

NOVIFAST: A Fast Algorithm for Accurate and Precise VFA MRI T_1 Mapping

Gabriel Ramos-Llordén¹, Member, IEEE, Gonzalo Vegas-Sánchez-Ferrero², Member, IEEE, Marcus Björk³, Floris Vanhevel, Paul M. Parizel, Raúl San José Estépar⁴, Member, IEEE, Arnold J. den Dekker, and Jan Sijbers

Abstract—In quantitative magnetic resonance T_1 mapping, the variable flip angle (VFA) steady state spoiled gradient recalled echo (SPGR) imaging technique is popular as it provides a series of high resolution T_1 weighted images in a clinically feasible time. Fast, linear methods that estimate T_1 maps from these weighted images have been proposed, such as DESPOT1 and iterative re-weighted linear least squares. More accurate, non-linear least squares (NLLS) estimators are in play, but these are generally much slower and require careful initialization. In this paper, we present NOVIFAST, a novel NLLS-based algorithm specifically tailored to VFA SPGR T_1 mapping. By exploiting the particular structure of the SPGR model, a computationally efficient, yet accurate and precise T_1 map estimator is derived. Simulation and *in vivo* human brain experiments demonstrate a twenty-fold speed gain of NOVIFAST compared with conventional gradient-based

NLLS estimators while maintaining a high precision and accuracy. Moreover, NOVIFAST is eight times faster than the efficient implementations of the variable projection (VARPRO) method. Furthermore, NOVIFAST is shown to be robust against initialization.

Index Terms— T_1 mapping, variable flip angle, SPGR, DESPOT1, relaxometry.

I. INTRODUCTION

QUANTITATIVE T_1 mapping is a Magnetic Resonance Imaging (MRI) technique that deals with the estimation of the spin-lattice relaxation time (T_1) in biological tissues [1]. The spin-lattice relaxation time has proved to be an excellent biomarker in a broad range of diseases, such as multiple sclerosis [2], epilepsy [3] and Alzheimer's disease [4]. Due to the large spectrum of potential applications, there is an increasing interest in turning quantitative T_1 mapping into a mature and robust MR modality that can be routinely used in clinical practice. In this effort, the longstanding goal is to provide accurate high resolution spatial maps of T_1 in a short time frame [5]. To that end, a plethora of T_1 mapping techniques has been proposed during the last decades (see [6] for a recent review). Among them, the Variable Flip Angle (VFA) technique, also known as the variable nutation angle method, has gained increasing popularity [5]. This is mainly due to its superior time efficiency compared to other T_1 mapping techniques, such as the traditional Inversion Recovery (IR) technique [6].

VFA T_1 mapping consists of the acquisition of a range of steady state spoiled gradient recalled echo (SPGR) MR images over a set of flip angles [7]–[9]. Since steady state MR sequences can use much shorter repetition times (TR) [10] than classical inversion/saturation recovery sequences, high resolution T_1 maps can be acquired in real-time clinical acquisition [5]. Importantly, the fact that the SPGR signal model can be easily linearized, an observation which dates back to 1977 [11], has encouraged researchers to develop fast linear T_1 estimation algorithms [5], [12], rendering the estimation time of the T_1 map negligible in comparison to the acquisition time. With such computationally inexpensive algorithms, real-time high resolution T_1 mapping can be achieved. The simplicity and efficiency of the T_1 estimation step are among the main reasons why VFA T_1 mapping has drastically grown in popularity, with DESPOT1 (Driven Equilibrium Single Pulse Observation of T_1) being the most widespread algorithm [5]. Unfortunately, the price to pay with such linear

Manuscript received March 22, 2018; revised April 26, 2018; accepted April 27, 2018. Date of publication June 4, 2018; date of current version October 29, 2018. The work of G. Ramos-Llordén was supported by the Research Foundation Flanders (FWO Belgium) through a long stay travel grant to Brigham and Women's Hospital, Harvard Medical School, Boston, USA, under Grant V440016N. The work of G. Vegas-Sánchez-Ferrero was supported in part by the Consejería de Educación, Juventud y Deporte of Comunidad de Madrid, and in part by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) for REA Grant under Agreement 291820. The work of R. San José Estépar was supported by the National Institutes of Health NHLBI under Awards R01HL116931, R01HL116473, and R21HL140422. The work of A. J. den Dekker and J. Sijbers was supported in part by the European Space Agency, BELSPO Prodex, and in part by FWO Belgium under Project G084217N. (Corresponding author: Gabriel Ramos-Llordén.)

G. Ramos-Llordén and J. Sijbers are with the imec-Vision Lab, Department of Physics, University of Antwerp, 2610 Antwerp, Belgium (e-mail: gabriel.ramos-llorden@uantwerpen.be).

G. Vegas-Sánchez-Ferrero is with the Applied Chest Imaging Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115 USA, and also with the Biomedical Image Technologies Group, CIBER-BBN, Universidad Politécnica de Madrid, 28040 Madrid, Spain.

M. Björk is with the Division of Systems and Control, Department of Information Technology, Uppsala University, 751 05 Uppsala, Sweden.

F. Vanhevel and P. M. Parizel are with the Department of Radiology, University of Antwerp, Antwerp University Hospital, 2650 Antwerp, Belgium.

R. San José Estépar is with the Applied Chest Imaging Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115 USA.

A. J. den Dekker is with the imec-Vision Lab, Department of Physics, University of Antwerp, 2610 Antwerp, Belgium, and also with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands.

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2018.2833288

estimators is a loss of accuracy, since the linearization of the SPGR model becomes inexact in the presence of noise, thereby introducing a noise-induced bias. For that reason, some researchers still adhere to more accurate non-linear least squares (NLLS) estimators, which can be shown to have optimal statistical properties for clinically achievable signal-to-noise ratio (SNR) [13]. However, NLLS estimators require the use of non-linear optimization algorithms, which are typically much slower than linear algorithms, can be difficult to implement, and are prone to convergence issues if not properly initialized.

Encouraged by such an apparent trade-off between speed and statistical optimality, we present a novel NLLS iterative optimization algorithm for VFA T_1 mapping which combines the best of both worlds: high accuracy and precision in a low computation time. The method, which we dubbed NOVIFAST (being an acronym for Non-linear Variable Flip Angle data based T_1 estimator), shares the same favorable statistical properties as the de facto standard NLLS optimization algorithm, the Levenberg-Marquard (LM) algorithm, and Variable Projection (VARPRO)-based methods, but with a much lower computation time, being comparable to that of DESPOT1.

In particular, the contributions of this work are: 1) VFA T_1 mapping is formulated as an NLLS optimization problem that can be iteratively solved as a two-by-two linear system, constituting a fixed-point algorithm (NOVIFAST). Due to its formulation, NOVIFAST turns out to be very easy to implement, and computationally highly efficient.

2) We prove that, with noiseless SPGR data, NOVIFAST is an exact method. That is, it provides the T_1 estimates with zero error in the first iteration, no matter which initial values are used. This is an important result since it explains the excellent convergence of the method at clinically realistic (high) SNR levels. 3) We provide evidence of the remarkable convergence characteristics of NOVIFAST even for low SNR under the framework of fixed-point algorithms theory.

Monte-Carlo (MC) based simulations are used to evaluate NOVIFAST's performance and compare it with that of NLLS optimization algorithms (i.e., the Gauss-Newton (GN) algorithm, the LM algorithm, and VARPRO-based methods), DESPOT1, and the Iterative Re-Weighted Linear Least Squares (IRWLLS) method proposed by Chang *et al.* [12]. Finally, NOVIFAST is validated with a set of in-vivo human brain SPGR MR images.

The structure of the remainder of the paper is as follows. In Section II, the SPGR signal model is described as well as the most popular VFA T_1 mapping algorithms. In Section III, the complete derivation of NOVIFAST is given and its convergence properties are presented. Experiments are described in Section IV and the corresponding results are presented in Section V. Next, extensions of NOVIFAST are described (Section VI), and the main conclusions are summarized at the end of the paper (Section VII).

II. BACKGROUND

A. SPGR Signal Model

The VFA SPGR method for T_1 mapping is based on the acquisition of a set of MR images where the images are

acquired with different/variable non-zero flip angles (FAs), but constant repetition time (TR) and echo time (TE). Since the images are differently T_1 -weighted, T_1 mapping can be performed by voxel-wise fitting a prescribed mathematical model to the set of MR images. A very popular model for the noiseless (magnitude) steady-state SPGR signal intensity $\{s_n\}_{n=1}^N$ is [14]:

$$s_n(K, T_1) = \frac{K(1 - e^{-\frac{TR}{T_1}}) \sin \alpha_n}{1 - e^{-\frac{TR}{T_1}} \cos(\alpha_n)}, \quad \text{for } n = 1, \dots, N, \quad (1)$$

where $\{\alpha_n\}_{n=1}^N$ denotes the FAs, with $\alpha_n \neq 180^\circ$ for every n , and N is the total number of FAs used (or images collected). The unknown parameter $K > 0$ includes multiplicative factors such as the longitudinal component of the net nuclear magnetization vector, and the attenuation due to T_2^* relaxation for a fixed TE [14]. In reality, the signal $\{s_n\}_{n=1}^N$ is always corrupted by noise. Therefore, the measured noisy SPGR samples y_n , for $n = 1, \dots, N$, can be considered as realizations of random variables whose distribution depends parametrically on K and T_1 . Given a set of samples $\{y_n\}_{n=1}^N$, inferring the values of the unknown underlying T_1 and K (assuming TR and α_n to be known) is an estimation problem. A wide variety of estimators may be defined. In this work, we focus on the ordinary, i.e., unweighed, NLLS estimator, as well as heuristic linear least squares (LLS) estimators, since these types of estimators are the most popular ones in VFA T_1 mapping. In the succeeding section, we review the most common NLLS optimization algorithms to solve the NLLS estimation problem, as well as the heuristic linearized variants. We discuss their pros and cons, after which we present our novel NLLS method.

B. NLLS Estimation Problem

For a given data set $\{y_n\}_{n=1}^N$, the ordinary NLLS estimator is defined as

$$\{\widehat{K}, \widehat{T}_1\} = \underset{K, T_1}{\operatorname{argmin}} \sum_{n=1}^N (y_n - s_n(K, T_1))^2. \quad (2)$$

Such an optimization problem cannot be solved analytically and hence one has to resort to numerical algorithms. A straightforward approach to find \widehat{K} and \widehat{T}_1 would be to discretize the search space of both K and T_1 , and then reformulate the problem as a discrete optimization problem. In other words, the cost function of Eq. (2) is evaluated in a sufficiently dense grid, given by a user-defined resolution, and then, the minimum value is retained. This technique is called brute force grid search. It should be noted that brute force grid search may be computationally inefficient since the range of values of the parameter K may be relatively large, thereby resulting in a prohibitive number of cost function evaluations. For model-based MRI relaxometry, such as VFA T_1 mapping, it is much more common to tackle the original optimization problem in its native continuous domain. By far, the most common line of action is to employ optimization algorithms that attempt to find the local minima with gradient-based

techniques. Perhaps, the main representative of this class of algorithms is the full Newton algorithm [15]. A downside of this algorithm is that it is considerably time consuming due to the need of line searches and the computation of the Hessian matrix. To solve NLLS estimation problems, it is often better to exploit the quadratic structure of the cost function, as is done by the celebrated Gauss-Newton (GN) and Levenberg-Marquardt (LM) algorithms [16], [17], which are, by far, the most popular NLLS optimization algorithms. Both algorithms are briefly reviewed below.

1) Gauss-Newton (GN): Let us define $\mathbf{r}(K, T_1) = (r_1(K, T_1), \dots, r_N(K, T_1))^T$, where $r_n(K, T_1) = y_n - s_n(K, T_1)$ are the so-called residuals of the NLLS problem. The GN algorithm can be derived by linearly approximating $\mathbf{r}(K, T_1)$ around a given estimate (K^k, T_1^k) [15],

$$\mathbf{r}(K, T_1) \approx \mathbf{r}(K^k, T_1^k) + \mathbf{J}_r(K^k, T_1^k)\Delta, \quad (3)$$

where $\mathbf{J}_r(K^k, T_1^k)$ is the Jacobian matrix of $\mathbf{r}(\cdot, \cdot)$, evaluated at (K^k, T_1^k) . The vector $\Delta = (K, T_1)^T - (K^k, T_1^k)^T$ is called the step vector. After substituting Eq.(3) into Eq.(2), and equating the gradient of the cost function with respect to Δ to zero, we arrive at the normal equations:

$$\tilde{\mathbf{J}}_r(K^k, T_1^k)\Delta = -\mathbf{J}_r(K^k, T_1^k)^T \mathbf{r}(K^k, T_1^k), \quad (4)$$

with $\tilde{\mathbf{J}}_r(K^k, T_1^k) = \mathbf{J}_r(K^k, T_1^k)^T \mathbf{J}_r(K^k, T_1^k)$. Provided $\tilde{\mathbf{J}}_r(K^k, T_1^k)$ is non-singular, a condition implicitly assumed in the GN algorithm, the step vector can be obtained analytically for each iteration, which, for the dimensions of the problem at hand, can be done with a negligible computational effort. The GN algorithm can also be interpreted as an inexact full Newton algorithm where the Hessian matrix is approximated to avoid the computational burden [15]. The GN algorithm works well close to the minimum, however, it may not converge at all if it is not initialized properly [17], [18]. Rigorous conditions to prove local convergence¹ are given in [19].

2) Levenberg-Marquardt (LM): The LM algorithm [16], [17] can be seen as a generalization of the GN algorithm. Like GN, the LM algorithm can be derived from the normal equations (Eq.(4)), but this time $\tilde{\mathbf{J}}_r(K^k, T_1^k)$ is modified to

$$\tilde{\mathbf{J}}_r(K^k, T_1^k) + \lambda \text{diag}(\tilde{\mathbf{J}}_r(K^k, T_1^k)), \quad (5)$$

where λ is a user-controlled parameter which may be updated at each iteration, and $\text{diag}(\tilde{\mathbf{J}}_r(K^k, T_1^k))$ is a diagonal matrix whose entries are the elements on the diagonal of $\tilde{\mathbf{J}}_r(K^k, T_1^k)$. When $\lambda \rightarrow 0$, the calculated step vector Δ approaches the GN step obtained by solving Eq.(4). When $\lambda \rightarrow \infty$, Δ approaches a gradient-descent step where each component is weighted according to $\text{diag}(\tilde{\mathbf{J}}_r(K^k, T_1^k))$. In a typical implementation, λ is set to a non-zero value at initial iterations. For finite λ values, it is not guaranteed that the

¹In optimization theory's parlance [19], [20], local convergence means that convergence is assured if the initial approximation is close enough to a stationary point. In contrast, global convergence implies convergence for any arbitrary initialization. Note that this terminology does not deal with the character of stationary points, e.g., local/global minima, but just refers to convergence to stationary points.

iterates $(K^{k+1}, T_1^{k+1})^T = (K^k, T_1^k)^T + \Delta$ decrease the cost function, therefore, this condition is always checked in the implementation of the LM method. If the descent-property is not fulfilled, λ is increased, and the normal equations are again solved till the step Δ assures a decrease in the cost function. If the descent property is fulfilled, iterates are updated and the value λ is reduced. Therefore, the LM algorithm behaves as a modified gradient-descent method at early iterations, but mimics GN as it gets closer to the minimum (λ is decreased towards zero). LM is slightly more computationally demanding than GN, but it converges for initializations that are far away from the solution, where GN often fails [17]. In this sense, LM shows global convergence properties and is therefore the preferred method of choice in common NLLS problems.

C. VARIABLE PROJECTION METHOD (VARPRO)

Let us write Eq. (1) as $s_n(K, T_1) = K g_n(T_1)$ with

$$g_n(T_1) = \frac{(1 - e^{-\frac{TR}{T_1}}) \sin \alpha_n}{1 - e^{-\frac{TR}{T_1}} \cos(\alpha_n)}, \quad \text{for } n = 1, \dots, N, \quad (6)$$

and let us focus on the optimization problem (Eq. (2)), which now takes the expression

$$\{\hat{K}, \hat{T}_1\} = \underset{K, T_1}{\text{argmin}} \sum_{n=1}^N (y_n - K g_n(T_1))^2. \quad (7)$$

By virtue of the Hilbert projection theorem [21], independently of the value of T_1 , one can prove that \hat{K} is always given by

$$\hat{K}(T_1) = \frac{\sum_{n=1}^N y_n g_n(T_1)}{\sum_{n=1}^N g_n^2(T_1)}. \quad (8)$$

By inserting Eq.(8) into Eq.(7), Eq.(7) reduces to a one-dimensional optimization problem:

$$\hat{T}_1 = \underset{T_1}{\text{argmin}} \sum_{n=1}^N (y_n - \hat{K}(T_1) g_n(T_1))^2. \quad (9)$$

It can be demonstrated that solving Eq. (9) and then calculating $\hat{K}(\hat{T}_1)$ provides the same result as that obtained by solving Eq.(7) directly [22], [23]. This approach is known as the VARIABLE PROJECTION method (VARPRO) [22]. The VARPRO method has been used in several MRI relaxometry applications, not only in VFA T_1 mapping [24]–[27], and is particularly appealing since it reduces the optimization problem to a simpler one-dimensional problem. While derivative-based algorithms that find the stationary points can be used [28], it is more common to use derivative-free algorithms that search for the global minimum directly, such as brute force grid search [24], [25], [27] or more efficient iterative methods such as the golden-section search, which require a substantially fewer number of cost function evaluations to get the same result [26].

D. Heuristic Linearized Variants

Both GN and LM make use of the quadratic expression of the NLLS cost function but, as general-purpose NLLS algorithms, they do not consider the particular structure of the

SPGR signal model described by Eq. (1). Indeed, by dividing both sides of Eq. (1) by $\sin \alpha_n$, and then rearranging the equation, it can be shown that Eq. (1) can be written as [11]:

$$\frac{s_n}{\sin \alpha_n} = c_2 \frac{s_n}{\tan \alpha_n} + c_1 \quad \text{for } n = 1, \dots, N, \quad (10)$$

with

$$c_1 = K(1 - e^{-\text{TR}/T_1}), \quad c_2 = e^{-\text{TR}/T_1}, \quad (11)$$

and where it has been assumed that $\alpha_n \neq 90^\circ$ and $\alpha_n \neq 270^\circ$ in order for $\tan \alpha_n$ to be well-defined. The linear relation of Eq. (10) (w.r.t. $c_1 > 0$ and $c_2 > 0$) is substantially exploited in several LLS-based algorithms, which we briefly review hereunder.

1) DESPOT1: The DESPOT1 algorithm, proposed by Deoni *et al.* [5], aims to estimate c_1 and c_2 in a linear regression framework. Indeed, in the absence of noise, $\frac{s_n}{\tan \alpha_n}$ may be considered as the regressor variables and $\frac{s_n}{\sin \alpha_n}$ as the regressand counterparts. In the noisy case, $\{s_n\}_{n=1}^N$ is not observable and hence DESPOT1 replaces $\{s_n\}_{n=1}^N$ by $\{y_n\}_{n=1}^N$, after which c_1 and c_2 are estimated by linear least squares (LLS) regression. DESPOT1 presents the lowest computational cost since it is an analytical estimator (i.e., it gives the estimates in a closed-form expression). However, as it strongly relies on the linear relation described by Eq. (10), and this relation no longer holds when $\{s_n\}_{n=1}^N$ is replaced by $\{y_n\}_{n=1}^N$, a bias is introduced, which becomes more pronounced at low SNR [29].

2) Iterative Re-Weighted Linear Least Squares (IRWLLS): To increase the accuracy of DESPOT1, Chang *et al.* [12] suggested to include a weighting function for each of the samples $n = 1, \dots, N$. The linear regression approach is then transformed into a weighted linear regression approach, where the weights are derived based on uncertainty propagation theory. Since the optimal weights depend on the parameters to be estimated, Chang *et al.* proposed an iterative approach where the parameters c_1 and c_2 are estimated using a weighted linear least squares estimator of which the weights are updated iteratively. This iterative re-weighted linear least squares (IRWLLS) method effectively outperforms DESPOT1 in terms of accuracy. However, there is no guarantee that the final estimates share the same optimal properties as a pure NLLS approach. Furthermore, convergence of the iterative procedure is not guaranteed. Consequently, the algorithm may in fact diverge.

III. METHOD

NOVIFAST is derived directly from the NLLS problem (Eq. (2)), and should thus be classified as an NLLS optimization algorithm. Interestingly, in its derivation, the particular structure of the SPGR model is exploited, which resembles the approach that was used in the linearized, but sub-optimal DESPOT1 and IRWLLS algorithms. This way, NOVIFAST combines the best of both worlds: the accuracy/precision of NLLS estimators and the computational speed of heuristic linear algorithms. Since NOVIFAST is not a general-purpose NLLS-based algorithm, we present its complete derivation below. In Section III-B, the main features of the algorithm are described and its pseudo-code is presented. Convergence

properties are briefly covered in Section III-B and thoroughly studied in Section I-B of the supplementary file which accompanies this paper.

A. Derivation of NOVIFAST: Exploiting the Structure of the SPGR Model

Let us denote by $\mathcal{L}(K, T_1)$ the cost function that is sought to be minimized in Eq. (2). We then formalize the change of variables given in Eq. (11) with the definition of a two-dimensional vector-valued mapping, $\Phi_{\text{TR}} : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+ \times (0, 1)$, given by

$$\Phi_{\text{TR}}(K, T_1) = \left(K(1 - e^{-\text{TR}/T_1}), e^{-\text{TR}/T_1} \right). \quad (12)$$

We call the first component of $\Phi_{\text{TR}}(K, T_1)$ as c_1 , whereas the second component is denoted as c_2 . The mapping $\Phi_{\text{TR}}(K, T_1)$ is a bijection between $\mathbb{R}^+ \times \mathbb{R}^+$ and $C = \mathbb{R}^+ \times (0, 1)$. In other words, given any pair of points c_1 and c_2 that lie in C , there is a unique combination of K and T_1 such that $(c_1, c_2) = \Phi_{\text{TR}}(K, T_1)$. The inverse mapping, $\Phi_{\text{TR}}^{-1}(c_1, c_2) : C \mapsto \mathbb{R}^+ \times \mathbb{R}^+$, has the expression of

$$\Phi_{\text{TR}}^{-1}(c_1, c_2) = \left(\frac{c_1}{1 - c_2}, -\frac{\text{TR}}{\log c_2} \right). \quad (13)$$

Let us then reparameterize $\mathcal{L}(K, T_1)$ in terms of the new variables c_1 and c_2 , that is,

$$\mathcal{L}(K, T_1)|_{(K, T_1) = \Phi_{\text{TR}}^{-1}(c_1, c_2)}, \quad (14)$$

where now $\mathcal{S}(\mathbf{c}) \triangleq \mathcal{L}(K, T_1)|_{(K, T_1) = \Phi_{\text{TR}}^{-1}(c_1, c_2)}$, with $\mathbf{c} = (c_1, c_2)^T$, is the new cost function, that is,

$$\mathcal{S}(\mathbf{c}) = \sum_{n=1}^N (y_n - s_n(\mathbf{c}))^2, \quad (15)$$

with

$$s_n(\mathbf{c}) = \frac{c_1 \sin \alpha_n}{1 - c_2 \cos(\alpha_n)}, \quad \text{for } n = 1, \dots, N. \quad (16)$$

With this change of variables, we can derive the NLLS estimate of \mathbf{c} by solving

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\text{argmin}} \mathcal{S}(\mathbf{c}), \quad (17)$$

and then obtain \hat{K} and \hat{T}_1 with $\Phi_{\text{TR}}^{-1}(\hat{c}_1, \hat{c}_2)$. Indeed, the equivalence of the global minimum of Eq. (2) and Eq. (17) can be demonstrated as follows: by definition of global minimum, we have that $\mathcal{S}(\hat{\mathbf{c}}) \leq \mathcal{S}(\mathbf{c})$ for all $\mathbf{c} \in C$. Let us call $(K^*, T_1^*) = \Phi_{\text{TR}}^{-1}(\hat{c}_1, \hat{c}_2)$, and let us pick any point (K, T_1) in $\mathbb{R}^+ \times \mathbb{R}^+$. Since we know that $\Phi_{\text{TR}}(K, T_1)$ is a bijection, we have that $(K, T_1) = \Phi_{\text{TR}}^{-1}(c_1, c_2)$ for some $\mathbf{c} = (c_1, c_2)^T$. Therefore, we have that

$$\mathcal{L}(K, T_1) = \mathcal{S}(\mathbf{c}) \geq \mathcal{S}(\hat{\mathbf{c}}) = \mathcal{L}(K^*, T_1^*), \quad (18)$$

which automatically implies that (K^*, T_1^*) is the global minimum, that is $K^* = \hat{K}$ and $T_1^* = \hat{T}_1$ in Eq. (2). Now that we have demonstrated that minimizing the cost function with respect to c_1 and c_2 is equivalent to minimizing it with respect to K and T_1 , let us put our efforts into the new optimization

problem of Eq.(17). The first-order conditions necessary to locate the stationary points of \mathcal{S} are:

$$\frac{\partial \mathcal{S}}{\partial c_1} = 2 \sum_{n=1}^N (y_n - s_n(\mathbf{c})) \frac{\sin \alpha_n}{1 - c_2 \cos \alpha_n} = 0 \quad (19)$$

$$\frac{\partial \mathcal{S}}{\partial c_2} = -2 \sum_{n=1}^N (y_n - s_n(\mathbf{c})) \frac{s_n(\mathbf{c}) \cos \alpha_n}{1 - c_2 \cos \alpha_n} = 0. \quad (20)$$

The NLLS estimate, by construction, must fulfill this system of non-linear equations. In principle, the complexity to solve Eqs. (19–20) makes this approach unattractive, and probably that is the main reason that existing NLLS optimization algorithms for VFA T_1 mapping do not attempt to follow this line of thinking. However, the particular structure of $s_n(\mathbf{c})$, a rational function of sines and cosines, in combination with certain algebraic rules, as those exposed by Dimitrov and Kamenski for rational functions in chemical kinetics [30], yields a surprisingly simple set of non-linear equations, which can be solved iteratively. Indeed, if we substitute $s_n(\mathbf{c})$ in Eq.(19) by its rational expression (Eq.(16)), and use the term $1 - c_2 \cos \alpha_n$ as common denominator, which is never zero, we get

$$\sum_{n=1}^N \frac{y_n(1 - c_2 \cos \alpha_n) - c_1 \sin \alpha_n}{(1 - c_2 \cos \alpha_n)^2} \sin \alpha_n = 0. \quad (21)$$

By rearranging terms, Eq.(19) can be written as

$$\begin{aligned} c_1 \sum_{n=1}^N \frac{\sin^2 \alpha_n}{(1 - c_2 \cos \alpha_n)^2} + c_2 \sum_{n=1}^N \frac{y_n \sin \alpha_n \cos \alpha_n}{(1 - c_2 \cos \alpha_n)^2} \\ = \sum_{n=1}^N \frac{y_n \sin \alpha_n}{(1 - c_2 \cos \alpha_n)^2}. \end{aligned} \quad (22)$$

Similarly, Eq.(20) can be expressed as

$$\begin{aligned} c_1 \sum_{n=1}^N \frac{s_n(\mathbf{c}) \sin \alpha_n \cos \alpha_n}{(1 - c_2 \cos \alpha_n)^2} + c_2 \sum_{n=1}^N \frac{y_n s_n(\mathbf{c}) \cos^2 \alpha_n}{(1 - c_2 \cos \alpha_n)^2} \\ = \sum_{n=1}^N \frac{y_n s_n(\mathbf{c}) \cos \alpha_n}{(1 - c_2 \cos \alpha_n)^2}. \end{aligned} \quad (23)$$

If we define the $(N \times 1)$ vectors \mathbf{a} , $\tilde{\mathbf{a}}$, \mathbf{b} , and \mathbf{z} as

$$\mathbf{a} = \left(\frac{y_1 \cos \alpha_1}{1 - c_2 \cos \alpha_1}, \frac{y_2 \cos \alpha_2}{1 - c_2 \cos \alpha_2}, \dots, \frac{y_N \cos \alpha_N}{1 - c_2 \cos \alpha_N} \right)^T, \quad (24)$$

$$\tilde{\mathbf{a}} = \left(\frac{s_1(\mathbf{c}) \cos \alpha_1}{1 - c_2 \cos \alpha_1}, \frac{s_2(\mathbf{c}) \cos \alpha_2}{1 - c_2 \cos \alpha_2}, \dots, \frac{s_N(\mathbf{c}) \cos \alpha_N}{1 - c_2 \cos \alpha_N} \right)^T, \quad (25)$$

$$\mathbf{b} = \left(\frac{\sin \alpha_1}{1 - c_2 \cos \alpha_1}, \frac{\sin \alpha_2}{1 - c_2 \cos \alpha_2}, \dots, \frac{\sin \alpha_N}{1 - c_2 \cos \alpha_N} \right)^T, \quad (26)$$

$$\mathbf{z} = \left(\frac{y_1}{1 - c_2 \cos \alpha_1}, \frac{y_2}{1 - c_2 \cos \alpha_2}, \dots, \frac{y_N}{1 - c_2 \cos \alpha_N} \right)^T, \quad (27)$$

the previous system of non-linear equations is concisely expressed as

$$\underbrace{\begin{pmatrix} \langle \mathbf{b}, \mathbf{b} \rangle & \langle \mathbf{b}, \mathbf{a} \rangle \\ \langle \mathbf{b}, \tilde{\mathbf{a}} \rangle & \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle \end{pmatrix}}_{\mathbf{A}(\mathbf{c})} \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\mathbf{c}} = \underbrace{\begin{pmatrix} \langle \mathbf{z}, \mathbf{b} \rangle \\ \langle \mathbf{z}, \tilde{\mathbf{a}} \rangle \end{pmatrix}}_{\mathbf{v}(\mathbf{c})}, \quad (28)$$

where $\langle \cdot, \cdot \rangle$ is the usual inner-product for vectors in Euclidean spaces. Note that even though we have omitted the dependence on \mathbf{c} in the definition of Eqs. (24–27), the 2×2 matrix $\mathbf{A}(\mathbf{c})$ depends on the linear coefficients, and the 2×2 vector $\mathbf{v}(\mathbf{c})$ does so as well. Eq.(28) is the basis of NOVIFAST. Although the equation is non-linear in \mathbf{c} , it resembles a purely linear system. If the variation of $\mathbf{A}(\mathbf{c})$ and $\mathbf{v}(\mathbf{c})$ w.r.t. \mathbf{c} were negligible, a simple 2×2 inversion technique would suffice to find the solution of Eq. (28). Since this is not the case, a natural approach is to solve it iteratively, thereby still exploiting the semi-linear structure. We propose an iterative technique where, given that $\mathbf{A}(\mathbf{c})$ and $\mathbf{v}(\mathbf{c})$ are known, \mathbf{c} is solved linearly, after which $\mathbf{A}(\mathbf{c})$ and $\mathbf{v}(\mathbf{c})$ are updated with the new guess. The repetition of those two steps constitutes our algorithm, which we dubbed NOVIFAST.

B. NOVIFAST: Algorithm Definition

Equations of the form $\mathbf{A}(\mathbf{c})\mathbf{c} = \mathbf{v}(\mathbf{c})$ often appear in discretization schemes for time differential equations [31]. A common method for its solution is what is called a semi-implicit technique [31]: an iterate \mathbf{c}^k derived from the k -th iteration is used to evaluate $\mathbf{A}(\mathbf{c})$ and $\mathbf{v}(\mathbf{c})$, and the following linear equation w.r.t. \mathbf{c}^{k+1} is solved:

$$\mathbf{A}(\mathbf{c}^k)\mathbf{c}^{k+1} = \mathbf{v}(\mathbf{c}^k). \quad (29)$$

Cramer's rule [32] allows us to obtain \mathbf{c}^{k+1} explicitly:

$$c_1^{k+1} = \frac{\begin{vmatrix} \langle \mathbf{z}, \mathbf{b} \rangle & \langle \mathbf{b}, \mathbf{a} \rangle \\ \langle \mathbf{z}, \tilde{\mathbf{a}} \rangle & \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle \end{vmatrix}}{\begin{vmatrix} \langle \mathbf{b}, \mathbf{b} \rangle & \langle \mathbf{b}, \mathbf{a} \rangle \\ \langle \mathbf{b}, \tilde{\mathbf{a}} \rangle & \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle \end{vmatrix}}, \quad c_2^{k+1} = \frac{\begin{vmatrix} \langle \mathbf{b}, \mathbf{b} \rangle & \langle \mathbf{z}, \mathbf{b} \rangle \\ \langle \mathbf{b}, \tilde{\mathbf{a}} \rangle & \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle \end{vmatrix}}{\begin{vmatrix} \langle \mathbf{b}, \mathbf{b} \rangle & \langle \mathbf{b}, \mathbf{a} \rangle \\ \langle \mathbf{b}, \tilde{\mathbf{a}} \rangle & \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle \end{vmatrix}}, \quad (30)$$

where $|\cdot|$ denotes the determinant of a matrix, and we have made the assumption that the denominator of both expressions is non-zero for every $\mathbf{c} \in C$. In Section II-A of the supplementary file we elaborate on this assumption. We are then ready to define the NOVIFAST algorithm through the following pseudo-code (Algorithm 1):

Below, we pinpoint some of the most interesting properties of NOVIFAST.

- **Convergence points are roots of Eqs. (19–20)**

If NOVIFAST converges to some \mathbf{c}^* , it is necessarily a root of Eqs. (19–20). To see this, one only needs to take limits when $k \rightarrow \infty$ on both sides of Eq.(29), and use the fact that both $\mathbf{A}(\mathbf{c})$ and $\mathbf{v}(\mathbf{c})$ are continuous functions of \mathbf{c} . Hence, limit point \mathbf{c}^* meets $\mathbf{A}(\mathbf{c}^*)\mathbf{c}^* = \mathbf{v}(\mathbf{c}^*)$, and it is a solution of Eq.(28), or equivalently a root of Eqs. (19–20). Therefore, the algorithm is well defined.

Algorithm 1 Pseudo-Code of NOVIFAST

```

1: Given parameters: TR and flip angles  $\{\alpha_n\}_{n=1}^N$ 
2: Initial values:  $K_{\text{ini}}$  and  $T_{1\text{ini}}$ 
3:  $c_1^0 \leftarrow K_{\text{ini}}(1 - \exp(-\text{TR}/T_{1\text{ini}}))$ 
4:  $c_2^0 \leftarrow \exp(-\text{TR}/T_{1\text{ini}})$ 
5:  $k \leftarrow 0$ 
6:  $\mathbf{c}^k \leftarrow (c_1^k, c_2^k)^T$ 
7: while convergence criterium is not met do
8:   Solve Eq. (29) with the solutions given in Eq. (30)
9:    $k \leftarrow k + 1$ 
10: end while
11: return  $\widehat{K} = c_1^k/(1 - c_2^k)$  and  $\widehat{T}_1 = -\text{TR}/\log c_2^k$ 

```

- **Good convergence properties**

With NOVIFAST, the only required stopping criterion is to check that the norm of the difference between consecutive iterates is below a fixed tolerance. There is no need to assure the descent property. This is an advantage in terms of computation time when compared to descent-based algorithms such as GM or LM. Indeed, we have corroborated with MC simulations that for realistic SNR values and typical ground-truth T_1 values in white and gray matter, the sequence \mathbf{c}^k lies in a closed subset of C with probabilities near to 100% (being 95% the worst case for extremely low SNR). This experiment is described in Section I-B.1 and II-B.2 of the supplementary file. A closed subset of C , together with the l_2 norm, define a complete metric space, which, by definition, implies that if $\lim_{k \rightarrow \infty} \|\mathbf{c}^{k+1} - \mathbf{c}^k\|_2 = 0$, then $\lim_{k \rightarrow \infty} \mathbf{c}^k = \mathbf{c}^*$, which, as we have seen, is a root of Eqs. (19–20). The conditions to meet $\lim_{k \rightarrow \infty} \|\mathbf{c}^{k+1} - \mathbf{c}^k\|_2 = 0$ are thoroughly studied under the framework of fixed-point algorithms theory. Convergence conditions are empirically checked in Section II-B of the supplementary file as well, where it is shown that they hold with probabilities near to 100% for realistic clinically achievable SNRs.

- **Low cost per iteration**

The cost per iteration of NOVIFAST is low, since it amounts to calculate two quotients of determinants of just 2×2 matrices. The cost per iteration is similar to that of IRWLLS, and also to the total cost of DESPOT1. Importantly, NOVIFAST aims to get the NLLS estimates (see first bullet point), whereas both IRWLLS and DESPOT1 are modified heuristic algorithms.

- **Robustness and high convergence rate**

NOVIFAST is rather insensitive to initial values K_{ini} and $T_{1\text{ini}}$, and convergence is usually reached within 2-4 iterations with the same tolerance criterion as LM or GN algorithms. This makes NOVIFAST an ideal algorithm to be used in practice. Like LM and VARPRO, NOVIFAST shows global convergence properties, but it converges considerably faster. This is in agreement with results provided in [33], and also with those of our previous work [34]. In Section I-B of the supplementary file, we provide evidence for these claims, and we

experimentally check the conditions with an MC analysis in Section II-B of the same document.

- **Simplicity**

NOVIFAST does not need any tuning parameter in contrast to LM, and its implementation is straightforward.

C. NOVIFAST as an Exact, Analytical Method

It is not difficult to demonstrate that in the noise-free case, that is, when $y_n = s_n$, the NLLS estimates coincide with the ground-truth values, and that they can be obtained analytically. Indeed, in the noiseless case, the ground-truth value $\mathbf{c}_{\text{GT}} = (c_{1\text{GT}}, c_{2\text{GT}})^T = (K_{\text{GT}}(1 - \exp(-\text{TR}/T_{1\text{GT}})), \exp(-\text{TR}/T_{1\text{GT}}))^T$ is the global minimum of \mathcal{S} since we have $y_n = s_n \triangleq s_n(\mathbf{c}_{\text{GT}})$, thus $\mathcal{S}(\mathbf{c}_{\text{GT}}) = 0$. Due to the linear relationship of Eq. (10), \mathbf{c}_{GT} can be linearly expressed in terms of just two samples from $\{s_n\}_{n=1}^N$. Since the NLLS estimate is equal to \mathbf{c}_{GT} , it can also be expressed this way, and thus \widehat{K} and \widehat{T}_1 can be derived as a closed-form expression. Therefore, in the absence of noise, DESPOT1 becomes an exact method, but it is not clear what happens with iterative algorithms, since GN, LM, and VARPRO, by design, do not exploit the particular structure of the SPGR signal. Remarkably, in the absence of noise, and due to the semi-linear structure of Eq. (28), NOVIFAST becomes an exact and analytical method as well, in the sense that it provides the ground-truth values with just one iteration. Observe that this is not true for GN, LM, and the VARPRO method. Indeed, for noiseless data, the following relations can be shown to be true:

$$\langle \mathbf{z}, \mathbf{b} \rangle = c_{1\text{GT}} \langle \mathbf{b}, \mathbf{b} \rangle + c_{2\text{GT}} \langle \mathbf{b}, \mathbf{a} \rangle \quad (31)$$

$$\langle \mathbf{z}, \tilde{\mathbf{a}} \rangle = c_{1\text{GT}} \langle \mathbf{b}, \tilde{\mathbf{a}} \rangle + c_{2\text{GT}} \langle \mathbf{a}, \tilde{\mathbf{a}} \rangle. \quad (32)$$

If we substitute $\langle \mathbf{z}, \mathbf{b} \rangle$ and $\langle \mathbf{z}, \tilde{\mathbf{a}} \rangle$ in the numerators of Eq. (30) by the expressions of Eqs. (31–32), and we make use of determinant properties, it is possible to prove that $c_1^{k+1} = c_{1\text{GT}}$ and $c_2^{k+1} = c_{2\text{GT}}$ for $k = 0, 1, \dots$. The interested reader may find the mathematical proof of this result in the supplementary file which accompanies this paper (Section III). In short, NOVIFAST shares with DESPOT1 that just one iteration is needed to provide the ground-truth parameters. This correspondence with DESPOT1 may be useful to better understand the convergence behavior of NOVIFAST in realistic conditions, when noise is present. Of course, we remark that in such real conditions, DESPOT1 is not optimal whereas NOVIFAST is a truly NLLS-based algorithm. For instance, as is shown in Section II-B of the supplementary file, NOVIFAST's convergence conditions are very likely to hold at clinically achievable SNR values, and the number of iterations required to reach convergence decreases with the SNR. This is not surprising since we already know that in the asymptotic case of an infinitely high SNR NOVIFAST *must* be an exact and analytical algorithm. In contrast, there is no theoretical reason why the GN algorithm, the LM algorithm or VARPRO method would require fewer iterations to converge when the SNR increases.

IV. EXPERIMENTS

We validated NOVIFAST with Monte Carlo (MC) simulation and in vivo human brain experiments. For the simulation experiments, we generated noisy SPGR magnitude data y_n , $n = 1, \dots, N$, as realizations of statistically independent random variables that follow a Rician distribution, where the signal parameter is given by $s_n(K_{GT}, T_{1GT})$, with K_{GT} and T_{1GT} the ground-truth K and T_1 values, respectively, and the noise standard deviation is denoted by σ . We employed the Rician distribution model since it is the most common statistical model for magnitude MR images, in both single and multi-coil acquisitions [13], [35], [36]. The noise standard deviation σ is parameterized as $\sigma = K_{GT}/\text{SNR}_{90^\circ}$, with SNR_{90° the maximum SNR per pixel for an image acquired with an FA of 90° and $\text{TR} > 6 \times T_1$ [37].

Such a definition of the SNR is commonly used in other works on VFA SPGR T_1 mapping [12], [37]. However, for better interpretation, we will additionally utilize the *conventional* SNR that is used in quantitative MRI:

$$\text{SNR} = \frac{1}{N} \sum_{n=1}^N \frac{s_n}{\sigma}. \quad (33)$$

We illustrate the performance of NOVIFAST with common sequence settings that are reported in the literature. We used a repetition time TR of 5 ms [37], and two sets of FAs, namely $\{2^\circ, 9^\circ, 19^\circ\}$ [37] and $\{2^\circ, 3^\circ, 4^\circ, 5^\circ, 7^\circ, 9^\circ, 11^\circ, 14^\circ, 17^\circ, 22^\circ\}$ [37], [38]. We denote the FA sets as $\mathcal{A}_{N=3}$ and $\mathcal{A}_{N=10}$, respectively.

A. MC Simulation Experiment With Synthetically Generated SPGR MR Signals

First, experiments with simulated SPGR MR signals were set up to compare the performance of NOVIFAST in terms of convergence and computational efficiency to that of GN, LM, VARPRO, DESPOT1, and IRWLLS. To investigate the convergence properties of GN, LM, and NOVIFAST, ideally we could check whether the final iterates provided by all of them truly correspond to the global minimum of the NLLS cost function. However, this approach is infeasible since we would need to do so for all possible noisy SPGR signals. Instead, we followed a more modest approach which is common practice when a new VFA T_1 mapping algorithm is introduced. Since convergence failure may introduce a bias in the final estimates, and as this bias is not intrinsic to the NLLS estimator but rather to the optimization algorithm, we could detect convergence pathologies by analyzing those estimates. By using realistic MC simulations, we assessed the statistical performance of GN, LM, and NOVIFAST, as if they were considered as different estimators.

All algorithms were run in Matlab, using both GN and LM implemented within the Matlab function `lsqnonlin`. Although both GN and LM have been described in the domain of K and T_1 , we implement them in the c domain, with an exact, analytical Jacobian. This provides a fair comparison with DESPOT1, IRWLLS, and NOVIFAST, which are genuinely conceived in the c domain. We set the initial

value of λ in the LM formulation to the recommended value in `lsqnonlin`. IRWLLS, GN, LM, and NOVIFAST were halted with the same tolerance criterion: the relative l_2 norm between consecutive iterates was set to 10^{-6} . The maximum number of iterations was set to 1000. We implemented two versions of the VARPRO method. In the first version, the global minimum \hat{T}_1 was obtained using a brute force grid-search approach, as implemented in [24] and [25]. We denote this approach as VARPRO-BRUTE. We used the acronym VARPRO-GSS for the second implementation, where \hat{T}_1 is attained with the golden-section search [26]. Both implementations used a grid of T_1 values between $T_{1\text{min-grid}} = 250$ ms and $T_{1\text{max-grid}} = 3000$ ms with a grid resolution (tol) of $\text{tol} = 1$ ms, as suggested in [26]. With this settings, VARPRO-BRUTE requires 2751 cost function evaluations.

The experiment setup for the MC simulation was the following.

- 1) We chose either $\mathcal{A}_{N=3}$ or $\mathcal{A}_{N=10}$ as FA set.
- 2) We fixed $K_{GT} = 1$ and generated ground-truth SPGR signals, $\{s_n\}_{n=1}^N$, with ten different values of T_{1GT} logarithmically spaced between 500 ms and 2500 ms.
- 3) For each of the ten ground-truth signals $\{s_n\}_{n=1}^N$, we generated $N_{MC} = 10^5$ realizations of Rician distributed noisy signals $\{y_n\}_{n=1}^N$. For each noisy dataset, SNR_{90° was selected among the following list of values: $\{80, 150, 250, 300\}$. Those values are within the range of SNR_{90° commonly used in similar MC-based analyses in the literature [12], [37].
- 4) The five algorithms were run with the same input datasets $\{y_n\}_{n=1}^N$. In order to check the influence of initialization, we chose a constant initialization for all range of T_{1GT} . We studied the robustness of NOVIFAST against the rest of the methods with respect to poor initialization. To that end, we selected two different configurations: 1) $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 1000$ ms, and 2) $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 500$ ms.

The computational time, in milliseconds, amounts the total CPU time of the algorithms' execution for given input data $\{y_n\}_{n=1}^N$. All algorithms were run on an Intel® Xeon® CPU E5-2680 v2 with 25 MB of cache clocked at 2.8 GHz.

B. MC Simulation With Synthetic 3D T_1 Phantom

Prior to validating NOVIFAST with in vivo SPGR MR images, we conducted an MC-based simulation with a synthetically generated set of SPGR MR images. We used realistic SNR values achievable in practice, realistic ground-truth T_{1GT} and K_{GT} maps (with a wider range of values than those in the previous experiment, including T_1 beyond that of white/gray matter), and we mimicked clinically realistic noise conditions, e.g., spatially variant noise maps. Specifically, several sets of noise-free 3D SPGR MR images (following Eq.(1) with $\text{TR} = 5$ ms and with FAs given by the $\mathcal{A}_{N=10}$ FA set) were created based on ground-truth T_{1GT} and K_{GT} maps. Those maps were estimated from a simulated IR gradient recalled echo sequence, with similar settings as those given in [39]. The size of both 3D maps was $111 \times 93 \times 71$ with an isotropic voxel size of 1.5 mm. Next, noisy Rician distributed images

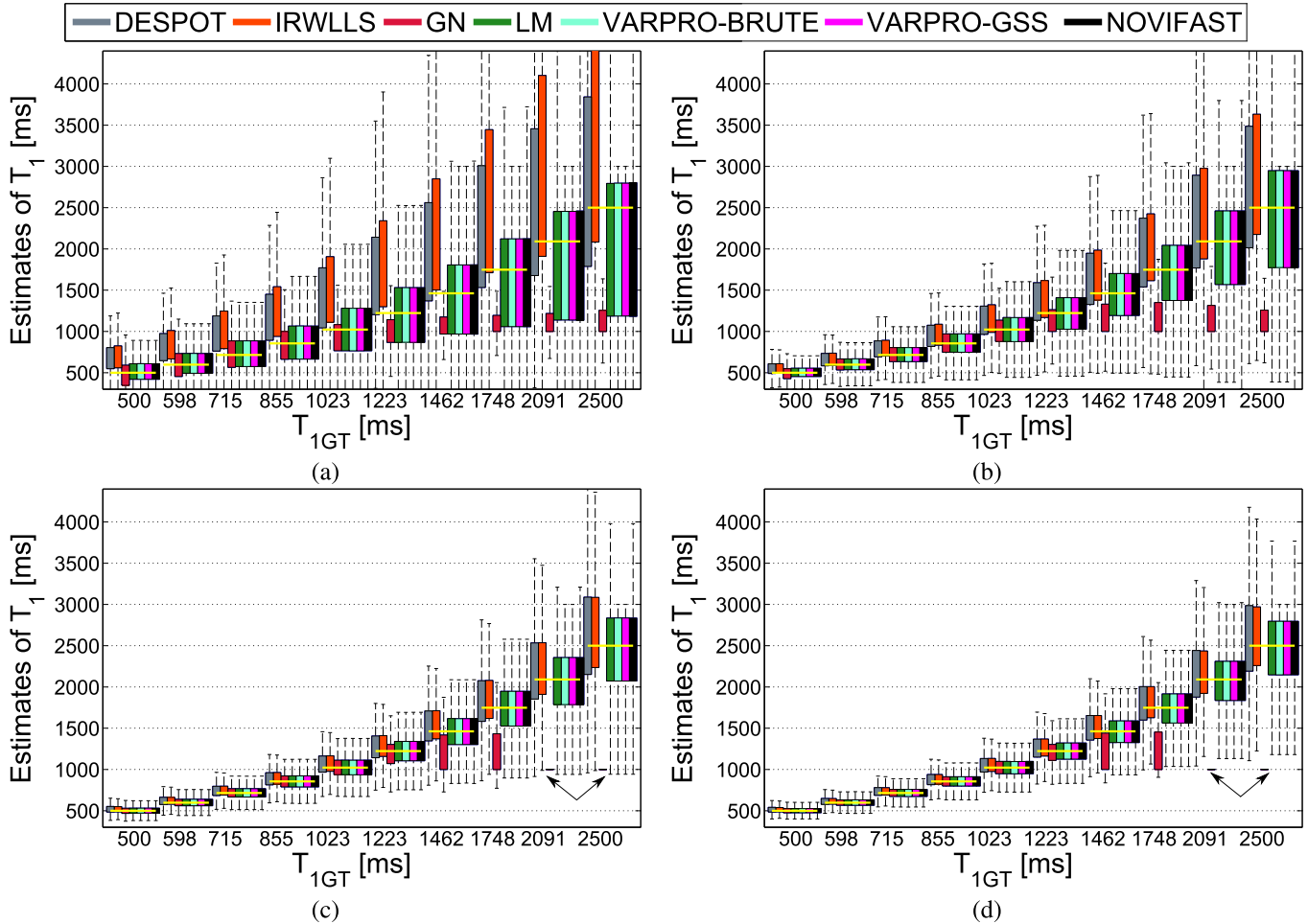


Fig. 1. Box-plots of the T_1 estimates that are obtained with the five SPGR VFA optimization algorithms. Tukey-style whiskers are shown that extend to a maximum of $1.5 \times \text{IQR}$ beyond each box, with IQR the interquartile range (corresponding with the length of each box) [45]. Ground-truth T_1 values are marked with horizontal yellow lines to ease interpretation (Case of $\mathcal{A}_{N=10}$ FA set and fixed initialization of $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 1000$ ms). (a) $\text{SNR}_{90^\circ} = 80$, $\text{SNR} \in [1.8, 6.9]$ & (b) $\text{SNR}_{90^\circ} = 150$, $\text{SNR} \in [3.3, 12.8]$. (c) $\text{SNR}_{90^\circ} = 250$, $\text{SNR} \in [5.6, 21.3]$ & (d) $\text{SNR}_{90^\circ} = 300$, $\text{SNR} \in [6.7, 25]$.

were generated, similarly as in Section IV-A, but this time with a spatially variant σ . We employed similar noise maps as those reported in [40], and we scaled them so as to get spatially averaged SNR_{90° values of 400 and 500 (and corresponding SNR values of 12.3 and 15, respectively). Those values are within the range of SNRs encountered in practice [41]. For each of those values of SNR_{90° , $N_{\text{MC}} = 10^4$ noisy realizations were generated.

T_1 maps were estimated by applying DESPOT1, IRWLLS, LM, VARPRO-GSS, and NOVIFAST in a voxel-wise manner, with a mask including only brain tissue voxels ($M_{\text{voxels}} = 212411$ in total).² Code parallelization was not performed, and all algorithms were run with the same parameters as in the first experiment, except VARPRO-GSS. Since the ground-truth T_1 map contains values beyond gray matter, we used a wider grid,

²Since VARPRO-GSS provided identical estimates as VARPRO-BRUTE, but with a much shorter computation time, we did not apply VARPRO-BRUTE in this experiment. Moreover, initial experiments showed that the GN algorithm did not converge for most of the voxels in the phantom. Therefore, we decided not to include GN in the MC experiment to avoid a drastic increase in total time, which would have rendered the experiment infeasible for $N_{\text{MC}} = 10^4$ repetitions.

namely $T_{1\text{min-grid}} = 250$ ms and $T_{1\text{max-grid}} = 8000$ ms with the same grid resolution (1 ms). The algorithms that require an initialization were initialized with constant K and T_1 maps. The initial values were $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 1000$ ms. Algorithms were stopped using the same criterion as in the first experiment. For each algorithm, the total computation time as well as the spatially averaged relative bias, standard deviation (Std.), and rMSE were reported.

C. In Vivo Human Brain Set of SPGR MR Images

A set of SPGR MR images of a healthy 29-year old male volunteer was acquired with a 1.5T MRI scanner (MAGNETOM Aera, Siemens) using a 20-channel head coil. For each FA given in the $\mathcal{A}_{N=10}$ set, a $256 \times 256 \times 30$ image (voxel size $1 \times 1 \times 5$ mm³) was acquired with a 3D FLASH sequence. The sequence parameters were: TR/TE = 7.8/3.48 ms and BW = 320 Hz. B_1 mapping was not performed. Magnitude images were reconstructed with the GRAPPA and adaptive combine method (acceleration factor of 2) [42], [43]. An SNR = 10 was estimated with the method of [44]. The total scan time per FA was

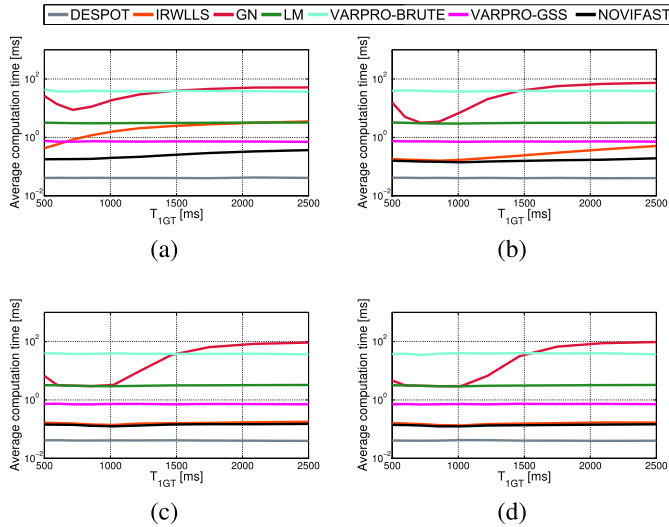


Fig. 2. Total computation time of each of the five optimization algorithms for the MC simulation-based experiment (Case of $\mathcal{A}_{N=10}$ FA set and fixed initialization of $K_{\text{ini}}=0.5$ and $T_{1\text{ini}}=1000$ ms). (a) $\text{SNR}_{90^\circ} = 80$, & (b) $\text{SNR}_{90^\circ} = 150$, $\text{SNR} \in [1.8, 6.9]$ & $\text{SNR} \in [3.3, 12.8]$. (c) $\text{SNR}_{90^\circ} = 250$, & (d) $\text{SNR}_{90^\circ} = 300$, $\text{SNR} \in [5.6, 21.3]$ & $\text{SNR} \in [6.7, 25]$.

45 seconds. In order to show that NOVIFAST does not require a careful initialization, and can be initialized with a constant map, we chose for all the algorithms that require initialization $T_{1\text{ini}} = 500$ ms and $K_{\text{ini}} = 1.9 \cdot 10^3$ (being the average value of the estimated K map obtained with DESPOT1). Algorithms were stopped according to the same criteria as used in the second experiment, and identical tuning parameters were used. Similar to the previous experiment, all algorithms were applied voxel-wise with a mask including only brain tissue voxels ($M_{\text{voxels}} = 632260$ in total). VARPRO-BRUTE was applied with the same setting parameters as VARPRO-GSS. As such, VARPRO-BRUTE requires 7751 cost function evaluations per voxel, that is, $M_{\text{voxels}} \cdot 7751 = 4.9 \cdot 10^9$ in total.

V. RESULTS

A. MC Simulation With Synthetically Generated SPGR MR Signals

To avoid manuscript overlength, we only present the results for the $\mathcal{A}_{N=10}$ FA set with initialization $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 1000$ ms. Results for $K_{\text{ini}} = 0.5$ and $T_{1\text{ini}} = 500$ ms, as well as results for the $\mathcal{A}_{N=3}$ FA set (with both initializations), are presented in the supplementary file. In Fig. 1, we show box-plots for the T_1 estimates obtained with the five algorithms under test. A box-plot-based visualization allows us to see whether there exist statistically significant differences in the population of the sample estimates of T_1 [45]. In that box-plot visualization, horizontal yellow lines, representing the values of $T_{1\text{GT}}$, are intentionally marked for ease of interpretation. To illustrate the speed performance, the average computation time is displayed in Fig. 2. Results for the second initialization, both regarding the statistical performance and the computation time, are presented in Fig. 9 and Fig. 10 of the supplementary file.

A first, general observation that we can already point out is that less accurate and precise T_1 estimates are obtained if heuristic linearized estimators, such as DESPOT1 and IRWLLS, are employed. This can be attributed simply to the statistical superiority of the NLLS estimator over modified linear versions. It is true, though, that the difference becomes less noticeable with increasing SNR, and is less drastic for the $\mathcal{A}_{N=3}$ FA data set. The most interesting observations are those related to GN, LM, VARPRO-BRUTE, VARPRO-GSS, and NOVIFAST. From Fig. 1 it is clear that GN is sensitive to initialization. The local-convergence behavior of GN evidently degrades for $T_{1\text{GT}} > 1000$ ms, yielding a systematic bias which persists over the whole range of SNRs. Observe that for $T_{1\text{GT}} > 1223$ ms, the GN interquartile ranges (IQRs), covering the middle 50% of the sample, do not cover $T_{1\text{GT}}$. It can also be observed that the corresponding IQRs are approximately clustered around 1000 ms, i.e., $T_{1\text{ini}}$, and the variability of the estimates greatly reduces for that regime. Note as well that the length of the GN boxes in Fig. 1(c) and Fig. 1(d) becomes very small for $T_{1\text{GT}} > 1748$ ms (see black arrows pointing the boxes). This behavior is due to convergence failure, since GN is stagnating at the initialization, as we observed in a large number of MC realizations. This is also reflected by the method's computation time in Fig. 2, which massively grows for high $T_{1\text{GT}}$, since GN was only stopped when it reached the maximum number of iterations.

The LM algorithm and, especially, NOVIFAST present a remarkable insensitivity to initialization, providing very similar results. Results presented in the supplementary file confirm that the same conclusions hold when the second initialization was employed. Indeed, in that case, NOVIFAST and LM also manifest a substantial robustness to poor initial values in contrast to GN, which starts to fail as soon as the value of $T_{1\text{GT}} = 715$ ms is reached. Both VARPRO-BRUTE and VARPRO-GSS, as expected, present identical results, and the statistical performance is nearly equal to that of LM and NOVIFAST.³

Having shown that both LM and NOVIFAST have good global convergence properties, and that their statistical performance is similar as VARPRO-BRUTE and VARPRO-GSS, what distinguishes them is, as we have already mentioned, their computational speed. Due to the negligible cost per iteration and the rapid convergence, the average computational time of NOVIFAST is overall more than one order lower than LM. Indeed, averaged over SNRs and $T_{1\text{GT}}$, NOVIFAST is 20 times faster. Furthermore, NOVIFAST is much faster than VARPRO-BRUTE (more than 200 times), and substantially faster than the efficient VARPRO-GSS, 8 times.

This speed gain increases with SNR, since the number of iterations needed for NOVIFAST to converge asymptotically decreases to one as the SNR approaches infinity, i.e., NOVIFAST, unlike LM, VARPRO-BRUTE, or VARPRO-GSS, asymptotically approaches an exact, analyt-

³Estimates with VARPRO-BRUTE and VARPRO-GSS show less variability for high $T_{1\text{GT}}$ (observe that the top part of the whiskers are shorter than those of LM and NOVIFAST) due to the maximum allowable value of the method's grid, 3000 ms in our implementation. It has nothing to do with any intrinsic feature of VARPRO that leads to statistical superior results.

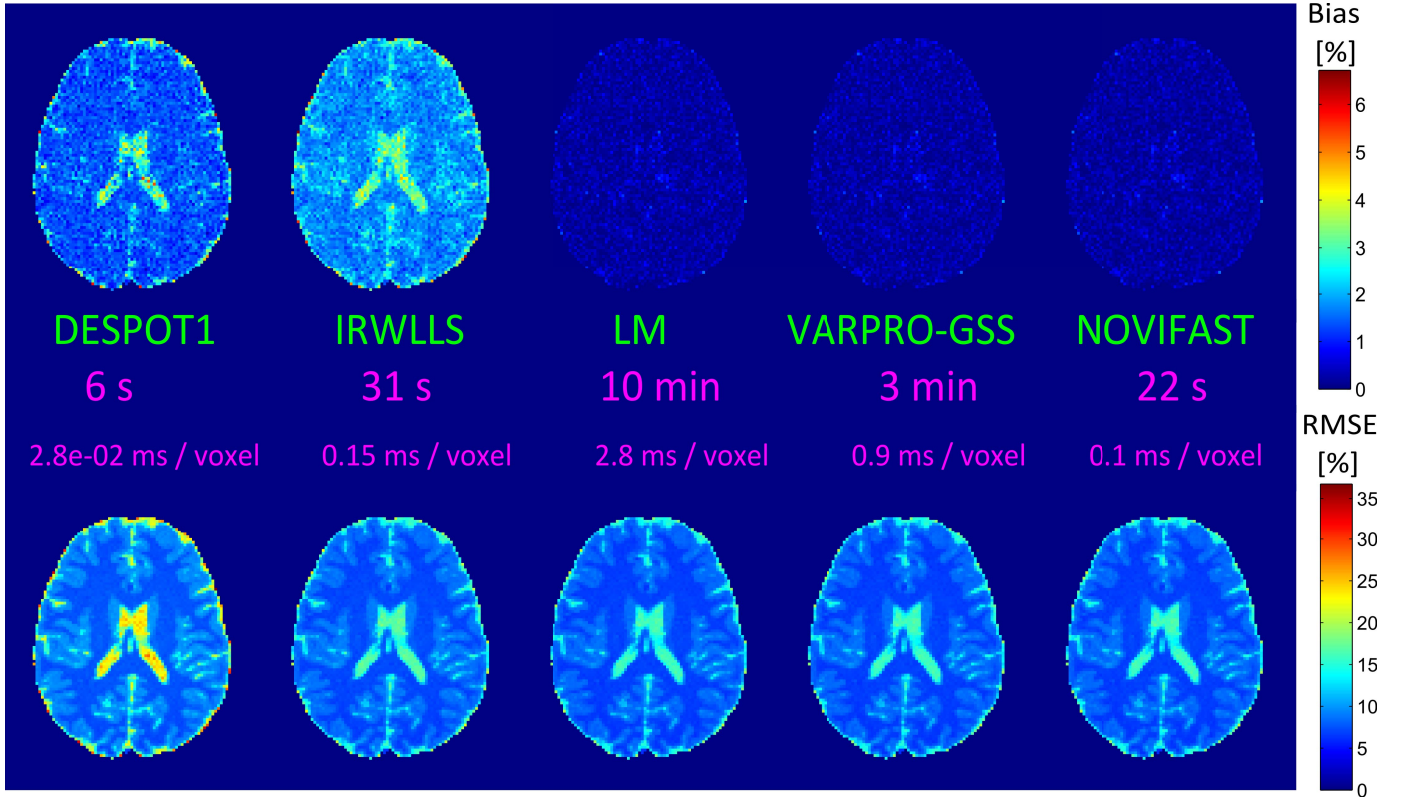


Fig. 3. Bias and RMSE maps of a mid-axial slice of a synthetic 3D T_1 map ($\text{SNR}_{90^\circ} = 500$) obtained with DESPOT1, IRWLLS, LM, VARPRO-GSS, and NOVIFAST. The computation times are shown as well (total time for the 3D T_1 map and time per voxel, $M_{\text{voxels}} = 212411$).

TABLE I

QUANTITATIVE RESULTS OBTAINED WITH DESPOT1, IRWLLS, LM, VARPRO-GSS, AND NOVIFAST, WHEN ESTIMATING A SYNTHETIC 3D T_1 MAP ($\text{SNR}_{90^\circ} = 400$)

	Bias [%]	Std. [%]	RMSE [%]	Time [s]
DESPOT1	2.48	13.45	13.68	6.2
IRWLLS	3.13	11.45	11.87	31.4
LM	0.32	11.03	11.03	640.2
VARPRO-GSS	0.32	11.03	11.03	157.8
NOVIFAST	0.32	11.03	11.03	23.2

TABLE II

QUANTITATIVE RESULTS OBTAINED WITH DESPOT1, IRWLLS, LM, VARPRO-GSS, AND NOVIFAST, WHEN ESTIMATING A SYNTHETIC 3D T_1 MAP ($\text{SNR}_{90^\circ} = 500$)

	Bias [%]	Std. [%]	RMSE [%]	Time [s]
DESPOT1	1.58	10.38	10.5	6.2
IRWLLS	1.99	9.00	9.23	30.2
LM	0.24	8.80	8.81	637.8
VARPRO-GSS	0.24	8.80	8.81	156.9
NOVIFAST	0.24	8.80	8.80	22.1

ical estimator. By observing Fig. 2, it is clear that NOVIFAST's computational time is nearly constant, around 0.16 ms.

The reported computational time of NOVIFAST seems similar to that of IRWLLS, but NOVIFAST's statistical performance is considerably higher. Indeed, being an heuristic algorithm, convergence of IRWLLS in l_2 norm does not imply that the NLLS cost function is effectively minimized. Readers can check that the same conclusions that are drawn in this subsection also hold for the $\mathcal{A}_{N=3}$ FA set (supplementary file).

B. MC Simulation With Synthetic 3D T_1 Phantom

The overall accuracy (bias), precision (Std.), RMSE and computational time are shown in Table I and II for $\text{SNR}_{90^\circ} = 400$ and $\text{SNR}_{90^\circ} = 500$, respectively. Bias and RMSE maps for one particular mid-axial slice are shown in Fig. 3 (for the case $\text{SNR}_{90^\circ} = 500$). As expected, LM, VARPRO-GSS, and NOVIFAST provide nearly identical accuracy, precision, and RMSE. Nevertheless, NOVIFAST provides the NLLS

estimated 3D T_1 map in about 23 s, whereas it takes around 10 min for LM and 3 min for VARPRO-GSS to do so. The accuracy of NLLS-based algorithms is drastically higher than that of DESPOT1 and IRWLLS (at least 6 \times) and the RMSE is lower as well, confirming the statistical superiority of the NLLS estimator over heuristic modifications. Finally, observe that NOVIFAST is consistently faster than IRWLLS.

C. In Vivo Human Brain Set of SPGR MR Images

Although we are not able to assess the statistical performance of the five methods with real data, due to the overarching issue of lacking a noise-free ground truth, interesting observations can be made about convergence and total computation time. Estimated T_1 maps of a mid-axial slice is shown in Fig. 4. A constant initial map is not an impediment for NOVIFAST to estimate a reliable T_1 map, which is not the case for GN (observe the yellow arrow on the lowest part of the brain, where the estimated T_1 value remains constant and

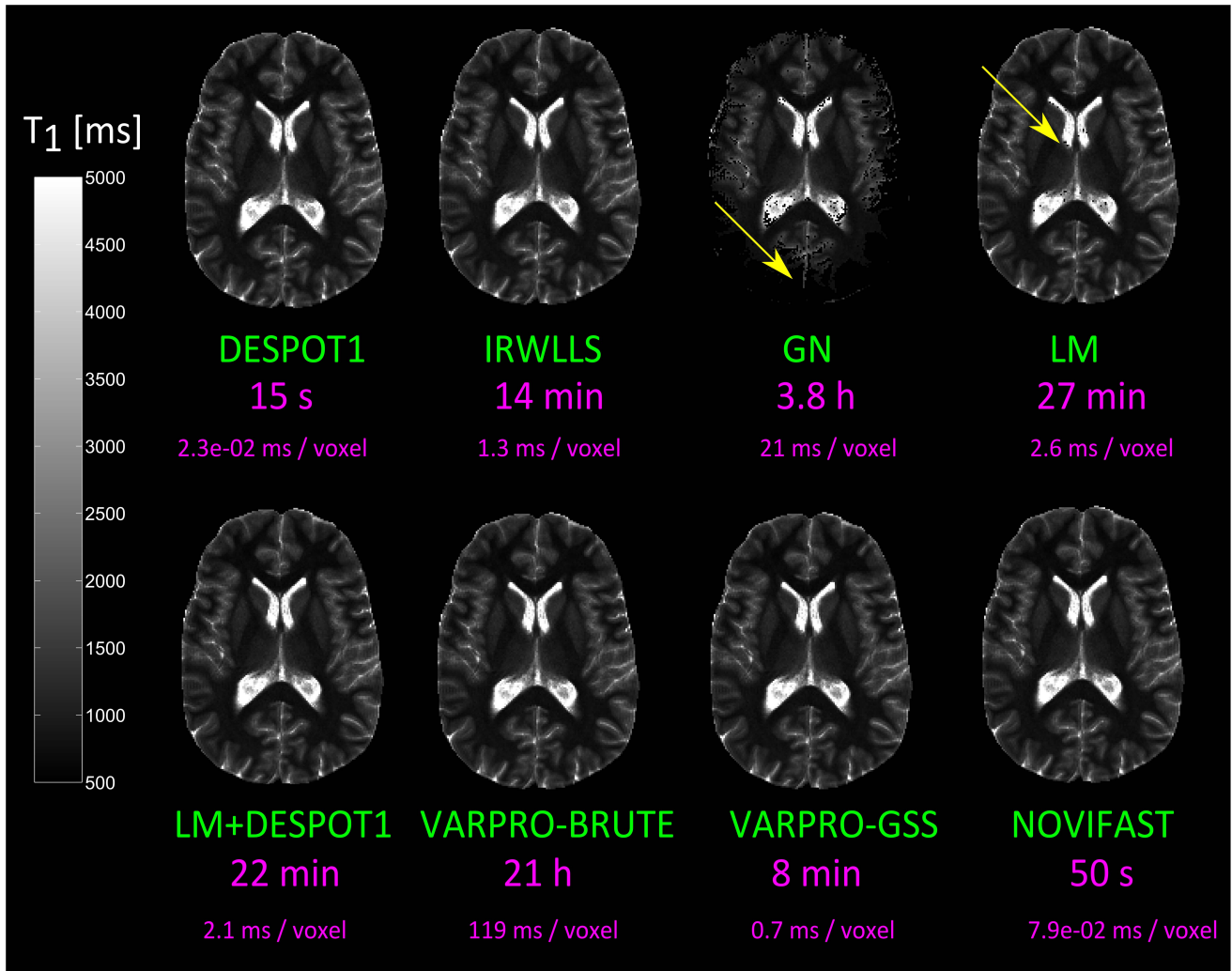


Fig. 4. Mid-axial slice of the estimated 3D T_1 map with the VFA T_1 estimation algorithms compared for the in vivo MR experiment. The computation times are shown as well (total time for the 3D T_1 map and time per voxel, $M_{\text{voxels}} = 632260$).

equal to 500 ms). Interestingly, several outliers are manifestly clear in the T_1 map that was estimated with LM. They are located on the boundary between cerebrospinal fluid and gray matter, and are due to poor initialization. Indeed, if LM is initialized with DESPOT1, those outliers are not present in the estimated T_1 map. No outliers were observed in the T_1 maps estimated with DESPOT1, IRWLLS, VARPRO-GSS, and NOVIFAST.

DESPOT1 and NOVIFAST are the only methods that can provide this high resolution 3D T_1 map in less than 1 min, where the NOVIFAST speed gain over IRWLLS, LM, VARPRO-BRUTE, and VARPRO-GSS, is about 16 \times , 27 \times , 1500 \times , and 9 \times , respectively. Although differences between the T_1 maps of DESPOT1 and NOVIFAST, are, as expected, visually indistinguishable, the T_1 estimates substantially differ when both methods are compared (see Fig. 5). It should be noted though that statistical claims can only be made based on the previous MC simulations, and not on difference maps from a single real data experiment.

Before finishing this section, it is worthwhile to point out that the SPGR model of Eq.(1) neglects partial volume

effects, which may be relevant for highly non-isotropic voxels, as those of the dataset used in this experiment. Furthermore, possible incomplete spoiling may introduce related T_2 decay effects [46]. However, such effects are intrinsic to the SPGR model at hand, and not to the choice of the algorithm. Indeed, NOVIFAST is always applicable wherever the rest of the compared methods in this manuscript are.

VI. EXTENSIONS

NOVIFAST can be extended in several ways. In this work, it has been presented as an *ordinary* NLLS optimization algorithm, but extension to the *weighted* NLLS estimator is possible (see [30] for more details).

NOVIFAST can be applied independently of the statistical distribution of the noisy data. However, it is worth mentioning that NOVIFAST equals the Maximum-Likelihood (ML) estimator when the data is Gaussian distributed [47], and as such, it provides statistically optimal estimates. This situation occurs when VFA T_1 mapping is directly performed on the real and imaginary part of complex-valued data that results from reconstructing single-coil k-space. It also arises in parallel

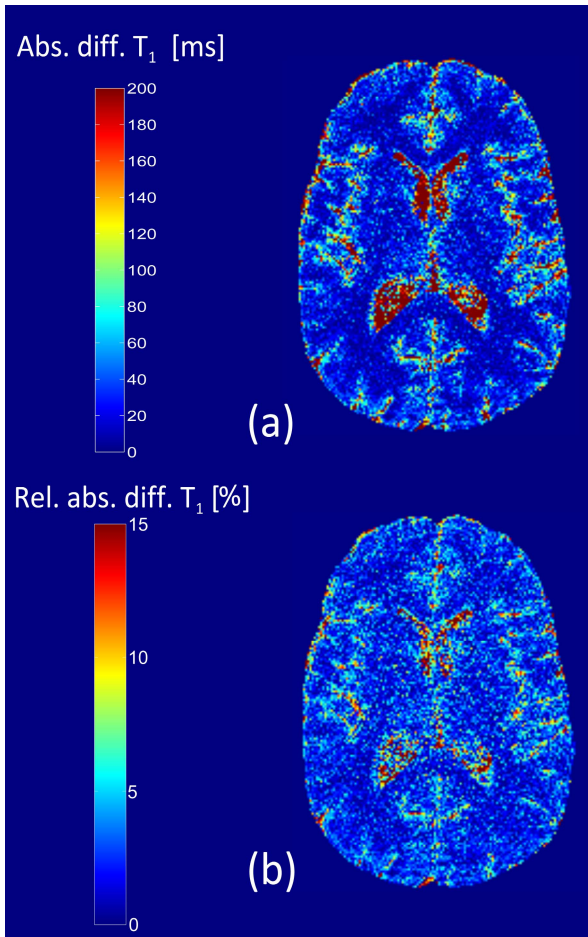


Fig. 5. Absolute difference T_1 map (a) and relative absolute difference T_1 map (w.r.t. DESPOT1) (b) between DESPOT1 and NOVIFAST for the mid-axial slice of Fig. 4.

acquisitions, when, for example, a complex MR image is obtained with the GRAPPA method. In this case, the Gaussian distribution also applies, and hence NOVIFAST provides the ML estimates [13]. Interestingly, when the data distribution at hand is not Gaussian, it is yet possible to embed NOVIFAST in a Maximum-Likelihood (ML) framework, aiming at fully exploiting the statistical knowledge of both single and multi-coil MR data [48], [49]. This is the case when the data follows a Rician or a non-central χ distribution, which are often encountered in typical parallel MRI acquisitions. Indeed, MR images reconstructed with the SENSE method follow a Rician distribution, and the magnitude of a complex image obtained with GRAPPA and the adaptive combine method does so as well [13], [36]. A non-central χ distribution typically applies when complex images from several coils are combined with the Sum of Squares (SoS) method [13], [36]. It has been shown recently that solving an ML estimation problem with non-central χ or Rician distributed data is equivalent to iteratively solving a collection of NLLS subproblems [39], [50]. As a result, NOVIFAST can be integrated into this approach, by solving each of the NLLS subproblems, thereby greatly boosting the speed of the overall ML estimation procedure.

Furthermore, NOVIFAST can be applied to other MR data sequences than SPGR. For example, it is well known that the

completely balanced TrueFISP sequence can be modeled as a quotient of continuous functions, where the (three) linear parameters encode the T_1 and T_2 values [51]. NOVIFAST can be reformulated for TrueFISP, since the semi-linear structure of Eq. (28) also appears when fitting a quotient of continuous functions [33].

Finally, NOVIFAST can be embedded into a unified framework where T_1 mapping and actual flip angle mapping are performed simultaneously, similarly to [52]. The NLLS optimization problem of Eq. (2) can be extended to include simultaneous estimation of the flip angle correction factor. As the nominal flip angles are known beforehand, this three-dimensional optimization problem also provides estimates of the actual flip angles, thereby performing actual flip-angle mapping [53]. The three-dimensional NLLS optimization problem may be solved with a cyclic coordinate descent algorithm [39], [54], being the optimization problem decomposed into two iterative subproblems. Indeed, given estimates of K and T_1 , the flip angle correction factors can be estimated with any conventional NLLS optimization algorithm. On the other hand, when estimates of the flip angle correction factors are known, the only remaining optimization variables are K and T_1 . Since this iterative subproblem is essentially the same as Eq. (2), it can be solved very fast and accurate with NOVIFAST. Cyclic iteration between these two subproblems constitutes the global optimization algorithm for simultaneously actual flip angle and T_1 mapping.

VII. CONCLUSIONS

In VFA T_1 mapping, it is common to use heuristic linear estimators such as the celebrated DESPOT1 method [5] or the IRWLLS method of Chang *et al.* [12]. They are preferred despite the superiority of NLLS estimators in terms of accuracy and precision, since optimization algorithms for NLLS estimators are much slower than linear estimators, and are also prone to initialization issues. In this work, we reconcile these two separate frameworks by proposing a novel NLLS method, NOVIFAST, which provides the NLLS estimates more than twenty times faster than conventional gradient-based NLLS optimization algorithms and about eight times faster than efficient VARPRO-based method implementations. Thus, our NOVIFAST algorithm merges the best of both approaches, i.e., accurate and precise T_1 mapping with a very short computation time. This is the consequence of conceiving NOVIFAST in a fundamentally different fashion than other VFA T_1 mapping methods. To derive NOVIFAST, our starting point was the reparameterization of the numerator and denominator of the SPGR model by a set of two linear coefficients (Eq. (11)). However, instead of applying the linearization of the SPGR model that DESPOT1 includes, we sought for the optimal linear coefficients in an NLLS sense, which is exactly equivalent to performing NLLS T_1 fitting directly, as we theoretically proved in this paper. NOVIFAST does not attempt to find the optimal linear coefficients by decreasing the NLLS criterion, as gradient-based descent algorithms, e.g., the LM algorithm, do. Moreover, with NOVIFAST, numerical evaluation of the NLLS cost function is not required, as other techniques such as VARPRO-based methods do so extensively.

Indeed, with NOVIFAST, the problem of finding the optimal linear coefficients in NLLS sense is transformed into a root-finding problem that arises when the first-order conditions for optimality are imposed. In this derivation, we made use of algebraic rules in the same spirit of those of [30], to come up with a set of non-linear equations, which is iteratively solved as a two-by-two linear system problem. The iterative procedure, which qualifies as a fixed-point algorithm, possesses very rapid convergence, usually between two and four iterations are required. The excellent convergence properties have been evidenced with MC simulations, in a framework of fixed-point algorithms theory. The computational time per iteration is rather low, similar to DESPOT1. Furthermore, as we have experimentally corroborated, NOVIFAST is quite robust to initialization. In fact, it can be initialized with constant T_1 maps. Therefore, minimal pre-processing is needed for NOVIFAST in a real implementation.

Another distinct feature of NOVIFAST in comparison to other NLLS-based algorithms is that, with noiseless SPGR data, it provides the T_1 estimates with zero error, with only one iteration, and for every initialization that is chosen. That is, NOVIFAST, despite being an iterative algorithm, becomes an analytical estimator when the SPGR model is noise-free. This result, which we prove theoretically, guarantees a very appealing behavior as the SNR increases in realistic noisy scenarios, since the number of iterations necessarily has to decrease to one so as to mimic the remarkable asymptotic behavior in absence of noise.

We believe that NOVIFAST is a good candidate to be included in every processing pipeline for high-quality, robust, and efficient VFA T_1 mapping.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

Note Added in Proof: The Matlab source code of NOVIFAST is available at <https://visielab.uantwerpen.be/research-topics/novifast>

REFERENCES

- [1] P. Tofts, *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*. Chichester, U.K.; Hoboken, NJ, USA: Wiley, 2004.
- [2] H. B. W. Larsson *et al.*, "Assessment of demyelination, edema, and gliosis by *in vivo* determination of T_1 and T_2 in the brain of patients with acute attack of multiple sclerosis," *Magn. Reson. Med.*, vol. 11, pp. 337–348, Sep. 1989.
- [3] P. Conlon, M. R. Trimble, D. Rogers, and C. Callicott, "Magnetic resonance imaging in epilepsy: A controlled study," *Epilepsy Res.*, vol. 2, pp. 37–43, Jan./Feb. 1988.
- [4] T. Erkinjuntti, L. Ketonen, R. Sulkava, J. Sipponen, M. Vuorioaho, and M. Iivanainen, "Do white matter changes on MRI and CT differentiate vascular dementia from Alzheimer's disease?" *J. Neurol. Neurosurg. Psychiatry*, vol. 50, pp. 37–42, Jan. 1987.
- [5] S. C. Deoni, B. K. Rutt, and T. M. Peters, "Rapid combined T_1 and T_2 mapping using gradient recalled acquisition in the steady state," *Magn. Reson. Med.*, vol. 49, no. 3, pp. 515–526, 2003.
- [6] N. Stikov, M. Boudreau, I. R. Levesque, C. L. Tardif, J. K. Barral, and G. B. Pike, "On the accuracy of T_1 mapping: Searching for common ground," *Magn. Reson. Med.*, vol. 73, no. 2, pp. 514–522, 2015.
- [7] K. A. Christensen, D. M. Grant, E. M. Schulman, and C. Walling, "Optimal determination of relaxation times of Fourier transform nuclear magnetic resonance. determination of spin-lattice relaxation times in chemically polarized species," *J. Phys. Chem.*, vol. 78, no. 19, pp. 1971–1977, 1974.
- [8] J. Homer and M. S. Beevers, "Driven-equilibrium single-pulse observation of T_1 relaxation. A reevaluation of a rapid 'new' method for determining NMR spin-lattice relaxation times," *J. Magn. Reson.*, vol. 63, no. 2, pp. 287–297, 1985.
- [9] E. K. Fram *et al.*, "Rapid calculation of T_1 using variable flip angle gradient refocused imaging," *Magn. Reson. Imag.*, vol. 5, no. 3, pp. 201–208, 1987.
- [10] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Optimizing MR scan design for model-based T_1 , T_2 estimation from steady-state sequences," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 467–477, Feb. 2017.
- [11] R. Gupta, "A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using Fourier transform NMR," *J. Magn. Reson.*, vol. 25, no. 1, pp. 231–235, 1977.
- [12] L. C. Chang, C. G. Koay, P. J. Basser, and C. Pierpaoli, "Linear least-squares method for unbiased estimation of T_1 from SPGR signals," *Magn. Reson. Med.*, vol. 60, no. 2, pp. 496–501, 2008.
- [13] A. J. den Dekker and J. Sijbers, "Data distributions in magnetic resonance images: A review," *Phys. Medica*, vol. 30, pp. 725–741, Nov. 2014.
- [14] R. P. A. G. Teixeira, S. J. Malik, and J. V. Hajnal, "Joint system relaxometry (JSR) and Crámer–Rao lower bound optimization of sequence parameters: A framework for enhanced precision of DESPOT T_1 and T_2 estimation," *Magn. Reson. Med.*, vol. 79, no. 1, pp. 234–245, 2018.
- [15] J. Nocedal and S. J. Wright, *Numerical Optimization*, 4th ed. New York, NY, USA: McGraw-Hill, 2006.
- [16] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quart. J. Appl. Math.*, vol. 2, no. 2, pp. 164–168, Jul. 1944.
- [17] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.
- [18] H. Dan, N. Yamashita, and M. Fukushima, "Convergence properties of the inexact Levenberg–Marquardt method under local error bound conditions," *Optim. Methods Softw.*, vol. 17, no. 4, pp. 605–626, 2002.
- [19] J. Chen and W. Li, "Convergence of Gauss–Newton's method and uniqueness of the solution," *Appl. Math. Comput.*, vol. 170, no. 1, pp. 686–705, 2005.
- [20] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in Neural Information Processing Systems*, Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, Eds. Vancouver, BC, Canada: Curran Associates, Inc., 2009, pp. 1759–1767.
- [21] D. G. Luenberger, *Optimization by Vector Space Methods*. 1st ed. New York, NY, USA: Wiley, 1969.
- [22] G. Golub and V. Pereyra, "Separable nonlinear least squares: The variable projection method and its applications," *Inverse Problems*, vol. 19, no. 2, p. R1, 2003.
- [23] D. P. O'Leary and B. W. Rust, "Variable projection for nonlinear least squares problems," *Comput. Optim. Appl.*, vol. 54, no. 3, pp. 579–593, 2013.
- [24] J. P. Haldar, J. Anderson, and S.-W. Sun, "Maximum likelihood estimation of T_1 relaxation parameters using VARPRO," in *Proc. Int. Soc. Mag. Reson. Med.*, vol. 15, 2007, p. 41.
- [25] J. K. Barral, E. Gudmundson, N. Stikov, M. Etezadi-Amoli, P. Stoica, and D. G. Nishimura, "A robust methodology for *in vivo* T_1 mapping," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1057–1067, 2010.
- [26] J. D. Trzasko, P. M. Mostardi, S. J. Riederer, and A. Manduca, "Estimating T_1 from multichannel variable flip angle SPGR sequences," *Magn. Reson. Med.*, vol. 69, no. 6, pp. 1787–1794, 2013.
- [27] B. Zhao, K. Setsompop, H. Ye, S. F. Cauley, and L. L. Wald, "Maximum likelihood reconstruction for magnetic resonance fingerprinting," *IEEE Trans. Med. Imag.*, vol. 35, no. 8, pp. 1812–1823, Aug. 2016.
- [28] J. Gong and J. P. Hornak, "A fast T_1 algorithm," *Magn. Reson. Imag.*, vol. 10, no. 4, pp. 623–626, 1992.
- [29] L. C. Chang, C. Koay, S. Deoni, and C. Pierpaoli, "Comparison of linear and non-linear fitting methods for estimating T_1 from SPGR signals," in *Proc. 15th Annu. Meeting ISMRM*, Berlin, Germany, 2007, p. 3715.
- [30] S. D. Dimitrov and D. I. Kamenski, "A parameter estimation method for rational functions," *Comput. Chem. Eng.*, vol. 15, no. 9, pp. 657–662, 1991.

- [31] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 398–410, Mar. 1998.
- [32] S. M. Robinson, "A short proof of Cramer's rule," *Math. Mag.*, vol. 43, no. 2, pp. 94–95, 1970.
- [33] S. D. Dimitrov and D. I. Kamenski, "Parameter estimation in complicated rational functions," *Comput. Chem.*, vol. 20, no. 3, pp. 331–337, 1996.
- [34] G. Ramos-Llordén, A. J. den Dekker, M. Björk, M. Verhoye, and J. Sijbers, "NOVIFAST: A fast non-linear least squares method for accurate and precise estimation of T_1 from SPGR signals," in *Proc. Int. Soc. Mag. Reson. Med.*, vol. 24, 2016, p. 2820.
- [35] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Mag. Reson. Med.*, vol. 34, no. 6, pp. 910–914, 1995.
- [36] S. Aja-Fernández and G. Vegas-Sánchez-Ferrero, *Statistical Analysis of Noise in MRI: Modeling, Filtering and Estimation*. 1st ed. Cham, Switzerland: Springer, 2016.
- [37] H. L. Cheng and G. A. Wright, "Rapid high-resolution T_1 mapping by variable flip angles: Accurate and precise measurements in the presence of radiofrequency field inhomogeneity," *Magn. Reson. Med.*, vol. 55, no. 3, pp. 566–574, 2006.
- [38] S. C. Deoni, T. M. Peters, and B. K. Rutt, "Determination of optimal angles for variable nutation proton magnetic spin-lattice, T_1 , and spin-spin, T_2 , relaxation times measurement," *Magn. Reson. Med.*, vol. 51, no. 1, pp. 194–199, 2004.
- [39] G. Ramos-Llordén *et al.*, "A unified maximum likelihood framework for simultaneous motion and T_1 estimation in quantitative MR T_1 mapping," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 433–446, Feb. 2017.
- [40] T. Pieciak, S. Aja-Fernández, and G. Vegas-Sánchez-Ferrero, "Non-stationary Rician noise estimation in parallel MRI using a single image: A variance-stabilizing approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 10, pp. 2015–2029, Oct. 2017.
- [41] M. Bouhrara and R. G. Spencer, "Rapid simultaneous high-resolution mapping of myelin water fraction and relaxation times in human brain using BMC-mcDESPOt," *NeuroImage*, vol. 147, pp. 800–811, Feb. 2017.
- [42] M. A. Griswold *et al.*, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, 2002.
- [43] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin, "Adaptive reconstruction of phased array MR imagery," *Magn. Reson. Med.*, vol. 43, no. 5, pp. 682–690, 2000.
- [44] S. Aja-Fernández, T. Pieciak, and G. Vegas-Sánchez-Ferrero, "Spatially variant noise estimation in MRI: A homomorphic approach," *Med. Image Anal.*, vol. 20, no. 1, pp. 184–197, 2015.
- [45] M. Krzywinski and N. Altman, "Points of significance: Visualizing samples with box plots," *Nature Methods*, vol. 11, no. 2, pp. 119–120, 2014.
- [46] R. Heule, C. Ganter, and O. Bieri, "Variable flip angle T_1 mapping in the human brain with reduced T_2 sensitivity using fast radiofrequency-spoiled gradient echo imaging," *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1413–1422, 2016.
- [47] A. van den Bos, *Parameter Estimation for Scientists and Engineers*, 1st ed. Hoboken, NJ, USA: Wiley, 2007.
- [48] J. J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck, "Maximum-likelihood estimation of Rician distribution parameters," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 357–361, Mar. 1998.
- [49] J. Sijbers and A. J. den Dekker, "Maximum likelihood estimation of signal amplitude and noise variance from MR data," *Magn. Reson. Med.*, vol. 51, no. 3, pp. 586–594, 2004.
- [50] D. Varadarajan and J. P. Haldar, "A majorize-minimize framework for Rician and non-central chi MR images," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 2191–2202, Oct. 2015.
- [51] K. Scheffler and J. Hennig, "Is TrueFISP a gradient-echo or a spin-echo sequence?" *Magn. Reson. Med.*, vol. 49, no. 2, pp. 395–397, 2003.
- [52] S. A. Hurley, V. L. Yarnykh, K. M. Johnson, A. S. Field, A. L. Alexander, and A. A. Samsonov, "Simultaneous variable flip angle-actual flip angle imaging method for improved accuracy and precision of three-dimensional T_1 and B_1 measurements," *Magn. Reson. Med.*, vol. 68, no. 1, pp. 54–64, 2012.
- [53] V. L. Yarnykh, "Actual flip-angle imaging in the pulsed steady state: A method for rapid three-dimensional mapping of the transmitted radiofrequency field," *Magn. Reson. Med.*, vol. 57, pp. 192–200, Jan. 2007.
- [54] J. A. Fessler and D. Kim, "Axial block coordinate descent (ABCD) algorithm for X-ray CT image reconstruction," in *Proc. Int. Mtg. Fully 3D Image Recon. Rad. Nucl. Med.*, 2011, pp. 262–265.