# Weierstraß-Institut
## für Angewandte Analysis und Stochastik
### Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855


# A physically oriented method for quantitative magnetic resonance imaging


Guozhi Dong[1,2], Michael Hintermüller[1,2], Kostas Papafitsoros[1]



submitted: July 31, 2018



[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: guozhi.dong@wias-berlin.de
michael.hintermueller@wias-berlin.de
kostas.papafitsoros@wias-berlin.de

[2] Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
E-Mail: hint@math.hu-berlin.de
guozhi.dong@hu-berlin.de



No. 2528
Berlin 2018



2010 *Mathematics Subject Classification.* 65K10, 65L09, 92C55, 49M15, 49N45.

*Key words and phrases.* Quantitative magnetic resonance imaging, Bloch equations, parameter identification, fingerprinting, dictionary, Newton method, Levenberg-Marquardt method.

This work has been supported by the Einstein Center Mathematics research project (ECMath CH12) in the framework of Mathematics in Clinical Research and Health Care.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

# A physically oriented method for quantitative magnetic resonance imaging

Guozhi Dong, Michael Hintermüller, Kostas Papafitsoros

**Abstract**

Quantitative magnetic resonance imaging (qMRI) denotes the task of estimating the values of magnetic and tissue parameters, e.g., relaxation times $T_1$, $T_2$, proton density $\rho$ and others. Recently in [Ma et al., Nature, 2013], an approach named Magnetic Resonance Fingerprinting (MRF) was introduced, being capable of simultaneously recovering these parameters by using a two step procedure: (i) a series of magnetization maps are created and then (ii) these are matched to parameters with the help of a pre-computed dictionary (Bloch manifold). In this paper, we initially put MRF and its variants in the perspective of optimization and inverse problems, providing some mathematical insights into these methods. Motivated by the fact that the Bloch manifold is non-convex, and the accuracy of the MRF type algorithms is limited by the discretization size of the dictionary, we propose here a novel physically oriented method for qMRI. In contrast to the conventional two step models, our model is dictionary-free and it is described by a single non-linear equation, governed by an operator for which we prove differentiability and other properties. This non-linear equation is efficiently solved via robust Newton type methods. The effectiveness of our method for noisy and undersampled data is shown both analytically and via numerical examples where also improvement over MRF and its variants is observed.

# 1 Introduction

## 1.1 Context

Current routine of magnetic resonance imaging (MRI) examinations typically provides qualitative images of nuclear magnetization of the tissue accompanied by contrast "weights". The physicians then typically inspect visually these images, which however being qualitative, may not provide enough information for a sufficiently accurate diagnosis. The idea of quantitative MRI (qMRI) is to provide additional measurements that can be used as a diagnostic tool. It aims not only to visualize the structure of the imaged object, but also to provide accurate values of parameters that characterize different types of tissues. Typical quantities that are used to discriminate different tissue types are for instance the proton density $\rho$ of Hydrogen atoms in water molecules, the longitudinal and the transverse relaxation parameters $T_1$, $T_2$, and others. The connection between the evolution of the net magnetization (indirect measurement) and the aforementioned magnetic parameters is done via the renowned Bloch equations [4], a system of time dependent ordinary differential equations

$$\begin{aligned} \frac{\partial m}{\partial t}(t) &= m(t) \times \gamma B(t) - \left(\frac{m_x(t)}{T_2}, \frac{m_y(t)}{T_2}, \frac{m_z - m_{eq}}{T_1}\right)^\top \\ m(0) &= m_0. \end{aligned} \tag{1}$$

These equations govern the macroscopic magnetization $m$ of the tissue under external magnetic fields $B$. The relaxation rates $T_1$ and $T_2$ are model parameters associated to the equations. Note that

DOI 10.20347/WIAS.PREPRINT.2528 Berlin 2018

in Section 2.1 we provide a general background to MRI and Bloch equations for the not so familiar readers.

Even though the techniques for qMRI are still in the stage of gestation, many ideas and methods have been proposed. Earlier approaches [20] are based on a set of spin echo or inversion recovery images that are reconstructed form the $k$-space data, with respect to various repetition times (TR) and echo times (TE). There, the acquisitions are designed for each parameter individually. This type of approach is usually called parametric mapping method [22] and it consists of two steps: (i) reconstruct a sequence of images as in qualitative MRI and (ii) for each pixel of the images, fit the intensity of the images to an ansatz curve characterized by the magnetic parameter associated to the tissue imaged at that pixel. Based on this idea, many improvements have been suggested in the literature, see for instance [15]. These types of approaches simplify the physical model and deal with tissue parameters separately, which are considered to be time consuming for the patient.

Another line of research, initiated by Ma et al. in [22] and named Magnetic Resonance Fingerprinting (MRF), has recently gained much attention. The basic idea of MRF is first to build a database (dictionary) consisting of all the trajectories (fingerprints) of the evolution of the magnetization corresponding to some indexed parameter values, typically those of $T_1$ and $T_2$ and also some other parameters e.g. off-resonance frequency. These trajectories are computed by solving the Bloch equations, as many times as the number of all the combination of the parameter values. The values of the parameters are obtained by discretizing their range in a feasible set $C_{ad}$ which corresponds to values that are naturally occurring in the human tissue. In this way, the dictionary establishes a discretized map $\mathrm{Dic}(C_{ad})$ (look-up table) between the all feasible parameters and the corresponding solutions of Bloch equations. Thus, given a magnetization trajectory that is assumed to be a solution of the Bloch equations for some parameters, with the help of this look-up table, one is able to find the parameter values that correspond to this trajectory. This is the main principle behind MRF and the one that allows for the simultaneous estimation of the tissue parameters.

The MRF workflow consists again of two steps. Focusing on a thin slice $\Omega$ of the tissue of interest, the first step is to reconstruct a sequence of $L$ images $\{X^{(\ell)}\}_{\ell=1}^{L}$ from data $\{D^{(\ell)}\}_{\ell=1}^{L}$ as in qualitative MRI, using a rich enough excitation process through $L$ fast radio pulses, see Section 2 for details. At every time step, the data consists of a sub-sampling of the Fourier coefficients of the magnetization. This sub-sampling occurs, because the time between each excitation is small. In a typical MRF routine the magnetization reconstruction process is by using the pseudo-inverse of the Fourier transform, something that leads to the occurence of artifacts in the magnetization images. However, looking at the evolution of the magnetization of a specific tissue element (voxel) along the series of the reconstructed magnetization images, it can be assumed that it should correspond, at least approximately, to the solution of the Bloch equations with parameters that correspond to this specific voxel. Hence, the second step of MRF is the matching of the recorded trajectory of each voxel to a fingerprint in the pre-computed dictionary, again typically through mimimizing a least square distance. The parameter values that correspond to the "best" fingerprint are then assigned to that voxel.

This MRF procedure can be formally described as follows:

$$\text{Solve each } X^{(l)} = \operatorname*{argmin}_{X} \left\| P^{(\ell)}\mathcal{F}X - D^{(\ell)} \right\|_2^2, \quad \ell = 1, \dots, L, \tag{MRF-step 1}$$

$$\text{in order to get } X^* = (X^{(1)}, \dots, X^{(L)})$$

$$\text{Solve } m^* = \operatorname*{argmin}_{m \in \mathrm{Dic}(C_{ad})} S(T_{x,y}m, X^*). \tag{MRF-step 2}$$

A few words about the notation of the above steps are in order. Starting from (MRF-step 1), with $D =$

$(D^{(1)}, \dots, D^{(L)})$ we denote the obtained data after each pulse. Each $D^{(\ell)}$ is a sub-sampling $P^\ell$ of the Fourier coefficients of the magnetization, in fact of its transverse component $T_{x,y}$, i.e., the first two components. The Fourier transform is denoted with $\mathcal{F}$. The first step consist then of $L$ reconstructions of the magnetization of this tissue slice. Note again that this reconstructions are done in a "naive" way, i.e., by considering the pseudo inverse of $\mathcal{F}$

$$X^{(\ell)} = \mathcal{F}^{-1}(P^{(\ell)})^\top D^{(\ell)}.$$

In (MRF-step 2), for every voxel (in practice, for every image pixel $r$) the matching is performed via exhausted search over the dictionary Dic$(C_{ad})$. Note that Dic$(C_{ad}) = \{m^\theta : \ \theta \in C_{ad}\}$, where $\theta$ is a vector of the tissue parameters – for simplicity we can consider $\theta = (T_1, T_2)$ – and $C_{ad}$ is an admissible domain for these parameters. By $m^\theta$ we denote the solution of the Bloch equations with parameter $\theta$, evaluated in the same time steps for which we have constructed the magnetization responses. Hence, every element of Dic$(C_{ad})$ is a vector sequence of length $L$. The function $S(\cdot, \cdot)$ is a Euclidean distance of normalized quantities, in order to avoid the multiplicative effect that the density $\rho$ has in the magnetization. Note that the local density $\rho$ of Hydrogen atoms, multiplies $m$, resulting in the net magnetization $\rho m$. The minimization in (MRF-step 2) is "voxel-wise", i.e., it is done as many times as the number of voxels (in practice pixels). Finally, the parameter spatial maps are formed by assigning the parameter values $\theta$ that correspond to the optimal matches $m^\theta$ to the corresponding voxels. A more detailed description of MRF is done in Section 2.3.

The first numerical results in [22] showed that MRF is a promising approach for qMRI. However, there are several open issues from a mathematical point of view. A basic question is whether two close enough trajectories would infer similar parameter values $\theta$. This refers to some stability analysis for the inversion of the Bloch map $\theta \mapsto m^\theta$. Assuming this stability, and that the number of pulses $L$ is fixed, then there are mainly two factors that influence the accuracy of MRF approaches, namely $(i)$ the quality of the $L$ magnetization reconstructions, and $(ii)$ the completeness (fineness) of the dictionary. In the literature, most of the works focus on improving the former factor, see e.g. [2, 7] while the latter will always impose limitations as it is intrinsically part of a dictionary method.

Moreover, as we have seen, both the parameter mapping method and MRF rely on a two-step strategy, that is first to reconstruct magnetization maps and then to fit the images to the parameters. It is clear that the first step partially determines the quality of the results in the second step. However, it can be observed that the second step procedure can also have a positive influence to the reconstruction in the first step, in particular when the data is under-sampled. For instance, the fact that the evolution of magnetization is dictated by the Bloch equations, can already be taken into account in the reconstruction step and thus improve its quality. In particular, as we will see in the next variant of MRF, it can serve as a constraint to assure uniqueness of solutions in the reconstruction step. Therefore the separation of the procedure into two steps does not seem optimal.

The question that now arises is whether we can estimate the tissue parameters in a single step from the Fourier space data, while in the same time making the modelling more robust. These considerations motivate our work.

An approach towards this direction was proposed by Davies et al., in [7], where a strategy based on a projected Landweber iteration was proposed, and significantly improved the result of MRF. The method was named BLoch response recovery through Iterated Projection (BLIP). An iterative approach was suggested to reconstruct the magnetization, with the key idea being, to project in each iteration the reconstructions to the Bloch manifold (dictionary). Note that this projection is still dictionary based. This leads to an improved solution for the magnetization especially in the case of poorly sub-sampled data. The BLIP algorithm can be roughly understood as a projected gradient descent method for solving the

following constrained optimization problem:

$$\min_{X} \quad \|P\mathcal{F}T_{x,y}X - D\|^2, \qquad \text{(BLIP)}$$
$$\text{subject to } X \in \mathbb{R}^+\text{Dic}(C_{ad}).$$

We refer the reader again to Section 2.3 for a complete description of the BLIP method. Even though, the BLIP method improves the MRF considerably there is still some issues remaining. An important one arises due to the fact that the convexity of the (positive cone of) Bloch manifolds $\mathbb{R}^+\text{Dic}(C_{ad})$, is not guaranteed (see Proposition 4.2), therefore every projection operation to the Bloch manifold may not be well-posed. This might cause serious errors especially in the case when the data are corrupted by noise. In addition, since the projection (matching) is still dictionary based, the method can be memory consuming, especially when the dictionary is highly refined in order to have higher accuracy.

## 1.2 Our contribution

Our work can be divided into two parts. In the first part, Section 3, using some sensitivity analysis we show that the matching process is a well-posed inverse problem when it is restricted to the Bloch manifold. This partially explains why the idea of dictionary in MRF has been successful. In particular in Theorem 3.6, we show that if two trajectories of the magnetization evolution as these are dictated by the Bloch equations, are close enough, the same holds for the parameters $\theta$ that dictate these trajectories. This follows from an estimate of a type

$$\|\theta - \theta^\delta\| \leq C\delta \;\text{ while }\quad \|m - m^\delta\| \leq \delta,$$

where $\theta,\ \theta^\delta$ are the inverted parameters from $m,\ m^\delta$ which are Bloch trajectories. Here, the constant $C$ is independent of $\delta$, and the norms will be clear later in the text. This explains why the idea of a dictionary works in practice. Furthermore, we also provide a mathematical understanding on why a large number of frames $L$, plays a positive role in the quality of the final result. This is discussed in Theorem 4.6.

In the second part of our work, we propose a novel method for qMRI. Our idea is to solve the parameter identification problem directly in the domain of the Bloch manifold, so that we can avoid the (ill-posed) projection as it is done in BLIP from a modelling point of view. With this in mind, we suggest the following single step model:

$$Q(\rho,\theta) := P\mathcal{F}(\rho T_{x,y}M(\theta)) = D, \quad \text{where } (\rho(r),\theta(r)) \in \mathbb{R}^+ \times C_{ad}, \quad \text{for every } r \in \Omega. \qquad (2)$$

The operator $Q$, named qMRI operator, incorporates directly the Bloch dynamics into the data acquisition, and by solving (2) we can recover both $\rho$ and $\theta$. However the non-linearity of $Q$ makes the problem rather challenging as also additional difficulties arise due to the setting of MRI, e.g., sub-sampling and noise. We propose Newton type methods for the solution of (2). In the case of sub-sampling and noise, we consider a variant of Gauss-Newton method, namely a Levenberg-Marquant type regularization approach. In order to do so, we show the necessary differentiability results for the operation that maps $\theta$ to the Bloch trajectories and as a result differentiability of the operator $Q$. As it is the case for many highly non-linear and non-convex problems, the initialization for these types of iterative methods turns out to be crucial. Here we suggest to use BLIP (or MRF) using very coarse dictionary in order to fastly produce a sensible initialization for our algorithm, in which case BLIP is competent. Hence, in the end

we are able to produce more accurate parameter maps with less computational power and time costs than the previous methods.

Note that a similar approach is followed in two recent papers [26] and [27] where also a single step model and a dictionary free method are considered respectively. In particular, the model in [27] abandons the Fourier space character of the data, and asks for relatively big amount of data frames to solve a very large non-linear system. As a result, relatively large memory and computational power is needed. The work of [26] mainly takes the experiment design point of view, aiming at optimizing the excitation pulse sequences as well as the repetition times.

#### Structure of the paper

The rest of the paper is planned as follows: In Section 2, we provide a general background to MRI, particularly to Bloch equations and MRF. In Section 3, we put the MRF-type algorithms in the perspective of inverse problems, where we perform a stability analysis for inversion of the Bloch mapping. In Section 4, we propose our new physically based, single step model for qMRI, which is described by a nonlinear operator equation. We analyse the differentiability of this operator, and prove the non-convexity of the Bloch manifold. Subsequently, we discuss a series of Newton type methods for its numerical solutions. We particularly focus on the case of undersampled and noisy data. Numerical tests and comparisons are presented in Section 5 to illustrate the efficiency of the proposed method for qMRI. A short description on solutions of Bloch equations in different cases is given in the Appendix.

# 2 Background on MRI and MRF

We provide here a brief, high level summary of the MRI principles. For more details, we refer the readers to [30].

## 2.1 Bloch equations

The Bloch Equations [4] characterize the key physical principles in nuclear magnetic resonance. We will discuss the Bloch equations in their classical form in the continuous time dependent case. We first describe our set up.

We are focusing on a thin slice of tissue which is modelled by a domain $\Omega \subseteq \mathbb{R}^2$. We denote every element point (or voxel) of $\Omega$ with $r$. The main principles of MRI lie on the interaction between an externally applied dynamic magnetic field $B = (B_x, B_y, B_z)^\top$ and the (net or bulk) magnetization which is equal to all the individual dipole moments of the proton spins within a voxel. This net magnetization is proportional to the hydrogen proton density $\rho$. We denote by $m = (m_x, m_y, m_z)^\top$ the magnetization per unit density element, hence the net magnetization in a voxel of density $\rho$ is equal to $\rho m$.

In the case of a static magnetic field $B_0$, which is typically regarded to lie on the $z$-direction, the net magnetization is aligned to that field, of with its longitudinal component $m_z$ reaching an equilibrium $m_{eq}$. This alignment is not achieved instantaneously but it is controlled by the *longitudinal relaxation time* $T_1$ (or $T_1(r)$ if we want to emphasize the dependence on a specific voxel), and the longitudinal magnetization evolves according to the following formula

$$m_z(t) = m_{eq}(1 - e^{-(t/T_1)}).$$

Furthermore, the part of the magnetization orthogonal to $B_0$, which is called the transverse magnetization $(m_x, m_y)^\top$, precesses about the $z$-axis at a frequency equal to $\gamma|B_0|$. The constant $\gamma$ is called the gyromagnetic ratio. This precession emits an electromagnetic signal which can be detected and measured by the coils of the MRI machine. The transverse magnetization decays also exponentially at a rate $T_2$, the *transverse relaxation time*.

The overall macroscopic dynamics that dictate the relation between the magnetization $m$, the magnetic field $B$ and the relaxation times $T_1$, $T_2$, are governed by the Bloch equations, which is a system of time dependent linear ODEs:

$$\begin{array}{rcl} \frac{\partial m(t,r)}{\partial t} & = & m(t,r) \times \gamma B(t,r) - \Theta(r) \bullet (m(t,r) - m_e), \\ m(0,r) & = & m_0(r). \end{array} \tag{3}$$

where $m_e = (0,0,m_{eq})^\top$ and "$\times$" denotes the outer product between vectors. Also, in order to simplify the presentation, we have introduce the variable $\Theta$

$$\Theta(r) = (\Theta_1(r), \Theta_2(r), \Theta_3(r))^\top := \left(1/T_2(r), 1/T_2(r), 1/T_1(r)\right)^\top,$$

The operation $\bullet$ in (3) denotes component-wise multiplication between vectors. As we have noted, here $m : (0,\tau) \times \Omega \to \mathbb{R}^3$, for some $\tau > 0$, denotes the magnetization in a unit volume per unit proton density. Notice that the dependence on $r$, is intrinsic here and does not enter in the dynamics of the equation. Furthermore, $m_e$ denotes the equilibrium state of the magnetization, which without loss of generality we can consider it to be unitary in magnitude, i.e. $m_e \equiv (0,0,1)^\top$. Since the equation is linear, in order to get the net magnetization, one simply has to multiply it by $\rho$.

The total magnetic field $B(t,r)$ can be typically decomposed as follows

$$B(t,r) = B_0(r) + B_1(t,r) + (0,0,G(t)\cdot r)^\top. \tag{4}$$

Here $B_0$ denotes the external constant magnetic field that points to the positive $z$ direction and in general assumed to be spatially homogeneous. However, we keep here the dependence on $r$ for generality. The component $B_1(t,r) = (B_{1,x}(t,r), B_{1,y}(t,r), 0)^\top$ denotes a radio frequency (RF) pulse, which is sent periodically and only lasts for a very short time. It is used in order to excite the magnetization from its equilibrium, by turning the magnetization precession away from the direction of the main magnetic field with an angle

$$\alpha(t) = \gamma \int_0^t |B_1(s)|\, ds,$$

which is called flip angle. These pulses usually last only very shortly in comparison with $T_1$ and $T_2$. Therefore RF sequences can be completely characterized by sequences of flip angles, and the time is normally omitted. The interval between two consecutive pulses is called repetition time (TR). As we will see later in the Section 4, we will consider a specific flip angle sequence pattern called *Inversion Recovery balanced Steady State Free Precession* (IR-bSSFP) [28]. By using these sequences the solution of the Bloch equations can be simulated by a discrete linear dynamical system, see Section 4.1. In the appendix, we provide some case discussions which help to understand the discrete Bloch dynamics and the solution of Bloch equations.

The component $G(t)$ in (4) is a magnetic gradient field that is designed to distinguish the point-wise information from the detected signal. Without getting into too much details, the measured signal can be described as

$$S(t) = \int_\Omega \rho(x,y)(T_{x,y} m(t,x,y)) e^{-\mathrm{i}\gamma|B_0|t} e^{-\mathrm{i}\gamma \int_0^t (xG_x + yG_y)d\tau} dxdy,$$

where we have introduced the notation $T_{x,y}m := m_x(x,y) + \mathrm{i}m_y(x,y)$ for the transverse magnetization, $\mathrm{i}$ being the imaginary unit. This type of presentation is a convention in the MRI community, but one can also think of $T_{x,y}m$ as a pair of real valued components. The third component of $m$ usually can not be measured because of the position of the coils. The MRI signal $D^{(t)}$can be mathematically modelled as some collected coefficients of a Fourier transform of the transverse magnetization up to a demodulation by $e^{\mathrm{i}\gamma|B_0|t}$. That is

$$P^{(t)}\mathcal{F}(\rho T_{x,y}m^{(t)}) = D^{(t)},$$

where $\mathcal{F}$ denotes the Fourier transform, and $P^{(t)}$ means a sub-sampling operator.

## 2.2 Sub-sampling

The idea to not fully sample the $k$-space in order to recover the signal, lies on the core of rapid MRI imaging, and also in MRF since, as we will see later, one does not wait for the signal to return to equilibrium and due to time constraints only a small proportion of the $k$-space is sampled between every excitation pulse. Reconstruction of the magnetization from sub-sampled measurements leads to the presence of aliasing artifacts especially when this reconstruction uses the basic (but fast) approach of pseudo inverse of $\mathcal{F}$.

In the context of MRF, three different types of sub-sampling are typically employed, namely, the spiral, the radial and the Cartesian sub-sampling pattern. Each of these, corresponds to a different variation in time, of the selection gradients $G_x$ and $G_y$. In the original version of MRF, the former two methods were preferred since the aliasing artifacts of these two patterns appear to be uncorrelated and can be roughly treated as random noise, something which is not the case with the Cartesian sub-sampling pattern, see for instance the corresponding numerical examples in [7]. However in the same paper, it was shown that the BLIP method, reviewed in the next section, perfectly fits in the framework of Cartesian sub-sampling, while in the same time, it improves the results of MRF. In this paper, since our starting point is the BLIP algorithm, we will also focus on Cartesian pattern for sub-sampling based on multishot echo-planar imaging (EPI) [23], see Section 5.2 for details. However, this does not limit our proposed method, and other sub-sampling patterns can be used.

## 2.3 Description of MRF and BLIP

In the section we describe the workflow of the MRF process. It consists of three, discrete steps, the construction of the dictionary, the reconstruction of the magnetizations and finally the matching to the magnetizations to the dictionary.

Initially, one considers a pre-designed excitation pattern that consists of $L$ flip angles $\{\alpha_\ell\}_{\ell=1}^{L}$ which are separated by a repetition time TR. Here, for simplicity, we consider TR to be constant but this is not necessary. Also a subset $C_{ad} \subset \mathbb{R}^m$ of the space of tissue parameters we want to estimate is predefined. For simplicity here we consider $C_{ad}$ to be the set of all the admissible $\theta = (T_1, T_2)$ values, i.e., $m = 2$, but more parameters can also be taken into account. The set $C_{ad}$ contains all values of $\theta$ that are expected to be found in the human body. For example values for $T_1$ would typically range from $685ms$ (white matter on brain) to $4880ms$ (cerebrospinal fluid), with the corresponding range for $T_2$ to be 65–550 ms [22]. As we will see later on, in our dictionary-free approach we will consider $C_{ad}$ to be simply a convex subset of $\mathbb{R}^+ \times \mathbb{R}^+$, in particular a box, allowing thus for all the values

between a minimum and a maximum value. In the context of dictionary method as it is the case in the original MRF, the set $C_{ad}$ is discretized, using fine enough discretization and $K$ pairs of parameters are selected, i.e., $\{\theta_k\}_{k=1}^K$. For this section we will use the same notation for $C_{ad}$ and its discretization, i.e., here $C_{ad} = \{\theta_k\}_{k=1}^K$ but later we will consider $C_{ad}$ to be the whole convex set. Using this set of parameters $\{\theta_k\}_{k=1}^K$, the specific excitation pattern, the sequence of flip angles $\{\alpha_\ell\}_{\ell=1}^L$ and the repetition time TR, one can simulate the Bloch equations by using a discrete linear dynamic system. The solutions of the Bloch equations are evaluated in discrete times $t_1, t_2, \ldots, t_L$, see Section 4.1 for details. This generates a dictionary $\mathrm{Dic}(C_{ad})$ of $K$ magnetization responses $\{m^{\theta_k}\}_{k=1}^K$ which can be seen as $K$ sequences of length $L$ with components in $\mathbb{R}^3$, that is $\mathrm{Dic}(C_{ad}) \subset \left((\mathbb{R}^3)^L\right)^K$. We name this creation of the dictionary as the *Step 0* of the MRF process:

**Step 0 of the MRF process: Construction of the dictionary**
*Given an excitation pattern of flip angles $\{\alpha_\ell\}_{\ell=1}^L$, repetition time $TR$, and a set of parameters $C_{ad} = \{\theta_k\}_{k=1}^K$, generate the dictionary $\mathrm{Dic}(C_{ad})$ of $K$ trajectories of the solutions of Bloch equations, evaluated at times $t_1,\ t_2, \ldots, t_L$*

$$\mathrm{Dic}(C_{ad}) = \{m^{\theta_k} : \theta_k \in C_{ad},\ k = 1, \ldots, K\} \subset \left(\left(\mathbb{R}^3\right)^L\right)^K$$

In the next phase, the MR data are collected at the respective $L$ read out times. Each component $D^{(\ell)}$ of the data $D = (D^{(1)}, \ldots, D^{(L)})$, corresponds, due to time constraints, to only a sub-sampling $P^{(\ell)}$ of the Fourier coefficients of the net magnetization $X^{(\ell)}$. Here the reconstruction of the transverse magnetization image is done via the least square solution and hence these images suffer from aliasing artifacts. So this step consists of solving $L$ least square solutions (using the pseudo-inverse Fourier transform $\mathcal{F}^{-1}(P^{(\ell)})^\top$) to obtain $X^* = (X^{(1)}, \ldots, X^{(L)})$, where $X^{(\ell)} : \Omega \to \mathbb{R}^2$. Note that, instead of $\mathbb{R}^2$, one can use the complex number representation for the reconstructed magnetization $X$ and the Bloch response $m$, that is $m = m_x + \mathrm{i}m_y$. Note that in this section $\Omega$ denotes a set of discrete voxels, which in practice are represented by pixels $i : 1, \ldots, N$. Summarizing

**Step 1 of the MRF process: Reconstruction of the magnetizations**
*Reconstruct the vector of $L$ net magnetizations $X^* = (X^{(1)}, \ldots, X^{(L)})$ by solving*

$$X^{(\ell)} \in \underset{X:\Omega\to\mathbb{R}^2}{\operatorname{argmin}} \|P^{(\ell)}\mathcal{F}X - D^{(\ell)}\|_2^2 \quad \textit{using}\ \ X^{(\ell)} = \mathcal{F}^{-1}(P^{(\ell)})^\top D^{(\ell)}, \quad \ell = 1, \ldots, L$$

The final step of the MRF, is to match the reconstructed magnetization of every voxel, i.e., a sequence of length $L$, to the transverse components of the element $m^{\theta_k}$ (denoted by $T_{x,y}m^{\theta_k}$) in the dictionary $\mathrm{Dic}(C_{ad})$, which is more likely to correspond to. Then the required parameter map $\theta : \Omega \to \mathbb{R}^2$, is formed by mapping every voxel $i$ to the $\theta$ value that corresponds to the matched $m^\theta$. Note that each reconstructed magnetization sequence of a $i$-th voxel $(X_i^\ell)_{\ell=1}^L$, has the contribution of the density $\rho_i$ of this particular tissue element. In order to compare with the candidate fingerprint sequence, minimal Euclidean distance is used to select the best fingerprint, which can be seen as an $\ell^2$ projection onto the discrete Bloch manifold. If each fingerprint is firstly normalized, i.e., $\tilde{m}^{\theta_k} = \frac{T_{x,y}m^{\theta_k}}{\|T_{x,y}m^{\theta_k}\|_2}$ and then this projection can be computed as a maximal inner product. Finally the density $\rho_i$ is computed by $\rho_i = \frac{\|X_i\|_2}{\|T_{x,y}m^{\theta_{k_i}}\|_2}$. Summarizing again

**Step 2 of the MRF process: Matching of the magnetizations to the dictionary**

*For every voxel $i = 1, \dots, N$, compute the projection of the magnetization $X_i = (X_i^\ell)_{\ell=1}^L$ to the discrete Bloch manifold $\mathrm{Dic}(C_{ad})$ by*

$$m^{\theta_{k_i}} = \operatorname*{argmin}_{m^\theta \in \mathrm{Dic}(C_{ad})} \left\| \frac{T_{x,y} m^\theta}{\|T_{x,y} m^\theta\|_2} - X_i \right\|_2^2,$$

*and use a look-up table to obtain the final parameter map $\{\theta_{k_i}\}_{i=1}^N = \{(T_1(i), T_2(i))\}_{i=1}^N$. Finally, compute the density map $\{\rho_i\}_{i=1}^N$ as*

$$\rho_i = \frac{\|X_i\|_2}{\|T_{x,y} m^{\theta_{k_i}}\|_2}.$$

One can observe in step 1, that the minimization problem $\min_X \left\|P^{(\ell)}\mathcal{F}X - D^{(\ell)}\right\|_2^2$ has very likely non-unique minimizers due to the sub-sampling operator $P^\ell$. In the original paper [22], the specific minimizer $X^{(\ell)} = \mathcal{F}^{-1}((P^{(\ell)})^\top D^{(\ell)})$ was chosen. However, this by no means, mean that this choice is optimal, see corresponding examples in [7].

We now introduce another approach similar to MRF, named *BLoch response recovery through Iterated Projection* (BLIP). This is of iterative nature, and it was recently proposed in [7]. We state the algorithm straight away and then we discuss it

**Algorithm 2.1: BLIP algorithm [7]**

- Generate the dictionary $\mathrm{Dic}(C_{ad})$ as in the Step 0 of MRF.
- Initialize the magnetization vector $X = 0$ and choose an initial step size $\mu_1$.
- Start an iterative procedure for $n = 1, 2, 3, \dots$ as follows:
  - ⊙ Do the following gradient decent step:

$$\left(X^{(\ell)}\right)_{n+1} = \left(X^{(\ell)}\right)_n - \mu_n \mathcal{F}^{-1}(P^{(\ell)})^\top (P^{(\ell)}\mathcal{F}\left(X^{(\ell)}\right)_n - D^{(\ell)}), \quad \ell = 1, \dots, L.$$

  - ⊙ Project each $(X_i)_{n+1} = \left(\left(X_i^{(\ell)}\right)_{\ell=1}^L\right)_{n+1}$ onto the dictionary $\mathrm{Dic}(C_{ad})$ as in the Step 2 of MRF, obtaining

$$\left(m^{\theta_{k_i}}\right)_{n+1} = \left(\left((m^{\theta_{k_i}})^{(\ell)}\right)_{\ell=1}^L\right)_{n+1} \text{ and } (\rho_i)_{n+1}$$

    for every voxel $i, = 1, \dots, N$.

  - ⊙ For every $\ell = 1, \dots, L$, update $\left(X^{(\ell)}\right)_{n+1}$ as follows

$$\left(X_i^{(\ell)}\right)_{n+1} \leftarrow (\rho_i)_{n+1} \left((T_{x,y} m^{\theta_{k_i}})^{(\ell)}\right)_{n+1}, \quad i = 1, \dots, N.$$

  - ⊙ Update the step size $\mu_n$ (see [7] for some rules).
- Construct parameter maps as in MRF, using look-up tables and the last iterates from above.

As we have mentioned in the introduction, BLIP approximates the solution of the following constrained minimization problem

$$\min_X \quad \|P\mathcal{F}T_{x,y}X - D\|_2^2, \qquad \text{(BLIP)}$$
$$\text{subject to } X \in \mathbb{R}^+\mathrm{Dic}(C_{ad}).$$

through a projected gradient descent iteration. The difference to the original MRF is that the constraint in (BLIP) enforces the solutions to belong to the scaled dictionary $\mathbb{R}^+\text{Dic}(C_{ad})$. Here we do not have the typical two step procedure of MRF, but rather the two steps, i.e., reconstruction and matching, are fused and benefit from each other. In [7], it was shown that particularly for the cartesian sub-samping that we focus here, the BLIP algorithm produces results of higher quality than the classical MRF.

From a mathematical point of view, there are several issues that remain for both methods above and also other MRF type methods. First we would like to know the how the discretization size of the dictionary would influence the matching of the parameters, which asks for a stability analysis for inverting the Bloch mapping. Assuming this stability, the accuracy of the above MRF-type algorithms is influenced by several factors, such as the accuracy of the reconstructed image from the Fourier space data, the fineness of the generated dictionary as well as the number of frames $L$. The BLIP algorithm greatly improves the accuracy of the reconstructed image by enforcing explicitly the constraint to the Bloch manifold as stated in the optimization model (BLIP). However, the Bloch manifolds constraint is not convex, see Proposition 4.2, making this projection step problematic. Moreover, the accuracy of the final solutions of these methods are unavoidably limited by the fineness of the dictionary.

Before we proceed with the introduction of our proposed method, we will give a mathematical insight to the dictionary matching procedure from inverse problems point of view.

# 3 MRF: Inverse Problems Point of View and Stability Analysis

## 3.1 Coupled inverse problems point of view

Here we will use the time continuous version for the Bloch equations by adopting a more general framework.

We start by describing our set up. We denote

$$\mathcal{Y} := [L^1(\Omega)]^3 \quad \text{and} \quad \mathcal{Z} := [L^\infty(\Omega)]^3.$$

We consider $m_0 \in \mathcal{Y}$ to be an initial vector of magnetization and $B \in L^\infty(0,\tau;\mathcal{Z})$ an external magnetic field for some $\tau > 0$. Recall that

$$L^\infty(0,\tau;\mathcal{Z}) = \{f : [0.\tau] \to \mathcal{Z} : \ \|f\|_{L^\infty(0,\tau;\mathcal{Z})} < \infty\} \quad \text{where} \quad \|f\|_{L^\infty(0,\tau;\mathcal{Z})} = \operatorname*{ess\,sup}_{0\le t\le\tau} \|f(t)\|_{\mathcal{Z}}.$$

Note that the space $L^1(0,\tau;\mathcal{Y})$ is defined similarly, as well as the space $W^{1,1}(0,\tau;\mathcal{Y})$. The latter space consists of all the functions $f : [0,\tau] \to \mathcal{Y}$ such that both $f$ and $\frac{\partial f}{\partial t}$ belong to $L^1(0,\tau;\mathcal{Y})$. We refer to [9] for this type of Bochner spaces.

For the parameter $\theta = (T_1, T_2)$ we have that $\theta \in [L^\infty(\Omega)]^2$, and we also consider this parameter to be bounded away from zero. As a result $\Theta = (1/T_2, 1/T_2, 1/T_1)^\top \in [L^\infty(\Omega)]^3$. Finally keep in mind that the equilibrium vector $m_e$ is constant and equal to $(0,0,1)^\top$. Recalling again the Bloch equations

$$\begin{array}{rcl} \frac{\partial m(t,r)}{\partial t} & = & m(t,r) \times \gamma B(t,r) - \Theta(r) \bullet (m(t,r) - m_e), \\ m(0,r) & = & m_0(r), \end{array} \tag{5}$$

It is convenient here to introduce the operator

$$\mathcal{B}_{m_0,B} : [L^\infty(\Omega)]^2 \to \{m : [0,\tau] \to \mathcal{Y}\},$$

where $\mathcal{B}_{m_0,B}(\theta)$ denotes the solution mapping of the Bloch equation up to time $\tau$.

With this notation, we can now write down the following family of inverse problems which can be thought of as a continuous version of the MRF process:

- Problem 1: For some $t_\ell \in (0,\tau)$, $\ell = 1,\dots,L$, solve for $X^{(t_\ell)} \in L^2(\Omega)$ the following linear equation
$$P^{(t_\ell)}\mathcal{F}X^{(t_\ell)} = D^{(t_\ell)}, \tag{6}$$
where $D^{(t_\ell)} \in [L^2(K)]^2$, $K$ being the $k$-space, $\mathcal{F} : [L^1(\Omega)]^2 \to [L^2(K)]^2$, $P^{(t_\ell)} : [L^2(K)]^2 \to [L^2(K)]^2$

- Problem 2: For every $r \in \Omega$, solve for $\theta = \theta(r) \in \mathbb{R}^+ \times \mathbb{R}^+$
$$\rho(r)T_{x,y}(\mathcal{B}_{m_0,B}(\theta))(\cdot, r) = X^{(\cdot)}(r), \tag{7}$$

where $\rho \in L^\infty(\Omega)$, and $T_{x,y}$ is the transverse projection. Note that, strictly speaking, the coupling of (6) and (7) makes sense only when $P^{(t_\ell)} = id$, i.e., there is no sub-sampling. This is because in the case of sub-sampling, uniqueness of solutions for (6) is not guaranteed, and $X^{(\cdot)}(r)$ may not belong to the Bloch manifold.

The first problem that corresponds to the first step in MRF, is the inversion of the Fourier transform for sub-sampled (and potentially noisy) data. This type of problems is the key mathematical problem in standard MRI, and as a result it has been extensively studied for a long time. In particular, sparse regularization methods, optimal weighted total variational methods have been successfully applied towards that [12, 14, 18, 21].

The second step in MRF is a parameter identification inverse problem involving the Bloch equations. From a rigorous mathematical point of view, a major issue concerning such an inverse problem is to study existence, uniqueness and stability of solutions. Especially, the stability gives a theoretical judgement on the feasibility of dictionary methods. Up to our best knowledge, these problems related to Bloch equations have not yet been carefully analysed in the literature. Due to its fundamental importance, we discuss it next.

In order to simplify the upcoming sensitivity analysis, we will focus on solving equation (7), without the effect of the density map $\rho$ and the transverse projection operator $T_{x,y}$. That means, in what follows, we analyse the stability of the inversion of the Bloch mapping by considering the following equation

$$\mathcal{B}_{m_0,B}(\theta) = m. \tag{8}$$

## 3.2 Stability analysis on inverting Bloch mapping

In this section, we focus on analysing the stability of problem (8). This is interesting to us since one would want to know how the deviations in the solutions of problem (6) will be propagated in the process of solving problem (7). The conclusion will serve as a mathematical basis for the approach MRF and BLIP algorithms.

In order to simplify the discussion, in this section we consider the time domain $(0,\tau)$, to be the time between each paired consecutive pulses. From a modelling point of view, $m_0$ will be the magnetization right after the first pulse, i.e., after the application of the flip angle displacement and, $m(\tau,\cdot)$ will be the magnetization right before the next pulse. In that case the magnetic field $B \neq 0$ is considered to

time independent which means that, $B$ is a constant function in $L^\infty(0,\tau;\mathcal{Y})$. Also the effect of the gradient field $G$ is ignored here, as it only encodes the MRI signal.

Given the setting introduced to the previous section and using a classical result for evolutionary equations in Banach spaces [3], we get that the solution of (5) $m$ belongs to the space $W^{1,1}(0,\tau;\mathcal{Y})$. In fact even more regularity holds, but this is enough for our purposes. Hence here,

$$\mathcal{B}_{m_0,B} : [L^\infty(\Omega)]^2 \to W^{1,1}(0,\tau;\mathcal{Y}),$$

Note that $m_0$ and $B$ are fixed here. To implement the analysis, we ask for the following assumptions on the magnetization $m$ and the time $\tau$, and also a condition for the feasible domain of the parameter $\theta$.

**Assumption 3.1.** *To be practical, we shall consider $\theta(r) \in C_{ad}$ for all $r \in \Omega$, where $C_{ad} \subset \mathbb{R}^+ \times \mathbb{R}^+$ denotes a feasible domain which is convex and bounded away from zero.*

We finally denote

$$R(\mathcal{B}_{m_0,B}) = \left\{ m : \ \exists\, \theta \in [L^\infty(\Omega)]^2 \text{ such that } \theta(r) \in C_{ad} \text{ for all } r \in \Omega \text{ and } m = \mathcal{B}_{m_0,B}(\theta) \right\},$$

to be the range of the Bloch mapping.

**Assumption 3.2.** *Let*

$$m \in R(\mathcal{B}_{m_0,B}) \subset W^{1,1}(0,\tau;\mathcal{Y}), \tag{9}$$

*be a solution of the Bloch equations* (5)*. Then the quantity $(\omega_\tau^1(r), \omega_\tau^2(r), \omega_\tau^3(r))^\top := \int_0^\tau m(t,r)dt - m_e\tau$ is bounded away from zero, i.e., there is a constant $c_\tau > 0$ such that*

$$\inf_{r\in\Omega} \left|\omega_\tau^i(\tau)\right| \geq c_\tau, \quad \text{for } i = 1,2,3. \tag{10}$$

**Remark 3.3.** *The Assumption 3.2 is in practice justified, since we consider $(0,\tau)$ as the time between two consecutive pulses which roughly equals to the repetition time. In this period, the net magnetization always satisfies $m_x > c'_\tau > 0$, $m_y > c''_\tau > 0$, and $m_z < m_e$, and these give the estimate* (10)*. Since in a MRI experiment, the time domain consists of repetition of periodic radio pulses, (the excitation time of the pulse is usually very short) then one can think the assumption will always hold in the whole experiment.*

**Theorem 3.4.** *Let Assumption 3.2 hold, the magnetic field $B \neq 0$, and let $m \in R(\mathcal{B}_{m_0,B})$ for some $\theta$. Then this $\theta$ is unique.*

*Proof.* Observe that by integrating the Bloch equations over the time domain $(0,\tau)$, we have

$$\Theta(r) = \frac{\left(m(0,r) - m(\tau,r) + \int_0^\tau m(t,r) \times \gamma B(t,r)dt\right)}{\int_0^\tau (m(t,r) - m_e(r))dt}, \tag{11}$$

where the dividing operation is defined component-wisely for vector valued functions. Note that the integrals above are well-defined, since for almost every $r$, $m(r,\cdot) \in L^1((0,\tau))$. Also due to Assumption 3.2 the denominator in (11) is not zero. The uniqueness of $\theta = (\frac{1}{\Theta_3}, \frac{1}{\Theta_1})^\top$ follows then directly.

□

We immediately have the next corollary:

**Corollary 3.4.1.** *Let Assumption 3.2 hold, $B \neq 0$. Then the Bloch mappingwe have*

$$\mathcal{B}_{m_0,B}(\theta_1) = \mathcal{B}_{m_0,B}(\theta_2) \iff \theta_1 = \theta_2$$

*Proof.* For every fixed $\theta(r) \in C_{ad}$, the Bloch mapping is well-defined under Assumption 3.2. The mapping is injective by Theorem 3.4. The other direction follows from Picard–Lindelöf Theorem (see e.g. [29]). □

**Remark 3.5.** *The uniqueness result indicates that to enforce the magnetization function to be in the range of a Bloch mapping will also enforce a unique parameter. This explains the idea of BLIP algorithm which helps to have a better solution from undersampled data by using projection steps to the Bloch manifold.*

We now proceed with the main stability result of this section.

**Theorem 3.6.** *Let the Assumption 3.1 be satisfied, and let $m, m^\delta \in R(\mathcal{B}_{m_0,B})$ with corresponding parameters $\theta, \theta^\delta$. If Assumption 3.2 holds for both $m$ and $m^\delta$, and*

$$\left\|m - m^\delta\right\|_{W^{1,1}(0,\tau;\mathcal{Y})} \leq \delta,$$

*where $\delta > 0$, then we have the estimate*

$$\left\|\theta - \theta^\delta\right\|_{[L^1(\Omega)]^2} \leq C(\tau, \theta, B)\delta,$$

*where $C(\tau, \theta, B)$ is a constant dependent on $\tau$, $\theta$ and $B$, but independent of $\delta$.*

*Proof.* Using equation (11), we have

$$\begin{aligned}
\Theta - \Theta^\delta &= \frac{\left(m_0 - m(\tau,r) + \int_0^\tau m(t,r) \times \gamma B(t,r) dt\right)}{\int_0^\tau (m(t,r) - m_e(r)) dt} \\
&\quad - \frac{\left(m_0^\delta - m^\delta(\tau,r) + \int_0^\tau m^\delta(t,r) \times \gamma B(t,r) dt\right)}{\int_0^\tau (m^\delta(t,r) - m_e^\delta(r)) dt} \\
&= \Theta \bullet \frac{\int_0^\tau m^\delta(t,r) - m_e^\delta(r) - (m(t,r) - m_e(r)) dt}{\int_0^\tau (m^\delta(t,r) - m_e^\delta(r)) dt} \\
&\quad - \frac{m_0^\delta - m_0 - m^\delta(\tau,r) + m(\tau,r)}{\int_0^\tau (m^\delta(t,r) - m_e^\delta(r)) dt} \\
&\quad - \frac{\int_0^\tau m^\delta(t,r) \times \gamma B(t,r) - m(t,r) \times \gamma B(t,r) dt}{\int_0^\tau (m^\delta(t,r) - m_e^\delta(r)) dt}.
\end{aligned}$$

Note that $\int_0^\tau m^\delta(t,r) - m_e^\delta(r) dt = \int_0^\tau m^\delta(t,r) dt - m_e^\delta \tau$, and $m_e^\delta = m_e$. Considering the condition given in (10), we have

$$\inf_{r\in\Omega} \left| \int_0^\tau m(t,r) - m_e(r) dt \right| \geq c_\tau, \quad \text{and } \inf_{r\in\Omega} \left| \int_0^\tau m^\delta(t,r) - m_e(r) dt \right| \geq c_\tau.$$

Then, we have the estimate for generic constants $C$ that depend on $B$ and $\Theta$

$$\begin{aligned}\left\|\Theta-\Theta^{\delta}\right\|_{[L^1(\Omega)]^3} &\leq \frac{1}{c_\tau}\int\left|\Theta(r)\bullet\int_0^\tau (m^\delta(t,r)-m(t,r))dt\right|dr\\ &+\frac{1}{c_\tau}\int\left|\int_0^\tau\left(\frac{\partial m(t,r)}{\partial t}-\frac{\partial m^\delta(t,r)}{\partial t}\right)dt\right|dr\\ &+\frac{1}{c_\tau}\int\left|\int_0^\tau (m^\delta(t,r)-m(t,r))\times\gamma B(t,r)dt\right|dr\\ &\leq \frac{1}{c_\tau}C\|m^\delta-m\|_{L^1(0,\tau;\mathcal{Y})}\\ &+\frac{1}{c_\tau}\left\|\frac{\partial m^\delta}{\partial t}-\frac{\partial m}{\partial t}\right\|_{L^1(0,\tau;\mathcal{Y})}\\ &+\frac{1}{c_\tau}C\|m^\delta-m\|_{L^1(0,\tau;\mathcal{Y})},\end{aligned}$$

where we have use the fact that $\Theta\in[L^\infty(\Omega)]^2$ and the fact the outer product with $B(t,r)$ can can be written as a linear operator with bounded norm, both in $t$ and $r$, due to the fact that $B$ is bounded in $L^\infty(0,\tau;\mathcal{Z})$ and also independent of time.

Finally note that, as $\theta_1=\frac{1}{\Theta_3}$ and $\theta_2=\frac{1}{\Theta_1}$ and by Assumption 3.1, we can find a constant $C_1>0$ such that

$$\left\|\theta-\theta^{\delta}\right\|_{[L^1(\Omega)]^2}\leq C_1\left\|\Theta-\Theta^{\delta}\right\|_{[L^1(\Omega)]^3}.$$

This follows from the fact that the function $h:[a,b]\to\mathbb{R}$, with $h(\beta)=1/\beta$, is Lipschitz in the domain $0<a<b<\infty$. The proof is complete by considering the last two inequalities.

□

The above result can be interpreted as follows. Theorem 3.6 shows that the inverse problem (8) is well-posed by restricting the right hand side to the range of the Bloch mapping. That is, if the reconstructed magnetization is in the (scaled) Bloch manifold, then the values of the tissue parameters $\theta$ recovered from the dictionary, should be in principle not too far away from the exact solutions. It also explains the improvement of BLIP over the classical MRF scheme, where in every iteration, a projection onto the Bloch manifold is performed.

The analytical property of the Bloch mapping and its inverse not only support the feasibility of the MRF type schemes, but also motivate us to find more accurate techniques for solving the quantitative MRI problem. We discuss this in the following section where we also introduce our proposed method.

# 4 A Physically Oriented Method for qMRI

We propose now a physically oriented, single step model for qMRI, instead of the previously discussed two step procedures. Without getting into too many details, we first state that our model is summarized by the following non-linear operator equation:

$$Q(\mathbf{x})=D, \tag{12}$$

where $\mathbf{x}(r)=(\rho(r),\theta(r))$, $D$ is the acquired MRI signal and $Q$ is an operator which is defined by

$$Q(\mathbf{x}):=P\mathcal{F}(\rho T_{x,y}M(\theta)) \text{ and } (\rho(r),\theta(r))\in\tilde{C}_{ad}:=\mathbb{R}^+\times C_{ad} \quad \text{for all } r\in\Omega. \tag{13}$$

We call this operator $Q$, the qMRI operator. It integrates the Bloch mapping within the data acquisition procedure. Note that here we use $M(\theta)$ to denote discrete Bloch dynamics which will be discussed below. That is, in this section, we return to the same time discrete version of MRF process.

## 4.1 Bloch mapping as discrete dynamics

We shall focus here on a specific type of discrete Bloch dynamics named Inversion Recovery balanced Steady-State Free Precession (*IR-bSSFP*) [28]. This is a specific MRI excitation pulse set-up widely used in applications and it allows for a simple approximation of the solutions of the Bloch equations at the read out times. From now on, in analysis and in the numerical examples, we will always use this type of dynamics instead of the continuous Bloch equations.

To simplify the presentation, we will ignore the factor of off resonance and only consider the homogeneous case of the flip angles and off resonance frequency. In this case, the magnetization after each $n$-th excitation pulse, at the middle of each $TR_\ell$ time, is simulated by the following recursion formula

$$\left\{\begin{array}{lcl} M_\ell &=& E_1(TR_\ell,\theta)R_\phi R_x(\alpha_\ell)R_\phi^\top M_{\ell-1} + E_2(TR_\ell,\theta)M_e, \\ M_e &=& (0,0,1)^\top, \\ M_0 &=& -M_e = (0,0,-1)^\top. \end{array}\right. \tag{14}$$

Here $\{\alpha_\ell\}_{\ell=1}^L$ and $\{TR_\ell\}_{\ell=1}^L$, are the flip angles and repetition time sequences, and $\{M_\ell\}_{\ell=1}^L$ are the magnetizations at the middle of each $TR_\ell$ time. Moreover we denote

$$E_1(TR_\ell,\theta) = \begin{pmatrix} e^{-\frac{TR_\ell}{T_2}} & 0 & 0 \\ 0 & e^{-\frac{TR_\ell}{T_2}} & 0 \\ 0 & 0 & e^{-\frac{TR_\ell}{T_1}} \end{pmatrix}, \qquad E_2(TR_\ell,\theta) = \left(1 - e^{-\frac{TR_\ell}{T_1}}\right)$$

and also

$$R_\phi = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \quad R_x(\alpha_\ell) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha_\ell) & \sin(\alpha_\ell) \\ 0 & -\sin(\alpha_\ell) & \cos(\alpha_\ell) \end{pmatrix}.$$

The angle $\phi$ is assumed to be a known quantity and it denotes phase shift by the gradient magnetic fields [28].

Writing again (14) in a compact form, we are able to derive the evolution of the discrete system for the magnetization vectors

$$\begin{aligned} M_\ell = &\left(\prod_{k=1}^{\ell} E_1(TR_k,\theta)R(\alpha_k)\right) M_0 + E_2(TR_\ell,\theta)M_e \\ &+ \sum_{k=1}^{\ell-1}\left(E_2(TR_k,\theta)\prod_{j=k+1}^{\ell} E_1(TR_j,\theta)R(\alpha_j)\right) M_e, \end{aligned} \tag{15}$$

where we have used the matrix notation $R := R_\phi R_x(\alpha_\ell)R_\phi^\top$ here and later on. We note that with (15) it establishes a mapping between $\theta$ and $\{M_\ell\}_{\ell=1}^L$.

We are interested in the differentiability of the above mapping since it would be important for developing algorithms to solve Equation (12). In the following section, we shall prove that $M(\theta)$ is Fréchet differentiable.

## 4.2 Properties of the Bloch mapping and the qMRI operator

This subsection is devoted to show some differentiability and additional properties of the discrete Bloch mapping

$$M(\theta) = \{M_\ell(\theta)\}_{\ell=1}^L.$$

defined in (14). We will also show among others the non-convexity of the Bloch manifold. Note that

$$M : [L^\infty(\Omega)]^2 \to \mathcal{Y}^L$$

where we recall that $\mathcal{Y} = [L^1(\Omega)]^3$. Observe that $M$ can be thought as the discrete time version of the operator $\mathcal{B}_{m_0,B}$ of the previous section, for this specified IR-bSSFP pulse sequence.

The differentiability of $M$ is shown below and as a consequence the differentiability of the qMRI operator (13) will follow afterwards.

**Proposition 4.1.** *Let $\{M_\ell(\theta)\}_{\ell=1}^L$ be the sequence as given by* (15)*, $\{\alpha_\ell\}_{\ell=1}^L$ the sequence of the flip angles with $\alpha_\ell \in (0,\pi)$ for every $\ell = 1,\dots,L$ and $\{TR_\ell\}_{\ell=1}^L$ the sequence of repetition times with $TR_\ell > 0$ for every $\ell$. Moreover let $M_e = (0,0,1)^\top$ and $M_0 = \pm M_e$ (means either of them). Consider the operator $M$ restricted to those $\theta \in [L^\infty(\Omega)]^2$ such that $\theta_r \in C_{ad}$ for every $r$, where $C_{ad}$ is convex subset of $\mathbb{R}^+ \times \mathbb{R}^+$, bounded, and bounded away from zero. We denote such a feasible set as $\mathcal{A}$ with $\mathcal{A} \subset [L^\infty(\Omega)]^2$. Then we have the following statements:*

*(i) The operator $M$ is injective, that is for $\theta^a,\ \theta^b \in \mathcal{A}$*

$$M(\theta^a) = M(\theta^b) \Longrightarrow \theta^a = \theta^b.$$

*(ii) $M$ is Fréchet differentiable with bounded derivative.*

*Proof.* Due to the recursion structure of $M_\ell$, it turns out that it is sufficient to analyse $M_1$.

(i) It suffices to show that the map $M_1 : [L^\infty(\Omega)]^2 \to \mathcal{Y}$ is injective. Since

$$M_1(\theta) = E_1(TR_1,\theta)R_\phi R_x(\alpha_1)R_\phi^\top M_0 + E_2(TR_1,\theta)M_e,$$

for some non-zero $\alpha_1$ and $TR_1$, we have that $R = R_\phi R_x(\alpha_1)R_\phi^\top$ is unitary, and $E_1(TR_1,\theta)$ and $E_2(TR_1,\theta)$ are contraction operators. Assume now towards contradiction that $M_1(\theta_a) = M_1(\theta_b)$ for $\theta^a \neq \theta^b$, then this implies that

$$\big(E_1(TR_1,\theta^a) - E_1(TR_1,\theta^b)\big)RM_0 + (E_2(TR_1,\theta^a) - E_2(TR_1,\theta^b))M_e = 0,$$

with the above equation to hold in $\mathcal{Y}$. Assume further that $T_1^a \neq T_1^b$, then in those points of $\Omega$ where this occurs we have (note that we suppress the spatial dependence $r$)

$$M_e = \begin{pmatrix} \frac{e^{-\frac{TR_1}{T_2^a}} - e^{-\frac{TR_1}{T_2^b}}}{e^{-\frac{TR_1}{T_1^a}} - e^{-\frac{TR_1}{T_1^b}}} & 0 & 0 \\ 0 & \frac{e^{-\frac{TR_1}{T_2^a}} - e^{-\frac{TR_1}{T_2^b}}}{e^{-\frac{TR_1}{T_1^a}} - e^{-\frac{TR_1}{T_1^b}}} & 0 \\ 0 & 0 & \frac{e^{-\frac{TR_1}{T_1^a}} - e^{-\frac{TR_1}{T_1^b}}}{e^{-\frac{TR_1}{T_1^a}} - e^{-\frac{TR_1}{T_1^b}}} \end{pmatrix} RM_0. \tag{16}$$

Supposing that $M_0 = -M_e$, then from the fact that $R$ is unitary, then (16) is satisfied if and only if $T_1^a = T_2^a$, $T_1^b = T_2^b$, $-RM_e = M_e$ and $\alpha_1 = \pi$. This contradicts our setting that $\alpha \in (0, \pi)$. Similarly we work if $M_0 = M_e$. If $T_1^a = T_1^b$ then $T_2^a \neq T_2^b$ and one can write down the inverse relation of (16) and arrive to the same conclusion. Thus we have injectivity for $M_1$ and hence also of $M$.

(ii) Again we start by considering the differentiability of

$$M_1(\theta) = E_1(TR_1, \theta)R(\alpha_1)M_0 + E_2(TR_1, \theta)M_e.$$

which is easy to derive by using the differentiability of function $e^{-\frac{TR}{x}}$ for $x > 0$. We denote by $M_1'(\theta)$ the Fréchet derivative of the map $M_1$ evaluated at $\theta$, that is $M_1'(\theta) : [L^\infty(\Omega)]^2 \to \mathcal{Y}$ bounded, linear such that

$$\lim_{h\to 0} \frac{\|M_1(\theta + h) - M_1(\theta) - M_1'(\theta)h\|_{\mathcal{Y}}}{\|h\|_{[L^\infty(\Omega)]^2}} = 0. \tag{17}$$

To simplify the formulas, for every $\ell = 1, \ldots, L$, we denote

$$U_1(\ell) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{TR_\ell}{(T_1)^2} e^{-\frac{TR_\ell}{T_1}} \end{pmatrix} R(\alpha_\ell)$$

and

$$U_2(\ell) := \begin{pmatrix} \frac{TR_\ell}{(T_2)^2} e^{-\frac{TR_\ell}{T_2}} & 0 & 0 \\ 0 & \frac{TR_\ell}{(T_2)^2} e^{-\frac{TR_\ell}{T_2}} & 0 \\ 0 & 0 & 0 \end{pmatrix} R(\alpha_\ell).$$

By calculations we compute

$$M_1'(\theta) = \big(M_{1,1}'(\theta),\ M_{1,2}'(\theta)\big) = \left(U_1(1)M_0 - \frac{TR_1}{T_1^2} e^{-\frac{TR_1}{T_1}} M_e,\ U_2(1)M_0\right) \tag{18}$$

Note that $M_1'(\theta) \in \mathcal{Y} \times \mathcal{Y} = [L^1(\Omega)]^3 \times [L^1(\Omega)]^3$. It can be regarded as a bounded linear operator from $[L^\infty(\Omega)]^2 \to \mathcal{Y}$ which is defined for every $h = (h_1, h_2) \in [L^\infty(\Omega)]^2$ as

$$\begin{aligned} M_1'(\theta)h &= M_{1,1}'(\theta)h_1 + M_{1,2}'(\theta)h_2 \\ &:= \big([M_{1,1}'(\theta)]_x h_1, [M_{1,1}'(\theta)]_y h_1, [M_{1,1}'(\theta)]_z h_1\big) \\ &+ \big([M_{1,2}'(\theta)]_x h_2, [M_{1,2}'(\theta)]_y h_2, [M_{1,2}'(\theta)]_z h_2\big) \end{aligned}$$

where above a multiplication between a $L^1(\Omega)$ and an $L^\infty(\Omega)$ function is regarded pointwise, resulting in a $L^1(\Omega)$ function. It is not difficult to see that (17) is satisfied for $M_1'(\theta)$ due to the differentiability of $e^{-\frac{1}{x}}$ for $x$ in strictly positive domain.

The derivative of $M_\ell(\theta)$ for $\ell > 1$ can then be calculated by applying chain rule to the recursion formulas (14),

$$M_\ell'(\theta) = \begin{pmatrix} \Big(U_1(\ell)M_{\ell-1}(\theta) + E_1(TR_\ell, \theta)R(\alpha_\ell)M_{\ell-1,1}'(\theta) - \frac{TR_\ell}{T_1^2} e^{-\frac{TR_\ell}{T_1}} M_e\Big)^\top \\ \big(U_2(\ell)M_{\ell-1}(\theta) + E_1(TR_\ell, \theta)R(\alpha_\ell)M_{\ell-1,2}'(\theta)\big)^\top \end{pmatrix}^\top. \tag{19}$$

We get the boundedness of the derivatives because all the quantities $U_a(\ell)$, $R(\alpha_\ell)$, $E_1(TR_\ell, \theta)$ and $e^{-\frac{TR_\ell}{T_a}} \frac{TR_\ell}{T_a^2}$ for $a = 1, 2$ and $\ell = 1, \ldots, L$ are uniformly bounded.

□

Regarding non-convexity of the Bloch manifold we are able to show the following result.

**Proposition 4.2.** *Under the same assumptions of Proposition 4.1, we have that the image $M[\mathcal{A}]$ of $M : \mathcal{A} \to \mathcal{Y}^L$ is a non-convex subset of $\mathcal{Y}^L$.*

*Proof.* Suppose towards contradiction that the image $M[\mathcal{A}]$ is a convex subset of $\mathcal{Y}^L$. This means that for arbitrary $\theta^a \neq \theta^b \in \mathcal{A}$, and for every $\lambda \in (0,1)$, there exist $\theta^\lambda \in \mathcal{A}$, such that

$$\lambda M_\ell(\theta^a) + (1-\lambda) M_\ell(\theta^b) = M_\ell(\theta^\lambda) \text{ for all } \ell \in \{1, \dots, L\}. \tag{20}$$

We focus on the first two components $M_1, M_2 \in \mathcal{Y}$. Recall

$$M_1(\theta) = E_1(TR_1, \theta) R(\alpha_1) M_0 + E_2(TR_1, \theta) M_e$$
$$M_2(\theta) = \left( \prod_{k=1}^{2} E_1(TR_k, \theta) R(\alpha_k) \right) M_0 + \left( E_2(TR_1, \theta) + E_2(TR_2, \theta) E_1(TR_1, \theta) R(\alpha_1) \right) M_e.$$

With some elementary calculations, the convexity condition (20) can be equivalently written as the following system of equations (where $\theta^a = (T_1^a, T_2^a)$, $\theta^b = (T_1^b, T_2^b)$, $\theta^\lambda = (T_1^\lambda, T_2^\lambda)$):

$$\lambda e^{-\frac{TR_1}{T_1^a}} + (1-\lambda) e^{-\frac{TR_1}{T_1^b}} = e^{-\frac{TR_1}{T_1^\lambda}}, \tag{21}$$
$$\lambda e^{-\frac{TR_1}{T_2^a}} + (1-\lambda) e^{-\frac{TR_1}{T_2^b}} = e^{-\frac{TR_1}{T_2^\lambda}}, \tag{22}$$
$$\lambda e^{-\frac{TR_2}{T_1^a}} e^{-\frac{TR_1}{T_1^a}} + (1-\lambda) e^{-\frac{TR_2}{T_1^b}} e^{-\frac{TR_1}{T_1^b}} = e^{-\frac{TR_2}{T_1^\lambda}} e^{-\frac{TR_1}{T_1^\lambda}}, \tag{23}$$
$$\lambda e^{-\frac{TR_2}{T_2^a}} e^{-\frac{TR_1}{T_2^a}} + (1-\lambda) e^{-\frac{TR_2}{T_2^b}} e^{-\frac{TR_1}{T_2^b}} = e^{-\frac{TR_2}{T_2^\lambda}} e^{-\frac{TR_1}{T_2^\lambda}}, \tag{24}$$
$$\lambda e^{-\frac{TR_1}{T_1^a}} \left(1 - e^{-\frac{TR_2}{T_1^a}}\right) + (1-\lambda) e^{-\frac{TR_1}{T_1^b}} \left(1 - e^{-\frac{TR_2}{T_1^b}}\right) = e^{-\frac{TR_1}{T_1^\lambda}} \left(1 - e^{-\frac{TR_2}{T_1^\lambda}}\right), \tag{25}$$
$$\lambda e^{-\frac{TR_1}{T_2^a}} \left(1 - e^{-\frac{TR_2}{T_1^a}}\right) + (1-\lambda) e^{-\frac{TR_1}{T_2^b}} \left(1 - e^{-\frac{TR_2}{T_1^b}}\right) = e^{-\frac{TR_1}{T_2^\lambda}} \left(1 - e^{-\frac{TR_2}{T_1^\lambda}}\right), \tag{26}$$

where the first two (21)–(22) come from (20) for $M_1$, and the rest (23)–(26) come from (20) for $M_2$. This system can further be simplified as:

$$\lambda e^{-\frac{TR_1}{T_1^a}} + (1-\lambda) e^{-\frac{TR_1}{T_1^b}} = e^{-\frac{TR_1}{T_1^\lambda}},$$
$$\lambda e^{-\frac{TR_1}{T_2^a}} + (1-\lambda) e^{-\frac{TR_1}{T_2^b}} = e^{-\frac{TR_1}{T_2^\lambda}},$$
$$\lambda e^{-\frac{TR_2+TR_1}{T_1^a}} + (1-\lambda) e^{-\frac{TR_2+TR_1}{T_1^b}} = e^{-\frac{TR_2+TR_1}{T_1^\lambda}}$$
$$\lambda e^{-\frac{TR_2+TR_1}{T_2^a}} + (1-\lambda) e^{-\frac{TR_2+TR_1}{T_2^b}} = e^{-\frac{TR_2+TR_1}{T_2^\lambda}}$$
$$\lambda e^{-\frac{TR_1}{T_2^a}} e^{-\frac{TR_2}{T_1^a}} + (1-\lambda) e^{-\frac{TR_1}{T_2^b}} e^{-\frac{TR_2}{T_1^b}} = e^{-\frac{TR_1}{T_2^\lambda}} e^{-\frac{TR_2}{T_1^\lambda}}.$$

Since $TR > 0$, this system has a solution only if $\theta^a = \theta^b = \theta^\lambda$ which gives the contradiction. □

What the Proposition 4.2 says is that the Bloch manifold is not a convex set. This has negative implications when one wants to project the magnetization reconstruction sequences onto this set as it is the case in the BLIP algorithm. Indeed in that case, uniqueness of these projections cannot be guaranteed.

Now we are in a position to further show the Fréchet differentiability for the qMRI operator $Q$.

**Lemma 4.3.** *Let $\mathbf{x} = (\rho, \theta) \in L^\infty(\Omega) \times [L^\infty(\Omega)]^2$, and $D \in ([L^2(K)]^2)^L$ where $K$ denotes the frequency domain. Then the qMRI operator*

$$Q : L^\infty(\Omega) \times [L^\infty(\Omega)]^2 \to \left([L^2(K)]^2\right)^L$$

*(further restricted so $\theta(r) \in C_{ad}$ for every $r \in \Omega$) is Fréchet differentiable.*

*Proof.* Recall that for $\mathbf{x} = (\rho, \theta)$ we have

$$Q(\mathbf{x}) = \{Q^{(\ell)}(\mathbf{x})\}_{\ell=1}^L = \left\{P^{(\ell)}\mathcal{F}(\rho T_{x,y} M_\ell(\theta))\right\}_{\ell=1}^L .$$

To see the Fréchet differentiability, we first notice that all $P^{(\ell)}$, $\mathcal{F}$ and $T_{x,y}$ are bounded, linear operators. Using the Fréchet differentiability of each $M_\ell$ from Proposition 4.1 , we have for every feasible $\mathbf{x} = (\rho, \theta)$ and $\mathbf{h} = (h_\rho, h_\theta)$ belonging to $L^\infty(\Omega) \times [L^\infty(\Omega)]^2$

$$\begin{aligned} Q^{(\ell)}(\mathbf{x}+\mathbf{h}) - Q^{(\ell)}(\mathbf{x}) &= P^{(\ell)}\mathcal{F}((\rho+h_\rho)T_{x,y}M_\ell(\theta+h_\theta)) - P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell(\theta+h_\theta)) \\ &\quad + P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell(\theta+h_\theta)) - P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell(\theta)) \\ &= P^{(\ell)}\mathcal{F}(h_\rho T_{x,y}M_\ell(\theta+h_\theta)) + P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell'(\theta)h_\theta) + o(h_\theta) \\ &= P^{(\ell)}\mathcal{F}(h_\rho T_{x,y}M_\ell(\theta)) + P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell'(\theta)h_\theta) + o(h_\rho h_\theta) + o(h_\theta) \\ &= A^{(\ell)}\mathbf{h} + o(\mathbf{h}). \end{aligned}$$

Note

$$A^{(\ell)}\mathbf{h} = P^{(\ell)}\mathcal{F}(T_{x,y}M_\ell(\theta)h_\rho) + P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell'(\theta)h_\theta)$$

is a bounded linear operator from $L^\infty(\Omega) \times [L^\infty(\Omega)]^2$ to $[L^2(K)]^2$. This shows that $Q$ is Fréchet differentiable. □

The above proof also presents a way on how to calculate the derivative of $Q^\ell$. In what follows, we will always use $Q'$ instead of $A$ to denote the Fréchet derivative of $Q$.

## 4.3 (Gauss-) Newton method for ideal data

In this section we discuss methods for the solution of the non-linear equation

$$Q(\mathbf{x}) = D, \tag{27}$$

which in fact consists of a system of equations

$$P^{(\ell)}\mathcal{F}(\rho T_{x,y}M_\ell(\theta)) = D^{(\ell)}, \quad \ell = 1, \ldots, L.$$

Because of the regularity of the operator $Q$, a first idea to solve the non-linear operator equation (27) is using a Newton type method, or more precisely Gauss-Newton method, which is based on the first order approximation:

$$Q(\mathbf{x}_{n+1}) \simeq Q(\mathbf{x}_n) + Q'(\mathbf{x}_n)\,(\mathbf{x}_{n+1} - \mathbf{x}_n) = D. \tag{28}$$

By letting $D_n := D - Q(\mathbf{x}_n) + Q'(\mathbf{x}_n)\mathbf{x}_n$, then (28) gives the equation, (Newton iteration)

$$Q'(\mathbf{x}_n)\mathbf{x}_{n+1} - D_n = 0. \tag{29}$$

Note that since $D = \{D^{(\ell)}\}_{\ell=1}^{L}$ is a sequence of data frames of length $L$, so is $D_n$, therefore (the space discrete version of) (29) is in general an overdetermined system of linear equations. Thus, one considers the equation (29) in a least square sense leading to the Gauss-Newton method:

$$\mathbf{x}_{n+1} = \operatorname*{argmin}_{\mathbf{x}} \ \|Q'(\mathbf{x}_n)\mathbf{x} - D_n\|^2, \tag{30}$$

whose solution is

$$\mathbf{x}_{n+1} = (Q')^{\dagger}(\mathbf{x}_n)D_n := \left((Q'(\mathbf{x}_n))^{\top}Q'(\mathbf{x}_n)\right)^{-1}(Q'(\mathbf{x}_n))^{\top}D_n.$$

Using the Fréchet differentiability, one can show superlinear convergence rate for the Newton type methods above, provided that the initial guess is sufficiently close to the exact solution. For completeness, we state the theorem below, for more details we refer to [16]

**Theorem 4.4.** *Let $\mathbf{x}^*$ be the exact solution of* (27)*, and suppose that there exists a neighbourhood of $\mathbf{x}^*$ such that $(Q')^{\dagger}(\mathbf{x})$ is uniformly bounded. Then there exists a potentially smaller neighbourhood such that for every initial guess $\mathbf{x}_0$ belonging there, for the iterates in* (30) *we have that $\mathbf{x}_n \to \mathbf{x}^*$ with superlinear convergence, i.e.,*

$$\|\mathbf{x}_{n+1} - \mathbf{x}^*\| = o(\|\mathbf{x}_n - \mathbf{x}^*\|) \ \text{ for all } n \in \mathbb{N}. \tag{31}$$

However, due to the non-linear and non-convex property of the function $Q$, the iteration in (30) will only converge for initial values $\mathbf{x}_0$ in a certain neighbourhood of the exact solution $\mathbf{x}^*$, provided that the data $D$ should contain no noise. We refer to the Newton-Kantorovich Theorem, e.g. [5], where more precise statements can be made.

Overall, for undersampled and noisy data, it is crucial to choose a good initial guess in order to obtain a robust and efficient numerical algorithm for solving the problem (12). This would be the main task of the next section.

## 4.4 Our main algorithm: A Levenberg-Marquardt method for undersampled noisy data

Undersampling is often unavoidable in the acquisition process of MRI due to the time constraint, with the issue being particularly important for fast MRI. The main problem caused by undersampling is ill-posedness of the equation (12) due to the highly compactness of the operator $P$ in the qMRI operator $Q$. In that case, the solution of (12) can be unreliable, especially if the data contains even small noise.

In order to resolve the problem of undersampling and noise, and to solve (12) robustly, we turn to a *Levenberg-Marquardt* (L-M) method instead of the aforementioned Gauss-Newton. There a Tikhonov type regularization is imposed on each iterates. Suppose that the ideal data $D$ have been corrupted with some noise, leading to some perturbed data $D^{\delta}$. The L-M iteration then reads:

$$\begin{aligned} \mathbf{h}_n^{\delta} &= \operatorname*{argmin}_{\mathbf{h}} \ \left\|Q'(\mathbf{x}_n)\mathbf{h} - \tilde{D}_n^{\delta}\right\|^2 + \lambda_n \|\mathbf{h}\|^2 \\ \mathbf{x}_{n+1}^{\delta} &= \mathbf{x}_n^{\delta} + \mathbf{h}_n^{\delta}. \end{aligned} \tag{32}$$

where

$$\tilde{D}_n^{\delta} = D^{\delta} - Q(\mathbf{x}_n).$$

From a regularization point of view, the L-M iteration (32) is nothing else but a iterative Tikhonov regularization for solving a non-linear equation [13, 17]. Note that if $\lambda_n = 0$ for every $n$, then the L-M method degenerates to the Gauss-Newton method. The convergence and convergence rates of L-M methods in the sense of regularization have been shown in many references, a standard one is [13]. There, general rules of choosing the parameter of a form $\lambda_n = \lambda_0 \beta^n$ for some $\lambda_0 > 0$, $\beta \in (0, 1)$ are discussed, as well as a discrepancy principle of terminating the iterations at step $n = n_e$ where $n_e$ is the first iteration index such that the condition

$$\left\| Q(\mathbf{x}_{n_e}^{\delta}) - D^{\delta} \right\| \leq \varrho\delta$$

holds. It is also shown that with these choices, the solution of the L-M method converges to a solution of the original non-linear equation. In our case that means $\mathbf{x}_{n_e}^{\delta} \to \mathbf{x}^*$ as $\delta \to 0$.

The local and global convergence and convergence rates of L-M algorithms have also been intensively studied, we refer to [8, 10, 31] for instance. Usually, in the absence of additive noise, with proper initial values, the optimal convergence rates of the algorithm are determined by the rates of the updated parameters $\lambda_n$,

$$\|\mathbf{x}_n - \mathbf{x}^*\| = \mathcal{O}(\lambda_n).$$

Instead, for non-zero residual case, i.e., in the presence of additive noise, it usually only achieves a linear convergence rate

$$\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq C \|\mathbf{x}_n - \mathbf{x}_{n-1}\|$$

for some constant $C < 1$.

Initialization is crucial for the L-M iteration as well. Proper initial values can not only theoretically prove the convergence rates, but as it is in the case of Gauss-Newton, they help stabilizing the algorithms and also accelerating the convergence to the final solutions.

There is no general way to produce good initial values but it is rather dependent on the specific considered problem. Here we suggest to use a very fast version of MRF and BLIP in order to produce a good initialization. By fast here, we mean that we use only a relatively coarse dictionary, in which case the MRF and its variant BLIP are computational efficient. The idea is that, by initializing the L-M algorithm in this way, then the iteration will converge to the expected solution without a need to refine the dictionary once again. Our main proposed algorithm is summarized below:

**Algorithm 4.1: Proposed algorithm for qMRI: L-M iteration with MRF/BLIP-based initialization**

*Input & setting:*
- MRI data $D^\delta \in ([L^2(K)]^2)^L$,
- Parameters for the physical setting of MRI, e.g., flip angle and repetition time sequences, $\{\alpha_\ell\}_{\ell=1}^L$, $\{TR_\ell\}_{\ell=1}^L$.
- A coarse discretization of the set $C_{ad} = [T_1^{\min}, T_1^{\max}] \times [T_2^{\min}, T_2^{\max}]$.

*Initialization:*
- Generate a dictionary $\mathrm{Dic}(C_{ad})$, using the coarse discretization of $C_{ad}$, the flip angles and the repetition times, with the help of the IR-bSSFP magnetization formula (15).
- Use the BLIP method, Algorithm 2.1, to produce an initialization $\mathbf{x}_0 = (\rho_0, \theta_0) \in L^\infty(\Omega) \times [L^\infty(\Omega)]^2$

*L-M iteration:*
- Run the Levenberg-Marquardt iteration (32), with initialization $\mathbf{x}_0$, and adaptive parameters $\lambda_n = \max\{\lambda_0 \beta^n, \mu_n\}$, for some $\lambda_0 \geq 1$, $\beta \in (0,1)$, and $\mu_n$ chosen according to the noise level.

*Output*: The estimated parameter map $\mathbf{x}_{n_e} = (\rho_{n_e}, \theta_{n_e})$, for some final $n = n_e$.

Here $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of parameters that depend on the data noise level, and $\lambda_0$ depends on sub-sampling rate. In the numerical examples later, we will set $\lambda_0 = s^2$, where $1/s$ is the under-sampling rate of the data. A typical choice for $\mu_n$ is $\mu_n = \epsilon \left\| Q(\mathbf{x}_n) - D^\delta \right\|$ where $\epsilon \in (0,1)$.

## 4.5 Why more data frames can help?

In the original paper of MRF, in order to deal with the problem of noisy data, it was proposed to use a large number $L$ of consecutive pulse sequences and acquisitions, with the idea that this will be able to average out the noise and thus obtain better reconstructions. The same idea holds for our approach and we shall show in the following that this idea can be theoretically justified.

We first recall the, so called Chebyshev's inequality for vector valued random variables (see e.g. [11, 25]).

**Lemma 4.5** (Chebyshev's inequality)**.** *Let $\phi = (\phi_1, \phi_2, \ldots, \phi_p)$ be a vector-valued random variable, for some $p \in \mathbb{N}$, with expected value and variance $E(\phi) = \chi = (\chi_1, \chi_2, \ldots, \chi_p)$, $V(\phi) = \Sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)$ respectively. Then we have*

$$\mathbb{P}(\|\phi - \chi\|_{\mathbb{R}^p} \leq \epsilon) \geq 1 - \frac{\|\Sigma^2\|_{\mathbb{R}^p}}{\epsilon^2}, \tag{33}$$

*for every $\epsilon > 0$, where here $\|\cdot\|_{\mathbb{R}^p}$ is the Euclidean norm in $\mathbb{R}^p$.*

We now show the main theorem of this section. It says that if a family of $L$ linear systems has a common solution, and the right hand sides are perturbed by noise, then by solving a least square problem one can get an approximation of the common solution, with a certain probability that improves as the number $L$ increases. Later we will see how this applies to our proposed approach for qMRI.

**Theorem 4.6.** *Let $A_\ell \zeta = b_\ell$, $\ell = 1, \dots, L$ be a family of $L$ linear systems of equations, where $\{b_\ell\}_{\ell=1}^L$ and $\{A_\ell\}_{\ell=1}^L$ are sequences of data and matrices respectively. We consider each $b_\ell$ to be a vector of dimension $p$, and every $A_\ell$ to be of size $d \times p$, where $p \leq d$ and of full rank, i.e., $\operatorname{rank}(A_\ell) = p$ for every $\ell = 1, \dots, L$. We assume that the singular values of $A_\ell$ have a uniform lower and upper bound $\sqrt{c}$, and $\sqrt{C}$ respectively, which are both independent of $L$. We further assume that this family of systems of equations has a common solution $\zeta^* \in \mathbb{R}^p$. Suppose that $\tilde{b}_\ell = b_\ell + \delta_\ell$, where $\{\delta_\ell\}_{\ell=1}^L$ are independently identically distributed (i.i.d.) random variables with expected value $(0, \dots, 0) \in \mathbb{R}^p$, and variance $(\sigma^2, \dots, \sigma^2) \in \mathbb{R}^p$. Then the least square solution*

$$\zeta_{ls} = \underset{\zeta \in \mathbb{R}^p}{\operatorname{argmin}} \left\| A\zeta - \tilde{b} \right\|_{\mathbb{R}^{Ld}}^2, \tag{34}$$

*where*

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_L \end{pmatrix} \quad \textit{and} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix}$$

*will approximate the real solution $\zeta^*$ with the following probability estimate*

$$\mathbb{P}(\|\zeta_{ls} - \zeta^*\|_{\mathbb{R}^p} > \epsilon) < \frac{\sigma^2}{\epsilon^2} \mathcal{O}\left(\frac{p}{L}\right), \qquad \textit{for every } \epsilon > 0. \tag{35}$$

*Proof.* From (34), we get $\zeta_{ls} = (A^*A)^{-1}A^*\tilde{b}$, which is also a random variable. Since $A$ is not random, we can compute the expected value of $\zeta_{ls}$ as follows

$$E(\zeta_{ls}) = E((A^*A)^{-1}A^*\tilde{b}) = (A^*A)^{-1}A^*E(\tilde{b}) = (A^*A)^{-1}A^*b = \zeta^*.$$

Therefore $E(\zeta_{ls} - \zeta^*) = 0$. Similarly for the variance (diagonal of the covariance matrix) we have

$$V(\zeta_{ls} - \zeta^*) = \sigma^2 \mathsf{diag}\left((A^*A)^{-1}\right),$$

where diag denotes the diagonal of a matrix. Denoting by $Tr$, the trace operator, i.e., the summation of the diagonal values and using Lemma 4.5, we get that for every $\epsilon > 0$

$$\begin{aligned} \mathbb{P}(\|\zeta_{ls} - \zeta^*\|_{\mathbb{R}^p} > \epsilon) &< \frac{\sigma^2 \|\mathsf{diag}\left((A^*A)^{-1}\right)\|_{\mathbb{R}^p}}{\epsilon^2} \\ &\leq \frac{\sigma^2 Tr\left((A^*A)^{-1}\right)}{\epsilon^2}. \end{aligned} \tag{36}$$

Here we have used the fact that the matrix $(A^*A)^{-1}$ is positive definite and hence it has strictly positive diagonal elements, together with the fact that $1$-norm in $\mathbb{R}^p$ is larger than the Euclidean one.

From the form of $A$ we have $A^*A = \sum_{\ell=1}^L A_\ell^* A_\ell$ with a trace $Tr(A^*A) = \sum_{\ell=1}^L Tr(A_\ell^* A_\ell)$. Since every $A_\ell^* A_\ell$ is positive definite, so is $A^*A$. Let $\{S_j\}_{j=1}^p$ be the eigenvalues of $A^*A$, then can have the decomposition

$$A^*A = USU^{-1} \quad \text{and} \quad (A^*A)^{-1} = US^{-1}U^{-1} \tag{37}$$

where $S$ is the diagonal matrix with $\{S_j\}_{j=1}^p$ on the diagonal, and $U$ is a unitary matrix. Then for the traces we have

$$Tr(A^*A) = \sum_{j=1}^p S_j \quad \text{and} \quad Tr\left((A^*A)^{-1}\right) = \sum_{j=1}^p \frac{1}{S_j}.$$

Due to the uniform lower and upper bounds on the singular values of $\{A_\ell\}_{\ell=1}^L$, we have a corresponding uniform bound of the eigenvalues of the matrices $\{A_\ell^* A_\ell\}_{\ell=1}^L$

$$cL \leq S_j \leq CL, \quad \text{for all } \; j = 1, 2, \ldots, p.$$

Consequently we have

$$\frac{1}{CL} \leq \frac{1}{S_j} \leq \frac{1}{cL} \quad \Rightarrow \frac{1}{S_j} = \mathcal{O}\left(\frac{1}{L}\right), \;\; \text{for all } \; j = 1, 2, \ldots, p.$$

This shows the following estimate

$$Tr\left((A^*A)^{-1}\right) = \sum_{j=1}^{p} \frac{1}{S_j} = \mathcal{O}\left(\frac{p}{L}\right),$$

and combined with (36), gives the final conclusion

$$\mathbb{P}(\|\zeta_{ls} - \zeta^*\| > \epsilon) = \frac{\sigma^2}{\epsilon^2}\mathcal{O}\left(\frac{p}{L}\right).$$

□

We discuss now how Theorem 4.6 can be put into context with our qMRI algorithm. Note initially that, regarding the setting of the qMRI problems, the noise in the data obtained after each pulse sequence can be considered as i.i.d. random variables. In the (Gauss-)Newton method, if there is no sub-sampling for the qMRI operator, then $Q'$ is non-degenerate in the effective domain $\Omega$. By effective we mean the part of the slices where the biological tissue is imaged. In this case, we can think $A_\ell = (Q^{(\ell)})'$, and $b_\ell = (D_k^{(\ell)})^\delta$, and they will satisfy the assumptions of Theorem 4.6 given that the data contains Gaussian noise. This indicates that the result of Theorem 4.6 can be applied to every Newton step for a given $\epsilon > 0$, and an initial value $\|\mathbf{x}_0 - \mathbf{x}^*\| \leq \epsilon$. Here the common solution $\zeta^*$ corresponding to (34) of Theorem 4.6 will be the exact solution $\theta$ of the least square problem (30).

In the case of Levenberg-Marquardt method, with the sub-sampling operators $P^{(\ell)}$, the results of Theorem 4.6 can still be applied as the matrices become $A_\ell = \left(((Q^{(\ell)})')^\top, \sqrt{\lambda_n}\mathrm{Id}\right)^\top$, and $b_\ell = \left(((\tilde{D}_n^{(\ell)})^\delta)^\top, 0\right)^\top$. Note that in this case $\zeta$ corresponds to $\mathbf{h}$. The matrices $A_\ell$ will be always of full rank with uniformly bounded singular values if we let the sequence $(\lambda_n)_{n\in\mathbb{N}}$ to be uniformly bounded from below, as it is typically the case when there is noise in the data. In this case the common solution of (32) will be $\zeta^* \equiv 0$.

The above analysis shows why by choosing multiple frames, i.e., many pulse sequences, one can get a better result. Even though, Theorem 4.6 cannot be applied to the original MRF framework, it gives a first understanding why these type of methods are benefited by a multiple frame strategy.

# 5 Numerical Results

In this section, we present numerical results implemented with a set of synthetic data using the proposed Algorithm 4.1. We will provide an extensive quantitative comparison with the BLIP method, Algorithm 2.1, which was shown in [7] to be superior to the original MRF in the setting we are considering.

We first explain the experimental set-up for our numerical algorithms.

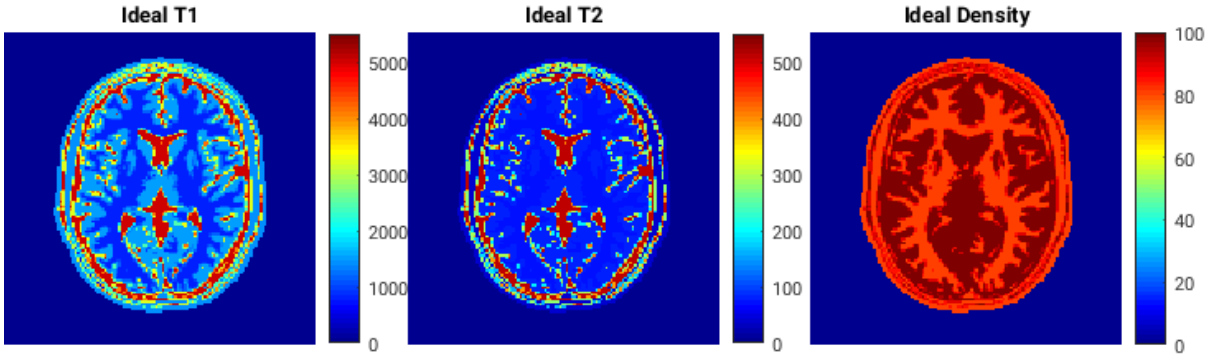


Figure 1: The interpolated parameters serve as the ground truth for our algorithm. From left to right: $T_1$, $T_2$ and $\rho$.

## 5.1 Tested data

We used synthetic anatomical brain phantom data, publicly available on the Brain Web Simulated Brain Database [1, 6]. We used a a $217 \times 181$ slice, which was subsequently, zero-filled in order to make a $256 \times 256$ pixels image. The range for the parameters $\theta = (T_1, T_2)^\top$, and $\rho$ were selected in a way that the correspond to the natural values encountered in the human body [7], where $T_1$ is ranged from $530ms$–$5012ms$, and $T_2$ is ranged from $41ms$–$512ms$, and values $\rho$ are between $80$–$100$. In order to make the testing data to be more realistic, we interpolated the values of each parameters $T_1$, $T_2$ and $\rho$ of the obtained $256 \times 256$ phantom by averaging the values of every four neighboured pixels with non-zero parameter values. This is based on the fact that the pixel units in practical images very likely contain multi-type of tissues instead of a pure one in a single volume of the observed pixels. This average process shrinks the $256 \times 256$ image to a $128 \times 128$ image. In Figure 1, we display the interpolated parameters of $T_1$, $T_2$ and $\rho$ as coloured images. These serve as the ground truth of our numerical tests.

We use a simulated IR-bSSFP pulse sequence scheme to generate the MRI data, where we considered constant flip angles $\alpha$ and repetition time $TR$ sequences. Both sequences are of length $L$. The data $D$ are generated by using the prescribed parameters $T_1$, $T_2$ and $\rho$ with the pulse sequences characterized by $\alpha$ and $TR$. With this setting, we first simulate the magnetization, and then use FFT to generate the Fourier space data from it. Sub-sampling is implemented using the scheme that is described in the next section. In order to generate the magnetization, we used formula (15), also taking advantage of the MATLAB code provided in [22].

The sub-sampling pattern of the Fourier space data is important for generating the synthetic data, and it is also a part of the qMRI operator. We discuss it in detail below.

## 5.2 Sub-sampling pattern

As we have mentioned before, here we focus on Cartesian sub-sampling which is quite standard in practice [23]. In what follows, we discuss the form of the sub-sampling operator $P^{(\ell)}$ for $\ell = 1, \ldots, L$. Note that since we are in the discrete setting, the full $k$-space data will be a full matrix of complex values, or equivalently two real valued full matrices, of size $N \times N$. In our examples, a full $k$-space data is then a $128 \times 128$ matrix of complex values, that is $N = 128$.

We use here an $n$ multishot Echo Planar Imaging (EPI) scheme, which means that in every read out

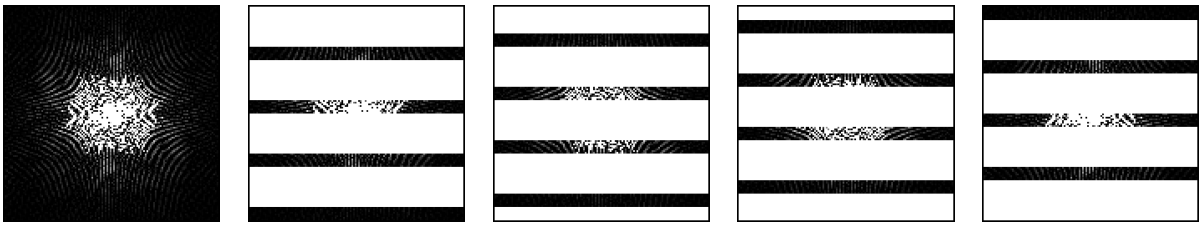

Figure 2: A sub-sampling pattern example for $s = 4$, $L = 4$, and $N = 16$. The first image depicts an example of a fully sampled $k$ space data. The second to the fifth images are example frames of the under-sampled data, where the information at the blank rows part is not collected in that frame. The sub-sampling pattern follows the order periodically if $L > s$.

time, $n$ rows of the $k$-space are simultaneously filled. Hence, in every acquisition, there will be $n < N$ rows of the matrix filled with data of Fourier coefficients. To simplify the discussion we consider $(N \mod n) \equiv 0$, and further set $s = N/n$, which means that the sub-sampling rate is $1/s$. The sampling pattern $P^{(\ell)}$ is described in detail as follows:

(i) For every $\ell$-th acquisition, define $\xi_\ell := (\ell \mod s)$ for $\ell = 1, \ldots, L$.

(ii) $P^\ell$ will include those rows in the full $k$ space matrices, indexed by numbers in a set $\iota$ where

$$\iota := \{i \in \{1, \ldots, N\}, \text{ such that } (i \mod s) \equiv \xi_\ell\}.$$

Thus, at every time, $P^{(\ell)}$ will sample $n$ rows from the full Fourier space to simulate the $n$ multishot EPI. A toy example of such a sub-sampling pattern is given in Figure 2.

We note that this sub-sampling strategy is different to the one described in [7]. There $\xi_\ell$ is defined as a uniformly distributed random number in $\{0, 1, \ldots, s-1\}$, whereas here we consider a deterministic periodical order. After experimentation we have observed that the deterministic order is more stable, and gives comparable or better results when the BLIP algorithm for Cartesian sub-sampled data is used, in comparison to the pseudo-random scheme in [7]. Therefore, in our implementation for the BLIP algorithm, we will use the deterministic strategy described above instead of the original one.

We consider different sub-sampling rates on the data using the above strategy. By taking into account that longer processing time is needed for higher sampling rates, it follows that the flip angles and the repetition times must be increased proportionally. We use constant repetition times $TR = (TR_1, TR_2, \ldots, TR_L)$ and also flip angles $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_L)$. In particular, we used the following values for each sub-sampling rate:

(a) For fully sampled data, we used: Repetition time $TR_\ell = 40ms$ and flip angles $\alpha_\ell = \frac{40\pi}{180}$ for all $\ell = 1, \ldots, L$.

(b) For $1/4$ sampled data (sub-sampling rate $25\%$), e.g., a 32 multi-shot EPI, we used: Repetition time $TR_\ell = 20ms$ and flip angles $\alpha_\ell = \frac{20\pi}{180}$ for all $\ell = 1, \ldots, L$.

(c) For $1/8$ sampled data (sub-sampling rate $12.5\%$), e.g., a 16 multi-shot EPI, a shorter repetition time $TR_\ell = 10ms$ and smaller flip angles $\alpha_\ell = \frac{10\pi}{180}$ are applied for all $\ell = 1, \ldots, L$.

It is reported in the literature [22] that adding random noise to the flip angles and to repetition times may improve the final results of MRF (and BLIP). However in our experiments, we did not find significant differences. Therefore, we do not add noise to the angles and repetition times in our numerical tests.

## 5.3 Results

In order to compare the proposed method and the BLIP algorithm, we will use for the latter a very fine dictionary where $T_1$ is discretized from $15ms$ to $5500ms$ with increments of $15ms$, and $T_2$ is discretized from $1.5ms$ to $550ms$ with increments of $1.5ms$. This means that the discretized feasible domain $C_{ad}$ for $\theta = (T_1, T_2)$ is (in MATLAB notation)

$$C_{ad} = \{[15 : 15 : 5500] \times [1.5 : 1.5 : 550]\}.$$

In this case the dictionary has $366 \times 366 = 133956$ entries, and it requires memory for a storage matrix of dimension $133956 \times L$. We used a computer with a CPU: Intel Core i5-7500, 3.40GHz, 2 cores, and an RAM: 8GB DDR4, 2400 MHz. All of our experiments run with MATLAB on a operator system openSUSE 42.3. We use the deterministic sub-sampling pattern in all the numerical examples of BLIP algorithms, including the generation of the initial values. The linear systems in L-M iterations were by MATLAB's backslash command.

### 5.3.1 Undersampled data with no additive noise

The first set of examples deals with noiseless under-sampled data (Cartesian sub-sampling at rate $1/8$), and totally $L = 80$ frames data. In this example, we use a coarse dictionary for initilizaing Algorithm 4.1. Here $T_1$ is discretized from $200ms$ to $5500ms$ with increments of $200ms$, and $T_2$ is discretized from $20ms$ to $550ms$ with increments of $20ms$. We compare our results to the solutions of the BLIP algorithm, and we also show the result of the initial guess given by BLIP. In this case, $20$ steps of Landweber iteration for the BLIP algorithm are applied, which according to [7] are sufficient. Correspondingly $25$ steps of Levenberg-Marquardt iterations were used for proposed algorithm and gave sufficient results. After that no significant change in the iterates was observed. We let the regularization parameters have the following values: $\mu = \mu_n = 0$, for every $n \in \mathbb{N}$, $\lambda_0 = s^2$, and $\beta = 0.01$.

The reconstructed parameters are presented in Figure 3. There we show in the first row the parameter maps $T_1$, $T_2$, $\rho$ of the BLIP algorithm, computed with the coarse dictionary, that are subsequently used for the initialization of our algorithm. In the second row, the corresponding results for the fine dictionary are shown. These are the ones that should be compared with the images of the third row, which are the results of our algorithm. In order to make the distinction more clear, we also provide the corresponding error maps in Figure 4. There we have computed the pointwise error maps

$$|\theta_{computed} - \theta_{gt}|,$$

where $\theta_{gt}$ are the ground truth parameter maps shown in Figure 1. We can observe that the accuracy of estimated parameters, especially for $T_1$, given by the proposed method is much higher than the BLIP algorithm. Note that the error in BLIP is actually larger than the dictionary mesh size, which indicates that this is not a matter of the fineness of dictionary, but it could also be due to the projection on a non-nonvex set as we have previously discussed.

Regarding the rate of convergence of the proposed algorithm, we see that in this case is linear, see Figure 5. There we have plotted the ratio $\frac{\|x_{n+1}-x_n\|_2}{\|x_n-x_{n-1}\|_2}$ versus the number of iterations. Note that here we denote $x = T_1,\ T_2$ or $\rho$.

#### 5.3.2 Undersampled data with additive noise

In this set of results, we present the case in which the data is $1/4$ sub-sampled and they are also corrupted with additive Gaussian noise. We also use here $L = 80$ data frames as the last set of examples. The noisy data were generated as follows: To each simulated magnetization frame, we added Gaussian noise of variance $\sigma^2 = 0.8$ and mean $0$, and then applied the discrete Fourier transform with a subsequent sub-sampling. The total signal to noise ratio between the data $D$ and the noise is $SNR = 35$.

The coarse dictionary mesh used in order to generate the initial value $\mathbf{x}_0$ is as follows: $T_1$ is discretized from $400ms$ to $5500ms$ with increments of $400ms$, and $T_2$ is discretized from $40ms$ to $550ms$ with increments of $40ms$. Therefore here, we have a dictionary which consists of $13 \times 13 = 169$ entries only, and ask for a matrix of dimension $169 \times L$ for complex values. The refined dictionary is the same as described previously, of dimension $133956 \times L$ for complex values. We use again $20$ iterations for the Landweber iterations in the BLIP algorithm, and $25$ iterations for our L-M algorithm. The regularization parameters are chosen as follows: $\mu_n = 10^{-5}$ for every $n \in \mathbb{N}$, $\lambda_0 = s^2$, and $\beta = 0.01$. Note that because of noise, here we used a fixed $\mu$ strictly larger than zero.

We depict the results in Figure 6 and the corresponding pointwise errors in Figure 7, using the same row system as in Figures 3 and 4. The result of the proposed algorithm again outperforms the refined BLIP algorithm, but not as significantly as in the noiseless case. However it still consumes much less memory and requires much less computational time, see Table 1.

In Figure 8, we plotted again the residual ratios, that again show a linear rate of convergence.

#### 5.3.3 Ideal data–fully sampled and no noise

We also show the results for the case of fully sampled and noise-free data. Here, we only need $L = 3$ data frames, which actually equals to the number of unknown parameters, and it gives a determined system. For both the BLIP and our algorithm, we chose the number of iterations to be $5$. The regularization parameters are chosen as follows: $\mu_n = 0$ for every $n \in \mathbb{N}$, $\lambda_0 = s^2 = 1$, and $\beta = 0$. Note that, as we have discussed, this choice makes the L-M iteration equivalent to Gauss-Newton method.

Here, we only show the error maps of the results in Figure 9. We observe that the Gauss-Newton algorithm essentially recovers the ground truth in this case, while the accuracy of BLIP algorithm, is limited by the discretization mesh of the dictionary.

In contrast to the previous case, as we verify numerically in Figure 10, the convergence rate of the algorithm is superlinear.

#### 5.3.4 Quantitative comparisons

We also provide in Table 1 a summary of qualitative comparisons of all the previous results. The index in our comparison is the computational time costs as well as the error rates given by the algorithms.

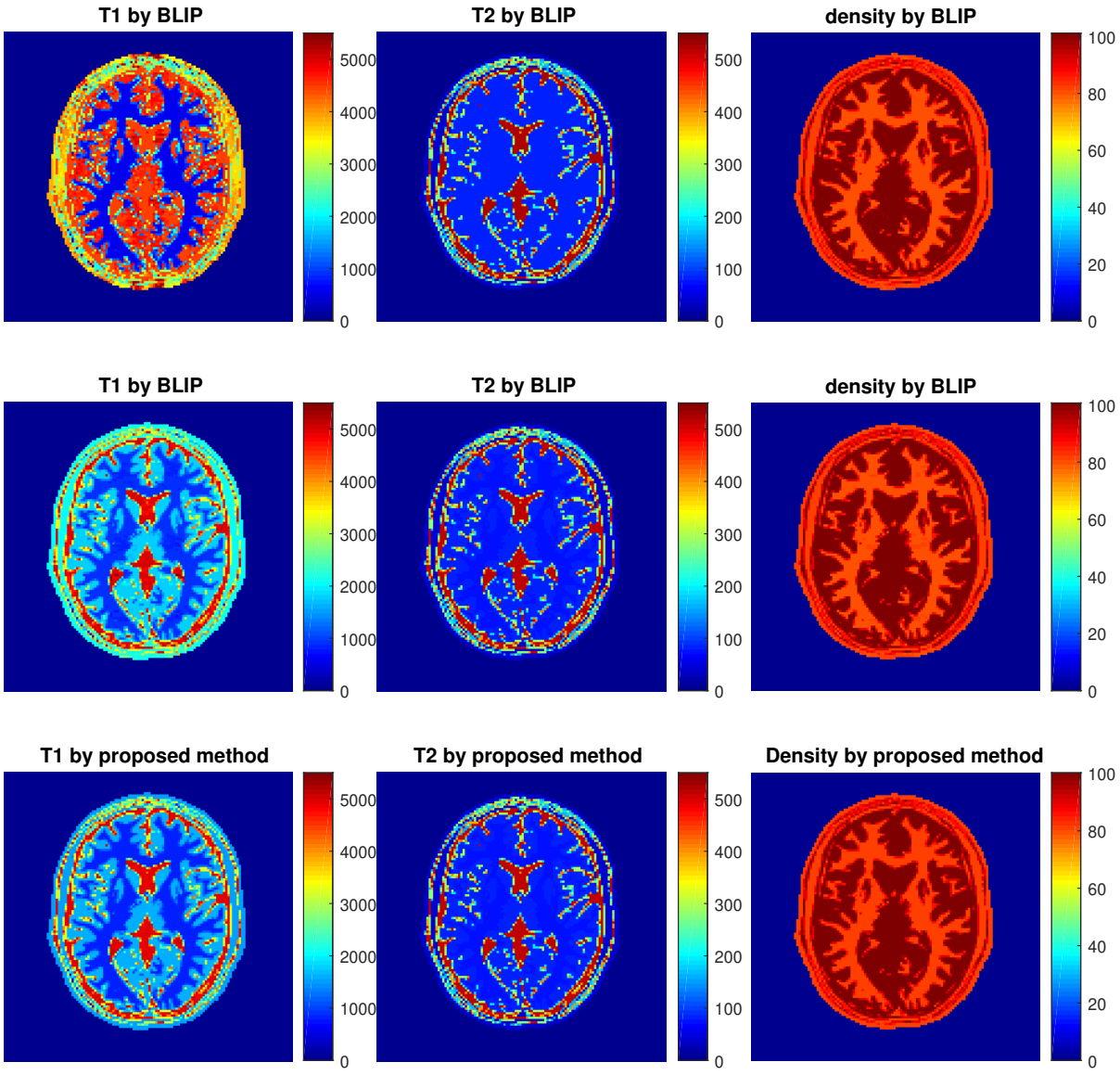


Figure 3: Experiment with noiseless undersampled data. First row: Initialization of our algorithm, computed by BLIP with a coarse dictionary. Middle row: Result by BLIP with fine dictionary. Last row: Solution by proposed algorithm.

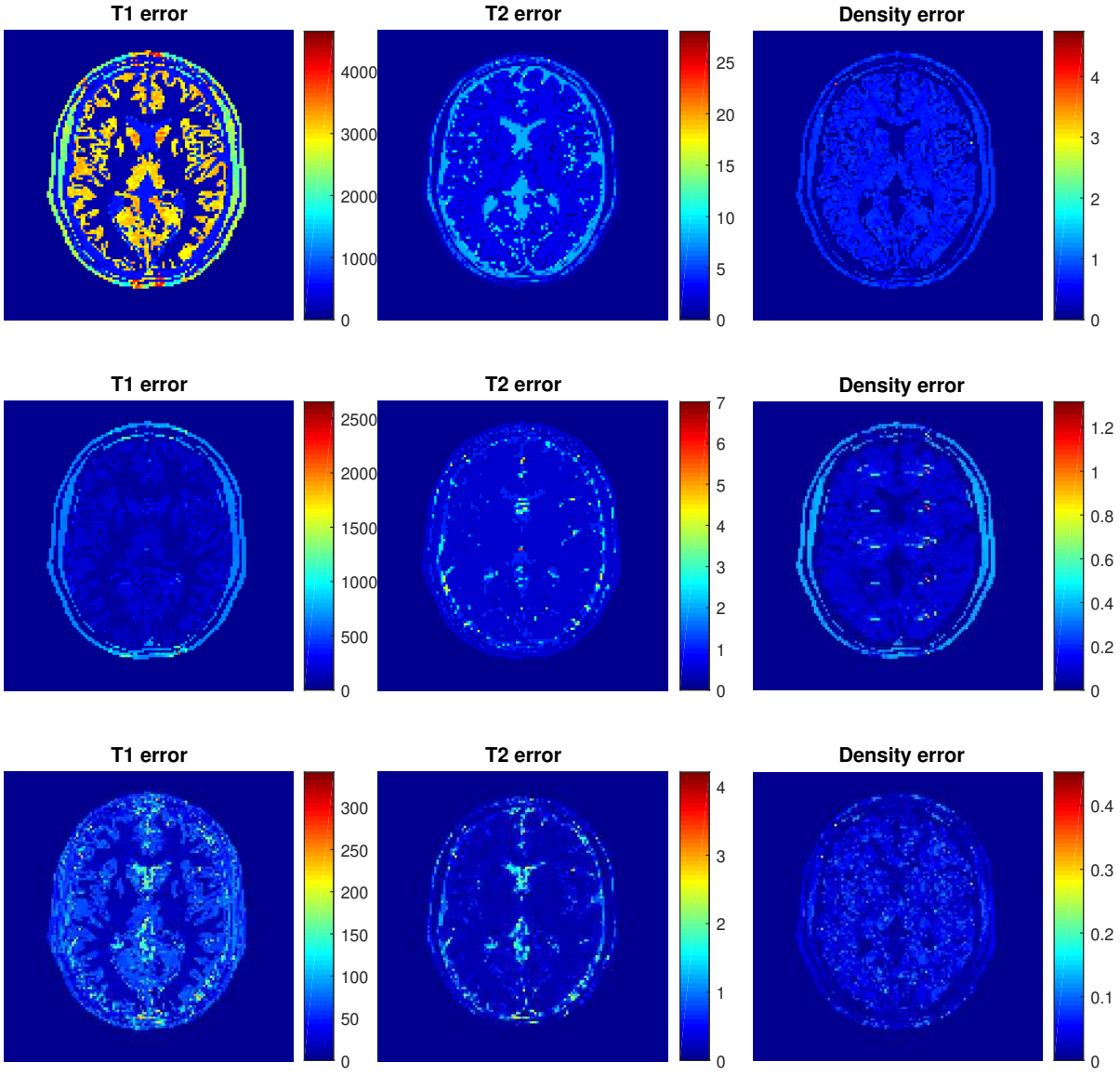


Figure 4: Experiment with noiseless undersampled data. Pointwise distance of the solutions of Figure 3, to the corresponding ground truths, in Figure 1. First row: Initial error of BLIP with a coarse dictionary. Middle row: error of BLIP with fine dictionary. Last row: Error of the proposed algorithm.

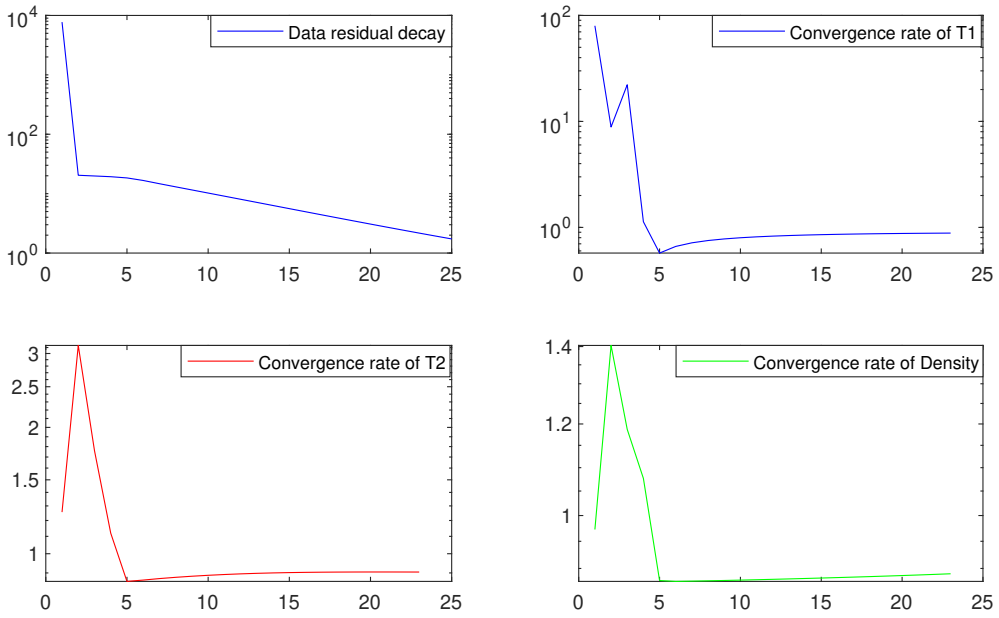


Figure 5: Experiment with undersampled data. From left to right and from above to bottom: Convergence of the data residual $\|Q(\mathbf{x}_n) - D\|$, convergence rates via plots of the iterate ratios $\frac{\|x_{n+1}-x_n\|_2}{\|x_n-x_{n-1}\|_2}$ for $x = T_1, T_2, \rho$ respectively.

The error rates are defined as

$$\frac{\|x_{computed} - x_{gt}\|_2}{\|x_{gt}\|_2}$$

where $x = T_1, T_2$ or $\rho$. We give also the corresponding results for the initialization $\mathbf{x}_0$. Note that the time costs for the proposed algorithm presented here include the time that was needed for the computation of initial value. From the table, we observe that the initial guess has been significantly

Table 1: Quantitative summary of the results: computational times and error rates

| | Full data | | | | $1/8$ sampled data | | | | $1/4$ sampled and noisy data | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | time $(s)$ | error rate $\frac{\|e\|}{\|x^*\|}$ | | | time $(s)$ | error rate $\frac{\|e\|}{\|x^*\|}$ | | | time $(s)$ | error rate $\frac{\|e\|}{\|x^*\|}$ | | |
| | | $T_1$ | $T_2$ | $\rho$ | | $T_1$ | $T_2$ | $\rho$ | | $T_1$ | $T_2$ | $\rho$ |
| Initialization | 1.20 | 0.036 | 0.009 | 0.008 | 15.26 | 0.472 | 0.010 | 0.003 | 13.67 | 0.148 | 0.088 | 0.188 |
| BLIP | 78.94 | 0.005 | 0.002 | 0.003 | 1064.04 | 0.072 | 0.002 | 0.001 | 1273.86 | 0.078 | 0.019 | 0.028 |
| Proposed | 8.41 | $2.2 * 10^{-13}$ | $2.6 * 10^{-15}$ | $6.4 * 10^{-16}$ | 519.41 | 0.009 | 0.001 | 0.0002 | 525.43 | 0.075 | 0.013 | 0.018 |

improved by both algorithms, but in the end our proposed algorithm outperforms the refined BLIP in all of the indexes.

Finally, we would like to verify the fact that larger sequences of frames can help for more accurate estimations, as we discussed in the end of Section 4.5. We do that in Table 2. There we performed a set of experiments, for data frames of increasing length $L$. All data are $1/4$ sub-sampled and corrupted with additive i.d.d. Gaussian noise as described before with variance $1$ and mean $0$. This gives a total signal noise ratio $SNR = 15$. We then run our L-M algorithm for 20 iterations using the same initial value $\mathbf{x}_0$, which was generated by using BLIP with $160$ frames and the coarse dictionary as described

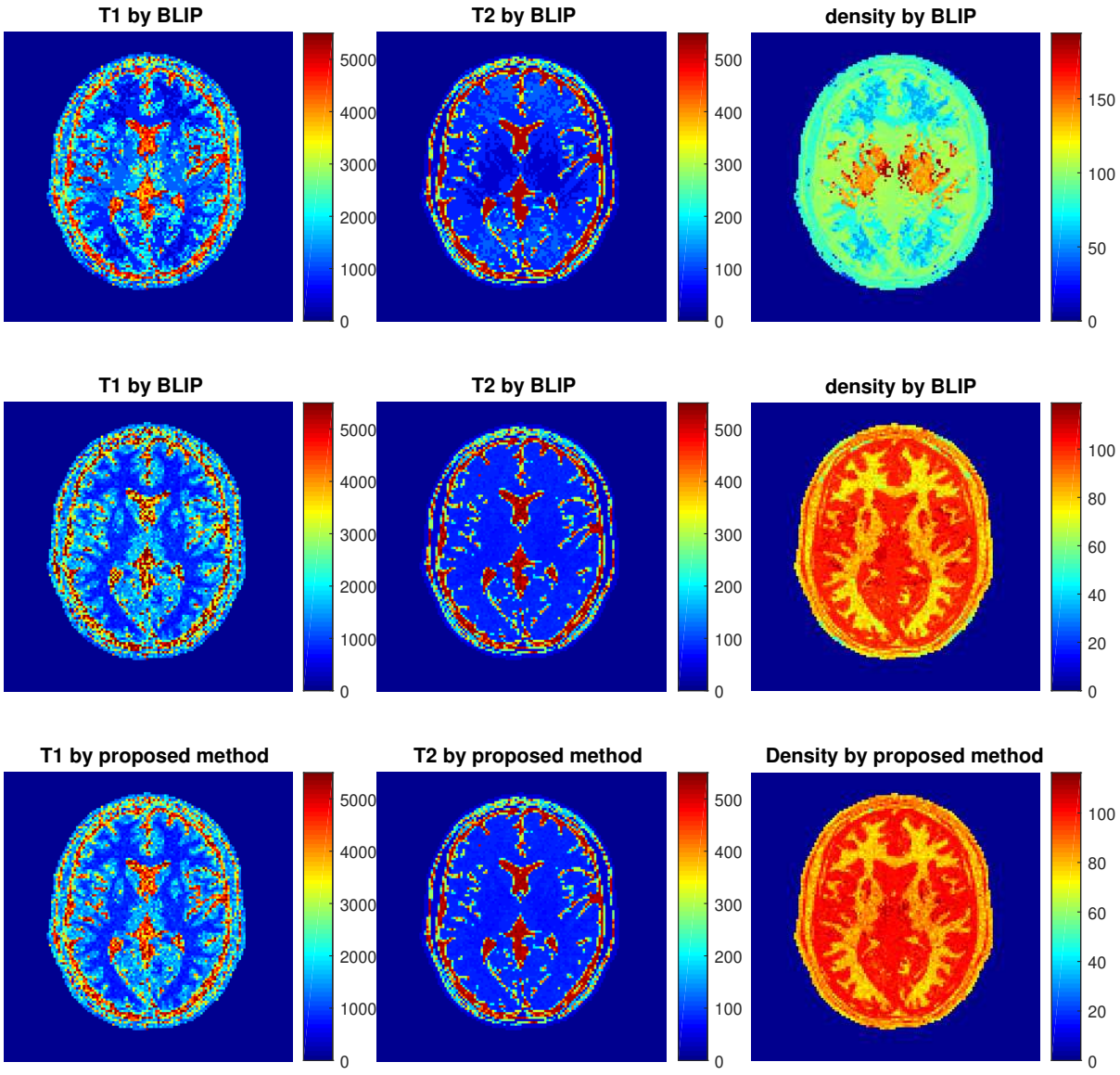


Figure 6: Experiment with undersampled and noisy data. First row: Initialization of our algorithm, computed by BLIP with a coarse dictionary. Middle row: Result by BLIP with fine dictionary. Last row: Solution by proposed algorithm.

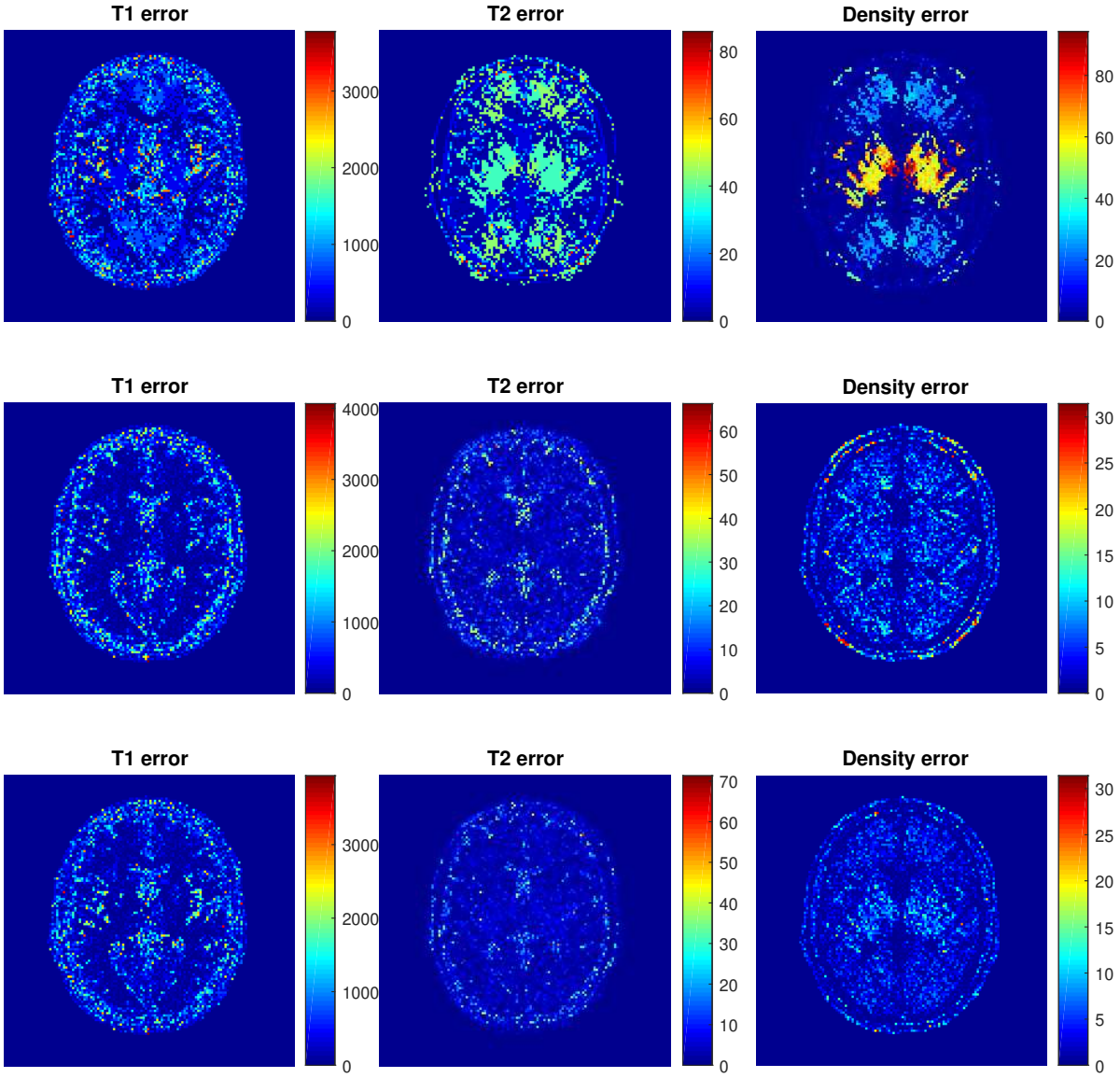


Figure 7: Experiment with undersampled and noisy data. Pointwise distance of the solutions of Figure 6, to the corresponding ground truths of Figure 1. First row: Initial error of BLIP with a coarse dictionary. Middle row: error of BLIP with fine dictionary. Last row: Error of the proposed algorithm.

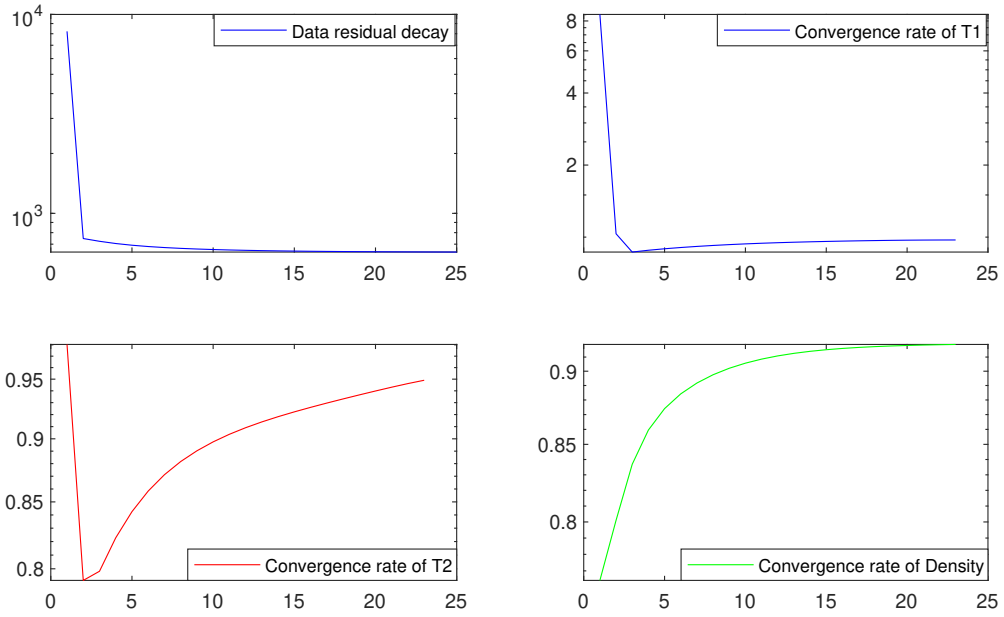


Figure 8: Experiment with noisy data. From left to right and from above to bottom: Convergence of the data residual $\|Q(\mathbf{x}_n) - D^\delta\|$, convergence rates via plots of the iterate ratios $\frac{\|x_{n+1}-x_n\|_2}{\|x_n-x_{n-1}\|_2}$ for $x = T_1, T_2, \rho$ respectively.

in Section 5.3.2. We choose the parameter simply as $\lambda_n = \mu_n = \lambda_0 * \left\|Q\mathbf{x}_n - D^\delta\right\|$ for all $n \in \mathbb{N}$, and $\lambda_0 = 10^{-8}$.

The results in Table 2 clearly indicate that increasing number of sequences improve accuracy of the estimated parameters. Note that the time costs that we show here do not include the initialization, which was the same for every case.

Table 2: The length of data sequences and its influence to the solution accuracy

| sequence length | error rate $T_1$ | error rate $T_2$ | error rate $\rho$ | time cost $(s)$ |
|---|---|---|---|---|
| $L = 5$ | $0.2004$ | $0.5131$ | $0.1622$ | $37.50$ |
| $L = 10$ | $0.1658$ | $0.1736$ | $0.0738$ | $60.37$ |
| $L = 20$ | $0.0986$ | $0.0409$ | $0.0315$ | $106.50$ |
| $L = 40$ | $0.0908$ | $0.0221$ | $0.0278$ | $198.68$ |
| $L = 80$ | $0.0791$ | $0.0175$ | $0.0270$ | $386.60$ |
| $L = 160$ | $0.0692$ | $0.0153$ | $0.0268$ | $756.54$ |

# 6 Concluding Remark and Discussions

In this paper, we analysed MRF from the perspective of inverse problems, and we were able to provide some mathematical insight in order to better understand it. Subsequently, we have proposed a novel model for quantitative MRI which is in accordance with the standard routine of MRI experiment setting.

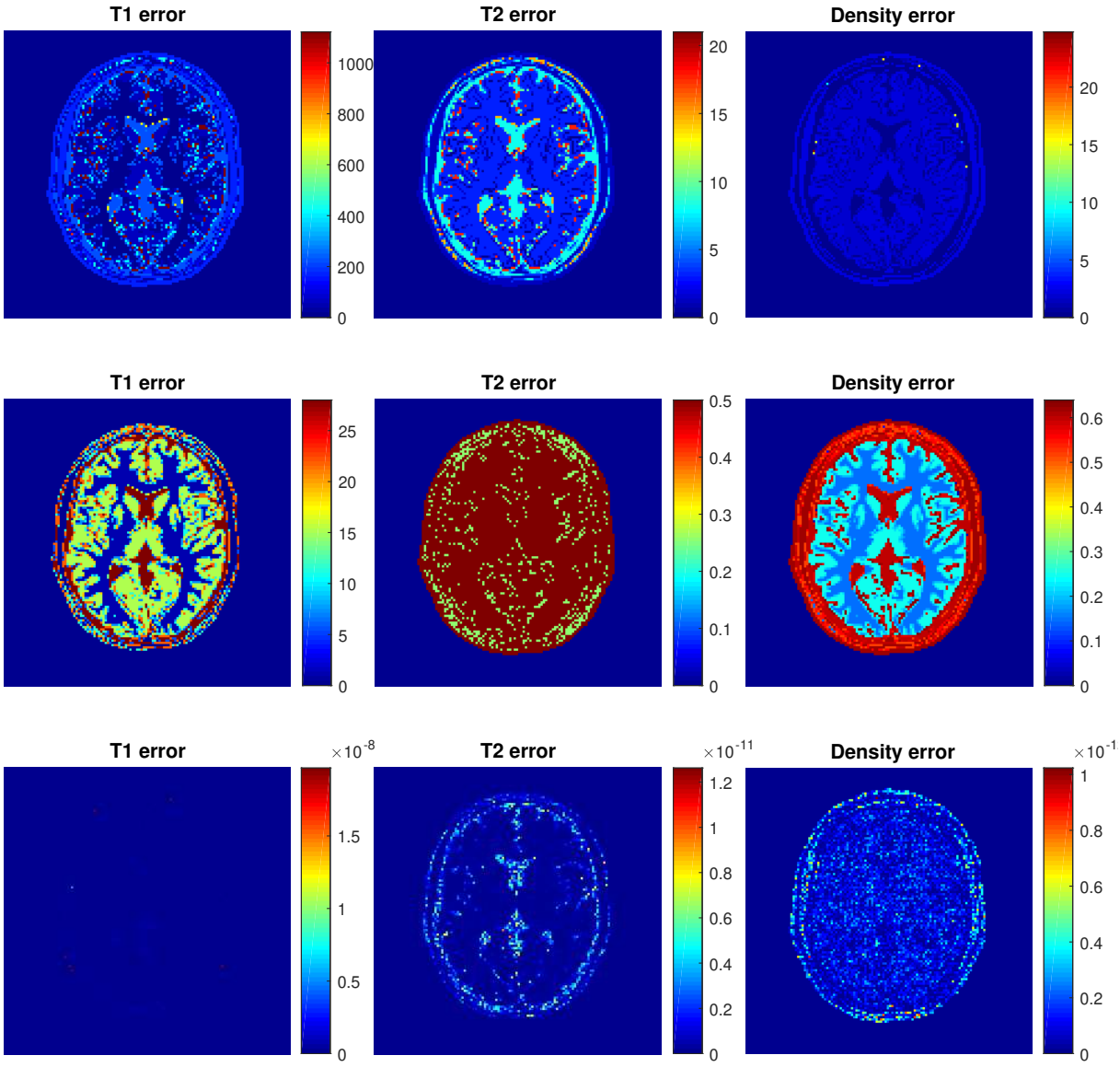


Figure 9: Fully sampled data. Pointwise distance of the solutions of BLIP algorithm and Newton algorithm to the corresponding ground truths. First row: Initial error of BLIP with a coarse dictionary. Middle row: error of BLIP with fine dictionary. Last row: Error of the proposed algorithm.

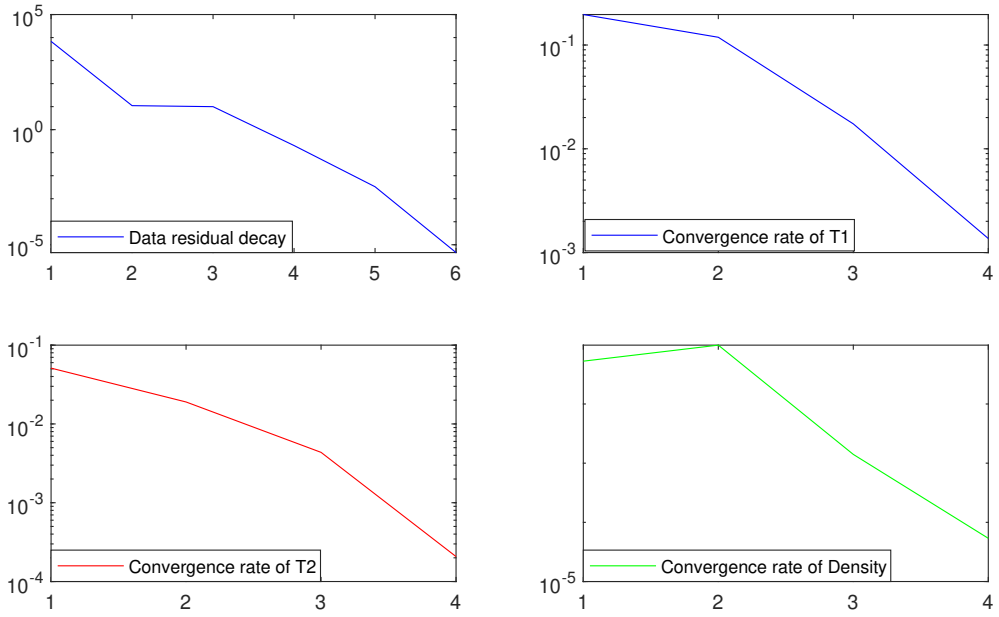


Figure 10: Fully sampled data and superlinear convergence of the Newton method. From left to right and from above to bottom: Convergence of the data residual $\|Q(\mathbf{x}_n) - D\|$, convergence rates via plots of the iterate ratios $\frac{\|x_{n+1}-x_n\|_2}{\|x_n-x_{n-1}\|_2}$ for $x = T_1, T_2, \rho$ respectively.

The model is dictionary-free and incorporates the physical setting of MRI in one single non-linear equation. We have proposed a robust algorithm, that was shown to be capable of estimating the tissue parameters in high precision. In contrast to the original MRF method and its variant BLIP, it does not rely on refining a dictionary to improve the accuracy. Even though our new algorithm is based on a specific Bloch dynamics called IR-bSSFP, this is not the limitation to the method itself. Other type of discrete dynamics or approximations to Bloch equation can be fitted to the current framework. Furthermore, we have considered as unknowns the relaxation parameters $T_1, T_2$ and the proton density $\rho$ in the current paper, but as long as parameters can be described into the Bloch dynamics, there is no difficulty to extend the algorithm to incorporate the new parameters into the current framework.

## Appendix: Solutions of Bloch equations with different cases

Here we briefly review several simplified cases towards the solutions of Bloch equations, which are helpful in order to understand the simulations by using the discrete dynamics. More detailed descriptions can be found in [19, 24]. Note that here we omit the position dependence in Bloch equations.

### Only main field with no relaxation

The Bloch equations in a setting which only takes into account the main magnetic field and with no relaxation is just an autonomous dynamical system. That is

$$\frac{\partial m}{\partial t} = m \times \gamma B_0. \tag{38}$$

The solution in this case is

$$m(t) = P_{\omega_0}(t)m(0),$$

where

$$P_{\omega_0}(t) = \begin{pmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and } \omega_0 = \gamma \left| B_0 \right|.$$

This says that the magnetization precesses about the main magnetic field at a frequency $\omega_0$, called Larmor frequency.

#### Main field with relaxation

This is the case of Bloch equations (1) with $B(t, r) = B_0$. This means that radio frequencies and the gradient fields are not considered here. After some change of variable and calculations, the solution turns out to be

$$m(t) = P_{\omega_0}(t)E(t)m(0) + (1 - e^{-\frac{t}{T_1}})m_e,$$

where

$$E(t) = \begin{pmatrix} e^{-\frac{t}{T_2}} & 0 & 0 \\ 0 & e^{-\frac{t}{T_2}} & 0 \\ 0 & 0 & e^{-\frac{t}{T_1}} \end{pmatrix}.$$

Note that the matrices $P_{\omega_0}(t)$ and $E(t)$ are commutable.

#### With perturbations and without relaxation

By perturbation we mean that there is a $B_1$ field which rotates at the Larmor frequency, and it is always orthogonal to the main field, such that $\langle B_0, B_1 \rangle = 0$. This models the excitation of radio pulses in the MRI machine. By convention, the direction of the $B_1$ field can be defined to be along the $x$-axis in space. Since in reality, the excitation pulse only lasts for a very short length of time in comparison with $T_1$ and $T_2$, we can ignore the relaxation terms. The solution of (1) in the case of no relaxation terms but with perturbation is

$$m(t) = P_{\omega_0}(t)R_x(\alpha(t))m(0),$$

where $\alpha(t) := \gamma \int_0^t |B_1(s)|\, ds$ is the called flip angle, and

$$R_x(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha(t)) & \sin(\alpha(t)) \\ 0 & -\sin(\alpha(t)) & \cos(\alpha(t)) \end{pmatrix}.$$

#### With perturbations and relaxation

Finally we are able to simulate the solutions of (1) in the case where both the perturbations of the main field and relaxation terms are taken into account. This is based on the assumption that the excitation pulse is turned on at the time period $(0, t_0)$, where $t_0$ is a very small number in comparison to the relaxation parameters. Therefore, we can estimate the solution of (1) with the following formula

$$m(t) = P_{\omega_0}(t)E(t)R_x(\alpha(t_0))m(0) + (1 - e^{-\frac{t}{T_1}})m_e.$$

The main tool in all of the above calculations is by changing variables to a rotating frame of reference, in order to match the Larmor precession.

# References


[1] Brainweb: Simulated brain database. *http://www.bic.mni.mcgill.ca/brainweb/*.

[2] J. Assländer, M.A. Cloos, F. Knoll, D.K. Sodickson, J. Hennig, and R. Lattanzi. Low rank alternating direction method of multipliers reconstruction for MR fingerprinting. *Magnetic Resonance Medcine*, 79:83–96, 2018.

[3] A. Bensoussan, G. Da Prato, M.C. Delfour, and S.K. Mitter. *Representation and Control of Infinite Dimensional Systems*. Systems and Control Foundations and Applications. Birkhauser, second edition, 2007.

[4] F. Bloch. Nuclear induction. *Physical Review*, 70:460–473, 1946.

[5] P.G. Ciarlet and C. Mardare. On the Newton-Kantorovich theorem. *Analysis and Applications*, 10(3):249–269, 2012.

[6] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.

[7] M. Davies, G. Puy, P. Vandergheynst, and Y. Wiaux. A compressed sensing framework for magnetic resonance fingerprinting. *SIAM Jounal of Imaging Sciences*, 7(4):2623–2656, 2014.

[8] P. Deuflhard. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, volume 35 of *Springer series in computational mathematics*. Springer, first edition, 2004.

[9] L.S. Evans. *Partial Differential Equations*, volume 19 of *Graduate studies in mathematics*. American Mathematical Society, second edition, 2010.

[10] J. Fan and Y. Yuan. On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. *Computing*, 74:23–39, 2005.

[11] K. Ferentinos. On Tcebycheff's type inequalities. *Trabajos de Estadistica y de Investigacion Operativa*, 33(1):125–132, 1982.

[12] W. Guo, Guohui Song, and Yue Zhang. PCM-TV-TFV: A novel two-stage framework for image reconstruction from fourier data. *SIAM Journal of Imaging Sciences*, 10(4):2250–2274, 2017.

[13] M. Hanke. A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13:79–95, 1997.

[14] M. Hintermüller, C.N. Rautenberg, T. Wu, and A. Langer. Optimal selection of the regularization function in a weighted total variation model. Part II: Algorithm, its analysis and numerical tests. *Journal of Mathematical Imaging and Vision*, 59(3):515–533, 2017.

[15] C. Huang, C.G. Graff, E.W. Clarkson, A. Bilgin, and M.I. Altbach. T2 mapping from highly undersampled data by reconstruction of principal component coefficient maps using compressed sensing. *Magnetic Resonance in Medicine*, 67:1355–1366, 2012.

[16] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics, 2008.

[17] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, volume 6 of *Radon Series on Computational and Applied Mathematics*. De Gruyter, 2008.

[18] F. Knoll, K. Bredies, T. Pock, and R. Stollberger. Second order total generalized variation (TGV) for MRI. *Magnetic Resonance in Medicine*, 65(2):480–491, 2011. http://dx.doi.org/10.1002/mrm.22595.

[19] V. Kuperman. *Magnetic Resonance Imaging: physical principles and applications*. Electromagnetism. Academic Press, first edition, 2000.

[20] H. Larsson, J. Frederiksen, J. Petersen, I. Nordenbo, A.and Zeeberg, O. Henriksen, and J. Olesen. 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. *Magnetic Resonance in Medicine*, 11:337–348, 1989.

[21] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly. Compressed sensing MRI. *IEEE Signal Processing Magazine*, 25:72–82, 2008.

[22] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. Sunshine, J.L. Duerk, and M.A. Griswold. Magnetic resonance fingerprinting. *Nature*, 495(187):187–193, 2013.

[23] G.C. McKinnon. Ultrafast interleaved gradient-echo-planar imaging on a standard scanner. *Magnetic Resonance in Medicine*, 30:609–616, 1993.

[24] D. Nishimura. *Principles of Magnetic Resonance Imaging*. Stanford University, second edition, 2010.

[25] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Series in Electrical Engineering. McGraw Hill Higher Education, third edition, 1991.

[26] A. Sbrizzi, T. Bruijnen, O. van der Heide, P. Luijten, and C.A.T. van den Berg. Dictionary-free MR fingerprinting reconstruction of balanced-GRE sequences. *arXiv preprint, arXiv:1711.08905v1*, 2017.

[27] A. Sbrizzi, O. van der Heide, and M. et. al Cloos. Fast quantitative MRI as a nonlinear tomography problem. *Magnetic Resonance Imaging*, 46:56–63, 2018.

[28] K. Scheffler. A pictorial description of steady-states in rapid magnetic resonance imaging. *Concepts in Magnetic Resonance*, 11(5):187–193, 1999.

[29] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, first edition, 2012.

[30] G.A. Wright. Magnetic resonance imaging. *IEEE Signal Processing Magzine*, 14:56–66, 1997.

[31] N. Yamashita and M. Fukushima. On the rate of convergence of the Levenberg-Marquardt method. In G. Alefeld and X. Chen, editors, *Topics in Numerical Analysis: With Special Emphasis on Nonlinear Problems*, Computing Supplementa book series, pages 239–249. Springer, 2001.