Gaussian mixture model Estimation using the Expectation Maximization algorithm for MRS inversion-recovery signals

H. Ratiney[1], A. Bucur[1], C. Cudalbu[1], and S. Cavassila[1]
[1]Creatis-LRMN, CNRS UMR 5220, Inserm U630, INSA-Lyon, Université Lyon 1, Villeurbanne, France

Introduction
The MR spectroscopic macromolecular signal usually considered as a nuisance contribution in the quantification of short echo time signals might reveal some interest as a disease marker by itself [1]. Once acquired an assumed metabolite-free macromolecular signal from an inversion recovery experiment, its quantitative analysis is problematic. Enabling its objective quantification and deriving a model function for its description is the purpose of the proposed method. For unknown prior model function, so called black box method based on singular value decomposition are often invoked to provide a description of MR spectroscopic signal. This approach works well for describing separate Lorentzian lineshape resonances such as singlet resonance from good shimmed acquisition. However, inversion recovery macromolecular spectrum coming from unresolved proteins, polypeptides, residual water and subcutaneous lipids shows broad patterns closer to a gaussian mixture model. The proposed method takes advantage of the Expectation Maximization (EM) algorithm [2] usually applied in machine learning or pattern recognition to provide a novel robust fitting procedure for broad patterns MR spectroscopic signal.

Method
***Experimental conditions*:** The signals used to demonstrate the feasibility of the method were acquired on a horizontal 4.7T and 7T Biospec BRUKER system. At 7T, 3 healthy adult rats (Sprague-Dawley), and at 4.7T, 3 healthy mice (Swiss, 4 weeks old) were anesthetized by a gas mixture with isoflurane. Acquisitions were performed using a short.echo time (TE=20ms,bandwidth of 4kHz, 4096 data-points, Number of Average of 512 for 7T and 1024 for 4.7T) including an inversion pulse prior to the PRESS sequence. All first- and second- order shim terms were adjusted using FASTMAP for each volume of interest centered in the hippocampus (left size of the brain, $3.2x2x3.2mm^3$ at 7T and $2.5x2x2mm^3$ at 4.7T). The inversion time after the inversion pulse was set to 700 ms TR=3.5s, for the acquisition at 4.7T and to 675 ms, TR=3.5s for the acquisition at 7T.
***Fitting Procedure Based on the EM algorithm*:** In the proposed method, the real part of the spectroscopic acquired spectrum is considered as a density probability function for which an estimator can be found using the Expectation Maximization algorithm [2]. Indeed, this algorithm is used for finding maximum likelihood estimates of parameters in probabilistic models with missing or hidden data. In case of Gaussian mixture model estimation, the unknown parameters are the means and variances and the probability for each Gaussian, which are respectively related to the frequency, linewidth and peak amplitude in MR spectroscopic terms. The EM alternates between two processes: an expectation (E) step, which computes an expectation assuming that unknown parameters were observed, and a maximization (M) step, which computes the maximum likelihood estimates of the wanted parameters. The parameters found during the M step are then used to begin another E step, and the process is repeated. Note that minimizing the distance between the original spectrum/signal data and a model function by a non linear least squares algorithm is usually done in MR spectroscopy quantification. This minimization also corresponds to a maximum-likelihood parameter estimation problem in case of assumed Gaussian distributed noise. The procedure consists in 5 successive steps implemented in MATLAB 7.4:
1) Zero-order phase correction of the signal after post-processing water removal.
2) Conversion of the spectrum data points into a density probability function, where each bin is a frequency value and with amplitude given by the spectrum intensity
3) Finding starting-values for the unknown parameter (mean, variance, probability) using the EM algorithm and a sliding window analysis of width 140 Hz at 7T, or 100Hz at 4.7T
4) Applying the EM algorithm on the frequency region of interest (0.2 to 4.5 ppm), 20 different results are obtained by randomly generates starting values around the ones obtained at step 3)
5) Determining the best parameter estimates (mean, variance, probability) by keeping the ones leading to the highest log likelihood value.
Results: Feasibility of the proposed procedure is illustrated on examples Figure 1 at 7T and Figure 2 at 4.7T. The Gaussian mixture model (blue) obtained with the proposed method was superimposed to the original spectrum (gray). The individual components (red) are also shown. Note that at 7T, one can retrieve published and identified resonances at 0.9, 1.24, 1.43, 1.72 and 2 ppm. [3]. The pattern from 2 to 2.3ppm is less reproducible. The region from 3.5 to 4 ppm shows also reproducible frequency estimates both at 4.7 and 7T.

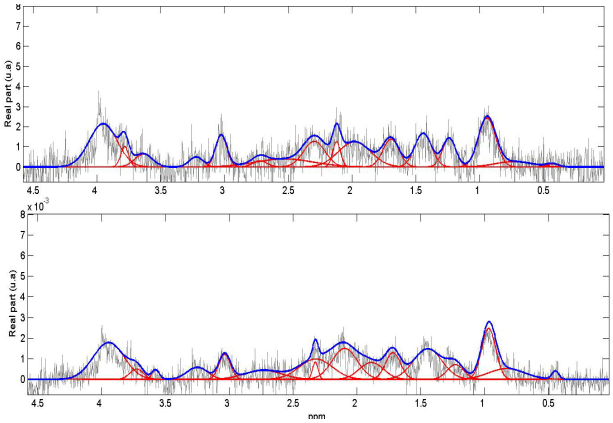

Figure1. 7T Gaussian mixture Model for rat brain macromolecular spectrum,

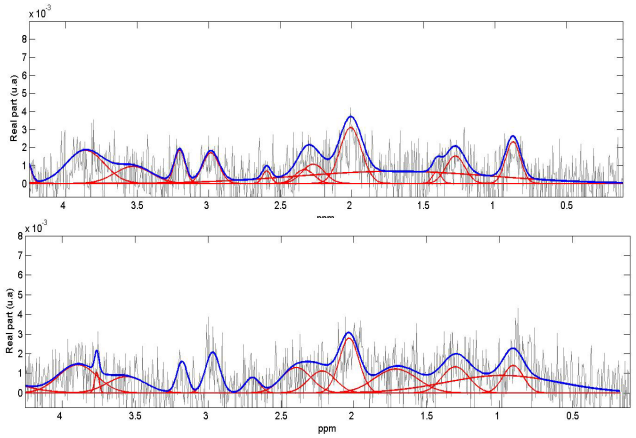

Figure2. 4.7T Gaussian mixture Model for mouse brain macromolecular spectrum

Discussion : A novel use of the EM algorithm has been proposed for the modeling and characterization of macromolecular signal obtained from an inversion recovery acquisitions.This method will provide a parameterized model helping background handling strategy in metabolite quantification method, that still needs some improvement [4]. Further refinement of this approach will also be developed by taking advantage from the analysis of the imaginary part of the spectrum. Other good properties of the EM algorithm will be investigated as it can provide an estimation strategy when the likelihood function is analytically intractable.

References:
[1] H. Ratiney *et al.*, Early and Progressive Disease Marker in MS; results from a large cross-sectional spectroscopic imaging study at 3T. In *American Academy of Neurology 59th Annual Meeting*, Boston, USA, May 2007.
[2] A.P.Dempster, *et al.,* Maximum-likelihood from incomplete data via the em algorithm. *J. Royal Statist. Soc. Ser. B.*, 39, 1977
[3] K.L.Behar *et al.,* Analysis of macromolecule resonances in 1H NMR spectra of human brain. *Magn Reson Med.* 1994 Sep;32(3):294-302.
[4] C. Cudalbu,*et al.,* Influence of the background-handling strategy on the metabolite concentration estimates. In *ISMRM proceedings*, Berlin, Germany, pages 3170, May 2007