

Classification de signaux EEG et filtrage spatial par CSP sparse

Marie Chavent, Frédérique Faïta-Aïnseba, Pierrick Legrand,
Laurent Vézard

Université de Bordeaux,
IMB, UMR 5251,
Inria Bordeaux Sud-Ouest, équipe CQFD

Plan

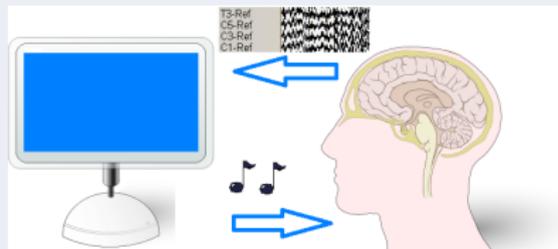
- 1 Acquisition de données EEG
 - Protocole d'acquisition
 - Données
- 2 Extraction de caractéristiques
 - Critère des pentes
 - Filtrage spatial par CSP et transformation logvar
- 3 Classification de signaux EEG
 - La méthodologie générale
 - Adaptation au problème multi-sujet
- 4 Sélection de variables
 - Filtrage par CSP sparse
 - Algorithme de filtrage par CSP sparse
 - Résultats

Plan

- 1 Acquisition de données EEG
 - Protocole d'acquisition
 - Données
- 2 Extraction de caractéristiques
 - Critère des pentes
 - Filtrage spatial par CSP et transformation logvar
- 3 Classification de signaux EEG
 - La méthodologie générale
 - Adaptation au problème multi-sujet
- 4 Sélection de variables
 - Filtrage par CSP sparse
 - Algorithme de filtrage par CSP sparse
 - Résultats

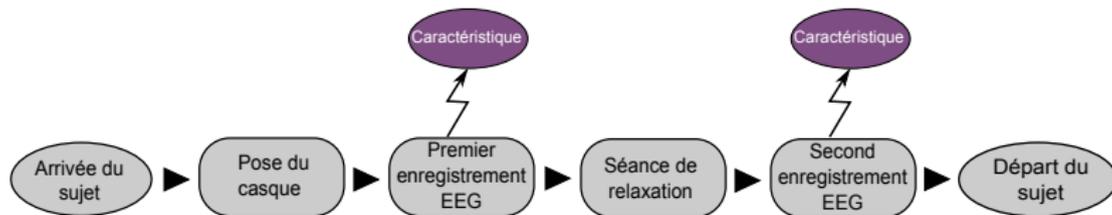
Objectifs

- **Caractériser l'état de vigilance d'une personne à partir de son électro-encéphalogramme (EEG).**
- Créer une interface cerveau-ordinateur.
- Influencer l'état de vigilance d'une personne à l'aide de musiques synthétisées.



Protocole d'acquisition

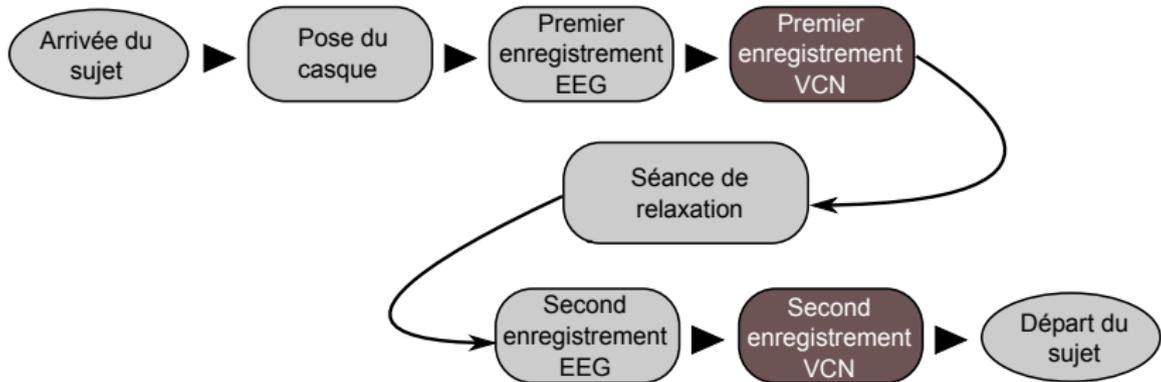
- Premier enregistrement EEG : sujet dans un état de vigilance haut dit "normal"
- Second enregistrement EEG : sujet dans un état de vigilance bas dit "relaxé"



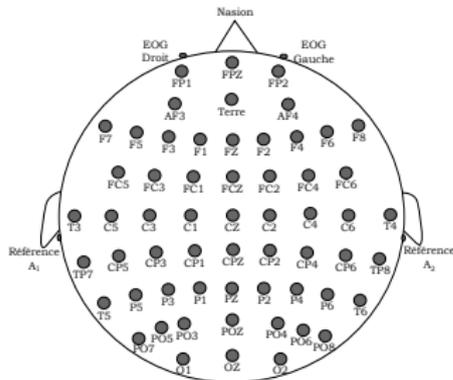
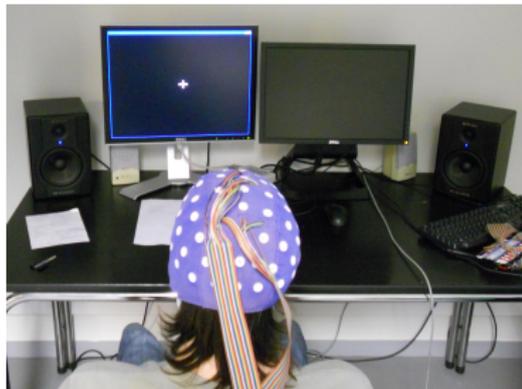
⇒ La séance de relaxation est elle efficace ?

Protocole d'acquisition

- Ajout de deux étapes pour effectuer un test VCN (Variation Contingence négative)



Protocole d'acquisition



- Durée de pose du casque EEG : **45 minutes**.
- Sujet immobile, avec les **yeux ouverts**.
- Fréquence d'échantillonnage de 256Hz.
- Durée d'enregistrement : 3 minutes (**46000 points de mesure**).

Protocole d'acquisition

- **58 électrodes** renumérotées de 1 à 58
- Sujets de moins de 35 ans, droitiers et non fumeurs
- **Deux campagnes** d'acquisition
 - Campagne 1 : 44 sujet \Rightarrow **13 conservés** selon la VCN
 - Campagne 2 : 14 sujets \Rightarrow **6 conservés** selon la VCN
- **Séance de relaxation** de 20 minutes par une voix enregistrée proposant 3 exercices :
 - Training autogène [Schultz, 1958] : répétition de phrases, auto-hypnose.
 - Relaxation musculaire progressive [Jacobson, 1974] : relaxation musculaire progressive.
 - Visualisation mentale (lieu familiés, odeurs, bruits).

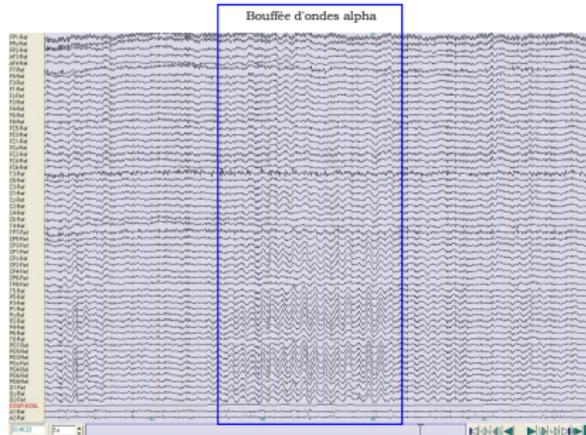
Protocole d'acquisition

3 minutes d'enregistrement EEG avant relaxation.



Protocole d'acquisition

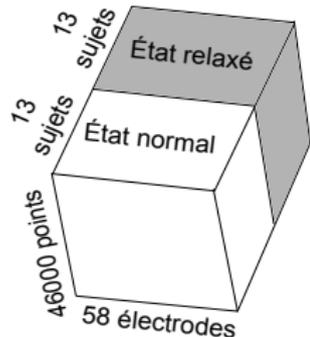
3 minutes d'enregistrement EEG après relaxation.



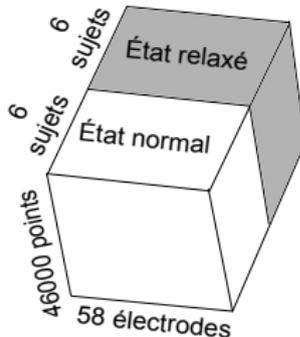
⇒ Oscillations caractéristiques d'un état d'éveil passif entre 8 et 12 Hz.

Les données

**Première campagne
d'acquisition**



**Seconde campagne
d'acquisition**



⇒ Trouver une **fonction de classification** qui permettrait de prédire l'état de vigilance d'un sujet en fonction de ses signaux EEG ?

Cadre général de l'apprentissage automatique

- Variables d'entrées X : variables explicatives
- Variable de sortie Y : variable à expliquer
 - Y quantitative \rightsquigarrow Problème de régression
 - Y qualitative \rightsquigarrow Problème de classification

Objectif

Définir une règle de classification c'est à une fonction $g : \mathcal{X} \rightarrow \mathcal{Y}$ à partir d'un ensemble de données d'apprentissage

Un échantillon d'apprentissage est un ensemble de n couples de variables aléatoires i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$ avec

- $X_i = (X_{i,1}, \dots, X_{i,p}) \in \mathcal{X} \subset \mathbb{R}^p$
- Y_i prend ses valeurs dans un ensemble fini $\mathcal{Y} = \{1, \dots, K\}$

Ici on a un problème de **discrimination en 2 classes** avec

- Y **binaire** à deux modalités **normal** et **relaxé**
- $p = 58$ variables **réelles** \Rightarrow extraire **une unique caractéristique pour chaque signal EEG**.

Deux approches envisagées pour **extraire les caractéristiques** :

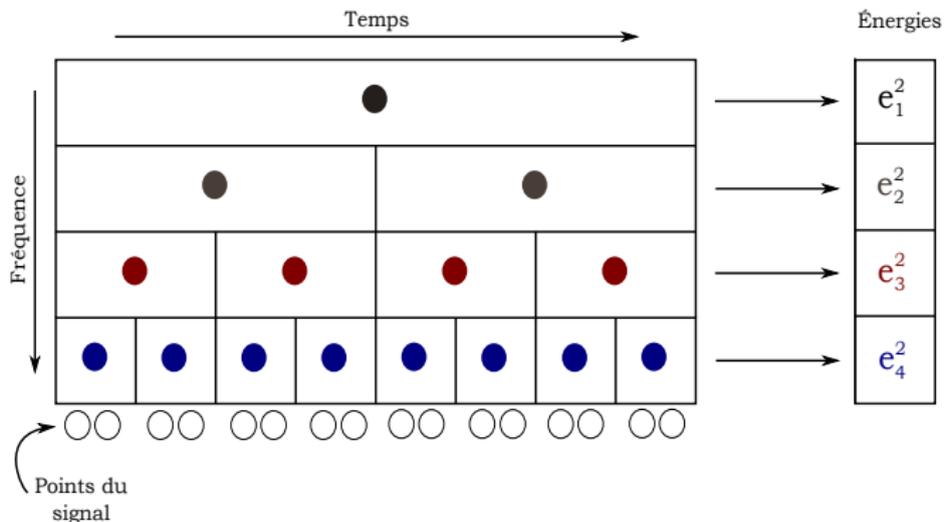
- Le critère des **pentés** : mesure la "pente" de l'énergie du signal dans les fréquences d'onde alpha.
- Filtrage spatial par **CSP** et **transformation logvar** :
 - Effectuer une combinaison linéaire entre les différentes électrodes en vue de maximiser la différence des variances du signal filtré entre les 2 classes .
 - Caractériser les signaux filtrés par le logarithme de leur variance

Plan

- 1 Acquisition de données EEG
 - Protocole d'acquisition
 - Données
- 2 Extraction de caractéristiques
 - Critère des pentes
 - Filtrage spatial par CSP et transformation logvar
- 3 Classification de signaux EEG
 - La méthodologie générale
 - Adaptation au problème multi-sujet
- 4 Sélection de variables
 - Filtrage par CSP sparse
 - Algorithme de filtrage par CSP sparse
 - Résultats

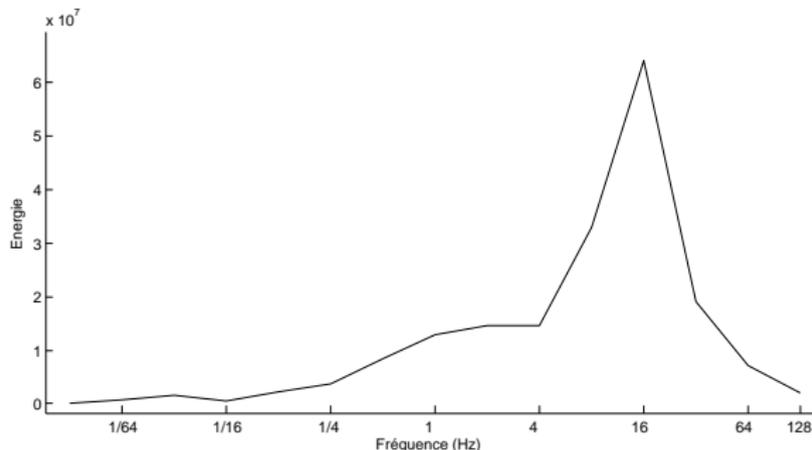
Critère des pentes

La **grille dyadique** donne une représentation spatio-fréquentielle de la décomposition en ondelettes discrète dyadique



Critère des pentes

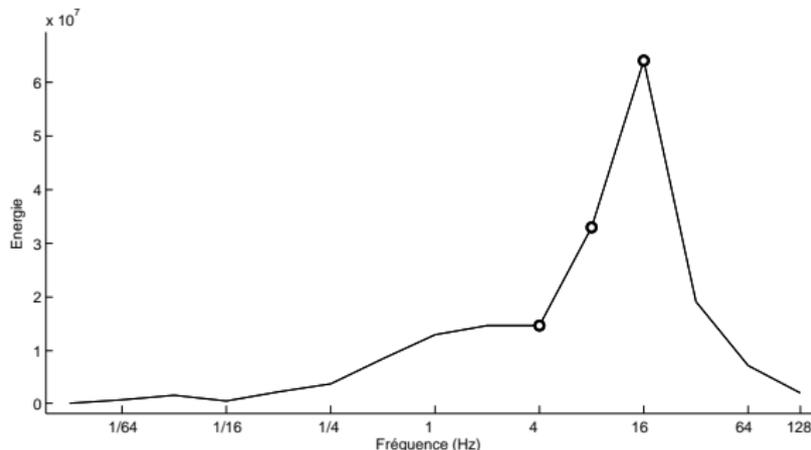
- Alpha : 8 – 12Hz.
- Ondes caractéristiques d'un état relaxé.



Régression linéaire entre 4 et 16Hz.

Critère des pentes

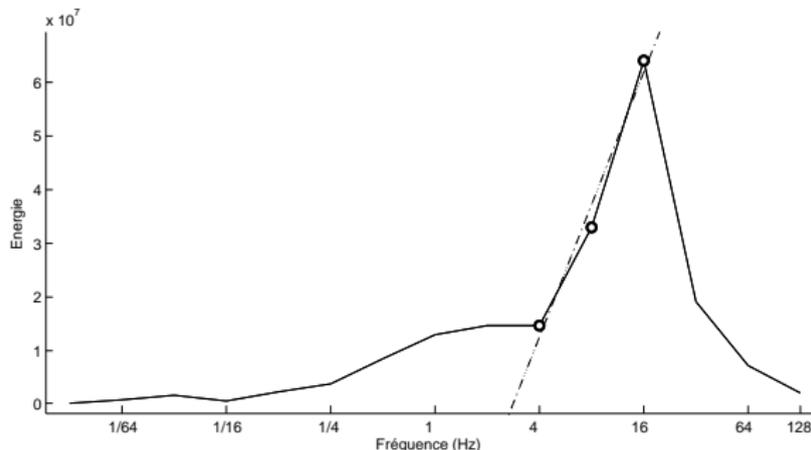
- Alpha : 8 – 12Hz.
- Ondes caractéristiques d'un état relaxé.



Régression linéaire entre 4 et 16Hz.

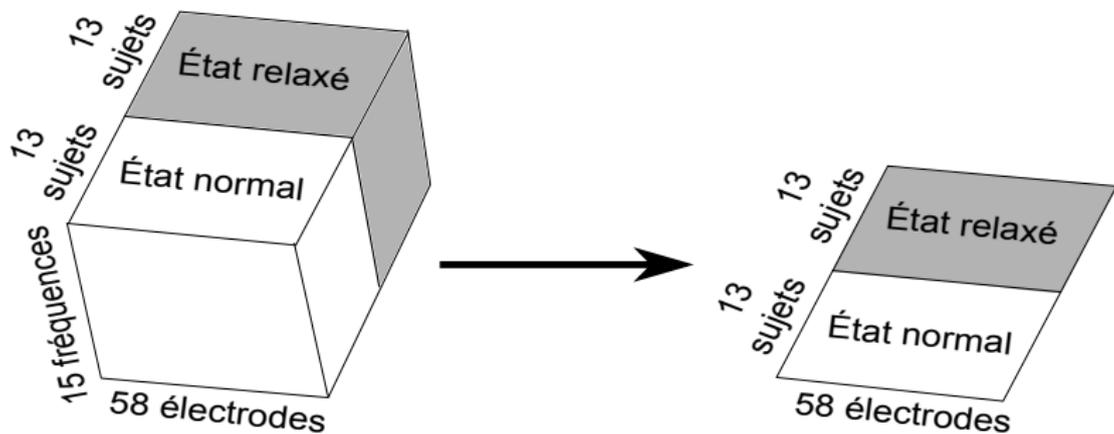
Critère des pentes

- Alpha : 8 – 12Hz.
- Ondes caractéristiques d'un état relaxé.

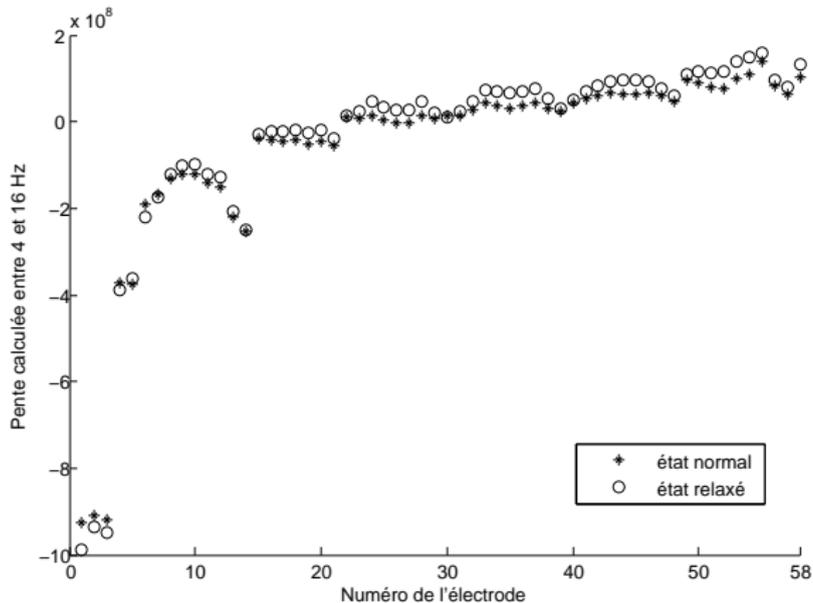


Régression linéaire entre 4 et 16Hz.

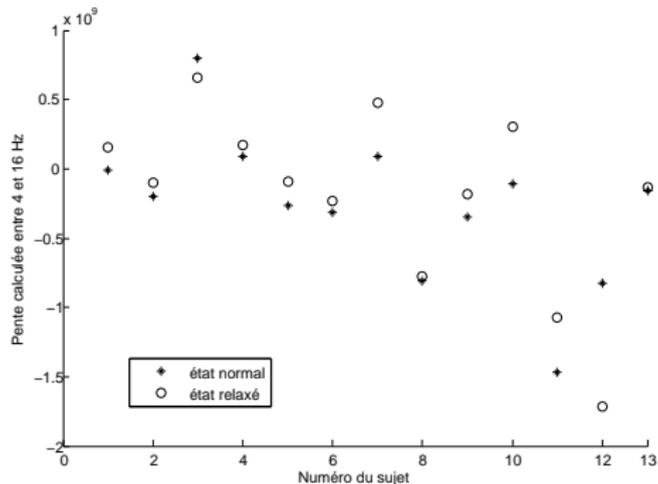
Critère des pentes



Critère des pentes



Critère des pentes



- ⇒ Très forte **variabilité inter-sujet**
- ⇒ Critère qui ne permet pas de construire un classifieur performant pour différents sujets.

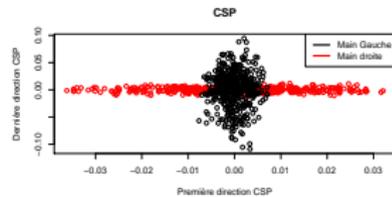
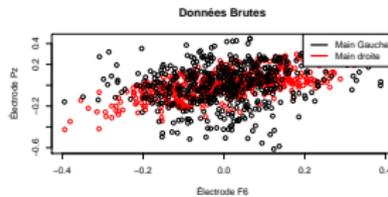
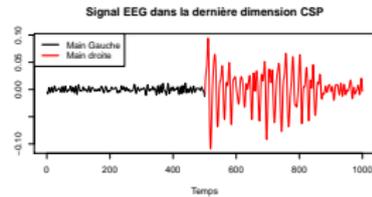
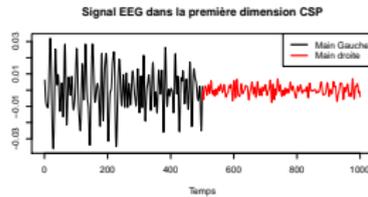
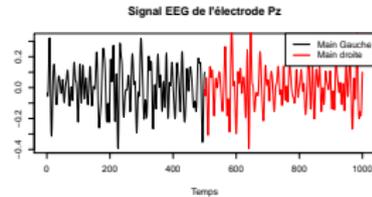
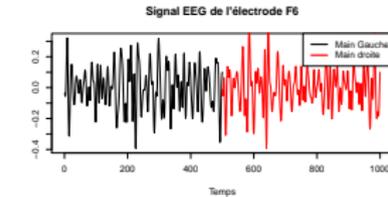
Filtrage spatial par CSP

- Il s'agit de trouver des **signaux synthétiques** (filtrés) de **variance maximale** dans une classe et simultanément de **variance minimale** dans l'autre classe.
- Les **signaux synthétiques** sont obtenus par combinaison linéaire des signaux initiaux.

$$\mathbf{f} = \mathbf{X}\mathbf{w}$$

où :

- \mathbf{X} est la matrice de données avec les signaux en colonne,
 - \mathbf{w} est le **filtre spatial**,
 - \mathbf{f} est le **signal filtré** (projeté).
- Illustration avec le jeu de données IIIa [Schlögl et al., 2005] de la compétition BCI III [Blankertz et al., 2006] représentant des signaux EEG enregistrés chez des sujets durant des **mouvements imaginés de la main droite et de la main gauche**.



- L'unité statistique est ici un **un essai** c'est à dire un enregistrement EEG sur p électrodes effectué durant une période de temps T .
- Les **données d'apprentissage** sont constituées de n essais dont n_1 appartiennent à la classe 1 et n_2 appartiennent à la classe 2.
- On note X_i la matrice de dimensions $T \times p$ des enregistrement EEG de l'essai i .
- On note \mathbf{X}_c la matrice de dimension $Tn_c \times p$ des enregistrement EEG des essais de la classe c .

Matrices de variances-covariances empiriques \mathbf{C}_1 et \mathbf{C}_2

$$\mathbf{C}_c = \frac{1}{Tn_c} \mathbf{X}_c^T \mathbf{X}_c = \frac{1}{n_c} \sum_{i \in \mathcal{I}_c} \frac{1}{T} X_i^T X_i = \frac{1}{n_c} \sum_{i \in \mathcal{I}_c} C_i$$

où $c \in \{1, 2\}$, \mathcal{I}_c est l'ensemble des indices correspondant aux essais dans la classe c .

- **Premier objectif** : trouver des signaux synthétiques de variance maximale dans la classe 1 et de variance minimale dans la classe 2.
- Trouver des vecteurs $\mathbf{w} \in \mathbb{R}^p$ qui **maximisent** $\text{var}(\mathbf{X}_1\mathbf{w})$ et **minimisent** $\text{var}(\mathbf{X}_2\mathbf{w})$ où

$$\text{Var}(\mathbf{X}_c\mathbf{w}) = \frac{1}{Tn_c} \mathbf{w}^T \mathbf{X}_c^T \mathbf{X}_c \mathbf{w} = \mathbf{w}^T \mathbf{C}_c \mathbf{w}$$

avec $c = 1, 2$

- **Premier problème d'optimisation**

$$\max_{\mathbf{w} \in \mathbb{R}^p} \frac{\mathbf{w}^T \mathbf{C}_1 \mathbf{w}}{\mathbf{w}^T \mathbf{C}_2 \mathbf{w}} \quad (1)$$

- **Solution**

- \mathbf{w} est le vecteur propre de $\mathbf{C}_2^{-1}\mathbf{C}_1$ associé à la plus grande valeur propre.
- Obtention d'un ensemble de k vecteurs $\mathbf{w}_1, \dots, \mathbf{w}_k$ maximisant (1) sous la contrainte $\mathbf{w}_j^T \mathbf{C}_2 \mathbf{w}_l = 0, \forall l \neq j$.

- **Second objectif** : trouver des signaux synthétiques de variance maximale dans la classe 2 et de variance minimale dans la classe 1.
- Trouver des vecteurs $\mathbf{v} \in \mathbb{R}^p$ qui **maximisent** $\text{var}(\mathbf{X}_2\mathbf{v})$ et **minimisent** $\text{var}(\mathbf{X}_1\mathbf{v})$.
- **Second problème d'optimisation**

$$\max_{\mathbf{v} \in \mathbb{R}^p} \frac{\mathbf{v}^T \mathbf{C}_2 \mathbf{v}}{\mathbf{v}^T \mathbf{C}_1 \mathbf{v}} \quad (2)$$

- **Solution**

- \mathbf{v} est le vecteur propre de $\mathbf{C}_1^{-1}\mathbf{C}_2$ associé à la plus grande valeur propre.
- Obtention d'un ensemble de k vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_k$ maximisant (2) sous la contrainte $\mathbf{v}_j^T \mathbf{C}_2 \mathbf{v}_l = 0, \forall l \neq j$.

On a vu que :

- Le premier ensemble de vecteurs $\mathbf{w}_1, \dots, \mathbf{w}_k$ est obtenu par la décomposition en éléments propres de $\mathbf{C}_2^{-1}\mathbf{C}_1$.
- Le second ensemble de vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_k$ est obtenu par la décomposition en éléments propres de $\mathbf{C}_1^{-1}\mathbf{C}_2$.

On peut montrer que

- Les vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_k$ sont aussi les k derniers vecteurs propres de $\mathbf{C}_2^{-1}\mathbf{C}_1$.
- La décomposition en éléments propres de $\mathbf{C}_2^{-1}\mathbf{C}_1$ donne directement les deux ensembles de vecteurs $\mathbf{w}_1, \dots, \mathbf{w}_k$ et $\mathbf{v}_1, \dots, \mathbf{v}_k$ en conservant les k premiers et les k derniers vecteurs propres.

Finalement, le filtrage par CSP fournit la matrice W de dimension $p \times 2k$ des k premiers et k derniers vecteurs propres de $\mathbf{C}_2^{-1}\mathbf{C}_1$:

$$W = [\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_k]$$

Transformation logvar

- 1 On utilise la matrice W pour construire les signaux EEG filtrés des essais X_i :

$$F_i = X_i W$$

F_i est une matrice de dimensions $T \times 2k$.

- 2 Pour chaque essai i on extrait **une unique caractéristique** par colonne j de F_i :

$$z_{ij} = \log(\text{var}(X_i \mathbf{w}_j)).$$

- ⇒ Obtention d'une **matrice Z** caractérisant les n essais en $2k$ dimensions.
- ⇒ Cette **matrice de caractéristiques** est utilisée pour faire de la **classification binaire des essais**.

Découpage de nos signaux EEG en essais

Les signaux EEG des sujets des campagnes 1 et 2 ont été **découpés en essais**

- Chaque signal EEG est découpé en n^* essais de longueur T soit $n^* = \left(\lfloor \frac{46000}{T} \rfloor \right)$.
- Changement d'unité statistique : **Sujet** \Rightarrow **Essais**

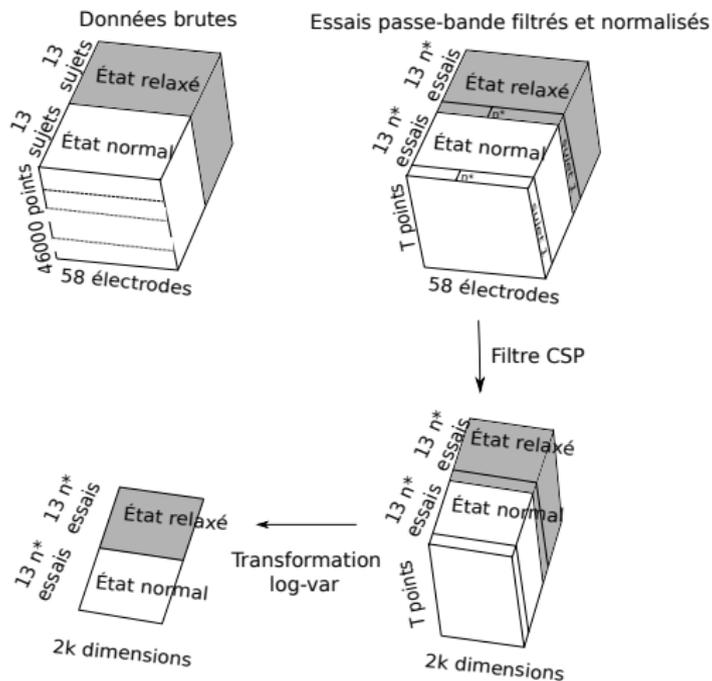
\Rightarrow Choix de T (longueur d'un essai)

\Rightarrow Choix de k (nombre de paires de signaux filtrés par CSP)

On a choisit :

- $T = 2048$ points soit $n^* = 22$ essais de 8 secondes.
- $k = 3$ paires de filtre CSP.

Extraction de caractéristiques sur nos données



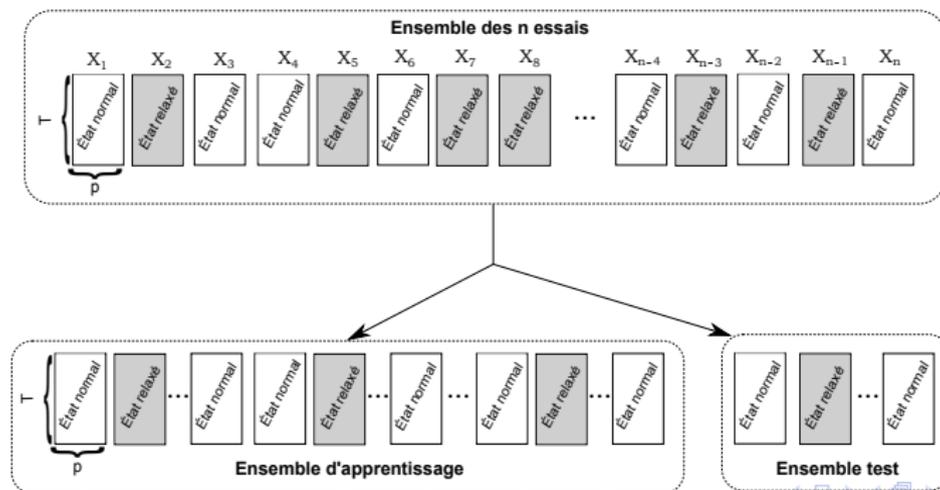
Plan

- 1 Acquisition de données EEG
 - Protocole d'acquisition
 - Données
- 2 Extraction de caractéristiques
 - Critère des pentes
 - Filtrage spatial par CSP et transformation logvar
- 3 Classification de signaux EEG
 - La méthodologie générale
 - Adaptation au problème multi-sujet
- 4 Sélection de variables
 - Filtrage par CSP sparse
 - Algorithme de filtrage par CSP sparse
 - Résultats

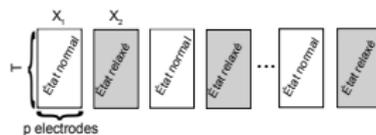
La méthodologie générale

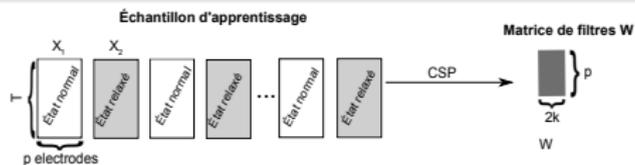
Découpage des essais en deux échantillons :

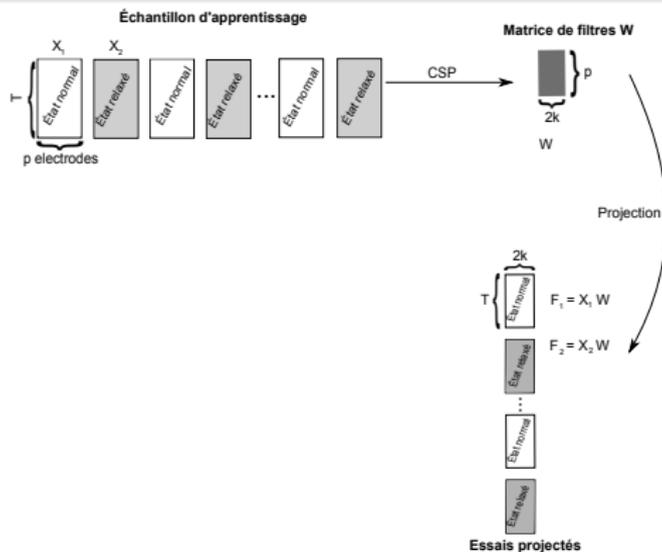
- Apprentissage de la **matrice W** des **filtres CSP** et de la **fonction de classification binaire** (LDA, QDA, régression logistique, forêt aléatoire...) sur l'**échantillon d'apprentissage**.
- Evaluation de la performance du classifieur en calculant un **taux de bon classement** sur l'**échantillon test**

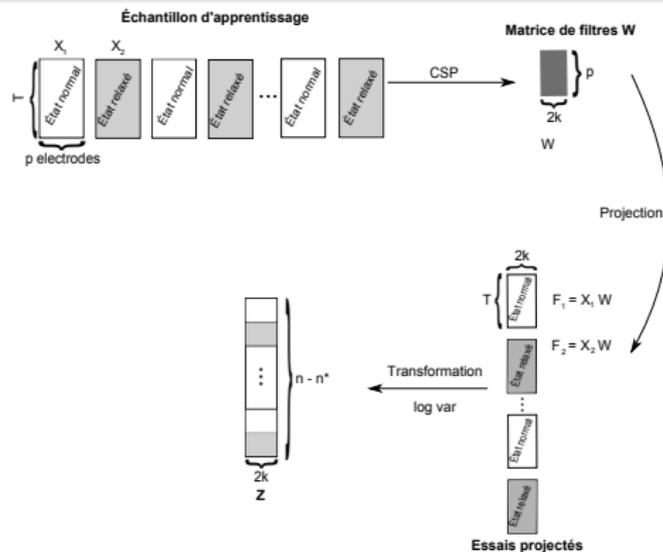


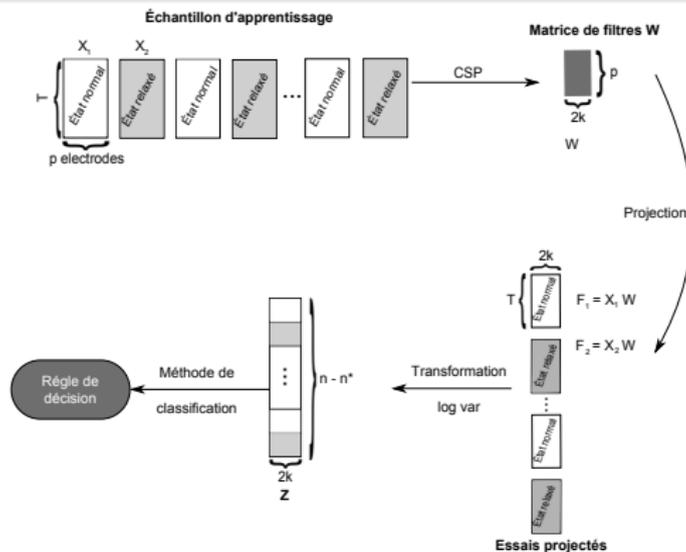
Échantillon d'apprentissage

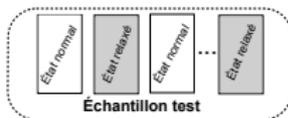
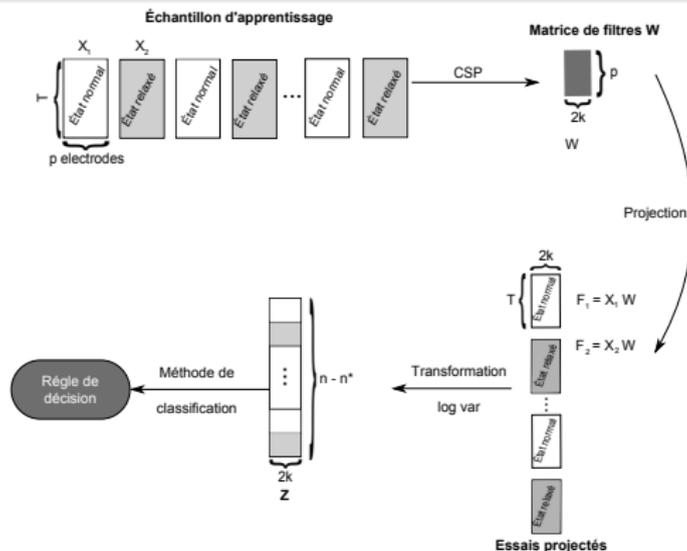


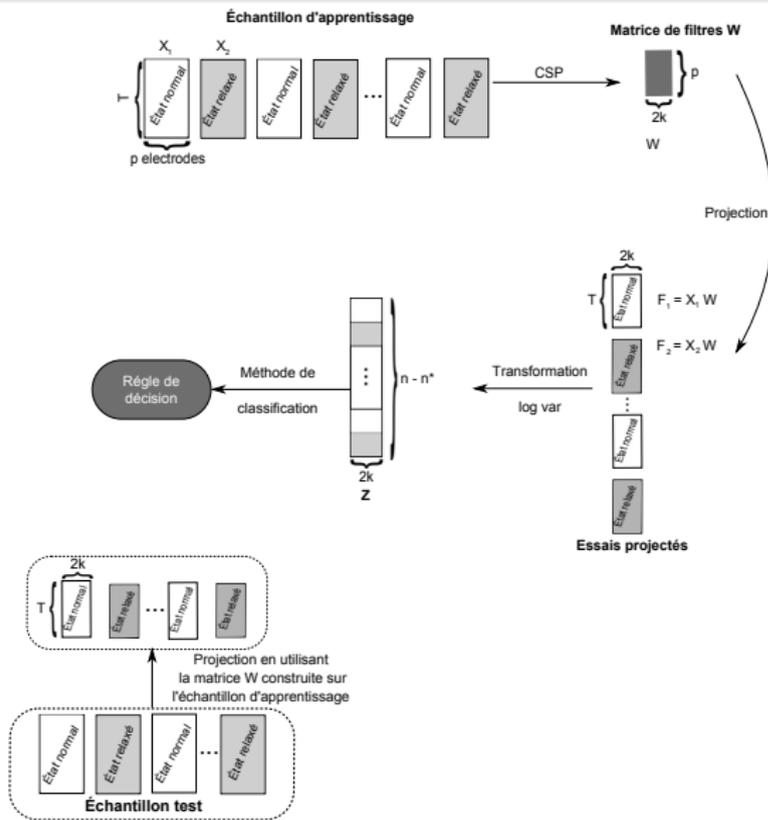


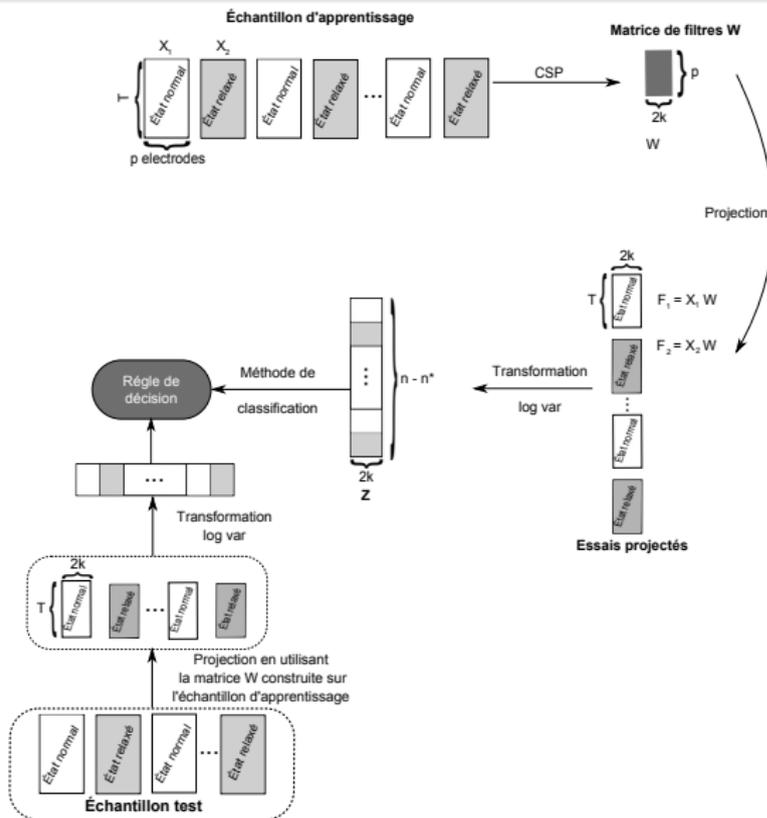


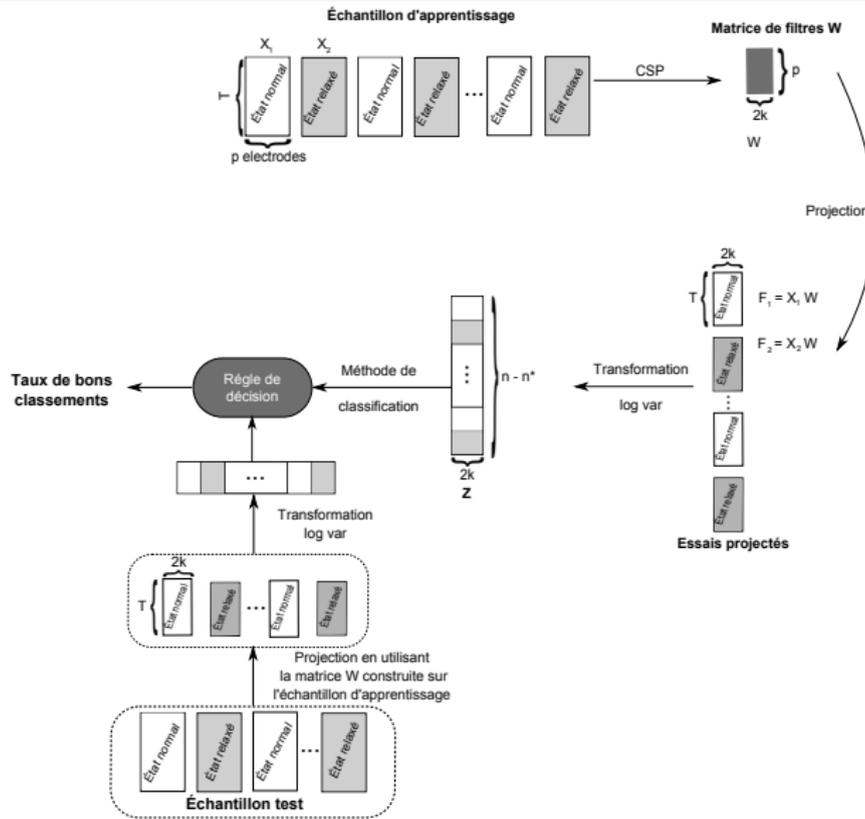








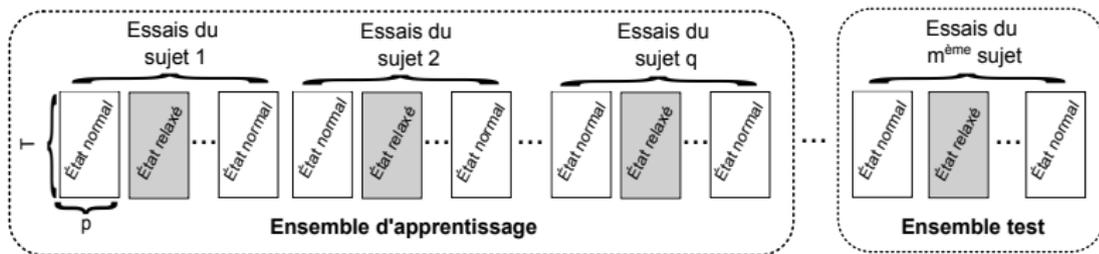




Adaptation au problème multi-sujet

Un sujet = plusieurs essais donc :

- Deux taux de bon classement :
 - le TBC des essais,
 - le TBC des sujets obtenu par "vote majoritaire".
- Dans l'estimation du TBC par LOO (Leave One Out), on retire **tous les essais d'un même sujet** pour former l'ensemble test lors de la m -ème itération (avec $m = 1, \dots, q$).



Plan

- 1 Acquisition de données EEG
 - Protocole d'acquisition
 - Données
- 2 Extraction de caractéristiques
 - Critère des pentes
 - Filtrage spatial par CSP et transformation logvar
- 3 Classification de signaux EEG
 - La méthodologie générale
 - Adaptation au problème multi-sujet
- 4 Sélection de variables
 - Filtrage par CSP sparse
 - Algorithme de filtrage par CSP sparse
 - Résultats

Un problème de sélection de variables

Objectif

Trouver la meilleure combinaison de p' électrodes (où $p = 58$ est le nombre d'électrodes et $p' < p$).

- $C_{58}^{p'} = \frac{58!}{p'!(58 - p)!}$ possibilités.
- La recherche exhaustive peut être très coûteuse en temps de calcul.
- Solution testées :
 - Algorithme génétique,
 - Méthode séquentielle de type backward et forward,
 - Sparsifier l'étape de filtrage par CSP.

Filtrage par CSP sparse

L'idée générale

- 1 Ecrire le CSP sous la forme de **deux ACP avec métriques**
- 2 Ecrire ces deux ACP avec métriques comme **deux ACP classiques**
- 3 Sélectionner les variables avec un algorithme d'**ACP sparse**.

Plusieurs algorithmes d'ACP sparse dans la littérature, en particulier :

- **Zou, Hastie et Tibshirani (2006)** → approche de type "bloc"
- Shen et Huang (2006) → approche de type "déflation"
- Journee et al. (2010)

Le filtrage par CSP : deux ACP avec métriques

⇒ Réécriture des deux problèmes d'optimisation du CSP :

$$\max_{\mathbf{w} \in \mathbb{R}^P} \frac{\text{var}(\mathbf{X}_1 \mathbf{w})}{\text{var}(\mathbf{X}_2 \mathbf{w})} \iff \max_{\mathbf{w} \in \mathbb{R}^P} \frac{\mathbf{w}^T \mathbf{C}_1 \mathbf{w}}{\mathbf{w}^T \mathbf{C}_2 \mathbf{w}} \iff \begin{array}{l} \max_{\mathbf{w} \in \mathbb{R}^P} \text{var}(\mathbf{X}_1 \mathbf{w}) \\ \text{s. c. } \mathbf{w}^T \mathbf{C}_2 \mathbf{w} = 1 \end{array}$$

$$\max_{\mathbf{v} \in \mathbb{R}^P} \frac{\text{var}(\mathbf{X}_2 \mathbf{v})}{\text{var}(\mathbf{X}_1 \mathbf{v})} \iff \max_{\mathbf{v} \in \mathbb{R}^P} \frac{\mathbf{v}^T \mathbf{C}_2 \mathbf{v}}{\mathbf{v}^T \mathbf{C}_1 \mathbf{v}} \iff \begin{array}{l} \max_{\mathbf{v} \in \mathbb{R}^P} \text{var}(\mathbf{X}_2 \mathbf{v}) \\ \text{s. c. } \mathbf{v}^T \mathbf{C}_1 \mathbf{v} = 1 \end{array}$$

⇒ En posant $\tilde{\mathbf{w}} = \mathbf{C}_2 \mathbf{w}$, on a :

$$\begin{aligned} \max_{\mathbf{w} \in \mathbb{R}^p} \text{var}(\mathbf{X}_1 \mathbf{w}) \\ \text{s. c. } \mathbf{w}^T \mathbf{C}_2 \mathbf{w} = 1 \end{aligned} \iff \begin{aligned} \max_{\tilde{\mathbf{w}} \in \mathbb{R}^p} \text{var}(\mathbf{X}_1 \mathbf{C}_2^{-1} \tilde{\mathbf{w}}) \\ \text{s. c. } \tilde{\mathbf{w}}_j^T \mathbf{C}_2^{-1} \tilde{\mathbf{w}}_j = 1 \end{aligned} \iff \text{ACP de } \mathbf{X}_1 \text{ avec} \\ \text{métrique } \mathbf{C}_2^{-1}$$

⇒ En posant $\tilde{\mathbf{v}} = \mathbf{C}_1 \mathbf{v}$, on a :

$$\begin{aligned} \max_{\mathbf{v} \in \mathbb{R}^p} \text{var}(\mathbf{X}_2 \mathbf{v}) \\ \text{s. c. } \mathbf{v}^T \mathbf{C}_1 \mathbf{v} = 1 \end{aligned} \iff \begin{aligned} \max_{\tilde{\mathbf{v}} \in \mathbb{R}^p} \text{var}(\mathbf{X}_2 \mathbf{C}_1^{-1} \tilde{\mathbf{v}}) \\ \text{s. c. } \tilde{\mathbf{v}}_j^T \mathbf{C}_1^{-1} \tilde{\mathbf{v}}_j = 1 \end{aligned} \iff \text{ACP de } \mathbf{X}_2 \text{ avec} \\ \text{métrique } \mathbf{C}_1^{-1}$$

Le CSP à partir de deux ACP classiques

- 1 L'ACP de \mathbf{X}_1 avec métrique s'obtient à partir de l'ACP classique de $\mathbf{Z}_1 = \mathbf{X}_1 \mathbf{C}_2^{-\frac{1}{2}}$:
 - recherche des vecteurs propres \mathbf{u}_j de la matrice de covariance de \mathbf{Z}_1 .
 - $\mathbf{w}_j = \mathbf{C}_2^{-\frac{1}{2}} \mathbf{u}_j$ donne le premier ensemble de vecteurs $\mathbf{w}_1, \dots, \mathbf{w}_k$
 - 2 L'ACP de \mathbf{X}_2 avec métrique s'obtient à partir de l'ACP classique de $\mathbf{Z}_2 = \mathbf{X}_2 \mathbf{C}_1^{-\frac{1}{2}}$ donne le second ensemble de vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_k$
- ⇒ Procédure équivalente à la procédure classique de CSP qui consiste à sélectionner les k premiers et les k derniers vecteurs propres de $\mathbf{C}_2^{-1} \mathbf{C}_1$.
- ⇒ On remplacera l'étape d'ACP classique par une ACP sparse pour sélectionner les électrodes.

L'ACP sparse

Le problème d'optimisation en ACP classique

$$\min_{F, \alpha} \|X - F\alpha^T\|^2 \quad (3)$$

- F est la **matrice** $n \times k$ des **scores** des n observations sur les k composantes principales ($F = X\alpha$)
- α est la **matrice** $p \times k$ des **loadings** des p variables sur les k composantes principales.
- $\|A\| = \text{Tr}(AA^T)$ est la norme de Frobenius de la matrice A ,

D'après [Eckart et Young, 1936], si $X = UDV^T$ est la **décomposition en valeurs singulières** de X avec

- U_k et V_k sont les matrices respectivement composées des k premières colonnes de U et V ,
- D_k est la matrices diagonale des s premières valeurs singulières de D ,

alors

$$F = U_k D_k \text{ et } \alpha = V_k.$$

Problème d'optimisation de l'ACP sparse de Zou, Hastie et Tibshirani (2006)

$$\begin{aligned} \min_{\alpha, \beta} & \|X - X\beta\alpha^T\|^2 + \sum_{j=1}^k \lambda^j \|\beta_j\|_1 \\ \text{s. c. } & \alpha^T \alpha = I_k \end{aligned} \quad (4)$$

où :

- β est la **matrice** des **loadings** **sparse**,
- α est la matrice des loadings orthogonaux,
- $\|\cdot\|_1$ est la norme l_1 ,
- $\beta_j \in \mathbb{R}^p$ est la $j^{\text{ème}}$ colonne de β ,
- λ^j est le **paramètre de régularisation**.

L'ACP sparse de Zou, Hastie et Tibshirani (2006) :

- Critère de type "block".
↳ Recherche **simultanée** des k vecteurs de loadings sparses.
- Implémenté dans le package R **elasticnet**.
- Comment **fixer les k paramètres de régularisation** $(\lambda^1, \dots, \lambda^k)$?

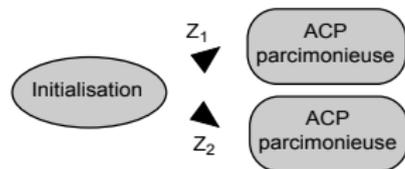
L'idée utilisée :

- Définir une grille matricielle de paramètres de régularisation qui est **une matrice** $\Lambda = (\lambda_i^j)_{r \times k}$.
- Augmenter i de 1 à r augmente globalement la sparsité sur les k composantes principales de l'ACP.
- Choisir "**le pas**" i qui optimise **un critère** en validation croisée par exemple.

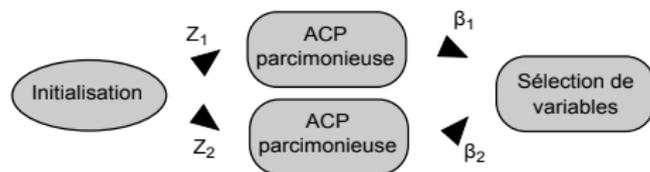
Algorithme de filtrage par CSP sparse

Initialisation

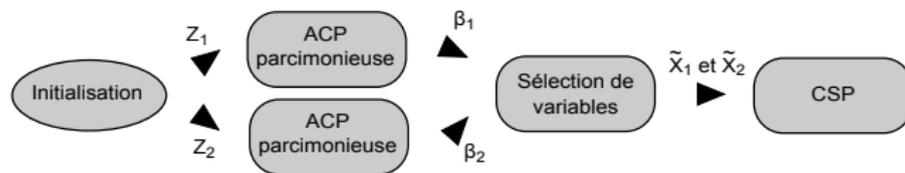
Algorithme de filtrage par CSP sparse



Algorithme de filtrage par CSP sparse



Algorithme de filtrage par CSP sparse



Algorithme de filtrage par CSP sparse



Illustration sur un exemple

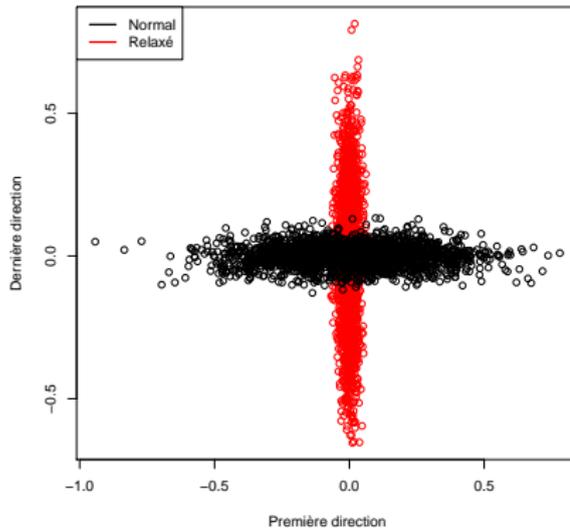
- Les 8 premières secondes d'enregistrement EEG du sujet 58 de l'expérience avant et après relaxation sont considérées.
- Le troisième pas de la grille de régularisation a été choisi et **7 variables** (électrodes) ont été sélectionnées pour construire $k = 2$ paires de composantes CSP.

On compare les signaux synthétiques obtenus

- avec les 58 électrodes
- avec les 7 électrodes sélectionnées.

Illustration sur un exemple

Paire 1 de composantes CSP



Paire 1 de composantes CSP parcimonieuses

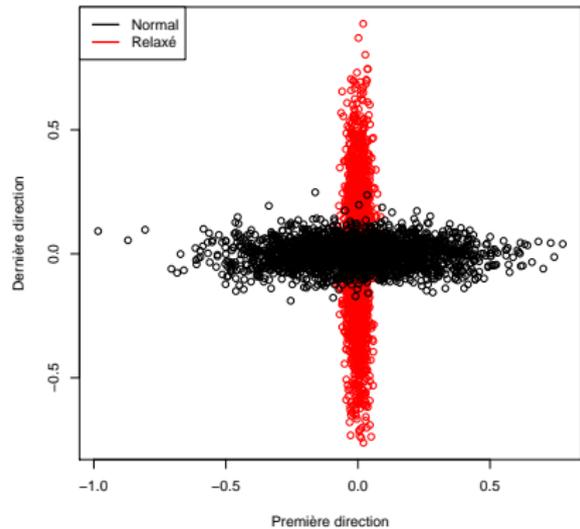
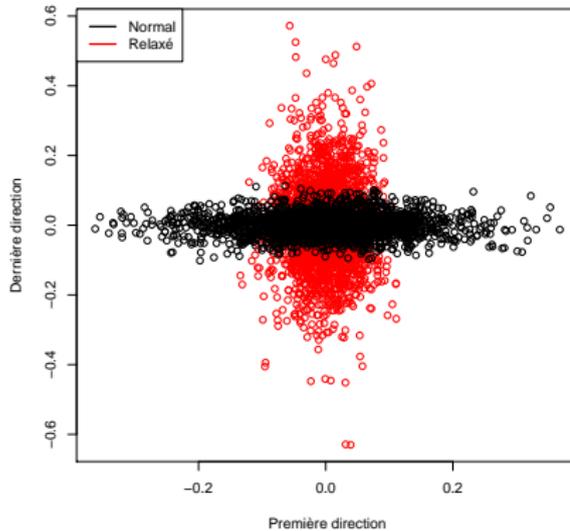
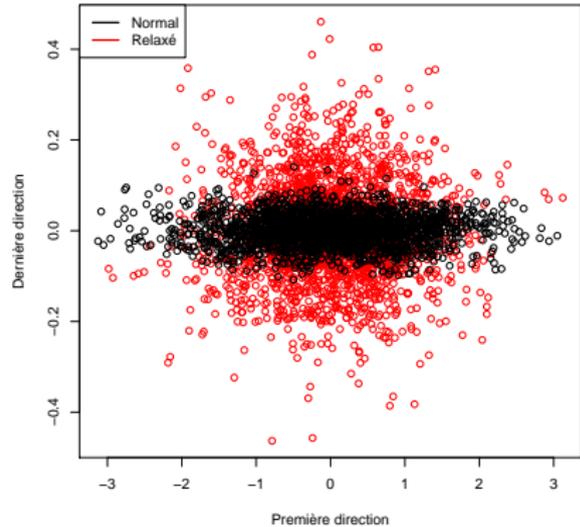


Illustration sur un exemple

Paire 2 de composantes CSP



Paire 2 de composantes CSP parcimonieuses



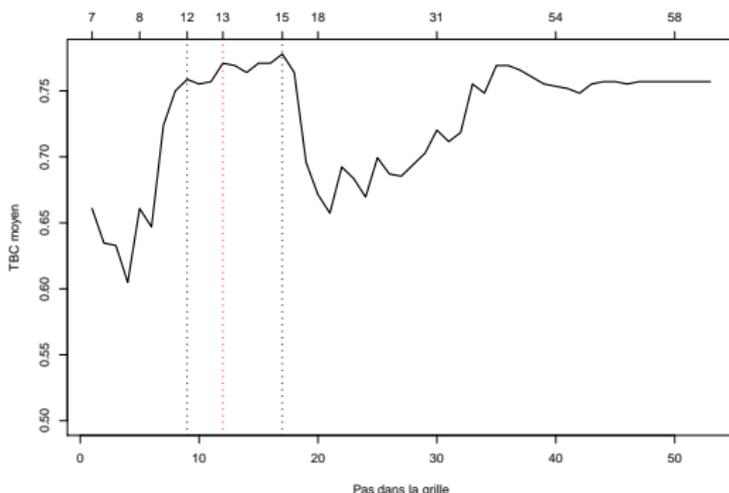
Résultats

Échantillons

- Échantillon d'apprentissage : Essais des sujets de la campagne 1.
- Échantillon test : Essais des sujets de la campagne 2.

Sélection des électrodes

Le TBC est calculé par LOO (de type sujet/essai) sur les **essais des sujets de la campagne 1**, avec $k = 3$ paires de filtres pour l'étape de CSP sparse et la méthode LDA pour l'étape de classification binaire.



↪ On choisit de conserver **13 électrodes**.

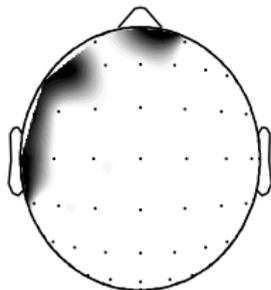
TBC sur l'échantillon test

- Sur l'échantillon d'apprentissage :
 - le filtrage CSP des 13 électrodes sélectionné donne la matrice W de taille 13×6 .
 - une transformation logvar donne la matrice Z de dimension 22×6 et un LDA appris sur Z fournit la règle de décision.
- Sur l'échantillon test, on calcule le TBC.

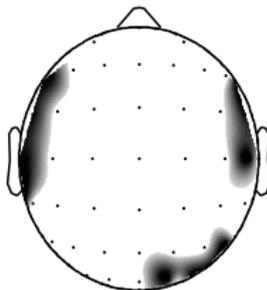
	Méthode de classification	Nombre d'électrodes dans le modèle prédictif	TBC estimé sur l'échantillon test (campagne 2)	
			essais	sujets
Algorithme génétique : CSP et transformation logvar	LDA	9	71.59	75
sCSP et transformation logvar	LDA	13	73.11	75

⇒ Résultats comparables mais sCSP est beaucoup plus rapide.

Sélection par algorithme génétique

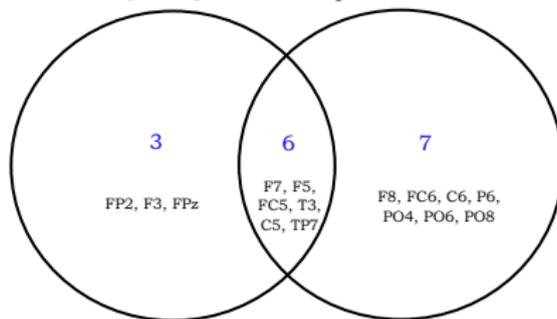


Sélection par CSP parcimonieux

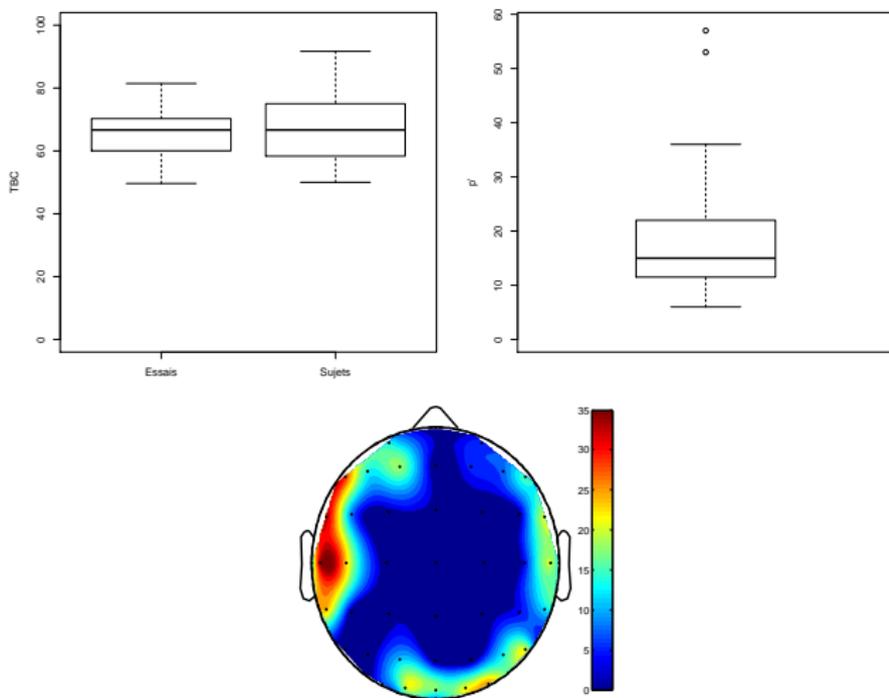


Algorithme
génétique

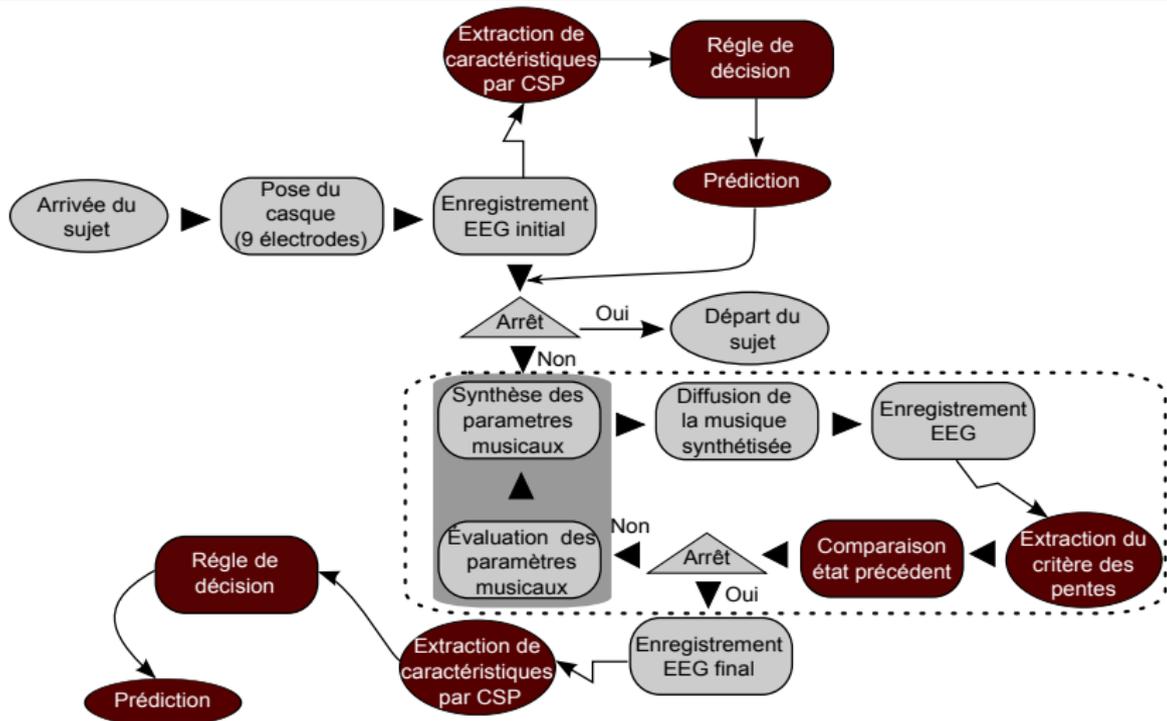
CSP
parcimonieux



Résultats pour 35 découpages



La suite...



Références I



B. Blankertz, K.R. Müller, D. Krusienski, G. Schalk, J.R. Wolpaw, A. Schlögl, G. Pfurtscheller, J.R. Millán, M. Schröder and N. Birbaumer
The BCI competition III: Validating alternative approaches to actual BCI problems
Transactions on Neural Systems and Rehabilitation Engineering, 2006, 14, 153-159



C. Eckart and G. Young
The approximation of one matrix by another of lower rank.
Psychometrika, Springer, 1936, 1, 211-218



B. Efron, T. Hastie, I. Johnstone and R. Tibshirani,
Least angle regression.
The Annals of Statistics, 2003, 32, 407-499



J. Friedman, T. Hastie and R. Tibshirani,
Regularization Paths for Generalized Linear Models via Coordinate Descent.
Journal of Statistical Software, 2010, 33, 1-22



K. Fukunaga and W. L. G. Koontz,
Application of the Karhunen-Loève Expansion to Feature Selection and Ordering.
IEEE Transactions on Computers, 1970, C-19, 311-318



E. Jacobson.
Biologie des émotions. Les bases théoriques de la relaxation.
E.S.P., Paris, 1974

Références II



M. Journée, Y. Nesterov, P. Richtarik and R. Sepulchre,
Generalized power method for sparse PCA.
Journal of Machine Learning Research, 2010, 11, 517-553



Z. J. Koles, M. S. Lazar and S. Z. Zhou
Spatial patterns underlying population differences in the background EEG.
Brain topography, 1990, vol. 2, 275-284



A. Schlögl, F. Lee, H. Bischof and G. Pfurtscheller
Characterization of four-class motor imagery EEG data for the BCI-competition 2005.
Journal of Neural Engineering, 2005, vol. 2, no. 4, pp. L14-L22



J. H. Schultz.
Le Training Autogène.
PUF, Paris, 1958



L. Vézard, P. Legrand, M. Chavent, F. Faïta-Aïnseba, and L. Trujillo
EEG classification for the detection of mental states.
To appear in Applied Soft Computing, 2015



L. Vézard,
Réduction de dimension en apprentissage supervisé : applications à l'étude de l'activité cérébrale.
Thèse de l'Université de Bordeaux, 2013



H. Zou, T. Hastie and R. Tibshirani,
Sparse Principal Component Analysis.
Journal of Computational and Graphical Statistics, 2006, vol. 15, no. 2, pp. 265-286