Non-Negative Canonical Polyadic Decomposition for Tissue Type Differentiation in Gliomas

H. N. Bharath[1,2], D. M. Sima[1,2], N. Sauwen[1,2], U. Himmelreich[3], L. De Lathauwer[1,4], S. Van Huffel[1,2]

***Abstract*—Magnetic resonance spectroscopic imaging (MRSI) reveals chemical information that characterizes different tissue types in brain tumors. Blind source separation techniques are used to extract the tissue-specific profiles and their corresponding distribution from the MRSI data. We focus on automatic detection of the tumor, necrotic and normal brain tissue types by constructing a 3-dimensional MRSI tensor from *in vivo* 2D-MRSI data of individual glioma patients. Non-negative canonical polyadic decomposition (NCPD) is applied to the MRSI tensor to differentiate various tissue types. An *in vivo* study shows that NCPD has better performance in identifying tumor and necrotic tissue type in glioma patients compared to previous matrix-based decompositions, such as non-negative matrix factorization and hierarchical non-negative matrix factorization.**

***Index Terms*—Non-negative canonical polyadic decomposition, Magnetic resonance spectroscopic imaging, Glioma, l_1 regularization.**

I. Introduction

Accurate characterisation and localization of pathologic tissue types play a key role in diagnosis and treatment planning of brain tumors. The tumor region of glioblastoma multiforme (GBM) could consist of several tissue types, which represent actively growing tumor, necrosis or normal brain tissue [1]. In recent years, many advanced MR modalities such as Magnetic Resonance Spectroscopic Imaging (MRSI), perfusion-weighted MRI (PWI) and diffusion weighted MRI (DWI) are being used to characterize brain tumors and detect full tumor extent [2]. Magnetic resonance spectroscopic imaging (MRSI) is a non-invasive imaging technique which provides spectral profiles in a two- or three- dimensional voxel grid, from which the spatial distribution of metabolite concentrations can be estimated. MRSI has been successfully applied to diagnosis and prognosis of brain tumors. There are many algorithms for MRSI data analysis available in the literature that aim at tissue characterisation, tumor localization and classification. In particular, nonnegative matrix factorization (NMF) and hierarchical NMF (hNMF) have shown potential to differentiate different tissue patterns in MRSI of GBM patients [1]. However, the performance of hNMF deteriorates in the presence of artifacts because it can handle only three tissue types (tumor, necrotic and normal).

[1]Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium.

[2]iMinds Medical Information Technologies, Leuven, Belgium.

[3]Biomedical MRI Unit/Molecular Small Animal Imaging Center, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium.

[4]The Group of Science, Engineering and Technology, KU Leuven Kulak, Kortrijk, Belgium.

Higher order tensors have certain properties that are not present in a matrix [3]. In recent years there is an increasing trend to convert the data represented in a matrix to a 3rd order tensor [4]. Tensorization of the matrix is mainly motivated by the fact that tensor decompositions can be unique under mild conditions without imposing additional constraints [5], [6]. The 2-D MRSI signal can naturally be represented as a 3-way tensor $\mathcal{P}$ as shown in Fig. 1. The mode-1 and mode-2 of the tensor $\mathcal{P}$ represent the spatial dimension of the 2-D MRSI signal and mode-3 represents the spectra from all voxels.

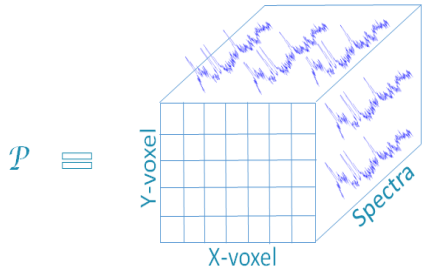


Fig. 1: 3-way spatial tensor representation of 2-D MRSI data.

To extract different tissue types from the spatial tensor $\mathcal{P}$, we can use block term decomposition (BTD) in R rank-$(L_r, L_r, 1)$ block terms [7]. The $(L_r, L_r, 1)$ BTD for a third-order tensor $\mathcal{P} \in \mathbb{R}^{I \times J \times K}$ can be written as:

$$\mathcal{P} \approx \sum_{r=1}^{R} (A_r B_r^T) \circ S_r \tag{1}$$

where S_r represents the tissue-specific spectral pattern, $A_r B_r^T$ having rank L_r represents the corresponding spatial distribution with $A_r \in \mathbb{R}^{I \times L_r}$, $B_r \in \mathbb{R}^{J \times L_r}$ and '$\circ$' represents outer product. The rank L_r for each tissue type plays an important role in the decomposition, which depends on the factors like size and shape of the tissue type. The rank L_r is not known a priori and it is difficult to estimate it from the MRSI data. Also there is no guarantee that the spatial distribution follows a low-rank structure so that it can be represented by a low rank matrix $A_r B_r^T$. Because of these problems, we developed a new tensorization of the 2-D MRSI signal which allows to exploit the low-rank structure.

In this article, we focus on automatic detection of the tumor, necrotic and normal tissue types by constructing a 3D-MRSI tensor from in vivo 2D-MRSI data and applying non-negative canonical polyadic decomposition (NCPD) with common factor in mode-1 and mode-2 on the tensor to differentiate various tissue types. Preliminary studies have

been previously presented in [8]. The paper is organized as follows: Section II explains the MRSI data acquisition protocol and the preprocessing steps that are performed. In Section III, the construction of the MRSI tensor and the NCPD algorithm for tissue type differentiation is explained. The performance evaluation of the proposed NCPD algorithm in comparison with one-level NMF and hNMF using short-echo time (TE) ^{1}H 2D-MRSI datasets from glioma patients is done in Section IV. Discussions are presented in Section V and finally the paper is concluded in Section VI.

II. MATERIALS

A. Data acquisition

A total of 28 (22 grade IV, 3 grade II and 3 grade II and III) 2D-^{1}H MRSI data was acquired from 17 glioma patients on a 3T MR scanner (Achieva, Philips, Best, The Netherlands) at the University Hospital of Leuven using the protocol [9]: A point-resolved spectroscopy (PRESS) sequence was used as the volume selection technique with a bandwidth of 1.3kHz for the conventional slice-selective pulses; repetition time (TR)/ echo time (TE): 2000/35ms; Field of view (FOV): 160×160mm^2; maximal volume of interest (VOI): 80×80mm^2; slice thickness: 10mm; acquisition voxel size: 10×10mm^2; reconstruction voxel size: 5×5mm^2; receiver bandwidth: 2000Hz; samples: 2048; number of signal averages: 1; water suppression method: multiple optimizations insensitive suppression train [10]; first- and second-order pencil beam shimming; parallel imaging: sensitivity encoding with reduction factors of 2 (left-right) and 1.8 (anterior-posterior); scan time: around 3min 30s. Automated pre-scanning optimized the shim in order to yield a peak width consistently under 20Hz full-width half-maximum (FWHM). The study and the experimental procedures involving human subjects have been approved by the ethical committee of the institute.

B. Data processing

The raw MRSI data were exported from the Philips platform after performing the following post-processing steps: zero filling in k space, transformation from k space to normal space, automatic phase correction and eddy current correction. The residual water component was removed from the MRSI data using Hankel Lanczos singular value decomposition with partial reorthogonalization (HLSVD-PRO) [11]. A model order of 30 and a passband of 0.25 to 4.2 parts per million (ppm) was used in HLSVD-PRO algorithm. After removing the water component, baseline correction and baseline offset correction was performed. All the pre-processing was done using the Matlab (The MathWorks, Inc., Natick, Massachusetts, United States) based software, SPID [12].

The spectra were aligned in frequency using a simulated reference spectrum, which was generated using the parameters given in [13]. The complex-valued pre-processed spectra were truncated to the region 0.25-4.2 ppm and the truncated spectra were normalized to unit norm (l_2). Voxels outside the MRSI PRESS excitation volume are excluded from the analysis.

III. METHOD

A. MRSI tensor construction

For each voxel in the MRSI grid, a reduced real-valued spectrum X is constructed from the corresponding complex-valued pre-processed spectrum. Elements of the vector X are obtained by moving an overlapping window over the spectrum, where the i^{th} element of X is the sum of squares of absolute value of all the elements in the i^{th} window segment,

$$X(i) = \sum_{j=1}^{L} s_i(j) s_i^*(j) \tag{2}$$

s_i is the spectrum at the i^{th} segment, s_i^* is its complex conjugate and L is the length of the window segment. Fig. 2a shows the construction of a vector X from the spectrum. The resulting vector X can be considered as a denoised and reduced-length version of the original spectrum. The window length L is chosen such that it covers the widest peak in the spectra and step value is chosen as $\frac{L}{4}$. Use of vector X has the following advantages:

1) It reduces the length of spectra without losing vital information required for tumor tissue type differentiation.
2) It gives more weight to the peaks and makes the signal smoother and non-negative.

A 3-way MRSI tensor $\mathcal{T}$ is constructed by stacking XX^T from all the voxels in the MRSI grid as shown in Fig. 2(b).

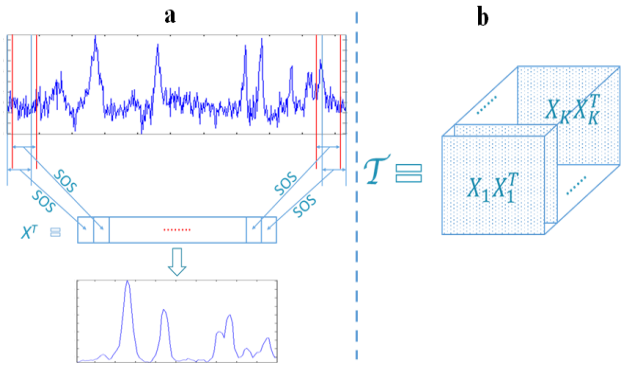


Fig. 2: (a) Construction of reduced spectrum, X from the pre-processed spectra. SOS: sum of squares. (b) Construction of the MRSI tensor $\mathcal{T}$ from the reduced spectra, X. K is the total number of voxels in the 2D MRSI excitation volume.

B. Non-negative CPD

Non-negative canonical polyadic decomposition (NCPD) is a tensor decomposition method, where the tensor is decomposed into a sum of rank-one tensors with non-negativity constraints on the factor matrices [14].

$$\mathcal{X} \approx [\![ABC]\!] \equiv \sum_{r=1}^{R} a_r \circ b_r \circ c_r, \qquad A,B,C \geq 0 \tag{3}$$

where $A = [a_1, a_2, ..., a_R] \in \mathbb{R}_+^{I \times R}$, $B = [b_1, b_2, ..., b_R] \in \mathbb{R}_+^{J \times R}$ and $C = [c_1, c_2, ..., c_R] \in \mathbb{R}_+^{K \times R}$ are non-negative factor matrices. R is the rank, defined as the number of rank-one terms.

In the MRSI tensor $\mathcal{T}$, the frontal slices are symmetrical, therefore we constrain the frontal slices of each NCPD rank-one term to be symmetric. To maintain symmetry, a common factor matrix is used for mode-1 and mode-2 in the NCPD as shown in Fig. 3. After performing the NCPD on the MRSI tensor $\mathcal{T}$ we obtain two factor matrices S and H, where S represents the tissue-specific patterns of the reduced spectra and H represents the spatial distribution of each tissue type.

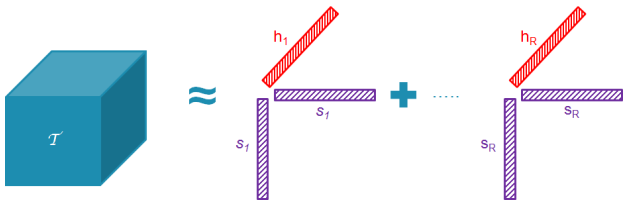


Fig. 3: Non-negative CPD of MRSI tensor $\mathcal{T}$: MRSI tensor $\mathcal{T}$ is decomposed into R rank-1 tensors. Common factor S is used in mode-1 and mode-2 to maintain symmetry of frontal slices. Each s_i gives a tissue-specific reduced spectral pattern and the corresponding h_i gives the spatial distribution of the respective tissue type, upon reshaping. Non-negativity of S and H is imposed in the decomposition.

Each rank-one term obtained from the NCPD of the MRSI tensor $\mathcal{T}$ is expected to correspond to a particular tissue type. Here, we assume that in most of the voxels the spectra belong to a particular tissue type, with the exception of a few voxels whose spectra may contain a mixture of at most two tissue types. Therefore, the factor matrix H will be sparse, meaning that each row will mostly have only one high value. A further refinement of the NCPD method exploits the sparsity assumption in the factor matrix H by imposing a l_1 regularization on it. The NCPD with l_1 regularization (NCPD-$l1$) can be written as

$$[S^*,H^*] = \arg\min_{S\geq 0,H\geq 0} \|\mathcal{T} - \sum_{i=1}^{R} S(:,i)\circ S(:,i)\circ H(:,i)\|_2^2 + \lambda\|Vec(H)\|_1 \quad (4)$$

where S and H are the aforementioned factor matrices and λ is the parameter which controls the sparsity. In this work, the tensor decomposition was performed using the Tensorlab Matlab package [15]. Non-negativity constraints, common factors to maintain symmetry and l_1 regularization are applied using the structured data fusion method [16] available in Tensorlab.

C. Spectral recovery and non-negative least squares

The NCPD of the MRSI tensor $\mathcal{T}$ gives the factor matrix S, which contains the reduced spectral patterns specific to different tissue types. However, it is desirable to recover tissue-specific spectral sources as vectors of the same length as the pre-processed MRSI spectra, which are more interpretable since they can be directly compared to the original spectra. To this end, a least squares problem is solved, and the matrix of spectral sources representing the tissue-specific spectral patterns W is:

$$W = (H^{\dagger}Y^T)^T \quad (5)$$

where $H^{\dagger}$ is the Moore-Penrose pseudoinverse of the NCPD factor matrix H and Y is the matrix containing the unit normalized spectra from all voxels as its columns. The source spectra W can be estimated as complex-valued or magnitude vectors by using complex-valued or magnitude spectra Y in equation (5), respectively.

In the NCPD of MRSI tensor $\mathcal{T}$, the factor matrix H corresponds to the weights of the linear combination of $S(:,i)S(:,i)^T$ and not the linear combination of source spectra W. Also, the tensor is constructed using the normalised spectra, therefore voxels having relatively small values compared to other spectra will also get enhanced and will have higher abundances. The abundances of the sources in the original unnormalised spectra are more meaningful and represent the true distribution. Also we want the abundances to represent the weights in the linear combination of source spectra W. To address these problems, spatial distributions, H_D of the different tissue types is calculated using non-negative least squares with l_1 regularization:

$$H_D(:,i) = \arg\min_{x\geq 0} \|Wx - Y_{un}(:,i)\|_2^2 + \lambda_1\|x\|_1 \quad (6)$$

where, $H_D(:,i)$ is the distribution of source spectra in each voxel, W contains the estimated source spectra and $Y_{un}(:,i)$ is the original unnormalised spectrum of each voxel. The problem in equation (6) is solved for all the voxels in the MRSI grid using a Matlab based large-scale l_1-regularized least squares problem solver [17]. When the estimated source spectra W and the MRSI voxel spectra are complex-valued, the real and imaginary part are concatenated to form a single real-valued matrix, which is then used as input to the non-negative least squares problem with l_1 regularization. Tissue distribution maps are obtained by reshaping the rows of H_D to the 2D-MRSI grid.

D. Initialization

The NCPD algorithm needs initial values for S and H and the non-negative least squares [17] algorithm needs initial values for H_D. Initializing S, H and H_D with uniformly distributed pseudorandom numbers between 0 and 1 gives good solution, but the results are not exactly the same between different runs. Although the solutions are similar between different runs, poor initial values may result in sub-optimal solutions. To find good initialization values, first we take the singular value decomposition (SVD) of the matrix Y, $Y = U\Sigma V^H$, where the columns of the matrix Y are the complex spectra from each voxel. Reduced spectra S_{init} are constructed from R dominant left-singular vectors as explained in Section III-A and are used as initial value for S in the NCPD algorithm. Initial values for H are obtained through least squares as $H_{init} = (S_{init}^{\dagger}M)^T$, where M is a matrix whose columns are the reduced spectra X from each voxel. Least squares may introduce negative values in H_{init}. However, these negative values are typically

rare and small in amplitude. Moreover, the NCPD algorithm in Tensorlab can handle negative initial values. A vector of all ones was used to initialize H_D in the final non-negative least squares step.

E. Source number estimation

The NCPD algorithm needs the number of sources (i.e. decomposition rank) as input. Estimating the rank from the input spectra/ tensor $\mathcal{T}$ is a difficult problem. The literature on estimation of decomposition rank from the tensor is limited. Tensorlab package [15] has a method *rankest*, which estimates the rank based on the L-curve of the number of rank-one terms in a CPD. However, this method gives good results when the noise is low or when the decomposition is exact, which is not the case for our MRSI tensor $\mathcal{T}$. The estimated rank R from this method is much higher than the required number of sources for a good tissue type differentiation. In [18] a Bayesian model based on automatic relevance determination is proposed for NMF, which also estimates the model order R along with non-negative factor matrices. In this method, the estimated model order is dependent on the choice of the dispersion parameter, which represents the tradeoff between the data fidelity and the regularization terms. Selecting the optimal dispersion parameter for each MRSI dataset is difficult and it is as hard as selecting the rank itself. In this paper we will use the covariance matrix based approach to estimate the number of sources.

Let A be the data matrix of size $K \times N$ where the rows represent K spectra of length N from all the voxels in the MRSI grid. Then the $K \times K$ covariance matrix is estimated as $C = \mathrm{E}[(A - \mathrm{E}[A])(A - \mathrm{E}[A])^T]$, where E[] is the expectation operator. The eigenvalues of the covariance matrix C are denoted by $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_K$. The number of sources is estimated as the minimum number R such that the cumulative sum of eigenvalues is greater than 99% of the sum of all eigen values.

$$R^* = \arg\min_R [\frac{\sum_{i=1}^{i=R} \lambda_i}{\sum_{i=1}^{i=K} \lambda_i} \geq 0.99] \tag{7}$$

When the data matrix A is constructed from the original complex valued spectra, the estimated number of sources is high. Therefore, we use the reduced spectra in A to calculate the covariance matrix C. Reduced spectra suppress noise and small variations present in the original spectra resulting in fewer eigenvalues significantly larger than zero in the covariance matrix C. Therefore, using reduced spectra to estimate R provides a good estimate for the number of sources in many MRSI datasets. However, when the MRSI data contain spectra of less quality or having more artifacts the estimated number of sources is still too high. To overcome this problem we incorporate prior knowledge about the maximum number of sources (includes tissue types + artifacts) for estimating the number of sources. Let the maximum number of sources be P. If the estimated R from (7) is less than or equal to P ($R \leq P$), then the number of sources is set to R and is used in the NCPD algorithm. When the estimated R from (7) is greater than P ($R > P$), only the largest $P+1$ eigenvalues of C are retained and the remaining ones are set to zero. Then the number of sources is estimated as in equation (7) with K set to $P+1$, i.e the set of eigenvalues is reduced to the largest $P+1$ values only.

F. Source and distribution correlation

The performance evaluation of the algorithms in the *in vivo* study was analyzed using two measures:

1) Source correlation: In this paper we have defined two types of source correlation:
   (a) Source correlation Type I ($SC1$): The source correlation is calculated as Pearson's linear correlation coefficient between the estimated source spectrum and the tissue-specific spectrum based on expert labeling of the *in vivo* MRSI voxels [1]. The tissue-specific spectrum based on expert labeling (*src_expt*) for a particular tissue type is computed as the average of all the spectra from the voxels labelled by the expert as belonging to that tissue type. The construction of expert spectra, *src_expt* for the tumor tissue is shown in Fig. 4.

$$src_expt = \frac{Y_1 + Y_2 + ... + Y_n}{n}$$
$$SC1 = r(W(:,T), src_expt)$$

   where $SC1$ is the Type I source correlation, $W(:,T)$ is the estimated source spectrum corresponding to a particular tissue type, *src_expt* is the expert labeled spectrum corresponding to that particular tissue type, $Y_1, Y_2, ..., Y_n$ are the spectra from the voxel that are marked by the expert as belonging to that particular tissue type and r is Pearson's linear correlation coefficient.
   (b) Source correlation Type II ($SC2$): First Pearson's linear correlation coefficients are calculated between the estimated source spectrum and all the spectra in the voxels that are marked by the expert as belonging to a particular tissue type. Then the source correlation $SC2$ is calculated by taking the average of Pearson's linear correlation coefficients.

$$SC2 = \frac{r(W(:,T), Y_1) + ... + r(W(:,T), Y_n))}{n}$$

   where $SC2$ is the Type II source correlation, $W(:,T)$ is the estimated source spectrum corresponding to a particular tissue type.
2) Distribution correlation (DC): Distribution correlation is calculated as Pearson's linear correlation coefficient between the estimated distribution map corresponding to a particular tissue type and the distribution map based on expert labeling (*dist_expt*). For each tissue type, a distribution map based on expert labeling is obtained by using values equal to the l_2 norm of the corresponding spectra for all voxels labeled as a certain tissue class, and values of 0 for the other voxels as shown in Fig. 4.

$$DC = r(H_D(:,T), dist_expt)$$

   where DC is the distribution correlation, $H_D(:,T)$ is the estimated distribution vector corresponding to a particular

tissue type, *dist_expt* is the expert labeled distribution vector corresponding to that particular tissue type.

Because of heterogeneity, the tumor tissue is modelled by more than one source spectrum in many patients. In this case, the average of source spectra and sum of the corresponding distribution maps are used in the calculation of $SC1$ and DC, respectively. Whereas, for $SC2$, only the maximum correlation among the source spectra is retained for averaging. In NCPD algorithm the estimated source spectra and the MRSI voxel spectra are complex signals. Therefore, the real and imaginary part of the complex spectra are concatenated to form a real signal which is then used in the calculation of correlation. Since absolute spectra were used in NMF and hNMF, source correlation was calculated on the spectra directly.

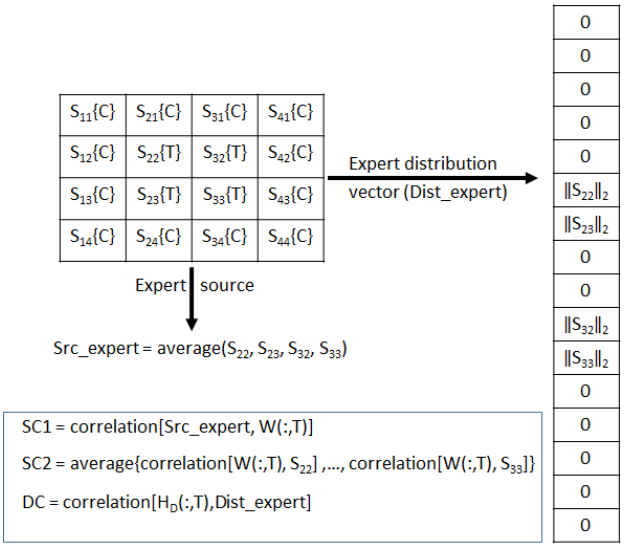
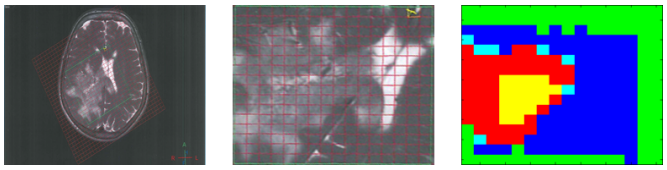


Fig. 4: Generation of expert labeled tissue-specific (Tumor) spectrum and distribution vector. Calculation of source and distribution correlation is shown in the box. S_{ij} is the spectra at i^{th} column and j^{th} row. The tissue type T-tumor or C-control is shown between braces.

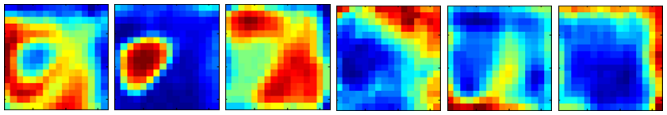
IV. Results on Brain Tumor Dataset

To test the feasibility of spatial tensor representation as shown in Fig.1 in tissue type differentiation, we constructed a spatial tensor $\mathcal{P}$ using magnitude spectra and applied $(L_r, L_r, 1)$ BTD on the spatial tensor for one high grade (grade IV) MRSI dataset. The spatial tensor $\mathcal{P}$ is decomposed into 6 rank $(L_r, L_r, 1)$ terms with rank $L_r = 10, 10, 9, 8, 5, 5$. The corresponding distribution maps of the different sources are shown in Fig. 5b. Non-negative constraints are applied to the source mode (mode-3) only. The rank L_r is chosen manually by trying different combinations and selecting the one which gives the best results. For comparison, we have shown the distribution maps obtained from the MRSI tensor $\mathcal{T}$ (shown in Fig.2) of the same dataset using NCPD-l_1 algorithm with 6 sources in Fig. 5c. The MRSI grid superimposed on anatomical image and the expert labelling is shown in Fig. 5a. Comparing the distribution maps in Fig. 5b and 5c with the expert labelling in Fig. 5a, we can observe that the tissue type differentiation is not good using a spatial tensor representation with $(L_r, L_r, 1)$ BTD compared to an XX^T based tensor representation with NCPD. The output of spatial tensor with $(L_r, L_r, 1)$ BTD is sensitive to choice of L_r and choosing a suitable L_r for all the R sources is difficult. Therefore the results are demonstrated for only one high grade glioma patient.

(a) MRSI voxel grid and expert label



(b) Distribution maps: $(L_r, L_r, 1)$ BTD of spatial tensor $\mathcal{P}$



(c) Distribution maps: NCPD of MRSI tensor $\mathcal{T}$

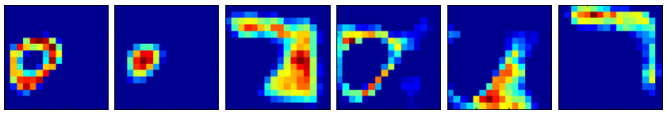


Fig. 5: (a) Left to right, First image: T2-weighted anatomical MR image of a brain tumor with areas of necrosis. Second image: voxels within the MRSI excitation volume superimposed on anatomical image. Third image: expert labeling, where yellow indicates necrotic, red indicates tumor, dark blue indicates normal, light blue indicates normal/tumor and green indicates spectra of poor quality. (b) Tissue distribution maps obtained from an $(L_r, L_r, 1)$ block term decomposition of spatial tensor $\mathcal{P}$. First three images from the left correspond to tumor, necrotic and normal tissue distribution respectively. (c) Tissue distribution maps obtained using the NCPD-l_1 algorithm on the MRSI tensor $\mathcal{T}$. First three images from the left correspond to tumor, necrotic and normal tissue distribution respectively.

In order to evaluate the performance and validate the tissue differentiation ability, three algorithms, NCPD (with and without l_1 regularization), single stage NMF and hNMF were applied on 28 *in vivo* ^{1}H MRSI datasets (22 grade IV, 3 grade II and 3 grade II and grade III) from 17 patients with gliomas. The Type I, Type II source correlation and the distribution correlation for tumor and necrotic tissue obtained from NCPD without regularization, NCPD with l_1 regularization, single stage NMF and hNMF are shown in Fig. 6. From Fig. 6 it is clearly evident that source and distribution correlation values are higher and less scattered when using NCPD-l_1 compared to single stage NMF and hNMF. The NCPD algorithm is unable to extract tumor tissue in 2 out of 28 datasets, whereas the single stage NMF and hNMF algorithms do not estimate tumor tissue in 5 datasets. The correlation values of the tissue types which are not recovered are set to zero (Fig. 6). A summary of the results i.e, mean, standard deviation (std dev), median,

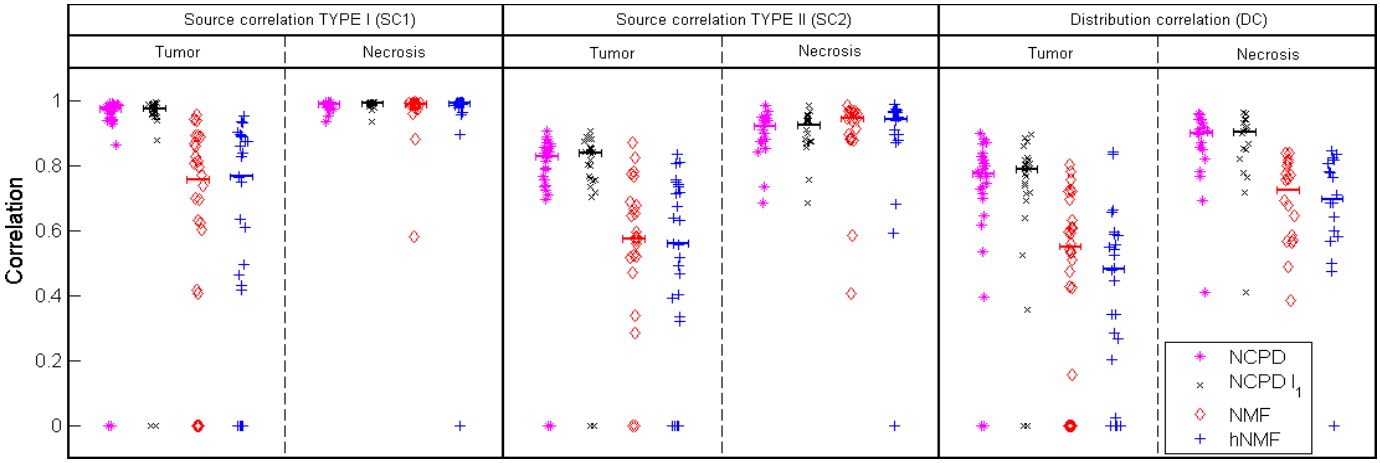


Fig. 6: Source and distribution correlation from 28 MRSI datasets. The plot shows Type I source correlation, Type II source correlation and distribution correlation values using NCPD without regularization (star:'*'), NCPD with l_1 regularization (times:'×'), single stage NMF (diamond:'◊') and hNMF (plus:'+') algorithms. The solid horizontal lines with whiskers in each correlation plot represent the median of the respective correlation. Zero correlation values indicate that the specific tissue type was not recovered.

TABLE I: Mean, standard deviation, median, MAD and range of source correlation (Type I and Type II) and distribution correlation for 28 glioma datasets. The results are shown on tumor and necrotic tissue type for NCPD without regularization, NCPD with l_1 regularization, single stage NMF and hNMF algorithms. P-values indicate the statistical significance of higher median for the NCPD-$l1$ algorithm (* indicates statistical significance for p < 0.01)

		Spectral correlation Type I				Spectral correlation Type II				Distribution correlation			
		NCPD	NCPD-l_1	NMF	hNMF	NCPD	NCPD-l_1	NMF	hNMF	NCPD	NCPD-l_1	NMF	hNMF
Tumor	Mean	0.891	0.898	0.619	0.617	0.752	0.759	0.496	0.496	0.699	0.702	0.477	0.403
	std dev	0.264	0.265	0.339	0.346	0.229	0.231	0.281	0.287	0.234	0.236	0.273	0.269
	median	0.975	0.978	0.757	0.767	0.829	0.841	0.574	0.560	0.775	0.789	0.552	0.482
	MAD	0.015	0.015	0.134	0.146	0.041	0.041	0.111	0.179	0.057	0.046	0.135	0.178
	Range	0.994-0.866	0.997-0.878	0.960-0.406	0.956-0.417	0.908-0.696	0.908-0.704	0.871-0.286	0.837-0.321	0.900-0.396	0.896-0.357	0.803-0.155	0.844-0.024
	p-value	0.1753	-	0.0000*	0.0000*	0.3269	-	0.0000*	0.0000*	0.4745	-	0.0000*	0.0000*
Necrosis	Mean	0.984	0.988	0.960	0.929	0.896	0.899	0.891	0.855	0.852	0.855	0.687	0.661
	std	0.017	0.015	0.098	0.233	0.079	0.076	0.151	0.238	0.132	0.132	0.134	0.203
	median	0.991	0.993	0.991	0.992	0.922	0.926	0.946	0.945	0.900	0.905	0.726	0.698
	MAD	0.005	0.004	0.003	0.004	0.037	0.037	0.025	0.032	0.044	0.045	0.089	0.114
	Range	0.998-0.935	0.998-0.936	0.996-0.581	0.998-0.897	0.987-0.686	0.987-0.686	0.987-0.405	0.989-0.593	0.961-0.411	0.967-0.409	0.840-0.384	0.847-0.473
	p-value	0.2238	-	0.075	0.3879	0.3639	-	0.0844	0.1965	0.4622	-	0.0001*	0.0001*

median absolute deviation (MAD) and range is shown in Table I. In case of tumor tissue, NCPD-l_1 has the highest mean and median values for source and distribution correlation. To check whether there is a significant increase in the median, we have performed a one-sided Wilcoxon rank sum test with 1% significance level ($\alpha = 0.01$) [19]. The Wilcoxon rank sum test was performed between the correlations obtained from NCPD-l_1 and other algorithms and the corresponding p-values are shown in Table I. There was a significant increase in the median of $SC1$, $SC2$ and DC from NCPD-$l1$ compared to single stage NMF and hNMF, which is evident from the p-values (p < 0.01). However, the increase in the median was not significant compared to NCPD without regularization (p > 0.05). In case of necrotic tissue, the performance of all the algorithms is good. NCPD-l_1 has significantly higher mean and median values compared to single stage NMF and hNMF for DC (p < 0.01). On the other hand, single stage NMF and hNMF has slightly better median values for $SC2$ but the increase is not significant (p > 0.01). For the necrotic sources NCPD-l_1 has slightly better median values for correlation than NCPD (no regularization), but the differences are not significant (p > 0.01).

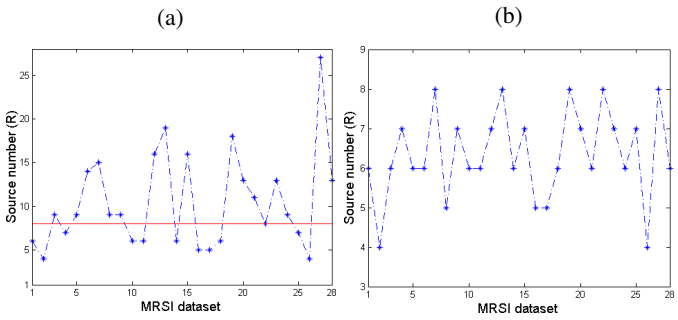


Fig. 7: (a) Estimated number of sources from the covariance matrix for 28 MRSI datasets without using the maximum number of tissue types prior knowledge. Red line indicate the maximum number of tissue types used in the analysis ($P = 8$) (b) Estimated number of sources by constraining the maximum number of tissue types to $P = 8$ for 28 MRSI datasets.

For all the MRSI datasets, the rank was automatically estimated using the method explained in section III-E. Fig.

7a shows the estimated ranks for 28 MRSI datasets without using any prior knowledge and Fig. 7b shows the estimated ranks after incorporating the prior knowledge. From Fig 7 we can observe that the estimated rank was higher in many MRSI datasets and after applying the maximum possible tissue types prior knowledge in the second stage the estimated ranks are reduced. The tissue types are assigned to the sources manually by visualizing the estimated distribution maps, expert labelling and the estimated source spectra.

The result of an *in-vivo* example is shown in Fig. 8. NCPD-l_1, single stage NMF and hNMF methods are applied to a 16×16 voxel grid shown in the second row of Fig. 8a. The spectra which are truncated to the region 0.25-4.2 ppm are of length 517. The reduced spectra are constructed using a window length $L = 20$ and the window is moved with a step-size of 5 samples. Therefore the size of the MRSI tensor $\mathcal{T}$ with 16×16 ($K = 256$) voxels is $100 \times 100 \times 256$. For this dataset the number of sources was estimated as $R = 7$, the same rank was used for the single stage NMF algorithm. The seven sources and their corresponding distribution maps obtained from NCPD-l_1 and NMF methods are shown in Fig. 8b, 8c, 8d and 8e. Using the hNMF method, only three sources are obtained as shown in Fig. 8f and 8g. Fig. 8g shows that the hNMF method identifies the normal and necrotic ($SC1 = 0.9971$) tissue properly but fails to recover the tumor tissue. Single-stage NMF identifies normal and necrotic tissue, but only the necrotic source is good ($SC1 = 0.9941$) and the normal source deviates a lot from the expert as shown in 8d (first row). In the single stage NMF method the recovered tumor tissue ($SC1 = 0.6041$) and its corresponding distribution ($DC = 0.4262$) are bad and it is difficult to identify it as tumor tissue from the source spectrum. The NCPD method identifies all three tumor ($SC1 = 0.9875$), necrotic ($SC1 = 0.9854$) and normal tissue types. Fig. 8b and 8c show that the estimated tissue sources and their corresponding spatial distribution are accurate when compared to expert labeling. In this example we have estimated seven sources and their corresponding distributions from rank-7 NCPD-l_1. Three sources correspond to tumor, necrosis and normal tissue type, the other four sources correspond to artifacts (Fig. 8b and 8c: 7^{th} row) and spectra from the outer edges of the voxel grid (Fig. 8b and 8c: 4^{th}, 5^{th} and 6^{th} row) which are contaminated by the chemical shift displacement artifact.

V. Discussion

The 2D-MRSI data can be directly represented as a third-order tensor. A $(L_r, L_r, 1)$ BTD based approach can be used to extract the different tissue types from this spatial tensor. The problem with this approach is that it is difficult to find the rank L_r of the factor matrices. The rank L_r is patient specific and it is different for different tissue types. Even when the ranks L_r are approximately known, this method does not perform better than single stage NMF or hNMF algorithms. This has motivated us to find a new way to represent 2-D MRSI data in a tensor.

In this paper we have proposed a method to represent the 2D-MRSI data in a tensor using a reduced format of the spectra. A novel tissue type differentiation algorithm based on non-negative canonical polyadic decomposition with l_1 regularization was developed. This study explored the feasibility and efficiency of the proposed algorithm (NCPD-$l1$) in recovering the normal, tumor and necrotic tissue patterns for patients with glioma using short-TE MRSI data. The previous matrix-based algorithms, NMF and hNMF failed consistently in extracting tissue-specific spectral patterns. NMF failed because the tumor spectral profile is not sufficiently uncorrelated from a linear combination of other tissue patterns: normal and necrosis [1]. Also, sometimes the NMF algorithm extracts individual peaks as sources and these sources do not represent the tissue-specific spectral patterns. The problem with the hNMF [1] is that the algorithm is designed for a maximum of three sources and cannot handle artifacts. Therefore, in [1] the voxels at the outer edge of the PRESS excitation volume are removed to minimize the effect of artifacts. By doing this we can lose the voxels belonging to clinically relevant tissue types. In the MRSI grid shown in Fig. 8, if we remove 2 or 3 outer rows or columns of voxels the necrotic tissue is almost lost. Also, due to heterogeneity of the tissue some datasets require more than one source to model that tissue type. In this case, hNMF fails to model all the tissue types with only 3 sources.

The hNMF algorithm is mainly designed to handle GBMs. When the hNMF algorithm is used on low grade gliomas, the second stage of hNMF is not applied and hNMF reduces to single stage NMF with two sources. NMF with more than two sources performs better compared to hNMF in MRSI data which does not contain necrotic tissue type (low grade gliomas). In high grade gliomas containing necrotic tissue type, hNMF performs better than single stage NMF because the second stage in the hNMF separates tumor and necrotic tissue type. In NCPD, MRSI data with more tissue types as in high grade glioma with necrotic tissue can be modelled using a higher number of sources and low grade gliomas with less artifacts can be modelled using a lower number of sources. The proposed NCPD algorithm can better separate tumor and necrotic tissue type than hNMF in high grade gliomas and better separate tumor tissue from other ones in low grade gliomas than NMF.

We have proposed an initialization scheme for the NCPD algorithm based on SVD of the complex-valued spectra and spectral reduction of singular vectors. Random initialization is used in most of the non-negative tensor factorization applications. In non-negative RESCAL tensor factorization [20], the factor matrices are initialized using an NMF initialization method Non-negative Double Singular Value Decomposition method (NNDSVD) [21]. When NCPD is initialised using NNDSVD on the reduced specta, the initial factor matrix S_{init} contains many zero values which are also retained in the decomposed factor matrix S. Therefore, the estimated reduced source spectra are unrealistic and deviate from the actual tissue-specific reduced spectra. Other initialization methods based on clustering will have more realistic source spectra compared to SVD-based methods. The problem with these methods is that they require some initialization and are computationally intensive [22], [23]. Although our initialization method is based on SVD, it does not suffer from too many

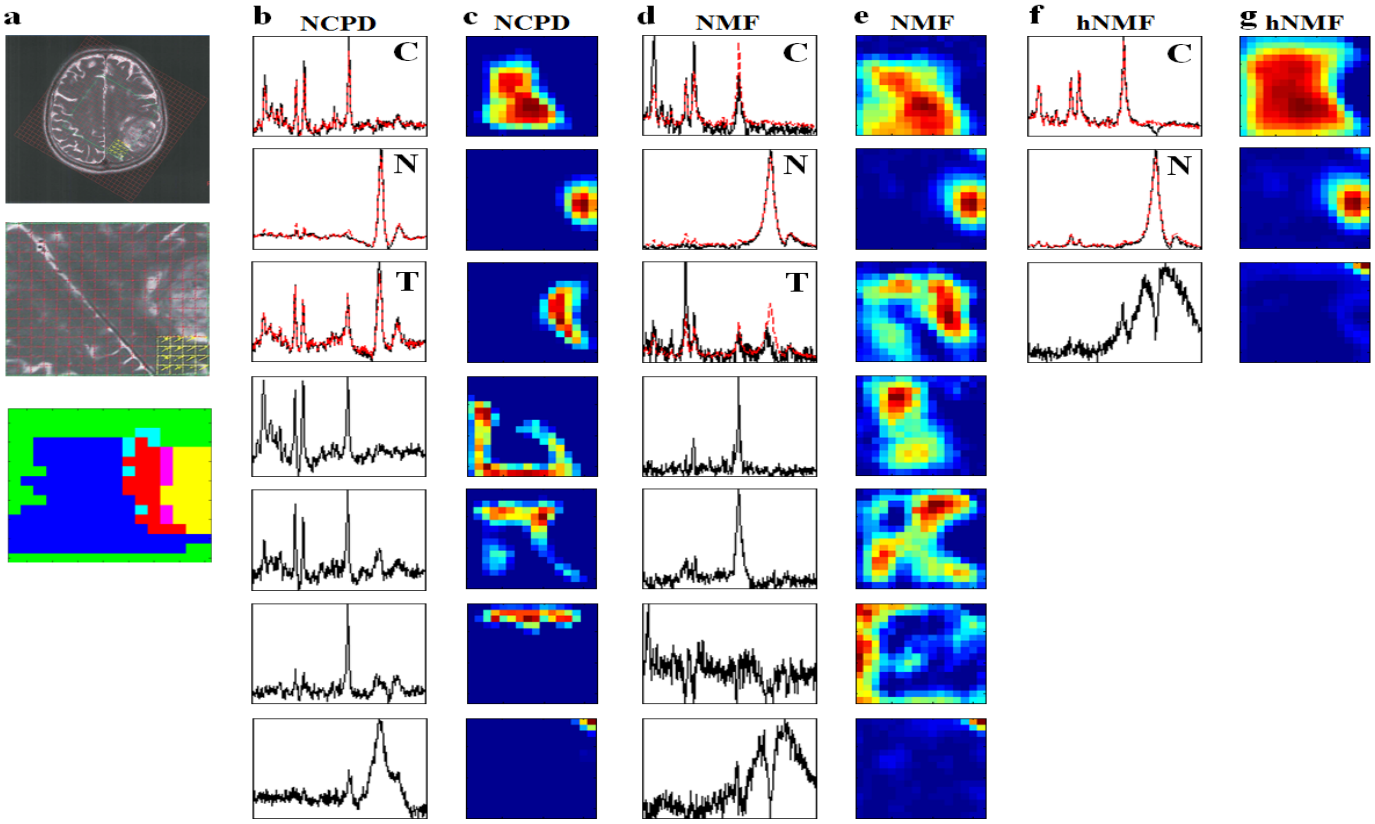


Fig. 8: Tissue pattern differentiation using ^{1}H MRSI: C, T and N represent normal, tumor and necrosis, respectively. (a) First row: T2-weighted anatomical MR image of a brain tumor with areas of necrosis. Second row: voxels within the MRSI excitation volume superimposed on anatomical image. Third row: expert labeling, where yellow indicates N, red indicates T, magenta indicates T/N, dark blue indicates C, light blue indicates C/T and green indicates spectra of poor quality. (b, c) results of NCPD. (b) The recovered sources from the NCPD-l_1 method are shown in black. First three rows represent C, N and T spectral sources in black, with tissue-specific spectra based on expert labeling overlaid in red. The remaining four rows represent artifacts and spectra from outer edges. (c) Distribution maps corresponding to spectral profiles in (b). (d, e) Results of single-stage NMF. (d) The recovered sources are shown in black, overlaid with the expert-based tissue-specific spectra in red. First three rows show normal, necrotic and tumor spectra and the remaining rows show other spectra obtained using rank-7 NMF. (e) Distribution maps corresponding to (d). (f, g) Results of hNMF. (f) Recovered sources shown in black and expert-based tissue-specific spectra in red. First two rows show control and necrotic spectra. (g) Distribution maps corresponding to (f).

zero values in the source initialization because the SVD is applied on the complex-valued spectra and the singular vectors are made positive by constructing the reduced spectra from them. The initialized sources are more realistic and close to the reduced spectra found in the MRSI voxels.

The number of sources for the NCPD algorithm is estimated using the eigenvalues of reduced spectra covariance matrix. The number of sources is overestimated when no prior knowledge is used as shown in Fig. 7a. Use of prior knowledge about the maximum number of tissue types prevents this overestimation and results in better number of sources estimation as shown in Fig. 7b. In this method we have used a cut-off of 99% on the cumulative sum of eigenvalues of C. Whereas, information theoretic criteria based methods such as in [24], [25] can determine the number of sources adaptively without the need for a cut-off value. However, these methods do not perform well in the presence of artifacts or when a linear model is not strictly satisfied. These methods highly overestimate the number of sources in our MRSI datasets.

The advantage of this tensor method is that the construction of the MRSI tensor couples the peaks in the spectra because of the XX^T in the frontal slices. Therefore, in the spectral sources obtained from the NCPD algorithm the peaks will be coupled, i.e. we will not get individual peaks as sources. The difference in the results between NCPD without regularization and NCPD-l_1 is negligible because the construction of the tensor and the extra sources already introduce sparsity in H. But NCPD-l_1 algorithm gives more stable results and sometimes models the tissue types with less sources. Also, the computational time is much less in NCPD-l_1 compared to NCPD (without regularization) as it converges in fewer iterations.

VI. Conclusion

The NCPD-l_1 algorithm outperforms the existing tissue type differentiation methods based on NMF and hNMF. The worse performance of the hNMF is due to the fact that the voxels in the outer edge of MRSI excitation volume are included in the assessment. By contrast, NCPD can account for artifacts and bad voxels present in the outer edges because more sources ($R \geq 3$) are used in the decomposition. NCPD is also able to separate artifacts from tissue sources, but NMF fails to separate these properly even after using more sources in the decomposition. The NCPD algorithm has the potential to replace the hNMF method in unsupervised nosologic imaging for brain tumors [26], which can be used as a tool to assist brain tumor diagnosis. Recently, instead of using only the MRSI signal, a multiparametric (MRSI, cMRI, DWI, PWI) approach based on a modified hierarchical non-negative matrix factorization (hNMF) has been used to characterize brain tumor heterogeneity [27]. In future research we would like to extend our proposed NCPD algorithm to multiparametric data and explore whether it has the potential to improve over hNMF [27].

ACKNOWLEDGMENT

The Authors would like to thank the University Hospitals of Leuven and in particular radiologist Dr. Sofie Van Cauter for data acquisition and labeling the MRSI voxels. This research was supported by: Flemish Government FWO project G.0869.12N (Tumor imaging), G.0830.14N (Block term decompositions); IWT IM 135005; Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, 'Dynamical systems, control and optimization', 2012-2017); Research Council KUL: CoE PFV/10/002 (OPTEC); EU: European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013): EU MC ITN TRANSACT 2012, #316679 and ERC Advanced Grant, #339804 BIOTENSORS. This paper reflects only the authors views and the Union is not liable for any use that may be made of the contained information.

References

[1] Y. Li, D. M. Sima, S. V. Cauter, Croitor Sava A. R., U. Himmelreich, Y. Pi, and S. Van Huffel, "Hierarchical non-negative matrix factorization (hNMF): a tissue pattern differentiation method for glioblastoma multiforme diagnosis using MRSI," *NMR in Biomedicine*, vol. 26, no. 3, pp. 307–319, 2013.

[2] S. Dimou, R. Battisti, D. Hermens, and J. Lagopoulos, "A systematic review of functional magnetic resonance imaging and diffusion tensor imaging modalities used in presurgical planning of brain tumour resection," *Neurosurgical review*, vol. 36, no. 2, pp. 205–214, 2013.

[3] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *Signal Processing Magazine, IEEE*, vol. 32, no. 2, pp. 145–163, March 2015.

[4] O. Debals, M. Van Barel, and L. De Lathauwer, "Lowner-based blind signal separation of rational functions with applications," *Signal Processing, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[5] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors—part I: Basic results and uniqueness of one factor matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 855–875, 2013.

[6] ——, "On the uniqueness of the canonical polyadic decomposition of third-order tensors—part II: Basic results and uniqueness of one factor matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 876–903, 2013.

[7] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(L_r, L_r, 1)$ terms," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 4, pp. 1451–1474, Dec. 2011.

[8] H. N. Bharath, D. M. Sima, N. Sauwen, U. Himmelreich, L. De Lathauwer, and S. Van Huffel, "Tensor based tumor tissue type differentiation using magnetic resonance spectroscopic imaging," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, Aug 2015, pp. 7003–7006.

[9] S. Van Cauter, D. M. Sima, J. Luts, L. ter Beek, A. Ribbens, R. R. Peeters, M. I. Osorio Garcia, Y. Li, S. Sunaert, S. W. Van Gool *et al.*, "Reproducibility of rapid short echo time CSI at 3 tesla for clinical applications," *Journal of Magnetic Resonance Imaging*, vol. 37, no. 2, pp. 445–456, 2013.

[10] R. J. Ogg, R. Kingsley, and J. S. Taylor, "WET, a T1-and B1-insensitive water-suppression method for in vivo localized ^{1}H NMR spectroscopy," *Journal of Magnetic Resonance, Series B*, vol. 104, no. 1, pp. 1–10, 1994.

[11] T. Laudadio, N. Mastronardi, L. Vanhamme, P. Van Hecke, and S. Van Huffel, "Improved Lanczos algorithms for blackbox MRS data quantitation," *Journal of Magnetic Resonance*, vol. 157, no. 2, pp. 292–297, 2002.

[12] J. B. Poullet, "Quantification and classification of magnetic resonance spectroscopic data for brain tumor diagnosis," Ph.D. dissertation, PhD Thesis. Leuven, Belgium, 2008.

[13] V. Govindaraju, K. Young, and A. A. Maudsley, "Proton NMR chemical shifts and coupling constants for brain metabolites," *NMR in Biomedicine*, vol. 13, no. 3, pp. 129–153, 2000.

[14] A. Cichocki, R. Zdunek, A. H. Phan, and S. i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[15] L. Sorber, V. Barel, and L. D. Lathauwer, "Tensorlab v2.0," Available online, January 2014.

[16] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 9, no. 4, pp. 586–600, June 2015.

[17] K. Koh, S. Kim, and S. Boyd, "l1_ls: A matlab solver for large-scale L1-Regularized least squares problems," http://web.stanford.edu/~boyd/l1_ls/, 2008.

[18] V. Y. Tan and C. Févotte, "Automatic relevance determination in nonnegative matrix factorization with the β-divergence," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 7, pp. 1592–1605, 2013.

[19] J. D. Gibbons and S. Chakraborti, *Nonparametric statistical inference*. Springer, 2011.

[20] D. Krompaß, M. Nickel, X. Jiang, and V. Tresp, "Non-negative tensor factorization with rescal," in *Tensor Methods for Machine Learning, ECML workshop*, 2013.

[21] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognition*, vol. 41, no. 4, pp. 1350–1362, 2008.

[22] S. Wild, J. Curry, and A. Dougherty, "Improving non-negative matrix factorizations through structured initialization," *Pattern Recognition*, vol. 37, no. 11, pp. 2217–2232, 2004.

[23] G. Casalino, N. Del Buono, and C. Mencar, "Subtractive clustering for seeding non-negative matrix factorizations," *Information Sciences*, vol. 257, pp. 369–387, 2014.

[24] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 33, no. 2, pp. 387–392, 1985.

[25] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, vol. 19, no. 6, pp. 716–723, 1974.

[26] Y. Li, D. M. Sima, S. Van Cauter, U. Himmelreich, A. R. C. Sava, Y. Pi, Y. Liu, and S. Van Huffel, "Unsupervised nosologic imaging for glioma diagnosis," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 6, pp. 1760–1763, 2013.

[27] N. Sauwen, D. M. Sima, S. Van Cauter, J. Veraart, A. Leemans, F. Maes, U. Himmelreich, and S. Van Huffel, "Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multiparametric MRI," *NMR in Biomedicine*, vol. 28, no. 12, pp. 1599–1624, 2015.