Multi-exponential Transverse Relaxation Times Estimation from Magnetic Resonance Images under Rician Noise and Spatial Regularization

Christian El Hajj, Said Moussaoui, Guylaine Collewet, Maja Musse

Multi-exponential Transverse Relaxation Times Estimation from Magnetic Resonance Images under Rician Noise and Spatial Regularization

Christian EL HAJJ, Saïd MOUSSAOUI, Guylaine COLLEWET and Maja MUSSE

***Abstract*—Relaxation signal inside each voxel of magnetic resonance images (MRI) is commonly fitted by a multi-exponential decay curve. The estimation of a discrete multi-component relaxation model parameters from magnitude MRI data is a challenging nonlinear inverse problem since it should be conducted on the entire image voxels under non-Gaussian noise statistics. This paper proposes an efficient algorithm allowing the joint estimation of relaxation time values and their amplitudes using different criteria taking into account a Rician noise model, combined with a spatial regularization accounting for low spatial variability of relaxation time constants and amplitudes between neighboring voxels. The Rician noise hypothesis is accounted for either by an adapted nonlinear least squares algorithm applied to a corrected least squares criterion or by a majorization-minimization approach applied to the maximum likelihood criterion. In order to solve the resulting large-scale non-negativity constrained optimization problem with a reduced numerical complexity and computing time, an optimization algorithm based on a majorization approach ensuring separability of variables between voxels is proposed. The minimization is carried out iteratively using an adapted Levenberg-Marquardt algorithm that ensures convergence by imposing a sufficient decrease of the objective function and the non-negativity of the parameters. The importance of the regularization alongside the Rician noise incorporation is shown both visually and numerically on a simulated phantom and on magnitude MRI images acquired on fruit samples.**

***Index Terms*—Multi-exponential decay, Rician noise, Maximum Likelihood estimation, Majoration-Minimization, Levenberg-Marquardt, T2 Relaxation times**

I. Introduction

MAGNETIC Resonance Imaging (MRI) is frequently used in several domains for sample structure and composition monitoring. MRI is a reference modality for medical applications, however it is also used in non-medical contexts such as food science [1] and civil engineering [2]. Multiple Spin Echo (Multi-SE) is an acquisition technique used in MRI in order to get accurate characterization of samples by analyzing their transverse relaxation times (T_2). Multiple T_2-weighted images, at a fixed sampling rate, are acquired and an exponential decay curve is fitted to the measured relaxation signal at each voxel (a 3D pixel where the third dimension corresponds to the thickness of the MRI slice). These data are used for the evaluation of the spatial distribution (maps) of the relaxation times T_2 from the decay curves and their corresponding amplitudes A_0 [3]. Actually, the measured relaxation signal in each voxel can be modeled, either by a mono-exponential decay, or, in a more general case, by a multi-exponential decay. In the latter case, extracting the multi-exponential parameters at each voxel provides relevant information on the micro-structure and the composition of the tissue. Indeed, the voxel (usually of the order of 1 to few mm^3) generally contains different water pools with specific properties and/or different tissues, water and fat for example. In the case of plants, T_2 decay components correspond to water in the principal sub-cellular compartments (vacuole, cytoplasm, cell wall) of the tissue inside each voxel [4]. MRI can therefore be used to access information about water status and distribution at the subcellular level. This is useful for example for the assessment of fruit characteristics during the monitoring of the postharvest ripening of tomato fruit [5] or during apple storage [6]. Multi-exponential T_2 maps are also used for the quantification of myelin in the brain [7], where three major components can be identified; water trapped between the myelin, the intra/extracellular water and the cerebro-spinal fluid.

The relaxation time values inside each voxel can be modeled by a non-parametric T_2 distribution model [8], [9] which is sampled on a pre-defined grid of T_2 values. This approach has been used extensively for myelin quantification [10], [11]. Alternative models use parametric representation based on a finite number of T_2 components inside each voxel. Each component in this parametric model is represented by either a continuous distribution function (Gaussian [12] , inverse Gaussian [13] or gamma shape [14]) described by few parameters (location, width and amplitude) or by a single T_2 peak with unknown value and amplitude [15], [16]. We will focus this study on the latter case, assuming that the number of T_2 components is known while their locations and amplitudes are unknown. Indeed, it has been shown that the parametric discrete-model gives more accurate description in applications where the number of components within each voxel can be considered as known and of low value (bi and tri-exponential models are mostly used) [17]. However, such model requires the solving of a nonlinear estimation problem [18].

The acquisition of MRI data is carried out in the complex domain, where the real and the imaginary parts of the relaxation signal are corrupted by an independent and identically distributed Gaussian noise. Due to various factors (magnetic field inhomogeneity, thermal noise, eddy currents...), the phase

Christian EL HAJJ is with INRAE, OPAALE, F-35044, Rennes, France and Ecole Centrale Nantes, LS2N, F-44321, Nantes, France. Saïd MOUSSAOUI is with Ecole Centrale Nantes, LS2N, F-44321, Nantes, France. Guylaine COLLEWET and Maja MUSSE are with INRAE, OPAALE, F-35044, Rennes, France. Corresponding author: said.moussaoui@ec-nantes.fr

of the complex data usually presents both time and space variability. As a consequence, parameter estimation in the complex domain is a challenging problem [19], and the general approach used for estimating T_2 and A_0 parameters is to exploit the magnitude of the complex data [20], [21]. This nonlinear transformation of the data yields a modification of the noise statistics. The noise in conventional magnitude MRI data can be described by a Rician distribution [22], [23], which, when not properly accounted for by the estimation algorithm, introduces a bias on the estimated relaxation times and signal component amplitudes [24]. For a high signal to noise ratio (SNR) the Rician distribution approaches the Gaussian one [25] and the maximum likelihood (ML) estimator can be approached by minimizing a least squares (LS) criterion. However, in T_2 MRI relaxometry, signal samples at the end of the decay curve inevitably have low SNR and therefore the noise statistics deviate from the Gaussian distribution.

Several correction strategies have been originally suggested on the LS criterion in order to deal with the Rician distribution in the context of MRI signal denoising [26], [27], [21].

Their application for discrete multi-exponential model parameter estimation [28], [29] leads to what we call hereafter *Corrected Least Squares* (CLS) criteria. Finally, a ML estimator exploiting the Rician probability density function has been exploited by several works for signal denoising [19], [20], [30], T_1 relaxation parameter estimation [21] and discrete multi-exponential model parameters estimation [28], [29]. The ordinary LS and the CLS criteria are usually minimized using gradient-based descent methods or nonlinear least squares algorithms in order to estimate the parameters in each voxel separately [29] and their extension to the image level would benefit from the adding of a spatial regularization. Indeed, approaches based on processing each voxel separately do not take advantage of the image structure and suffer from high estimation variability [10], [31]. However, adding the spatial regularization makes the solution for each voxel dependent on the solution of the other voxels. Thus, solving the inversion problem by direct implementation of spatial regularization would be intractable due to heavy necessary amount of memory and computation time. In the case of non-parametric T_2 distributions, spatial regularization has been proposed [10], [9], [11], [32] as well as in combined diffusion and relaxation estimation [33], with adapted optimization algorithms for non-negative least squares estimation. It has also been proposed for a parametric multi-Gaussian model [31]. To our knowledge spatial regularization has not been proposed for the parametric discrete model framework.

The main purpose of this paper is to propose an efficient algorithm allowing the estimation of the discrete multi-exponential relaxation model parameters at the entire image level and to take into account the Rician noise statistics. In that respect, a Majorization-Minimization (MM) approach based on using a quadratic majorant of the ML criterion is used. This majoration strategy has been initially proposed by [30] in the context of MRI signal estimation under Rician and non-central chi distribution noise. The proposed adaptation to handle the case of a discrete multi-exponential model estimation is to carry out an approximate minimization of the underlying non-linear least squares criterion using an early stopped Levenberg-Maquardt (LM) algorithm. We provide a comparison between the ML approach and the minimization of CLS criteria using this adapted variant of the LM algorithm. The second contribution is to propose an efficient optimization algorithm which deals with the minimization of the criterion resulting from the inclusion of the spatial regularization. The MM approach is also adopted according to the strategy proposed in [34] for the majorization using separable functions and their early-stopped minimization using the modified version of the LM algorithm.

The rest of this paper is organized as follows. Section II introduces the estimation problem by presenting both the signal decay model and the noise statistics. Different criteria that can be minimized for the multi-exponential decay model parameter estimation are presented in section III. These criteria are discussed regarding the way how the non-Gaussian noise is accounted for and the proposed approach for incorporating spatial information. The proposed estimation algorithm is presented in section IV, and finally sections V and VI provide, respectively, a description of the materials used for the evaluation of the methods and a comparison of the different estimation approaches. The data used here are related to fruits which are the main target of our work and for which we have expertise regarding the expected values of T_2 and A_0. However, the proposed approach can be applied in medical MRI imaging as well.

II. Relaxation decay modeling

MRI data, measured with a Multi-SE sequence, is composed of N_τ images, each having a dimension of $[N_x \times N_y]$ which represent $N_v = N_x N_y$ voxels. The minimum time ΔTE between two successive echoes τ_t and τ_{t+1} is limited by the image acquisition sequence, and the duration τ_{N_τ} is chosen in order to measure the complete relaxation decay in all voxels.

A. Signal model

The measured signal y_{jt} in each voxel j at time τ_t is represented by a multi-exponential decay model s_{jt} given by:

$$s_{jt}\left(\boldsymbol{\theta}_j\right) = \sum_{c=1}^{N_c} A_{0_{(c,j)}} e^{-\frac{\tau_t}{T_{2_{(c,j)}}}}, \quad \text{for } t = 1,\ ...,\ N_\tau \tag{1}$$

where N_c represents the number of exponential decay components, the parameters (relaxation time T_2 and amplitude A_0) correspond to different components inside each voxel of the MRI image. The vector $\boldsymbol{\theta}_j = [A_{0_{(1,j)}}\ ,\ T_{2_{(1,j)}}\ ...\ A_{0_{(N_c,j)}}\ ,\ T_{2_{(N_c,j)}}]$ contains the unknown parameters at voxel j and is of length $N_p = 2N_c$. This model is based on the hypothesis that indirect and stimulated echoes do not interfere with the acquired signal [35].

For a given voxel j, the real and imaginary parts of the complex MRI data at time τ_t can be modeled as follows:

$$y_{jt}^{re} = s_{jt}\cos\left(\phi_{jt}\right) + n_{jt}^{re} \tag{2}$$

$$y_{jt}^{im} = s_{jt}\sin\left(\phi_{jt}\right) + n_{jt}^{im} \tag{3}$$

here s_{jt} is the relaxation signal amplitude and ϕ_{jt} the phase of the complex signal at voxel j at time τ_t. n_{jt}^{re} and n_{jt}^{im}

represent the additive Gaussian noise at voxel j on the real and imaginary parts, respectively. In order to overturn the problem of estimating the time dependent phase [36], parameters estimation is carried out on the magnitude of the complex data:

$$y_{jt} = \sqrt{(y_{jt}^{re})^2 + (y_{jt}^{im})^2}. \quad (4)$$

B. Noise statistics

In conventional MRI, the noise in the magnitude data is modeled using a non-central chi distribution which reduces to a Rician distribution in certain cases. Conventional MRI is considered here by assuming that the data are acquired using a regular cartesian sampling, that the different sources of noise are independent and that no post-processing schemes are applied [37]. This excludes for example parallel MRI which is used to reduce the acquisition time. The probability density function of a Rician distribution is given by:

$$P_R\left(y_{jt} \mid s_{jt}, \sigma_j\right) = \frac{y_{jt}}{\sigma_j^2} e^{-\frac{y_{jt}^2 + s_{jt}^2}{2\sigma_j^2}} I_0\left(\frac{y_{jt} s_{jt}}{\sigma_j^2}\right) \mathbb{I}_{y_{jt}>0}, \quad (5)$$

where I_0 is the zeroth order modified Bessel function of the first kind and $\mathbb{I}_{x\geq 0}$ stands for the indicator function. σ_j^2, the noise variance, does not depend on τ_t. Figure 1 illustrates the Rician distribution shape for $\sigma_j = 7$ and different SNR levels, expressed by (s_{jt}/σ_j). The dashed vertical lines represent the projection of the actual value of (s_{jt}/σ_j) on its corresponding distribution function showing its dependance on the distribution mode. It can be clearly observed that as the SNR becomes higher, the Rician distribution approaches the Gaussian distribution. However, as mentioned in the introduction, low SNR level always occurs in relaxometry for high values of τ_t. In this case, the noisy data deviates from the underlying relaxation signal as illustrated in Figure 2 and the data distribution becomes asymmetric.

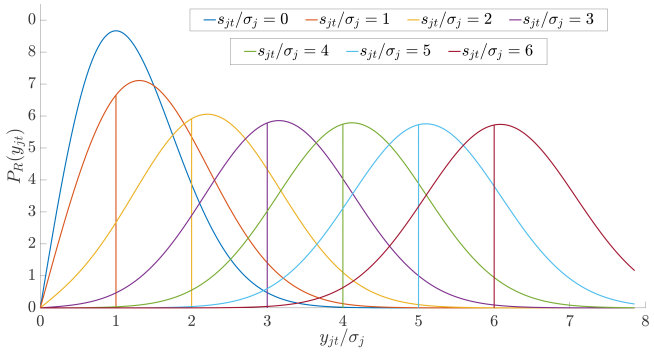

Fig. 1. Rician distribution shape for different values of $\frac{s_{jt}}{\sigma_j}$ and $\sigma_j = 7$.

III. Parameter Estimation Criteria

Different criteria can be used for the estimation of the relaxation model parameters (relaxation times and amplitudes) from the measured signals. Actually, these criteria differ in the way how the noise statistics are accounted for in order to overcome the limitation of the Gaussian noise hypothesis.

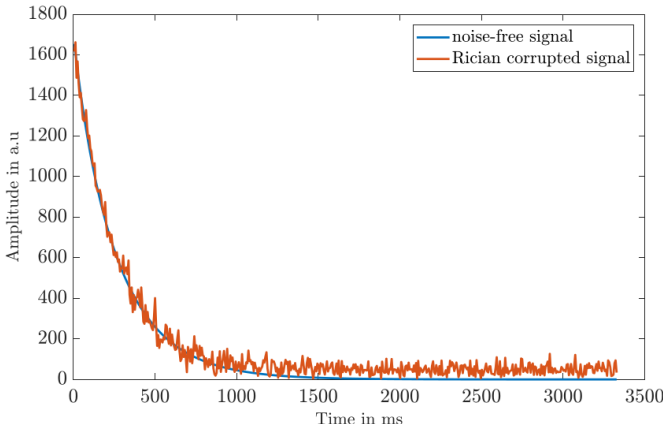

Fig. 2. Effect of the Rician distribution on the data. The underlying signal is shown in blue and the noisy data is shown in red.

A detailed analysis of the bias and the variance of these estimators for MRI signal estimation is given by [36]. In the sequel, it will be assumed that the noise variance σ_j^2, in each voxel is known or has been previously estimated using a dedicated algorithm (See for instance the survey of [38] and the references therein).

A. Ordinary Least Squares

The classical hypothesis of Gaussian noise comes down to minimizing the LS criterion which is the squared difference between the fitted signal according to the multi-exponential model and the measured data. For each voxel j the quadratic criterion C_{LS_j} to be minimized is given by:

$$C_{LS_j}(\boldsymbol{y}_j, \boldsymbol{\theta}_j) = \frac{1}{2\sigma_j^2} \sum_{t=1}^{N_\tau} \left[y_{jt} - s_{jt}\left(\boldsymbol{\theta}_j\right)\right]^2, \quad (6)$$

where $\boldsymbol{y}_j$ corresponds to a vector containing the N_τ samples of the decay signal in voxel j.

B. Corrected Least Squares

Different modifications have been introduced to the LS criterion in order to reduce the bias due to the Rician distribution of the noise [20]. One technique consists of adding a constant offset to the imposed model in order to account for the deviation from the underlying signal at low SNR [39]. However, this technique assumes a constant shift over the scope of the decay curve which is not precisely the case for high SNR. Another approach consists in weighting the LS squares criterion [40] in order to assign smaller weights for signal samples at low SNR. A special case of this approach is to set weights to zero for signal samples below a given SNR threshold. However, the difficulty lies in the choice of an appropriate weighting strategy whatever the signal level. Alternatively, there are some modifications of the LS criterion that take explicitly the Rician statistics into account. We will present the ones that have been used in the literature in the context of T_2 estimation.

1) McGibney Corrected LS: In order to cope with the Rician distribution, McGibney and Smith [26] and Miller and Joseph [41] proposed independently to use the second order moment of the decay signal which is given by (a detailed formulation is provided in appendix B1):

$$E_{Rice}\left[y_{jt}^2\right] = s_{jt}^2 + 2\sigma_j^2. \tag{7}$$

Based on this equation, the so-called McGibney corrected LS (MGCLS) criterion is given by:

$$C_{MGCLS_j}(\boldsymbol{y}_j, \boldsymbol{\theta}_j) = \sum_{t=1}^{N_\tau} \left[\, y_{jt}^2 - \left(s_{jt}\left(\boldsymbol{\theta}_j\right)^2 + 2\sigma_j^2\right)\right]^2. \tag{8}$$

2) Rician Expectation Corrected LS: By considering the first order moment of the noise distribution, the estimation bias caused by the Rician distribution can be reduced by minimizing the squared difference between the measured signal and the Rician expectation value of the multi-exponential decay model [21], [24]. The Rician expectation value of the measured signal y_{jt}, knowing the noise standard deviation σ_j and the signal amplitude s_{jt}, is given by:

$$\begin{aligned} &E_{Rice}[y_{jt}] \\ &= \sigma_j \sqrt{\frac{\pi}{2}} e^{-\alpha_{jt}}[(1+2\alpha_{jt})I_0(\alpha_{jt}) + 2\alpha_{jt} I_1(\alpha_{jt})], \end{aligned} \tag{9}$$

with $\alpha_{jt} = \frac{s_{jt}^2}{4\sigma_j^2}$ and I_1 the modified Bessel function of the first order [21]. Therefore, the Rician expectation corrected LS criterion, noted C_{RECLS_j}, is given by:

$$C_{RECLS_j}(\boldsymbol{y}_j, \boldsymbol{\theta}_j) = \frac{1}{2\sigma_j^2} \sum_{t=1}^{N_\tau} \left[y_{jt} - f\left(s_{jt}\left(\boldsymbol{\theta}_j\right), \sigma_j\right)\right]^2, \tag{10}$$

with $f\left(s_{jt}\left(\boldsymbol{\theta}_j\right), \sigma_j\right) = E_{Rice}[y_{jt}]$.

C. Maximum Likelihood

Under the noise stationarity and the Rician distribution assumption, the Likelihood function L_j is given by:

$$L_j(\boldsymbol{y}_j, \boldsymbol{\theta}_j) = \prod_{t=1}^{N_\tau} P_R\left(y_{jt} \mid s_{jt}\left(\boldsymbol{\theta}_j\right), \sigma_j\right). \tag{11}$$

Its maximization is equivalent to the minimization of the negative log-likelihood function:

$$\begin{aligned} -\log\left(L_j(\boldsymbol{y}_j, \boldsymbol{\theta}_j)\right) &= \sum_{t=1}^{N_\tau} \left[\frac{y_{jt}^2}{2\sigma_j^2} - \log\left(\frac{y_{jt}}{\sigma_j^2}\right)\right] \\ &+ \sum_{t=1}^{N_\tau} \left[\frac{s_{jt}\left(\boldsymbol{\theta}_j\right)^2}{2\sigma_j^2} - \log\left(I_0\left(z_{jt}\right)\right)\right], \end{aligned} \tag{12}$$

where $z_{jt} = \dfrac{y_{jt} s_{jt}\left(\boldsymbol{\theta}_j\right)}{\sigma_j^2}$.

Therefore, maximizing the likelihood for parameter estimation from the measured relaxation signal comes down to minimizing the ML criterion given by:

$$C_{ML_j}\left(\boldsymbol{y}_j, \boldsymbol{\theta}_j\right) = \sum_{t=1}^{N_\tau} \left[\frac{s_{jt}\left(\boldsymbol{\theta}_j\right)^2}{2\sigma_j^2} - \log\left(I_0\left(z_{jt}\right)\right)\right]. \tag{13}$$

Although this estimator presents desirable statistical properties, the minimization of the underlying criterion is more challenging than the LS based criteria.

D. Accounting for the spatial dimension

The signals inside different voxels of the image are not totally mutually independent. Actually, coherent regions are found inside MRI images that share similarities in decay curve parameters. Taking into account the image structure and constitution is crucial for both computational time reduction and ensuring parameter estimation stability.

1) Considering simultaneously the image voxels: The criteria discussed in section III are classically minimized for each voxel independently from others. This does not take into account the image structure of the MRI data. We suggest expanding the criteria by adding a summation over the entire scope of the image. The criteria to be minimized, thus becomes, for any of the previously defined criteria C_j:

$$C\left(\boldsymbol{y}, \boldsymbol{\theta}\right) = \sum_{j=1}^{N_v} C_j\left(\boldsymbol{y}_j, \boldsymbol{\theta}_j\right), \tag{14}$$

where $\boldsymbol{y}$ is a vector containing the measurement at all the voxels of the MRI image and $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ ..., \ \boldsymbol{\theta}_{N_v}]$. In a numerical computing point of view, this reformulation may render the estimation process faster than by minimizing the criteria at each voxel independently. However, the main advantage of this reformulation is the possibility to deal with the ill-posed inverse problem more efficiently by adding spatial regularization.

2) Spatial regularization: Finding the relaxation model parameters by minimizing any of the previously described criteria is a highly ill-posed inverse problem. The low SNR and the high number of parameters implies a solution space with many local minima which affect the stability of the estimation process. Spatial regularization methods proved to be efficient in enforcing numerical stability of the algorithm and reducing the estimation uncertainty [10], [9]. The coherent regions usually found in the MRI images imply slow variation of the relaxation model parameters between adjacent voxels with the exception of voxels located at borders between two different tissues. Spatial regularity can be accounted for by adding a regularization term to the criteria in such a way that high differences in parameter values between adjacent voxels are disadvantaged. Thus, for each voxel j we define a neighboring voxel set V_j (for example : a window of 3×3 voxels) and impose a penalty function ψ on the difference between the parameters:

$$R(\boldsymbol{\theta}) = \sum_{p=1}^{N_p} \beta_p \sum_{j=1}^{N_v} \sum_{k \in V_j} \psi\left(\theta_j\left(p\right) - \theta_k\left(p\right)\right), \tag{15}$$

where $(\beta_1, \ldots, \beta_{N_p})$ is a set of regularization parameters allowing to adjust the influence of the regularization term in the global criterion. In order to equally penalize the roughness of the different parameters, a smaller β_p for parameters with high magnitudes might be chosen.

Accounting for a data fitting measure, expressed by a least squares or a maximum likelihood criterion, and the regularization term favoring spatial smoothness, the penalized estimation approach consists in solving the optimization problem:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\geq 0} F(\boldsymbol{y}, \boldsymbol{\theta}) \tag{16}$$

with a penalized criterion $F(\boldsymbol{y}, \boldsymbol{\theta}) = \mathcal{C}(\boldsymbol{y}, \boldsymbol{\theta}) + R(\boldsymbol{\theta})$

IV. Proposed Estimation Algorithm

The minimization of the penalized criterion F by solving problem (16) does not admit an analytic solution. Hence, an iterative optimization approach must be adopted for the estimation of the relaxation model parameters. Moreover, the minimization over the whole image leads to a large-scale optimization problem. We propose to use a quadratic majorization-minimization approach coupled with the LM algorithm. Adaptations are introduced to ensure the positivity of the estimated parameters and to perform the optimization on the whole image. More precisely, the majorization-minimization approach is used to overturn the non-quadratic nature of the criterion (13) and to reduce the computation load by ensuring separability of the variables during the minimization of the regularization term (15).

A. Levenberg-Marquardt Algorithm

The LM algorithm [42], [43] is widely used for solving nonlinear least squares problems. It is an iterative descent method that exploits the gradient of the criterion and a local approximation of its Hessian at each iterate. Firstly, a residual vector $\boldsymbol{r}(\boldsymbol{\theta})$ is expressed in such a way that the criterion F can be written as $F(\boldsymbol{\theta}) = \frac{1}{2}||\boldsymbol{r}(\boldsymbol{\theta})||_2^2$. At each iteration ℓ of the LM algorithm, the parameters are updated in the direction $\boldsymbol{d}^{(\ell)}$:

$$\boldsymbol{\theta}^{(\ell+1)} = \boldsymbol{\theta}^{(\ell)} + \alpha^{(\ell)}\, \boldsymbol{d}^{(\ell)}, \tag{17}$$

where $\alpha^{(\ell)}$ is a step-size allowing to control the algorithm convergence. The LM descent direction at each iteration ℓ is calculated as follows:

$$\boldsymbol{d}^{(\ell)} = -\left[\boldsymbol{J}(\boldsymbol{\theta}^{(\ell)})^T \boldsymbol{J}(\boldsymbol{\theta}^{(\ell)}) + \lambda^{(\ell)}\, \boldsymbol{I}\right]^{-1} \boldsymbol{g}(\boldsymbol{\theta}^{(\ell)}), \tag{18}$$

where $\boldsymbol{J}(\boldsymbol{\theta})$ and $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{J}(\boldsymbol{\theta})^T \boldsymbol{r}(\boldsymbol{\theta})$ are respectively the Jacobian matrix of $\boldsymbol{r}(\boldsymbol{\theta})$ and the gradient of F at $\boldsymbol{\theta}$. The damping parameter $\lambda^{(\ell)}$ is updated at each iteration in order to ensure an adequate balance between gradient-like descent (stability) and Gauss-Newton update (quadratic convergence rate) [44], [45].

The detailed expressions of the gradient and the Jacobian in the case of the least squares criteria and the multi-exponential decay model are given in appendix A, B2 and C.

The applications of the LM algorithm to the analysis of multi-exponential decay model in the whole image space is carried out according to the following improvements:

- **Accounting for image structure:** In the LM algorithm the heaviest computation step is the inversion of the first term in equation (18). Actually, by applying the LM algorithm on the whole image at once, we obtain a Jacobian of size $N_\tau \times (N_p N_v)$, thus a matrix of size $N_p N_v \times N_p N_v$ to inverse. However, by considering the fact of variable separability (the solutions are independent from one voxel to another) the matrix to inverse becomes block diagonal. The inversion is therefore performed on N_v blocks of size $N_p \times N_p$. This formulation reduces considerably the computational heaviness and memory usage [46].
- **Choice of the step-size:** A linesearch strategy based on a backtracking technique to satisfy the Armijo condition [45] is used in order to compute an adequate step-size $\alpha^{(\ell)}$ that ensures a sufficient decrease of the criterion $F(\boldsymbol{\theta})$:

  $$F(\boldsymbol{\theta}^{(\ell+1)}) \leq F(\boldsymbol{\theta}^{(\ell)}) + c_1 \alpha^{(\ell)} \boldsymbol{g}^{(\ell)^T} \boldsymbol{d}^{(\ell)} \text{ with } c_1 \in]0, 1/2].$$

  A typical value for c_1 is 10^{-3}. The backtracking is started from the largest step-size α_{max} that preserves the non-negativity of the parameters. It is obtained using the following expression:

  $$\alpha_{\text{max}} = \min_{k\in\mathbb{K}} -\theta_k^{(\ell)}/d_k^{(\ell)} \text{ with } \mathbb{K} = \{k; d_k^{(\ell)} < 0\}.$$

The main steps of the modified LM algorithm for multi-exponential decay model parameter estimation for the whole image voxels simultaneously is given by Algorithm 1. The algorithm is considered to have converged either if the relative variation of the cost function between two successive iterations is less then a small value ϵ_f or when the gradient norm is below some threshold ϵ_g [45].

Algorithm 1: Summary of the main steps of the modified LM algorithm including an Armijo linesearch

```
Input: θ^(0), ρ ∈]0;1[ ;
for ℓ ← 0 to ℓ_max or until convergence do
    Compute r(θ^(ℓ)) and J(θ^(ℓ));
    Update λ^(ℓ);
    Compute d^(ℓ);
    Compute the maximum step size α_max ;
    Set α^(0) = α_max;
    for n ← 0 to n_max or until Armijo condition is
     met do
        Set α^(n+1) ← ρ α^(n);
        Check Armijo condition;
    end
    Set α^(ℓ) = α^(n+1);
    Update θ^(ℓ+1) = θ^(ℓ) + α^(ℓ) d^(ℓ);
end
Output: θ^(ℓ+1)
```

B. Likelihood Majorization-Minimization

The minimization of a given criterion using the MM approach is carried out by a succession of approximate minimization of majorant functions easier to minimize than the initial criterion [47], [48]. MM algorithms operate by exploiting the properties of the original criterion in terms of convexity and concavity. Let us for instance consider a convex criterion F

defined for $\boldsymbol{x} \in \mathrm{dom}_f$. At each iteration m of the MM algorithm [48], there exists a function $M^{(m)}$ such that:

$$\begin{aligned} &M^{(m)}(\boldsymbol{x}^{(m)}) = F(\boldsymbol{x}^{(m)}), \\ &M^{(m)}(\boldsymbol{x}) \geq F(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \mathrm{dom}_f \end{aligned} \tag{19}$$

$M^{(m)}$ is called a tangent majorant of F at $\boldsymbol{x}^{(m)}$. The next iterate $\boldsymbol{x}^{(m)}$ is generally taken as the minimizer of $M^{(m)}(\boldsymbol{x})$. However, the exact minimization of the majorant function is not necessary since the main requirement for theoretical convergence is to ensure the iterative decrease of the minimized criterion [49], [50].

Actually, the criterion C_{MLj} can be rewritten as $\sum_{t=1}^{N_\tau}\left(C_1\left(s_{jt}\right)+C_2\left(s_{jt}\right)\right)$, a sum of two functions; a strictly convex quadratic function:

$$C_1\left(s_{jt}\right) = \frac{s_{jt}^2}{2\sigma_j^2}, \tag{20}$$

and a strictly concave function:

$$C_2\left(s_{jt}\right) = -\log\left(I_0\left(z_{jt}\right)\right). \tag{21}$$

The later function can be majorized [30] by its tangent at any point $\tilde{s}_{jt}$ according to:

$$\begin{aligned} -\log\left(I_0\left(z_{jt}\right)\right) \leq &-\log\left(I_0\left(\tilde{z}_{jt}\right)\right) \\ &-\frac{y_{jt}}{\sigma_j^2} R\left(\tilde{z}_{jt}\right)\left(s_{jt}-\tilde{s}_{jt}\right), \end{aligned} \tag{22}$$

where $R(\cdot)=\dfrac{I_1(\cdot)}{I_0(\cdot)}$ and $\tilde{z}_{jt}=\dfrac{y_{jt}\tilde{s}_{jt}}{\sigma_j^2}$. Such majorization leads to:

$$\begin{aligned} C_1\left(s_{jt}\right)+C_2\left(s_{jt}\right) \leq &\frac{s_{jt}^2}{2\sigma_j^2} - \frac{\tilde{y}_{jt}}{\sigma_j^2}\, s_{jt} \\ &-\log\left(I_0\left(\tilde{z}_{jt}\right)\right) + \frac{\tilde{y}_{jt}}{\sigma_j^2}\tilde{s}_{jt}. \end{aligned} \tag{23}$$

where $\tilde{y}_{jt} = R\left(\tilde{z}_{jt}\right) y_{jt}$. The last two terms are independent from $\boldsymbol{\theta}_j$ and consequently, by setting $\tilde{s}_{jt} = s_{jt}(\boldsymbol{\theta}_j^{(m)})$, the majorant criterion M_{ML} of the maximum likelihood criterion C_{ML} resulting from the Rician noise distribution model can be approximated at each iteration m of the MM algorithm by:

$$M_{ML}^{(m)}\left(\boldsymbol{y},\boldsymbol{\theta}\right) = \sum_{t=1}^{N_\tau}\sum_{j=1}^{N_v} \frac{1}{2\sigma_j^2}\left[\tilde{y}_{jt} - s_{jt}\left(\boldsymbol{\theta}_j\right)\right]^2. \tag{24}$$

It can be noticed that the MM strategy leads to a nonlinear least squares criterion whose minimization can be performed using few steps of the LM algorithm, as the exact minimization is not required. Our proposal is to alternate iteratively between updating the majorant function and few steps of the LM algorithm. The resulting Majorize-Minimize Maximum Likelihood (MM-ML) algorithm is summarized by Algorithm 2.

Algorithm 2: The main steps of the MM-LM algorithm for ML estimation

Input: $\boldsymbol{\theta}^{(0)}, \boldsymbol{y}$;
for $m \leftarrow 0$ **to** $m_{\max}$ or until convergence **do**
  Set $\tilde{s}_{jt} = s_{jt}(\boldsymbol{\theta}_j^{(m)})$, $\forall j = 1,\ldots,N_v$;
  Compute $\tilde{y}_{jt} = R\left(\tilde{z}_{jt}\right) y_{jt}$, $\forall j = 1,\ldots,N_v$;
  Minimize $M_{ML}^{(m)}\left(\boldsymbol{y},\boldsymbol{\theta}\right)$ using Algorithm 1 with the following settings:
  - Set ℓ_{max} to a small number;
  - Input : $\boldsymbol{\theta}^{(m)}$;
  - Output : $\boldsymbol{\theta}^{(m+1)}$;
end
Output $\boldsymbol{\theta}^{(m+1)}$

C. Separable Majorization-Minimization

Adding the spatial regularization introduces a coupling between the parameters, which breakdowns the separability of the optimisation with respect to different voxels of the image. Actually, the separability is important to reduce the computation burden for Jacobian evaluation and matrix inversion for LM direction calculation in Algorithm 1. The MM approach can be used to alleviate this limitation by setting a voxel wise separable majorant functions. Based on convexity properties [47], by setting a symmetric convex penalization function ψ, a separable majorant function can be obtained. For any variables (u, v) from the parameters vector $\boldsymbol{\theta}$, it has been established by [34] that:

$$\psi\left(u-v\right) \;\leq\; \frac{1}{2}\psi\left(2u-\delta_{uv}^{(m)}\right) + \frac{1}{2}\psi\left(2v-\delta_{uv}^{(m)}\right), \tag{25}$$

with $\delta_{uv}^{(m)} = \left(u^{(m)}+v^{(m)}\right)$. The majorant function is thus a sum of two separable convex functions. Applying this strategy to the regularization criterion on the whole image voxels leads to the following majorant:

$$M_{Reg}^{(m)}\left(\boldsymbol{\theta}\right) = \sum_{p=1}^{N_p}\beta_p \sum_{j=1}^{N_v}\sum_{k\in V_j} \psi\left(2\theta_j\left(p\right) - \delta_{jk}^{(m)}(p)\right), \tag{26}$$

where $\delta_{jk}^{(m)}(p) = \left(\theta_j^{(m)}\left(p\right)+\theta_k^{(m)}\left(p\right)\right)$ and $\theta_j^{(m)}(p)$ stands for the p-th parameter inside the j-th voxel at iteration m of the MM algorithm. When a quadratic function ψ is used, the separable majorant M_{Reg} will also be quadratic. An additional reformulation will be involved in the case of non-quadratic function in order to get a nonlinear least squares form according to:

$$M_{Reg}^{(m)}\left(\boldsymbol{\theta}\right) = \sum_{p=1}^{N_p}\sum_{j=1}^{N_v}\sum_{k\in V_j} \left(\beta_p^{\frac{1}{2}}\psi^{\frac{1}{2}}\left(2\theta_j\left(p\right) - \delta_{jk}^{(m)}(p)\right)\right)^2, \tag{27}$$

with the requirement that ψ takes non-negative values, which is the case in classical regularization functions [51].

D. Optimization algorithm using the MM-LM approach

By combining either the least squares-based criteria (C_{LS}, C_{MGCLS} or C_{RECLS}) or the majorant function of the ML criterion M_{ML} and the majorant function M_{Reg} of the regularization term, the Minimization-Majorization algorithm for penalized criteria optimization is summarized in Algorithm 3.

Algorithm 3: The proposed MM-LM algorithm for penalized criteria minimization

Initialize $\boldsymbol{\theta}^{(0)}$;
for $m \leftarrow 0$ **to** $m_{\max}$ or until convergence **do**
  Compute $\delta_{jk}^{(m)}(p)$ for all j and k ;
  Define $M_{Reg}^{m}(\boldsymbol{\theta})$;
  if $C = F_{ML}$ **then**
    Compute $\tilde{y}_{jt} = R\big(z_{jt}^{(m)}\big) y_{jt}$;
    Set $M_{Q}^{(m)}(\boldsymbol{y}, \boldsymbol{\theta}) = M_{ML}^{(m)}(\boldsymbol{y}, \boldsymbol{\theta})$;
  end
  else
    Set $M_{Q}^{(m)}(\boldsymbol{y}, \boldsymbol{\theta}) = C(\boldsymbol{y}, \boldsymbol{\theta})$
  end
  Minimize $M_{Q}^{(m)}(\boldsymbol{y}, \boldsymbol{\theta}) + M_{Reg}^{(m)}(\boldsymbol{y}, \boldsymbol{\theta})$ using Algorithm 1 with following settings:
  - Set ℓ_{max} to a small number
  - Input : $\boldsymbol{\theta}^{(m)}$;
  - Output : $\boldsymbol{\theta}^{(m+1)}$;
end
Output $\boldsymbol{\theta}^{(m+1)}$

V. Performance Evaluation Setup

The performances of the multi-exponentional model estimation using different criteria are assessed using synthetic and experimental MRI images which are presented in this section. A three-exponential model ($N_c = 3$) is used since such model seems to be relevant to describe the different water pools in plant tissues [6]. The same number of components is used in myelin fraction quantification [52]. However, the proposed estimation algorithm can be applied for different number of components.

A. Simulated data

The empirical analysis of the algorithm performance is performed using a phantom constituted by a circular disk (Region 2) within an image of 128×128 voxels, with an outer ring (Region 1) and 9 inner disks simulating three internal regions as represented in table I. Within each of the five regions of the phantom, a different three-exponential model was generated. The values of T_2 and A_0 parameters are set to be close to typical values found in tomato fruits. A stationary Rician noise acquisition model with $\sigma_j = \sigma \forall j$ and with different σ values is set in order to evaluate the algorithms under different noise levels. The SNR evaluated for the images acquired at the first echo time is computed as follows:

$$\text{SNR} = \frac{\sum_{j=1}^{N_v} y_{j1}}{\sigma N_v}. \tag{28}$$

Depending on the number of averaged scans in the k-space, typical values of SNR in real experiments varies from 150 to 500.

B. Experimental data

The experimental MRI settings have been obtained using a Multi-SE sequence on a 1.5T MRI scanner (Magnetom, Avanto, Siemens, Erlangen, Germany) [3], with inter-echo spacing (ΔTE) of 6.5 ms, bandwidth of 260 Hz/pixel, 512 echoes per echo train and a repetition time of 10s. The median planes of fruit (transverse section at middle height of fruit) were imaged with a total of 128×128 voxels and a slice thickness of 5 mm, resulting in voxel size of $1.19 \times 1.19 \times 5$ mm^3. A head-array coil composed of 4 elements was used, the images of which were combined using a spatial matched filter [53]. Data were acquired in a regular cartesian scheme and no post-processing was applied. A first set of images was acquired with only one scan, which corresponds to a SNR of 145 and to realistic acquisition times for the targeted application (21 minutes). In order to get reference maps we could use as "ground truth", a set of images with high SNR (687) was acquired using 32 scans. The average of the multiple scans was computed in the complex space thus maintaining the same noise statistic for 1 and 32 scans.

TABLE I
Simulated phantom image and values of the (A_0, T_2) parameters in each region.

No.	Region	A_{01}	T_{21}	A_{02}	T_{22}	A_{03}	T_{23}
1	Outer ring	95	88	459	356	1015	716
2	Inner medium	96	76	410	433	1024	870
3	Central disc	71	50	339	218	908	627
4	Left-right discs	109	78	375	303	1004	685
5	Top-down discs	108	50	482	202	756	508

C. Algorithm settings

In order to implement the regularization, we had to choose the regularization function ψ and the regularization weight vector $\boldsymbol{\beta}$. The choice of the regularization function affects the inter-tissues smoothness, whilst the regularization weight affects the degree of smoothness on the whole image. In this study, we used a quadratic penalization function with a weight $\boldsymbol{\beta} = [0.4,\ 0.4,\ 0.06,\ 0.06,\ 0.01,\ 0.01]$ chosen by trial and error so as to preserve a good separation between image components. The algorithm stopping criteria was set to $\epsilon_f = \epsilon_g = 10^{-6}$.
Results were evaluated using the normalized root mean square error, NRMSE, computed as follows:

$$\text{NRMSE} = 100 \sqrt{\frac{1}{N_v N_p} \sum_{j=1}^{N_v} \sum_{p=1}^{N_p} \frac{\left(\theta_j(p) - \theta_j^*(p)\right)^2}{\theta_j^*(p)^2}} \tag{29}$$

with $\theta_j^*(p)$ the reference parameters at voxel j for parameter p.

D. Noise Variance estimation

In practice, in the case of a Rician noise stationary in the spatial domain, the noise standard-deviation can be estimated from regions where no signal is present such as the background or, in the relaxation acquisition framework, the images

acquired at the highest τ_t [54]. Actually, if $s_{jt} = 0$, the Rician distribution leads to the Rayleigh distribution:

$$P_R(y_{jt}|s_{jt} = 0, \sigma_j) = \frac{y_{jt}}{\sigma_j^2} e^{-\frac{y_{jt}^2}{2\sigma_j^2}} \mathbb{I}_{y_{jt}>0}, . \quad (30)$$

Therefore, the maximum likelihood estimator of the noise standard deviation σ_j in each voxel can be obtained according to:

$$\sigma_j = \sqrt{\frac{2}{\pi}} \frac{1}{N} \sum_{t=N_\tau-N+1}^{N_\tau} y_{jt}, \quad (31)$$

where N is the number of decay samples y_{jt} with signal value theoretically equal to zero. More sophisticated noise estimation strategies can be envisaged such as those based on local statistics in MRI data [55]. In the acquisition configuration considered in this paper, and without loss of generality, the noise stationarity was checked and the noise standard deviation was estimated from the background voxels.

E. Estimation criteria

The proposed optimization algorithm has been firstly applied in the case of non-regularized criteria to discuss the effectiveness of the proposed optimization approach in addition to a comparative analysis of the impact of how the Rician distribution is accounted for. In a second step, the spatial regularization was tested on the two approaches that yielded the lowest NRMSE, which turned out to be the RECLS and the ML approach (defined hereafter).

The different algorithms that were evaluated on the phantom are summarized below:

1) LS: LM algorithm for minimizing the LS criterion C_{LS},
2) MGCLS: LM algorithm for minimizing the McGibney Corrected LS criterion C_{MGCLS},
3) RECLS: LM algorithm for minimizing the LS corrected by taking the Rician Expectation value C_{RECLS},
4) ML: MM-LM algorithm for minimizing the ML criterion C_{ML},
5) PRECLS: MM-LM algorithm for minimizing the penalized LS corrected by taking the Rician Expectation value F_{RECLS},
6) PML: MM-LM algorithm for minimizing the penalized ML criterion F_{ML}.

The initial values of the parameters are taken from a preliminary estimation from a region of interest manually delimited on a tomato where the $(A_0,\ T_2)$ values were estimated on the average signal. Based on these results the following initial vector is chosen:

$$\boldsymbol{\theta}_j^{(0)} = [(95, 75),\ (425, 365),\ (600, 605)], (\forall\ j = 1, \ldots, N_v).$$

VI. Performance Evaluation Results

A. Algorithm convergence

The evolution of the F_{ML} criterion during the LM iterations of the optimization algorithm is shown in figure 3. A comparison is made between choosing a high number of LM iterations (inner loop with $\ell_{\max} = 10$) and choosing a low number ($\ell_{\max} = 2$). This choice does not affect the overall convergence of the algorithm, and subsequently the results, but it however impacts the speed of convergence of the algorithm. As a result, we can clearly see that by choosing a high number of internal LM iteration, the algorithm converges with a lower number of MM steps but with a higher computation time. Furthermore, it is clearly shown that by decreasing the majorant criterion, $(M_{ML} + M_{reg})$, the descent of the majorized criterion is guaranteed which validates the choice of the majorant functions.

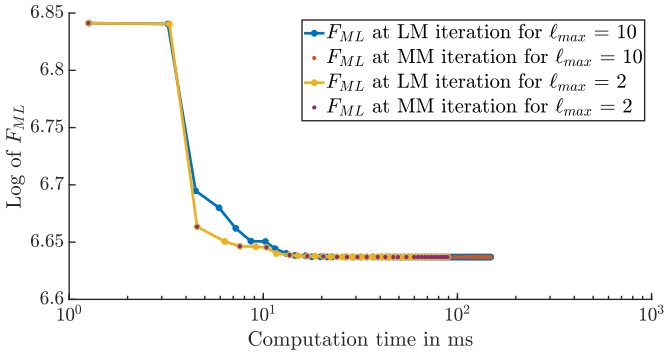

Fig. 3. Evolution of the F_{ML} criterion across the external MM steps (orange and purple points) and the internal LM steps (blue and yellow points)

B. Results on the virtual phantom

1) Estimation without regularization: Table II summarizes the NRMSE of the estimation using the four non regularized criteria.

TABLE II
NRMSE ON THE PHANTOM FOR EACH OF THE FOUR NON-REGULARIZED ALGORITHMS LS, MGCLS, RECLS AND ML FOR DIFFERENT SNR

	SNR=150	SNR=300	SNR=450
LS	125.1	40.0	17.4
MGCLS	36.4	24.9	23.3
RECLS	35.2	15.1	11.5
ML	34.8	15.2	11.7

As expected, the LS criterion leads to the highest error especially at low SNR since it does not take the Rician noise into account. MGCLS method gave better results than LS except for high SNR but under-performed compared to ML and RECLS methods. These latter two methods gave very close results and provided the lowest error between the non-regularized methods across the entire scope of the image. It can be explained by the fact that these methods explicitly account for the noise statistics. In the sequel, the analysis will be focused on these these two methods and an improvement of performances will be achieved by accounting for the spatial regularization.

2) Spatial regularization: Figure 4 summarizes the performances of the RECLS, ML, PRECLS and PML methods in terms of estimation quality (NRMSE per voxel at the image level) for different SNR values and Figure 5 details the results for the fast, medium and slow decaying components. One can notice that there is no significant difference between RECLS and ML performances. However, an important reduction of the

error is obtained by PML and the PRECLS methods, which highlights the importance of the spatial regularization. It is interesting to note that the performance of the regularized versions depends very slightly on the SNR. Furthermore, we can clearly see that the estimation error was higher for classes with higher parameter couple (A_0, T_2). Concerning the computing time, the PML algorithm was faster than the PRECLS algorithm, since its computing time on a work desktop was $84s$ whilst the PRECLS took on average $251s$.

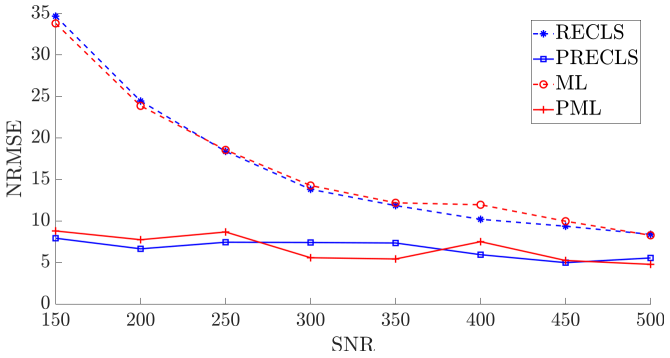

Fig. 4. NRMSE of the RECLS, ML, PRECLS and PML methods for different SNR values.

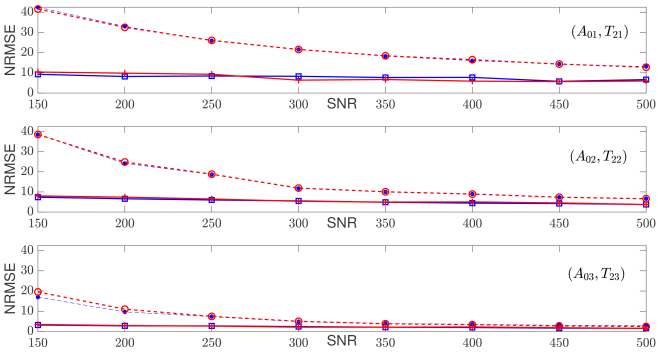

Fig. 5. NRMSE for the different parameters couples (A_0, T_2) estimated on the phantom for the different methods. The legend is the same as the one in figure 4.

Figures 6 and 7 represent the T_2 and A_0 maps obtained on the phantom in the case of an SNR of 150. The positive impact of the regularization is clearly shown since the reconstructed maps are no longer drowned in noise and the separation between different regions of the phantom are more pronounced. Moreover, there is no significant difference between PRECLS and PML results. Therefore, for display convenience, we will show the results of the PML method for the rest of the study since it is faster to converge.

C. Experimental results

Figures 8 and 9 show the results obtained on the real MRI images of a tomato with acquisitions realized with 1 scan and 32 scans (long acquisition time) with the PML method. This validates the robustness of the method since the reconstructed maps from low and high SNR data were quite similar for all the different parameters. We used the parameters estimated from the high SNR image as a reference and obtained a low NRMSE equal to 6%. This shows that even with low SNR the algorithm converged to values close to those estimated from high SNR data This confirms the results obtained on the phantom. As the SNR is directly linked to the number of scans, this result shows that the proposed PML algorithm can save crucial acquisition time since it leads to satisfying results by its application to data acquired with only one scan.

VII. Conclusion

An efficient method for the estimation of T_2 and A_0 maps from noisy MRI magnitude data has been proposed in the context of a discrete parametric multi-exponential model. It is based on the minimization of a penalized criterion that takes into account the Rician distribution by using either a penalized maximum likelihood approach or by minimizing a penalized least squares criterion based on the expectation value of the Rician distribution. Both approaches gave similarly satisfying results by explicitly taking into account the noise statistics and the spatial regularity of the parameters. The first main challenge is the proposal of an efficient and convenient way to carry out the minimization on the whole image at once based on a majorization-minimization approach coupled with an adapted LM algorithm. We showed that visually and numerically the proposed method yields good results both on a simulated data and on real scans of a tomato.

One perspective to be studied is related to the choice of the regularization parameters. An automatic approach for these choices would be ideal, yielding a totally user independent algorithm. One solution would be to adapt the regularization function and the regularization weights to the experimental parameters for each application (fruits, brain studies, liver studies) using databases. Another prospect to study is the choice of the number of components inside each voxel, this is important in studies where the number of components might not be the same between different voxel inside the same MRI image.

Acknowledgments

This work was supported by the French GDR ISIS of the CNRS (PEPS 2017 IRMmultiT2 Poject) and The Région Bretagne (TIRMIV Project).

Appendix

In this appendix we will present both the computation steps for obtaining the different criteria and computing their Jacobians.

A. Least Squares Criterion

In order to minimize the least squares criterion using the LM algorithm, the Jacobian matrix setting requires the computing of $\dfrac{\partial s_{jt}(\boldsymbol{\theta}_j)}{\partial \theta_j(p)}$

for every component $\theta_j(p)$ of the vector parameter $\boldsymbol{\theta}_j = [A_{0_{(1,j)}}\ ,\ T_{2_{(1,j)}}\ \ldots\ A_{0_{(N_c,j)}}\ ,\ T_{2_{(N_c,j)}}]$. It can be noted that $\theta_j(p)$ correspond to an amplitude A_0 for an odd value of p

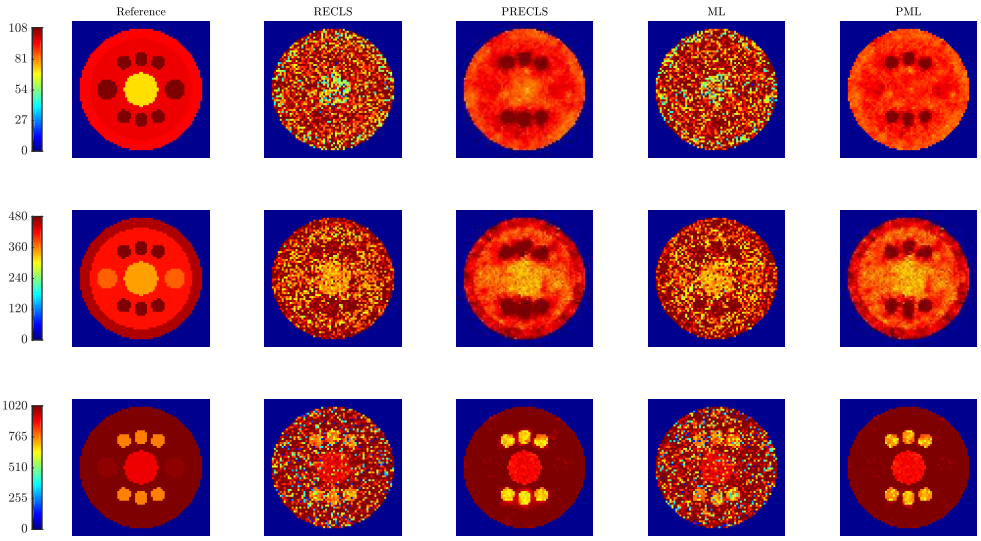

Fig. 6. Reconstructed intensities A_0 maps for the fast (first row), medium (second row) and slow (third row) relaxation times. The Reference map (first column) is compared to the results of the RECLS (second column), PRECLS (third column), the ML (fourth column) and the PML (fifth column) algorithm.

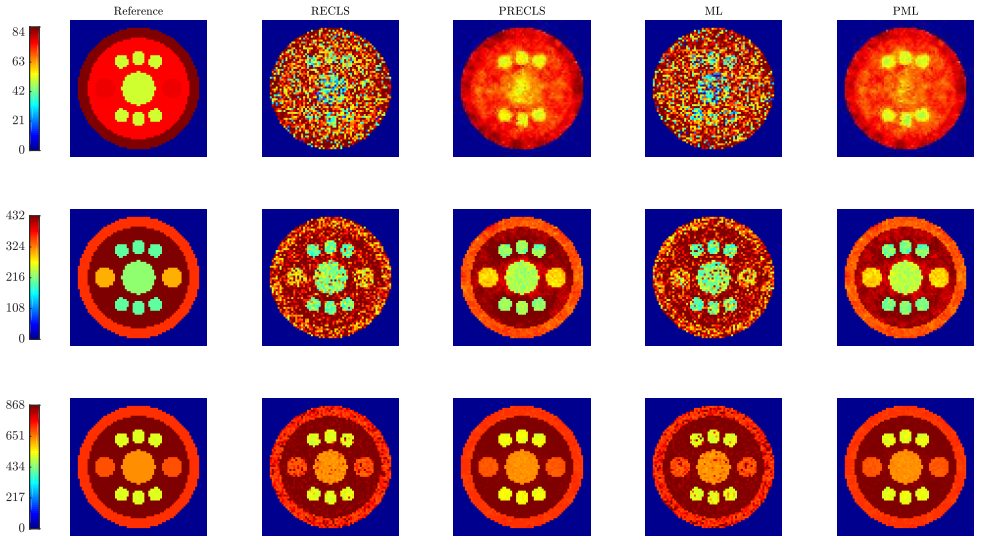

Fig. 7. Reconstructed T_2 maps from a tri-exponential model with T_{2_1} representing the shortest relaxation time (first row), T_{2_2} representing the medium relaxation time(second row) and T_{2_3} representing the longest relaxation time (third row). The Reference map (first column) is compared to the results of the RECLS (second column), PRECLS (third column), the ML (fourth column) and the PML (fifth column) algorithm.

and to a T_2 value for an even value of p.
For odd p:

$$\frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \theta_j\left(p\right)} = e^{-\tau_t/\theta_j(p+1)}. \tag{32}$$

For even p:

$$\frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \theta_j\left(p\right)} = \frac{\theta_j\left(p-1\right)\tau_t}{\theta_j\left(p\right)^2} e^{-\tau_t/\theta_j(p)}. \tag{33}$$

B. McGibney Corrected Least Squares

1) Criterion: The measured signal being expressed in the complex domain by $y_{jt}^{re} = s_{jt}\cos\phi_{jt} + n_{jt}^{re}$ and $y_{jt}^{im} = s_{jt}\sin\phi_{jt} + n_{jt}^{im}$, the second order moment of its magnitude is expressed by

$$\begin{aligned} E\left[y_{jt}^2\right] = & \\ & s_{jt}^2\cos^2\phi_{jt} + s_{jt}^2\sin^2\phi_{jt} \\ & + 2s_{jt}\cos\left(\phi_{jt}\right)E\left[n_{jt}^{re}\right] \\ & + 2s_{jt}\sin\left(\phi_{jt}\right)E\left[n_{jt}^{im}\right] \\ & + E\left[(n_{jt}^{re})^2\right] + E\left[(n_{jt}^{im})^2\right] = s_{jt}^2 + 2\sigma_j^2, \end{aligned} \tag{34}$$

since the additive noise is of zero-mean and of variance σ_j^2.

2) Derivative computation: The derivative to be computed in the case of the MGCLS is given by:

$$\begin{aligned} \frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)^2}{\partial \theta_j\left(p\right)} &= \frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)^2}{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)} \frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \theta_j\left(p\right)}, \\ &= 2s_{jt}\left(\boldsymbol{\theta}_j\right)\frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \theta_j\left(p\right)}. \end{aligned} \tag{35}$$

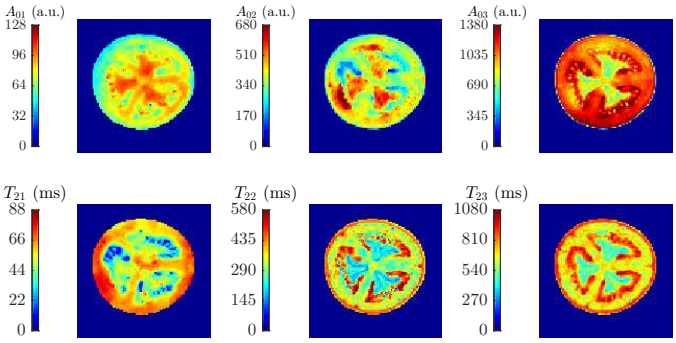

Fig. 8. Estimated intensities A_0 and relaxation times T_2 using PML from the shortest (left) to the longest (right) from MRI data acquired with 1 scan.

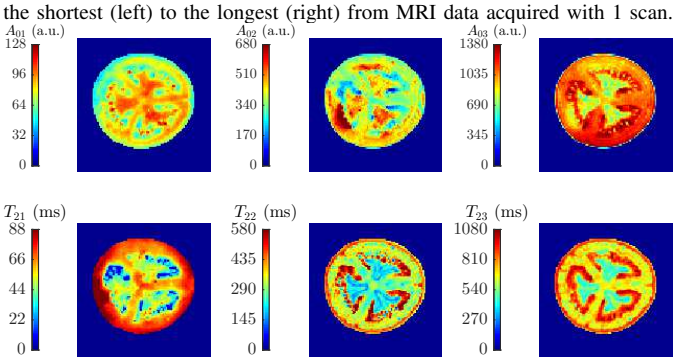

Fig. 9. Estimated intensities A_0 and relaxation times T_2 using PML from the shortest (left) to the longest (right) from MRI data acquired with 32 scans.

C. Rician Expectation Corrected Least Squares

This criterion involves the first order moment

$$E_{Rice}[y_{jt}] = \sigma_j \sqrt{\frac{\pi}{2}} e^{-\alpha_{jt}} [(1+2\alpha_{jt}) I_0(\alpha_{jt}) + 2\alpha_{jt} I_1(\alpha_{jt})]$$

with $\alpha_{jt} = \frac{s_{jt}(\boldsymbol{\theta}_j)^2}{4\sigma_j^2}$. Its derivative with respect to parameters $\boldsymbol{\theta}_j$ can be reformulated as follows

$$\frac{\partial E_{Rice}\left[y_{jt}\right]}{\partial \boldsymbol{\theta}_j} = \frac{\partial E_{Rice}[y_{jt}]}{\partial \alpha_{jt}} \frac{\partial \alpha_{jt}}{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)} \frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \boldsymbol{\theta}_j}. \tag{36}$$

Moreover, since $I_0'(x) = I_1(x)$ and $I_1'(x) = I_0(x) - \frac{1}{x} I_0$,

$$\begin{aligned}
&\frac{\partial E_{Rice}\left[s_{jt}\left(\boldsymbol{\theta}_j\right)\right]}{\partial \alpha_{jt}} = \\
&\sigma_j \sqrt{\frac{\pi}{2}} \big(- e^{-\alpha_{jt}} I_0\left(\alpha_{jt}\right) + e^{-\alpha_{jt}} I_1\left(\alpha_{jt}\right) + 2e^{-\alpha_{jt}} I_0\left(\alpha_{jt}\right) \\
&- 2\alpha_{jt} e^{-\alpha_{jt}} I_0\left(\alpha_{jt}\right) + 2\alpha_{jt} e^{-\alpha_{jt}} I_1\left(\alpha_{jt}\right) \\
&+ 2I_1\left(\alpha_{jt}\right) e^{-\alpha_{jt}} + 2\alpha_{jt} e^{-\alpha_{jt}} \big(I_0\left(\alpha_{jt}\right) - \frac{1}{\alpha_{jt}} I_1\left(\alpha_{jt}\right) \big) \\
&- 2\alpha_{jt} e^{-\alpha_{jt}} I_1\left(\alpha_{jt}\right) \big) \\
&= \sigma_j \sqrt{\frac{\pi}{2}} e^{-\alpha_{jt}} \left[I_0\left(\alpha_{jt}\right) + I_1\left(\alpha_{jt}\right) \right].
\end{aligned} \tag{37}$$

Therefore:

$$\frac{\partial E_{Rice}[y_{jt}]}{\partial \boldsymbol{\theta}_j} = \frac{s_{jt}(\boldsymbol{\theta}_j)}{2\sigma_j} \sqrt{\frac{\pi}{2}} e^{-\alpha_{jt}} \times \left[I_0\left(\alpha_{jt}\right) + I_1\left(\alpha_{jt}\right) \right] \frac{\partial s_{jt}\left(\boldsymbol{\theta}_j\right)}{\partial \theta_j\left(p\right)}. \tag{38}$$

REFERENCES

[1] Mariette F., Rodts S., Faure P., Moucheront P., Musse M., Davenel A., Collewet G., and Lucas T., "MRI ventures into non-medical world," Actualite Chimique, , no. 364, pp. 104–110, 2012.

[2] A. P. Lehoux, S. Rodts, P. Faure, E. Michel, D. Courtier-Murias, and P. Coussot, "Magnetic resonance imaging measurements evidence weak dispersion in homogeneous porous media," Phys. Rev. E, vol. 94, pp. 053107, Nov 2016.

[3] H. Adriaensen, M. Musse, S. Quellec, A. Vignaud, M. Cambert, and F. Mariette, "MSE-MRI sequence optimisation for measurement of bi- and tri-exponential T2 relaxation in a phantom and fruit," Magnetic Resonance Imaging, vol. 31, no. 10, pp. 1677–1689, 2013.

[4] H. Van As, "Intact plant MRI for the study of cell water relations, membrane permeability, cell-to-cell and long distance water transport," Journal of Experimental Botany, vol. 58, no. 4, pp. 743–756, 2007.

[5] M. Musse, S. Quellec, M. Cambert, M.F. Devaux, M. Lahaye, and F. Mariette, "Monitoring the postharvest ripening of tomato fruit using quantitative MRI and NMR relaxometry," Postharvest Biology and Technology, vol. 53, no. 1-2, pp. 22–35, 2009.

[6] G. Winisdorffer, M. Musse, S. Quellec, A. Barbacci, S. Le Gall, F. Mariette, and M. Lahaye, "Analysis of the dynamic mechanical properties of apple tissue and relationships with the intracellular water status, gas distribution, histological properties and chemical composition," Postharvest Biology and Technology, vol. 104, pp. 1–16, 2015.

[7] C. Laule, I Vavasour, S Kolind, D Li, T Traboulsee, G. Moore, and A. MacKay, "Magnetic resonance imaging of myelin," Neurotherapeutics, vol. 4, no. 3, pp. 460–484, 2007.

[8] K. Whittall and A. MacKay, "Quantitative interpretation of NMR relaxation data," Journal of Magnetic Resonance, vol. 84, no. 1, pp. 134–152, 1989.

[9] D. Hwang and Y. Du, "Improved myelin water quantification using spatially regularized non-negative least squares algorithm," Journal of Magnetic Resonance Imaging, vol. 30, no. 1, pp. 203–208, 2009.

[10] D. Kumar, T. Nguyen, S. Gauthier, and A. Raj, "Bayesian algorithm using spatial priors for multiexponential T2 relaxometry from multiecho spin echo MRI," Magnetic Resonance in Medicine, vol. 68, no. 5, pp. 1536–1543, 2012.

[11] Y. Yoo and R. Tam, Non-Local Spatial Regularization of MRI T2 Relaxation Images for Myelin Water Quantification, vol. 8149 of Lecture Notes in Computer Science, pp. 614–621, 2013.

[12] X. Shen, T. Nguyen, S. Gauthier, and A. Raj, "Robust myelin quantitative imaging from multi-echo T2 MRI using edge preserving spatial priors," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2013, pp. 622–630.

[13] A. Akhondi-Asl, O. Afacan, M. Balasubramanian, R. Mulkern, and S. Warfield, "Fast myelin water fraction estimation using 2D multislice CPMG," Magnetic Resonance in Medicine, vol. 76, no. 4, pp. 1301–1313, 2016.

[14] Sudhanya Chatterjee, Olivier Commowick, Onur Afacan, Simon K Warfield, and Christian Barillot, "Multi-Compartment Model of Brain Tissues from T2 Relaxometry MRI Using Gamma Distribution," in ISBI 2018 - IEEE International Symposium on Biomedical Imaging, Washington DC, United States, 2018, pp. 141–144, IEEE.

[15] J. L. Lancaster, T. Andrews, L. J. Hardies, S. Dodd, and P. T. Fox, "Three-pool model of white matter," Journal of Magnetic Resonance Imaging, vol. 17, no. 1, pp. 1–10, 2003, Lancaster, JL Andrews, T Hardies, LJ Dodd, S Fox, PT.

[16] H. Chung, Y. Nam, D. Kim, D. Hwang, and Ieee, Three-Pool Model vs. Nonnegative Least Squares Algorithm for Myelin Water Quantification, Midwest Symposium on Circuits and Systems Conference Proceedings. 2011, Chung, Hyunjin Nam, Yoonho Kim, Donghyun Hwang, Dosik Mwscas 54th IEEE International Midwest Symposium on Circuits and Systems (MWSCAS) Aug 07-10, 2011 Yonsei Univ, Seoul, SOUTH KOREA IEEE, CAS, IEEK, SK Telecom, Korea Tourism Org, Seoul Tourism Org.

[17] T. Andrews, J. L. Lancaster, S. J. Dodd, C. Contreras-Sesvold, and P. T. Fox, "Testing the three-pool white matter model adapted for use with t-2 relaxometry," Magnetic Resonance in Medicine, vol. 54, no. 2, pp. 449–454, 2005, Andrews, T Lancaster, JL Dodd, SJ Contreras-Sesvold, C Fox, PT.

[18] K. J. Layton, M. Morelande, D. Wright, P. M. Farrell, B. Moran, and L. A. Johnston, "Modelling and estimation of multicomponent t-2 distributions," Ieee Transactions on Medical Imaging, vol. 32, no. 8, pp. 1423–1434, 2013, Layton, Kelvin J. Morelande, Mark Wright, David Farrell, Peter M. Moran, Bill Johnston, Leigh A.

[19] J. M. Bonny, J. P. Renou, and M. Zanca, "Optimal measurement of magnitude and phase from MR data," Journal of Magnetic Resonance, Series B, vol. 113, no. 2, pp. 136 – 144, 1996.

[20] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck, "Maximum-likelihood estimation of Rician distribution parameters," IEEE Transactions on Medical Imaging, vol. 17, no. 3, pp. 357–361, June 1998.

[21] O. T Karlsen, R. Verhagen, and W.W.J. Bovée, "Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: application to T1 and perfusion measurements," Magnetic Resonance in Medicine, vol. 41, no. 3, pp. 614–623, 1999.

[22] R. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," Medical Physics, vol. 12, no. 2, pp. 232–233, 1985.

[23] O. Dietrich, J. Raya, S. Reeder, M. Reiser, and S. Schoenberg, "Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters," Journal of Magnetic Resonance Imaging, vol. 26, no. 2, pp. 375–385, 2007.

[24] J. Raya, O. Dietrich, A. Horng, J. Weber, M. Reiser, and C. Glaser, "T2 measurement in articular cartilage: impact of the fitting method on accuracy and precision at low SNR," Magnetic Resonance in Medicine, vol. 63, no. 1, pp. 181–193, 2010.

[25] S. Rice, "Mathematical analysis of random noise," Bell System Technical Journal, vol. 24, no. 1, pp. 46–156, 1945.

[26] G. McGibney and MR. Smith, "An unbiased signal-to-noise ratio measure for magnetic resonance images," Medical physics, vol. 20, no. 4, pp. 1077–1078, 1993.

[27] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," Magnetic Resonance in Medicine, vol. 34, no. 6, pp. 910–914, 1995.

[28] M. Bouhrara, D. Reiter, H. Celik, J.M. Bonny, V. Lukas, K. Fishbein, and R. Spencer, "Incorporation of Rician noise in the analysis of biexponential transverse relaxation in cartilage using a multiple gradient echo sequence at 3 and 7 Tesla," Magnetic Resonance in Medicine, vol. 73, no. 1, pp. 352–366, 2015.

[29] T. Yokoo, Q. Yuan, J. Sénégas, A. Wiethoff, and I. Pedrosa, "Quantitative R2* MRI of the liver with Rician noise models for evaluation of hepatic iron overload: Simulation, phantom, and early clinical experience," Journal of Magnetic Resonance Imaging, vol. 42, no. 6, pp. 1544–1559, 2015.

[30] D. Varadarajan and J. Haldar, "A majorize-minimize framework for Rician and non-central chi MR images," IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 2191–2202, 2015.

[31] A. Raj, S. Pandya, X. Shen, E. LoCastro, T. Nguyen, and S. Gauthier, "Multi-compartment T2 relaxometry using a spatially constrained multi-Gaussian model," PLoS One, vol. 9, no. 6, pp. e98391, 2014.

[32] D. Kumar, S. Siemonsen, C. Heesen, J. Fiehler, and J. Sedlacik, "Noise robust spatially regularized myelin water fraction mapping with the intrinsic b-1-error correction based on the linearized version of the extended phase graph model," Journal of Magnetic Resonance Imaging, vol. 43, no. 4, pp. 800–817, 2016, Kumar, Dushyant Siemonsen, Susanne Heesen, Christoph Fiehler, Jens Sedlacik, Jan.

[33] D. Kim, E. K. Doyle, J. L. Wisnowski, J. H. Kim, and J. P. Haldar, "Diffusion-relaxation correlation spectroscopic imaging: A multidimensional approach for probing microstructure," Magnetic Resonance in Medicine, vol. 78, no. 6, pp. 2236–2249, 2017, Kim, Daeun Doyle, Eamon K. Wisnowski, Jessica L. Kim, Joong Hee Haldar, Justin P.

[34] H. Erdogan and J. Fessler, "Ordered subsets algorithms for transmission tomography," Physics in Medicine & Biology, vol. 44, no. 11, pp. 2835, 1999.

[35] Kelly C McPhee and Alan H Wilman, "Transverse relaxation and flip angle mapping: evaluation of simultaneous and independent methods using multiple spin echoes," Magnetic Resonance in Medicine, vol. 77, no. 5, pp. 2057–2065, 2017.

[36] M. Dylan Tisdall, Richard A. Lockhart, and M. Stella Atkins, "The bias/variance trade-off when estimating the MR signal magnitude from the complex average of repeated measurements," Magnetic Resonance in Medicine, vol. 66, no. 5, pp. 1456–1467, 2011.

[37] Santiago Aja-Fernandez and Gonzalo Vegas-Sanchez-Ferrero, Statistical Analysis of Noise in MRI: Modeling, Filtering and Estimation, Springer Publishing Company, Incorporated, 1st edition, 2016.

[38] Soumya V, A. Varghese, and Neetha N, "A review on noise estimation methods in multi-coil mri," in 2015 International Conference on Computing and Network Communications (CoCoNet), Dec 2015, pp. 859–866.

[39] John C Wood, Cathleen Enriquez, Nilesh Ghugre, J Michael Tyzka, Susan Carson, Marvin D Nelson, and Thomas D Coates, "MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients," Blood, vol. 106, no. 4, pp. 1460–1465, 2005.

[40] S. S. Vasanawala, H. Z. Yu, A. Shimakawa, M. Jeng, and J. H. Brittain, "Estimation of liver t*2 in transfusion-related iron overload in patients with weighted least squares t*2 ideal," Magnetic Resonance in Medicine, vol. 67, no. 1, pp. 183–190, 2012, Vasanawala, Shreyas S. Yu, Huanzhou Shimakawa, Ann Jeng, Michael Brittain, Jean H.

[41] A. Miller and P. Joseph, "The use of power images to perform quantitative analysis on low SNR MR images," Magnetic Resonance Imaging, vol. 11, no. 7, pp. 1051–1056, 1993.

[42] Levenberg Kenneth, "A method for the solution of certain non-linear problems in least squares," Quarterly of Applied Mathematics, vol. 2, no. 2, pp. 164–168, 1944.

[43] Donald W Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," Journal of the society for Industrial and Applied Mathematics, vol. 11, no. 2, pp. 431–441, 1963.

[44] Jinyan Fan and Jianyu Pan, "A note on the Levenberg-Marquardt parameter," Applied Mathematics and Computation, vol. 207, no. 2, pp. 351 – 359, 2009.

[45] J. Nocedal and S. J. Wright, Numerical Optimization, Springer-Verlag, New York, NY, 1999.

[46] E. Chouzenoux, M. Legendre, S. Moussaoui, and J. Idier, "Fast constrained least squares spectral unmixing using primal-dual interior-point optimization," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 1, pp. 59–69, Jan. 2012.

[47] A. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," IEEE Transactions on Medical Imaging, vol. 14, no. 1, pp. 132–137, 1995.

[48] D. Hunter and K. Lange, "A tutorial on MM algorithms," The American Statistician, vol. 58, no. 1, pp. 30–37, 2004.

[49] M. W. Jacobson and J. A. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," IEEE Transactions on Image Processing, vol. 16, no. 10, pp. 2411–2422, Oct. 2007.

[50] M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms," IEEE Transactions on Image Processing, vol. 15, pp. 1130 – 1142, 06 2006.

[51] J. Idier, "Convex half-quadratic criteria and interacting auxiliary variables for image restoration," IEEE Transactions on Image Processing, vol. 10, no. 7, pp. 1001–1009, July 2001.

[52] M. Björk, D. Zachariah, J. Kullberg, and P. Stoica, "A multicomponent T2 relaxometry algorithm for myelin water imaging of the brain," Magnetic Resonance in Medicine, vol. 75, no. 1, pp. 390–402, 2016.

[53] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin, "Adaptive reconstruction of phased array mr imagery," Magnetic Resonance in Medicine, vol. 43, no. 5, pp. 682–690, 2000, Walsh, DO Gmitro, AF Marcellin, MW.

[54] R. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," IEEE Transactions on Image Processing, vol. 8, no. 10, pp. 1408–1419, 1999.

[55] S. Aja-Fernandez, C. Alberola-Lopez, and C. Westin, "Noise and signal estimation in magnitude MRI and Rician distributed images: A LMMSE approach," IEEE Transactions on Image Processing, vol. 17, no. 8, pp. 1383–1398, Aug 2008.