

M/EEG Decoding and Brain-Computer Interfacing

Moritz Grosse-Wentrup

Max Planck Institute for Intelligent Systems
Department Empirical Inference
Tübingen, Germany

April 17, 2014



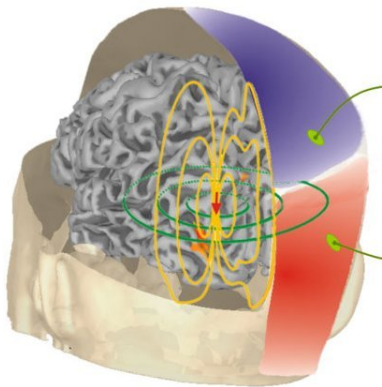
MAX-PLANCK-GESELLSCHAFT



- 1 M/EEG Decoding Models
- 2 Brain-Computer Interfacing

- 1 M/EEG Decoding Models
- 2 Brain-Computer Interfacing

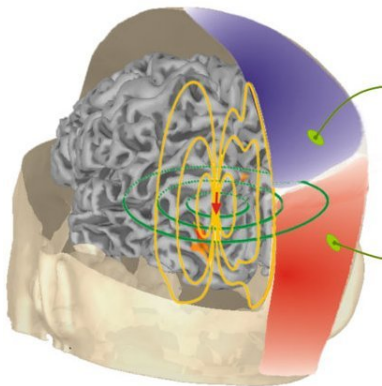
The brain's electromagnetic field (EMF)



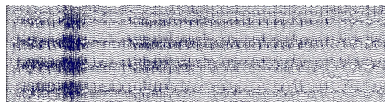
<http://www.canada-meg-consortium.org/>

(P.L. Nunez. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press, 2005.)

The brain's electromagnetic field (EMF)



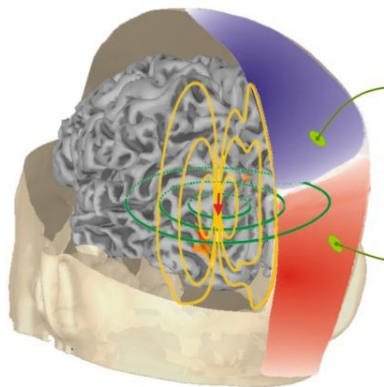
<http://www.canada-meg-consortium.org/>



Time

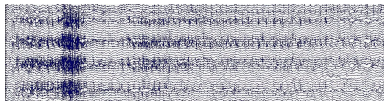
(P.L. Nunez. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, 2005.)

The brain's electromagnetic field (EMF)



Notation:

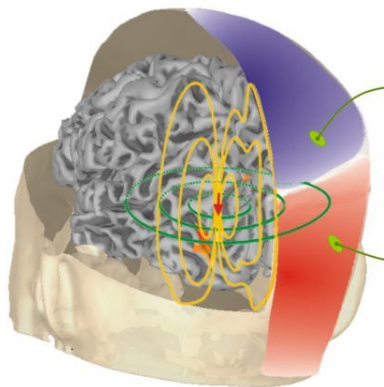
<http://www.canada-meg-consortium.org/>



Time

(P.L. Nunez. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press, 2005.)

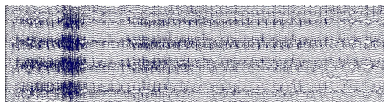
The brain's electromagnetic field (EMF)



Notation:

- M/EEG signal $\mathbf{x}_j[t] \in \mathbb{R}^N$

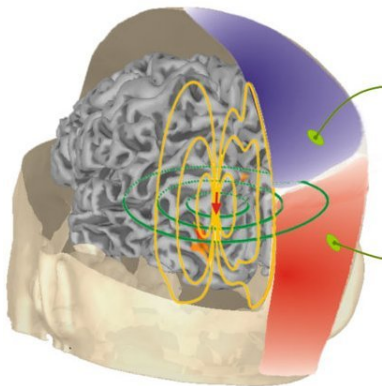
<http://www.canada-meg-consortium.org/>



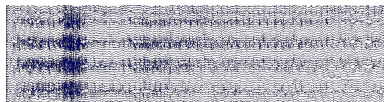
Time

(P.L. Nunez. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, 2005.)

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>

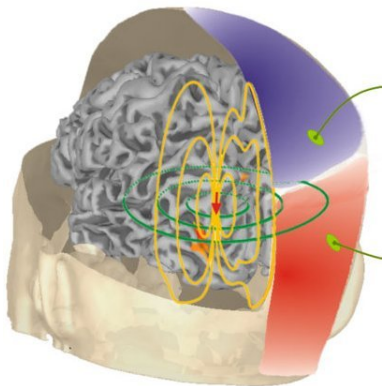


Time

Notation:

- M /EEG signal $\mathbf{x}_j[t] \in \mathbb{R}^N$
- N electrodes (32–256)

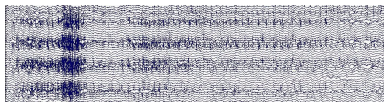
The brain's electromagnetic field (EMF)



Notation:

- M/EEG signal $\mathbf{x}_j[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$

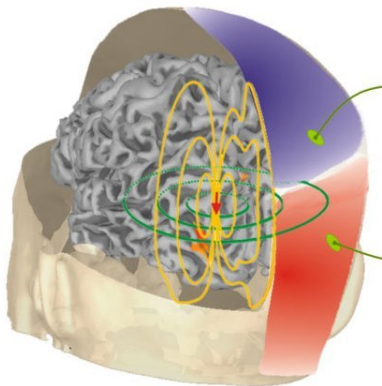
<http://www.canada-meg-consortium.org/>



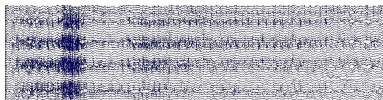
Time

(P.L. Nunez. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press, 2005.)

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>

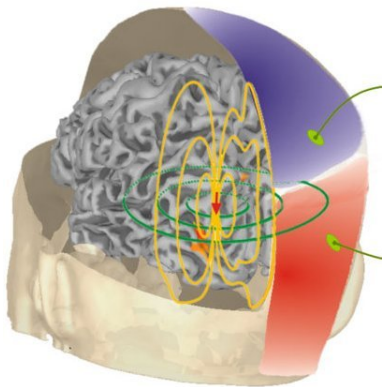


Time

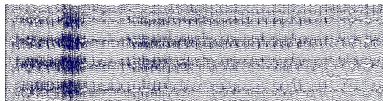
Notation:

- M /EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>

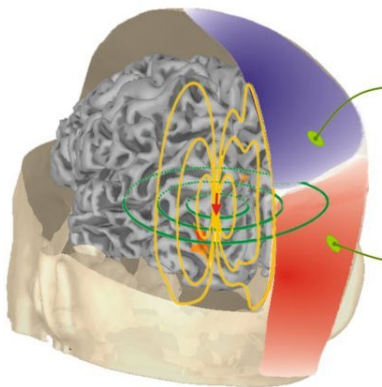


Time

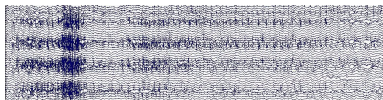
Notation:

- M/EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$
- Exp. condition $c_i \in \{-1, +1\}$

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>



Time

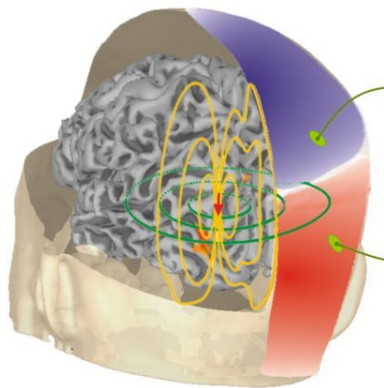
Notation:

- M /EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$
- Exp. condition $c_i \in \{-1, +1\}$

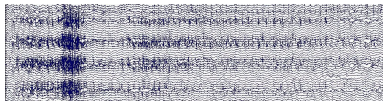
Experimental data:

(P.L. Nunez. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press, 2005.)

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>



Time

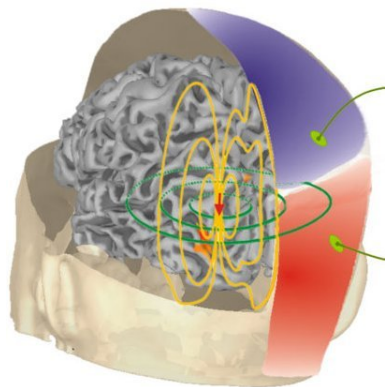
Notation:

- M/EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$
- Exp. condition $c_i \in \{-1, +1\}$

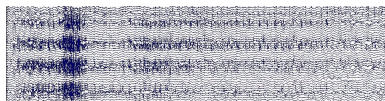
Experimental data:

- $(c_i, X_i) \sim p(c, X)$

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>



Time

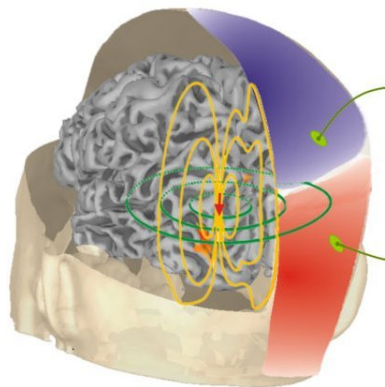
Notation:

- M/EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$
- Exp. condition $c_i \in \{-1, +1\}$

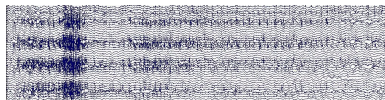
Experimental data:

- $(c_i, X_i) \sim p(c, X)$
- $\mathcal{D} = \{(c_1, X_1), \dots, (c_M, X_M)\}$

The brain's electromagnetic field (EMF)



<http://www.canada-meg-consortium.org/>



Time

Notation:

- M/EEG signal $\mathbf{x}_i[t] \in \mathbb{R}^N$
- N electrodes (32–256)
- Time $t = 1, \dots, T$
- Trial $i = 1, \dots, M$
- Exp. condition $c_i \in \{-1, +1\}$

Experimental data:

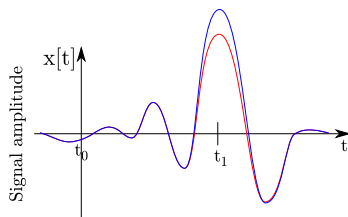
- $(c_i, X_i) \sim p(c, X)$
- $\mathcal{D} = \{(c_1, X_1), \dots, (c_M, X_M)\}$
- Typically i.i.d. sampling is assumed

Encoding models

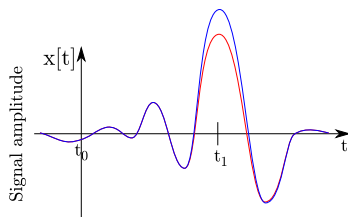
- How does the brain respond to an experimental condition: $p(X|c)$?

- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials

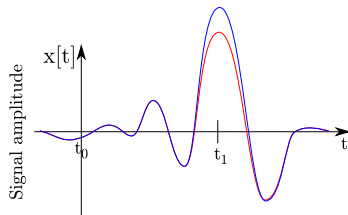
- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials



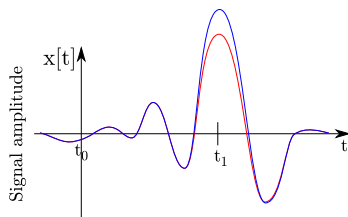
- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials $E(x[t_1]|c = +1) > E(x[t_1]|c = -1)$?



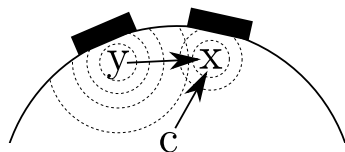
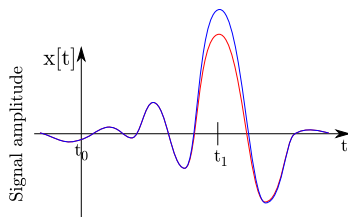
- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials $E(x[t_1]|c = +1) > E(x[t_1]|c = -1)$?
- Disadvantage I: Hypothesis-driven



- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials $E(x[t_1]|c = +1) > E(x[t_1]|c = -1)$?
- Disadvantage I: Hypothesis-driven
- Disadvantage II: Not robust against noise.

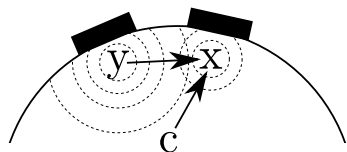
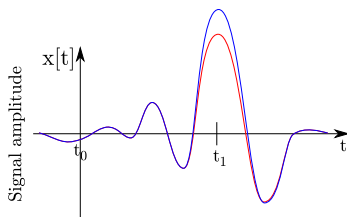


- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials $E(x[t_1]|c = +1) > E(x[t_1]|c = -1)$?
- Disadvantage I: Hypothesis-driven
- Disadvantage II: Not robust against noise.



$$x_i[t_1] = \alpha c_i + \beta y_i[t_1]$$

- How does the brain respond to an experimental condition: $p(X|c)$?
- Example: Event related potentials $E(x[t_1]|c = +1) > E(x[t_1]|c = -1)$?
- Disadvantage I: Hypothesis-driven
- Disadvantage II: Not robust against noise. If $|\beta| \gg |\alpha|$ then the task-irrelevant source y may make it hard to find any effect of c on x .



$$x_i[t_1] = \alpha c_i + \beta y_i[t_1]$$

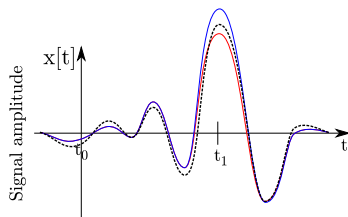
Decoding models

Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?

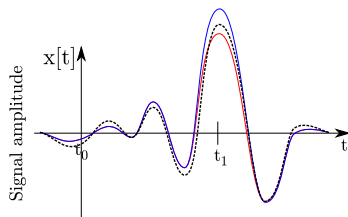
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?



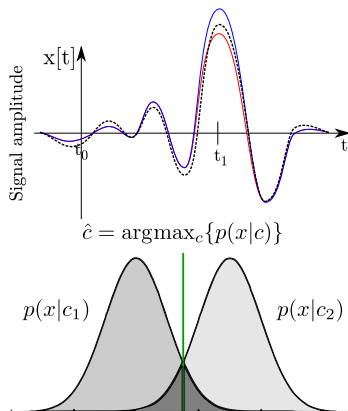
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$



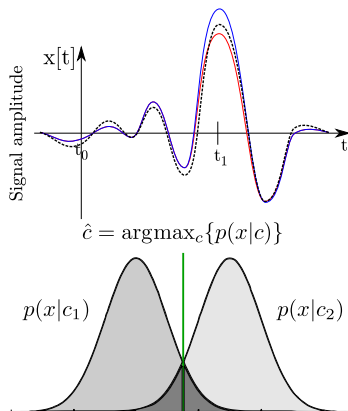
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$



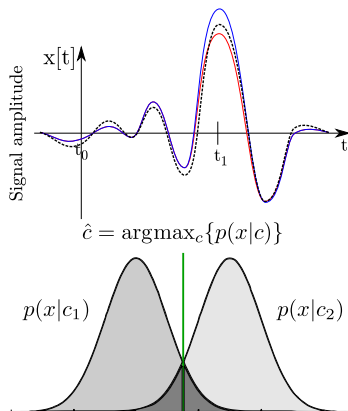
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$



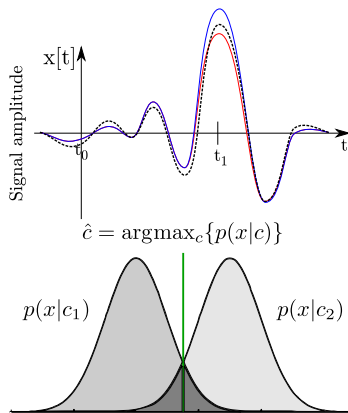
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - ▶ Linear discriminant analysis (LDA)
 - ▶ Support Vector Machines (SVM)
 - ▶ Gaussian Processes (GP)



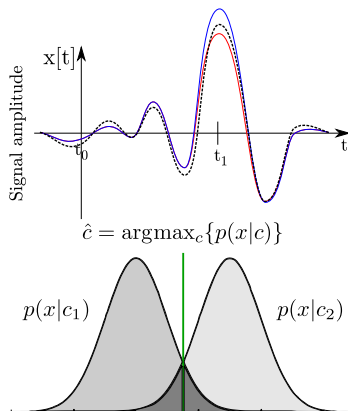
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - ▶ Linear discriminant analysis (LDA)
 - ▶ Support Vector Machines (SVM)
 - ▶ Gaussian Processes (GP)
- Prediction error $P_e := 1 - P(\hat{c} = c)$



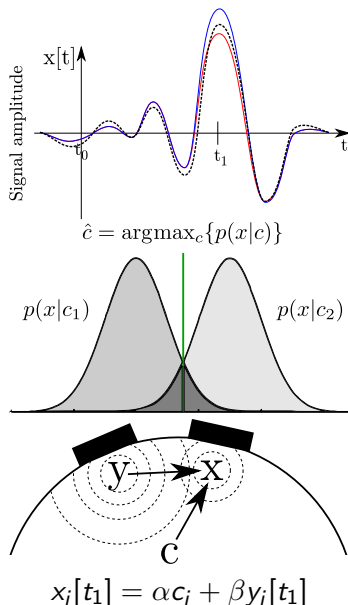
Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - ▶ Linear discriminant analysis (LDA)
 - ▶ Support Vector Machines (SVM)
 - ▶ Gaussian Processes (GP)
- Prediction error $P_e := 1 - P(\hat{c} = c)$
- Advantage: Robust against (measurable) noise.



Decoding models

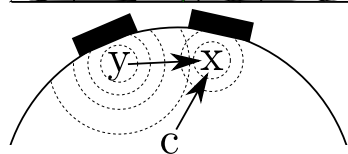
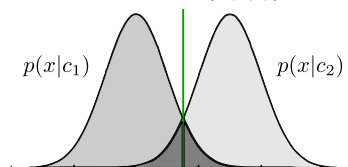
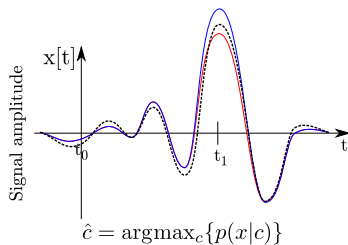
- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - Linear discriminant analysis (LDA)
 - Support Vector Machines (SVM)
 - Gaussian Processes (GP)
- Prediction error $P_e := 1 - P(\hat{c} = c)$
- Advantage: Robust against (measurable) noise.



Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - Linear discriminant analysis (LDA)
 - Support Vector Machines (SVM)
 - Gaussian Processes (GP)
- Prediction error $P_e := 1 - P(\hat{c} = c)$
- Advantage: Robust against (measurable) noise.

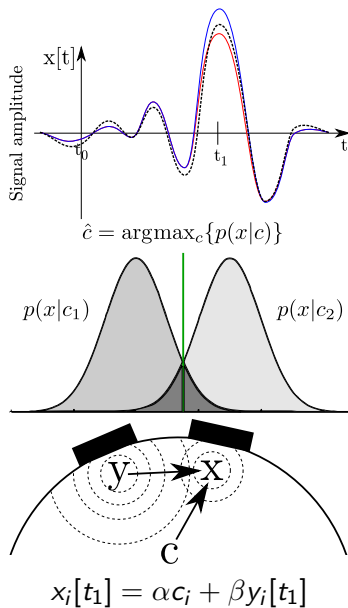
$$\hat{c}_i = \underbrace{\operatorname{sign}\{x_i[t] - \beta y[t_1]\}}_f = c_i$$



$$x_i[t_1] = \alpha c_i + \beta y_i[t_1]$$

Decoding models

- Given a certain brain-state, what is the probability of an experimental condition: $p(c|X)$?
- The (optimal) Bayes classifier:
 $\hat{c}_i = \operatorname{argmax}_c \{p(X_i|c)\}$
- Practical approach: $\hat{c}_i = f(X_i)$
 - Linear discriminant analysis (LDA)
 - Support Vector Machines (SVM)
 - Gaussian Processes (GP)
- Prediction error $P_e := 1 - P(\hat{c} = c)$
- Advantage: Robust against (measurable) noise.
 $\hat{c}_i = \underbrace{\operatorname{sign}\{x_i[t] - \beta y[t_1]\}}_f = c_i$
- Disadvantage $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$ needs to be learned from \mathcal{D} .



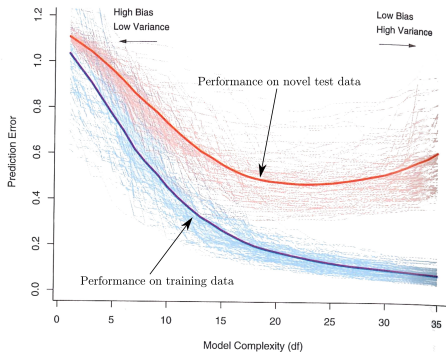
Learning decoding models I

(Hastie, Tibshirani, & Friedman. The Elements of Statistical Learning. Springer, 2009)

The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?

Learning decoding models I

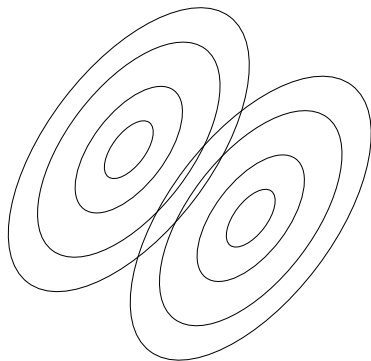
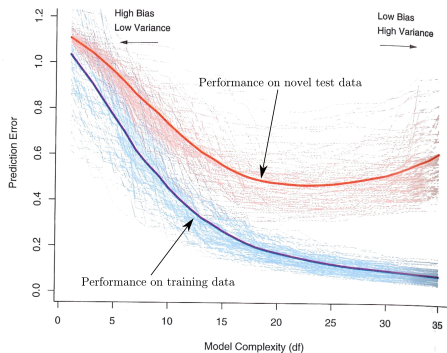
The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



(Hastie, Tibshirani, & Friedman. The Elements of Statistical Learning. Springer, 2009)

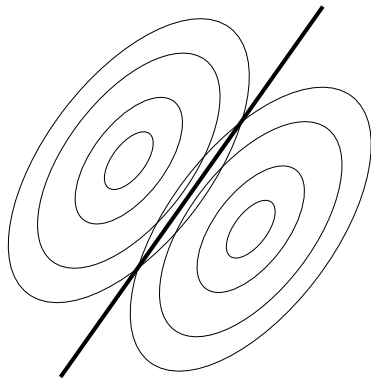
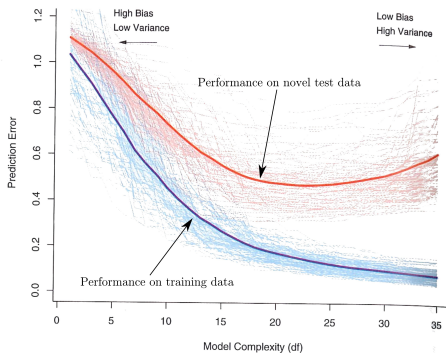
Learning decoding models I

The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



Learning decoding models I

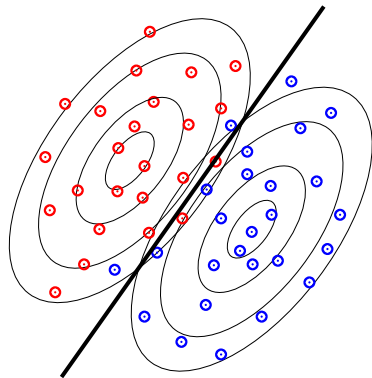
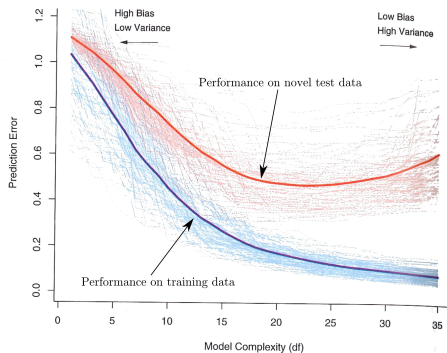
The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



(Hastie, Tibshirani, & Friedman. The Elements of Statistical Learning. Springer, 2009)

Learning decoding models I

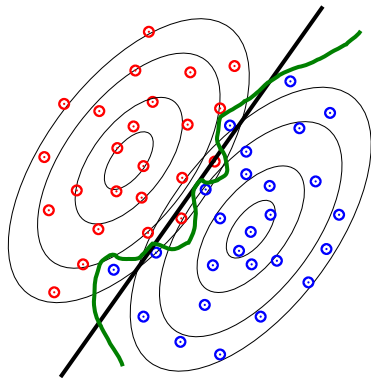
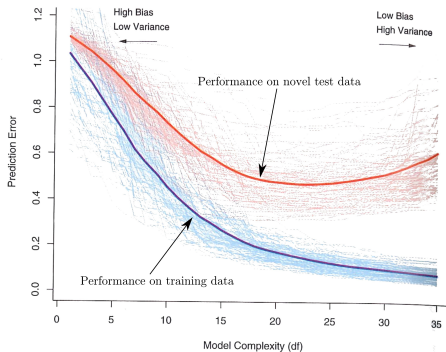
The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



(Hastie, Tibshirani, & Friedman. The Elements of Statistical Learning. Springer, 2009)

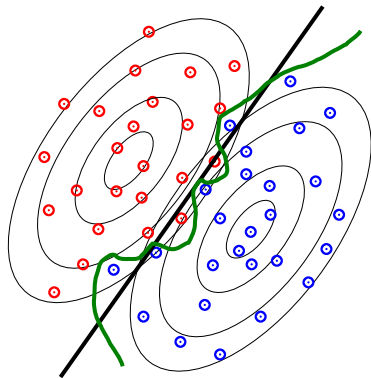
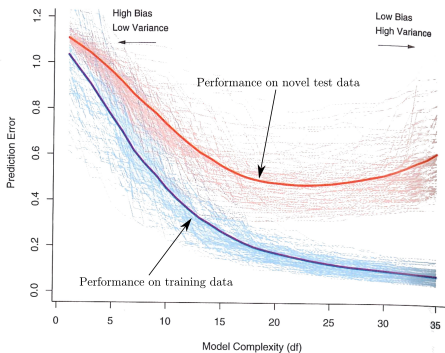
Learning decoding models I

The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



Learning decoding models I

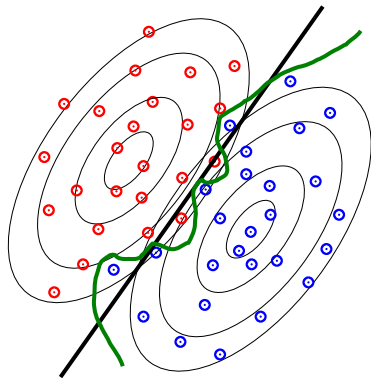
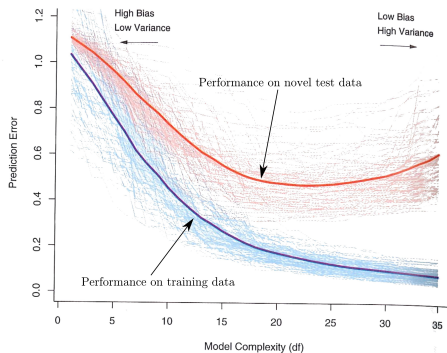
The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



How do we determine which model generalizes best?

Learning decoding models I

The decoder f has to be chosen from a model class \mathcal{F} . How to choose \mathcal{F} ?



How do we determine which model generalizes best? Cross-validation!

Learning decoding models II

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} .

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

- the model complexity,

Learning decoding models II

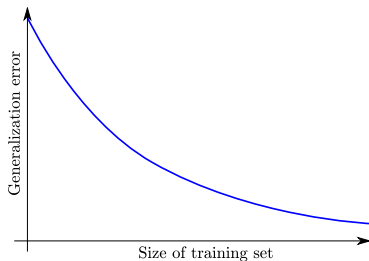
In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

- the model complexity,
- and the number of features $N \times T$.

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

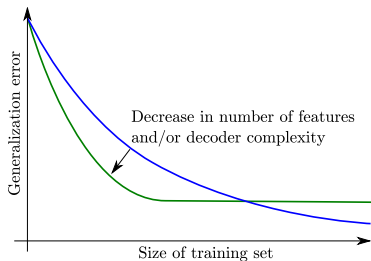
- the model complexity,
- and the number of features $N \times T$.



Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

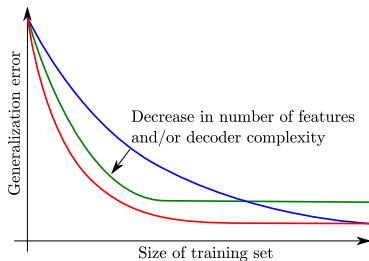
- the model complexity,
- and the number of features $N \times T$.



Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

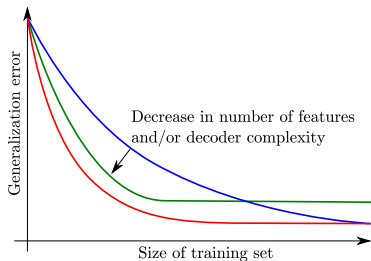
- the model complexity,
- and the number of features $N \times T$.



Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

- the model complexity,
- and the number of features $N \times T$.

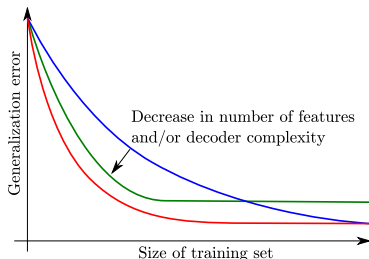


To learn a good classifier with limited training data, we should

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

- the model complexity,
- and the number of features $N \times T$.



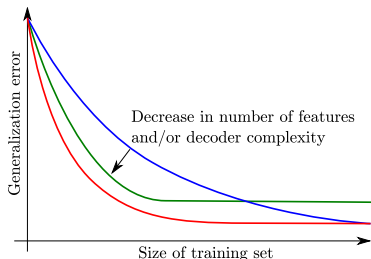
To learn a good classifier with limited training data, we should

- reduce N and T without discarding information relevant for c ,

Learning decoding models II

In order to learn $f : \mathbb{R}^{N \times T} \mapsto \{-1, +1\}$, we require training data \mathcal{D} . The number of training samples needed to find the best $f \in \mathcal{F}$ scales with

- the model complexity,
- and the number of features $N \times T$.



To learn a good classifier with limited training data, we should

- reduce N and T without discarding information relevant for c ,
- and find a representation of X that allows us to use a simple model class, e.g. a linear decoder.

Reducing the number of M/EEG channels (N)

(Baillet, Mosher & Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 2001)

Reducing the number of M/EEG channels (N)

Spatial filtering of M/EEG data:

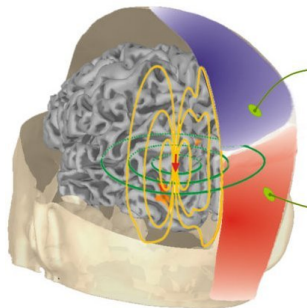
(Baillet, Mosher & Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 2001)

Reducing the number of M/EEG channels (N)

Spatial filtering of M/EEG data:

$$\mathbf{x}[t] = \mathbf{L}\mathbf{s}[t]$$

- Source vector $\mathbf{s}[t] \in \mathbb{R}^K$
- Leadfield matrix $\mathbf{L} \in \mathbb{R}^{N \times K}$



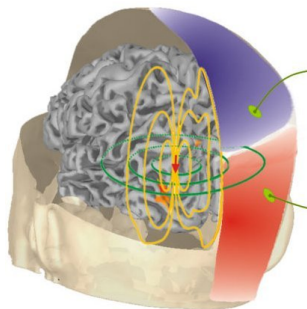
<http://www.canada-meg-consortium.org/>

Reducing the number of M/EEG channels (N)

Spatial filtering of M/EEG data:

$$y[t] = \mathbf{w}^T \mathbf{x}[t] = \mathbf{w}^T \mathbf{L} \mathbf{s}[t]$$

- Source vector $\mathbf{s}[t] \in \mathbb{R}^K$
- Leadfield matrix $\mathbf{L} \in \mathbb{R}^{N \times K}$
- Spatial filter $\mathbf{w} \in \mathbb{R}^N$



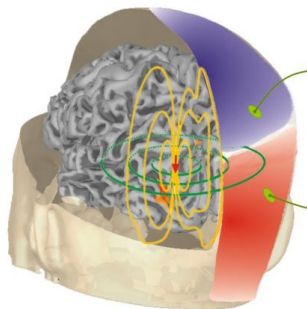
<http://www.canada-meg-consortium.org/>

Reducing the number of M/EEG channels (N)

Spatial filtering of M/EEG data:

$$y[t] = \mathbf{w}^T \mathbf{x}[t] = \mathbf{w}^T \mathbf{L} \mathbf{s}[t] = \mathbf{g}^T \mathbf{s}[t]$$

- Source vector $\mathbf{s}[t] \in \mathbb{R}^K$
- Leadfield matrix $\mathbf{L} \in \mathbb{R}^{N \times K}$
- Spatial filter $\mathbf{w} \in \mathbb{R}^N$
- Gain vector $\mathbf{g} \in \mathbb{R}^K$



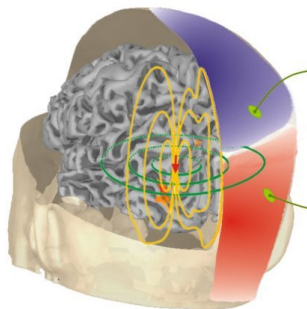
<http://www.canada-meg-consortium.org/>

Reducing the number of M/EEG channels (N)

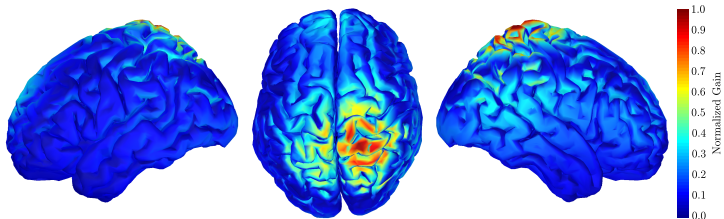
Spatial filtering of M/EEG data:

$$y[t] = \mathbf{w}^T \mathbf{x}[t] = \mathbf{w}^T \mathbf{L} \mathbf{s}[t] = \mathbf{g}^T \mathbf{s}[t]$$

- Source vector $\mathbf{s}[t] \in \mathbb{R}^K$
- Leadfield matrix $\mathbf{L} \in \mathbb{R}^{N \times K}$
- Spatial filter $\mathbf{w} \in \mathbb{R}^N$
- Gain vector $\mathbf{g} \in \mathbb{R}^K$



<http://www.canada-meg-consortium.org/>



(Baillet, Moshier & Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 2001)

Beamforming

(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Beamforming

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

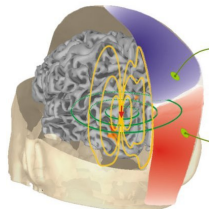
Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*

Beamforming

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*

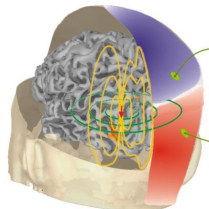


<http://www.canada-meg-consortium.org/>

Beamforming

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

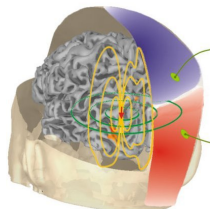
- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$



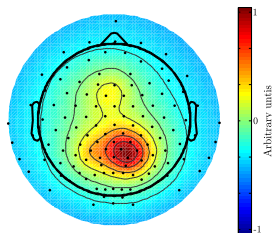
<http://www.canada-meg-consortium.org/>

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$



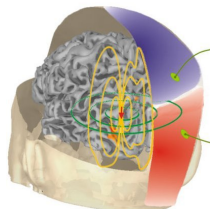
<http://www.canada-meg-consortium.org/>



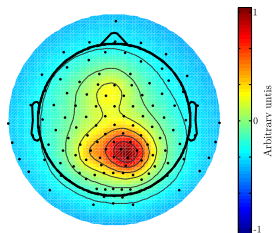
(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$



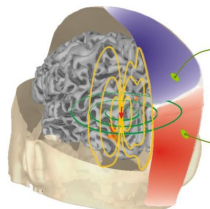
<http://www.canada-meg-consortium.org/>



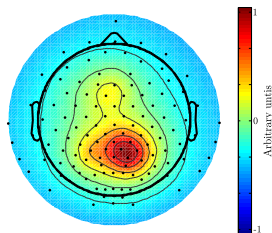
(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$
- Solve the optimization problem $\mathbf{w} = \operatorname{argmin}_{\mathbf{w}} \{\mathbf{w}^T \Sigma \mathbf{w}\}$ s.t. $\mathbf{w}^T \mathbf{a} = 1$



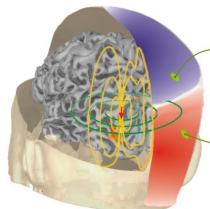
<http://www.canada-meg-consortium.org/>



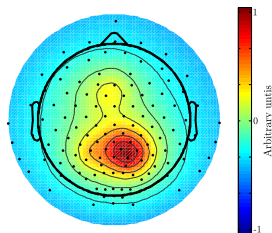
(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$
- Solve the optimization problem $\mathbf{w} = \operatorname{argmin}_{\mathbf{w}} \{\mathbf{w}^T \Sigma \mathbf{w}\}$ s.t. $\mathbf{w}^T \mathbf{a} = 1$

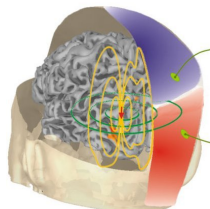


<http://www.canada-meg-consortium.org/>

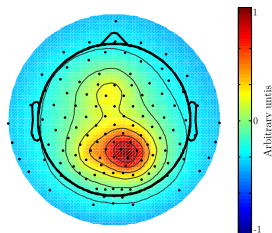


Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{I}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$
- Solve the optimization problem $\mathbf{w} = \operatorname{argmin}_{\mathbf{w}} \{\mathbf{w}^T \Sigma \mathbf{w}\} \text{ s.t. } \mathbf{w}^T \mathbf{a} = 1$
- $\mathbf{w} = \mathbf{a}^T \Sigma^{-1} / (\mathbf{a}^T \Sigma^{-1} \mathbf{a})$



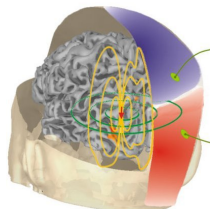
<http://www.canada-meg-consortium.org/>



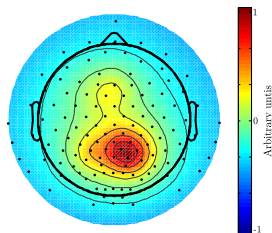
(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{L}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$
- Solve the optimization problem $\mathbf{w} = \operatorname{argmin}_{\mathbf{w}} \{\mathbf{w}^T \Sigma \mathbf{w}\}$ s.t. $\mathbf{w}^T \mathbf{a} = 1$
- $\mathbf{w} = \mathbf{a}^T \Sigma^{-1} / (\mathbf{a}^T \Sigma^{-1} \mathbf{a})$
- Check the gain vector $\mathbf{g} = \mathbf{w}^T \mathbf{L}$



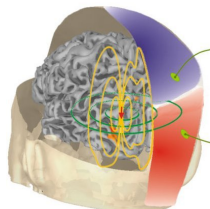
<http://www.canada-meg-consortium.org/>



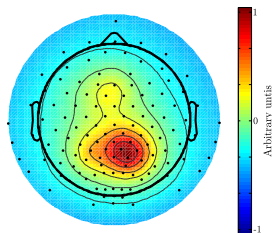
(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Unsupervised method based on a-priori knowledge of the spatial origin of relevant sources:

- Pick a cortical target source s^*
- Compute the forward solution $\mathbf{a} = \mathbf{L}s^*$
- Compute the M/EEG covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$
- Solve the optimization problem $\mathbf{w} = \operatorname{argmin}_{\mathbf{w}} \{\mathbf{w}^T \Sigma \mathbf{w}\} \text{ s.t. } \mathbf{w}^T \mathbf{a} = 1$
- $\mathbf{w} = \mathbf{a}^T \Sigma^{-1} / (\mathbf{a}^T \Sigma^{-1} \mathbf{a})$
- Check the gain vector $\mathbf{g} = \mathbf{w}^T \mathbf{L}$
- Apply the spatial filter: $y[t] = \mathbf{w}^T \mathbf{x}[t]$



<http://www.canada-meg-consortium.org/>



(Van Veen et al. Localization of brain electrical activity via LCMV spatial filtering. *IEEE TBME*, 1997)

Common Spatial Patterns (CSP)

(Ramoser et al. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE TBME*, 2000)

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

(Ramoser et al. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE TBME*, 2000)

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$
- Matlab: $[V, D] = \operatorname{eig}(\Sigma_{c=+1}, \Sigma_{c=-1})$

Common Spatial Patterns (CSP)

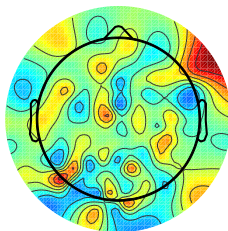
Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$
- Matlab: $[V, D] = \operatorname{eig}(\Sigma_{c=+1}, \Sigma_{c=-1})$
- Columns of V are spatial filters

Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

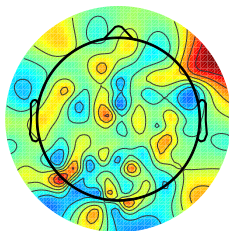
- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$
- Matlab: $[V, D] = \operatorname{eig}(\Sigma_{c=+1}^{-1} \Sigma_{c=-1})$
- Columns of V are spatial filters



Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

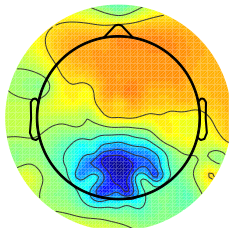
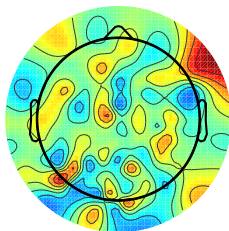
- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$
- Matlab: $[V, D] = \operatorname{eig}(\Sigma_{c=+1}^{-1} \Sigma_{c=-1})$
- Columns of V are spatial filters
- Rows of V are spatial patterns



Common Spatial Patterns (CSP)

Supervised method to find spatial filters that discriminate between two conditions:

- Covariance matrices $\Sigma_{c=+1}$ & $\Sigma_{c=-1}$
- Optimization problem:
$$\mathbf{w} = \operatorname{argmax}_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \Sigma_{c=+1} \mathbf{w}}{\mathbf{w}^T \Sigma_{c=-1} \mathbf{w}} \right\}$$
- Solution: Largest λ for which
$$\Sigma_{c=+1}^{-1} \Sigma_{c=-1} \mathbf{w} = \lambda \mathbf{w}$$
- Matlab: $[V, D] = \operatorname{eig}(\Sigma_{c=+1}, \Sigma_{c=-1})$
- Columns of V are spatial filters
- Rows of V are spatial patterns



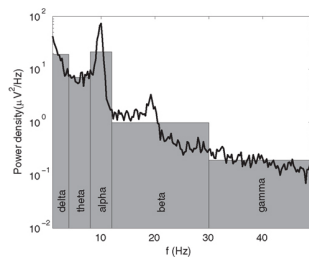
Reducing the number of time points (T)

Reducing the number of time points (T)

The brain cares about oscillations:

Reducing the number of time points (T)

The brain cares about oscillations:

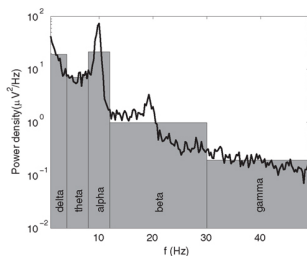


(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)

Reducing the number of time points (T)

The brain cares about oscillations:

- $\text{DTFT}(y_i[t], \omega) = 1/T \sum_{t=0}^{T-1} y_i[t] e^{-j\omega t}$

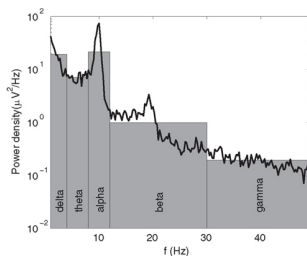


(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)

Reducing the number of time points (T)

The brain cares about oscillations:

- $\text{DTFT}(y_i[t], \omega) = 1/T \sum_{t=0}^{T-1} y_i[t] e^{-j\omega t}$
- $z_i[\omega] = \log|\text{DTFT}(y_i[t], \omega)|$

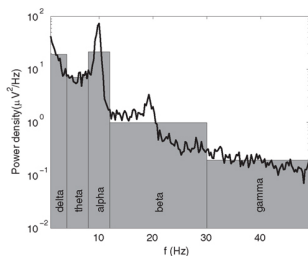


(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)

Reducing the number of time points (T)

The brain cares about oscillations:

- $\text{DTFT}(y_i[t], \omega) = 1/T \sum_{t=0}^{T-1} y_i[t] e^{-j\omega t}$
- $z_i[\omega] = \log|\text{DTFT}(y_i[t], \omega)|$
- $\mathcal{Z}_i = \{z_i[\delta], z_i[\theta], z_i[\alpha], z_i[\beta], z_i[\gamma]\}$

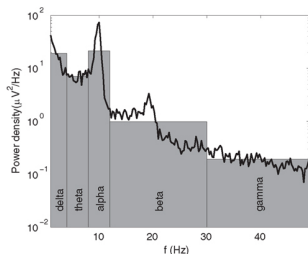


(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)

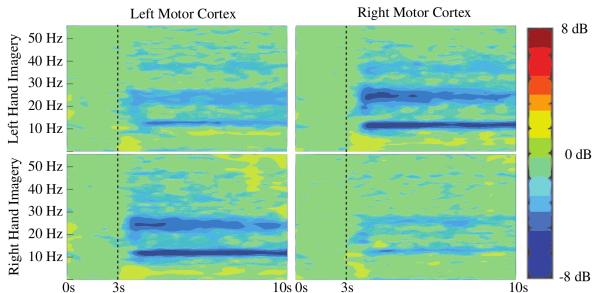
Reducing the number of time points (T)

The brain cares about oscillations:

- $\text{DTFT}(y_i[t], \omega) = 1/T \sum_{t=0}^{T-1} y_i[t] e^{-j\omega t}$
- $z_i[\omega] = \log|\text{DTFT}(y_i[t], \omega)|$
- $\mathcal{Z}_i = \{z_i[\delta], z_i[\theta], z_i[\alpha], z_i[\beta], z_i[\gamma]\}$



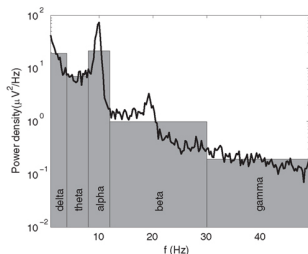
(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)



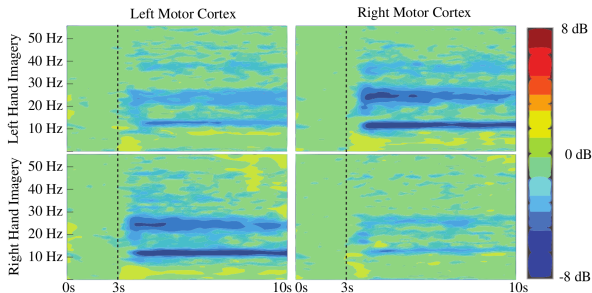
Reducing the number of time points (T)

The brain cares about oscillations:

- $\text{DTFT}(y_i[t], \omega) = 1/T \sum_{t=0}^{T-1} y_i[t] e^{-j\omega t}$
- $z_i[\omega] = \log|\text{DTFT}(y_i[t], \omega)|$
- $\mathcal{Z}_i = \{z_i[\delta], z_i[\theta], z_i[\alpha], z_i[\beta], z_i[\gamma]\}$
- For log-bandpower features, linear decoders appear sufficient.



(van Albada & Robinson, *Frontiers in Human Neuroscience*, 2013)



Interpretation of encoding- and decoding models

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically		
Example		

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)$?	
Example		

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$		
$\mathcal{Z} \rightarrow$ Response		

Interpretation of encoding- and decoding models

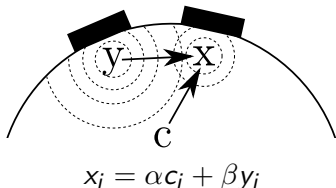
- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$	$c \rightarrow z$	$c - z$
$\mathcal{Z} \rightarrow$ Response		

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$	$c \rightarrow z$	$c - z$
$\mathcal{Z} \rightarrow$ Response		

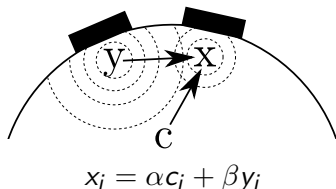


(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$	$c \rightarrow z$	$c - z$
$\mathcal{Z} \rightarrow$ Response	$c - z$	$c - z$

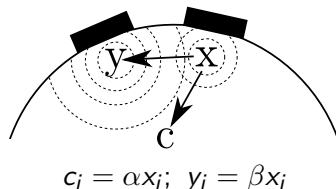
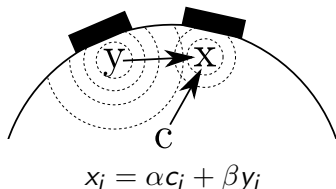


(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

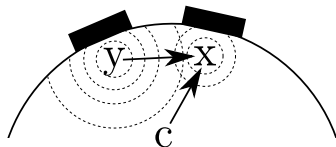
Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$	$c \rightarrow z$	$c - z$
$\mathcal{Z} \rightarrow$ Response	$c - z$	$c - z$



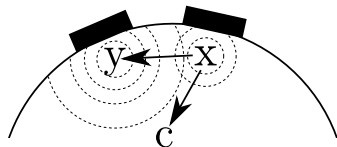
Interpretation of encoding- and decoding models

- How do we determine whether a feature $z \in \mathcal{Z}$ is relevant in an experimental setting?
- If a feature is relevant, what does this mean?

Type of model	Encoding model	Decoding model
Theoretically	$p(z c = 1) \neq p(z c = -1)?$	$p(c \mathcal{Z}) \neq p(c \mathcal{Z}\setminus z)?$
Example	$\mu(z_i c_i = 1) = \mu(z_i c_i = -1)?$	$Pe(\mathcal{Z}) > Pe(\mathcal{Z}\setminus z)?$
Stimulus $\rightarrow \mathcal{Z}$	$c \rightarrow z$	$c - z$
$\mathcal{Z} \rightarrow$ Response	$c - z$	$c - z$



$$x_i = \alpha c_i + \beta y_i$$



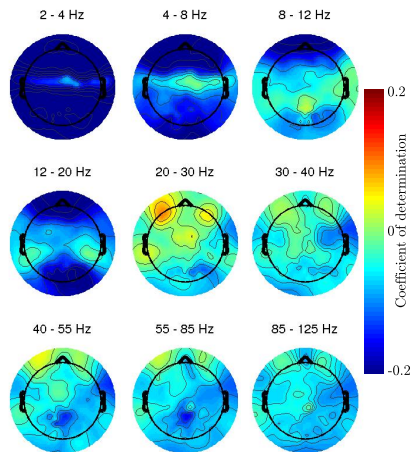
$$c_i = \alpha x_i; y_i = \beta x_i$$

videlectures.net/bbci2014_grosse_wentrup_causal_inference

(Weichwald et al. Causal and anti-causal learning in pattern recognition for neuroimaging. *PRNI*, 2014)

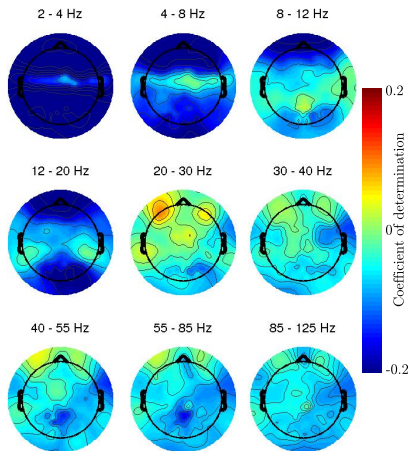
Confounding by EOG-artifacts

Eye-blinking

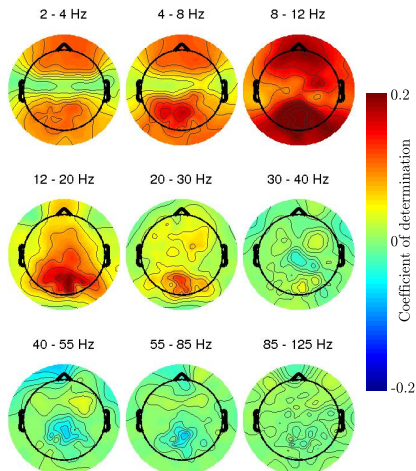


Confounding by EOG-artifacts

Eye-blinking

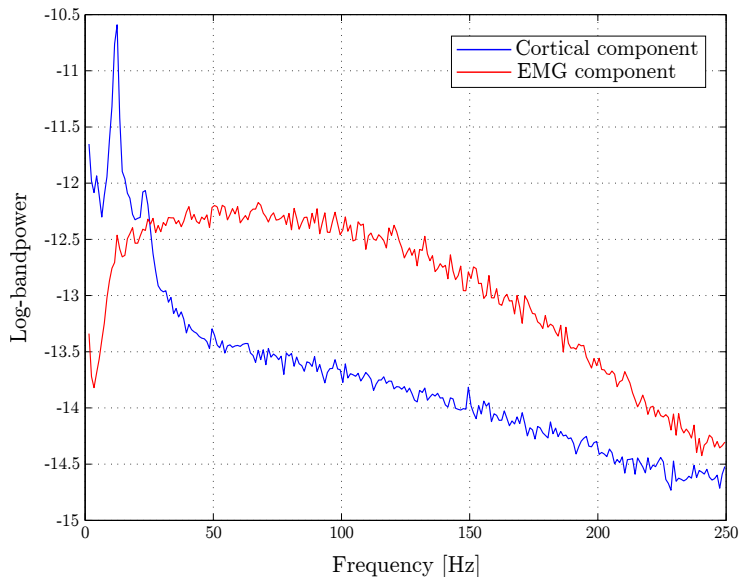


Horizontal eye-tracking



Confounding by EMG-artifacts

Confounding by EMG-artifacts



Attenuation of non-cortical artifacts by ICA

(Grosse-Wentrup et al. How to Test the Quality of Reconstructed Sources in ICA of EEG/MEG Data. *PRNI*, 2013)

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$

EEG mixing model:

- $\mathbf{x}[t] = A\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, A \in \mathbb{R}^{N \times N}$

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, \mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, \mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, \mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$

Independent Component Analysis (ICA):

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$

Independent Component Analysis (ICA):

- Find $\mathbf{W} \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = \mathbf{W}\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

- $\mathbf{x}[t] = A\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, A \in \mathbb{R}^{N \times N}$
- A : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$

Independent Component Analysis (ICA):

- Find $W \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = W\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$
- Then $\hat{\mathbf{s}}[t] = \mathbf{s}[t]$ and $W^{-1} = A$ up to permutation and scaling.

EEG mixing model:

- $\mathbf{x}[t] = A\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N, A \in \mathbb{R}^{N \times N}$
- A : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$

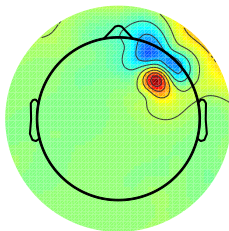
Independent Component Analysis (ICA):

- Find $W \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = W\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$
- Then $\hat{\mathbf{s}}[t] = \mathbf{s}[t]$ and $W^{-1} = A$ up to permutation and scaling.
- Identify non-cortical components

Attenuation of non-cortical artifacts by ICA

EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$



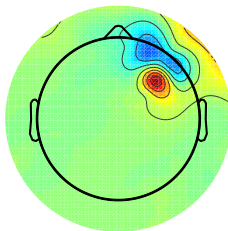
Independent Component Analysis (ICA):

- Find $\mathbf{W} \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = \mathbf{W}\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$
- Then $\hat{\mathbf{s}}[t] = \mathbf{s}[t]$ and $\mathbf{W}^{-1} = \mathbf{A}$ up to permutation and scaling.
- Identify non-cortical components

Attenuation of non-cortical artifacts by ICA

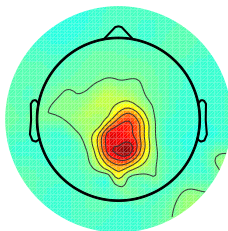
EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$



Independent Component Analysis (ICA):

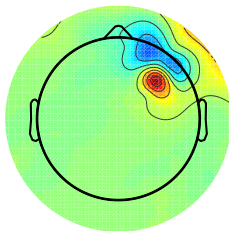
- Find $\mathbf{W} \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = \mathbf{W}\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$
- Then $\hat{\mathbf{s}}[t] = \mathbf{s}[t]$ and $\mathbf{W}^{-1} = \mathbf{A}$ up to permutation and scaling.
- Identify non-cortical components



Attenuation of non-cortical artifacts by ICA

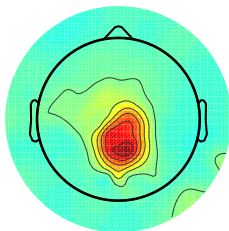
EEG mixing model:

- $\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t]$
- $\mathbf{x}[t], \mathbf{s}[t] \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times N}$
- \mathbf{A} : Mixing matrix, with the column \mathbf{a}_i the source topography of $s_i[t]$
- $\forall i, j : s_i \perp s_j$



Independent Component Analysis (ICA):

- Find $\mathbf{W} \in \mathbb{R}^{N \times N}$ such that $\hat{\mathbf{s}}[t] = \mathbf{W}\mathbf{x}[t]$ with $\forall i, j : \hat{s}_i \perp \hat{s}_j$
- Then $\hat{\mathbf{s}}[t] = \mathbf{s}[t]$ and $\mathbf{W}^{-1} = \mathbf{A}$ up to permutation and scaling.
- Identify non-cortical components
- Only reproject cortical components



- 1 M/EEG Decoding Models
- 2 Brain-Computer Interfacing

Studies with severely paralyzed patients in Tübingen

Patient	Diagnosis	Age ^a	Sex	Duration of participation/year of study entry	Level of impairment	Type of BCI and average CRR ^b			Level of success ^c	CRR published in
						SCP	SMR	P300		
HPS	ALS spinal	41	m	Present/1996	4	87	73		4 (SCP), 3 (P300) ^f	Kübler et al. (1999)
JB	ALS bulbar	49	m	2 years/1997	4	86			4 (SCP)	Birbaumer et al. (1999)
MP	ALS spinal	37	m	2 years/1997	3	66			3 (SCP)	Kübler et al. (1999)
MW	Brain stem stroke	26	f	Months/1995	4	X ^d			2 (SCP)	Kuebler et al. (1998)
HE	ALS spinal	42	m	Present/1998	3	94			4 (SCP)	Neumann and Birbaumer (2003)
EK	ALS spinal	66	m	Months/1998	2	57			2 (SCP)	Neumann and Birbaumer (2003)
MZ	ALS spinal	31	m	Months/2000	4	70			3 (SCP)	Kübler et al. (2001)
LB	ALS	63	m	Months/1999	5	48			1 (SCP)	Neumann and Birbaumer (2003)
NB	ALS	40	m	Months/2000	5	59			2 (SCP)	Neumann and Birbaumer (2003)
KI	Cerebral paresis	33	m	Months/1998	3	50			1 (SCP)	Kübler (2000)
TK	Muscular dystrophy	33	m	Months/1998/2006	4	43		59	1 (SCP), 2 (P300) ^f	Kübler (2000)
RCS	ALS spinal	56	m	1 year/2003	4		77	69	3 (SMR), 2 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
HAC	ALS bulbar	67	m	2 years/2003	2		78	86	3 (SMR), 4 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
UBA	ALS spinal	47	f	Present/2004	3		81	82	3 (SMR), 4 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
HM	ALS spinal	53	m	3 years/2002	2	67	76	32	2 (SCP), 3 (SMR), 2 (P300)	Kübler et al. (2004), Kübler et al. (2005) and Nijboer et al. (2008)
JAK	ALS spinal	49	m	2 years/2005	3			50	2 (P300)	Nijboer et al. (2008)
LEK	ALS spinal	49	f	Present/2004	3			80	3 (P300)	Nijboer et al. (2008)
IR	ALS spinal	42	f	Present/2003	5	54	43		1 (SCP), 1 (SMR) ^f	Kübler et al. (2004)
SM	ALS spinal	35	m	Months/2002	2	84			3 (SCP)	Kübler et al. (2004)
KW	ALS spinal	47	f	Months/2002	2	78			3 (SCP)	Kübler et al. (2004)
GW	ALS bulbar	59	f	Months/2002	2	70			3 (SCP)	Kübler et al. (2004)
GB	ALS spinal	62	f	Months/2002	2	79			3 (SCP)	Kübler et al. (2004)
KR	ALS spinal	35	f	Present/2002	3	62		87	2 (SCP), 4 (P300)	Kübler et al. (2004) and Nijboer et al. (2008)
HJZ	ALS	60	m	Months/2002	3	74			3 (SCP)	Kübler et al. (2004)
RB	ALS spinal	64	f	Months/2002	1	68			2 (SCP)	Kübler et al. (2004)
JF	ALS spinal	50	m	Months/2002	1	61			2 (SCP)	Kübler et al. (2004)
GR	ALS spinal	37	m	Present/2005	4			74	3 (P300) ^f	
PR	Heart attack	55	m	Years ^g /2002	5	50	X ^e	X ^g	1 (SCP) ^f , 1 (SMR), 1 (P300) ^f	Hill et al. (2006)
AG	Chronic GBS ^l	42	f	Years ^g /2000	5	50	X ^e	X ^g	1 (SCP) ^f , 1 (SMR), 1 (P300) ^f	Hill et al. (2006)
WER	ALS spinal	63	m	Days/2005	5		X ^e	X ^g	1 (SMR), 1 (P300) ^f	Hill et al. (2006)
G	Stroke	61	m	One session/2005	4		X ^e		1 (SMR)	Hill et al. (2006)
WEW	ALS spinal	46	m	Months/2004	4			X ^g	1 (P300)	Nijboer et al. (2008)
EM	ALS spinal	58	m	Weeks/2002	5	62			2 (SCP)	Hinterberger et al. (2005)
VWI	ALS spinal	57	f	3 sessions/2007	4			63	2 (P300) ^f	
UB	ALS spinal		m	Months/1997	2	X ^h				

(adapted from Kübler et al., 2008)

Studies with severely paralyzed patients in Tübingen

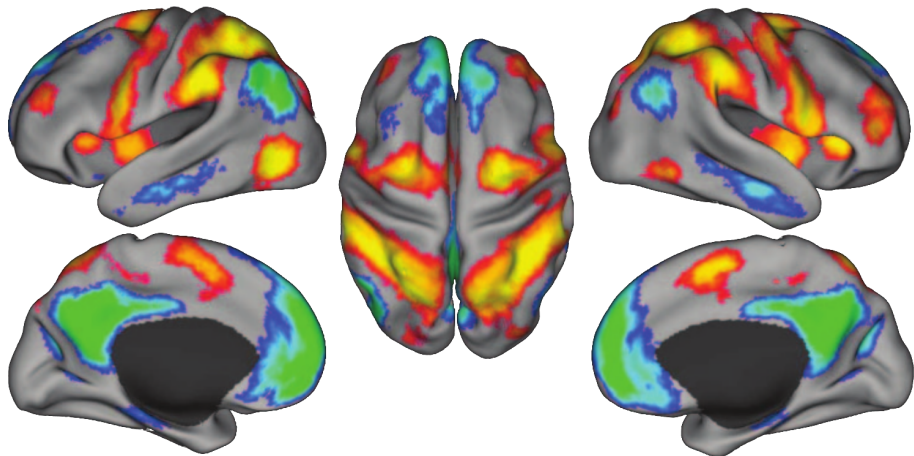
Patient	Diagnosis	Age ^a	Sex	Duration of participation/year of study entry	Level of impairment	Type of BCI and average CRR ^b			Level of success ^c	CRR published in
						SCP	SMR	P300		
HPS	ALS spinal	41	m	Present/1996	4	87	73		4 (SCP), 3 (P300) ^f	Kübler et al. (1999)
JB	ALS bulbar	49	m	2 years/1997	4	86			4 (SCP)	Birbaumer et al. (1999)
MP	ALS spinal	37	m	2 years/1997	3	66			3 (SCP)	Kübler et al. (1999)
MW	Brain stem stroke	26	f	Months/1995	4	X ^d			2 (SCP)	Kuebler et al. (1998)
HE	ALS spinal	42	m	Present/1998	3	94			4 (SCP)	Neumann and Birbaumer (2003)
EK	ALS spinal	66	m	Months/1998	2	57			2 (SCP)	Neumann and Birbaumer (2003)
MZ	ALS spinal	31	m	Months/2000	4	70			3 (SCP)	Kübler et al. (2001a)
LB	ALS	63	m	Months/1999	5	48			1 (SCP)	Neumann and Birbaumer (2003)
NB	ALS	40	m	Months/2000	5	59			2 (SCP)	Neumann and Birbaumer (2003)
KI	Cerebral paresis	33	m	Months/1998	3	50			1 (SCP)	Kübler (2000)
TK	Muscular dystrophy	33	m	Months/1998/2006	4	43			1 (SCP), 2 (P300) ^f	Kübler (2000)
RCS	ALS spinal	56	m	1 year/2003	4		77	69	3 (SMR), 2 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
HAC	ALS bulbar	67	m	2 years/2003	2		78	86	3 (SMR), 4 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
UBA	ALS spinal	47	f	Present/2004	3		81	82	3 (SMR), 4 (P300)	Kübler et al. (2005) and Nijboer et al. (2008)
HM	ALS spinal	53	m	3 years/2002	2	67	76	32	2 (SCP), 3 (SMR), 2 (P300)	Kübler et al. (2004), Kübler et al. (2005) and Nijboer et al. (2008)
JAK	ALS spinal	39	m	2 years/2005	3			50	2 (P300)	Nijboer et al. (2008)
LEK	ALS spinal	49	f	Present/2004	3			80	3 (P300)	Nijboer et al. (2008)
IR	ALS spinal	42	f	Present/2003	5	54	43		1 (SCP), 1 (SMR) ^f	Kübler et al. (2004)
SM	ALS spinal	35	m	Months/2002	2	84			3 (SCP)	Kübler et al. (2004)
KW	ALS spinal	47	f	Months/2002	2	78			3 (SCP)	Kübler et al. (2004)
GW	ALS bulbar	59	f	Months/2002	2	70			3 (SCP)	Kübler et al. (2004)
GB	ALS spinal	62	f	Months/2002	2	79			3 (SCP)	Kübler et al. (2004)
KR	ALS spinal	35	f	Present/2002	3	62		87	2 (SCP), 4 (P300)	Kübler et al. (2004) and Nijboer et al. (2008)
HJZ	ALS	60	m	Months/2002	3	74			3 (SCP)	Kübler et al. (2004)
RB	ALS spinal	64	f	Months/2002	1	68			2 (SCP)	Kübler et al. (2004)
JF	ALS spinal	50	m	Months/2002	1	61			2 (SCP)	Kübler et al. (2004)
GR	ALS spinal	37	m	Present/2005	4			74	3 (P300) ^f	
PR	Heart attack	55	m	Years/2002	5	50	X ^e	X ^s	1 (SCP) ^f , 1 (SMR), 1 (P300) ^f	Hill et al. (2006)
AG	Chronic GBS ^l	42	f	Years/2000	5	50	X ^e	X ^s	1 (SCP) ^f , 1 (SMR), 1 (P300) ^f	Hill et al. (2006)
WER	ALS spinal	63	m	Days/2005	5		X ^e	X ^s	1 (SMR), 1 (P300) ^f	Hill et al. (2006)
G	Stroke	61	m	One session/2005	4		X ^e		1 (SMR)	Hill et al. (2006)
WEW	ALS spinal	46	m	Months/2004	4			X ^s	1 (P300)	Nijboer et al. (2008)
EM	ALS spinal	58	m	Weeks/2002	5	62			2 (SCP)	Hinterberger et al. (2005)
VWI	ALS spinal	57	f	3 sessions/2007	4			63	2 (P300) ^f	
UB	ALS spinal		m	Months/1997	2	X ^h				

(adapted from Kübler et al., 2008)

Studies with severely paralyzed patients outside Tübingen

Diagnosis	Level of impairment	BCI type	Performance (CRR)	Level of success	Publication
SCI ^a (T7)	2	SMR	96/89	3	McFarland et al. (2005) and Wolpaw and McFarland (2004)
SCI (C6)	2	SMR	58/92/>60-<90 ^b	3	McFarland et al. (2005), Wolpaw and McFarland (2004) and McFarland et al. (2003)
Cerebral palsy	2	SMR	>60-<90 ^b		McFarland et al. (2003)
ALS	1	SMR	70-80 ^b >70 ^b	3	Wolpaw et al. (1997) and Miner et al. (1998)
ALS	2	P300	80 ^c	3	Sellers and Donchin (2006)
ALS	2	P300	73 ^c	3	Sellers and Donchin (2006)
ALS	3	P300	62 ^c	2	Sellers and Donchin (2006)
SCI (C5)	2	SMR	73	3	Müller-Putz et al. (2005)
SCI (TH8)	2	SMR	95 ^c	3	Krausz et al. (2003)
SCI (L1, incomplete)	2	SMR	72 ^c	3	Krausz et al. (2003)
SCI (L1)	2	SMR	80 ^c	3	Krausz et al. (2003)
SCI (TH12 (incomplete), L1, L4 (complete))	2	SMR	80 ^c	3	Krausz et al. (2003)
Cerebral palsy	3	SMR	70	3	Neuper et al. (2003)
SCI (C5)	2	SMR	≤100/grasp function ^d	3	Pfurtscheller et al. (2000), Pfurtscheller et al. (2003)
ALS	3	SMR	83	3	Müller-Putz et al. (2004)
ALS	3	P300	80 ^c	3	Piccione et al. (2006)
Brain stem stroke	4	P300	63 ^c	2	Piccione et al. (2006)
SCI (C4)	2	P300	76 ^c	3	Piccione et al. (2006)
GBS	2	P300	67 ^c	2	Piccione et al. (2006)
Multiple sclerosis	3	P300	58 ^c	2	Piccione et al. (2006)
SCI (C4 or C5 complete)	2	SMR	87 ^c	3	Kauhanen et al. (2006)
SCI (C4 or C5 complete)	2	SMR	88 ^c	3	Kauhanen et al. (2006)
SCI (C4 or C5 complete)	2	SMR	69 ^c	2	Kauhanen et al. (2006)
Cerebral palsy	3	P300	100 ^c	3	Hoffmann et al. (2008)
Multiple sclerosis	3	P300	100 ^c	3	Hoffmann et al. (2008)
ALS	3	P300	100 ^c	3	Hoffmann et al. (2008)
Traumatic brain and spinal cord injury, C4 level	3	P300	100 ^c	3	Hoffmann et al. (2008)
Post-anoxic encephalopathy	3	P300	Not reported	1	Hoffmann et al. (2008) (adapted from Kübler et al., 2008)

Large-scale brain networks

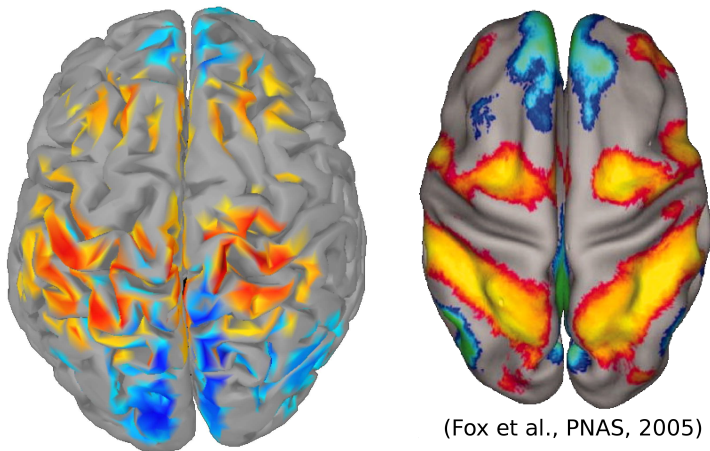


(Adapted from Fox et al., 2005)

Can large-scale cortical networks be observed in the EEG?

(Grosse-Wentrup & Schölkopf, High Gamma-Power Predicts Performance in SMR BCIs, *Journal of Neural Engineering*, 2012)

Can large-scale cortical networks be observed in the EEG?



(Grosse-Wentrup & Schölkopf, High Gamma-Power Predicts Performance in SMR BCIs, *Journal of Neural Engineering*, 2012)

Neurofeedback of γ -power in superior parietal cortex (SPC)

Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

- 19 healthy subjects

Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

- 19 healthy subjects
- 121-channel EEG @ 500 Hz

Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline

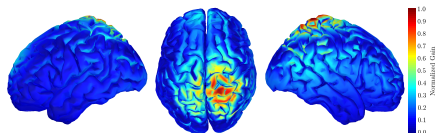
Experimental setup:

- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex

Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

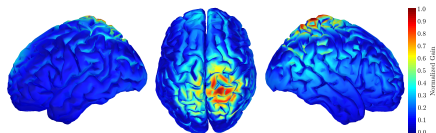
- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex



Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

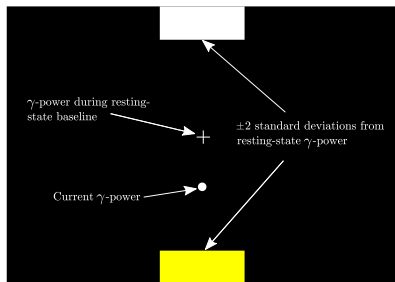
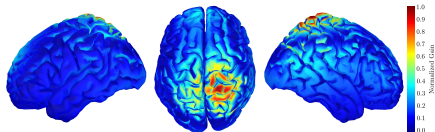
- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex
- 3x20 min feedback sessions of parietal γ -power (55-85 Hz)



Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

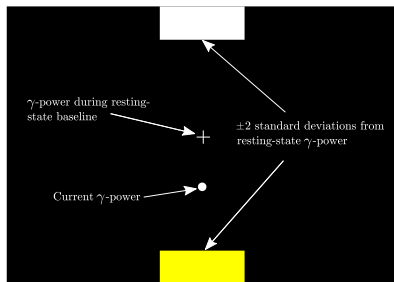
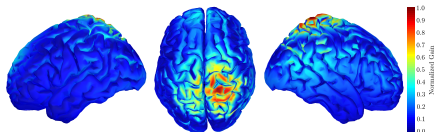
- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex
- 3x20 min feedback sessions of parietal γ -power (55-85 Hz)



Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

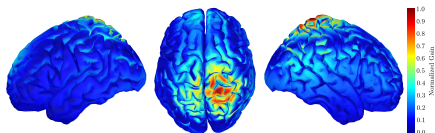
- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex
- 3x20 min feedback sessions of parietal γ -power (55-85 Hz)
- 8 subjects excluded due to EMG contamination



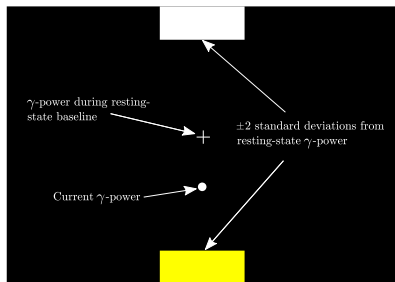
Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex
- 3x20 min feedback sessions of parietal γ -power (55-85 Hz)
- 8 subjects excluded due to EMG contamination



Average decoding accuracy:



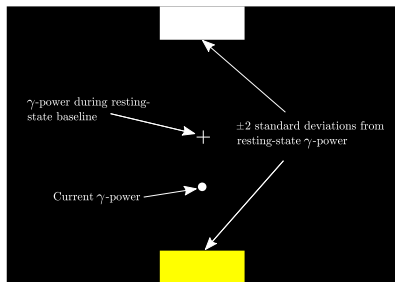
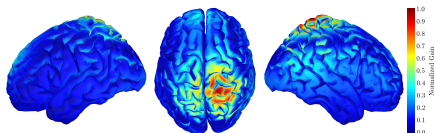
Neurofeedback of γ -power in superior parietal cortex (SPC)

Experimental setup:

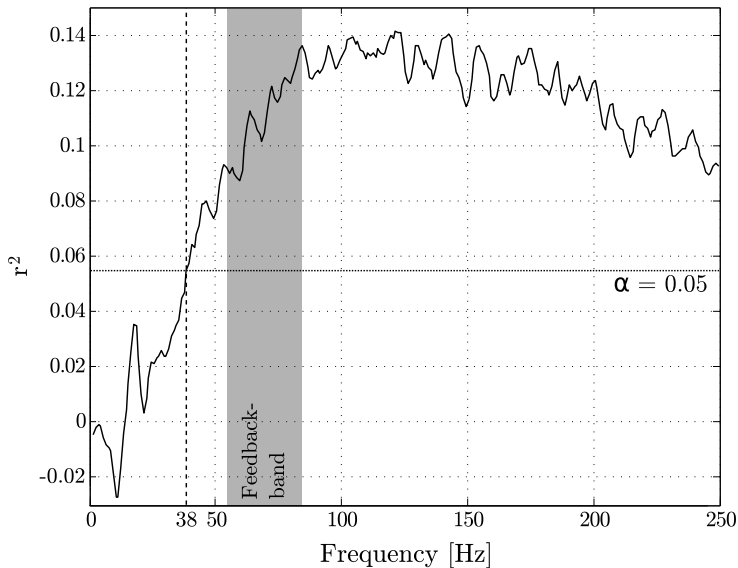
- 19 healthy subjects
- 121-channel EEG @ 500 Hz
- 5 min resting-state baseline
- Online beamforming targeting parietal cortex
- 3x20 min feedback sessions of parietal γ -power (55-85 Hz)
- 8 subjects excluded due to EMG contamination

Average decoding accuracy:

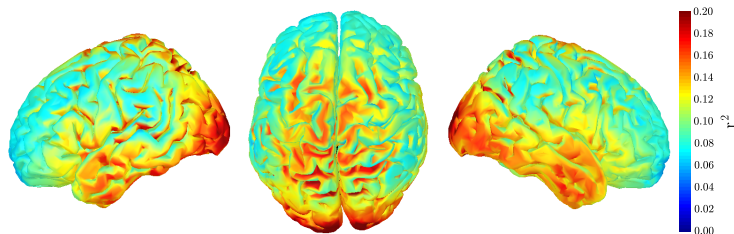
- 70.3% ($p = 0.002$)



Spectral specificity of bandpower-regulation in SPC

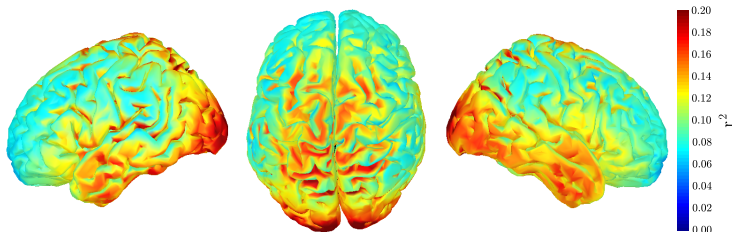


Spatial specificity of bandpower-regulation in the γ -range

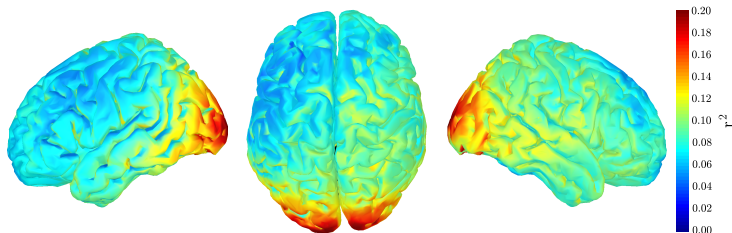


r^2 -map of cortical sources in the γ -range

Spatial specificity of bandpower-regulation in the γ -range

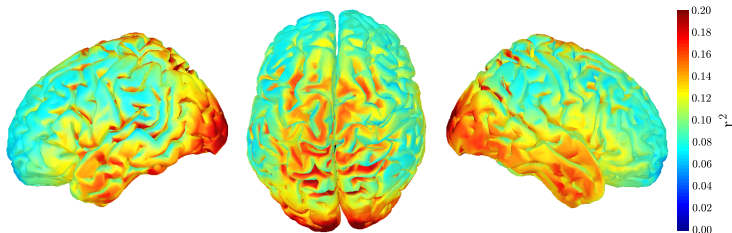


r^2 -map of cortical sources in the γ -range

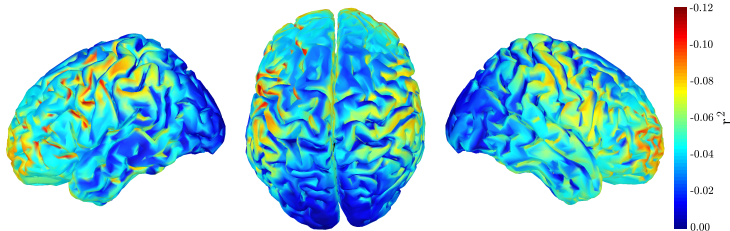


r^2 -map of electromyogenic sources in the γ -range

Self-regulation of γ -power in SPC modulates μ -rhythms



r^2 -map of cortical sources in the γ -range



r^2 -map of cortical sources in the μ -range

Enhancement of γ -power in SPC was achieved by

Enhancement of γ -power in SPC was achieved by

- *just doing it*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Attenuation of γ -power in SPC was achieved by

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Attenuation of γ -power in SPC was achieved by

- *just doing it*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Attenuation of γ -power in SPC was achieved by

- *just doing it*
- *not thinking*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Attenuation of γ -power in SPC was achieved by

- *just doing it*
- *not thinking*
- *feeling my legs to be heavy*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

Attenuation of γ -power in SPC was achieved by

- *just doing it*
- *not thinking*
- *feeling my legs to be heavy*
- *relaxing*

Enhancement of γ -power in SPC was achieved by

- *just doing it*
- *solving math problems*
- *doing free-style skating in my mind*
- ... and other tasks requiring focused attention

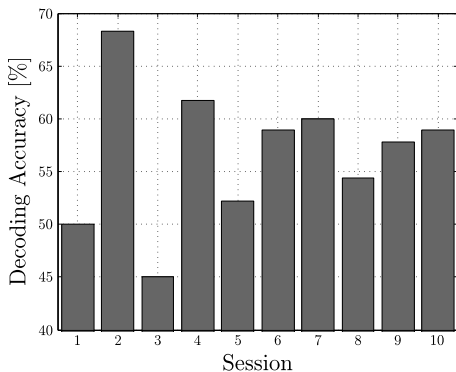
Attenuation of γ -power in SPC was achieved by

- *just doing it*
- *not thinking*
- *feeling my legs to be heavy*
- *relaxing*
- ... and other states-of-mind related to relaxed wakefulness

Patient LEK: Decoding results



Patient LEK: Decoding results



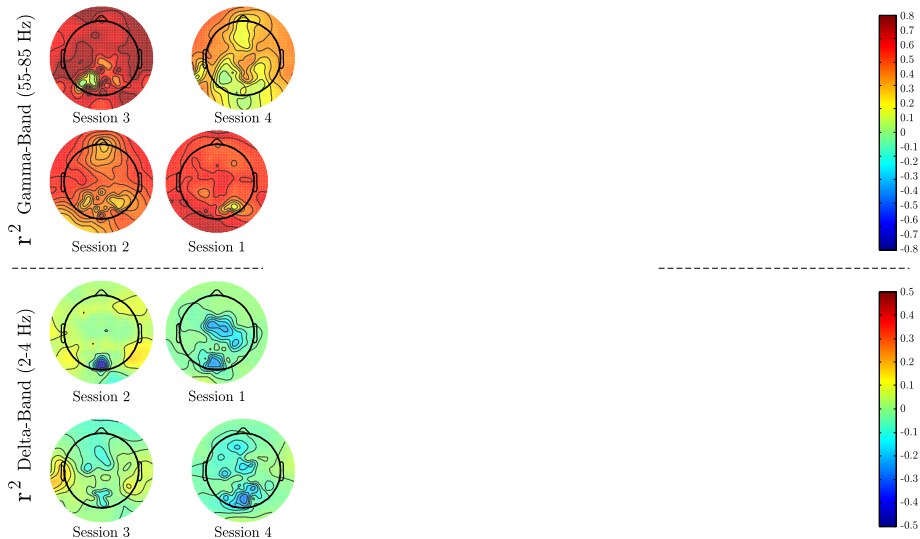
Average: 56.7% ($p = 0.02$, $N = 780$)

Patient GH: Decoding results

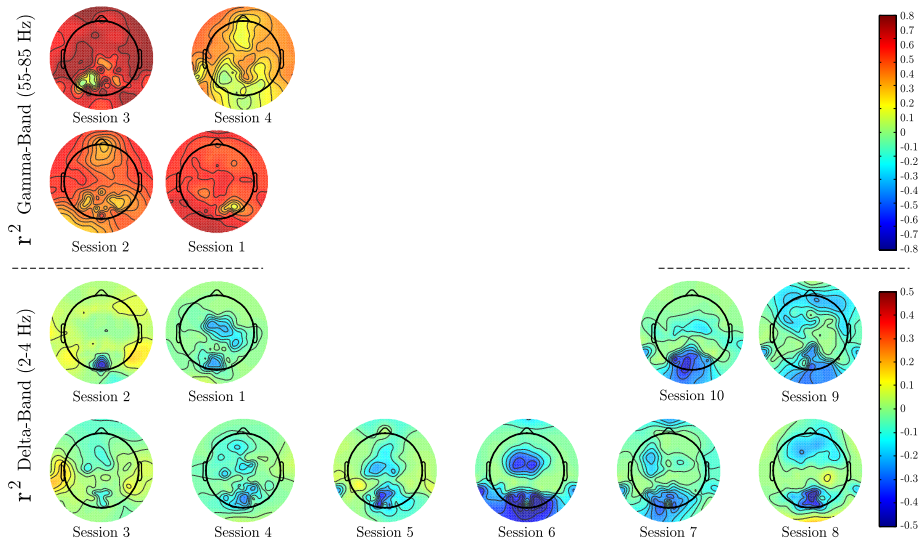
Patient GH: Decoding results



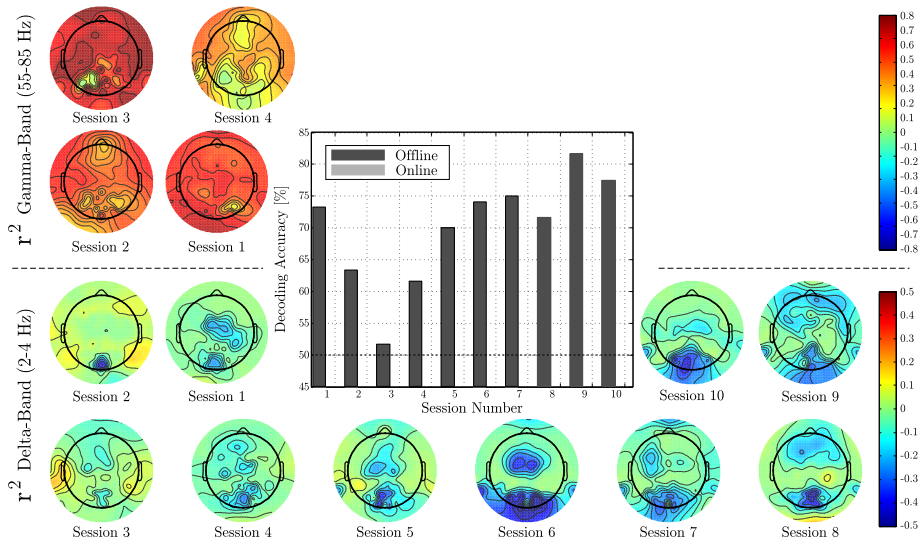
Patient GH: Decoding results



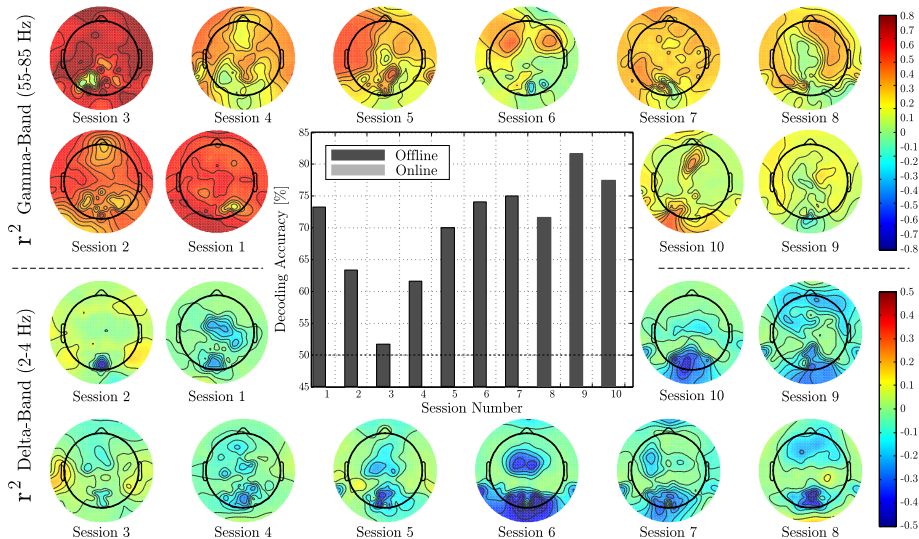
Patient GH: Decoding results



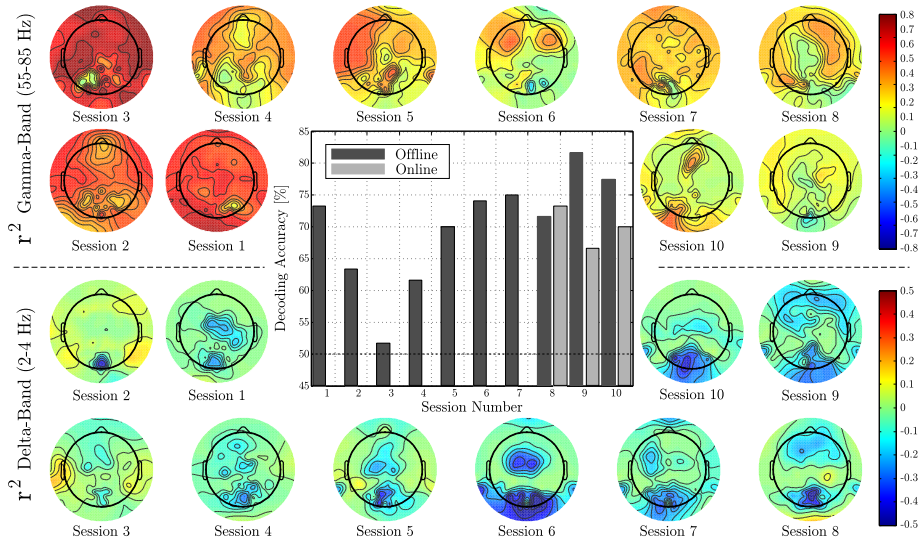
Patient GH: Decoding results



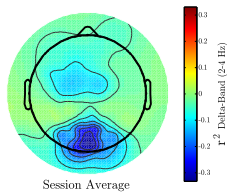
Patient GH: Decoding results



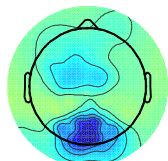
Patient GH: Decoding results



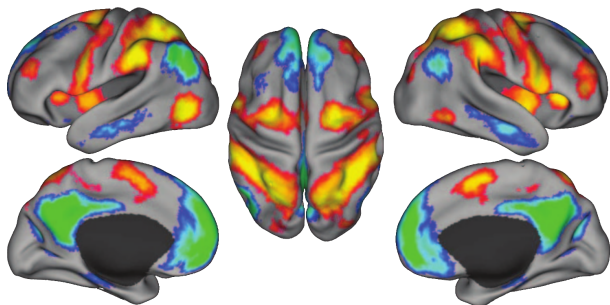
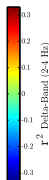
Patient GH: fMRI-study



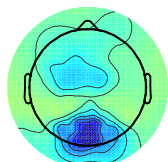
Patient GH: fMRI-study



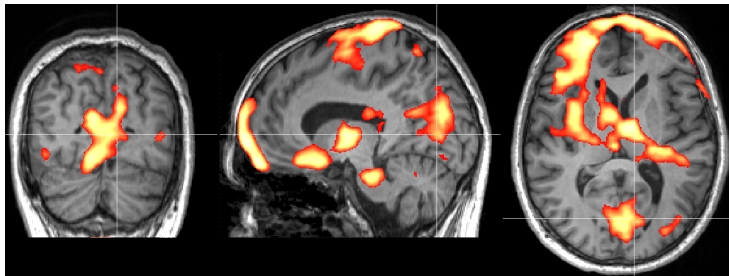
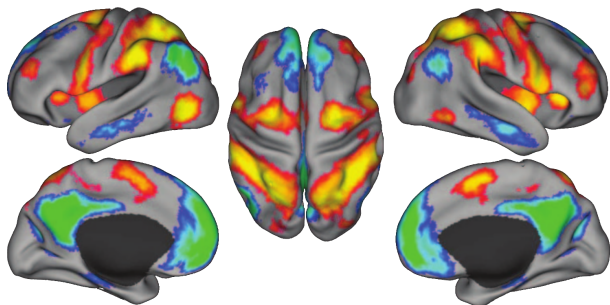
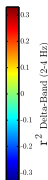
Session Average



Patient GH: fMRI-study



Session Average



Max Planck Institutes:

- Bernd Battaes
- Alexander Bretin
- Tatiana Fomina
- Christian Förster
- Marius Klug
- Gabriele Lohmann
- Natalie Widmann
- Bernhard Schölkopf

- Nadine Simon

- Sebastian Weichwald

University of Tübingen:

- Michael Erb

- Thomas Ethofer

<http://brain-computer-interfaces.net>