

Supplementary material to the article “EEG extended source localization: tensor-based vs. conventional methods”

H. Becker^{a,e,b,c}, L. Albera^{b,c,d}, P. Comon^e, M. Haardt^f, G. Birot^{b,c}, F. Wendling^{b,c}, M. Gavaret^{g,h,i}, C. G. Bénar^{g,h}, I. Merlet^{b,c}

^a*Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, F-06900 Sophia Antipolis, France*

^b*INSERM, U1099, Rennes, F-35000, France*

^c*Université de Rennes 1, LTSI, Rennes, F-35000, France*

^d*Centre INRIA Rennes-Bretagne Atlantique, Rennes, F-35042, France*

^e*GIPSA-Lab, CNRS UMR5216, Grenoble Campus BP.46, F-38402 St Martin d'Herès Cedex*

^f*Ilmenau University of Technology, Communications Research Laboratory, P. O. Box 10 05 65, D-98684 Ilmenau, Germany*

^g*INSERM, UMR 1106, F-13005 Marseille, France*

^h*Aix-Marseille Université, F-13005 Marseille, France*

ⁱ*AP-HM, Hopital Timone, F-13005 Marseille, France*

S.1. Trilinear approximation

In the following, we provide a theoretical analysis of the STF and STWV techniques. In order to treat both tensor methods simultaneously, we use a different notation for the matrices as in the paper to avoid confusion that may be caused by discrepancies from the data model of equation (??). Please note that in the following, for the STF method, the matrix \mathbf{D} replaces the data matrix \mathbf{X} , the matrix \mathbf{U} corresponds to the spatial mixing matrix \mathbf{H} that we want to extract and the matrix \mathbf{P} corresponds to the signal matrix \mathbf{S} . For the STWV method, \mathbf{D} replaces the transpose of the data matrix, \mathbf{X}^T , \mathbf{U}

Email address: laurent.albera@univ-rennes1.fr (I. Merlet)

corresponds to the transpose of the signal matrix \mathbf{S}^T that is to be identified and \mathbf{P} corresponds to the transpose of the spatial mixing matrix \mathbf{H}^T .

If a time-frequency or space-wave-vector transform is applied to the second dimension of the matrix $\mathbf{D} = \mathbf{U}\mathbf{P}$ where $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{N \times R}$ is the matrix of interest and $\mathbf{P} \in \mathbb{R}^{R \times K}$, one obtains a tensor with the following structure:

$$\mathbf{T} = \sum_{r=1}^R \mathbf{u}_r \circ \mathbf{M}_r \quad (1)$$

where $\mathbf{M}_r \in \mathbb{C}^{K \times J}$, $r = 1, \dots, R$, are matrices of rank L_r and \circ denotes the outer (tensor) product. Equation (1) corresponds to a block-decomposition into rank(1, L_r , L_r)-components, which is unique up to scale and permutation indeterminacies for rank-deficient matrices \mathbf{M}_r under certain conditions on N , K , J , L_r and R (De Lathauwer, 2008). However, in practice, the matrices \mathbf{M}_r generally have full rank. In this case, it is not possible to identify \mathbf{u}_r and \mathbf{M}_r from the given tensor \mathbf{T} . In order to restore identifiability, the matrices \mathbf{M}_r need to be approximated by matrices $\tilde{\mathbf{M}}_r$ of lower rank \tilde{L}_r such that one obtains a model of the form:

$$\tilde{\mathbf{T}} = \sum_{r=1}^R \mathbf{u}_r \circ \tilde{\mathbf{M}}_r. \quad (2)$$

For $\tilde{L}_r = 1$, $r = 1, \dots, R$, the tensor $\tilde{\mathbf{T}}$ can then be decomposed using the CP decomposition, which permits to uniquely identify the vectors of interest \mathbf{u}_r up to scale and permutation ambiguities.

The objective thus consists in transforming equation (1) into equation (2). This can, under certain conditions, be achieved by a truncated SVD

in one or several modes of the tensor \mathbf{T} .¹ This procedure can be viewed as some kind of PCA applied to the data in the transformed (time-frequency or space-wave-vector) domain.

S.1.1. Sufficient conditions for perfect recovery of \mathbf{U}

In the following, we determine the conditions under which the SVD permits to obtain the model of (2) for $\tilde{L}_r = 1$. For simplicity, we limit the considerations in the remainder of this section to the case of $R = 2$ components. Nevertheless, we believe that it is possible to extend our analysis to cases where $R > 2$. We define the following notation for the SVD of \mathbf{M}_1 and \mathbf{M}_2 :

$$\begin{aligned} \mathbf{M}_1 &= \sum_{l=1}^{L_1} \sigma_l \mathbf{v}_l \mathbf{w}_l^T = \begin{bmatrix} \mathbf{v}_1 & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \mathbf{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 & \mathbf{W}_2 \end{bmatrix}^T \\ \mathbf{M}_2 &= \sum_{l=1}^{L_2} \lambda_l \mathbf{x}_l \mathbf{y}_l^T = \begin{bmatrix} \mathbf{x}_1 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 & \mathbf{Y}_2 \end{bmatrix}^T \end{aligned}$$

where $\mathbf{V}_2 = [\mathbf{v}_2, \dots, \mathbf{v}_L]$, $\mathbf{W}_2 = [\mathbf{w}_2, \dots, \mathbf{w}_L]$, $\mathbf{X}_2 = [\mathbf{x}_2, \dots, \mathbf{x}_L]$, $\mathbf{Y}_2 = [\mathbf{y}_2, \dots, \mathbf{y}_L]$, $\sigma_1 > \sigma_2 > \dots > \sigma_{L_1}$, and $\lambda_1 > \lambda_2 > \dots > \lambda_{L_2}$. Moreover, without loss of generality, we assume that $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$. For simplicity, we subsequently base our considerations on the mode-2 unfolding of the tensor \mathbf{T} . The same consideration can be conducted for the mode-3 unfolding in an analogous way.

¹The truncated SVD in mode n is obtained by calculating the SVD of the n -mode unfolding matrix and setting all but the R greatest singular values to 0. Please note that a truncated SVD of the first mode does not change the data because the mode-1 unfolding matrix inherently has rank R .

With the above definitions, and \otimes denoting the Kronecker product, the mode-2 unfolding of the tensor \mathbf{T} can be written as

$$\begin{aligned} [\mathbf{T}]_{(2)} &= \sigma_1 \mathbf{v}_1 (\mathbf{w}_1 \otimes \mathbf{u}_1)^T + \lambda_1 \mathbf{x}_1 (\mathbf{y}_1 \otimes \mathbf{u}_2)^T \\ &\quad + \mathbf{V}_2 \boldsymbol{\Sigma}_2 (\mathbf{W}_2 \otimes \mathbf{u}_1)^T + \mathbf{X}_2 \boldsymbol{\Lambda}_2 (\mathbf{Y}_2 \otimes \mathbf{u}_2)^T \end{aligned} \quad (3)$$

$$= \sigma_1 \mathbf{v}_1 (\mathbf{w}_1 \otimes \mathbf{u}_1)^T + \lambda_1 \mathbf{x}_1 (\mathbf{y}_1 \otimes \mathbf{u}_2)^T + \mathbf{R}. \quad (4)$$

We would like to obtain the matrix

$$[\tilde{\mathbf{T}}]_{(2)} = \sigma_1 \mathbf{v}_1 (\mathbf{w}_1 \otimes \mathbf{u}_1)^T + \lambda_1 \mathbf{x}_1 (\mathbf{y}_1 \otimes \mathbf{u}_2)^T, \quad (5)$$

which corresponds to the CP model

$$\tilde{\mathbf{T}} = \sigma_1 \mathbf{u}_1 \circ \mathbf{v}_1 \circ \mathbf{w}_1 + \lambda_1 \mathbf{u}_2 \circ \mathbf{x}_1 \circ \mathbf{y}_1 \quad (6)$$

and would therefore permit us to recover the vectors \mathbf{u}_1 and \mathbf{u}_2 from the mode-2 unfolding matrix $[\mathbf{T}]_{(2)}$ by means of a truncated SVD. This is possible if (4) corresponds to the SVD of $[\mathbf{T}]_{(2)}$, which is generally not the case. Our objective now consists in finding conditions under which the SVD of $[\mathbf{T}]_{(2)}$ takes the form of (4) and under which truncation of (4) leads to (5).

Let us consider the case that $\mathbf{v}_1^T \mathbf{X}_2 = \mathbf{0}^T$, $\mathbf{x}_1^T \mathbf{V}_2 = \mathbf{0}^T$, $\mathbf{w}_1^T \mathbf{Y}_2 = \mathbf{0}^T$, and $\mathbf{y}_1^T \mathbf{W}_2 = \mathbf{0}^T$. The columns of the matrices $\sigma_1 \mathbf{v}_1 (\mathbf{w}_1 \otimes \mathbf{u}_1)^T$ and $\lambda_1 \mathbf{x}_1 (\mathbf{y}_1 \otimes \mathbf{u}_2)^T$ are then pairwise orthogonal to the columns of \mathbf{R} and the columns of the matrices $\sigma_1 (\mathbf{w}_1 \otimes \mathbf{u}_1) \mathbf{v}_1^T$ and $\lambda_1 (\mathbf{y}_1 \otimes \mathbf{u}_2) \mathbf{x}_1^T$ are pairwise orthogonal to the columns of \mathbf{R}^T . Due to the correlation between the vectors \mathbf{v}_1 and \mathbf{x}_1 , the vectors \mathbf{u}_1 and \mathbf{u}_2 , and the vectors \mathbf{w}_1 and \mathbf{y}_1 , the two associated mode-2 vectors $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{x}}_1$ that are obtained by the SVD correspond to a linear combination of \mathbf{v}_1 and \mathbf{x}_1 . Furthermore, the vectors $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{x}}_1$ are associated

with two new singular values, $\mu_1 \geq \max(\sigma_1, \lambda_1)$ and $\mu_2 \leq \min(\sigma_1, \lambda_1)$. These singular values can be computed as the square roots of the eigenvalues of

$$[\sigma_1 \mathbf{v}_1(\mathbf{w}_1 \otimes \mathbf{u}_1)^T + \lambda_1 \mathbf{x}_1(\mathbf{y}_1 \otimes \mathbf{u}_2)^T] [\sigma_1 \mathbf{v}_1(\mathbf{w}_1 \otimes \mathbf{u}_1)^T + \lambda_1 \mathbf{x}_1(\mathbf{y}_1 \otimes \mathbf{u}_2)^T]^T$$

and are given by

$$\mu_{1,2} = \sqrt{\frac{\sigma_1^2 + \lambda_1^2 + 2\sigma_1\lambda_1c_1c_2c_3}{2} \pm \sqrt{\frac{(\sigma_1^2 + \lambda_1^2 + 2\sigma_1\lambda_1c_1c_2c_3)^2}{4} - \sigma_1^2\lambda_1^2(1 - c_1^2c_3^2)(1 - c_2^2)}}$$

with $c_1 = \mathbf{u}_1^T \mathbf{u}_2$, $c_2 = \mathbf{v}_1^T \mathbf{x}_1$, and $c_3 = \mathbf{w}_1^T \mathbf{y}_1$. If $\mu_2 > \epsilon_1$, where ϵ_1 is the highest singular value of \mathbf{R} (which can, depending on the correlation of vectors of \mathbf{X}_2 and \mathbf{V}_2 or \mathbf{W}_2 and \mathbf{Y}_2 be greater than $\max(\lambda_2, \sigma_2)$), the truncation of the SVD of $[\mathbf{T}]_{(2)}$ yields the matrix $[\tilde{\mathbf{T}}]_{(2)}$ of equation (6) and permits therefore to identify \mathbf{u}_1 and \mathbf{u}_2 using the CP decomposition.

Please note that in the special case where \mathbf{u}_1 and \mathbf{u}_2 are orthogonal, the columns of the matrices $\sigma_1 \mathbf{v}_1(\mathbf{w}_1 \otimes \mathbf{u}_1)^T$ and $\lambda_1 \mathbf{x}_1(\mathbf{y}_1 \otimes \mathbf{u}_2)^T$ in equation (4) are also pairwise orthogonal to the columns of \mathbf{R} if only $\mathbf{v}_1^T \mathbf{X}_2 = \mathbf{0}^T$ and $\mathbf{x}_1^T \mathbf{V}_2 = \mathbf{0}^T$. In this case, the conditions $\mathbf{w}_1^T \mathbf{Y}_2 = \mathbf{0}^T$ and $\mathbf{y}_1^T \mathbf{W}_2 = \mathbf{0}^T$ are thus not needed.

As a consequence, since the DIAG algorithm (Luciani and Albera, 2011) is based on a truncated SVD in one mode of the tensor, it permits to perfectly recover \mathbf{u}_1 and \mathbf{u}_2 if it is based on the mode-2 unfolding and the conditions

$$\text{C1) } \mathbf{v}_1^T \mathbf{X}_2 = \mathbf{0}^T, \mathbf{x}_1^T \mathbf{V}_2 = \mathbf{0}^T, \mathbf{w}_1^T \mathbf{Y}_2 = \mathbf{0}^T, \mathbf{y}_1^T \mathbf{W}_2 = \mathbf{0}^T, \text{ and } \mu_2 > \epsilon_1 \text{ or}$$

$$\text{C2) } \mathbf{v}_1^T \mathbf{X}_2 = \mathbf{0}^T, \mathbf{x}_1^T \mathbf{V}_2 = \mathbf{0}^T, \mathbf{u}_1^T \mathbf{u}_2 = 0, \text{ and } \mu_2 > \epsilon_1$$

are fulfilled. Similar conditions can be derived for a truncated SVD in the third mode. If the DIAG decomposition is based on the mode-1 unfolding, the truncated SVD does not change the unfolding matrix, which is already

of rank $R = 2$, and therefore does not lead to a loss of information. But contrary to the unfolding matrix of a tensor that follows the CP model, the right signal subspace of the mode-1 unfolding matrix does not have a Kronecker structure. However, this assumed structure is exploited in the following steps of the DIAG algorithm (more particularly during the JET) and its absence generally causes errors on the estimated vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$. These errors are difficult to quantify because they depend on the iterative optimization of the JEVD algorithm and their analysis is out of the scope of this paper.

S.1.2. Discrepancies from the above conditions

If the conditions on orthogonality are not fulfilled, which is usually the case in practice, the vectors \mathbf{u}_1 and \mathbf{u}_2 cannot be correctly recovered, leading to errors of the estimated vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$. For small correlation coefficients between \mathbf{v}_1 and \mathbf{X}_2 , \mathbf{w}_1 and \mathbf{Y}_2 , \mathbf{x}_1 and \mathbf{V}_2 , and \mathbf{y}_1 and \mathbf{W}_2 or correlation of \mathbf{v}_1 , \mathbf{w}_1 , \mathbf{x}_1 , and \mathbf{y}_1 with vectors that are associated with very small singular values, the errors on the estimated vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ can be regarded as negligible. In this case, the STF and STWV methods yield good results for the space or time characteristics of each patch. On the other hand, for large correlation coefficients between the singular vectors of \mathbf{M}_1 and \mathbf{M}_2 and especially in the case where the condition on the singular values ($\mu_2 > \epsilon_1$) is not fulfilled (which occurs, for example, if the singular values of \mathbf{M}_1 and \mathbf{M}_2 do not decrease quickly or if one source is much stronger than the other source), the result of the CP decomposition can be seriously perturbed (up to containing only information about one of the sources) and does not permit to obtain an adequate estimate of the vectors \mathbf{u}_1 and \mathbf{u}_2 . In this case, the

STF or STWV analysis fails.

S.1.3. Interpretation of the mathematical conditions with respect to the STF and STWV analyses

In the following, we consider three types of conditions that are involved in C1) and C2) and point out how they intervene in the STF and STWV analyses of EEG data.

$\mu_2 > \epsilon_1$ The validity of this condition depends on the one hand on the singular value profiles of the time-frequency or space-wave-vector matrices of the patches (matrices \mathbf{M}_1 and \mathbf{M}_2) and on the other hand on the source strengths. For slowly decreasing singular values, it requires the source strengths to be approximately equal whereas quickly decreasing singular value profiles enable the STF and STWV techniques to tolerate a certain difference in source strength, which may be due to different patch sizes, different patch locations, or different signal amplitudes. This is the case for the STF analysis, for oscillatory signals, where one can assume that there is one dominant frequency characteristic for each source, yielding time-frequency matrices \mathbf{M}_r with only one great singular value. In a similar way, superficial patches generate focused spatial distributions that can be described by one dominant spatial component per patch, leading to a quickly decreasing singular value profile of the space-wave-vector matrix.

$\mathbf{u}_1^T \mathbf{u}_2 = 0$ In case of the STF analysis, the vectors \mathbf{u}_1 and \mathbf{u}_2 correspond to the spatial mixing vectors of the patches. This condition thus requires the spatial mixing vectors to be uncorrelated. The correlation of the spatial mixing vectors is related to the patch distance and is generally

small for distant patches and high for close patches. For the STWV method, the source time signals are required to be uncorrelated as the vectors \mathbf{u}_1 and \mathbf{u}_2 characterize the time courses of the patch amplitudes. In practice, small correlation coefficients are usually sufficient to obtain reasonably good results (cf. previous paragraph).

$\mathbf{v}_1^T \mathbf{X}_2 = \mathbf{0}^T$, $\mathbf{x}_1^T \mathbf{V}_2 = \mathbf{0}^T$, $\mathbf{w}_1^T \mathbf{Y}_2 = \mathbf{0}^T$, $\mathbf{y}_1^T \mathbf{W}_2 = \mathbf{0}^T$ These orthogonality conditions concern correlations of the time-frequency or space-wave-vector profiles of the two patches and are difficult to interpret in practice. For the STF analysis, this is the case for sufficiently different time and frequency characteristics of two sources (for example sources with uncorrelated time signals involving different frequency bands) whereas this is achieved for sufficiently distant patches giving rise to different dominant spatial components in the case of the STWV analysis. The influence of each of these correlation coefficients also depends on the associated singular values. Quickly decreasing singular value profiles of the time-frequency or space-wave vector matrix considerably reduce the importance of a large number of correlation coefficients.

S.1.4. Theoretical analysis of selected two patch scenarios

In the following, we establish a link between the theoretical findings described above and the simulation results of the STWV-DA algorithm presented in Section ?? of the paper. To this end, we analyze what happens when applying the STWV analysis to two examples of two patch scenarios and explain the consequences on the source localization results. More particularly, we are interested in the impact that the application of the DIAG

algorithm for the CP decomposition has on the STWV tensor when the model is not exactly trilinear. As explained above, the first step of DIAG consists in truncating the SVD of, e.g., the mode-2 unfolding matrix, which ideally leads to a trilinear model where each component corresponds to one source (cf. equations (4) to (6)). In the following, we examine whether this step is successful for the STWV data of our simulation examples. This determines whether the patches are correctly separated and thus has a high impact on the performance of the source localization.

In order to avoid perturbations that are not directly related to the STWV preprocessing and would complicate the evaluation of the results, we generate realistic simulation data as described in Section ??, but without background activity or noise. Furthermore, we attribute the same signal to all dipoles that belong to the same patch. In a first step, we then compute the STWV tensors \mathbf{F}_1 and \mathbf{F}_2 separately for each of the two patches. For each of these tensors, we determine the two dominant left singular vectors of the space-wave-vector matrices (vectors \mathbf{v}_1 and \mathbf{v}_2 for tensor \mathbf{F}_1 , and \mathbf{x}_1 and \mathbf{x}_2 for tensor \mathbf{F}_2), which contain information about the spatial distribution. In a second step, we calculate the SVD of the mode-1 unfolding matrix of the combined data tensor $\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2$ and truncate it to obtain a rank-2 matrix (for $R = 2$ patches). If the condition C1) or C2) of Section ?? is fulfilled, the resulting two left singular vectors \mathbf{z}_1 and \mathbf{z}_2 should (at least approximately) span the same subspace as the vectors \mathbf{v}_1 and \mathbf{x}_1 . Otherwise, the separation of the two patches using the STWV analysis fails.

Figure 1 corresponds to the scenario of two distant sources and shows the absolute value of the interpolated spatial distributions described by the

two dominant singular vectors \mathbf{v}_1 and \mathbf{v}_2 of the patch Frsup-rost and the two dominant singular vectors \mathbf{x}_1 and \mathbf{x}_2 of the patch Occsup, as well as the left singular vectors \mathbf{z}_1 and \mathbf{z}_2 recovered from the truncated SVD of the mode-1 unfolding matrix of the tensor \mathbf{F} . Obviously, the first singular vector \mathbf{z}_1 corresponds to the dominant \mathbf{x}_1 of the patch Occsup, while the second singular vector \mathbf{z}_2 corresponds to the dominant vector vector \mathbf{v}_1 of the patch Frsup-rost. Therefore, the STWV analysis leads to a separation of the two patches and allows for an accurate localization (see Section ??).

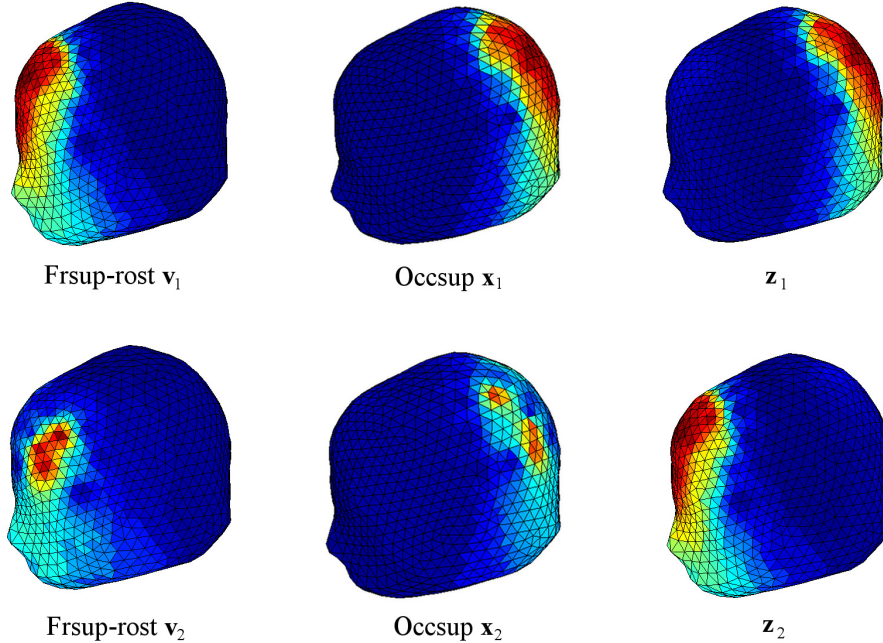


Figure 1: Dominant components of the patch Frsup-rost (left), dominant components of the patch Occsup (middle) and components recovered with the truncated SVD (right).

Figure 2 shows the corresponding interpolated spatial distributions for the scenario of deep patches MidTe and Hipp. In this case, the left singular vectors \mathbf{z}_1 and \mathbf{z}_2 look like slightly perturbed versions of the two dominant

vectors \mathbf{v}_1 and \mathbf{v}_2 of the patch MidTe, which leads to the conclusion that the patch MidTe yields observations with higher amplitudes than the patch Hipp. This means that the condition on the singular values μ_2 and ϵ_1 is not fulfilled. The slight perturbation of the vectors \mathbf{v}_1 and \mathbf{v}_2 could be explained by an additional violation of the orthogonality conditions. In short, the STWV analysis fails in this case because it loses the information about the patch Hipp. This explains the bad performance of STWV-DA for this scenario (cf. Section ??).

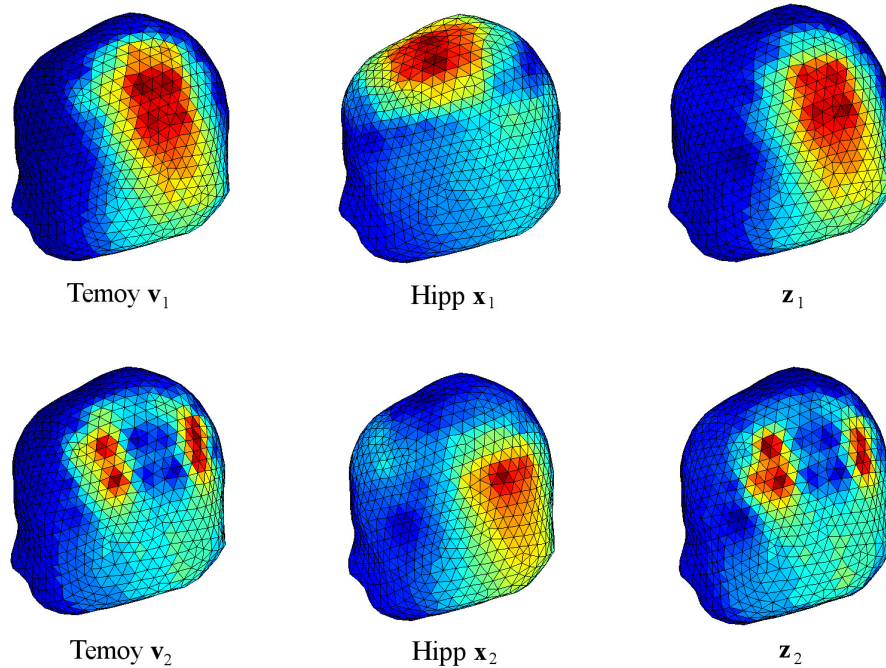


Figure 2: Dominant components of the patch Temoy (left), dominant components of the patch Hipp (middle) and components recovered with the truncated SVD (right).

S.2. Computational complexity

	Number of real-valued multiplications
STF analysis	
Tensor construction	$N_t^2 N_r N_f$
Tensor decomposition	$(12R + 7)N_r N_t N_f$ $+ (N_r + N_t + N_f)N_{\text{sw}}R(R^2 + \frac{1}{2}R - \frac{3}{2})$
STWV analysis	
Tensor construction	$2N_r' \bar{N} N_t N_k$
Tensor decomposition	$(48R + 28)N_r' N_t N_k$ $+ 56R(R - 1)N_{\text{sw}}(N_r' + N_t + N_k)$ $+ 9N_{\text{iter}}RN_r' N_t N_k$
Disk algorithm	$3RN_d D_{\text{max}} N_r$
sLORETA	
Generalized inverse comp.	$\frac{3}{2}N_r^2 N_d + \frac{1}{6}N_r^3$
Application	$N_r N_d$
cLORETA	
Generalized inverse comp.	$\frac{1}{2}N_r^2 N_d + \frac{1}{6}N_r^3$
Application	$N_r N_d$
4-ExSo-MUSIC	
Cumulant estimation + EVD	$\frac{1}{24}N_r^4 N_t + \frac{1}{6}N_r^6$
Spectrum estimation	$\frac{1}{4}N_d D_{\text{max}} N_r^3 + \frac{5}{4}N_r^2 N_d D_{\text{max}}$

Table 1: Computational complexity in terms of real-valued multiplications for the tensor-based preprocessing methods and different source localization algorithms.

References

- De Lathauwer, L., 2008. Decompositions of a higher-order tensor in block terms-part ii: Definitions and uniqueness. *SIAM J. Matrix Anal. Appl.* 30 (3), 1033–1066.
- Luciani, X., Albera, L., 2011. Semi-algebraic canonical decomposition of multi-way arrays and joint eigenvalue decomposition. In: *IEEE Proc. ICASSP. Prague, Czech Republic*, pp. 4104 – 4107.