Bayesian uncertainty quantification for magnetic resonance fingerprinting

Selma Metzner1, Gerd Wübbeler1, Sebastian Flassbeck2,3, Constance Gatefait1, Christoph Kolbitsch1 and Clemens Elster1

1 Physikalisch-Technische Bundesanstalt, Abbestraße 2-12, D-10587 Berlin, Germany
2 Department of Radiology, Center for Biomedical Imaging, New York University School of Medicine, New York, United States of America
3 Center for Advanced Imaging Innovation and Research, New York University School of Medicine, New York, United States of America

E-mail: selma.metzner@ptb.de

Keywords: MRF, Bayesian inference, uncertainty

Abstract

Magnetic Resonance Fingerprinting (MRF) is a promising technique for fast quantitative imaging of human tissue. In general, MRF is based on a sequence of highly undersampled MR images which are analyzed with a pre-computed dictionary. MRF provides valuable diagnostic parameters such as the $T_1$ and $T_2$ MR relaxation times. However, uncertainty characterization of dictionary-based MRF estimates for $T_1$ and $T_2$ has not been achieved so far, which makes it challenging to assess if observed differences in these estimates are significant and may indicate pathological changes of the underlying tissue. We propose a Bayesian approach for the uncertainty quantification of dictionary-based MRF which leads to probability distributions for $T_1$ and $T_2$ in every voxel. The distributions can be used to make probability statements about the relaxation times, and to assign uncertainties to their dictionary-based MRF estimates. All uncertainty calculations are based on the pre-computed dictionary and the observed sequence of undersampled MR images, and they can be calculated in short time. The approach is explored by analyzing MRF measurements of a phantom consisting of several tubes across which MR relaxation times are constant. The proposed uncertainty quantification is quantitatively consistent with the observed within-tube variability of estimated relaxation times. Furthermore, calculated uncertainties are shown to characterize well observed differences between the MRF estimates and the results obtained from high-accurate reference measurements. These findings indicate that a reliable uncertainty quantification is achieved. We also present results for simulated MRF data and an uncertainty quantification for an in vivo MRF measurement. MATLAB® source code implementing the proposed approach is made available.

1. Introduction

Magnetic resonance imaging (MRI) (Nishimura 1996) is a non-invasive medical imaging technique. It is applied in clinical routine for diagnosing diseases and monitoring therapies. While standard qualitative MRI produces images that show good contrasts, it does not yield quantitative results for tissue-related parameters like the spin relaxation times. It has been shown that such quantitative information could be beneficial for detecting diseases (Tofts 2005). Nevertheless, many methods for quantitative MRI (qMRI) (Schmitt et al 2004, Warntjes et al 2007, 2008), require long acquisition times which can be difficult to achieve in clinical practice.

In the past few years, Magnetic Resonance Fingerprinting (MRF) (Ma et al 2013) has evolved into a promising and fast method for qMRI. The main idea is to acquire and analyze a sequence of images from data that are highly undersampled in the spatial Fourier domain. Application of a pseudo-inverse Fourier transformation such as Greengard and Lee (2004) yields a sequence of images with undersampling (aliasing) artifacts. From these images individual time series of the dynamic magnetization are extracted for each voxel and
subsequently correlated with a dictionary. The dictionary contains time series for different tissue types calculated for the applied pulse sequence using a physical model. For each voxel tissue-related parameters ($T_1$ and $T_2$ relaxation times) are then determined as those parameters of the physical model for which the corresponding dictionary entry shows highest correlation with the time series obtained from the measured data. While MRF has originally been used for qMRI of the brain (Ma et al. 2013), recent applications include the measurement of blood flow (Flassebeck et al. 2019), cerebral blood volume (Christen et al. 2014) and cardiac tissue characterization (Hamilton et al. 2017).

Several improvements have been proposed for MRF recently. For example, to accelerate the dictionary matching method a singular value decomposition of the dictionary has been used in McGivney et al. (2014). In Hoppe et al. (2017), Virtue et al. (2017), Balsiger et al. (2018), Cohen et al. (2018), Fang et al. (2019) machine learning has been applied to overcome the discrete nature of the dictionary. Another route of development has been to adopt statistical approaches that fit the MRF data directly in the Fourier domain (Zhao et al. 2016, Wübbeler and Elster 2017, Metzner et al. 2019). In this way aliasing effects caused by the pseudo-inverse Fourier transform used in the dictionary matching method are avoided. The drawback, however, is that fitting the data in the Fourier domain leads to a non-convex, large-scale optimization task since the tissue parameters of all voxels need to be adjusted at the same time. This optimization task is challenging and it is very time-consuming to solve. Furthermore, good starting values have to be assigned in order to reach the correct solution.

The dictionary matching method usually restricts itself to the calculation of estimates for the tissue parameters. As part of the matching process the correlation between the measured course of magnetization at a voxel and all entries of the chosen dictionary is calculated. The maximum correlation provides a measure of the quality of the fit. However, it does not yield an uncertainty of the tissue-related parameters. Uncertainties that reliably characterize the accuracy of the tissue-related parameter estimates (Polders et al. 2012) are essential to assess the significance of observed differences in single qMRI results, which is particularly relevant in applications such as therapy monitoring. While the statistical approaches (Zhao et al. 2016, Wübbeler and Elster 2017, Metzner et al. 2019) that fit the data directly in the Fourier domain allow for such an uncertainty quantification, their practical applicability is challenged due to the involved large-scale optimization tasks.

The goal of this paper is to develop an uncertainty characterization of the estimates obtained by the dictionary matching method and to demonstrate its practical feasibility using MRF phantom measurements. The approach applies Bayesian statistics (Gelman et al. 2013) and yields posterior probability distributions for the tissue-related parameters at all voxels. These probability distributions allow probability statements about the tissue-related parameters to be made conditional on the observed data, and to quantify the uncertainties associated with the estimates of the tissue-related parameters obtained by the dictionary matching method. The proposed Bayesian approach employs so-called noninformative priors (Kass and Wasserman 1996) combined with an informative prior for the relaxation times that utilizes as prior knowledge the ranges chosen for the relaxation times when calculating the dictionary. The calculation of results is facilitated through the analytical derivation of the marginal posterior for the relaxation times and can be carried out using numerical integration. Since these calculations can effectively be carried out using the pre-computed dictionary, they are also fast.

The practicability of the approach is validated by analyzing MRF measurements of a phantom that consists of several tubes. Relaxation times can be assumed constant within each tube, which allows for a consistency check by comparing the within-tube variability of their estimates obtained by the dictionary matching method with the proposed uncertainty quantification for these estimates. In addition, the estimates determined by the dictionary matching method are compared with estimates obtained by reference measurements using long acquisition times in terms of the proposed uncertainty quantification for the former. Validity of the proposed uncertainty evaluation for MRF is also assessed in terms of simulated data with known ground truth, and applicability to an in vivo measurement is demonstrated as well.

The proposed approach for uncertainty estimation in MRF fingerprinting is fast and can be applied in combination with any dictionary matching method. MATLAB® source code will be shared to facilitate its application (Metzner and Wübbeler 2021).

The paper is organized as follows. Section 2 briefly describes the structure of MRF data and introduces the dictionary matching method. Then the proposed Bayesian uncertainty quantification is developed. In section 3 the experiment is described for recording MRF data. Results achieved by the proposed Bayesian uncertainty quantification for numerical simulations and phantom experiments are then presented in section 4. The results obtained for the phantom measurements are also compared with additional reference measurements of the relaxation times using long acquisition times. In addition, results for an in vivo measurement are presented. In section 5 then a discussion is provided and some conclusions are finally given in section 6.
2. Methods

This section briefly outlines the structure of MRF data and the dictionary matching method. Then the proposed Bayesian approach is developed. A description of the physical model used for compiling the dictionary as well as mathematical details of the Bayesian approach are given in the appendices.

2.1. MRF data

In MRI nuclear spins are excited in a selected slice by applying a high frequency pulse in the presence of a static magnetic field. The acquired data are samples in the Fourier domain of the image. The measured magnetization depends on the tissue parameters, namely relaxation times $T_1, T_2$ and proton spin density $\rho$, as well as on the MR sequence parameters (e.g. echo time, repetition time, RF-pulse design). In MRF a sequence of pulses with varying flip angles is applied, and measurements are conducted between subsequent pulses. These measurements do not cover the whole Fourier domain and can thus be carried out fast. The dynamics of the magnetization depends on the whole sequence of pulses and can be modeled through a three-dimensional linear discrete time system, see appendix A. Relevant parameters that are controlled in the design of the sequence of pulses are flip angles, repetition times and echo times.

The dependency of the dynamics of the magnetization on the flip angles and repetition times is not indicated in the following. To keep notations simple, we also suppress the presence of measurements using multiple receiver coils (as used in our experiments) and neglect the spatially varying received sensitivities of the multiple receiver coils in the following description. This simplification does not affect our treatment and complete relationships are given in appendix A.

Let $N$ denote the number of considered voxels in the selected slice, and $\theta_1 = (T_1^1, \ldots, T_1^N)^T$, $\theta_2 = (T_2^1, \ldots, T_2^N)^T$, $\theta_3 = (\rho^1, \ldots, \rho^N)^T$ the sought vectors of relaxation times $T_1, T_2$ and proton spin density $\rho$, respectively, for $N$ voxels. As is common in MRI, we model the proton spin density in terms of a complex number which can account for phase changes in the measured data (e.g. due to susceptibility artefacts). Let $m_l = m_l(\theta_1, \theta_2, \theta_3), l = 1, \ldots, L$ denote the $N \times 1$ complex-valued vector of magnetizations at the $N$ voxels after applying the $l$th excitation pulse, and let $m^i$ denote the $L \times 1$ complex-valued vector of magnetizations at the $i$th voxel after applying all $L$ pulses, see figure 1. The sequence of magnetizations at the $i$th voxel, $m^i$, depends only on the unknown tissue parameters at that voxel (and the known parameters of the applied pulse sequence), i.e. $m^i = m^i(T_1^i, T_2^i, \rho^i)$. Furthermore, the model $m_l = m_l(\theta_1, \theta_2, \theta_3)$ depends linearly on $\theta_3$ (and hence $m^i$ also depends linearly on $\rho^i$), which will be utilized in the proposed Bayesian treatment.

The measured data consist of the complex-valued vectors

$$z_l = F_l(m_i), \quad l = 1, \ldots, L.$$  

(2.1)

$F_l(m)$ denotes the $n_l \times 1$ vector of measurements in the Fourier domain after the $l$th pulse, where typically $n_l \ll N$ holds. Hence, the magnetization $m_l$ cannot be reconstructed exactly. The mapping $F_l$ is indexed by $l$ as the locations of measurements in Fourier domain are generally different for the different pulses. Note that each entry of $z_l$ depends on the tissue parameters of all voxels.

![Figure 1. Concept of MRF: sequence of images which are obtained by applying a pseudo-inverse to the undersampled measurements in Fourier space (left) and temporal development of magnetization at a single voxel (blue, right). The red curve corresponds to the matched dictionary entry, the so-called ‘fingerprint’. Tissue parameters at each voxel are extracted from the fingerprint.](image)
2.2. Dictionary matching MRF
The principles of the dictionary matching method according to Ma et al (2013) are briefly described in the following. In a first step, the magnetizations $m_l$, $l = 1, \ldots , L$, are approximately reconstructed from (2.1) through
\begin{equation}
\tilde{m}_l = F^*_l (z_l), \quad l = 1, \ldots , L,
\end{equation}
where $F^*_l$ denotes a pseudo-inverse of the Fourier transform mapping $F_l$ in (2.1). From the sequence of estimated magnetizations, $\tilde{m}_1, \ldots, \tilde{m}_L$, the estimated course of magnetization $\tilde{m}^i$ for each voxel $i$ is then obtained and subsequently denoted by
\begin{equation}
\tilde{m}^i = (((\tilde{m}_1^i), \ldots, (\tilde{m}_L^i))^T,
\end{equation}
see also figure 1. These estimates are generally not accurate as they contain aliasing errors caused by the imperfect inversion $F^*_l$. In our calculations, we employed the pseudo-inverse Fourier transformation from Greengard and Lee (2004). The basic idea of MRF is to choose flip angles and repetition times of the pulse sequence randomly in such a way that the sequence
\begin{equation}
e^i = (((\tilde{m}^i - m^i)_1, (\tilde{m}^i - m^i)_2, \ldots, (\tilde{m}^i - m^i)_L)^T
\end{equation}
of (complex-valued) temporal errors in the reconstruction of the magnetizations (2.2) for each voxel $i$ can be modeled as noise. The reconstruction errors $e^i$ are usually dominated by aliasing effects caused by the imperfect inverse Fourier mapping, and these effects depend on the course of the magnetization itself and the undersampling scheme which both are determined through the choice of the pulse sequence parameters. Since the magnetization at each voxel depends on the tissue parameters of only that voxel, estimates of those tissue parameters can then be obtained by fitting the modeled dynamics of the magnetization of each voxel to the observed approximate dynamics (2.3) at that voxel. In Ma et al (2013) this fitting process is done by pre-computing a dictionary of (normalized) courses of magnetizations on a discrete set of possible values of the sought relaxation times and then selecting that dictionary entry which shows maximum correlation to the observed dynamics (2.2). The spin density, which enters linearly into the model for the magnetization, is
\begin{equation}
\text{given by minimizing}
\begin{equation}
\| \tilde{m}^i - m^i(T_{T_1}^i, T_{T_2}^i, \rho^i) \|^2
\end{equation}
with respect to $T_{T_1}^i, T_{T_2}^i, \rho^i$, provided that the resolution of the dictionary is made arbitrarily fine; in other words, the dictionary matching method basically fits the modeled dynamics of the magnetization independently at each voxel to the observed rough estimates (2.3) at that voxel. In (2.5) $\cdot \cdot$ stands for the usual 2-norm in $\mathbb{C}^L$.  

2.3. Bayesian inference for dictionary matching method
A Bayesian inference is carried out to quantify the uncertainty associated with the tissue-related parameters calculated by the dictionary matching method. The Bayesian inference is based on a statistical model for the observed magnetizations in (2.3) and the choice of prior distributions. In order to improve readability, some additional notations are introduced first.

2.3.1. Notations
Let the $2L \times 1$ real-valued vector $\tilde{y}^i$ denote real part and imaginary part of the observed estimates $\tilde{m}^i$ of the course of magnetizations at voxel $i$, i.e.
\begin{equation}
\tilde{y}^i = (\Re (\tilde{m}^i)^T, \Im (\tilde{m}^i)^T)^T,
\end{equation}
and the $2 \times 1$ real-valued vector $\mu^i$ the corresponding proton spin density
\begin{equation}
\mu^i = (\Re (\rho^i)^T, \Im (\rho^i)^T)^T,
\end{equation}
where $\Re (\cdot)$ and $\Im (\cdot)$ stand for real and imaginary part, respectively. Furthermore, let the $2L \times 2$ real-valued matrix $M^i = M^i(T_{T_1}^i, T_{T_2}^i)$ be defined such that
\begin{equation}
\mu^i(T_{T_1}^i, T_{T_2}^i, \rho^i) = M^i \mu^i
\end{equation}
contains real and imaginary part of the modeled course of magnetizations $m^i(T_{T_1}^i, T_{T_2}^i, \rho^i)$ at voxel $i$, see appendix A for details.

In accordance with the dictionary matching method the subsequent statistical analysis is carried out independently at each voxel, and we will hence omit the reference to a particular voxel $i$ in the following, e.g. we use $M = M(T_{T_1}, T_{T_2})$ to denote $M^i = M(T_{T_1}^i, T_{T_2}^i)$ for some considered voxel $i$.

Probability distributions will be loosely denoted through their arguments, e.g. $\pi(T_{T_1}, T_{T_2}, \mu)$ stands for a probability density function of $T_{T_1}, T_{T_2}$ and $\mu$.  

4
2.3.2. Statistical model
Following the rationale of the dictionary matching method, the errors in the observed course of magnetization in (2.3) are modeled as independently and identically distributed with zero mean and variance $\sigma^2$. In addition, we assume normality of this distribution. More precisely, we assume

$$ y \sim N(M\mu, \sigma^2I_2), $$

where $M = M(T_1, T_2)$ depends on $T_1$ and $T_2$ only, $\mu$ denotes real and imaginary part of the spin density $\rho$, and $I_2$ stands for the identity matrix of dimension 2L, i.e. the observed data $y$ are modeled as a realization of a Gaussian distribution with mean vector $M\mu$ and covariance matrix $\sigma^2I_2$. Model (2.9) is a homoscedastic Gaussian model with a mean vector that depends linearly on the spin density $\mu$ and nonlinearly on the relaxation times $T_1$ and $T_2$. Note that the variance $\sigma^2$ in (2.9) is allowed to differ for different voxels.

2.3.3. Bayesian inference
A Bayesian inference combines the prior knowledge about the unknowns with the information contained in the data. The result is the posterior probability distribution which is obtained via Bayes theorem according to

$$\pi(T_1, T_2, \mu, \sigma^2 | y) \propto \pi(T_1, T_2, \mu, \sigma^2) l(T_1, T_2, \mu, \sigma^2; y).$$

(2.10)

$\pi(T_1, T_2, \mu, \sigma^2)$ denotes the prior beliefs about the unknowns $T_1$, $T_2$, $\mu$, $\sigma^2$, and $l(T_1, T_2, \mu, \sigma^2; y)$ is the likelihood function which is determined through the statistical model (2.9) for the data.

As a prior distribution we choose

$$\pi(T_1, T_2, \mu, \sigma^2) = \pi(T_1, T_2) \pi(\mu) \pi(\sigma^2) \propto I_\Omega(T_1, T_2) / \sigma^2,$$

(2.11)

where $\Omega$ denotes the domain for $T_1$ and $T_2$ in the dictionary, and the function $I_\Omega(T_1, T_2)$ equals 1 for $(T_1, T_2) \in \Omega$ and 0 otherwise. The prior $\pi(T_1, T_2) \propto I_\Omega(T_1, T_2)$ thus represents the prior belief about the range of possible values for $T_1$ and $T_2$, without a priori preferring any single value from this range. The spin density acts as a location parameter in the likelihood, and the prior $\pi(\mu) \propto 1$ is a standard noninformative prior in such a case, see Kass and Wasserman (1996). Similarly, $\pi(\sigma^2) \propto 1/\sigma^2$ is a standard noninformative prior for the scale parameter $\sigma^2$. Hence, we combine the available prior knowledge about the relaxation times with standard noninformative priors for the proton spin density and for the variance $\sigma^2$ in (2.9). Note that albeit the prior (2.11) is improper, i.e. cannot be normalized, propriety of the posterior (2.10) can be ensured, see appendix B.

From the joint posterior $\pi(T_1, T_2, \mu, \sigma^2 | y)$ in (2.10) marginal posterior distributions such as, e.g. $\pi(T_1 | y)$ are obtained by integrating out the remaining unknowns. Once the marginal posterior for one of the tissue parameters has been determined, the uncertainty $u(T_1)$ associated with the estimate $\hat{T}_1$ obtained by the dictionary matching method can be calculated as

$$u^2(\hat{T}_1) = \int\pi(T_1 | y)(T_1 - \hat{T}_1)^2 dT_1,$$

(2.12)

In addition, a 95% credible interval $I$ can be determined which satisfies

$$0.95 = \int_I \pi(T_1 | y) dT_1.$$

(2.13)

Relation (2.13) does not uniquely determine an interval $I$, and an additional condition needs to be posed, for example the interval of minimal length satisfying (2.13) or the choice of a probabilistically symmetric interval. One can also use the marginal posterior to make probability statements, for example to determine the probability that $T_1$ exceeds a certain limit, $\bar{T}_1$ say, according to

$$P(T_1 > \bar{T}_1 | y) = \int_{\bar{T}_1}^\infty \pi(T_1 | y) dT_1.$$

(2.14)

In appendix B we derive the marginal joint posterior for $T_1, T_2$ as

$$\pi(T_1, T_2 | y) \propto \frac{1}{\sqrt{\det(M^TM)}} \times \frac{I_\Omega(T_1, T_2)}{\|y - M\mu(T_1, T_2)\|^2} \times \sigma^{2L-1},$$

(2.15)

with

$$\mu(T_1, T_2) = [M^TM]^{-1}M^Ty$$

(2.16)

and $M = M(T_1, T_2)$. Furthermore,

$$\mu(T_1, T_2, y) \sim t_{2L-2}\left(\mu(T_1, T_2), \frac{\|y - M\mu(T_1, T_2)\|^2}{2L - 2} [M^TM]^{-1}\right).$$

(2.17)

i.e. the conditional posterior distribution for $\mu$ given $T_1, T_2$ is a scaled and shifted 2-dimensional $t$-distribution with $2L - 2$ degrees of freedom. Since $L \gg 1$, a very good approximation is given by the Gaussian distribution.
\[ \mu | T_1, T_2, y \sim \mathcal{N} \left( \mu (T_1, T_2), \frac{||y - M\mu (T_1, T_2)||^2}{2L - 2} [M^T M]^{-1} \right) \]  

(2.18)

where, again, \( M = M(T_1, T_2) \). The marginal posterior for \( \mu \) is then obtained as

\[
\pi (\mu | y) = \int \pi (\mu | T_1, T_2, y) \pi (T_1, T_2 | y) dT_1 dT_2.
\]  

(2.19)

The \( 2 \times 2 \) covariance matrix \( U (\bar{\mu}) \) associated with the estimate \( \bar{\mu} \) obtained by the dictionary matching method is simply given through

\[
U (\bar{\mu}) = \int \left\{ (\bar{\mu} - \mu (T_1, T_2)) (\bar{\mu} - \mu (T_1, T_2))^T + \frac{||y - M\mu (T_1, T_2)||^2}{2L - 2} [M^T M]^{-1} \right\} \pi (T_1, T_2 | y) dT_1 dT_2.
\]  

(2.20)

where again \( M = M(T_1, T_2) \).

Expressions (2.15), (2.18) and (2.20) allow for a practical calculation of results using 2-d approximate numerical integration, which is most efficiently carried out on the dictionary itself. Note that the columns of matrix \( M \) can be pre-calculated and also stored in the dictionary for the chosen set of values for \( T_1, T_2 \). Once the marginal posterior distributions for \( T_1, T_2, \mu \) have been determined, uncertainties such as (2.12), or probability statements as in (2.14) can be determined again using numerical integration. Details are given in appendix C.

The shared MATLAB® source code (Metzner and Wübbeler 2021) provides the proposed Bayesian uncertainty quantification using numerical integration. The required input consists of the employed dictionary, together with the observed sequence of magnetizations in (2.3) for each voxel as well as estimates for \( T_1 \) and \( T_2 \). Note that it is generally possible to use other estimates than the dictionary-based MRF estimates. Its output are the discretized marginal posteriors for \( T_1 \) and \( T_2 \) and the uncertainties for \( T_1 \) and \( T_2 \) calculated according to (2.12). Due to the use of the pre-computed dictionary, calculations are carried out fast. The computation time depends on the size of the dictionary and on the number of voxels of interest. The calculation time of the uncertainties shown in this paper was about 90 minutes, whereas calculating the estimates by the dictionary matching method took about 6 minutes. When the size of the dictionary is reduced by a factor of 4, also the computation time for the uncertainties is reduced by a factor of 4.

3. Experiments and simulations

The phantom data was acquired on a 3T Verio MR Scanner (Siemens Healthineers, Erlangen, Germany) with a Fast Imaging with Steady State Precession (FISP) sequence starting with an inversion pulse using a radial readout, each with 384 data points per radial line. Repetition times \((TR)\) were chosen to be 12.54 ms, the inversion time \((TI)\) as 42 ms, echo times \((TE)\) as 7 ms, and the flip angle follows the magnitude of a sinusoidal curve with maximum 70 degree as proposed in Flasbeck et al (2019). A total number of 1000 radial lines were acquired and for the parameter estimation one image per radial line was reconstructed using the pseudo-inverse Fourier transformation (Greengard and Lee 2004). Figure 2 shows two reconstructed magnetization images and the sum of all the reconstructions. Data were acquired with an 8-channel head coil, and the coil sensitivity maps required for image reconstruction were obtained from the data themselves. The inplane resolution of the MRF scan was 0.78 \( \times \) 0.78 mm\(^2\) with a slice thickness of 5.0 mm and a field of view (FoV) of 300 \( \times \) 300 mm\(^2\). The calculation of the dictionary was carried out using the Bloch equations (see appendix A) assuming 201 isochromats in each voxel.
Reference values for $T_1$ were determined by a Cartesian inversion recovery spin echo sequence acquisition with 7 inversion times ($TI = 25–2400\,ms, TE/TR = 12/8000\,ms$). Reference values for $T_2$ were determined using a Cartesian spin echo sequence with multiple echo times ($TE = 24–1000\,ms, TR = 3000\,ms$). For both acquisitions, inplane resolution was $0.83 \times 0.83\,mm^2$ with a slice thickness of $5.0\,mm$ and a FoV of $159 \times 144\,mm^2$. A 2-parameter nonlinear least squares fitting algorithm was used to determine the reference $T_1$ and $T_2$ values.

The simulation closely follows the phantom measurement. All scanning parameters ($TR, TE, TI, \text{flip angles}$) and $k$-space locations for the radial points are set to the values chosen for the phantom measurements. Again, a total number of 1000 radial lines were used, and for the dictionary matching method parameter estimation was carried out using one image per radial line. The same complex coil sensitivity maps were chosen as for the phantom scan. We assigned values for $T_1$, $T_2$ and the complex proton density that match those obtained in the analysis of the phantom data. Magnetization during the simulated data acquisition was modeled by the same FISP model used for the calculation of the dictionary (see appendix A). The modeled magnetization was then transformed into the Fourier domain using Fourier transformation (Greengard and Lee 2004), and Gaussian noise with similar signal-to-noise ratio as for the real scanner data was added with a standard deviation of $3.43 \times 10^{-6}$ (the maximum value of the real and the imaginary part of the simulated magnetizations are of the order $10^{-5}$).

The in vivo data acquisition was performed on a 3T Verio MR Scanner (Siemens Healthineers, Erlangen, Germany) using a 32-channel cardiac coil (Siemens Healthineers, Erlangen, Germany) and the coil sensitivity maps required for image reconstruction were obtained from the data themselves. A cardiac scan was realized in a short-axis orientation utilizing a FISP sequence with a total acquisition time of 12 s (allowing for an acquisition during a single breathhold). A radial readout was applied, each with 320 data points per radial line. Repetition times ($TR$) were chosen to be 8.2 ms, the inversion time ($TI$) as 21 ms, echo times ($TE$) as 4 ms, and the flip angle follows the magnitude of a sinusoidal curve with maximum 70 degree. A total number of 1500 radial lines were acquired and for the parameter estimation one image per radial line was reconstructed using the pseudo-inverse Fourier transformation. The resolution of the MRF scan was $1.3 \times 1.3\,mm^2$ with a FOV of $320 \times 320\,mm^2$ and a slice thickness of 8 mm. To reduce the impact of the slice profile, a bandwidth time product of 8 was chosen (SINC pulse length $= 1920\,ms$). ECG triggering was used to start the acquisition during diastole, but data were then acquired continuously without further triggering.

4. Results

Real and simulated MRF measurements of the phantom (see figure 2) were conducted and analyzed by the proposed uncertainty quantification for the dictionary matching method. In addition, reference measurements for $T_1$ and $T_2$ were carried out to assess these results. The dictionary was calculated according to the physical model given in appendix A for $T_1 = (200, 205, \ldots, 3000)\,ms$ and $T_2 = (20, 22, \ldots, 300)\,ms$. The phantom consists of 9 different tubes and the values for $T_1$ and $T_2$ within each tube can be assumed to be constant. The variation of $T_1$ and $T_2$ between the tubes almost covers their range in the chosen dictionary, see figure 4.

4.1. Phantom and in vivo measurements

Figure 3 illustrates the (estimated) residuals in (2.4) for a typical voxel. The residuals largely follow a Gaussian distribution, supporting the statistical model chosen for the Bayesian uncertainty quantification. In figure 4 calculated uncertainties are compared with the observed variation of estimated values for $T_1$ and $T_2$ within each tube. More precisely, the figure shows mean values of the estimates obtained by the dictionary matching method for $T_1$ and $T_2$ within the different tubes; in addition, standard deviations of the estimates within the tubes and within-tube means of uncertainties are displayed for the different tubes. Mean uncertainties reflect the within-tube variability of the estimates well, yet they are slightly larger than the observed variability of the estimates.

In order to analyze the impact of the range of the values in the dictionary we also evaluated the phantom data with a larger and a smaller dictionary in $T_2$. When taking a larger dictionary with $T_2 = (20, 22, \ldots, 500)\,ms$, the within-tube standard deviations of the estimates differed from those obtained for the original dictionary by at most 10%. This is also true for the mean uncertainties, except for tube nine in which the mean uncertainty is 50% higher. Results for a smaller dictionary with $T_2 = (20, 22, \ldots, 100)\,ms$ are shown in figure 5. In this case the $T_2$ estimates can no longer cover the whole range of the underlying $T_2$ values, and for tubes with a higher $T_2$ than allowed by the dictionary the corresponding estimates equal the maximum value of $T_1$ in the dictionary. Note that the resulting uncertainties indicate that the true $T_2$ values are beyond the upper bound induced by the (restricted) dictionary.
Figure 3. Results from phantom data. Left: residuals of real (top) and imaginary (bottom) part of the magnetization in one voxel obtained by the dictionary matching method for the phantom data. Right: histogram of all residuals on the left.

Figure 4. Mean values for the estimates obtained by the dictionary matching method within the tubes \( t_1 \), \( t_9 \) for \( T_1 \) (left) and \( T_2 \) (right), together with their standard deviation (blue errorbars) and mean uncertainty (red errorbars). Note that the errorbars are in ascending order in terms of the mean values for the estimates, respectively for \( T_1 \) and \( T_2 \).

Figure 5. Results for a dictionary with a smaller range: mean values for the estimates obtained by the dictionary matching method within the tubes \( t_1 \), \( t_9 \) for \( T_1 \) (left) and \( T_2 \) (right), together with their standard deviation (blue errorbars) and mean uncertainty (red errorbars). Note that the errorbars are in ascending order in terms of the mean values for the estimates, respectively for \( T_1 \) and \( T_2 \).
In figure 6 the same quantities are plotted as in figure 4, but for data where 5 radial lines in $k$-space are acquired per image. That means, the amount of data is increased by a factor of 5. It can be seen that the mean uncertainties and the standard deviation of the estimates in each tube are smaller now (roughly a factor 5). The ratio of mean uncertainties and standard deviations of the estimates remains unchanged.

Figure 7 compares the reference measurements with the results obtained by the dictionary matching method in terms of the uncertainties calculated for the latter. For $T_1$ estimation the calculated uncertainties appear to describe well the deviations between results of the dictionary matching method and the reference measurements. Regarding $T_2$ estimation this is true for the smaller values of $T_2$, while for larger values of $T_2$ a slight bias in the estimates produced by the dictionary matching method can be observed. Yet, the calculated uncertainties still characterize the size of the errors well in most cases.

In figure 8 marginal distributions for $T_1$ and $T_2$ in two voxels of different tubes are shown. The dotted red lines illustrate the bounds for $T_1$ and $T_2$. Especially for $T_2$ some of the distributions are truncated at these
bounds, i.e. the bounds are truly informative. The derived marginal distributions can be used, for example, to calculate the probability that $T_1$ is smaller than a certain limit. In figure 8 the probabilities for $T_1 < 1600$ ms of the two voxels indicated by the arrows are 0.54 and almost 1, respectively.

Estimates for $T_1$ and $T_2$ as well as their uncertainties of a cardiac in vivo measurement are shown in figure 9. Note that in some areas the uncertainty increases towards the centre of the body, which is reasonable since the coil sensitivity is lower there.

4.2. Simulated MRF

In figure 10 calculated uncertainties are compared with the observed variation of the estimates for $T_1$ and $T_2$ within each tube. The uncertainties reflect the different within tube variabilities of the estimates well. Similarly to the results for the real MRF measurements the calculated uncertainties appear to be slightly conservative.

For higher $T_1$ and $T_2$ values the associated uncertainty is also higher. Nonetheless, the relative uncertainty is roughly constant. We simulated further MRF data by using the same simulation setup, changing only repetition times ($TR$ was set to 11 ms) or flip angle (the maximum of the magnitude of the sinusoidal curve was set to 60 degree). In both cases, the relative uncertainties were again similar, the mean uncertainties in all nine tubes differed by a maximum of 10% when using the different repetition times, and by 20% when applying the different flip angle.

Figure 11 compares the estimates obtained by the dictionary matching method with the known ground truth in terms of the uncertainties calculated for the former. Here only the results for one tube are shown which are representative also for the other tubes. Calculated uncertainties describe well the size of the observed errors in
the estimates obtained by the dictionary matching method for both $T_1$ and $T_2$. Again, the calculated uncertainties appear to be slightly conservative.

5. Discussion

The residuals in figure 3 obtained for the course of the reconstructed magnetizations from the phantom data in a single voxel appear to largely follow a Gaussian distribution. This finding supports the statistical model assumed for the Bayesian uncertainty quantification. Note that the results shown in figure 3 are typical and residuals in other voxels and other tubes look similarly. In figure 2 two single reconstructed magnetization images are shown
with severe aliasing artifacts. When looking at the sum of all these individual reconstructions, the aliasing artifacts appear to be canceled out which supports the assumption that aliasing artifacts act like noise over time.

The simulated MRF data were constructed by adding in $k$-space Gaussian noise of realistic size. When carrying out the simulations without this additional measurement noise, results for the estimates of $T_1$ and $T_2$ essentially remain the same. This implies that the (unavoidable) aliasing errors in the course of the observed magnetizations at each voxel are the dominant uncertainty source. We have shown that the dictionary matching method is equivalent to fitting the modeled magnetizations to the observed ones, which can be recommended only under the assumption that the (aliasing) errors appear like random noise, ideally following a Gaussian distribution. The statistical model employed for the Bayesian uncertainty quantification is based on the very same assumption. Hence, when the design of the MRF data acquisition has been chosen properly to meet the assumptions underlying the dictionary matching method, one can be confident that the assumptions of the proposed Bayesian uncertainty quantification are met as well.

The squared uncertainty has been taken as the expected quadratic loss of the estimate obtained by the dictionary matching method under the derived posterior. It describes the a posteriori belief about the size of the errors. We observed that the within-tube variances of the estimates are slightly smaller than these uncertainties for both the real MRF data (figure 4) and the simulated MRF data (figure 10). If the within-tube variances can be taken as a reference for the size of the errors, then this size is slightly overestimated by our Bayesian inference. One possible source of error is the presence of temporal and/or spatial correlations of the residuals. Although the assumption of a Gaussian distribution for the errors in the sequence of magnetization is in line with the assumption underlying the dictionary matching method, the statistical model will generally provide an approximation to the true distribution of errors only, and represents a limitation of the proposed approach. Future work could address more sophisticated statistical models. However, one benefit of the proposed method is that calculations can be carried out by numerical integration using only the dictionary and the observed data in an efficient way, which is hardly possible when more sophisticated statistical models, containing additional unknowns, are employed. Another possible source of error in the proposed approach is the calculation of the Bayesian uncertainty quantification itself. The calculation of integrals such as (2.12) are carried out on the dictionary and might be prone to numerical approximation errors. For the results shown in this paper this error source is not relevant since the results remain essentially unchanged when doubling the resolution of the dictionary. However, basically this is a potential source of error in the uncertainty calculation (as well as in the dictionary matching method itself), especially when dictionaries with low resolution are taken.

While for the simulated data the set of errors in the estimates produced by the dictionary matching method are well characterized by the calculated uncertainties (figure 11), this does not hold for all results obtained for the measured MRF data (figure 7). In the latter case a significant bias can be observed for the estimates obtained by the dictionary matching method for the larger $T_2$ values, albeit the size of single errors of the estimates are still largely captured by the calculated uncertainties. Since the reference measurements for $T_2$ can be trusted, and as their uncertainty is small compared to the observed discrepancies to the dictionary matching method, the results of the dictionary matching method appear to have some bias for large $T_2$ values. The reason likely is some error in the underlying dynamic model used to create the dictionary. The Bayesian uncertainty quantification is carried out assuming the employed physical model to be correct and therefore underrates calculated uncertainties in cases when a significant model error occurs. One reason for an error in the employed physical

![Figure 11. Simulation results: differences of the simulated ground truth $\tilde{T}_1$, $\tilde{T}_2$ and the estimates $\hat{T}_1$, $\hat{T}_2$ obtained by the dictionary matching method for one tube. Errorbars show calculated uncertainties. For clarity, results are shown only for a representative subset of voxels from the selected tube.](image)
model are inhomogeneities of the $B_1$ field. The RF excitation in our physical model is assumed to be perfect (which would only be achieved with an infinite sinc pulse), but this is a simplification which can lead to errors in $T_2$.

The fact that the unavoidable (large) aliasing errors are the dominant uncertainty source for the dictionary matching method makes this method—as well as the proposed uncertainty quantification—robust. That is, as long as other error sources like measurement errors, incorrect $k$-space positioning, or errors in the dynamic model of magnetization, are small compared to the aliasing errors, both the dictionary matching method and the proposed uncertainty quantification can be expected to produce reasonable results.

We used an isochromat simulation with 201 isochromats per voxel to model the magnetization (see appendix A). It is also possible to apply the Extended Phase Graph formalism (Weigel 2015) for which no further parameters such as the number of isochromats are needed. When comparing results from both models, the mean values of the estimates of all nine tubes have a deviation less than 1%, and the same holds for within-tube standard deviations and for mean uncertainties.

The results of the estimates as well as of the uncertainties depend on the range of the dictionary. When taking a dictionary with a larger range of values for the relaxation times, the mean uncertainties (across single tubes) behave similarly compared to the (within-tube) standard deviations of the estimates. When taking a dictionary with a too small range for the relaxation times, biased estimates result and calculated uncertainties can no longer be expected to provide a reasonable characterization of the errors in the estimates. Hence, the range of the dictionary plays an important role both for the results achieved by the dictionary matching method and the proposed uncertainty quantification, and it should therefore be carefully chosen.

In Zhao et al (2018) the Cramér–Rao bound (CRB) was used to characterize the accuracy of estimates obtained in an MRF experiment, and to optimize its protocol (in Zhao et al 2018) the design parameters are chosen as the sequence parameters repetition times and flip angle). The proposed Bayesian uncertainty quantification also depends on the chosen protocol, and it could likewise be used for protocol optimization. Future work could address this feature of our method and compare resulting protocols with those achieved by CRB minimization.

It was also possible to apply the proposed uncertainty calculation to in vivo data. Reasonable results for the estimates of the dictionary matching method as well as for their uncertainties could be achieved.

One important advantage of Bayesian inference is that a probability distribution is obtained. Basically, the whole distribution can serve as an uncertainty quantification. In some cases, a single summary derived from it such as the uncertainty as defined in (2.12) may be sufficient. In other cases, when the distributions deviate significantly from a Gaussian distribution (see the truncated posteriors in figure 8 for example), the whole distribution should be considered. The posterior distribution can be used to calculate probabilities about a hypothesis such as $T_i > \tilde{T}_i$. In figure 12 a probability map of the hypothesis $T_i < 500$ ms is shown for the phantom data. It can be seen that $T_i$ is almost certainly smaller than 500 ms in all voxels of the tube on the bottom right, while the opposite is true for, e.g. the tube on the top left. In other tubes a clear decision cannot be made, which is represented through probabilities for this hypothesis which neither close to 1 nor to 0. We believe that this sort of information can be useful in clinical follow-up measurements when trying to monitor possible changes in the properties of the tissue.

![Figure 12. Probability map for the hypothesis $T_i < 500$ ms. Approximately true values for $T_i$ as obtained from the reference measurement are from left to right and top to bottom: $\tilde{T}_i = (1834, 562, 587, 1432, 427, 435, 1088, 339, 295)$ ms](image-url)
Usually, Bayesian inferences require extensive numerical calculations, typically performed by Markov chain Monte Carlo methods (MCMC) (Robert and Casella 2013). Results achieved by MCMC methods are stochastic in nature, and need some form of convergence checks. An advantage of the proposed approach is that its results are deterministic and can be calculated efficiently. Furthermore, as input only the data (i.e. sequence of magnetization at every voxel) and the dictionary are required, which provides a convenient interface for applications.

6. Conclusions

Reliable quantification of uncertainties in quantitative magnetic imaging is beneficial as it allows a meaningful assessment of single results. A Bayesian uncertainty quantification has been proposed for estimates obtained by a dictionary matching method. The proposed uncertainty quantification has been tested through its application to measurements of a suitably designed phantom. Comparison with additional reference measurements, together with consistency checks, as well as results obtained for simulated data, show that the proposed approach yields a reasonable uncertainty quantification. The proposed uncertainty quantification has also been applied to an in vivo measurement. The Bayesian approach offers the possibility to also derive probability statements which can increase confidence in conclusions drawn from MRF results. While Bayesian inference typically requires extensive calculations, our analytical derivations allowed for a calculation of uncertainties based on the pre-calculated dictionary. As a consequence, uncertainty calculations can be carried out efficiently. We believe that applications of the dictionary-based MRF can benefit from the proposed uncertainty quantification scheme.

Acknowledgments

This work was supported by EMPIR project 18HLT05 QUIERO. This project has received funding from the EMPIR program co-financed by the Participating States and from the European Union’s Horizon 2020 research and innovation program.

Appendix A. MRF data and physical model

The dictionary used for the dictionary matching method is based on a physical model describing the magnetization process of the MRI sequence. We will use the Bloch equations to derive a three dimensional linear system of equation for the employed FISP sequence, for more details we refer to Hargreaves et al. (2001). In the FISP sequence a dephasing gradient is applied in the middle of each repetition time $TR$. The signal is therefore an average of the spins in one voxel. We will simulate this by taking different spins in one voxel. The following Matlab code is used to explain the implemented sequence. The flip angles will be denoted by $\text{flip}$, echo times by $TE$ and the inversion time by $TI$.

```matlab
function Msig = FISP_sequence(T1, T2, TE, TI, TR, flip, Nf)
flip = pi*flip/180; % degree to radians
Nex = length(TR);
phi = linspace(-0.5,0.5,801)*8*pi;
Msig = zeros(1,Nex);
M0 = [zeros(2,Nf);-ones(1,Nf)];
[Ati,Bti] = freeprecess(T1, T1, T2);
M = Ati*M0 + Bti;
on = ones(1,Nf);
for n = 1:Nex;
[At,e,bit] = freeprecess(TR(n),T1,T2);
[M,e,bit] = freeprecess(TR(n)-TE(n),T1,T2);
A = At+*xrot(flip(n));
B = Bt;
M = A*M+B+*on;
Msig(n) = mean(squeeze(M(1,:)+1i*M(2,:)));
M = At+*M+B+*on;
for k = 1:Nf
M(:,k) = xrot(phi(k))*M(:,k);
end
end
function [Afp,Bfp] = freeprecess(T,T1,T2)
```
The matrix $M'$ containing the real and imaginary part of the modeled magnetization course $m'(T_1^i, T_2^i, \rho')$ is introduced in (2.8). The modeled magnetization course is given as $m'(T_1^i, T_2^i, \rho') = (\Re(\rho') + i\Im(\rho'))(\Re(\tilde{m}^i) + i\Im(\tilde{m}^i))$ where $\tilde{m}^i$ is the part of the whole magnetization without the linear factor $\rho^i$.

It follows that the real and imaginary part of $m'$ are $\Re(m'(T_1^i, T_2^i, \rho')) = \Re(\rho')\Re(\tilde{m}^i) - \Im(\rho')\Im(\tilde{m}^i)$ and $\Im(m'(T_1^i, T_2^i, \rho')) = \Im(\rho')\Re(\tilde{m}^i) + \Re(\rho')\Im(\tilde{m}^i)$. $M'$ is defined as

$$(\Re(m'(T_1^i, T_2^i, \rho'))^T, \Im(m'(T_1^i, T_2^i, \rho'))^T) = M'(\Re(\rho'), \Im(\rho'))^T.$$ 

This leads to

$$M' = \begin{pmatrix} \Re(\tilde{m}^i(T_1^i, T_2^i)) & -\Im(\tilde{m}^i(T_1^i, T_2^i)) \\ \Im(\tilde{m}^i(T_1^i, T_2^i)) & \Re(\tilde{m}^i(T_1^i, T_2^i)) \end{pmatrix}.$$ 

In section 2 the physical model is introduced. We highlighted that this model can easily be extended to an acquisition with several receiver coils. Let $\{C_i\}_{i=1}^{n_c}$ be the $n_c$ different $\sqrt{N} \times \sqrt{N}$ coil sensitivity matrices. We define $\tilde{C}_i = \text{diag}(\{C_i\}_{i=1}^{N}) \in \mathbb{R}^{N \times N}$. Equation (2.1) will then change to

$$z_l = F_l(\tilde{C}_i m_l), \ l = 1, \ldots, L,$$

and 2.2 will become

$$\tilde{m}_l = \frac{1}{n_c} \sum_{i=1}^{n_c} \tilde{C}_i^{-1} F_l^i(z_l), \ l = 1, \ldots, L.$$ 

The entries in $\tilde{C}_i$ can be very close to zero or even be zero, therefore it is better to work with the complex conjugate which is denoted by $\tilde{C}_i^*$. We will then get

$$\tilde{m}_l = (\tilde{C}_i^* \tilde{C}_i)^{-1} \tilde{C}_i^* F_l^i(z_l), \ l = 1, \ldots, L, i = 1, \ldots, n_c.$$ 

and

$$\tilde{C}_i^* \tilde{m}_l = \tilde{C}_i F_l^i(z_l), \ l = 1, \ldots, L, i = 1, \ldots, n_c.$$ 

Equation (2.2) will then become

$$\tilde{m}_l = \left( \sum_{i=1}^{n_c} \tilde{C}_i^* \tilde{C}_i \right)^{-1} \left( \sum_{i=1}^{n_c} \tilde{C}_i F_l^i(z_l), \ l = 1, \ldots, L. \right. \right) \left. \right.$$ 

All the other derivations follow easily.

**Appendix B. Bayesian inference**

The dictionary matching solution (see (Davies et al 2014)) for $(\hat{T}_1^i)$ and $(\hat{T}_2^i)$, in voxel $i$ is chosen to be the entry $\hat{k}_i$ of the dictionary $\{D_k\}_{k=1}^{K}$ which maximizes

$$\hat{k}_i = \text{argmax}_k \frac{\langle \hat{m}', D_k \rangle}{\|D_k\|}.$$
The complex proton density is then calculated as
\[ \hat{\rho}_i = \frac{<\hat{m}_i, D_k>}{||D_k||^2}. \]

In the following, the reference to a specific voxel \( i \) is omitted as the treatment is the same for all voxels.

The dictionary is based on the physical model \( m(T_1, T_2, \rho) \) without the linear factor \( \rho \), which we will denote by \( \hat{m}(T_1, T_2) \). The above equations are therefore equivalent to
\[
(\hat{T}_1, \hat{T}_2) = \text{argmax}_{\hat{T}_1, \hat{T}_2} \left[ \frac{<\hat{m}, \hat{m}(\hat{T}_1, \hat{T}_2)>}{||\hat{m}(\hat{T}_1, \hat{T}_2)||^2} \right],
\]
provided that the resolution of dictionary is infinitely small.

Let \( m_R \) and \( m_I \) denote the real and imaginary part of \( \hat{m}(T_1, T_2) \), and \( \hat{m}_R \) and \( \hat{m}_I \) real and imaginary part of \( \hat{m} \). It follows
\[
(\hat{T}_1, \hat{T}_2) = \text{argmax}_{\hat{T}_1, \hat{T}_2} \sqrt{\left( \frac{m_R^2 + m_I^2}{m_R + m_I} \right)^2 + \left( \frac{\hat{m}_R^2 + \hat{m}_I^2}{\hat{m}_R + \hat{m}_I} \right)^2},
\]
for \( h = (\hat{m}_R^2 + \hat{m}_I^2)^2 + (\hat{m}_R^2 + \hat{m}_I^2)^2 + (\hat{m}_R^2 + \hat{m}_I^2)^2 + 2(\hat{m}_R \hat{m}_I \hat{m}_R + \hat{m}_I \hat{m}_I \hat{m}_R) - 2(\hat{m}_R \hat{m}_I \hat{m}_R + \hat{m}_I \hat{m}_I \hat{m}_R) \) and
\[
\hat{\rho} = \frac{<\hat{m}, \hat{m}(\hat{T}_1, \hat{T}_2)>}{||\hat{m}(\hat{T}_1, \hat{T}_2)||^2} = \frac{\hat{m}_R^2 + \hat{m}_I^2}{\hat{m}_R + \hat{m}_I} - \frac{\hat{m}_R \hat{m}_I}{\hat{m}_R + \hat{m}_I}.
\]
We will now consider the minimization of (2.5). It is
\[
\hat{\rho} = \text{argmin}_\rho ||\hat{m} - M\hat{\mu}||^2
\]
\[ = \text{argmin}_\rho -2\hat{m}^T M\hat{\mu} + \mu^T M^\dagger M\hat{\mu}
\]
\[ = (M^\dagger M)^{-1} M^\dagger \hat{m}
\]
\[ = \begin{pmatrix}
   \hat{m}_R^T & -\hat{m}_I^T \\
   -\hat{m}_I^T & \hat{m}_R^T
\end{pmatrix}
\]
\[ = \begin{pmatrix}
   m_R & m_I \\
   -m_I & m_R
\end{pmatrix}
\]
\[ = \begin{pmatrix}
   \hat{m}_R & \hat{m}_I \\
   -\hat{m}_I & \hat{m}_R
\end{pmatrix}
\]
\[ = \frac{1}{\hat{m}_R^2 + \hat{m}_I^2} \begin{pmatrix}
   \hat{m}_R^2 & \hat{m}_I^2 \\
   \hat{m}_I^2 & \hat{m}_R^2
\end{pmatrix}
\]
which corresponds to the dictionary matching solution B.1.

Let \( c = (\hat{m}_R^2 \hat{m}_R + \hat{m}_I^2 \hat{m}_R, -\hat{m}_R^2 \hat{m}_I + \hat{m}_I^2 \hat{m}_I)^\dagger \), then
\[
\text{argmin}_{\hat{T}_1, \hat{T}_2} ||\hat{m} - M\hat{\mu}||^2 = \text{argmin}_{\hat{T}_1, \hat{T}_2} \begin{pmatrix}
   2 <\hat{m}, M\hat{\mu}> + \frac{<\hat{m}, M\hat{\mu}>}{||\hat{m}||^2}
\end{pmatrix}
\]
\[ = \text{argmax}_{\hat{T}_1, \hat{T}_2} \begin{pmatrix}
   2 <\hat{m}, M\hat{\mu}> - <\hat{m}, M\hat{\mu}>
\end{pmatrix}
\]
\[ = \text{argmax}_{\hat{T}_1, \hat{T}_2} 2 <\hat{m}, M\hat{\mu}>
\]
\[ = \text{argmax}_{\hat{T}_1, \hat{T}_2} \begin{pmatrix}
   \frac{2}{m_R^2 + m_I^2} \hat{m}_R^2 + \hat{m}_I^2 \\
   \frac{2}{m_R^2 + m_I^2} \hat{m}_I^2 + \hat{m}_R^2
\end{pmatrix}
\]
\[ = \text{argmax}_{\hat{T}_1, \hat{T}_2} \begin{pmatrix}
   \frac{2}{m_R^2 + m_I^2} \hat{m}_R^2 + \hat{m}_I^2 \\
   \frac{2}{m_R^2 + m_I^2} \hat{m}_I^2 + \hat{m}_R^2
\end{pmatrix}
\]
\[ = \text{argmax}_{\hat{T}_1, \hat{T}_2} \begin{pmatrix}
   \frac{1}{m_R^2 + m_I^2} \hat{m}_R^2 + \hat{m}_I^2 \\
   \frac{1}{m_R^2 + m_I^2} \hat{m}_I^2 + \hat{m}_R^2
\end{pmatrix}
\]
Hence, the dictionary matching method yields estimates which are given by minimizing (2.5).
In order to derive the posterior distribution in (2.15), we need two prior statements:
For $x \in \mathbb{R}^p$, $\mu \in \mathbb{R}^p$, $\nu > 1$ and symmetric, positive-definite $\Sigma \in \mathbb{R}^{p \times p}$ we have
\[
\int \frac{1}{\det(\Sigma)^{1/2}} \left(1 + \frac{1}{\nu} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)^{-\frac{(\nu+p)}{2}} \, dx \propto 1
\]
and therefore
\[
\int \left(1 + \frac{1}{\nu} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)^{-\frac{(\nu+p)}{2}} \, dx \propto \det(\Sigma)^{1/2}.
\]
This follows from the PDF of the multivariate t-distribution.

For $A \in \mathbb{R}^{n \times n}$ and $a \in \mathbb{R}$ it is
\[
\det \left(\frac{1}{a^2} A\right)^{-1/2} \propto \det \left(\frac{1}{a^2} A\right)^{-1/2} = \left(\frac{1}{a^2}\right)^{\frac{1}{2}} \det(A)^{-1/2}.
\]

We then get for $\hat{\mu} = \mu(T_1, T_2) = [M^T M]^{-1} M^T y$:
\[
\pi(T_1, T_2 | y) \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(T_1, T_2, \mu, \sigma^2 y) \, d\sigma^2 d\mu
\]
\[
\propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_{11}(T_1, T_2) \frac{1}{\sigma^{2+1}} \exp\left(-\frac{1}{2\sigma^2} \|y - M\mu\|^2\right) \, d\sigma^2 d\mu
\]
\[
= I_{11}(T_1, T_2) \int_{-\infty}^{\infty} \int_{0}^{\infty} \frac{1}{\sigma^{1+1}} \exp\left(-\frac{1}{2\sigma^2} \|y - M\mu\|^2\right) \, d\sigma^2 d\mu
\]
\[
= I_{11}(T_1, T_2) \int_{-\infty}^{\infty} \int_{0}^{\infty} \frac{1}{(\sigma+1)^2} \exp\left(-\frac{1}{2\sigma} \|y - M\mu\|^2\right) \, d\sigma^2 d\mu
\]
\[
\propto I_{11}(T_1, T_2) \int_{-\infty}^{\infty} \frac{1}{(\|y - M\hat{\mu}\|^2)^2} \, d\mu
\]
\[
= I_{11}(T_1, T_2) \frac{1}{(\|y - M\hat{\mu}\|^2)^2} \int_{-\infty}^{\infty} \frac{1}{\left(1 + (\mu - \hat{\mu}) M^T M (\mu - \hat{\mu})^{-1}\right)^{1/2}} \, d\mu
\]
(B.2)
\[
= I_{11}(T_1, T_2) \frac{1}{\det(M^T M)^{1/2}} \frac{1}{(\|y - M\hat{\mu}\|^2)^2}.
\]
(B.3)

The posterior (B.3) exists if we make the following two assumptions: $\det(M^T M) > \delta_1 > 0$ and $\min_{y, T_1} \|y - M\hat{\mu}\|^2 > \delta_2 > 0$. Together with the compact set of $T_1$ and $T_2$, the propriety of the posterior is then guaranteed.

A similar derivation with marginalization w.r.t. $\sigma^2$ yields the expression over which is integrated in (B.2), from it follows that the conditional posterior for $\mu$ is that given in (2.17).

Appendix C. Calculation of posterior and derived results

The joint posterior distribution for $T_1$ and $T_2$ is given in (2.15). This function is calculated on a grid of $T_1$ and $T_2$. By taking numerical integrals, the marginal posterior distribution for $T_1$ and $T_2$ can then be approximated. Note that if the stepsize of the grid is infinitely small, the integral is exactly calculated. We used a trapezoidal method with unit spacing given via
\[
\int_{a}^{b} f(x) \, dx \approx \frac{b - a}{2N} \sum_{n=1}^{N} \left(f(x_n) + f(x_{n+1})\right)
\]
\[
= \frac{b - a}{2N} \left(f(x_1) + 2f(x_2) + \ldots + 2f(x_N) + f(x_{N+1})\right),
\]
where the scalar value $\frac{b - a}{N}$ is equal to one. This is implemented in the Matlab® function trapz.
ORCID iDs

Selma Metzner  
https://orcid.org/0000-0001-8879-1895

References

Balsiger F, Konar A S, Chikop S, Chandran V, Scheidegger O, Geethanath S and Reyes M 2018 Magnetic resonance fingerprinting reconstruction via spatiotemporal convolutional neural networks International Workshop on Machine Learning for Medical Image Reconstruction (Berlin: Springer) pp 39–46

Christen T, Pannetier N, Ni W W, Qiu D, Moseley M E, Schuff N and Zaharchuk G 2014 Mr vascular fingerprinting: A new approach to compute cerebral blood volume, mean vessel radius, and oxygenation maps in the human brain Neuroimage 89 262–70

Cohen O, Zhu B and Rosen M S 2018 Mr fingerprinting deep reconstruction network (drono) Magn. Reson. Med. 80 885–94

Davies M, Puy G, Vandergheynst P and Wiaux Y 2014 A compressed sensing framework for magnetic resonance fingerprinting SIAM J. Imaging Sci. 7 2623–56

Fang Z, Chen Y, Liu M, Xiang L, Zhang Q, Wang Q, Lin W and Shen D 2019 Deep learning for fast and spatially constrained tissue quantification from highly accelerated data in magnetic resonance fingerprinting IEEE Trans. Med. Imaging 38 2364–74

Flasbeck S, Schmidt S, Bachert P, Ladd M E and Schmitter S 2019 Flow mr fingerprinting Magn. Reson. Med. 81 2536–50

Gelman A, Carlin J B, Stern H S, Dunson D B, Vehtari A and Rubin D B 2013 Bayesian Data Analysis 3rd edn. (London/Boca Raton, FL: Chapman and Hall/CRC Press)

Greengard L and Lee J-Y 2004 Accelerating the nonuniform fast Fourier transform SIAM Rev. 46 443–54

Hamilton J L, Jiang Y, Chen Y, Ma D, Lo W-C, Griswold M A, Jakob P M, Kotas M, Gulani V, Flentje M and Haase A 2004 Inversion recovery true fi

Hoppe E, Korfadloter G, Wübbeler G and Elster C 2019 Approximate large-scale bayesian spatial modeling with application to quantitative magnetic resonance imaging AStA Adv. Stat. Anal. 103 333–55

Metzner S and Wübbeler G 2021 Software - Working Group 8.42 (https://www.ptb.de/cms/en/ptb/fachabteilungen/abt8/fb-84/ag-842/software.html)

Nishimura D 1996 Principles of Magnetic Resonance Imaging (Stanford, CA: Stanford University)

Polders D L, Leemans A, Luijten P R and Hoogduin H 2012 Uncertainty estimations for quantitative in vivo mri t1 mapping J. Magn. Reson. 224 53–60

Robert C and Casella G 2013 Monte Carlo Statistical Methods (New York: Springer) (https://doi.org/10.1007/978-1-4757-4145-2)

Schmitter S, Schmidt S, Bachert P, Ladd M E and Schmitter S 2019 Flow mr fingerprinting Magn. Reson. Med. 81 2536–50

Weigel M 2015 Extended phase graphs: dephasing, rf pulses, and echoes-pure and simple J. Magn. Reson. Imaging 41 266–95