A unified Maximum Likelihood framework for simultaneous motion and $T_1$ estimation in quantitative MR $T_1$ mapping

Gabriel Ramos-Llordén*, Member, IEEE, Arnold J. den Dekker, Gwendolyn Van Steenkiste, Ben Jeurissen, Floris Vanhevel, Johan Van Audekerke, Marleen Verhoye and Jan Sijbers

Abstract—In quantitative MR $T_1$ mapping, the spin-lattice relaxation time $T_1$ of tissues is estimated from a series of $T_1$-weighted images. As the $T_1$ estimation is a voxel-wise estimation procedure, correct spatial alignment of the $T_1$-weighted images is crucial. Conventionally, the $T_1$-weighted images are first registered based on a general-purpose registration metric, after which the $T_1$ map is estimated. However, as demonstrated in this paper, such a two-step approach leads to a bias in the final $T_1$ map. In our work, instead of considering motion correction as a preprocessing step, we recover the motion-free $T_1$ map using a unified estimation approach. In particular, we propose a unified framework where the motion parameters and the $T_1$ map are simultaneously estimated with a Maximum Likelihood (ML) estimator. With our framework, the relaxation model, the motion model as well as the data statistics are jointly incorporated to provide substantially more accurate motion and $T_1$ parameter estimates. Experiments with realistic Monte Carlo simulations show that the proposed unified ML framework outperforms the conventional two-step approach as well as state-of-the-art model-based approaches, in terms of both motion and $T_1$ map accuracy and mean-square error. Furthermore, the proposed method was additionally validated in a controlled experiment with real $T_1$-weighted data and with two in vivo human brain $T_1$-weighted data sets, showing its applicability in real-life scenarios.

Index Terms—$T_1$ mapping, Maximum Likelihood, motion correction, dynamic MRI, registration

I. INTRODUCTION

Quantitative $T_1$ mapping is a Magnetic Resonance Imaging (MRI) technique in which the spin-lattice relaxation time $T_1$ of tissues is measured [1]. Because $T_1$ depends on biophysical properties, it is used as biomarker in a broad range of diseases, such as multiple sclerosis [2], epilepsy [3] and Alzheimer’s disease [4], as well as in the measurement of perfusion [5] and blood flow [6]. Hence, its accurate and precise estimation is of utmost importance [1], [7]. In order to quantify $T_1$, a set of $T_1$-weighted images with different sequence settings needs to be acquired [1], [8], [9]. From this set, a spatial map of $T_1$ values can be calculated by fitting a known relaxation model at every voxel. Evidently, to obtain a meaningful $T_1$ map, spatial correspondence between the images in the acquired series is crucial [10]. However, due to patient motion and/or apparent spatial shifts introduced by the scanner (e.g., scanner drift [11]), $T_1$-weighted images are often misaligned.

To deal with this problem, $T_1$-weighted images are commonly spatially registered prior to the estimation of the $T_1$ map [12], [13]. This is often done by choosing one $T_1$-weighted image as a target and subsequently registering the remaining $T_1$-weighted images to this target image by using a similarity measure such as Mutual Information (MI) [14], [15].

Such an approach, however, suffers from inherent problems. First, the specific relation between the intensity value as a function of time of the (aligned) voxels is ignored. Second, the registration is not driven by a global optimization criterion that considers all $T_1$-weighted images simultaneously. Even more problematic is the fact that current motion correction is a preprocessing step prior to the estimation of the $T_1$ values. Such a two-step processing pipeline lacks a feedback mechanism between the image registration and the $T_1$ map estimation step. As a result, registration errors will propagate to the estimation step, leading to biased estimates [16].

Recently, progress in registration of $T_1$-weighted images was made by the introduction of model-based approaches. Such techniques integrate the signal model connecting the series of images (such as a $T_1$ relaxation model) into the registration step. State-of-the-art model-based methods have shown to outperform the conventional two-step approach in terms of accuracy, for example, in myocardial $T_1$ mapping [17], [18]. Unfortunately, they all come with serious limitations for precise and accurate $T_1$ mapping, mainly because different criteria for registration and estimation are heuristically combined [17], [18]. Since they do not constitute a truly unified framework, the output of the algorithms cannot be related to the optimal value of a given global information-based criterion. As a consequence, it is doubtful whether all
the information gathered in the series of $T_1$-weighted images, including the data statistics [7], [19], is optimally exploited.

In our work, we propose an integrated model-based image registration and $T_1$ estimation approach, where the motion parameters and $T_1$ map are jointly estimated using a unified global information criterion, more specifically, the maximum likelihood (ML) criterion [19]–[21]. By combining models of $T_1$ relaxation, motion, and noise into one statistical model of the $T_1$-weighted images, we are able to restore the original motion-free $T_1$ map using a joint ML estimator. The unified ML framework allows to account for the statistical noise model, the relaxation model and the motion model simultaneously, exploiting, in addition to the temporal information, knowledge on data statistics. The large-scale ML optimization problem is solved by alternating between the estimation of motion and relaxation parameters in an efficient and robust manner, making use of block coordinate descent [22] and Majorize-Minimize (MM) algorithms [23]. Exact convergence properties of the algorithm are presented, demonstrating that the proposed iterative procedure leads to the ML estimates in a computationally efficient way.

We thoroughly validate the proposed joint maximum likelihood estimator (MLE) with realistic Monte Carlo (MC) simulations and compare it with the conventional two-step approach as well as the newest state-of-the-art model-based approach of Hallack [18]. We show that substantially more accurate $T_1$ maps as well as motion parameters can be obtained with our proposed joint MLE. Additionally, the $T_1$ maps estimated with the joint MLE are superior in terms of the root-mean-square error (RMSE). Apart from simulation experiments, we also quantitatively evaluate the performance of the joint MLE in a controlled experiment involving real $T_1$-weighted data. Further, we validate it with two in vivo human brain $T_1$-weighted data sets corrupted by patient motion, showing its applicability in real-life scenarios.

The remainder of the paper is organized as follows. In Section II, the image model used to construct the joint MLE is presented. Section III is devoted to the joint MLE algorithm. Section IV describes the experiments of which the results are presented in Section V, which is followed by a discussion in Section VI. Finally, conclusions are drawn in Section VII.

II. THEORY

The derivation of the joint MLE requires a parametric statistical model of the images. This section is devoted to the derivation of such a model, which comprises a relaxation signal model, a motion model, and a statistical noise model.

A. Relaxation signal model

In the absence of noise, the evolution of the magnitude MRI signal in each voxel of a series of $N$ $T_1$-weighted images can be described by a parametric model $\{f_n(\kappa, T_1)\}_{n=1}^N$, where $\kappa$ denotes a vector of nuisance parameters. The exact expression for this $T_1$-relaxation model depends on the pulse sequence that is used. In this work, we will use the Inversion Recovery (IR) sequence, being the gold standard for $T_1$-mapping [1]. Note that signal models corresponding with other sequences, such as SPoiled Gradient Recalled echo (SPGR) [12] or MOdified Look-Locker Inversion recovery (MOLLI) [24], can be accommodated within our framework as well. For the IR sequence, a common magnitude relaxation model is given by [25]:

$$f_n(a, b, T_1) = |a + be^{-\frac{T_n}{T_1}}|, \quad \text{with } n = 1, \ldots, N,$$

where $\{T_n\}_{n=1}^N$ are the inversion times. The other parameters $a$ and $b$ are related to, among others, the repetition time (TR), radio frequency (RF) pulse angles and the tissue-dependent proton density. The mathematical expressions that relate $a$ and $b$ to these quantities for the case of gradient-recalled-echo (GRE) IR and spin-echo (SE) IR can be found in [25].

To model the noiseless $T_1$-weighted images, we use a vector notation for the spatially varying parameters $T_1$, $a$ and $b$. Let $r = (x, y, z)^T$ be a vector in the Cartesian coordinate system in which they are defined. Then, a 3D spatial $T_1$ map of $M$ voxels can be defined as a column vector, $T_1 \in \mathbb{R}^{M \times 1}$, where $[T_1]_{lm}$ represents $T_1$ defined at the spatial point $r_m$, indexed by voxel $m$. Similarly, we define $a \in \mathbb{R}^{M \times 1}$ and $b \in \mathbb{R}^{M \times 1}$ as the parameter maps of $a$ and $b$. For ease of readability and to alleviate the notation, we introduce the parameter vector $\kappa = (a^T, b^T)^T \in \mathbb{R}^{2M \times 1}$. The relaxation model for the noiseless $n$-th $T_1$-weighted image is then given by

$$f_n(\kappa, T_1) = |a + b \odot e^{-\frac{T_n}{T_1}}|,$$

where $\odot$ and $| \cdot |$ denote the point-wise multiplication and point-wise modulus operator, respectively.

B. Motion model

In what follows, we will restrict the motion model of the unified ML framework to inter-image motion, that is, motion between the 3D $T_1$-weighted images, as in [18]. In the discussion section (Section VI), we further elaborate on extensions of the unified ML framework in which intra-image motion is incorporated, in particular, motion between the slices of a multi-slice $T_1$-weighted image.

The effect of inter-image motion is modeled by assuming that $f_n(\kappa, T_1)$ is observed in a different Cartesian coordinate system $r''$ for each acquisition $n = 1, \ldots, N$. In this work, we illustrate the joint MLE with rigid motion. Hence, the spatial point $r''_m$, with $m = 1, \ldots, M$, is related to the reference-system point $r_m$ through a rigid transformation matrix, $M_{\theta_n} \in \mathbb{R}^{4 \times 4}$ (in homogeneous coordinates), parameterized by

$$\theta_n = (t_{xn}, t_{yn}, t_{zn}, \alpha_n, \beta_n, \gamma_n)^T,$$

with $t_{xn}, t_{yn}, t_{zn}$ the translation parameters and $\alpha_n, \beta_n, \gamma_n$ are the Euler angles of the three elementary rotation matrices around axis $x$, $y$, and $z$, respectively [26]. In our work, the reference system $r$ is defined similarly as the intrinsic coordinate system which MATLAB uses to represent 3D images. That is, axis $x$ points in the direction of increasing column index while $y$ points in the direction of increasing row index. Finally, the axis $z$ is aligned with the direction of increasing index of the third dimension. The origin of this coordinate
system is the center of the 3D image. Furthermore, in multislice acquisitions the axis $z$ is aligned with the slice-encoding direction.

The noisefree $T_1$-weighted image observed at $r^n$ can be modeled as the output of a linear operator that performs rigid motion, $H_{\theta_n}\{\cdot\}$, and whose input is the unobserved $f_n(\kappa, T_1)$. Because $H_{\theta_n}\{\cdot\}$ is linear, the input-output relation can be concisely written in matrix form as:

$$f_n(\theta_n, \kappa, T_1) = H_{\theta_n} f_n(\kappa, T_1),$$  \hspace{1cm} (4)

where $f_n(\theta_n, \kappa, T_1)$ is the motion-corrupted noisefree $T_1$-weighted image acquired at $T_1^n$ and $H_{\theta_n} \in \mathbb{R}^{M \times M}$ is the matrix representation of the linear motion operator $H_{\theta_n}\{\cdot\}$.

To design $H_{\theta_n} \in \mathbb{R}^{M \times M}$, we use the method proposed in [27], where it was demonstrated that each of the rotation matrices of $M_{\theta_n}$ can be decomposed as the product of three shear matrices. Each of the shearings is implemented very efficiently with Fast Fourier Transforms (FFT). Translation is efficiently with FFT approach, the motion operator $H_{\theta_n}$, with the FFT approach, the motion operator $H_{\theta_n}$, can be concisely written in matrix form as:

$$H_{\theta_n} = \begin{pmatrix} H_{\phi_n} & 0 \\ 0 & 1 \end{pmatrix},$$

with $H_{\phi_n}$ the Hermitian conjugate. Hence, the motion operator $H_{\theta_n}$ is reversible, i.e., when applied to an image, this image can be retrieved by applying $H_{\theta_n}^H$ to the output of this operation. The unitarity property of the motion operator will turn out to be useful in the derivation of the joint MLE algorithm. Details of the exact analytical expression of $H_{\theta_n}$ and the proof of the unitarity property are provided in Section I of the additional document which is included as part of the downloadable supplementary material which accompanies this paper.

C. Statistical noise model

In practice, acquired $T_1$-weighted images are inherently disturbed by noise. A typical data distribution for (single-coil) magnitude $T_1$-weighted images is the Rice distribution [28]:

$$p_s(s|\mu, \sigma) = \frac{s}{\sigma^2} e^{-\frac{(s^2 + \mu^2)}{2\sigma^2}} I_0\left(\frac{s \mu}{\sigma^2}\right) u(s),$$

with $\mu$ the noisefree magnitude signal in a voxel, $s$ the noise disturbed signal, $I_0(\cdot)$ the zeroth order modified Bessel function of the first kind, and $\sigma$ the standard deviation of the Gaussian noise disturbing the underlying complex data [21]. The unit step function $u(\cdot)$ is used to indicate that the expression for the probability density function (PDF) of $s$ is valid for nonnegative values of $s$ only. Note that for high signal-to-noise ratio $\frac{\mu}{\sigma} > 3$, the Rician PDF becomes quasi Gaussian [29]. If multiple - instead of just one - receiving coils are used to acquire the data and the k-space is fully sampled (by each coil), the magnitude image that is reconstructed using the Sum of Squares (SoS) method obeys a noncentral chi (nc-$\chi$) distribution, being the natural extension of the Rician distribution for the single-coil case [30]. When parallel MRI techniques that undersample the k-space to decrease the acquisition time are performed, such as SENSE or GRAPPA, other distributions may apply [31]. For a recent review on data distributions in MRI, the reader is referred to [21]. In this work, we will illustrate the proposed joint MLE by deriving it for the case of independent Rician distributed voxels, with different noise standard deviation $\sigma$ for each voxel $m$ and for each acquisition $n$. This is an accurate noise model for magnitude images that are reconstructed with SENSE [31]. It is also a valid noise model for magnitude images that are reconstructed with GRAPPA jointly with a spatially-matched-filter (SMF) data combination [32]. If, instead of SMF, SoS is used in combination of GRAPPA, the data distribution can be well approximated at high SNR by a Gaussian distribution with a spatially variant variance [33]. The derivation of the joint MLE for Gaussian distributed data will be covered in subsection III-E.

III. JOINT MLE

Let $s_n \in \mathbb{R}^{M \times 1}$, with $n = 1, \ldots, N$, denote an actual, noisy $T_1$-weighted image acquired at inversion time $T_1^n$. Assuming Rician distributed data, it follows from Eq. (5) and the motion-corrupted noisefree $T_1$-weighted model Eq. (4) that the PDF of the voxels $|s_n|_m$, $m = 1, \ldots, M$, of this image is given by:

$$p_{|s_n|_m}(|s_n|_m f_n, |f_n|_m, |f_n|_m, |s_n|_m) = \frac{|s_n|_m f_n}{|s_n|_m} e^{-\frac{|s_n|_m^2 f_n^2}{2|s_n|_m^2}}$$

$$\times I_0\left(\frac{|s_n|_m f_n^2}{|s_n|_m |s_n|_m} \right) u(|s_n|_m).$$

Further, if all voxels are assumed to be independent, the joint PDF of the voxels constituting the image $s_n$ is given by the product of the PDFs of the individual voxels, i.e.,

$$p_{s_n} (s_n | f_n, \sigma_n) = \prod_{m=1}^{M} p_{|s_n|_m} (|s_n|_m | f_n, \sigma_n).$$

Similarly, the joint PDF of the supposedly independent voxels of a set of $N T_1$-weighted images $\{s_n\}_{n=1}^{N}$ is given by

$$p_s(s | f, \sigma) = \prod_{n=1}^{N} p_{s_n} (s_n | f_n, \sigma_n)$$

with $s = (s_1^T, \ldots, s_N^T)^T$, $f = (f_1^T, \ldots, f_N^T)^T$ and $\sigma = (\sigma_1^T, \ldots, \sigma_N^T)^T$. Note that this joint PDF depends on the unknown parameters $\theta = (\theta_1^T, \theta_2^T, \ldots, \theta_N^T)^T$, $\kappa$ and $T_1$ via $f$ and can hence be written as $p_s(s | \theta, \kappa, T_1, \sigma)$. To construct the MLE of these parameters, the likelihood function must be derived. The likelihood function is obtained from the joint PDF, Eq. (7), by replacing the independent variables $s$ by the actual acquired voxel intensity values - that is, by numbers - and the supposedly fixed, exact parameters $\theta, \kappa$ and $T_1$ by independent variables. The likelihood function is, therefore, a function of the parameters considered as independent variables and is parametric in the acquired voxel intensities, from now on called observations [7]. To express this, the likelihood function is written as $L(\theta, \kappa, T_1 | s)$. Strictly speaking, the likelihood function also depends on $\sigma$. However, in our work, we assume that $\sigma$ can be estimated prior to the construction of the joint MLE using tailored noise estimation techniques [33], [34]. Hence, we omit the explicit $\sigma$-dependence in the notation.

To simplify the notation, let us define the parameter vector $\tau = (\theta^T, \kappa^T, T_1^T)^T$. The joint MLE $\hat{\tau}_M$ of $\tau$ from the observations $s$ is that value of $\tau$ that maximizes the likelihood
function $L(\tau|s)$, or equivalently, minimizes the so-called negative log-likelihood function $\mathcal{L}_s(\tau|s) \triangleq -\log L(\tau|s)$ with respect to $\tau$, i.e.,

$$\hat{\tau}_{\text{ML}} = \arg\min_{\tau} \mathcal{L}_s(\tau|s).$$ (8)

It follows from Eq. (7) that $\mathcal{L}_s(\tau|s)$ can be written as

$$\mathcal{L}_s(\tau|s) = \sum_{n=1}^{N} \mathcal{L}_{s_n}(\theta_n, \kappa, T_1|s_n),$$ (9)

with $\mathcal{L}_{s_n}(\theta_n, \kappa, T_1|s_n) = -\log p_{s_n}(s_n|\tilde{f}_n, \sigma_n)$.

As the joint MLE cannot be found analytically, one has to resort to numerical optimization algorithms. In order to solve this very-large-scale optimization problem, a cyclic block-coordinate descent (cBCD) method was used [22]. cBCD methods work by iteratively minimizing the cost function $\mathcal{L}_s(\theta, \kappa, T_1|s)$ with respect to a subset of the optimization variables, holding the remaining variables fixed [35], where in each iteration, the roles of the optimization and fixed variables are reversed [35]. The utility of the cBCD algorithm relies on a smart selection of the subset of optimization variables. In the case of our joint MLE, this subset is chosen to contain the motion parameters or the relaxation parameters. In this way, the very-large-scale optimization problem is separated into more easily solvable problems.

Indeed, alternating between the motion estimation problem and the relaxation estimation problem, the joint MLE is found in an efficient way. Moreover, the cBCD method asserts that $\mathcal{L}_s(\theta, \kappa, T_1|s)$ decreases at every iteration [22]. Therefore, convergence to at least a local minimum is guaranteed [36]. In summary, the cBCD-based joint MLE is obtained by the following iterative recursive procedure:

$$\hat{\theta}^{(t+1)} = \arg\min_{\theta} \sum_{n=1}^{N} \mathcal{L}_{s_n}(\theta_n, \hat{\kappa}^{(t)}, T_1^{(t)}|s_n),$$ (P.1)

$$\hat{\kappa}^{(t+1)}, \hat{T}_1^{(t+1)} = \arg\min_{\kappa, T_1} \sum_{n=1}^{N} \mathcal{L}_{s_n}(\hat{\theta}^{(t+1)}_n, \kappa, T_1|s_n),$$ (P.2)

with $\hat{\theta}^{(0)} = \theta_{\text{init}}, \hat{\kappa}^{(0)} = \kappa_{\text{init}}$ and $\hat{T}_1^{(0)} = T_{1\text{init}}$ the initial values of the parameters $\theta$, $\kappa$ and $T_1$, respectively. This procedure is terminated when the number of iterations exceeds $t_{\text{max}}$ or the relative decrease $\mathcal{E}^{(t)}$ of $\mathcal{L}_s(\theta, \kappa, T_1|s)$ between two consecutive iterations is below a fixed tolerance, $\mathcal{E}_{\text{min}}$.

A detailed description of the problems (P.1) and (P.2) is provided in subsections III-A and III-B, respectively. Furthermore, a pseudo-code of the joint MLE algorithm is shown in subsection III-C, whereas its implementation is described in subsection III-D.

A. Problem 1 (P.1): estimation of the motion parameters

The motion estimation problem adopts a particularly simple structure when the relaxation parameters are fixed. If no dependence of $\{\theta_n\}_{n=1}^{N}$ through index $n$ is assumed, as is done here, the minimization can be decoupled into $N$ optimization problems, which can be implemented very efficiently by parallel operations. The parameters $\{\theta_n\}_{n=1}^{N}$ enter the linear motion operator in a non-trivial way, which renders the analytical calculation of the derivatives infeasible. Fortunately, the low dimensionality of the $N$ minimization problems, involving only six variables each, allows us to use a derivative-free optimization method. In our approach, simulated annealing (SA) minimization is performed [37], which is known for its ability to avoid being trapped in local minima and its robustness to functions with complex structure [37]–[40]. Each time, motion estimates from the previous iteration $t$ are used as initial guesses.

B. Problem 2 (P.2): estimation of the relaxation parameters

In contrast to P.1, the relaxation parameter estimation problem is a very-large-scale minimization problem. Derivative-free algorithms such as SA are therefore impractical to use. Quasi-Newton algorithms, on the other hand, may produce critical memory storage problems, due to the large dimensionality of the Hessian matrix approximation. Furthermore, line searches dramatically slow down the algorithm [41].

The optimization method of choice for solving this minimization problem is the use of an MM framework [42], which yields a voxel-wise independent algorithm, allowing a computationally efficient implementation. Here, we sketch the basics of the MM algorithm applied to the estimation problem at hand. The reader is referred to [42] for further details.

Let $J(\kappa, T_1) = \sum_{n=1}^{N} \mathcal{L}_{s_n}(\hat{\theta}^{(t+1)}_n, \kappa, T_1|s_n)$ be the cost function of P.2 that we seek to minimize w.r.t. $\kappa$ and $T_1$. MM algorithms are defined through the following recursive minimization problem:

$$\kappa^{k+1}, T_1^{k+1} = \arg\min_{\kappa, T_1} G(\kappa, T_1|\kappa^k, T_1^k),$$ (10)

where $G(\kappa, T_1|\kappa^k, T_1^k)$ is a new user-designed cost function. It can be demonstrated that the sequence of iterates $\kappa^k, T_1^k$ obtained from Eq. (10) converges to a local minimum of $J(\kappa, T_1)$ if $G(\kappa, T_1|\kappa^k, T_1^k)$ is what is called in the optimization literature a surrogate function of $J(\kappa, T_1)$. The properties that characterize a surrogate function are 1) $J(\kappa, T_1) \leq G(\kappa, T_1|\kappa^k, T_1^k) \forall \kappa, T_1$ and 2) $J(\kappa^k, T_1^k) = G(\kappa^k, T_1^k|\kappa^k, T_1^k)$.

Obviously, to really benefit from the MM algorithm, the surrogate function $G(\kappa, T_1|\kappa^k, T_1^k)$ should be easier to minimize than the original cost function $J(\kappa, T_1)$. A key result is presented by Varadarajan and Haldar [23], showing that the following function is a valid surrogate function of

$$\frac{1}{2} \frac{1}{|\sigma_n|^2_{m}} (|\hat{f}_{n}^{(t+1)}(\hat{\theta}_{n}^{(t+1)}, \kappa, T_1)|_{m} - |\hat{s}_{n}^{k}|_{m}^2)^2 + C_m(k)$$ (11)

with $C_m(k)$ a constant independent of $\kappa$ and $T_1$,

$$\hat{f}_{n}^{(t+1)}(\hat{\theta}_{n}^{(t+1)}, \kappa, T_1) = H_{\theta_{n}^{(t+1)}} f_{n}(\kappa, T_1),$$ (12)

and

$$|\hat{s}_{n}^{k}|_{m} = |s_{n}|_{m} I_0(\frac{|\hat{s}_{n}^{k}|_{m}^2}{|\sigma_{n}|_{m}^2})$$ (13)

with $I_1(\cdot)$ the first order modified Bessel function of the first kind.
Note that \( \mathcal{G}^{(t+1)}(\hat{\theta}^{(t+1)}, \kappa, T_1) \) describes the motion-corrupted synthetic \( T_1 \)-weighted image, whereas \( \hat{s}^k \), from now on called the Bessel image, is the actual acquired image \( s_n \) corrected with a Bessel correction factor. A surrogate function for \( \mathcal{L}_n(\theta, \kappa, T_1|s_n) \) is now obtained by summing Eq.(11) over \( m \), i.e.,

\[
\mathcal{G}_n(\kappa, T_1|\kappa^k, T^k_1) = ||W_n^{1/2}(H\hat{\theta}^{(t+1)} | f_n(\kappa, T_1) - \hat{s}^k) ||_2^2 + C_n(\kappa) \tag{14}
\]

with \( W_n = \text{diag}(\frac{1}{\sigma_n^2}) \). By summing \( \mathcal{G}_n(\kappa, T_1|\kappa^k, T^k_1) \) over \( n \), we would obtain a global surrogate function for \( J(\kappa, T_1) \). At this point, the main benefit of applying the MM framework is that the relaxation estimation problem has now been transformed into a collection of weighted non-linear least squares (NLLS) problems, avoiding complicated minimization with Bessel functions. However, we still can further simplify the problem and convert it into a fully separable (voxel-wise independent) NLLS problem. To that end, we apply another surrogate function \( \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) \) to each \( \mathcal{G}_n(\kappa, T_1|\kappa^k, T^k_1) \). It is easy to demonstrate that if \( \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) \) is a surrogate function for \( \mathcal{G}_n(\kappa, T_1|\kappa^k, T^k_1) \), it is also a valid surrogate function for \( \mathcal{L}_n(\theta, \kappa, T_1|s_n) \). Therefore, we finally define \( G(\kappa, T_1|\kappa^k, T^k_1) \) as

\[
G(\kappa, T_1|\kappa^k, T^k_1) = \sum_{n=1}^{N} \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1). \tag{15}
\]

The choice for \( \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) \) is a separable quadratic surrogate (SQS) function [43], which, when applied to our problem at hand, takes the expression,

\[
\mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) = ||f_n(\kappa, T_1) - \rho_n(\kappa^k, T^k_1)||_2^2 + C_n^*(\kappa), \tag{16}
\]

with

\[
\rho_n(\kappa^k, T^k_1) = f_n(\kappa^k, T^k_1) + \sigma^* H_{\hat{\theta}^{(t+1)}} W_n(\hat{s}_n - H_{\hat{\theta}^{(t+1)}} f_n(\kappa^k, T^k_1)) \tag{17}
\]

and \( \sigma^* = 2 \min_{n,m}(\sigma_n)^2 \). The complete derivation of \( \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) \) can be found in Section II of the additional document which is included as part of the downloadable supplementary material. With \( \mathcal{G}_n^*(\kappa, T_1|\kappa^k, T^k_1) \), minimization of Eq.(15) is nothing more than fitting the relaxation model \( f_n(\kappa, T_1) \) to the “residual” images \( \rho_n(\kappa^k, T^k_1) \) with \( n = 1, ..., N \) in a least squares sense. Therefore, it is a completely separable optimization problem and hence it can be implemented in parallel for every voxel \( m \). This is the main distinct characteristic of the joint MLE that we present in this work, which makes it an efficient method to be used in practice. Once the model-fitting is performed, the new iterate serves to update the “residual” images. This process is repeated until \( k > k_{\text{max}} \) or \( \mathcal{E}(\kappa^k) < \mathcal{E}(\kappa) \) where \( \mathcal{E}(\kappa) \) is the analogues of \( \mathcal{E}(\theta) \) for \( J(\kappa, T_1) \). The final iterate yields the new \( \kappa^{(t+1)} \) and \( T_1^{(t+1)} \), which are then used as input in the motion estimation problem P.1.

### C. Initialisation

Although convergence to a local minimum is guaranteed, convergence to the global minimum, which corresponds with the MLE estimate, cannot be proved, since \( \mathcal{L}^*(\theta, \kappa, T_1|s_n) \) is non-convex. To increase the chances of finding the global minimum, providing good initial values is crucial. In our approach, initial values were obtained using the conventional approach (CA), which consists of image registration prior to voxel-wise relaxation model fitting. A sequential estimation to initialize unified motion model-based approaches was also used in [44], with very good results, and we found it a robust method to initialize our joint MLE. In particular, firstly, the initial motion parameters \( \theta_{\text{ini}} \) were obtained by registering the set of \( T_1 \)-weighted images based on maximization of MI between the images [45], [46]. All images were pairwise registered to the reference system \( r \) with a pyramidal multi resolution scheme of three levels. The number of iterations of the internal optimization algorithm was set to a very high value (> 900) to ensure convergence of the motion parameter estimation. Next, the relaxation parameters \( \kappa_{\text{ini}} \) and \( T_{1\text{ini}} \) were voxel-wise estimated from the registered images using the MLE based on Rician distributed data [20]. To compute the MLE, a Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-newton algorithm [41] was used with exact analytical derivatives. The spatially variant standard deviation, which is required for the MLE, was estimated with the method of [33]. By using the CA as initialization, we have invariably found that the estimated \( T_1 \) maps, \( T_{1\text{MLE}} \), are superior in terms of accuracy and rMSE compared to those obtained with the CA.

Furthermore, simulation results have shown that the joint MLE is stable and robust to occasional inaccuracies in the CA-based initial motion parameters.

Pseudo-code of the joint MLE algorithm is presented in Fig.1 and an illustrative flow-chart is shown in Fig.2. In practice, the joint MLE requires as input, apart from the initial motion and relaxation parameters, an estimate of the spatially variant standard deviation \( \sigma, \sigma^* \). Such an estimate is obtained with the method of [33] directly applied on the acquired images \( \{s_n\}_{n=1}^{N} \).

### D. Parameters selection, code implementation and computational cost

The proposed joint MLE was implemented in MATLAB and run on a computer with an Intel i7-4770K processor consisting of four cores at 3.5 GHz. The machine had 32 GB of RAM. The SA algorithm of P.1 was implemented using the MATLAB routine simulannealbnd with the default parameters. The NLLS fitting of P.2 was performed by the MATLAB routine lsqnonlin using the Levenberg-Marquardt (LM) [41] method, also with the default parameters. The tolerance criteria and the maximum number of iterations to halt the algorithm were chosen to be \( \varepsilon_{\text{J}} = 10^{-2} \) and \( \varepsilon_{\text{min}} = 10^{-3} \), and \( t_{\text{max}} = k_{\text{max}} = 10 \), respectively. To exploit the highly parallelizable structure of the joint MLE, MATLAB parallel computing tools were used to estimate \( \theta_n \) for each value of \( n \) separately. Similarly, the voxel-wise NLLS relaxation model fitting was performed in a parallel
1: initialize \( t = 0, \sigma = \hat{\sigma}, \hat{\theta}^{(0)} = \theta_{\text{ini}}, T_1^{(0)} = T_{1\text{ini}} \) and \( \kappa^{(0)} = \kappa_{\text{ini}} \)
2: while \( \mathcal{E}^{(t)} \geq \mathcal{E}_{\text{min}} \) and \( t < t_{\text{max}} \) do
3: Solve (P.1) to get \( \hat{\theta}^{(t+1)} \)
4: Set \( k = 0 \) (P.2 begins)
5: \( \kappa^k = \kappa^{(t)} \) and \( T_1^k = T_1^{(t)} \)
6: while \( \mathcal{J}^{(k)}(\kappa, s^k) > \mathcal{J}^{(k)}(\hat{\kappa}^{(t+1)}, s^k) \) and \( k < k_{\text{max}} \) do
7: Calculate \( f_n(\kappa^k, T_1^{k}) \) with Eq. (12)
8: Calculate \( \rho_n(\kappa^k, T_1^{k}) \) with Eq. (17)
9: Voxel-wise NLLS fitting of \( f_n(\kappa, T_1) \) to \( \rho_n(\kappa^k, T_1^{k}) \) so as to get \( \kappa^{k+1} \) and \( T_1^{k+1} \)
10: Calculate \( \mathcal{E}^{(k)}(\kappa, T_1) \) and set \( k = k + 1 \)
11: end while
12: Set \( \kappa^{(t+1)} = \kappa^k \) and \( T_1^{(t+1)} = T_1^k \) (P.2 ends)
13: Calculate \( \mathcal{E}^{(t)} \) and set \( t = t + 1 \)
14: end while
15: \( \theta_{\text{ML}} = \hat{\theta}^{(t)} \) and \( T_1^{\text{ML}} = \hat{T}_1^{(t)} \)

Fig. 1: Pseudo-code of the joint MLE algorithm.

manner by dividing the spatial grid into 8 non-overlapping 3D blocks. Finally, a mask was used to avoid calculation of the relaxation parameters in background areas, hence speeding up the implementation.

The computational time per iteration of the joint MLE algorithm is dominated by the voxel-wise \( T_1 \) fitting, which depends linearly on the number of voxels \( M \), depending in turn on the Field-of-View (FOV) and the voxel size. With relatively little effort to optimize our code, and using the MATLAB parallel tools mentioned earlier, the voxel-wise \( T_1 \) fitting took approximately 8 min to process a series of \( N = 8 \) \( T_1 \)-weighted images with \( M \approx 10^5 \) voxels. Overall, with the tolerance criterion described above, the average number of total iterations (external plus internal) were roughly 15, providing precise and accurate \( \theta_{\text{ML}} \) and \( T_{1\text{ML}} \) in an average time of 2.2 hours. Note that migration of the MATLAB code to C++ would produce a much faster implementation, especially if multi-threading is used for the highly parallelizable relaxation estimation problem [44].

E. Gaussian approximation for GRAPPA+SoS

When GRAPPA reconstructed data is combined with SoS, the statistical distribution of the composite magnitude image can be well approximated with a non-stationary nc-\( \chi \) distribution, where both the variance and the (effective) degrees-of-freedom parameter, \( L_{\text{eff}} \) are spatially non-stationary (i.e., vary from voxel to voxel) [31]. Since the MM framework was originally developed for the nc-\( \chi \) distribution [23], the application of the proposed joint MLE is straightforward provided an estimate of \( L_{\text{eff}} \) for every voxel is available. Unfortunately, practical estimators of spatial maps of \( L_{\text{eff}} \) are, to the authors’ knowledge, not yet available in the literature. Nevertheless, for high SNR, a Gaussian distribution with spatially variant \( \sigma \) has been proved to be an accurate model in replacement of the nc-\( \chi \) model [33]. In this case, the joint MLE is even simpler than it was for the Rician case. Indeed, it can easily be shown that

\[
\mathcal{L}_{s_n}(\theta_n, \kappa^{(t)}|T_1^{(t)}) = ||W_n^{1/2}(H_{\theta_n}f_n(\kappa^{(t)}, T_1^{(t)}) - s_n)||_2^2
\]

and

\[
\mathcal{L}_{s_n}(\hat{\theta}^{(t+1)}, \kappa, T_1|s_n) = ||W_n^{1/2}(H_{\hat{\theta}^{(t+1)}}f_n(\kappa, T_1) - s_n)||_2^2
\]

The same SA optimization algorithm as before can be used to minimize Eq. (18) for solving the motion estimation problem (P.1). To simplify the minimization of Eq. (19) for solving the relaxation parameter estimation problem (P.2), we can apply directly the SQS function on \( \mathcal{L}_{s_n}(\hat{\theta}^{(t+1)}, \kappa, T_1|s_n) \), avoiding the Bessel correction step. Indeed, the relaxation parameter estimation problem is again a NLLS fitting of \( f_n(\kappa, T_1) \) to a different \( \rho_n(\kappa^k, T_1^{k}) \) where just \( s_n^{k} \) in Eq. (17) has to be replaced by \( s_n \), the actual acquired images.

IV. EXPERIMENTS

The proposed joint motion and \( T_1 \) MLE was validated using both simulated and real data. Moreover, its performance was compared to that of the CA with MI-based registration [46], and a recently proposed model-based approach of Hallack et al. [18].

A. Simulated \( T_1 \)-weighted data

A set of 3D IR-SE \( T_1 \)-weighted images \( \{s_n\}_{n=1}^N \) affected by inter-image motion (as in Eq. (4)) and noise was simulated from ground truth \( T_1 \) and proton density maps. The ground truth \( T_1 \) map was created from the BrainWeb anatomical model, using reported \( T_1 \) values in human brain tissue at 3T [47], [48]. For the three main brain tissues, white matter, grey matter and cerebrospinal fluid (CSF), the reference values were 838 ms, 1607 ms, and 4300 ms, respectively. The ground truth proton density map was created in a similar fashion. The size of both 3D maps data was \( 111 \times 93 \times 71 \) with an isotropic voxel size of 1.5 mm. From these maps, a set of IR-SE \( T_1 \)-weighted images was simulated based on [25] with TR/TE = 10000/14 ms, and \( N = 8 \) logarithmically equidistant inversion times.

Fig. 2: Flow-chart of the joint MLE algorithm.
\{T_n\}_{n=1}^N \} between 20 ns and 8000 ms. The three consecutive RF pulse angles were set to 180°, 90° and 180°. In the next step, we randomly generated ground truth motion parameters \{\theta_{1n}\}_{n=1}^N. Each of the six rigid motion parameters followed an independent Gaussian Random Walk (RW) [44] along the temporal dimension \(n\).

More precisely, the motion parameters were generated as

\[ \theta_n = c + \theta_{n-1} + w_n, \]

where \(c \in \mathbb{R}^{6 \times 1}\) denotes the motion drift and \(w_n \in \mathbb{R}^{6 \times 1}\) a vector valued, zero mean, Gaussian random variable with covariance matrix \(\Sigma = \sigma_{RW}^2 I\), with \(\sigma_{RW}\) the standard deviation of each of the elements of \(w_n\) and \(I\) the \(6 \times 6\) identity matrix. The reference system \(r\) was chosen to be \(r_1\), hence \(\theta_1 = 0\). Finally, to account for noise, Rician distributed images \(\{s_{1n}\}_{n=1}^N\) were simulated [49] with spatially variant noise maps. Synthetic spatially variant noise maps were generated based on a realistic noise pattern that was presented in [33]. This pattern was derived from a real parallel MRI acquisition [50].

The proposed joint motion and \(T_1\) MLE was compared to that of the CA with MI-based registration [46] (the initialization technique for the joint MLE) and a recently proposed inter-image model-based approach of Hallack et al. [18]. The MI-based rigid registration step of the CA was implemented using the first image of the series as a reference. Details of the implementation were already given in subsection III-C. The remaining MI registration parameters were set to those provided in the MATLAB built-in code. Hallack’s method was implemented by following the guidelines provided in [18]. Just like the joint MLE, it was initialized with the CA. The parameters \(\kappa\) and \(T_1\) were estimated with the LM algorithm. Hallack’s algorithm was stopped when either the decrease of the cost function between iterations was below \(\mathcal{E}_{\text{min}}\), or the number of iterations exceeded \(t_{\text{max}}\).

Two types of simulation experiments were performed:

1) Exp.1: Performance as a function of SNR: In a first set of experiments (Exp.1), the performance of the joint MLE as a function of the SNR of the \(T_1\)-weighted image data set was tested. To that end, motion parameters \(\{\theta_{1n}\}_{n=1}^N\) were generated with \(\sigma_{RW} = 0.4 \text{ mm/degree}\) and no drift. After fixing the motion parameters, \(T_1\)-weighted image data sets with SNR \(\{n\}_{n=1}^N\) were simulated [49] with spatially variant noise maps. An overall RMSE measure was obtained by taking the spatial mean of these relative sample RMSE values, again within the same brain mask.

to assess the ability of each method to estimate the \(T_1\) map, the following performance measures were used:

(a) Relative bias. The bias quantifies the accuracy of the estimator. For each voxel, the relative sample bias was calculated as \((\hat{T}_1 - T_1)/T_1\), where \(\hat{T}_1\) is the sample mean of the \(N_{MC}\) estimates \(\hat{T}_1\) and \(T_1\) is the true value. A measure of the overall accuracy of the \(T_1\) map was obtained by calculating the spatial mean of the absolute value of the relative sample bias, using a brain mask to avoid the skull.

(b) Relative standard deviation. The standard deviation quantifies the precision of the estimator. For each voxel the relative sample standard deviation was calculated as \(\text{std}(\hat{T}_1)/T_1\), and an overall precision measure was obtained by taking the spatial mean of these relative sample standard deviations, using the same brain mask.

c) Relative root-mean-square error (relative RMSE). The RMSE is a measure that incorporates both accuracy and precision. For each voxel, the relative sample RMSE was calculated as \(\sqrt{(\hat{T}_1 - T_1)^2/T_1}\). An overall RMSE measure was obtained by calculating the spatial mean of these relative sample RMSE values, again within the same brain mask.

To assess the ability of each method to estimate motion, the following performance measures were used:

d) Relative motion error, defined as

\[ ||\hat{\theta} - \theta||_2/||\theta||_2, \]

e) Motion component relative bias, defined as

\[ \frac{1}{N} \sum_{n=1}^N \frac{||\hat{\theta}_{1n} - \theta_{1n}||}{||\theta_{1n}||}, \]

with \(\theta_{1n}\) the \(j\)th component of \(\theta\) and \(\hat{\theta}_{1n}\) the sample mean of the \(N_{MC}\) estimates \(\hat{\theta}_{1n}\).

B. Ground-truth based real experiment

In order to assess the performance of the joint MLE with an actual \(T_1\)-weighted data set corrupted by motion, we performed a controlled experiment. The experiment comprised
the acquisition of two data sets. Firstly, we acquired an IR $T_1$-weighted data set of a (static) watermelon. In a second step, we deliberately introduced motion between the acquisition of each of the acquired 2D multi-slice $T_1$-weighted images. In particular, we manually translated and rotated the watermelon after the complete acquisition for a fixed TI. From this data, estimated $T_1$ maps were obtained with the CA, Hallack’s method and the joint MLE. We then quantitatively compared these $T_1$ maps to the estimated $T_1$ map of the first dataset, which was unaffected by motion and hence can be considered as a reasonable ground-truth.

Both IR $T_1$-weighted data sets were acquired with a 3T MRI scanner (MAGNETOM Prisma, Siemens) using a 32-channel head coil. The IR $T_1$-weighted data sets comprised $N = 8$ $T_1$-weighted multi-slice images whose inversion times were logarithmically spaced between 300 and 6000 ms. For each inversion time, we acquired a 2D multi-slice image with a 2D interleaved multi-slice IR Turbo Spin Echo (TSE) sequence [51]. The Echo Train Length (ETL) was 10 and TR/TE = 7920/8.8 ms. Each multi-slice image was acquired within approximately 3 min. The total scan time was about 24 minutes. The acquisition plane was axial and the acquisition matrix was $128 \times 128 \times 3$ and no slice gap. Magnitude data were reconstructed with the GRAPPA method of De Vore [53] adapted for local estimation [33] to work with one single image ($3 \times 3$ neighborhoods) assuming a $\chi^2$ distribution. Due to the high SNR, we relied on results of [33] and used the version of the joint MLE algorithm adapted for spatially variant Gaussian noise (subsection III-E). Noise maps were obtained with the method of [33]. The CA was implemented with a Gaussian MLE where the noise standard deviation was estimated with the method described in [33].

C. In vivo $T_1$-weighted data

We validated the joint MLE with two in vivo human brain data sets suffering from involuntary patient motion.

1) In vivo axial human brain data: An IR $T_1$-weighted data set of a healthy 26-year old male volunteer was acquired with a 3T MRI scanner (MAGNETOM Prisma, Siemens) using a 20-channel head coil. For each inversion time, we acquired a 2D multi-slice image with an interleaved 2D multi-slice IR TSE sequence [51], [54]. The sequence parameters were: ETL = 4 and TR/TE = 8040/18 ms. Each multi-slice image was acquired within 2.5 min approximately. The IR $T_1$-weighted data set comprised $N = 7$ $T_1$-weighted multi-slice images whose inversion times were logarithmically spaced between 50 and 3200 ms. The total scan time was about 19 minutes. The acquisition plane was axial and the acquisition matrix was $128 \times 128 \times 25$ with an anisotropic voxel size of $1.9 \times 1.9 \times 6 \text{ mm}^3$ and a slice gap of 10%. The SENSE method was employed to reconstruct the magnitude data with an acceleration factor of 3. Noise maps were obtained with the method of [33]. We estimated an SNR of 24.3 with the method of [53]. In this case, the reference image was the one with lowest inversion time.

2) In vivo sagittal human brain data: An IR $T_1$-weighted data set of a healthy 26-year old male volunteer was acquired with a 3T MRI scanner (MAGNETOM Prisma, Siemens) using a 32-channel head coil. As in the previous in vivo experiment, we acquired for each inversion time a 2D multi-slice image with an interleaved 2D multi-slice IR TSE sequence [51]. The sequence parameters were: ETL = 10 and TR/TE = 5000/4.8 ms. Each multi-slice image was acquired within 2 min approximately. The IR $T_1$-weighted data set comprised $N = 14$ $T_1$-weighted images whose inversion times were logarithmically spaced between 100 and 3000 ms, giving a total acquisition time of 28 min. The acquisition plane was sagittal and the acquisition matrix was $256 \times 256 \times 40$ with an anisotropic voxel size of $1 \times 1 \times 4 \text{ mm}^3$ and no slice gap. Magnitude data were reconstructed with the GRAPPA method with SoS reconstruction (acceleration factor of 3) [55]. The image with the lowest inversion time was chosen as a reference. We estimated an SNR of 55 with a locally adapted ML estimator ($3 \times 3$ neighborhoods) assuming a $\chi^2$ distribution. Due to the high SNR, we relied on results of [33] and used the version of the joint MLE algorithm adapted for spatially variant Gaussian noise (subsection III-E). Noise maps were obtained with the method of [33]. The CA was implemented with a Gaussian MLE where the noise standard deviation was estimated with the method described in [33].

V. RESULTS

A. Simulated $T_1$-weighted data

1) Exp.1: Performance as a function of SNR:

Overall relative $T_1$ bias, standard deviation and RMSE results are shown in Fig.3(a-c). For the whole range of SNR, the joint MLE allows a much more accurate estimation of the $T_1$ map than Hallack’s method and especially than the CA.

In terms of precision, CA obtains the best result, followed by the joint MLE and Hallack’s method. However, in terms of the overall RMSE, the joint MLE performs best for all values of the SNR. Furthermore, the box-plot shown in Fig.3(d) demonstrates the superiority of the proposed joint MLE in terms of motion estimation.

To complement the results, maps of the absolute value of the relative sample bias for the three methods are shown in Fig.4(b-d), along with the simulated ground truth in Fig.4(a). A close look at the bias maps corroborates the poor performance of the CA compared to the joint MLE. It is also clearly seen that the bias map of Hallack’s method presents much higher values than that of the joint MLE, especially in white/grey matter surroundings.

2) Exp.2: Performance as a function of the type of motion:

Bar charts representing the overall $T_1$ accuracy, precision and rMSE for the four cases of motion and the no-motion scenario are shown in Fig.5.

In light of these results, it can be concluded that the joint MLE yields the most accurate $T_1$ maps in all the four considered motion scenarios, followed by Hallack’s method. Furthermore, the performance of all methods seems to be
fairly insensitive to the type of motion considered. Even though the highest precision was consistently obtained with the CA, its overall relative RMSE is much higher compared to Hallack’s method and especially to the joint MLE, which again produces the best \( T_1 \) maps in RMSE sense. The case of no-motion is particularly interesting. In such a scenario, the CA performance drastically improves in terms of accuracy and RMSE, though its precision decreases. In the absence of motion, the error propagation of the CA approach is negligible, hence not contributing to a reduction in the variability of the estimates.

The motion component relative bias for each of the six components are reported in Table I. For the no-motion case, the motion component relative bias is not well-defined (division by zero). Instead, we report the motion component absolute bias.

The best results are highlighted in shaded gray. In 27 of 30 cases, the joint MLE achieved the highest accuracy in both the translation and rotation parameters. Sometimes, this improvement is even more than 5-fold compared to the CA. In general, Hallack’s method provides more accurate \( T_1 \) estimates than CA, which is in agreement with previously reported results [10], [18]. Nonetheless, further substantial improvement can be obtained if the joint MLE is used. To illustrate the quality of the motion estimation, we have shown graphs, as a function of \( n \), of the ground-truth and estimated motion parameters for one of the rotational motion (R-m) simulations in the additional document (Fig.1) of the downloadable supplementary material.

It is also important to notice that the occasionally poor CA-based motion initialization does not prevent the joint MLE from producing the most accurate motion estimates. This highlights another feature of the joint MLE: it is fairly robust to scenarios where the CA-based motion initialization is relatively poor.
B. Ground-truth based real experiment

A top-axial and a mid-axial slice of the estimated $T_1$ maps for the CA, Hallack’s method and the joint MLE are shown in Fig. 6(c-e) and Fig. 6(m-o), whereas the ground truth $T_1$ map and the $T_1$ map estimated without motion correction are displayed in Fig. 6(a-b) for the top-axial slice and in Fig. 6(k-l) for the mid-axial slice. From this experiment, it can be observed that more detailed $T_1$ maps can be obtained with the joint MLE in comparison to Hallack’s method and especially to the CA. Aside from the presence of a large number of outliers in the $T_1$ maps obtained with the CA and Hallack’s method, which are drastically reduced with the joint MLE, fine structural details seem better preserved with our proposed method. This observation is confirmed by inspecting the magnified regions. The heterogeneity of the $T_1$ values in those regions, as noticed from the ground-truth $T_1$ map, is better maintained with the joint MLE. See for instance the delineation of low $T_1$ value structures in Fig. 6(f-j). Note as well that artifacts in the $T_1$ maps, as shown in Fig. 6(p-t) (green arrow), are considerably mitigated with the joint MLE. Quantitative validation of the estimated $T_1$ maps was based on spatial maps of the absolute value of the relative errors [%](Fig. 7).

In accordance to our previous discussion, the spatial distributions of the relative errors further indicate the good performance of the joint MLE in comparison to competing methods. It is manifestly clear that the error maps of Hallack’s method and the CA present much higher values than that of the joint MLE. To complement the quantitative analysis, we calculated an overall relative error, within a mask neglecting the background, in a similar fashion as done in Exp.1. Numerical results are in agreement with the observation made from the spatial maps. We found that the joint MLE produced the $T_1$ map with the lowest overall relative error. Indeed, the overall relative error for the without motion correction case, the CA, Hallack’s method and the joint MLE was 87%, 20.2%, 19.7% and 15.1%, respectively. To further complement the

Copyright (c) 2016 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
TABLE I: Results of Exp.2: for four types of motion (column 1), the maximum and mean values of the motion parameters (column 2), and the motion component relative bias for each of the six components for CA, Hallack’s and the joint MLE method (columns 3-5) are shown. For the no-motion case, the motion component absolute bias is reported instead of the motion component relative bias since the latter metric is not well-defined for parameters that are equal to zero.

<table>
<thead>
<tr>
<th>Type</th>
<th>Motion (max / mean)</th>
<th>CA</th>
<th>Hallack</th>
<th>Joint MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-m</td>
<td>(t_x (2.5 / 0.98 \text{ mm}))</td>
<td>8.4 %</td>
<td>4 %</td>
<td>11.1 %</td>
</tr>
<tr>
<td></td>
<td>(t_y (1.4 / 0.88 \text{ mm}))</td>
<td>1.4 %</td>
<td>2.1 %</td>
<td>0.4 %</td>
</tr>
<tr>
<td></td>
<td>(t_z (1.1 / 0.58 \text{ mm}))</td>
<td>1.5 %</td>
<td>3.9 %</td>
<td>5.3 %</td>
</tr>
<tr>
<td></td>
<td>(\alpha (0.2 / 0.003 \text{ degree}))</td>
<td>63.5 %</td>
<td>20.6 %</td>
<td>9.5 %</td>
</tr>
<tr>
<td></td>
<td>(\beta (-0.2 / -0.06 \text{ degree}))</td>
<td>90 %</td>
<td>45.8 %</td>
<td>19.1 %</td>
</tr>
<tr>
<td></td>
<td>(\gamma (0.5 / 0.25 \text{ degree}))</td>
<td>75.4 %</td>
<td>37.5 %</td>
<td>6.8 %</td>
</tr>
<tr>
<td>HA-m</td>
<td>(t_x (-9.4 / -4.4 \text{ mm}))</td>
<td>2.3 %</td>
<td>3.9 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td></td>
<td>(t_y (-2.3 / -0.98 \text{ mm}))</td>
<td>30 %</td>
<td>12.9 %</td>
<td>2.4 %</td>
</tr>
<tr>
<td></td>
<td>(t_z (1.8 / 1.2 \text{ mm}))</td>
<td>19.5 %</td>
<td>13.5 %</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(\alpha (2.1 / 0.51 \text{ degree}))</td>
<td>4.9 %</td>
<td>1.0 %</td>
<td>6.4 %</td>
</tr>
<tr>
<td></td>
<td>(\beta (3.5 / 1.9 \text{ degree}))</td>
<td>2.3 %</td>
<td>2 %</td>
<td>0.9 %</td>
</tr>
<tr>
<td></td>
<td>(\gamma (1.3 / 0.5 \text{ degree}))</td>
<td>4.9 %</td>
<td>18.2 %</td>
<td>19.6 %</td>
</tr>
<tr>
<td>R-m</td>
<td>(t_x (-0.8 / -0.4 \text{ mm}))</td>
<td>5.5 %</td>
<td>3.4 %</td>
<td>1.2 %</td>
</tr>
<tr>
<td></td>
<td>(t_y (1.1 / 0.8 \text{ mm}))</td>
<td>2.3 %</td>
<td>3.1 %</td>
<td>1.1 %</td>
</tr>
<tr>
<td></td>
<td>(t_z (0.7 / 0.32 \text{ mm}))</td>
<td>100 %</td>
<td>100 %</td>
<td>29.8 %</td>
</tr>
<tr>
<td></td>
<td>(\alpha (3.2 / 1.5 \text{ degree}))</td>
<td>2.1 %</td>
<td>1.3 %</td>
<td>0.7 %</td>
</tr>
<tr>
<td></td>
<td>(\beta (5.1 / 1.5 \text{ degree}))</td>
<td>1.6 %</td>
<td>0.9 %</td>
<td>1.1 %</td>
</tr>
<tr>
<td></td>
<td>(\gamma (1.9 / 1.1 \text{ degree}))</td>
<td>2.9 %</td>
<td>0.5 %</td>
<td>0.13 %</td>
</tr>
<tr>
<td>T-m</td>
<td>(t_x (2.9 / 1.4 \text{ mm}))</td>
<td>1.3 %</td>
<td>1.9 %</td>
<td>0.8 %</td>
</tr>
<tr>
<td></td>
<td>(t_y (2.8 / 1.9 \text{ mm}))</td>
<td>1.2 %</td>
<td>3 %</td>
<td>1.1 %</td>
</tr>
<tr>
<td></td>
<td>(t_z (2.1 / 1.2 \text{ mm}))</td>
<td>2 %</td>
<td>2.6 %</td>
<td>0.8 %</td>
</tr>
<tr>
<td></td>
<td>(\alpha (-0.3 / -0.2 \text{ degree}))</td>
<td>25.7 %</td>
<td>64.7 %</td>
<td>20.7 %</td>
</tr>
<tr>
<td></td>
<td>(\beta (-0.8 / -0.3 \text{ degree}))</td>
<td>121.15 %</td>
<td>172 %</td>
<td>47.7 %</td>
</tr>
<tr>
<td></td>
<td>(\gamma (-0.5 / -0.2 \text{ degree}))</td>
<td>91.1 %</td>
<td>2.1 %</td>
<td>21.7 %</td>
</tr>
<tr>
<td>No-m</td>
<td>(t_x (0 / 0 \text{ mm}))</td>
<td>0.012</td>
<td>0.032</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(t_y (0 / 0 \text{ mm}))</td>
<td>0.011</td>
<td>0.077</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(t_z (0 / 0 \text{ mm}))</td>
<td>0.0078</td>
<td>0.159</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>(\alpha (0 / 0 \text{ degree}))</td>
<td>0.0281</td>
<td>0.0135</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(\beta (0 / 0 \text{ degree}))</td>
<td>0.0183</td>
<td>0.0157</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(\gamma (0 / 0 \text{ degree}))</td>
<td>0.00201</td>
<td>0.00185</td>
<td>0.011</td>
</tr>
</tbody>
</table>

C. In vivo T1-weighted data

1) In vivo axial human brain data: A mid-axial and a top-axial slice of the estimated \(T_1\) map for the three methods are shown in Fig.8(b-d) and Fig.8(f-h), respectively. The estimated \(T_1\) maps without motion correction are shown in Fig.8(a) and Fig.8(e). The presence of outliers in the \(T_1\) map when motion correction is not applied is not completely avoided with the CA. Indeed, it can be clearly observed that outliers are still present, in particular at the interfaces between ventricles and white matter (green arrow) for the mid-axial slice, and in the interfaces between white and grey matter for the top-axial slice Fig.8(f). The \(T_1\) map produced by Hallack’s method seems free from outliers in the mid-axial but not in top-axial slice Fig.8(g) (green arrow). The joint MLE provides \(T_1\) maps which does not suffer from this issue. The motion parameter estimates obtained with the three methods are shown in the additional document (Fig. 3) which is included as part of the supplementary material which accompanies this paper.

2) In vivo sagittal human brain data: Two mid-sagittal slice of the estimated \(T_1\) map for the three methods are shown in Fig.9(b-d) and Fig.9(j-l), respectively. In Fig.9(a) and Fig.9(i) estimated \(T_1\) maps for the no motion correction case are presented.

In this case, the estimated \(T_1\) map with the CA is almost free from outliers, which are widespread when no motion correction is accomplished, see magnified regions in Fig.9(e-f) and Fig.9(m-n). However, the CA sacrifices the final resolution is accomplished, see magnified regions in Fig.9(e-f) and Fig.9(m-n). However, the CA sacrifices the final resolution is accomplished, see magnified regions in Fig.9(e-f) and Fig.9(m-n). However, the CA sacrifices the final resolution is accomplished, see magnified regions in Fig.9(e-f) and Fig.9(m-n). However, the CA sacrifices the final resolution
VI. DISCUSSION

We presented a unified model-based approach for simultaneous motion correction and $T_1$ mapping that jointly estimates the motion parameters and the $T_1$ map using a maximum likelihood estimator (MLE). The proposed joint MLE possesses optimal statistical properties, which are shared by neither the conventional two-step approach (image registration prior to $T_1$ estimation) nor other heuristic integrated model-based methods.

Using realistic MC simulation experiments, it was shown that the proposed joint MLE outperforms existing $T_1$ mapping methods in terms of both accuracy and RMSE, next to providing more accurate motion parameter estimates. Results of the controlled experiment based on ground-truth real data are in line with the findings of MC simulations. We have shown that detailed and meaningful $T_1$ map can be recovered with the joint MLE under the influence of manually induced, severe motion. Quantitative comparison against a ground-truth $T_1$ map demonstrates the superior quality in $T_1$ map restoration compared to CA and Hallack’s method.

Furthermore, the optimal unified ML framework has been validated with in vivo human brain data experiments, suffering from involuntary motion. From these experiments, it has become evident that motion correction is indispensable in $T_1$ mapping, even when subject motion is relatively small. Interestingly, yet recognizing the limitation of visual assessment in quantitative MRI, some of the rigorously derived statistical conclusions from the MC simulations can be noticed in the two whole brain human data sets. For instance, the arguably poorer motion estimation performance of the CA compared to model-based registration approaches, already reported by [10], [18] and confirmed by our MC results (see Fig. 3(d) and Table I), may be the cause of the presence of outliers in the estimated $T_1$ map and the loss of fine details. The presence of small number of outliers in the estimated $T_1$ map with Hallack’s method can be attributed to its non-optimal/heuristic design. While improvement in motion estimation compared to CA has been demonstrated, optimality in terms of $T_1$ estimation cannot be theoretically and empirically guaranteed. In contrast, our ML framework combines, in a single integrated approach, the benefits of model-based registration techniques with optimal $T_1$ map restoration based on statistical theory, where the noise statistics are properly accounted for.

On top of that, our careful algorithm design avoids heavy computational burden. Note that the voxel-wise $T_1$ fitting task, which contributes most to the computational cost of the proposed joint MLE, is often also included in the iterative loop of other model-based integrated methods. Consequently, the computation time per iteration is comparable.

The proposed ML framework can be extended in different ways without compromising its optimal statistical properties. Extension to other MRI sequences or modalities is straightforward by substituting a properly modified parametric signal model for Eq. (1). Such potential extensions include $T_2$ and $T^*_2$ mapping [29]. Moreover, multi-component $T_2$ mapping would benefit as well from the proposed ML framework [56]. Furthermore, it is worthwhile mentioning that when the joint

folds (see Fig. 9(g-h)) seems better defined with the joint MLE. In addition, yet a (reduced) number of outliers can be detected in the $T_1$ map estimated with the Hallack’s method. See for example, Fig. 9(o). Such outliers, as in the in vivo axial experiment, are not present provided the joint MLE is applied. In Fig. 4 of the additional document, which is part of the downloadable supplementary material, the motion parameter estimates of this experiment are shown.
MLE proposed in this work is applied to the particular case of $T_1$ mapping using the spoiled gradient recalled echo (SPGR) sequence, its computational efficiency can even be further improved by using the recently proposed fast non-linear least squares $T_1$ estimator NOVIFAST [57] for solving the voxel-wise NLLS problems in P.2.

Extensions towards the inclusion of different types of motion, e.g., non-rigid motion, are also possible but require further study, which is considered future work. In this work, we have assumed a motion model which accounts for inter-image motion, that is, motion between each of the 3D $T_1$-weighted images. Although this model left aside intra-motion effects, which is known to affect the k-space reconstruction process, we have not observed any derived ghosting artifacts in the in vivo reconstructed $T_1$-weighted images. It should be noted that, to deal with such kind of motion, navigators [58] or advanced k-space reconstruction methods with motion correction [59] can be applied in conjunction with the joint MLE. However, though intra-image motion may be alleviated with such techniques, image registration, that is, inter-image motion correction, will remain necessary, which further emphasizes the relevance of the unified ML framework. Finally, it is relatively straightforward to extend the ML framework to cope with inter-slice motion, that is, motion occurring between the acquisition of 2D slices of a $T_1$-weighted dataset. An outlook to such an extension, which is especially relevant for image acquisition methods that acquire 3D volumes slice by slice sequentially, such as Echo Planar Imaging (EPI) sequences [44], is given in the additional document (Section IV) which is part of the downloadable supplementary material.

VII. CONCLUSION

In quantitative MR $T_1$ mapping, it is common practice to register the $T_1$-weighted images prior to $T_1$ map estimation. However, as demonstrated in this paper, this conventional two-step approach lacks high accuracy motion estimation and leads to biased $T_1$ estimates. Hence, we have proposed a rigorous unified framework for simultaneous motion and $T_1$ estimation using a Maximum Likelihood (ML) estimator. It has been demonstrated that the proposed joint MLE outperforms the conventional approach as well as a recently proposed model-based method [18] in terms of motion and $T_1$ estimation accuracy and RMSE. Our ML framework, which uses an efficient algorithm, has been validated in a controlled experiment with real $T_1$-weighted data and also with two in vivo human brain data sets. We believe that the unified ML framework possesses serious advantages over the conventional approach to replace it in clinical scenarios where precise and accurate $T_1$ estimates are the ultimate goal.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. The authors would also like to thank Santiago Aja-Fernández and Sayuan Liang for interesting discussions on noise estimation and acquisition settings, respectively.
REFERENCES


Copyright (c) 2016 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

The full version of this paper is available at:

http://dx.doi.org/10.1109/TMI.2016.2611653
Additional document for “A unified Maximum Likelihood framework for simultaneous motion and $T_1$ estimation in quantitative MR $T_1$ mapping”

Gabriel Ramos-Llordén, Member, IEEE, Arnold J. den Dekker, Gwendolyn Van Steenkiste, Ben Jeurissen, Floris Vanhevel, Johan Van Audekerke, Marleen Verhoye and Jan Sijbers

Abstract—In Section I of this additional document we provide details on the motion operator $H_{\theta_n}$. In Section II, we prove Eqs.(16-17) of the main body of the paper, which gives a complete description of the MM framework used in P.2. In Section III, we show graphs of estimated motion parameters for the in vivo experiments and one of the simulation experiments described in the main body of the paper. Finally, Section IV discusses an extension of the joint MLE to account for intra-image motion.

I. DETAILS ON MOTION OPERATOR $H_{\theta_n}$

In this section, an explicit expression of the motion operator $H_{\theta_n}$ is derived. Furthermore, we sketch the proof for its unitarity property, i.e., $H_{\theta_n}^H H_{\theta_n} = H_{\theta_n} H_{\theta_n}^H = I$, with $I$ the identity matrix. Specific details can be found in [1], [2] and especially in [3].

Let $r_m^n$ be a spatial point related to the reference-system point $r_m$ through a rigid transformation matrix $M_{\theta_n} \in \mathbb{R}^{4 \times 4}$:

$$
\begin{pmatrix}
    r_m^n \\
    1
\end{pmatrix}
= M_{\theta_n} \begin{pmatrix}
    r_m \\
    1
\end{pmatrix}.
\quad \text{(S1)}
$$

The rigid transformation matrix $M_{\theta_n} \in \mathbb{R}^{4 \times 4}$, which includes 3D rotation and translation, can then be written as [4]:

$$
M_{\theta_n} = \begin{pmatrix}
R(\alpha_n, \beta_n, \gamma_n) & t_n \\
0^T & 1
\end{pmatrix},
\quad \text{(S2)}
$$

with $t_n = (t_{xn}, t_{yn}, t_{zn})^T$ a vector of translation parameters, $0^T$ a $1 \times 3$ zero vector, and $R(\alpha_n, \beta_n, \gamma_n) \in \mathbb{R}^{3 \times 3}$ the product of three elementary rotation matrices ($R_z(\alpha_n)$, $R_y(\beta_n)$ and $R_z(\gamma_n)$) describing rotations around the $x$, $y$ and $z$ axis, with angles $\alpha_n$, $\beta_n$ and $\gamma_n$, respectively. With such parametrization, we get

$$
r_m^n = R_z(\alpha_n) R_y(\beta_n) R_z(\gamma_n) r_m + t_n.
\quad \text{(S3)}
$$

Let $f(\cdot)$ be spatially-continuous function (a relaxation model in our problem). Then, to calculate $f(r_m^n)$, the following spatial transformations on $f(\cdot)$ are consecutively applied.

1) 3D translation: $f_T(r) = f(r + t_n)$
2) Rotation around x axis $f_{Rot-x}(r) = f_T(R_z(r))$
3) Rotation around y axis $f_{Rot-y}(r) = f_{Rot-x}(R_y(r))$
4) Rotation around z axis $f_{Rot-z}(r) = f_{Rot-y}(R_z(r))$

Indeed, by evaluating $f_{Rot-z}(\cdot)$ at $r_m$, we get $f(r_m^n)$.

In a discrete domain, each of the previous four operations is represented by linear operators, hence matrices, that we denote as $H_T$, $H_{Rot-x}$, $H_{Rot-y}$, and $H_{Rot-z}$, respectively. Note that we have omitted the dependence on the motion parameter for the sake of notational convenience.

As a consequence, the motion operator, $H_{\theta_n}$, can be written as

$$
H_{\theta_n} = H_{Rot-z} H_{Rot-y} H_{Rot-x} H_T
\quad \text{(S4)}
$$

and its Hermitian transpose as

$$
H_{\theta_n}^H = H_T^H H_{Rot-z}^H H_{Rot-y}^H H_{Rot-x}^H.
\quad \text{(S5)}
$$

It is clear that if $H_T$, $H_{Rot-x}$, $H_{Rot-y}$ and $H_{Rot-z}$ are unitary, $H_{\theta_n}$ is unitary as well.

Sketch of Proof 1: $H_T$ is unitary

The translation operator $H_T$ consists of 1) a 3D FFT, 2) a voxel-wise multiplication with a purely complex exponential whose phase depends linearly on the translation parameters, and 3) an inverse 3D FFT [3]. By noting that the multidimensional FFT is a unitary operator [2], the translation operator can be succinctly written as

$$
H_T = F_{3D}^H \Delta F_{3D},
\quad \text{(S6)}
$$

where $F_{3D}$ is the 3D unitary Discrete Fourier Transform (DFT) matrix and $\Delta$ is a diagonal matrix whose entries are purely complex exponentials.

It is known that $\Delta$ is a unitary matrix if and only if the modulus of each diagonal entries is one. Since this is always true for purely complex exponentials, it demonstrates that $H_T$ is unitary.

Sketch of Proof 2: $H_{Rot-x}$, $H_{Rot-y}$, and $H_{Rot-z}$ are unitary.

For brevity, we present the proof only for $H_{Rot-x}$. The proof for $H_{Rot-y}$ and $H_{Rot-z}$ is completely similar.

Because $R_x(\alpha_n)$ can be decomposed as the product of three one-dimensional shear matrices [3], it is possible to write

$$
H_{Rot-x} = S_x S_y S_x,
\quad \text{(S7)}
$$

where $S_x$ and $S_y$ are fractional delay filters [1], which model the shearings in the $x$ and $y$ dimension, respectively. Note that these filters can be implemented efficiently with FFT [3]. If both $S_x$ and $S_y$ are unitary, $H_{Rot-x}$ is unitary as well. Indeed, $S_x$ has essentially the same diagonal expression as Eq.(S6), where the role of the 3D DFT matrices is fulfilled by a (unitary) Fourier matrix which applies an FFT only along the $x$ direction. The phase of the complex exponential
in the diagonal matrix now depends linearly on the shearing parameter [3], which is a real value. Therefore, the associated diagonal matrix is unitary. The unitarity property of $S_x$ follows immediately.

The proof for $S_y$ is equivalent, with the exception that the unitary Fourier matrix now represents an FFT along the $y$ direction. We can prove then that $S_y$ is unitary and thus $H_{\text{Rot-}x}$ is unitary.

As already mentioned, the proof for $H_{\text{Rot-}y}$ and $H_{\text{Rot-}z}$ are analogous. Combining Proof 1 and Proof 2, the unitary property of $H_{\theta_n}$ is demonstrated.

II. SEPARABLE QUADRATIC SURROGATE (SQS) FUNCTION DERIVATION FOR THE JOINT MLE

In order to get the final version of the joint MLE algorithm, a necessary step was to obtain a surrogate function for $S$. The following expression: $G_n(\kappa, T_1|\kappa^k, T_1^k) = ||W_n^{1/2}(H_{\theta_n} f_n(\kappa, T_1) - \hat{s}_n^k)||_2^2 + C_n(k), \quad (S8)$

with $W_n = \text{diag}\{\frac{1}{2\sigma_n^2}\}$.

The choice we made in this work was as a SQS function [5], that when applied to Eq.(S8), yields Eq.(16-17). Here, we present the proof of these equations. To that end, we build on results presented in [5]. In that work, a SQS function was applied to a generic quadratic form $\frac{1}{2}||y - Ax||_2^2$. Such SQS function had the following expression:

$$\frac{1}{2}||x - (x^k - D_f^{-1}A^H(Ax^k - y))||_2^2 \leq \xi, \quad (S9)$$

with $\xi$ a constant independent of $x$ and where the matrix $D_f$ is defined in such way that it satisfies $D_f \succeq A^H A$, that is, $D_f - A^H A$ is a positive-semidefinite matrix.

We can easily identify the terms of the quadratic form at hand, i.e., $G_n(\kappa, T_1|\kappa^k, T_1^k)$, with the terms of $\frac{1}{2}||y - Ax||_2^2$, and hence easily define our SQS function, $G_n(\kappa, T_1|\kappa^k, T_1^k)$, as

$$G_n(\kappa, T_1|\kappa^k, T_1^k) = ||f_n(\kappa, T_1) - \rho_n(\kappa^k, T_1^k)||_2^2 + C_n^*(k), \quad (S10)$$

with

$$\rho_n(\kappa^k, T_1^k) = f_n(\kappa^k, T_1^k) - D_f^{-1}A^H(Af_n(\kappa^k, T_1^k) - y), \quad (S11)$$

$C_n^*(k)$ a constant independent of $\kappa$ and $T_1$, and where $A = W_n^{1/2} H_{\theta_n} f_n$ and $y = W_n^{1/2} \hat{s}_n^k$.

After some algebra, we obtain

$$\rho_n(\kappa^k, T_1^k) = f_n(\kappa^k, T_1^k) - D_f^{-1}H_{\theta_n}^H W_n (\hat{s}_n - H_{\theta_n} f_n(\kappa^k, T_1^k)). \quad (S12)$$

Before giving an expression for $D_f$ satisfying $D_f \succeq A^H A$, first we recognize that

$$A^H A = H_{\theta_n}^H W_n^{1/2} H_{\theta_n}^{1/2} H_{\theta_n}^H W_n^{1/2} H_{\theta_n}^{1/2} = H_{\theta_n}^H W_n H_{\theta_n}^{1/2}, \quad (S13)$$

Furthermore, it is easy to show that the diagonal matrix $W_n$ fulfills $W_n \succeq (\sigma_n^*)^{-1} I$ with $(\sigma_n^*)^{-1}$ being the maximum value along its diagonal, which is

$$\frac{1}{2} \min_{m} (\sigma_n^*)^2. \quad (S14)$$

Thus, if $W_n \succeq (\sigma_n^*)^{-1} I$, it follows that $H_{\theta_n}^H W_n H_{\theta_n}^{1/2} \succeq (\sigma_n^*)^{-1} H_{\theta_n}^H H_{\theta_n}^{1/2} = (\sigma_n^*)^{-1} I$, since the motion operator is unitary. Therefore, by defining $D_f$ as $D_f \succeq (\sigma_n^*)^{-1} I$, $D_f \succeq A^H A$ holds.

Note that previous $D_f$ definition depends on $n$ and hence the NNLS problem (Eq.(S9)) is weighted differently along dimension $n$. To provide an unweighted NNLS problem, that is, the version we have presented in the main body of the paper, we set $\sigma_n^* = 2 \min_{m} (\sigma_n^*)^2$. Clearly $W_n \succeq (\sigma_n^*)^{-1} I$ for all $n$.

Hence, we redefine $D_f$ as $D_f \succeq (\sigma_n^*)^{-1} I$, and trivially we get $D_f \succeq A^H A$ as desired. By substituting $D_f$ into Eq.(S11), we arrive at the final expressions which are shown in the main body of the paper.

III. GRAPHS OF MOTION

![Graphs of the ground-truth and estimated motion parameters for one realization of the simulation experiment with rotational motion: (a) $\epsilon_x$, (b) $\epsilon_y$, (c) $\epsilon_z$, (d) $\alpha$, (e) $\beta$, (f) $\gamma$.](chart.png)
IV. EXTENSION OF THE JOINT MLE TO ACCOUNT FOR INTRA-IMAGE MOTION

As mentioned in the discussion section of the main body of the paper, the joint MLE can be extended to include intra-image motion, in particular, motion between the acquisition of the different slices of a multi-slice image. A brief outlook to such an extension is given here. The implementation of the extended algorithm should take into account the following considerations. First, given an inversion time $T_{1,n}$, the $z$th-noiseless and motion-corrupted 2D slice $T_{1,n,z}$ is related to the unobserved 3D image, $f_n(\kappa, T_1)$, through the motion parameters $\theta_{n,z}$. Note that the number of motion parameters scales with $M_z \times N$, where $M_z$ is the number of slices. Second, the mapping between a 2D slice $T_{1}$-weighted image and the noiseless unobserved 3D image also requires a slice-selective profile filter, which can be included as a matrix $\delta_z$ [6], just after the motion operator, that is, $f_{n,z}(\theta_{n,z}, \kappa, T_1) = \delta_z H_{\theta_{n,z}} f_n(\kappa, T_1)$. The final details of the derivation of the MM algorithm are beyond the scope of this paper and therefore not presented here.

REFERENCES

Fig. 4: Graphs of the estimated motion parameters for in vivo sagittal human brain data experiment. (a) $t_x$, (b) $t_y$, (c) $t_z$, (d) $\alpha$, (e) $\beta$, (f) $\gamma$. 