

# BCI signal classification using a Riemannian-based kernel

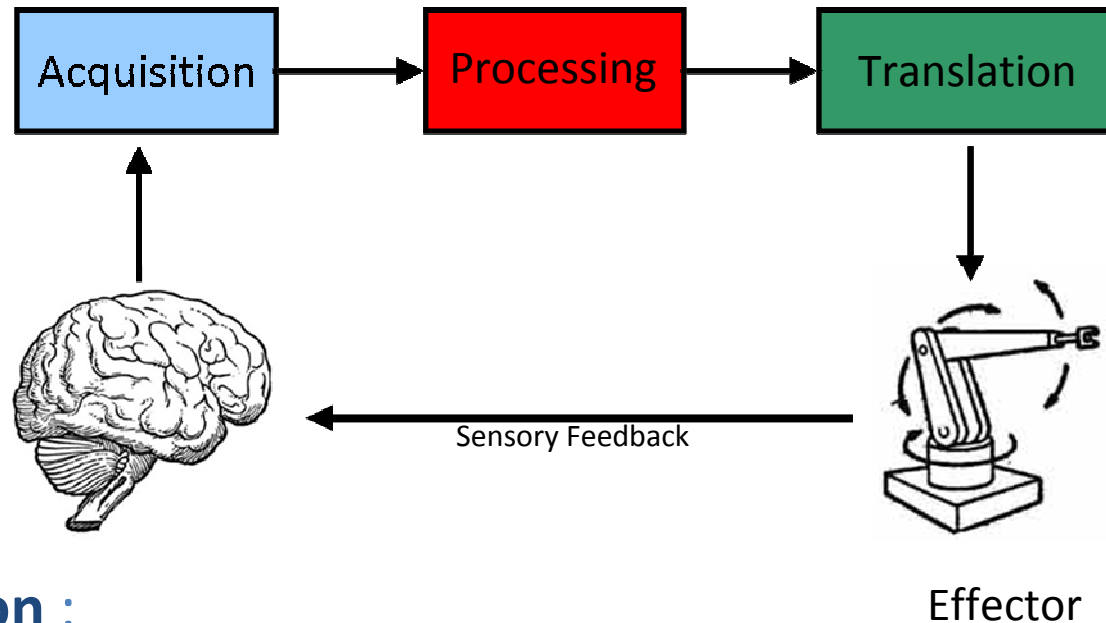
A. Barachant<sup>1,2</sup>, [S. Bonnet](#)<sup>1</sup>, M. Congedo<sup>2</sup>, C. Jutten<sup>2</sup>

<sup>1</sup> CEA-LETI/DTBS, Grenoble, France

<sup>2</sup> GIPSA-lab, Grenoble, France

25-April-2012

# Brain-Computer Interface (BCI)



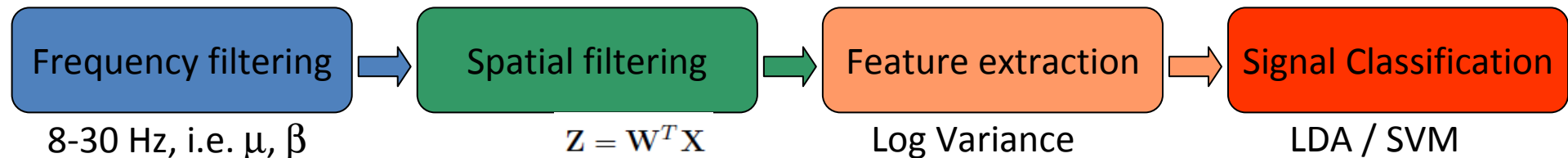
- **Acquisition :**

- EEG system with scalp electrodes (system 10-20)
- BCI paradigm using motor imagery (asynchronous BCI)

- **Signal Processing :**

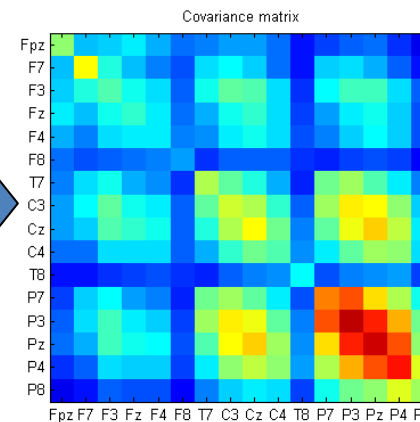
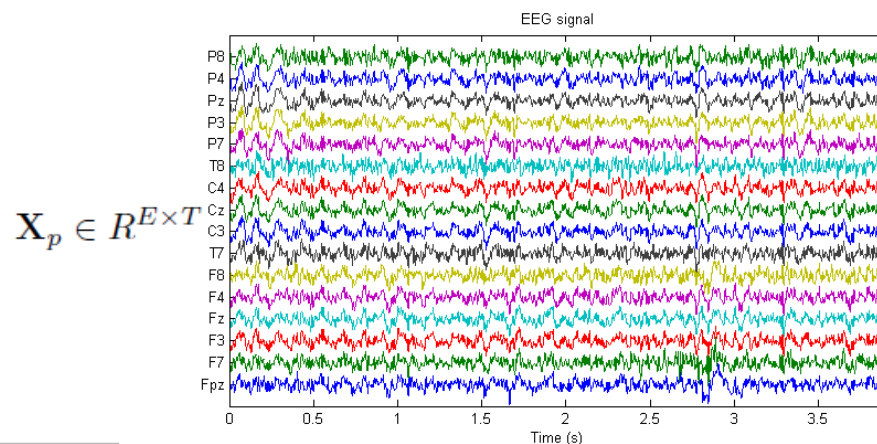
- Pre-processing : artefact removal, EEG signal band-pass filtering
- Two-class classification : Discriminate between two MI tasks

# BCI Motor imagery (standard approach)



## ■ Spatial Filtering step (usually data-driven)

- CSP Criterion : promote **variance difference** between two classes
- Joint **diagonalization** of class-conditional mean spatial covariance matrices
- Select  $N_f$  spatial filters : Loss of information



$$\begin{aligned}
 C_p &= \Sigma(X_p) \\
 &= \frac{1}{T-1} X_p X_p^T
 \end{aligned}$$

# Can we avoid the spatial filtering?

- Re-interpret CSP-based linear classification

- CSP decision function :

$$h(\mathbf{X}) = v_0 + \sum_{n=1}^{N_f} v_n \log [\mathbf{W}^T \Sigma(\mathbf{X}) \mathbf{W}]_{n,n}$$

- Omitting log() operator, this yields a feature dimension space  $E^*=E (E+1)$

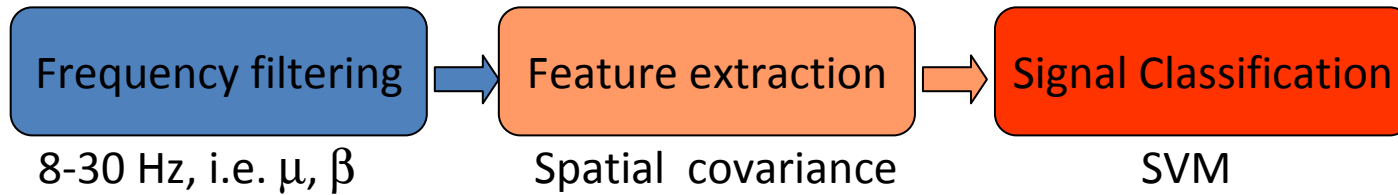
$$h(\mathbf{X}) = u_0 + \mathbf{u}^T \text{vec} [\Sigma(\mathbf{X})]$$

- Spatial filter matrix  $\mathbf{W}$  can be estimated from  $\mathbf{u}$ , since  $\mathbf{u} = \text{vec} [\mathbf{W} \text{diag}(\mathbf{v}) \mathbf{W}]$

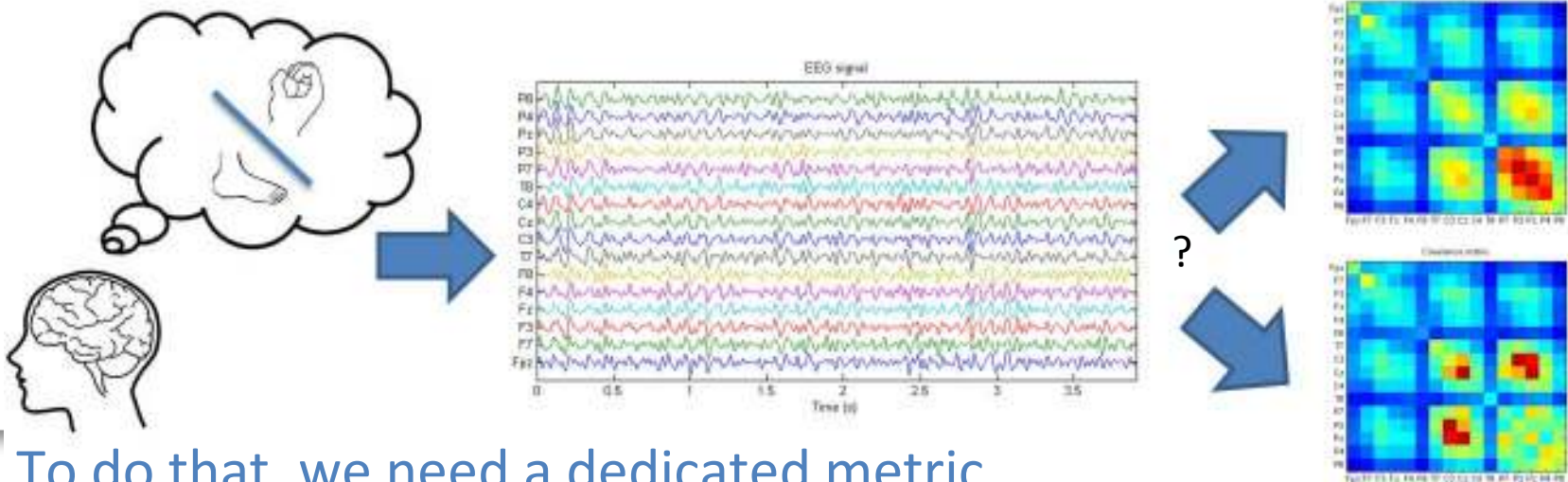
- [Reuderink, 2011] Direct covariance classification on whitened trials  $\tilde{\mathbf{X}} = \mathbf{P}^T \mathbf{X}$

$$h(\mathbf{X}) = u_0 + \mathbf{u}^T \text{vec} [\mathbf{P}^T \Sigma(\mathbf{X}) \mathbf{P}]$$

# Can we go further ?

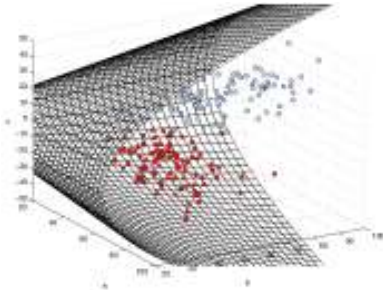


- How to classify covariance matrices in MI-based BCI ?
  - Bayesian Framework (Wishart distribution)
  - Algebraic Framework (geometric approach)
  - Information Geometry Framework



# Riemannian manifold of SPD matrices

- Space of Symmetric Positive-Definite (SPD) matrices
- Differentiable manifold (dimension  $E^*$ )
  - Covariance matrices are points in this manifold and Riemannian distance can be computed between two points [Barachant, 2012].
  - At each point  $C$  ( i.e. each covariance matrix), a **scalar product** can be defined in the associated tangent space



$$\langle S_1, S_2 \rangle_C = \text{tr}(S_1 C^{-1} S_2 C^{-1}).$$

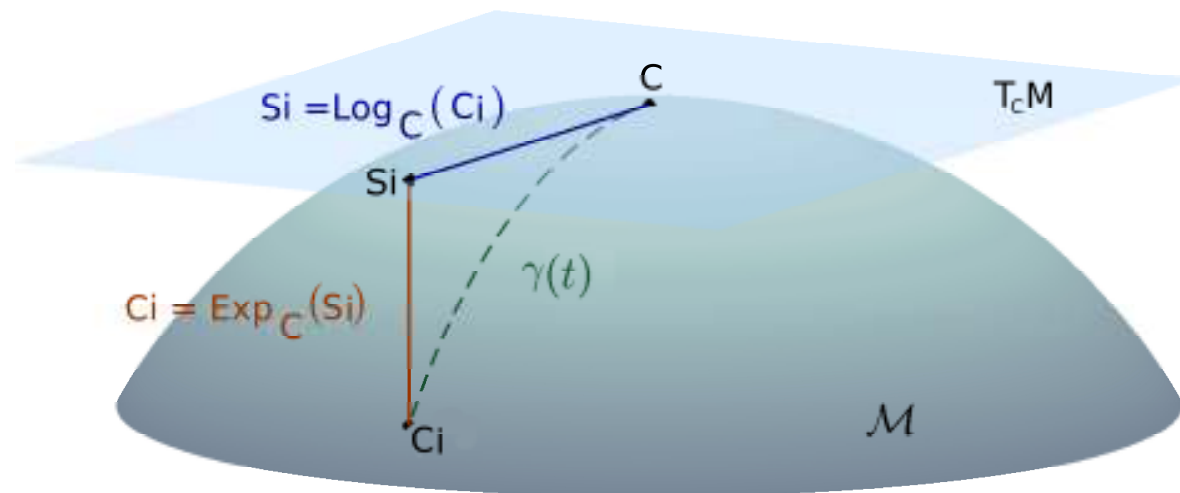
- Distance between two SPD matrices (along the geodesic)

$$d_R(C_i, C_j) = \|\logm(C_i^{-1} C_j)\|_F$$

# Logarithmic/Exponential map of SPD matrices

- Project locally a covariance matrix  $\mathbf{C}_p$  onto the tangent plane

$$\mathbf{S}_p = \text{Log}_{\mathbf{C}}(\mathbf{C}_p) = \mathbf{C}^{1/2} \logm \left( \mathbf{C}^{-1/2} \mathbf{C}_p \mathbf{C}^{-1/2} \right) \mathbf{C}^{1/2}$$





# Proposed kernel

- An usual approach consists in mapping data in another feature space (usually with higher dimensionality)
  - Empirical kernel choice
  - Most employed : RBF kernel  $k(\mathbf{x}_i, \mathbf{x}_j) = \exp[-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2]$
- Riemannian geometry provides a natural kernel to deal with covariance matrices
  - Mapping function  $\phi(\mathbf{C}) = \text{Log}_{\mathbf{C}_{\text{ref}}}(\mathbf{C})$

$$\begin{aligned} k(\text{vec}(\mathbf{C}_i), \text{vec}(\mathbf{C}_j)) &= \langle \phi(\mathbf{C}_i), \phi(\mathbf{C}_j) \rangle_{\mathbf{C}_{\text{ref}}} \\ &= \text{tr} \left[ \logm \left( \mathbf{C}_{\text{ref}}^{-1/2} \mathbf{C}_i \mathbf{C}_{\text{ref}}^{-1/2} \right) \logm \left( \mathbf{C}_{\text{ref}}^{-1/2} \mathbf{C}_j \mathbf{C}_{\text{ref}}^{-1/2} \right) \right]. \end{aligned}$$

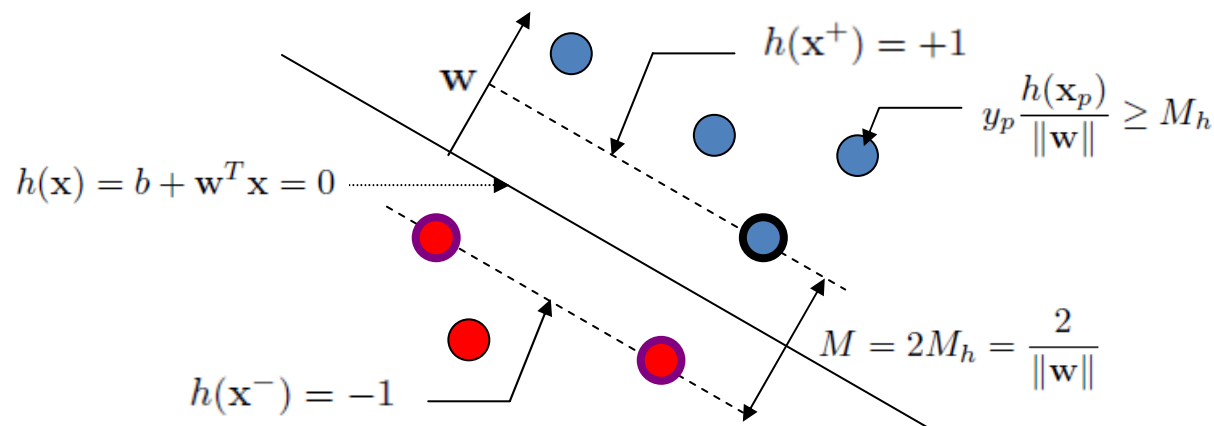
- Application to SVM classification



# SVM classification

- Linear (separable) two-class SVM

- Supervised classification with a set of labelled feature vectors  $\{(\mathbf{x}_p, y_p)\}_{p=1}^P$
- Seeks to linearly separate data by finding an hyperplane maximising the margin  $M$



Cost function:

$$\operatorname{argmin}_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \text{ s.t. } y_p h(\mathbf{x}_p) \geq 1$$

- Decision function

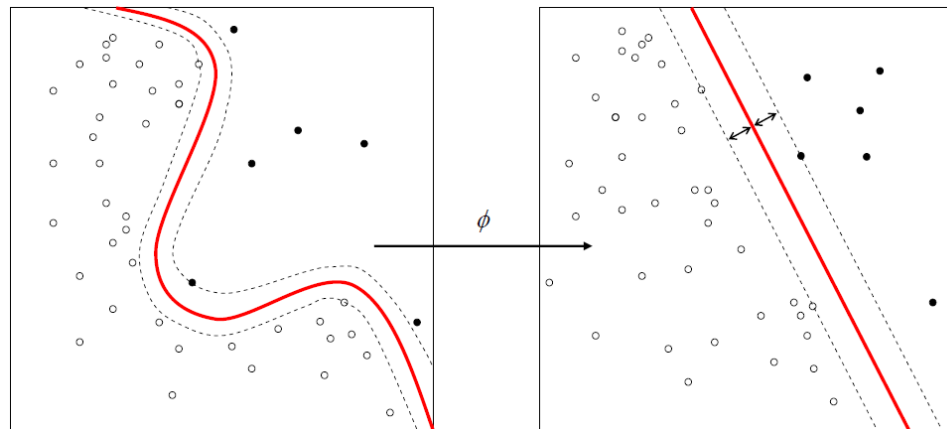
$$h(\mathbf{x}) = b + \sum_{p=1}^P \alpha_p y_p \langle \mathbf{x}_p, \mathbf{x} \rangle$$

non-zero for support vectors

# SVM (kernel trick)

- Kernel trick :

- Map data in a new feature space (where hopefully separable)
- The mapping function is rarely expressed, the kernel function is key in the computation  $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle_{\mathcal{H}}$



- Decision function :

$$h(\mathbf{x}) = b + \sum_{p=1}^P \alpha_p y_p \langle \phi(\mathbf{x}_p), \phi(\mathbf{x}) \rangle_{\mathcal{H}}$$

- Use new kernel based on Riemannian geometry

$$k(\text{vec}(\mathbf{C}_i), \text{vec}(\mathbf{C}_j)) = \langle \phi(\mathbf{C}_i), \phi(\mathbf{C}_j) \rangle_{\mathcal{C}_{\text{ref}}}$$

# Choice of free parameter $\mathbf{C}_{\text{ref}}$

- $\mathbf{C}_{\text{ref}}$  := Point in SPD space where the tangent plane is computed
- @ Identity (log-Euclidean kernel)

$$\mathbf{C}_{\text{ref}} = \mathbf{I}_E$$

$$\begin{aligned} \longrightarrow k(\text{vec}(\mathbf{C}_i), \text{vec}(\mathbf{C}_j)) &= \text{tr} [\text{logm}(\mathbf{C}_i) \text{logm}(\mathbf{C}_j)] \\ &= \langle \text{logm}(\mathbf{C}_i), \text{logm}(\mathbf{C}_j) \rangle_F \end{aligned}$$

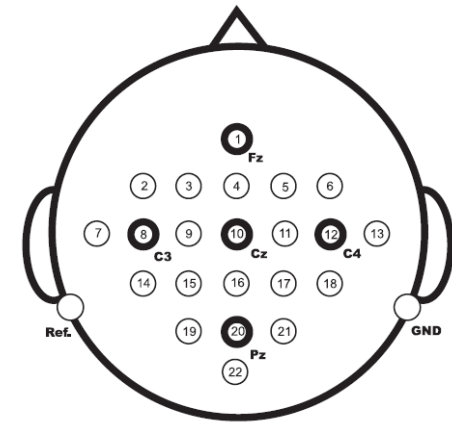
- @ geometric mean of the  $P$  labeled covariance matrices

$$\mathbf{C}_{\text{ref}} = \underset{\mathbf{C}}{\text{argmin}} \sum_{p=1}^P d_R^2(\mathbf{C}, \mathbf{C}_p)$$

# Experiments

- **Asynchronous MI-based BCI** (BCI competition IV, dataset 2a)

- 9 subjects
- 22 electrodes
- Reference electrode on the left mastoid
- 8-35 Hz (general) band-pass filter
- 4-class dataset : RH, LH, TO, BF (144 trials per class)



- **Objectives :**

- Average performance across subjects and across all pairs of binary classification
- Performance comparison
  - Standard CSP method
  - SVM applied on vectorized covariance matrices
  - Covariance kernel-SVM @ identity
  - Covariance kernel-SVM @ geometric mean
- 30-fold cross-validation

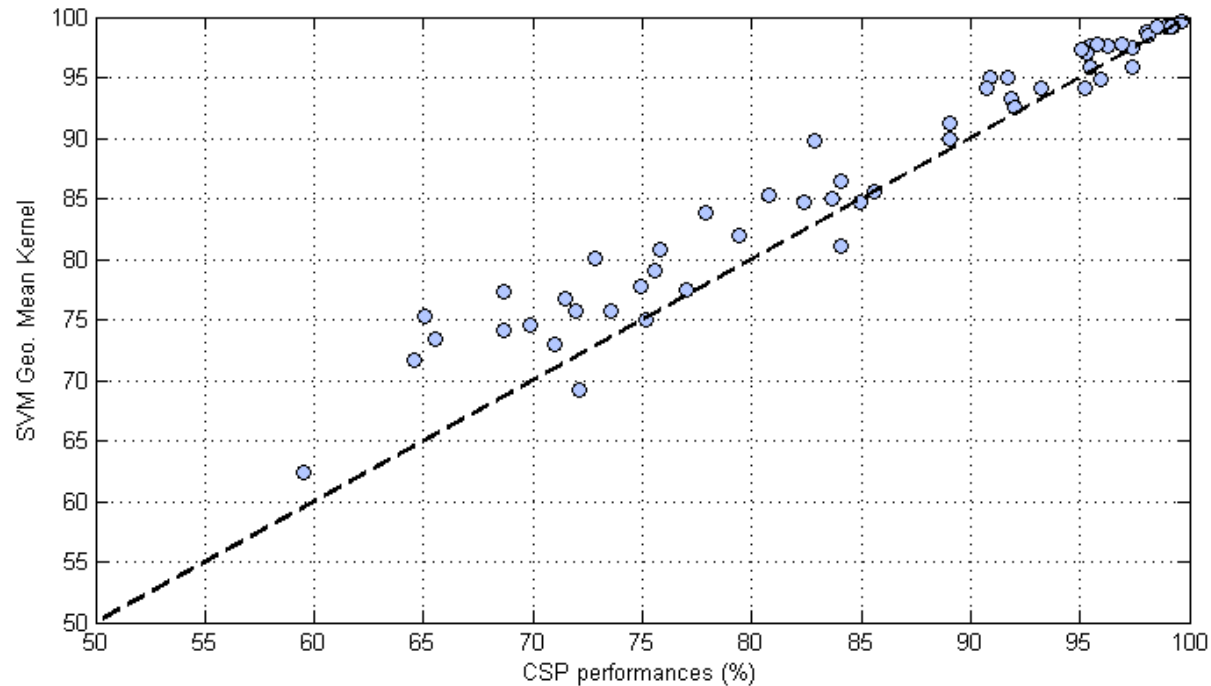
# Results (1/3)

	SVM kernel Geometric Mean	SVM kernel Identity	SVM plain	CSP
LH/RH	83.4	82.8	76.9	81.4
LH/BF	89.9	89.5	82.2	86.9
LH/TO	89.3	88.8	84	87.3
RH/BF	88.8	88.2	81.1	86.3
RH/TO	88.7	88	83.4	85.7
BF/TO	82.1	81.4	74.9	80.8
mean	87	86.4	80.4	84.7

Table 1: Average classification accuracy across the 9 subjects for 6 pairs of mental tasks.

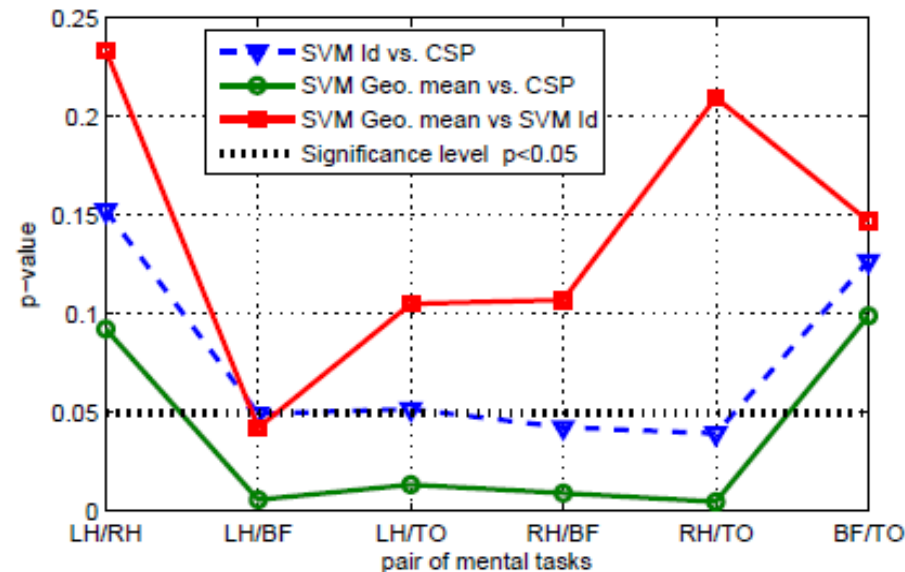
- R-based kernel SVM **outperforms** CSP+LDA in all cases,
- Direct covariance classification gives poor results,
- Geometric mean of the P SPD matrices is a good location to compute the tangent plane.

## Results (2/3)



- Individual sessions are shown
- RkSVM is superior especially in almost all sessions, especially in difficult situations

# Results (3/3)



- Statistical analysis (9 subjects)
  - P-values for the 6 pairs of mental tasks in a subject-independent manner.
  - Hypothesis  $H_0 : \mu_1 > \mu_2$ , one-tailed dependent t-test (8 df) for paired samples
- RkSVM classification is significantly better in almost all pairs of mental task.



# Conclusion

- New covariance kernel for directly handling covariance matrices in classification methods
  - No need for explicit spatial filtering
  - Simple to implement (just add a new kernel in SVM toolbox!)
- New framework of R. geometry in BCI
- Successful application on a BCI competition dataset
  - outperforms significantly the conventional CSP method (two-class)
- Future work
  - Multi-class classification
  - Online application w/ location update of the tangent space between BCI sessions
  - Investigate the minimum # of trials required to properly estimate the classifier
  - Use regularized version of SVM to deal with high-dimensional ( $E^*$ ) features

# leti

LABORATOIRE D'ÉLECTRONIQUE  
ET DE TECHNOLOGIES  
DE L'INFORMATION

CEA-Leti

MINATEC Campus, 17 rue des Martyrs

38054 GRENOBLE Cedex 9

Tel. +33 4 38 78 36 25

[www.leti.fr](http://www.leti.fr)



## Merci de votre attention



énergie atomique • énergies alternatives

