

# Multicompartment Magnetic Resonance Fingerprinting

Sunli Tang<sup>\*1</sup>, Carlos Fernandez-Granda<sup>\*1,2</sup>, Sylvain Lannuzel<sup>2,3</sup>, Brett Bernstein<sup>1</sup>, Riccardo Lattanzi<sup>4,5</sup>, Martijn Cloos<sup>4,5</sup>, Florian Knoll<sup>4,5</sup>, and Jakob Assländer<sup>4,5</sup>

<sup>1</sup>Courant Institute of Mathematical Sciences, New York University

<sup>2</sup>Center for Data Science, New York University

<sup>3</sup>École CentraleSupélec

<sup>4</sup>Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine

<sup>5</sup>Center for Advanced Imaging and Innovation Research (CAI<sup>2</sup>R) Department of Radiology, New York University School of Medicine

February 2018; Revised June 2018

## Abstract

Magnetic resonance fingerprinting (MRF) is a technique for quantitative estimation of spin-relaxation parameters from magnetic-resonance data. Most current MRF approaches assume that only one tissue is present in each voxel, which neglects intravoxel structure, and may lead to artifacts in the recovered parameter maps at boundaries between tissues. In this work, we propose a multicompartment MRF model that accounts for the presence of multiple tissues per voxel. The model is fit to the data by iteratively solving a sparse linear inverse problem at each voxel, in order to express the measured magnetization signal as a linear combination of a few elements in a precomputed fingerprint dictionary. Thresholding-based methods commonly used for sparse recovery and compressed sensing do not perform well in this setting due to the high local coherence of the dictionary. Instead, we solve this challenging sparse-recovery problem by applying reweighted- $\ell_1$ -norm regularization, implemented using an efficient interior-point method. The proposed approach is validated with simulated data at different noise levels and undersampling factors, as well as with a controlled phantom-imaging experiment on a clinical magnetic-resonance system.

**Keywords.** Quantitative MRI, magnetic resonance fingerprinting, multicompartment models, parameter estimation, sparse recovery, coherent dictionaries, reweighted  $\ell_1$ -norm.

---

\*Sunli Tang and Carlos Fernandez-Granda contributed equally to this paper.

# 1 Introduction

## 1.1 Quantitative magnetic resonance imaging

Magnetic resonance imaging (MRI) is a medical-imaging technique, which has become a key technology for non-invasive diagnostics due to its excellent soft-tissue contrast. MRI is based on the nuclear magnetic resonance phenomenon, in which the nuclear spin of certain atoms absorb and emit electromagnetic radiation. To perform MRI, subjects are placed in a strong magnetic field, which generates a macroscopic net magnetization of the spin ensemble. Radio-frequency (RF) pulses are used to manipulate this magnetization, causing it to precess. As a result, an electro-magnetic signal is emitted, which can be measured and processed. The dynamics of the magnetization are commonly described by the Bloch equation [9]:

$$\frac{\partial}{\partial t} \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} = \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \times \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} - \begin{pmatrix} \frac{M_x}{T_2} \\ \frac{M_y}{T_2} \\ \frac{M_z - 1}{T_1} \end{pmatrix}, \quad (1.1)$$

where  $\partial/\partial t$  is a partial derivative with respect to time.  $M_x$ ,  $M_y$  and  $M_z$  are functions of time which denote the magnetization along the three spatial dimensions. The Larmor frequency  $\omega_z$  is the frequency at which the spins precess. Magnetic fields, called *gradient fields* because they vary linearly along different spatial dimensions, are used to modify  $\omega_z$  and select the specific region to be imaged, as well as to encode the magnetization in the Fourier domain (see [57] for more details). The frequencies  $\omega_x$  and  $\omega_y$  are varied over time using RF pulses in order to create a specific spin evolution that makes spin-relaxation effects accessible. This evolution is governed by the time constants  $T_1$  and  $T_2$ , which are tissue specific and serve as valuable biomarkers for various pathologies.

In traditional MRI techniques, a sequence of RF pulses, and hence of frequencies  $\omega_x$  and  $\omega_y$ , is designed so that the resulting signal at each voxel of an image is predominantly weighted by the values of  $T_1$  or  $T_2$ . MR images obtained in this way are qualitative in nature, in contrast to other medical-imaging modalities such as computed tomography or positron emission tomography. The images consist of gray values that capture relative signal intensity changes between tissues, caused by their different  $T_1$  and  $T_2$  values, which are then interpreted by radiologists to detect pathologies. The specific numerical value at each voxel is typically subject to variations between different MRI systems and data-acquisition settings. As a result, it is very challenging to use data from current clinical MRI systems for longitudinal studies, early detection and progress-tracking of disease, and computer-aided diagnosis.

The goal of quantitative MRI is to measure physical parameters such as  $T_1$  and  $T_2$  quantitatively, in a way that is *reproducible* across different MRI systems [23,35,67,71]. Quantitative MRI techniques can be used to extract quantitative biomarkers [53,56,63] and synthesize images with standardized contrasts [26,59]. Unfortunately, existing data-acquisition protocols for quantitative MRI often lead to long measurement times that are challenging to integrate in the clinical work-flow and as a consequence are not widely deployed.

## 1.2 Magnetic resonance fingerprinting

Magnetic resonance fingerprinting (MRF) [48] is a recently-proposed quantitative MRI technique to measure tissue-specific parameters within scan times that are clinically feasible. In many traditional MRI techniques, the same RF pulse is applied repeatedly, driving the magnetization into a steady state in which the signal has a fixed weighting of  $T_1$  and  $T_2$  effects. In MRF, the amplitudes of the RF pulses are varied over time to deliberately avoid a steady state of the magnetization, creating a time-dependent signal weighting. This makes it possible to estimate the relaxation times  $T_1$  and  $T_2$  quantitatively by fitting the corresponding signal model, governed by the Bloch equation (1.1). Unfortunately, it is very challenging to fit this model by directly minimizing the fitting error, as this results in a highly non-convex cost function. Instead, MRF methodology fits the model by comparing the observed time evolution of the magnetization signal to a dictionary of *fingerprints*, which are precomputed by solving the Bloch equation (1.1) numerically for a range of  $T_1$  and  $T_2$  relaxation times.

Current MRF methods usually operate under the assumption that only one type of tissue is present in each volume element, called voxel. However, this assumption is violated at boundaries between tissues. As mentioned previously, spatial information in MR signals is typically encoded in the Fourier domain. This domain can only be measured up to a certain cut-off frequency, which results in a limited spatial resolution, typically of the order of a millimeter. Thus, sharp boundaries between different tissues may be blurred, so that voxels close to the boundary contain several tissue compartments (see Figure 1 for a concrete example). Fitting a single-compartment model to data obtained from a voxel with multiple tissues yields erroneous parameter estimates that may not correspond to any of the contributing compartments, a phenomenon known as the *partial-volume* effect [72]. Consequently, single-compartment MRF methods often do not accurately characterize boundaries between tissues, as discussed by the original developers of MRF [26]. This limitation is particularly problematic in applications that require geometric measurements, such as the diagnostic of Alzheimers disease, where cortical thickness provides a promising biomarker for early detection [40].

The cellular microstructure of biological tissues results in multiple tissue compartments being present even in voxels that are not located at tissue boundaries. An important example is white matter in the brain, which contains extra- and intra-axonal water with similar relaxation times, as well as water trapped between the myelin sheets, which has substantially shorter relaxation times [44]. If a disease causes demyelination, the *myelin-water fraction*, which is the fraction of the proton density in the voxel corresponding to myelin-water, is reduced [45]. Multicompartment models capable of evaluating the myelin-water fraction are therefore useful for the diagnosis of neurodegenerative diseases such as multiple sclerosis [44, 45]. Single-compartment models cannot be applied in such settings because they fail to account for this microstructure, as illustrated in Figure 2.

It is important to note that the presence of different tissues does not necessarily result in entirely separated compartments within a voxel. In fact, one can observe chemical exchange of water molecules between compartments [1]. However, in the case of tissue boundaries, these effects play a subordinate role due to the macroscopic nature of the interfaces. Similarly, in the case of myelin-water imaging, neglecting chemical exchange is a reasonable and commonly-used assumption in literature [45]. In this work, we consider a multicompartment model that incorporates this

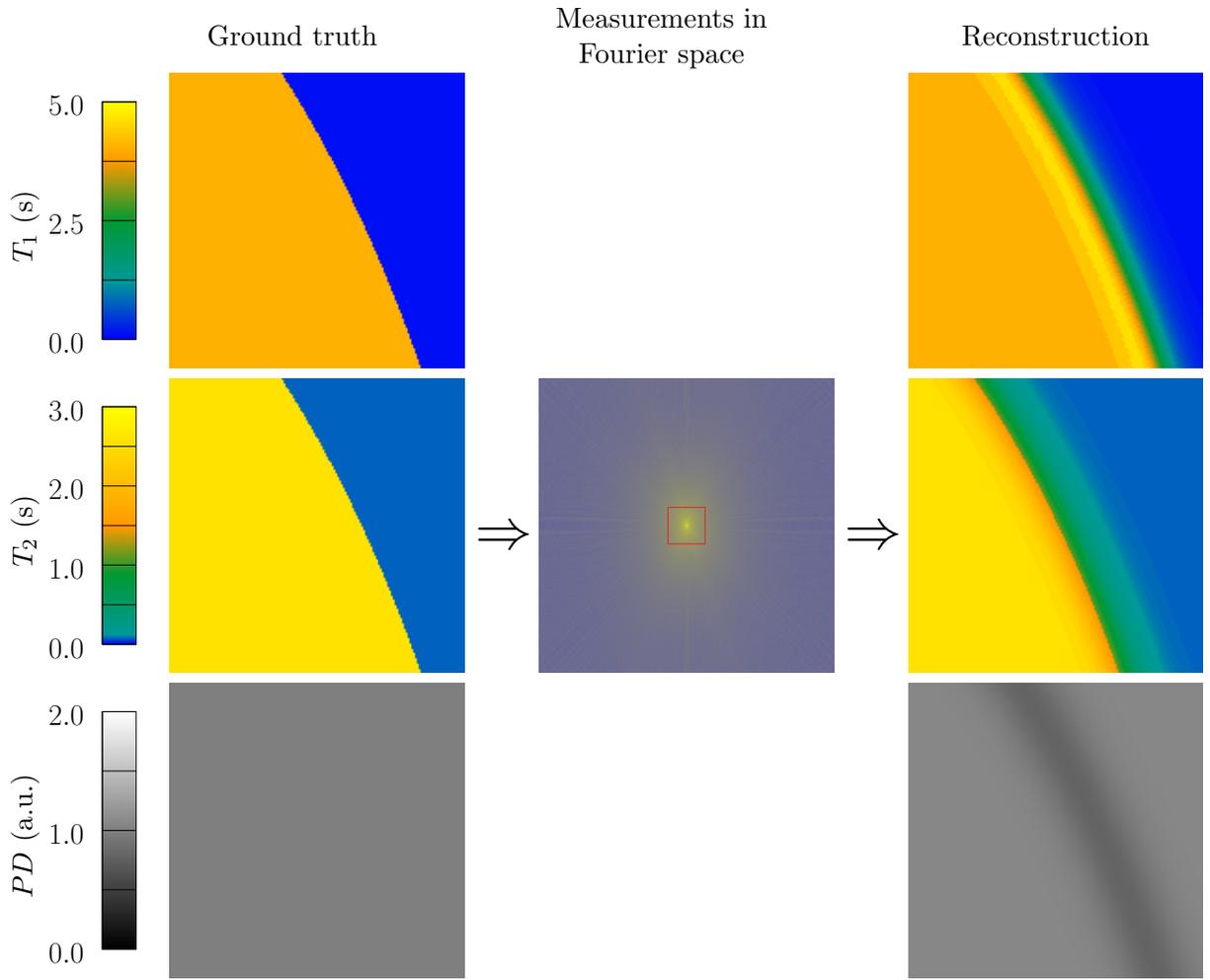


Figure 1: MRF reconstruction of a sharp boundary between two tissue regions with constant proton density (PD) and relaxation times corresponding to gray matter ( $T_1 = 1123$  ms,  $T_2 = 88$  ms) and cerebrospinal fluid ( $T_1 = 4200$  ms,  $T_2 = 2100$  ms) [29]. Due to the limited resolution caused by the low-pass measurements depicted by a red rectangle in the image at the center, voxels close to the boundary contain signals from the two tissues. This leads to erroneous parameter estimations near the boundary when using the standard MRF single-compartment reconstruction. A multicompartement model would be required to identify the correct relaxation times of the two contributing compartments.

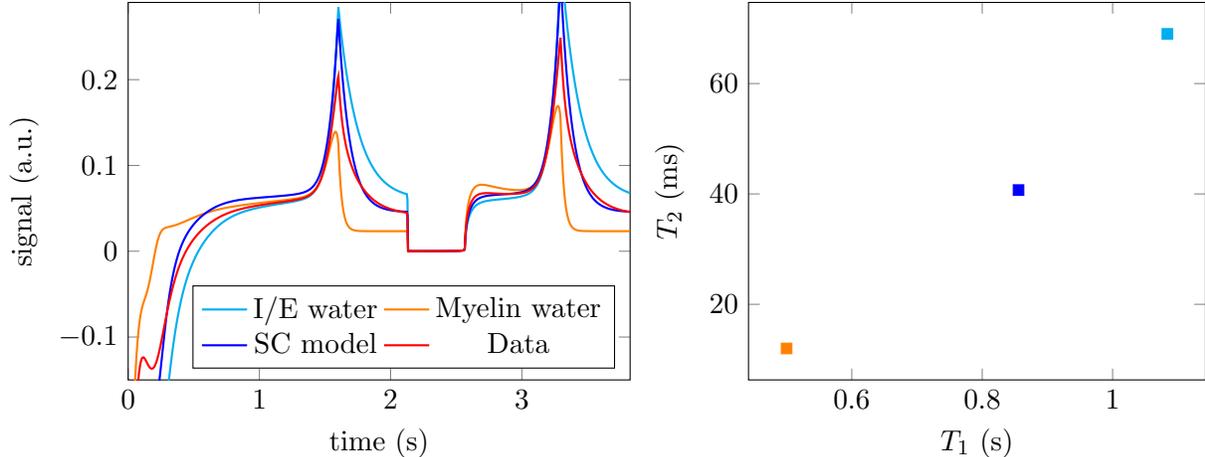


Figure 2: The left image shows the simulated signals corresponding to two different tissues: intra/extra (I/E) axonal water (light blue) and myelin water (orange). The signal in a voxel containing both tissues corresponds to the sum of both signals (red). A single-compartment (SC) model (blue) is not able to approximate the data (left) and results in an inaccurate estimate of the relaxation parameters  $T_1$  and  $T_2$  even in the absence of noise (right). The MRF dictionary is generated using the approach described in [3], which produces real-valued fingerprints.

assumption. The measured signal at a given voxel is modeled as the sum of the signals corresponding to the individual tissues present in the voxel, as in Ref. [54].

### 1.3 Related work

Previous works applying multicompartment estimation in the context of myelin-water imaging, applied bi-exponential fitting to data obtained from a multi-echo experiment [44,45]. Other works used non-convex optimization to fit a multicompartment model to measurements from a multi-steady-state experiment [24,25]. In the context of MR fingerprinting, the multicompartment problem has been tackled by exhaustive search in [39]. The work that lies closer to our proposed approach is [54], which also proposes to fit a multicompartment MRF model by solving sparse-recovery problems at each voxel. Ref. [54] reports that solving these sparse-recovery problems using a first-order method for  $\ell_1$ -norm minimization does not yield sparse estimates (which is consistent with our experiments, see Figure 6). They instead propose to use a reweighted-least-squares method based on a Bayesian framework to fit the multicompartment parameters to MRF data, which produces superior results, but still does not yield a sparse estimate (see Figure 9 for an example). As a result, the method tends to produce an estimate of a set of possible values for the relaxation times at each compartment, which must be post-processed to produce the parameter estimates. In contrast, we show reweighted- $\ell_1$ -methods tend to yield sparse solutions, which can be used to estimate the relaxation parameters directly (see Section 2.4 for more details).

Sparse linear models are a fundamental tool in statistics [33,41], signal processing [51] and machine learning [49], as well as across many applied domains [47,78]. Most works that apply these models to recover sparse signals focus on incoherent measurements [8, 14, 17, 20, 27, 28, 37, 55, 73]. As

explained in more detail in Section 2.4, this setting is not relevant to MRF where neighboring dictionary columns are highly correlated. Instead, the present work is more related to recent advances in optimization-based methods for sparse recovery from coherent dictionaries [7, 15, 30, 31, 50, 68, 69]. Related applied work includes the application of these techniques to source localization in electroencephalography [65, 80], analysis of positron-emission tomography data [38, 42] and radar imaging [61]. In addition, our work combines insights from recovery methods based on reweighting convex sparsity-inducing norms [19, 76] and techniques for distributed optimization [11].

## 1.4 Contributions

The main contribution of this work is a method for fitting multicompartment MRF models. These models represent the magnetization signal in each voxel as a linear combination of a small number of precomputed basis functions or *fingerprints*, which requires solving a sparse linear inverse problem. We combine the following insights to solve this challenging inverse problem both accurately and efficiently:

- Fitting the multicompartment model can be decoupled into multiple sparse-recovery problems—one for each voxel—using the alternating-direction of multipliers framework (see Section 2.2).
- The correlations between atoms in the dictionary make it possible to compress the dictionary to decrease the computational cost (see Section 2.3).
- Thresholding-based algorithms, typically used for sparse recovery and compressed sensing, do not perform well on the multicompartment MRF problem, due to the high correlation between dictionary atoms. However,  $\ell_1$ -norm minimization does succeed in achieving exact recovery in the absence of noise as long as the problem is solved using higher-precision second-order methods (see Section 2.4).
- The sparse-recovery approach can be made robust to noise and model imprecisions by using reweighted- $\ell_1$ -norm methods (see Section 2.5).
- These reweighted methods can be efficiently implemented using a second-order interior-point solver described in Section 2.6.

We validate our proposed method via numerical experiments on simulated data (see Section 3.1), as well as on experimental data from a custom-built phantom using a clinical MR system (see Section 3.2).

## 2 Methods

### 2.1 Multicompartment model

In this work, we consider the problem of estimating the relaxation times  $T_1$  and  $T_2$  via magnetic-resonance fingerprinting (MRF) [48] in situations where a voxel may contain several tissue compartments. As described briefly in Section 1.1,  $T_1$  and  $T_2$  determine the time evolution of the measured

nuclear-spin magnetization signal. MR systems measure the magnetization component that is perpendicular to the external field, which is usually assumed to be aligned with the  $z$  axis [57]. For a voxel which contains only one tissue, the time evolution of this component can be approximated as a function of the values of  $T_1$  and  $T_2$

$$M_x(t) + iM_y(t) = \phi(t; T_1, T_2), \quad (2.1)$$

where the two-dimensional vector is represented as a complex number following the usual convention in the MRI literature. The main insight underlying MRF is that even though the mapping  $\phi$  cannot be computed explicitly, it can be *evaluated numerically* by solving the Bloch equations (1.1). This makes it possible to build a dictionary of possible time evolutions or fingerprints for a discretized set of values of  $T_1$  and  $T_2$ . We denote such a dictionary by  $D \in \mathbb{C}^{n \times m}$

$$D_{kl} := \phi(t_k; \tau_1(l), \tau_2(l)), \quad (2.2)$$

where  $t_1, t_2, \dots, t_n$  are the times at which the magnetization signal is sampled and  $\{\tau_1(1), \tau_2(1)\}, \{\tau_1(2), \tau_2(2)\}, \dots, \{\tau_1(m), \tau_2(m)\}$  are the discretized  $T_1$  and  $T_2$  values respectively (there are  $m$  fingerprints in total). Small variations of  $T_1$  and  $T_2$  in Eq. (2.1) result in small changes in the magnetization signal. As a result, neighboring columns of  $D$  (i.e. fingerprints corresponding to similar  $T_1$  and  $T_2$  values) are highly correlated.

When several tissues are present in a single voxel, we assume that the different magnetization components combine additively (see the discussion at the end of Section 1.2). In that case, the magnetization vector  $x^{[j]} \in \mathbb{C}^n$  measured at voxel  $j$  is given by

$$x_k^{[j]} = \sum_{s=1}^S \rho(\text{tissue } s) \phi(t_k; T_1(\text{tissue } s), T_2(\text{tissue } s)), \quad 1 \leq k \leq n, \quad (2.3)$$

where  $\rho(\text{tissue } s)$  denotes the proton density of tissue  $s$  and  $S$  the number of tissues in the voxel. If the values of the relaxation times in each voxel are present in the dictionary, the magnetization vector can be expressed as

$$x^{[j]} = Dc^{[j]}, \quad 1 \leq j \leq N, \quad (2.4)$$

where  $N$  is the number of voxels in the volume of interest. The vectors of coefficients  $c^{[1]}, c^{[2]}, \dots, c^{[N]}$  are assumed to be sparse and nonnegative: each nonzero entry in  $c^{[j]}$  corresponds to the proton density of a tissue that is present in voxel  $j$ . The nonnegativity assumption implies that the magnetization signal from all compartments has the same complex phase, which is unwound prior to the fitting process. This neglects the effect of Gibbs ringing, due to the truncation of MR measurements in the frequency space (the lobes of the convolution kernel corresponding to the truncation may be negative). However, this issue can be mitigated by applying appropriate filtering, which makes the nonnegativity assumption reasonable in most cases.

Raw MRI data sampled at a given time do not directly correspond to the magnetization signal of the different voxels, but rather to samples from the spatial Fourier transform of the magnetization over all voxels in the volume of interest [46], a representation known as *k-space* in the MRI literature [74]. Staying within clinically-feasible scan times in MRF requires subsampling the image k-space and therefore violating the Nyquist-Shannon sampling theorem [48]. The k-space is typically subsampled along different trajectories, governed by changes in the magnetic-field gradients

used in the measurement process. We define a linear operator  $F_{t_i} \in \mathbb{C}^{d \times n}$  that represents the linear operator mapping the magnetization at time  $t_i$  to the measured data  $y_{[t_i]} \in \mathbb{C}^d$ , where  $d$  is the number of k-space samples at  $t_i$ . This operator encodes the chosen subsampling pattern in the frequency domain. The model for the data is consequently of the form

$$y_{t_i} = F_{t_i} \begin{bmatrix} x_{t_i}^{[1]} \\ x_{t_i}^{[2]} \\ \vdots \\ x_{t_i}^{[N]} \end{bmatrix}, \quad 1 \leq i \leq n, \quad (2.5)$$

where we have ignored noise and model inaccuracies. If the data are acquired using multiple receive coils, the measurements can be modeled as samples from the spectrum of the pointwise product between the magnetization and the sensitivity function of the different coils [62, 66]. In that case the operator  $F_{t_i}$  also includes the sensitivity functions and  $d$  is the number of k-space samples multiplied by the number of coils.

## 2.2 Parameter Map Reconstruction via Alternating Minimization

For ease of notation, we define the magnetization and coefficient matrices

$$X := [x^{[1]}, x^{[2]}, \dots, x^{[N]}], \quad (2.6)$$

$$C := [c^{[1]}, c^{[2]}, \dots, c^{[N]}], \quad (2.7)$$

so that we can write (2.4) as

$$X = DC. \quad (2.8)$$

In addition we define the data matrix  $Y \in \mathbb{C}^{d \times n}$

$$Y := [y_{t_1}, y_{t_2}, \dots, y_{t_n}] \quad (2.9)$$

and a linear operator  $\mathcal{F}$  such that

$$Y = \mathcal{F}X. \quad (2.10)$$

Note that by (2.5) the  $i$ th column of  $Y$  is the result of applying  $F_{t_i}$  to the  $i$ th row of  $X$ .

In principle, we could fit the coefficient matrix by computing a sparse estimate such that  $Y \approx \mathcal{F}DC$ . Unfortunately, this would require solving a sparse-regression problem of intractable dimensions. Instead we incorporate a variable  $\tilde{X}$  to represent the magnetization and formulate the model-fitting problem as

$$\underset{\tilde{X}, \tilde{C}}{\text{minimize}} \quad \left\| Y - \mathcal{F}\tilde{X} \right\|_F^2 + \mathcal{R}(\tilde{C}) \quad (2.11)$$

$$\text{subject to} \quad \tilde{X} = D\tilde{C}, \quad (2.12)$$

$$\tilde{C}_{ij} \geq 0, \quad 1 \leq i \leq m, 1 \leq j \leq N, \quad (2.13)$$

where  $\mathcal{R}$  is a regularization term to promote sparse solutions. In order to alternate between updating the magnetization and coefficient variables, we follow the framework of the alternative-direction method of multipliers (ADMM) [11]. Consider the augmented Lagrangian with respect to the constraint (2.12),

$$\mathcal{L}_\mu(\tilde{X}, \tilde{C}, \Lambda) := \left\| Y - \mathcal{F}\tilde{X} \right\|_F^2 + \mathcal{R}(\tilde{C}) + \langle \Lambda, \tilde{X} - D\tilde{C} \rangle + \frac{\mu}{2} \left\| \tilde{X} - D\tilde{C} \right\|_F^2, \quad (2.14)$$

where the constant  $\mu > 0$  is a parameter and the dual variable  $\Lambda$  is an  $N \times n$  matrix. ADMM alternates between minimizing the augmented Lagrangian with respect to each primal variable sequentially and updating the dual variable. We now describe each of the updates in more detail. We denote the values of the different variables at iteration  $l$  by  $C^{(l)}$ ,  $X^{(l)}$  and  $\Lambda^{(l)}$ . More details on the implementation can be found in Ref. [2].

### Updating $\tilde{X}$

If we fix  $\tilde{C}$  and  $\Lambda$ , minimizing  $\mathcal{L}_\mu$  over  $\tilde{X}$  is equivalent to solving the least-squares problem

$$X^{(l+1)} := \arg \min_{\tilde{X}} \left\| Y - F\tilde{X} \right\|_F^2 + \frac{\mu}{2} \left\| \tilde{X} - DC^{(l)} - \frac{1}{\mu}\Lambda^{(l)} \right\|_F^2. \quad (2.15)$$

This step amounts to estimating a magnetization matrix that fits the data while being close to the magnetization corresponding to the current estimate of dictionary coefficients. The optimization problem has a closed-form solution, but due to the size of the matrices it is more efficient to solve it using an iterative algorithm such as the conjugate-gradients method [64].

### Updating $\tilde{C}$

If we fix  $\tilde{X}$  and  $\Lambda$ , minimizing  $\mathcal{L}_\mu$  over  $\tilde{C}$  decouples into  $N$  subproblems. In more detail, the goal is to compute a sparse nonnegative vector of coefficients  $\tilde{c}^{[j]}$  for each voxel  $j$ ,  $1 \leq j \leq N$ , such that

$$\left( X^{(l)} - \frac{1}{\mu}\Lambda^{(l)} \right)_j \approx D\tilde{c}^{[j]}, \quad (2.16)$$

where the left hand side corresponds to the  $j$ th column of  $X^{(l)} - \frac{1}{\mu}\Lambda^{(l)}$ . This collection of sparse-recovery problems is very challenging to solve due to the correlations between the columns of  $D$ . In Sections 2.4, 2.5 and 2.6 we present an algorithm to tackle them efficiently.

### Updating $\Lambda$

The dual variable is updated by setting

$$\Lambda^{(l+1)} := \Lambda^{(l)} + \mu \left( X^{(l)} - \hat{D}C^{(l)} \right). \quad (2.17)$$

We refer to [11] for a justification based on dual-ascent methods.

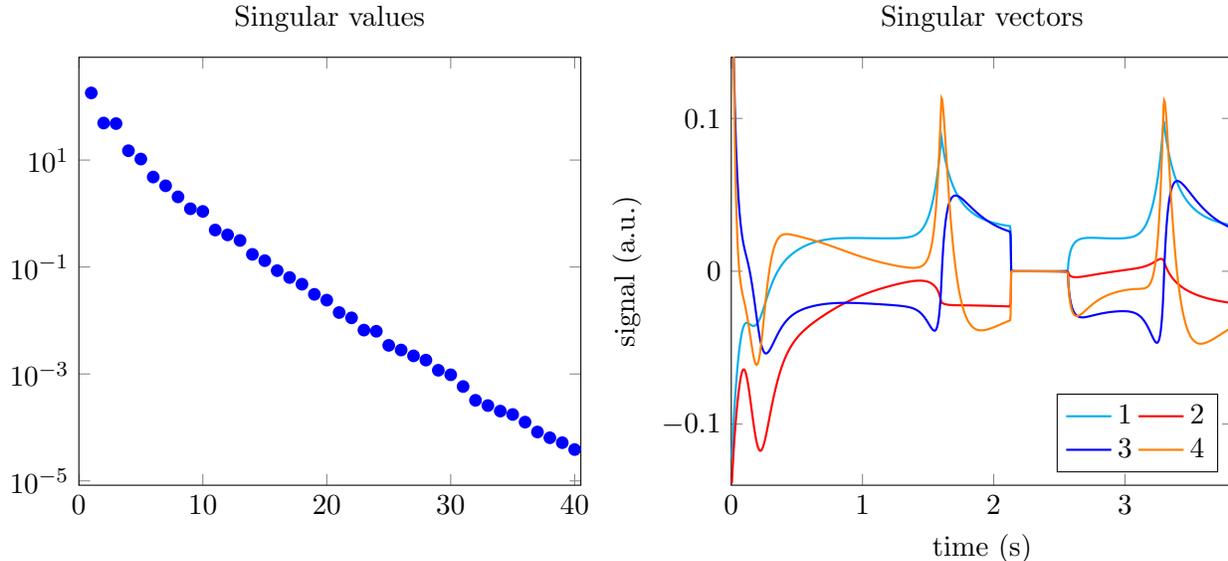


Figure 3: Singular values (left) and corresponding left singular vectors (right) of an MRF dictionary generated using the approach in [3, 4].

### 2.3 Dictionary compression

As explained in Section 2.1 nearby columns in the dictionary of fingerprints  $D$  are highly correlated. In fact, fingerprint dictionaries tend to be approximately low rank [2, 54], which means that their singular values present a rapid decay (see Figure 3 for an example). Consider the singular value decomposition

$$D = U\Sigma V^*, \quad (2.18)$$

where the superscript  $*$  denotes conjugate transpose. If  $D$  is approximately low rank, then its column space is well approximated as the span of the first few left singular vectors (Figure 3 shows some of these singular vectors for a concrete example). Following [2, 54], we exploit this low-rank structure to compress the dictionary. Let  $U_k \Sigma_k V_k^*$  denote the rank- $k$  truncated SVD of  $D$ , where  $U_k$  contains the first  $k$  left singular vectors and  $\Sigma_k$  and  $V_k$  the corresponding singular values and right singular vectors. If the measured data follow the linear model  $Y \approx \mathcal{F}DC$ , then

$$Y \approx \mathcal{F}U_k \Sigma_k V_k^* C \quad (2.19)$$

$$= \bar{\mathcal{F}} \bar{D} C, \quad (2.20)$$

where  $\bar{\mathcal{F}} := \mathcal{F}U_k$  is a modified linear operator and

$$\bar{D} := \Sigma_k V_k^* \quad (2.21)$$

$$= U_k^* D. \quad (2.22)$$

$\bar{D}$  can be interpreted as a compressed dictionary with dimensions  $k \times m$  obtained by projecting each fingerprint onto the  $k$ -dimensional subspace spanned by the columns of  $U_k$ , which are depicted

on the right image of Figure 3. In our experiments,  $D$  is well approximated by  $U_k \Sigma_k V_k^*$  for values of  $k$  between 10 and 15. As a result, replacing  $\mathcal{F}$  by  $\overline{\mathcal{F}}$  and  $D$  by  $\overline{D}$  within the ADMM framework in Section 2.2 dramatically decreases the computational cost.

## 2.4 Sparse estimation via $\ell_1$ -norm minimization

In this section, we consider the problem of estimating the tissue parameters at a fixed voxel from the time evolution of its magnetization signal. As described in Section 2.2, this is a crucial step in fitting the MRF multicompartment model. To simplify the notation, we denote the discretized magnetization signal at an arbitrary voxel by  $x \in \mathbb{C}^n$  (where  $n$  is the number of time samples) and the sparse nonnegative vector of coefficients by  $c \in \mathbb{R}^m$ . In Eq. (2.16),  $x$  represents the column of  $X^{(l)} - \frac{1}{\mu} \Lambda^{(l)}$  corresponding to the voxel at a particular iteration  $l$  of the alternating scheme.

Let us first consider the sparse recovery problem assuming that the magnetization estimate is exact. In that case there exists a sparse nonnegative vector  $c$  such that

$$x = Dc. \tag{2.23}$$

Even in this simplified scenario, computing  $c$  from  $x$  is very challenging. The dictionary is overcomplete: there are more columns than time samples because the number of fingerprints  $m$  is larger than the number of time samples  $n$ . As a result the linear system is underdetermined: there are infinite possible solutions. However, we are not interested in arbitrary solutions, but rather in sparse nonnegative solutions, where  $c$  contains a small number of nonzero entries corresponding to the tissues present in the voxel. We would like to emphasize that the level of sparsity (i.e. the number of nonzeros in  $c$ ) is not assumed to be known. As a result, the number of estimated nonzero coefficients produces an estimate for the number of compartments in the voxel.

There is a vast literature on the recovery of sparse signals from underdetermined linear measurements. Two popular techniques are greedy methods that select columns from the dictionary sequentially [52, 60] and optimization-based approaches that minimize a sparsity-promoting cost function such as the  $\ell_1$  norm [16, 21, 27]. Theoretical analysis of these methods has established that they achieve exact recovery and are robust to additive noise if the overcomplete dictionary is *incoherent*, meaning that the correlation between the columns in the dictionary is low [8, 14, 17, 27, 28, 37, 55, 73]. Some of these works assume stronger notions of incoherence such as the restricted-isometry condition [17] or the restricted-eigenvalue condition [8].

Recall that the mapping (2.1) between the magnetization signal and the parameters of interest is smooth. As a result, when the parameter space is discretized finely to construct the dictionary, the correlation between fingerprints corresponding to similar parameters is very high. This is illustrated by Figure 4, which shows the correlations between columns for a simple MRF dictionary simulated following [3, 4]. For ease of visualization, the only parameter that varies in the dictionary is  $T_1$  (the value of  $T_2$  is fixed to 62 ms), so that the correlation in the depicted dictionary has a very simple structure: neighboring columns correspond to similar values of  $T_1$  and are, therefore, more correlated, whereas well-separated columns are less correlated. For comparison, the figure also shows the correlations between columns for a typical incoherent compressed-sensing dictionary generated by sampling each entry independently from a Gaussian distribution [18, 27]. In contrast, to the compressed-sensing dictionary, MRF dictionaries are highly coherent, which means that the aforementioned theoretical guarantees for sparse recovery do not apply.

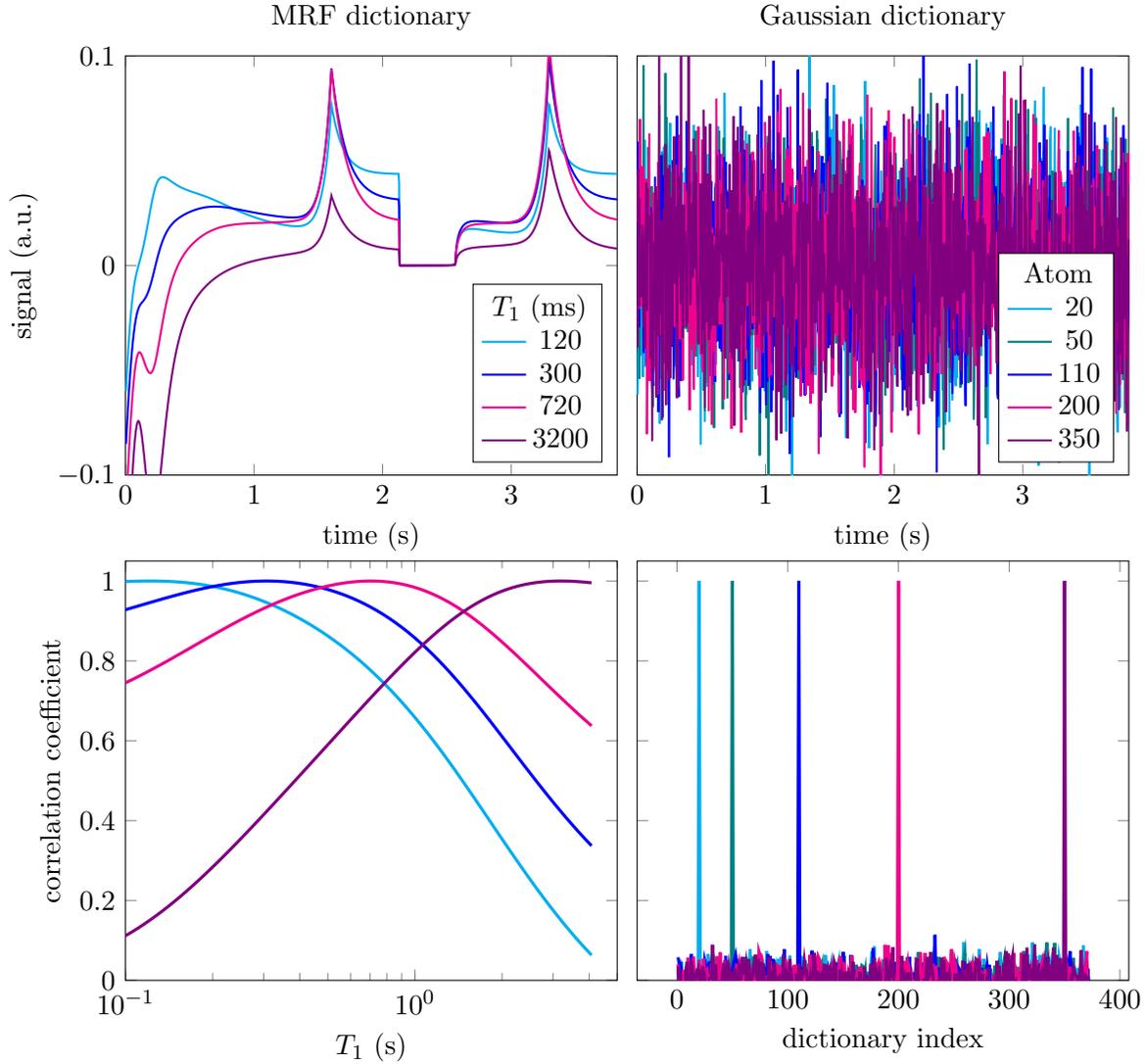


Figure 4: The figure shows a small selection of columns in a magnetic-resonance fingerprinting (MRF) dictionary (top left) and an i.i.d. Gaussian compressed-sensing dictionary (top right). In the MRF dictionary, generated using the approach in [3, 4] which produces real-valued fingerprints, each fingerprint corresponds to a value of the relaxation parameter  $T_1$  ( $T_2$  is fixed to 62 ms). In the bottom row we show the correlation of the selected columns with every other column in the dictionary. This reveals the high local coherence of the MRF dictionary (left), compared to the incoherence of the compressed-sensing dictionary (right).

Coherence is not only a problem from a theoretical point of view. Most fast-estimation methods for sparse recovery exploit the fact that for incoherent dictionaries like the Gaussian dictionary in Figure 4 the Gram matrix  $D^*D$  is close to the identity. As a result, multiplying the data by the transpose of the dictionary yields a noisy approximation to the sparse coefficients

$$D^*y = D^*Dc \approx c, \quad (2.24)$$

which can be *cleaned up* using some form of thresholding. This concept is utilized by greedy approaches such as orthogonal matching pursuit [60] or iterative hard-thresholding methods [10], as well as fast algorithms for  $\ell_1$ -norm minimization such as proximal methods [22] and coordinate descent [34], which use iterative soft-thresholding instead. In our experiments with MRF dictionaries, for which  $D^*y$  is not approximately sparse due to the correlations between columns (see the second row of Figure 5), these techniques fail to produce accurate estimates. As an example, Figure 5 shows the result of applying iterative hard thresholding (IHT) [10] to an MRF dictionary. The estimated coefficients are sparse, but closer to the maximum of  $D^*y$  than to the true values. In contrast, the method is very effective when applied in a compressed-sensing setting, where the dictionary is incoherent (see right column of Figure 5).

Recently, theoretical guarantees for sparse-decomposition methods have been established for dictionaries arising in super-resolution [15,31] and deconvolution problems [7]. As in the case of MRF, these dictionaries have very high local correlations between columns. These papers show that although robust recovery of *all sparse* signals in such dictionaries is not possible due to their high coherence,  $\ell_1$ -norm minimization achieves exact recovery of a more restricted class of signals: sparse signals with nonzero entries corresponding to dictionary columns that are not highly correlated. To be clear, arbitrarily high local correlations may exist in the dictionary, as long the columns that are actually present in the measured signal are relatively uncorrelated with each other. Although this theory does not apply directly to MRF, it suggests that recovery may be possible for multicompartment models where the fingerprints of the tissues present in each voxel are sufficiently different. Indeed, our numerical results confirm this intuition, as illustrated by Figure 5 where  $\ell_1$ -norm minimization achieves exact recovery also for the coherent MRF dictionary. Solving the convex program

$$\underset{\tilde{c} \in \mathbb{R}^m}{\text{minimize}} \quad \|\tilde{c}\|_1 \quad (2.25)$$

$$\text{subject to} \quad D\tilde{c} = x \quad (2.26)$$

$$\tilde{c} \geq 0. \quad (2.27)$$

makes it possible to successfully fit the sparse multicompartment model, as long as the parameters of each compartment are not too close (see Figure 12). Interestingly, in our experiments we observe that for these dictionaries it is often necessary to use high-precision second-order methods [36] in order to achieve exact recovery. In contrast, fast  $\ell_1$ -norm minimization algorithms based on soft-thresholding or coordinate descent tend to converge very slowly due to the coherence of the dictionary, even for small-scale problems (see Figure 6).

## 2.5 Robust sparse estimation via reweighted- $\ell_1$ -norm minimization

In the previous section, we discuss exact recovery of the sparse coefficients in a multicompartment model under the assumptions that (1) there is no noise and (2) the fingerprints corresponding to the

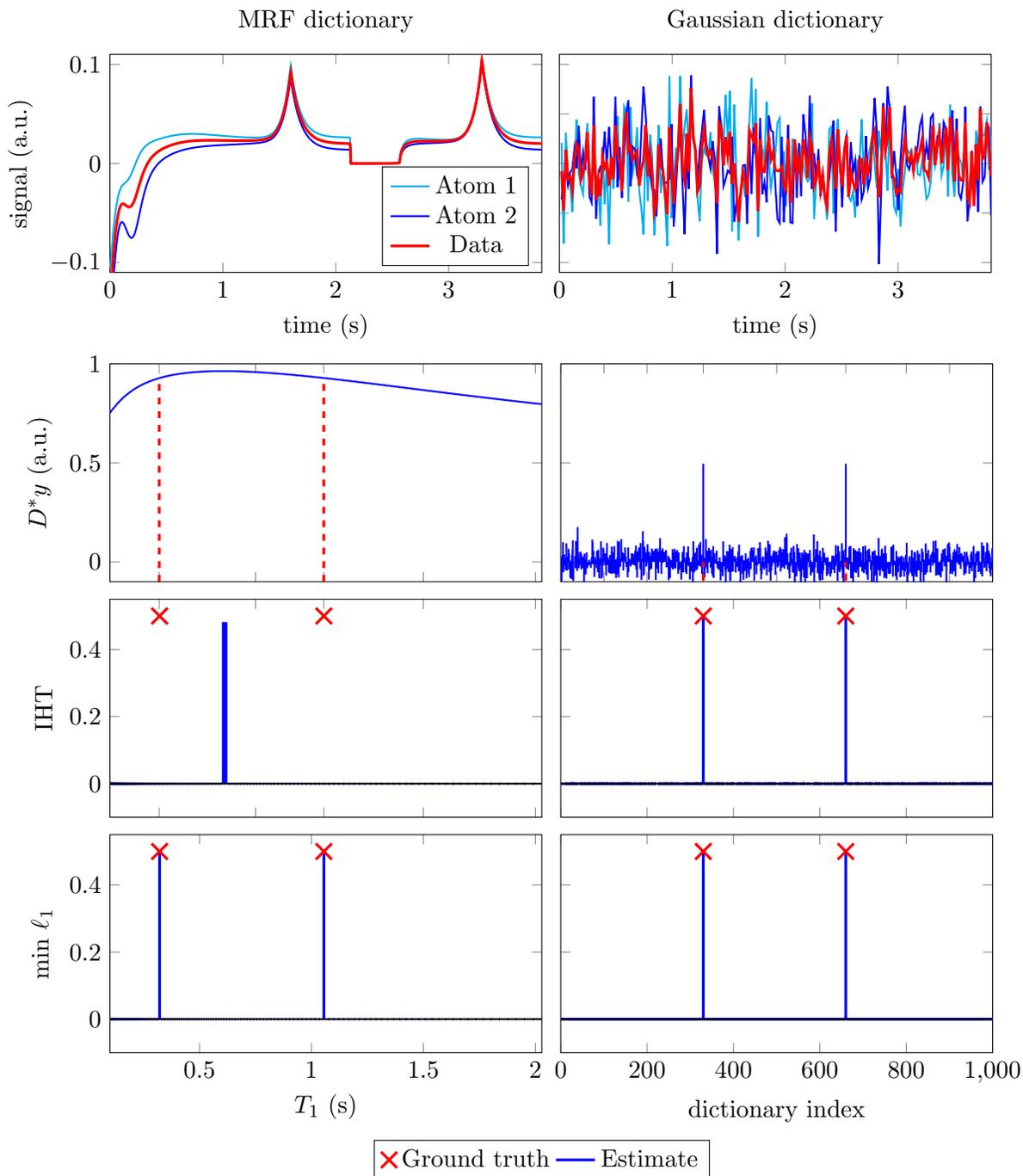


Figure 5: The data in the top row are an additive superposition of two columns from the MRF (left) and the compressed-sensing (right) dictionaries depicted in Figure 4. The correlation between the data and each dictionary column (second row) is much more informative for the incoherent dictionary (right) than for the coherent one (left). Iterative hard thresholding (IHT) [10] (third row) achieves exact recovery of the true support for the compressed-sensing problem (right), but fails for MRF (left) where it recovers two elements that are closer to the maximum of the correlation than to the true values. In contrast, solving (2.27) using a high-precision solver [36] achieves exact recovery in both cases (bottom row).

true parameters are present in the dictionary. In practice, neither of these assumptions holds: noise is unavoidable and the tissue parameters may not fall exactly on the chosen grid. In fact, the vector  $x$  represents a column in  $X^{(l)} - \frac{1}{\mu}\Lambda^{(l)}$  at iteration  $l$  of the alternating scheme described in Section 2.2 and is therefore also subject to aliasing artifacts. A standard way to account for perturbations and model imprecisions in sparse-recovery problems is to relax the  $\ell_1$ -norm minimization problem (2.27) to a regularized least-squares problem of the form

$$\begin{aligned} & \underset{\tilde{c} \in \mathbb{R}^m}{\text{minimize}} && \frac{1}{2} \|D\tilde{c} - x\|^2 + \lambda \|\tilde{c}\|_1 \\ & \text{subject to} && \tilde{c} \geq 0 \end{aligned} \tag{2.28}$$

where  $\lambda \geq 0$  is a regularization parameter that governs the trade-off between the least-squares data-consistency term and the  $\ell_1$ -norm regularization term. Solving (2.28) to perform sparse recovery is a popular sparse-regression method known as the lasso in statistics [70] and basis pursuit in signal processing [21].

The top row of Figure 6 shows the results of estimating the components of a noisy MRF signal by solving problem (2.28) using a high-precision interior-point method for different values of the parameter  $\lambda$ . For large values of  $\lambda$ , the solution is sparse, but the fit to the data is not very accurate and the true support is not well approximated. For small values of  $\lambda$ , the solution to the convex program correctly locates the vicinity of the true coefficients, but is contaminated by small spurious spikes. Perhaps surprisingly, a popular first-order method based on soft-thresholding called FISTA [5] produces solutions that are quite dense, a phenomenon previously reported in [54]. This suggests that higher-precision solvers may be needed to fit sparse linear models for coherent dictionaries, as opposed to incoherent dictionaries for which low-precision methods such as FISTA are very effective.

Reweighted- $\ell_1$ -methods [19, 76] are designed to enhance the performance of  $\ell_1$ -norm-regularized problems by promoting sparser solutions that are close to the initial estimate. In the case of the MRF dictionary, these methods are able to remove spurious small spikes, while retaining an accurate support estimate. This is achieved by solving a sequence of weighted  $\ell_1$ -norm regularized problems of the form

$$\begin{aligned} & \underset{\tilde{c} \in \mathbb{R}^m}{\text{minimize}} && \frac{1}{2} \|D\tilde{c} - x\|^2 + \lambda \sum_{i=1}^m \omega_i^{(k)} \tilde{c}_i \\ & \text{subject to} && \tilde{c} \geq 0 \end{aligned} \tag{2.29}$$

for  $k = 1, 2, \dots$ . We denote the solution of this optimization problem by  $\hat{c}^{(k)}$ . The entries in the initial vector of weights  $\omega^{(1)}$  are initialized to one, so  $\hat{c}^{(1)}$  is the solution to problem (2.28). The weights are then updated depending on the subsequent solutions to problem (2.29).

We consider two different reweighting schemes. The first, proposed in [19], simply sets the weights to be inversely proportional to the previous estimate  $\hat{c}^{(k-1)}$

$$\omega_i^{(k)} := \frac{1}{\epsilon + |\hat{c}^{(k-1)}|}, \quad 1 \leq i \leq m, \tag{2.30}$$

where  $\epsilon$  is a parameter typically fixed to a very small value with respect to the expected magnitude of the nonzero coefficients (e.g.  $10^{-8}$ ) in order to preclude division by zero (see [19] for more details).

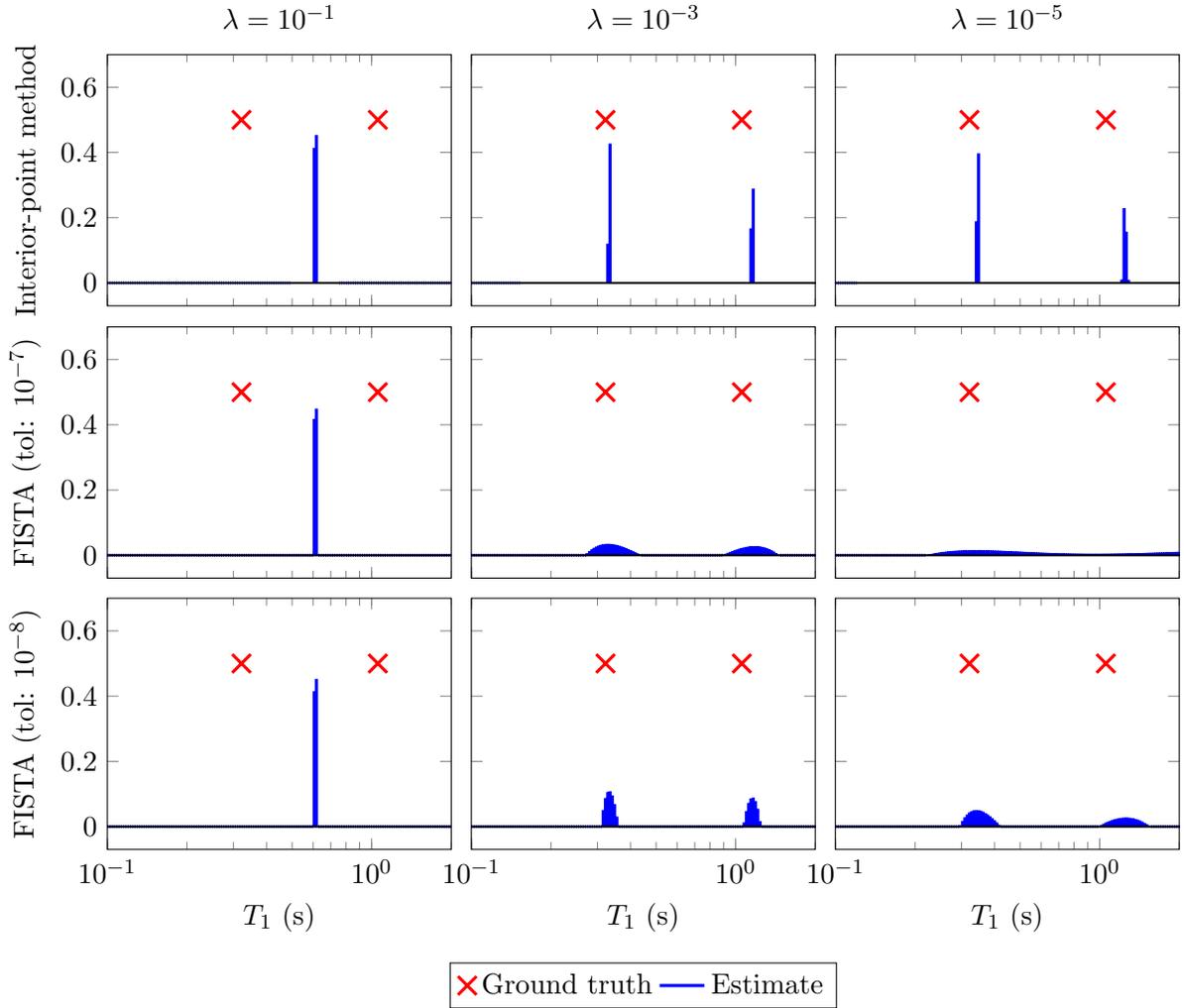


Figure 6: Estimates for the components of an MRF signal obtained by solving problem (2.28) for different values of the parameter  $\lambda$  using an interior point solver [36] and an implementation of FISTA [5] from the TFOCS solver [6]. The data are generated by adding i.i.d. Gaussian noise to the two-compartment MRF signal shown in Figure 5 (top left). The signal-to-noise ratio is equal to 100 (defined as the ratio between the  $\ell_2$  norm of the signal and the noise) or equivalently 40 dB. FISTA is terminated at an iteration  $k$  where the relative change in the coefficients  $\|c^{(k)} - c^{(k-1)}\| / \max(1, \|c^{(k)}\|)$  falls below the chosen tolerance. It takes between 5 and 10 minutes for a single instance with a tolerance of  $10^{-8}$  and 20-40 seconds with a tolerance of  $10^{-7}$ . CVX is run with the default precision and requires 2 seconds per instance.

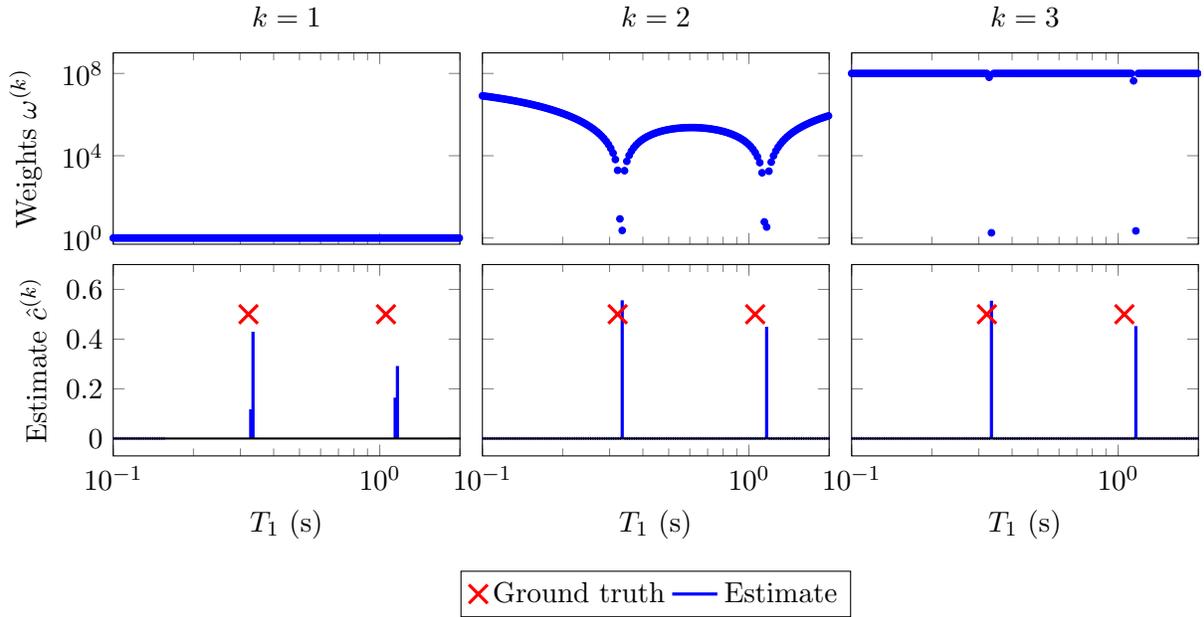


Figure 7: Solutions to problem (2.29) (bottom) and corresponding weights (top) computed following the reweighting scheme (2.30), with  $\lambda := 10^{-3}$  and  $\epsilon := 10^{-8}$ . The data are the same as in Figure 6.

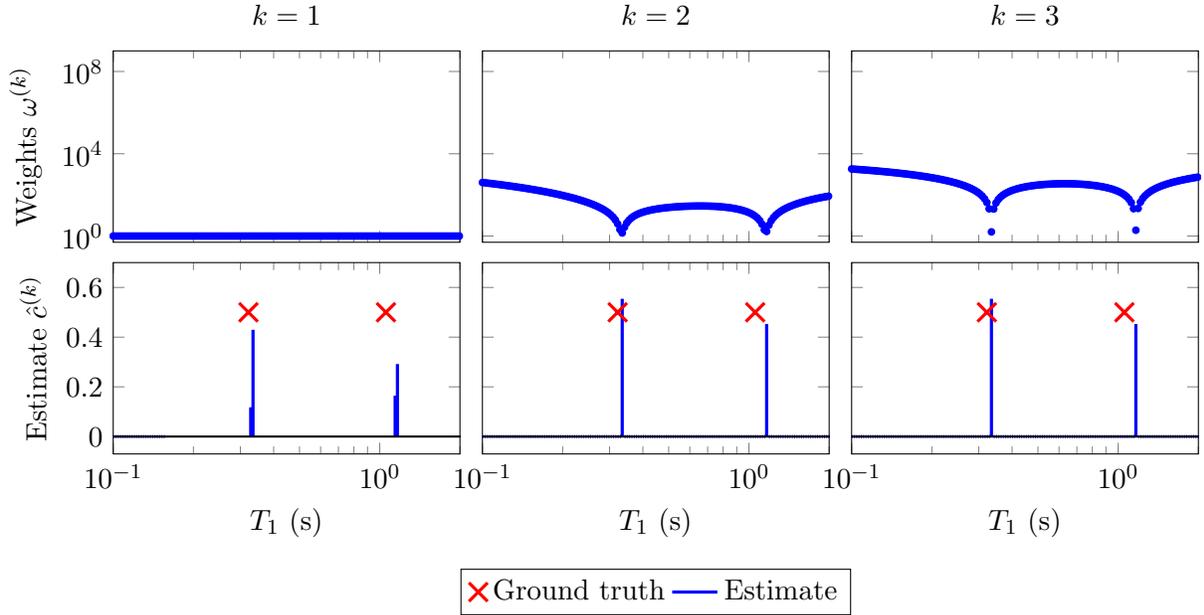


Figure 8: Solutions to problem (2.29) (bottom) and corresponding weights (top) computed following the reweighting scheme (2.31), with  $\lambda := 10^{-3}$  and  $\epsilon := 10^{-8}$ . The data are the same as in Figure 6.

Intuitively, the weights penalize the coefficients that do not contribute to the fit in previous iterations, promoting increasingly sparse solutions. Mathematically, the algorithm can be interpreted as a majorization-minimization method for a non-convex sparsity-inducing penalty function [19]. The second updating scheme, proposed in [76] and inspired by sparse Bayesian learning methods [77], is of the form

$$\omega_i^{(k)} := \left( D_i^* (\epsilon I + DCW^{-1}D^*)^{-1} D_i \right)^{\frac{1}{2}}, \quad (2.31)$$

$$C := \text{diag} \left( \hat{c}^{(k-1)} \right), \quad (2.32)$$

$$W := \text{diag} \left( \omega^{(k-1)} \right). \quad (2.33)$$

The parameter  $\epsilon$  precludes the matrix  $\epsilon I + DCW^{-1}D^*$  from being singular and can be set to a very small fixed value (e.g.  $10^{-8}$ ). The reweighting scheme sets  $\omega_i^{(k)}$  to be small if there is any large entry  $\hat{c}_j^{(k-1)}$  such that the corresponding columns of the dictionary  $D_i$  and  $D_j$  are correlated. This produces smoother weights than (2.30) providing more robustness to initial errors in the support estimate.

Figures 7 and 8 show the results obtained by applying reweighted- $\ell_1$ -norm regularization to the same problem as in Figure 6. As expected, the update (2.31) yields smoother weights. Convergence to a sparse solution is achieved in just two iterations. The estimate is an accurate estimate of the parameters despite the presence of noise and the gridding error (the true parameters do not lie on the grid used to construct the dictionary).

Finally, we would like to mention reweighted least-squares, a sparse-recovery method applied to an MRF multicompartment model in [54]. Similarly to the reweighted- $\ell_1$  norm method given by (2.31), this technique is based on a nonconvex cost function derived by Bayesian principles. The resulting scheme is not as accurate as the reweighting  $\ell_1$ -norm based methods (see Figure 9) and often yields solutions that are not sparse, as reported in [54]. However, it is extremely fast, since it just requires solving a sequence of regularized least-squares problems. Unfortunately, this also means that the nonnegativity assumption on the coefficients cannot be easily incorporated (the intermediate problems then become nonnegative least-squares problems which no longer have closed-form solutions). Developing fast methods that incorporate such constraints and are effective for coherent dictionaries is an interesting topic for future research. In this work, our method of choice for the sparse-recovery problem arising in multicompartment MRF is least squares with reweighted- $\ell_1$ -norm regularization, implemented using an efficient interior-point solver described in the following section.

## 2.6 Fast interior-point solver for $\ell_1$ -norm regularization in coherent dictionaries

In this section we describe an interior-point solver for the  $\ell_1$ -norm regularized problem

$$\begin{aligned} \underset{\tilde{c} \in \mathbb{R}^m}{\text{minimize}} \quad & \frac{1}{2} \|D\tilde{c} - x\|^2 + \lambda \sum_{i=1}^m \tilde{c}_i \\ \text{subject to} \quad & \tilde{c} \geq 0, \end{aligned} \quad (2.34)$$

which allows us to apply the reweighting schemes described in Section 2.5 efficiently. Interior-point methods enforce inequality constraints by using a barrier function [12, 58, 79]. In this case we use

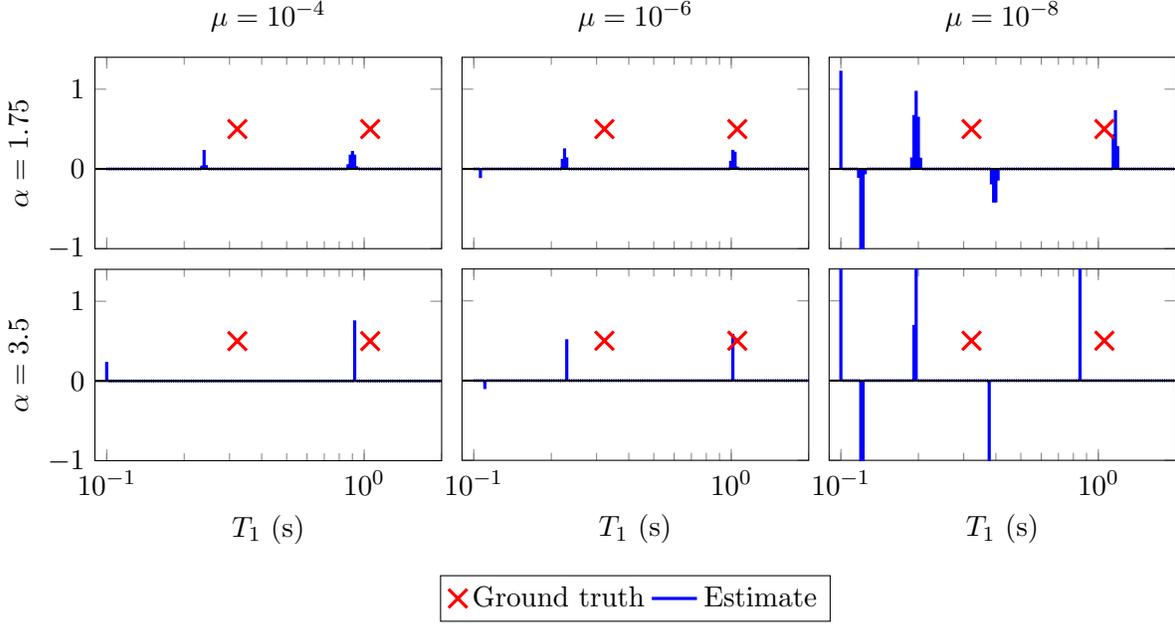


Figure 9: Results of applying the method in [54] to the data used in Figure 6. The underlying gamma distribution is parameterized with  $\alpha \in \{1.75, 3.5\}$  and  $\beta = 0.01$  as recommended by the authors. Reconstruction is shown for a range of values of an additional parameter ( $\mu$ ).

a logarithmic function that forces the coefficients to be nonnegative. Parametrizing the modified optimization problem with  $t > 0$ , we obtain a sequence of cost functions of the form

$$\phi_t(\tilde{c}) := \frac{1}{2}t\|D\tilde{c} - x\|^2 + t\lambda \sum_{i=1}^n \tilde{c}_i - \sum_{i=1}^n \log \tilde{c}_i. \quad (2.35)$$

The *central path* of the interior-point scheme consists of the minimizers of  $\phi_t$  as  $t$  varies from 0 to  $\infty$ . To find a solution for problem 2.34, we find a sequence of points  $c^{(1)}, c^{(2)}, \dots$  in the central path by iteratively solving the Newton system

$$(Z + 2tD^*D) \left( c^{(k)} - c^{(k-1)} \right) = -\nabla \phi_t \left( c^{(k-1)} \right) \quad (2.36)$$

where

$$Z := \text{diag} \left( \frac{1}{\tilde{c}_1^2}, \dots, \frac{1}{\tilde{c}_m^2} \right). \quad (2.37)$$

This is essentially the approach taken in [13, 43] to tackle  $\ell_1$ -norm regularized least squares. Solving (2.36) is the computational bottleneck of this method. In the case of the MRF dictionary,  $D^*D$  is a matrix of rank  $k$ , where  $k$  is the order of the low-rank approximation to the dictionary described in Section 2.3. As a result, solving the Newton system (2.36) directly is extremely slow for large values of  $t$ . Ref. [43] suggests applying a preconditioner of the form

$$P := 2t \text{diag}(D^*D) + Z \quad (2.38)$$

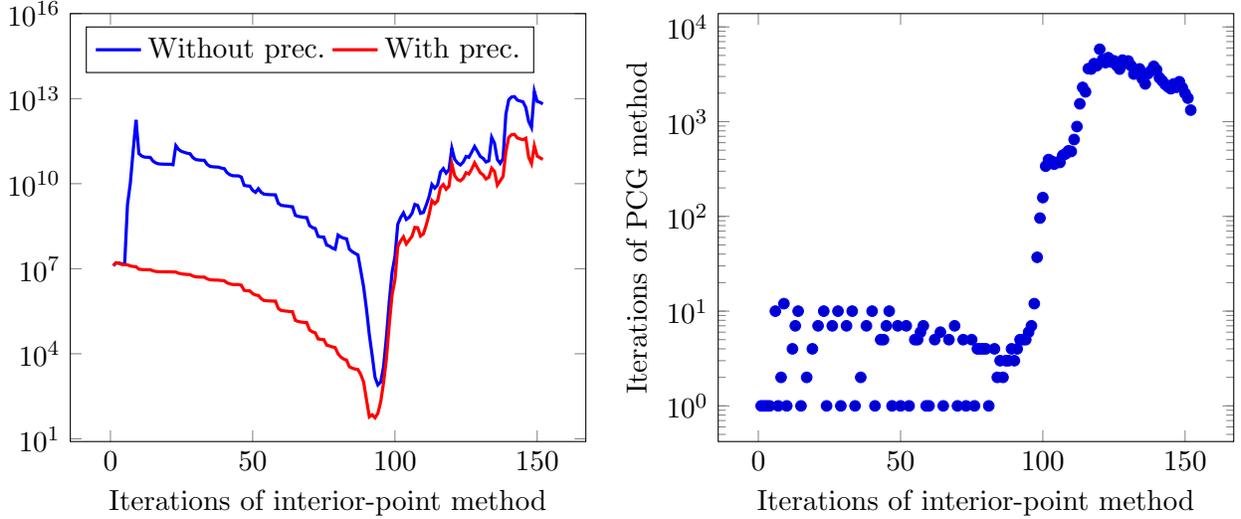


Figure 10: The left image shows the condition number of the matrix in the Newton system (2.36) with and without the preconditioning defined in (2.38), as the interior-point solver proceeds. Preconditioning does not succeed in significantly reducing the condition number, especially as the method converges. The right image shows the number of conjugate-gradient iterations needed to solve the preconditioned Newton system, which become impractically large after 100 iterations. The experiment is carried out by fitting a two-compartment model using a dictionary containing  $10^4$  columns.

combined with an iterative conjugate-gradient method to invert the system. However, in our case the preconditioner is not very effective in improving the conditioning of the Newton system, as shown in Figure 10.

In order to solve the Newton system (2.36) efficiently, we take a different route: exploiting the low-rank structure of the Hessian by applying the Woodbury inversion lemma, also known as Sherman-Morrison-Woodbury formula.

**Lemma 2.1** (Woodbury inversion lemma). *Assume  $A \in \mathbb{C}^{n \times n}$ ,  $B, G \in \mathbb{C}^{n \times r}$  and  $L \in \mathbb{C}^{r \times r}$ . Then,*

$$(A + BLG^*)^{-1} = A^{-1} - A^{-1}B(L^{-1} + GA^{-1}B)^{-1}GA^{-1} \quad (2.39)$$

Setting  $A := Z$ ,  $B := D$ ,  $G := D^*$  and  $L := I \in \mathbb{C}^k$  in the lemma yields the following expression for the inverse of the system matrix

$$(Z + 2tD^*D)^{-1} = Z^{-1} - 2tZ^{-1}D^*(I^{-1} + 2tDZ^{-1}D^*)^{-1}DZ^{-1}. \quad (2.40)$$

Since in practice  $k$  is very small (typically between 10 and 15) due to the dictionary-compression scheme described in Section 2.3 and  $Z$  is diagonal, using this formula accelerates the inversion of the Newton system dramatically.

Figure 11 compares the computational cost of our method based on Woodbury-inversion with a general-purpose convex-programming solver based on an interior-point method [36] and an interior-point method tailored to large-scale problems, which applies the preconditioner in Eq. (2.38) and

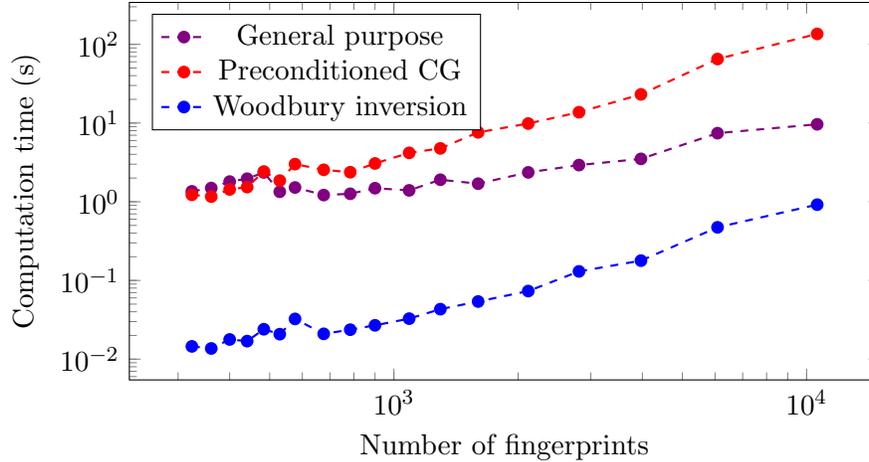


Figure 11: Computation times of the proposed Woodbury-inversion method compared to a general-purpose solver [36] and an approach based on preconditioned conjugate-gradients (CG) [43]. MRF dictionaries containing different numbers of fingerprints in the same range ( $T_1$  values from 0.1 s to 2 s,  $T_2$  values from 0.005 s to 0.1 s) were generated using the approach described in Refs. [3,4]. Each method was applied to solve 10 instances of problem (2.34) for each dictionary size on a computer cluster (Four AMD Opteron 6136, each with 32 cores, 2.4 GHz, 128 GB of RAM). In all cases, the interior-point iteration is terminated when the gap between the primal and dual objective functions is less than  $10^{-5}$ .

uses an iterative method to solve the Newton system [43]. Our method is almost an order of magnitude faster for dictionaries containing up to  $10^4$  columns.

To conclude, we note that this algorithm can be directly applied to weighted  $\ell_1$ -norm-regularization least squares problems of the form

$$\begin{aligned} & \underset{\tilde{c} \in \mathbb{R}^m}{\text{minimize}} && \frac{1}{2} \|D\tilde{c} - x\|^2 + \lambda W\tilde{c} \\ & \text{subject to} && \tilde{c} \geq 0, \end{aligned} \tag{2.41}$$

where  $W$  is a diagonal matrix containing the weights. We just need to apply the change of variable  $\tilde{c}' = W\tilde{c}$  and  $D' = DW^{-1}$ . Since the weights are all positive, the constraint  $\tilde{c} \geq 0$  is equivalent to  $\tilde{c}' \geq 0$ .

## 3 Results

### 3.1 Numerical simulations

In this section we evaluate the performance of the sparse recovery algorithm described in Sections 2.5 and 2.6 on simulated data.

### 3.1.1 Single-voxel data

In this section we consider simulated data from a voxel that contains two compartments with equal relative contribution. The parameters of the first compartment are fixed ( $T_1 = 0.21$  s,  $T_2 = 9$  ms), whereas the parameters of the second compartment vary on a grid ( $T_1$  between 0.1 s and 2.0 s and  $T_2$  between 5 ms and 100 ms). To make the setting more realistic and incorporate discretization error, the dictionary used for recovery does not match the one used to simulate the data. The data are generated by adding the fingerprints from the two compartments and perturbing the result with i.i.d. Gaussian noise. The signal-to-noise ratio is defined as the ratio between the  $\ell_2$  norm of the signal and the noise.

Figure 12 shows the error when fitting the multicompartment model using the reweighting scheme defined in Eq. (2.31). The error of the fixed compartment is depicted as a function of the relaxation times of the second compartment. The closer the two compartments are in  $T_1$ - $T_2$ -space, the more correlated their corresponding fingerprints are, which results in a more challenging sparse-recovery problem. In the noiseless case, exact recovery occurs (even without reweighting) except for combination of tissues with almost identical relaxation times. For the noisy data, the algorithm tends to detect the correct number of compartments as long as the relaxation times of the two compartments are sufficiently separated (b, c). When the relaxation times of the two compartments are very similar, the algorithm cannot separate the compartments and returns, as expected, a single compartment (cf. red region around the green square in Subfigure b). Increasing the noise increases the size of this region (c). Within the region, one can observe that the relaxation times of the reconstructed single compartment lie between the two original compartments (e, f, h, i). Outside of the region, where multiple compartments are detected, the observed error is mostly noise-like and increases as the noise level increases (e, f, h, i). In some cases, a spurious third compartment can be observed (b, c).

### 3.1.2 Boundary estimation

When applying the proposed approach to the numerical experiment sketched in Figure 1, we find that the two contributing compartments are correctly identified (Figure 13). The relaxation times of both compartments are constant over space, in accordance with the ground truth. The proton density, which reflects the relative contributions of each compartment to the signal, exhibits a gradual variation that correctly reflects the effect of the low-pass filtering of the measurements in Fourier space.

### 3.1.3 Simulated phantom

In this section we evaluate our methods on a numerical phantom that mimics a realistic MRF experiment with radial Fourier, or  $k$ -space sampling. The numerical phantom has 16 different  $19 \times 19$  voxel regions. The voxels in each region consist of either one or two compartments with relaxation times in the same range as biological tissues such as myelin-water (A), gray matter (B), white matter (C), and cerebrospinal fluid (C). Figure 14 depicts the structure of the phantom along with the corresponding relaxation times. For each compartment, a fingerprint was calculated with Bloch simulations utilizing the flip-angle pattern described in [4] (see Fig. 3k). These fingerprints

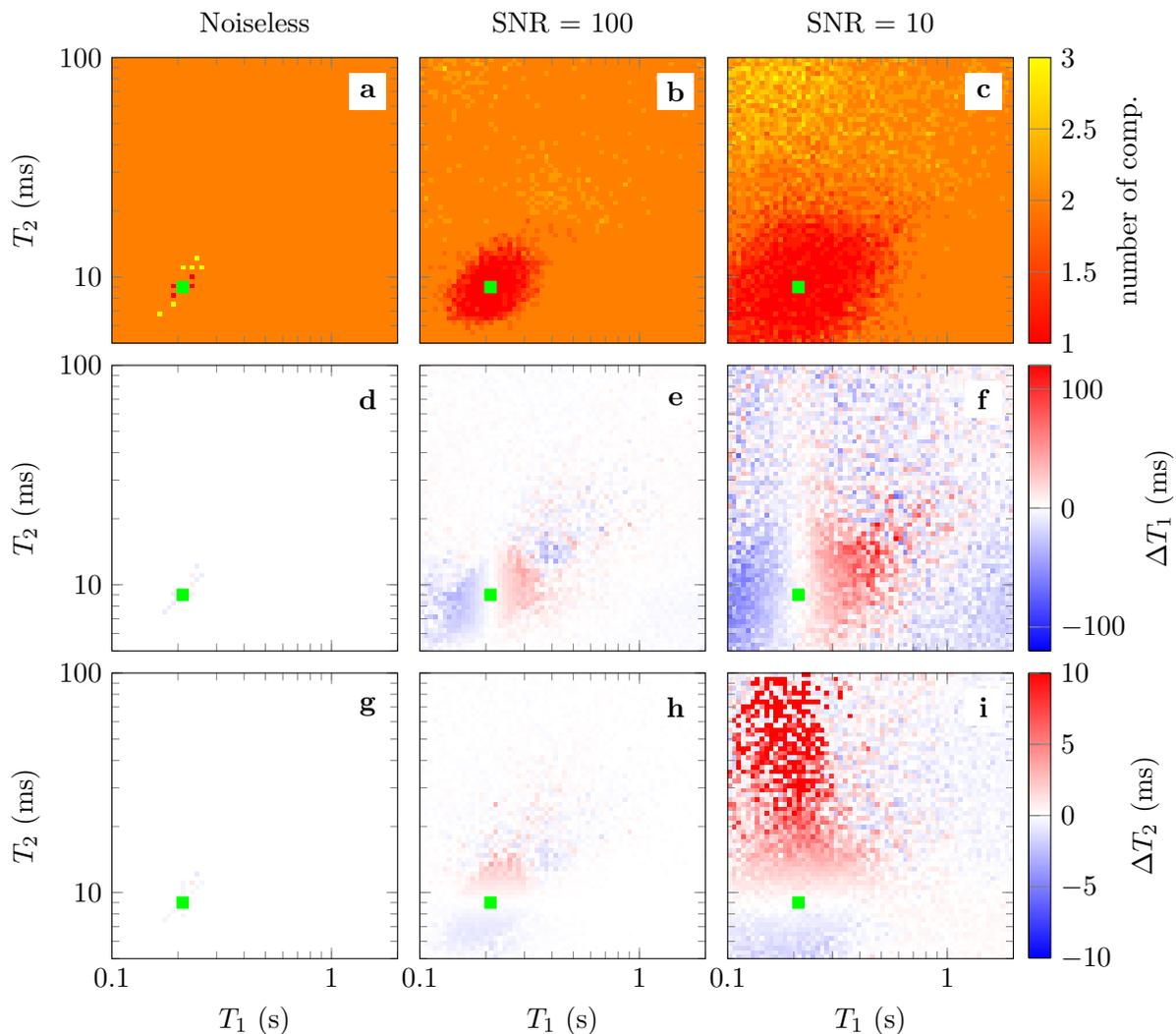


Figure 12: Results of fitting a multicompartment model where one compartment is fixed to  $T_1 = 0.21$  s and  $T_2 = 9$  ms (marked by the green square), and the second compartment has different  $T_1$  (between 0.1 s and 2.0 s) and  $T_2$  values (between 5 ms and 100 ms). The proton density of each compartment is fixed to 0.5. The fingerprints are generated following Refs. [3, 4]. The data are perturbed by adding i.i.d. Gaussian noise to obtain two different SNRs (second and third columns). The multicompartment model is fit using the reweighting scheme defined in Eq. (2.31) combined with the method in Section 2.6. The heat maps show the results for each different  $T_1$  and  $T_2$  value of the second compartment at the corresponding position in  $T_1$ - $T_2$  space, color-coded as indicated by the colorbars. The first row depicts the number of compartments resulting from the reconstruction (the ground truth is two). The second and third row show the  $T_1$  and  $T_2$  errors for the first compartment respectively. These errors correspond to the difference between the relaxation times of the fixed compartment ( $T_1 = 0.21$  s and  $T_2 = 9$ ) and the closest relaxation time of the reconstructed atoms. The results are averaged over 5 repetitions with different noise realizations.

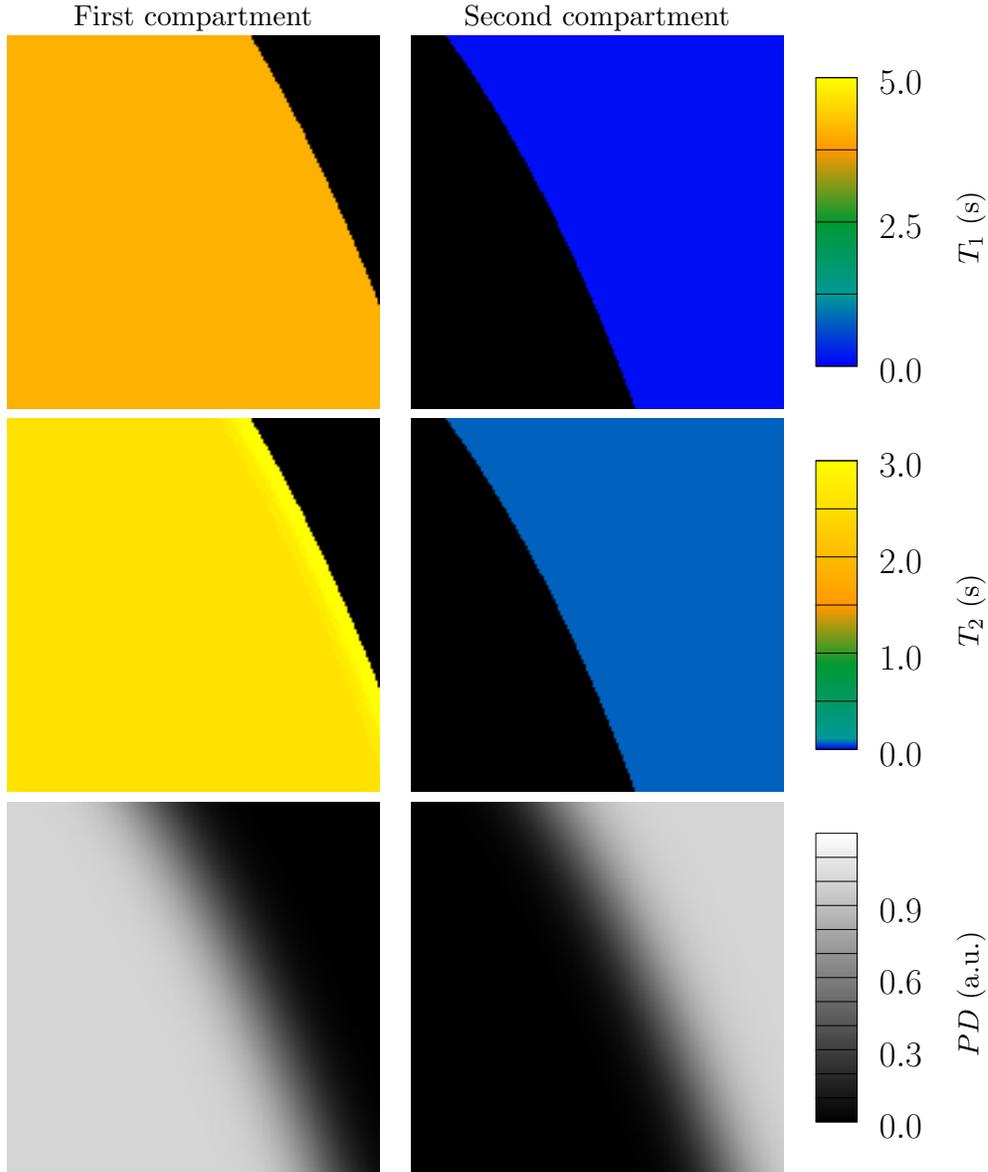


Figure 13: Results of fitting our multicompartment model, using the algorithm described in Section 2.5 and Section 2.6, to the data in Figure 1. The proposed multicompartment reconstruction correctly identifies the relaxation times of the two compartments, as well as their relative contributions, quantified by the proton density (PD) of each compartment.  $T_1$  and  $T_2$  maps of all voxels with proton density less than 0.01 are depicted in black. Since the simulated data do not contain noise, we set the regularization parameter  $\lambda$  to a very small value ( $10^{-8}$ ). The reweighting parameter  $\epsilon$  is set to  $10^{-10}$ .

are then combined additively to yield the multicompartment fingerprints of each voxel. The data correspond to undersampled k-space data of each time frame, calculated with a non-uniform fast Fourier transform [2, 32] with 16 radial k-space spokes per time frame. The fingerprints consist of 850 time frames, and the trajectory of successive time frames are rotated by 16 times the golden angle with respect to each other [75]. Noise with an i.i.d Gaussian distribution was added to the simulated  $k$ -space data to achieve different signal-to-noise ratios.

When two different tissues are present in a voxel, two conditions are necessary so that the problem of distinguishing their contributions to the signal is well posed. First, their corresponding fingerprints should be sufficiently distinct, i.e. have a low correlation coefficient (see Section 2.4). Second, the weighted sum of their fingerprints should be different from any other fingerprint in the dictionary. Otherwise, a single-compartment model will fit the data well. Figure 15 shows the correlations between the fingerprints corresponding to each of the tissues present in the numerical phantom and the rest of the fingerprints in the dictionary. For tissues that have similar  $T_1$  and  $T_2$  values, the correlation is very high. In particular, the fingerprints corresponding to tissue B and C are extremely similar. Furthermore, their additive combination is almost indistinguishable from another fingerprint also present in the dictionary, as illustrated in Figure 16. It will consequently be almost impossible to distinguish a voxel containing a single tissue with that particular fingerprint and a two-compartment voxel containing tissue B and C if there is even a very low level of noise in the data. Fortunately, this is not the case for the other combinations of the four relaxation times of interest.

To estimate the relaxation times  $T_1$  and  $T_2$  from the simulated data we apply the method described in Section 2.2 while fixing the number of ADMM iterations to 10. We compress the dictionary by truncating its SVD to obtain a rank-10 approximation, as described in Section 2.3. The magnetization vector  $\tilde{X}$  is updated by applying 10 conjugate-gradient iterations. The coefficient variable  $\tilde{C}$  is updated by applying the reweighted- $\ell_1$  method described in Section 2.5 over 5 iterations with the reweighting scheme in equation (2.30) (the reweighting scheme (2.31) yields similar results). For each iteration, the interior-point method described in Section 2.6 is run until the gap between the primal and dual objective functions is less than  $10^{-5}$ .

Figure 17 shows the reconstructed parameter maps for a signal-to-noise ratio (SNR) of  $10^3$  (the SNR is defined as the ratio between the  $\ell_2$ -norm of the signal and the added noise), and using 16 radial k-space spokes. The multicompartment reconstruction recovers the two compartments accurately, except for the voxels that contain the highly-correlated tissues B and C. This behavior is expected due to the high correlations of the fingerprints of those tissues (see Figure 16). The model also yields an accurate estimate of the proton density of the compartments that are present in each voxel (except, again, for voxels containing tissue B and C). In most voxels, the reconstruction does not detect additional compartments, even though this is not explicitly enforced. Figure 18 shows that the performance degrades gracefully when the SNR is decreased to 100: the model still achieves similar results, only noisier, and a spurious third compartment with a relatively low proton density appears in some voxels.

To evaluate the performance of the method at different undersampling factors, we fit the multicompartment model to the simulated phantom data using different number of radial k-space spokes. Figure 19 shows the decrease in the normalized root-mean-square-error (NRMSE) of the  $T_1$  and  $T_2$  estimates averaged over all voxels in the phantom. As expected, the NRMSE decreases when acquiring more k-space spokes. As explained above, the largest contributor to the NRMSE are the

tissue	A	B	C	D
$T_1(s)$	0.21	0.66	0.91	2.32
$T_2(s)$	0.009	0.042	0.062	0.416

	Column 1	Column 2	Column 3	Column 4
Row 1	100% A	50% A + 50% B	50% A + 50% C	50% A + 50% D
Row 2	50% B + 50% A	100% B	50% B + 50% C	50% B + 50% D
Row 3	50% C + 50% A	50% C + 50% B	100% C	50% C + 50% D
Row 4	50% D + 50% A	50% D + 50% B	50% D + 50% C	100% D

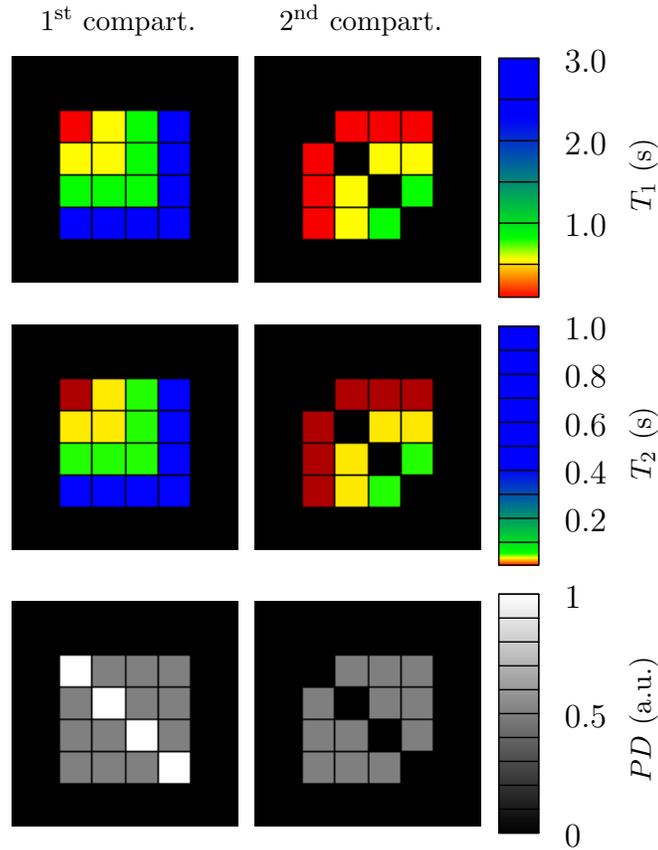


Figure 14: The numerical phantom consists of four different synthetic tissues with relaxation times in the range commonly found in biological brain tissue. Each voxel contains one or two tissue compartments, as depicted in the figure. The relaxation times of the tissues A, B, C and D are in the range found in myelin-water, gray matter, white matter, and cerebrospinal fluid respectively.

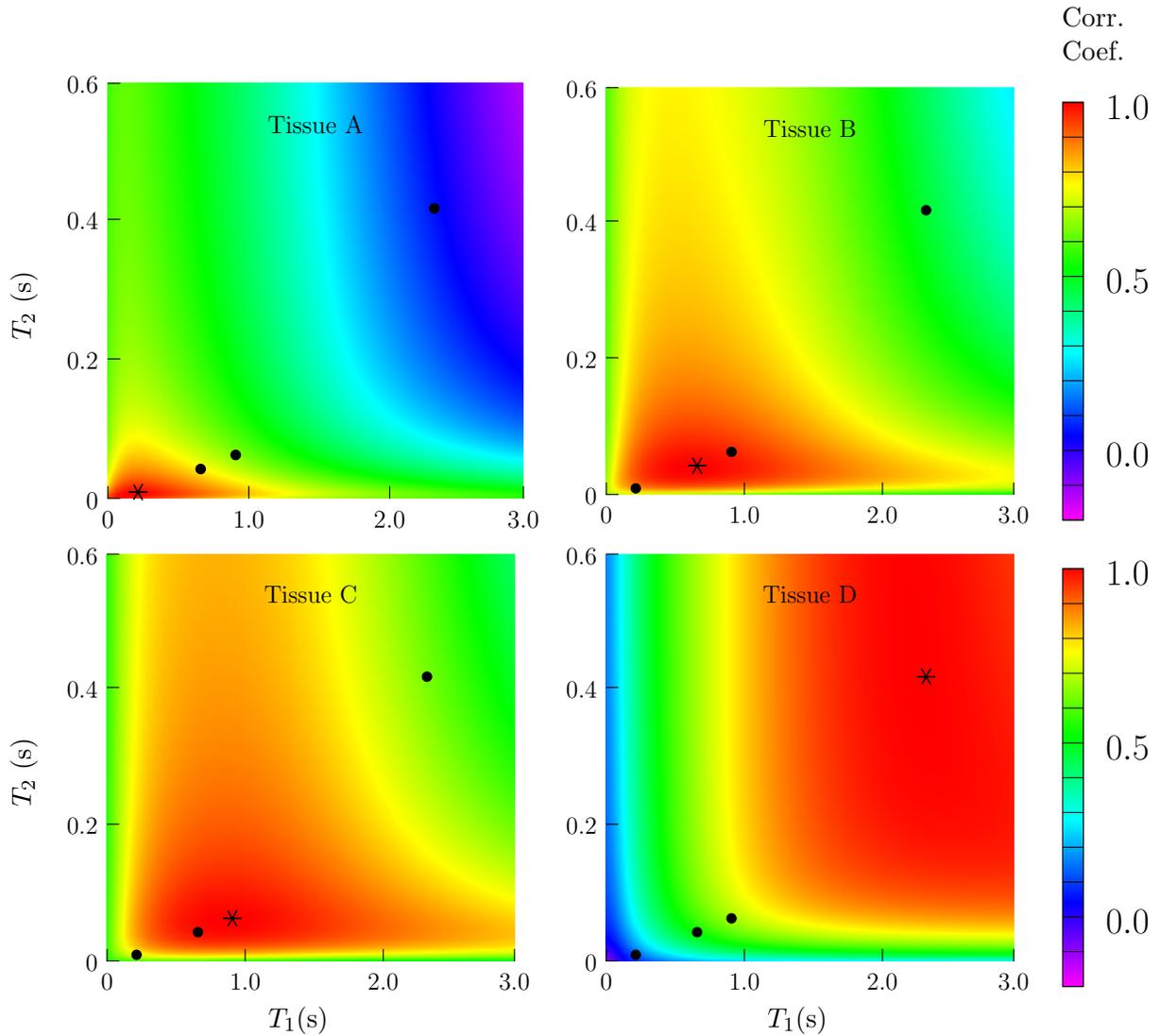


Figure 15: Correlation between the fingerprints corresponding to the four tissues present in the numerical phantom with every other fingerprint in the dictionary, indexed by the corresponding  $T_1$  and  $T_2$  values. The tissue corresponding to the fixed fingerprint is marked with a star, the position of the other three tissues also present in the phantom are marked with dots.

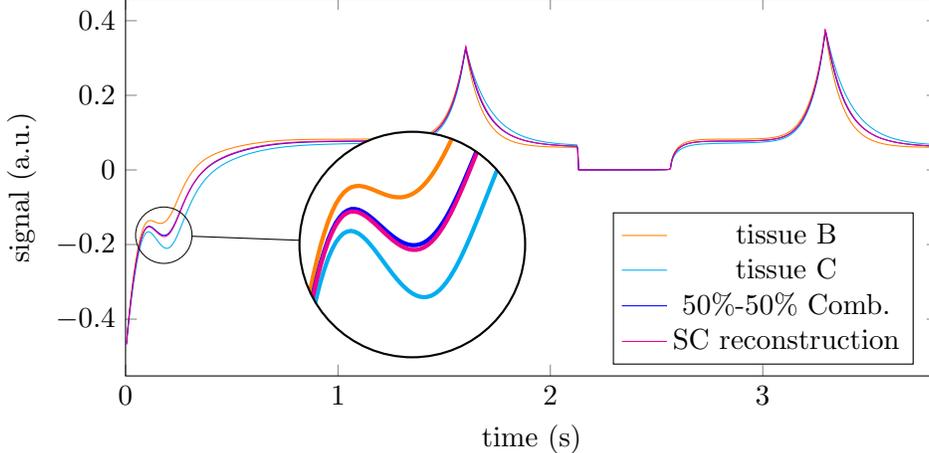


Figure 16: Simulated fingerprints of tissue B and C, and a 50%-50% combination of those tissues are shown, along with a single-compartment (SC) reconstruction, i.e. the fingerprint in the dictionary that is closest to the 50%-50% combination. In this particular example, the fingerprints of those tissues are highly correlated, and the single-compartment reconstruction represents the data well (the relative error in the  $\ell_2$ -norm is 1.42%). As a result, the multicompartiment reconstruction cannot distinguish the two compartments if the data contain even a small level of noise.

voxels that contain both tissue B and C, so the NRMSE excluding these voxels is also depicted. The scatter plots in Figure 20 show the results at each voxel containing a 50%-50% mixture of tissues A and C. The two compartments are separated accurately, even when only two radial k-space spokes are measured. As more samples are gathered, the two clusters corresponding to the two tissues concentrate more tightly around the ground truth. Similarly, when keeping the number of k-space spokes fixed and increasing the SNR, the NRMSE decreases as expected (Figure 21) and the estimated parameters converge to the ground truth (Figure 22). The dependence of the reconstruction quality on the correlation of the fingerprints is visualized in Figure 23, where scatter plots for all tissue combinations are shown. One can observe a dependency of the reconstruction quality on the distance of the two compartments in parameter space. In particular, the combination of tissues B and C, which have highly correlated fingerprints, results in a single compartment fit (see Figure 16).

### 3.2 Phantom experiment

For validation on real data, we performed an imaging experiment on a clinical MR system with a field strength of 3 Tesla (Prisma, Siemens, Erlangen, Germany). We designed a dedicated phantom and manufactured it using a 3D printer (Figure 24). The structure of the phantom is effectively the same as that of the simulated phantom described in Section 3.1 and Figure 14: the diagonal elements contain only one compartment, whereas the off-diagonal elements contain two compartments. The compartments were filled with different solutions of doped water in order to achieve relaxation times in the range of biological tissues. MRF scans were performed with the sequence described in Fig. 3k of Ref. [4] with 8 radial k-space spokes per time frame and the same parameters otherwise.

Figure 25 shows the parameter maps reconstructed with a single-compartment model and a low-

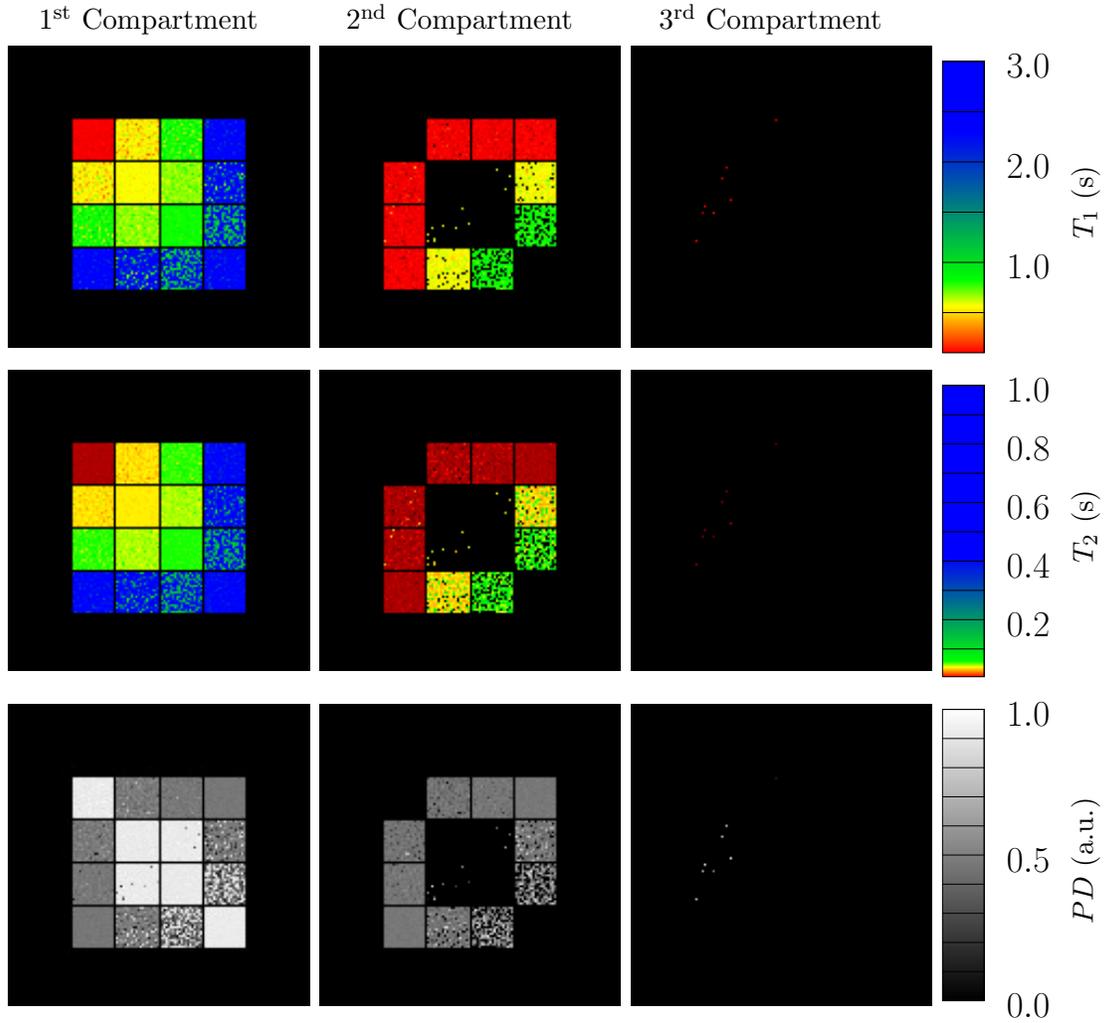


Figure 17: The depicted parameter maps were reconstructed from data simulated with an SNR =  $10^3$  and using 16 radial k-space spokes. The proton density (PD, third row) indicates the fraction of the signal corresponding to the recovered compartments. The compartments in each voxel are sorted according to their  $\ell_2$ -norm distance to the origin in  $T_1$ - $T_2$  space. Compare to Figure 14, which shows the ground truth.

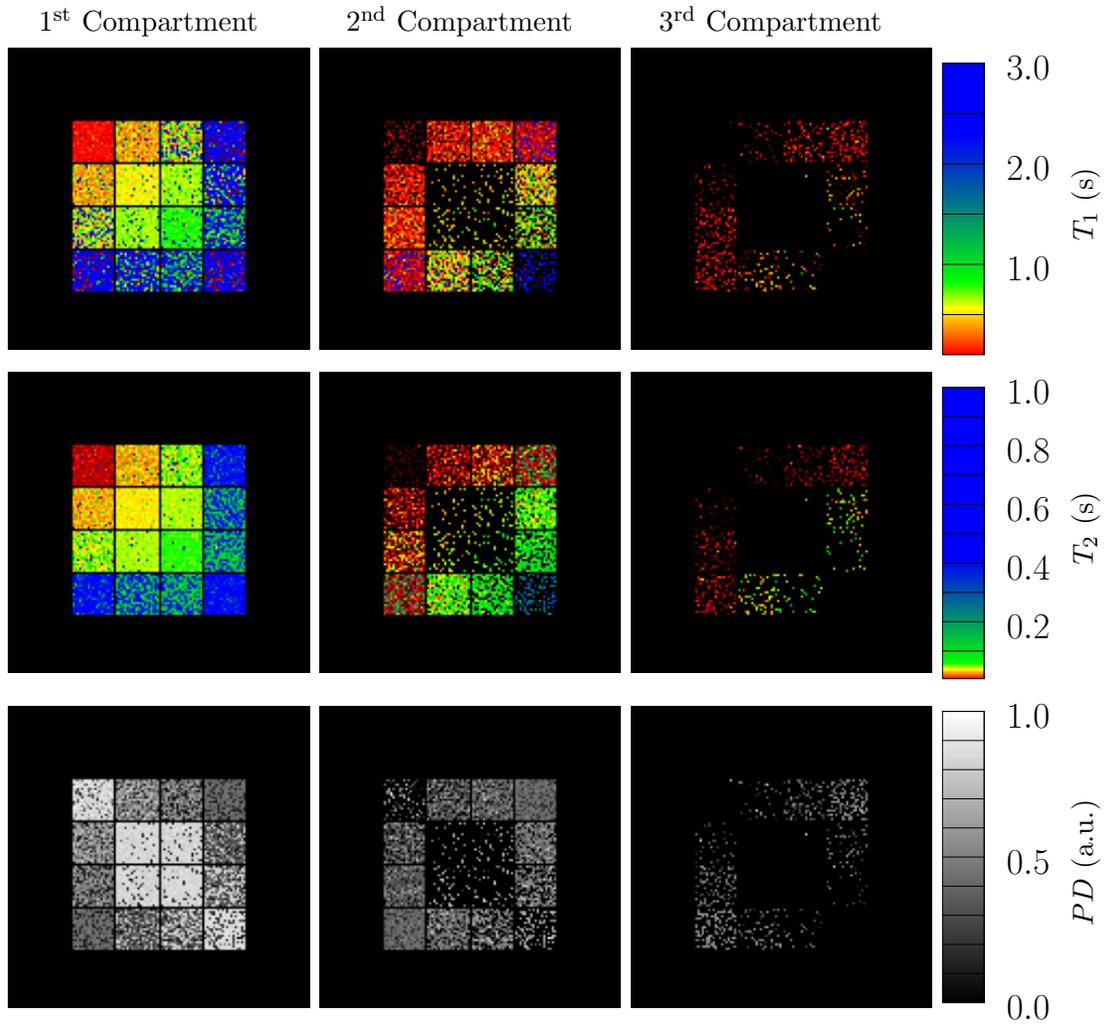


Figure 18: The depicted parameter maps were reconstructed from data simulated with an SNR = 100 and using 16 radial k-space spokes. The proton density (PD, third row) indicates the fraction of the signal corresponding to the recovered compartments. The compartments in each voxel are sorted according to their  $\ell_2$ -norm distance to the origin in  $T_1$ - $T_2$  space. Compare to Figure 14, which shows the ground truth.

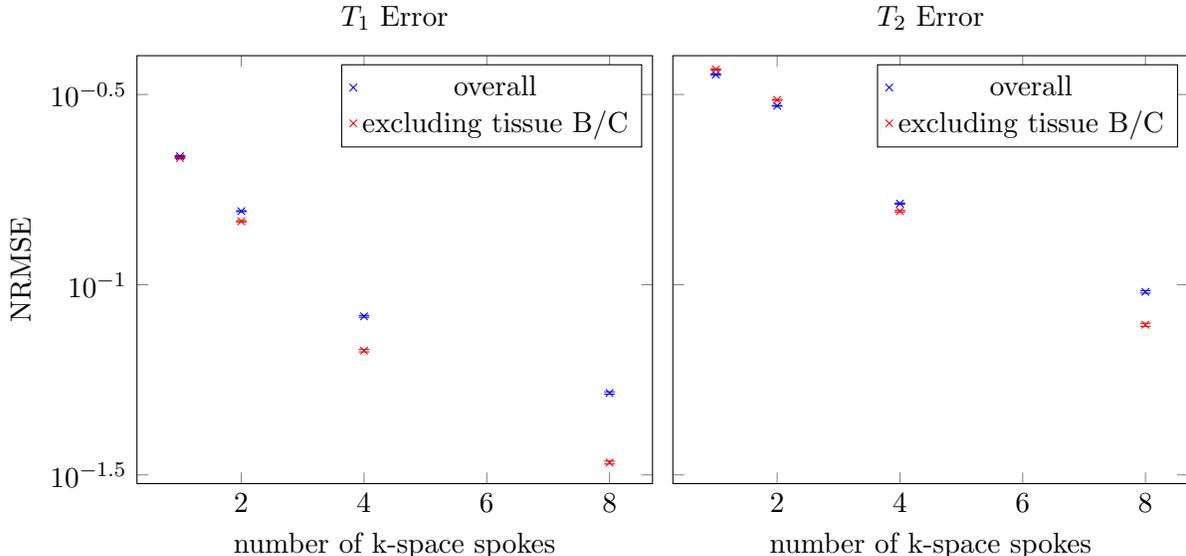


Figure 19: The normalized root-mean-square errors (NRMSE) in the  $T_1$  and  $T_2$  estimates are plotted as a function of the number of radial k-space spokes at a fixed SNR = 1000. The NRMSE is defined as the root-mean square of the  $\Delta T_{1,2}/T_{1,2}$ . The results are averaged over 10 realizations of the noise, and the hardly-visible error bars depict one standard deviation of the variation between different noise realizations. The NRMSE was also computed excluding voxels containing a combination of tissue B and C, because the corresponding fingerprints are highly correlated.

rank ADMM algorithm [2]. For the diagonal elements of the grid phantom, which only contain signal from a single compartment, we consider the identified  $T_1$  and  $T_2$  values listed in Table 1 as the reference gold standard. In the off-diagonal elements, the single-compartment model fails to account for the presence of two compartments, so the apparent relaxation times lie in between the true values of the two compartments. Note that the square shape of the phantom causes significant variations in the main magnetic field towards the edges of the phantom. Since we do not account for those variations, substantial artifacts can be observed in those areas (Figure 25).

Figure 26 shows the parameter maps reconstructed with the proposed multicompartment reconstruction method. At the diagonal elements, results are consistent with the single-compartment reconstruction, except for some voxels where a spurious second compartment appears with low proton density. This indicates that the method correctly identifies voxels containing only a single compartment. In the off-diagonal grid elements, two significant compartments are detected for most voxels, and the relaxation times coincide with the gold-standard within a standard deviation (Table 2), even though the noise level is quite severe at the given experimental conditions. A spurious third compartment with low proton density is detected in some voxels. Many voxels containing the combination of the doped water solutions B and C are reconstructed as a single compartment with apparent relaxation times between the gold-standard values. This is expected, since the corresponding fingerprints are almost indistinguishable, as in the case of the simulated tissues B and C (see Fig. 12). Errors are larger for  $T_2$  than for  $T_1$ , as expected from the Cramér-Rao bound estimation performed in [4]. The artifacts towards the edge of the phantom are due to variations in the main magnetic field (this effect is also visible in the single-compartment reconstruction in

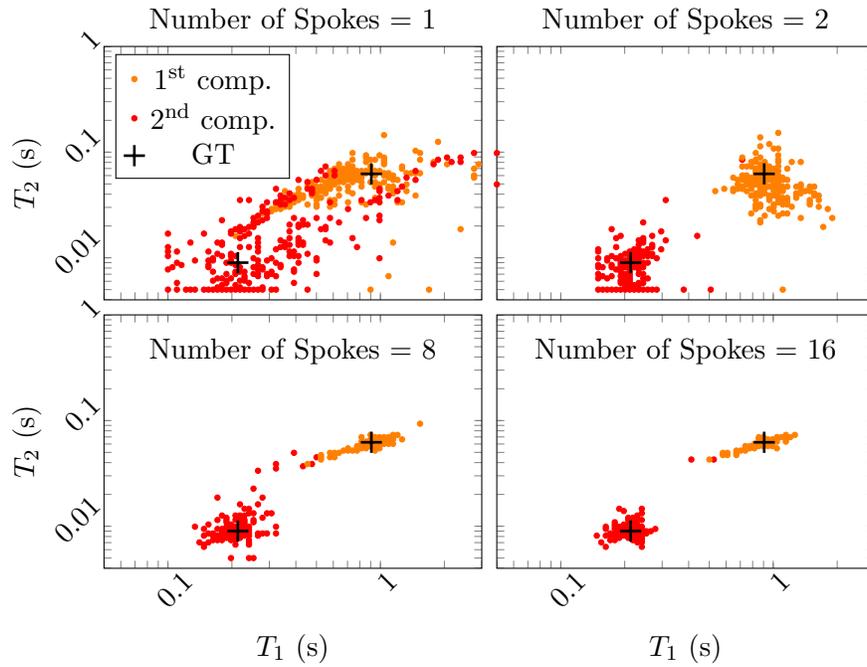


Figure 20: Scatter plots showing the estimated  $T_1$  and  $T_2$  values in the multicompartiment model for voxels containing tissue A and C. Different number of radial k-spaces spokes are tested at a fixed SNR of  $10^3$ . The two clusters corresponding to the two tissues concentrate more tightly around the ground truth (GT, marked with black crosses) as the number of spokes increases.

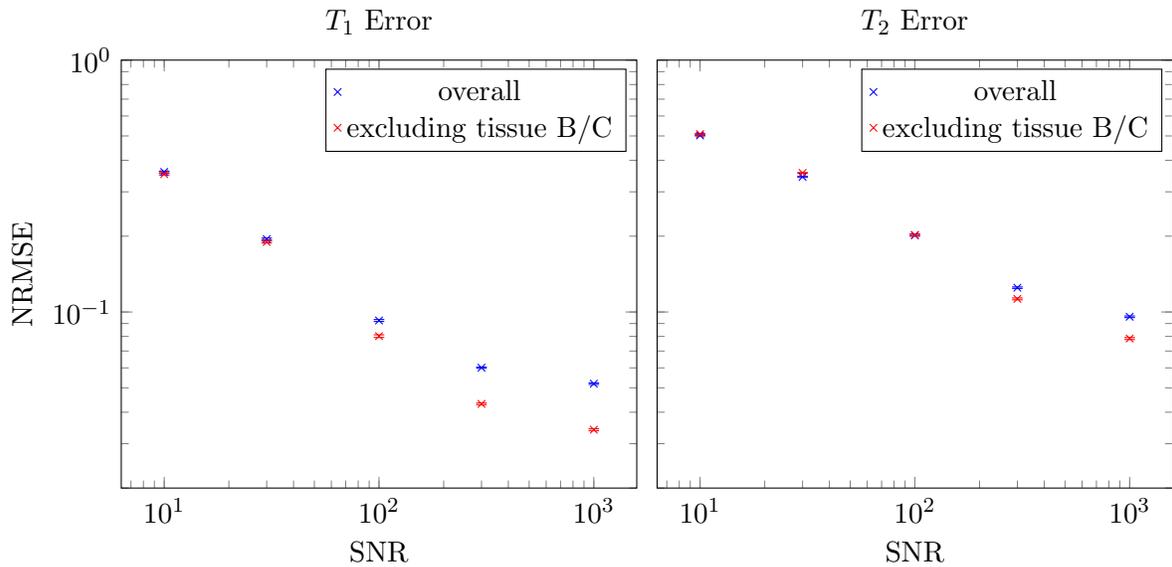


Figure 21: The normalized root-mean-square errors (NRMSE) in the  $T_1$  and  $T_2$  estimates are plotted as a function of the SNR when acquiring 8 radial k-space spokes. The NRMSE is defined as the root-mean square of the  $\Delta T_{1,2}/T_{1,2}$ . The results are averaged over 10 realizations of the noise, and the hardly-visible error bars depict one standard deviation of the variation between different noise realizations. The NRMSE was also computed excluding voxels containing a combination of tissue B and C, because the corresponding fingerprints are highly correlated.

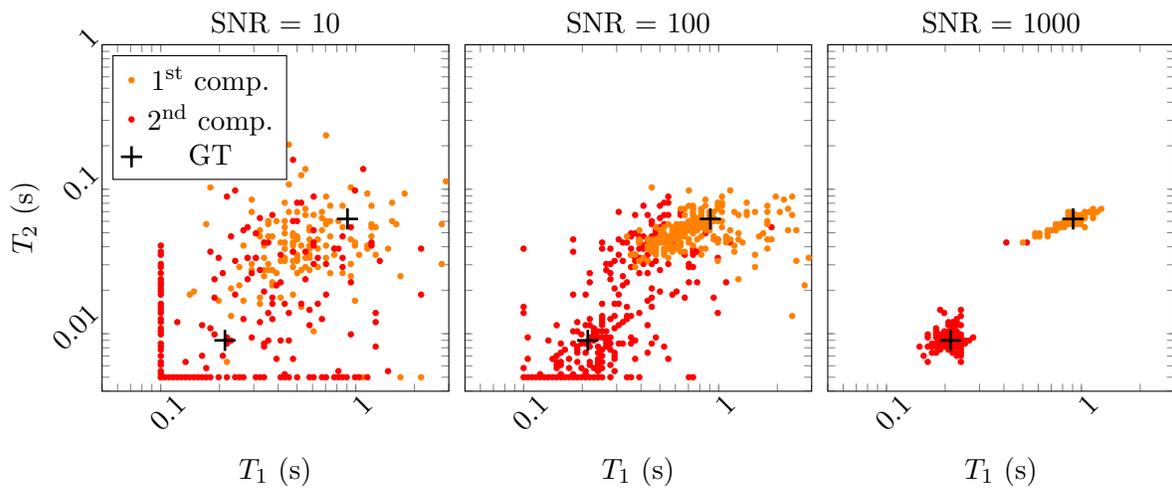


Figure 22: Scatter plots showing the estimated  $T_1$  and  $T_2$  values in the multicompartement model for voxels containing tissue A and C. The data was reconstructed from 16 radial k-space spokes. The two clusters corresponding to the two tissues concentrate more tightly around the ground truth (GT, marked with black crosses) as the SNR increases.

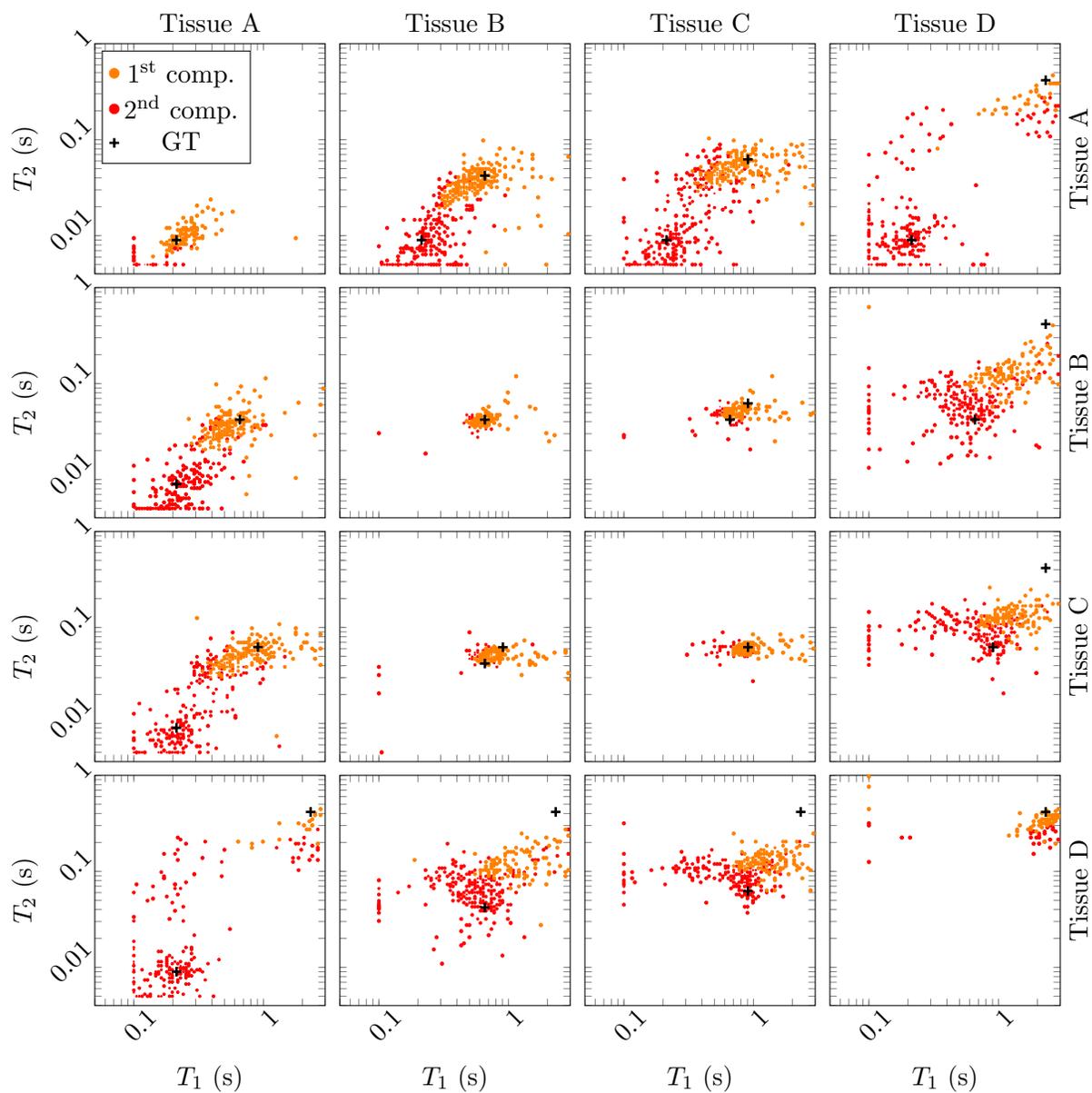


Figure 23: Scatter plots showing the estimated  $T_1$  and  $T_2$  values in the multicompartement model for simulated data with an SNR = 100 and using 16 radial k-space spokes. The tissues contained in those voxels are indicated above and to the right. The ground truth value of their respective relaxation times is marked with black crosses.

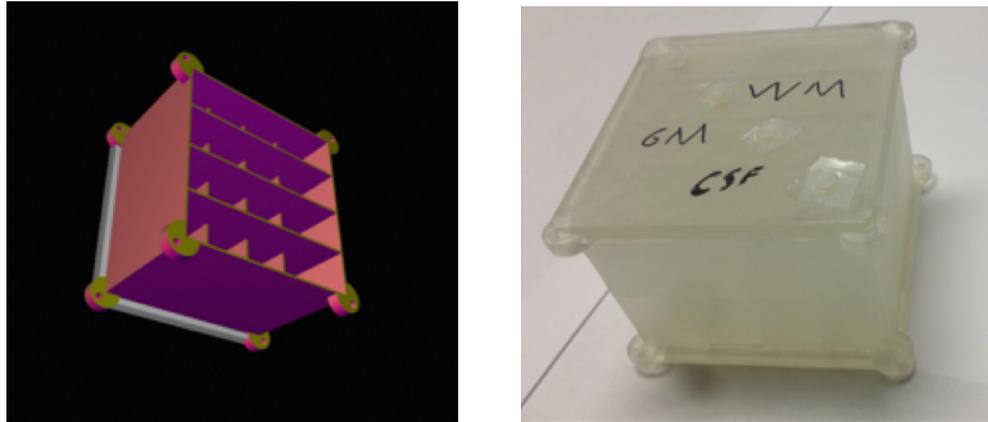


Figure 24: Computer-aided design (left) and the corresponding 3D-printed phantom (right) used for the phantom evaluation. It contains two layers, one with four horizontal and one with four vertical bars. Imaging a slice that contains both layers results in the same effective structure as the numerical phantom shown in Figure 14.

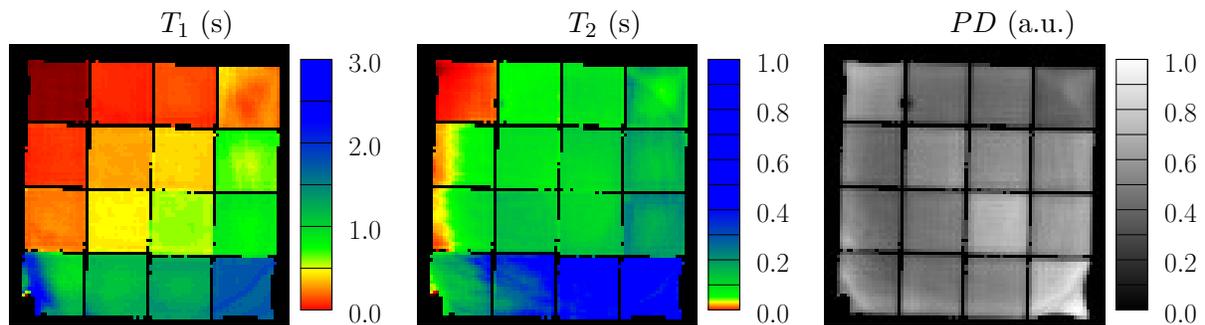


Figure 25: The depicted parameter maps were reconstructed from data measured in a phantom and using a single-compartment model. The data were acquired using 8 radial k-space spokes. The colorbar is adjusted so that distinct colors correspond to the parameters of the gold-standard reference measurements: red, yellow green and blue represent the solutions A-D.

	A	B	C	D
$T_1$ (s)	$0.1262 \pm 0.0075$	$0.5075 \pm 0.015$	$0.7640 \pm 0.055$	$1.831 \pm 0.061$
$T_2$ (s)	$0.0174 \pm 0.0046$	$0.1019 \pm 0.0084$	$0.1328 \pm 0.018$	$0.66 \pm 0.21$

Table 1: The gold-standard (GS) values measured by single-compartment reconstruction on the diagonal grids of phantom described in Figure 24, where only a single compartment is present.

	A and C		A and D		B and D	
	A	C	A	D	B	D
$T_1$ (s)	$0.35 \pm 0.25$	$0.58 \pm 0.26$	$0.24 \pm 0.26$	$1.40 \pm 0.57$	$0.66 \pm 0.18$	$1.49 \pm 0.36$
$T_2$ (s)	$0.068 \pm 0.071$	$0.20 \pm 0.20$	$0.063 \pm 0.24$	$1.20 \pm 0.58$	$0.080 \pm 0.11$	$1.25 \pm 0.44$

Table 2: Mean values and standard deviations of the multicompartment estimates for the selected combinations appearing in Figure 27.

Figure 25).

Further insights are provided by the scatter plots depicted in Figure 27, which show the estimated relaxation times of the voxels containing three combinations of tissues. For these combinations, the multicompartment reconstruction correctly detects the presence of both tissues and the estimated relaxation times cluster around the gold-standard values. In contrast, the single-compartment algorithm results in a single cluster of apparent relaxation times at a position in parameter space that provides little information about the underlying tissue composition.

## 4 Conclusions and Future Work

In this work we propose a framework to perform quantitative estimation of parameter maps from magnetic resonance fingerprinting data, which accounts for the presence of more than one tissue compartment in each voxel. This multi-compartment model is fit to the data by solving a sparse linear inverse problem using reweighted  $\ell_1$ -norm regularization. Numerical simulations show that an efficient interior-point method is able to tackle this sparse-recovery problem, yielding sparse solutions that correspond to the parameters of the tissues present in the voxel.

The proposed method is validated through simulations, as well as with a controlled phantom imaging experiment on a clinical MR system. The results indicate that incorporating a sparse-estimation procedure based on reweighted  $\ell_1$ -norm regularization is a promising avenue towards achieving multicompartment parameter mapping from MRF data. However, the proposed method still requires a high SNR, which entails long scan times for in vivo applications. A promising research direction is to incorporate additional prior assumptions on the structure of the parameter maps in order to leverage the method in the realm of clinically-feasible scan times. Future work will also include the application of the proposed framework to the quantification of the myelin-water fraction in the human brain, which is an important biomarker for neurodegenerative diseases.

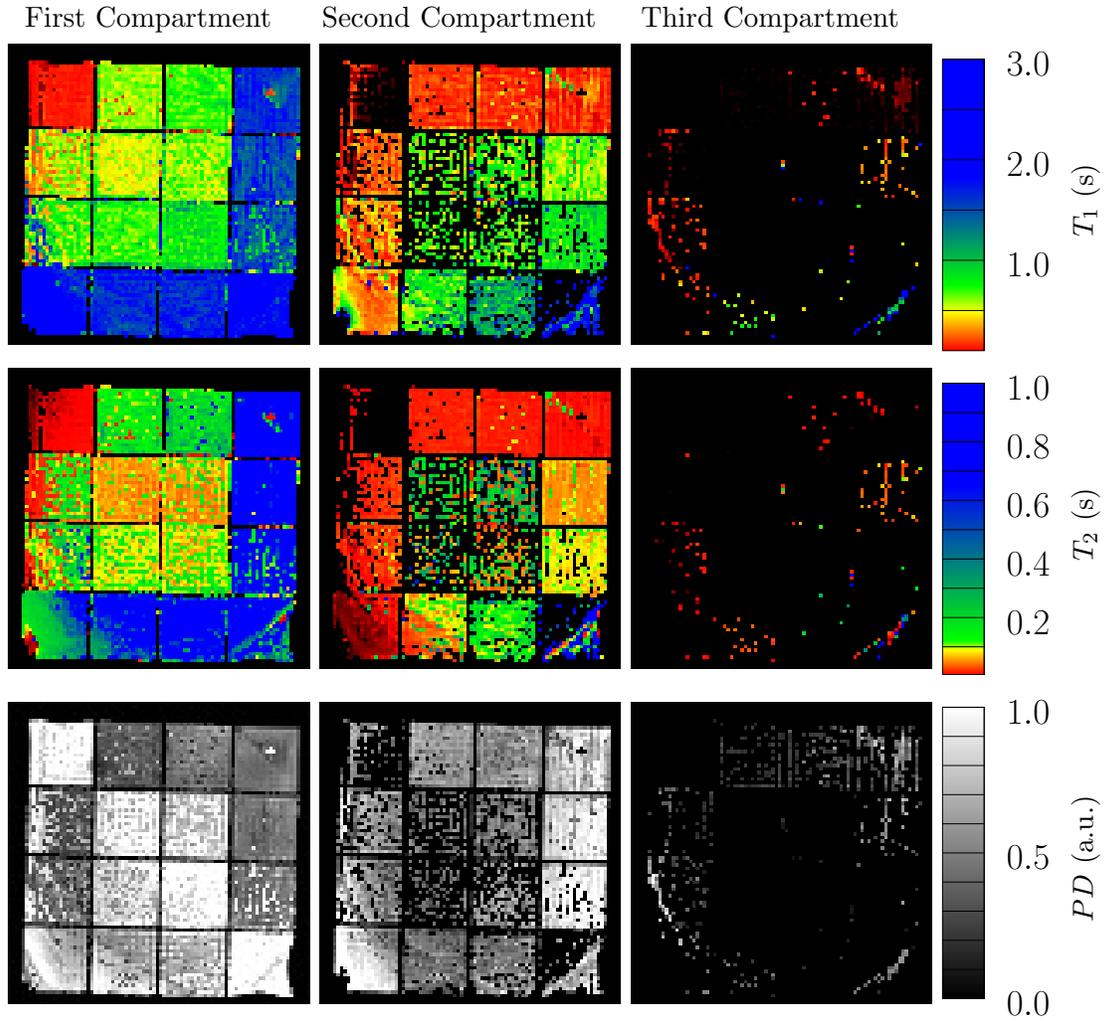


Figure 26: The parameter maps of the scanned phantom was reconstructed using the proposed multicompartment method. The data were acquired using eight radial k-space spokes. The three compartments with the highest detected proton density are shown here, and are sorted in each voxel according to their  $\ell_2$ -norm distance to the origin in  $T_1$ - $T_2$  space. The colorbar is adjusted so that distinct colors correspond to the parameters of the gold-standard reference measurements: red, yellow green and blue represent the four solutions.

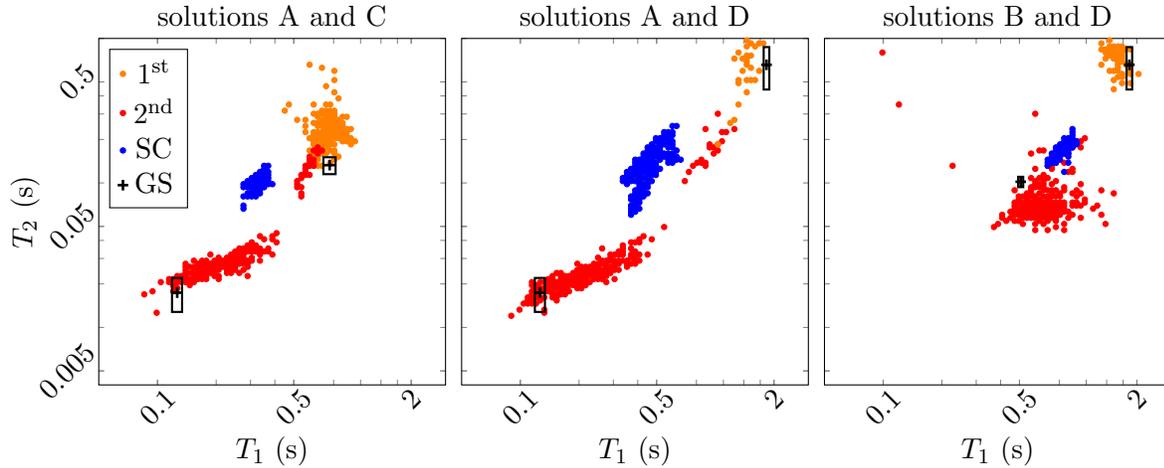


Figure 27: The scatter plot shows the estimated relaxation times of the measured phantom at the example of voxels containing different combinations of compartments. For comparison, the estimates obtained from the single-compartment (SC) are shown, as well as the gold-standard (GS) values measured on the diagonal, where only a single compartment is present. The rectangles represent the standard deviation of the gold-standard measurement.

## Acknowledgements

C.F. is supported by NSF award DMS-1616340. J.A., F.K., R.L. and M.C. are supported by NIH/NIBIB R21 EB020096, NIH/NIAMS R01 AR070297 and NIH P41 EB017183. S. L. is supported by a seed grant from the Moore-Sloan Data-Science Environment at NYU. S. T. is supported in part by the Office of Naval Research under Award N00014-17-1-2059.

## References

- [1] A. Allerhand and H. S. Gutowsky. Spin-Echo NMR Studies of Chemical Exchange. I. Some General Aspects. *The Journal of Chemical Physics*, 41:2115–2126, Oct. 1964.
- [2] J. Assländer, M. A. Cloos, F. Knoll, D. K. Sodickson, J. Hennig, and R. Lattanzi. Low rank alternating direction method of multipliers reconstruction for mr fingerprinting. *Magnetic Resonance in Medicine*, 79(1):83–96, 2018.
- [3] J. Assländer, S. J. Glaser, and J. Hennig. Pseudo steady-state free precession for mr-fingerprinting. *Magnetic resonance in medicine*, 77(3):1151–1161, 2017.
- [4] J. Assländer, R. Lattanzi, D. K. Sodickson, and M. A. Cloos. Relaxation in Spherical Coordinates: Analysis and Optimization of pseudo-SSFP based MR-Fingerprinting. *arXiv*, 1703.00481(v1), 2017.
- [5] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202, 2009.
- [6] S. R. Becker, E. J. Candès, and M. C. Grant. Templates for convex cone problems with applications to sparse signal recovery. *Mathematical programming computation*, 3(3):165–218, 2011.
- [7] B. Bernstein and C. Fernandez-Granda. Deconvolution of Point Sources: A Sampling Theorem and Robustness Guarantees. *preprint*, 2017.
- [8] P. J. Bickel, Y. Ritov, and A. B. Tsybakov. Simultaneous analysis of lasso and  $\{D\}$ antzig selector. *The Annals of Statistics*, pages 1705–1732, 2009.
- [9] F. Bloch. Nuclear induction. *Physical review*, 70(7-8):460, 1946.
- [10] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. *Applied and computational harmonic analysis*, 27(3):265–274, 2009.
- [11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.
- [12] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [13] E. Candes and J. Romberg. l1-magic: Recovery of sparse signals via convex programming. *URL: www.acm.caltech.edu/l1magic/downloads/l1magic.pdf*, 4:14, 2005.
- [14] E. J. Candes. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9-10):589–592, 2008.
- [15] E. J. Candès and C. Fernandez-Granda. Towards a Mathematical Theory of Super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.
- [16] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.
- [17] E. J. Candes and T. Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.
- [18] E. J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.
- [19] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier analysis and applications*, 14(5):877–905, 2008.
- [20] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.

- [21] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.
- [22] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.
- [23] S. C. Deoni. Quantitative Relaxometry of the Brain. *Topics in Magnetic Resonance Imaging*, 21(2):101–113, 2010.
- [24] S. C. L. Deoni, L. Matthews, and S. H. Kolind. One component? two components? three? the effect of including a nonexchanging free water component in multicomponent driven equilibrium single pulse observation of t1 and t2. *Magnetic Resonance in Medicine*, 70(1):147–154, 2013.
- [25] S. C. L. Deoni, B. K. Rutt, T. Arun, C. Pierpaoli, and D. K. Jones. Gleaning multicomponent T1 and T2 information from steady-state imaging data. *Magnetic Resonance in Medicine*, 60(6):1372–1387, 2008.
- [26] A. Deshmane, D. McGivney, C. Badve, A. Yu, Y. Jiang, D. Ma, and M. A. Griswold. Accurate synthetic FLAIR images using partial volume corrected MR fingerprinting. In *Proceedings of the 24th Annual Meeting of ISMRM*, page 1909, 2016.
- [27] D. L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.
- [28] D. L. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via  $\ell_1$  minimization. *Proceedings of the National Academy of Sciences*, 100(5):2197–2202, 2003.
- [29] P. Ehses, N. Seiberlich, D. Ma, F. A. Breuer, P. M. Jakob, M. A. Griswold, and V. Gulani. Ir truefisp with a golden-ratio-based radial readout: Fast quantification of t1, t2, and proton density. *Magnetic resonance in medicine*, 69(1):71–81, 2013.
- [30] E. Esser, Y. Lou, and J. Xin. A method for finding structured sparse solutions to nonnegative least squares problems with applications. *SIAM Journal on Imaging Sciences*, 6(4):2010–2046, 2013.
- [31] C. Fernandez-Granda. Super-resolution of point sources via convex programming. *Information and Inference*, 5(3):251–303, 2016.
- [32] J. Fessler and B. Sutton. Nonuniform fast fourier transforms using min-max interpolation. *IEEE Trans. Signal Process.*, 51(2):560–574, feb 2003.
- [33] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.
- [34] J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.
- [35] J. N. Giedd, J. W. Snell, N. Lange, J. C. Rajapakse, B. J. Casey, P. L. Kozuch, A. C. Vaituzis, Y. C. Vauss, S. D. Hamburger, D. Kaysen, and Others. Quantitative magnetic resonance imaging of human brain development: ages 4–18. *Cerebral cortex*, 6(4):551–559, 1996.
- [36] M. Grant, S. Boyd, M. Grant, S. Boyd, V. Blondel, S. Boyd, and H. Kimura. Cvx: Matlab software for disciplined convex programming, version 2.1. *Recent Advances in Learning and Control*, pages 95–110.
- [37] R. Gribonval and M. Nielsen. Sparse representations in unions of bases. *IEEE transactions on Information theory*, 49(12):3320–3325, 2003.
- [38] R. N. Gunn, S. R. Gunn, F. E. Turkheimer, J. A. Aston, and V. J. Cunningham. Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. *Journal of Cerebral Blood Flow & Metabolism*, 22(12):1425–1439, 2002.

- [39] J. I. Hamilton, A. Deshmane, S. Hougen, M. Griswold, and N. Seiberlich. Magnetic Resonance Fingerprinting with Chemical Exchange (MRF-X) for Quantification of Subvoxel T1, T2, Volume Fraction, and Exchange Rate. In *Proc. Intl. Soc. Mag. Reson. Med. 23*, page 329, 2015.
- [40] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany, P. Maguire, D. Rosas, N. Makris, A. Dale, B. Dickerson, and B. Fischl. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, 32(1):180–194, 2006.
- [41] T. Hastie, R. Tibshirani, and M. Wainwright. *Statistical learning with sparsity*. CRC press, 2015.
- [42] P. Heins, M. Moeller, and M. Burger. Locally sparse reconstruction using the  $\ell_{1,\infty}$ -norm. *arXiv preprint arXiv:1405.5908*, 2014.
- [43] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale  $\ell_1$ -regularized least squares. *IEEE journal of selected topics in signal processing*, 1(4):606–617, 2007.
- [44] C. Laule, E. Leung, D. K. Li, A. L. Traboulsee, D. W. Paty, A. L. MacKay, and G. R. Moore. Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. *Multiple Sclerosis Journal*, 12(6):747–753, 2006.
- [45] C. Laule, I. M. Vavasour, G. R. W. Moore, J. Oger, D. K. B. Li, D. W. Paty, and A. L. MacKay. Water content and myelin water fraction in multiple sclerosis: A T2 relaxation study. *Journal of Neurology*, 251(3):284–293, 2004.
- [46] S. Ljunggren. A simple graphical representation of Fourier-based imaging methods. *Journal of Magnetic Resonance (1969)*, 54(2):338–343, 1983.
- [47] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed Sensing {MRI}. *IEEE Signal Processing Magazine*, 25(2):72–82, mar 2008.
- [48] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold. Magnetic resonance fingerprinting. *Nature*, 495(7440):187–192, mar 2013.
- [49] J. Mairal, F. Bach, and J. Ponce. Sparse modeling for image and vision processing. *Foundations and Trends® in Computer Graphics and Vision*, 8(2-3):85–283, 2014.
- [50] D. Malioutov, M. Çetin, and A. S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE transactions on signal processing*, 53(8):3010–3022, 2005.
- [51] S. Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.
- [52] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on signal processing*, 41(12):3397–3415, 1993.
- [53] S. J. Matzat, J. van Tiel, G. E. Gold, and E. H. Oei. Quantitative MRI techniques of cartilage composition. *Quantitative imaging in medicine and surgery*, 3(3):162, 2013.
- [54] D. McGivney, A. Deshmane, Y. Jiang, D. Ma, C. Badve, A. Sloan, V. Gulani, and M. Griswold. Bayesian estimation of multicomponent relaxation parameters in magnetic resonance fingerprinting. *Magnetic Resonance in Medicine*, 2017.
- [55] D. Needell and J. A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.
- [56] M. Neema, J. Stankiewicz, A. Arora, V. S. Dandamudi, C. E. Batt, Z. D. Guss, A. Al-Sabbagh, and R. Bakshi. T1-and T2-based MRI measures of diffuse gray matter and white matter damage in patients with multiple sclerosis. *Journal of Neuroimaging*, 17(s1), 2007.
- [57] D. G. Nishimura. *Principles of magnetic resonance imaging*. Stanford University, 1996.

- [58] J. Nocedal and S. Wright. *Numerical optimization*. Springer, 2006.
- [59] U. Nöth, E. Hattingen, O. Bähr, J. Tichy, and R. Deichmann. Improved visibility of brain tumors in synthetic MP-RAGE anatomies with pure T1 weighting. *NMR in Biomedicine*, 28(7):818–830, 2015.
- [60] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 40–44. IEEE, 1993.
- [61] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, 98(6):1006–1020, 2010.
- [62] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.
- [63] D. S. Reich, S. A. Smith, K. M. Zackowski, E. M. Gordon-Lipkin, C. K. Jones, J. A. Farrell, S. Mori, P. C. van Zijl, and P. A. Calabresi. Multiparametric magnetic resonance imaging analysis of the corticospinal tract in multiple sclerosis. *Neuroimage*, 38(2):271–279, 2007.
- [64] J. R. Shewchuk. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*. 1994.
- [65] C. Silva, J. Maltez, E. Trindade, A. Arriaga, and E. Ducla-Soares. Evaluation of l1 and l2 minimum norm performances on EEG localizations. *Clinical neurophysiology*, 115(7):1657–1668, 2004.
- [66] D. K. Sodickson and W. J. Manning. Simultaneous Acquisition of Spatial Harmonics (SMASH): Fast Imaging with Radiofrequency Coil Arrays. *Magnetic Resonance in Medicine*, 38(4):591–603, 1997.
- [67] R. L. Suddath, M. F. Casanova, T. E. Goldberg, D. G. Daniel, J. R. Kelsoe, D. R. Weinberger, and Others. Temporal lobe pathology in schizophrenia: a quantitative magnetic resonance imaging study. *Am J Psychiatry*, 146(4):464–472, 1989.
- [68] G. Tang, B. N. Bhaskar, and B. Recht. Sparse recovery over continuous dictionaries-just discretize. In *Signals, Systems and Computers, 2013 Asilomar Conference on*, pages 1043–1047. IEEE, 2013.
- [69] G. T. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11):7465–7490, 2013.
- [70] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- [71] P. Tofts. *Quantitative MRI of the brain: measuring changes caused by disease*. John Wiley & Sons, 2005.
- [72] J. Tohka. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review. *World journal of radiology*, 6(11):855, 2014.
- [73] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information theory*, 50(10):2231–2242, 2004.
- [74] D. B. Twieg. The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Medical Physics*, 10(5):610–621, 1983.
- [75] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel. An optimal radial profile order based on the golden ratio for time-resolved MRI. *IEEE Transactions on Medical Imaging*, 26(1):68–76, 2007.
- [76] D. Wipf and S. Nagarajan. Iterative reweighted  $\ell_1$  and  $\ell_2$  methods for finding sparse solutions. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):317–329, 2010.
- [77] D. P. Wipf and B. D. Rao. Sparse Bayesian learning for basis selection. *IEEE Transactions on Signal processing*, 52(8):2153–2164, 2004.

- [78] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan. Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98(6):1031–1044, 2010.
- [79] S. J. Wright. *Primal-dual interior-point methods*. SIAM, 1997.
- [80] P. Xu, Y. Tian, H. Chen, and D. Yao. Lp norm iterative sparse solution for EEG source localization. *IEEE transactions on biomedical engineering*, 54(3):400–409, 2007.