



Heriot-Watt University
Research Gateway

Greedy Approximate Projection for Magnetic Resonance Fingerprinting with Partial Volumes

Citation for published version:

Duarte, R, Repetti, A, Gómez, PA, Davies, M & Wiaux, Y 2020, 'Greedy Approximate Projection for Magnetic Resonance Fingerprinting with Partial Volumes', *Inverse Problems*, vol. 36, no. 3, 035015. <https://doi.org/10.1088/1361-6420/ab356d>

Digital Object Identifier (DOI):

[10.1088/1361-6420/ab356d](https://doi.org/10.1088/1361-6420/ab356d)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

Published In:

Inverse Problems

Publisher Rights Statement:

This is an author-created, un-copyedited version of an article accepted for publication/published in Inverse Problems. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at <https://doi.org/10.1088/1361-6420/ab356d>

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Greedy Approximate Projection for Magnetic Resonance Fingerprinting with Partial Volumes

Roberto Duarte¹, Audrey Repetti^{1,2}, Pedro A. Gómez³,
Mike Davies⁴ and Yves Wiaux¹

¹Institute of Sensors Signals and Systems (ISSS), Heriot-Watt University, Edinburgh, EH14 4AS, UK

²Department of Actuarial Mathematics & Statistics (AMS), Heriot-Watt University, Edinburgh, EH14 4AS, UK

³Munich School of Bioengineering, Technische Universität München, Munich, Germany

⁴Institute for Digital Communications (IDCom), The University of Edinburgh, EH9 3JL, UK

E-mail: rd8@hw.ac.uk, a.repetti@hw.ac.uk, pedro.gomez@tum.de,
Mike.Davies@ed.ac.uk, y.wiaux@hw.ac.uk

April 2017

Abstract. In quantitative Magnetic Resonance Imaging, 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 memory efficient, greedy approximate projected gradient descent algorithm, dubbed GAP-MRF. Our method adaptively finds the regions of interest on the manifold of fingerprints defined by the MRF sequence. We generalise our method to compensate for phase errors appearing in the model, using an alternating minimisation approach. We show, through simulations on synthetic data with PVE, that our algorithm outperforms state-of-the-art methods in reconstruction quality. Our approach is validated on the EUROSPIN phantom and on *in vivo* datasets.

1. 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 it 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

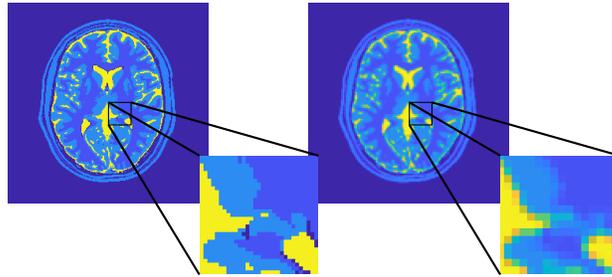


Figure 1. Partial volume effect in a T_1 parameter map. Left: true T_1 parameter map. Right: low resolution reconstruction.

under investigation, such as T_1 and T_2 relaxation times [1]. Unfortunately, due to prohibitively long acquisition times, qMRI is not the standard for diagnosis. To overcome this difficulty, Magnetic Resonance Fingerprinting (MRF) was introduced to accelerate qMRI acquisitions [2], inspired by Compressive Sensing (CS) theory [3]. MRF uses a combination of random excitation pulse sequences and k -space (i.e. Fourier space) undersampling to simultaneously acquire all relevant quantitative information. These random excitation sequences are used to produce unique temporal patterns called *fingerprints*, which are compared to the ones predicted by the model to extract the parameters of interest, similar to dictionary based methods such as [4]. More recently, a full CS strategy was formulated in [5] for MRF. In this work, the authors developed an iterative projection algorithm (also known as projected gradient descent, or forward-backward algorithm [6, 7]), dubbed BLoch response recovery via Iterative Projection (BLIP), reconstructing MRF signal with less acquisitions than the traditional MRF method [2]. Note that increasing the number of acquisitions can significantly increase the acquisition time and potentially induce additional modelling inaccuracies (e.g., resulting from motion or timing).

In general, qMRI techniques, particularly MRF-based methods [8, 9, 10, 11, 12, 5, 13], assume that a voxel contains at most one type of tissue, e.g. white matter (WM), grey matter (GM), etc. This assumption is not suitable in practice. Consequently, voxels containing multiple tissue types may be assigned with incorrect parameters. This problem is known as the Partial Volume Effect (PVE) and appears in all medical imaging modalities with limited spatial resolution [14]. An example of PVE is given in Fig. 1. The left image shows a spatial distribution of T_1 in a simulated brain. The right image shows a reconstruction using voxels four times bigger and assuming a single tissue per voxel. All low resolution voxels at the edge between tissues contain partial volumes, which implies a wrong estimate (single wrong value of T_1 rather than multiple values).

The PVE has been analysed in the supplementary material of [2]. In this work, using a least-squares method, the signal is decomposed as a weighted sum of at most three distinct signals, each representing a different tissue. Although this method was shown to be robust to noise for long sequences, since it necessitates both information about the spatial distribution of the PV voxels and the true components of the original signal

(which are unknown in practice), it is not adapted to handle *in vivo* data. An extension of this approach has been proposed in [15], where the tissue parameters are learnt using a clustering approach on the parameter maps, obtained by the match filter. Then the data is matched with a PV dictionary varying the tissue proportion. This method has shown good results when considering long sequences, but the reconstruction quality is limited by the precision of the dictionary elements obtained by the MRF solution and the precision of the tissue proportions used to generate the PV dictionary. Moreover, the size of the PV dictionary increases exponentially with the maximum number of tissues allowed in a voxel and, the parameters for the PV dictionary are manually selected from the clustering results. In addition, all the reconstructions are performed with high aliased images requiring more acquisitions for accurate results. Additionally, for short sequences, the noise in the measurements and the sampling of the manifold of fingerprints describing the signal can significantly affect the estimations. More recently, a Bayesian method was proposed in [16], to tackle the PVE in MRF (we will refer to this method as Bayesian-MRF). The authors show that their approach estimates the parameters of the PV voxels. However, due to the high aliasing effect encountered with undersampled noisy data, this estimation comes at the cost of an increased acquisition time with respect to traditional MRF based reconstructions (*i.e.* three times longer sequences than traditional MRF). 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 both reconstruction time and memory requirement). While this method was formulated as a convex optimisation problem in the Bayesian framework, the algorithm in [16] removes dictionary entries at each iteration, and consequently does not ensure that the algorithm converges to the maximum *a posteriori* estimate. Finally, inspired by the re-weighted ℓ_1 -norm regularisation for sparse recovery, a novel algorithm was proposed in [17]. This method is based on the alternative-direction method of multipliers (ADMM). At each iteration, the tissue proportions are updated through the re-weighting ℓ_1 -norm in a voxel-wise fashion. As shown in [17], this algorithm is able to estimate partial volumes in almost noiseless scenarios. Even with the re-weighting, the support estimation is extremely ill-posed, in consequence, the estimations are very sensitive to noise. In addition, the full dictionary is used in the re-weighting process introducing a high computational cost.

In this paper, we propose to tackle the PVE in MRF by reformulating the problem as a non-convex constrained least-squares minimisation problem. In our approach, we assume that the number of independent tissues in the imaged volume is upper bounded, and that there exists at least a region of the total volume with only pure voxels for each tissue. To solve the resulting non-convex constrained minimisation problem, we develop a greedy approximate projected gradient descent method, dubbed GAP-MRF. It can be seen as a generalisation of BLIP method for PVE. It consists in a projected gradient descent algorithm, where the projection is computed inexactly, through a memory efficient greedy approach. The proposed method is also generalised to compensate for phase errors in the model, due to timing or coil sensitivity errors,

using an alternating minimisation approach [18, 19, 20, 21, 22]. Through simulations on a simulated PV phantom, we show that our approach outperforms state-of-the-art methods. Our method is afterward validated on the EUROSPIN phantom and on *in vivo* MRF datasets.

The remainder of the paper is organised as follows: In Section 2 we introduce the notation used throughout the paper, and we define the MRF inverse problem introducing the proposed PV model. In Section 3 we give the proposed algorithm to solve the PV problem. Finally, in Sections 5 and 6, we investigate the behaviour of the proposed method on simulated data and show the results on the *in vivo* datasets, respectively. We conclude in Section 7.

2. Notation and MRF problem description

2.1. Notation

In this section, we introduce the notation we will use in the remainder of the paper. We refer the reader to [23, 24] for additional details about optimisation. To have a compact notation when selecting a specific row $n \in \{1, \dots, N\}$ of a matrix $\mathbf{M} \in \mathbb{C}^{N \times L}$, we use the notation $\mathbf{M}_{n,:} = (\mathbf{M}_{n,l})_{1 \leq l \leq L}$. Similarly, to select a specific column $l \in \{1, \dots, L\}$ of this matrix, we use $\mathbf{M}_{:,l} = (\mathbf{M}_{n,l})_{1 \leq n \leq N}$. More generally, this notation is also used to select subparts of tensors. The operator $\text{real}(\cdot)$ gives the real part of its complex argument, the operator $\text{Diag}(\cdot)$ builds a diagonal matrix whose diagonal elements are given by its argument, and $(\cdot)^\dagger$ gives its adjoint. The adjoint of a linear operator $g: \mathbb{C}^L \rightarrow \mathbb{C}^N$ is denoted by g^\dagger . The cardinality of a countable set \mathcal{T} is given by $\text{card}(\mathcal{T})$. The ℓ_p norm ($p \in [0, +\infty]$) is denoted by $\|\cdot\|_p$. The ℓ_0 pseudo-norm [3], counting the non-zero entries of its argument, is defined as $(\forall \mathbf{x} \in \mathbb{R}^N) \|\mathbf{x}\|_0 = \sum_{n=1}^N (\mathbf{x}_n)^0$, with the convention $0^0 = 0$. By abuse of notation, the ℓ_p norms and the ℓ_0 pseudo-norm will be used for tensors by reshaping them into vectors. Finally, the projection of a vector $\bar{\mathbf{x}} \in \mathbb{C}^N$ onto a non-empty closed subset \mathcal{S} of \mathbb{C}^N is given by $\mathcal{P}_{\mathcal{S}}(\bar{\mathbf{x}}) = \text{argmin}_{\mathbf{x} \in \mathcal{S}} \frac{1}{2} \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2$ [23]. The same notation is used for projections of tensors.

2.2. 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

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

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, $\mathbf{S} \in \mathbb{C}^{N \times N \times C}$ 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 $\mathbf{Y} = h(\mathbf{M}) + \boldsymbol{\eta}$.

For each voxel $n \in \{1, \dots, N\}$, the magnetisation response $\mathbf{M}_{n,:}$ is modelled through the smooth non-linear mapping $B : \mathcal{M} \rightarrow \mathbb{C}^{1 \times L}$ (commonly Bloch equations or Extended Phase Graphs (EPG) model [25]) scaled by the unknown proton density $\rho_n \in \mathbb{R}_+$, $\mathbf{M}_{n,:} = \rho_n B(\hat{\boldsymbol{\theta}}_{n,:}, \boldsymbol{\Gamma})$, where $\boldsymbol{\Gamma} \in \mathbb{R}^{A \times 1}$ represents the concatenation of A known acquisition parameters (*e.g.*, flip angles α , repetition times TR) chosen such that $\mathbf{M}_{n,:}$ is only sensitive to the P parameters $\hat{\boldsymbol{\theta}}_{n,:} \in \mathcal{M}$ under investigation, where $\mathcal{M} \subset \mathbb{R}^{1 \times P}$ denotes the subset of feasible parameters. In the remainder, we fix $P = 2$ and choose \mathcal{M} corresponding to T_1 and T_2 relaxation times.

2.3. Proposed partial volume model

The model described in the previous section considers that each voxel contains at most one element. PV voxels are introduced due to the spatial discretisation in the acquisition process. The magnetisation sequence can be described as $\mathbf{M} = \mathbf{X}\boldsymbol{\Phi}$, where $\mathbf{X} \in \mathbb{R}_+^{N \times D}$ is a sparse mixing matrix (each line of \mathbf{X} represents the proton densities associated with a specific voxel, and would contain more than a nonzero value only for voxels with partial volumes), and $\boldsymbol{\Phi} \in \mathbb{C}^{D \times L}$ is the over-complete dictionary of fingerprints, introduced in [2], as a discrete sampling of the low dimensional manifold B . $\boldsymbol{\Phi}$ is constructed from D samples of \mathcal{M} , stored in a matrix $\boldsymbol{\theta} \in \mathbb{R}^{D \times P}$. Due to the smoothness of B , $\boldsymbol{\Phi}$ is highly coherent. Consequently, the estimation of \mathbf{X} from highly undersampled noisy data is expected to fail without additional priors. Leveraging CS theory [3, 26, 27, 28], the sparsest matrix \mathbf{X} , fitting the measurement model, can be found by solving:

$$\underset{\mathbf{X} \in \mathbb{R}_+^{N \times D}}{\text{minimise}} \|\mathbf{X}\|_0 \text{ subject to } \|\mathbf{Y} - h(\mathbf{X}\boldsymbol{\Phi})\|_2 \leq \epsilon, \quad (2)$$

where $\epsilon > 0$ is a bound chosen according to the acquisition noise level. Since this function is non-convex and non-differentiable, problem (2) is difficult to solve in practice, in particular in the context of high dimensional problems (usually, $D \sim 10^6$ and $L \sim 10^3$). Greedy methods that deal directly with the ℓ_0 pseudo-norm, such as [29, 30], rely on a fixed sparsity level. Besides the computational burden of dealing with the huge dictionary $\boldsymbol{\Phi}$, these methods are not suited for partial volume estimations. The main reason is that the problem is intrinsically ill-posed, resulting in noise sensitive estimations without additional constraints. Note that the inclusion of any other constraint to these methods is not straightforward. The non-convexity of the ℓ_0 pseudo-norm is often relaxed by the use of the ℓ_1 -norm [31]. Nevertheless, $\boldsymbol{\Phi}$ being highly coherent, this convex relaxation cannot be used to correctly estimate the coefficients of \mathbf{X} [32, 17]. To overcome these difficulties, similarly to the BLIP approach, we propose to

$$\underset{\mathbf{M} \in \mathcal{B}_{S^+}(\boldsymbol{\Phi})}{\text{minimise}} \frac{1}{2} \|\mathbf{Y} - h(\mathbf{M})\|_2^2 \quad (3)$$

where

$$\mathcal{B}_{\mathcal{S}_+}(\Phi) = \{ \mathbf{M} \in \mathbb{C}^{N \times L} \mid \mathbf{M} = \mathbf{X}\Phi \text{ with } \mathbf{X} \in \mathcal{S}_+ \}, \quad (4)$$

$$\mathcal{S}_+ = \bigcap_{s=1}^4 \mathcal{S}_s, \quad (5)$$

and, for every $s \in \{1, \dots, 4\}$, \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 below.

2.3.1. 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:

$$\mathcal{S}_1 = \mathbb{R}_+^{N \times D}. \quad (6)$$

2.3.2. 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. 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 $\mathcal{D}_{\mathbf{X}} = \{d \in \{1, \dots, D\} \mid (\exists n \in \mathcal{G}_{\mathbf{X}}) \mathbf{X}_{n,d} \neq 0\}$, where $\mathcal{G}_{\mathbf{X}} = \{n \in \{1, \dots, N\} \mid \|\mathbf{X}_{n,:}\|_1 > \xi\}$. 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 upper bounding the cardinality of this set by K :

$$\mathcal{S}_2 = \{ \mathbf{X} \in \mathbb{R}^{N \times D} \mid \text{Card}(\mathcal{D}_{\mathbf{X}}) \leq K \}. \quad (7)$$

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

$$\begin{aligned} \mathcal{N}_v(d) = & \{d' \in \{1, \dots, D\} \setminus \{d\} \mid \\ & (\forall p \in \{1, \dots, P\}) \mid \boldsymbol{\theta}_{d',p} - \boldsymbol{\theta}_{d,p} \mid < v \boldsymbol{\theta}_{d,p} \}, \end{aligned} \quad (8)$$

where $v > 0$. We 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. Precisely, we constrict all the neighbour columns of each element in $\mathcal{D}_{\mathbf{X}}$ to be the null element $\mathbf{0}$ of \mathbb{R}^N :

$$\mathcal{S}_3 = \{ \mathbf{X} \in \mathbb{R}^{N \times D} \mid (\forall d' \in \bigcup_{d \in \mathcal{D}_{\mathbf{X}}} \mathcal{N}_v(d)) \mathbf{X}_{:,d'} = \mathbf{0} \} \quad (9)$$

2.3.4. Constraint on the pure voxels Due to the additive noise in model (1), 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 $\kappa > 0$

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  $i = 0, 1, \dots$  do
4:    $\mu = 2N/Q$ ,  $\nu = 0$ 
5:   while  $\mu > \nu$  do
6:      $\mu = \mu/2$ 
7:     Gradient Step:
8:      $\overline{\mathbf{M}}^{(i)} = \mathbf{M}^{(i)} - \mu h^\dagger (h(\mathbf{M}^{(i)}) - \mathbf{Y})$ 
9:     Projection Step:
10:     $\mathbf{M}^{(i+1)} \approx \mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}(\overline{\mathbf{M}}^{(i)})$ 
11:    Backtracking step
12:     $\nu = \zeta \frac{\|\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\|_2^2}{\|h(\mathbf{M}^{(i+1)}) - h(\mathbf{M}^{(i)})\|_2^2}$ 
13:   end while
14: end for

```

rows (i.e. voxels) of \mathbf{X} contain only one non-zero value for each active column of \mathbf{X} . These rows identify the *pure voxels*. This constraint can be formulated as follows:

$$\mathcal{S}_4 = \{\mathbf{X} \in \mathbb{R}^{N \times D} \mid (\forall d \in \mathcal{D}_{\mathbf{X}}) \|\mathbf{X}_{n,d}\|_0 \geq \kappa\} \quad (10)$$

where $\mathcal{V}_{\mathbf{X}} = \{n \in \{1, \dots, N\} \mid \|\mathbf{X}_{n,:}\|_0 = 1\}$.

3. Greedy Approximate Projection for MRF

3.1. Proposed iterative projected gradient descent algorithm

To solve problem (3), we use an iterative projected gradient descent method [33]. At each iteration $i \in \mathbb{N}$, this method updates $\mathbf{M}^{(i+1)}$ by computing a gradient step followed by a projection step:

$$\mathbf{M}^{(i+1)} = \mathcal{P}_{\mathcal{B}_{S_+}(\Phi)}(\mathbf{M}^{(i)} - \mu h^\dagger(h(\mathbf{M}^{(i)}) - \mathbf{Y})), \quad (11)$$

where $\mu > 0$. In [5], it is shown that choosing $\mu \approx N/Q$ is theoretically justifiable. However, in order to ensure the stability of the iterative projected gradient descent algorithm and accelerate convergence, in [34, 5] 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 (11). The resulting method, named Greedy Approximate Projection for MRF (GAP-MRF), is described in Algorithm 1. It can be noticed that the GAP-MRF method and BLIP are solving similar problems, using the same algorithmic structure. In this context, as in [5], 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.

3.2. Approximate projection

For every $\overline{\mathbf{M}} \in \mathbb{C}^{N \times L}$, we have:

$$\begin{aligned} \mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\Phi)}(\overline{\mathbf{M}}) &= \operatorname{argmin}_{\mathbf{M} \in \mathcal{B}_{\mathcal{S}_+}(\Phi)} \frac{1}{2} \|\mathbf{M} - \overline{\mathbf{M}}\|_2^2 \\ &= \operatorname{argmin}_{\mathbf{M}=\mathbf{X}\Phi, \mathbf{X} \in \mathcal{S}_+} \frac{1}{2} \|\mathbf{X}\Phi - \overline{\mathbf{M}}\|_2^2 \\ &= \left(\operatorname{argmin}_{\mathbf{X} \in \mathcal{S}_+} \frac{1}{2} \|\mathbf{X}\Phi - \overline{\mathbf{M}}\|_2^2 \right) \Phi, \end{aligned} \quad (12)$$

Note that $\mathcal{S}_2, \mathcal{S}_3$ and \mathcal{S}_4 can be handled through the definition of Φ . Let $\mathbf{M} = \mathbf{X}\Phi \in \mathcal{B}_{\mathcal{S}_+}(\Phi)$ and $\overline{T} \in \{1, \dots, K\}$ (K is the upper bound defined in (7)). Let $\mathbf{U} \in \mathbb{R}^{N \times \overline{T}}$ be a subpart of \mathbf{X} with non-zero columns and $\Delta \in \mathbb{C}^{\overline{T} \times L}$ the corresponding subpart of Φ such that $\mathbf{M} = \mathbf{U}\Delta$. Then we have

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\Phi)}(\overline{\mathbf{M}}) = \left(\operatorname{argmin}_{\mathbf{U} \in \mathbb{R}_+^{N \times \overline{T}}} \frac{1}{2} \|\mathbf{U}\Delta - \overline{\mathbf{M}}\|_2^2 \right) \Delta. \quad (13)$$

In (13), the dictionary Δ is defined as

$$\Delta = \operatorname{argmin}_{\Delta \in \mathcal{C}} \left(\min_{\overline{\mathbf{U}} \in \mathbb{R}_+^{N \times \overline{T}}} \frac{1}{2} \|\overline{\mathbf{U}}\Delta - \overline{\mathbf{M}}\|_2^2 \right), \quad (14)$$

where \mathcal{C} is the set given by

$$\begin{aligned} \mathcal{C} &= \left\{ \Delta \in \mathbb{C}^{\overline{T} \times L} \mid (\exists \mathbf{X} \in \mathcal{S}_+) \mathbf{X} = \mathcal{Z}(\overline{\mathbf{U}}) \right. \\ &\quad \left. \text{with } \overline{\mathbf{U}} = \operatorname{argmin}_{\mathbf{U} \in \mathbb{R}_+^{N \times \overline{T}}} \frac{1}{2} \|\mathbf{U}\Delta - \overline{\mathbf{M}}\|_2^2 \right\}. \end{aligned} \quad (15)$$

with $\mathcal{Z}: \mathbb{R}_+^{N \times \overline{T}} \rightarrow \mathbb{R}_+^{N \times D}$ 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. Recent advances in reconstruction methods have introduced neural networks to efficiently approximate the projection or proximal operators within model based iterative algorithms [35, 36, 37, 38]. A major challenge with such methods is obtaining sufficient accurate training data. In consequence, they can only accelerate the techniques where a prior computational solution to provide ground truth already exists. To overcome this difficulty, we propose a greedy approach to approximate the projection by finding a reduced dictionary $\tilde{\Delta} \in \mathbb{C}^{\overline{T} \times L}$ and its corresponding mixing matrix $\tilde{\mathbf{U}} \in \mathbb{C}^{N \times \overline{T}}$, with $\overline{T} \leq K$, such that $\mathbf{U}\Delta \approx \tilde{\mathbf{U}}\tilde{\Delta}$. Then the projection in step 10 of Algorithm 1 can be approximated as

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\Phi)}(\overline{\mathbf{M}}) \approx \left(\operatorname{argmin}_{\tilde{\mathbf{U}} \in \mathbb{R}_+^{N \times \overline{T}}} \frac{1}{2} \|\tilde{\mathbf{U}}\tilde{\Delta} - \overline{\mathbf{M}}\|_2^2 \right) \tilde{\Delta}. \quad (16)$$

As mentioned in [5], it is a common practice to allow the proton density to be complex-valued in order to absorb phase terms correcting for timing and coil sensitivity errors. We incorporate a vector $\lambda \in \mathbb{C}^N$ to compensate for these errors. Let $\tilde{\mathcal{B}}_{\mathcal{S}_+}(\Phi)$ be the set of magnetisation sequences of the form $\mathbf{M} = \operatorname{Diag}(\lambda)\mathbf{X}\Phi$ such that $\mathbf{X} \in \mathcal{S}_+$ and

$\boldsymbol{\lambda} \in \mathbb{C}^N$ satisfies $(\forall n \in \{1, \dots, N\}) |\boldsymbol{\lambda}_n| = 1$. The approximate projection with the phase compensation is given by:

$$\mathcal{P}_{\tilde{\mathcal{B}}_{S_+}(\Phi)}(\overline{\mathbf{M}}) \approx \text{Diag}(\boldsymbol{\lambda}) \tilde{\mathbf{U}} \tilde{\boldsymbol{\Delta}}, \quad (17)$$

where $(\boldsymbol{\lambda}, \tilde{\mathbf{U}})$ are obtained by solving:

$$\begin{aligned} & \text{minimise } \frac{1}{2} \|\text{Diag}(\boldsymbol{\lambda}) \tilde{\mathbf{U}} \tilde{\boldsymbol{\Delta}} - \overline{\mathbf{M}}\|_2^2 \\ & \boldsymbol{\lambda} \in \mathbb{C}^N, \tilde{\mathbf{U}} \in \mathbb{R}_+^{N \times T} \\ & \text{subject to } (\forall n \in \{1, \dots, N\}) |\boldsymbol{\lambda}_n| = 1. \end{aligned} \quad (18)$$

It is worth mentioning that in (16) and (18), all the rows of $\tilde{\mathbf{U}}$ can be computed independently in parallel.

On the one hand, forward-backward based algorithms [39, 7, 40] can be used to solve problem (16) (in particular, in our simulations, we use the built-in MATLAB function of non-negative least-squares, that is an implementation of [41]). On the other hand, to solve problem (18), to jointly estimate $\boldsymbol{\Lambda}$ and \mathbf{U} , block coordinate approaches must be considered (e.g. Gauss-Seidel approaches [18], alternating forward-backward methods [19, 20, 21, 22]). Note that in comparison with the traditional MRF methods which densely sample the manifold, our approach reduces the memory requirements, by using the dictionary $\tilde{\boldsymbol{\Delta}}$ containing at most K elements, without the inaccuracies related to the manifold discretisation.

3.3. Greedy dictionary estimation

The GAP-MRF algorithm takes advantage of the dictionary coherence and the constraints imposed on \mathbf{X} (described in Section 2.3) to approximate the projection onto $\mathcal{B}_{S_+}(\Phi)$ in line 10 of Algorithm 1. As described in Section 3.2, this projection can be approximated at each iteration $i \in \mathbb{N}$, by solving (16), which necessitates to estimate the dictionary $\tilde{\boldsymbol{\Delta}}^{(i)}$. 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 , \mathcal{S}_3 and \mathcal{S}_4 (note that the constraint \mathcal{S}_1 is handled directly in (16)). The proposed approach is described in details in this section.

The process to obtain $\tilde{\boldsymbol{\Delta}}^{(i)}$ 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 the set \mathcal{B}^+ defined as:

$$\begin{aligned} \mathcal{B}^+ &= \{\mathbf{M} \in \mathbb{C}^{N \times L} \mid (\forall n \in \{1, \dots, N\}) \mathbf{M}_{n,:} = \rho \mathbf{m}, \\ & \text{with } \rho \in \mathbb{R}_+ \text{ and } \mathbf{m} \in B(\mathcal{M}, \boldsymbol{\Gamma})\}. \end{aligned} \quad (19)$$

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 parameters that are too close to each other are discarded (\mathcal{S}_3 constraint) by using a Non-Maximum Suppression based method [42]. 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

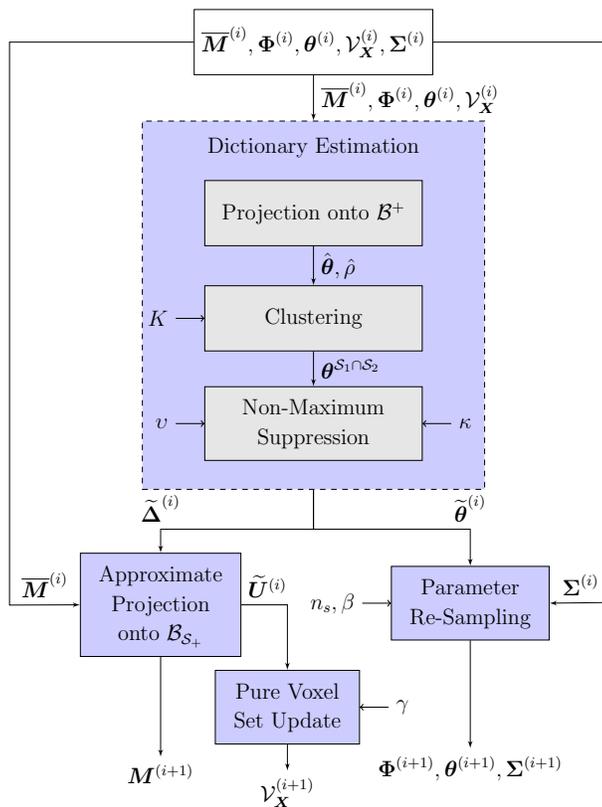


Figure 2. Greedy approximate projection diagram. The blue boxes represent the main steps in the approximate projection, the gray boxes represent the intermediate steps for the dictionary estimation and the arrows show the input and output variables.

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 $\tilde{\mathbf{U}}^{(i)}$ and the magnetisation sequence $\mathbf{M}^{(i+1)}$ using equation (16) with the resulting dictionary $\tilde{\Delta}^{(i)}$. Then, we update the pure voxel set $\mathcal{V}_{\mathbf{X}}$ using the mixing matrix $\tilde{\mathbf{U}}^{(i)}$. Finally, the dictionary Φ is refined by randomly sampling around the parameters $\tilde{\theta}^{(i)}$. The complete method is described in Algorithm 2 and explained in the following paragraphs.

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

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+(\Phi)}}(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}) = \mathcal{P}_{\mathcal{B}^+(\Phi)}(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}), \quad (20)$$

where \mathcal{B}^+ is the set defined in equation (19), and $\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:} = (\overline{\mathbf{M}}_{n,:})_{n \in \mathcal{V}_{\mathbf{X}}^{(i)}}$, $\mathcal{V}_{\mathbf{X}}^{(i)}$ 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 $\mathcal{V}_{\mathbf{X}}^{(0)} = \{1, \dots, N\}$, and it is updated during the greedy process (see Algorithm 2, step 20). Note that that the region containing pure voxel does not need to be known *a priori* (it is automatically estimated), and does not need to be large to be detected (depending on the noise).

Algorithm 2 Greedy Approximate Projection

-
- 1: **Input:** $\overline{\mathbf{M}}^{(i)}, \Phi^{(i)}, \theta^{(i)}, \mathcal{V}_{\mathbf{X}}^{(i)}, \Sigma^{(i)}, K, \Gamma, \kappa, v, \gamma, \beta, \xi, n_s$
 - 2: **Dictionary Estimation:**
 - 3: **Projection onto \mathcal{B}^+**
 - 4: **for** $n = 1, 2, \dots, N$ **do**
 - 5: $\hat{d}_n = \underset{d}{\operatorname{argmax}} \operatorname{real}(\overline{\mathbf{M}}_{n,:}^{(i)} \Phi_{d,:}^{\dagger(i)}) / \|\Phi_{d,:}^{(i)}\|_2$
 - 6: $\hat{\rho}_n = \max(\operatorname{real}(\overline{\mathbf{M}}_{n,:}^{(i)} \Phi_{\hat{d}_n,:}^{\dagger(i)}) / \|\Phi_{\hat{d}_n,:}^{(i)}\|_2, 0)$
 - 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: $[\theta^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}] = \text{k-means}(\hat{\theta}_{\mathcal{I},:}, K)$
 - 12: **Non-Maximum Suppression**
 - 13: $\tilde{\theta}^{(i)} = \text{NonMaximumSuppression}(\theta^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}, v, \kappa)$
 - 14: $\tilde{\Delta}^{(i)} = B(\tilde{\theta}^{(i)}, \Gamma)$
 - 15: **Approximate Projection onto $\mathcal{B}_{\mathcal{S}_+}$**
 - 16: $\tilde{\mathbf{U}}^{(i)} = \underset{\tilde{\mathbf{U}} \in \mathbb{R}_+^{N \times T}}{\operatorname{argmin}} \frac{1}{2} \|\tilde{\mathbf{U}} \tilde{\Delta}^{(i)} - \overline{\mathbf{M}}^{(i)}\|_2^2$
 - 17: $\mathbf{M}^{(i+1)} = \tilde{\mathbf{U}}^{(i)} \tilde{\Delta}^{(i)}$
 - 18: **Pure Voxel Set Update**
 - 19: $\mathcal{G}_{\mathbf{X}}^{(i)} = \{n \in \{1, \dots, N\} \mid \|\tilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\}$
 - 20: $\mathcal{V}_{\mathbf{X}}^{(i+1)} = \{n \in \mathcal{G}_{\mathbf{X}}^{(i)} \mid \max(\tilde{\mathbf{U}}_{n,:}^{(i)}) \geq \gamma \|\tilde{\mathbf{U}}_{n,:}^{(i)}\|_1\}$
 - 21: **Parameter Re-sampling**
 - 22: $\theta^{(i+1)} = \text{ParameterReSampling}(\tilde{\theta}^{(i)}, \Sigma^{(i)}, n_s)$
 - 23: $\Phi^{(i+1)} = B(\theta^{(i+1)}, \Gamma)$
 - 24: $\Sigma^{(i+1)} = \Sigma^{(i)} \beta$
 - 25: **Output:** $\theta^{(i+1)}, \Phi^{(i+1)}, \Sigma^{(i+1)}, \mathcal{V}_{\mathbf{X}}^{(i+1)}$ and $\mathbf{M}^{(i+1)}$
-

From (20), we can estimate the parameters $\hat{\theta}$ and the proton density $\hat{\rho}$ of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ using the projection onto \mathcal{B}^+ with a dictionary $\Phi^{(i)}$ (see steps 4-8 of Algorithm 2). $\Phi^{(i)}$ 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 (Left).

3.3.2. Clustering In order to find K centers approximating the parameters of interest, we propose to use the k-means algorithm [43]. The objective of k-means is to find K centers that minimise the squared distance from all points to its closest center. 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 (16) 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 $\mathcal{Z}(\mathbf{U}^{\mathcal{S}_1 \cap \mathcal{S}_2}) \in \mathcal{S}_1 \cap \mathcal{S}_2$.

3.3.3. 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 [42], we use the number of pure voxels assigned to each center to remove the neighbours defined in equation (8). 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. 3 (Center)). We repeat the process until all values of \mathbf{c} are 0. Finally, we use the resulting parameters $\tilde{\boldsymbol{\theta}}^{(i)} \in \mathbb{R}^{\bar{T} \times P}$ to construct $\tilde{\Delta}^{(i)} \in \mathbb{C}^{\bar{T} \times L}$.

3.3.4. Inexact projection onto $\mathcal{B}_{\mathcal{S}_+}$ Once the dictionary $\tilde{\Delta}^{(i)}$ is approximated, computing the three steps described above, the magnetisation sequence $\mathbf{M}^{(i+1)}$ can be updated. To this aim, we use equation (16), where the minimisation problem is solved using MATLAB built-in function for non-negative least-squares problems [41].

3.3.5. Pure voxel set update In order to avoid noisy voxels, we re-define the set $\mathcal{G}_{\mathbf{X}}$, introduced in Section 2.3.2, for $\tilde{\mathbf{U}}^{(i)}$. Note that $\mathcal{Z}(\tilde{\mathbf{U}}^{(i)})$ is a matrix of the size of \mathbf{X} filling the missing values of $\tilde{\mathbf{U}}^{(i)}$ with zeros, and thus we can re-define the set $\mathcal{G}_{\mathbf{X}}^{(i)}$ in terms of $\tilde{\mathbf{U}}^{(i)}$ as:

$$\mathcal{G}_{\mathbf{X}}^{(i)} = \{n \in \{1, \dots, N\} \mid \|\tilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\}. \quad (21)$$

Then, we update the pure voxel set as:

$$\mathcal{V}_{\mathbf{X}}^{(i+1)} = \{n \in \mathcal{G}_{\mathbf{X}}^{(i)} \mid \max(\tilde{\mathbf{U}}_{n,:}^{(i)}) \geq \gamma \|\tilde{\mathbf{U}}_{n,:}^{(i)}\|_1\}, \quad (22)$$

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. Note that the parameter γ is defined as a proportion of the total proton density in the voxels, and it is used as a threshold to determine if a voxel is pure or not.

3.3.6. 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 $\tilde{\boldsymbol{\theta}}^{(i)}$ using a Gaussian distribution with a diagonal covariance matrix $\Sigma^{(i)}$ (see Fig. 3 (Right)). 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 $\tilde{\Delta}$ 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.

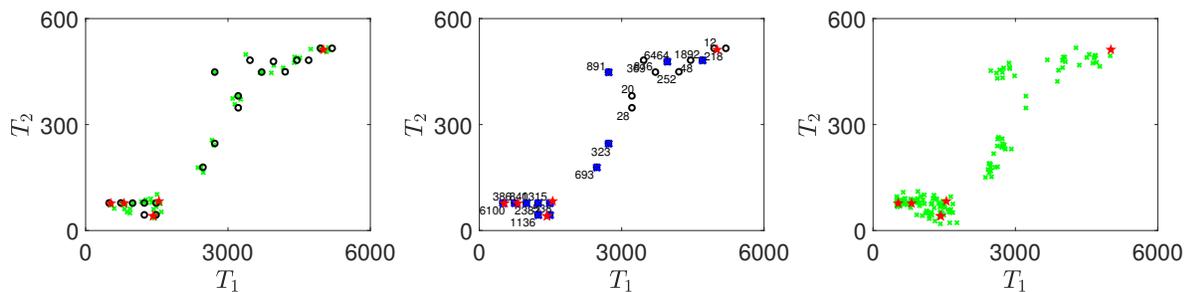


Figure 3. Examples of the clustering, non-maximum suppression and parameter re-sampling. For all examples the red stars represent the true phantom parameters. (Left) Clustering. The parameters of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ which its corresponding proton density is greater than ξ (green crosses) are the input of the k-means algorithm and the output are the centers (black circles). (Center) Non-maximum suppression. The centers obtained by the k-means (black circles) and the filtered centers are output of the Non-Maximum Suppression (blue crosses). (Right) Parameter re-sampling. The parameters of the dictionary $\Phi^{(i+1)}$ are obtained by randomly sampling around the parameters obtained by the Non-Maximum Suppression (green crosses).

4. Choice of the parameters and initialisation

Since \mathcal{S}_+ is a non-convex set, the choice of the initialisation is important. If the initial magnetisation sequence or the dictionary are not close to the desired values, the greedy approximate projection can fail. In this section, we will describe the initialisation for our algorithm.

4.1. Choice of the parameters

The choice of ξ , setting the minimum proton density, is related to the background noise, the ideal ξ is a value between the background noise and the signal in the volume of interest. If ξ is too small, empty voxels will affect the clustering process. If it is too big, the tissue voxels will not be considered in the clustering process.

As mentioned before, the dictionary $\Phi^{(i)}$ is updated through the iterations to reduce the complexity of the algorithm. We fix $\Phi^{(0)}$ to all possible combinations of 20 values of T_1 and 20 values of T_2 , equally spaced in \mathcal{M} .

Concerning the number of random samples n_s , on the one hand if we choose it too big, we increase the complexity of our pure voxel projection. On the other hand if we set n_s too small, more iterations will be needed to find the elements of interest. In all our simulations (simulated and *in vivo* data) we fix $n_s = 10$.

For the diagonal elements of the covariance matrix (i.e. $\Sigma_{1,1}^{(0)}$ and $\Sigma_{2,2}^{(0)}$) associated to the resampling of the dictionary, if they are chosen too big, the parameter sampling will be far from the parameters of interest, increasing the number of iterations required to find them. If they are too small, the algorithm may not find the parameter of interest. $\Sigma^{(0)}$ should be chosen based on the parameter separation of $\Phi^{(0)}$. In all the reconstructions we fix $\Sigma_{1,1}^{(0)} = 40$ and $\Sigma_{2,2}^{(0)} = 10$.

N	Number of voxels in the volume of interest
Q	Number of measurements per excitation and per coil
L	Number of excitation instances
$\mathbf{Y} \in \mathbb{C}^{Q \times L \times C}$	Measurement matrix
$\mathbf{M} \in \mathbb{C}^{N \times L}$	Magnetisation Response of the volume of interest, introduced in (1)
h	Linear operator from $\mathbb{C}^{N \times L}$ to $\mathbb{C}^{Q \times L \times C}$ defining the acquisition process
P	Number of the tissue parameters
$\mathcal{M} \subset \mathbb{R}^{1 \times P}$	Subset describing the feasible parameter space
B	Non-linear smooth operator from \mathcal{M} to $\mathbb{C}^{1 \times L}$ describing the magnetic resonance experiment, introduced in Section 2.2
$\Phi \in \mathbb{C}^{D \times L}$	Discretisation of B with D elements
\mathcal{B}^+	Set describing \mathbf{M} as a magnetisation response of pure voxels
$\mathbf{X} \in \mathbb{R}_+^{N \times D}$	Mixing matrix used to describe the PVE, introduced in Section 2.3
$\mathcal{B}_{S_+}(\Phi)$	Set describing all possible \mathbf{M} satisfying the proposed PV model, defined in (4)
S_+	Set describing the intersection of the sets S_1, S_2, S_3 and S_4 , defined in (5)
S_1	Positive orthant, defined in (6)
S_2	Set describing the maximum number of active dictionary elements, defined in (7) with: $\mathcal{G}_{\mathbf{X}}$ Set indicating the columns of \mathbf{X} contributing to the magnetisation response ξ Minimum voxel proton density $\mathcal{D}_{\mathbf{X}}$ Set giving the voxels with significant contribution to the magnetisation response K Maximum number of active dictionary elements
S_3	Set describing the constraint on the distinct active dictionary elements, defined in (8) with: v Constant used to define neighbourhoods for the dictionary elements in the parameter space \mathcal{N}_v Set describing the neighbour dictionary elements for given dictionary element d
S_4	Set giving the minimum number of pure voxels per active dictionary element, defined in (10) with: $\mathcal{V}_{\mathbf{X}}$ Set describing the pure voxels in \mathbf{X} κ Minimum number of pure voxels per active dictionary element
μ	Step size in Algorithm 1
$\mathbf{U} \in \mathbb{R}^{N \times \bar{T}}$	Sub-matrix of \mathbf{X} (see eq. (13))
$\Delta \in \mathbb{C}^{\bar{T} \times L}$	Sub-matrix of Φ (see eq. (14))
\mathcal{Z}	Operator from $\mathbb{R}_+^{N \times \bar{T}}$ to $\mathbb{R}_+^{N \times D}$ that maps a matrix \mathbf{U} to the corresponding \mathbf{X}
$\lambda \in \mathbb{C}^N$	Vetor used to compensate for the complex phase errors in the model
γ	Tolerance parameter for a pure voxels (see (22))
$\tau, \tau_k, \tau_v, \tau_\kappa$	Tolerance parameters for the initialisation process (see Section 4)

Table 1. Table of symbols.

Similarly, for the decreasing parameter β of the covariance matrix (see step 24 in Algorithm 2), if it is chosen too big, the algorithm will need more iterations to find the correct elements while if it is too small the algorithm may not explore the true parameters. We fix $\beta = 0.9$ in the considered scenarios.

The choice of the pure voxel tolerance γ is related to the noise and the accuracy of the dictionary during the iterations of the algorithm. If it is too big, the elements of interest could be eliminated through the iterations since pure voxels may be considered

as PV voxels, if it is too small, the PV voxels may be considered as pure affecting the clustering process. We found in our simulations that $\gamma = 0.85$ is a suitable choice.

The choice of the different parameters K , ν and κ has been investigated during preliminary work. In particular, we observed a significant increase in the residual $\|\mathbf{Y} - h(\mathbf{M})\|_2$ when K is not sufficiently large. For ν and κ , we see a significant increase in the residual when they are chosen too large (i.e. merging proton density maps of the true tissues), and an increase of noisy proton density maps when they are chosen too low.

We propose to automatically choose K , ν and κ by analysing the residual. Precisely, we choose a tolerance value on the residual, denoted by $\tau > 0$. This value, indicates the minimum contribution of an element of the dictionary in the residual. If τ is chosen too big, our solution will contain noisy elements. While if it is chosen too small our elements of interest will be removed from the reconstruction.

4.2. Initialisation

The global GAP-MRF method, including the initialisation process, is described in Algorithm 3. It describes the process to choose the parameters K , ν and κ . Firstly, the estimation of K is described in steps 2-10. Fixing all the other parameters, K is estimated by running multiple times the GAP-MRF iterations given in Algorithm 1. We assume that we have a suitable estimate of K when the stopping criteria given in step 10 of Algorithm 3 is reached. The same process is adopted for the estimation of ν described (steps 11-19) and κ (steps 20-28). For these two estimates, we allow for a small tolerance ($\tau_\nu > 0$ and $\tau_\kappa > 0$, respectively), for robustness purposes. Note that each new run of Algorithm 1 uses the previous estimated of \mathbf{M} , $\mathbf{\Delta}$ and $\mathbf{\theta}$, in order to accelerate the global method.

5. Simulations and results

In this section, we present the procedure used to evaluate the reconstruction with simulated data using a simulated PV phantom. For the sake of simplicity, we consider a particular case of model (1), with only one coil (i.e. $C = 1$) and the corresponding sensitivity map $\mathbf{S}_{:,1}$ to be the identity matrix. An Echo-planar Imaging (EPI) undersampling scheme is used [44, 45]. The Bloch Equations are used for the non-linear mapping, with the random flip angles α and fixed repetition times TR as described in [5]. We compare the BLIP algorithm [5] to the proposed GAP-MRF method, considering two different experiments (we also implemented the Bayesian method in [16] but the results were not meaningful due to the reduced number of acquisitions). In the first experiment, we investigate the effect of measurement noise by varying the input SNR (iSNR in dB), defined as $\text{iSNR} = 20 \log (\|h(\mathbf{M})\|_2 / (\sqrt{QLC}\sigma_{\mathbf{Y}}))$, where $\sigma_{\mathbf{Y}}$ is the standard deviation of the noise. We vary the iSNR from 10dB to 50dB. In the second experiment, we investigate the effect of the magnetisation sequence length $L \in [200, 600]$, affecting

Algorithm 3 GAP-MRF global method

```

1: Input:  $\mathbf{Y}$ ,  $\mathbf{\Gamma}$ ,  $\mathbf{\Phi}$ ,  $\boldsymbol{\theta}$ ,  $\xi$ ,  $\tau$ ,  $\boldsymbol{\Sigma}_{1,1}^{(0)} = 40$ ,  $\boldsymbol{\Sigma}_{2,2}^{(0)} = 10$ ,  $\zeta = 0.99$ ,  $\mathcal{V}_{\mathbf{X}}^{(0)} = \{1, \dots, N\}$ ,  $\beta = 0.9$ ,  $n_s = 10$ ,
    $\mathbf{M}^{(0)} = \mathbf{0}$ ,  $(\tau_K, \tau_v, \tau_\kappa) = (10, 0.02, 10)$ 
2: Estimation of  $K$ :
3:   Input:  $(\gamma, K, v, \kappa) = (0, 0, 0, 0)$ ,  $\tilde{\boldsymbol{\Delta}}^{(0)} = \{\}$ ,  $\tilde{\boldsymbol{\theta}}^{(0)} = \{\}$ ,  $j = 0$ .
4:   Do
5:      $\boldsymbol{\Phi}^{(0)} = \left[ \mathbf{\Phi}, \tilde{\boldsymbol{\Delta}}^{(j)} \right]$ ,  $\boldsymbol{\theta}^{(0)} = \left[ (\boldsymbol{\theta})^T, (\tilde{\boldsymbol{\theta}}^{(j)})^T \right]^T$ 
6:      $K = K + \tau_K$ 
7:      $\left[ \mathbf{M}^{(j+1)}, \tilde{\boldsymbol{\Delta}}^{(j+1)}, \tilde{\boldsymbol{\theta}}^{(j+1)} \right] = \text{Algorithm1}(\mathbf{Y}, \zeta, \mathbf{M}^{(j)})$ 
8:      $j=j+1$ 
9:   while  $\|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 > \tau$ .
10:  Output:  $K^* = K$ ,  $j = j - 1$ 
11: Estimation of  $v$ :
12:  Input:  $(\gamma, K, v, \kappa) = (0.85, K^*, 0, 0)$ 
13:  Do
14:     $\boldsymbol{\Phi}^{(0)} = \tilde{\boldsymbol{\Delta}}^{(j)}$ ,  $\boldsymbol{\theta}^{(0)} = \tilde{\boldsymbol{\theta}}^{(j)}$ 
15:     $v = v + \tau_v$ 
16:     $\left[ \mathbf{M}^{(j+1)}, \tilde{\boldsymbol{\Delta}}^{(j+1)}, \tilde{\boldsymbol{\theta}}^{(j+1)} \right] = \text{Algorithm1}(\mathbf{Y}, \zeta, \mathbf{M}^{(j)})$ 
17:     $j=j+1$ ;
18:  while  $\|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 > \tau$ .
19:  Output:  $v^* = v - 2\tau_v$ ,  $j = j - 1$ 
20: Estimation of  $\kappa$ :
21:  Input:  $(\gamma, K, v, \kappa) = (0.85, K^*, v^*, 0)$ 
22:  Do
23:     $\boldsymbol{\Phi}^{(0)} = \tilde{\boldsymbol{\Delta}}^{(j)}$ ,  $\boldsymbol{\theta}^{(0)} = \tilde{\boldsymbol{\theta}}^{(j)}$ 
24:     $\kappa = \kappa + \tau_\kappa$ 
25:     $\left[ \mathbf{M}^{(j+1)}, \tilde{\boldsymbol{\Delta}}^{(j+1)}, \tilde{\boldsymbol{\theta}}^{(j+1)} \right] = \text{Algorithm1}(\mathbf{Y}, \zeta, \mathbf{M}^{(j)})$ 
26:     $j=j+1$ ;
27:  while  $\|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 > \tau$ .
28:  Output:  $\kappa^* = \kappa - 2\tau_\kappa$ 
29: GAP-MRF Global Iterations:
30:  Input:  $(\gamma, K, v, \kappa) = (0.85, K^*, v^*, \kappa^*)$ ,  $\boldsymbol{\Phi}^{(0)} = \tilde{\boldsymbol{\Delta}}^{(j-1)}$ ,  $\boldsymbol{\theta}^{(0)} = \tilde{\boldsymbol{\theta}}^{(j-1)}$ 
31:   $[\bar{\mathbf{M}}, \bar{\boldsymbol{\Delta}}, \bar{\boldsymbol{\theta}}] = \text{Algorithm1}(\mathbf{Y}, \zeta, \mathbf{M}^{(j-1)})$ 
32:   $\bar{\mathbf{U}} = \underset{\mathbf{U} \in \mathbb{R}_+^{N \times T}}{\text{argmin}} \frac{1}{2} \|\mathbf{U} \bar{\boldsymbol{\Delta}} - \bar{\mathbf{M}}\|_2^2$ 
33:  Output:  $\bar{\mathbf{M}}$ ,  $\bar{\boldsymbol{\Delta}}$ ,  $\bar{\boldsymbol{\theta}}$  and  $\bar{\mathbf{U}}$ 

```

directly the acquisition time. In both the cases, we choose the undersampling ratio $N/Q = 16$ to simulate the EPI *in vivo* data in Section 6.

The BLIP algorithm and Algorithm 1 are stopped when the following stopping criterion is satisfied $|E^{(i+1)} - E^{(i)}| < 10^{-4} E^{(i+1)}$, where $E^{(i)} = \|h(\mathbf{M}^{(i)}) - \mathbf{Y}\|_2^2$, and $(\mathbf{M}^{(i)})_{i \in \mathbb{N}}$ is a sequence generated by the algorithms. In all simulations, GAP-MRF takes at most 120 iterations of Algorithm 1 to converge taking the initialisation into account. Both algorithms were implemented in MATLAB. For the longest test, BLIP takes around

30 minutes and GAP-MRF takes around 3 hours using a computer with a 3rd generation Quad Core Intel i5 processor. The computational time can be significantly improved with a parallel implementation of both algorithms.

5.1. Partial volume simulated phantom

We create a simulated phantom according to [46], with five tissues: adipose, WM, GM, muscle and cerebrospinal fluid (CSF). More precisely, to introduce the PVE, we use blocks of 2×2 voxels to form a lower resolution phantom containing PV voxels. The resulting volume is resized to 256×256 voxels, and we generate the magnetisation sequence from this volume using the Bloch equations. All the reconstructions are performed with the same resolution. In the first column of Fig. 4 proton density maps and the voxel distribution of the simulated phantom are shown. Using this representation we can see the structure of the tissues of interest. Traditionally in qMRI, individual parameter maps are evaluated since only a tissue per voxel is considered but in a PV scenario this is not meaningful since several parameter maps would be needed and visually do not show the tissue structures. We also compute the dominant tissue (highest proton density in the voxel) parameter maps for a traditional evaluation. The phantom dominant tissue parameter maps can be seen in the first column of Fig. 6. Note that for the construction of the phantom, we only consider in-plane PV, while in reality through-plane PV and in-plane PV occurs. Both kind of PV are modelled the same way and should not make any difference in the reconstructions.

5.2. Evaluation

In order to evaluate the algorithms, we use the Signal-to-Noise Ratio (SNR in dB) defined as $\text{SNR}(\mathbf{U}_{:,t}, \tilde{\mathbf{U}}_{:,t}) = 10 \log \left(\sum_{n=1}^N (\mathbf{U}_{n,t})^2 / \sum_{n=1}^N (\mathbf{U}_{n,t} - \tilde{\mathbf{U}}_{n,t})^2 \right)$, where $t \in \{1, \dots, T\}$ is the index of the evaluated tissue, \mathbf{U} is the mixing matrix ground truth and $\tilde{\mathbf{U}}$ is the estimation. Similarly for the magnetisation sequence SNR, we sum for all values in the matrix. To construct the matrix $\tilde{\mathbf{U}}$, a tolerance of 15% from the ground truth parameter values is used (i.e. for $T_1 = 530$ and $T_2 = 77$ milliseconds (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. The same definition of SR is used for both pure and PV voxels (considering only the corresponding phantom voxels). Due to noise, there could be small values in $\tilde{\mathbf{U}}$ that could significantly affect the SR. 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.

Table 2. Parameter values of example in Fig. 4 corresponding to Experiment 1 with an iSNR of 30dB. The relaxation times are in ms.

	Ground Truth		BLIP		GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	[460-590]	[74-84]	531.1	77.0
White Matter	811	77	[690-930]	[66-80]	811.1	77.0
Muscle	1425	41	[1220-1630]	[36-46]	1424.0	41.0
Gray Matter	1545	83	[1320-1610]	[74-86]	1544.3	83.1
CSF	5012	512	[4400-5000]	500	5013.1	512.1

5.3. Experiment 1 - Impact of the iSNR

In this experiment, we investigate the behaviour of both the BLIP and the GAP-MRF algorithms while changing the input noise. We fix the magnetisation sequence length $L = 1000$. The dictionary for BLIP is defined as in [5] with $D = 16170$. The results correspond to an average (with standard deviation) over 10 runs of each choice of iSNR.

The results of the proton density maps are shown in Fig. 7 (Left). GAP-MRF significantly outperform BLIP when the iSNR is greater than 30dB. We can notice that GAP-MRF estimates correctly the number of true atoms when the iSNR is 30dB or greater. The reconstruction of adipose tissue is more affected 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 PVE and also because of the dictionary inaccuracy. GAP-MRF magnetisation sequence SNR has a linear behaviour with respect to the iSNR. In Fig. 7 (Center), the SR with respect to the iSNR can be seen. We can observe that the SR is significantly affected by the iSNR.

The results for the dominant tissue parameter maps SNR can be seen in Fig. 7 (Right). GAP-MRF outperforms BLIP reconstructing the dominant tissue parameter maps. It is important to mention that GAP-MRF is more affected by noise because the linear combination of dictionary elements overfits the noise.

We show an example of the proton density maps for each tissue in Fig. 4 when the iSNR is 30dB. By visual inspection, we can observe that the GAP-MRF method outperforms the BLIP method for PV reconstructions for moderate noise scenarios. The values of BLIP in Table 2 are given in a range because multiple parameters were assigned to the corresponding ground truth tissue. On the contrary, GAP-MRF has a single value because only one value was assigned to the corresponding ground truth tissue. In this example, for BLIP and GAP-MRF respectively, the SNR values are as follows: 9.70dB and 11.94dB for Adipose, 9.14dB and 19.52dB for WM, 17.66dB and 39.29dB for Muscle, 8.31dB and 31.87dB for GM, 5.72dB and 52.60dB for CSF and for the magnetisation sequence 23.84dB and 48.18dB. The SR: 0.9944 and 0.9745 for pure voxels, and for PV voxels 0 and 0.9465. The GAP-MRF correctly estimates the manifold

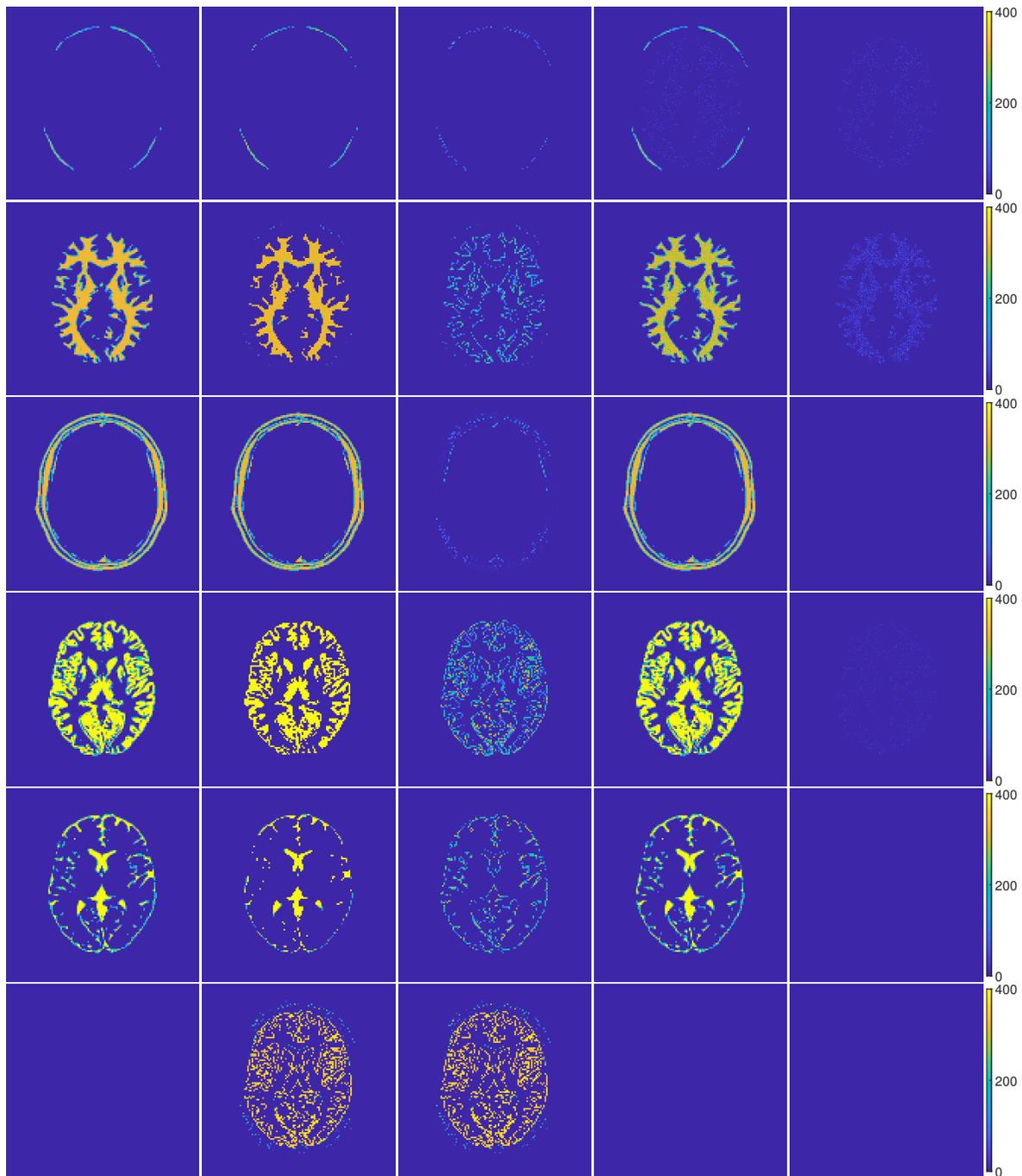


Figure 4. Experiment 1 - Example of the proton density maps with $L = 1000$ and an iSNR of 30dB. From first to last column: Ground truth images, BLIP reconstructions, absolute difference between BLIP and the Ground Truth, GAP-MRF reconstructions and absolute difference between GAP-MRF and the Ground Truth. From first to fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum 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 are given in Table 2.

regions of interest. BLIP has a residual map formed by all the elements that are not sufficiently close to the true elements (see last row of Fig. 4). Note that the residual

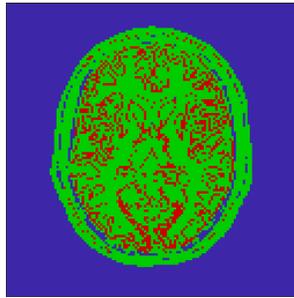


Figure 5. Voxel distribution map of the simulated phantom showing the pure voxels (green) and the PV voxels (red).

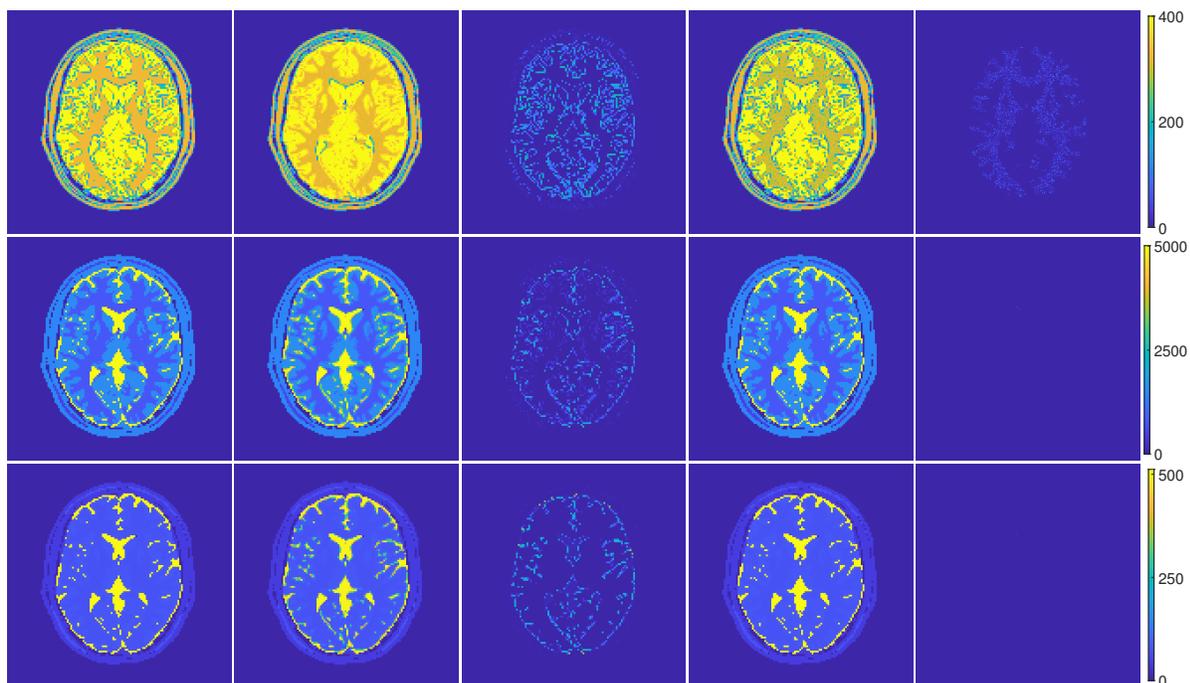


Figure 6. Experiment 1 - Example of the dominant tissue parameters with $L = 1000$ and an iSNR of 30dB. From first to last column: Ground truth images, BLIP reconstructions, absolute difference between BLIP and the Ground Truth, GAP-MRF reconstructions and absolute difference between GAP-MRF and the Ground Truth. From first to last row: Proton density, T_1 and T_2 parameter maps.

map is quite similar to the distribution of the PV voxels shown in Fig. 5, this shows that the parameter mismatch is due to the PVE. In the GAP-MRF reconstructions, the WM and adipose tissue are slightly mixed due to the noise since their parameters are close one to each other. By choosing a better $\mathbf{\Gamma}$ we can make the atoms of the dictionary more distant in the ℓ_2 -norm sense, this would provide noise robustness to the reconstructions. The dominant tissue parameter maps are shown in Fig. 6. The T_1 and T_2 maps reconstructed by BLIP show a smooth transition from one tissue to another due to the partial volume. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps delimiting the tissues. This is expected since each tissue is modelled with a unique set of parameters. We can observe that the dominant

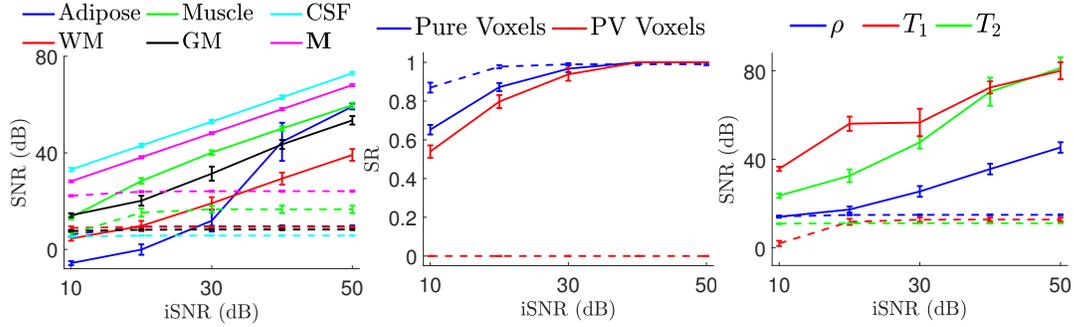


Figure 7. Experiment 1 - Simulation results obtained with BLIP (dashed lines) and GAP-MRF (solid lines). Left: Tissue proton density maps (Adipose, WM, Muscle, GM, CSF) and magnetisation sequence (\mathbf{M}) evaluation. Center: SR evaluation for the pure and PV voxels. Right: Dominant tissue parameter maps (ρ , T_1 , T_2) evaluation.

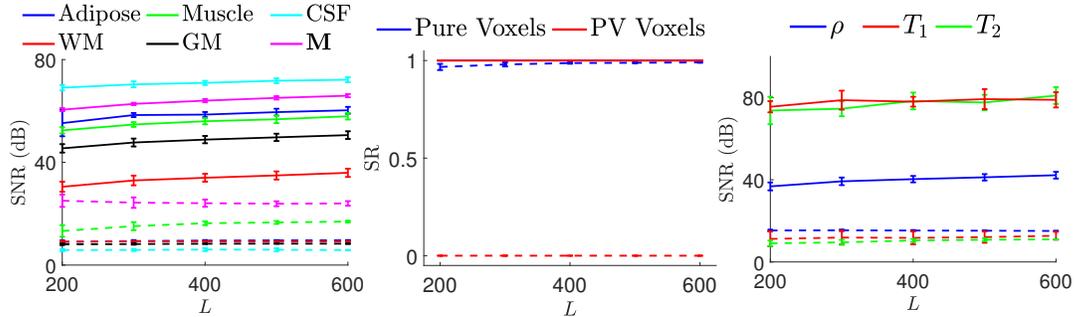


Figure 8. Experiment 2 - Simulation results obtained with BLIP (dashed lines) and GAP-MRF (solid lines). Left: Tissue proton density maps (Adipose, WM, Muscle, GM, CSF) and magnetisation sequence (\mathbf{M}) evaluation. Center: SR evaluation for the pure and PV voxels. Right: Dominant tissue parameter maps (ρ , T_1 , T_2) evaluation.

tissue proton density reconstruction of GAP-MRF is significantly affected by the noise. Nevertheless, thanks to the constraint \mathcal{S}_+ handled by the proposed method, the T_1 and T_2 parameter maps are accurate.

5.4. Experiment 2 - Impact of L

In this subsection, we compare the proposed GAP-MRF algorithm with the BLIP algorithm, for different number of excitation instances L . The iSNR is set to 50dB. The dictionary for BLIP is defined as in [5] with $D = 16170$. The results correspond to an average (with standard deviation) over 10 runs of each choice of L .

Fig. 8 (Left) shows the evaluation of the proton density maps for each tissue (Adipose, WM, Muscle, GM, and CSF) and the magnetisation sequence. Note that GAP-MRF results are taken directly from the matrix $\tilde{\mathbf{U}}$ without using any post-processing. We can observe that GAP-MRF outperforms BLIP in reconstructing \mathbf{U} . This can be explained by the fact that BLIP is restricted to the input dictionary, while our method estimates the dictionary. In addition, we can observe that the SNR values of magnetisation sequence reconstructed with BLIP slightly decreases while L increases

(while it is not the case for the proton density maps). This is expected since the linear combination of short fingerprints are less distinctive, hence it is easier to approximate it with fingerprints of other elements (allowing BLIP to fit better PV voxels with other elements). This behaviour is not observed with the proposed GAP-MRF method for which accurate proton density map estimates result in accurate magnetisation sequence reconstructions.

Fig. 8 (Center) gives the SR for both pure and PV voxels. Since BLIP can only reconstruct one element per voxel, its SR for PV is always equal to 0. In a low noise scenario, GAP-MRF can identify the correct voxel elements even for short sequences as can be seen in Fig. 8 (Center), where the SR of GAP-MRF for pure and PV voxels is 1. An important remark is that due to PV, the dictionary sampling and the number of excitation instances, the BLIP algorithm can mis-reconstruct pure voxels even in a low noise scenario.

For the dominant tissue parameter maps in the low noise scenario, GAP-MRF outperforms BLIP as shown in Fig. 8 (Right). The proton density map of BLIP is affected by the PV since it is not able to distinguish between the voxel tissues. The T_1 and T_2 maps are affected by the PV voxels and the dictionary inaccuracies.

BLIP reconstructions show a variation on T_1 and T_2 for the same tissue while GAP-MRF reconstructions are accurate. The GAP-MRF has the additional advantage that it simultaneously estimates the manifold regions of interest, resulting in better reconstructions.

6. Real data results

In this section, we show the reconstructions on the EUROSPIN phantom and on two *in vivo* datasets. The first and second datasets were acquired using spiral sampling scheme and the third dataset was acquired using EPI sampling scheme [47]. The parameters were chosen as discussed in Section 4. The obtained proton density maps were normalised as $\tilde{U}/\max(\tilde{U})$ and only the proton densities greater than the 10% of $\max(\tilde{U})$ are shown in the figures. The normalised proton density is in arbitrary units (a.u.) and the relaxation times are in ms. Note that for the spiral datasets, a single spiral interleaf is acquired for each excitation instance. For the next excitation instance, the interleaf is rotated a fixed angle given by the total number of interleaves (e.g. 377 interleaves corresponds to $360/377 \approx .9549^\circ$).

6.1. EUROSPIN phantom dataset with spiral sampling

In this subsection, we show the results obtained with the proposed approach and BLIP method, considering a dataset from 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 377 interleaves using FISP based α [48] and a constant $TR = 10\text{ms}$. The excitation sequence length is $L = 1000$. In this experiment, we have

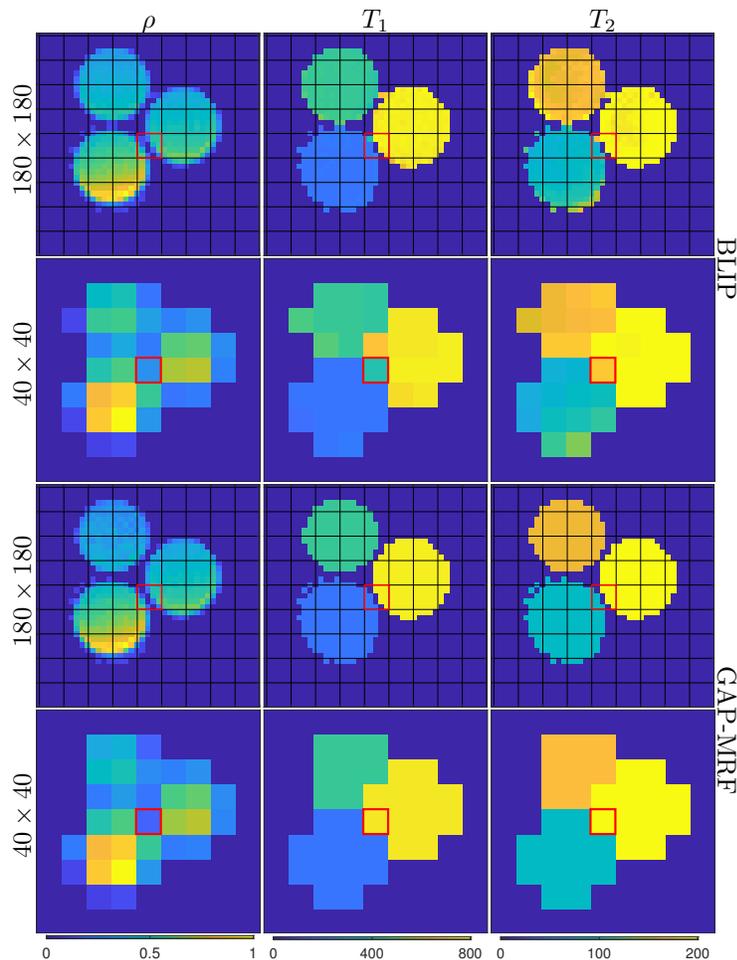


Figure 9. Dominant tissue parameter maps corresponding to the EUROSPIN phantom dataset. From first to last column: 180×180 BLIP reconstruction trimmed to 41×41 voxels, 40×40 BLIP reconstruction trimmed to 9×9 voxels, 180×180 GAP-MRF reconstruction trimmed to 41×41 voxels, 40×40 GAP-MRF reconstruction trimmed to 9×9 voxels. From first to last row: normalised proton density, T_1 , and T_2 .

FOV = $22.5 \times 22.5\text{cm}^2$ with a 5mm slice thickness. The EPG model is used for the reconstructions with an inversion time (TI) of 18ms and an Echo Time $TE = 1.902\text{ms}$. The scanned objects are the tubes 1, 5 and 9 of the EUROSPIN phantom. We reconstruct the parameter maps with two spatial resolutions: the first one at 180×180 with an undersampling ratio of $N/Q = 44.8753$, and the second one at 40×40 with an undersampling ratio of $N/Q = 20.6869$ to introduce the PV. Note that for the 40×40 reconstruction, only the Fourier samples corresponding to the target resolution are used. Reconstructing for higher spatial resolution would introduce high frequency artefacts as shown in [49]. An acquisition without the tubes is performed to estimate $\sigma_{\mathbf{Y}}$ and compute a lower bound on the iSNR. More precisely, using the triangle inequality, since $\|\mathbf{Y}\|_2 \geq \|\boldsymbol{\eta}\|_2$, we have $\text{iSNR} \geq 20 \log \left(\frac{(\|\mathbf{Y}\|_2 - \|\boldsymbol{\eta}\|_2)}{(\sqrt{QLC}\sigma_{\mathbf{Y}})} \right) = 64.73\text{dB}$, where \mathbf{Y} corresponds to the measurements with the tubes, $\boldsymbol{\eta}$ corresponds to the measurements without the tubes, and the value $\sigma_{\mathbf{Y}}$ is the standard deviation of $\boldsymbol{\eta}$.

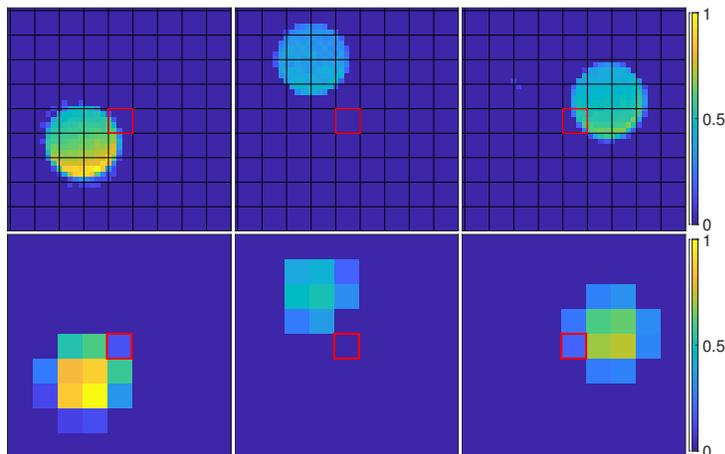


Figure 10. Normalised proton density maps corresponding to the EUROSPIN phantom dataset. The first and second row correspond to the 180×180 reconstruction trimmed to 41×41 , and the 40×40 reconstruction trimmed to 9×9 , respectively. The corresponding T_1 and T_2 can be seen in Table 3. From left to right the columns correspond to Tube 1, Tube 5 and Tube 9 of the EUROSPIN phantom.

Table 3. Comparison between the parameters obtained with GAP-MRF and the EUROSPIN phantom values.

	Phantom Values		180×180		40×40	
	T_1	T_2	T_1	T_2	T_1	T_2
Tube 1	200 ± 6	52 ± 1.6	197.0	93.9	195.4	96.0
Tube 5	450 ± 13.5	94 ± 2.8	455.9	159.8	459.5	168.1
Tube 9	754 ± 22.6	116 ± 3.5	766.3	199.2	757.5	199.9

The box in red shows a PV voxel artificially created by reconstructing a lower resolution image. As predicted by the corresponding high resolution maps, this voxel is formed by a linear combination of the Tubes 1 and 9.

In Fig. 9, we show a comparison of the reconstructions with two different spatial resolutions. The T_1 and T_2 lower resolution maps of BLIP show a variation introduced by the PVE. A clear example of the PVE is the voxel in the red box where two tissues appear, the BLIP method shows a parameter mismatch. Note that the parameters predicted by BLIP suggest that the voxel contains the same substance as Tube 5, contrary to the true composition (Tubes 1 and 9). GAP-MRF reconstructions do not show this behaviour since we take the PV into account in the model. Note that GAP-MRF is more sensitive to noise as shown in the simulations, this may explain small artefacts in the proton density maps.

The T_1 values in Table 3 are in agreement with the values of the EUROSPIN phantom. The T_2 values are higher than expected. As seen in Fig. 9, BLIP results show the same increased T_2 , suggesting that the errors may be related to the acquisition parameters specifically to the constant TR as shown in [50].

In Fig. 10, the normalised proton density maps reconstructed by GAP-MRF with two different resolutions can be seen. As highlighted by the red box, the PV voxel in

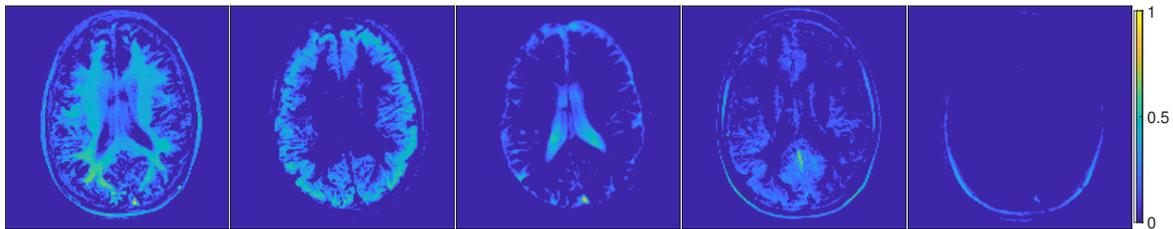


Figure 11. Normalised proton density maps corresponding to the brain dataset with spiral sampling. The corresponding T_1 and T_2 can be seen in Table 4. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right and top to bottom, the figures correspond to WM, GM, CSF, muscle and fat.

Table 4. Comparison between the parameters obtain with GAP-MRF for the brain dataset with spiral sampling and the reported values in [51].

	Values reported in [51]		GAP-MRF	
	T_1	T_2	T_1	T_2
WM	781 ± 61	65 ± 6	758.7	42.1
GM	1193 ± 65	109 ± 11	872.4	67.3
CSF			1658.5	799.8
Muscle	1100 ± 59	44 ± 9	1218.0	23.2
Fat	253 ± 42	68 ± 4	325.5	68.1

the low resolution reconstruction has values different than 0 in the maps corresponding to Tube 1 and 9, this is in agreement with the high resolution maps.

6.2. In vivo brain dataset with spiral sampling

This dataset was acquired by self-experimentation on our team members (all experts in MRI). Since these experiments are not intended to be qualified as a clinical investigation, they do not require any formal IRB approval according to the German Act on Medical Devices (Medical Device Act, MDA). The self-experiments were performed on a device that has already met the requirements of the assessment procedure of conformity, certifying its safety and functionality for the intended purpose (aka “CE marking of MR scanner”). The experiments were neither invasive nor stressful, therefore, they fully comply with internal GE and German/EU regulations. 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 [48]. The excitation sequence length is $L = 1000$. In this experiment, we have $FOV = 22.5 \times 22.5 \text{cm}^2$ and the spatial resolution is 180×180 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 reconstruction for BLIP and GAP-MRF was accelerated with the SVD compression in the time domain described in [13, 49] using 30 eigenvectors.

In Fig. 11, we can observe the resulting proton density maps provided by the GAP-MRF algorithm and the Table 4 shows a comparison between the parameters reported

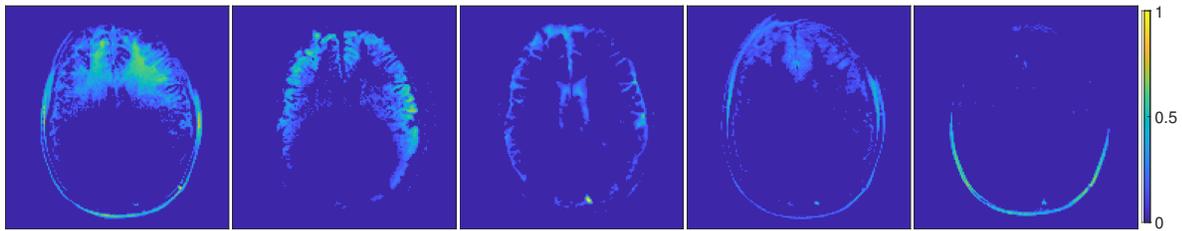


Figure 12. Normalised proton density maps corresponding to the brain dataset with spiral sampling. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right and top to bottom, the figures correspond to WM, GM, CSF, muscle and fat.

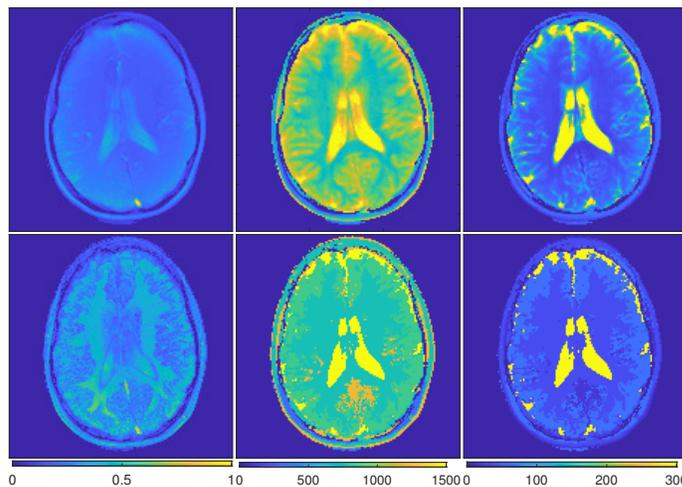


Figure 13. Dominant tissue parameter maps corresponding to the brain dataset with spiral sampling. The first row corresponds to the BLIP reconstructions and the second row to the GAP-MRF reconstructions. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right, the columns correspond to the normalised proton density, T_1 and T_2 . The values of T_1 and T_2 are capped to 1500ms and 300ms respectively.

in [51] and the parameters obtained by GAP-MRF. The WM, GM and Fat parameters obtained by GAP-MRF slightly differ from those reported for MRF sequences but the values are in agreement with the parameters of other qMRI methods reported in [51]. The muscle parameters are far from the expected values. 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 $\mathbf{\Gamma}$ to make the elements of the dictionary more distant in the ℓ_2 -norm sense can significantly improve the accuracy of the parameters. Also, inaccuracies in the model such as calibration or motion in the acquisition can produce artefacts in the reconstruction. In order to show the importance of the phase compensation in the real data, we present the results without phase compensation. As seen in Fig. 12, due to the phase errors there are voxels within the brain without proton density.

In Fig. 13, the T_1 and T_2 maps reconstructed by BLIP show a smooth transition from one tissue to another (similar to the simulated phantom). Moreover, the proton

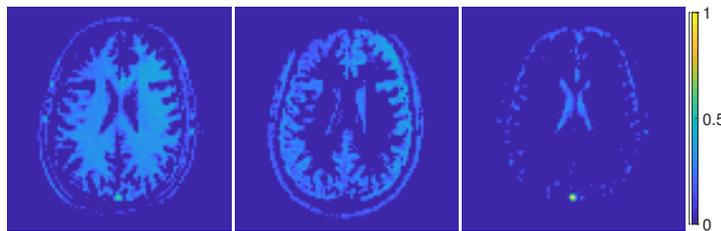


Figure 14. Normalised proton density maps corresponding to the brain dataset with EPI sampling. The reconstructions were trimmed from 128×128 to 89×89 voxels. The corresponding T_1 and T_2 can be seen in Table 5. From left to right, the figures correspond to WM, GM and CSF.

Table 5. Comparison between the parameters obtain with GAP-MRF for the brain dataset with EPI sampling and the reported values in [51].

	Values reported in [51]		GAP-MRF	
	T_1	T_2	T_1	T_2
WM	781 ± 61	65 ± 6	762.6	67.2
GM	1193 ± 65	109 ± 11	1116.6	107.1
CSF			2391.1	856.2

density map reconstructed by BLIP does not provide any information on the tissue distribution. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps of the dominant tissues. In addition, the proton density map shows more structure than BLIP, but not all the tissue structures are appreciated compared to the normalised proton density maps. Note that the voxels with higher proton density values indicate the pure voxels, and the voxels with reduced values, which are observed at tissue interfaces, indicate the dominant tissue that occupying only a fraction of the voxel.

6.3. In vivo brain dataset with EPI sampling

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, ranging from 1° to 70° . The excitation sequence length is $L = 500$. The repetition time TR was set to 16ms. In [50], it was shown to be as effective at estimating the MRF parameters but had better sensitivity than the FISP sequence in [48]. 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. A reference scan with null Gy gradient was performed for phase correction of EPI raw data.

In Fig. 14, we can observe the resulting proton density maps provided by the GAP-

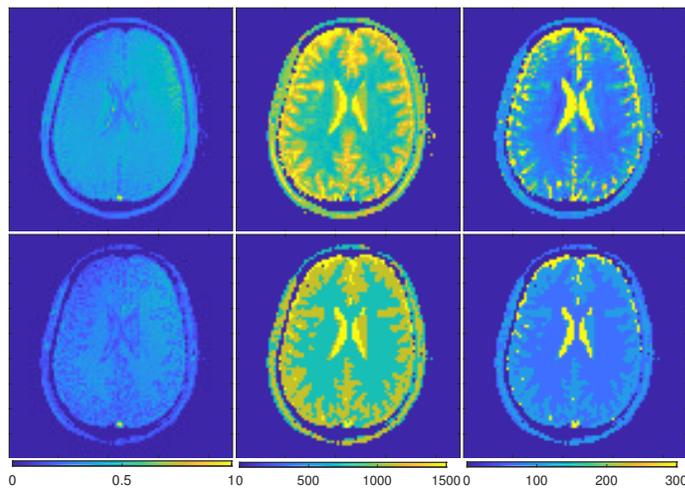


Figure 15. Dominant tissue parameter maps corresponding to the brain dataset with EPI sampling. The first row corresponds to the BLIP reconstructions and the second row to the GAP-MRF reconstructions. The reconstructions were trimmed from 128×128 to 89×89 voxels. From left to right, the columns correspond to the normalised proton density, T_1 and T_2 . The values of T_1 and T_2 are capped to 1500ms and 300ms respectively.

MRF algorithm and the Table 5 shows a comparison between the parameters reported in [51] for MRF FISP sequences and the parameters obtained by GAP-MRF. CSF values are not reported for the MRF FISP sequence. The WM parameters are similar to the ones reported in [51] and the GM T_1 is slightly lower than the reported one. We believe that the lack of pure voxels (due the spatial resolution) made the approach unable to find the other tissues.

In Fig. 15, the T_1 and T_2 maps reconstructed by BLIP shows a smooth transition from one tissue to another. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps of the dominant tissues.

7. Conclusions and future work

We have presented an extension of the model in [5] to PV reconstructions in the context of MRF. Our algorithm provides a way to explore the manifold of magnetic resonance fingerprints without densely sampling \mathcal{M} . For this reason, the algorithm is memory efficient and the algorithmic structure allows parallel implementations. The proposed model assumes that the number of independent tissues in the imaged volume is upper bounded, and that each tissue has a minimum number of pure voxels. Also, the parameters of each tissue should be sufficiently different to be distinguished. Finally, we assume that the combination of the sampling patterns should cover most of the k -space to avoid high frequency artefacts.

The simulation results presented in Section 5 show that the proposed GAP-MRF method can achieve accurate reconstructions with very short pulse sequences in the low input noise scenario. It also performs well when the iSNR is greater than 30dB. We also

present in Section 6 the results obtained with *in vivo* datasets. 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 PV reconstructions with a real PV phantom in the scanner and a full brain reconstruction to provide enough pure voxels to accurately estimate the true parameters. In particular, an interesting point would be to evaluate the behaviour of GAP-MRF in presence of a pathology. A pathology can be seen as a distinct additional tissue. Therefore, since the number of tissues is estimated along the iterations, if the pathology is represented by enough pure voxels, it should be detected by the algorithm exactly in the same way as for the other tissues. In addition, we plan to incorporate spatial regularisation in the objective function to improve the robustness of the method. A joint calibration and imaging problem will also be developed in order to provide both phase estimation and compensation. Finally, we acknowledge that deep learning is an interesting research direction to accelerate the projection onto \mathcal{B}_{S_+} .

Acknowledgements

R. D. would like to thank CONACYT and Heriot-Watt University for the PhD funding. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC, grants EP/M019306/1 and EP/M019802/1). We would like to thank GE for the access to the *in vivo* datasets and Benjamin Arnold for acquiring the EPI dataset. We also would like to thank Zhouye Chen and Mohammad Golbabaei for their help in the code and the BASP Group in Heriot-Watt University for the insightful discussions.

References

- [1] Jara H 2013 *Theory of Quantitative Magnetic Resonance Imaging* (World Scientific) ISBN 9789814295239 URL https://books.google.co.uk/books?id=w_Y8QwAACAAJ
- [2] Ma D, Gulani V, Seibelich N, Liu K, Duerk J L and Griswold M A 2013 Magnetic resonance fingerprinting *Nature* **465** 187–192 URL <http://dx.doi.org/10.1038/nature11971>
- [3] Donoho D L 2006 Compressed sensing *IEEE Trans. Inform. Theory* **52**(4) 1289–1306
- [4] Doneva M, Brnert P, Eggers H, Stehning C, Sngas J and Mertins A 2010 Compressed sensing reconstruction for magnetic resonance parameter mapping *Magnetic Resonance in Medicine* **64**(4) 1114–1120 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.22483> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22483>
- [5] Davies M, Puy G, Vandergheynst P and Wiaux Y 2014 A Compressed Sensing Framework for Magnetic Resonance Fingerprinting *SIAM J. Imaging Sci.* **7**(4) 2623
- [6] Beck A and Teboulle M 2009 A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems *SIAM J. Imaging Sci.* **2**(1) 183–202 URL <http://dx.doi.org/10.1137/080716542>
- [7] Tseng P 2000 A modified forward-backward splitting method for maximal monotone mappings *SIAM J. Control Optim.* **38**(2) 431–446 <https://doi.org/10.1137/S0363012998338806> URL <https://doi.org/10.1137/S0363012998338806>
- [8] Pierre E Y, Ma D, Chen Y, Badve C and Griswold M A 2016 Multiscale reconstruction for mr fingerprinting *Magnetic Resonance in Medicine* **75**(6) 2481–2492 <https://doi.org/10.1002/mrm.24811>

- [//onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.25776](https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.25776) URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.25776>
- [9] Zhao B, Setsompop K, Ye H, Cauley S F and Wald L L 2016 Maximum likelihood reconstruction for magnetic resonance fingerprinting *IEEE Transactions on Medical Imaging* **35**(8) 1812–1823
- [10] Zhao B, Setsompop K, Adalsteinsson E, Gagoski B, Ye H, Ma D, Jiang Y, Ellen Grant P, Griswold M A and Wald L L 2017 Improved magnetic resonance fingerprinting reconstruction with low-rank and subspace modeling *Magnetic Resonance in Medicine* **79**(2) 933–942 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.26701> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26701>
- [11] Doneva M, Amthor T, Koken P, Sommer K and Brnert P 2017 Matrix completion-based reconstruction for undersampled magnetic resonance fingerprinting data *Magnetic Resonance Imaging* **41** 41 – 52 URL <http://www.sciencedirect.com/science/article/pii/S0730725X17300383>
- [12] Mazor G, Weizman L, Tal A and Eldar Y C 2018 Low-rank magnetic resonance fingerprinting *Medical Physics* **45**(9) 4066–4084 <https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.13078> URL <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13078>
- [13] McGivney D F, Pierre E, Ma D, Jiang Y, Saybasili H, Gulani V and Griswold M A 2014 Svd compression for magnetic resonance fingerprinting in the time domain *IEEE transactions on medical imaging* **33**(12) 2311–2322
- [14] Tohka J 2014 Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review *World J Radiol* **6**(11) 855–864 URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4241492/>
- [15] Deshmane A, McGivney D F, Ma D, Jiang Y, Badve C, Gulani V, Seiberlich N and Griswold M A 2019 Partial volume mapping using magnetic resonance fingerprinting *NMR in Biomedicine* **32**(5) e4082 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.4082> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4082>
- [16] McGivney D, Deshmane A, Jiang Y, Ma D, Badve C, Sloan A, Gulani V and Griswold M 2018 Bayesian estimation of multicomponent relaxation parameters in magnetic resonance fingerprinting *Magnetic Resonance in Medicine* **80**(1) 159–170 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.27017> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27017>
- [17] Tang S, Fernandez-Granda C, Lannuzel S, Bernstein B, Lattanzi R, Cloos M, Knoll F and Assländer J 2018 Multicompartment magnetic resonance fingerprinting *Inverse problems* **34**(9) 094005
- [18] Bertsekas D P 1999 *Nonlinear programming* (Athena scientific Belmont)
- [19] Tseng P 2001 Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization *Journal of Optimization Theory and Applications* **109**(3) 475–494 URL <https://doi.org/10.1023/A:1017501703105>
- [20] Bolte J, Sabach S and Teboulle M 2014 Proximal Alternating Linearized Minimization for Nonconvex and Nonsmooth Problems *Math. Program.* **146**(1-2) 459–494 URL <http://dx.doi.org/10.1007/s10107-013-0701-9>
- [21] Frankel P, Garrigos G and Peypouquet J 2015 Splitting Methods with Variable Metric for Kurdyka–Lojasiewicz Functions and General Convergence Rates *Journal of Optimization Theory and Applications* **165**(3) 874–900 URL <https://doi.org/10.1007/s10957-014-0642-3>
- [22] Chouzenoux E, Pesquet J C and Repetti A 2016 A block coordinate variable metric forward-backward algorithm *Journal of Global Optimization* pp. 1–29
- [23] Bauschke H and Combettes P L 2017 *Convex Analysis and Monotone Operator Theory in Hilbert Spaces* (Springer)
- [24] Rockafellar R T and Wets R J B 2009 *Variational analysis* vol 317 (Springer Science & Business Media)
- [25] Weigel M 2015 Extended phase graphs: dephasing, RF pulses, and echoes-pure and simple *Journal of Magnetic Resonance Imaging* **41**(2) 266–295

- [26] Candes E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inform. Theory* **52**(2) 489–509
- [27] Donoho D L and Huo X 2001 Uncertainty principles and ideal atomic decomposition *IEEE Trans. Inform. Theory* **47**(7) 2845–2862
- [28] Donoho D L and Elad M 2003 Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization *Proceedings of the National Academy of Sciences* **100**(5) 2197–2202 <http://www.pnas.org/content/100/5/2197.full.pdf> URL <http://www.pnas.org/content/100/5/2197>
- [29] Davenport M A, Needell D and Wakin M B 2013 Signal space cosamp for sparse recovery with redundant dictionaries *IEEE Transactions on Information Theory* **59**(10) 6820–6829
- [30] Giryes R and Elad M 2013 Iterative hard thresholding with near optimal projection for signal recovery *10th Intl. Conf. on Sampling Theory and Appl.(SAMPTA)*
- [31] Chen S S, Donoho D L and Saunders M A 2001 Atomic decomposition by basis pursuit *SIAM J. Sci. Comput.* **43** 129–159
- [32] Candes E J, Eldar Y C, Needell D and Randall P 2011 Compressed sensing with coherent and redundant dictionaries *Applied and Computational Harmonic Analysis* **31**(1) 59 – 73 URL <http://www.sciencedirect.com/science/article/pii/S1063520310001156>
- [33] Blumensath T 2011 Sampling and Reconstructing Signals From a Union of Linear Subspaces *IEEE Trans. Inform. Theory* **57**(7) 4660–4671
- [34] Blumensath T and Davies M E 2010 Normalized Iterative Hard Thresholding: Guaranteed Stability and Performance *IEEE J. Selected Topics Signal Process.* **4**(2) 298–309
- [35] Golbabaee M, Chen D, Gómez P A, Menzel M I and Davies M E 2018 A deep learning approach for magnetic resonance fingerprinting *CoRR* **abs/1809.01749** 1809.01749 URL <http://arxiv.org/abs/1809.01749>
- [36] Cohen O, Zhu B and Rosen M S 2018 Mr fingerprinting deep reconstruction network (drone) *Magnetic resonance in medicine* **80**(3) 885–894
- [37] Hoppe E, Körzdörfer G, Würfl T, Wetzl J, Lugauer F, Pfeuffer J and Maier A K 2017 Deep learning for magnetic resonance fingerprinting: A new approach for predicting quantitative parameter values from time series. *GMD S* pp. 202–206
- [38] Virtue P, Stella X Y and Lustig M 2017 Better than real: Complex-valued neural nets for mri fingerprinting *2017 IEEE International Conference on Image Processing (ICIP)* (IEEE) pp. 3953–3957
- [39] Chen G H G and Rockafellar R T 1997 Convergence Rates in Forward–Backward Splitting *SIAM J. Optim.* **7**(2) 421–444 <https://doi.org/10.1137/S1052623495290179> URL <https://doi.org/10.1137/S1052623495290179>
- [40] Combettes P L and Wajs V R 2005 Signal Recovery by Proximal Forward-Backward Splitting *Multiscale Modeling & Simulation* **4**(4) 1168–1200 <https://doi.org/10.1137/050626090> URL <https://doi.org/10.1137/050626090>
- [41] Lawson C L and Hanson R J 1995 *Solving least squares problems Classics in Applied Mathematics*, vol 15 (Philadelphia, PA: SIAM) ISBN 0-89871-356-0
- [42] Canny J 1986 A Computational Approach to Edge Detection *IEEE Trans. Pattern Anal. Mach. Int.* **PAMI-8**(6) 679–698
- [43] Lloyd S 1982 Least squares quantization in PCM *IEEE Trans. Inform. Theory* **28**(2) 129–137
- [44] Rieger B, Zimmer F, Zapp J, Weingrtnr S and Schad L R 2017 Magnetic resonance fingerprinting using echo-planar imaging: Joint quantification of T1 and T2* relaxation times *Magnetic Resonance in Medicine* **78**(5) 1724–1733 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.26561> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26561>
- [45] Rieger B, Akçakaya M, Pariente J C, Llufríu S, Martínez-Heras E, Weingärtner S and Schad L R 2018 Time efficient whole-brain coverage with mr fingerprinting using slice-interleaved echo-planar-imaging *Scientific Reports* **8**(1) 6667 URL <https://doi.org/10.1038/s41598-018-24920-z>

- [46] Collins D L, Zijdenbos A P, Kollokian V, Sled J G, Kabani N J, Holmes C J and Evans A C 1998 Design and construction of a realistic digital brain phantom *IEEE Trans. Med. Imag.* **17**(3) 463–468
- [47] Benjamin A, Gómez P, Golbabaee M, Mahbub Z, Sprenger T, Menzel M, Davies M and Marshall I 2018 Balanced multi-shot EPI for accelerated Cartesian MR Fingerprinting: An alternative to spiral MR Fingerprinting *Proc Intl Soc Mag Reson Med* p. 4265
- [48] Jiang Y, Ma D, Seiberlich N, Gulani V and Griswold M A 2015 MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout *Magnetic Resonance in Medicine* **74**(6) 1621–1631
- [49] Cline C C, Chen X, Mailhe B, Wang Q, Pfeuffer J, Nittka M, Griswold M A, Speier P and Nadar M S 2017 AIR-MRF: Accelerated iterative reconstruction for magnetic resonance fingerprinting *Magnetic Resonance Imaging* **41** 29 – 40 URL <http://www.sciencedirect.com/science/article/pii/S0730725X1730125X>
- [50] Gómez P A, Bounincontri G, Molina-Romero M, Sperl J I, Menzel M I and Menze B H 2017 Accelerated parameter mapping with compressed sensing: an alternative to MR Fingerprinting *Proc Intl. Soc. Mag. Reson. Med*
- [51] Bojorquez J Z, Bricq S, Acquitte C, Brunotte F, Walker P M and Lalande A 2017 What are normal relaxation times of tissues at 3T? *Magnetic Resonance Imaging* **35** 69 – 80 URL <http://www.sciencedirect.com/science/article/pii/S0730725X16301266>