Nonnegative Canonical Polyadic Decomposition for Tissue-Type Differentiation in Gliomas

H. N. Bharath, D. M. Sima, N. Sauwen, U. Himmelreich, L. De Lathauwer, and S. Van Huffel

***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 3D MRSI tensor from *in vivo* 2D-MRSI data of individual glioma patients. Nonnegative 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 nonnegative matrix factorization and hierarchical nonnegative matrix factorization.**

***Index Terms*—Glioma, magnetic resonance spectroscopic imaging (MRSI), nonnegative canonical polyadic decomposition (NCPD), l_1 regularization.**

I. Introduction

ACCURATE characterization 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 magnetic resonance (MR) modalities such as magnetic resonance spectroscopic imaging (MRSI), perfusion-weighted magnetic resonance imaging (MRI) (PWI) and diffusion-weighted MRI (DWI) are being used to characterize brain tumors and detect full tumor extent [2]. MRSI is a noninvasive imaging technique that provides spectral profiles in a 2- or 3-D voxel grid, from which the spatial distribution of metabolite concentrations can be estimated. Each voxel in the MRSI grid has a spectrum composed of several peaks corresponding to the metabolites present in that grid. The area under the peak is proportional to the metabolite concentration. 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 characterization, 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).

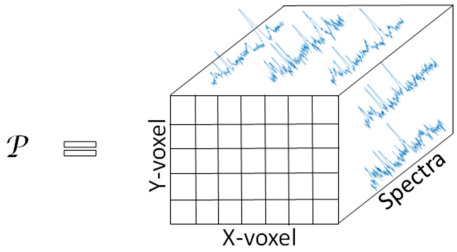

Fig. 1. Three-way spatial tensor representation of 2-D MRSI data.

Higher order tensors have certain properties that are not present in a matrix [3]. Tensor decompositions are now being used in various biomedical applications like genomics [4], electroencephalogram and functional MRI data analysis [5], and smart patient monitoring [6]. In recent years, there is an increasing trend to convert the data represented in a matrix to a third-order tensor [7]. Tensorization of the matrix is mainly motivated by the fact that tensor decompositions can be unique under mild conditions without imposing additional constraints [8], [9]. The 2-D MRSI signal can naturally be represented as a three-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 2D-MRSI signal and mode-3 represents the spectra from all voxels.

Manuscript received March 21, 2016; revised May 20, 2016; accepted June 14, 2016. Date of publication July 14, 2016; date of current version June 29, 2017. This work was supported in part by Flemish Government FWO project G.0869.12N (Tumor imaging) and G.0830.14N (Block term decompositions), in part by IWT IM 135005, in part by Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, 'Dynamical systems, control and optimization', 2012-2017), in part by Research Council KUL: CoE PFV/10/002 (OPTEC), in part by EU: European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013), and in part by EU MC ITN TRANSACT 2012, #316679, and ERC Advanced Grant: BIOTENSORS (n° 339804). This paper reflects only the author's views and the Union is not liable for any use that may be made of the contained information.

H. N. Bharath, D. M. Sima, N. Sauwen, and S. V. Huffel are with the Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium, iMinds Medical Information Technologies, Leuven, Belgium (e-mail: bhalandu@esat.kuleuven.be; diana.sima@esat.kuleuven.be; nicolas.sauwen@esat.kuleuven.be; sabine.vanhuffel@esat.kuleuven.be).

L. De. Lathauwer is with the Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium, The Group of Science, Engineering and Technology, KU Leuven Kulak, Kortrijk, Belgium (e-mail: lieven.delathauwer@kuleuven-kulak.be).

U. Himmelreich is with the Biomedical MRI Unit/Molecular Small Animal Imaging Center, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium (e-mail: uwe.himmelreich@med.kuleuven.be).

Digital Object Identifier 10.1109/JBHI.2016.2583539

2168-2194 © 2016 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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 [10]. 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 that allows to exploit the low-rank structure.

In this paper, we propose a novel algorithm for the detection of tumor, necrotic, and normal tissue types from MRSI signals. The algorithm first applies a window method to enhance the peaks and reduce the length of the spectra, and then, constructs a 3-D MRSI tensor. Decomposing this tensor using nonnegative canonical polyadic decomposition (NCPD) with common factor in mode-1 and mode-2, allows to retrieve the tissue-specific spectral profiles from the NCPD factor matrices. Preliminary studies have been previously presented in [11]. This 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, this 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 astrocytoma with focal progression to a grade III glioma) 2D-^{1}H MRSI data were acquired from 17 glioma patients on a 3-T MR scanner (Achieva, Philips, Best, The Netherlands) at the University Hospital of Leuven using the protocol [12]: A point-resolved spectroscopy (PRESS) sequence was used as the volume selection technique with a bandwidth of 1.3 kHz for the conventional slice-selective pulses; repetition time (TR)/ echo time (TE): 2000/35 ms; field of view (FOV): 160 mm $\times$ 160 mm; maximal volume of interest (VOI): 80 mm $\times$ 80 mm; slice thickness: 10 mm; acquisition voxel size: 10 mm $\times$ 10 mm; reconstruction voxel size: 5 mm $\times$ 5 mm; receiver bandwidth: 2000 Hz; samples: 2 048; number of signal averages: 1; water suppression method: multiple optimizations insensitive suppression train [13]; 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 3 min 30 s. Automated prescanning optimized the shim in order to yield a peak width consistently under 20-Hz full-width half-maximum. 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 postprocessing 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) [14]. A model order of 30 and a passband of 0.25–4.2 parts per million (ppm) was used in the HLSVD-PRO algorithm. After removing the water component, baseline correction and baseline offset correction was performed. All the preprocessing was done using the MATLAB (The MathWorks, Inc., Natick, MA, USA) based software, SPID [15].

The spectra were aligned in frequency using a simulated reference spectrum, which was generated using the parameters given in [16]. The complex-valued preprocessed 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 preprocessed spectrum. Elements of the vector X are obtained by moving an overlapping window over the spectrum, where the ith element of X is the sum of squares of absolute value of all the elements in the ith window segment.

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

where s_i is the spectrum at the ith segment, s_i^* is its complex conjugate, and L is the length of the window segment. Fig. 2(a) 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 nonnegative.

A three-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).

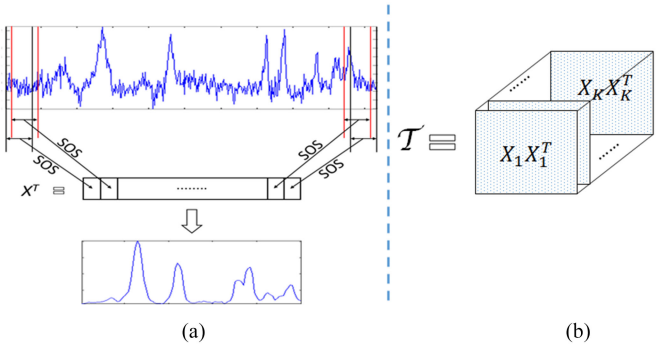

Fig. 2. (a) Construction of reduced spectrum, X from the preprocessed spectra. SOS: sum of squares. (b) Construction of the MRSI tensor $\mathcal{T}$ from the reduced spectra, X. K is the total number of voxels in the 2-D MRSI excitation volume.

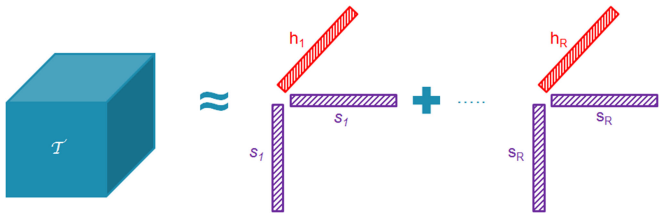

Fig. 3. NCPD of MRSI tensor $\mathcal{T}$: MRSI tensor $\mathcal{T}$ is decomposed into R rank-1 tensors. Common factor S is used in mode-1 and mode-2 to maintain symmetry of frontal slices. Each s_i gives a tissue-specific reduced spectral pattern and the corresponding h_i gives the spatial distribution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

B. NCPD

NCPD is a tensor decomposition method, where the tensor is decomposed into a sum of rank-one tensors with nonnegativity constraints on the factor matrices [17].

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

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 nonnegative 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.

Each rank-one term obtained from the NCPD of the MRSI tensor $\mathcal{T}$ is expected to correspond to a particular tissue type, although it is not always guaranteed with the current formulation. 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 an l_1 regularization on it. The NCPD with l_1 regularization (NCPD-$l1$) can be written as

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

where S and H are the aforementioned factor matrices and λ is the parameter that controls the sparsity. In this study, the tensor decomposition was performed using the Tensorlab MATLAB package [18]. Nonnegativity constraints, common factors to maintain symmetry and l_1 regularization are applied using the structured data fusion method [19] available in Tensorlab.

C. Spectral Recovery and Nonnegative 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 preprocessed 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

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

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 (5), respectively.

In the NCPD of the 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 normalized 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 unnormalized 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 nonnegative least squares with l_1 regularization

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

where $H_D(:,i)$ is the distribution of source spectra in each voxel, W contains the estimated source spectra, and $Y_{\mathrm{un}}(:,i)$ is the original unnormalized spectrum of each voxel. The problem in (6) is solved for all the voxels in the MRSI grid using a MATLAB-based large-scale l_1-regularized least-squares problem solver [20]. 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 nonnegative 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 nonnegative least-squares [20] 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 suboptimal solutions. To find good initialization values, first we take the 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 $H_{\text{init}} = (S_{\text{init}}^{\dagger} M)^T$, 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 nonnegative least-squares step.

E. Source ENumber 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 [18] has a method *rankest*, which estimates the rank based on the *L*-curve of the number of rank-one terms in a canonical polyadic decomposition. 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 [21], a Bayesian model based on automatic relevance determination is proposed for the NMF, which also estimates the model order R along with nonnegative 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 $K \times K$ sample covariance matrix is estimated as $C = \frac{1}{N-1}[(A - \bar{A}1_N^T)(A - \bar{A}1_N^T)^T]$, where $\bar{A} = \frac{1}{N-1}\sum_{i=1}^{N} A_i$ and 1_N is a vector of all ones with length N. 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.

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

where R^* is the estimated number of sources. 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 R^* obtained 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 R^* obtained 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 (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 following two measures.

1) *Source correlation:* In this paper, we have defined following 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 labeled 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.

$$\text{src_expt} = \frac{Y_1 + Y_2 + \cdots + Y_n}{n}$$

$$\text{SC1} = r(W(:,T), \text{src_expt})$$

   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 be-

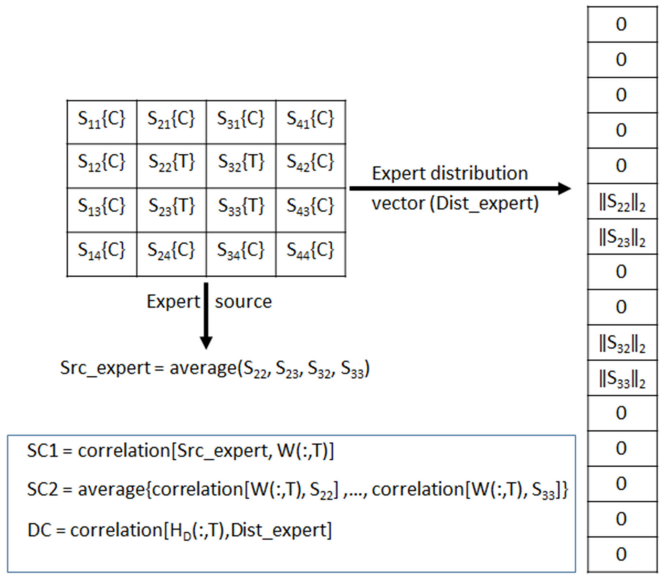

Fig. 4. Generation of expert labeled tissue-specific (Tumor) spectrum and distribution vector. Calculation of source and distribution correlation is shown in the box. S_{ij} is the spectra at ith column and jth row. The tissue type T-tumor or C-control is shown between braces.

tween 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.

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

where SC2 is the Type II source correlation and $W(:,T)$ 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.

$$\text{DC} = r(H_D(:,T), \text{dist_expt})$$

where DC is the distribution correlation, $H_D(:,T)$ is the estimated distribution vector corresponding to a particular tissue type, and dist_expt is the expert labeled distribution vector corresponding to that particular tissue type.

Because of heterogeneity, the tumor tissue is modeled 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 the 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.

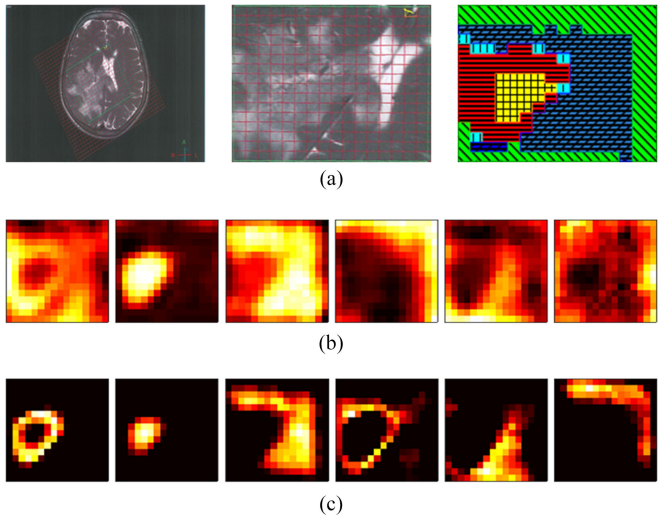

Fig. 5. (a) (Left to right) First image: T2-weighted anatomical MR image of a brain tumor with areas of necrosis; second image: voxels within the MRSI excitation volume superimposed on anatomical image; third image: expert labeling, where yellow (horizontal + vertical line) indicates necrotic, red (horizontal line) indicates tumor, dark blue (slanted + horizontal line) indicates normal, light blue (vertical line) indicates normal/tumor and green (slanted line) 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. Remaining three images correspond to bad spectra/artifact. (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. Remaining three images correspond to tumor, normal tissue, and bad spectra/artifact distribution, respectively. In this dataset, the tumor tissue is modeled by two sources. For all the distribution maps, hot colormap was used, where dark area represents lower values and light area represents higher values.

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 six 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. 5(b). Nonnegative 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 that 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 six sources in Fig. 5(c). The MRSI grid superimposed on anatomical image and the expert labeling is shown in Fig. 5(a). Comparing the distribution maps in Fig. 5(b) and 5(c) with the expert labeling in Fig. 5(a), 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

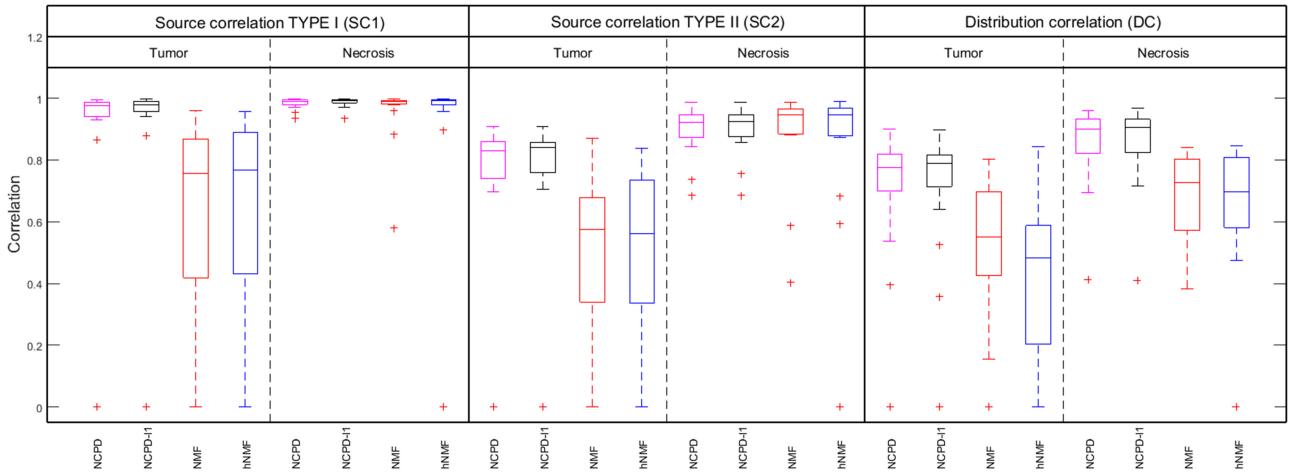

Fig. 6. Box plot of Type I source correlation, Type II source correlation and distribution correlation values obtained from NCPD without regularization, NCPD with l_1 regularization , and single-stage NMF and hNMF algorithms. Zero correlation values indicate that the specific tissue type was not recovered.

TABLE I
MEAN, STANDARD DEVIATION, MEDIAN, MAD, AND RANGE OF SOURCE CORRELATION (TYPE I AND TYPE II) AND DISTRIBUTION CORRELATION FOR 28 GLIOMA DATASETS

		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	-	<0.0001*	<0.0001*	0.3269	-	<0.0001*	<0.0001*	0.4745	-	<0.0001*	<0.0001*
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*

The results are shown on tumor and necrotic tissue type for NCPD without regularization, NCPD with l_1 regularization, and single-stage NMF and hNMF algorithms. p-values indicate the statistical significance of higher median for the NCPD-$l1$ algorithm (* indicates statistical significance for $p < 0.01$)

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.

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 astrocytoma with focal progression to a grade III glioma) from 17 patients with gliomas. The Type I and Type II source correlations 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 as box plots 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 five datasets. The correlation values of the tissue types that are not recovered are set to zero (Fig. 6). A summary of the results, i.e, mean, standard deviation (std. dev.), median, 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 ($\alpha = 0.01$) [22]. 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 ($p > 0.01$).

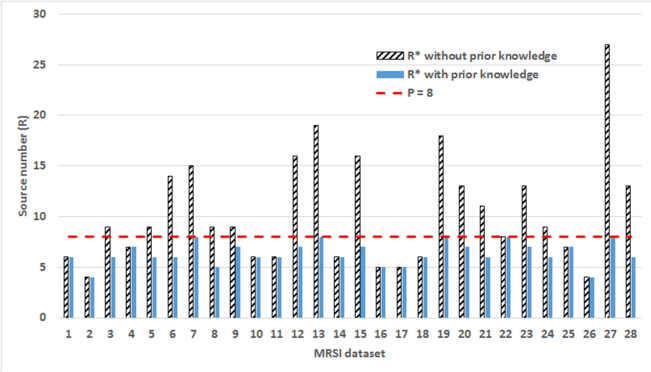

Fig. 7. Estimated number of sources from the covariance matrix for 28 MRSI datasets with and without using the maximum number of tissue types as prior knowledge. The horizontal line indicates the maximum number of tissue types used as prior knowledge in the analysis ($P = 8$).

For the necrotic sources, NCPD-l_1 has slightly better median values for correlation than NCPD (no regularization), but the differences are not significant ($p > 0.01$).

For all the MRSI datasets, the rank was automatically estimated using the method explained in Section III-E. Fig. 7 shows the estimated ranks for 28 MRSI datasets with and without using 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 as 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 labeling, 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. 8(a). The spectra that 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 five 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. 8(b)–(e). Using the hNMF method, only three sources are obtained as shown in Fig. 8(f) and 8(g). Fig. 8(g) shows that the hNMF method identifies the normal and necrotic ($\mathrm{SC1} = 0.9971$) tissue properly but fails to recover the tumor tissue. Single-stage NMF identifies normal and necrotic tissue, but only the necrotic source is good ($\mathrm{SC1} = 0.9941$) and the normal source deviates a lot from the expert as shown in Fig. 8(d) (first row). In the single-stage NMF method, the recovered tumor tissue ($\mathrm{SC1} = 0.6041$) and its corresponding distribution ($\mathrm{DC} = 0.4262$) are bad and it is difficult to identify it as tumor tissue from the source spectrum. The NCPD method identifies all three tumor ($\mathrm{SC1} = 0.9875$), necrotic ($\mathrm{SC1} = 0.9854$,) and normal tissue types. Fig. 8(b) and 8(c) shows 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 (see Fig. 8(b) and 8(c): seventh row) and spectra from the outer edges of the voxel grid (see Fig. 8(b) and 8(c): fourth, fifth, and sixth 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. An $(L_r, L_r, 1)$ 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 2D-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 NCPD 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 three 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 modeled using a higher number of sources and low-grade gliomas with less artifacts can be modeled 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.

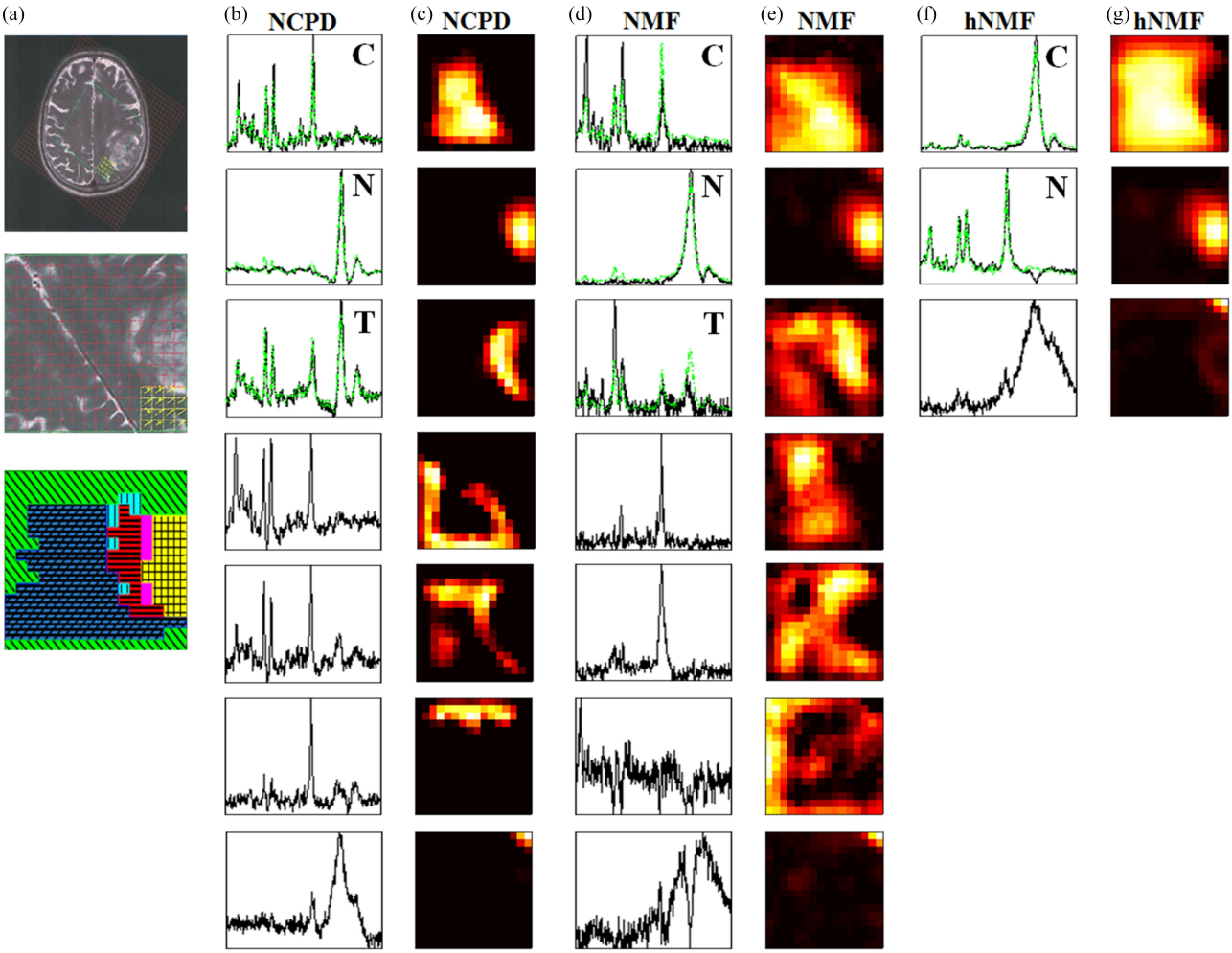

Fig. 8. Tissue pattern differentiation using ^{1}H MRSI: C, T, and N represent normal, tumor, and necrosis, respectively. (a) First row: T2-weighted anatomical MR image of a brain tumor with areas of necrosis. Second row: voxels within the MRSI excitation volume superimposed on anatomical image. Third row: expert labeling, where yellow (horizontal + vertical line) indicates N, red (horizontal line) indicates T, magenta (no pattern) indicates T/N, dark blue (slanted + horizontal line) indicates C, light blue (vertical line) indicates C/T and green (slanted line) indicates spectra of poor quality. (b), (c) Results of NCPD. (b) Recovered sources from the NCPD-l_1 method are shown in black (solid line). First three rows represent C, N, and T spectral sources in black (solid line), with tissue-specific spectra based on expert labeling overlaid in green (dash-dot line). 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) Recovered sources are shown in black (solid line), overlaid with the expert-based tissue-specific spectra in green (dash-dot line). 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 (solid line) and expert-based tissue-specific spectra in green (dash-dot line). First two rows show control and necrotic spectra. (g) Distribution maps corresponding to (f). For all the distribution maps, hot colormap was used, where dark area represents lower values and light area represents higher values.

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 nonnegative tensor factorization applications. In nonnegative RESCAL tensor factorization [23], the factor matrices are initialized using an NMF initialization method nonnegative double SVD method (NNDSVD) [24]. When NCPD is initialized using NNDSVD on the reduced spectra, 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 [25], [26]. Although our initialization method is based on SVD, it does not suffer from too many 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 by striped bars in Fig. 7. Use of prior knowledge about the maximum number of tissue types prevents this overestimation and results in better number of sources estimation as shown by solid bars in Fig. 7. In this method, we have used a cutoff of 99% on the cumulative sum of eigenvalues of C. Whereas, information theoretic criteria-based methods such as in [27] and [28] can determine the number of sources adaptively

without the need for a cutoff 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 the 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 [29], 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 hNMF has been used to characterize brain tumor heterogeneity [30]. 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 [30].

Acknowledgment

The Authors would like to thank the University Hospitals of Leuven, and in particular, radiologist Dr. S. Van Cauter for data acquisition and labeling the MRSI voxels.

References

[1] Y. Li *et al.*, "Hierarchical non-negative matrix factorization (hNMF): A tissue pattern differentiation method for glioblastoma multiforme diagnosis using MRSI," *NMR Biomed.*, 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," *Neurosurg. Rev.*, vol. 36, no. 2, pp. 205–214, 2013.

[3] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[4] C. Muralidhara, A. M. Gross, R. R. Gutell, and O. Alter, "Tensor decomposition reveals concurrent evolutionary convergences and divergences and correlations with structural motifs in ribosomal RNA," *PloS One*, vol. 6, no. 4, p. e18768, 2011.

[5] M. Mørup, L. K. Hansen, S. M. Arnfred, L.-H. Lim, and K. H. Madsen, "Shift-invariant multilinear decomposition of neuroimaging data," *NeuroImage*, vol. 42, no. 4, pp. 1439–1450, 2008.

[6] S. Van Huffel, "Tensor Decompositions in Smart Patient Monitoring," *SIAM News*, vol. 48, no. 7, p. 1, Sep. 2015.

[7] O. Debals, M. Van Barel, and L. De Lathauwer, "Löwner-based blind signal separation of rational functions with applications," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1909–1918, Apr. 2016.

[8] 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 J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.

[9] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors—Part II: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.

[10] 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.

[11] H. N. Bharath *et al.*, "Tensor based tumor tissue type differentiation using magnetic resonance spectroscopic imaging," in *Proc. IEEE 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2015, pp. 7003–7006.

[12] S. Van Cauter *et al.*, "Reproducibility of rapid short echo time CSI at 3 tesla for clinical applications," *J. Magn. Reson. Imag.*, vol. 37, no. 2, pp. 445–456, 2013.

[13] 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," *J. Magn. Reson.*, vol. 104, no. 1, pp. 1–10, 1994.

[14] T. Laudadio, N. Mastronardi, L. Vanhamme, P. Van Hecke, and S. Van Huffel, "Improved Lanczos algorithms for blackbox MRS data quantitation," *J. Magn. Reson.*, vol. 157, no. 2, pp. 292–297, 2002.

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

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

[17] 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*. New York, NY, USA: Wiley, 2009.

[18] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Available online, URL: http://www.tensorlab.net/, Mar. 2016.

[19] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 586–600, Jun. 2015.

[20] K. Koh, S. Kim, and S. Boyd. (2008). l1_ls: A MATLAB solver for large-scale L1-regularized least squares problems. [Online]. Available: http://web.stanford.edu/ boyd/l1_ls/

[21] V. Y. Tan and C. Févotte, "Automatic relevance determination in nonnegative matrix factorization with the β-divergence," *IEEE Trans., Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1592–1605, Jul. 2013.

[22] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*. Berlin, Germany: Springer, 2011.

[23] D. Krompaß, M. Nickel, X. Jiang, and V. Tresp, "Non-negative tensor factorization with RESCAL," presented at the ECML Workshop Tensor Methods for Machine Learning, Prague, Czech Republic, 2013.

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

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

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

[27] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[28] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[29] Y. Li *et al.*, "Unsupervised nosologic imaging for glioma diagnosis," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 6, pp. 1760–1763, Jun. 2013.

[30] N. Sauwen *et al.*, "Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI," *NMR Biomed.*, vol. 28, no. 12, pp. 1599–1624, 2015.

Authors' photographs and biographies are not available at the time of publication.