Magnetic Resonance Spectrum Separation Using Sparse Representations and Wavelet Filters

Yu Guo, Su Ruan, Jérôme Landré, Jean-Marc Constants

Magnetic Resonance Spectrum Separation Using Sparse Representations and Wavelet Filters

Yu Guo, Su Ruan, Jérôme Landré
Univ. Reims-Champagne-Ardenne, CReSTIC
IUT, 9 rue de Québec, 10 026 Troyes cedex, France
Email: yu.guo@etudiant.univ-reims.fr, su.ruan@univ-reims.fr
Tel.: +33 (0)3 25 42 71 68, Fax: +33 (0)3 25 42 46 43

Jean-Marc Constans
Dept. de radiologie, CHU de Caen
Hôpital Côte de Nacre 14 033 Caen cedex, France
Email: constans-jm@chu-caen.fr
Tel.: +33 (0)2 31 06 49 71, Fax: +33 (0)2 31 06 45 52

***Abstract*—Magnetic Resonance spectroscopy (MRS) provides a "frequency-signal intensity" spectrum of multiple peaks that reflect the biochemical composition of a localized region in the body. The peak intensity or the area under each peak is proportional to the concentration of that assigned metabolite. Accurate quantification of in vivo MRS (measuring peak intensity or area) is very important to diagnose certain metabolic disorders. However, strongly overlapping metabolite peaks, poor knowledge about background component (the baseline), and low signal-to-noise ratio (SNR) make the task difficult. In this paper, a novel spectrum separation method using sparse representations and wavelet filters is proposed to separate baseline and spectra of different metabolites and finally achieves an accurate MRS quantification. With the proposed method, the accuracy and the robustness of MRS quantification are improved, from simulation data, compared with a commonly used frequency-domain MRS quantification method. The quantification on tumor metabolism with in vivo brain MR spectra is also demonstrated.**

I. Introduction

Over the years MRS (Magnetic resonance spectroscopy) has clearly proven to be an important non-invasive tool in biomedical research to study humans and animals [1]. The observed MRS spectrum is a mixture of several spectra respectively corresponding to different metabolites in a localized region under study. Each biochemical or metabolite is identified by its unique position or chemical shift along the frequency axis of a spectrum [2]. The biochemical quantities (concentration, PH) of a certain metabolite can be obtained by accurately quantifying the corresponding MR spectrum (computing the peak intensity or the area under the spectrum peak). Accurate quantification of in vivo MRS is very important to diagnose certain metabolic disorders, especially those affecting the brain. However, when a large background component (the baseline) is present along with a low signal to noise ratio, accurate quantification of MR spectra is a challenging problem [3]. In this paper, a method using sparse representations and wavelet filters is proposed to separate baseline and spectra of different metabolites for achieving more accurate MRS quantification.

In recent years sparse representation is attracting more and more interest in signal processing domain. In the study of sparse representation, in literature [4]–[8] lots of pursuits algorithms are proposed for finding the sparsest representation of a signal in a given dictionary. Taking into account the a priori knowledge about the model of MR spectra and the peak frequencies of different metabolites, the representation of a mixed MR spectrum in a dictionary constructed with a set of Lorentzian and Gaussian basis functions will be sparse and disjoint. Therefore, a non-negative pursuit algorithm which is based on the regulative FOCUSS algorithm in [8] is used to estimate the sparse representation and separate spectra of different metabolites.

Besides, for eliminating the influence of the baseline to the estimation of sparse representation coefficients, a wavelet filter is used to remove the smooth component of an observed spectrum. Because of the smoothness and broadness of baseline, baseline will be eliminated with the wavelet filters. However, the overlapping component of metabolite spectra with the baseline will also be removed with the filter. By processing each basis function in the given dictionary with the same filter, the estimation of representation coefficients will not be influenced by the removed component of metabolite spectra.

The proposed method is evaluated with simulation experiments and compared with a commonly used frequency-domain MRS quantification method in [9]. The results of these experiments show the effective performance of the proposed method. The quantifications of in vivo brain MR spectra are confirmed by the experts and demonstrated in this paper.

II. Theory and Method

A. Signal model

Generally, the mathematical model of a mixed metabolite spectrum S can be expressed in the following form:

$$S = \sum_{k=1}^{K} s_i = \sum_{k=1}^{K} [L_k(f) + G_k(f)] \tag{1}$$

with

$$L_k(f) = \frac{a_{Lk}}{1 + ((f - f_{Lk})/d_{Lk}^2)}$$

and

$$G_k(f) = a_{Gk} e^{-((f - f_{Gk})/d_{Gk})^2}$$

where L_k and G_k denote Lorentzian and Gaussian lineshape respectively, f is the frequency of each data point, K stands

for the number of lines used to build up the spectrum and the spectrum of single metabolite is corresponding to one or several lines [10]. $p_k = [a_{Lk}, d_{Lk}, f_{Lk}, a_{Gk}, d_{Gk}, f_{Gk}]$ is a model parameter vector, where a is the intensity, d the linewidth and f the central frequency. The spectra of different metabolites have different central frequencies which are known and can be used as an a priori knowledge in a MRS quantification method. When a possible baseline contribution and a noise which is often assumed as Gaussian distributed are considered, an observed MR spectrum can be modelized as

$$X = S + B + e \tag{2}$$

In most of MRS quantification approaches, especially the data obtained by quantum mechanics or from measured in vitro metabolite solutions are unavailable such as the method in [11], quantification can be achieved by using a nonlinear least-squares algorithm to estimate the nonlinear parameter vector p. To deal with the baseline problem, baseline is generally approximated by a mathematical function (a sum of splines, wavelets, sinusoids or polynomials with adjustable parameters) then included in the parametric nonlinear least-squares fit or separately accommodated in a pre-processing step [12].

For these MRS quantification methods based on nonlinear optimization, some nonlinear methods have the risk of converging to a local minimum while the other have the disadvantage of poor computational efficiency, and their performances deteriorate with the increase of the number of model parameters. For a better application, the quantification accuracy needs to be improved, especially in the condition that large baseline components overlap with metabolite spectra. For avoiding the disadvantage caused by nonlinear optimization and improving the quantification accuracy, a linear non-parameter method proposed in this paper.

B. Sparse representation

A sparse representation model can be described as

$$x = Dw \tag{3}$$

where $x_{N\times 1}$ is the signal vector which can be sparsely represented as a linear combination of columns (often called "basis vectors") of a dictionary matrix $D_{N\times M} = [d_1, d_2, \ldots, d_M]$, and when $M > N$, it is a overcomplete dictionary. $w_{M\times 1}$ is the representation coefficient vector, minimal elements of which are nonzero.

The study of sparse representation has concentrated mainly on the design of dictionaries and pursuit algorithms. Designing dictionaries to better fit the sparse representation model can be done by either selecting one from a prespecified set of linear transforms or adapting the dictionary to a set of training signals [13]. When the dictionary is known and fixed, pursuit algorithms can be used for the extraction of the sparsest representation. The main pursuit algorithms in literature are matching pursuit [4] and the orthogonal matching pursuit (OMP) algorithms [5] which are greedy algorithms that select basis functions sequentially by involving the computation of inner products between the signal and dictionary columns, the basis pursuit algorithm which suggests a convexification of the problems by replacing l_0-norm with l_1-norm [6], the focal underdetermined system solver (FOCUSS) which using the l_p-norm with $p \leq 1$ as a replacement for the l_1-norm [7,8].

In this paper sparse representation is used for separating MR spectra. Based on the character and the available a priori knowledge of the particular signal, the dictionary is designed and a non-negative pursuit algorithm is used to finding the sparsest representation.

C. Dictionary construction

According to the model of MR spectra in (1), a mixed metabolite spectrum can be represented as the linear combination of several Gaussian and Lorentizian functions. Hence, an overcomplete dictionary D, which contains enough basis vectors designed as the sample vectors of normalized Gaussian and Lorentizian functions, can will fit the sparse representation model in (3). Then, a mixed metabolite spectrum vector S can be represented as

$$S = Dw = \sum_{m=1}^{M} d_m w_m$$

where d_m is one of the basis vectors which has the same length as S, M is the number of basis vectors in D. For a detailed illumination of the construction of the dictionary, all the basis vectors in D are divided into K groups $\{D_1, D_2, \ldots, D_K\}$, where K is the number of the spectrum peaks under study. The basis vectors d_{ij} $(j = 1 \ldots L)$ in group D_i $(i = 1 \ldots K)$ have the same central frequencies with the known central frequency of one of the metabolite peaks under study and different widths $d_{ij} = d_{0j} + j\Delta d$. Because the spectrum peaks of different metabolites have different central frequencies, the representations of spectra of different metabolites in the constructed dictionary are disjoint. If we can find w the sparsest solution of $S = Dw$, we can decompose the spectrum S into several components $S_i = D_i w_i$ $(i = 1 \ldots K)$ which relate to spectra of different metabolites.

However, because of the presence of baseline B in an observed spectrum, the sparest solution w of $X = Dw$ will certainly depart from the expected solution of $S = Dw$. The smoothness of baseline is usually used to deal with baseline problem and a wavelet filter can be used to eliminate baseline [12]. However, because of the overlapping between baseline and the metabolite spectra, when a wavelet filter is used to process observed spectra, it is difficult to only remove baseline and reserve the whole mixed metabolite spectrum.

Here, for accurate estimation of the sparest solution w, each basis vector of the constructed dictionary is also processed with the same wavelet filter, then a basis vector d in D can be denoted as $d = d_h + d_l$, where d_l is the removed part of d with the wavelet filter, and d_h the reserved parts. Based on the wavelet filter processing results of all the basis vectors in D, the dictionary D can be divided into D_h and D_l, which

are corresponding to d_h and d_l respectively. The reserved part S_h of a mixed metabolite spectrum S with the wavelet filter can be represented as $S_h = D_h w$. When the wavelet filter removes the whole baseline, there is $S_h = x_h$, where x_h the reserved part of an observed spectrum x. Furthermore, if each column of D_h is not zeros or too small value vector, the sparest solution w of $x_h = D_h w$ is equal to the sparse representation of a mixed metabolite spectrum S in the constructed dictionary D.

Therefore, MR spectra separation can be achieved by finding the sparsest solution w of $x_h = D_h w$.

D. Non-negative FOCUSS pursuit algorithm

Theoretically, the sparsest representation of $y_{N\times 1}$ with respect to an overcomplete dictionary $D_{N\times M}$ is the following optimization problem:

$$\min_{w} \| w \|_0 \quad \text{subject to} \quad y = Dw \tag{4}$$

or

$$\min_{w} \| w \|_0 \quad \text{subject to} \quad \| y - Dw \|_2 \leq \epsilon \tag{5}$$

where $\| . \|_0$ is the l_0 norm, for counting the nonzero entries of a vector; $w_{M\times 1}$ is the coefficient vector of basic functions. However, it is a NP-hard problem. Thus, approximate solutions are considered instead, the FOCUSS minimizes $l_{(p\leq 1)}$-norm in place of l_0-norm to obtain sparse solution. Then, the sparse representation becomes the solution of

$$\min_{w} \sum_{i=1}^{M} sgn(w(i)) \mid w(i) \mid^p \quad \text{subject to} \quad y = Dw \tag{6}$$

When the noise exists, an exact minimum norm solution can not be sought. Instead, a solution that minimizes $l_{(p\leq 1)}$-norm and approximately satisfies the set of constraints is found. The solution is

$$w = arg\min_{w} J(w) \tag{7}$$

where

$$J(w) = [\| Dw - y \|^2 + \gamma E^{(p)}(w)]$$

For the problem in this paper, taking into account of the non-negativity of the representation coefficients, a nonnegative constraint is added. So the representation coefficient vector w can be estimated as

$$w = arg\min_{w} J(w) \tag{8}$$

where

$$J(w) = [\| Dw - y \|^2 + \gamma E^{(p)}(w)] \quad \text{and} \quad \forall i : w_i \geq 0$$

At each iteration step, we set the negative values of the solution to zero for guaranteeing the nonnegative character of coefficient vector w. Based on the basic iterative form of the regularized FOCUSS algorithm, the iterative form of the nonnegative pursuit algorithm becomes:

$$W_{k+1} = diag(w_{ki}^{1-(p/2)}) \tag{9a}$$

$$w_{k+1} = W_{k+1} D_{k+1}^T (D_{k+1} D_{k+1}^T + \lambda I)^{-1} y \tag{9b}$$

$$w_{k+1}(i) = \begin{cases} 0 & \text{if} \quad w_{k+1}(i) < 0 \\ w_{k+1}(i) & \text{if} \quad w_{k+1}(i) \geq 0 \end{cases} \tag{9c}$$

The parameter λ controls the tradeoff between quality of fit $\| y - Dw \|$ and the degree of sparsity, and the value of λ should increase with the level of noise. The iterative algorithm in (9) is used to estimate the sparsest solution w of $x_h = D_h w$. With the representation coefficient vector w, spectrum of different metabolites can be separated by computing $s_i = D_i x_i$ $(i = 1 \ldots K)$. Finally, MRS quantification can be achieved easily benefiting from the separation results.

TABLE I
PARAMETERS OF SIMULATED METABOLITE SPECTRA AND BASELINE.

Parameters of metabolic spectra				Parameters of baseline			
Metab.	f_k (ppm)	d_k	a_k	k	f_k (ppm)	d_k	a_k
Cr	3.91	8.0	8.0	1^G	2.02	85	2.33
Glu/Gln	3.74	4.0	4.0	2^G	2.35	100	0.33
mI	3.56	4.0	5.5	3^G	2.50	70	0.67
Tau	3.42	3.0	3.2	4^G	3.00	70	0.67
Cho	3.22	6.0	2.0	5^G	3.29	100	1.0
Cr/PCr	3.03	6.0	4.0	6^G	3.50	80	2.00
GABA	2.37	4.0	1.5	7^G	4.00	70	0.67
Glu/ Gln	2.12	6.0	9.5	8^L	0.90	300	1.33
NAA	2.02	5.0	13.0	9^L	1.30	300	1.67
Lac	1.33	5.0	1.5				
Lac	1.26	3.0	1.5				

III. EVALUATION AND RESULT

A. Evaluation with simulated spectra

The proposed approach is tested with simulated ^1H human brain MRS signals. A simulated MR spectrum consisting of 512 data points is composed of eleven single metabolite spectra modelized as Gaussian functions, a baseline mixing of nine models and a certain level of Gaussian noise. The parameters of simulated metabolite spectra and baseline are shown in Table I. For each noise level, a set of 100 simulated MR spectra is created in order to give reliable values for the uncertainties of the estimation results and check the robustness of the proposed method. The value of the SNR is defined as the ratio of the amplitude of the NAA resonance to the standard deviation of noise e.

As an a priori knowledge, the basis vectors in the constructed dictionary supposed to represent spectra of main metabolites under study have the known central frequencies as shown in Table I and the unknown linewidths taken from the range $[d_1, d_2] = [0.5, 10]$ with the sample step $\Delta_d = 0.5$. In this paper, the Coiflet discrete wavelet decomposition of the observed MR spectra using COIF5 at level 5, is computed and the detail coefficients of the wavelet decomposition are reserved to reconstruct x_h, and the same wavelet decomposition

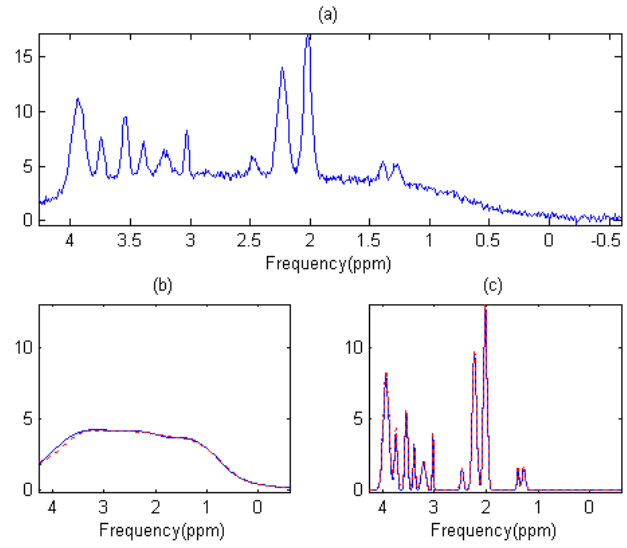

Fig. 1. Separation results of a simulated MR spectrum (SNR=17dB) using the proposed method (a) raw spectrum; (b) the estimated and the true baseline (dashed line); (c) the estimated and the true spectrum (dashed line).

of each basis vectors in D and reconstruction are computed to get D_h.

B. Results

Figure 1 displays a simulated MR spectrum with SNR=17dB as in vivo conditions, and the corresponding separation results. The comparison of separation results and real spectra in Fig. 1 shows that baseline is well separated from metabolite spectra of interest.

Figure 2 plots the RRMSE (Relative Root Mean Square Error) of estimated metabolite peak areas in different noise levels (SNR =12dB, 14dB, 16dB, 18dB, 20dB). RRMSE is defined as the ratio of RMSE (Root Mean Square Error) of estimated results to the real value. Because of the different peak intensities, a noise influences the quantification of these metabolites in different degrees. As demonstrated in Fig. 2, for the metabolites with relatively big peaks, such as NAA, Cr, the estimation results are less influenced with the increase of the noise, while for the weak metabolite peaks, especially the metabolites Lac and Cho, the noise have relatively big influence to the estimation result. Like most of the other MRS quantification methods, the estimation accuracy decreases with the increase of noise level.

Using simulated MR spectra with SNR=17dB, the proposed method is compared with a commonly used frequency-domain MRS quantification method in which Levenberg-Marquardt algorithm is used to estimate the nonlinear model parameters of metabolite spectra and a wavelet filter is used to remove the baseline component in an iterative subtraction manner [9]. In simulation experiments, the same a priori knowledge is used in both methods. For the method in [9], the choice of the initial parameters can deeply influence the estimation result, while

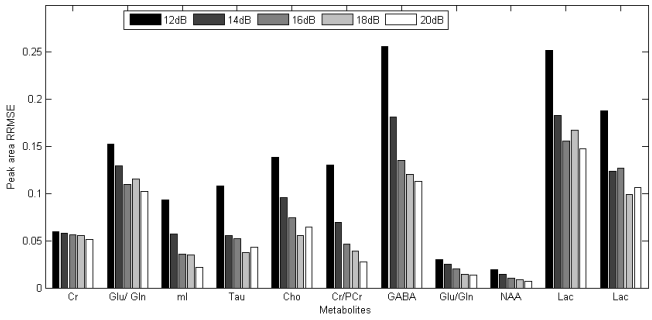

Fig. 2. The RRMSE of estimated metabolite peak areas in different noise levels with the proposed method.

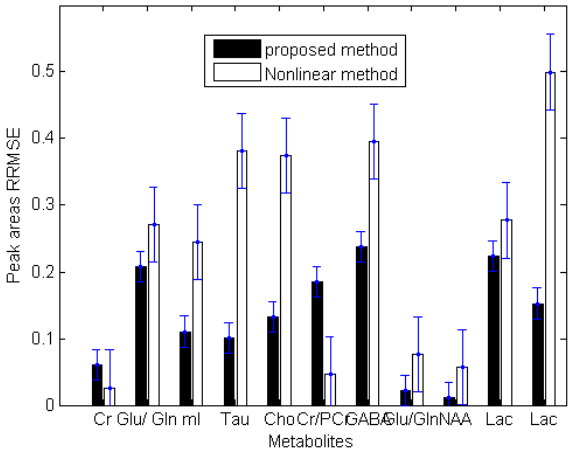

Fig. 3. Comparison results. RRMSE of estimated metabolite peak areas and the error bars of the proposed method and the nonlinear one (SNR=17dB).

in the proposed method, a small random positive value vector is used for the initialization of the representation coefficients and different initializations have no influence to estimation results. The RRMSE of metabolite peak areas and the error bars (the standard deviation of relative estimation errors) of the both methods are illustrated in Fig. 3. Obviously, the relatively small estimation errors and error bars of the proposed method show the superiority of our method.

C. In vivo MRS data

The method was also tested on in vivo ^1H human brain MR spectra. Two human brain MR spectra from a brain tumor patient obtained with PRESS and an echo-time of 35 ms were quantified by our method. Fig. 4 (a) and Fig. 4(c) are the MR spectra of tumor tissue and normal tissue respectively. Fig. 4 (b) and Fig. 4 (d) show the separation results of the baseline and metabolite spectra of interest. According to the spectrum separation results, the peak areas of the main metabolites in each MR spectra were computed. Comparing the different MRS quantification results of normal tissue and tumor tissue, doctors can analyse the kind of tumor more easily.

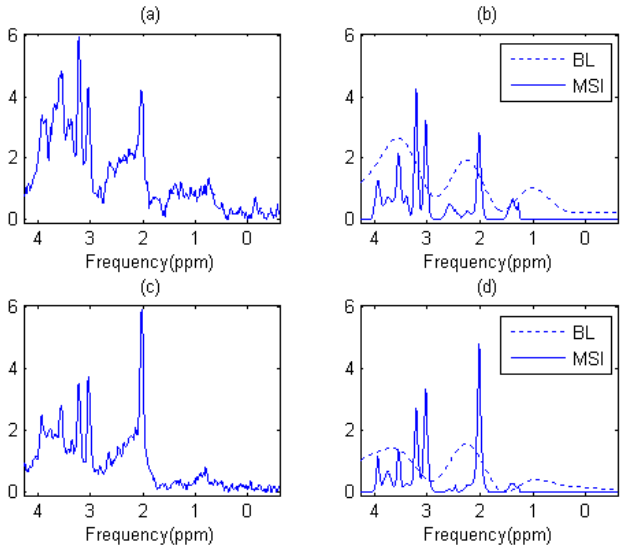

Fig. 4. The separation results of in vivo 1H human brain MR spectra. (a) spectrum from a tumor tissue; (b) separation results of the MR spectrum in (a); (c) spectrum from a normal tissue; (d) separation results of the MR spectrum in (c) (BL=baseline, MSI=spectra of interest).

IV. Conclusions

The new MRS quantification method proposed in this paper combines a wavelet filter and a linear non-parametric algorithm based on sparse representation to remove baseline and separate spectra of different metabolites and finally achieves an accurate MRS quantification. The method can separate the overlapping components of baseline and metabolite spectra of interest and well deals with the challenging baseline problem. In addition, it effectively uses the a priori knowledge and achieves satisfying MR spectrum separation. Compared with commonly used nonlinear parametric optimization algorithms, the proposed method has a better computational efficiency, avoid the choice of initial parameters and are more robust. In future work, the method will be tested on large data of in vivo brain spectra.

References

[1] HJA in't Zandt, M van der Graaf, et al., *Common processing of in vivo MR spectra*, NMR Biomed, pp. 224–232, 2001.

[2] Jeff A. Stanley, Jay W. Pettegrew, and Matcheri S. Keshavan, *Magnetic Resonance Spectroscopy in Schizophrenia: Methodological Issues and Findings-Part I*, Society of Biological Psychiatry, 48, pp. 357–368, 2000.

[3] R. Bartha, D.J. Drost and P.C. Williamson, *Factors affecting the quantification of short echo in-vivo 1H MR spectra: prior knowledge, peak elimination, and filtering*, NMR Biomed, pp. 205–216, 1999.

[4] S. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. Signal Process., vol. 41, no. 12, pp. 3397–3415, 1993.

[5] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, *Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition*, in Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput., vol. 1., pp. 40–44, 1993

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by basis pursuit*, SIAM Rev., vol. 43, no. 1, pp. 129–159, 2001.

[7] I. F. Gorodnitsky and B. D. Rao, *Sparse signal reconstruction from limited data using FOCUSS: A re-weighted norm minimization algorithm*, IEEE Trans. Signal Process., vol. 45, pp. 600–616, 1997.

[8] B.D. Rao, K. Engan, S.F. Cotter, et al., *Subset selection in noise based on diversity measure minimization*, IEEE Trans. Sig. Process., pp. 760–770, 2003.

[9] P. Gillies, I. Marshall, M. Asplund, P. Winkler, and J. Higinbotham, *Quantification of MRS data in the frequency domain using a wavelet filter, an approximated Voigt lineshape model and prior knowledge*, NMR Biomed, 19, pp. 617–626, 2006.

[10] S. Mierisova, M. Ala-korpela, *MR spectroscopy quantitation: a review of frequency domain methods*, NMR Biomed, pp. 247–259, 2001.

[11] L. Vanhamme, A. van den Boogaart and S. Van Huffel, *Improved method for accurate and efficient quantification of MRS data with use of prior knowledge*, J. Magn. Reson., pp. 35–43, 1997.

[12] H. Ratiney, M. Sdika, and Y. Coenradie, *Time-domain semi-parametric estimation based on a metabolite basis set*, NMR Biomed, pp.1–13, 2005.

[13] M. Aharon, M. Elad, and A. Bruckstein, *K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation*, IEEE Trans. Sig. Process., pp. 4311–4322, 2006.