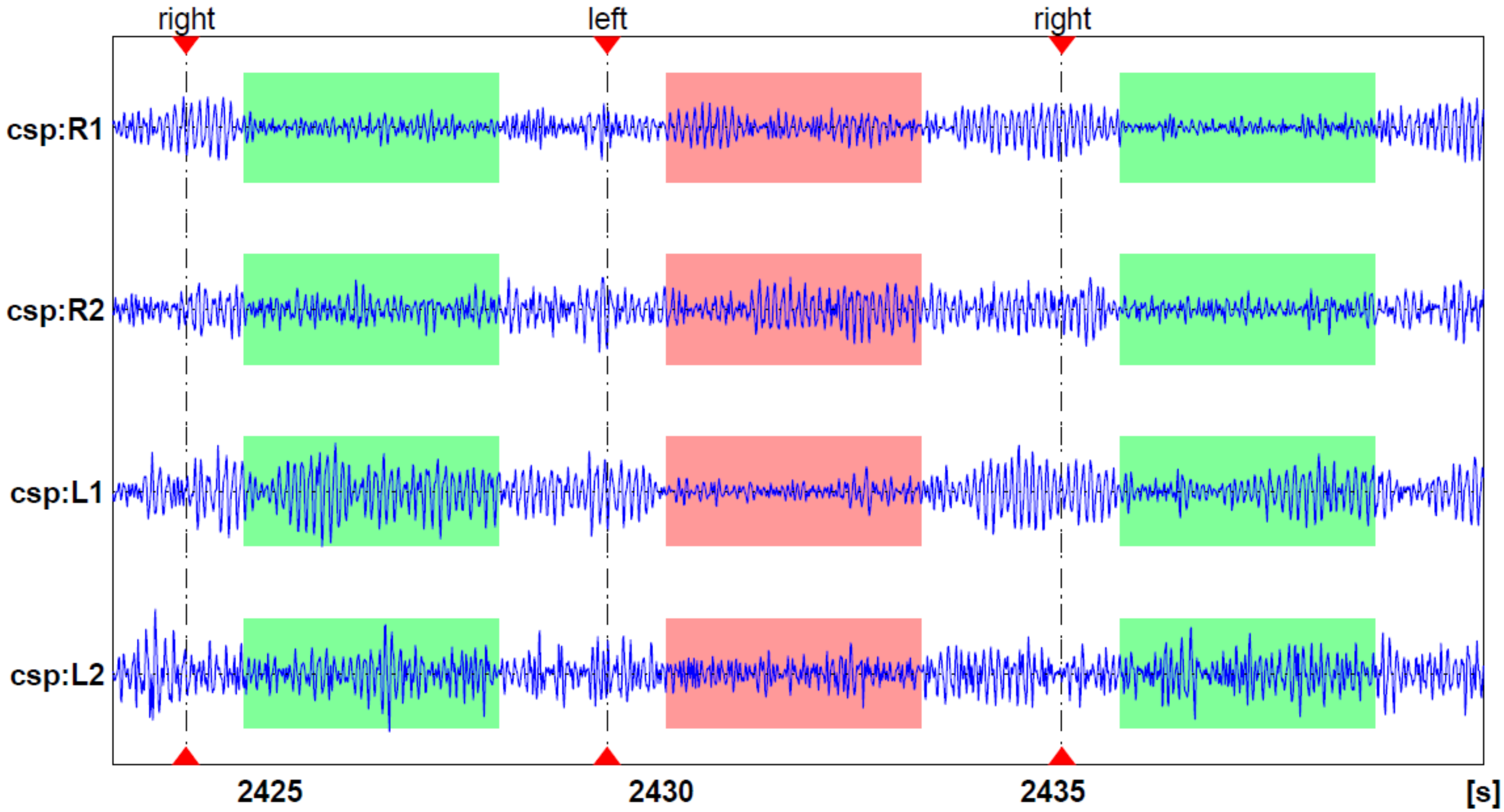
Average feature value for class 2



Toy data set representing EEG signals (not features!) from 2 channels

Blankertz et al, “Optimizing spatial filters for robust EEG single-trial analysis”, IEEE Signal Processing Magazine, 2008

CSP in action

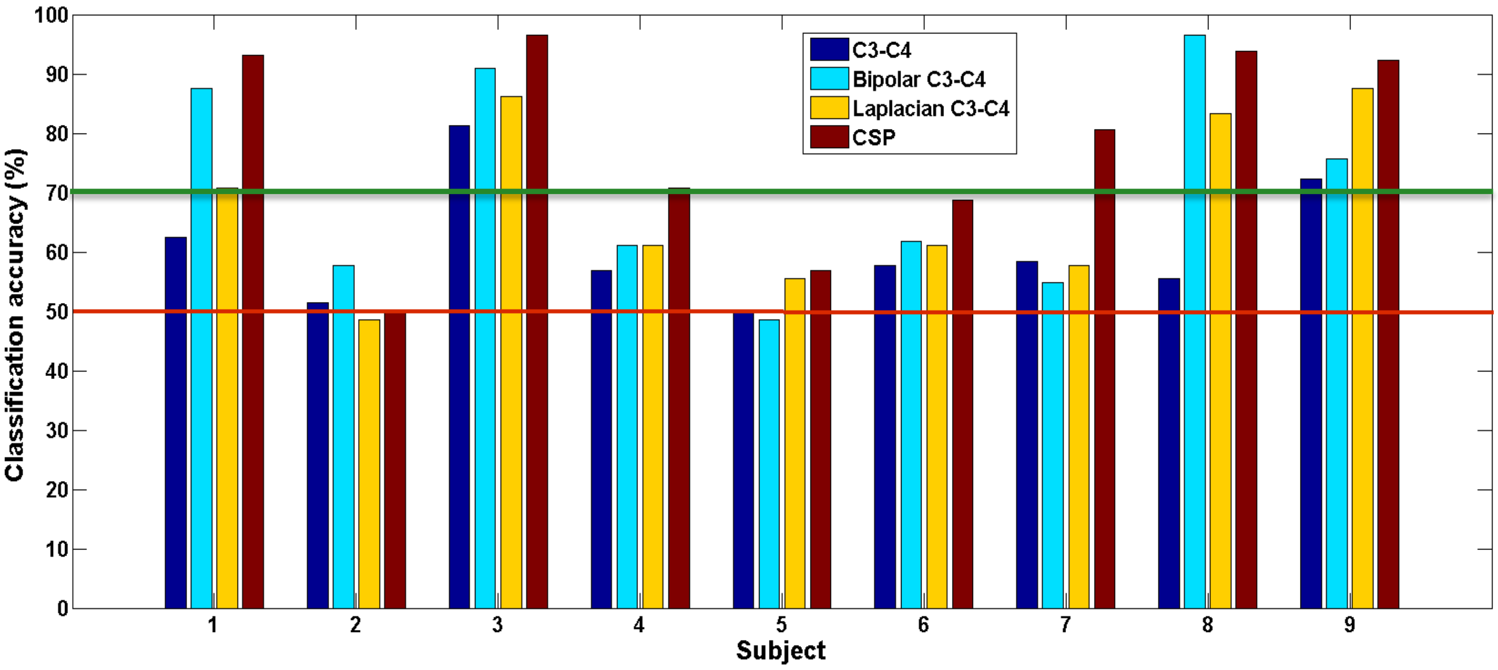


Examples of 4 CSP filtered signals, during left and right motor imagery

(from Blankertz et al, IEEE Sig. Proc. Mag., 2008)

CSP performance examples

- BCI competition IV, data set IIa
  - 9 subjects, left vs right hand motor imagery
  - 8-30 Hz band power+LDA, different spatial filters



Average Accuracy:

C3-C4: **60,7%**

Bipolar C3-C4: **70,5%**

Laplacian C3-C4: **68%**

CSP: **78,1%**

Pros and cons of CSP

Pros

- Lead to high classification performances
- Computationally efficient & simple to implement

➔ One of the most popular & efficient approach

Cons

- Non robust to noise and non-stationarities [Grosse-Wentrup08]
- Prone to overfitting [Reuderink09]
- Requires many training examples which leads to long calibration times [Blankertz08]
- Only for classification

3

Common Spatial Patterns Extensions

Towards a more robust BCI?

How to make BCI robust and stable?

- With limited training data
- With noisy and non-stationary training data

Idea: add a-priori information into the learning process

- Use a regularization framework to penalize unlikely and/or undesired solutions (e.g., unlikely spatial filters)
- Add a priori information to stabilize statistical estimates

→ **Using a Regularized CSP**

Regularized CSP (RCSP)

CSP	RCSP
Goal: extremizing $\frac{wC_1w^T}{wC_2w^T}$	Goal: maximizing $\frac{w\tilde{C}_1w^T}{w\tilde{C}_2w^T \boxed{+\alpha P(w)}}$ and $\frac{w\tilde{C}_2w^T}{w\tilde{C}_1w^T \boxed{+\alpha P(w)}}$
	Penalty term
	with $\tilde{C}_i = (1-\beta)C_i \boxed{+\beta G_i}$
	Stabilization term

Lotte & Guan, *IEEE Trans. BioMed. Eng.*, 2011

What prior knowledge to use?
Spatial knowledge to deal with noise

- Neighboring neurons are responsible for similar brain functions + EEG is smeared due to volume conduction

  => (close) neighboring electrodes should measure similar brain signals and thus have similar contributions

$$P(w) = \sum_{i,j} \underbrace{G(i,j)}_{\text{proximity of two electrodes}} \underbrace{(w_i - w_j)^2}_{\text{weight difference between electrodes}}$$

- For a given task, not all brain regions are involved

$$P(w) = w^T D w \;\; \text{with} \;\; D(i,j) = \begin{cases} \text{channel } i \text{ "uselessnes"} & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

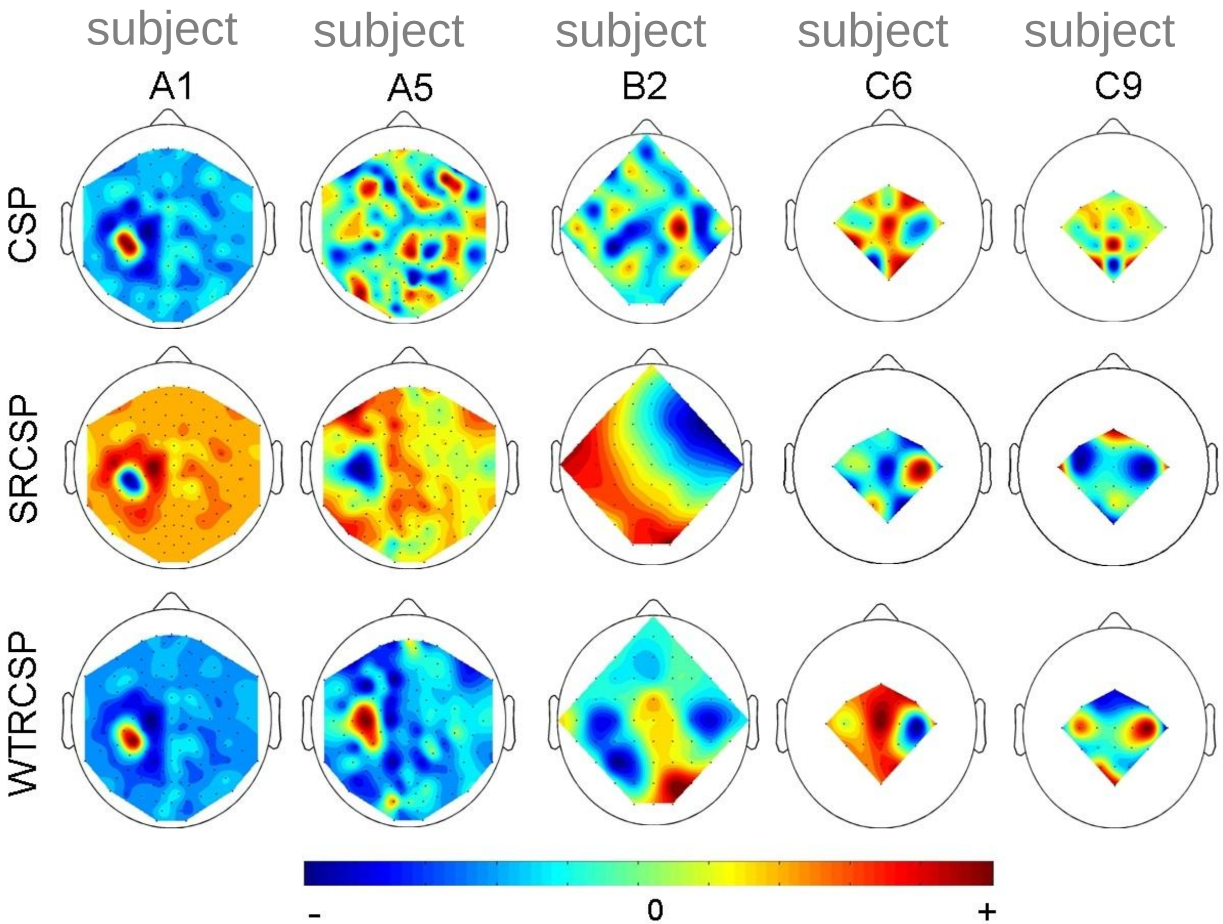
- Can be used for subject-to-subject transfer!
- Caution: CSP spatial filters are optimized for discriminability only

$\Rightarrow$ Their weights do not have to match neuroscience knowledge

Haufe et al, NeuroImage, 2014

Inria

Spatial filters obtained



Average classification accuracy (%)
(n = 12 subjects, 2 classes)

CSP	73.1 %
SRCSP	**78.7 %**
WTRCSP	77.6 %

Lotte & Guan, *IEEE Trans. on Biomedical Engineering*, 2011

Regularization terms to deal with non-stationarities

Invariant CSP [Blankertz, NIPS 09]

- To find filters invariant to a given non-stationary noise source (e.g., occipital alpha activity)

More on dealing with non-stationarities with Klaus-Robert Müller's lecture on Friday!

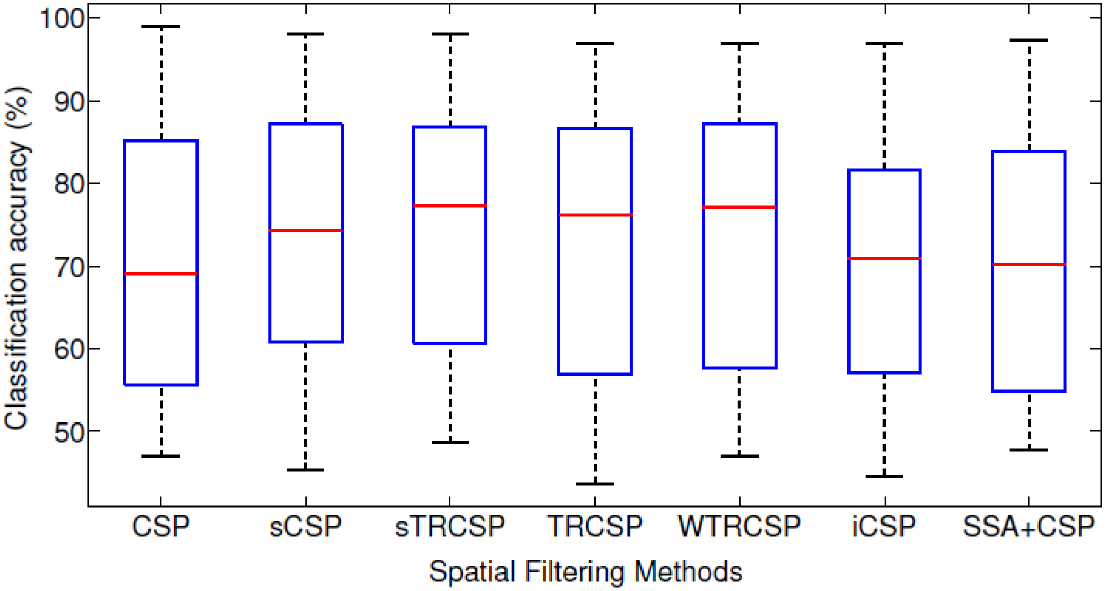
time/trials

- $P(w) = w^T(\Delta_1 + \Delta_2)w$ with $\Delta_i = \frac{1}{N}\sum_{i=1}^{N} P(C_i^k - C_i)$

Samek *et al*, *IEEE Reviews on Biomedical Engineering*, 2014

Combining multiple regularization terms

$$P(w) = \sum_i \lambda_i P_i(w)$$



Combining a Stationary + a Tikhonov regularization term (sTRCSP)

[Samek et al, JNE, 2012]

Regularization terms to reduce calibration time

- Automatic covariance matrix shrinkage (Ledoit & Wolf 2004)
    - Statistical tool dedicated to small sample size problems

$$\tilde{C}_i = (1-\beta)C_i + \beta I$$

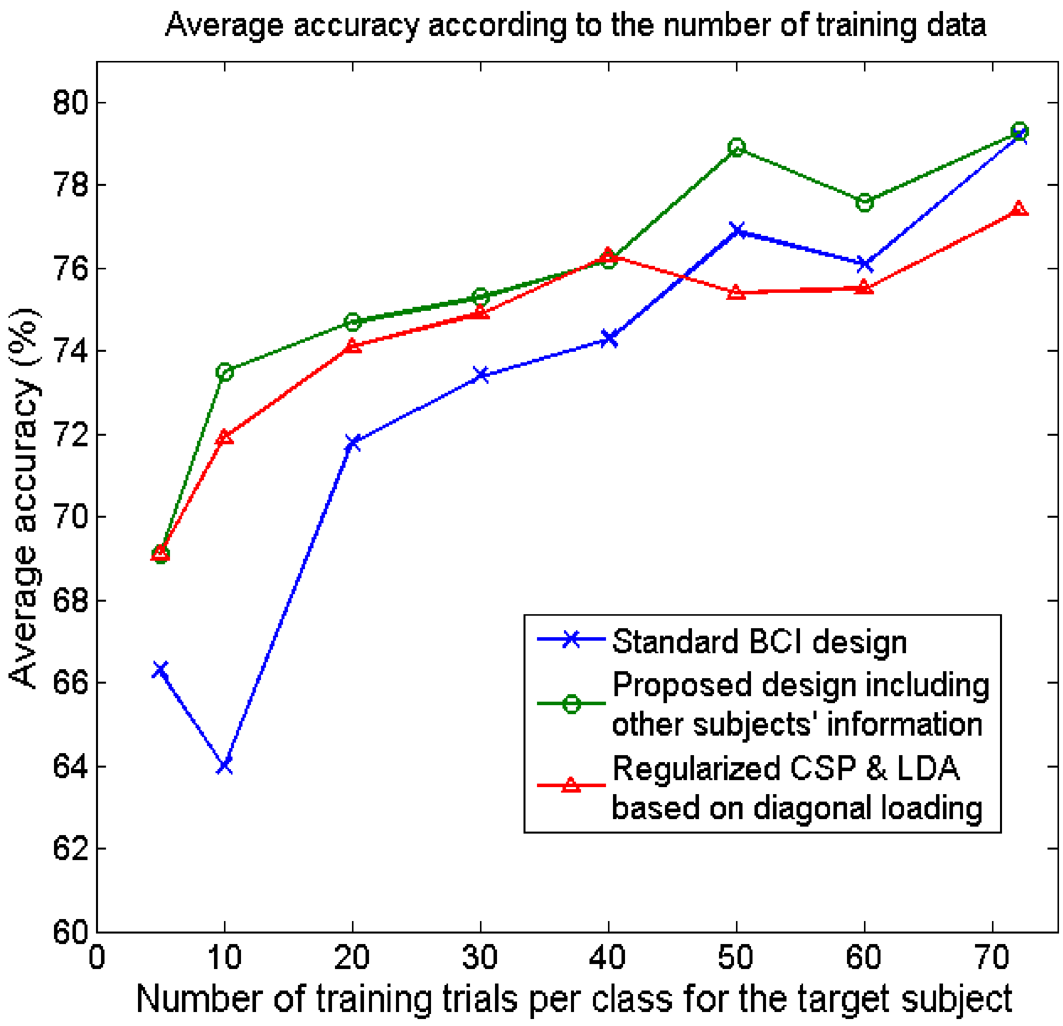
- Using data from other subjects (previously recorded) as the stabilisation term
    - Advantage: enable learning with less training data
      => calibration time reduction!

$$\tilde{C}_i = (1-\beta) \quad C_i \quad + \beta \quad G_i$$

$= (1-\beta)$

$+\beta$

Evaluation



Lotte & Guan, ICASSP 2010

Sparse CSP for channel selection

Using less EEG channels is

- More comfortable and less bulky
- Faster to set-up

Sparse CSP

- Enforce the use of few channels
- Also deals with noise (by removing noisy channels)

$$P(w) = \sum_i |w_i| = \|w\|_1$$

[Arvaneh et al, *IEEE TBME*, 2011][Farquhar et al, BCI workshop, 2006]

Using a-priori knowledge for CSP

Covariance estimation bias
(deal with calibration time)

Other subjects data
(deal with calibration time)

$$\frac{w\tilde{C}_1 w^T}{w\tilde{C}_2 w^T + \alpha P(w)}$$

Sparse solution
(deal with convenience, comfort & noise)

Spatial smoothness
(deal with noise)

Noise variance
(deal with known noise)

Channel usefulness
(deal with noise)

Inter-trial variance
(deal with non-stationarities)

Spatial filters for relating EEG band power to a target variable

- CSP are optimal spatial filters for Band Power-based **classification**

- What if you want to do **regression**?
  - E.g., to find EEG oscillatory activity features that correlate to a given continuous stimulus value (ex: auditory stimulus intensity) or a cognitive state level (ex: workload or attention)

- Spatial filter would still be useful, but CSP is not designed for that

- Fortunately, there is **SPoC** – it does exactly this
  - Source Power Comodulation
  - Can be seen as an extension of CSP for continuous variables

Dähne et al, NeuroImage, 2014

SPoC: Source Power Comodulation

- Find a spatial filter w such that the resulting source (i.e., spatially filtered signals) power is maximally correlated with the target variable z:

$$\underset{w}{\operatorname{argmax}} \frac{Cov(\overbrace{w^T(C(e)-\bar{C})w}^{\text{Spatially filtered signal variance} = \text{source power}}, z(e))^2}{Var(w^T(C(e)-\bar{C})w)Var(\underbrace{z(e)}_{\text{Target variable for epoch e}})}$$

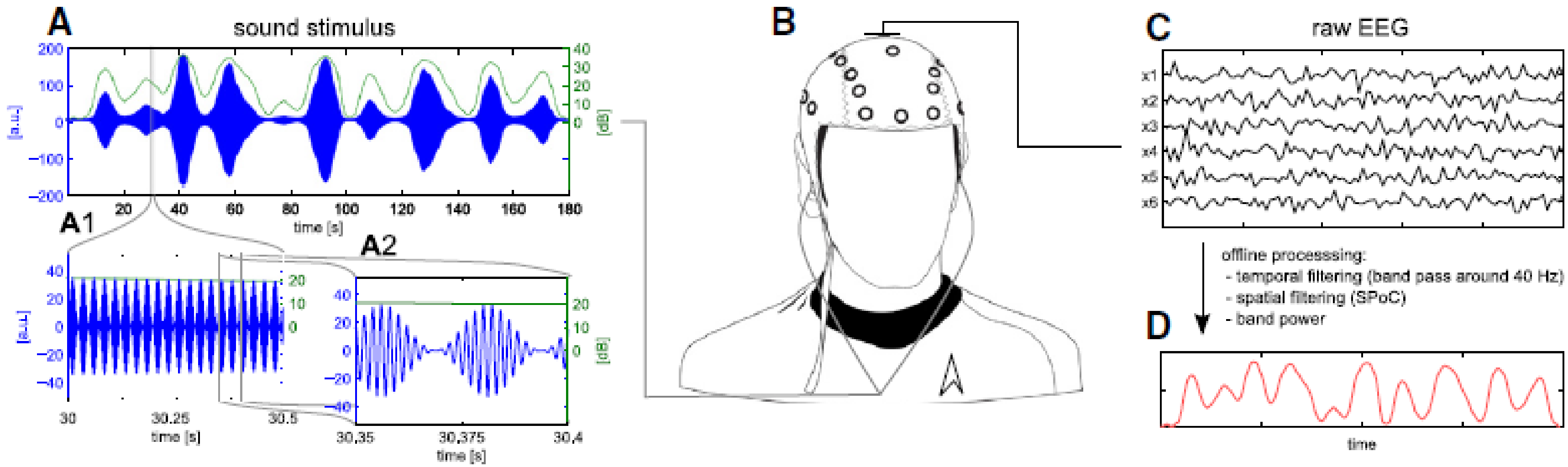
- Can be approximated (and more easily computed) using the covariance between the source power and the target

$$\underset{w}{\operatorname{argmax}} \frac{Cov(w^T(C(e)-\bar{C})w, z(e))}{w^T\bar{C}w} = \frac{w^T\langle C(e)z(e)\rangle w}{w^T\bar{C}w}$$

- Solved by GEVD (again)

SPoC example

- Estimating the intensity modulation of a sound stimulus from EEG signals source power using SPoC



- SPoC proved more efficient (higher correlation) than ICA (Independent Component Analysis) or channel-wise regression

Dähne et al, NeuroImage, 2014

In short, spatial filters are really useful…
But do we really need them?

For ERD/ERS linear classification:

$$y = wf + b \quad = w\, diag(W C_x W^T) + b \approx \hat{w}\, vec(C_x) + b$$

Classifier weights **Features** **Spatial filter weights**

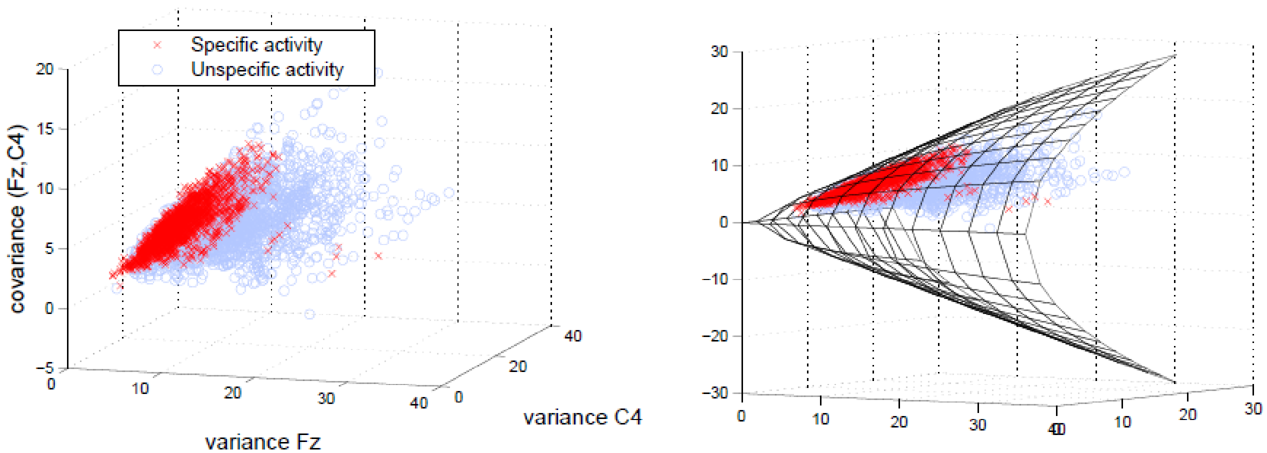
All the weights can be learned in the (vectorized) covariance matrix space

- Enough data and a good regularized classifier is necessary

[Tomioka & Muller, NeuroImage, 2009] [Farquhar, IEEE TNN, 2009]

Classifying Covariance matrices directly

- Using distance measures between covariance matrices
  - Ex: Riemannian distance, Stein Kernel, etc.



From [Barachant et al, BCI conference, 2010]

- Can be kernelized

$$y = \sum_i \alpha_i K(C, C_i) + b$$

[Barachant et al, IEEE TBME, 2012][Barachant et al, ESANN, 2012][Yger MLSP 2014]

Summary of CSP spatial filters extensions

- Using appropriate regularization, CSP can be mode robust to
  - noise
  - non-stationarity
  - overfitting with limited training data
- CSP can be extended to the continuous variable case: SPoC
- CSP spatial filters can be learnt implicitly

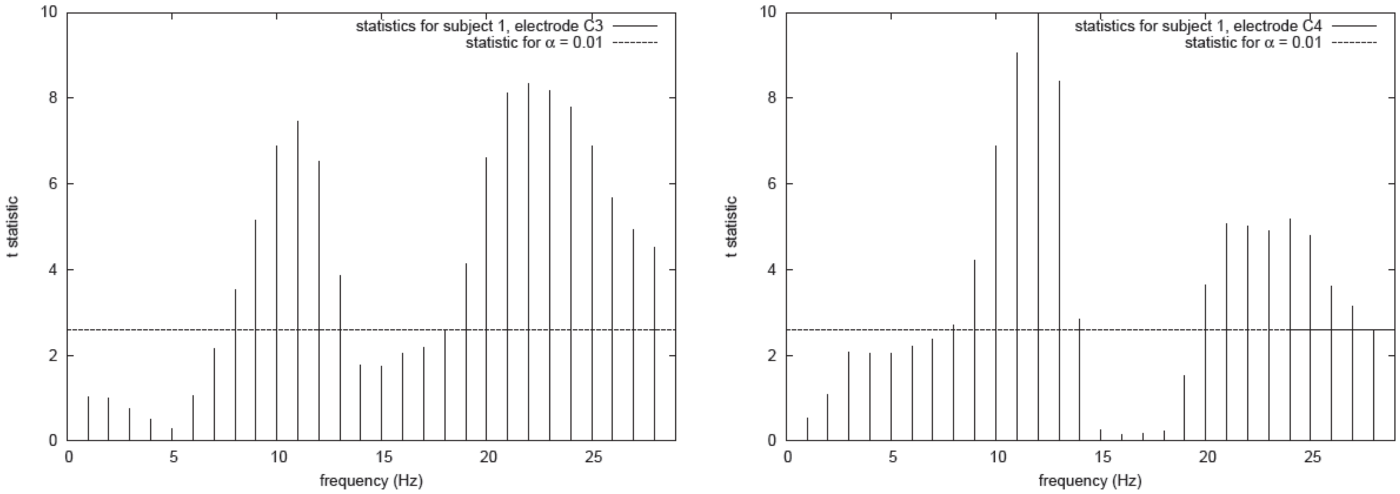
4

Optimizing Spectral Filtering

Using Subject-specific frequency bands

How to find the optimal bands?

- Manually (trial-and-errors)…
- Looking at the average spectrum in each class
- Computing statistics (R^2, Fisher score, etc.) on the spectrum



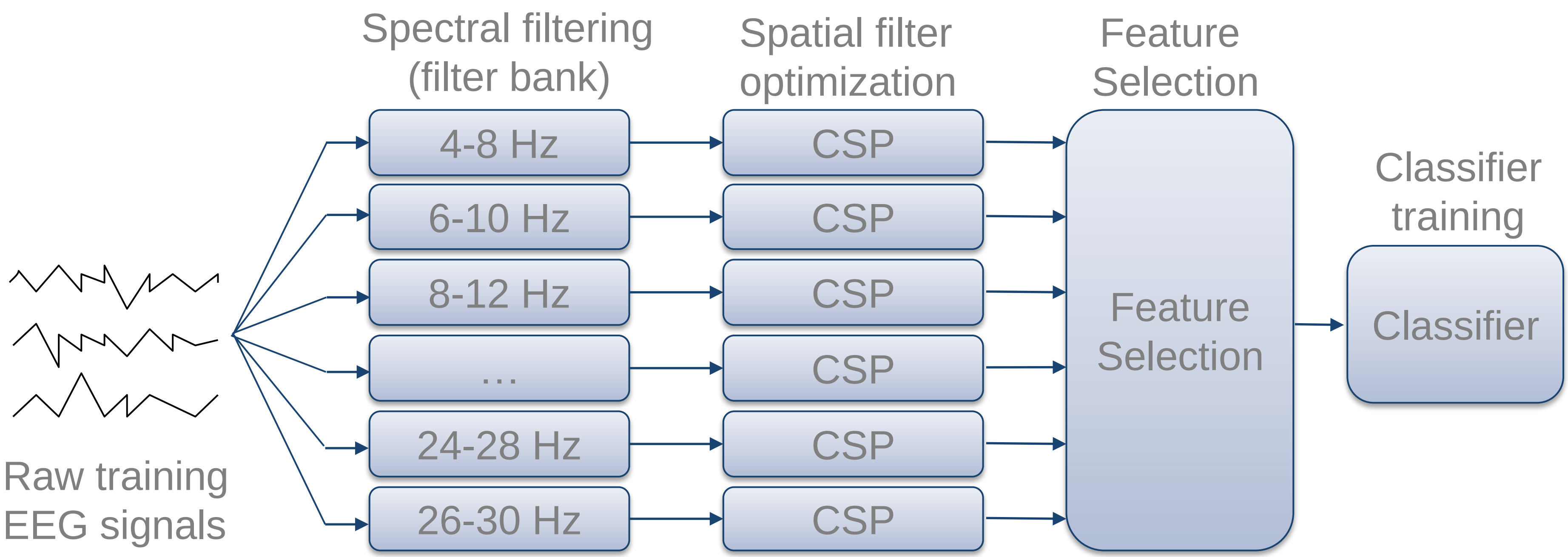
From Zhong et al, Patt. Rec. Let., 2008

Optimizing Spatio-Spectral Filters for BCI

- Common Spatio-Spectral Patterns (CSSP) [Lemm et al, TBME, 2005]
- Common Sparse Spectral Spatial Pattern (CSSSP) [Dornhege et al, TBME 2006]
- Spectrally Weighted CSP [Tomikia et al, 2006]
- Discriminative CSP [Thomas et al, TBME, 2009]
- Filter Bank Common Spatial Pattern (FBCSP) [Ang et al, PR, 2011]
- Discriminative FBCSP [Higashi et al, TBME, 2013]

All more efficient than basic CSP

The Filter Bank CSP (FBCSP)



Ang et al, « Filter Bank Common Spatial Patterns in Brain-Computer Interfaces », IJCNN, 2008

FBCSP Results

Method	Classification accuracy (%)	
	Data Set I (5 subjects)	Data Set II (11 subjects)
CSP	86.6	73.3
FBCSP	90.3	81.1

Efficiency of FBCSP (from Ang et al, IJCNN, 2008)

Winning algorithm of BCI competition 2008 on all EEG data sets
Ang et al, Pattern Recognition, 2011

Optimizing temporal filters for BCI (1)

- Finite Impulse Response (FIR) temporal filters:

$$\hat{x}_t = \sum_{k=0}^{N} h_k x_{t-k}$$

Filter order

Temporally filtered signal at time t

Original signal

Temporal filter weights

- In matrix form:

Optimizing temporal filters for BCI (2)

- Optimizing temporal FIR filters similarly as we optimize CSP spatial filters for maximally discriminant band power features

$$J_{FIR}(h) = \frac{\overbrace{(T_{x_1}h)^T(T_{x_1}h)}^{\text{Variance of the temporally filtered EEG signals from class 1}}}{\underbrace{(T_{x_2}h)^T(T_{x_2}h)}_{\text{... from class 2}}} = \frac{h^T T_{x_1}{}^T T_{x_1} h}{h^T T_{x_2}{}^T T_{x_2} h}$$

CSP Reminder:

$$J_{CSP}(w) = \frac{w^T X_1^T X_1 w}{w^T X_2^T X_2 w}$$

D. Devlaminck, PhD Thesis, Ghent University, 2011

Discriminative Filter Bank CSP (DFBCSP)

- Combining CSP and FIR objective function

$$J(w,h) = \frac{w^T h^T {T_{x_1}}^T T_{x_1} h w}{w^T h^T {T_{x_2}}^T T_{x_2} h w}$$

- Can be solved by alternating FIR filter and CSP spatial filter optimizations, using appropriate GEVD for each
- Can extract a bank of temporal/spatial filter pairs

Higashi & Tanaka, IEEE Trans. Biomed. Eng., 2013

DFBCSP results

- Data Set Iva, BCI Competition III, [Blankertz 2006]
  - 5 subjects, right hand versus right foot motor imagery

Method	Results (%)
CSP	88,9
FBCSP	90,90
DFCSP	**92,80**

Higashi & Tanaka, IEEE Trans. Biomed. Eng., 2013

5

Alternative Features for Oscillatory activity based BCI

Alternative features

Temporal representations [Vidaurre09]

- Ex: Hjorth or Time Domain Parameters

Connectivity measures [Grosse-Wentrup09]

- Ex: Coherence, phase locking value, causality

Complexity measures [Balli11, Brodu12]

- Ex: Entropy, predictive complexity, Fractal dimension, multifractal

…

Using multiple features

Provides complementary information

- Likely to increase performances

[Dornhege et al, IEEE TBME, 2004]

- Features that are weak alone may be efficient together

[Brodu, Lotte, Lécuyer, Neurocomputing, 2012]

May provide a kind of redundancy

- Provide robustness to different kinds of noise

[Fatourechi, JNE, 2008]

Spatial filters for alternative features?

- Band-power (BP) are not the only valuable features for oscillatory activity-based BCI
  - BUT -

  They are almost the only ones with dedicated spatial filters, namely CSP
  - CSP are optimized for BP features only

  $\Rightarrow$ Using other features with CSP filters is suboptimal

- Could spatial filters for alternative features be useful?

  $\Rightarrow$ YES!

Example 1: Time Domain Parameters (TDP)

- A measure of temporal variations

$$TDP^{(k)} = \left\|\Delta x^{(k)}\right\| = \sum_i \left\|\Delta x_{i+1}^{(k-1)} - \Delta x_i^{(k-1)}\right\|$$

$$\Delta x_{i+1}^{(k)} = \Delta x_{i+1}^{(k-1)} - \Delta x_i^{(k-1)} \quad \Delta x_{i+1}^{(1)} = x_{i+1} - x_i$$

- Ex: with k=1 and the l_1-norm
  - TDP=waveform length
- TDP have been proved a valuable alternative to band-power, being more efficient that BP on average
  [Vidaurre et al, Neural Networks 2009; Bruner et al, BCI conference, 2011]

Spatial filters for TDP

- Objective function

$$J_{TDP^{(k)}} = \frac{\left\| w\Delta_{X_1}^{(k)} \right\|}{\left\| w\Delta_{X_2}^{(k)} \right\|} = \frac{w\Delta_{X_1}^{(k)}{}^{T} \Delta_{X_1}^{(k)} w^{T}}{w\Delta_{X_2}^{(k)}{}^{T} \Delta_{X_2}^{(k)} w^{T}}$$

- Like CSP, this is solved by GEVD

Evaluation

Evaluation for k=1 (waveform length) on N=15 subjects

Features	Classification accuracy
Band Power	68 %
TDP	66.7 %
CSP	77 %
TDP-Spatial Filter	78.7 %
TDP-SF + CSP	80.1 %

Lotte, ICPR 2012

Example 2: Phase synchronization

- Any oscillatory signal can be entirely described by its amplitude, frequency and phase

Amplitude

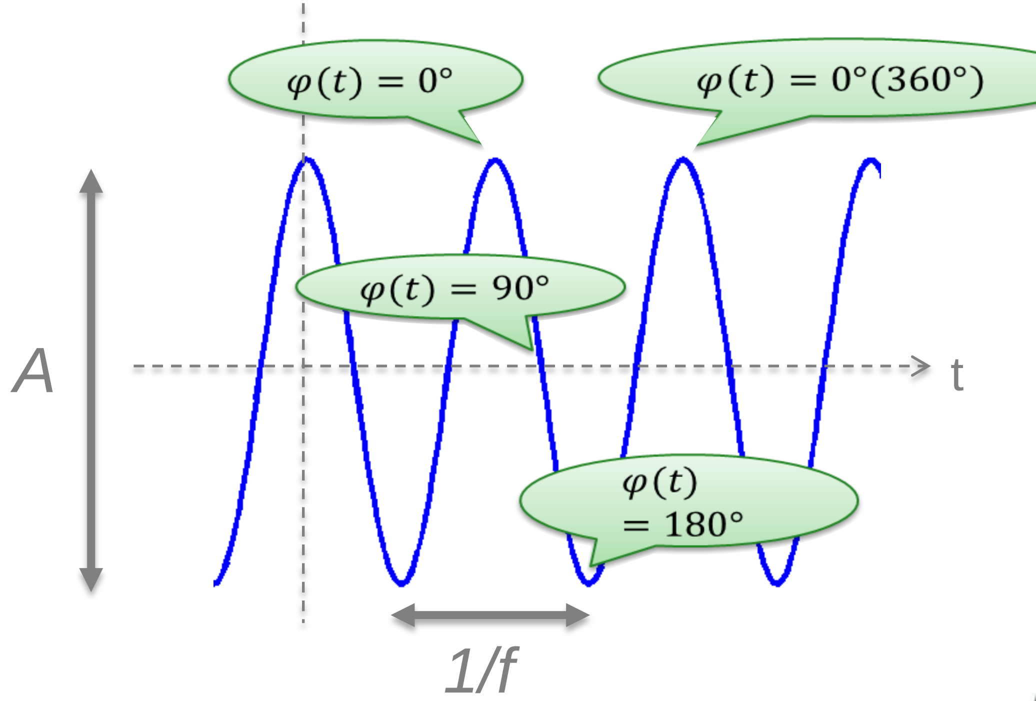
Frequency

$$x(t) = A.e^{i(2\pi f t+\theta)}$$

$$\varphi(t) = 2\pi f t + \theta$$

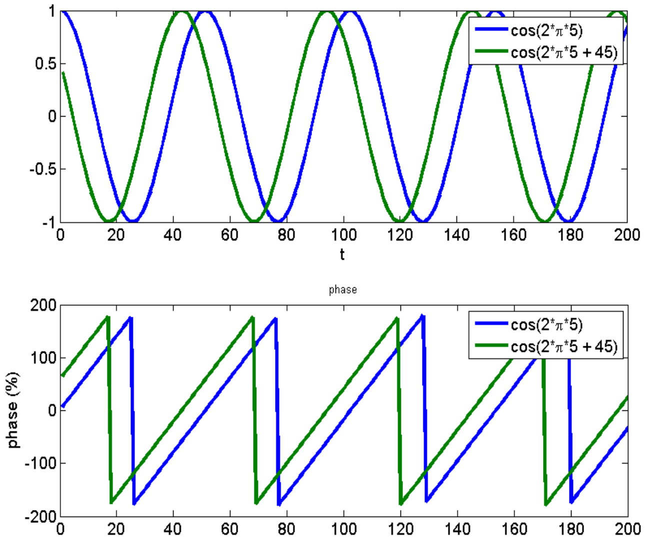
Instantaneous phase

Phase offset



Phase synchronisation (2)

The phase of EEG oscillations provides unique information about cognitive states, in particular the **phase synchronization** between distant brain areas [Varela et al, Nature Reviews Neuroscience, 2001; Sauseng & Klimesh, Neuroscience & Behavioral Reviews, 2008]



Phase synchronization can notably be measured with the Phase Locking Value (PLV) [Lachaux 1999]

$$PLV = \frac{1}{N}\sum_{t=1}^{N} e^{i\left|\varphi_{1(t)} - \varphi_{2(t)}\right|}$$

Phase from area 1 (e.g., one channel)

Phase from area 2 (e.g., another channel)

High PLV => Cooperating brain areas

Phase synchronization for BCI

PLV has been successfully used as a feature for oscillatory activity-based BCI, in particular for Motor Imagery

- Usually not as efficient as BP features
- Generally a good complement to BP features though, often boosting the overall performances

Gysels, Sig. Proc., 2005
Brunner, IEEE TBME, 2006
Wei, JNE, 2007
Krusienski, Brain. Res. Bull., 2012
Daly, Pat. Rec., 2012

Remarks:

- PLV between close EEG channels does not really measure distant synchronization due to volume conduction
- There are many other ways to compute distant synchronizations

Spatial Filter for PLV-based BCI

- PLV features, as any EEG features, may benefit from spatial filtering
- Again, CSP is not optimal for PLV features
  $\Rightarrow$ Need for PLV-specific spatial filters

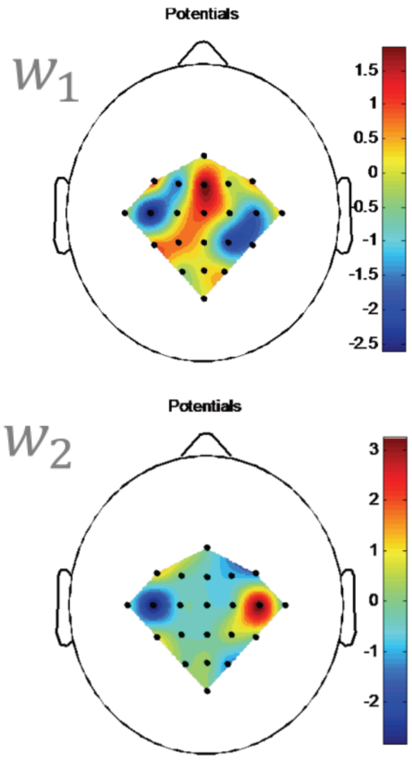
- Optimization of a spatial filter pair to compute the phase synchronization between the 2 areas targeted by such filters, e.g.,

$$J_{PLV}(w_1, w_2) = \left| \underbrace{PLV(w_1^T X_1, w_2^T X_1)}_{\text{Average phase synchronization between the 2 spatially filtered signals for class 1}} - \underbrace{PLV(w_1^T X_2, w_2^T X_2)}_{\text{... for class 2}} \right| + \underbrace{\lambda w_1^T w_2}_{\text{Regularization to ensure distinct filters (and thus areas)}}$$

Spatial Filter for PLV-based BCI: results

- Motor Imagery (left VS right hand), 9 subjects, 8-24Hz band

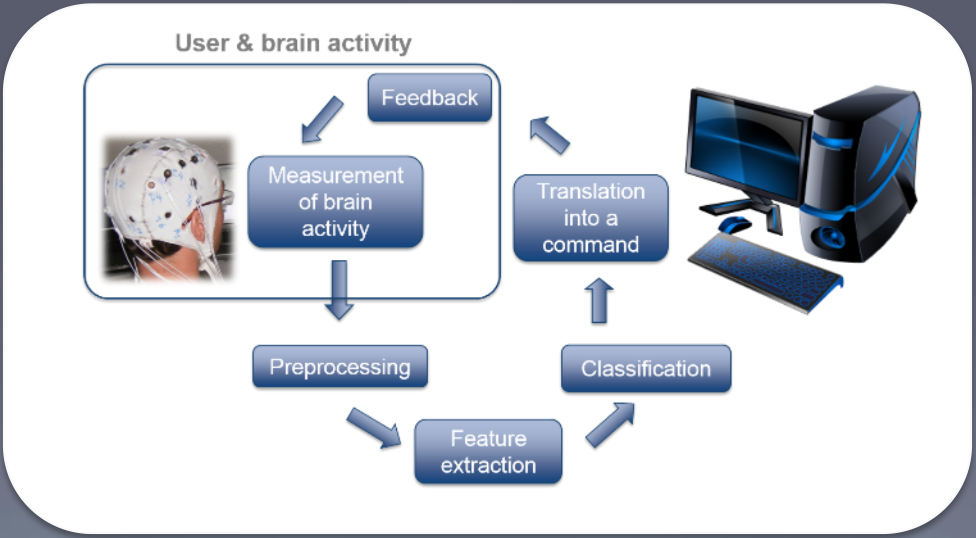
Method	Mean accuracy (%)
Best raw PLV feature (1 feat.)	64,1
PLV (1 feat.) with Spatial Filter (SF)	73,9
CSP (1 pair)	75,8
CSP (3 pairs)	76
PLV-SF (1 feat.) + CSP (1 pair)	**78**
PLV-SF (1 feat.) + CSP (3 pairs)	76,6



Caramia, Lotte, Ramat, ICASSP 2014

6

Feedback & User training



A BCI is a co-adaptive system

- A successful Brain-Computer Interfaces requires the successful interaction between
  - The user (and his/her brain) who produces brain activity patterns
  - The computer, which recognizes such patterns

$\Rightarrow$ Thus, making a good BCI system implies
  - Making good EEG signal processing/machine learning tools
  - Making a good user

BCI skills

« BCI use is a skill »

- BCI performances increase with user training
- Brain activity patterns become more stable and distinct with training

A BCI is unlikely to be efficient and robust without good user's skill in BCI control

- 20-30% of users cannot control BCI (illiteracy/deficiency)

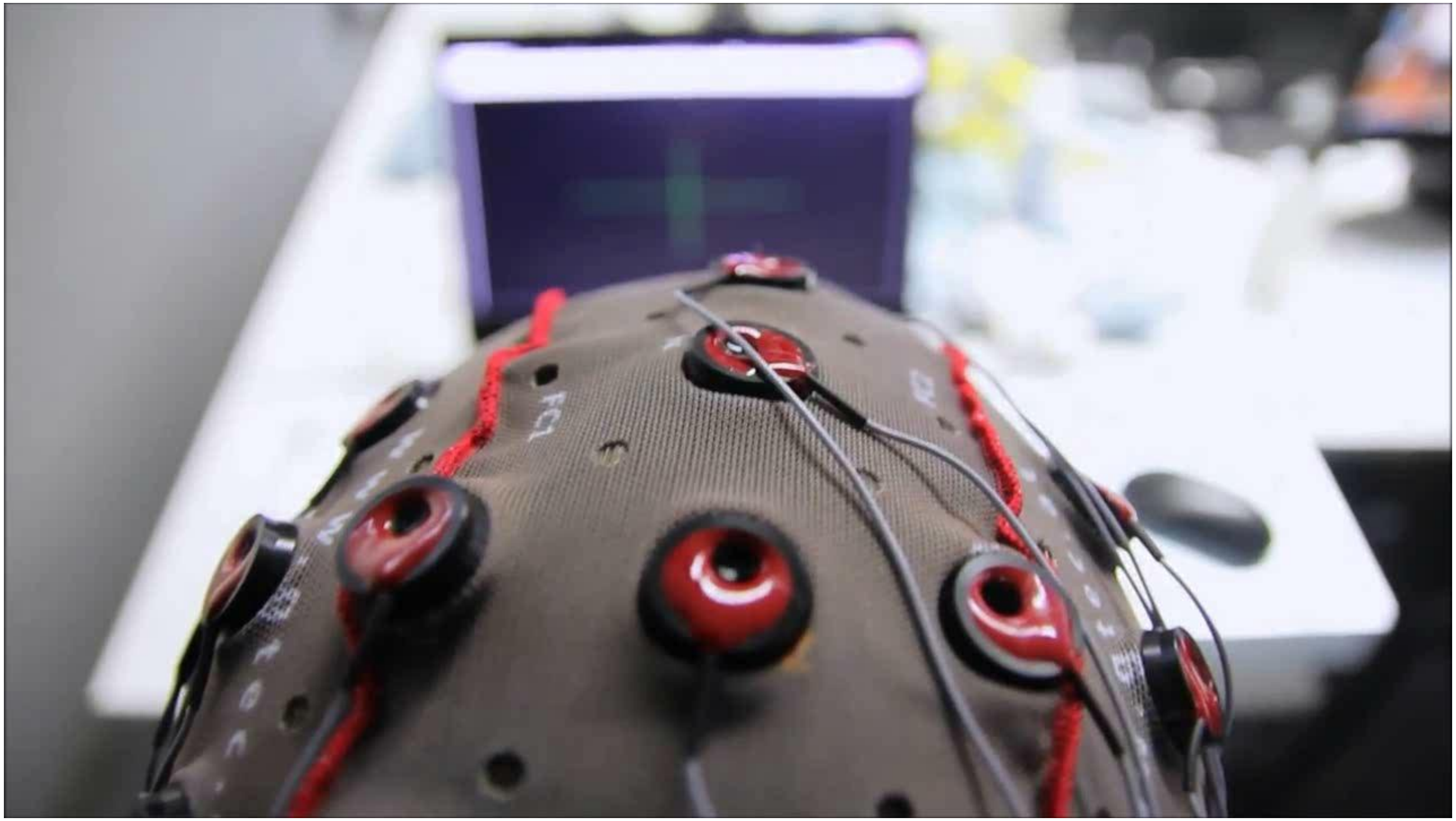
BCI users have to be trained to BCI control

- In particular mental imagery-based BCI
- Feedback is a necessary tool to ensure this training

Wolpaw et al, *Brain-computer interfaces for communication and control,* Clin. Neurophys., 2002
Neuper & Pfurtscheller, *Neurofeedback Training for BCI Control*, The Frontiers Collection, 2010
Allison & Neuper, *Could Anyone Use a BCI?,* Springer London, 2010

What does classical BCI training looks like?

Example for Motor Imagery training



Some remarks on standard BCI feedback

- To give feedback you need a trained classifier
  - BCI training generally starts with a calibration session without feedback

- The difference between calibration (no feedback) and training (with feedback) context often leads to a bias in the classifier
  - Need to correct this bias [Krauledat et al, 2007]

- The user will learn how to control the BCI

  ⇒ The filters/classifier will become out-of-date

  ⇒ Need for
  - Regular retraining of the classifier/filters [Pfurtscheller 2001]
  - Online adaptation of the classifier/filters [Shenoy 2006][Millan 2001][Vidaurre 2011]

On BCI standard user training approaches

- Such standard training approaches, although relatively old, are used a lot
- Comparatively to signal processing, there is relatively little work on user training approaches in BCI
- Nevertheless, some works suggested that user training approaches can be improved to make BCI even more robust

Huggins et al, *“Workshops of the Fifth International Brain-Computer Interface Meeting: Defining the Future”,* Brain-Computer Interfaces journal, 2014

Improving BCI feedback

- Biased and positive feedback increases BCI performance of naive users for motor imagery

  Barbero-Jimenez & Grosse-Wentrup, *Biased Feedback in Brain-Computer Interfaces* Journal of NeuroEngineering and Rehabilitation, 2010

- Rich feedback based on inverse solution improves BCI performances for motor imagery

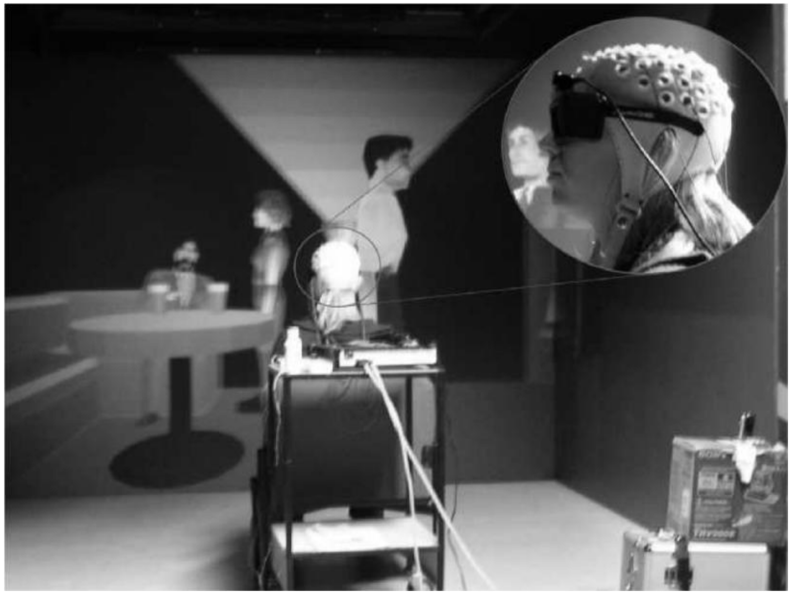
  Hwang et al, *Neurofeedback-based motor imagery training for brain-computer interface (BCI),* Journal of Neuroscience Methods, 2009

- Multimodal feedback (haptic + visual) improves motor imagery BCI performances

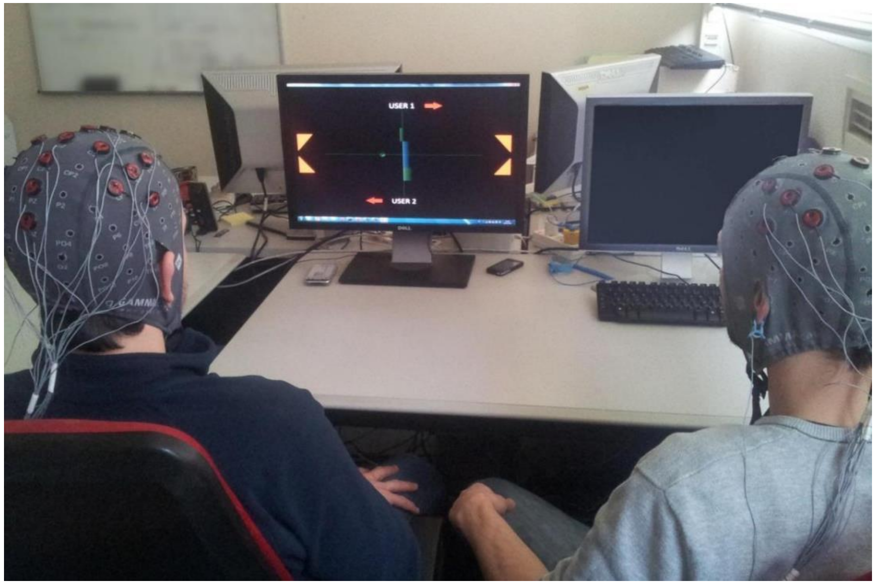
  Gomez-Rodriguez et al, *Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery,* Journal of neural engineering, 2011

Improving BCI training environment

Using motivating and appealing training environment
- Notably virtual reality and gaming



Immersive Virtual Reality
Leeb et al, Presence, 2006



Multiplayer BCI-gaming
Bonnet, Lotte & Lécuyer, *IEEE Trans. Comp. Int. & AI in Games*, 2013

F. Lotte, J. Faller, C. Guger, Y. Renard, G. Pfurtscheller, A. Lécuyer, R. Leeb, *"Combining BCI with Virtual Reality: Towards New Applications and Improved BCI"*, Towards Practical Brain-Computer Interfaces, Springer, 2013

Improving training tasks

- Adaptive training tasks improve sensorimotor BCI

  McFarland et al, *Electroencephalographic (EEG) control of three-dimensional movement,* Journal of Neural Engineering, 2010

- Training tasks that enables to identify and select the best mental tasks for each user improve performance

  Friedrich, Neuper, & Scherer, *Whatever Works: A Systematic User-Centered Training Protocol to Optimize Brain-Computer Interfacing Individually*. PloS one, 2013

  Fruitet et al, *Automatic motor task selection via a bandit algorithm for a brain-controlled button,* Journal of neural engineering, 2013

- Co-Adaptive BCI improve performances

  Vidaurre et al, *Machine-learning-based coadaptive calibration for brain-computer interfaces,* . Neural computation, 2011

The point of view of instructional design

For many years, instructional design and educational psychology research have conducted extensive studies about how to teach someone a skill, and derive guidelines from them:

What BCI training should provide	What standard BCI training actually provides
Explanatory feedback	Corrective feedback only
Multimodal feedback	Unimodal feedback
Motivating/Appealing learning environment	Plain and boring environment
Progressive and adaptive training tasks	Fixed and identical training tasks
Self paced training tasks	Synchronous training tasks

Shute, *Focus on Formative Feedback*, Review of Educational Research, 2008
Sweller et al, *Cognitive Architecture and Instructional Design*, Educational Psychology Review, 1998
Hattie & Timperley, *The Power of Feedback,* Review of Educational Research, 2007
Merrill, *First principles of instruction: a synthesis,* Trends and issues in instructional design and technology, 2007

Summary on user training for BCI

- For oscillatory activity-based BCI, user training and feedback is necessary
- They are known ways to improve user training
  - Unfortunately they are generally not used in standard BCI training protocols
- They are relatively little research work on user training but promising opportunities according to instructional design

F. Lotte, F. Larrue, C. Mühl, *Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design*, Frontiers in Human Neurosciences, vol. 7, no. 568, 2013

Summary
and
Conclusion

Summary on oscillatory activity-based BCI-design

1. Oscillatory activity-based BCI exploit the power of EEG oscillations (spectral information) in some specific channels (spatial information)
2. Spatial filters, and notably CSP are essential for efficient BCI design
3. CSP can be made more robust with regularization and extended to the continuous case
4. Spectral filtering can (and should be) optimized for each user
5. Alternative features (e.g., TDP, synchronization) are valuable complements to band power and need dedicated spatial filters
6. User training and feedback is necessary

Some related open research challenges

Signal processing: finding features and filters that are

- More Informative
  - ➔ To reach better classification performances
- Robust to noise & artifacts
  - ➔ To use outside laboratories and/or with moving users
- Invariant
  - ➔ To deal with non-stationarity (within and between sessions)
  - ➔ To deal with between subject variability

User training

- Finding optimal feedback and training tasks to ensure efficient BCI control skills for all users

Take home messages for oscillatory activity EEG-based BCI design

A form of spatial filtering is essential

- Whether explicit (e.g., CSP), or implicit (covariance matrix space)
- The right filter must be used for the right features

User training is also essential

- The user must learn the BCI control skill
- User training approaches deserve more research

Thank you for your attention!

Any question?

Fabien Lotte
web: http://sites.google.com/site/fabienlotte/
Potioc team web: http://team.inria.fr/potioc
e-mail: fabien.lotte@inria.fr