Greedy Approximate Projection for Magnetic Resonance Fingerprinting with Partial Volumes

Greedy Approximate Projection for Magnetic Resonance Fingerprinting with Partial Volumes

Roberto Duarte, Audrey Repetti, Pedro A. Gómez, Mike Davies and Yves Wiaux

Abstract—In the context of quantitative Magnetic Resonance Imaging (qMRI), traditional methods suffer from the so-called Partial Volume Effect (PVE) due to spatial resolution limitations. As a consequence of PVE, the parameters of the voxels containing more than one tissue are not correctly estimated. Magnetic Resonance Fingerprinting (MRF) is not an exception. The existing methods addressing PVE are neither scalable nor accurate. We propose to formulate the recovery of multiple tissues per voxel as a non-convex constrained least-squares minimisation problem. To solve this problem, we develop a greedy approximate projection algorithm called GAP-MRF. Our method adaptively finds the regions of interest on the manifold of fingerprints defined by the MRF sequence, and is memory efficient. In addition, we generalise our method, using an alternating minimisation approach, to compensate for phase errors appearing in the model. We show, through simulations on synthetic data with PVE, that our algorithm outperforms state-of-the-art methods. Our approach is also validated on in vivo data using two different MRF protocols.

Index Terms—MRI, qMRI, MRF, PVE, non-convex, manifold, greedy, iterative projection.

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a powerful tool for diagnosis in medicine. Its main advantage over other medical imaging modalities is that MRI acquisitions are non-ionising and non-invasive. Nevertheless, the main drawback of MRI is that produces qualitative images whose intensity values are a nonlinear response to underpinning physical parameters. Quantitative MRI (qMRI) is a particular modality that aims to produce spatial quantitative maps of parameters related to the tissues under investigation such as T_1 and T_2 relaxation times [1]. Due to prohibitively long acquisition times, this modality is not a standard tool in diagnosis, despite the advantages of quantitative imaging. Motivated by Compressive Sensing (CS) theory, Magnetic Resonance Fingerprinting (MRF) was introduced to accelerate qMRI acquisitions [2]. MRF uses a combination of random excitation pulse sequences and k-space undersampling to simultaneously acquire all relevant quantitative information at once. These random excitation sequences are used to produce unique temporal patterns called *fingerprints*. These patterns are compared to the ones predicted by the model to extract the parameters of interest.

R. Duarte, A. Repetti and Y. Wiaux are with the Institute of Sensors Signals and Systems, Heriot-Watt University, Edinburgh, UK (e-mail: rd8@hw.ac.uk; a.repetti@hw.ac.uk; y.wiaux@hw.ac.uk).

P. A. Gómez is with the Munich School of Bioengineering, Technische Universität München, Munich, Germany (e-mail: pedro.gomez@tum.de).

M. Davies is with the Institute for Digital Communications (IDCom), The University of Edinburgh, EH9 3JL, UK (e-mail: mike.davies@ed.ac.uk)

In general, qMRI techniques, particularly MRF based reconstructions, assume that a voxel contains at most one type of tissue. This assumption is not feasible in practice, and in consequence, voxels containing multiple tissue types may be assigned with completely different parameters. This problem is known as the Partial Volume Effect (PVE) and is present in all the modalities of medical imaging with limited spatial resolution [3]. In Fig. 1, an example of PVE can be seen. The left image shows a spatial distribution of T_1 in a simulated brain. The right image shows the reconstruction using voxels four times bigger. We can observe in this example that the partial volume voxels are not correctly reconstructed.
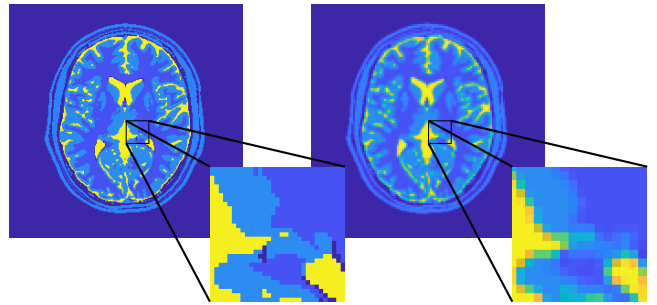

Fig. 1: Partial volume effect in a T_1 parameter map. The image of the left represents the true T_1 parameter map and image on the right is a lower resolution reconstruction.

The PVE has been analysed in the supplementary material of [2]. In this work, the authors propose to decompose the signal using a least-squares method as a weighted sum of at most three distinct signals each representing a different tissue. While the method was shown to be robust to noise for long sequences, it is not well adapted to handle real data. Indeed, in practice we neither have information about the spatial distribution of the partial volume voxels nor the true components of the original signal. Additionally, for short sequences, the noise in the measurements and the sampling of the manifold of fingerprints that describes the signal can significantly affect the estimations. More recently, a Bayesian method was proposed in [4], to tackle the PVE in MRF. The method proposed there is shown to be able to estimate the parameters of the partial volume voxels. However, this estimation comes at the cost of an increased acquisition time with respect to traditional MRF based reconstructions. Indeed, results are presented for three times longer sequences than traditional MRF. This is due to the high aliasing effect encountered with undersampled noisy data. Furthermore, to obtain accurate results, this method relies on a high sampling of the fingerprint manifold, resulting in a high computational cost (in terms of reconstruction time and

memory requirements).

In this paper, we propose to tackle the PVE in the context of the MRF by reformulating the problem as a non-convex constrained least-squares minimisation problem. In particular, we assume that the number of independent tissues in the imaged volume is upper bounded, and that each tissue has a minimum number of pure voxels. In addition, since the parameters of each tissue should be sufficiently different, we add a constraint to distinguish them using information on the neighborhood of each element of the dictionary. To solve the resulting constrained minimisation problem, we develop a greedy approximate projection algorithm called GAP-MRF to solve the non-convex problem. This method works as a projected gradient descent algorithm (also known as iterative projection, or forward-backward algorithm [5, 6]), where the projection is computed inexactly leveraging a greedy approach. In particular, the proposed method for the inexact projection is adaptive, memory efficient, and is a core contribution to our work. Finally, we generalise our method to compensate for phase errors in the model due to timing or coil sensitivity errors, using an alternating minimisation approach [7–11]. We show that our approach outperforms state-of-the-art methods using a simulated partial volume phantom. In addition, we validate our results on in vivo MRF data.

The remainder of the paper is organised as follows: Section II introduces the MRF inverse problem considering only pure voxels and the state-of-the-art method. Section III formulates the target partial volume problem. Section IV proposes the algorithm to solve the partial volume problem. Section V investigates the behaviour of the proposed method on simulated and in vivo data. We conclude in Section VII.

II. MRF PROBLEM DESCRIPTION

A. Inverse problem for single tissue recovery

In the context of MRF, the objective is to estimate the parameters of each voxel in the imaged volume from degraded undersampled measurements. Let $\mathbf{Y} \in \mathbb{C}^{Q \times L \times C}$ be the measurement matrix, where L is the excitation sequence length, C is the number of coils and Q is the number of measurements at each excitation and each coil. Let $\mathbf{M} \in \mathbb{C}^{N \times L}$ be the response of the imaged volume of interest with N voxels. For every $(l, c) \in \{1, \dots, L\} \times \{1, \dots, C\}$, the corresponding observation $\mathbf{Y}_{:,l,c} \in \mathbb{C}^Q$ is given by

$$\mathbf{Y}_{:,l,c} = \mathbf{\Omega}_{:,:,l} \mathbf{F} \mathbf{S}_{:,:,c} \mathbf{M}_{:,l} + \boldsymbol{\eta}_{:,l,c}, \tag{1}$$

where $\mathbf{\Omega} \in \{1, 0\}^{Q \times N \times L}$ is the concatenation of L selection matrices, $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the 2-dimensional discrete Fourier transform, and $\mathbf{S} \in \mathbb{C}^{N \times N \times C}$ is the concatenation of C spatial sensitivity coil diagonal matrices, and $\boldsymbol{\eta} \in \mathbb{C}^{Q \times L \times C}$ is a realisation of a random i.i.d. Gaussian noise. Let $h : \mathbb{C}^{N \times L} \rightarrow \mathbb{C}^{Q \times L \times C}$ be the linear mapping defining the complete acquisition process such that:

$$\mathbf{Y} = h(\mathbf{M}) + \boldsymbol{\eta}. \tag{2}$$

For each voxel $n \in \{1, \dots, N\}$, the magnetisation response $\mathbf{M}_{n,:}$ is modelled through the smooth non-linear mapping $B :$

$\Theta \rightarrow \mathbb{C}^{1 \times L}$ (commonly Bloch equations or Extended Phase Graphs) scaled by the unknown proton density $\rho_n \in \mathbb{R}_+$:

$$\mathbf{M}_{n,:} = \rho_n B(\hat{\theta}_{n,:}, \Gamma), \tag{3}$$

where $\Gamma \in \mathbb{R}^{A \times 1}$ represents the concatenation of A known acquisition parameters (eg., flip angles α, repetition times TR) that are chosen in such a way that $\mathbf{M}_{n,:}$ is only sensitive to the P parameters $\hat{\theta}_{n,:} \in \Theta$ under investigation, where $\Theta \subset \mathbb{R}^{1 \times P}$ denotes the subset of feasible parameters. In the remainder of this paper, we focus on the case when $P = 2$ and Θ corresponds to T_1 and T_2 relaxation times.

B. BLoch response recovery via Iterative Projection (BLIP)

Recently, a full Compressive Sensing (CS) strategy was formulated in [12]. In this work, the authors proposed to solve:

$$\underset{\mathbf{M} \in \mathcal{B}^+}{\text{minimise}} \frac{1}{2} \| \mathbf{Y} - h(\mathbf{M}) \|_2^2, \tag{4}$$

where

$$\mathcal{B}^+ = \left\{ \mathbf{M} \in \mathbb{C}^{N \times L} \mid (\forall n \in \{1, \dots, N\}) \, \mathbf{M}_{n,:} = \rho \mathbf{m}, \right.$$
$$\left. \text{with } \rho \in \mathbb{R}_+ \text{ and } \mathbf{m} \in B(\Theta, \Gamma) \right\}. \tag{5}$$

In [12], an iterative projection algorithm called BLIP has been developed to solve (4). This method imposes that $\mathbf{M}$ belongs to $\mathcal{B}^+$ by computing at iteration i:

$$\mathbf{M}^{(i+1)} = \mathcal{P}_{\mathcal{B}^+} \left(\mathbf{M}^{(i)} - \mu h^\dagger \left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y} \right) \right), \tag{6}$$

where $\mu > 0$ is a step size chosen using a backtracking method and $h^\dagger$ is the adjoint operator of h, and $\mathcal{P}_{\mathcal{B}^+}$ represents the projection onto the set $\mathcal{B}^+$. The authors also derived a technical condition on both L and the undersampling ratio N/Q, for the Restricted Isometry Property (RIP) to be satisfied by the operator h which guarantees robust recovery. In this work, the authors interpreted the dictionary $\Phi \in \mathbb{C}^{D \times L}$ of fingerprints, introduced in [2], as a discrete sampling of the low dimensional manifold B, and the matched filter as the projection $\mathcal{P}_{\mathcal{B}^+}$. Φ is constructed from D samples of Θ, stored in a matrix $\theta \in \mathbb{R}^{D \times P}$. The matched filter is computed for each voxel n, by finding the closest sample in the dictionary $\Phi_{\hat{d}_n,:}$:

$$\hat{d}_n = \underset{d}{\text{argmax}} \frac{\text{real}\left(\mathbf{M}_{n,:} [\Phi_{d,:}]^\dagger \right)}{\|\Phi_{d,:}\|_2}, \tag{7}$$

and scaling it using the proton density calculated by:

$$\hat{\rho}_n = \max \left(\frac{\text{real}\left(\mathbf{M}_{n,:} \left[\Phi_{\hat{d}_n,:}\right]^\dagger \right)}{\|\Phi_{\hat{d}_n,:}\|_2^2}, 0 \right). \tag{8}$$

As mentioned in [12], the proton density ρ is non-negative and real valued. However, it is a common practice to allow the proton density to absorb phase terms to correct for timing or coil sensitivity errors. If ρ_n is allowed to be complex, then the projection can be reformulated as:

$$\hat{d}_n = \underset{d}{\operatorname{argmax}} \frac{\left| \mathbf{M}_{n,:} \left[\Phi_{d,:} \right]^\dagger \right|}{\| \Phi_{d,:} \|_2}, \qquad (9)$$

and

$$\hat{\rho}_n = \frac{\mathbf{M}_{n,:} \left[\Phi_{\hat{d}_n,:} \right]^\dagger}{\| \Phi_{\hat{d}_n,:} \|_2^2}. \qquad (10)$$

In order to obtain the parameters of interest $\hat{\theta} \in \mathbb{R}^{N \times P}$, the matrix θ is used as Look-Up Table (LUT). Inaccuracies related to the manifold discretisation can significantly affect the quality of the reconstructions. Since the exact parameter values for the tissues are unknown, a common approach to achieve good quality reconstruction is to densely sample the manifold. However, this strategy introduces memory and computational issues due to the dictionary size.

In reality, volumetric images have a finite resolution introducing voxels that contain more than one element [3]. This leads to a signal consisting in a linear combination of the signals produced by each tissue in the voxel. This is called the PVE and is present in all the modalities of medical imaging. By not taking PVE into account in the model results in parameter mismatch.

III. PROPOSED PARTIAL VOLUME MODEL

The model described in the previous section considers that each voxel contains at most one element. Partial volume voxels are introduced due to the spatial discretisation in the acquisition process. When partial volume voxels are involved, the magnetisation sequence can be described as:

$$\mathbf{M} = \mathbf{X}\Phi, \qquad (11)$$

where $\mathbf{X} \in \mathbb{R}_+^{N \times D}$ is a mixing matrix containing the proton densities of the elements constituting each voxel, and $\Phi \in \mathbb{C}^{D \times L}$ is the over-complete dictionary described in Section II-B. Since only few elements per voxel are expected, the mixing matrix $\mathbf{X}$ is sparse.

As a consequence of the smoothness of B, Φ is highly coherent. In this context, the estimation of $\mathbf{X}$ from highly undersampled noisy data is expected to fail without additional priors. Leveraging CS theory [13–16], we can find the sparsest matrix $\mathbf{X}$ that is sufficiently close to the measurements by solving:

$$\underset{\mathbf{X} \in \mathbb{R}_+^{N \times D}}{\operatorname{minimise}} \|\mathbf{X}\|_0 \text{ subject to } \|\mathbf{Y} - h(\mathbf{X}\Phi)\|_2 \leq \epsilon, \qquad (12)$$

where $\epsilon > 0$ is a bound chosen according to the acquisition noise level. In (12), $\|.\|_0$ denotes the ℓ_0 pseudo-norm [13], counting the non-zero entries of its argument, and is defined as:

$$\|\mathbf{X}\|_0 = \sum_i \sum_j \left(\mathbf{X}_{i,j} \right)^0, \qquad (13)$$

with the convention $0^0 = 0$. Since this function is non-convex and non-differentiable, problem (12) is difficult to solve in practice, in particular in the context of high dimensional problems (usually, D is of the order of 10^6 and L of the

order of 10^3). Thus, the non-convexity of the ℓ_0 pseudo-norm is often relaxed by the use of the ℓ_1-norm [17]. Nevertheless, Φ being highly coherent, this convex relaxation cannot be used to correctly estimate the coefficients of $\mathbf{X}$ [18]. To overcome these difficulties, similarly to the BLIP approach, we propose to reformulate problem (12) as:

$$\underset{\mathbf{M} \in \mathcal{B}_{\mathcal{S}_+}(\Phi)}{\operatorname{minimise}} \frac{1}{2} \|\mathbf{Y} - h(\mathbf{M})\|_2^2 \qquad (14)$$

where

$$\mathcal{B}_{\mathcal{S}_+}(\Phi) = \left\{ \mathbf{M} \in \mathbb{C}^{N \times L} \,|\, \mathbf{M} = \mathbf{X}\Phi \text{ with } \mathbf{X} \in \mathcal{S}_+ \right\}, \quad (15)$$

$$\mathcal{S}_+ = \underset{s=1}{\overset{4}{\cap}} \mathcal{S}_s, \qquad (16)$$

and, for every $s \in \{1, \ldots, 4\}$, $\mathcal{S}_s$ is a closed non-empty subset of $\mathbb{R}^{N \times D}$, used to impose feasibility constraints on $\mathbf{X}$. These sets are defined in the following subsections.

A. Positivity Constraint

Since the proton densities of the imaged volume must be non-negative, we can restrict our solution to be in the positive orthant:

$$\mathcal{S}_1 = \mathbb{R}_+^{N \times D}. \qquad (17)$$

B. Constraint on the number of tissues

Commonly MRF aims to obtain quantitative values of a small set of tissues. In practice, only $T \ll D$ elements of the dictionary Φ are necessary to characterise $\mathbf{M}$. While T is unknown, we have a reasonable estimate for it. We propose to introduce a loose upper bound K, such that $T \leq K \leq D$, to limit the number of active dictionary elements (although K appears to be a free parameter, in practice we choose it to be at least four times greater than T for our method to work for low noise scenarios and at least eight times greater than T for high noise scenarios). Let us define a set $\mathcal{D}_\mathbf{X}$ that is formed by the column indices of $\mathbf{X}$ with non-zero coefficients. To avoid noisy voxels, only rows with proton density greater than $\xi > 0$ (chosen according to the noise level) will be considered. Formally, this set is defined as:

$$\mathcal{D}_\mathbf{X} = \{d \in \{1, \ldots, D\} \,|\, (\exists n \in \mathcal{G}_\mathbf{X}) \, \mathbf{X}_{n,d} \neq 0\}, \qquad (18)$$

where

$$\mathcal{G}_\mathbf{X} = \{n \in \{1, \ldots, N\} \,|\, \|\mathbf{X}_{n,:}\|_1 > \xi\} \qquad (19)$$

and $\|.\|_1$ denotes the ℓ_1-norm. The set $\mathcal{D}_\mathbf{X}$ indicates the columns of $\mathbf{X}$ contributing to the magnetisation sequence. We can limit the number of used elements of the dictionary by constricting the cardinality of this set, denoted by $\operatorname{Card}(\mathcal{D}_\mathbf{X})$ to be upper bounded by K:

$$\mathcal{S}_2 = \left\{ \mathbf{X} \in \mathbb{R}^{N \times D} \,|\, \operatorname{Card}(\mathcal{D}_\mathbf{X}) \leq K \right\}. \qquad (20)$$

C. Constraint on voxel's neighbourhoods

The tissues of interest are unique and need to be sufficiently different to be distinguished. In order to incorporate this prior information in the reconstruction process, we define the neighbour set associated to each element $d \in \{1, \ldots, D\}$ of the dictionary as:

$$\mathcal{N}_v(d) = \{d' \in \{1, \ldots, D\}\backslash\{d\}\,| \\ (\forall p = \{1, \ldots, P\})\,|\theta_{d',p} - \theta_{d,p}| < \upsilon\theta_{d,p}\}, \quad (21)$$

where $\upsilon > 0$. We can define a set of all possible $\mathbf{X}$ such that, the parameters of each element in $\mathcal{D}_{\mathbf{X}}$ are sufficiently far from each other. This can be achieved by constricting all the neighbour columns of each element in $\mathcal{D}_{\mathbf{X}}$ to be the null element of $\mathbb{R}^N$ denoted by $\mathbf{0}$:

$$\mathcal{S}_3 = \left\{\mathbf{X} \in \mathbb{R}^{N \times D}\,|\,\left(\forall d' \in \underset{d\in\mathcal{D}_{\mathbf{X}}}{\cup}\mathcal{N}_v(d)\right)\mathbf{X}_{:,d'} = \mathbf{0}\right\} \tag{22}$$

D. Constraint on the pure voxels

Due to the additive noise in model (2), some elements of $\mathbf{X}$ corresponding to non-used dictionary elements take non-zero values. In order to avoid these noisy elements in the reconstructions, we impose that at least $\kappa > 0$ rows (i.e. voxels) of $\mathbf{X}$ contain only one non-zero value for each active column of $\mathbf{X}$ (our method uses these voxels to explore the manifold of fingerprints, and we heuristically choose it depending on the noise level). These rows identify the *pure voxels*. This constraint can be formulated as follows:

$$\mathcal{S}_4 = \left\{\mathbf{X} \in \mathbb{R}^{N \times D}\,|\,(\forall d \in \mathcal{D}_{\mathbf{X}})\,\|\,(\mathbf{X}_{n,d})_{n\in\mathcal{V}_{\mathbf{X}}}\,\|_0 \geq \kappa\right\} \tag{23}$$

where

$$\mathcal{V}_{\mathbf{X}} = \{n \in \{1, \ldots, N\}\,|\,\|\mathbf{X}_{n,:}\|_0 = 1\}. \tag{24}$$

IV. GREEDY APPROXIMATE PROJECTION FOR MRF

A. Proposed iterative projection algorithm

In order to solve problem (14), as for BLIP algorithm, we rely on an iterative projection method [19]. At each iteration $i \in \mathbb{N}$, this method consists in updating $\mathbf{M}^{(i+1)}$ by computing a gradient step followed by a projection step:

$$\mathbf{M}^{(i+1)} = \mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}\left(\mathbf{M}^{(i)} - \mu h^\dagger\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)\right), \tag{25}$$

where $\mu > 0$. In [12], it is shown that choosing $\mu \approx N/Q$ is theoretically justifiable. However, in order to ensure the stability of the iterative projection algorithm and accelerate convergence, in [12, 20] the authors proposed to choose μ using a backtracking method. In order to handle efficiently the constraint $\mathcal{B}_{S_+}(\Phi)$, we propose to compute inexactly the projection onto this set in (25). The resulting method, named Greedy Approximate Projection for MRF (GAP-MRF), is described in Algorithm 1. We can notice that the GAP-MRF method and BLIP are solving similar problems, using the same algorithmic structure. In this context, as in [12], a condition on both L and the undersampling ratio N/Q might be derived for recovery guarantee. However, the investigation of such condition is beyond the scope of this article.

Algorithm 1 GAP-MRF global iterations

1: **Input:** $\mathbf{Y} \in \mathbb{C}^{Q \times L \times C}$, $\zeta < 1$, $\mathbf{M}^{(0)} \in \mathbb{C}^{N \times L}$
2: **Iterations:**
3: **for** $i = 0, 1, \ldots$ **do**
4: $\qquad \mu = 2N/Q$, $\nu = 0$
5: $\qquad$ **while** $\mu > \nu$ **do**
6: $\qquad\qquad \mu = \mu/2$
7: $\qquad\qquad$ **Gradient Step:**
8: $\qquad\qquad \overline{\mathbf{M}}^{(i)} = \mathbf{M}^{(i)} - \mu h^\dagger\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)$
9: $\qquad\qquad$ **Projection Step:**
10: $\qquad\qquad \mathbf{M}^{(i+1)} \approx \mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}\left(\overline{\mathbf{M}}^{(i)}\right)$
11: $\qquad\qquad$ **Backtracking step**
12: $\qquad\qquad \nu = \zeta\dfrac{\|\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\|_2^2}{\|h\left(\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\right)\|_2^2}$
13: $\qquad$ **end while**
14: **end for**

B. Approximate projection

To compute the projection onto $\mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}$, one can notice that, for every $\overline{\mathbf{M}} \in \mathbb{C}^{N \times L}$:

$$\mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}\left(\overline{\mathbf{M}}\right) = \underset{\mathbf{M}\in\mathcal{B}_{S_+}(\Phi)}{\operatorname{argmin}}\frac{1}{2}\|\mathbf{M} - \overline{\mathbf{M}}\|_2^2$$

$$= \underset{\mathbf{M}=\mathbf{X}\Phi,\,\mathbf{X}\in\mathcal{S}_+}{\operatorname{argmin}}\frac{1}{2}\|\mathbf{X}\Phi - \overline{\mathbf{M}}\|_2^2$$

$$= \left(\underset{\mathbf{X}\in\mathcal{S}_+}{\operatorname{argmin}}\frac{1}{2}\|\mathbf{X}\Phi - \overline{\mathbf{M}}\|_2^2\right)\Phi \tag{26}$$

Note that $\mathcal{S}_2, \mathcal{S}_3$ and $\mathcal{S}_4$, can be handled through the definition of Φ. Let $\mathbf{M} = \mathbf{X}\Phi \in \mathcal{B}_{S_+}(\Phi)$. For $T \leq K$ (K being the upper bound defined in (20)), let $\mathbf{U} \in \mathbb{R}^{N \times T}$ be a matrix constructed with the columns of $\mathbf{X}$ containing at least a value different from 0, and $\Delta \in \mathbb{C}^{T \times L}$ be the corresponding dictionary elements such that $\mathbf{X}\Phi = \mathbf{U}\Delta$. Then (26) can be rewritten as follows:

$$\mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}\left(\overline{\mathbf{M}}\right) = \left(\underset{\mathbf{U}\in\mathbb{R}_+^{N\times T}}{\operatorname{argmin}}\frac{1}{2}\|\mathbf{U}\Delta - \overline{\mathbf{M}}\|_2^2\right)\Delta. \tag{27}$$

In (27), the dictionary Δ is defined as

$$\Delta = \underset{\overline{\Delta}\in\mathcal{C}}{\operatorname{argmin}}\left(\underset{\overline{\mathbf{U}}\in\mathbb{R}_+^{N\times T}}{\min}\frac{1}{2}\|\overline{\mathbf{U}}\,\overline{\Delta} - \overline{\mathbf{M}}\|_2^2\right), \tag{28}$$

where $\mathcal{C}$ is the set given by

$$\mathcal{C} = \left\{\Delta \in \mathbb{C}^{T\times L}\,|\,(\exists\mathbf{X}\in\mathcal{S}_+)\,\mathbf{X} = \mathcal{Z}(\overline{\mathbf{U}})\right. \\ \left.\text{with }\overline{\mathbf{U}} = \underset{\mathbf{U}\in\mathbb{R}_+^{N\times T}}{\operatorname{argmin}}\frac{1}{2}\|\mathbf{U}\Delta - \overline{\mathbf{M}}\|_2^2\right\}, \tag{29}$$

with $\mathcal{Z}\colon \mathbb{R}_+^{T\times L} \to \mathbb{R}_+^{D\times L}$ the linear operator defined such that $\mathcal{Z}(\mathbf{U})\Phi = \mathbf{U}\Delta$.

As mentioned earlier, Φ is an over-complete dictionary, which makes the exact projection practically impossible to compute. To overcome this difficulty, we propose a greedy algorithm to approximate the projection by finding a dictionary

$\widetilde{\Delta}$ and a mixing matrix $\widetilde{\mathbf{U}}$ such that $\mathbf{U}\Delta \approx \widetilde{\mathbf{U}}\widetilde{\Delta}$. Then the projection in line 10 of Algorithm 1 can be approximated as follows:

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\Phi)}\left(\overline{\mathbf{M}}\right) \approx \left(\underset{\widetilde{\mathbf{U}}\in\mathbb{R}_+^{N\times T}}{\arg\min} \frac{1}{2}\|\widetilde{\mathbf{U}}\widetilde{\Delta} - \overline{\mathbf{M}}\|_2^2\right)\widetilde{\Delta}. \quad (30)$$

As discussed in the previous section, in practice phase errors are appearing in the model, we propose to incorporate an orthogonal diagonal matrix $\Lambda \in \mathbb{C}^{N\times N}$ to compensate for these errors. The approximate projection with the phase compensation is then given by:

$$\mathcal{P}_{\widetilde{\mathcal{B}}_{\mathcal{S}_+}(\Phi)}\left(\overline{\mathbf{M}}\right) \approx \Lambda\widetilde{\mathbf{U}}\widetilde{\Delta}, \quad (31)$$

where Λ and $\widetilde{\mathbf{U}}$ are obtained by solving:

$$\underset{\Lambda\in\mathbb{C}^{N\times N}, \widetilde{\mathbf{U}}\in\mathbb{R}_+^{N\times T}}{\arg\min} \frac{1}{2}\|\Lambda\widetilde{\mathbf{U}}\widetilde{\Delta} - \overline{\mathbf{M}}\|_2^2$$
$$\text{subject to } |\Lambda_{n,n'}| = \begin{cases} 1, & \text{if } n = n', \\ 0, & \text{otherwise.} \end{cases} \quad (32)$$

It is worth mentioning that in (30) and (32), all the rows of $\widetilde{\mathbf{U}}$ can be computed independently in parallel.

On the one hand, forward-backward based algorithms [6, 21, 22] can be used to solve problem (30) (in particular, in our simulations, we use the built-in Matlab function of non-negative least-squares, that is an implementation of [23]). On the other hand, to solve problem (32), to jointly estimate Λ and $\mathbf{U}$, block coordinate approaches must be considered (e.g. Gauss-Seidel approaches [7], alternating forward-backward methods [8–11]). Note that in comparison with the traditional MRF methods which densely sample the manifold, our approach reduces the memory requirements, by using the dictionary $\widetilde{\Delta}$ containing at most K elements, without the inaccuracies related to the manifold discretisation.

The greedy algorithm to estimate $\widetilde{\Delta}$ is described in the next section.

C. Greedy Dictionary Estimation

The GAP-MRF algorithm takes advantage of the dictionary coherence and the constraints imposed on $\mathbf{X}$ (described in Section III) to approximate the projection onto $\mathcal{B}_{\mathcal{S}_+}(\Phi)$ in line 10 of Algorithm 1. As described in Section IV-B, this projection can be approximated at each iteration $i \in \mathbb{N}$, by solving (30), which necessitates to estimate the dictionary $\widetilde{\Delta}^{(i)}$. We propose to estimate it using a greedy approach, leveraging both the knowledge of $\overline{\mathbf{M}}^{(i)}$ and the properties of the sets $\mathcal{S}_{2\leq s\leq 4}$ (note that the constraint $\mathcal{S}_1$ is handled directly in (30)). The proposed approach is described in details in this section.

The process to obtain $\widetilde{\Delta}^{(i)}$ consists in three main steps leveraging the set of pure voxels. The first step consists in approximating the parameters of the pure voxels ($\mathcal{S}_4$ constraint) using the projection onto $\mathcal{B}^+$ (the definition of $\mathcal{B}^+$ is given in 5). The objective of the second step is to find K regions of interest ($\mathcal{S}_2$ constraint) of the manifold by exploiting its smoothness. Finally, in the third step, the
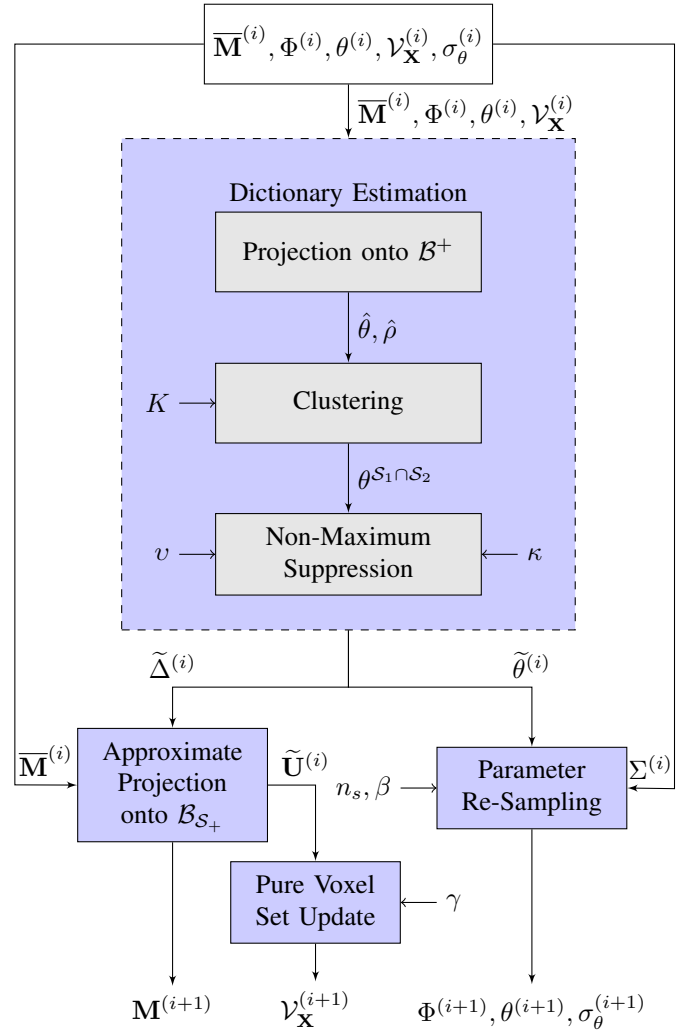

Fig. 2: Greedy approximate projection diagram. The blue boxes represent the main steps in the approximate projection, the gray boxes represent the intermediate steps for the dictionary estimation and the arrows show the input and output variables.

parameters that are too close to each other are discarded ($\mathcal{S}_3$ constraint) by using a Non-Maximum Suppression based method [24]. This method acts on the number of voxels that corresponds to each parameter and keeps only the elements which have enough pure voxels to satisfy the $\mathcal{S}_4$ constraint. This process is summarised in the dictionary estimation step on Fig. 2. The remaining blue blocks in the diagram are used to update the variables in the greedy approximate projection. More precisely, we compute the mixing matrix $\widetilde{\mathbf{U}}^{(i)}$ and the magnetisation sequence $\mathbf{M}^{(i+1)}$ using equation (30) with the resulting dictionary $\widetilde{\Delta}^{(i)}$. Then, we update the pure voxel set $\mathcal{V}_{\mathbf{X}}$ using the mixing matrix $\widetilde{\mathbf{U}}^{(i)}$. Finally, the dictionary Φ is refined by randomly sampling around the parameters $\widetilde{\theta}^{(i)}$. The complete method is described in Algorithm 2 and explained in the following paragraphs.

1) Projection onto $\mathcal{B}^+$: At iteration $i \in \mathbb{N}$, we have:

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\Phi)}\left(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}\right) = \mathcal{P}_{\mathcal{B}^+(\Phi)}\left(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}\right), \quad (33)$$

Algorithm 2 Greedy Approximate Projection

1: **Input:** $\overline{\mathbf{M}}^{(i)}, \Phi^{(i)}, \theta^{(i)}, \mathcal{V}_{\mathbf{X}}^{(i)}, \Sigma^{(i)}, K, \Gamma, \kappa, \upsilon, \gamma, \beta, \xi, n_s$

2:   **Dictionary Estimation:**

3:     **Projection onto $\mathcal{B}^+$**

4:       **for** $n = 1, 2, ..., N$ **do**

5:         $\hat{d}_n = \underset{d}{\operatorname{argmax}} \ \operatorname{real}\left(\overline{\mathbf{M}}_{n,:}^{(i)} \Phi_{d,:}^{\dagger(i)}\right) / \|\Phi_{d,:}^{(i)}\|_2$

6:         $\hat{\rho}_n = \max\left(\operatorname{real}\left(\overline{\mathbf{M}}_{n,:}^{(i)} \Phi_{\hat{d}_n,:}^{\dagger(i)}\right) / \|\Phi_{\hat{d}_n,:}^{(i)}\|_2^2, 0\right)$

7:         $\hat{\theta}_{n,:} = \theta_{\hat{d}_n,:}^{(i)}$

8:       **end for**

9:   **Clustering**

10:     $\mathcal{I} = \{n \in \mathcal{V}_{\mathbf{X}}^{(i)} \mid \hat{\rho}_n > \xi\}$

11:     $[\theta^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}] = \text{k-means}\left(\hat{\theta}_{\mathcal{I},:}, K\right)$

12:   **Non-Maximum Suppression**

13:     $\widetilde{\theta}^{(i)} = \text{NonMaximumSuppression}\left(\theta^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}, \upsilon, \kappa\right)$

14:     $\widetilde{\Delta}^{(i)} = B(\widetilde{\theta}^{(i)}, \Gamma)$

15:   **Approximate Projection onto $\mathcal{B}_{\mathcal{S}_+}$**

16:     $\widehat{\mathbf{U}}^{(i)} = \underset{\overline{\mathbf{U}} \in \mathbb{R}_+^{N \times T}}{\operatorname{argmin}} \frac{1}{2}\|\overline{\mathbf{U}}\widetilde{\Delta}^{(i)} - \overline{\mathbf{M}}^{(i)}\|_2^2$

17:     $\mathbf{M}^{(i+1)} = \widetilde{\mathbf{U}}^{(i)}\widetilde{\Delta}^{(i)}$

18:   **Pure Voxel Set Update**

19:     $\mathcal{G}_{\mathbf{X}}^{(i)} = \left\{n \in \{1, \ldots, N\} \mid \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\right\}$

20:     $\mathcal{V}_{\mathbf{X}}^{(i+1)} = \left\{n \in \mathcal{G}_{\mathbf{X}}^{(i)} \mid \max\left(\widetilde{\mathbf{U}}_{n,:}^{(i)}\right) \geq \gamma\|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1\right\}$

21:   **Parameter Re-sampling**

22:     $\theta^{(i+1)} = \text{ParameterReSampling}\left(\widetilde{\theta}^{(i)}, \Sigma^{(i)}, n_s\right)$

23:     $\Phi^{(i+1)} = B(\theta^{(i+1)}, \Gamma)$

24:     $\sigma_{\theta}^{(i+1)} = \sigma_{\theta}^{(i)}\beta$

25: **Output:** $\theta^{(i+1)}, \Phi^{(i+1)}, \sigma_{\theta}^{(i+1)}, \mathcal{V}_{\mathbf{X}}^{(i+1)}$ and $\mathbf{M}^{(i+1)}$
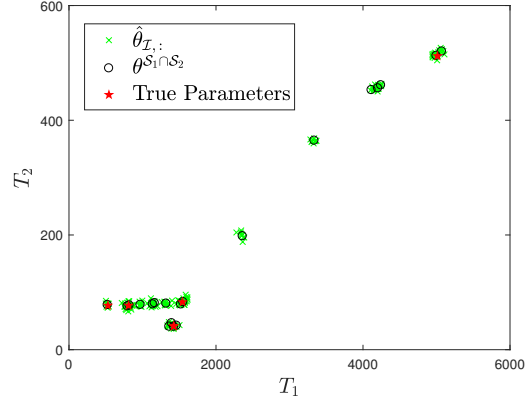

Fig. 3: Clustering example. The green crosses represent the parameters of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ which its corresponding proton density is greater than ξ. The black circles represent the centers obtained by the k-means algorithm. The red stars are the true phantom parameters.

where $\mathcal{B}^+$ is the set defined in equation (5), and $\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:} = \left(\overline{\mathbf{M}}_{n,:}\right)_{n \in \mathcal{V}_{\mathbf{X}}^{(i)}}$, $\mathcal{V}_{\mathbf{X}}^{(i)}$ corresponding to an estimate of the pure voxel positions in $\mathbf{X}^{(i)}$ at iteration i (the true set $\mathcal{V}_{\mathbf{X}}$ corresponding to the pure voxels of the original $\mathbf{X}$ being unknown). At the first iteration, we choose $\mathcal{V}_{\mathbf{X}}^{(0)} = \{1, \ldots, N\}$, and the set is updated during the greedy process (see Algorithm 2, step 20).

From (33), we can estimate the parameters $\hat{\theta}$ and the proton density $\hat{\rho}$ of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ using equations (7)-(8) with a dictionary $\Phi^{(i)}$ (we construct $\Phi^{(0)}$ by coarsely sampling Θ). $\Phi^{(i)}$ is an adaptive dictionary that is refined at each iteration to reduce the computational cost, the simulations suggest that the accuracy of the reconstructions is preserved. Since there are at least κ pure voxels for each active element in Φ and the value of the proton density is at least ξ, we expect that the voxel parameters in $\mathcal{V}_{\mathbf{X}}^{(i)}$ with $\hat{\rho} > \xi$ will form clusters around the true values of the dictionary elements, an example can be seen in Fig. 3.

2) Clustering: In order to find K centers approximating the parameters of interest, we propose to use the k-means

algorithm [25]. It is important to emphasize that the parameter K should be large enough to avoid partial volume voxels to significantly affect the estimation. The objective of k-means is to find K centers that minimise the squared distance from all points to its closest center. Our simulations suggest to choose K at least four times larger than the expected number of tissues to avoid partial volume voxels to shift the centers. The centers obtained by solving the k-means problem $\theta^{\mathcal{S}_1 \cap \mathcal{S}_2} \in \mathbb{R}^{K \times P}$ can be used to compute a dictionary $\Delta^{\mathcal{S}_1 \cap \mathcal{S}_2} \in \mathbb{C}^{K \times L}$. By solving equation (30) with $\Delta^{\mathcal{S}_1 \cap \mathcal{S}_2}$, we would obtain a $\mathbf{U}^{\mathcal{S}_1 \cap \mathcal{S}_2} \in \mathbb{R}^{N \times K}$ such that $\mathcal{Z}\left(\mathbf{U}^{\mathcal{S}_1 \cap \mathcal{S}_2}\right) \in \mathcal{S}_1 \cap \mathcal{S}_2$.

3) Non-Maximum Suppression: The k-means algorithm also provides a label to each voxel corresponding to the matched center. We define $\mathbf{c} \in \mathbb{R}^{K \times 1}$ to be the vector containing the number of voxels associated with each center. Inspired by the Non-Maximum Suppression method in [24], we use the number of pure voxels assigned to each center to remove the neighbours defined in equation (21). We first take the parameters of the highest value of $\mathbf{c}$, and we add all the $\mathbf{c}$ values of the neighbours to the maximum value of $\mathbf{c}$ if it is greater than κ we keep the parameters, if not we discard them and set the corresponding values of $\mathbf{c}$ to 0 (see Fig. 4). We repeat the process until all values of $\mathbf{c}$ are 0. Finally, we use the resulting parameters $\widetilde{\theta}^{(i)} \in \mathbb{R}^{T \times P}$ to construct $\widetilde{\Delta}^{(i)} \in \mathbb{C}^{T \times L}$.

4) Inexact Projection onto $\mathcal{B}_{\mathcal{S}_+}$: Once the dictionary $\widetilde{\Delta}^{(i)}$ is approximated, computing the three steps described above, the magnetisation sequence $\mathbf{M}^{(i+1)}$ can be updated. To this aim, we use equation (30), where the minimisation problem is solved using Matlab built-in function for non-negative least-squares problems [23].

5) Pure Voxel Set Update: In order to avoid noisy voxels, we use the set defined in (19) for $\widetilde{\mathbf{U}}^{(i)}$. Note that $\mathcal{Z}\left(\widetilde{\mathbf{U}}^{(i)}\right)$ is a matrix of the size of $\mathbf{X}$ filling the missing values of $\widetilde{\mathbf{U}}^{(i)}$ with zeros, and thus we can re-define the set $\mathcal{G}_{\mathbf{X}}^{(i)}$ in terms of $\widetilde{\mathbf{U}}^{(i)}$ as:

$$\mathcal{G}_{\mathbf{X}}^{(i)} = \left\{n \in \{1, \ldots, N\} \mid \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\right\}. \tag{34}$$
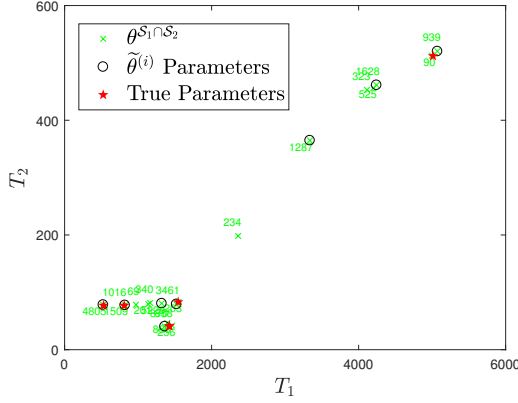
Fig. 4: Non-maximum suppression example. The green crosses represent the centers obtained by the k-means and the green numbers represent the corresponding number of voxels assigned to that center. The black circles represent the parameters obtained by the Non-Maximum Suppression. The red stars are the true phantom parameters.

Then, we update the pure voxel set as:

$$\mathcal{V}_{\mathbf{X}}^{(i+1)} = \left\{ n \in \mathcal{G}_{\mathbf{X}}^{(i)} \mid \max \left(\widetilde{\mathbf{U}}_{n,:}^{(i)} \right) \geq \gamma \| \widetilde{\mathbf{U}}_{n,:}^{(i)} \|_1 \right\}, \quad (35)$$

where $0 < \gamma < 1$ is a relaxation factor used to compensate both for the noise and for the fact that the true dictionary elements are not guaranteed to be present. In our simulations, this parameter is fixed to be $\gamma = 0.85$.

6) Parameter Re-sampling: We update $\Phi^{(i)}$ to refine the manifold elements of interest. For this process, we produce n_s random samples around the elements in $\widetilde{\theta}^{(i)}$ using a Gaussian distribution with a diagonal covariance matrix $\Sigma^{(i)}$ (see Fig. 5). The values of the covariance matrix $\Sigma^{(i)}$ are reduced by a factor $0 < \beta < 1$ at each iteration. When the values of $\Sigma^{(i)}$ are sufficiently small, the dictionary $\widetilde{\Delta}$ will not change anymore and after a fixed number of iterations the sequences generated by Algorithm 1 will stabilise. Since the samples are randomly Gaussian distributed, the parameter values are not limited to a given resolution.
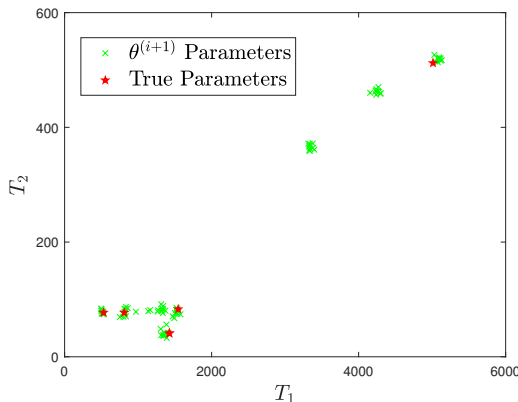

Fig. 5: Parameter re-sampling example. The green crosses represent the parameters of the dictionary $\Phi^{(i+1)}$ obtained by randomly sampling around the parameters obtained by the Non-Maximum Suppression. The red stars represent the true phantom parameters.

7) Initialisation: Since the set $\mathcal{S}_+$ is non-convex, the choice of the initialisation is important in practice. If the initial magnetisation sequence or the dictionary are not close to the desired values, the greedy approximate projection can fail. To tackle this problem, we propose to run the algorithm with $\kappa = 0$ and $\upsilon = 0$ and fixing $\mathcal{V}_{\mathbf{X}}^{(i)} = \{1, \ldots, N\}$ for few iterations (we choose 30 iterations in our simulations). Note that the values of Σ should be large enough to explore Θ (for our simulations, we choose $\Sigma_{1,1}^{(0)} = 200$, $\Sigma_{2,2}^{(0)} = 50$). Then we use the resulting dictionary Φ, the corresponding parameters θ and the magnetisation sequence $\mathbf{M}$ as initialisation for the GAP-MRF algorithm.

V. SIMULATIONS AND RESULTS

In this section, we present the procedure used to evaluate the reconstruction in the context of simulated data with a partial volume phantom. To keep the experiments simple, we use the simplified model (1) with only one coil (i.e. $C = 1$) and the corresponding sensitivity map to be the identity matrix. An Echo-planar Imaging (EPI) undersampling scheme is used. The Bloch Equations are used for the non-linear mapping, with the random flip angles α and fixed repetition times TR as described in [12]. We compare the BLIP algorithm to our proposed algorithm GAP-MRF in two experiments. In the first experiment, we investigate the effect of the magnetisation sequence length L (this parameter affects directly the acquisition time) by varying from $L = 200$ to $L = 600$ acquisition instances. For the second experiment, we investigate the effect of measurement noise by varying the input SNR defined as:

$$\text{input SNR} = 20 \log \left(\frac{\|\mathbf{Y}\|_F}{\sqrt{ML}\sigma_{\mathbf{Y}}} \right), \quad (36)$$

where $\sigma_{\mathbf{Y}}$ is the standard deviation of the noise. We vary the input SNR from 10dB to 50dB. We choose the undersampling ratio $N/Q = 16$ to simulate the EPI in vivo data in section VI. The BLIP algorithm is stopped when either a fixed maximum iteration number is reached or the following stopping criterion is satisfied:

$$\frac{|E^{(i+1)} - E^{(i)}|}{E^{(i+1)}} < 10^{-4}, \quad (37)$$

where $E^{(i)} = \|h(\mathbf{M}^{(i)}) - \mathbf{Y}\|_F^2$, $(\mathbf{M}^{(i)})_{i \in \mathbb{N}}$ being a sequence generated by the BLIP algorithm. Additionally for Algorithm 1, we choose:

$$\forall p \in \{1, \ldots, P\}, \quad \Sigma_{p,p}^{(i)} < \sigma_p, \quad (38)$$

where $\sigma \in \mathbb{R}^P$ is the tolerance for each parameter.

A. Partial Volume Phantom

We create a simulated phantom based on the work in [26]. Precisely, to introduce the PVE, we use blocks of 2×2 voxels to form a lower resolution phantom containing partial volume voxels. The resulting volume is resized to 256×256 voxels. In Fig. 6(a)-6(e) the five tissues are presented as they appear in the resulting volume. The phantom's voxel distribution is shown in Fig. 6(f). Using this representation we can see the structure of the tissues of interest. A range of parameter values
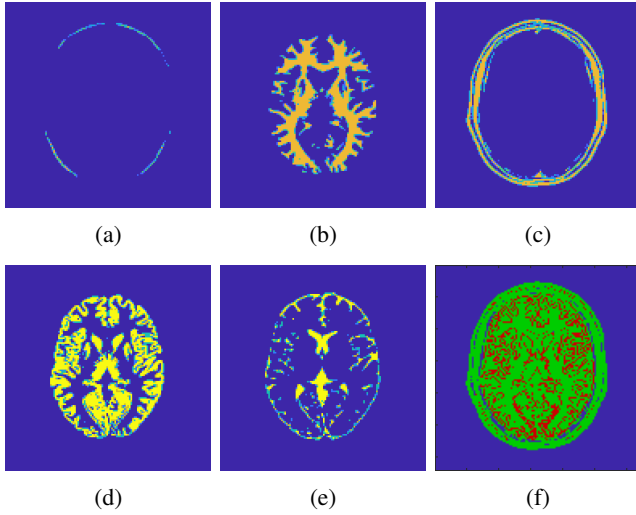
Fig. 6: Proton Density Maps. (a) Adipose ($T_1 = 530$ms and $T_2 = 77$ms). (b) White Matter ($T_1 = 811$ms and $T_2 = 77$ms). (c) Muscle ($T_1 = 1425$ms and $T_2 = 41$ms). (d) Gray Matter ($T_1 = 1545$ms and $T_2 = 83$ms). (e) CSF ($T_1 = 5012$ms and $T_2 = 512$ms). (f) Voxel distribution, the empty voxels are in blue, the pure voxels are in green and the partial volume voxels are in red.

of T_1 and T_2 are used to construct the proton density maps in order to evaluate the reconstructions.

B. Evaluation

In order to evaluate the algorithms, we use the Signal-to-Noise Ratio (SNR) defined as:

$$\text{SNR}(\mathbf{U}_{:,t}, \widetilde{\mathbf{U}}_{:,t}) = 10 \log \left(\frac{\sum_{n=1}^{N} (\mathbf{U}_{n,t})^2}{\sum_{n=1}^{N} \left(\mathbf{U}_{n,t} - \widetilde{\mathbf{U}}_{n,t}\right)^2} \right), \tag{39}$$

where $t \in \{1, \dots, T\}$ is the index of the evaluated tissue, $\mathbf{U}$ is the mixing matrix ground truth and $\widetilde{\mathbf{U}}$ is the estimation. Similarly for the magnetisation sequence SNR, we sum for all values in the matrix. To construct the matrix $\widetilde{\mathbf{U}}$, a tolerance of 15% from the ground truth parameter values is used (i.e. for $T_1 = 530$ms and $T_2 = 77$ms all the dictionary elements that fall for T_1 in the range of $[450.5 - 609.5]$ms and simultaneously for T_2 in the range of $[65.45 - 88.55]$ms are considered). In order to evaluate if the tissues are correctly identified, we define the success rate (SR) index as the proportion of voxels where the number of elements are correctly identified and its corresponding parameters fall within the 15% of the true parameters. Due to noise, there could be small values in $\widetilde{\mathbf{U}}$ that could significantly affect the success rate. In consequence, we choose not to consider values that are smaller than 30, given that the range of the proton densities is from 80 to 400.

C. Experiment 1 - Impact of the magnetisation sequence length

In this subsection, we compare the proposed GAP-MRF algorithm with the BLIP algorithm, for different number of acquisition instances L. The input SNR is set to 50dB. The dictionary for BLIP is defined as in [12] with $D = 16170$.

For GAP-MRF, the initial dictionary is defined by all the combinations of 20 equally spaced values for T_1 and T_2 with $D = 400$. The number of clusters is set to $K = 40$ and the number of random samples per cluster is chosen to be $n_s = 10$. In addition, we choose the parameters in Algorithms 1 and 2 as follows: $\upsilon = 0.15$, $\beta = 0.9$, $\gamma = 0.85$, $\xi = 200$, $\Sigma_{1,1}^{(0)} = 12$, $\Sigma_{2,2}^{(0)} = 3$, $\kappa = 50$, $\sigma_1 = 1$ and $\sigma_2 = 1$. We run the experiment 10 times for each value of L and we show the average of the results with its corresponding standard deviation.
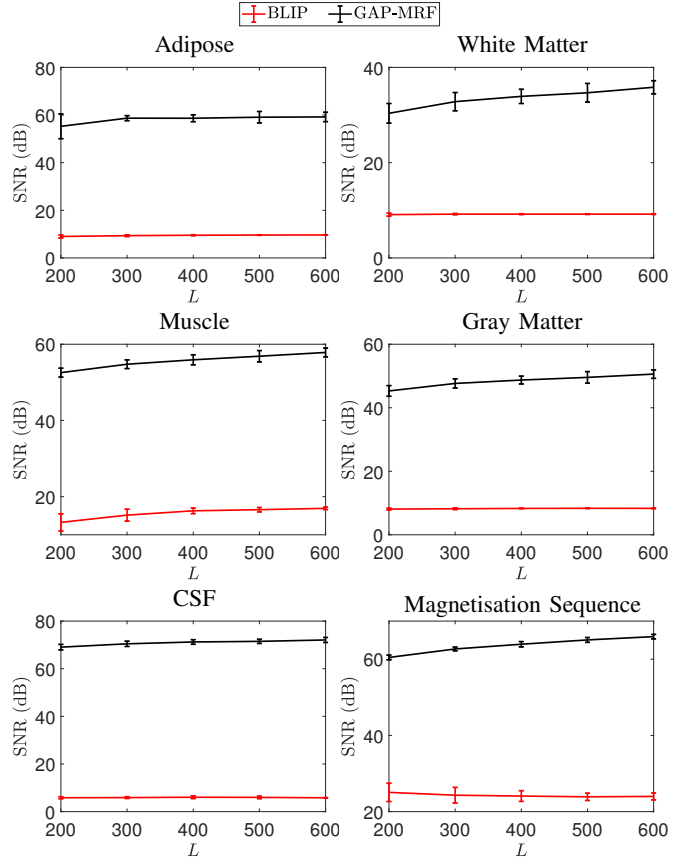

Fig. 7: Experiment 1 - Tissue proton density maps evaluation. The red and black lines correspond to the average results (and the corresponding standard deviation) for the BLIP and GAP-MRF algorithms, respectively.

Fig. 7 shows the evaluation of the proton density maps for each tissue and the magnetisation sequence. Note that GAP-MRF results are taken directly from the matrix $\widetilde{\mathbf{U}}$ without using any post-processing. We can observe that GAP-MRF outperforms BLIP in reconstructing $\mathbf{U}$. One of the advantages of GAP-MRF over BLIP is the dictionary estimation since BLIP is restricted to the input dictionary. As a consequence of the accurate proton density maps, the magnetisation sequence is also accurate. In Fig. 8, the success rate for pure and partial volume voxels can be seen. Since BLIP can only reconstruct one element per voxel, its success rate for partial volume is always equal to 0. In a low noise scenario, GAP-MRF can identify the correct voxel elements even for short sequences. An important remark is that due to partial volume, the dictionary sampling and the number of acquisitions, the BLIP algorithm can mis-reconstruct pure voxels even in a low
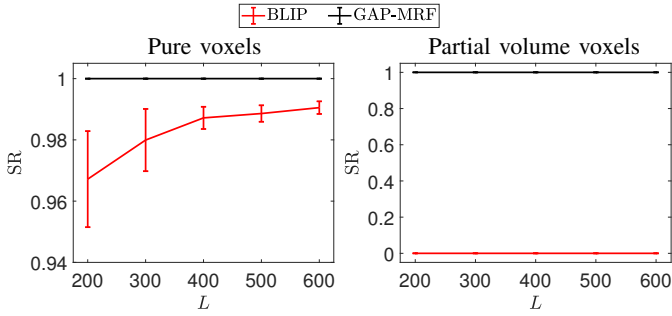
Fig. 8: Experiment 1 - Success rate evaluation. The red and black lines correspond to the average results (and the corresponding standard deviation) for the BLIP and GAP-MRF algorithms, respectively.

noise scenario.

TABLE I: Parameter values of example in Fig. 9 corresponding to Experiment 1 with $L = 400$. The relaxation times are in milliseconds.

	Ground Truth		BLIP		GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	[460-590]	[74-88]	530.2	77.0
White Matter	811	77	[690-930]	[66-88]	811.1	77.0
Muscle	1425	41	[1220-1630]	[36-46]	1425.1	41.0
Gray Matter	1545	83	[1320-1620]	[72-94]	1545.0	83.0
CSF	5012	512	[4400-5000]	500	5011.4	512.0

We show an example of the proton density maps for $L = 400$ in Fig. 9. By visual inspection, we can observe that the GAP-MRF method outperforms the BLIP method for partial reconstructions even for short sequences in a low noise scenario. The proton density SNR and the SR for the example in Fig. 9 are given in Table II. The GAP-MRF has the additional advantage that it simultaneously estimates the manifold regions of interest, resulting in better reconstructions.

TABLE II: Proton density map SNR and SR of the example in Fig. 9 corresponding to Experiment 1 with $L = 400$. The SNR values are in decibels.

	BLIP	GAP-MRF
Adipose SNR	9.09	58.64
White Matter SNR	9.16	34.40
Muscle SNR	14.97	58.14
Gray Matter SNR	8.24	49.61
CSF SNR	5.75	71.31
Pure Voxel SR	0.9804	1
Partial Volume SR	0	1

D. Experiment 2 - Impact of the input SNR

In this experiment, we investigate the behaviour of both the BLIP and the GAP-MRF algorithms while changing the input noise. We fixed the magnetisation sequence length $L = 1000$. The dictionary for BLIP is defined as in [12] with $D = 16170$. For GAP-MRF, the initial dictionary is defined by all the combinations of 20 equally spaced values for T_1 and T_2 with $D = 400$. We choose the number of clusters to be $K = 40$ and the number of random samples per cluster $n_s = 10$. In addition, we choose the parameters of Algorithms 1 and 2 as follows: $\upsilon = 0.15$, $\beta = 0.9$, $\gamma = 0.85$, $\xi = 200$, $\Sigma_{1,1}^{(0)} =$
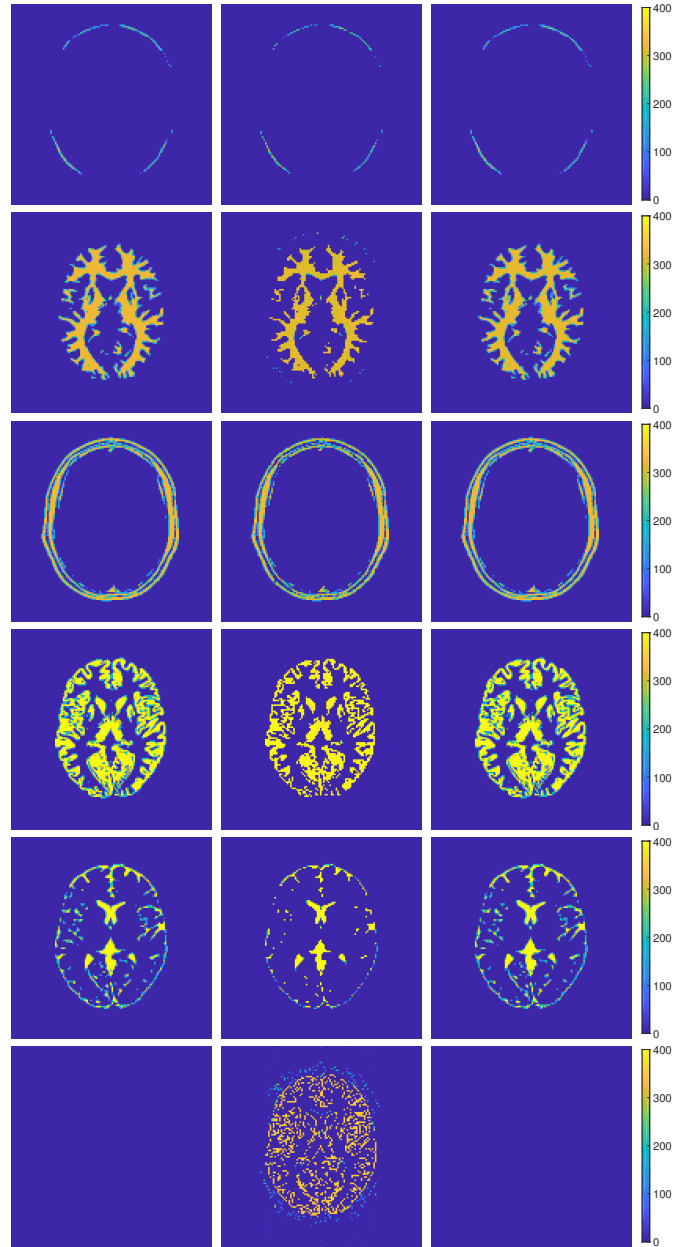

Fig. 9: Example of the proton density maps with $L = 400$ and an input SNR of 50dB corresponding to Experiment 1. The first column images are the Ground Truth, the second column images are the BLIP reconstructions and the third column images are the GAP-MRF reconstructions. From first to fifth row, the tissues Adipose, White Matter, Muscle, Gray Matter and CSF are shown. The last row corresponds to the sum of proton density of all other matched elements that are not in the 15% range of the Ground Truth elements. The corresponding T_1 and T_2 values can be seen in Table I and the SNR and SR values can be seen in Table II.

12, $\Sigma_{2,2}^{(0)} = 3$, $\kappa = 50$, $\sigma_1 = 1$ and $\sigma_2 = 1$. We run the experiment 10 times for each choice of input SNR and we show the average of the results with its corresponding standard deviation.

Fig. 10 shows the results for the proton density maps for each tissue. GAP-MRF significantly outperform BLIP when the input SNR is greater than 30dB. It should be noted that GAP-MRF estimates correctly the number of true atoms when
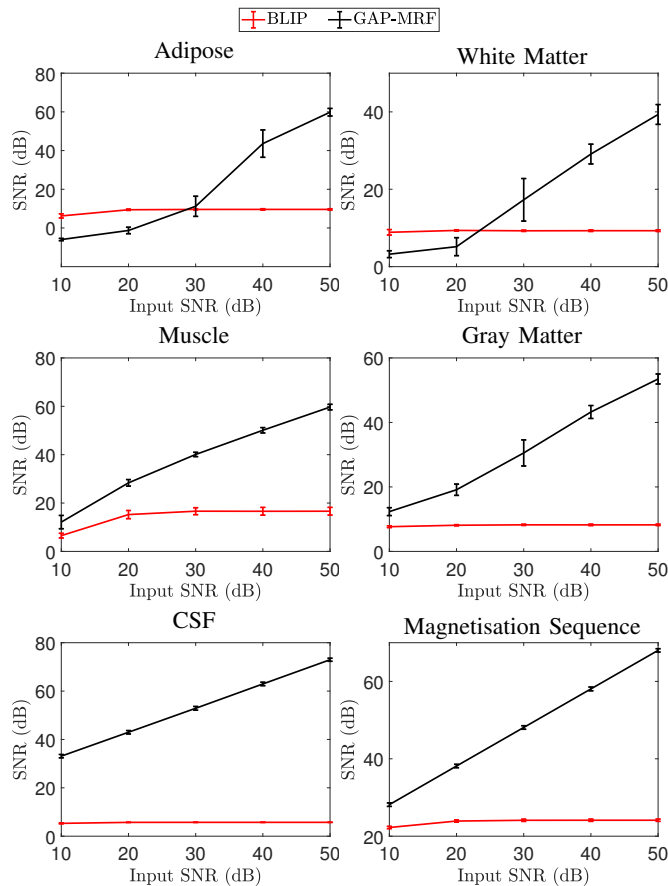
Fig. 10: Experiment 2 - Tissue proton density maps evaluation. The red and black lines correspond to the average results (and the corresponding standard deviation) for the BLIP and GAP-MRF algorithms, respectively.

the input SNR is 30dB or greater. The reconstruction of Adipose is affected more by the noise since there are significantly less pure voxels of this tissue. GAP-MRF magnetisation sequence reconstruction is significantly more accurate than BLIP reconstruction, because BLIP does not consider the partial volume effect and also because of the dictionary inaccuracy. GAP-MRF magnetisation sequence SNR has a linear behaviour with respect to the input SNR. In Fig. 11, the success rate with respect to the input SNR can be seen. We can observe that the success rate is significantly affected by the input SNR.

TABLE III: Parameter values of example in Fig. 12 corresponding to Experiment 2 with an input SNR of 30dB. The relaxation times are in milliseconds.

	Ground Truth		BLIP		GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	[460-590]	[74-84]	531.6	77.0
White Matter	811	77	[690-930]	[66-80]	810.9	77.1
Muscle	1425	41	[1220-1630]	[36-46]	1423.6	41.1
Gray Matter	1545	83	[1320-1610]	[74-86]	1543.4	83.0
CSF	5012	512	[4400-5000]	500	5011.8	512.0

We show an example of the proton density maps for each tissue in Fig. 12 when the input SNR is 30dB. By visual
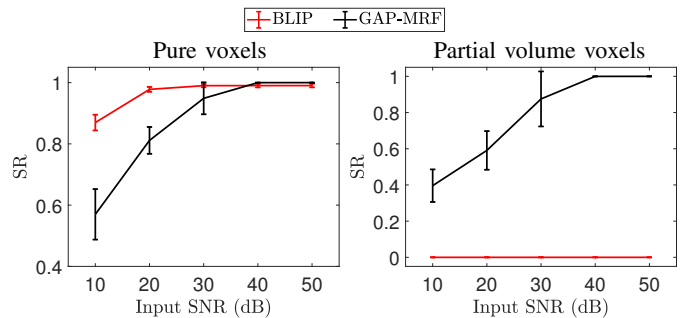

Fig. 11: Experiment 2 - Success rate evaluation. The red and black lines correspond to the average results (and the corresponding standard deviation) for the BLIP and GAP-MRF algorithms, respectively.

TABLE IV: Proton density map SNR and SR of the example in Fig. 12 corresponding to Experiment 1 with an input SNR of 30dB. The SNR values are in decibels.

	BLIP	GAP-MRF
Adipose SNR	9.70	12.15
White Matter SNR	9.14	19.62
Muscle SNR	17.15	39.61
Gray Matter SNR	8.30	32.10
CSF SNR	5.71	52.55
Pure Voxel SR	0.9984	0.9766
Partial Volume SR	0	0.9459

inspection, we can observe that the GAP-MRF method outperforms the BLIP method for partial volume reconstructions for moderate noise scenarios. The proton density SNR and the SR for the example in Fig. 12 are given in Table IV. The GAP-MRF correctly estimates the manifold regions of interest. BLIP due the PVE and the input noise, has a residual map formed by all the elements that are not sufficiently close to the true elements. In the GAP-MRF reconstructions, the White Matter and Adipose tissue are slightly mixed due to the noise since their parameters are close to each other. By choosing a better Γ we can make the atoms of the dictionary more distant in the ℓ_2 sense, this would provide noise robustness to the reconstructions.

VI. RESULTS ON REAL DATA

In this section, we show the reconstructions obtained considering two in vivo datasets. The first dataset was acquired using EPI sampling scheme [27] and the second dataset was acquired using spiral sampling scheme. The parameters are tuned using the following heuristic. During the initialisation, we choose $K = 20$, and increase this value by 10 until the residual stabilises, the idea is to have enough centers such that the partial volume voxels does not shift the values of true parameters. The value of ξ is chosen to only consider the voxels containing the tissues. We fixed $n_s = 10$ and $\beta = 0.9$ to avoid long reconstruction time. We initialise the algorithm as described in the Section IV-C with the obtained K. We solve the problem increasing parameter υ from 0.15 (with steps of 0.05) until we can see the structures in the parameter maps setting $\kappa = 20$ and $\gamma = 0.85$. Finally, we increase κ until we remove all the noisy proton density maps. We normalise the proton density maps to reflect the proportion of the tissue by computing, for each voxel $n \in \{1, \ldots, N\}$, $\widetilde{\mathbf{U}}_{n,:}/\|\widetilde{\mathbf{U}}_{n,:}\|_2$.
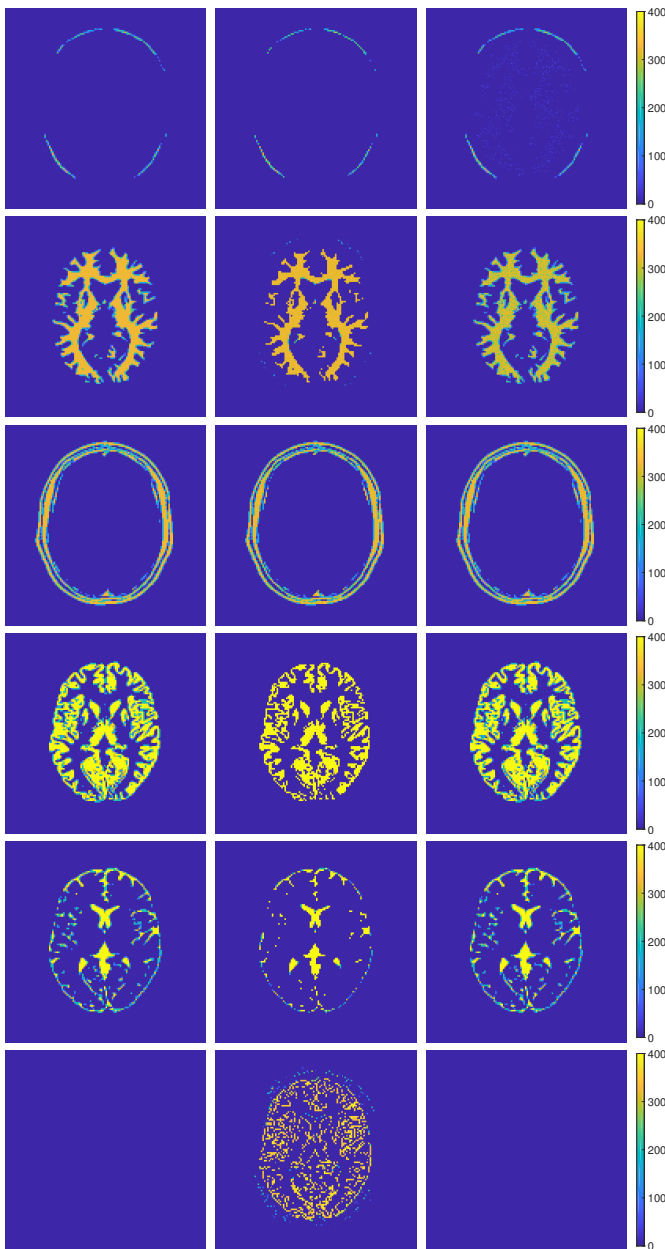
Fig. 12: Example of the proton density maps with $L = 1000$ and an input SNR of 30dB corresponding to Experiment 2. The first column images are the Ground Truth, the second column images are the BLIP results and the third column images are GAP-MRF results. From first to fifth row, the tissues Adipose, White Matter, Muscle, Gray Matter and CSF are shown. The last row corresponds to the sum of proton density of all other matched elements that are not in the 15% range of the Ground Truth elements. The corresponding T_1 and T_2 values can be seen in Table III. The corresponding T_1 and T_2 values can be seen in Table III and the SNR and SR values can be seen in Table IV.

A. EPI sampling dataset

The scanning for this dataset has been performed on a 3T GE MR750w scanner with a 12 channel receive only head RF coil (GE Medical Systems, Milwaukee, WI). The study was approved by the local ethics committee. The used acquisition scheme was 16-shot EPI-MRF on a healthy volunteer using a variable flip angle α ramp (see Fig. 13), ranging from $1°$ to $70°$. The excitation sequence length is $L = 500$. The repetition

time TR was set to 16ms. The acquisition bandwidth (BW) = 5kHz and the Field of View (FOV) = $22.5 \times 22.5\text{cm}^2$. The spatial resolution is 128×128 voxels, with a 5mm slice thickness. The undersampling ratio is $N/Q = 16$. The EPG model is used for the reconstructions with an Inversion Time (TI) of 18ms and an Echo Time (TE) of 3.5ms. The acquisition time for the slice was 9s. We choose the GAP-MRF parameters to be $\kappa = 35$, $\xi = 10^8$, $K = 40$, $\upsilon = 0.4$, $n_s = 10$, $\Sigma_{1,1}^{(0)} = 1$, $\Sigma_{2,2}^{(0)} = 1$, $\gamma = 0.85$, $\sigma_1 = 0.01$ and $\sigma_2 = 0.01$. In this experiment we have chosen smaller values of Σ than the synthetic data, because the noise was shifting the parameters when the noisy elements were removed. In
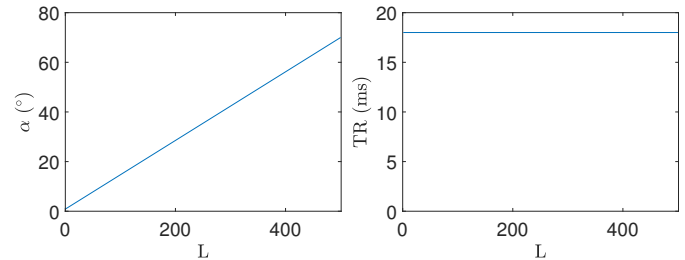

Fig. 13: The image of the left represents the flip angles of the acquisition and the image of the right represents the repetition times of the dataset acquisition.
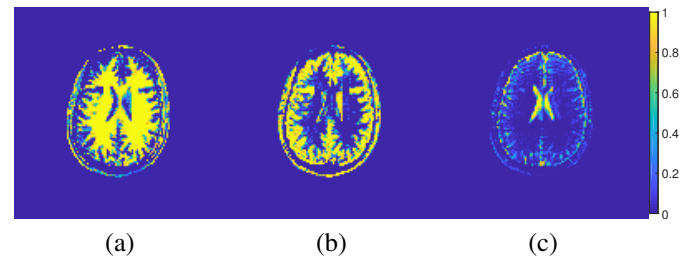

Fig. 14: Normalised proton density maps of the tissues corresponding to the EPI sampling dataset. The values are in arbitrary units (a.u.). The corresponding T_1 and T_2 can be seen in Table V. The figures correspond to (a) White Matter. (b) Gray Matter. (c) CSF.

TABLE V: Comparison between the parameters obtain with GAP-MRF and the reported values in [28] (MRF FISP sequence) for the EPI sampling dataset. The relaxation times are in milliseconds.

	Values reported in [28]		GAP-MRF	
	T_1	T_2	T_1	T_2
White Matter	781 ± 61	65 ± 6	762.6	67.2
Gray Matter	1193 ± 65	109 ± 11	1116.6	107.1
CSF			2391.1	856.2

Fig. 14, we can observe the resulting proton density maps provided by the GAP-MRF algorithm and the Table V shows a comparison between the parameters reported in [28] for MRF FISP sequences and the parameters obtained by GAP-MRF. CSF values are not reported for the MRF FISP sequence. The White Matter parameters are similar to the ones reported in [28] and the Gray Matter T_1 is slightly lower than the reported one.

B. Spiral sampling dataset

The scanning for this dataset was performed on a GE HDx MRI system with an 8 channel receive only head RF coil (GE Medical Systems, Milwaukee, WI). The acquisition scheme uses a variable density spiral with 89 interleaves using FISP based α and TR as in [29] (see Fig. 15). The excitation sequence length is $L = 1000$. The repetition time TR was set to 16ms. In this experiment, we have FOV = $22.5 \times 22.5\text{cm}^2$ and the spatial resolution is 256×256 voxels, with a 5mm slice thickness. The undersampling ratio is $N/Q = 89.53$. The EPG model is used for the reconstructions with a TI of 18ms and a TE of 2ms. The GAP-MRF parameters are $\kappa = 60$, $\xi = 1.1 \times 10^7$, $K = 30$, $\upsilon = 0.4$, $n_s = 10$, $\Sigma_{1,1}^{(0)} = 20$, $\Sigma_{2,2}^{(0)} = 5$, $\gamma = 0.85$, $\sigma_1 = 0.01$ and $\sigma_2 = 0.01$.
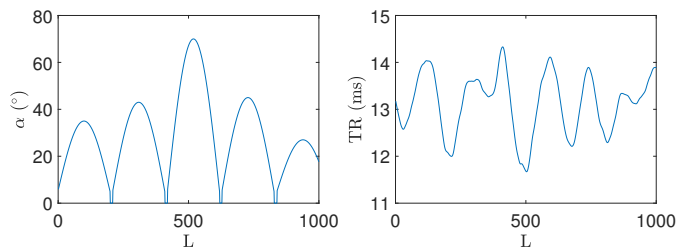
Fig. 15: The image of the left represents the flip angles of the acquisition and the image of the right represents the repetition times of the dataset acquisition.

TABLE VI: Comparison between the parameters obtain with GAP-MRF and the reported values in [28] (MRF FISP sequence) for the spiral sampling dataset. The relaxation times are in milliseconds.

	Values reported in [28]		GAP-MRF	
	T_1	T_2	T_1	T_2
White Matter	781±61	65±6	742.0	41.6
Gray Matter	1193±65	109±11	1246.3	85.3
CSF			1841.5	1480.9
Muscle	1100±59	44±9	2326.5	35.8
Fat	253±42	68±4	503.6	41.5

In Fig. 16, we can observe the resulting proton density maps provided by the GAP-MRF algorithm and the Table VI shows a comparison between the parameters reported in [28] and the parameters obtained by GAP-MRF. The White and Gray Matter parameters obtained by GAP-MRF differ slightly to the ones reported in [28]. But the other tissues are far from the expected tissues in the brain. This could be due to the small number of pure voxels that are not sufficient to accurately estimate the parameters. We believe that choosing better acquisition parameters Γ to make the elements of the dictionary more distant in the ℓ_2 sense can significantly improve the accuracy of the parameters. Also, inaccuracies in the model such as calibration or motion in the acquisition can produce artifacts in the reconstruction.

VII. CONCLUSIONS AND FUTURE WORK

We have presented an extension of the model in [12] to partial volume reconstructions in the context of MRF. The simulations presented in Section V show that the proposed technique GAP-MRF can achieve accurate reconstructions with very short pulse sequences in the low input noise scenario. It also performs well when the input SNR is greater than 30dB. Our algorithm provides a way to explore the manifold of magnetic resonance fingerprints without densely sampling Θ. In comparison to BLIP algorithm, the memory requirements are significantly reduced since only few dictionary elements are stored.

We also present in Section VI the results obtained with real data. Some parameters differ slightly to the reported in the literature, but the structure seen in the proton densities maps suggests that this approach can provide additional information that can be useful for diagnosis. The next step is to evaluate the partial volume reconstructions with a real partial volume phantom in the scanner and a full brain reconstruction to provide enough pure voxels to accurately estimate the true parameters.

For future work, incorporating spatial regularisation in the objective function should improve the robustness of the method. A joint calibration and imaging problem should also be considered in order to provide both phase estimation and compensation.

REFERENCES

[1] H. Jara, *Theory of Quantitative Magnetic Resonance Imaging*. World Scientific, 2013.

[2] D. Ma, V. Gulani, N. Seibelich, K. Liu, J. L. Duerk, and M. A. Griswold, "Magnetic resonance fingerprinting," *Nature*, vol. 465, pp. 187–192, 2013.

[3] J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review," *World J Radiol*, vol. 6, no. 11, pp. 855–864, Nov. 2014, 25431640[pmid].

[4] D. McGivney, A. Deshmane, Y. Jiang, D. Ma, C. Badve, A. Sloan, V. Gulani, and M. Griswold, "Bayesian estimation of multicomponent relaxation parameters in magnetic resonance fingerprinting," *Magnetic Resonance in Medicine*, vol. 80, no. 1, pp. 159–170, 2018.

[5] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.

[6] P. Tseng, "A modified forward-backward splitting method for maximal monotone mappings," *SIAM J. Control Optim.*, vol. 38, no. 2, pp. 431–446, 2000.
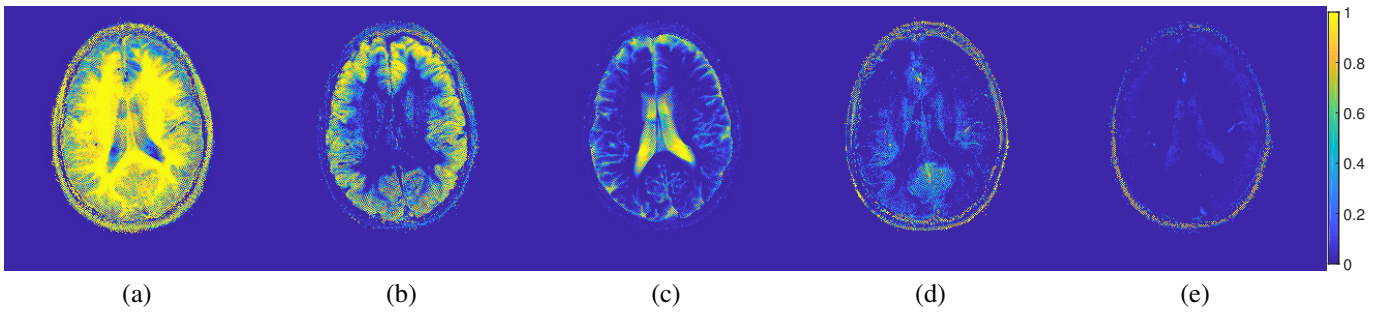
Fig. 16: Normalised proton density maps of the tissues corresponding to the spiral sampling dataset. The values are in arbitrary units (a.u.). The corresponding T_1 and T_2 can be seen in Table VI. The figures correspond to (a) White Matter. (b) Gray Matter. (c) CSF. The figures (d) and (e) are not estimated accurately enough to distinguish tissues.

[7] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[8] P. Tseng, "Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization," *Journal of Optimization Theory and Applications*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[9] J. Bolte, S. Sabach, and M. Teboulle, "Proximal Alternating Linearized Minimization for Nonconvex and Nonsmooth Problems," *Math. Program.*, vol. 146, no. 1-2, pp. 459–494, Aug. 2014.

[10] P. Frankel, G. Garrigos, and J. Peypouquet, "Splitting Methods with Variable Metric for Kurdyka–Łojasiewicz Functions and General Convergence Rates," *Journal of Optimization Theory and Applications*, vol. 165, no. 3, pp. 874–900, Jun. 2015.

[11] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forwardbackward algorithm," *Journal of Global Optimization*, pp. 1–29, Feb. 2016.

[12] M. Davies, G. Puy, P. Vandergheynst, and Y. Wiaux, "A Compressed Sensing Framework for Magnetic Resonance Fingerprinting," *SIAM J. Imaging Sci.*, vol. 7, no. 4, p. 2623, Nov. 2014.

[13] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[14] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[15] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[16] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 43, pp. 129–159, 2001.

[18] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59 – 73, 2011.

[19] T. Blumensath, "Sampling and Reconstructing Signals From a Union of Linear Subspaces," *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4660–4671, Jul. 2011.

[20] T. Blumensath and M. E. Davies, "Normalized Iterative Hard Thresholding: Guaranteed Stability and Performance," *IEEE J. Selected Topics Signal Process.*, vol. 4, no. 2, pp. 298–309, Apr. 2010.

[21] G. H.-G. Chen and R. T. Rockafellar, "Convergence Rates in Forward–Backward Splitting," *SIAM J. Optim.*, vol. 7, no. 2, pp. 421–444, 1997.

[22] P. L. Combettes and V. R. Wajs, "Signal Recovery by Proximal Forward-Backward Splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[23] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, ser. Classics in Applied Mathematics. Philadelphia, PA: SIAM, 1995, vol. 15.

[24] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Int.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[25] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[26] 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 Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, Jun. 1998.

[27] A. Benjamin, P. Gómez, M. Golbabaee, Z. Mahbub, T. Sprenger, M. Menzel, M. Davies, and I. Marshall, "Balanced multi-shot EPI for accelerated Cartesian MR Fingerprinting: An alternative to spiral MR Fingerprinting," Feb. 2018.

[28] J. Z. Bojorquez, S. Bricq, C. Acquitter, F. Brunotte, P. M. Walker, and A. Lalande, "What are normal relaxation times of tissues at 3T?" *Magnetic Resonance Imaging*, vol. 35, pp. 69 – 80, 2017.

[29] Y. Jiang, D. Ma, N. Seiberlich, V. Gulani, and M. A. Griswold, "MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout," *Magnetic Resonance in Medicine*, vol. 74, no. 6, pp. 1621–1631.