# 1 Dependent component extraction

## 1.1 The task

In this competition data are constructed which are as close as possible to real EEG data with minor changes to ensure some ground truth which is, in principle, detectable. While independent sources are often a useful assumption recent research has focused on the analysis of brain connectivity. Therefore, the objective of the research is the detection of dependent sources.

The competion data were constructed as a superposition of $N = 19$ sources measured in as many sensors. Out of the 19 sources two were dependent and all others were mutually independent. The construction was done in the following way. We decomposed EEG data $\mathbf{x}(t) = (x_1(t), ..., x_N(t))^T$ from one subject using the TDSEP algorithm [Ziehe et al., 1998] in the standard way as

$$\mathbf{x}(t) = A\mathbf{s}(t) \tag{1}$$

with $A$ being the mixing matrix and $\mathbf{s}(t)$ being the estimated sources. Two of the 19 sources were selected as being dependent. For notational simplicity we denote the respective source indices as 1 and 2.

The challenge data $\mathbf{y}(t)$ were then constructed as

$$\mathbf{y}(t) = \sum_{i=1}^{2} \mathbf{a}_i s_i(t) + \sum_{i=3}^{N} \mathbf{a}_i \tilde{s}_i(t) \tag{2}$$

where $\mathbf{a}_i$ is the $ith.$ column of the mixing matrix $A$. The time series $\tilde{s}_i(t)$ were all taken from real data from mutually different subjects. For each subject the data were decomposed using ICA and the $i.th$ original source $s_i(t)$ (for $i > 2$) was replaced by the $i.th$ source of the $i.th$ subject with ordering according to magnitude of the ICA-components.

The task was to recover from $\mathbf{y}(t)$ the space spanned by the two columns $\mathbf{a}_1$ and $\mathbf{a}_2$. It was not the task to recover the two columns separately because for an interacting system the information given was not sufficient. It would have been necessary to make additional, e.g. spatial assumptions, on the nature of the sources to uniquely decompose the subspace into separate sources.

A 'distance' between two subspaces can be defined in terms of the respective projectors on the respective subspace. If $\hat{A} = (\mathbf{a}_1, \mathbf{a}_2)$ then

$$P_A = \hat{A}\left(\hat{A}^T\hat{A}\right)^{-1}\hat{A}^T \tag{3}$$

is a projector onto the space spanned by the columns $\mathbf{a}_1$ and $\mathbf{a}_2$, i.e. $P_A$ is a projector onto the true subspace. Let, similarly, $P_B$ be the projector on the estimated 2-dimensional subspace. We then calculate the eigenvalues of

$$D = P_A P_B P_A \tag{4}$$

Writing the eigenvalues in desceding order gor 2-dimensional subspaces only the first two eigenvalues can be non-vanishing and all eigenvalues are in the interval $[0, 1]$. If the first eigenvalue is equal to 1 a direction exists which is common to to both subspaces. If also the second eigenvalue is equal to 1 another such direction, orthogonal to the first one, exists. In particular, the subspaces are identical, of the second eigenvalue is equal to 1. For the data competition the value of this second eigenvalue was used to asses how accurately the true subspace was recovered.

# References

[Ziehe et al., 1998] Ziehe A, Müller KR. TDSEP-an efficient algorithm for blind separation using time structure Proceedings of the 8th International Conference on Artificial Neural Networks—ICANN 98, 2:675-80.