

# This item is the archived peer-reviewed author-version of:

Joint Maximum Likelihood Estimation of Motion and T1 Parameters from Magnetic Resonance Images in a Super-resolution Framework: a Simulation Study

## **Reference:**

Beirinckx Quinten, Ramos-Llorden Gabriel, Jeurissen Ben, Poot Dirk H.J., Parizel Paul M., Verhoye Marleen, Sijbers Jan, den Dekker Arjan. 2020. Joint Maximum Likelihood Estimation of Motion and T-1 Parameters from Magnetic Resonance Images in a Super-resolution Framework: a Simulation Study, Fundamenta Informaticae – 172(2):105-128.

Full text (Publisher's DOI): https://doi.org/10.3233/fi-2020-1896

# Joint Maximum Likelihood Estimation of Motion and $T_1$ Parameters from Magnetic Resonance Images in a Super-resolution Framework: a Simulation Study

#### Quinten Beirinckx

imec - Vision Lab Department of Physics University of Antwerp, Antwerp, Belgium

#### **Ben Jeurissen**

imec - Vision Lab Department of Physics University of Antwerp, Antwerp, Belgium

#### Paul M. Parizel

David Hartley Chair of Radiology Royal Perth Hospital & University of Western Australia Medical School Perth, WA, Australia

#### Jan Sijbers

imec - Vision Lab Department of Physics University of Antwerp, Antwerp, Belgium

#### **Gabriel Ramos-Llordén**

Psychiatry Neuroimaging Laboratory Brigham and Women's Hospital Harvard Medical School, Boston, MA, USA

#### **Dirk H.J. Poot** Biomedical Imaging Group Rotterdam Departments of Medical informatics and Radiology Erasmus MC, Rotterdam, The Netherlands

Marleen Verhoye Bio-Imaging Lab Department of Biomedical Sciences University of Antwerp, Antwerp, Belgium

#### Arnold J. den Dekker

imec - Vision Lab Department of Physics University of Antwerp, Antwerp, Belgium and Delft Center for Systems and Control Delft University of Technology, Delft, The Netherlands

Address for correspondence: Quinten Beirinckx, imec - Vision Lab, Department of Physics, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium. E-mail: quinten.beirinckx@uantwerpen.be

**Abstract.** Magnetic resonance imaging (MRI) based  $T_1$  mapping allows spatially resolved quantification of the tissue-dependent spin-lattice relaxation time constant  $T_1$ , which is a potential biomarker of various neurodegenerative diseases, including Multiple Sclerosis, Alzheimer disease, and Parkinson's disease. In conventional  $T_1$  MR relaxometry, a quantitative  $T_1$  map is obtained from a series of  $T_1$ -weighted MR images. Acquiring such a series, however, is time consuming. This has sparked the development of more efficient  $T_1$  mapping methods, one of which is a super-resolution reconstruction (SRR) framework in which a set of low resolution (LR)  $T_1$ -weighted images is acquired and from which a high resolution (HR)  $T_1$  map is directly estimated.

In this paper, the SRR  $T_1$  mapping framework is augmented with motion estimation. That is, motion between the acquisition of the LR  $T_1$ -weighted images is modeled and the motion parameters are estimated simultaneously with the  $T_1$  parameters. Based on Monte Carlo simulation experiments, we show that such an integrated motion/relaxometry estimation approach yields more accurate  $T_1$  maps compared to a previously reported SRR based  $T_1$  mapping approach.

**Keywords:** quantitative magnetic resonance imaging, super-resolution,  $T_1$  mapping, maximum likelihood estimation, motion correction

### 1. Introduction

 $T_1$  mapping is a quantitative Magnetic Resonance Imaging (MRI) technique that generates maps of the tissue-specific spin-lattice relaxation time  $T_1$  [1]. There is growing evidence that  $T_1$  mapping can be applied to detect subtle microscopic tissue damage, with potential for earlier diagnosis of various brain diseases including multiple sclerosis [2–4], epilepsy [5] and Alzheimer's disease [6]. Despite these promising results,  $T_1$  mapping currently remains a research tool and is not yet part of routine clinical assessment. The main obstacle for clinical adaptation of  $T_1$  mapping is that conventional  $T_1$  mapping techniques require long scan times to achieve adequate accuracy, precision and spatial resolution.

The gold standard method for  $T_1$  mapping, the inversion recovery method, is presented in Fig. 1. When an object is placed in a strong magnetic field, its nuclear spins align to this magnetic field, resulting in a net magnetic moment oriented in the so-called longitudinal direction (i.e., the direction parallel to the external magnetic field, corresponding with the z-direction in Fig. 1). Next, this equilibrium state is disturbed by applying a 180° radio frequency (RF) pulse, which inverts the longitudinal magnetization. After this pulse, the spins start to relax back towards the equilibrium state with a time constant  $T_1$  [7]. After inversion time TI, the longitudinal component is tipped into the transverse plane by a 90° RF pulse, after which the ( $T_1$ -weighted) MR signal is measured. In this way,  $T_1$ -weighted images are acquired at different inversion times. Subsequently, a  $T_1$  map is estimated by voxel-wise fitting a parametric model to these images. Since many images are required in such an acquisition scheme, conventional  $T_1$  mapping suffers from long acquisition times.

A simple way to reduce acquisition time is to lower the number of  $T_1$ -weighted images. This, however, results in a loss of precision in the estimated  $T_1$  map. Alternatively, a large number of  $T_1$ -weighted images can be acquired in a short acquisition time by reducing the acquisition time of each individual  $T_1$ -weighted image by lowering their spatial resolution. Commonly, this is done by acquiring multi-slice images where the slice-thickness is much larger than the spatial resolution within



Figure 1. A: Inversion Recovery Sequence: The longitudinal net nuclear magnetization vector is inverted by a 180° pulse. After inversion time TI, the longitudinal component is tipped into the transverse plane by a 90° pulse, after which the ( $T_1$ -weighted) MR signal is measured. By acquiring multiple MR signals (images) at different inversion times, the recovery of the longitudinal magnetization towards its equilibrium value can be sampled. The time between two repetitions of the sequence, i.e. the time between the inversion pulses, is called the repetition time TR. B: Effect of the inversion recovery sequence on the net nuclear longitudinal magnetization vector  $M_z$  as seen in the RF-rotating frame.  $B_0$  represents the external magnetic field vector. (a) Initial net nuclear longitudinal magnetization in alignment with  $B_0$ , (b) the 180° inverts the longitudinal magnetization  $M_z$ , (c)-(d) the longitudinal magnetization  $M_z$  is tipped into the transverse plane by a 90° pulse before readout.

the slice, i.e., the through-plane resolution is much lower than the in-plane resolution. Additionally, increasing the slice thickness increases the signal-to-noise ratio (SNR) of the  $T_1$ -weighted images, as signal strength scales linearly with imaged volume. However, thicker slices also suffer from increased partial volume effects, which arise when different tissues occur within a single voxel. In summary, reducing the acquisition time in conventional  $T_1$  mapping is clearly a trade-off in which faster scanning comes at the cost of either a lower precision or a lower spatial resolution and increased partial volume effects, of the resulting  $T_1$  map.

To improve this trade-off, a super-resolution reconstruction (SRR) method was recently proposed that estimates a 3D high resolution  $T_1$  map with isotropic voxel size from a set of low resolution  $T_1$ weighted multi-slice images with different slice orientations and anisotropic voxel size [8]. These low resolution images are acquired at a high in-plane resolution and a low through-plane resolution. It was shown that this method indeed provides a better trade-off between resolution, precision and acquisition time than direct high-resolution acquisition [9]. In this approach, motion was compensated for by adjusting the transformation parameters constituting the motion operator in a preprocessing step, and fixing these parameters in the SR-T1 estimation routine that followed. Fixing the motion parameters, however, may lead to inaccurate (i.e., biased)  $T_1$  maps since no feedback mechanism is present in the SR-T1 estimation routine that can undo incorrect fixation of motion parameters. As such, errors that potentially exist in the motion estimation step might propagate into the T1 estimation. At the same time, another recent work has proposed a unified Maximum Likelihood framework for simultaneous motion and  $T_1$  estimation in non-super-resolution  $T_1$  mapping [10]. It was demonstrated that the joint incorporation of the relaxation model, the motion model as well as the data statistics provide substantially more accurate motion and  $T_1$  parameter estimates. In the present paper, we explore, by means of simulation experiments, the potential of combining both approaches, resulting into joint Maximum Likelihood estimation of  $T_1$  and motion in a super-resolution framework.

The remainder of this paper is organized as follows. In Section 2, the image acquisition model, the proposed joint Maximum Likelihood estimator (MLE) and its implementation are described. Section 3 describes the simulation experiments, of which the results are presented and discussed in Section 4. Finally, in Section 5 conclusions are drawn.

## 2. Theory

The proposed method starts from a set of  $N T_1$ -weighted multi-slice images, each with a different slice direction. The multi-slice images, which are assumed to have a high in-plane resolution and a low through-plane resolution, will be referred to as the low resolution (LR) images. That is, the slice thickness, or through-plane voxel size, is larger than the in-plane voxel size, leading to anisotropic voxels. The method then estimates a high-resolution (HR)  $T_1$  map with isotropic voxels from a set of LR multi-slice  $T_1$ -weighted images and simultaneously estimates the motion between the acquisition of these LR images.

In the derivation of the imaging model, we will assume that the LR  $T_1$ -weighted images are acquired with a multi-slice inversion recovery (IR) conventional spin echo (SE) sequence, being the gold standard sequence for  $T_1$  mapping [11–13].

#### 2.1. MR imaging model

Let  $T_1 = (T_1(j)) \in \mathbb{R}^{N_r \times 1}$  be the vector containing the values of the unknown  $T_1$  map at the HR grid points  $\{x_j\}$  (with  $x_j \in \mathbb{R}^{3 \times 1}$  and j the HR voxel index,  $j = 1, \ldots, N_r$ ). Furthermore, let  $s_n \in \mathbb{R}^{N_s \times 1}$ , with  $n = 1, \ldots, N$ , denote the vector containing the intensities of the noiseless LR  $T_1$ -weighted multi-slice image with slice direction n, acquired with inversion time  $TI_n$ , and sampled at the LR grid points  $\{y_{nl}\}$  (with  $y_{nl} \in \mathbb{R}^{3 \times 1}$  and l the LR voxel index,  $l = 1, \ldots, N_s$ ). To derive

the mathematical relation between the HR  $T_1$  map of interest,  $T_1$ , and the LR image  $s_n$ , we now first introduce the virtual, noise free, HR  $T_1$ -weighted image  $r_n = (r_n(j)) \in \mathbb{R}^{N_r \times 1}$ , which is assumed to be acquired with the same inversion time  $TI_n$  as  $s_n$  and sampled at the (nonrotated) HR grid points of  $T_1$ .

Then,  $r_n$  can be modeled as a function of  $T_1$  and a quantity  $\rho = (\rho_j) \in \mathbb{R}^{N_r \times 1}$ , which is proportional to the proton density [14]:

$$r_n(j) = \rho(j) \left( 1 - 2e^{-\frac{\mathrm{TI}_n}{T_1(j)}} \right), \tag{1}$$

where we have assumed a perfect inversion pulse of  $180^{\circ}$  and a repetition time TR  $\gg T_1$ . In the remainder of this paper,  $T_1$  and  $\rho$  will be referred to as *relaxation model parameters* to be estimated.

Mathematically, we can now express the LR image  $s_n$  as the result of applying a sequence of operators on the virtual HR image  $r_n$ :

$$\boldsymbol{s}_n = \boldsymbol{D}\boldsymbol{B}\boldsymbol{G}_n\boldsymbol{M}_{\boldsymbol{\theta}_n}\boldsymbol{r}_n,\tag{2}$$

with  $M_{\theta_n} \in \mathbb{R}^{N_r \times N_r}$ ,  $G_n \in \mathbb{R}^{N_r \times N_r}$ ,  $B \in \mathbb{R}^{N_r \times N_r}$  and  $D \in \mathbb{R}^{N_s \times N_r}$  linear operators that describe, respectively, unintended motion, a known geometric transformation, spatially invariant blurring, and downsampling. In this work, we assume the unintended motion  $M_{\theta_n}$  to be rigid, parameterized by

$$\boldsymbol{\theta}_n = (t_{xn}, t_{yn}, t_{zn}, \alpha_n, \beta_n, \gamma_n)^T, \tag{3}$$

with  $t_{xn}, t_{yn}, t_{zn}$  the translation parameters and  $\alpha_n$ ,  $\beta_n$ ,  $\gamma_n$  the Euler angles of three elementary rotation matrices that describe rotation around the x, y and z axis, respectively. The superscript T in Eq. (3) denotes the transpose operation. In the present work, we use the same implementation of the rigid motion operator  $M_{\theta}$  as in [10], where they used the fact that rotation matrices can be decomposed as the product of three shear matrices. Each of the shearings can be implemented efficiently with Fast Fourier Transforms (FFT). Translation is implemented using the FFT as well. The operator  $G_n$  applies a known geometric transformation that models the image acquisition with a specific slice direction. More specifically, operator  $G_n$  models the SRR acquisition, in which multiple LR  $T_1$ -weighted images at different orientations are acquired by rotation of the acquisition plane for each image around one fixed encoding axis. In our implementation, operator  $G_n$  is a simplified version of  $M_{\theta}$  that models rotation around one fixed encoding axis. However, whereas  $M_{\theta_n}$  models the effect of unintended motion, which is unknown and has to be estimated from the data, the geometric transformation  $G_n$  is known and determined by the prescribed slice direction of the LR image  $s_n$ .

The blurring operator B describes the point spread function (PSF) of the MRI acquisition process, which can be modeled as tensor product of the PSF in three orthogonal directions: the through-plane (i.e., slice-selection) direction and the two in-plane directions, which are known as the phase- and frequency-encoding direction. We currently consider the PSF in the through-plane direction only. This through-plane PSF depends on the slice selection method. In this incipient work, the operators B and D describing spatially invariant blurring and downsampling, respectively, are combined into one operator  $\overline{D}$  that performs downsampling by averaging along the through-plane direction [15]. Details about this specific  $\overline{D}$  operator (and corresponding adjoint operator) are given in Appendix A.

For convenience of expression, we define  $A_n = \overline{D}G_n M_{\theta_n}$ , and rewrite Eq. (2) as

$$\boldsymbol{s}_n = \boldsymbol{A}_n \boldsymbol{r}_n,\tag{4}$$

with  $A_n = (a_n(l, j)) \in \mathbb{R}^{N_s \times N_r}$ . By combining Eqs. (1) and (4), the noiseless signal in voxel l of the LR  $T_1$ -weighted image can be described in terms of the HR maps  $T_1 = (T_1(j))$  and  $\rho = (\rho(j))$ :

$$s_n(l; T_1, \rho) = \sum_{j=1}^{N_r} a_n(l, j) \rho(j) \left( 1 - 2e^{-\frac{\mathrm{TI}_n}{T_1(j)}} \right).$$
(5)

In this work, we will assume that magnitude images are acquired, as is common for spin echo IR sequences. The voxel intensities of magnitude images reflect only the magnitude of the longitudinal magnetization, disregarding polarity [16]. Hence, in the absence of noise, the magnitude images are described by  $|s_n|$ , with  $|\cdot|$  the point-wise modulus operator.

Obviously, real-world images will be subject to noise. In this work, the noise is assumed to be additive, zero mean Gaussian noise. It has been shown that this is a valid assumption when the signal-to-noise ratio of the magnitude data is sufficiently high (> 3) [17–20], which is typically the case for the LR images. Hence, if we denote the acquired LR magnitude images by  $\tilde{s}_n \in \mathbb{R}^{N_s \times 1}$ , our image acquisition model can be described as

$$\tilde{\boldsymbol{s}}_n = |\boldsymbol{A}_n \boldsymbol{r}_n| + \boldsymbol{e}_n, \quad n = 1, \dots N, \tag{6}$$

with  $e_n \in \mathbb{R}^{N_s \times 1}$  a vector containing zero mean Gaussian noise contributions.

#### 2.2. The joint Maximum Likelihood estimator

Having derived the imaging model in subsection 2.1, this subsection will describe the model-based SRR framework that estimates an HR  $\rho$  and  $T_1$  map ( $\rho$  and  $T_1$ , respectively) simultaneously with the motion parameters  $\theta_n$ , n = 1, ..., N, from a set of LR images  $\{\tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_N\}$ , using a joint Maximum Likelihood estimator (MLE). The MLE is chosen because it is asymptotically unbiased, efficient (i.e., most precise) and consistent [21]. The MLE fully exploits prior knowledge on the statistical distribution of the data. In our case, the data consists of the voxels of the LR images. These voxels can be modeled as random variables that, due to the presence of noise, fluctuate about their expected values which are described by the model  $|A_n r_n|$  that was derived in subsection 2.1. Assuming additive, zero-mean Gaussian distributed noise, the PDF of a voxel  $\tilde{s}_n(l)$ , with  $l = 1, ..., N_s$ , of the image  $\tilde{s}_n$  is given by:

$$p_{\tilde{s}_n(l)}(\tilde{s}_n(l); \boldsymbol{T_1}, \boldsymbol{\rho}, \boldsymbol{\theta}_n) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\tilde{s}_n(l) - s_n(l; \boldsymbol{T_1}, \boldsymbol{\rho}, \boldsymbol{\theta}_n))^2}{2\sigma^2}},\tag{7}$$

with  $\sigma$  the standard deviation of the noise, which in this work is assumed to be spatially and temporally invariant. Assuming all voxels of all  $T_1$ -weighted LR images statistically independent, the joint PDF of all voxels is given by:

$$p_{\tilde{\mathbf{s}}}(\tilde{\mathbf{s}}; \boldsymbol{T_1}, \boldsymbol{\rho}, \boldsymbol{\theta}) = \prod_{n=1}^{N} \prod_{l=1}^{N_s} p_{\tilde{s}_n(l)}(\tilde{s}_n(l); \boldsymbol{T_1}, \boldsymbol{\rho}, \boldsymbol{\theta}_n),$$
(8)

with  $\tilde{\mathbf{s}} = (\tilde{\mathbf{s}}_1^T, \dots, \tilde{\mathbf{s}}_N^T)^T$  and  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_N^T)^T$ . To simplify the notation, let us define the parameter vector  $\boldsymbol{\tau} = (\boldsymbol{T}_1^T, \boldsymbol{\rho}^T, \boldsymbol{\theta}^T)^T$ .

To construct the MLE of  $\tau$ , the *likelihood function*  $L(\tau|\tilde{s})$ , which is the joint PDF of Eq. (8) regarded as a function of the unknown parameter vector  $\tau$  (with  $\tilde{s}$  fixed), is needed. The MLE  $\hat{\tau}_{ML}$  of the parameter vector  $\tau$  from measured data  $\tilde{s}$  is that value of  $\tau$  that maximizes the likelihood function  $L(\tau|\tilde{s})$ , or equivalently, the so-called log-likelihood function  $\mathcal{L}_{\tilde{s}}(\tau|\tilde{s}) \triangleq \log L(\tau|\tilde{s})$ , with respect to  $\tau$ , i.e.,

$$\hat{\boldsymbol{\tau}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{\sigma}} \mathcal{L}_{\tilde{\mathbf{s}}}(\boldsymbol{\tau}|\tilde{\mathbf{s}}). \tag{9}$$

It follows from Eq. (8) that the log-likelihood function can be written as

$$\mathcal{L}_{\tilde{\mathbf{s}}}(\boldsymbol{\tau}|\tilde{\mathbf{s}}) = -NN_{s}\ln\left(\sqrt{2\pi}\sigma\right) - \frac{1}{2\sigma^{2}}\sum_{n=1}^{N}\sum_{l=1}^{N_{s}}\left(\tilde{s}_{n}(l) - s_{n}(l;\boldsymbol{T}_{1},\boldsymbol{\rho},\boldsymbol{\theta}_{n})\right)^{2}.$$
 (10)

Hence, the ML estimator  $\hat{\tau}_{ML}$  is equal to the ordinary (unweighted) least-squares estimator:

$$\hat{\tau}_{\rm ML} = \arg\min_{\boldsymbol{\tau}} J(\boldsymbol{\tau}),\tag{11}$$

with

$$J(\boldsymbol{\tau}) = \sum_{n=1}^{N} \sum_{l=1}^{N_s} \left( \tilde{s}_n(l) - s_n(l; \boldsymbol{T}_1, \boldsymbol{\rho}, \boldsymbol{\theta}_n) \right)^2.$$
(12)

The non-linear optimization problem (11) can be solved using the alternating minimization method, also known as the cyclic block-coordinate descent (cBCD) method [22,23]. In this method, the parameter vector  $\tau$  is split into blocks and the cost function  $J(\tau)$  is successively minimized with respect to each block in a cyclic order. In our case, we use a split into blocks that contain the motion parameters and relaxation model parameters, respectively. In this way, the large-scale optimization problem (11) is separated into more easily solvable problems [10]. Moreover, it can be shown that this cBCD method assures a convergence property where  $J(\tau)$  decreases at every iteration [22]. Convergence to at least a local minimum is guaranteed [24]. In summary, the joint MLE is obtained by the following iterative recursive procedure:

$$\hat{\boldsymbol{\theta}}^{(t+1)} = \arg\min_{\boldsymbol{\theta}} J(\hat{\boldsymbol{T}}_{1}^{(t)}, \hat{\boldsymbol{\rho}}^{(t)}, \boldsymbol{\theta})$$
(P.1)

$$\hat{T}_{1}^{(t+1)}, \hat{\rho}^{(t+1)} = \arg\min_{T_{1}, \rho} J(T_{1}, \rho, \hat{\theta}^{(t+1)})$$
(P.2)

with  $\hat{\theta}^{(0)} = \theta_{\text{ini}}$ ,  $\hat{\rho}^{(0)} = \rho_{\text{ini}}$  and  $\hat{T}_1^{(0)} = T_{1\text{ini}}$  the initial values of the parameters  $\theta$ ,  $\rho$  and  $T_1$ , respectively. By its definition, this procedure produces a nonincreasing sequence of cost function values [23]. The procedure is terminated when the number of iterations exceeds  $t_{\text{max}}$  or when  $\mathcal{E}^{(t)} < \mathcal{E}_{\text{min}}$ , where  $\mathcal{E}^{(t)} = J(\hat{\tau}^{(t-1)}) - J(\hat{\tau}^{(t)})$ , and consecutive iterations are started from  $\mathcal{E}^{(0)} = r\mathcal{E}_{\text{min}}$ , with  $r \in \mathbb{R}_{>1}$ . The pseudo-code of the joint MLE algorithm is presented in Algorithm 1.

#### Algorithm 1: Joint MLE

Input: LR images  $\tilde{s}$  and initial values  $\theta_{\text{ini}}$ ,  $\rho_{\text{ini}}$  and  $T_{1\text{ini}}$ Output: ML estimates  $\hat{\theta}_{\text{ML}}$ ,  $\hat{T}_{1\text{ML}}$  and  $\hat{\rho}_{\text{ML}}$ Set  $t \leftarrow 0$  and  $\hat{\theta}^{(0)}$ ,  $\hat{T}_{1}^{(0)}$ ,  $\hat{\rho}^{(0)} \leftarrow \theta_{\text{ini}}$ ,  $\rho_{\text{ini}}$ ,  $T_{1\text{ini}}$ ;  $\mathcal{E}^{(0)} = r\mathcal{E}_{\text{min}}$ , with  $r \in \mathbb{R}_{>1}$ ; while  $\mathcal{E}^{(t)} \geq \mathcal{E}_{\text{min}}$  and  $t < t_{\text{max}}$  do Solve (P.1) to get  $\hat{\theta}^{(t+1)}$ :  $\hat{\theta}^{(t+1)} = \arg\min_{\theta} J(\hat{T}_{1}^{(t)}, \hat{\rho}^{(t)}, \theta)$ , started from  $\theta \leftarrow \hat{\theta}^{(t)}$ ; Solve (P.2) to get  $\hat{\rho}^{(t+1)}$ ,  $\hat{T}_{1}^{(t+1)}$ :  $\hat{T}_{1}^{(t+1)}, \hat{\rho}^{(t+1)} = \arg\min_{T_{1},\rho} J(T_{1}, \rho, \hat{\theta}^{(t+1)})$ , started from  $T_{1}, \rho \leftarrow \hat{T}_{1}^{(t)}, \hat{\rho}^{(t)}$ ; Calculate  $\mathcal{E}^{(t+1)} = J(\hat{T}_{1}^{(t)}, \hat{\rho}^{(t)}, \hat{\theta}^{(t)}) - J(\hat{T}_{1}^{(t+1)}, \hat{\rho}^{(t+1)}, \hat{\theta}^{(t+1)})$ ; Set  $t \leftarrow t + 1$ ; end  $\hat{\theta}_{\text{ML}} = \hat{\theta}^{(t)}, \hat{T}_{1\text{ML}} = \hat{T}_{1}^{(t)}$  and  $\hat{\rho}_{\text{ML}} = \hat{\rho}^{(t)}$ ; return  $\hat{\theta}_{\text{ML}}, \hat{T}_{1\text{ML}}, \hat{\rho}_{\text{ML}}$ ;

#### 2.3. Implementation

For the proposed joint MLE algorithm, in which the non-linear optimization problem is solved in alternating fashion between problems (P.1) and (P.2), the motion estimation problem (P.1) adopts a particularly simple structure when the relaxation model parameters are fixed. Assuming no dependence of  $\{\theta_n\}_{n=1}^N$  through index n, as is done here, the motion estimation problem can be decoupled into N independent minimization problems that can be evaluated in parallel. In the absence of additional information, a natural choice for the initialization of the motion parameters in the first iteration is a zero-motion initial condition such that the rigid motion operator  $M_{\theta_{ini}} = I$ , with I the identity matrix. Initial values  $\rho_{ini}$  and  $T_{1ini}$  were obtained by voxel-wise NLLS fitting the modulus of the relaxation model in Eq. (1) to the upsampled LR images with a Levenberg-Marquardt [25] algorithm, using the MATLAB routine lsqnonlin. Upsampling was performed using the adjoint operator sequence  $G_n^T \overline{D}^T$  acting on the LR images  $\tilde{s}_n$ , followed by application of  $|\cdot|$ , the point-wise modulus operator, to regain magnitude images.

The cost functions of LS problems (P.1) and (P.2) were minimized with a trust-region Newton algorithm using the MATLAB routine fminunc. Explicit analytical gradients were supplied for LS problem (P.1), while both explicit analytical gradients and implicit Hessian matrix elements (in the form of a matrix multiplication routine) were supplied for LS problem (P.2). The matrix multiplications in Eq. (2) were implemented by splitting the transformation operators  $M_{\theta_n}$  and  $G_n$  in sets of shear operations, each of which can be efficiently applied as a filtering operation in the frequency domain [10].

The computational complexity of the joint MLE algorithm is primarily defined by the FFT operations that are part of the implementation of operators  $G_n$  and  $M_{\theta_n}$  in equation Eq. (2). Using that a Q element 1D FFT has computational complexity of  $\mathcal{O}(Q \log_2(Q))$ , we derived that a single step of problem (P.1) required  $\mathcal{O}(69M^3 \log_2(M^2) + 5M^6 \log_2(M^6))$  floating point operations, given that  $G_n$  and  $M_{\theta_n}$  operate on HR images with isotropic dimensions  $M \times M \times M$ . This includes the operations introduced by the explicit analytical expressions for the gradient of the objective function of problem (P.1) w.r.t. motion parameters  $\theta_n$ . Furthermore, the given number of floating point operations should be multiplied by a factor N - 1, since problem (P.1) is optimized in parallel manner. In addition, a single step of problem (P.2) requires  $\mathcal{O}(N \cdot (48M^3 \log_2(M^2) + 4M^6 \log_2(M^6)))$  floating point operations. This number also includes the additional FFT operations introduced by the analytical gradient and Hessian expression for the objective function w.r.t. the  $T_1$  parameters. The computational requirements of operator  $\overline{D}$  were less demanding with one call of operator  $\overline{D}$  calculated up to 10 times faster than  $G_n$ , and up to 50 times faster than  $M_{\theta_n}$ , for the considered phantom size.

## 3. Simulation experiments

In this section, we describe the Monte Carlo (MC) simulation experiments that were carried out to evaluate the performance of the proposed joint MLE and to compare it with:

- SRR-T1: SR least squares (LS) estimation without motion correction. In this approach, the motion is simply ignored. That is, the least-squares criterion (12) is minimized with respect to the relaxation model parameters only, while fixing the motion parameters at  $\theta = 0$ .
- SRR-T1-MI: Mutual Information (MI) based registration prior to SR LS estimation. In this approach, the LR images  $\{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_N\}$  are first upsampled by applying the adjoint operator  $A_n^T$ . Next, the upsampled (HR) images are registered using a mutual information image similarity metric [26,27]. The motion parameters that result from this procedure are then substituted in the LS criterion (12), which is then minimized with respect to the relaxation model parameters only.
- SRR-T1-PRE: SRR T1 mapping described in [8]. In this approach, the motion parameters are estimated in a preprocessing step prior to the estimation of the HR  $T_1$  and  $\rho$  map. For this purpose, an iterative model-based motion correction scheme is used. First, the LR images  $\{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_N\}$  are upsampled by applying the adjoint operator  $A_n^T$ . The first time that this upsampling is performed, the motion operator  $M_{\theta_n}$  that co-constitutes A is set equal to the identity matrix. Next, a  $T_1$  and  $\rho$  map are estimated by voxel-wise fitting the modulus of model (1) to these upsampled images, using the Levenberg-Marquardt algorithm. Based on the thus obtained HR  $T_1$  and  $\rho$  maps, LR images are generated using model (5). These LR images are then rigidly aligned with the LR images  $\{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_N\}$  by minimizing their mean squared difference. The resulting motion parameter estimates are then used to update  $M_{\theta_n}$ . All steps are repeated until the stopping criterion is met with  $\mathcal{E}_{\min} = 10^{-4}$ . The motion parameters are then fixed and a HR  $T_1$  and  $\rho$  map is estimated using a LS estimator, with as initial values for  $T_1$  and  $\rho$  the values that resulted from the motion correction procedure.

Details about the specific simulation settings for each of the described MC simulation experiments are summarized in Table 1. Information about the initialization of (P.1) and (P.2) is also summarized in

Table 1. Simulation settings for the different Monte Carlo experiments: SR LS estimation without motion correction (SRR-T1), Mutual information based registration prior to SR LS estimation (SRR-T1-MI), SRR T1 mapping with preprocessing loop (SRR-T1-PRE), the proposed Joint MLE (SRR-T1-JMLE).

	SRR-T1	SRR-T1-MI	SRR-T1-PRE	SRR-T1-JMLE
Dimension of HR maps	$12\times12\times12$	$12 \times 12 \times 12$	$12 \times 12 \times 12$	$12 \times 12 \times 12$
Dimension of LR images	$12\times12\times6$	$12 \times 12 \times 6$	$12 \times 12 \times 6$	$12\times12\times6$
Spatial SNR	$snr \in [20, 30, \dots, 100]^{a}$			
Number of images $N$	14	14	14	14
Inversion times $TI_n$ [s]	$\mathrm{TI}_n \in [0.1, \ldots, 8]^{b}$			
# slice orientations	7	7	7	7
# TI per slice orientation	2	2	2	2
Slice orientation angles [°]	0:(180/7):154.28	0:(180/7):154.28	0:(180/7):154.28	0:(180/7):154.28
Initialization of (P.1)	n/a	MI registration <sup>c</sup>	preprocessing loop <sup>d</sup>	zero-motion IC <sup>e</sup>
Initialization of (P.2)	vw-NLLS-LM <sup>f</sup>	vw-NLLS-LM <sup>f</sup>	vw-NLLS-LM <sup>f</sup>	vw-NLLS-LM <sup>f</sup>
Optim. algorithm of (P.1)	n/a	n/a, fixed <sup>g</sup>	n/a, fixed <sup>g</sup>	trust-region Newton <sup>h</sup>
Optim. algorithm of (P.2)	trust-region Newton <sup>i</sup>	trust-region Newton <sup>g,i</sup>	trust-region Newton <sup>g,i</sup>	trust-region Newton <sup>i</sup>

<sup>a</sup>During simulations, nine different SNR values were studied ranging from 20 to 100, sampled with step size 10.

<sup>b</sup>Each of the LR images  $\{\theta_n\}_{n=1}^N$ , with N = 14, had a unique inversion time  $TI_n$ . The logarithms of these inversion times  $\{TI_n\}_{n=1}^N$  were equidistantly spaced between  $T_1 = 0.1s$  and  $T_{14} = 8s$ .

<sup>c</sup>The upsampled LR images are pairwise registered using MATLAB's imregtform function [28]. During the rigid registration process, the number of multi-level image pyramid levels is equal to two, and the first image of the series is chosen as a reference, hence  $\theta_1 = 0$ . The one-plus-one evolutionary optimizer configuration is used, for which the number of iterations is set to a very high value (> 5000) to ensure convergence of the motion parameter estimation. The remaining MI registration parameters are set to the default values of the MATLAB built-in code. Assuming no dependence of the motion parameters  $\{\theta_n\}_{n=1}^N$  through the index n, the different pairwise registration problems can be decoupled into N - 1 sub-problems, which can be implemented very efficiently with MATLAB parallel computing tools.

<sup>d</sup>Preprocessing loop: the iterative model-based motion correction scheme is used, as described in [8].

<sup>e</sup>zero-motion IC: In the absence of additional information, a natural choice for the initialization of the motion parameters in the first iteration is a zero-motion initial condition (IC) such that the rigid motion operator  $M_{\theta_{ini}} = I$ , with I the identity matrix.

<sup>f</sup>vw-NLLS-LM: voxel-wise NLLS fitting the modulus of relaxation model in Eq. (1) to upsampled LR images with Levenberg-Marquardt algorithm, using MATLAB's lsqnonlin routine, with the initial estimate per voxel chosen equal to  $[\rho, T_1] = [0.5, 1.5]$ . Upsampling was performed using the adjoint operator sequence  $G_n^T \overline{D}^T$  followed by application of  $|\cdot|$ , the point-wise modulus operator.

<sup>g</sup>Problem (P.2) is solved for the relaxation model parameters only, while fixing the motion parameters at those that result from the initialization procedure.

<sup>h</sup>MATLAB's fminunc routine, implemented with explicit analytical expressions for the gradient of the objective function. <sup>i</sup>MATLAB's fminunc routine, implemented with explicit analytical expressions for the gradient of the objective function and implicit Hessian matrix elements (in the form of a matrix multiplication routine).

Table 1 for each simulation experiment. To perform realistically adequate simulation experiments, a set of 2D multi-slice IR-SE  $T_1$ -weighted LR images affected by inter-image motion (as in Eq. (6)) and

10

noise was modeled from the ground truth  $T_1$  and proton density maps. These ground truth maps were based on a simple cubic  $12 \times 12 \times 12$  numerical phantom that was adopted from [8]. The phantom maps consisted of distinct regions, representing grey and white matter tissue parameters that were characterized using reported  $T_1$  and  $\rho$  (proton density) values in human brain tissue at 3T [29]. The ground truth  $T_1$  values for grey and white matter were 1607 ms and 838 ms, respectively, while those for  $\rho$  were set at 0.86 and 0.77 for the respective regions. An overview of this numerical phantom, which we refer to as Phantom 1, is shown in Fig. 2.



Figure 2. Phantom 1: Overview of the ground truth HR maps, and visualization of the downsampling along the slice dimension for one LR image.

From these ground truth  $T_1$  and  $\rho$  maps, a set of N = 14 noiseless LR  $T_1$ -weighted magnitude images was simulated using the forward model (6). The dimensions of each LR image were equal to  $12 \times 12 \times 6$ . The anisotropy factor (AF) was equal to 2. This AF is defined as the ratio between the through-plane slice thickness and the (isotropic) in-plane voxel size in the frequency encoding and phase encoding direction, see also Fig. 3. For each LR image N, the translational shifts  $t_{xn}, t_{yn}, t_{zn}$ and Euler angles  $\alpha_n, \beta_n, \gamma_n$  that define the rigid motion parameters  $\{\theta_n\}_{n=1}^N$  in Eq. (3) were generated randomly from a uniform distribution on the interval [-1, 1] (voxel units) and [-5, 5] degrees respectively. The reference image was chosen to be  $\tilde{s}_1$ , hence  $\theta_1 = 0$ . The same set of randomly generated rigid motion parameters was used for all simulation experiments. Furthermore, MATLAB's intrinsic coordinate system is used to represent 3D images, for which the origin is chosen at the center of the 3D image.

To account for wraparound artefacts that stem from the use of FFT to perform rotations, appropriate zeropadding was performed in each direction. Furthermore, it should be noted that since in our simulations the  $M \times M \times M$  discrete sampled image volume to be rotated had an even number of sampling points in each direction (i.e., M was even), the FFT based procedure required an extra multiplication with an exponential phase factor to obtain real values after rotation [30].

To fully recover the HR information, the LR images need to contain complementary information about the phantom. Rotation in image space corresponds to a rotation in frequency domain. As previously argued [31], acquiring the LR images with different slice orientations ensures that each LR image covers a different part of k-space (Fig. 3). In this way, the LR data will contain high spatial frequencies in all three dimensions. This approach results in more effective sampling of k-space than shifting the LR images by subpixel distances along the slice selection direction. In the latter case, the SR reconstruction result relies heavily on the success of recovering the aliased high frequency in the slice direction, since the narrow slice selection frequency band covers exactly the same part



Figure 3. Schematic comparison between image space and k-space (2D and 3D view) for a multi-orientation low resolution acquisition. The anisotropy of the voxels in image space is defined by the anisotropy factor,  $AF = \frac{b}{a}$  with b, the slice thickness, and a, the voxel size in the frequency encoding direction (and phase encoding direction). Since we choose to rotate only around the phase encoding axis, the k-space can only be sampled in a cylinder with diameter  $\frac{1}{a}$ .

of the k-space for each LR image. Similar to the acquisition protocol in [8, 32], the rotation of the LR images was performed around the virtual phase encoding axis in increments of  $180/N_o$  degrees, with  $N_o$  the number of slice orientations. Aiming at a short acquisition time, the number of slice orientations was kept low, but sufficiently high to ensure that the k-space is maximally covered. In particular, the number of slice orientations was fixed to 7, corresponding to having two different LR  $T_1$ -weighted magnitude images per slice orientation. An overview of the slice orientations and corresponding inversion times (TI), combined with their coverage of k-space is shown in Fig. 3. Each LR  $T_1$ -weighted magnitude image had a unique inversion time, with  $TI_n \in [0.1, \ldots, 8]s$ , where the TIs were sampled equidistantly in log-space.

After fixing the motion parameters and acquisition geometry, downsampling by averaging and the application of the point-wise modulus operator for magnitude images in accordance with (6), zero mean white Gaussian noise was added to the LR images. The noise level was chosen to obtain SNR values between 20 and 100, where SNR is the ratio of the spatial mean of the LR image with the highest TI and the standard deviation of the noise. For each SNR,  $N_{MC} = 140$  realizations of sets of LR images were generated. For each realization, a HR  $T_1$  and  $\rho$  map as well as the motion parameters  $\theta$  were estimated.

The different simulations experiments were implemented in MATLAB [28], and run on a computer with an Intel i7-6850K hexa-core CPU @ 3.6 GHz and 32 GB of RAM. The proposed joint MLE simulation experiment for the numerical phantom required around 4 GB RAM (allocated memory usage), and for a fixed tolerance of  $\mathcal{E}_{min} = 10^{-4}$ , on average about 14 alternating MLE iterations were needed to ensure convergence of the optimization process. In order to run the series of MC simulations quickly and efficiently, the University of Antwerp's High Performance Computing core (HPC) facility CalcUA was used.



Overview of the slice orientations of the LR images and corresponding inversion times  $TI_n$  (top row) and the k-space sampling strategy (middle and bottom row). The middle row shows the k-space sampling of seven individual LR images, each having a different slice orientation, whereas the bottom row shows the overlap in k-space when those images are combined. The shaded area denotes the sampled k-space, while the white region is not sampled.

To assess the performance of each method to estimate the  $T_1$  map, the following performance measures were used [10]:

- (a) *Relative bias*. The bias quantifies the *accuracy* or, equivalently, the systematic error of the estimator [21]. 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.
- (b) Relative standard deviation. The standard deviation quantifies the precision, or, equivalently, the non-systematic error of the estimator [21]. For each voxel, the relative sample standard deviation was calculated as  $\operatorname{std}(\hat{T}_1)/T_1$ , and an overall precision measure was obtained by taking the spatial mean of these relative sample standard deviations.
- (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.

By substitution of  $\rho$  for  $T_1$ , the performance of each method to estimate the  $\rho$  map was assessed in an identical way using measures (a)-(c).

To assess the ability of the proposed method to estimate motion, the following performance measure was used:

#### (d) Motion component root-(mean)-mean-square-error (RMMSE), defined as

$$\sqrt{\frac{1}{N-1}\sum_{n=2}^{N}\overline{\left(\left[\hat{\boldsymbol{\theta}}_{n}\right]_{j}-\left[\boldsymbol{\theta}_{n}\right]_{j}\right)^{2}}},$$
(13)

with  $[\theta_n]_j$  the *j*th component of  $\theta_n$  and  $\overline{[\hat{\theta}_n]_j}$  the sample mean of the  $N_{\text{MC}}$  estimates  $[\hat{\theta}_n]_j$ .

To supplement the results of Phantom 1, also a second, more challenging phantom was created (Fig. 4), which consisted of distinct grey and white matter tissue regions that included both uniform regions and checkerboard patterns, combined with horizontal and vertical planar structures. This second phantom will be referred to as Phantom 2. The same ground truth  $T_1$  values as for Phantom 1 were used to characterize grey and white matter voxels. The different Monte Carlo experiments were repeated for Phantom 2 for a fixed spatial SNR = 50, keeping the other simulation settings identical as for Phantom 1. The reconstruction results of both phantoms were visually compared w.r.t. their respective ground truth  $T_1$  parameter maps. For each phantom and for each method, an average  $T_1$  map was calculated by voxel-wise averaging over all  $N_{\rm MC}$  reconstruction results for SNR = 50.



Figure 4. Phantom 2: Overview of the ground truth HR maps, and visualization of the downsampling along the slice dimension for one LR image.

## 4. Results and Discussion

Figs. 5-7 summarize the statistical performance results that were obtained from the simulation experiments for Phantom 1. Fig. 5(a-c) and Fig. 6(a-c) show the overall relative sample bias, standard deviation and RMSE for all considered estimators of, respectively,  $T_1$  and  $\rho$ , as a function of the SNR. The results clearly show that for the SRR-T1 method without motion correction, the estimation of the relaxation model parameters performs poorly in terms of accuracy and precision, with high relative bias and high relative standard deviation, thereby underlining the importance of a proper motion estimation framework. Furthermore, SRR-T1-MI also shows a poor performance. In terms of precision, SRR-T1-MI performs even worse than SRR-T1. This observation is quite remarkable. It may reflect the limitations of intensity-based image registration when it comes to registering images of largely different contrast, which have been reported earlier [33, 34]. The poor performance of the SRR-T1-MI method may also be partly due to its implementation. In our implementation, the LR magnitude

images were naively upsampled using the adjoint operator  $A_n^T$  followed by application of the pointwise modulus operator. The latter preserves the magnitude characteristic of the resulting HR images after upsampling. However, information loss is inherent and unavoidable in this upsampling process, partially due to the non-existence of an adjoint modulus operation, and this might impede correct intensity-based registration of these HR images in the next step. Finally, the small size of the images considered in our simulation experiment may also play a role, as SRR-T1-MI may be more sensitive to the image size than the other methods considered. Next, the SRR-T1-PRE method clearly improves the estimation results, both in terms of accuracy and precision. For low SNR values, this method shows even a better relative standard deviation than the SRR-T1-JMLE approach, as can be observed from Fig. 5(b) and Fig. 6(b). However, this lower precision of the proposed SRR-T1-JMLE is more than compensated by its higher accuracy, resulting in the superior performance of SRR-T1-JMLE in terms of relative RMSE over the whole range of SNR values.

Fig. 7 shows the motion component RMMSE for each of the six rigid motion components as a function of the SNR. SRR-T1-JMLE clearly outperforms the other methods (SRR-T1, SRR-T1-MI, and SRR-T1-PRE) in terms of the motion component RMMSE. This is particularly visible for the rotation parameter components  $\alpha$ ,  $\beta$ ,  $\gamma$ , in the lower bottom half of Fig. 7. It is worth emphasizing once more that the motion parameter problem in the SRR-T1-JMLE method was initialized from a zero-motion initial condition, i.e.  $M_{\theta_{ini}} = I$ , with I the identity matrix. This also highlights the robustness of the SRR-T1-JMLE method for poor motion initialization scenarios. The quantitative performance measures for Phantom 2, for a fixed spatial SNR = 50, are summarized in Table 2. Results with this new phantom are very similar to those obtained with Phantom 1, underlining the superior performance of SRR-T1-JMLE in terms of relative RMSE and motion component RMMSE.

Finally, to provide a visual comparison of the performance of the different methods, for each method an average  $T_1$  map was calculated by voxel-wise averaging over all  $N_{MC}$  reconstruction results for SNR = 50. Fig. 8 shows the resulting average  $T_1$  maps for three orthogonal slices (sagittal, axial and coronal planes) of Phantom 1, accompanied by histograms of the voxel data distribution of the corresponding full 3D average  $T_1$  maps. The number of histogram bins was specified at 200. Note that the grey and white matter phantoms considered have histograms consisting of two distinct peaks corresponding with the ground truth  $T_1$  values of both tissues. From Fig. 8, for Phantom 1, it is clear that only with SRR-T1-PRE and SRR-T1-JMLE, these peaks can be distinguished. Moreover, Fig. 8 clearly shows that the block-wise homogeneous structure of the phantom is best reconstructed by SRR-T1-JMLE. The same observations can be made for Phantom 2, for which the results of the visual comparison are shown in Fig. 9.

Overall, the results clearly demonstrate the superior performance of the SRR-T1-JMLE method in terms of accuracy and relative RMSE for both motion and  $T_1$  and  $\rho$  estimation. This is also supported by the visual comparison presented in Fig. 8 and Fig. 9, where SRR-T1-JMLE clearly outperforms the other estimation frameworks.

In the outline of the MR imaging model under paragraph 2.1, the voxel intensity values of the HR images  $r_n$  were modeled by a three-parameter  $T_1$  model (Eq. (1)) that depends on  $T_1$  and  $\rho$ , for given inversion times TI<sub>n</sub> [14]. It should be noted that if the assumptions that substantiate the choice of this model, i.e. perfect inversion pulse of 180° and a repetition time TR  $\gg T_1$ , are invalid in practice, the proposed joint MLE method can still be used, but the model should be extended so as to avoid biased

results. Such an extension may include the introduction of additional unknown model parameters to be estimated from the data [35], which may have a negative influence on the precision. This well-known trade-off between bias and precision should always been taken into account in model selection.

According to the computational complexity of the SRR-T1-JMLE algorithm described in section 2.3, increasing the volume size of the images, i.e. choosing a larger region-of-interest (ROI), will result in longer computation times, as the number of floating point operations increases. As a possible solution to this problem, the ROI could be split in several blocks (with overlap to avoid edge artifacts), where the HR relaxation parameters are reconstructed in each block separately. This would allow parallelization of estimation problem (P.2), which would considerably reduce the computational complexity, memory consumption and computation time of the SRR-T1-JMLE method.



Figure 5. Results of the simulation experiments for Phantom 1: (a) relative  $T_1$  bias, (b) relative  $T_1$  standard deviation, (c) relative  $T_1$  RMSE, as a function of SNR. Error bars correspond with the standard error of the spatial mean, but are omitted when their size matches the order of the graph symbol size.

16



Figure 6. Results of the simulation experiments for Phantom 1: (a) relative  $\rho$  bias, (b) relative  $\rho$  standard deviation, (c) relative  $\rho$  RMSE, as a function of SNR. Error bars correspond with the standard error of the spatial mean, but are omitted when their size matches the order of the graph symbol size.

Table 2. Quantitative performance measures for Phantom 2, calculated over all  $N_{MC} = 140$  reconstruction results for SNR = 50, for the different Monte Carlo experiments: SR LS estimation without motion correction (SRR-T1), Mutual information based registration prior to SR LS estimation (SRR-T1-MI), SRR T1 mapping with preprocessing loop (SRR-T1-PRE), the proposed Joint MLE (SRR-T1-JMLE).

	SRR-T1	SRR-T1-MI	SRR-T1-PRE	SRR-T1-JMLE
Overall relative $T_1$ bias [%]	$71\pm3$	$42\pm2$	$6.3\pm0.1$	$1.83\pm0.03$
Overall relative $T_1$ standard deviation [%]	$8.0\pm0.5$	$46\pm2$	$0.51\pm0.01$	$2.50\pm0.03$
Overall relative $T_1$ RMSE [%]	$72 \pm 3$	$67 \pm 3$	$6.4\pm0.1$	$3.21\pm0.04$
Overall relative $\rho$ bias [%]	$16.1\pm0.3$	$15.2\pm0.3$	$2.23\pm0.04$	$0.309 \pm 0.005$
Overall relative $\rho$ standard deviation [%]	$1.44\pm0.05$	$9.6\pm0.1$	$0.204 \pm 0.001$	$0.686 \pm 0.005$
Overall relative $\rho$ RMSE [%]	$16.4\pm0.3$	$19.1\pm0.2$	$2.25\pm0.04$	$0.772 \pm 0.006$
RMMSE of $t_x$ [voxel units]	n/a	$0.59\pm0.06$	$0.06\pm0.01$	$0.014\pm0.002$
RMMSE of $t_y$ [voxel units]	n/a	$0.53\pm0.06$	$0.10\pm0.02$	$0.011\pm0.002$
RMMSE of $t_z$ [voxel units]	n/a	$0.54\pm0.05$	$0.05\pm0.02$	$0.010\pm0.001$
RMMSE of $\alpha$ [degrees]	n/a	$2.1\pm0.1$	$0.93\pm0.07$	$0.055 \pm 0.004$
RMMSE of $\beta$ [degrees]	n/a	$1.9\pm0.1$	$0.91\pm0.07$	$0.068 \pm 0.004$
RMMSE of $\gamma$ [degrees]	n/a	$2.4\pm0.1$	$1.54\pm0.09$	$0.060\pm0.005$



Figure 7. Results of the simulation experiments for Phantom 1, showing the motion component RMMSE for each of the six rigid motion components, as a function of SNR. Error bars are omitted when their size matches the order of the graph symbol size.



Figure 8. Visual comparison of the performance of the different SRR methods: SRR-T1 (row 1), SRR-T1-MI (row 2), SRR-T1-PRE (row 3), and our proposed SRR-T1-JMLE (row 4). On the left, three orthogonal slices of the  $T_1$  map obtained by voxel-wise averaging over all  $N_{MC} = 140$  reconstruction results for SNR=50. On the right, histograms showing the voxel data distribution of the corresponding full 3D  $T_1$  parameter map estimates for the different simulation experiments. The ground truth  $T_1$  values for grey and white matter, 1607 ms and 838 ms respectively, are marked with vertical lines.



Figure 9. SRR-T1-JMLE reconstruction result for Phantom 2: Ground Truth (row 1), SRR-T1-PRE (row 2), SRR-T1-JMLE (row 3). On the left, three orthogonal slices of the  $T_1$  map obtained by voxel-wise averaging over all  $N_{MC} = 140$  reconstruction results for SNR=50. On the right, histograms showing the voxel data distribution of the corresponding full 3D  $T_1$  parameter map estimate. The ground truth  $T_1$  values for grey and white matter, 1607 ms and 838 ms respectively, are marked with vertical lines.

## 5. Conclusion

Ouantitative MR  $T_1$  mapping suffers from long acquisition times with high risk for patient motion artefacts, resulting in poor accuracy of estimated  $T_1$  relaxometry parameters. In this work, we explored the potential of augmenting a recently proposed super-resolution reconstruction method for MRI  $T_1$  mapping with simultaneous motion estimation in a maximum likelihood framework. Superresolution reconstruction provides a better trade-off between resolution, precision and acquisition time than conventional direct high-resolution acquisition. By extending super-resolution reconstruction with simultaneous motion estimation, potential bias in the estimated  $T_1$  map caused by motion can be substantially reduced compared to motion correction by preprocessing. By means of Monte Carlo simulation experiments, our newly proposed method was quantitatively compared against a ground-truth  $T_1$  map together with three other approaches. The results were analysed using statistical performance measures and by performing a visual comparison. In conclusion, the results of these simulation experiments demonstrate that our newly proposed joint relaxometry and motion estimation approach yields more accurate  $T_1$  maps than a previously reported SRR based  $T_1$  mapping approach, in which motion registration is applied as a preprocessing step prior to  $T_1$  mapping. Future work will focus on the validation of the proposed joint MLE method on real data scenarios, the development of advanced blurring operators for the acquisition model, and the extension of the motion model to non-rigid or affine motion. In addition, it is also worthwhile to investigate intra-image motion correction strategies and to further customize the alternating optimization scheme.

## Appendix A Downsampling and upsampling operators

In the buildup of the forward model (4) encapsulated in operator A, operators B and D describing spatially invariant blurring and downsampling, respectively, are combined into one operator  $\overline{D}$  that performs downsampling by averaging [15]. Conventionally, downsampling keeps one sample out of a block and discards the remaining samples, whereas blurring takes into consideration the point spread functions of the MRI acquisition process. Downsampling by averaging is used here. For example, downsampling by a factor of 2 in 1D has matrix form,

$$\begin{bmatrix} \ddots & \ddots & & & \\ & \frac{1}{2} & \frac{1}{2} & & \\ & & & \frac{1}{2} & \frac{1}{2} & \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \boldsymbol{x}[0] \\ \boldsymbol{x}[1] \\ \boldsymbol{x}[2] \\ \boldsymbol{x}[3] \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{\vdots \\ \frac{\boldsymbol{x}[0] + \boldsymbol{x}[1]}{2} \\ \frac{\boldsymbol{x}[2] + \boldsymbol{x}[3]}{2} \\ \frac{\boldsymbol{x}[2] + \boldsymbol{x}[3]}{2} \\ \vdots \end{bmatrix}.$$
(A.14)

In the spatial (i.e. image) domain, this downsampling from a discrete vector  $\boldsymbol{x}[n]$  to  $\boldsymbol{y}[n]$  can be written in more compact form as

$$\boldsymbol{y}[n] = \overline{\boldsymbol{D}}\boldsymbol{x}[n] = \frac{1}{M} \sum_{m=0}^{M-1} \boldsymbol{x}[nM+m], \qquad (A.15)$$

where M corresponds with the anisotropy factor AF, defined as the ratio of the through-plane slice thickness and the (isotropic) in-plane voxel size.

Implementing the iterative recursive procedure described by problems (P.1) and (P.2) benefits from having adjoint operators. Interestingly, upsampling is the adjoint of downsampling. Conventionally, upsampling with zero insertion is used, we use upsampling with replication. For example, upsampling by a factor of 2 in 1D has matrix form,

$$\frac{1}{2} \begin{bmatrix} \ddots & & & \\ \ddots & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 & \\ & & & 1 & \ddots \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{y}[0] \\ \mathbf{y}[1] \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \frac{1}{2} \mathbf{y}[0] \\ \frac{1}{2} \mathbf{y}[0] \\ \frac{1}{2} \mathbf{y}[1] \\ \frac{1}{2} \mathbf{y}[1] \\ \vdots \end{bmatrix}.$$
(A.16)

The upsampling matrix is the transpose of the downsampling matrix in (A.15). We can write the upsampled  $\boldsymbol{x}[n]$  from  $\boldsymbol{y}[n]$  as

$$\boldsymbol{x}[n] = \overline{\boldsymbol{D}}^T \boldsymbol{y}[n] = \frac{1}{M} \boldsymbol{y}\left[\left\lfloor \frac{n}{M} \right\rfloor\right],$$
 (A.17)

where the floor function  $|\cdot|$  gives the largest integer less than or equal to its argument.

## Acknowledgements

The authors gratefully acknowledge support of the Industrial Research Fund of the Antwerp University Association, the European Space Agency and BELSPO Prodex (BrainDTI), the Research Foundation Flanders (FWO Belgium) through project funding G084217N and 12M3116N.

## References

- [1] Deoni SCL, Rutt BK, Peters TM. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. *Magn. Reson. Med.*, 2003. **49**(3):515–526. doi:10.1002/mrm.10407.
- [2] Larsson HBW, et al. Assessment of demyelination, edema, and gliosis by in vivo determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis. *Magn. Reson. Med.*, 1989. 11(3):337– 348. doi:10.1002/mrm.1910110308.
- [3] Vrenken H, Geurts JJG, Knol DL, van Dijk LN, Dattola V, Jasperse B, van Schijndel RA, Polman CH, Castelijns JA, Barkhof F, Pouwels PJW. Whole-brain T1 mapping in multiple sclerosis: global changes of normal-appearing gray and white matter. *Radiology*, 2006. 240(3):811–820. doi:10.1148/radiol. 2403050569. PMID: 16868279.
- [4] Papadopoulos K, Tozer DJ, Fisniku L, Altmann DR, Davies G, Rashid W, Thompson AJ, Miller DH, Chard DT. T1-relaxation time changes over five years in relapsing-remitting multiple sclerosis. *Mult. Scler. J.*, 2010. 16(4):427–433. doi:10.1177/1352458509359924.
- [5] Conlon P, Trimble MR, Rogers D, Callicott C. Magnetic resonance imaging in epilepsy: a controlled study. *Epilepsy Res.*, 1988. 2(1):37–43. doi:10.1016/0920-1211(88)90008-3.
- [6] Erkinjuntti T, Ketonen L, Sulkava R, Vuorialho M, Livanainen M. Do white matter changes on MRI and CT differentiate vascular dementia from Alzheimer's disease? J. Neurol. Neurosurg. Psychiatry, 1987. 50(1):37–42. doi:10.1136/jnnp.50.1.37.
- [7] Taylor AJ, Salerno M, Dharmakumar R, Jerosch-Herold M. T1 Mapping Basic Techniques and Clinical Applications. JACC: Cardiovasc. Imaging, 2016. 9(1, SI):67–81. doi:10.1016/j.jcmg.2015.11.005.
- [8] Van Steenkiste G, Poot DHJ, Jeurissen B, den Dekker AJ, Vanhevel F, Parizel PM, Sijbers J. Superresolution T1 estimation: Quantitative high resolution T1 mapping from a set of low resolution T1weighted images with different slice orientations. *Magn. Reson. Med.*, 2017. **77**(5):1818–1830. doi: 10.1002/mrm.26262.
- [9] Poot DHJ, Van Meir V, Sijbers J. General and efficient super-resolution method for multi-slice MRI. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010, volume 6361. 2010 pp. 615–622. doi:10.1007/978-3-642-15705-9\_75.
- [10] Ramos-Llordén G, den Dekker AJ, Van Steenkiste G, Jeurissen B, Vanhevel F, Van Audekerke J, Verhoye M, Sijbers J. A unified maximum likelihood framework for simultaneous motion and T<sub>1</sub> estimation in quantitative MR T<sub>1</sub> mapping. *IEEE Trans. Med. Imag.*, 2017. **36**(2):433–446. doi:10.1109/TMI.2016. 2611653.
- [11] Hahn EL. An accurate nuclear magnetic resonance method for measuring spin-lattice relaxation times. *Phys. Rev.*, 1949. **76**(1):145–146. doi:10.1103/physrev.76.145.
- [12] Drain LE. A direct method of measuring nuclear spin-lattice relaxation times. Proc. Phys. Soc. London, Sect. A, 1949. 62:301–306. doi:10.1088/0370-1298/62/5/306.

- [13] Crawley AP, Henkelman RM. A comparison of one-shot and recovery methods in T1 imaging. Magn. Reson. Med., 1988. 7(1):23–34. doi:10.1002/mrm.1910070104.
- [14] Bernstein MA, King KF, Zou XJ. Handbook of MRI pulse sequences. Elsevier Academic press, London, 2004.
- [15] Li Y, Matej S, Metzler SD. Image reconstructions from super-sampled data sets with resolution modeling in PET imaging. *Med. Phys.*, 2014. 41(12):121912. doi:10.1118/1.4901552.
- [16] Tofts P. Quantitative MRI of the brain: measuring changes caused by disease. Chichester, England: John Wiley & Sons, 2004.
- [17] Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magn. Reson. Med.*, 1995. 34(6):910–914. doi:10.1002/mrm.1910340618.
- [18] Andersen AH, Kirsch JE. Analysis of noise in phase contrast MR imaging. Med. Phys., 1996. 23(6):857– 869. doi:10.1118/1.597729.
- [19] Constantinides CD, Atalar E, McVeigh ER. Signal-to-noise measurements in magnitude images from NMR phased arrays. *Magn. Reson. Med.*, 1997. **38**(5):852–857. doi:10.1002/mrm.1910380524.
- [20] den Dekker AJ, Sijbers J. Data distributions in magnetic resonance images: A review. *Physica Med.*, 2014. 30(7):725–741. doi:10.1016/j.ejmp.2014.05.002.
- [21] van den Bos A. Parameter Estimation for Scientists and Engineers. Hoboken, New Yersey, USA: John Wiley & Sons, 2007. doi:10.1002/9780470173862.
- [22] Fessler JA, Kim D. Axial block coordinate descent (ABCD) algorithm for X-ray CT image reconstruction. In: Proc. Intl. Mtg. on Fully 3D image Recon. in Rad. and Nuc. Med. 2011 pp. 262–5.
- [23] Beck A, Tetruashvili L. On the Convergence of Block Coordinate Descent Type Methods. SIAM J. Optim., 2013. 23(4):2037–2060. doi:10.1137/120887679.
- [24] Fan J, Farmen M, Gijbels I. Local maximum likelihood estimation and inference. J. R. Stat. Soc. Series B. Stat. Methodol., 1998. 60(3):591–608. doi:10.1111/1467-9868.00142.
- [25] Nocedal J, Wright S. Numerical optimization. New York, NJ, USA: Springer-Verlag, 2006.
- [26] Mattes D, et al. PET-CT image registration in the chest using free-form deformations. *IEEE Trans. Med. Imag.*, 2003. 22(1):120–128. doi:10.1109/tmi.2003.809072.
- [27] Pluim JPW, Maintz JB, Viergever MA. Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imag.*, 2003. 22(8):986–1004. doi:10.1109/tmi.2003.815867.
- [28] The Mathworks, Inc., Natick, Massachusetts, United States. MATLAB and Image Processing Toolbox Release 2017b.
- [29] Wright PJ, Mougin OE, Totman JJ, Peters AM, Brookes MJ, Coxon R, Morris PE, Clemence M, Francis ST, Bowtell RW, Gowland PA. Water proton T1 measurements in brain tissue at 7, 3, and 1.5 T using IR-EPI, IR-TSE, and MPRAGE: results and optimization. *Magn. Reson. Mater. Phys., Biol. Med.*, 2008. 21(1-2):121–130. doi:10.1007/s10334-008-0104-8.
- [30] Larkin KG, Oldfield MA, Klemm H. Fast Fourier method for the accurate rotation of sampled images. Opt. Commun., 1997. 139(1-3):99–106. doi:10.1016/s0030-4018(97)00097-7.

- [31] Plenge E, Poot DHJ, Bernsen M, Kotek G, Houston G, Wielopolski P, Van Der Weerd L, Niessen WJ, Meijering E. Super-resolution methods in MRI: Can they improve the trade-off between resolution, signalto-noise ratio, and acquisition time? *Magn. Reson. Med.*, 2012. **68**(6):1983–1993. doi:10.1002/mrm. 24187.
- [32] Van Steenkiste G, Jeurissen B, Veraart J, den Dekker AJ, Parizel PM, Poot DHJ, Sijbers J. Super-resolution reconstruction of diffusion parameters from diffusion-weighted images with different slice orientations. *Magn. Reson. Med.*, 2016. **75**(1):181–195. doi:10.1002/mrm.25597.
- [33] Xue H, Shah S, Greiser A, Guetter C, Littmann A, Jolly MP, Arai AE, Zuehlsdorff S, Guehring J, Kellman P. Motion correction for myocardial T1 mapping using image registration with synthetic image estimation. *Magn. Reson. Med.*, 2012. **67**(6):1644–1655. doi:10.1002/mrm.23153.
- [34] Roujol S, Foppa M, Weingartner S, Manning W J, Nezafat R. Adaptive registration of varying contrastweighted images for improved tissue characterization (ARCTIC): Application to T1 mapping. *Magn. Reson. Med.*, 2015. **73**(4):1469–1482. doi:10.1002/mrm.25270.
- [35] Barral JK, et al. A robust methodology for in vivo T1 mapping. Magn. Reson. Med., 2010. 64(4):1057– 1067. doi:10.1002/mrm.22497.