Fakultät für Informatik
Technische Universität München

Disentangling Tissue Microstructure with Magnetic Resonance Imaging

Miguel Molina Romero

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr.-Ing. Darius Burschka

Prüfende der Dissertation:

1. Prof. Dr. rer. nat. Bjoern H. Menze
2. Priv.-Doz. Dr. rer. nat. Marion I. Menzel

Die Dissertation wurde am 29.03.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 14.07.2018 angenommen.

DISENTANGLING TISSUE MICROSTRUCTURE WITH MAGNETIC RESONANCE IMAGING

MIGUEL MOLINA ROMERO

Act as if

— William James

To María

ABSTRACT

Magnetic Resonance Imaging (MRI) has significantly contributed to modern medicine, especially neurology. It is a noninvasive and harmless imaging modality, capable of producing a wide variety of tissue contrasts, being the preferred modality for soft-tissue imaging. MRI is sensitive to a diversity of biophysical effects and thus, the generated images are a mixture of them. In particular, Diffusion Magnetic Resonance Imaging (dMRI) is influenced by the thermal motion of water molecules, and has clinical value in the diagnostic of stroke, and the surgical planning for brain tumors. Moreover, dMRI contains information that goes beyond the image resolution of millimeters, down to the underlying tissue microstructure.

This thesis takes MRI a step further to study brain tissue microstructure. To that end, the dMRI signal is reformulated in a Blind Source Separation (BSS) framework, enabling the disentanglement of subvoxel tissue signal components, and the estimation of multiple tissue parameters. Furthermore, this unsupervised machine learning approach is transformed into a supervised deep learning model, tackling the partial volume contamination caused by Cerebrospinal Fluid (CSF) or free-water in brain dMRI. The impact of this method is assessed in glioblastoma cases, yielding corrected diffusion indexes and better delineation of edema and tumor infiltration. Finally, Quantitative Transient-state Imaging (QTI), an ultra-fast acquisition and reconstruction scheme for multiparameter mapping, is extended to a tissue multicompartment model.

Summarizing, this thesis contributes with three strategies to extract more relevant tissue microstructure information from the measured MRI data.

ZUSAMMENFASSUNG

Die Magnetresonanztomographie (MRT) hat wesentlich zur modernen Medizin Diagnostik, insbesondere im Bereich der Neurobildgelung, beigetragen. Es ist eine nichtinvasive Bildgebungsmodalität, die in der Lage ist, eine Vielzahl von Gewebekontrasten zu erzeugen. Dies macht sie zur bevorzugten Modalität für die Weichteilbildgebung. Die MRT ist empfindlich gegenüber einer Vielfalt von boiphysiken Effekten und die erzeugten räumlichen aufgelösten Signale enthalten die Superposition dieser Effekten. Insbesondere misst die Diffusionsbildgebung (dMRI) die thermische Bewegung von Wassermolekülen. dMRI ist von hohem Klinische wert, z.B. in der Diagnostik von Schlaganfall und der chirurgische Planung von Hirntumoren. Darüber hinaus enthält dMRI Informationen, die über die Bildauflösung hinaus bis zur zugrunde liegenden Gewebemikrostruktur reichen.

Diese Arbeit führt die MRT einen Schritt weiter, um Hirngewebemikrostrukturen zu untersuchen. Zu diesem Zweck wird das dMRI-Signal in ein Blind Source Separation (BSS) -Framework umformuliert, das die Entflechtung von Subvoxel- Gewebe-Signalkomponenten und die Schätzung von mehreren Gewebeparametern ermöglicht. Darüber hinaus wird dieser maschinelles Lernansatz in ein Deep-Learning-Modell umgewandelt. So kann die durch Liquor cerebrospinalis (CSF) verursachte Partialvolumenkontamination oder das freie Wasser in dMRI bestimmt und separiert wurden. Der Einfluss dieser Methode wird im Glioblastom untersucht. Die korrigierten Diffusionsindizes ermöglichen eine bessere Abgrenzung von Ödemen und Tumorinfiltrationen. Schließlich wird das Quantitative Transient-state Imaging (QTI), ein ultraschnelles Akquisitions- und Rekonstruktionsschema für multiparametrische-Quantifizierung, auf ein Gewebe-Multikompartiment-Modell erweitert.

Zusammenfassend liefert diese Arbeit drei Ansätze, um relevantere Gewebe-Mikrostrukturinformationen aus den gemessenen MRI-Daten zu extrahieren.

PUBLICATIONS

This publication-based thesis contains the following publications:

TISSUE MICROSTRUCTURE CHARACTERIZATION WITH DIFFUSION MRI

1. **M. Molina-Romero**, PA. Gómez, JI. Sperl, M. Czisch, PG. Sämann, DK. Jones, MI. Menzel, BH. Menze. A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization. In: *Magnetic Resonance in Medicine*, 00:1–18. https://doi.org/10.1002/mrm.27181 (2018) [52].

MULTIPARAMETRIC MAPPING IN THE TRANSIENT-STATE

3. PA. Gómez, C. Ulas, JI. Sperl, T. Sprenger, **M. Molina-Romero**, MI. Menzel, BH. Menze. Learning a Spatiotemporal Dictionary for Magnetic Resonance Fingerprinting with Compressed Sensing. In: *MICCAI Patch-MI Workshop* LNCS 9467 (2015), pp. 112–119 [25].

4. PA. Gómez, **M. Molina-Romero**, C. Ulas, G. Buonincontri, JI. Sperl, DK. Jones, MI. Menzel, BH. Menze. Simultaneous Parameter Mapping, Modality Synthesis, and Anatomical Labeling of the Brain with MR Fingerprinting. In: *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention* LNCS 9902 (2016), pp. 579–586 [26].

The following relevant contributions to this dissertation were published as Peer-reviewed abstracts at the annual conference of the International Society for Magnetic Resonance in Medicine:

TISSUE MICROSTRUCTURE CHARACTERIZATION WITH DIFFUSION MRI

1. **M. Molina-Romero**, PA. Gómez, JI. Sperl, DK. Jones, MI. Menzel, BH. Menze. Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis. In: *ISMRM Workshop on Breaking the Barriers of Diffusion MRI* (2016) [50].

2. **M. Molina-Romero**, PA. Gómez, JI. Sperl, AJ. Stewart, DK. Jones, MI. Menzel, BH. Menze. Theory, Validation and Aplication of Blind Source Separation to Diffusion MRI for Tissue Characterisation and Partial Volume Correction. In: *Proc Intl Soc Mag Reson Med* (2017) [51].

3. **M. Molina-Romero**, PA. Gómez, S. Albarqouni, JI. Sperl, MI. Menzel, BH. Menze. Deep learning with synthetic data for free water elimination in diffusion MRI. In: *Proc Intl Soc Mag Reson Med* (2018) [53].

MULTIPARAMETRIC MAPPING IN THE TRANSIENT-STATE

4. PA. Gómez, G. Buonincontri, **M. Molina-Romero**, C. Ulas, JI. Sperl, MI. Menzel, BH. Menze. 3D Magnetic Resonance Fingerprinting with a Clustered Spatiotemporal Dictionary. In: *Proc Intl Soc Mag Reson Med* (2016) [26].

5. PA. Gómez, G. Buonincontri, **M. Molina-Romero**, JI. Sperl, MI. Menzel, BH. Menze. Accelerated Parameter Mapping with Compressed Sensing: an Alternative to MR Fingerprinting. In: *Proc Intl Soc Mag Reson Med* (2017) [27].

The annex of this thesis contains the following manuscript currently under review for publication and the summary of a relevant co-supervised master thesis:

MANUSCRIPT UNDER REVIEW AND MASTER THESIS

1. **M. Molina-Romero**, B. Wiestler, PA. Gómez, MI. Menzel, BH. Menze. Deep learning with synthetic diffusion MRI data for free-water elimination in glioblastoma cases. *MICCAI*. Under review (2018).

2. MP. Orihuela, **M. Molina-Romero**, PA. Gómez, BH. Menze. Brain Microstructure Quantification from Transient-State Magnetic Resonance Imaging.

ACKNOWLEDGMENTS

I want to sincerely thank my supervisor Prof. Björn Menze for his support and continuous push towards my limits, that made me advance scientifically and personally. To my mentor at GE, P.D. Dr. Marion Menzel, thank you for your encouragement, and overall, for bringing me into this adventure. Thank you to Prof. Axel Haase for your trust. Finally, a big thank you to Prof. Derek Jones for his warm welcome to Cardiff and for enlightening my stay at CUBRIC.

This all started in 2013 in Madrid, when Prof. Juan Tamames gave me the opportunity to enter research. Thank you Juan, you paved my way towards today. And thank you to José Pineda and Anabea Solana, I will never forget that you gave me the first push to start walking along this path.

My gratitude to the GE Global Research team, Martin, Anne, Anabea, Rolf, Dirk, Florian, Silke, André, Guido, Matt, Cristina, Brice, Jonathan, Marion, and Timo for your openness and help. The uncountable cakes, social events, and skiing trips will be always in my memory, but also everything I learned from you.

To our partners at the Max Plank Institute of Psychology Dr. Michael Czisch and Dr. Philipp Sämann, thank you for your patience and help with the volunteers in the stressful days. Also my recognition to Jana Lipkova and the collaborators at ETH Zurich, Georgios Arampatzis, Panagiotis Angelikopoulos and Panagiotis Chatzidoukas for showing me the secrets of the Bayesian inference.

My sincere acknowledgment to my colleagues in CUBRIC, Greg Parker, Alberto Merola, Lorenzo Magazzini, Mark Drakesmith, Joe Whittaker, John Evans, Hannah Furby, Angela Reardon, Silvia de Santis, Yaniv Assaf and Derek Jones. Also to Andrew Stewart for enabling and easing the dozens of scanning hours.

Thank you to my fellow PhD and master students, Eduardo, Tim, Markus, Uli, Eugen, Cagdas, Dhritman, Carolin, Paula, Jonathan, and Nico for launching my kicker career and creating the best possible working atmosphere. Naturally, my special recognition to my BERTI, Teresa, Marwan, Marta, Aurelien, Saheed, Fatih, Ming, Bea, Yash, Suat, Xin, Jaber, Shufang, Pedro, Petra, Katharina, and Andrea. We learned how to swim together, and you made of our training, kick-offs, summer schools, and social events an enjoyable intercultural exchange that cultivated ourselves. BERTI was funded by the European Commission with the Grant Agreement Number 605162.

Por último, gracias a mi familia por aceptar que un hijo y hermano se aleje. Vuestro cariño, apoyo y persistencia me sostuvo cuando fue necesario, incluso sin saberlo. Esta tesis es vuestra.

CONTENTS

ACRONYMS

AD Axial Diffusivity

ALS Alternating Least Squares

ANN Artificial Neural Network

BSS Blind Source Separation

cALS Constrained Alternating Least Squares

CSF Cerebrospinal Fluid

CS Compressed Sensing

CPMG Carr-Purrcell-Meiboom-Gill

DTI Diffusion Tensor Imaging

dMRI Diffusion Magnetic Resonance Imaging

EPI Echo-planar Imaging

EPG Extended Phase Graphs

FA Fractional Anisotropy

FLASH Fast Low Angle Shot

FT Fourier Transform

FWE Free-Water Elimination

GM Gray Matter

ICA Independent Component Analysis

ILT Inverse Laplace Transform

ISMRM International Society for Magnetic Resonance in Medicine

LR Low Rank

MCMC Markov Chain Monte Carlo

MICCAI Medical Imaging Computing and Computer Assisted Intervention

MRI Magnetic Resonance Imaging

MRF Magnetic Resonance Fingerprinting

MD Mean Diffusivity

NMR Nuclear Magnetic Resonance

NNLS Non-Negative Least Squares

NMF Nonnegative Matrix Factorization

NSC Nonnegative Sparse Coding

PCA Principal Component Analysis

PDF Probability Density Function

PD Proton Density

PGSE Pulsed Gradient Spin Echo

QTI Quantitative Transient-state Imaging

RD Radial Diffusivity

RF Radiofrequency

SCA Sparse Component Analysis

SSFP Steady-State Free Precession

TE Echo Time

TMCMC Transitional Markov Chain Monte Carlo

TR Repetition Time

WM White Matter

Part I

INTRODUCTION AND SUMMARY OF CONTRIBUTIONS

1 INTRODUCTION

The decade of 1970s saw a revolution in Radiology with the introduction of important non invasive image modalities. Computer tomography [30], single photon emission computer tomography [12], positron emission tomography [66], and magnetic resonance imaging (MRI) [24, 42] were developed. Among these image methods, MRI became the method of choice in many clinical application due to its versatility and innocuity. The first commercial MRI scanners were available during the 1980s, reaching 10000 scanners worldwide by 1996. Important research and engineering efforts were required during previous decades. Nuclear Magnetic Resonance (NMR) was first observed in 1938 by Rabi [61] and measured and formally described by Purcell [60] and Bloch [10] in 1946. It was not until 1971 that NMR was considered a breakthrough in medical technology. Raymond Damadian discovered that some tumors presented distinct relaxation times compared with normal tissue [17]. Two years later, in 1973, Paul Lauterbur published in Nature an article proposing for the first time to use magnetic field gradients to localize NMR signals, creating an image from projections (as in computer tomography). Shortly after, Sir Peter Mansfield and his group in Nottingham introduced the selective excitation that made MRI a tomography method [24]. Whilst, in 1975 Richard Ernst presented the concept of phase encoding and two-dimensional Fourier Transform. The contributions of Purcell, Bloch, Lauterbur, Mansfield, and Ernst established the basis of modern MRI and were awarded with the Nobel Price.

Classical MRI is driven by the generation of images weighted by different contrasts. These type of images are subject to a diversity of biophysical effects and do not describe independent physical properties of the tissue. To extract tissue quantitative properties it is necessary create biophysically inspired signal models that explain the observed data. This is the case of Diffusion Magnetic Resonance Imaging (dMRI) [14, 64], which enables to measure the water diffusion parameters in the brain tissue. However, the unit of volume in MRI, the *voxel*, is not an homogeneous object. It contains the complexity of the tissue microstructure organization [18, 20, 38, 58, 63, 65, 72]. Diffusion MRI is sensitive to this complexity but models considering multiple tissue components are only of interest for the research community, far from clinical application. This thesis introduces plausible solutions with minimum modeling to extract important information beyond the voxel limit.

1.1 INTRODUCTION

This section introduces the theoretical concepts of NMR and MRI that are used in this thesis. The aim is to cover the principles to the extent that is necessary to follow this work. The reader can find further theoretical details in [48] and [36].

1.1.1 *Magnetic Resonance Imaging*

When Lauterbur [42] and Mansfield [24] proposed MRI, it was accepted that the *spin* (angular momentum) of a proton placed in a magnetic field precesses around the field direction. The precession frequency was established by Sir Joseph Larmor, and it is known as Larmor frequency:

$$\omega_0 = \gamma B_0, \tag{1}$$

where ω_0 is the angular frequency of the precessing protons, and B_0 is the magnetic field strength. The constant γ is known as the gyromagnetic ratio and depends on the element. In water (^{1}H), $\gamma = 2.68 \times 10^8$ rad/s/Tesla for the hydrogen proton, or $\gamma = \gamma/2\pi = 42.6$ MHz/Tesla. The experiments in this thesis were conducted in 3 Tesla scanners, setting the reference Larmor frequency to $\omega_0 = 127.8$ MHz.

The classical mechanics approach is based on the aggregation of the magnetic moment ($\vec{\mu}$) of the spins contained in an unit of volume V, the *voxel*. The macroscopic magnetization is known as $\vec{M}$:

$$\vec{M} = \frac{1}{V}\sum_{V} \vec{\mu_i} \tag{2}$$

The evolution of $\vec{M}$ with time in the presence of an external magnetic field, $\vec{B}_{ext}$ was described by Bloch:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B}_{ext} + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}\vec{M}_\perp, \tag{3}$$

where $\vec{M}$ is composed by the longitudinal magnetization (M_z) and the transverse magnetization ($\vec{M}_\perp = M_x\hat{x} + M_y\hat{y}$). In the equilibrium position $M_z(\infty) = M_0$ and $M_x(\infty) = M_y(\infty) = 0$. The relaxation times ($T_1$ and T_2) describe the return to equilibrium after excitation, and appear as a consequence of the interaction of the spins with the environment. The spin-lattice or longitudinal relaxation time (T_1) is driven by the energy exchange between the protons and the environment, and describes the recovery of M_z towards M_0. The spin-spin or transverse relaxation time (T_2) is induced by local variations of B_0 in

the microenvironment of the protons that lead to small differences in the Larmor frequency, which cause dephasing of the transverse magnetization pushing M_x and M_y to zero. Interestingly, T_1 and T_2 are tissue dependent and their differences across the body are the major source of contrast in MRI.

The solution of the Bloch equation (Equation 3) for a constant field ($\vec{B}_{ext} = B_0\hat{z}$) generates the set of equations:

$$M_x(t) = e^{-t/T_2}\,(M_x(0)cos(\omega_0 t) + M_y(0)sin(\omega_0 t)) \tag{4}$$

$$M_y(t) = e^{-t/T_2}\,(M_y(0)cos(\omega_0 t) - M_x(0)sin(\omega_0 t)) \tag{5}$$

$$M_z(t) = M_z(0)e^{-t/T_1} + M_0(1 - e^{-t/T_1}) \tag{6}$$

Only the transverse components are measurable by Radiofrequency (RF) coils. However, these components do not exist in the equilibrium state, requiring the system to by excited. The excitation of a NMR spin system happens when an RF (B_1) pulse is played at the Larmor frequency to meet the resonance condition. In such a case, the external field is composed by the static magnetic field and the RF field: $\vec{B}_{ext} = B_0\hat{z} + \vec{B_1}$.

When the system is not in equilibrium $\vec{M}$ rotates at the Larmor frequency around $\vec{B}_{ext}$. To simplify the mathematical framework a new reference system is defined, the rotating frame, where the new coordinates $(\hat{x}', \hat{y}')$ rotate in the fixed frame defined by $(\hat{x}, \hat{y})$ counterclockwise at frequency ω around $\hat{z}$:

$$\vec{B}_{eff} = (B_0 - \frac{\omega}{\gamma})\hat{z} + B_1\hat{x}'. \tag{7}$$

To solve the Bloch equations in the rotating frame of reference, the RF field $\vec{B_1}$ is assumed to be much smaller in magnitude than B_0. Then,

$$\left(\frac{dM_{x'}}{dt}\right)' = \Delta\omega M_{y'} - \frac{M_{x'}}{T_2}, \tag{8}$$

$$\left(\frac{dM_{y'}}{dt}\right)' = -\Delta\omega M_{x'} + \omega_1 M_z - \frac{M_{y'}}{T_2}, \tag{9}$$

$$\left(\frac{dM_z}{dt}\right)' = -\omega_1 M_{y'} + \frac{M_0 - M_z}{T_1}, \tag{10}$$

where ω_1 is the spin frequency due to B_1. The term $\Delta\omega = \omega_0 - \omega$ are the *off-resonance* contributions that represent non-ideal conditions.

If the time duration of B_1 is short, the relaxation terms can be ignored relative to the frequency terms during that period, resulting in:

$$\left(\frac{\vec{M}}{dt}\right)' = \gamma\vec{M} \times \vec{B_{eff}} \tag{11}$$

After B_1 is turned off, the magnetization evolution is as expressed by Equation 4 – 6, but in the rotating frame:

$$M_{x'}(t) = e^{-t/T_2}\left(M_{x'}(0)\cos(\Delta\omega t) + M_y(0)\sin(\Delta\omega t)\right) \tag{12}$$

$$M_{y'}(t) = e^{-t/T_2}\left(M_{y'}(0)\cos(\Delta\omega t) - M_x(0)\sin(\Delta\omega t)\right) \tag{13}$$

$$M_z(t) = M_z(0)e^{-t/T_1} + M_0(1 - e^{-t/T_1}) \tag{14}$$

Let M_+ be the transverse magnetization $M_\perp$ in the rotating frame of reference when it rotates at the Larmor frequency: $\Delta\omega = 0$. When RF pulses, that flip M from the equilibrium state $\pi/2$ rad to the transverse plane, are applied at periods of Repetition Time (TR), and the NMR signal are measured at Echo Time (TE), the solution to Equation 3 following the RF pulse is

$$M_+(TE, TR) = M_0\left(1 - e^{-TR/T_1}\right)e^{-TE/T_2} \tag{15}$$

Equation 15 introduces the concept of contrast in MRI. Images dominated by Proton Density (PD) (proportional to M_0) can be generated for long TR and short TE. T_1 weighted images require short TR and TE. Whereas T_2 weighting needs long TR and TE.

The image formation based on NMR requires the spatial encoding of $\vec{M}$. This is achieved by the use of *imaging gradients* (g_x, g_y, g_z) that produce a variation of the static magnetic field along each direction (measured in mT/m), inducing differences in the Larmor frequency at any spatial location,

$$\omega_g(\vec{r}) = \gamma(B_0 + g(\vec{r})), \tag{16}$$

which causes the spins to accumulate phase during time t:

$$\omega_g(\vec{r}, t) = -\gamma\int_0^t g(\vec{r}, t)dt. \tag{17}$$

The frequency encoding of the space induced by the imaging gradients (Equations 16 and 17) is known as k-space (spatial frequencies):

$$k = -\gamma \int_0^t g(t)dt. \tag{18}$$

and relates to the image space by a spatial multidimensional Fourier Transform (FT):

$$\rho(\vec{r}) = \int_{\vec{k}} y(\vec{k}) e^{i2\pi\vec{k}\vec{r}}, \tag{19}$$

where $y(\vec{k})$ is the measured signal in the k-space ($\vec{k} = [k_x\ k_y\ k_z]$), and $\rho(\vec{r})$ is the actual image ($\vec{r} = [r_x\ r_y\ r_z]$).

1.1.2 *Diffusion MRI*

The term diffusion makes reference to the random translational motion of particles due to the thermal energy. It was first observed in 1827 by the botanist Robert Brown on pollen grains large enough to be seen in the optical microscope, and small enough to be subject of the thermal motion of the surrounding water molecules. Albert Einstein [22], using Fick's laws of flux for a net flow of zero, demonstrated that the diffusion coefficient, D, depends on the absolute temperature, T, the viscosity of the medium, η, and the radius of a spherical particle, R:

$$D = \frac{k_B T}{6\pi\eta R}. \tag{20}$$

Given that diffusion is a random process, Einstein also derived the conditional probability that a particle in $\vec{r}$ at time zero moves to $\vec{r}'$ after time t:

$$P(\vec{r}|\vec{r}', t) = \frac{1}{(4\pi D t)^{-3/2}} e^{-\frac{\vec{r}'-\vec{r}}{4Dt}}, \tag{21}$$

leading to the mean-squared distance traveled by particles along one direction in time t:

$$\langle X^2 \rangle = \langle (\vec{r}' - \vec{r})^2 \rangle = 2Dt. \tag{22}$$

To quantify translational motion in NMR, as for imaging, magnetic filed gradients are introduced causing linear spatial variations of the Larmor frequencies:

$$\omega(\vec{r}) = \gamma B_0 + \gamma \vec{g} \cdot \vec{r} \tag{23}$$

where $\vec{g}$ indicates the magnetic profile variation induced that is parallel to the main field $\vec{B}_0$.

Flow ($\vec{v}$) and diffusion (D) terms were introduced in the Bloch equation [67] to form the Bloch-Torrey equation. Thus, in absence of an RF field, the transverse magnetization in the rotating frame (M_+) is:

$$\frac{\partial M_+}{\partial t} = -i\gamma\vec{g}\cdot\vec{r} - \frac{M_+}{T_2} + D\nabla^2 M_+ - \vec{v}\cdot\nabla M_+ \tag{24}$$

The most common technique to make NMR sensitive to diffusion is called Pulsed Gradient Spin Echo (PGSE) and was introduced by Stejskal and Tanner in 1965 [64]. It works under the principles of *dephase* and *rephase*. Two squared magnetic gradients are introduced, called *diffusion gradients* ($\vec{g}$), to encode for diffusion. The first one is played during a time δ causing particles to gain phase or *dephase* during that time. The second one is played after time Δ during exactly δ seconds but with reversed polarity, causing a loss of phase or *rephase*. The underlying principle is that only the particles that remain static during time Δ can rephase, while those that moved due to diffusion cannot. The reason is that these moving particles experience different Larmor frequencies during the first and second diffusion gradient (Equation 23) leading to a residual phase that causes a signal attenuation described by:

$$A = e^{-\gamma^2 g^2 \delta^2 D(\Delta-\delta/3)} e^{i\gamma\delta\vec{v}\cdot\vec{g}\Delta}. \tag{25}$$

The attenuation in the first term of Equation 25 is the source of contrast in Diffusion Magnetic Resonance Imaging (dMRI). The phase in the second term caused by flow is generally neglected –it may cause flow and motion artefacts–. The effects of the diffusion gradients are summarized in the parameter b [9]:

$$b = \gamma^2 \int_0^{TE} \left[\int_0^t \vec{g}(t')dt'\right]^2 dt, \tag{26}$$

which in the PGSE case takes the form:

$$b = \gamma^2 g^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) \tag{27}$$

simplifying the attenuation expression to:

$$A = e^{-bD}. \tag{28}$$

Diffusion MRI is sensitive to the displacement of particles in the order of the size of a cell. This makes dMRI the only non-invasive

imaging modality to study tissue microstructure. A plethora of methods and models exist for this matter (see [1–3, 23, 74]), but the most extended one with clinical relevance is Diffusion Tensor Imaging (DTI) [4].

DTI models the diffusion, D, in Equation 28 as a tensor describing the diffusion coefficient in each direction:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}. \tag{29}$$

To estimate the six unknown variables of the tensor (Equation 29) at least six diffusion directions must be encoded by the diffusion gradients, although a minimum of 30 are advisable [37]. Besides, given that the diffusion contrast is based on attenuation, an extra measurement with no diffusion encoding is necessary for reference ($b = 0$ s/mm^2).

Only one tissue compartment is considered in the formulation of DTI, oversimplifying tissue complexity. Nevertheless, the factorization in eigenvalues (λ) and eigenvectors of diffusion tensor (Equation 29) reveals microscopic information encoded in the tensor. Scalar indexes can be derived, for instance the Mean Diffusivity (MD)

$$MD = \langle\lambda\rangle = \frac{1}{3}\sum_{i=1}^{3}\lambda_i. \tag{30}$$

The Fractional Anisotropy (FA) is an index between 0 and 1 that informs about the anisotropy of the tissue, where 0 indicates a pure isotropic structure

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{\sum_{i=1}^{3}(\lambda_i - \langle\lambda\rangle)^2}{\sum_{i=1}^{3}\lambda_i^2}}. \tag{31}$$

The Axial Diffusivity (AD) indicates the maximum diffusivity along the direction of the associated eigenvector

$$AD = \lambda_1 \tag{32}$$

Finally, the Radial Diffusivity (RD) is the mean diffusivity in the plane perpendicular to the main diffusion direction

$$RD = \frac{1}{2}\sum_{i=2}^{3}\lambda_i. \tag{33}$$

1.1.3 *Multiparametric mapping in the transient state*

The estimation of T_1 and T_2 parameters in MRI is slow. This process requires repeated measures for different TE or TR and the fit of the data to Equation 15. The acceleration alternatives are based on undersampled acquisition, fast imaging sequences, or their combination.

The Compressed Sensing (CS) based techniques measure data under the Nyquist limit, which turns the image reconstruction into an ill-posed problem that requires iterative regularized reconstructions [8, 21, 33, 69, 75]. Fast imaging sequences include Echo-planar Imaging (EPI) [49], Fast Low Angle Shot (FLASH) [28], Look-Locker [39], or Steady-State Free Precession (SSFP) based schemes [15]. In the recent years, Magnetic Resonance Fingerprinting (MRF) [45] introduced a mixed solution that combines undersampled image coverage with a SSFP like acquisition. However, unlike SSFP the measures are made in the transient-state of the magnetization, rather than the steady-state.

The acquisition of an MRI signal repeated times produces a signal evolution in k-space with temporal variation:

$$y(\vec{k}, t) = \int_{\vec{r}} \rho(\vec{r}) l_t(\vec{r}) e^{-2\pi \vec{k}(t)\cdot\vec{r}} \tag{34}$$

where $\vec{k}(t)$ is the trajectory used to cover the k-space. At the spatial position $\vec{r}$, ρ is the proton density, and l_t is the temporal evolution of the signal, which can be expressed in a recursive manner:

$$l_t(\vec{r}) = l_{t-1}(\vec{r}) g(\alpha(t), TR(t), TE(t), T_1(\vec{r}), T_2(\vec{r})). \tag{35}$$

Equations 34 and 35 show a dependence on the temporal prescription of the acquisition parameters: $\alpha(t)$, $TR(t)$, and $TE(t)$; and the tissue local properties: $\rho(\vec{r})$, $T_1(\vec{r})$, and $T_2(\vec{r})$. The evolution is modulated by $g(\cdot)$, which accounts for gradient effects, RF excitation, and relaxation. This can be solved with Extended Phase Graphs (EPG) [70, 71] or Bloch simulators [10].

1.1.4 *Tissue microstructure and partial volume effects*

The scale size of cells is in the order of tens of micrometers, creating pools of water organized in patterns of variable complexity and properties. These pools of water or compartments are permeable allowing water to filtrate from one to another. However, this is a slow process that is neglected in this thesis [2, 54].

In clinical MRI the size of a voxel is in the millimeter scale, making unfeasible to detect the complexity of the underlying tissue microstructure. As a result, the estimates of the tissue parameters (T_1, T_2, and diffusion indexes) are averages over the volume of the voxel.

This issue is known as partial volume effects. Several approaches have been developed to address it, and gain sub-voxel information that reflects microstructural organization.

The tissue microstructure characterization have been addressed in three different manners: relaxometry, diffusion, and their combination. All of them acknowledge the existence of multiple water compartment within a voxel, having each different properties.

The dMRI microstructure theory [2, 38, 41, 63, 65, 74] models tissue as a composition of compartments with different geometrical shapes and diffusion properties: free, hindered, and restricted. Moreover, specific models yielding metrics as axon diameter [1, 3], or neurite orientation dispersion [74] are used in clinical research.

From the relaxometry perspective, data intensive techniques like the Inverse Laplace Transform (ILT) with Non-Negative Least Squares (NNLS) are used to fit multiexponential decays to the measures [47, 72]. More recently, a technique based on steady-sate acquisition was developed to yield information of tissue microstructure up to three compartments [19].

Nonetheless, dMRI and relaxometry can be combined to survey the tissue underlying structure. Studies using a diffusion-weighted Carr-Purrcell-Meiboom-Gill (CPMG) sequence [20, 58] showed correlation between diffusivity and relaxation. Traditional diffusion-relaxation correlation studies are highly data demanding [13], but more recent techniques have presented smart regularization techniques [5, 40] that bring them closer to in vivo applications

In some cases, the averaging outcome of the partial volume effects is considered contamination[59]. This is the case of the diffusion estimates and the bias that Cerebrospinal Fluid (CSF) or free-water introduces. As it is discussed in the next section (2) Free-Water Elimination (FWE) and tissue microstructure characterization are the main themes of this work.

2 OUTLINE AND CONTRIBUTION OF THIS THESIS

This thesis focuses on methods to gain information about the tissue composition beyond the resolution limit of MRI. Many have addressed this problem before using cumbersome models with plenty of assumptions, or ill-posed model-free approaches (see section 1.1.4). This dissertation opens three new to ways that minimize the assumptions, without compromising convergence. Moreover, these are less data demanding than existing state-of-the-art solutions.

For the first time, the Blind Source Separation (BSS) formulation and solutions are applied to the dMRI compartmental problem (section 3.1). BSS provides the theory and tools to separate signals that have been linearly entangled. Thus, it is applicable to the tissue microstructure partial volume effects issue. In section 4.2.1, the dMRI signal is for the first time expressed in a BSS framework [50]. A solution based on Sparse Component Analysis (SCA) was proposed. However, SCA limits the clinical usability due to protocol constraints. A more general solution based on Nonnegative Sparse Coding (NSC) was introduced in section 4.2.2 [51]. Finally, peer-reviewed journal publication based on Nonnegative Matrix Factorization (NMF) (section 4.1.1) proved the stability and robustness of this approach to disentangle dMRI signals from tissue microstructure components [52].

Quantitative Transient-state Imaging (QTI) technique for ultra-fast quantification of multiple tissue parameter [27] is introduced in chapter 5. This imaging method enables simultaneous voxel-based quantification of T_1, T_2, and PD. However, as explained in section 1.1.4 partial volume effects are also induced in this technique. In appendix A.2.1 a disentangling approach different from BSS is explored. This is based on tissue compartmental modeling and parameter estimation, based on a Transitional Markov Chain Monte Carlo (TMCMC). A Bayesian estimation and uncertainty quantification framework.

Finally, the most recurrent issue in partial volume effects for dMRI is CSF or free-water contamination, which induces undesired biases in the diffusion indexes. Literature proved the ill-posed nature of fitting the DTI model accounting for two microstrucutre compartments [59]. Several solutions to this problem were introduced in form of spatial regularization [57] or protocol optimization [31]. Section 4.2.3 and annex A.1.1 introduce a new perspective to this issue. The ill-posed regularization problem is addressed with a pattern matching solution based on deep learning [53]. This approach enables an Artificial Neural Network (ANN) to be trained with synthetically generated dMRI

data. Unlike the state-of-the-art solutions, it does not require spatial regularization or specific protocols, and it is at least 55-fold faster.

3

METHODOLOGY

This chapter provides with a theoretical background of the signal processing techniques applied in this thesis. An introduction to the specific application is made. For further information read the referenced sections containing the publication that developed the work.

3.1 BLIND SOURCE SEPARATION FOR DINSENTANGING DIFFUSION MRI

As a general definition BSS [73] comprises the set of techniques that try to disentangle signals sources that have been linearly mixed together. All these techniques share a common working principles: there is some prior information about the signal sources, and there is redundancy in the entangled measures. In this thesis, BSS is applied to dMRI data in order to separate signals coming from independent microstructure compartments in brain tissue.

Let $X_i(\cdot)\ \forall i \in [1, N]$ be the measured signals that are the mixture of the signal sources $S_j(\cdot)\ \forall j \in [1, M]$:

$$X_i(\cdot) = \sum_{j=1}^{N} a_{ij} S_j(\cdot), \tag{36}$$

where a_{ij} are the mixing coefficient that weight the contribution of S_j to the measure X_i. The variation of a measuring parameter of X_i can induce different linear combinations of the sources, leading to:

$$\begin{bmatrix} X_1(\cdot) \\ \vdots \\ X_N(\cdot) \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{bmatrix} \begin{bmatrix} S_1(\cdot) \\ \vdots \\ S_M(\cdot) \end{bmatrix}, \tag{37}$$

which can be written as a matrix factorization: $\mathbf{X} = \mathbf{AS}$. Where the measures matrix, $\mathbf{X} \in \mathbb{R}^{N \times B}$, is known, and decomposed into the mixing matrix, $\mathbf{A} \in \mathbb{R}^{N \times M}$, and the sources matrix, $\mathbf{S} \in \mathbb{R}^{M \times B}$.

The dMRI signal can be expressed as a linear combination of the signals produced in the M tissue microstructure compartments, S_j. Moreover, a TE dependence of the measures X appears as a result of the T_2 differences in the sub-voxel compartments. Thus a compartmental solution of the Equation 24 is

$$X_i(TE_i, b, \vec{g}) = S_0 \sum_{j=1}^{N} f_j e^{-TE_i/T_{2_j}} S_j(b, \vec{g}). \tag{38}$$

From Equation 38 one can infer that the variation of TE yields distinct combinations of S_j resulting in different X_i. Therefore, mimicking Equation 37

$$\begin{bmatrix} X_1(TE_1,\cdot) \\ \vdots \\ X_N(TE_N,\cdot) \end{bmatrix} = S_0 \begin{bmatrix} f_1 e^{\frac{-TE_1}{T_{2_1}}} & \dots & f_M e^{\frac{-TE_1}{T_{2_M}}} \\ \vdots & \ddots & \vdots \\ f_1 e^{\frac{-TE_N}{T_{2_1}}} & \dots & f_M e^{\frac{-TE_N}{T_{2_M}}} \end{bmatrix} \begin{bmatrix} S_1(\cdot) \\ \vdots \\ S_M(\cdot) \end{bmatrix}, \quad (39)$$

where the influence of the compartmental volume (f_j) and relaxation (T_{2_j}) properties are encoded in the mixing matrix, while the diffusion properties are exclusively located in the sources matrix.

The direction of the columns of $\mathbf{A}$ is determined by the ratio between TE and T_2. Hence, after characterizing $\mathbf{A}$ and knowing the TE values used in the protocol

$$T_{2_i} = \frac{TE_k - TE_l}{\log(\frac{a_{li}}{a_{ki}})} \quad (40)$$

where a_{ki} and a_{li} are elements in the ith column of $\mathbf{A}$ for $TE_k < TE_l$. Once the T_2 values are computed as in Equation 40, and given that $S_j(b = 0, \vec{g}) = 1$, Equation 39 simplifies to

$$\begin{bmatrix} X_1(TE_1, b = 0) \\ \vdots \\ X_N(TE_N, b = 0) \end{bmatrix} = S_0 \begin{bmatrix} e^{\frac{-TE_1}{T_{2_1}}} & \dots & e^{\frac{-TE_1}{T_{2_M}}} \\ \vdots & \ddots & \vdots \\ e^{\frac{-TE_N}{T_{2_1}}} & \dots & e^{\frac{-TE_N}{T_{2_M}}} \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_M \end{bmatrix}, \quad (41)$$

where $\sum_{j=1}^{M} f_j = 1$, enables the estimation of f_j and S_0.

The family of BSS can be separated in four big groups. Principal Component Analysis (PCA) [35], Independent Component Analysis (ICA) [34], Sparse Component Analysis (SCA) [11], and Nonnegative Matrix Factorization (NMF) [6]. During this work all of them were considered. However, PCA and ICA were discarded in an early stage, since the properties of dMRI signals do not meet the requirements of these techniques (see section 4.1.1).

3.1.1 *Sparse Component Analysis*

The working principles of SCA [11] require the signals sources ($\mathbf{S}$) to be sparse and disjoint in a given transformed domain. When these conditions are met, the measured samples in the transformed space ($\mathbf{X}'$) organize themselves along lines. The direction of these lines are defined by the mixing coefficients (a_{ij}), enabling the characterization of the mixing matrix and thus, allowing the disentangling of

the sources. In different way, when $\mathbf{S}'$ are sparse and disjoint, only independent mixing coefficients are active at a time in the columns of $\mathbf{A}$, enabling its estimation.

The advantage of SCA over other BSS techniques falls in its independence from the number of signal sources: each line in the transformed domain represents a source. However, the full characterization of $\mathbf{A}$ requires the number of measures to be equal or lager to the number of sources ($M \geqslant N$), limiting the maximum number of detectable sources.

The disentangling effect of dMRI signal sources with SCA is investigated in section 4.2.1. The most important challenge for the usability of SCA is to find the right transformed domain that meets the sparsity and disjoint requirements. There is a dependence of the signal and the prescribed diffusion protocol that makes imposible to generalize a unique transformation. A solution based on wavelet transform was found for a protocol consisting of only one diffusion direction and an incremental b-value. Sections 3.1.2 and 3.1.3 present the generalization to any prescribed diffusion protocol.

3.1.2 *Nonnegative Matrix Factorization*

The theory of Nonnegative Matrix Factorization (NMF) [56] tackles the problem expressed in Equation 37 when the elements of $\mathbf{A}$ and $\mathbf{S}$ are nonnegative. Since only the magnitude of the dMRI signal is considered, and given that it is an attenuation based contrast, the nonnegativity requirement is satisfied by the data. A wide variety of solutions exist for this problem (see [73]). The Alternating Least Squares (ALS) algorithm [6] is chosen over others for its faster convergence. Let F be the target function

$$F = \frac{1}{2} \sum_{ij} \left(x_{ij} - a_{ij} s_{ij} \right)^2 . \tag{42}$$

The iterative solution for Equation 42 is given by algorithm 1.

However, infinite number of solutions exist for this problem due to a scaling factor that can shift from the columns of $\mathbf{A}$ to $\mathbf{S}$. To guarantee that only physically possible solutions are reached, extra constraints must be added based on Equation 39. This is shown in section 4.1.1 where relaxometry and diffusion constraints are incorporated to the algorithm 1 to produce the Constrained Alternating Least Squares (cALS). Specifically for two compartments (tissue and CSF), large convergence areas are found when the $T_{2_{tissue}} \in [0, 300]$ ms, $T_{2_{CSF}} = 2$ seconds, and $S_{CSF} = e^{-bD_{CSF}}$ with $D_{CSF} = 3 \times 10^{-3}$ mm^2/s.

The results shown in section 4.1.1 prove that tissue microstructure information can be extracted from dMRI data using BSS, and more pre-

Algorithm 1 Alternating Least Squares

```
procedure ALS(X)
    Randomize the positive defined A_{t=0} and S_{t=0}.
    while until convergence do
        Solve for S_{t+1} in A_t^T A_t S_{t+1} = A_t^T X
        Set all negative elements of S_{t+1} to 0.
        Solve for A_{t+1} in S_{t+1} S_{t+1}^T A_{t+1}^T = S_{t+1} X^T
        Set all negative elements of A_{t+1} to 0.
        t = t + 1
    end while
end procedure
```

cisely a physically constrained NMF algorithm. In this sense, PD, volume, and relaxometry compartmental effects can be separated. More important, diffusion signals from independent sub-voxel component are disentangled, paving the way for independent analysis. For instance. in a two compartments model, this approach introduces a new method for FWE.

There are two main disadvantages of this method. The first one comes from hardware limitations in clinical scanners. The minimum reachable TE 60 ms, impeding the inclusion of a third compartment corresponding to myelin water [47]. The second limitation comes from the condition of the mixing matrix. When $\mathbf{A}$ is bad-conditioned –due to similar compartmental T_{2_i} or similar TE_i– the cALS algorithm might be unstable due to numerical errors derived from the iterative inversion of $\mathbf{A}$.

3.1.3 *Nonnegative Sparse Coding*

Nonnegative Sparse Coding (NSC) [32, 43] was explored as a way to generalize the results achieved with SCA. The idea is to use sparseness as a regularization factor, reducing the number of physical restrictions. This technique imposes two constraints: nonnegativity (as in section 3.1.2), and sparseness (differently from section 3.1.1), defining the objective function

$$F = \frac{1}{2}\sum_{ij}\left(x_{ij} - a_{ij}s_{ij}\right)^2 + \lambda\sum_{kj} s_{kj} \tag{43}$$

where $a_{ik} \geqslant 0$, $s_{kj} \geqslant 0$, $\forall k$, $\sum_i a_{ik} = 1$, and $\lambda \geqslant 0$. The sparseness is control through λ.

The iterations over $\mathbf{A}$ and $\mathbf{S}$ to solve Equation 43 are defined by the algorithm 2.

However, signal diffusion sources are not sparse in the diffusion direction. Thus, a previous sparsifying step is needed for NSC to yield

Algorithm 2 Nonnegative sparse coding algorithm

procedure NSC($\mathbf{X}$)
  Randomize the positive defined $\mathbf{A}_{t=0}$ and $\mathbf{S}_{t=0}$.
  while until convergence **do**
    $\mathbf{A}' \leftarrow \mathbf{A}_t - \lambda(\mathbf{A}_t\mathbf{S}_t - \mathbf{X})\mathbf{S}_t^T$
    Set all negative elements of $\mathbf{A}'$ to 0.
    Normalize the columns of $\mathbf{A}'$
    $\mathbf{A}_{t+1} = \mathbf{A}'$
    $\mathbf{S}_{t+1} \leftarrow \mathbf{S}_t . * (\mathbf{A}_{t+1}^T\mathbf{X})./(\mathbf{A}_{t+1}^T\mathbf{A}_{t+1}\mathbf{S}_t + \lambda)$
    $t = t + 1$
  end while
end procedure

relevant results. In this work, instead of using an specific transform (e.g. FT), it is learnt from the data [62], allowing for protocol independence. The combination of NSC and sparsifying transform learning is discussed in section 4.2.2. The results in phantom and in vivo experiments indicate that this method is suitable for disentangling up to two compartments.

3.2 PATTERN MATCHING WITH DEEP LEARNING FOR FREE-WATER ELIMINATION

The partial volume contamination, specifically the one coming from free-water (section 1.1.4), poses a problem in dMRI. In some areas of the brain, this issue influences the diffusion signal, introducing a bias in the derived diffusion indexes (e.g. FA, MD, AD) [59].

This problem has been typically tackled with a two compartments approach: tissue and free-water (Equation 44). However, for the DTI model, this problem is ill-posed for acquisitions containing only one b-value (besides $b = 0$ mm^2/s). The state-of-the-art solutions regularize this problem [31, 57].

$$S(b, \vec{g}) = f_{tissue}S_{tissue}(b, \vec{g}) + (1 - f_tissue)S_{CSF}(b, \vec{g}). \quad (44)$$

In section 4.2.3 and appendix A.1.1 a deep learning pattern matching approach is developed, shifting the perspective of this problem. This is possible due to the good definition of the isotropic CSF or free-water: $S_{CSF}(b, \vec{g}) = e^{-bD_{CSF}}$ with $D_{CSF} = 3 \times 10^{-3}$ mm^2/s, which can be analytically calculated and mixed with random signals ($S_{tissue}(b, \vec{g})$) for random volume fractions (f_{tissue}), creating a theoretically infinite amount of training data. An ANN can learn from these synthetic data to recognize the free-water induced patterns and estimate the tissue volume fraction from the data.

After training the ANN with synthetically generated diffusion data, clinical measures are processed voxelwise with the trained model.

The knowledge of f_{tissue} and S_{CSF} allow to compute the diffusion signal of the tissue

$$S_{tissue}(b, \vec{g}) = \frac{S(b, \vec{g}) - (1 - f_{tissue})S_{CSF}(b, \vec{g})}{f_{tissue}}, \tag{45}$$

and fit it to any tissue model that do not account for isotropic diffusion, e.g. DTI. This enables the computation of free-water corrected maps of FA, which are shown to be relevant for patients with glioblastoma (annex A.1.1).

3.3 BAYESIAN METHODS FOR MULTIPARAMETER MICROSTRUCTURE ESTIMATION

Variations of the signal evolution described in Equation 34 measured in steady-state are use to extract tissue microstructure information. However, acquiring enough data to fit the mcDESPOT [19] model is a lengthy process. The introduction of QTI as an ultra-fast quantitative scheme paved the way to yield microstructural parameters from the transient-state evolution of the NMR signal. In this sense, Equation 34 can be extended for N compartments

$$y(\vec{k}, t) = \sum_{i=0}^{M} f_i y_i(\vec{k}, t), \tag{46}$$

where f_i and y_i are the volume fraction and the signal evolution of the ith component. Given that the prescribed acquisition protocol is the same for all the compartments ($\alpha(t)$, $TR(t)$, and $TE(t)$), it is necessary that the compartmental relaxation properties differ ($T_{2_i} \neq T_{2_j}$ and $T_{1_i} \neq T_{1_j}$ $\forall$ $i \neq j$) for them to be detected.

The estimation of the volume and relaxation parameters for one, two and three compartments is developed in Appendix A.2.1. A parallel high performance computing framework, named Π4U [29], quantifies the model parameters and uncertainty. To that end, Π4U uses a stochastic Bayesian inference algorithm, TMCMC [7, 16], that computes the posterior Probability Density Function (PDF) of the parameters ($\vec{\theta}$) given the data (D) and the model class (M).

$$p(\vec{\theta}|D, M) = \frac{p(D|\vec{\theta}, M)\pi(\theta|\vec{M})}{p(D|M)}, \tag{47}$$

where the prior PDF of the parameters, $\pi(\vec{\theta}|M)$, includes information on the uncertainty of the model parameters. The likelihood of observing the data from the model class is $p(D|\vec{\theta}, M)$, and the evidence of the mode is $p(D|M)$.

Let $\vec{g}(\vec{\theta}|M)$ be the predictions of the model and the measured data $D = \{\vec{h}\}$ fulfilling

$$\vec{h} = \vec{g}(\vec{\theta}|M) + \vec{e}, \tag{48}$$

where $\vec{e}$ accounts for all sources of error (measuring, computational, and modeling), and follows a Gaussian distribution with zero mean and Σ covariance matrix. Under this error condition, the likelihood is

$$p(D|\vec{\theta}, M) = \frac{|\Sigma(\vec{\theta})|^{-1/2}}{(2\pi)^{n/2}} exp\left[\frac{-J(\vec{\theta}, M)}{2}\right] \tag{49}$$

where

$$J(\vec{\theta}, M) = \left[\vec{h} - \vec{g}(\theta|\vec{M})\right]^T \Sigma^{-1}(\vec{\theta}) \left[\vec{h} - \vec{g}(\theta|\vec{M})\right] \tag{50}$$

quantifies the weighted error between the model fit and the measured data.

Stochastic methods based on Markov Chain Monte Carlo (MCMC) enable to draw samples from the posterior PDF, from which one can derive the marginal distribution of the parameters $\vec{\theta}$, allowing the computation of the evidence of the model.

To disentangle tissue microstructure, three model classes were defined by increasing M from one to three in Equation 46. The parameter space comprises $\vec{\theta} = \{T_{1_i}, T_{2_i}, f_i, \sigma^2\}\ \forall\ i \in [0, M]$, where σ^2 represents the noise variance, redefining the likelihood

$$p(D|\vec{\theta}, M) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\eta(D|\vec{\theta}, M)}{2\sigma^2}}, \tag{51}$$

where the error is defined as

$$\eta(D|\vec{\theta}, M) = ||\vec{h} - \vec{g}(\vec{\theta}|M)||_2^2 \tag{52}$$

The prior PDF of the parameters ($\pi(\vec{\theta}|M)$) was defined uniformly in a range of values informed by literature [19, 45, 47, 72]. In this way, the posterior PDF can be sampled in the parameter space, and marginal maximum likelihood values and uncertainty can be computed for each parameter.

The results show a smaller error variance for three compartments than for one and two, agreeing with previous findings [19, 47]. Moreover, although this approach showed some bias coming from the monocompartmental acquisition and reconstruction design, and noise model, the distribution and estimates of the components volume fraction and relaxometry parameters where compatible with the brain anatomy.

Part II

PUBLICATIONS

4

TISSUE MICROSTRUCTURE CHARACTERIZATION WITH DIFFUSION MRI

4.1 PEER-REVIEWED PUBLICATIONS

This chapter deals with compartmental tissue microstructure characterization from dMRI. The journal paper with title A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization, develops the theoretical framework that merges BSS and dMRI and shows simulations of convergence along with phantom and in vivo analysis. The results prove the clinical potential and limitations of this method, and establish the constrained ALS algorithm here developed as a robust technique to disentangle diffusion signals from sub-voxel tissue compartments.

4.1.1 *A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization*

Peer-reviewed Journal Paper

Authors: M. Molina-Romero, PA. Gómez, JI. Sperl, M. Czisch, PG. Sämann, DK. Jones, MI. Menzel, BH. Menze.

In: *Magn Reson Med.*, 00:1–18. https://doi.org/10.1002/ mrm.27181 (2018) [52].

Abstract: *Purpose*: The compartmental nature of brain tissue microstructure is typically studied by diffusion MRI, MR relaxometry or their correlation. Diffusion MRI relies on signal representations or biophysical models, while MR relaxometry and correlation studies are based on regularized inverse Laplace transforms (ILTs). Here we introduce a general framework for characterizing microstructure that does not depend on diffusion modeling and replaces ill-posed acILTs with blind source separation (BSS). This framework yields proton density, relaxation times, volume fractions and signal disentanglement, allowing for separation of the free-water component. *Theory and Methods*: Diffusion experiments repeated for several different echo times, contain entangled diffusion and relaxation compartmental information. These can be disentangled by BSS using a physically constrained nonnegative matrix factorization. *Results*: Computer simulations, phantom studies, together with repeatability and reproducibility experiments demonstrated that BSS is capable of estimating proton density, compartmental volume fractions and transverse relaxations. In vivo results proved its potential to correct for free-water contamination and to estimate tissue parameters. *Conclusion*: Formulation of the diffusion-relaxation dependence as a BSS problem introduces a new framework for studying microstructure compartmentalization, and a novel tool for free water elimination.

Contribution of thesis author: Model development and implementation, experimental analysis, manuscript preparation and editing.

A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization

Miguel Molina-Romero[1,2,*] Pedro A. Gómez[1,2] Jonathan I. Sperl[2] Michael Czisch[3] Philipp G. Sämann[3] Derek K. Jones[4,5] Marion I. Menzel[2] Bjoern H. Menze[1]

1 Computer Science, Technical University Munich, Munich, Germany

2 GE Global Research Europe, Garching, Germany

3 Max Planck Institute of Psychiatry, Munich, Germany

4 CUBRIC, Cardiff University, Cardiff, UK

5 School of Psychology, Faculty of Health Sciences, Australian Catholic University, Victoria, Australia

***** Corresponding author:

Name	Miguel Molina Romero
Department	Computer Science
Institute	Technical University Munich
Address	Boltzmannstr 11
	85748 Garching
	Germany
E-mail	miguel.molina@tum.de

Manuscript word count:	5000
Abstract word count:	200

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Abstract

Purpose: The compartmental nature of brain tissue microstructure is typically studied by diffusion MRI, MR relaxometry or their correlation. Diffusion MRI relies on signal representations or biophysical models, while MR relaxometry and correlation studies are based on regularized inverse Laplace transforms (ILTs). Here we introduce a general framework for characterizing microstructure that does not depend on diffusion modeling and replaces ill-posed ILTs with blind source separation (BSS). This framework yields proton density, relaxation times, volume fractions and signal disentanglement, allowing for separation of the free-water component.

Theory and Methods: Diffusion experiments repeated for several different echo times, contain entangled diffusion and relaxation compartmental information. These can be disentangled by BSS using a physically constrained non-negative matrix factorization.

Results: Computer simulations, phantom studies, together with repeatability and reproducibility experiments demonstrated that BSS is capable of estimating proton density, compartmental volume fractions and transversal relaxations. In vivo results proved its potential to correct for free-water contamination and to estimate tissue parameters.

Conclusion: Formulation of the diffusion-relaxation dependence as a BSS problem introduces a new framework for studying microstructure compartmentalization, and a novel tool for free water elimination.

Keywords: brain microstructure, diffusion MRI, blind source separation, free-water elimination, MR relaxometry, non-negative matrix factorization

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Introduction

More than fifty years have passed since Stejskal and Tanner published their early research on pulsed gradient spin-echo (PGSE) (1). Thereafter, diffusion weighted imaging (DWI) became an essential tool for non-destructive tissue microstructure characterization. The pioneering studies on ex vivo tissue and simulations of Krägger (2), Latour et al. (3), Szafer et al. (4), and Stanisz et al.(5) established the theoretical basis of the compartmental model of neural tissue.

These early contributions were later translated to target specific biomarkers for in vivo human studies. White matter (WM) anisotropy became fiber orientation with the introduction of diffusion tensor imaging (DTI) (6). The composite hindered and restricted model of diffusion MR imaging (CHARMED) (7) extended DTI to two compartments with restricted and hindered diffusion behavior. Using the same principles, the neurite orientation dispersion and density imaging (NODDI) model (8) introduced fiber orientation dispersion metrics and added an isotropic compartment. Additionally, axon diameter was addressed by AxCaliber (9) and ActiveAx (10). These and other approaches rely on diffusion signal representations or a variety of geometric biophysical assumptions about the underlying tissue compartments, producing a wide range of possible configurations (11).

In parallel with the development of multicomponent diffusion tissue models, relaxometry addressed the compartmental nature of tissue microstructure from a different perspective (12). Multi-echo spin echo (SE) experiments combined with regularized inverse Laplace transforms (ILTs) for multi-exponential fitting showed the presence of multiple water compartments in the tissue. Non-negative least squares (NNLS) (13) is the current gold standard for computing a regularized discrete ILTs for several components (14, 15). Alternatively, the exponential analysis via system identification using Steiglitz–McBride (EASI-SM) for multicomponent estimation was introduced by Stoika et al. (16, 17). Additionally, mcDESPOT (18), used a spoiled gradient-recalled echo and a balanced steady-state free precession to yield relaxation, volume fraction, and water exchange parameters for three compartments.

Nevertheless, the paths of diffusion MRI and MR relaxometry have become entangled over the years. Studies on ex vivo nerves with a diffusion-weighted Carr-Purcell-Meiboom-Gill (CPMG) sequence (19, 20) showed the relationship that existed between compartmental T_2 decay and diffusivity. However, diffusion-weighted CPMG experiments need long acquisition times and high specific absorption rates, which makes them unsuitable for human in vivo studies. Typically, two-dimensional ILTs were used to fit the data, but this approach is highly ill-posed and requires large

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

amounts of data for stabilization. Recently, Benjamini et al. (21) introduced the marginal distributions constrained optimization (MADCO), a non-CPMG compressed-sensing based solution that reduced the amount of data necessary for NMR diffusion-relaxation correlation experiments. Kim et al. translated diffusion-relaxation correlation spectroscopy (DR-COSY) (22, 23) into imaging (DR-CSI) (24) using spatial regularization to reduce the amount of necessary data and stabilize the ILTs. However, they require specific diffusion protocols with increasing b-values along a unique diffusion direction and repeated echoes or inversion times. Other alternatives combine diffusion models with multicompartmental relaxation. For instance, inversion recovery DWI has been used to identify fiber populations (25, 26), and WM integrity has been characterized using the axonal stick model and multiple echo times (TE) (27).

Compartmental analysis of the diffusion signal is intimately related to a recurring issue: cerebrospinal fluid (CSF) contamination (28, 29). All the existing contributions agree on using a bi-tensor signal model: parenchyma and CSF. However, this is an ill-posed problem for a single-shell and ill-conditioned for multiple-shell acquisitions (30). Spatial regularization was proposed by Pasternak et al. (31), relying on the local smoothness of the diffusion tensor. Later, a protocol optimization for multiple shells was presented by Hoy et al. (32), eliminating such a constraint. Other solutions regularize the problem by adding priors (33) or finding the best fit to the model (34). Nevertheless, the CSF contribution to the diffusion signal depends on the TE. Thus, disentangling the tissue CSF volume fraction requires an approach that includes T_2 compartmental dependencies (33, 35, 36).

We propose a general framework for studying diffusion and relaxation characteristics in tissue microstructures. We call it general because it does not model the compartmental diffusion behavior. It replaces the ILTs by a blind source separation (BSS) technique, reducing the minimum number of distinct echo times required to the number of compartments in the tissue, less than for ILTs-based methods. Other than the requirement to measure at more than one echo time, this framework is diffusion protocol-agnostic, and can be used in combination with any protocol of interest. Our approach quantifies proton density (PD), compartmental volume fractions, and transverse relaxation times. Importantly, it handles diffusion signals from each compartment independently, allowing for individual analyses, and thus performs CSF partial volume correction as a direct application.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Theory

Following the Bloch-Torrey equation, we describe the diffusion signal as a weighted sum of the signals from the compartments comprising the tissue:

$$X(TE, b, \mathbf{g}) = S_0 \sum_{i=1}^{M} f_i e^{-\frac{TE}{T_{2_i}}} S_i(b, \mathbf{g}). \quad [1]$$

Where b summarizes the gradient effects (1, 37) and $\mathbf{g}$ defines the gradient directions. Here, the compartmental diffusion sources $S_i(b, \mathbf{g})$ are weighted by their volume fraction, f_i, TE and T_{2_i}. The exponent (the ratio between TE and T_{2_i}) scales the contribution of each compartment to the acquired signal. Therefore, measuring at different TEs produces distinct diffusion signals (38) with different weights from the compartmental signal sources.

As a result, the signal of a single voxel measured with a protocol that accounts for multiple echoes can be formulated as:

$$\begin{bmatrix} X_1(TE_1, b, \mathbf{g}) \\ \vdots \\ X_N(TE_N, b, \mathbf{g}) \end{bmatrix} = S_0 \begin{bmatrix} f_1 e^{\frac{-TE_1}{T_{2_1}}} & \cdots & f_M e^{\frac{-TE_1}{T_{2_M}}} \\ \vdots & \ddots & \vdots \\ f_1 e^{\frac{-TE_N}{T_{2_1}}} & \cdots & f_M e^{\frac{-TE_N}{T_{2_M}}} \end{bmatrix} \begin{bmatrix} S_1(b, \mathbf{g}) \\ \vdots \\ S_M(b, \mathbf{g}) \end{bmatrix}, \quad [2]$$

where X_j ($j \in [1, N]$) are the diffusion signals acquired for the N TEs. f_i and T_{2_i} ($i \in [1, M]$) are the volume fraction and T_2 decay for the ith compartment, respectively, and M is the number of compartments.

Equation 2 can be expressed in matrix form as $\mathbf{X}$=$\mathbf{AS}$. This is a matrix factorization of the measurements, $\mathbf{X} \in \mathbb{R}_{\geq 0}^{N \times n}$, into two new matrices: the mixing matrix, $\mathbf{A} \in \mathbb{R}_{\geq 0}^{N \times M}$, which is defined by the experimental TEs, the compartmental volume fractions f, and T_2 decays; and the sources matrix, $\mathbf{S} \in \mathbb{R}_{\geq 0}^{M \times n}$, representing the diffusion sources in each sub-voxel compartment. Interestingly, we noticed from the definition of $\mathbf{A}$ that the ratio between the experimental TEs and T_{2_i} determines the direction (or slope for $N = 2$) of the ith column vector of the mixing matrix. Therefore:

$$T_{2_i} = \frac{TE_k - TE_l}{log(\frac{a_{li}}{a_{ki}})}, \quad [3]$$

where $TE_k < TE_l$, and a_{ki} and a_{li} are the kth and lth elements of the ith column of the mixing matrix, respectively.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Additionally, diffusion is an attenuation contrast and as such, $S(b=0)=1$, allowing Eq. 2 to be rewritten as

$$\begin{bmatrix} X_1(TE_1, b=0, \mathbf{g}) \\ \vdots \\ X_N(TE_N, b=0, \mathbf{g}) \end{bmatrix} = S_0 \begin{bmatrix} e^{\frac{-TE_1}{T_{2_1}}} & \cdots & e^{\frac{-TE_1}{T_{2_M}}} \\ \vdots & \ddots & \vdots \\ e^{\frac{-TE_N}{T_{2_1}}} & \cdots & e^{\frac{-TE_N}{T_{2_M}}} \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_M \end{bmatrix}, \quad [4]$$

which, together with $\sum_{i=1}^{M} f_i = 1$, allows us to solve for the volume fractions and proton density (f_i and S_0) when the number of measurements matches the number of compartments ($M = N$). Contrary, when there are more compartments than measurements ($M > N$), Eq. 4 is undetermined and f_i and S_0 cannot be estimated.

Factorizing $\mathbf{X}$ into $\mathbf{A}$ and $\mathbf{S}$ is known as blind source separation (BSS) (39) of mixed measurements into their generating sources (Figure 1). For BSS to identify these sources, they have to be distinct: $S_i \neq S_j \; \forall \; i \neq j$. Therefore, based on previous work (19, 20), we assumed them to be different.

There are four main approaches to BSS: principal component analysis (PCA) (40), independent component analysis (ICA) (41), non-negative matrix factorization (NMF) (42) and sparse component analysis (SCA) (43). PCA is not an applicable solution for this problem because the diffusion sources are not orthogonal. ICA assumes, as prior knowledge, that the signal sources are statistically independent and have non-Gaussian distributions. However, diffusion MRI signals are correlated with the tissue structure and temperature and they present non-Gaussian distributions only in restricted compartments, meaning that ICA is not suitable either. We previously explored SCA (44) and found that even though the results for simulations and real data for specific diffusion protocols were encouraging, finding a sparse and disjoint domain to meet the method's requirements was not always possible for arbitrary protocols. We observed the same issue for a version of NMF that enforces sparsity similarly (36).

In the present work, we took a BSS approach based on NMF (assuming $\mathbf{X}$, $\mathbf{A}$, and $\mathbf{S}$ are non-negative). Instead of depending on sparsity, we used a popular NMF solver: the alternating least squares algorithm (ALS) (42, 45, 46). We chose ALS instead of the multiplicative update algorithm (47) due to its faster convergence (48). We extended ALS to account for physically plausible limitations, resulting in Algorithm 1, which we refer to as constrained alternating least squares (cALS). Compartmental T_2 values available from the literature (15) allowed us to limit the solution space of the columns of $\mathbf{A}$ (Eq. 3). Additionally, for in vivo data, the diffusion behavior

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

of CSF is known to be approximately isotropic with 3×10^{-3} mm^2/s diffusivity (28), adding extra prior information. These constraints and priors make cALS converge toward physically realistic solutions (Figure 1).

Algorithm 1 Constrained Alternating Least Squares (cALS)

procedure cALS($\mathbf{X}$)
  Use priors on T_2 and experimental TEs to initialize the direction of the columns of $\mathbf{A}$ at the central T_2 value of the solution space of each column.
  while iter $<$ maximum iterations **do**
    Solve for $\mathbf{S}$ in $\mathbf{A}^T\mathbf{AS}=\mathbf{A}^T\mathbf{X}$. ▷ Least Squares.
    Set all negative elements of $\mathbf{S}$ to 0. ▷ Non-negativity.
    [Fix the one element of $\mathbf{S}$ to a known signal.] ▷ If analytical expression is known.
    Solve for $\mathbf{A}$ in $\mathbf{SS}^T\mathbf{A}^T=\mathbf{SX}^T$. ▷ Least Squares.
    Set all negative elements of $\mathbf{A}$ to 0. ▷ Non-negativity.
    Constrain the directions of the columns of $\mathbf{A}$. ▷ T_2 consistency.
    $error_i = \|A - SX\|^2$
    if $error_i < tolerance$ **then**
      break ▷ Check for data consistency.
    end if
    if $error_i >= error_{i-1}$ **then**
      break ▷ Check for convergence.
    end if
  end while
  return $\mathbf{A}$
end procedure

Constrained ALS initializes the column vectors of $\mathbf{A}$ at the central T_2 of their given constraints, avoiding random initializations in regions that are not physically feasible and increasing the stability. After each iteration, cALS verifies that the resulting T_2 of each column vector is between its boundaries, and sets it back to the center of its constrained solution space otherwise.

Following the factorization of $\mathbf{A}$, we estimated T_2 and f for each compartment, (Eqs. 3 and 4), and recalculated the real $\mathbf{A}$. This is important since the column norms of the factorized $\mathbf{A}$ do not tell us about the volume fractions. Then, $\mathbf{S}=\mathbf{A}^{-1}\mathbf{X}$ is calculated.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

An iterative algorithm like cALS inverts **A** repeatedly, requiring it to be non-singular and introducing a new condition. From Eq. 2, **A** is non-singular when $T_{2_i} \neq T_{2_j}\ \forall\ i \neq j$. Hence, in accordance with the literature (19, 20), we assumed that the transverse relaxation times for each compartment were distinct.

An open source implementation can be found in https://github.com/mmromero/dwybss.

Methods

Simulations

NMF is known for converging to local minima (45). Thus, it is necessary to asses the impact of the constraints. We ran simulations with Rician noise for signal-to-noise ratio (SNR) levels of 50, 100, and 150 at the non-diffusion weighted volume and minimum TE. We accounted for T_2 values, volume fractions, and diffusivities supported by literature (15, 28).

Two compartments

Two compartments were simulated mimicking IE and CSF water. The diffusion protocol included one non-diffusion weighted volume and 30 directions. We modeled diffusion as a Gaussian process (see Figure S4). For all the simulations we used $T_{2_{CSF}} = 2000$ ms, and varied $T_{2_{IE}}$ from 50–150 ms in 30 increments (15). Values of $f_{IE} = 0.25$, 0.5 and, 0.75 were used. We fixed $\text{TE}_1 = 60$ ms, and explored TE_2 from 70–150 ms in 31 increments. We defined $\Delta\text{TE} = \text{TE}_2 - \text{TE}_1$. The performance of the cALS algorithm was tested under the following conditions:

1. **Overlapped T_2 constraints:** $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded from 0–1000 and 0–3000 ms respectively, and no assumption on S_{CSF} was made (Figures 2 and S5).

2. **Overlapped T_2 constraints and prior S_{CSF}:** $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded from 0–1000 and 0–3000 ms respectively. CSF diffusivity was assumed to be isotropic with value 3×10^{-3} mm^2/s (Figure S10).

3. **Separated T_2 constraints:** $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded from 0–300 and 300–3000 ms respectively, and no assumption on S_{CSF} was made (Figure S11).

4. **Separated T_2 and prior S_{CSF}:** $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded from 0–300 and 300–3000 ms respectively. CSF diffusivity was assumed to be isotropic with value 3×10^{-3} mm^2/s (Figure

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

S13).

5. **Fixed $T_{2_{CSF}}$:** $T_{2_{IE}}$ was bounded from 0–300 ms. $T_{2_{CSF}}$ was fixed to 2000 ms. No assumption on S_{CSF} was made (Figure S12).

6. **Fixed $T_{2_{CSF}}$ and prior S_{CSF}:** $T_{2_{IE}}$ was bounded from 0–300 ms. $T_{2_{CSF}}$ was fixed to 2000 ms. CSF diffusivity was assumed to be isotropic with value 3×10^{-3} mm^2/s (Figures 3 and S6).

We repeated the last simulation for values of $f_{IE} = 0$ and 1, accounting only for IE or CSF (Figures 4 and S7).

Finally, intra-cellular (IC) and extra-cellular (EC) T_2 values are similar (15). We assessed the potential of BSS to separate them. Two diffusion signals were generated (see Figure S14). We used $f_{IC} = 0.25$, 0.5, and 0.75. The $T_{2_{IC}}$ vales ranged from 50–90 ms in 30 increments, and $T_{2_{EC}} = 100$ ms. TE_1 was fixed to 60 ms and TE_2 was varied between 70–150 ms in 31 increments. No assumption was made on the diffusion signals, and T_2 constraints were defined between 0–150 and 0–200 ms for IC and EC respectively (Figures 5 and S8).

We simulated 1000 times each combination of parameters, and reported the mean value of the absolute error of f, the relative error of T_2, and their standard errors (SEM).

Three compartments: searching for myelin

We incorporated a fast decaying component to model myelin, and fixed the T_2 of myelin (T_{2_M}) to 15 ms (15). $T_{2_{IE}}$ was varied from 50–150 ms in 30 increments, and $T_{2_{CSF}} = 2000$ ms. To account for short T2 components we needed to reduce the minimum TE of our simulations (see phantom experiments in the supporting material). Therefore, we fixed $TE_1 = 10$ ms, $TE_3 = 150$ ms, and varied TE_2 from 20–140 ms in 31 increments. We defined $\Delta TE = TE_2 - TE_1$. Three cases were explored: 1) $f_M = 0.1$, $f_{IE} = 0.6$; 2) $f_M = 0.2$, $f_{IE} = 0.5$; and 3) $f_M = 0.3$, $f_{IE} = 0.4$; keeping $f_{CSF} = 0.3$ for all of them. Simulations were run for two cases:

1. **Overlapped T_2 constraints:** T_{2_M}, $T_{2_{IE}}$, and $T_{2_{CSF}}$ were bounded from 0–40, 0–300, and 0–3000 ms respectively. No assumption on S_{CSF} was made.

2. **Separated T_2 constraints, fixed $T_{2_{CSF}}$ and prior S_{CSF}:** T_{2_M} and $T_{2_{IE}}$ were bounded from 0–40 and 41–300 ms respectively, while $T_{2_{CSF}} = 2000$ ms. CSF diffusivity was assumed to be isotropic with value 3×10^{-3} mm^2/s (Figures 6 and S9).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Each combination of parameters was simulated 1000 times. The mean value of the absolute error of f, the relative error of T_2, and their SEM were reported.

In vivo clinical data: free-water elimination

We aim to show that BSS has potential applications in clinical settings. To this end, we ran an experiment to analyze its performance for estimating tissue parameters and correcting for CSF contamination.

Data acquisition

Two volunteers, a male (age 28 years) and a female (age 24 years) were scanned in a 3.0 T GE MR750w (GE Healthcare, Milwaukee, WI). The in vivo study protocol was approved by our institutional review board and prior informed consent was obtained. We acquired seven diffusion PGSE echo planar imaging (EPI) volumes for TE values from 75.1–135.1 ms in 10 ms increments. The following parameters were constant: FOV = 240 mm; 4 mm slice thickness; TR = 6000 ms; 96 × 96 matrix size; ASSET = 2; and 30 directions. Additionally, we measured fluid-attenuated inversion recovery (FLAIR) SE EPIs for 17 equally-spaced TEs ranging from 20–260 ms. The same imaging parameters were used as for the diffusion experiments but with no acceleration (ASSET = 0).

Data analysis

Diffusion data for all TEs were first registered with FSL FLIRT (49) to the shortest TE volume. We then processed them with BSS in pairs ($M = N = 2$) with a fixed short TE of 75.1 ms. The long TE was increased from 85.1 to 135.1 ms for a total ΔTE of 60 ms (Figures 7 and 8). We used literature CSF values ($T_{2_{CSF}} = 2$ s and $D_{CSF} = 3 \cdot 10^{-3}$ mm^2/s) as the prior knowledge, and constrained the possible values of $T_{2_{IE}}$ between 0–200 ms (15, 28). We report maps of the BSS relative factorization error (Figure 7a, 7b and 7g), CSF volume fraction (Figure 7c and 7h), proton density (Figure 7d and 7i), $T_{2_{IE}}$ (Figure 7e and 7j) and number of compartments (Figure 7f and 7k).

For reference, FLAIR multi-echo EPI data were also registered with FLIRT to the shortest TE non-diffusion weighted volume. The signal decay for each voxel was then matched to a dictionary of mono-exponential decays from 0–300 ms with a grid of 1 ms. We compared this map against the BSS $T_{2_{IE}}$ map (Figure 8).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

We defined the relative error of the matrix factorization for the in vivo data as follows:

$$\epsilon = \frac{|\mathbf{X} - S_0\mathbf{AS}|_2}{|\mathbf{X}|_2}. \quad [5]$$

This is a measure of the performance of BSS for each voxel. Given that we calculated $\mathbf{S}{=}\mathbf{A}^{-1}\mathbf{X}$, this error formulation is sensitive to: 1) breaches of the BSS conditions due to artifacts, and 2) numerical instabilities due to the condition of $\mathbf{A}$. Point one is the result of B_0 drift, subject motion, flow, and eddy currents. These effects produce a violation of the BSS condition, making the signal sources different between TE measurements. The second point is the error amplification factor. A high ϵ denotes that the factorization could not find a solution within the constrained space and thus, results might not be trustworthy.

Finally, BSS does not model the compartmental diffusion signal. However, to demonstrate a simple way to perform compartment-independent analysis and correct for CSF contamination, we fitted the measured and disentangled signals to the DTI model (6). We fitted the measured diffusion volumes at the shortest TE, and the BSS separated signals for the IE and CSF compartments to a mono-exponential model using standard linear regression (FSL FDT Toolbox (http://www.fmrib.ox.ac.uk/fsl)). For comparison, bi-exponential models using Pasternak's and Collier's methods were used (Figures 9, S15 and 10). Fractional anisotropy (FA) and mean diffusivity (MD) maps were derived for each fit.

Results

Simulations

Two Compartments

The convergence area is the region where the mean relative error of $T_{2_{IE}}$ is lower than 0.1 per unit (p.u). Its shape for all the simulations (Figures 2, 3, 4, 5, S5, S6, S7, S8, S10, S11, S12, and S13) follows two effects. First, the condition number of the mixing matrix limits the lower bound of ΔTE – similar TE values produce more linearly dependent column vectors of $\mathbf{A}$ –. And second, the SNR plays a double role, it increases the error regions where $\mathbf{A}$ is bad-conditioned (small ΔTE), and limits the maximum ΔTE due to the T_2 decay of the signals. Thus, when the SNR increases the convergence area grows and the region of minimum SEM, denoting an improvement on the stability of the algorithm. The convergence area also depends on the IE volume fraction. The larger is the

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

contribution of IE, the better is th $T_{2_{IE}}$ estimate.

Adding priors on S_{CSF} improves the $T_{2_{IE}}$ estimate, even at SNR = 50 (Figure S10). Bounding the solution space into non-overlapping regions also improves the results of $T_{2_{IE}}$ (Figure S11), although less than combining it with CSF prior knowledge (Figure S13). The $T_{2_{CSF}}$ estimate shows a 0.17 p.u. due to the small variation of S_{CSF} along the acquired TEs (4.4 %). This is corrected when relaxometry prior is incorporated (Figures 3 and S12). The comparison between Figure 2 and 3, show the benefit of including prior knowledge into the factorization algorithm, specially at low SNR. Then, the accuracy of the estimates will be influenced by the selection of ΔTE, the T_2 boundaries, the S_{CSF} prior, and the expected $T_{2_{IE}}$ and f_{IE} values. We used literature values for $T_{2_{IE}}$, $T_{2_{CSF}}$ (15), and S_{CSF} (28). According to Figures 3a and 3b one needs a minimum ΔTE of 26 ms for an accurate f_{IE} estimate. Interestingly, f_{IE} is a reliable parameter that tell us about the bias of $T_{2_{IE}}$, the larger f_{IE} is, the more accurate $T_{2_{IE}}$ becomes (3a and 3c).

For one tissue compartment BSS is able to precisely (SEM < 0.01) estimate the volume fraction with mean absolute error below 0.1 when ΔTE > 35 ms (Figure 4a and 4b). When $f_{IE} = 1$ the area of mean convergence of the $T_{2_{IE}}$ estimate is almost independent from ΔTE (4c and 4d). We found an equivalent result for the mean relative error of $T_{2_{CSF}}$ when $f_{IE} = 0$ (4e and 4f), although in this case it comes from the $T_{2_{CSF}}$ prior. Notice the large error and instability of $T_{2_{IE}}$ and $T_{2_{CSF}}$ in the opposite cases, $f_{IE} = 0$ and $f_{IE} = 1$ respectively (Figures 4c and 4e). This results when BSS tries to find a component that is not in the tissue and thus, cannot be estimated.

For two components with similar T_2 values and little priors (IC and EC) cALS losses efficiency. The volume fraction estimates are biased (Figure 5a), and $T_{2_{IC}}$ shows a narrow convergence region that is almost independent of ΔTE. The lower bound of this region is limited by the proximity of $T_{2_{IC}}$ and $T_{2_{EC}}$ that worses the condition of $\mathbf{A}$. The upper bound results of the lack of prior on the signal of one of the compartments, in contrast with the S_{CSF} prior used before (compare Figures 2 and S10) that increased the convergence area towards lower T_2 values.

Three Compartments: searching for myelin

The convergence area is the one where the errors of f_M, f_{IE}, T_{2_M}, and $T_{2_{IE}}$ are lower than 0.1 in absolute value for the volume fractions and per unit for T_2. Figures 6a, 6c, 6e, and 6g show and optimal ΔTE = 36 ms. Notice that when ΔTE increases the error of the myelin parameters grows due to the reduction of the myelin contribution to the second TE, worsening the SNR of that component (Figures 6a and 6e). Since all the volume fractions add up to one, errors on f_M increase

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

the error on f_{IE} (Figures 6a and 6c). The estimate of $T_{2_{IE}}$ is dependent on SNR and its volume fraction, compounding its calculation for SNR < 50 and $f_{IE} < 0.4$ (Figure S9g lower left corner).

One should notice that including a third compartment increases the condition number of $\mathbf{A}$, rising the instability of the factorization (Figure 6f). See the phantom experiments in the supporting material.

In vivo clinical data: free-water elimination

We observed that the mean relative error for the whole brain ($\langle\epsilon\rangle$) decreased as ΔTE increased (Figures 7a, 7b, and 7g), in agreement with phantom findings (see supporting material) and the results of the simulations for two compartments. Interestingly, for the maximum ΔTE, we can see that the number of compartments is two in regions next to the ventricles and the cortex, but one inside the ventricles and in some deep WM areas (Figure 7k). It is also noteworthy that the pure CSF areas (e.g., the ventricles) have been removed from the $T_{2_{IE}}$ map (Figures 7e and 7j), while the opposite is observed in the CSF volume fraction (Figures 7c and 7h), indicating a successful disentangling effect.

We compared the BSS-estimated $T_{2_{IE}}$ maps for increasing ΔTE values with the reference map obtained from the FLAIR multi-echo SE data. We noted how the structural similarity index (50) increased and the mean relative error decreased as ΔTE grew (Figure 8a and 8b). Additionally, the histograms for both subjects tended toward the reference as the difference between the short and long TEs grew. This reflects an underestimation of $T_{2_{IE}}$ for small ΔTE values that can be explained by Eq. 3 and Figure S1c. Moreover, the FLAIR T_2 map showed high values in the ventricles, possibly indicating imperfect CSF suppression and, thus, slightly increased reference values (Figure 8a, 8c, and 8d).

FA and MD maps and histograms were calculated from the BSS IE and CSF disentangled signals for both subjects (Figures 9, S15, and 10). These maps displayed an overestimation of the CSF volume fraction for low ΔTE values (the low FA peak in Figures 9b and S15b was removed). This resulted in a compensation effect for the previously shown underestimation of $T_{2_{IE}}$. Additionally, the FA histograms (Figures 9b and S15b) showed a tendency toward higher FA values and a reduction of the low FA peak associated with free water. At long ΔTE values, FA seems to tend toward a stable distribution. We also observed an enlargement of the corpus callosum and a general recovery of peripheral WM tracts and the fornix in the colored FA maps (Figures 9a and S15a).

Additionally, on the MD histograms for IE water (Figures 9d and S15d) we found a reduced

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

number of voxels with diffusivities greater than $1{\times}10^{-3}$ mm^2/s. In contrast, the main peak at $0.7{\times}10^{-3}$ mm^2/s, associated with the parenchyma, remained in its original position, indicating that IE water represents a non-CSF tissue. This MD reduction was also visible in the maps (Figure 9c and S15c). Finally, the MD histograms for CSF water (Figure 10) showed a tendency toward 3×10^{-3} mm^2/s as ΔTE increased, in agreement with the literature (28). All these findings agreed with a disentangling of IE and CSF signals and thus, a correction of the free water partial volume effect in the diffusion signal.

Discussion

Stability

Four main approaches exist for the BSS problem (ICA, PCA, NMF, and SCA). Choosing the appropriate method depends on the prior knowledge of the signal sources. In our experiments, we relied on NMF, using a constrained version of the ALS algorithm (cALS). Others explored these algorithms before. Pauca et al. (51) used low-rank and sparsity constraints to distinguish semantic features in text mining, and later (52) smoothness regularization to identify space objects from spectral data. Gao and Church (53) also employed sparseness for cancer class discovery through gene clustering, which was later extended by Kim and Park (54) improving the balance between accuracy and sparseness through regularization. They also introduced a variation based on the active set method (55) and low-rank approximation (56). Liu et al. (57) incorporated label information to create a semi-supervised matrix decomposition method. Sun and Févotte (58) introduced a version based on the alternating direction method of multipliers (59) (ADMM), that was further stabilized by Zhang et al. (60).

Supported by previous work, we presented a biophysical inspired solution to constrain the diffusion-relaxometry NMF compartmental problem. Essentially, our cALS algorithm imposes two constraints: 1) the rows of $\mathbf{A}$ must follow exponential relationships (relaxometry); and 2) when the analytical expression of one component is known (i.e. CSF) the corresponding row in $\mathbf{S}$ is fixed (diffusion). The stability of cALS is linked to the condition of $\mathbf{A}$ and SNR; an ill-conditioned mixing matrix will lead to error propagation due to numerical instability. We optimized the experimental TEs to reduce the condition number of $\mathbf{A}$ for literature values of T_2. However, further research based on ADMM might yield better results.

We ran extensive simulations for two compartments at clinical TE values with different priors,

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

and three compartments at lower TEs. These simulations highlighted the importance of choosing literature supported priors to improve the convergence, especially at low SNR. Constrained ALS converges when the number of compartments in tissue is equal or lower than the expected, but it looses performance for species with similar T_2.

Phantom experiments (see supporting material) agreed with simulation results, validating that BSS was able to accurately estimate T_2 for one compartment and separate diffusion signal sources and estimate T_2 and f for two compartments. However, they also showed that scaling the cALS algorithm to three compartments, including fast T_2 decaying species, is unstable in the range of the clinically available TE values.

Finally, repeatability and reproducibility analyses (see supporting material) show that cALS yield consistent results across repetitions and subjects, highlighting its stability.

Relaxation time and volume fraction estimates

BSS provides the means to estimate T_2 relaxation values and volume fractions. Interestingly, only a number of TE repetitions equal to the number of compartments that are assumed to be in the tissue is necessary. This results of the substitution of the ILTs by BSS, in comparison to other techniques (15, 17, 21, 24). We found a good agreement between the $T_{2_{IE}}$ estimates of the FLAIR multi-echo SE for 17 TEs and those of BSS for 2 TEs. In this sense, all the measurements along the diffusion space are considered for both TEs, incorporating redundancy and reinforcing the estimation of T_2. The SNR for the in vivo data were 147 and 104 for subjects one and two. According to the simulations at $\Delta\text{TE} = 60$ ms, the expected absolute error for the volume fraction estimate is below 0.03, meaning that $T_{2_{IE}}$ is highly reliable in white matter areas, and lesser in the CSF borders.

Myelin detection

Simulations proved that our method has the potential to disentangle three compartments by reducing the minimum TE in diffusion experiments. As a result, myelin water could be incorporated into the model (Figure 6). However, we are prevented from conducting such experiments by gradient performance on clinical scanners.

Disentangling the diffusion sources and free water elimination

Unlike other multicompartment diffusion models (2, 7, 8, 11) or more recent contributions (27, 35), our approach does not model compartmental diffusion. Our framework instead relies on three

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

assumptions: 1) microstructural water compartments have distinct T_2 relaxation times (14, 15); 2) each have different diffusion characteristics (19, 20); and 3) the effects of the water exchange are negligible on the timescale of our experiments (9, 61). Furthermore, our solution is diffusion protocol-agnostic (only two TEs and one non-diffusion weighted volume are necessary), allowing for flexibility in the design of the acquisition protocol, which might include any number of diffusion directions and b-values. This gives it an advantage over diffusion-relaxation correlation techniques based on regularized inverse Laplace transforms (21, 24).

A promising application of the protocol-agnostics nature of our framework is correcting for free water contamination. Recently Collier et al. (35) included TE dependence in their bi-exponential diffusion tensor model to regularize the fitting problem. However, they fitted the bi-exponential DTI model directly. Contrary, our solution does not assume any particular diffusion model, we instead separated the signal from each compartment, allowing more flexible and independent study. In this regard, analysis of the signal associated with the CSF compartment can be seen as a disentanglement quality assurance metric (Figures 9, S15, and 10), or in brain tissue applications, a general indicator of the goodness-of-fit for IE and CSF.

We fitted our data to Collier's model (35) without reaching convergence, which resulted due to our single-shelled dataset. Comparison of BSS with Pasternak's free-water elimination (FWE) method (31) is show in Figures 9 and S15. We observed a good agreement between BSS for ΔTE = 60 ms and Pasternak's FWE for FAs between 0–0.2 and 0.8–1. In the middle FA range both methods disagree, BSS shows an homogeneous correction, while Pasternak's results follow the standard DTI fit from 0.2 to 0.4 and shows a correcting effect from 0.4 to 1 (Figures 9a, 9b, S15a, and S15b). It is impossible to determine which method is better (no ground-truth). However, there are two indicators that BSS might be performing better: 1) the BSS FA curve runs in parallel to the standard DTI fit from 0.2 to 0.8, denoting an stable correction without favoring any FA range; and 2) Pasternak's MD is spatially over-regularized (Figures 9c, 9d, S15c, and S15d), while BSS's MD keeps its maximum at 0.7 mm^2/s, the reference for parenchyma (28).

Long ΔTE values benefit our framework, which is not surprising and agrees with the findings of Collier et al. (35). This is not only due to the relationship between $\mathbf{A}$ and T_2 (Eq. 3 and Figure S1c) but also because longer differences between TEs produce more distinct levels of mixing and thus better codification of the information from each source. That is to say, the short TE contains more information about the fast-relaxing species, while the long TE is dominated by CSF.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Conclusions

We have introduced for the first time a blind source separation framework for expressing the relationships between diffusion signals acquired at different TEs. This new approach does not rely on diffusion modeling or the inverse Laplace transform. Our results show that, with the current hardware, blind source separation allows for disentangling the diffusion signal sources generated by each sub-voxel compartment up to two compartments, making it a suitable tool for free-water elimination. Moreover, it simultaneously estimates proton density, volume fractions, relaxation times and the number of compartments in the underlying microstructure, paving the way for tissue microstructure characterization when the hardware constraints are relieved.

Acknowledgments

The authors want to thank Dr. Ofer Pasternak for his support in the comparison of the methods. This work was supported by the TUM Institute of Advanced Study, funded by the German Excellence Initiative, and the European Commission (Grant Agreement Number 605162). DKJ was supported by a New Investigator Award from the Wellcome Trust (096646/Z/11/Z).

References

[1] Stejskal EO, Tanner JE. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. J. Chem. Phys. 1965. 42:288–292.

[2] Kärger J. Der Einfluss der Zweibereichdiffusion auf die Spinechodaempfung unter Beruecksichtigung der Relaxation bei Messungen mit der Methode der gepulsten Feldgradienten. Annalen der Physik. 1971. 7:107–109.

[3] Latour LL, Svoboda K, Mitra PP, Sotak CH. Time-dependent diffusion of water in a biological model system. Proc. Natl. Acad. Sci. U. S. A. 1994. 91:1229–1233.

[4] Szafer a, Zhong J, Gore JC. Theoretical model for water diffusion in tissues. Magn. Reson. Med. 1995. 33:697–712.

[5] Stanisz GJ, Szafer A, Wright GA, Henkelman RM. An Analytical Model of Restricted Diffusion in Bovine Optic Nerve. Magn. Reson. Med. 1997. 37:103–111.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

[6] Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. Biophys. J. 1994. 66:259–267.

[7] Assaf Y, Basser P. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage. 2005. 27:48–58.

[8] Zhang H, Schneider T, Wheeler-Kingshott CAM, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage. 2012. 61:1000–1016.

[9] Assaf Y, Blumenfeld-Katzir T, Yovel Y, Basser PJ. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. Magn. Reson. Med. 2008:1347–1354.

[10] Alexander DC, Hubbard PL, Hall MG, Moore EA, Ptito M, Parker GJM, Dyrby TB. Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage. 2010. 52:1374–1389.

[11] Ferizi U, Schneider T, Panagiotaki E, Nedjati-Gilani G, Zhang H, Wheeler-Kingshott CAM, Alexander DC. A ranking of diffusion MRI compartment models with in vivo human brain data. Magn. Reson. Med. 2014. 72:1785–92.

[12] Kimmich R. NMR : Tomography, Diffusometry, Relaxometry. Springer Berlin Heidelberg. 1997.

[13] Lawson CL, Hanson RJ. Solving Least Squares Problems (Classics in Applied Mathematics). 1987.

[14] Whittall KP, Mackay AL, Graeb DA, Nugent RA, Li DKB, Paty DW. In vivo measurement of T_2 distributions and water contents in normal human brain. Magn. Reson. Med. 1997. 37:34–43.

[15] MacKay AL, Laule C, Vavasour I, Bjarnason T, Kolind S, Mädler B. Insights into brain microstructure from the T_2 distribution. Magn. Reson. Med. 2006. 24:515–25.

[16] Stoica P, Babu P. Parameter estimation of exponential signals: A system identification approach. Digit. Signal Process. 2013. 23:1565–1577.

[17] Björk M, Stoica P. New approach to phase correction in multi-echo T2 relaxometry. J. Magn. Reson. 2014. 249:100–107.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

[18] Deoni SCL, Rutt BK, Arun T, Pierpaoli C, Jones DK. Gleaning multicomponent T_1 and T_2 information from steady-state imaging data. Magn. Reson. Med. 2008. 60:1372–1387.

[19] Peled S, Cory DG, Raymond SA, Kirschner DA, Jolesz FA. Water diffusion, T_2, and compartmentation in frog sciatic nerve. Magn. Reson. Med. 1999. 42:911–8.

[20] Does MD, Gore JC. Compartmental study of diffusion and relaxation measured in vivo in normal and ischemic rat brain and trigeminal nerve. Magn. Reson. Med. 2000. 43:837–44.

[21] Benjamini D, Basser PJ. Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry. J. Magn. Reson. 2016. 271:40–45.

[22] Callaghan PT, Arns CH, Galvosas P, Hunter MW, Qiao Y, Washburn KE. Recent Fourier and Laplace perspectives for multidimensional NMR in porous media. Magn. Reson. Imaging. 2007. 25:441–444.

[23] Hurlimann MD, Venkataramanan L, Flaum C. The diffusion-spin relaxation time distribution function as an experimental probe to characterize fluid mixtures in porous media. J. Chem. Phys. 2002. 117:10223–10232.

[24] Kim D, Doyle EK, Wisnowski JL, Kim JH, Haldar JP. Diffusion-Relaxation Correlation Spectroscopic Imaging: A Multidimensional Approach for Probing Microstructure. Magn. Reson. Med. 2017. doi:10.1002/mrm.26629.

[25] De Santis S, Barazany D, Jones DK, Assaf Y. Resolving relaxometry and diffusion properties within the same voxel in the presence of crossing fibres by combining inversion recovery and diffusion-weighted acquisitions. Magn. Reson. Med. 2016. 75:372–380.

[26] De Santis S, Assaf Y, Jeurissen B, Jones DK, Roebroeck A. T1 relaxometry of crossing fibres in the human brain. Neuroimage. 2016. 141:133–142.

[27] Veraart J, Fieremans E, Novikov DS. Quantifying neuronal microstructure integrity with TE dependent Diffusion Imaging (TEdDI). In: Proc. 25th Annu. Meet. ISMRM, Honolulu. 2017. p. 0836.

[28] Pierpaoli C, Jones DK. Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model. In: Proc. 12th Annu. Meet. ISMRM, Kyoto. 2004. p. 1215.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

[29] Metzler-Baddeley C, O'Sullivan MJ, Bells S, Pasternak O, Jones DK. How and how not to correct for CSF-contamination in diffusion MRI. Neuroimage. 2012. 59:1394–1403.

[30] Bergmann O, Westin CF, Pasternak O. Challenges in solving the two-compartment free-water diffusion MRI model. In: Proc. 24th Annu. Meet. ISMRM, Singapore. 2016. p. 0793.

[31] Pasternak O, Sochen N, Gur Y, Intrator N, Assaf Y. Free Water Elimination and Mapping from Diffusion MRI. Magn. Reson. Med. 2009. 62:717–730.

[32] Hoy AR, Koay CG, Kecskemeti SR, Alexander AL. Optimization of a Free Water Elimination Two-Compartment Model for Diffusion Tensor Imaging. Neuroimage. 2014:323–333.

[33] Vallée E, Douaud G, Monsch AU, Gass A, Wu W, Smith SM, Jbabdi S. Modelling free water in diffusion mri. In: Proc. 23rd Annu. Meet. ISMRM, Toronto. 2015. p. 0474.

[34] Henriques RN, Bergmann Ø, Rokem A, Pasternak O, Correia MM. Exploring the potentials and limitations of improved free-water elimination DTI techniques. In: Proc. 25th Annu. Meet. ISMRM, Honolulu. 2017. p. 1787.

[35] Collier Q, Veraart J, Dekker AJD, Vanhevel F, Parizel PM, Sijbers J. Solving the free water elimination estimation problem by incorporating T_2 relaxation properties. In: Proc. 25th Annu. Meet. ISMRM, Honolulu. 2017. p. 1783.

[36] Molina-Romero M, Gómez PA, Sperl JI, Stewart AJ, Jones DK, Menzel MI, Menze BH. Theory, validation and application of blind source separation to diffusion MRI for tissue characterisation and partial volume correction. In: Proc. 25th Annu. Meet. ISMRM, Honolulu. 2017. p. 3462.

[37] Bihan DL, Breton E. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. Comptes-Rendus l'Académie des Sci. 1985. 93:27–34.

[38] De Santis S, Assaf Y, Jones DK. The influence of T2 relaxation in measuring the restricted volume fraction in diffusion MRI. In: Proc. 24th Annu. Meet. ISMRM, Singapore. 2016. p. 1998.

[39] Yu X, Hu D, Jindong X. Blind source separation: theory and applications. Science Press. 2014.

[40] Jolliffe I. Principal component analysis. New York: Springer Verlag. 2002.

[41] Hyvarinen A, Oja E. Independent Component Analysis : A Tutorial. Neural Networks. 1999. 1:1–30.

[42] Paatero P, Tapper U. Positive Matrix Factorization - A Nonnegative Factor Model With Optimal Utilization of Error-Estimates of Data Values. Environmetrics. 1994. 5:111–126. doi: 10.1002/env.3170050203.

[43] Bofill P, Zibulevsky M. Underdetermined blind source separation using sparse representations. Signal Processing. 2001. 81:2353–2362.

[44] Molina-Romero M, Gómez PA, Sperl JI, Jones DK, Menzel MI, Menze BH. Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis. In: Workshop on Breaking the Barriers of Diffusion MRI, Lisbon. 2016. p. 17.

[45] Berry MW, Browne M, Langville AN, Pauca VP, Plemmons RJ. Algorithms and applications for approximate nonnegative matrix factorization. Comput. Stat. Data Anal. 2007. 52:155–173.

[46] Lin CJ. Projected Gradient Methods for Nonnegative Matrix Factorization. Neural Comput. 2007. 19:2756–2779. doi:10.1162/neco.2007.19.10.2756.

[47] Lee D, Seung H. Algorithms for non-negative matrix factorization. In: Adv. Neural Inf. Process. Syst.. 1. 2001. p. 556–562. doi:10.1109/IJCNN.2008.4634046. `0408058v1`.

[48] On the Convergence of Multiplicative Update Algorithms for Nonnegative Matrix Factorization. IEEE Trans. Neural Networks. 2007. 18:1589–1596. doi:10.1109/TNN.2007.895831.

[49] Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 2001. 5:143–156.

[50] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 2004. 13:600–612.

[51] Pauca V, Shahnaz F, Berry M, Plemmons R. Text mining using non-negative matrix factorizations. In: Proc. SIAM Int. Conf. Data Min. 2004. p. 452–456.

[52] Pauca VP, Piper J, Plemmons RJ. Nonnegative matrix factorization for spectral data analysis. Linear Algebra Appl. 2006. 416:29–47. doi:10.1016/j.laa.2005.06.025.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

[53] Gao Y, Church G. Improving molecular cancer class discovery through sparse non-negative matrix factorization. Bioinformatics. 2005. 21:3970–3975. doi:10.1093/bioinformatics/bti653.

[54] Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics. 2007. 23:1495–1502. doi:10.1093/bioinformatics/btm134.

[55] Nocedal J, Wright SJ, SpringerLink (Online service). Numerical Optimization. 2006. doi:10.1007/978-0-387-40065-5. NIHMS150003.

[56] Kim H, Park H. Nonnegative Matrix Factorization Based on Alternating Nonnegativity Constrained Least Squares and Active Set Method. SIAM J. Matrix Anal. Appl. 2008. 30:713–730. doi:10.1137/07069239X.

[57] Liu H, Wu Z, Cai D, Huang TS. Constrained Nonnegative Matrix Factorization for Image Representation. IEEE Trans. Pattern Anal. Mach. Intell. 2012. 34:1299–1311. doi:10.1109/TPAMI.2011.217.

[58] Sun DL, Févotte C. Alternating direction method of multipliers for non-negative matrix factorization with the beta-divergence. IEEE Int. Conf. Acoust. Speech Signal Process. 2014:6201–6205. doi:10.1109/ICASSP.2014.6854796.

[59] Boyd S, Parikh N, E Chu BP, Eckstein J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Found. Trends Mach. Learn. 2011. 3:1–122. doi:10.1561/2200000016.

[60] Zhang S, Huang D, Xie L, Chng ES, Li H, Dong M. Non-negative Matrix Factorization using Stable Alternating Direction Method of Multipliers for Source Separation. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. 2015:222–228.

[61] Sakka L, Coll G, Chazal J. Anatomy and physiology of cerebrospinal fluid. Eur. Ann. Otorhinolaryngol. Head Neck Dis. 2011. 128:309–316.

[62] Laubach HJ, Jakob PM, Loevblad KO, Baird AE, Bovo MP, Edelman RR, Warach S. A phantom for diffusion-weighted imaging of acute stroke. J Magn Reson Imaging. 1998. 8:1349–1354.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

[63] Andersson JL, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage. 2003. 20:870–888. doi:10.1016/S1053-8119(03)00336-7.

[64] Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004. 23:S208–S219. doi: 10.1016/j.neuroimage.2004.07.051.

[65] Andersson JL, Sotiropoulos SN. Non-parametric representation and prediction of single- and multi-shell diffusion-weighted MRI data using Gaussian processes. Neuroimage. 2015. 122:166–176. doi:10.1016/j.neuroimage.2015.07.067.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Figures and Tables

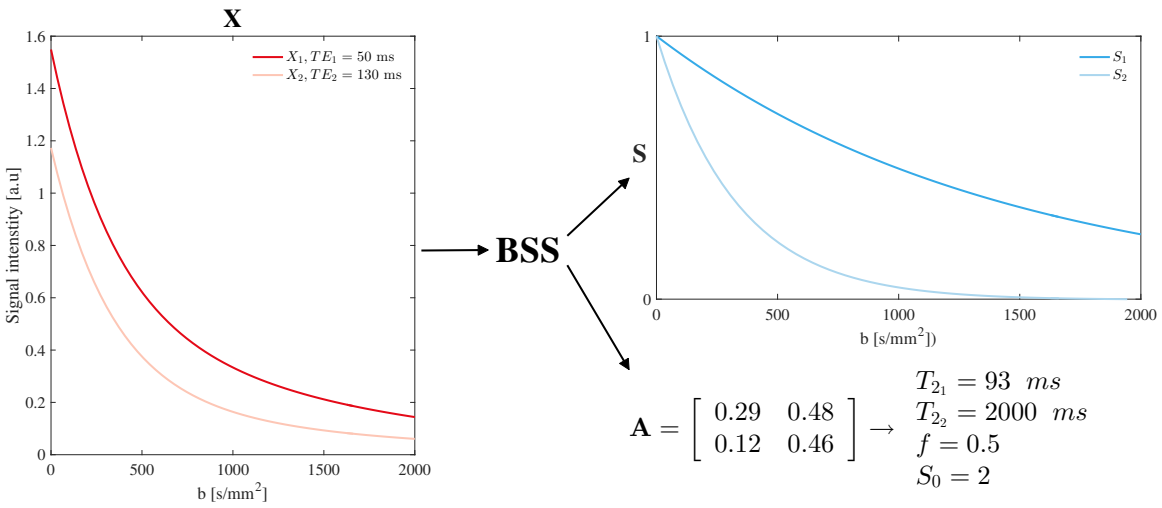

Figure 1: **Factorization of measurements, X, into the sources, S, and mixing matrix, A.** Example of a BSS operation for two mono-exponential sources ($M = 2$) and two TE measurements ($N = 2$). In this illustration, the measurements, **X**, show a bi-exponential decay profile. BSS is capable of separating these two independent exponential source functions, **S**; and calculating their mixing matrix, **A**. The parameters that determine the degree of mixing (T_{2_1}, T_{2_2} and f), and the scaling factor, S_0, were estimated as described in Eqs. 3 and 4. We showed an exponential case for simplicity, but BSS is not limited to this choice; any signal can be processed in the same manner.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

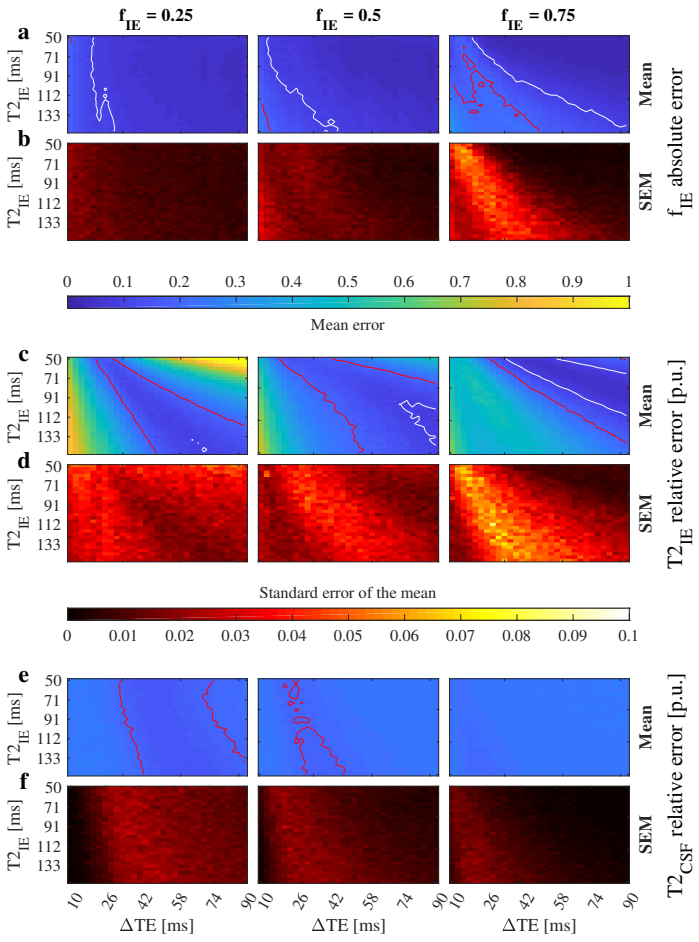

Figure 2: **Convergence for two compartments (IE and CSF) with overlapping T_2 constraints and no S_{CSF} prior (SNR = 50).**

The mean of f_{IE} absolute error and its standard error (SEM) (a and b), and the mean of $T_{2_{IE}}$ (c) and $T_{2_{CSF}}$ (e) relative errors per unit (p.u.), and their standard error (d and f).Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded between 0–1000 ms and 0–3000 ms respectively, and no prior was imposed on S_{CSF}. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. The bias of f_{IE} and $T_{2_{IE}}$ decreases for long ΔTEs as f_{IE} increases. See Figure S5 for more SNR levels.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

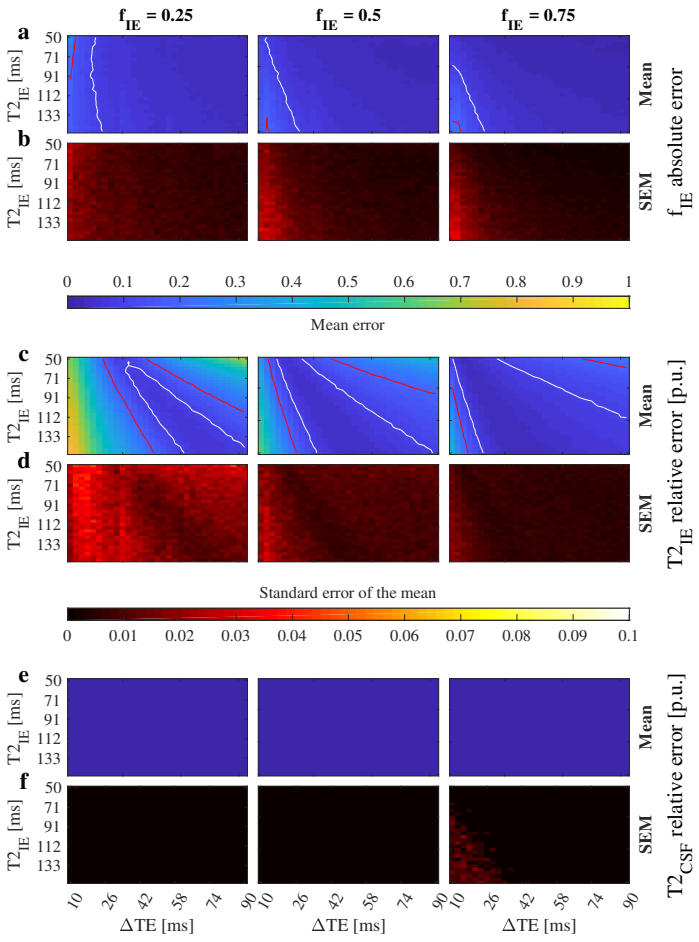

Figure 3: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constraints and S_{CSF} prior (SNR = 50).**

The mean of f_{IE} absolute error and its standard error (SEM) (a and b), and the mean of $T_{2_{IE}}$ (c) and $T_{2_{CSF}}$ (e) relative error per unit (p.u.), and their standard errors (d and f). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded between 0–300 ms and 2000 ms respectively, and S_{CSF} was set to have isotropic diffusivity with value 3 $\times 10^{-3}$ mm^2/s. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. This area is larger than for Figure 2 stressing the importance of priors. See Figure S6 for more SNR levels.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

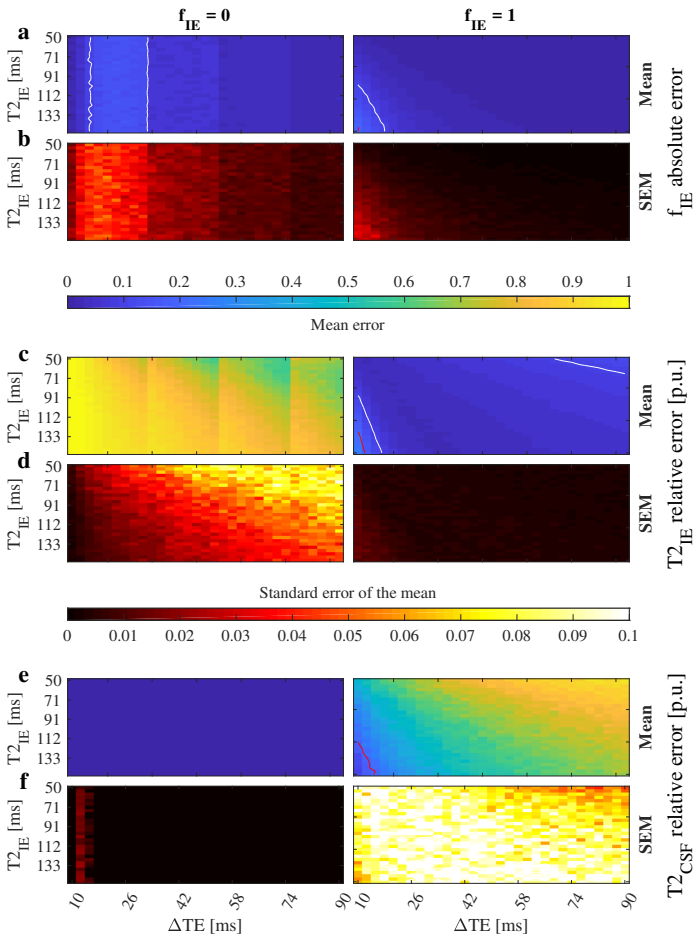

Figure 4: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constraints and S_{CSF} prior when only one is actually present in the tissue (SNR = 50).** The mean of f_{IE} absolute error and its standard error (SEM) (a and b), and the mean of $T_{2_{IE}}$ (c) and $T_{2_{CSF}}$ (e) relative error per unit (p.u.), and their standard errors (d and f). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bounded between 0–300 ms and 2000 ms respectively, and S_{CSF} was set to have isotropic diffusivity with value 3 $\times 10^{-3}$ mm^2/s. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. Estimates of f_{IE} are reliable for ΔTE $>$ 45 ms (a and b). Estimates of $T_{2_{IE}}$ and $T_{2_{CSF}}$ are accurate for each case. See Figure S7 for more SNR levels.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

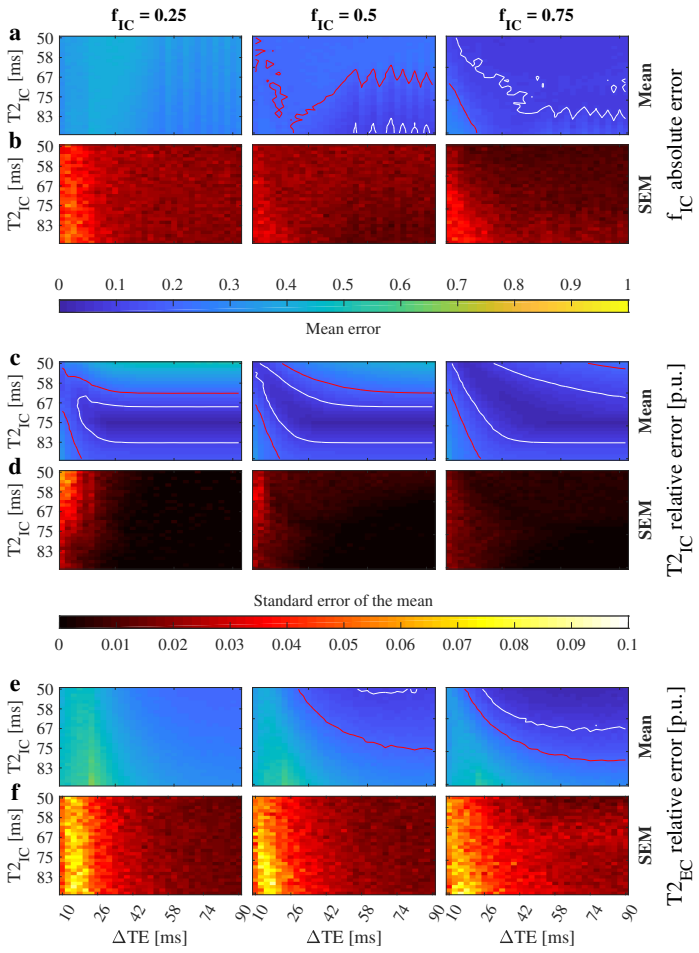

Figure 5: **Convergence for two compartments (IC and EC) with overlapping T_2 constraints and no other priors (SNR = 50).**

The mean of f_{IE} absolute error and its standard error (SEM) (a and b), and the mean of $T_{2_{IE}}$ (c) and $T_{2_{CSF}}$ (e) relative error per unit (p.u.), and their standard errors (d and f). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of f_{IC}, $T_{2_{IC}}$, and ΔTE. $T_{2_{IC}}$ and $T_{2_{EC}}$ were bounded between 0–150 ms and 0–200 ms respectively, and no other prior was imposed in the signal sources. We define the convergence area as the one with error lower than 0.1 for f_{IC}, $T_{2_{IC}}$, and $T_{2_{EC}}$. Estimate of f_{IC} is biased for all f_{IC} levels. T_2 estimates show a narrow band of convergence limited by the lack of prior knowledge (see Figures 2, S5 and S10) and the condition of **A** when the T_2 values are similar. See Figure S8 for more SNR levels.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

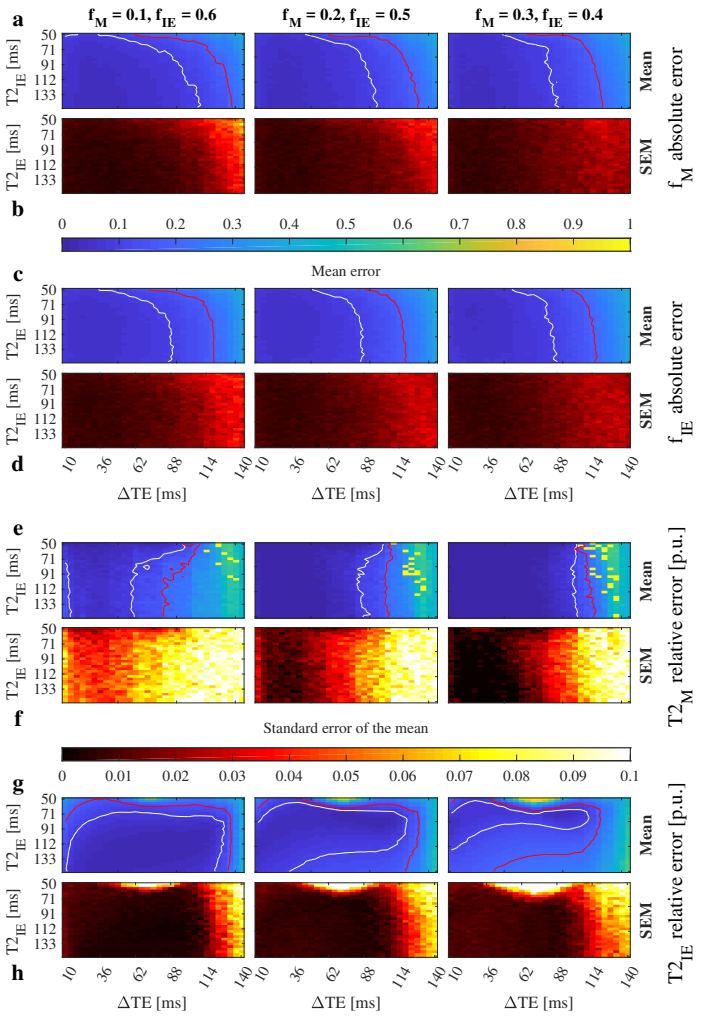

Figure 6: **Convergence for three compartments (myelin, IE, and CSF) with non-overlapping T_2 constraints and S_{CSF} prior (SNR = 50).**

The mean absolute errors of the volume fraction estimates and their standard errors (SEM) (a, b, c, and d); and the mean of T_{2_M} (e) and $T_{2_{IE}}$ (g) relative error per unit (p.u.), and their standard errors (f and h). Red and white lines mark the 0.2 and 0.1 contour respectively. There is a large convergence area when TE_1 = 10 ms, TE_2 = 46 ms, and TE_3 = 150 ms, which is not reachable with current clinical hardware. See Figure S9 for more SNR levels.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

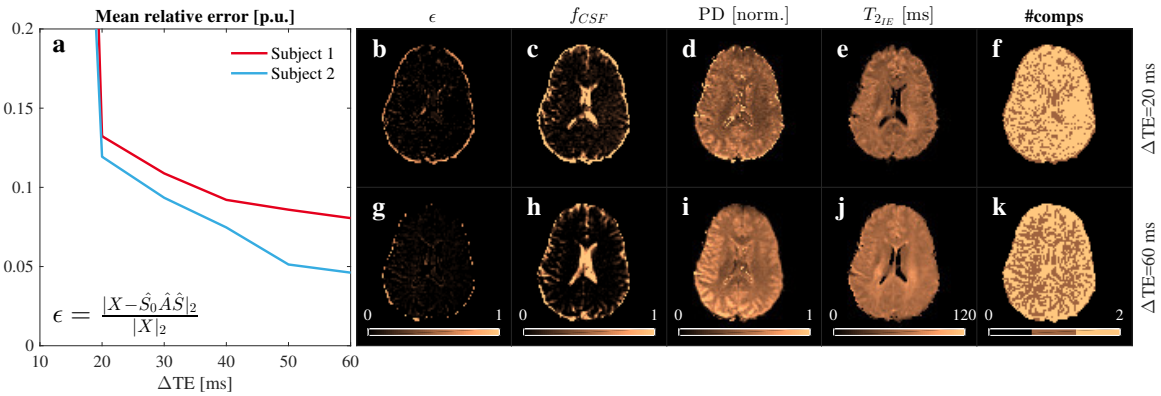

Figure 7: **BSS relative factorization error for increasing ΔTE values.**

The evolution of the relative factorization error with ΔTE, averaged over the whole brain, is shown in (a). As an example of how this error reduction affects BSS estimates we also show the relative error maps (b) and (g), CSF volume fractions (c) and (h), PDs (d) and (i), $T_{2_{IE}}$ values (e) and (j) and the number of compartments (f) and (k) for ΔTEs values of 20 and 60 ms. The mean relative factorization error decreases as ΔTE increases, improving the parameter estimates.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

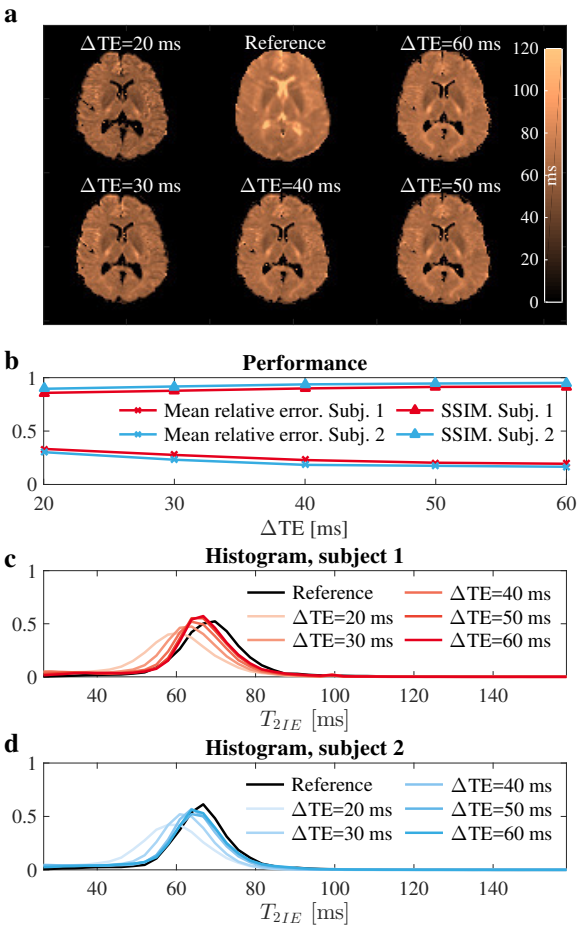

Figure 8: **Comparison of the BSS-estimated $T_{2_{IE}}$ values against a FLAIR reference.** A comparison of the reference (a, upper middle), for subject one with the BSS $T_{2_{IE}}$ estimate is shown for increasing values of ΔTE. The visual comparison was quantified by SSIM (50) and mean relative error (b). Histograms of the BSS-estimated $T_{2_{IE}}$ values are plotted against the reference (c) and (d). High T_2 values in the ventricles for the reference indicate that the suppression of the CSF signal in the FLAIR experiment was not perfect, although they appeared dark in the images (see supplementary Figure **??**). This might have induced a positive bias for the reference. Finally, the BSS-estimated of $T_{2_{IE}}$ values for ΔTE above 50 ms showed good agreement with the reference.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

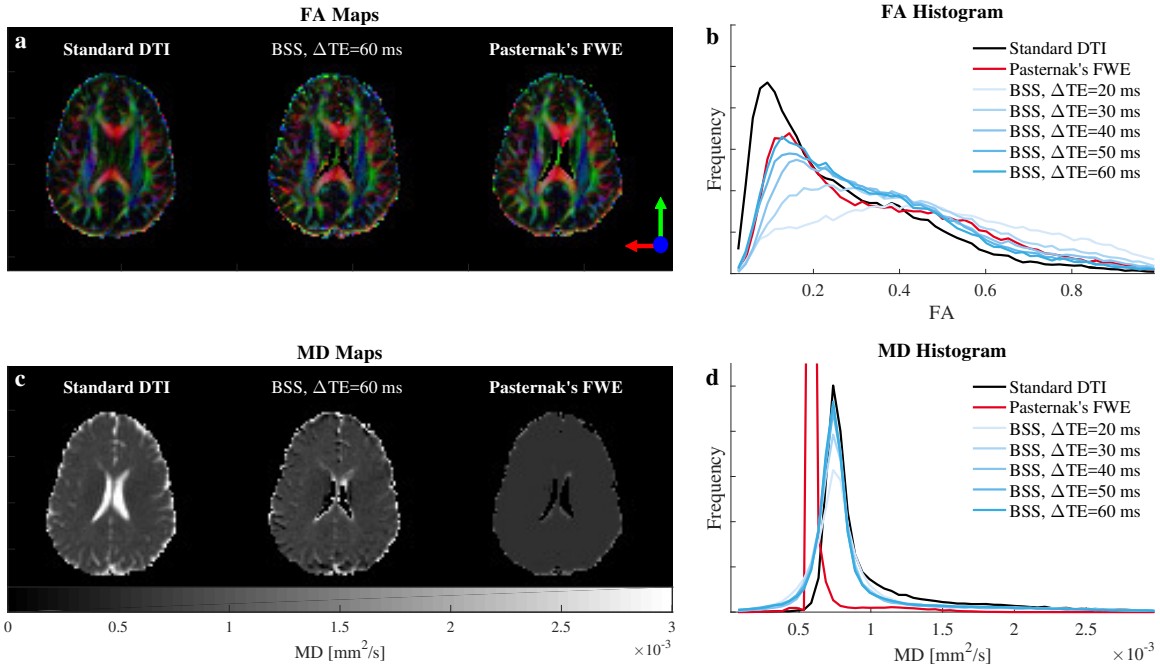

Figure 9: **FA and MD of the BSS-disentangled IE signal against the standard DTI and Pasternak's free-water elimination (FWE) for subject two.**

Comparisons of the FA (b) and MD (d) histograms calculated from the separated IE signals are plotted against the standard DTI fit and Pasternak's method for the short TE measured data. MD (c) and colored FA (a) maps are also included for comparison. We observed a CSF correction effect in the long ΔTE BSS for FA in agreement with Pasternak's FWE. However, both method disagree for MD, where Pasternak's introduces spatial over-regularization. See Figure S15 for the subject one.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

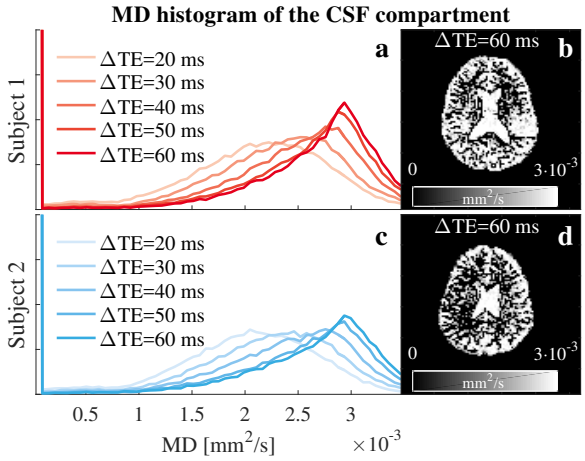

Figure 10: **Evolution of the MD histogram of the BSS-disentangled CSF component with ΔTE.**

The MD histograms, calculated from the the DTI fits for the signals disentangled for the CSF compartment, are plotted in (a) and (c). MD maps (b) and (d) are shown for anatomical inspection. The CSF MD histograms tends towards $3{\times}10^{-3}$ mm^2/s, in agreement with the literature.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Supporting Material

Phantom experiment

Methods

We built a phantom based on pure water and eleven different concentrations of agar and sucrose, producing eleven unique combinations of T_2 and diffusivity (Table S1) (62). We scanned the phantom (see below) and defined regions of interest (ROIs) in the tubes containing the eleven concentrations. Each ROI was independently processed with BSS to study the one compartment case. We also mixed the signals from two ROIs to generate a pair of two-compartment datasets and fed these to our BSS solver. Finally, for the three-compartment case we combined three ROIs and separated them with BSS. We were aiming to demonstrate that our framework was able to yield T_2 estimates for one compartment; and volume fraction, T_2 estimates and diffusion signal separation for two and three compartments.

For reference, we measured multi-echo SE acquisitions (Signa HDx 3T, GE Healthcare, Milwaukee, WI) for TE values from 10–640 ms in 10 ms increments. The following values were constant: TR = 3460 ms; NEX = 2; 128 × 128 matrix size; FOV = 240 mm; and 7 mm slice thickness. Eleven diffusion experiments were undertaken for TE values from 77.5–127.5 ms in 5 ms increments. the following parameters were constant: FOV = 240 mm; 7 mm slice thickness; 64 × 64 matrix size; TR = 4000 ms; ASSET = 2; A/P diffusion direction; and 41 equally spaced b-values from 0–2000 s/mm^2.

The multi-echo SE signals were averaged within each ROI. Each signal was fitted with NNLS (13) using a log-scaled grid with T_2 values at 500 points between 10–2000 ms. We used the maximum values of the NNLS T_2 spectra as ROI reference values (Figure S1) and fitted the signal from each ROI with EASI-SM (17) for reassurance.

One compartment

For one compartment ($M = 1$), we processed the diffusion data from ten pairs of TE measurements ($N = 2$) with BSS to include the relaxation effects in the dataset. For each pair, the short TE was fixed at 77.5 ms, while the long TE was increased from 82.5–127.5 ms along with the measured echo times. We constrained the solution space for the estimated $\hat{T}_2$ values to 10–2000 ms to account for all the ROIs. No other prior information was considered. We report the evolution of the T_2

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

values estimated using BSS for each ROI and the differences between the short and long TEs (ΔTE) compared with their reference values (Figure S1).

Two compartments

For two compartments ($M = N = 2$), we created two different datasets. First, we used the diffusion data measured at the shortest TE for ROI_6 and ROI_{11} as the sources, $\mathbf{S}$. These signals did not contain relaxation information (Figure S2a). Thus, to mix them together, we had to compute the mixing matrix ($\mathbf{A}$) as in in Eq. 2. We used their reference T_2 values, the experimental TEs, and a volume fraction of $f_{ROI_6} = 0.7$ (Figure S2c). We called this the *simulated* dataset, given that the signals were mixed under ideal conditions. Second, we normalized the measured data for each ROI and TE to its maximum value at the shortest TE to allow for later comparison of the volume fractions. In this case, the signals already contained the relaxation information (Figure S2b), so we did not need to compute $\mathbf{A}$. We scaled the normalized measured signals using the given volume fraction and added them together to create the mixed measurements, $\mathbf{X}$. We called this the *measured* dataset (Figure S2c). It accounts for system imperfections like signal drift, imperfect non-diffusion weighting, and eddy currents. To constrain the solution of the cALS algorithm we used $T_{2_{ROI_{11}}}$ and $S_{ROI_{11}}$ as the prior knowledge and searched for $\hat{T}_{2_{ROI_6}}$ between 0–200 ms.

Three compartments

We extended the two-compartments experiment to three ($M = N = 3$) by adding ROI_5. *Simulated* and *measured* datasets were created as for the two-compartments case (Figures S3a, S3b, S3c, S3d and S3e). This time, we used the volume fractions $f_{ROI_5} = 0.2$ and $f_{ROI_6} = 0.6$. To limit the solution space of the cALS algorithm, we assumed $T_{2_{ROI_{11}}}$ and $S_{ROI_{11}}$ to be prior the knowledge. We also constrained the $\hat{T}_{2_{ROI_5}}$ and $\hat{T}_{2_{ROI_6}}$ values to be between 0–50 ms, and 50–200 ms, respectively. For the two- and three-compartment experiments we report the stability of the framework, the relative error of the parameters and the disentangling capability of the method.

Results

One compartment

There was a correlation between the estimated T_2 values for one compartment obtained using multi-echo SE for 17 TEs and BSS for 2 TEs (Figures S1a, S1b, and Table S1). The T_2 estimates from

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

ROI_2 to ROI_{10} showed relative errors below 0.1 p.u. for a ΔTE of 50 ms (Figures S1a and S1b). The decreasing error trend is due to the relationship between the slope of a column of $\mathbf{A}$ and its T_2 value (Figure S1c). As ΔTE increased, the dynamic range of the slope of $\mathbf{A}$'s columns expanded, yielding better T_2 estimates. On the other hand, in Figures S1a and S1b, ROI_1 and ROI_{11} showed increasing errors as ΔTE increased. In the case of ROI_1, this was due to the low SNRs of the measurements at the experimental TEs. The noise floor caused changes in the signals for longer TEs that biased the T_2 estimates. The effect observed in ROI_{11} cannot be explained by SNR or T_2-slope dependence. We attribute this result to an underestimation of the reference T_2 value due to incomplete recovery of the longitudinal magnetization, which is caused by the short experimental TR (TR = 3460 ms) compared to the T_1 value of ROI_{11} ($T_{1_{11}}$ = 2200 ms). Finally, the error between the NNLS and BSS T_2 estimate for ROI_4, ROI_6, ROI_7, and ROI_8 is larger than for the others (Figure S1a) at ΔTE = 50 ms, except for ROI_1 and ROI_{11} already discussed. For these ROIs, NNLS converges to a bi-exponential decay (See Figure **??** and Table S1) increasing the value of the long T_2 coefficient compared to BSS and EASI-SM.

Two compartments

The disentangled signals for the *simulated* dataset replicated the profiles of the reference sources (Figure S2d). Moreover, the maximum relative errors for $\hat{f}_{S,ROI_6}$ and $\hat{T}_{2_{S,ROI_6}}$ were below 0.01 p.u. for all the possible ΔTE values. Interestingly, BSS was also able to separate the signal sources of the *measured* dataset (Figure S2d). This data accounted for non-ideal conditions due to system imperfections, such as signal drift, eddy currents, or imperfect non-diffusion weighting (Figure S2b, S2c, S2d, and S2f). In that case, the relative error in the $\hat{T}_{M_{S,ROI_6}}$ estimate remained under 0.1 p.u. for all ΔTEs above 10 ms. We believe that the 0.15 p.u. error in $\hat{f}_{M,ROI_6}$ is due to the differences between the *simulated* and *measured* signals at b = 0 s/mm^2, their influence on Eq. 4, and propagation of the error in the $\hat{T}_2$ estimate. Finally, we also observed a small stabilization effect in the volume fraction estimates as ΔTE increased (Figure S2f). This behavior is due to reductions in $\mathbf{A}$'s condition number improving the cALS algorithm's numerical stability (Figure S2e).

Three compartments

The condition number of $\mathbf{A}$ significantly increased compared with the two-compartment model (Figures S2e and S3g). Results for the *simulated* data (Figures S3a, S3e and S3f) showed that the signals for compartments ROI_6 and ROI_{11} had been separated, in agreement with their references.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Likewise, the relative errors in the $\hat{T}_{2_{S,ROI_6}}$ and $\hat{f}_{S,ROI_6}$ estimates were below 0.01 p.u., confirming the good separation. It is worth noting that the signal for the fast-decaying compartment (ROI_5) was detected, despite being heavily contaminated by the ROI_6. We believe that this result is due to the comparably large experimental TE, reducing its contribution to the observed signal. Equivalently, we found a 0.15 p.u. error in the $\hat{f}_{S,ROI_5}$ estimate and 0.45 p.u. in the $\hat{T}_{2_{S,ROI_5}}$ estimate.

Results for the *measured* data when **A**'s condition number was lowest showed that the signals from ROI_6 and ROI_{11} had still been separated, in agreement with the references (Figure S3f). However, the signal from ROI_5 was lost due to acquisition imperfections, bad conditioning of **A**, and small contributions of this compartment at the measured TEs. On the other hand, the $\hat{T}_{2_{M,ROI_6}}$ estimate was stability with a relative error of 11%. In contrast, $\hat{f}_{M,ROI_6}$ was more unstable due to the bad conditioning of the system and propagation of the error in the $\hat{T}_{2_{M,ROI_6}}$ estimate.

Repeatability and reproducibility

Following the simulations, phantom experiment, and in vivo studies for incremental ΔTE, repeatability and reproducibility analyses were conducted to demonstrate the stability and reliability of our method. In this supporting section we described the experimental setup and results, while the discussion remained in the main body of the paper.

Methods

Repeatability

A healthy volunteer (male, 28 years old) was scanned six times in a 3.0T GE MR750w scanner (GE Healthcare, Milwaukee, WI). For each repetition we acquired two diffusion PGSE EPI volumes with TE values 75.3 and 135.3 ms (ΔTE = 60 ms); FOV = 225 mm; 4 mm slice thickness; 22 slices; TR = 8000 ms; 96 × 96 matrix size; ASSET = 2; 30 directions; and one non-diffusion-weighted volume. Besides, one non-diffusion-weighted volume was acquired with reversed polarity at each TE. Finally, a FLAIR multi-echo sequence was acquired with the same geometrical prescription for TE = 20 – 260 ms in 30 ms increments; ASSET = 0; and TR = 8000 ms. An extra volume was acquired with reverse polarity at TE = 20 ms.

Diffusion and FLAIR data were processed with FSL Topup (63, 64) and Eddy (65) to correct for distortions. The long TE diffusion volume was registered to the short TE one with FLIRT and processed with BSS for two compartments (IE and CSF). We used literature CSF values ($T_{2_{CSF}} = 2$

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

s and $D_{CSF} = 3 \times 10^{-3}$ mm^2/s) as the prior knowledge, and constrained $T_{2_{IE}}$ between 0 – 200 ms. Then, the resulting tissue volume was fitted to the DTI model using standard linear regression (FSL FDT). For comparison, the distortion free short TE diffusion volume was also fitted to the DTI model and free-water corrected with Pasternak's method. Finally, FLAIR data were matched to a dictionary of mono-exponential decays from 0 – 300 ms with 1 ms increments.

We reported the FA and MD histograms for the six repetitions of the standard DTI fit for the short TE, IE BSS, and Pasternak's method (Fig S16a and b). The free-water correction effects were quantified dividing the histograms in sectors, and computing relative change per sector in the number of voxels of BSS and Pasternak's method reference to the standard DTI fitting (Fig S16d and e). FA histograms were split in four quarters, while MD in two sectors with threshold in the IE literature value (MD = 0.7×10^{-3} mm^2/s) (28). Statistical t-test analyses were conducted to determine the differences between BSS and Pasternak's FWE. Histograms of BSS $T_{2_{IE}}$ and FLAIR T_2 (Fig S16c) were compared by their peak and full width half maximum (FWHM) values (Fig S16f).

Reproducibility

Twenty healthy volunteers (8 females, 26 years old in average) were scanned in a 3.0T GE MR750 scanner (GE Healthcare, Milwaukee, WI) at the Max Planck Institute of Psychiatry in Munich, Germany. Two diffusion PGSE EPI volumes with TE values 60.1 and 120.1 ms (ΔTE = 60 ms) with TR = 5000 ms were acquired. All the other acquisition parameters and data processing steps were as described for the repeatability experiment. Due to scanner availability FLAIR data was only acquired for half of the subjects. Histograms of FA, MD, and T_2, along with their statistical analyses were reported in Figure S17.

Results

Repeatability

The histograms of FA, MD and T_2 (Figure S16a, b, and c) showed highly overlapping curves for each repetition and method, denoting good repeatability for all of them. After splitting the FA histogram in four sectors and computing the relative change in the area per sector for BSS and Pasternak's method (Figure S16d), we found that the lowest ratio between the mean and the standard deviation for BSS was 5.3 (sector IV) and 11.4 for Pasternak's (sector II). The fact that the mean is 5.3 larger

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

than the standard deviation suggested that BSS produces highly stable free-water correction. This result held also for MD, where we found ratios of 9.9 and 65.4 for BSS and Pasternak's method in sector II; and T_2, with mean to standard deviation ratios of 62.3 and 10.4 in the peak and FWHM values for BSS.

Interestingly, we found that BSS and Pasternak's free-water correction yield statistically different results for FA > 0.25 (sectors II, III, and IV, Figure S16d and e) and both MD sectors, with a significance level $\alpha = 0.01$. Finally, BSS $T_{2_{IE}}$ and FLAIR $T_{2_{IE}}$ histogram peaks were not statistically different, while the FWHM values were with a small effect. These findings indicate a good agreement of BSS with the FLAIR reference (Figure S16f).

Reproducibility

The FA, MD and T_2 histograms showed larger inter-subject variability (Figure S17a, b, and c) compared to the intra-subject one (Figure S16a, b, and c)) . The lowest ratios between the mean and the standard deviation of the free-water correction factor were 3.9 for BSS FA (sector IV) and 6.8 for Pasternak's FA (sector II); and 4.9 for BSS MD (sector II) and 20.5 for Pasternak's MD (sector II). Furthermore, mean to standard deviation ratios of BSS $T_{2_{IE}}$ were 47.2 and 8.9 for peak and FWHM values. These results suggested that BSS experiments are highly reproducible among subjects.

The statistical differences found in the repeatability study in FA sectors III and IV, and both MD sectors were still present in this analysis, indicating consistent differences between BSS and Pasternak's method (Figure S17d and e).

We found a statistically significant difference ($\alpha = 0.01$) between the means of the histogram peaks of the BSS and FLAIR $T_{2_{IE}}$ but with a small size effect (Figure S17f). Which indicates that BSS might yield a small bias in group comparisons compared to FLAIR multi-echo.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

Supporting Figures (for publication)

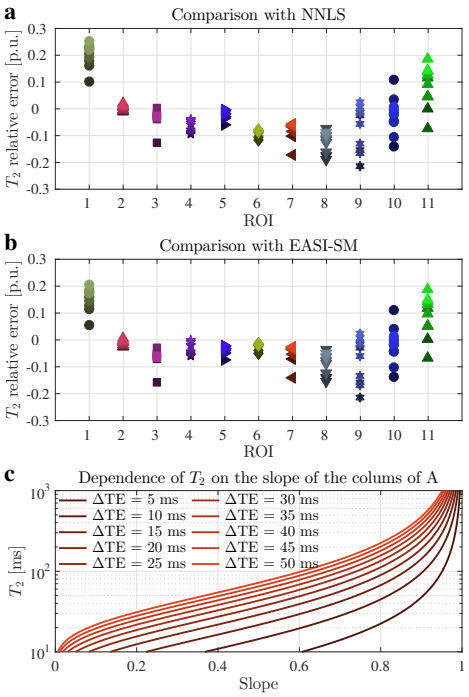

Figure S1: **Evolution of the relative error in the T_2 estimate with ΔTE for one compartment.**

The mean relative error of T_2 estimated using BSS is shown in (a) for NNLS and in (b) for EASI-SM references. ΔTE goes from 5 ms (darker colors) to 50 ms (lighter colors). The dependence of T_2 on the direction (slope) of the columns of $\mathbf{A}$ (Eq. 3) is shown in (c), where it can be seen how increasing ΔTE improves the dynamic range of the slope of $\mathbf{A}$, resulting in a better estimate for T_2. Except for ROI_1 and ROI_{11}, the remaining ones reduce the T_2 mean relative error as ΔTE increases (a and b, lighter colors are closer to zero), in agreement with plot c.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

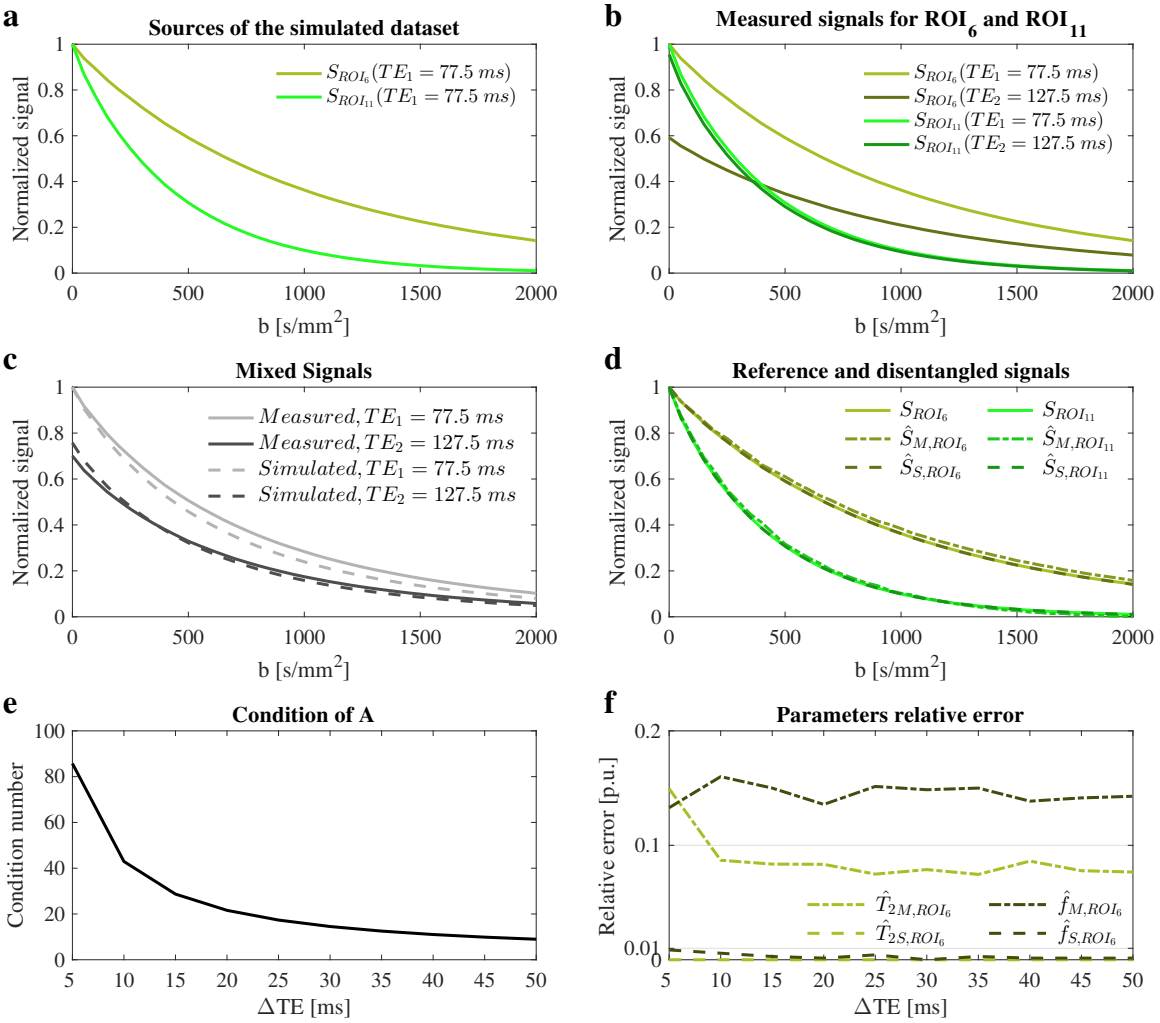

Figure S2: **Separation of two compartments and parameter estimation for the phantom data.**

The signal sources of the *simulated* dataset are plotted in (a), and the *measured* data generated from the sources in (b). The resulting mixtures for both datasets are shown in (c). We use the subscripts M and S to refer to estimates for the *measured* and *simulated* datasets, respectively. Measurement errors are highlighted by the differences between the *measured* and *simulated* signals, shown in (c). BSS disentangled the original sources for both datasets, as shown in (d). We chose a ΔTE of 50 ms to minimize the condition of $\mathbf{A}$ (shown in (e)) and increase the numerical stability of the framework. Finally, the relative errors in the estimated parameters, $\hat{T}_{2_{ROI_6}}$ and $\hat{f}_{ROI_6}$, are plotted in (f) for all possible values of ΔTE. We observed good agreement between the reference signals and those disentangled with BSS.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

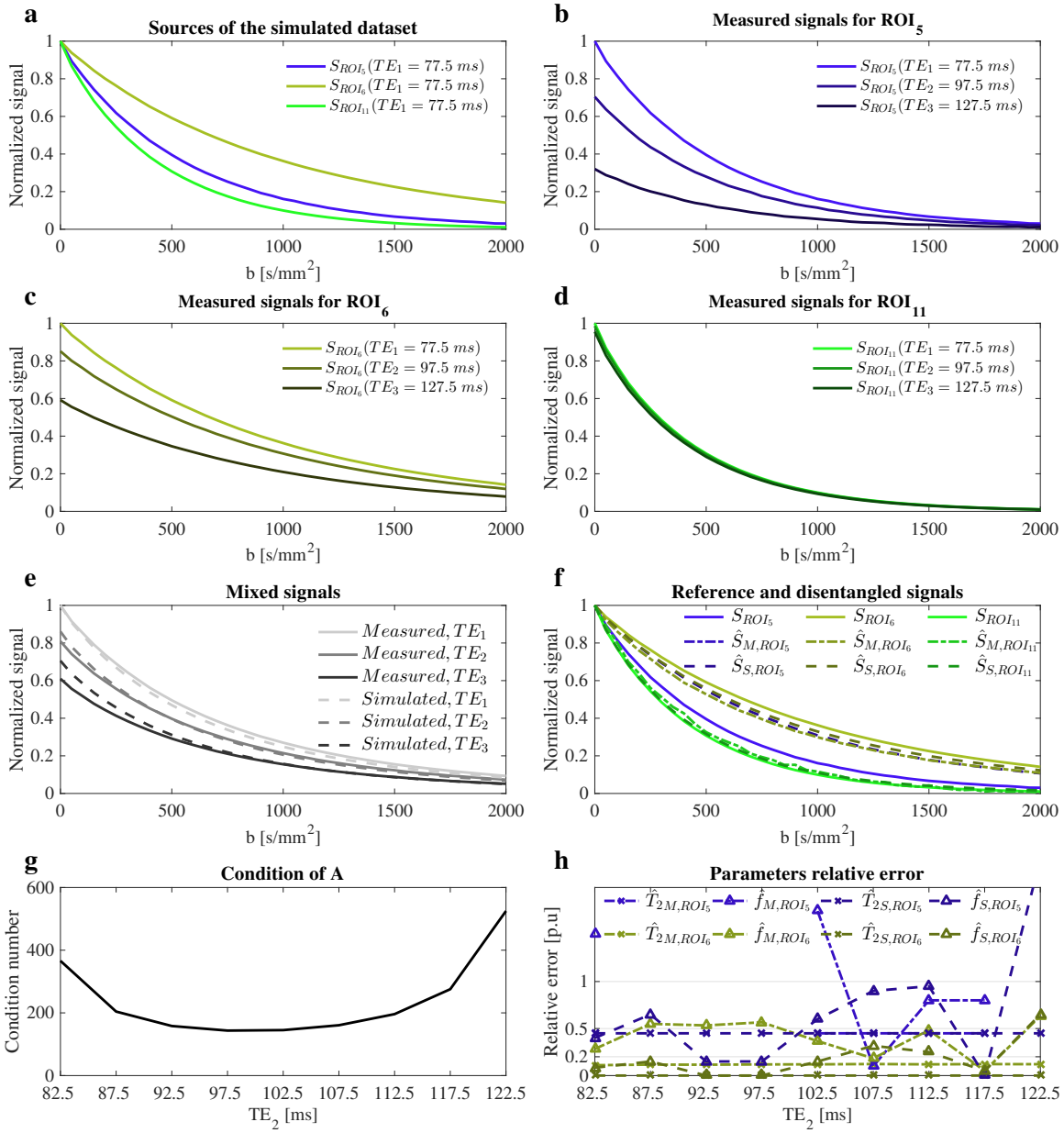

Figure S3: **Separation of three compartments and parameter estimation for the phantom data.**

The *simulated* dataset was generated from the signal sources in (a). The *measured* datasets were calculated from the measured signals for ROI_5 (b), ROI_6 (c), and ROI_{11} (d). The mixed signals for both datasets (shown in (e)) show a mismatch due to measurement errors. They were disentangled with BSS, as shown in (f). We fixed $TE_1 = 77.5$ ms and $TE_3 = 127.5$ ms, and varied TE_2 to minimize the condition number of $\mathbf{A}$ (shown in (g)). The relative errors of the estimated parameters are plotted for different values of the TE_2 in (h).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

ROI	Agar [%]	Sucrose [%]	$T_{2_{EASI-SM}}[ms]$	$T_{2_{NNLS}}$ [ms]	$T_{2_{BSS}}$ [ms]	ϵ_{NNSL} [%]	$\epsilon_{EASI-SM}$ [%]
1	5	15	25.02	23.88 ± 1.92	29.9	25.37	19.61
2	5	5	31.59	31.13 ± 2.19	31.6	1.43	0.04
3	3	30	37.68	36.50 ± 3.04	35.4	2.95	5.99
4	3	15	106.23	110.07 ± 7.93	106.0	3.70	0.22
5	3	5	45.40	44.66 ± 2.85	44.5	0.40	2.02
6	1	30	95.46	102.19 ± 10.30	93.9	8.13	1.66
7	1	15	222.22	228.94 ± 12.15	216.3	5.53	2.67
8	1	5	225.19	233.85 ± 13.84	213.4	8.76	5.25
9	0	30	457.08	456.37 ± 26.50	467.6	2.47	2.31
10	0	15	395.95	397.56 ± 21.17	401.0	0.87	1.28
11	0	0.5	876.97	881.23 ± 64.07	1008.6	14.46	15.01

Table S1: **Phantom reference values and BSS estimates.**

The ROIs in the phantom experiment was built using the concentrations of agar and sucrose shown here. Signal decays along the diffusion dimension were compared to each other to ensure that they were all different, as required by BSS (see supplementary Figure **??**). For reference, the T_2 values were characterized using an NNLS fit. Confidence intervals were taken at the half maxima of the NNLS spectral peaks. In addition, a second method, EASI-SM (17), was used to confirm the validity of the fits. Finally, the $T_{2_{BSS}}$ values were estimated for ΔTE = 50 ms and compared with the NNLS and EASI-SM references (where ϵ refers to the relative error).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

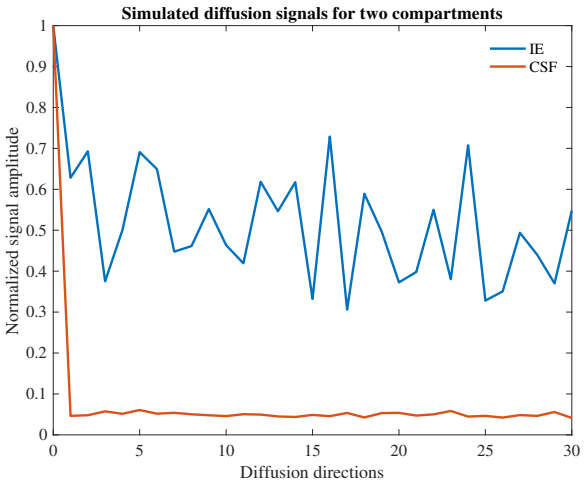

Figure S4: **Simulated diffusion signals for IE and CSF.**

Synthetically generated diffusion signals for 30 directions (b = 1000 s/mm^2) and one non-diffusion weighted measurement. We modeled diffusion as a Gaussian process with MD of IE and CSF equal to 0.7×10^{-3} and 3×10^{-3} mm^2/s respectively (28), and standard deviations of 0.3×10^{-3} and 0.1×10^{-3} mm^2/s respectively to distinguish between hindered anisotropic (IE) and free isotropic (CSF) diffusivity.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

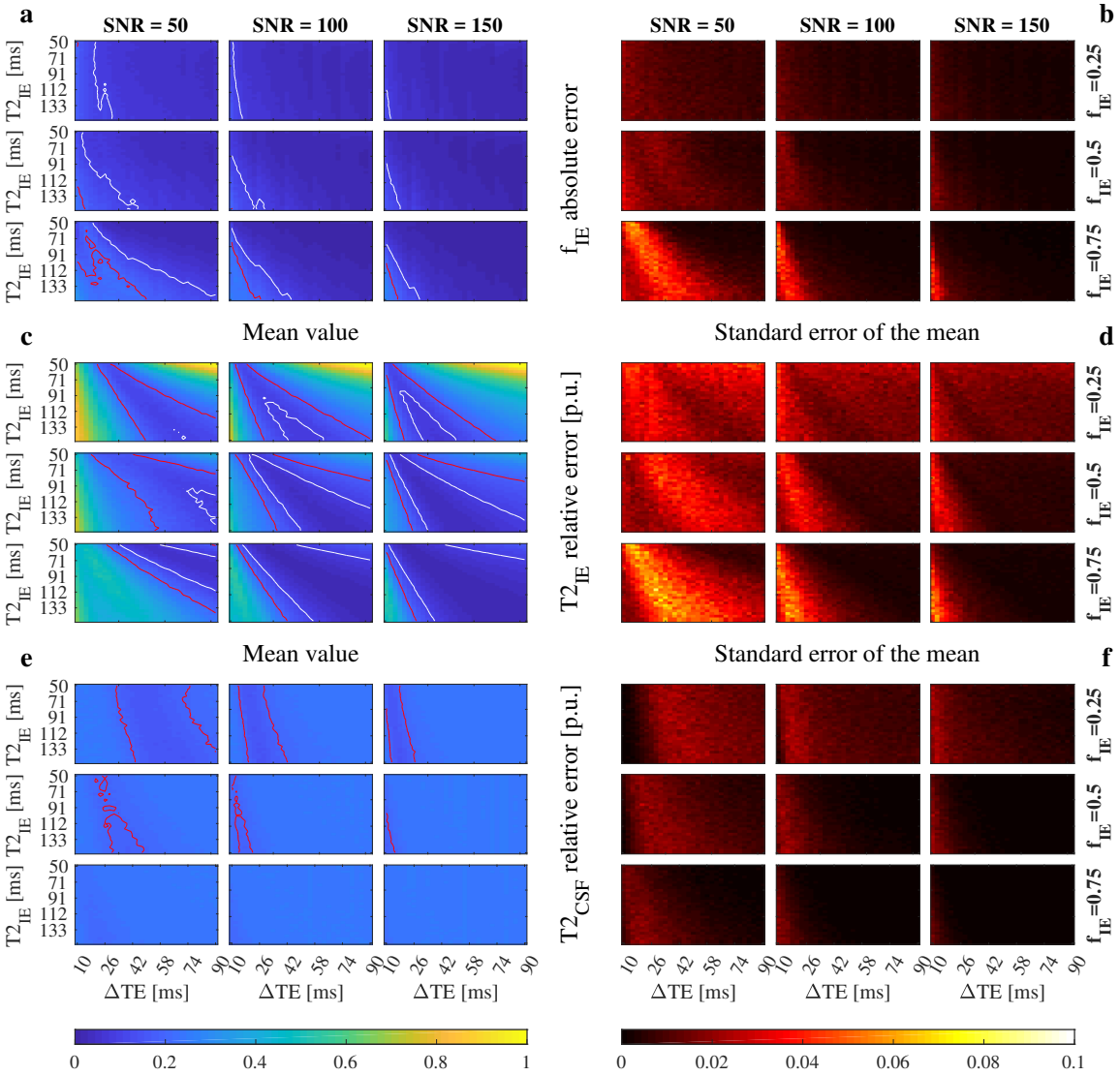

Figure S5: **Convergence for two compartments (IE and CSF) with overlapping T_2 constraints and no S_{CSF} prior.**

This figure extends the analysis of Figure 2 for SNR = 100 and 150. The stability for f_{IE} increases with SNR (a and b) and with f_{IE} for $T_{2_{IE}}$ (c and d).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

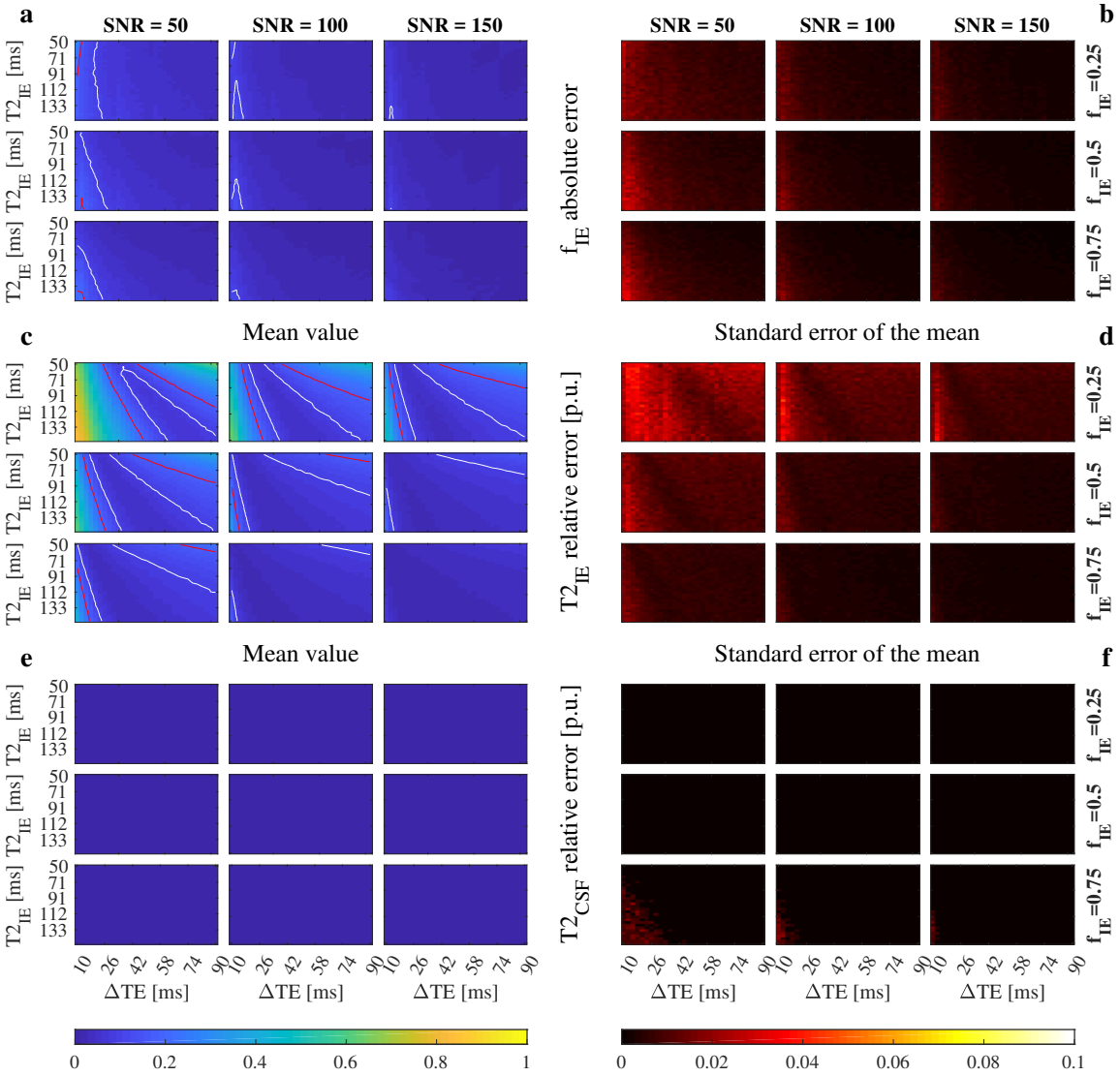

Figure S6: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constraints and S_{CSF} prior.**

This figure extends the analysis of Figure 3 for SNR = 100 and 150. The size and stability of the convergence area for f_{IE} and $T_{2_{IE}}$ increase with SNR.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

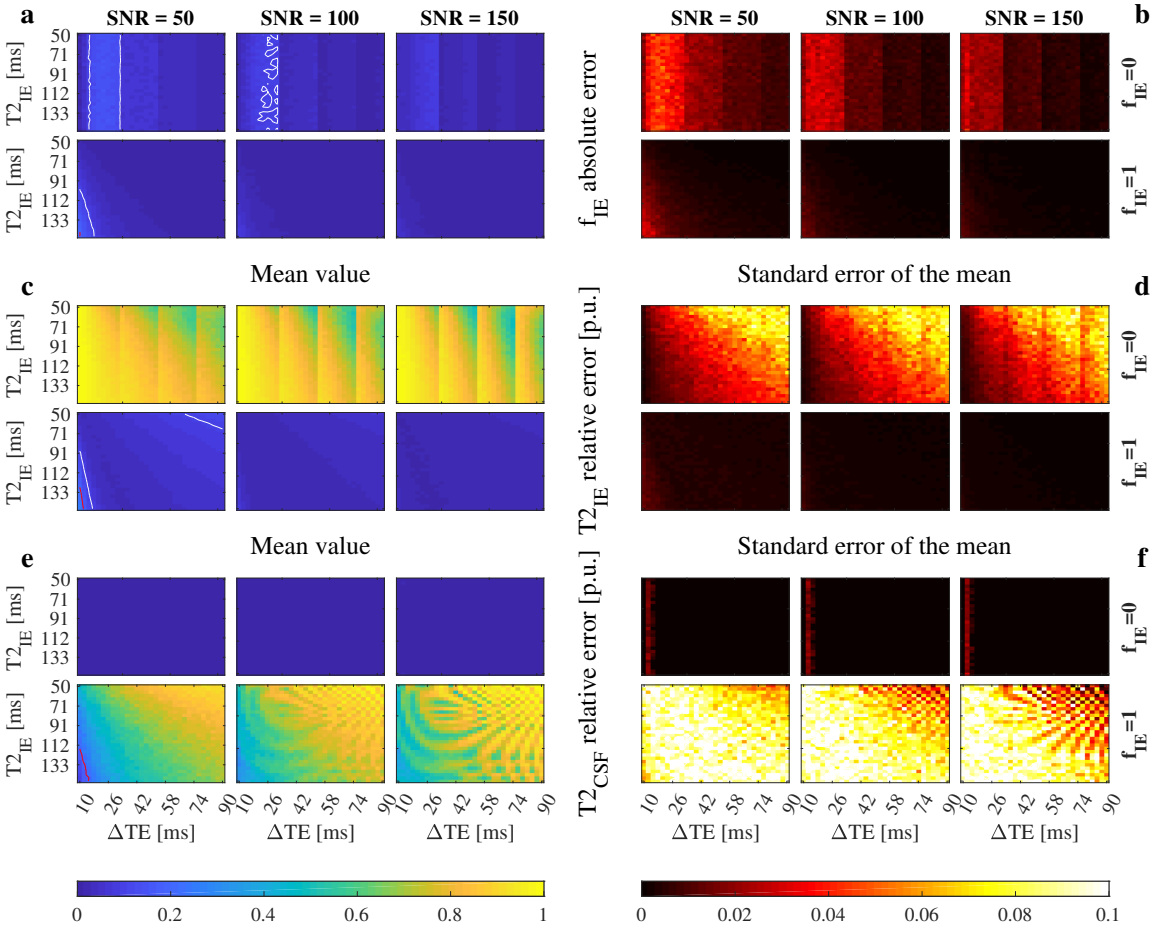

Figure S7: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constraints and S_{CSF} prior when only one is actually present in the tissue.**

This figure extends the analysis of Figure 4 for SNR = 100 and 150. The SNR does not play an important role in the definition of the convergence area.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

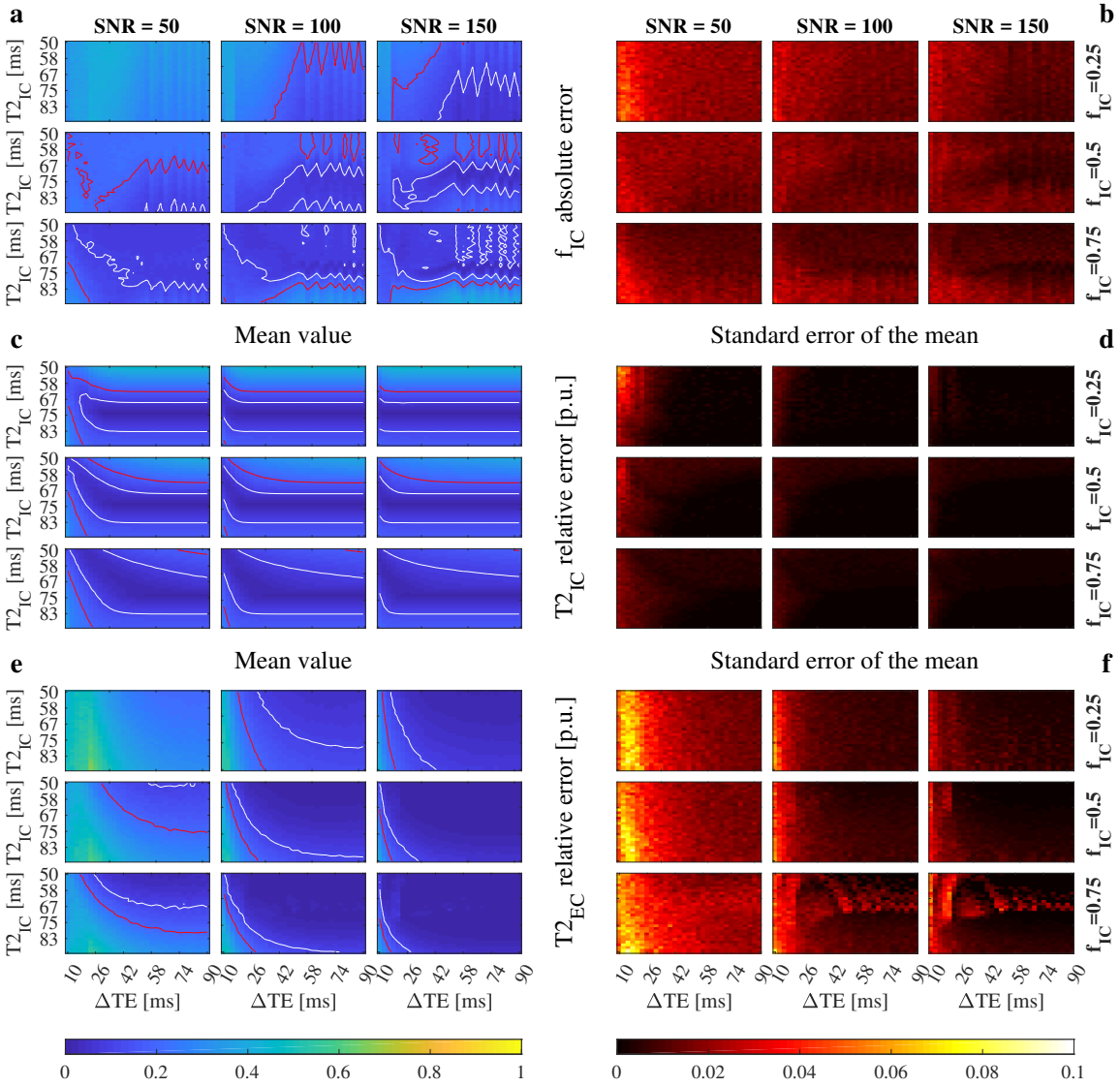

Figure S8: **Convergence for two compartments (IC and EC) with overlapping T_2 constraints and no other priors.**

This figure extends the analysis of Figure 5 for SNR = 100 and 150. The influence of SNR on f and $T_{2_{IC}}$ is small.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

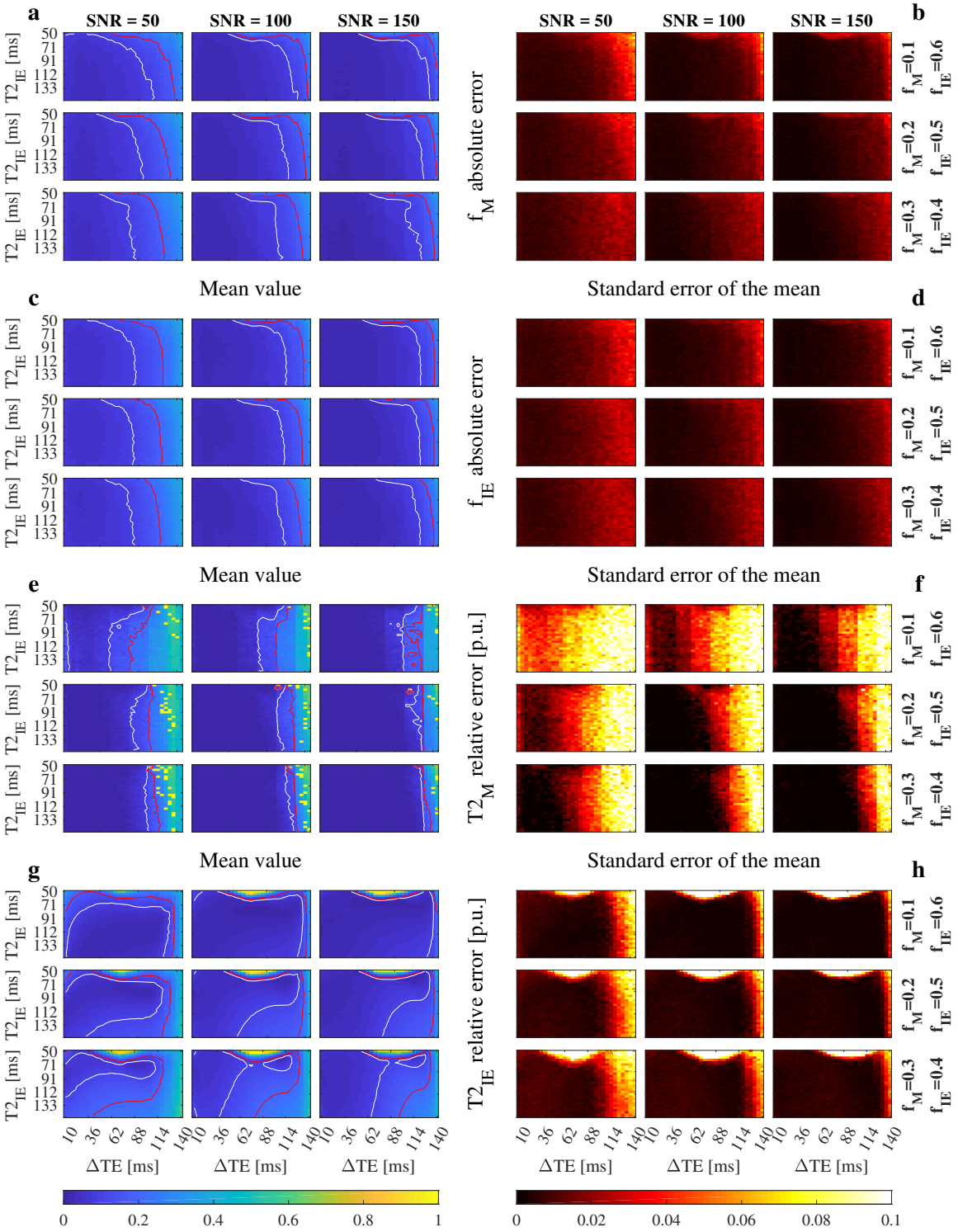

Figure S9: **Convergence for three compartments (myelin, IE, and CSF) with non-overlapping T_2 constraints and S_{CSF} prior.**

This Figure extends the analysis of Figure 6 for SNR = 100 and 150.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

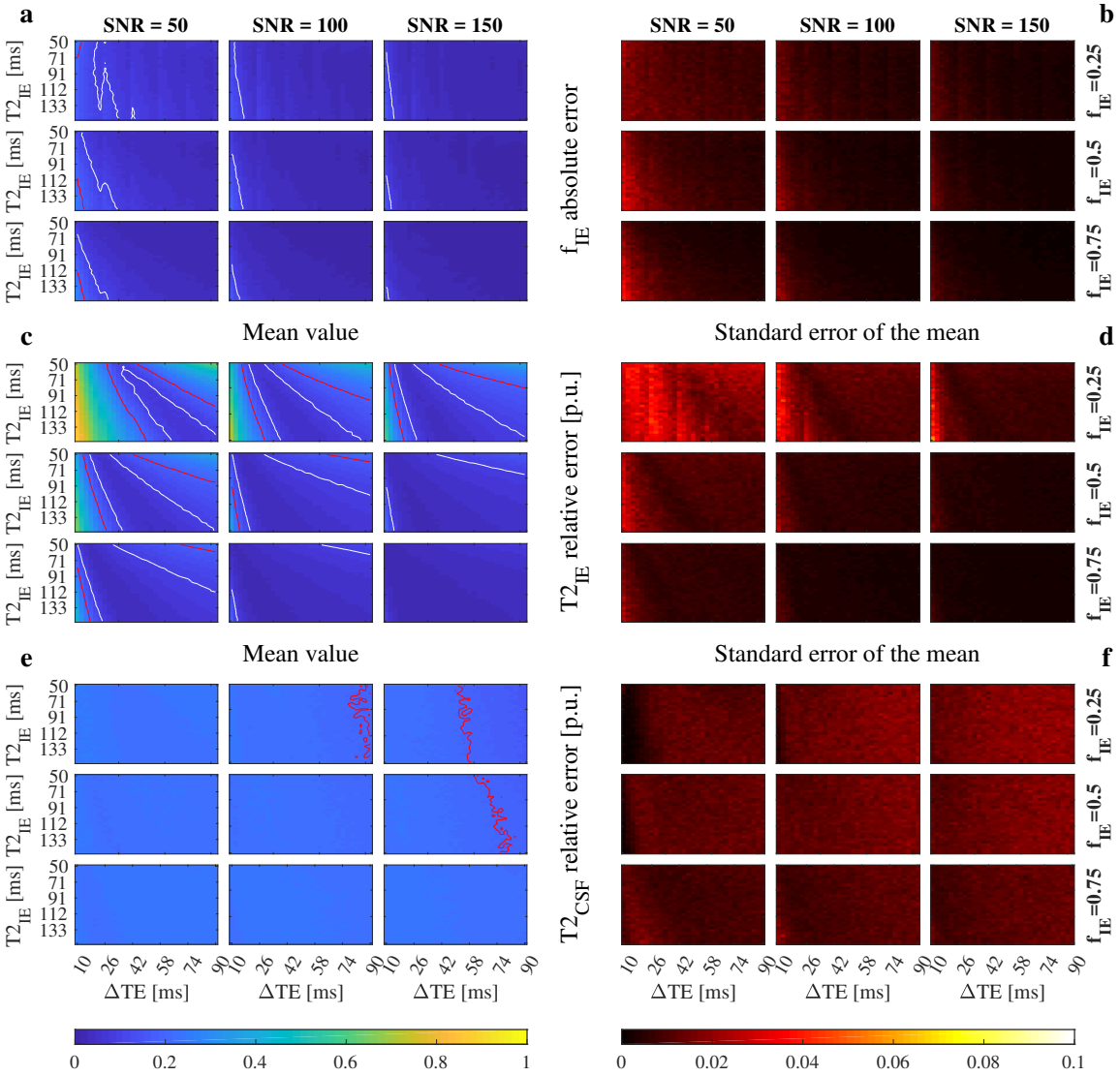

Figure S10: **Convergence for two compartments (IE and CSF) with overlapping T_2 constraints and S_{CSF} prior.**

The mean and the standard error of f_{IE} absolute error (a and b), and the mean and the standard error of $T_{2_{IE}}$ (c and d), and $T_{2_{CSF}}$ (e and f) relative error per unit (p.u.). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of SNR, f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bound between 0–1000 ms and 0–3000 ms respectively. S_{CSF} was set to have isotropic diffusivity with value 3×10^{-3} mm^2/s. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. Notice the growth of the converge area compared to the lack of priors (Figures 2 and S5).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

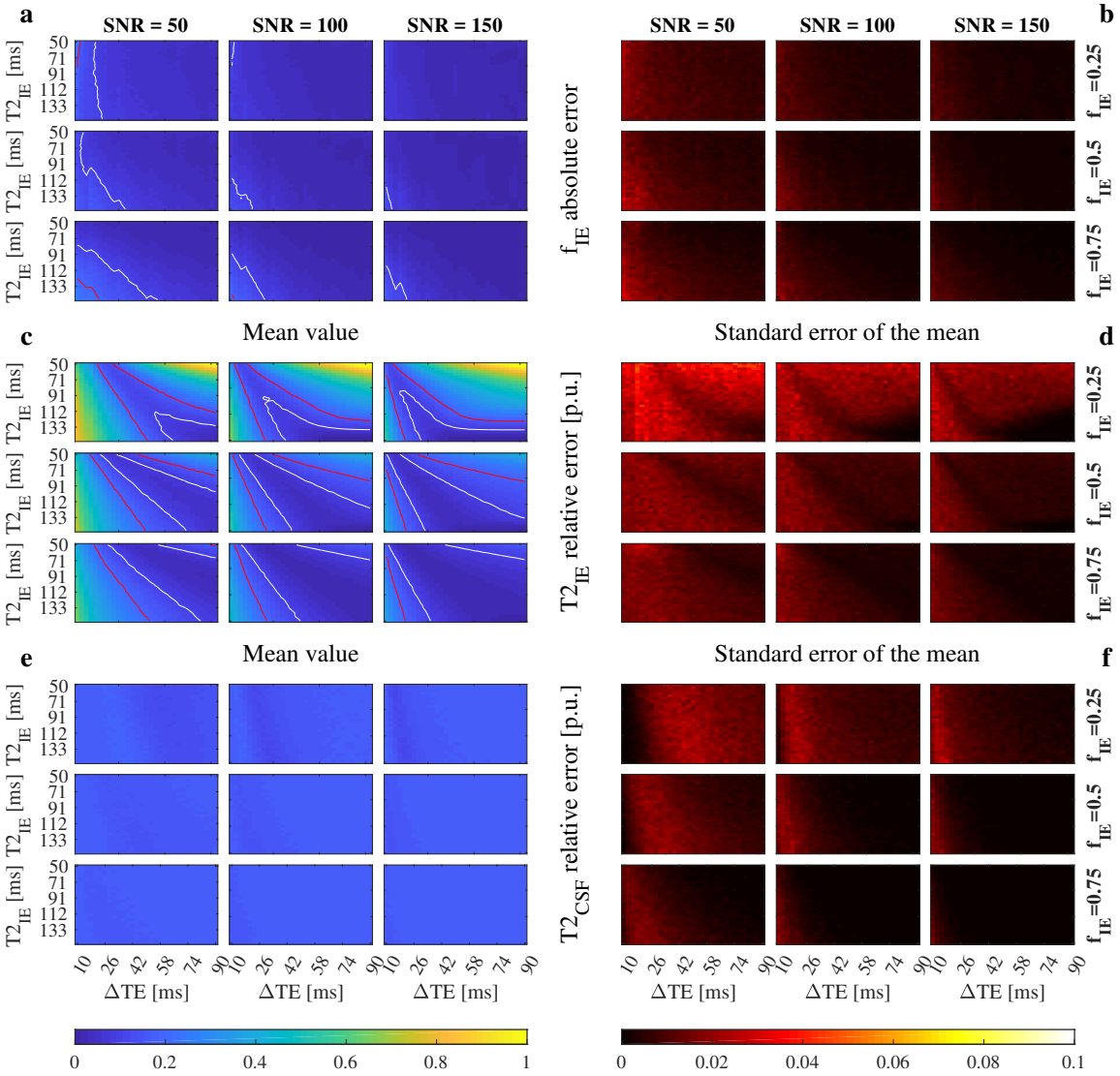

Figure S11: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constrained and no S_{CSF} prior.**

The mean and the standard error of f_{IE} absolute error (a and b), and the mean and the standard error of $T_{2_{IE}}$ (c and d), and $T_{2_{CSF}}$ (e and f) relative error per unit (p.u.). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of SNR, f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bound between 0–300 ms and 300–3000 ms respectively. No prior was imposed on S_{CSF}. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. Non-overlapping T_2 bounds stabilize the factorization, compared to Figures 2 and S5, although not as much as using priors on the signal sources (Figure S10).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

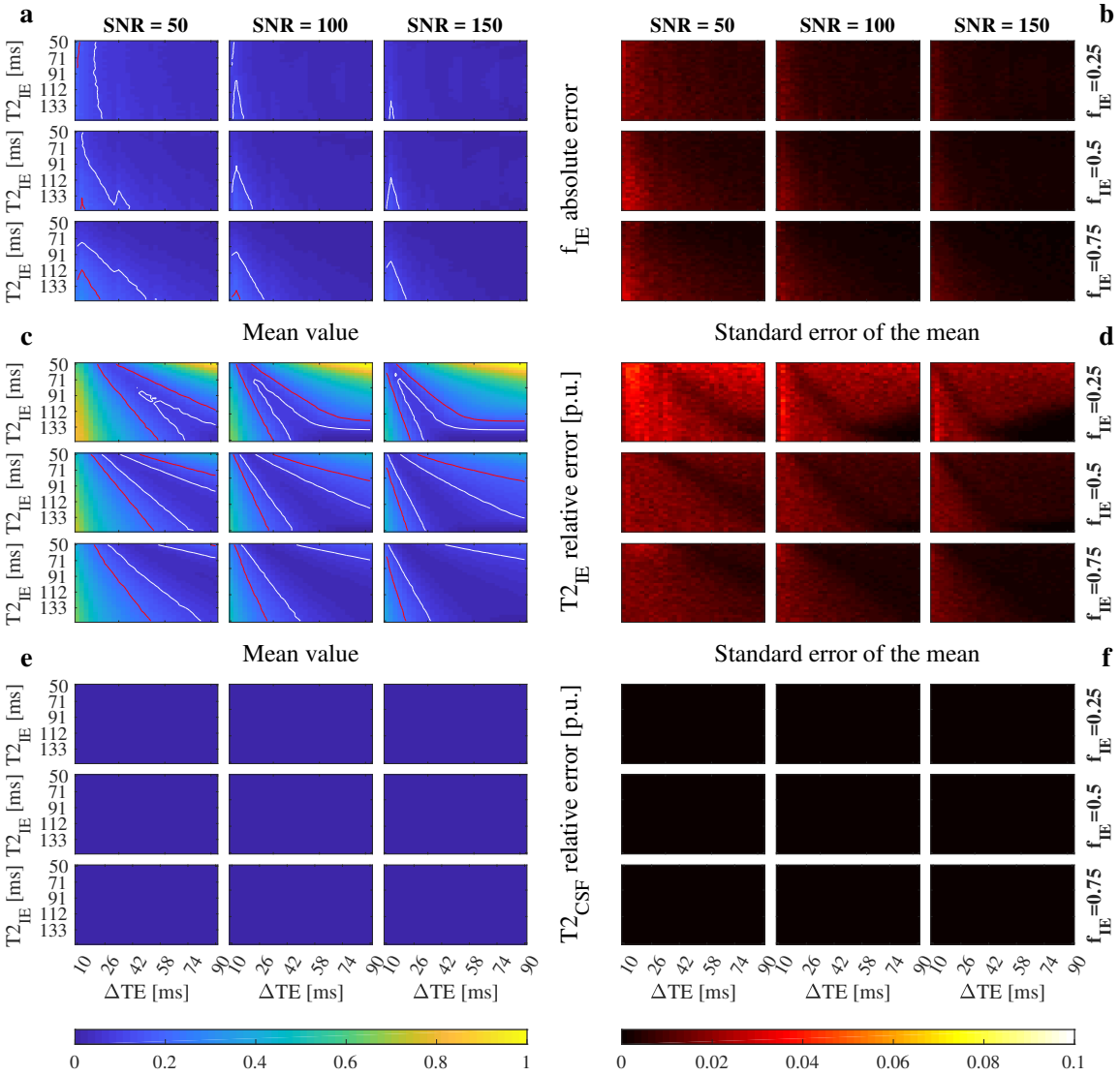

Figure S12: **Convergence for two compartments (IE and CSF) with fixed $T_{2_{CSF}}$ and no S_{CSF} prior.**

The mean and the standard error of f_{IE} absolute error (a and b), and the mean and the standard error of $T_{2_{IE}}$ (c and d), and $T_{2_{CSF}}$ (e and f) relative error per unit (p.u.). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of SNR, f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ was bound between 0–300 and $T_{2_{CSF}}$ fixed to 2000 ms. No prior was imposed on S_{CSF}. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. Fixing the value of $T_{2_{CSF}}$ does not have any effect on the size of the convergence area, while bounding $T_{2_{IE}}$ does it (see Figure S11).

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

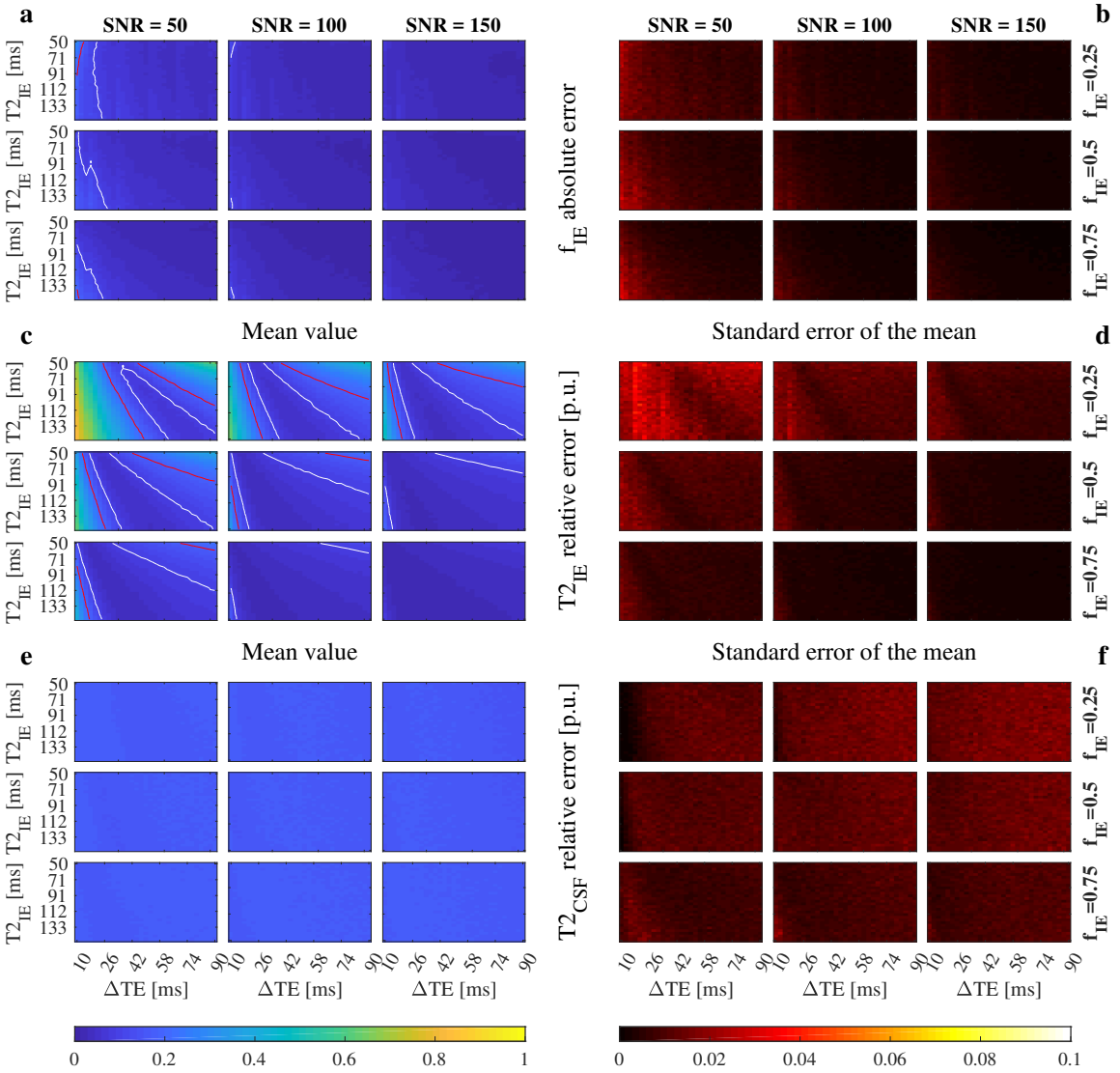

Figure S13: **Convergence for two compartments (IE and CSF) with non-overlapping T_2 constraints and S_{CSF} prior.**

The mean and standard error of f_{IE} absolute error (a and b), and mean and standard error of $T_{2_{IE}}$ (c and d), and $T_{2_{CSF}}$ (e and f) relative error per unit (p.u.). Red and white lines mark the 0.2 and 0.1 contour respectively. One thousand simulations were run for each combination of SNR, f_{IE}, $T_{2_{IE}}$, and ΔTE. $T_{2_{IE}}$ and $T_{2_{CSF}}$ were bound between 0–300 ms and 300–3000 ms respectively. S_{CSF} was set to have isotropic diffusivity with value 3×10^{-3} mm^2/s. We defined the convergence area as the one with error lower than 0.1 for f_{IE} and $T_{2_{IE}}$. Incorporating prior knowledge on the behavior of the signal sources (as CSF) improves convergence and stability more than bounding T_2 (Compare with Figures S10 and S11)

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

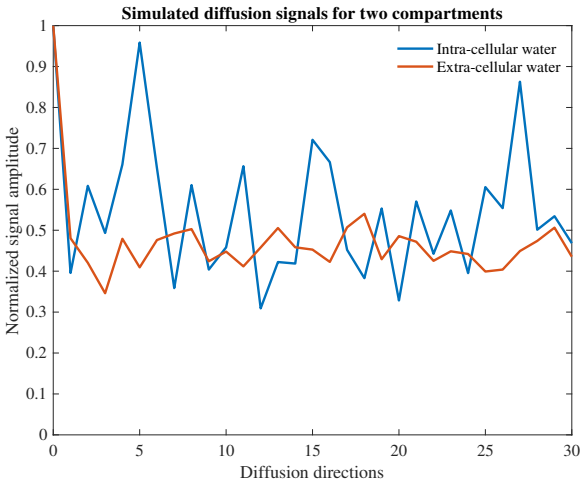

Figure S14: **Simulated diffusion signals for intra and extra-cellular water compartments.** Synthetically generated diffusion signals for 30 directions (b = 1000 s/mm^2) and one non-diffusion weighted measurement. We modeled diffusion as a Gaussian process with MD of intra-cellular (IC) and extra-cellular (EC) equal to 0.6×10^{-3} and 0.8×10^{-3} mm^2/s respectively (to keep the MD of parenchyma equals to 0.7×10^{-3} mm^2/s (28)) and standard deviations of 0.3×10^{-3} and 0.1×10^{-3} mm^2/s respectively to distinguish between a more (IC) and less (EC) hindered anisotropic diffusivity.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

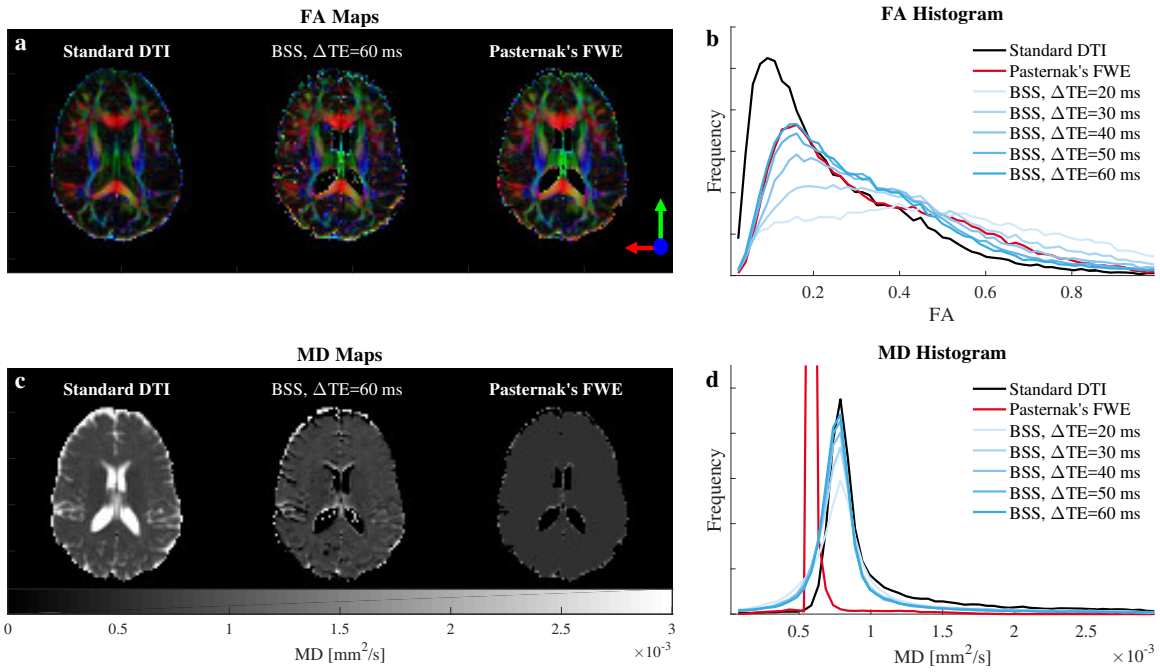

Figure S15: **FA and MD of the BSS-disentangled IE signal against the standard DTI and Pasternak's free-water elimination (FWE) for subject one.**

Comparisons of the FA (b) and MD (d) histograms calculated from the separated IE signals are plotted against the standard DTI fit and Pasternak's method for the short TE measured data. MD (c) and colored FA (a) maps are also included for comparison. We observed a CSF correction effect in the long ΔTE BSS for FA in agreement with Pasternak's FWE. However, both method disagree for MD, where Pasternak's introduces spatial over-regularization. See Figure 9 for subject two.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

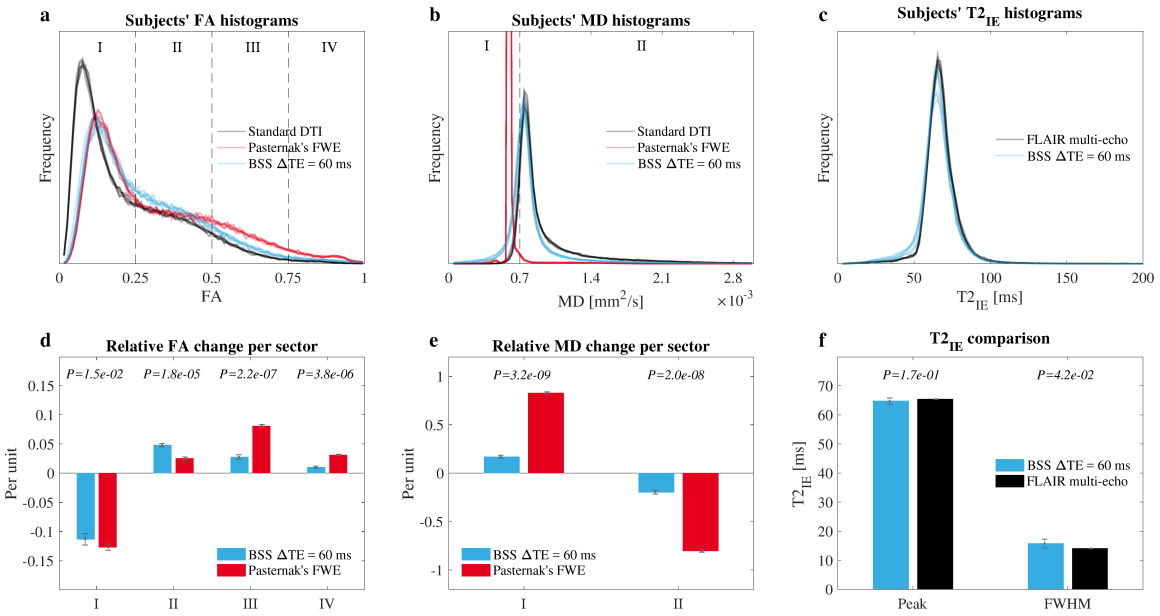

Figure S16: **Repeatability analysis showing intra-subject variability.**

A healthy volunteer was scanned six times. The FA (a) and MD (b) histograms for standard DTI, BSS and Pasternak's method are shown. These histograms were fragmented in sectors and the relative changes in number of voxels per sector and repetition for BSS and Pasternak's methods were computed. Statistical t-tests were run per sector to determine the level of significance of the differences between BSS and Pasternak's results (d and e). BSS and FLAIR $T_{2_{IE}}$ histograms (c) showed good agreement. Their peak and the full width half maximum (FWHM) were used for t-test comparison between BSS and FLAIR (f) highlighting the concordance.

DOI: 10.1002/MRM.27181 ACCEPTED IN MAGNETIC RESONANCE IN MEDICINE

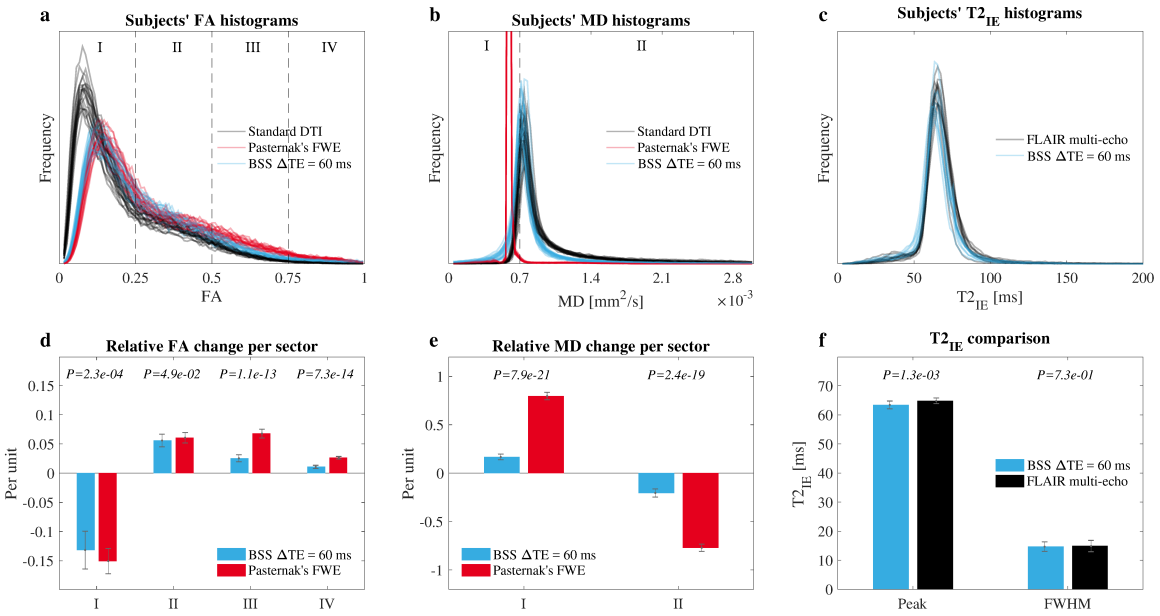

Figure S17: **Reproducibility analysis showing inter-subject variability.**

Twenty healthy volunteers were scanned. The FA (a) and MD (b) histograms for standard DTI, BSS and Pasternak's method are shown. These histograms were fragmented in sectors and the relative changes in number of voxels per sector and repetition for BSS and Pasternak's methods were computed. Statistical t-tests were run per sector to determine the level of significance of the differences between BSS and Pasternak's results (d and e). Notice that the inter-subject variability is larger than intra-subject (Figure S16). BSS and FLAIR $T_{2_{IE}}$ histograms (c) were depicted. Their peak and the full width half maximum (FWHM) were used for t-test comparison between BSS and FLAIR (f).

4.2 PEER-REVIEWED ABSTRACTS

This section contains the abstracts accepted at the International Society for Magnetic Resonance in Medicine (ISMRM) annual conference. These abstract explore tissue microstructure disentangling in two manners: BSS and deep learning. First abstract Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis introduces a BSS solution based on SCA. However, this approach has protocol acquisition limitations that are alleviated by using NSC in the abstract Theory, Validation and Aplication of Blind Source Separation to Diffusion MRI for Tissue Characterisation and Partial Volume Correction. Finally, a deep learning approach is introduced to tackle the FWE problem in the abstract Deep learning with synthetic data for free water elimination in diffusion MRI.

4.2.1 *Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis*

Peer-reviewed Conference Abstract

Authors: M. Molina-Romero, PA. Gómez, JI. Sperl, DK. Jones, MI. Menzel, BH. Menze

In: *ISMRM Workshop on Breaking the Barriers of Diffusion MRI* (2016) [50]

Abstract: Here we present Blind Source Separation (BSS) as a new tool to analyze multi-echo diffusion data. This technique is designed to separate mixed signals and is widely used in audio and image processing. Interestingly, when it is applied to diffusion MRI, we obtain the diffusion signal from each water compartment, what makes BSS optimal for partial volume effects correction. Besides, tissue characteristic parameters are also estimated. Here, we first state the theoretical framework; second, we optimize the acquisition protocol; third, we validate the method with a two compartments phantom; and finally, show an in-vivo application of partial volume correction.

Contribution of thesis author: Model development and implementation, experimental design, abstract revision and editing.

Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis

Miguel Molina-Romero[1,2,3], Pedro A. Gómez[1,2,3], Jonathan I. Sperl[2], Derek K. Jones[3], Marion I. Menzel[2], and Bjoern H. Menze[1]

[1] Computer Science, Technischen Universität München, Munich, Germany
[2] GE Global Research Europe, Munich, Germany
[3] CUBRIC School of Psychology, Cardiff University, Cardiff, United Kingdom.

Abstract. Brain tissue microstructure characterisation through diffusion MRI and relaxometry have high scanning time requirements and need for regularisation to separate tissue components. We present a new approach that does not require regularisation and is less acquisition time demanding. To this end, we use sparse component analysis of the diffusion signal to estimate the number of compartments present in the tissue, their T_2 decays, volume fractions and diffusivities, as well as the proton density for each voxel.

Keywords: Diffusion MRI, multicompartment, microstructure, Sparse Component Analysis, Blind Source Separation.

Introduction

Brain tissue microstructure characterisation through diffusion MRI[1][2][3][4] and relaxometry [5] have been a topic of interest for the last 20 years. However, only few works have considered both methods together [6][7]. These techniques have high scanning time requirements and need for regularisation to separate tissue components. In this abstract we integrate the work on tissue characterization from the relaxometry and diffusion perspectives. We present a new approach that does not require regularisation and is less acquisition time demanding. To this end, we use sparse component analysis (SCA)[8] of the diffusion signal to estimate the number of compartments present in the tissue, their T_2 decays, volume fractions and diffusivities, as well as the proton density (PD) for each voxel. For the signal model, we assume that 1) the brain tissue contains distinct water pools; 2) there is no water exchange between them; and 3) each pool has a different T_2 and diffusivity.

Methods

When a diffusion protocol is acquired for a given echo time (TE) the measurement is a linear combination of the diffusion signals from each compartment

II

($E_i, i = 1$ to N compartments) scaled by the volume fractions (f_i) and the inverse exponential of the ratio between the TE and each T_{2_i} (Eq. 1).

$$S(TE, \Delta, q) = S_0 \sum_{i=1}^{N} f_i e^{\frac{-TE}{T_{2_i}}} E_i(\Delta, q). \tag{1}$$

If the same diffusion experiment is repeated for different values of TE, the mixtures of the linear combination of the sources change, producing different signal strengths, according to Eq. 2.

$$\begin{bmatrix} S(TE_1, \Delta, q) \\ \vdots \\ S(TE_M, \Delta, q) \end{bmatrix} = S_0 \begin{bmatrix} f_1 e^{\frac{-TE_1}{T_{2_1}}} & \cdots & f_M e^{\frac{-TE_1}{T_{2_N}}} \\ \vdots & \ddots & \vdots \\ f_1 e^{\frac{-TE_M}{T_{2_1}}} & \cdots & f_M e^{\frac{-TE_M}{T_{2_N}}} \end{bmatrix} \begin{bmatrix} E_1(\Delta, q) \\ \vdots \\ E_N(\Delta, q) \end{bmatrix}, \tag{2}$$

or equally $X = AS$.

Only the noisy measurements are known (X). However, we are interested in the mixing matrix (A), that only depends on the volume fraction and relaxation properties of the tissue, and the compartmental diffusion sources (S). Typically, this is a blind source separation problem. Approaches based on independent component analysis (ICA) cannot be used since the diffusion sources are not statistically independent. Principal component analysis (PCA) does not offer a good alternative given that the sources are not orthogonal. Therefore, non-negative matrix factorization (NMF) and sparse component analysis (SCA) are the two suitable solutions. NMF is discounted as it requires prior knowledge of the number of compartments. Therefore, SCA is used here. It relies on finding a transform domain where the sources (rows of S) are sparse and disjoint. When these requirements are met, only a few elements in S are non-zero. Then, only one of the sources is active at a time and therefore, the contribution of each specific source to the measured signal can be estimated.

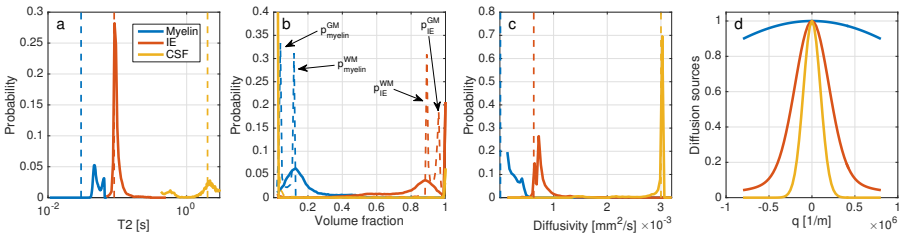

Fig. 1. Histograms for T_2 values (a), volume fractions (b) and diffusivities (c). Dashed lines mark the reference values while solid lines correspond to the estimates. Diffusion signal sources are shown in (d).

We study this approach by simulating a three compartment tissue made of myelin, intra-extra (IE) axonal water and CSF [5]. The volume fraction for each compartment are based on a fuzzy segmentation with tissue probabilities for grey matter (p_{GM}), white matter (p_{WM}) and CSF(p_{CSF}) [9]. We split the WM and GM probabilities into myelin and IE probabilities by multiply them by a volume of myelin of 11.3% for WM and 3.1% for GM [10] ($p^{WM}_{myelin} = 0.113 p_{WM}$, $p^{WM}_{IE} = 0.887 p_{WM}$, $p^{GM}_{myelin} = 0.031 p_{GM}$, $p^{GM}_{IE} = 0.969 p_{GM}$). Data was simulated using three one-dimension diffusion protocol comprising for 32 equally spaced q-values from 0 to 0.8×10^6 m^{-1}, for TE values of 40, 70 and 80 ms. The diffusion sources were generated with Camino [11] using restricted, hindered and free diffusion models for the myelin, IE water and CSF respectively (Fig. 1d). Finally, the T_2 decays for each compartment were 30 ms, 90 ms and 2 s for myelin, IE and CSF [5]. Signals were mixed using Eq.1 to generate the measurements (X). Then, Rician noise was added up to an SNR of 30 dB at b=0. Finally, the signals were disentangled with SCA using a Gaussian wavelet transform. This type of wavelet has a Gaussian envelop that is similar to the measured 1D diffusion signal thus, transformed coefficients are sparse and, as disjoint as different their diffusion coefficients are.

Results

Fig. 2 presents the volume fractions for CSF (a), IE (c) and myelin (e) compared to their reference tissue type probabilities (b, d, f). The normalised PD is shown (Fig. 2g) in comparison to the ground truth (Fig. 2 h). The distribution of the values in Fig. 2(a, c, e) are presented in Fig. 1b along with the disentangled T_2 (Fig. 1a) and diffusivities (Fig. 1c) for each compartment. The probability values in Fig. 1 correspond to the histograms of the parameter for each compartment and cannot be compared to others.

Discussion

High correlation levels between the estimates and the references for CSF, IE and PD are shown in Fig. 2(a-b, c-d and g-h) indicating a good level of separation for the volume fractions. However, myelin correlation level (Fig. 2e-f) is lower, pointing out the difficult detection of short T_2 decay components within the limits of feasible TE values. Fig. 1b shows the histogram distribution volume fractions. The myelin and IE fractions are correctly extracted for WM, while for GM the peaks are shifted to 0 and 1 respectively, due to the low level of myelin present in that tissue. The T_2 distribution (Fig. 1a) presents a bias between 15 to 30 ms for myelin, while accurately capture the reference value for IE and CSF. The diffusivity also shows a good agreement between estimates and reference for the three compartments. Nevertheless, the small amount of myelin signal in the mixture yields a low SNR disentangled source preventing for a reliable diffusivity estimate for this compartment.

IV

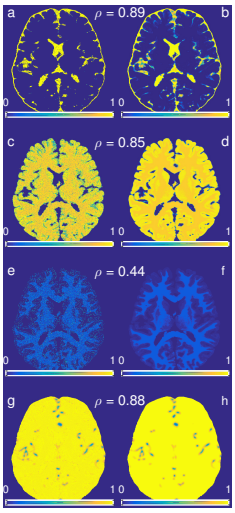

Fig. 2. Comparison between the noisy estimates (left) and the noise free ground truth (right).

Conclusions

To the best of our knowledge, this is the first attempt to use SCA to characterise tissue microstructure using diffusion MRI. Although this proof-of-concept is subject to further improvement, the results presented here indicate that this technique is able to disentangle multiple compartments and thus, can be used to study the T_2, volume fractions, diffusion and PD properties of the tissue microstructure simultaneously. An alternative approach to disentangle relaxometry and diffusion information is presented in [7]. Although, unlike SCA, it relies on heavy regularisation. Finally, since A is independent of the diffusion signal, this technique can be applied to any diffusion protocol. This paves the way for potential applications such as free water elimination or estimation of the compartmental propagators.

Acknowledgments

This work was funded by the European Commission under Grant Agreement Number 605162.

References

1. Greg J Stanisz, Aaron Szafer, Graham A Wright, and R Mark Henkelman. An Analytical Model of Restricted Diffusion in Bovine Optic Nerve. *Magn. Reson. Med.*, 37:103–111, 1997.

V

2. Y Assaf and PJ Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *Neuroimage*, 27:48–58, 2005.
3. H Zhang, T Schneider, Claudia A M Wheeler-Kingshott, and Daniel C Alexander. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61:1000–1016, 2012.
4. Uran Ferizi, Torben Schneider, Eleftheria Panagiotaki, Gemma Nedjati-Gilani, Hui Zhang, Claudia A M Wheeler-Kingshott, and Daniel C Alexander. A ranking of diffusion MRI compartment models with in vivo human brain data. *Magn. Reson. Med.*, 72(6):1785–92, dec 2014.
5. Alex L MacKay, Cornelia Laule, Irene Vavasour, Thorarin Bjarnason, Shannon Kolind, and Burkhard Mädler. Insights into brain microstructure from the T_2 distribution. *Magn. Reson. Med.*, 24(4):515–25, may 2006.
6. S Peled, D G Cory, S A Raymond, D A Kirschner, and F A Jolesz. Water diffusion, T_2, and compartmentation in frog sciatic nerve. *Magn. Reson. Med.*, 42(5):911–8, nov 1999.
7. Daeun Kim, Eamon K Doyle, Jessica L Wisnowski, Joong Hee Kim, and Justin P Haldar. Diffusion-Relaxation Correlation Spectroscopic Imaging: A Multidimensional Approach for Probing Microstructure. *Magn. Reson. Med.*, 2017.
8. Pau Bofill and Michael Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal Processing*, 81(11):2353–2362, 2001.
9. Collins et al. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging*, 17(3):463–468, 1998.
10. Kenneth P Whittall, Alex L Mackay, Douglas A Graeb, Robert A Nugent, David K B Li, and Donald W Paty. In vivo measurement of T_2 distributions and water contents in normal human brain. *Magn. Reson. Med.*, 37(1):34–43, 1997.
11. P A Cook, Y Bai, K K Seunarine, M G Hall, G J Parker, and D C Alexander. Camino: Open-Source Diffusion-MRI Reconstruction and Processing. In *ISMRM 14th Annual Meeting & Exhibition*, volume 14, page 2759, 2006.

4.2.2 *Theory, Validation and Aplication of Blind Source Separation to Diffusion MRI for Tissue Characterisation and Partial Volume Correction*

Peer-reviewed Conference Abstract

Authors: M. Molina-Romero, PA. Gómez, JI. Sperl, AJ. Stewart, DK. Jones, MI. Menzel, BH. Menze

In: *Proc Intl Soc Mag Reson Med* (2017) [51]

Abstract: Brain tissue microstructure characterisation through diffusion MRI and relaxometry have high scanning time requirements and need for regularisation to separate tissue components. We present a new approach that does not require regularisation and is less acquisition time demanding. To this end, we use sparse component analysis of the diffusion signal to estimate the number of compartments present in the tissue, their T_2 decays, volume fractions and diffusivities, as well as the proton density for each voxel.

Contribution of thesis author: Model development and implementation, experimental design, abstract revision and editing.

Theory, Validation and Application of Blind Source Separation to Diffusion MRI for Tissue Characterisation and Partial Volume Correction

Miguel Molina-Romero[1,2], Pedro A Gómez[1,2], Jonathan I Sperl[2], Andrew J Stewart[3], Derek K Jones[4], Marion I Menzel[2], Bjoern H Menze[1]

[1]Computer Science, Technische Universität München, Munich, Germany
[2]GE Global Research, Munich, Germany
[3]EMRIC, Cardiff University, Cardiff, UK
[4]CUBRIC School of Psychology, Cardiff University, Cardiff, UK

Abstract. Here we present blind source separation (BSS) as a new tool to analyse multi-echo diffusion data. This technique is designed to separate mixed signals and is widely used in audio and image processing. Interestingly, when it is applied to diffusion MRI, we obtain the diffusion signal from each water compartment, what makes BSS optimal for partial volume effects correction. Besides, tissue characteristic parameters are also estimated. Here, we first state the theoretical framework; second, we optimise the acquisition protocol; third, we validate the method with a two compartments phantom; and finally, show an in-vivo application of partial volume correction.

1 Purpose

The compartmental nature of tissue is generally accepted [1,7,11,14,17,19]. The diffusion-weighted MRI (dMRI) signal depends on the relaxation times of the compartments ($T2_i$), their diffusivities (D_i), volume fractions (f_i) and proton density (S_0). The simultaneous contribution of these parameters results in a lack of specificity to each independent effect and induces a bias [13,16] on the diffusion metrics known as partial volume contamination. Specificity and partial volume correction problems have been addressed independently [2,6,9,13,14]. Here we present blind source separation (BSS) as a new approach in dMRI that separates mixed signals and yields tissue microstructure parameters, tackling both problems at once.

2 Methods

2.1 Theory

This method is based on three assumptions: 1) tissue is made of water compartments with different diffusivities [6,14]; 2) there is no water exchange [1]; and 3) each compartment has a different $T2$ [6,11,14]. Hence, we can describe the

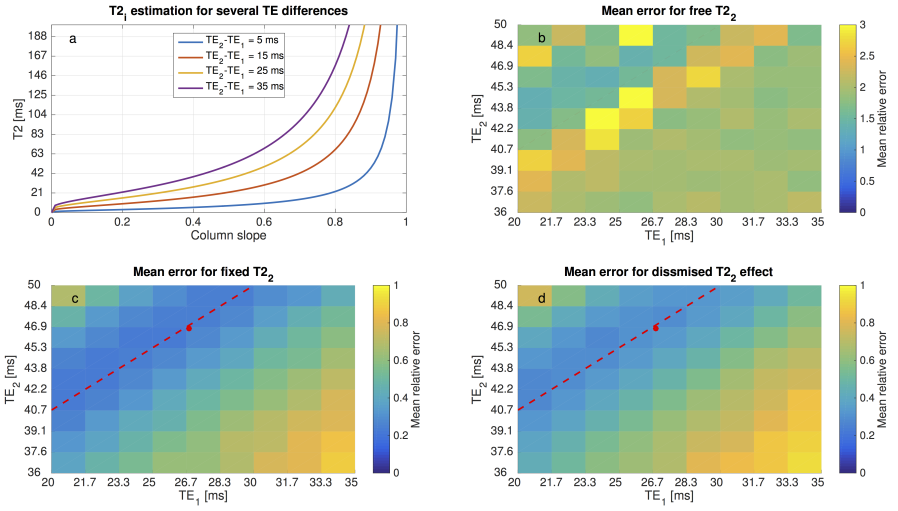

Fig. 1. (b-d) Mean error of the parameter estimations. (a) Relationship between the slope of the columns of $\mathbf{A}$ and the estimation of $T2$ for several TE differences. When the slope of the columns tends towards 1 ($T2 \gg TE$), the estimation of $T2$ is in the asymptotic region and thus uncertain. This uncertainty can be observed in (b) where the minimum error is larger than in (c,d) for fixed $T2_2$ and dismissed $T2_2$ effect. Notice that the optimal TE pairs are marked by the red dashed lines. The red dots mark the TE pair used for phantom validation experiment.

measured diffusion signal as the weighted sum of the compartmental sources. These weights depend only on the volume fraction (f) and the ratio between the compartmental $T2_i$ and the experimental TE_j. Therefore, varying TE modifies the weights and the system can be expressed as a BSS problem:

$$\begin{bmatrix} X(TE_1, \Delta, q) \\ \vdots \\ X(TE_M, \Delta, q) \end{bmatrix} = \begin{bmatrix} f_1 e^{TE_1/T2_1} & \cdots & f_N e^{TE_1/T2_N} \\ \vdots & \ddots & \vdots \\ f_1 e^{TE_M/T2_1} & \cdots & f_N e^{TE_M/T2_N} \end{bmatrix} \begin{bmatrix} S_1(\Delta, q) \\ \vdots \\ S_N(\Delta, q) \end{bmatrix} S_0 \qquad (1)$$

$$\mathbf{X} = \mathbf{AS}, \qquad (2)$$

where $\mathbf{X}$ are the measurements for several TEs, $\mathbf{A}$ the mixing matrix, $\mathbf{S}$ the compartmental diffusion source, M the number of measurements, and N the number of compartments. Here, among the possible BSS solutions [18], and unlike in [12], we use a sparsifying transform [15] followed by non-negative sparse coding [8].

Here we focus on two-compartment environments ($N = M = 2$). Besides, when $T2_i$ is larger than the TEs (i.e. CSF), the exponential term can be dismissed ($\exp(TE_j/T2_i) \approx 1$) and thus the $T2_i$. Alternatively, $T2_i$ can be fixed to

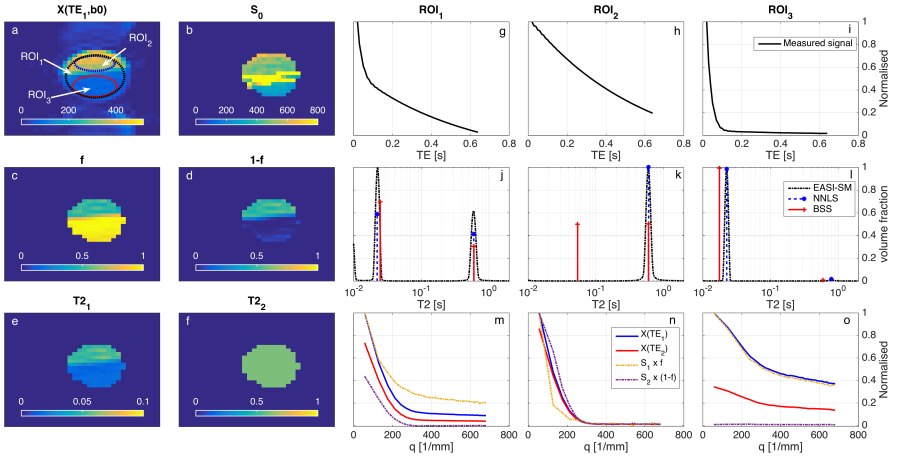

Fig. 2. (a) b0 image at $TE_1 = 26\ ms$ with ROIs overlaid. Each ROI represents a possible case: ROI_1 ($f \approx 0.5$), whole phantom; ROI_2 ($f \approx 0$), water; ROI_3 ($f \approx 1$), yeast. (b) Signal intensity at $TE = 0\ ms$. Volume fractions for the associated intra-cellular (c) and extra-cellular (d) compartments. $T2$ for the intra-cellular (e) and extra-celullar (f) cell compartments. Averaged multi-echo signal for each ROI (g,h,i) and the corresponding $T2$ spectral fitting with NNLS and EASI-SM (j,k,l) compared with the volume fractions and $T2$s estimated by BSS ($T2_2$ fixed at 0.6 s according to NNLS and EASI-SM). Measured and separated diffusion signals for each ROI (m,n.o).

an expected value if prior knowledge is available (i.e. $\text{T2}_{CSF} \approx 2\ s$ 6). We study the effect both approximations on the error of the parameter estimation.

We perform three experiments to: 1) find the range of optimal TEs; 2) validate our method; and 3) show an application. Figure 4 contains the experimental details.

2.2 Optimisation simulations

Tissue with two compartments was simulated with known $T2$s (22 and 597 ms) for restricted and free diffusion signals [4]. We ran a simulation experiment varying TE and f (11 points) to calculate the mean error for all the parameter combinations and find the optimal TE region for free, fixed and dismissed $T2_2$.

2.3 Phantom validation

For validation, we used a phantom made of yeast and water (1:1) as a two compartments sample [5]. A multi-echo experiment was acquired and $T2$s fitted with NNLS [10] and EASI-SM [3]. Besides, BSS was applied on the diffusion dataset fixing $T2_2 = 0.6\ s$ (NNLS). Finally, results from the three methods were compared.

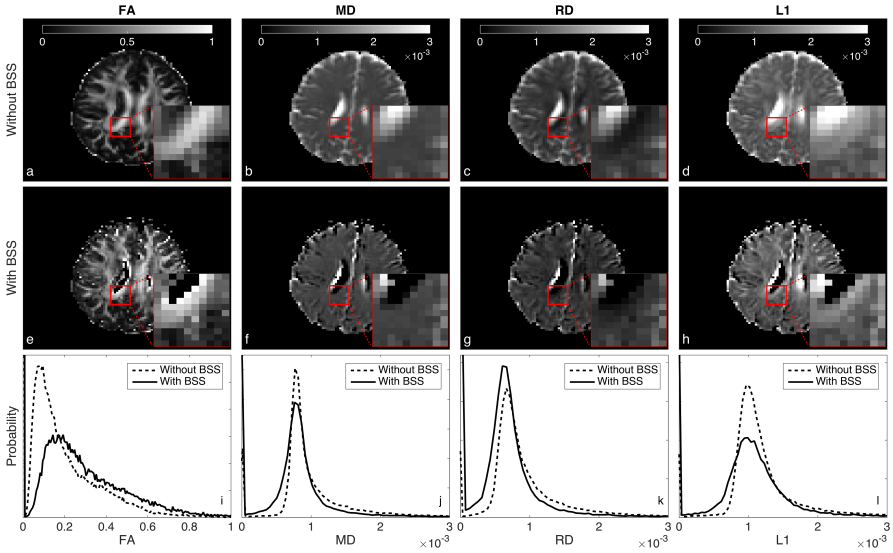

Fig. 3. Comparison of DTI metrics with and without CSF contamination correction by BSS. Histograms of values for the whole brain (i-l) show an increase of FA, and a decrease of MD, RD and L1. Both effects are consistent with the elimination of the CSF contribution. Besides, we observe a significant increase of FA in the borders of the ventricles (zoomed area), where the contamination is expected to be high. Notice that BSS mostly crops the ventricles and the external CSF and increases the contrast of the white matter.

2.4 In vivo

A young female volunteer went under a DTI acquisition. CSF signal was extracted from the data using BSS, fixing $T2_2 = 2\ s$ [11]. Finally, DTI metrics with and without correction were compared.

3 Results and discussion

3.1 Optimisation simulations

Fig 1a depicts $T2$ versus the slope of a column of $\mathbf{A}$. As the slope tends towards 1, the estimation falls into an asymptotic region increasing the uncertainty on the $T2$ estimation. Therefore, fixing its value or dismissing its contribution reduces the mean error of the parameter estimations (Fig. 1b-d). Moreover, fixing the $T2$ value performs slightly better than dismissing its effect (Fig. 1c-d).

3.2 Phantom validation

Fig.2g-o compare the results of BSS against NNLS and EASI-SM in a ROI-based analysis. Fig. 2j,l show agreement of $T2_1$ and f with NNLS and EASI-SM for

		Optimisation	Validation	In-Vivo
Diffusion acquisition	TR	Inf	3 s	4 s
	TE_1	20-35 ms (10 points)	26 ms	73 ms
	TE_2	36-50 ms (10 points)	46 ms	110 ms
	#Diffusion directions	1	1	41
	#b-values	33	33	2
	Range of b-values	0-13504 s/mm^2	70-10048 s/mm^2	0-1000 s/mm^2
	Reverse polarity	No	No	Yes
	Resolution	1 voxel	32x32	96x96
	Sequence	PGSE	STEAM	PGSE
T2 mapping	TR		2.5 s	
	#TEs		64	
	Range of TEs		10-640 ms	
	NEX		4	
	Resolution		32x32	
	Sequence		SE	
	Scanner	Simulated	9.4T Biospec 94/20 (Bruker, Ettlingen, Germany)	3T HDx MRI (GE, Milwaukee, WI)

Fig. 4. Experimental setups for the optimisation simulation, the phantom validation and the in-vivo experiment.

ROI_1 and ROI_3. Besides, in Fig. 1m, S_1 (associated with intra-cellular space) describes a restricted diffusion signal similar as in Fig 2o, and S_2 (associated with extra-cellular space) shows a free diffusion behaviour as in Fig. 2n. Both findings are in agreement with the simulations and indicate that BSS successfully separates signals from two compartments. Interestingly, BSS disentangles measurements from ROI_2 into two similar and equally scaled sources (Fig. 2n) indicating that only one source exists. For illustration, Fig. 2b-f show that the voxel-based maps generated with BSS are consistent with the ROI based analysis.

3.3 In vivo

In Fig. 3, with BSS, we observe an increase of the fractional anisotropy (FA) (a,e,i) and a reduction of the mean diffusivity (MD) (b,f,j), radial diffusivity (RD) (c,g,k), and tensor's main eigenvalue (L1) (d,h,l). This is consistent with the elimination of the CSF contribution. Also, we notice that with BSS the ventricles are extracted and white matter structures are better defined, especially the voxels at the border of the ventricles (zoomed area).

4 Conclusions

Here we show that BSS of diffusion data is a suitable technique to separate compartmental sources. We demonstrate that this method is appropriate for partial volume correction. Besides, tissue volume fraction, relaxation and diffusivity parameters are estimated allowing for simultaneous tissue characterisation.

5 Acknowledgments

With the support of the TUM Institute for Advanced Study, funded by the German Excellence Initiative and the European Commission under Grant Agreement Number 605162.

References

1. Assaf, Y., Basser, P.J.: Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. NeuroImage 27, 48–58 (2005)
2. Benjamini, D., Basser, P.J.: Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry. Journal of Magnetic Resonance 271, 40–45 (2016)
3. Björk, M., Zachariah, D., Kullberg, J., Stoica, P.: A multicomponent T 2 relaxometry algorithm for myelin water imaging of the brain. Magnetic Resonance in Medicine (2015)
4. Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., Alexander, D.C.: Camino: Open-Source Diffusion-MRI Reconstruction and Processing. Proc Intl Soc Mag Reson Med (2014)
5. Cory, D.G.: Measurement of translational displacement probabilities by NMR: An indicator of compartmentation. Magnetic Resonance in Medicine 14(3), 435–444 (1990)
6. Does, M.D., Gore, J.C.: Compartmental study of diffusion and relaxation measured in vivo in normal and ischemic rat brain and trigeminal nerve. Magnetic resonance in medicine 43(6), 837–44 (2000)
7. Ferizi, U., Schneider, T., Witzel, T., Wald, L.L., Zhang, H., Wheeler-Kingshott, C.A., Alexander, D.C.: White matter compartment models for in vivo diffusion MRI at 300mT/m. NeuroImage 118, 468–483 (2015)
8. Hoyer, P.: Non-negative sparse coding. In: Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing. pp. 557–565. IEEE (2002)
9. Kim, D., Kim, J.H., Haldar, J.P.: Diffusion-Relaxation Correlation Spectroscopic Imaging (DR-CSI): An Enhanced Approach to Imaging Microstructure. Proc Intl Soc Mag Reson Med (2016)
10. Lawson, C.L., Hanson, R.J.: Solving Least Squares Problems. Society for Industrial and Applied Mathematics (1995)
11. Mackay, A., Laule, C., Vavasour, I., Bjarnason, T., Kolind, S., M7dler, B.: Insights into brain microstructure from the T 2 distribution. Magnetic Resonance Imaging 24, 515–525 (2006)
12. Molina-Romero, M., Gómez, P.A., Sperl, J.I., Jones, D.K., Menzel, M.I., Menze, B.H.: Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis. ISMRM Workshop on Breaking the Barriers of Diffusion MRI (2016)

13. Pasternak, O., Sochen, N., Gur, Y., Intrator, N., Assaf, Y.: Free water elimination and mapping from diffusion MRI. Magnetic Resonance in Medicine 62(3), 717–730 (2009)
14. Peled, S., Cory, D.G., Raymond, S.A., Kirschner, D.A., Jolesz, F.A.: Water diffusion, T(2), and compartmentation in frog sciatic nerve. Magnetic resonance in medicine 42(5), 911–8 (1999)
15. Ravishankar, S., Bresler, Y.: 0 Sparsifying Transform Learning with Efficient Optimal Updates and Convergence Guarantees. arXiv preprint arXiv:11501.02859 pp. 1–16 (2015)
16. Santis, S.D., Assaf, Y., Jones, D.: The influence of T2 relaxation in measuring the restricted volume fraction in diffusion MRI. Proc Intl Soc Mag Reson Med (2016)
17. Stanisz, G.J., Wright, G.A., Henkelman, R.M., Szafer, A.: An analytical model of restricted diffusion in bovine optic nerve. Magnetic Resonance in Medicine 37(1), 103–111 (1997)
18. Yu, X., Hu, D., Xu, J.: Blind source separation: theory and applications. Wiley (2014)
19. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C.: NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61(4), 1000–1016 (2012)

4.2.3 *Deep learning with synthetic data for free water elimination in diffusion MRI*

Peer-reviewed Conference Abstract

Authors: M. Molina-Romero, PA. Gómez, S. Albarqouni, JI. Sperl, MI. Menzel, BH. Menze

In: *Proc Intl Soc Mag Reson Med* (2018) [53]

Abstract: Brain tissue microstructure characterization through diffusion MRI and relaxometry have high scanning time requirements and need for regularization to separate tissue components. We present a new approach that does not require regularization and is less acquisition time demanding. To this end, we use sparse component analysis of the diffusion signal to estimate the number of compartments present in the tissue, their T_2 decays, volume fractions and diffusivities, as well as the proton density for each voxel.

Contribution of thesis author: Discussion of algorithmic implementation and experimental design, manuscript revision and editing.

Deep learning with synthetic data for free water elimination in diffusion MRI

Miguel Molina-Romero[1,2], Pedro A. Gómez[1,2], Shadi Albarqouni[1], Jonathan I. Sperl[2], Marion I. Menzel[2], and Bjoern H. Menze[1]

[1] Computer Science, Technischen Universität München, Munich, Germany
[2] GE Global Research Europe, Munich, Germany

Abstract. Diffusion metrics are typically biased by Cerebrospinal fluid (CSF) contamination. In this work, we present a deep learning based solution to remove the CSF contribution. First, we train an artificial neural network with synthetic data to estimate the tissue volume fraction. Second, we use the resulting network to predict estimates of the tissue volume fraction for real data, and use them to correct for CSF contamination. Results show corrected CSF contribution which, in turn, indicates that the tissue volume fraction can be estimated using this joint data generation and deep learning approach.

Keywords: Diffusion MRI, Deep Learning, Free-water elimination, Cerebrospinal fluid, Partial volume contamination.

Introduction

Cerebrospinal fluid (CSF) partial volume contamination poses a problem for detecting changes in tissue microstructure [1], biasing the diffusion measurements and derived metrics. CSF is mostly composed of free water, with isotropic diffusion and diffusivity three times bigger than parenchyma [2].

FLAIR DWI [3] tackles the problem suppressing the CSF signal during acquisition, at the cost of low SNR and longer acquisition times. Post-processing solutions have focused on fitting a bi-tensor model; yet, this is an ill-posed problem with several regularizations [1][2][4][5][6][7].

In this work, we hypothesize and show that artificial neural networks (ANN) can estimate the tissue volume fraction from the diffusion signal. Then the CSF contribution can be corrected.

Methods

Theory: CSF has isotropic diffusion with diffusivity $D_{CSF} = 3 \times 10^{-3}$ mm^2/s [2] and can be computed from b:

$$S_{CSF} = e^{-bD_{CSF}}, \tag{1}$$

II

The measured signal is the contribution of CSF and tissue (parenchyma) components:

$$S = f \cdot S_{tissue} + (1 - f)S_{CSF}. \tag{2}$$

Eq. 2 is ill-posed since S_{tissue} and its volume fraction, f, are unknowns. In this work, we present a deep learning approach that uses ANNs to estimate f, regularizing the problem:

$$S_{tissue} = \frac{S - (1 - f)S_{CSF}}{f}. \tag{3}$$

Generation of synthetic data: The training dataset were designed to teach the ANN to detect CSF-like components mixed with a random signal (Fig. 1). CSF signal was derived from Eq.1 and acquisition parameter b. Tissue signal was randomly generated to simulate undetermined directions. The generation steps were:

1. $S_{CSF}^{training}$ was computed (Eq. 1).
2. $S_{tissue}^{training}$ was randomly created simulating arbitrary directions: $U(0, 1)$.
3. f was randomly generated: $U(0, 1)$.
4. $S^{training}$ was computed (Eq. 2).
5. The ANN was trained with input to match the output (Fig. 2).

Free water elimination: For comparison, we trained [8] five ANN architectures in MATLAB (MathWorks, Natick, MA) for datasets with 32 directions (one shell) and 64 directions (two shells), (Fig. 2). We chose the best performing ANNs and compared them against Pasternak's [4] and Hoy's [6][9] methods.

Data acquisition: A volunteer went under a diffusion acquisition (GE 3T MR750w, Milwaukee, WI) with 30 directions; 2 shells: b = 500, 1000 s/mm^2; four b = 0 s/mm^2; TR/TE = 8000/80 ms; FOV = 200 mm; resolution 128x128; ASSET = 2; and 25 slices with 3.6 mm thickness and no gap.

Pipeline:

1. Diffusion measurement.
2. Synthetic data generation from the experimental b (Fig. 1).
3. ANN training.
4. Volume fraction estimation: $ANN(S) \rightarrow f$.
5. Computation of S_{tissue} (Eq. 3).
6. Fitting of the tensor model [10][11] on S_{tissue}.

Results

The five ANN architectures (Fig. 2) showed similar performance (Fig. 3). ANNs trained for two shells (ANN2s) outperformed those for one shell (ANN1s), due to the better CSF encoding of two shells protocols. The best performing ANNs

were L=2 and L=3 for one and two shells respectively, suggesting a potential coupling between the number of hidden layers and shells.

DTI metrics after ANNs correction showed differences depending on the number of shells. ANN1s estimated larger volumes of CSF than ANN2s (Fig. 4c), that resulted in larger FA (Fig. 4a) and lower MD (Fig. 4b) estimates. This difference on the f estimate might be explained by the limited CSF information contained in the single shell protocol. MD values for ANN2s (Fig. 4b) agreed with the reference [2].

ANNs kept the anatomical integrity of the FA, MD, and f_{CSF} maps (Fig. 5). We observed the CSF correction in the enlargement of the corpus callosum and fornix, and a general increment of FA in white matter, compared to the standard DTI (Fig. 5a,c,e,f). CSF contribution was accurately removed from MD maps, especially for ANN2s (Fig. 5g,h,k,l). ANNs1 and ANNs2 differ on the f estimate in white matter (Fig. 5m,o), as previously explained.

Discussion

ANNs trained with synthetic data are capable of estimating the tissue volume fraction from the measured diffusion signal. Their correction is equivalent to well-established methods: Pasternak et al. and Hoy et al. (Fig. 4 and Fig. 5).

Using ANNs has a performance advantage. Their training time is in the order of ten minutes and once trained they can be used for any data acquired with the same protocol. CSF correction is faster than traditional methods. For one shell, Pasternaks method ran for 38.4s and ANN1s for 0.7s (55x). For two shells, Hoys method ran for 392.5s and ANN2s for 1.3s (302x). Besides, to improve the accuracy, one can carefully design the training dataset to mimic only tissue characteristics (here it is random), or incorporate prior knowledge of the bi-exponential problem and noise model into the learning process [12].

Conclusions

This is the first application of ANNs to remove CSF contamination. We proved that tissue volume fraction can be estimated by ANNs trained with synthetic data, creating a new tool for free water elimination.

Acknowledgments

With the support of the TUM Institute of Advanced Study, funded by the German Excellence Initiative and the European Commission under Grant Agreement Number 605162.

References

1. C. Metzler-Baddeley et al. How and how not to correct for CSF-contamination in diffusion MRI. *Neuroimage*, 59(2):1394–1403, 2012.

IV

2. C. Pierpaoli and D.K. Jones. Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model. In *ISMRM, Kyoto*, page 1215, 2004.
3. Guoying Liu, Peter Van Gelderen, Jeff Duyn, and Chrit T W Moonen. Single-shot diffusion MRI of human brain on a conventional clinical instrument. *Magn. Reson. Med.*, 35(5):671–677, 1996.
4. O. Pasternak et al. Free Water Elimination and Mapping from Diffusion MRI. *Magn. Reson. Med.*, 730:717–730, 2009.
5. Z. Eaton-Rosen et al. Beyond the Resolution Limit: Diffusion Parameter Estimation in Partial Volume. *MICCAI*, pages 605–612, 2016.
6. A.R. Hoy et al. Optimization of a Free Water Elimination Two-Compartment Model for Diffusion Tensor Imaging. *Neuroimage*, (103):323–333, 2014.
7. M. Molina-Romero et al. Theory, validation and application of blind source separation to diffusion MRI for tissue characterisation and partial volume correction. In *ISMRM, Honolulu*, page 3462, 2017.
8. Martin T. Hagan and Mohammad B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Trans. Neural Networks*, 5(6):989–993, 1994.
9. Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan van der Walt, Maxime Descoteaux, and Ian Nimmo-Smith. Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.*, 8, 2014.
10. P J Basser, J Mattiello, and D LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophys. J.*, 66(1):259–267, 1994.
11. Mark Jenkinson, Christian F. Beckmann, Timothy E.J. Behrens, Mark W. Woolrich, and Stephen M. Smith. FSL. *Neuroimage*, 62(2):782–790, 2012.
12. Jonas Adler and Ozan Öktem. Solving ill-posed inverse problems using iterative deep neural networks. (1):1–24, 2017.

V

Figures

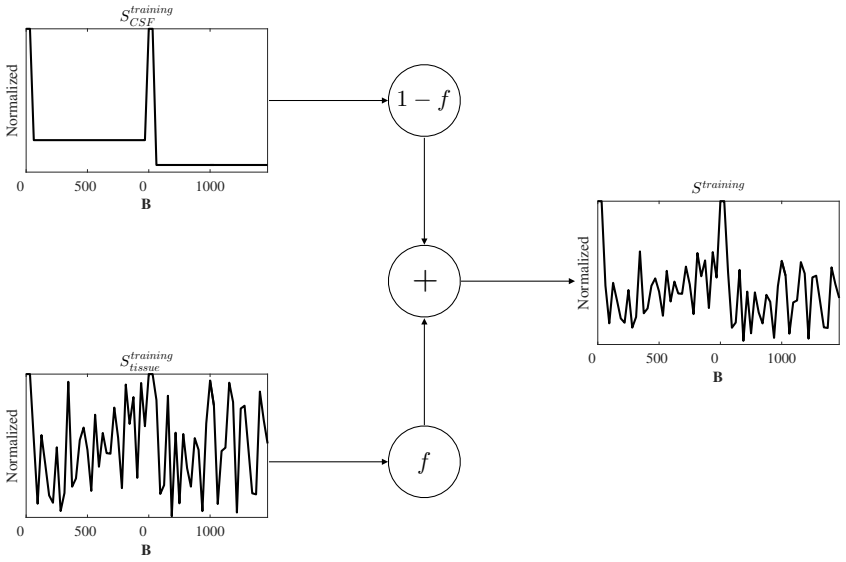

Fig. 1. Generation of synthetic data. The vectorization of the diffusion MRI signal along the diffusion directions (**B**) shows a tissue dependent pattern. S_{CSF} is characterized by Eq. 1 and can be calculated from the diffusion protocol (b values). S_{tissue} depends on the tissue anisotropy and acquired directions, thus it cannot be predicted. We represented S_{tissue} as a uniformly distributed signal, $U(0,1)$, with maximums where $b = 0$ s/mm^2. Tissue volume fraction, f, was also generated uniformly, $U(0,1)$. Finally, $S^{training}$ was computed as in Eq. 2, and presented to the input of the ANN, and f to the output for training (Fig. 2).

VI

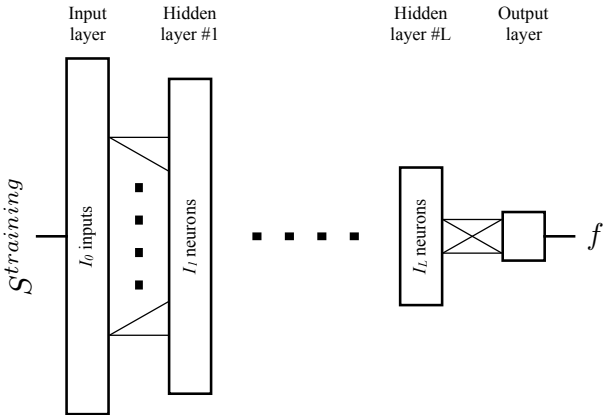

Fig. 2. ANN architectures. Five architectures with L = 1–5 were tested to determine their performance. For L = 1 hidden layer, I_1=$I_0/3$. For L = 2 hidden layers, $I_1 = I_0/2$ and $I_2 = I_0/4$. For L = 3 hidden layers, $I_1 = I_0/2$, $I_2 = I_0/3$ and $I_3 = I_0/4$. For L = 4 hidden layers, $I_1 = I_0/2$, $I_2 = I_0/3$, $I_3 = I_0/4$ and $I_4 = I_0/5$. For L = 5 hidden layers, $I_1 = 2 \times I_0/3$, $I_2 = I_0/2$, $I_3 = I_0/3$, $I_4 = I_0/4$ and $I_5 = I_0/5$. The number of inputs, I_0, matched the number of diffusion directions and non-diffusion-weighted volumes. In these experiments, we used $I_0 = 32$ for one shell and $I_0 = 64$ for two shells. One million signal combinations and volume fractions were generated for training, 20% were separated for validation and 20% for testing.

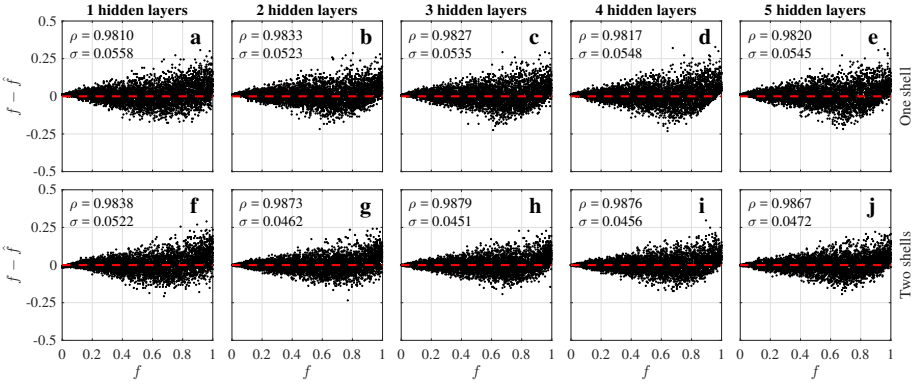

Fig. 3. Performance comparison for the five ANN architectures. We generated 5000 artificial diffusion signals for FA = 0–1 and f = 0–1. They were mixed as in Eq. 2 and Fig. 1 and presented to the trained ANNs to estimate $\hat{f}$. We plot the error $(f - \hat{f})$ of the estimated volume fraction $(\hat{f})$ against its true value (f), their correlation (ρ), and the standard deviation of the error (σ). For ANN1s, we found the largest ρ and minimum σ for L = 2; and L = 3 for ANN2s. We used L=2 for the one shell and L=3 for two shells for in vivo experiments.

VII

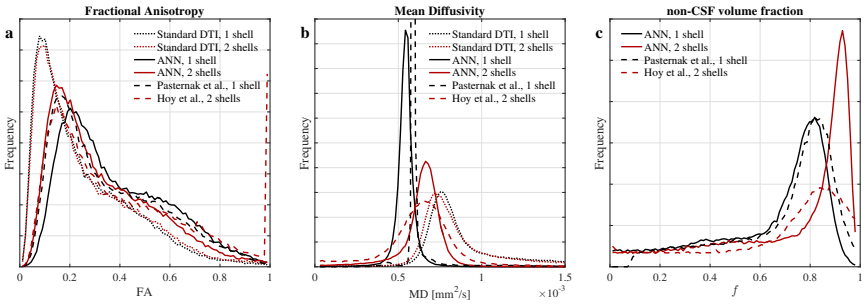

Fig. 4. Comparison of FA, MD and f histograms. FA (a) and MD (b) were consistent for standard DTI of one and two shells, fixing a common reference. ANN1s showed larger correction of FA. Hoys method did not correct FA = 0.15–0.45. ANN2s and Pasternaks method showed stable correction for all FA values. However, ANN1s and Pasternaks method suffered from over regularization of MD (b), with peaks off the reference (0.7 mm^2/s). Volume fraction estimates (c) for ANN1s and Pasternaks were similar, but the later struggled to estimate small f. ANN2s estimated less CSF volume (c) in white matter than other methods (Fig. 5o).

VIII

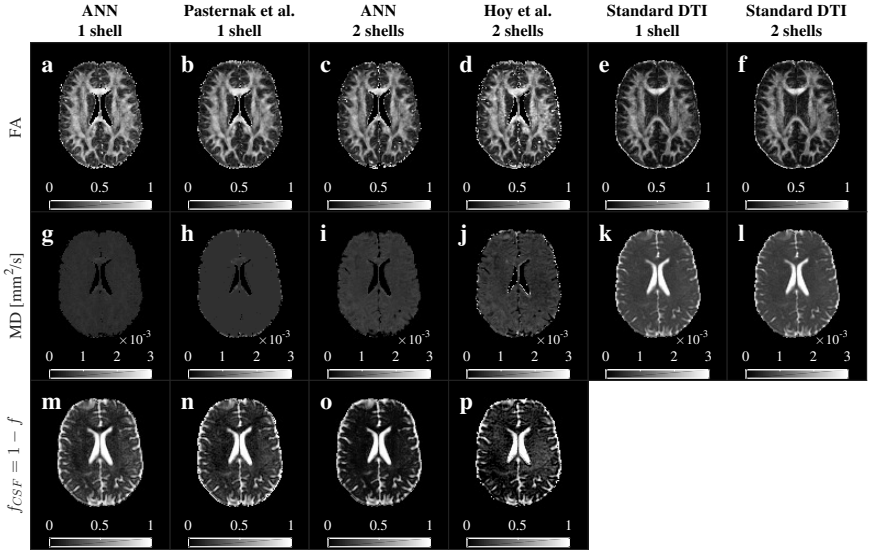

Fig. 5. Comparison of FA, MD and f_{CSF} maps. ANNs maps showed anatomical coherence with standard DTI, Pasternaks and Hoys methods. We observed an enlargement of the corpus callosum and recovery of the fornix in FA for all the methods (a-d) comparted to the standard (e-f). CSF contribution was removed from all the MD maps (g-j vs k-l). MD maps for two shells methods (i-j) contained more information than one shell (g-h). Single shell methods (m-n) showed a larger CSF volume estimate in white matter. ANN2s estimated a lower and more homogeneous CSF volume (o) than Hoys (p) in white matter.

5

MULTIPARAMETRIC MAPPING IN THE TRANSIENT-STATE

5.1 PEER-REVIEWED PUBLICATIONS

Two peer-reviewed scientific articles published at Medical Imaging Computing and Computer Assisted Intervention (MICCAI) are presented in this section. The benefits of an image reconstruction method merging CS and machine learning is shown in Learning a Spatiotemporal Dictionary for Magnetic Resonance Fingerprinting with Compressed Sensing. This concept is further developed incorporating spatiotemporal patches matching to an MRF acquisition in Simultaneous Parameter Mapping, Modality Synthesis, and Anatomical Labeling of the Brain with MR Fingerprinting. This enables the estimation of supplementary quantitative parameters.

5.1.1 *Learning a Spatiotemporal Dictionary for Magnetic Resonance Fingerprinting with Compressed Sensing*

Peer-reviewed Conference Paper

Authors: PA. Gómez, C. Ulas, JI. Sperl, T. Sprenger, **M. Molina-Romero**, MI. Menzel, BH. Menze

In: *MICCAI Patch-MI Workshop* LNCS 9467 (2015), pp. 112–119 [25]

Abstract: Magnetic Resonance Fingerprinting (MRF) is a novel technique that allows for the fast and simultaneous quantification of multiple tissue properties, progressing from qualitative images, such as T_1- or T_2-weighted images commonly used in clinical routines, to quantitative parametric maps. MRF consists of two main elements: accelerated pseudorandom acquisitions that create unique signal evolutions over time and the voxel-wise matching of these signals to a dictionary simulated using the Bloch equations. In this study, we propose to increase the performance of MRF by not only considering the simulated temporal signal, but a full spatiotemporal neighborhood for parameter reconstruction. We achieve this goal by first training a dictionary from a set of spatiotemporal image patches and subsequently coupling the trained dictionary with an iterative projection algorithm consistent with the theory of Compressed Sensing (CS). Using data from BrainWeb, we show that the proposed patch-based reconstruction can accurately recover T_1 and T_2 maps from highly undersampled k-space measurements, demonstrating the added benefit of using spatiotemporal dictionaries in MRF.

Contribution of thesis author: Discussion of algorithmic implementation and experimental design, manuscript revision and editing.

Learning a Spatiotemporal Dictionary for Magnetic Resonance Fingerprinting with Compressed Sensing

Pedro A. Gómez[1,2,3(✉)], Cagdas Ulas[1], Jonathan I. Sperl[3], Tim Sprenger[1,2,3], Miguel Molina-Romero[1,2,3], Marion I. Menzel[3], and Bjoern H. Menze[1]

[1] Computer Science, Technische Universität München, Munich, Germany
pedro.gomez@tum.de
[2] Biomedical Engineering, Technische Universität München, Munich, Germany
[3] GE Global Research, Munich, Germany

Abstract. Magnetic resonance fingerprinting (MRF) is a novel technique that allows for the fast and simultaneous quantification of multiple tissue properties, progressing from qualitative images, such as T1- or T2-weighted images commonly used in clinical routines, to quantitative parametric maps. MRF consists of two main elements: accelerated pseudorandom acquisitions that create unique signal evolutions over time and the voxel-wise matching of these signals to a dictionary simulated using the Bloch equations. In this study, we propose to increase the performance of MRF by not only considering the simulated temporal signal, but a full spatiotemporal neighborhood for parameter reconstruction. We achieve this goal by first training a dictionary from a set of spatiotemporal image patches and subsequently coupling the trained dictionary with an iterative projection algorithm consistent with the theory of compressed sensing (CS). Using data from BrainWeb, we show that the proposed patch-based reconstruction can accurately recover T1 and T2 maps from highly undersampled k-space measurements, demonstrating the added benefit of using spatiotemporal dictionaries in MRF.

1 Introduction

Quantitative magnetic resonance imaging (qMRI) techniques measure relevant biological parameters, providing a profound characterization of the underlying tissue. In contrast to conventional weighted MRI, where the image signal is represented by intensity values and different tissues are described relative to each other, qMRI generates parametric maps of absolute measures that have a physical interpretation, leading to reduced bias and reproducible diagnostic information. On the other hand, obtaining quantitative maps is a time consuming task. It requires the repeated variation of typical MR acquisition parameters, such as flip angle (FA) or repetition time (TR), and the fitting of the measured signal to a model in order to estimate the parameters of interest, including the MR specific longitudinal (T1) and transversal (T2) relaxation times. Long acquisition times, together with high sensitivity to the imaging device

G. Wu et al. (Eds.): Patch-MI 2015, LNCS 9467, pp. 112–119, 2015.
DOI: 10.1007/978-3-319-28194-0_14

and system setup, are the main restrictions to clinical applications of qMRI techniques.

A recently proposed qMRI method, magnetic resonance fingerprinting (MRF), aims to overcome these limitations through accelerated pseudorandom acquisitions [6]. It is based on the idea that pseudorandom variations on acquisition parameters cause the signal response for different tissue types to be unique. This unique signal evolution can be matched to a precomputed dictionary created from known combinations of the parameters of interest (e.g. T1 and T2). Therefore, by matching the measured signal to one atom in the dictionary, all of the parameters used to simulate the corresponding atom can be simultaneously extracted. Furthermore, since the form of the signal evolution used for pattern matching is known *a priori*, MRF is less sensitive to measurement errors, facilitating accelerated acquisitions through the undersampling of the measurement space (k-space). It should be noted that, so far, all matching is done for one-dimensional temporal signals only.

The notion of reconstructing signals from undersampled measurements comes from the theory of compressed sensing (CS) [5]. CS has been successfully applied to accelerate parameter mapping [4] and recently Davies et al. [3] demonstrated a CS strategy for MRF that does not rely on pattern matching for error suppression and has exact recovery guarantees, resulting in increased performance for shorter pulse sequences. The authors further extend their CS model to exploit global spatial structure by enforcing sparsity in the wavelet domain of the estimated density maps, slightly improving the performance of their approach.

Spatial information can also be incorporated locally by using image patches. Patch-based dictionaries have the advantage of being able to efficiently represent complex local structure in a variety of image processing tasks. Furthermore, the use of overlapping patches allows for averaging, resulting in the removal of both noise and incoherent artefacts caused by undersampling. Patch-based dictionaries have been previously used for the task of MR image reconstruction [7], where the sparsifying dictionary was learnt directly from the measured data, resulting in accurate reconstructions for up to six fold undersampling.

In this work, we propose to use a dictionary with both temporal and local spatial information for parametric map estimation. We create a training set by using the Bloch equations to simulate the temporal signal response over a predefined spatial distribution obtained from anatomical images and train a spatiotemporal dictionary by clustering similar patches. The trained dictionary is incorporated into a patch-based iterative projection algorithm to estimate T1 and T2 parametric maps. We see two main benefits of our approach:

1. Incorporating spatial data increases the atom length, i.e. the amount of descriptive information available per voxel, requiring less temporal points for an accurate reconstruction.
2. Training improves the conditioning of the dictionary by creating atoms distinct to each other, leading to a better signal matching.

The rest of this paper is structured as follows. In Sect. 2 we describe the method, in particular the proposed patch-based algorithm for MRF. Section 3

depicts the experiments and demonstrates the application of recovering parametric maps from undersampled data, and in Sect. 4 we offer conclusions.

2 Methods

The goal of MRF is to obtain parametric maps $\boldsymbol{\theta} \in \mathbb{R}^{N\times Q}$ from a sequence of undersampled measurements $\mathbf{Y} \in \mathbb{C}^{M\times T}$, where Q is the number of tissue relaxation parameters (T1 and T2), T is the sequence length, every map $\boldsymbol{\theta}_q \in \mathbb{R}^N$ has a total of N voxels, every measurement $\mathbf{y}_t \in \mathbb{C}^M$ is sampled M times, and $M \ll N$. This is achieved in three steps: image reconstruction, template matching, and parameter extraction.

Image reconstruction is the task of obtaining the image sequence $\mathbf{X} \in \mathbb{C}^{N\times T}$ from the measurements $\mathbf{Y}$. This is generally formulated as a inverse problem: $\mathbf{Y} = \mathbf{EX}$, where $\mathbf{E} \in \mathbb{C}^{M\times N}$ is the encoding operator. The reconstructed image is then matched to a precomputed dictionary $\mathbf{D} \in \mathbb{C}^{T\times L}$ of L atoms, to find the dictionary atom $\mathbf{d}_l \in \mathbb{C}^T$ that best describes it. This is done at every voxel location $\mathbf{x}_n \in \mathbb{C}^T$ by selecting the entry l_n that maximizes the modulus of the atom and the conjugate transpose of the signal:

$$\hat{l}_n = \underset{\substack{l \\ l=1,\dots,L}}{\arg\max} |\mathbf{x}_n^* \mathbf{d}_l| \tag{1}$$

where both, $\mathbf{d}_l$ and $\mathbf{x}_n$, were previously normalized to have unitary length. Finally, the T1 and T2 parameters used to construct the matching entry are assigned to the voxel n, creating $\boldsymbol{\theta}_n = \{T1_n, T2_n\}$. Thus, by repeating the matching over all voxels of the image, the parametric T1 and T2 maps are found.

Davies et al. [3] interpret the template matching as a projection of $\mathbf{x}_n$ onto the cone of the Bloch response manifold, and propose an iterative projection algorithm to accurately extract parametric maps. The algorithm, termed Bloch response recovery via iterated projection (BLIP), iteratively alternates between a gradient step, a projection step, and a shrinkage step to reconstruct the image sequence $\mathbf{X}$ and estimate the corresponding parameter maps $\boldsymbol{\theta}$.

2.1 Spatiotemporal Dictionary Design

Given a set of fully sampled 2D spatial parametric maps $\boldsymbol{\theta} \in \mathbb{R}^{N\times Q}$, where $N = N_i \times N_j$ and $Q = 2$, an image sequence $\mathbf{X} \in \mathbb{C}^{N\times T}$ of T temporal points can be created at each voxel using the Bloch equations to simulate the magnetization response of an inversion-recovery balanced steady state free-precession (IR-bSSFP) sequence with pseudorandomized acquisition parameters (see Fig. 1) [6]. $\mathbf{X}$ can be processed to create a spatiotemporal dictionary as follows.

Let $\mathbf{R}_n \in \mathbb{C}^{P\times N}$ be the operator that extracts 2D image patches of size $P = P_i \times P_j$, so that the spatiotemporal image patch $\tilde{\mathbf{x}}_n \in \mathbb{C}^{P\times T}$ at a given spatial location n is given by

$$\tilde{\mathbf{x}}_n = \mathbf{R}_n \mathbf{X}. \tag{2}$$

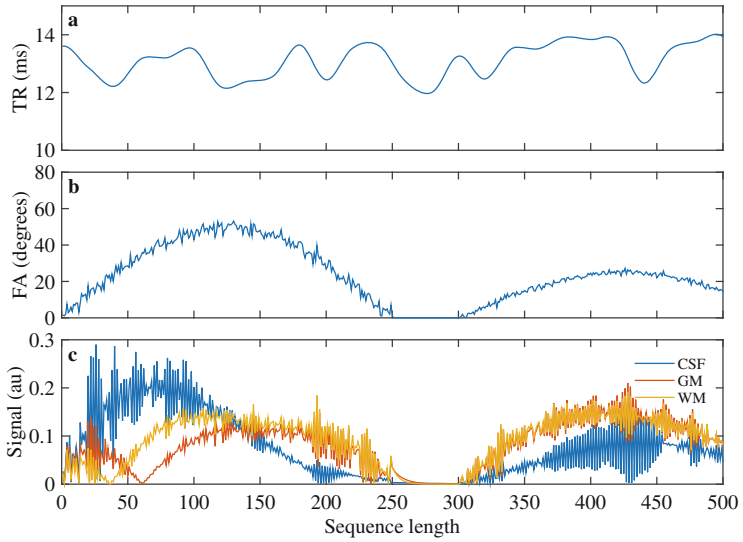

Fig. 1. Pseudorandom acquisition sequence and the corresponding signal response. **a**, TR values following a Perlin noise pattern. **b**, Flip angle series of repeating sinusoidal curves and added random values. **c**, Signal evolution for different tissue classes: white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF).

It is then possible to create the patch-based image matrix $\tilde{\mathbf{X}} \in \mathbb{C}^{PT \times N}$ by concatenating the vector representation of every spatiotemporal patch of dimension $P_i \times P_j \times T$ for each spatial location in $\mathbf{X}$. Repeating the operation on $\boldsymbol{\theta}$ creates the patch-based multiparametric matrix $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{PQ \times N}$. The spatiotemporal dictionary $\tilde{\mathbf{D}} \in \mathbb{C}^{PT \times K}$ is then constructed by using k-means to cluster atoms in $\tilde{\mathbf{X}}$ with similar signal values into K clusters, averaging the corresponding T1 and T2 values in $\tilde{\boldsymbol{\theta}}$ to create the clustered patch-based matrix $\boldsymbol{\Theta} \in \mathbb{C}^{PQ \times K}$, and simulating the signal evolution for each cluster. A new simulation of the signal evolution ensures that the atoms in $\tilde{\mathbf{D}}$ correspond exactly to the entries in $\boldsymbol{\Theta}$.

2.2 Patch-Based BLIP Reconstruction (P-BLIP)

The BLIP algorithm [3] reconstructs the image sequence $\mathbf{X}$ in an iterative fashion. Given an image sequence $\mathbf{X}^{(i)}$ at iteration i, the reconstructed sequence $\mathbf{X}^{(i+1)}$ in the next iteration is determined by

$$\mathbf{X}^{(i+1)} = \mathcal{P}_{\mathcal{A}}(\mathbf{X}^{(i)} + \mu \mathbf{E}^{H}(\mathbf{Y} - \mathbf{E}\mathbf{X}^{(i)})), \tag{3}$$

where $\mathcal{P}_{\mathcal{A}}$ represents the projection onto the signal model $\mathcal{A}$, $\mathbf{E}^{H}$ is the Hermitian adjoint of the encoding operator, and μ equals the step size. P-BLIP builds on this algorithm, incorporating the patch extraction operator in (2) and an update step to make (3) applicable to a spatiotemporal signal model.

At every iteration the updated sequence $\mathbf{X}$ is transformed into the patch-based matrix $\tilde{\mathbf{X}}$ by (2). $\tilde{\mathbf{X}}$ is related to the trained dictionary $\tilde{\mathbf{D}}$ by

$$\tilde{\mathbf{X}} = \tilde{\mathbf{D}}\mathbf{W}, \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{K \times N}$ represents the weights. Equation 4 can be readily solved using greedy algorithms that find sparse solutions to linear systems of equations by adding a sparsity constraint to the ℓ_0-norm of each column vector $\mathbf{w}_n$:

$$\hat{\mathbf{W}} = \arg\min_{\mathbf{W}} \|\tilde{\mathbf{X}} - \tilde{\mathbf{D}}\mathbf{W}\|_2^2, \text{ s.t. } \|\mathbf{w}_n\|_0 \leq \gamma, \ n = 1, ..., N. \tag{5}$$

We set the sparsity constraint to $\gamma = 1$, equivalent to finding one dictionary atom, as done in the template matching used in [3,6].

After estimating the weights, the patch-based image matrix is projected onto the dictionary by $\hat{\mathbf{X}} = \tilde{\mathbf{D}}\hat{\mathbf{W}}$. At this point, each voxel is overrepresented a total of P times, requiring an update step to return to the original image sequence $\mathbf{X}$. This update is achieved by averaging the P temporal signals that contribute to a given voxel location. Finally, the parametric maps $\boldsymbol{\theta}$ are estimated by applying the weights and patch-wise updates on $\boldsymbol{\Theta}$.

3 Experiments and Results

Image Data. Experiments were performed using twenty digital brain phantoms from BrainWeb [2]. Of these, ten were used to train the spatiotemporal dictionary and ten to test the performance of three different reconstruction algorithms: the original MRF reconstruction [6], BLIP [3], and the proposed P-BLIP. Experiments were designed to evaluate the performance of each algorithm as a function of sequence length and acceleration factors, and, for the case of P-BLIP, also as a function of spatial patch size. Ground truth datasets were generated by selecting a slice of crisp datasets labeled with different tissue classes, and resampling them to a matrix size of 256×256 to accelerate computations. Quantitative maps were then obtained by replacing the tissue labels with their corresponding T1 and T2 values. The values for the three main tissue types grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) were equaled to those reported in [6], while the values for the rest of the classes (fat, bone, muscle, vessels, dura matter, and connective tissue) were obtained directly from [1].

Modeling the Signal Evolution. At every voxel, the ground truth quantitative maps served as a basis to simulate the temporal evolution of the signal based on the IR-bSSFP pulse sequence with acquisition parameters displayed in Fig. 1, where the TRs follow a Perlin noise pattern, FAs are a series of repeating sinusoidal curves with added random values, and the radio frequency phase alternates between 0° and 180° on consecutive pulses. Off-resonance frequencies were not taken into account. This pulse sequence was combined with all possible combinations of a given range of T1 and T2 values to create a temporal dictionary used in both MRF and BLIP. The selected range was reported in [3], where T1 spans from 100 ms to 6000 ms and T2 from 20 ms to 1000 ms, both sampled at varying step sizes. Additionally, the dictionary included the exact T1 and T2 combinations corresponding to the different tissue classes.

Spatiotemporal Dictionary. To train the spatiotemporal dictionary used in P-BLIP, a region of interest that accounted for the entire head area was defined.

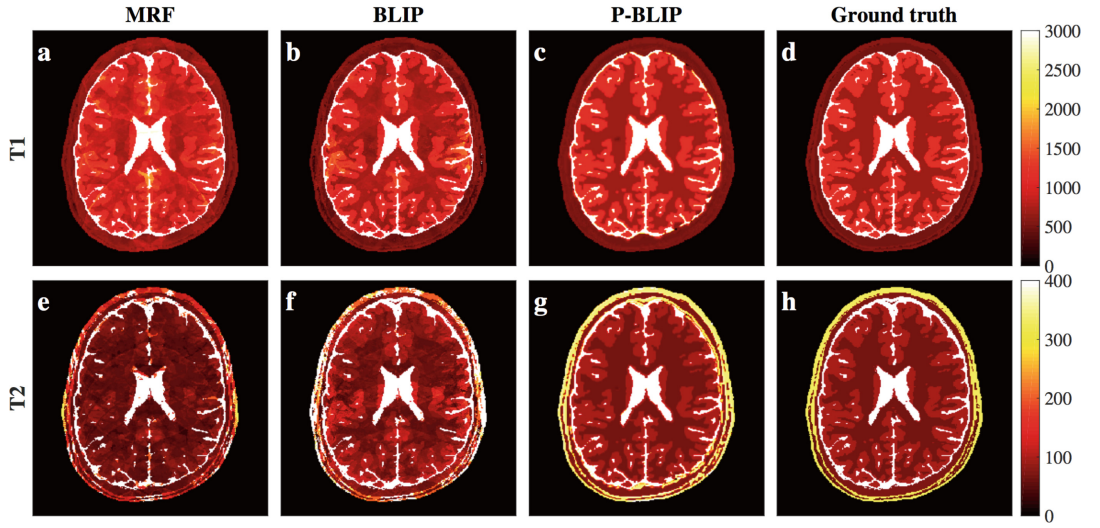

Fig. 2. Exemplary reconstruction results of one dataset with $T = 200$, $R = 10$, and $P = 3 \times 3$. The upper row shows T1 maps for all algorithms and the ground truth; and the bottom row the corresponding T2 maps. Most visible in T2 maps, subsampling artefacts can be effectively removed with P-BLIP.

The space covered by this region of interest was randomly and equally subsampled and each of the subsampled sets was assigned to a training subject. The selected parametric maps of each subject were then used as an input to train the dictionary as described in Sect. 2.1 with a total of $K = 200$ clusters.

Subsampling Strategy. We use a random EPI subsampling strategy for all experiments: the k-space is fully sampled in the read direction (k_x) and uniformly undersampled in the phase encoding direction (k_y) by an acceleration factor R. The sampling pattern is shifted by a random a number of k_y lines at every shot of the sequence.

Experimental Setup. An initial experiment was performed with spatiotemporal patches of size $3 \times 3 \times 200$ and an acceleration factor $R = 10$ to visually evaluate the reconstructed maps (see Fig. 2). Subsequently, three experiments assessed the reconstruction performance with respect to sequence length, acceleration factor and spatial patch size. The first experiment varied sequence lengths from 100 to 500 in step sizes of 100, the second experiment used acceleration factors of $R = \{2, 5, 10, 15, 20\}$, and the final experiment used spatial patches of sizes $P = \{1\times1, 3\times3, 5\times5, 7\times7\}$. The reconstruction error of the first two experiments was calculated using the signal-to-error ratio (SER) in decibels (dB), defined as $20 \log_{10} \frac{\|\mathbf{x}\|_2}{\|\mathbf{x}-\hat{\mathbf{x}}\|_2}$; and the third experiment with the SSIM values [8].

Results. Figure 2 displays the reconstructed parametric maps of an exemplary dataset. The MRF estimates show the characteristic ghosting artefacts caused by sub-Nyquist sampling. BLIP removes most of these artefacts from the T1 estimation, though they are still visible in the T2 maps. P-BLIP effectively removes these artefacts from both maps, resulting in reconstructions very close to the ground truth. These visual observations can be confirmed with quantitative

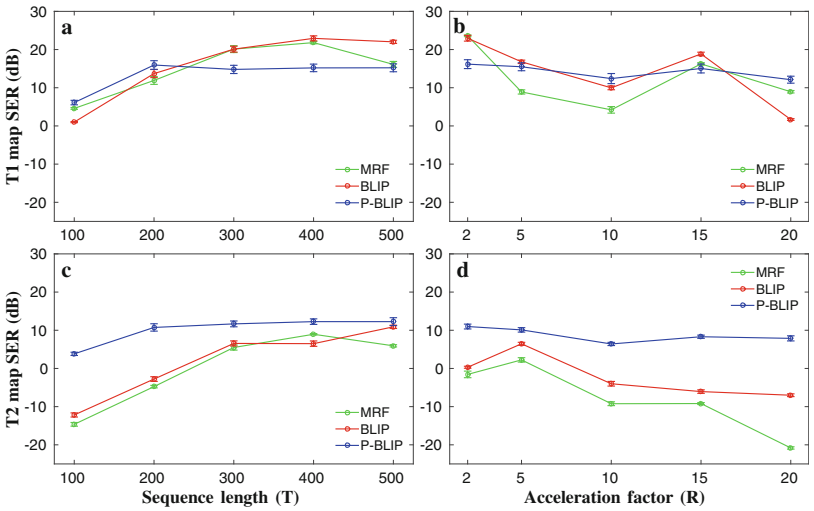

Fig. 3. **a**,**c**, Performance as a function of sequence length with $R = 10$, $P = 3 \times 3$; and **b**,**d**, as a function of acceleration factor with $T = 200$, $P = 3 \times 3$. P-BLIP is best in estimating T2 maps and shows better results for shorter sequences and higher acceleration for T1 maps.

Table 1. Average SSIM values for T1 and T2 map estimation with respect to different spatial patch sizes, $T = 200$ and $R = 10$.

Method	Baseline		Proposed: P-BLIP			
	MRF	BLIP	1×1	3×3	5×5	7×7
T1	0.761	0.814	0.848	**0.852**	0.691	0.625
T2	0.616	0.591	0.769	**0.857**	0.667	0.601

results. Figure 3c and d show how P-BLIP achieves better T2 estimates independently of the sequence length or acceleration factor. On the other hand, T1 maps for P-BLIP remain relatively constant for sequence lengths larger than 100 (Fig. 3a) and all acceleration factors (Fig. 3b), whilst the performance of MRF and BLIP increases with the sequence length and lower acceleration factors. The reason for these results is twofold. First, the IR-bSSFP sequence is mostly T1-weighted, favoring a better T1 matching over T2 matching for all methods. Second, a trained dictionary containing a longer sequence, but fixed K, is less flexible, and if the trained dictionary does not exactly contain the ground truth values, the quantitative error will be higher.

Table 1 indicates the performance of P-BLIP for different patch sizes in comparison to the performance of MRF and BLIP. A spatial patch size of $P = 1 \times 1$ implies that the training dataset was created from voxel-wise temporal evolutions and that the trained dictionary is a clustered version of the temporal dictionary. It can be seen that clustering a temporal dictionary alone improves the reconstruction with respect to MRF and BLIP, and that the spatiotemporal dictionary further improves these results for $P = 3 \times 3$. At larger spatial patch sizes the results begin to decline, indicating that the cluster size of $K = 200$ is not enough to capture the entire spatial variability of the parametric maps.

4 Conclusions

This work presents a novel patch-based reconstruction scheme for MRF consistent with the theory of CS. It is based on a spatiotemporal signal model and relies on the training of the corresponding dictionary from a set of examples. This patch-based scheme shows improved performance for shorter pulse sequences and at higher acceleration factors, leading to an increased efficiency of parameter mapping with MRF.

An important discussion point of our approach is the size of the dictionary in terms of space, time, and atoms. Larger spatial patches allow, in theory, for the acquisition of less temporal points, but the amount of atoms in the dictionary should in turn be large enough to account for large spatial variability. We have seen from our results that a dictionary size of $K = 200$ is not enough for spatial patch sizes larger than 3×3 for structures in the brain. A potencial solution to this shortcoming might be to make K dependant on the atom length or arbitrarily large at the cost of computational complexity. This point is currently under investigation and future work will focus on extending the method to incorporate 3D spatial patches and applying it to real datasets.

References

1. Aubert-Broche, B., Evans, A.C., Collins, L.: A new improved version of the realistic digital brain phantom. NeuroImage **32**, 138–145 (2006)
2. Aubert-Broche, B., Griffin, M., Pike, G.B., Evans, A.C., Collins, D.L.: Twenty new digital brain phantoms for creation of validation image data bases. IEEE Trans. Med. Imaging **25**(11), 1410–1416 (2006)
3. Davies, M., Puy, G., Vandergheynst, P., Wiaux, Y.: A compressed sensing framework for magnetic resonance fingerprinting. SIAM J. Imaging Sci. **7**(4), 2623–2656 (2014)
4. Doneva, M., Börnert, P., Eggers, H., Stehning, C., Sénégas, J., Mertins, A.: Compressed sensing reconstruction for magnetic resonance parameter mapping. Magn. Reson. Med. **64**, 1114–1120 (2010)
5. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theor. **52**, 1289–1306 (2006)
6. Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A.: Magnetic resonance fingerprinting. Nature **495**, 187–192 (2013)
7. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging **30**(5), 1028–1041 (2011)
8. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Proc. **13**, 600–612 (2004)

5.1.2 *Simultaneous Parameter Mapping, Modality Synthesis, and Anatomical Labeling of the Brain with MR Fingerprinting*

Peer-reviewed Conference Paper

Authors: PA. Gómez, **M. Molina-Romero**, C. Ulas, G. Buonincontri, JI. Sperl, DK. Jones, MI. Menzel, BH. Menze

In: *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention* LNCS 9902 (2016), pp. 579 – 586 [26]

Abstract: Magnetic Resonance Fingerprinting (MRF) quantifies various properties simultaneously by matching measurements to a dictionary of precomputed signals. We propose to extend the MRF framework by using a database to introduce additional parameters and spatial characteristics to the dictionary. We show that, with an adequate matching technique which includes an update of selected fingerprints in parameter space, it is possible to reconstruct parametric maps, synthesize modalities, and label tissue types at the same time directly from an MRF acquisition. We compare (1) relaxation maps from a spatiotemporal dictionary against a temporal MRF dictionary, (2) synthetic diffusion metrics versus those obtained with a standard diffusion acquisition, and (3) anatomical labels generated from MRF signals to an established segmentation method, demonstrating the potential of using MRF for multiparametric brain mapping.

Contribution of thesis author: Discussion of algorithmic implementation and experimental design, manuscript revision and editing.

Simultaneous Parameter Mapping, Modality Synthesis, and Anatomical Labeling of the Brain with MR Fingerprinting

Pedro A. Gómez[1,2,4( )], Miguel Molina-Romero[1,2,4], Cagdas Ulas[1,2], Guido Bounincontri[3], Jonathan I. Sperl[2], Derek K. Jones[4], Marion I. Menzel[2], and Bjoern H. Menze[1]

[1] Computer Science, Technische Universität München, Munich, Germany
pedro.gomez@tum.de
[2] GE Global Research, Munich, Germany
[3] INFN Pisa, Pisa, Italy
[4] CUBRIC School of Psychology, Cardiff University, Cardiff, UK

Abstract. Magnetic resonance fingerprinting (MRF) quantifies various properties simultaneously by matching measurements to a dictionary of precomputed signals. We propose to extend the MRF framework by using a database to introduce additional parameters and spatial characteristics to the dictionary. We show that, with an adequate matching technique which includes an update of selected fingerprints in parameter space, it is possible to reconstruct parametric maps, synthesize modalities, and label tissue types at the same time directly from an MRF acquisition. We compare (1) relaxation maps from a spatiotemporal dictionary against a temporal MRF dictionary, (2) synthetic diffusion metrics versus those obtained with a standard diffusion acquisition, and (3) anatomical labels generated from MRF signals to an established segmentation method, demonstrating the potential of using MRF for multiparametric brain mapping.

1 Introduction

Magnetic resonance fingerprinting (MRF) is an emerging technique for the simultaneous quantification of multiple tissue properties [7]. It offers absolute measurements of the T1 and T2 relaxation parameters (opposed to traditional weighted imaging) with an accelerated acquisition, leading to efficient parameter mapping. MRF is based on matching measurements to a dictionary of precomputed signals that have been generated for different parameters. Generally, the number of atoms in the dictionary is dictated by the amount of parameters, and the range and density of their sampling. As an alternative to continuous sampling of the parameter space, one could use measured training examples to learn the dictionary, reducing the number of atoms to only feasible parameter combinations [2]. In this work, we propose to use a database of multi-parametric datasets to create the dictionary, presenting two new features of MRF that can be achieved simultaneously with relaxation mapping: modality synthesis and automatic labeling of the corresponding tissue.

S. Ourselin et al. (Eds.): MICCAI 2016, Part III, LNCS 9902, pp. 579–586, 2016.
DOI: 10.1007/978-3-319-46726-9_67

In this extended application of MRF towards image synthesis and segmentation, we follow a direction that has recently gained attention in the medical image processing literature [1,3,5,6,9,10]. The working principle behind these methods is similar: given a source image and a multi-contrast database of training subjects, it is possible to generate the missing contrast (or label) of the source by finding similarities within the database and transferring them to create a new image. The search and synthesis strategy can take several forms: it could be iterative to incorporate more information [10]; can be optimized for multiple scales and features [1]; may include a linear combination of multiple image patches [9]; or be configured to learn a nonlinear transform from the target to the source [5]. There have been several applications of synthetic contrasts, including inter-modality image registration, super-resolution, and abnormality detection [3,5,6,9,10]. Furthermore, in addition to the creation of scalar maps in image synthesis, similar techniques can be used for mapping discrete annotations; for example, in the segmentation of brain structures [1].

Inspired by these ideas, we present a method for synthesizing modalities and generating labels from magnetic resonance fingerprints. It relies on the creation of a spatiotemporal dictionary [2] and its mapping to different parameters. Specifically, in addition to the physics-based mapping of MRF signals to the T1 and T2 relaxation parameters, we train empirical functions for a mapping of the signals to diffusion metrics and tissue probabilities. We show that we can achieve higher efficiency relaxation mapping, and demonstrate how the use of a spatiotemporal context improves the accuracy of synthetic mapping and labeling.

We see three main contributions to our work. (1) We present a framework for creating a spatiotemporal MRF dictionary from a multi-parametric database (Sect. 2.1). (2) We generalize fingerprint matching and incorporate a data-driven update to account for correlations in parameter space, allowing for the simultaneous estimation of M different parameters from any fingerprinting sequence (Sect. 2.2). (3) Depending on the nature of the m-th parameter, we call it a mapping, synthesis, or labeling, and show results for all three applications (Sect. 3.1). This is the first attempt - to the best of our knowledge - to simultaneously map parameters, synthesize diffusion metrics, and estimate anatomical labels from MR fingerprints.

2 Methods

Let $\mathcal{Q} = \{Q_s\}_{s=1}^{S}$ represent a database of spatially aligned parametric maps for S subjects, where each subject $Q_s \in \mathbb{R}^{N \times M}$ contains a total of $N = N_i \times N_j \times N_k$ voxels and M maps. Every map represents an individual property, and can originate from a different acquisition or modality, or even be categorical. Our database includes the quantitative relaxation parameters T1 and T2; a non-diffusion weighted image (S0); the diffusion metrics mean diffusivity (MD), radial diffusivity (RD), and fractional anisotropy (FA); and probability maps for three tissue classes: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Thus, for every subject $Q_s = \{\mathrm{T1}, \mathrm{T2}, \mathrm{S0}, \mathrm{MD}, \mathrm{RD}, \mathrm{FA}, \mathrm{GM}, \mathrm{WM}, \mathrm{CSF}\}$. We use this database to create a spatiotemporal MRF dictionary as follows.

2.1 Building a Spatiotemporal MRF Dictionary

With the relaxation parameters T1 and T2 and knowledge of the sequence variables, it is possible to follow the extended phase graph (EPG) formalism to simulate the signal evolution of a *fast imaging with steady state precession MRF* (FISP-MRF) pulse sequence [4]. In EPG the effects of a sequence on a spin system are represented by operators related to radio-frequency pulses, relaxation, and dephasing due to gradient moments. Therefore, for every voxel in all subjects, application of the EPG operators leads to a dictionary $D \in \mathbb{C}^{NS \times T}$ with a total of T temporal points (see Fig. 1).

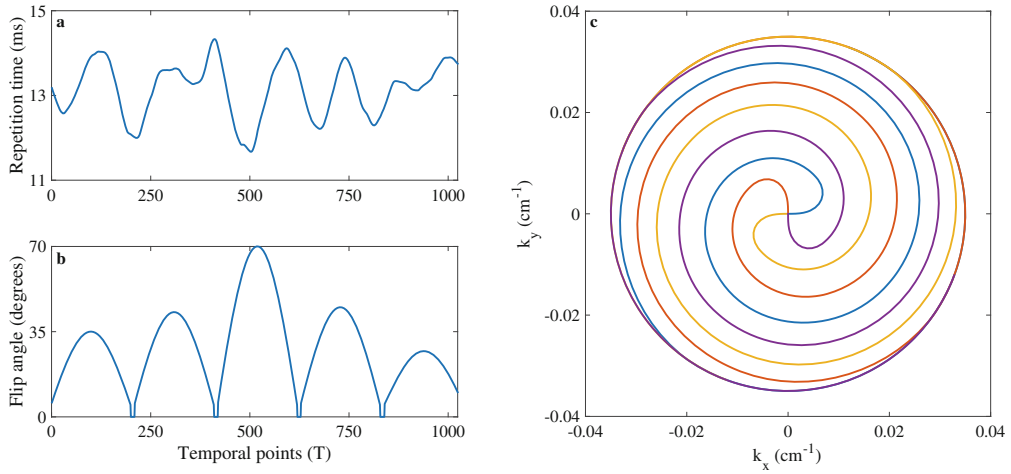

Fig. 1. FISP-MRF acquisition sequence. **a**, Repetition times following a Perlin noise pattern. **b**, Flip angles of repeating sinusoidal curves. **c**, k-space trajectory of four different spiral interleaves, 32 interleaves are required for full k-space coverage.

We further process the dictionary to incorporate spatial information by expanding each voxel with its 3D spatial neighborhood of dimension $P = P_i \times P_j \times P_k$ and compressing the temporal dimension into its first V singular vectors [8]. This results in a compressed spatiotemporal dictionary $\tilde{D} \in \mathbb{C}^{NS \times PV}$. Finally, we define a search window $W_n = W_i \times W_j \times W_k$ around every voxel n, limiting the dictionary per voxel to $\tilde{D}_n \in \mathbb{C}^{W_n S \times PV}$. The choice for a local search window has a two-fold motivation: it reduces the number of computations by decreasing the search space and it increases spatial coherence for dictionary matching [10].

Applying subject concatenation, patch extraction, and search window reduction on the database $\mathcal{Q}$ leads to a voxel-wise spatio-parametric matrix $\tilde{R}_n \in \mathbb{R}^{W_n S \times PM}$. For simplicity, we will use D and R instead of $\tilde{D}_n$ and $\tilde{R}_n$, where every dictionary entry $d_c \in \mathbb{C}^{PV}$ has its corresponding matrix entry $r_c \in \mathbb{R}^{PM}$.

2.2 Dictionary Matching and Parameter Estimation

MRF aims to simultaneously estimate several parametric maps from undersampled data. This is achieved by reconstructing an image series and matching it

to the dictionary. We reconstruct V singular images [8] and extract 3D patches from them to create the patch-based matrix $X \in \mathbb{C}^{N \times PV}$. At every voxel x_n, we find the set $\mathcal{M}_n$ of the C highest correlated dictionary entries $d_c,\ c = 1, .., C$, by:

$$\mathcal{M}_n = \{d_c \in D : \rho(x_n, d_c) > \tau_C\} \tag{1}$$

with the threshold value τ_C such that $|\mathcal{M}| = C$ and

$$\rho(x, d) = \frac{\langle x, d \rangle}{\|x\|_2 \|d\|_2}. \tag{2}$$

Making use of the selected entries d_c and the corresponding parametric vectors r_c, an estimated value $\tilde{q}_{n,m}$ at voxel location n in map m is determined by the weighted average of the correlation between every entry d_c and the signals x_p within Ω_n, the spatial neighborhood of n:

$$\tilde{q}_{n,m} = \frac{\sum_{p \in \Omega_n} \sum_c \rho(x_p, d_c) r_{c,pm}}{P \sum_c \rho(x_p, d_c)}, \tag{3}$$

where $r_{c,pm}$ indexes the quantitative value of voxel p centered around atom c in map m. Repeating this procedure for every voxel creates an estimate $\tilde{Q}$ of all of the parametric maps, including synthetic modalities and anatomical labels.

Data-Driven Updates. Ye et al. [10] proposed the use of intermediate results to increase spatial consistency of the synthetic maps. We take a similar approach, and define a similarity function relating image space and parameter space:

$$f(x, d, r, q, \alpha) = (1 - \alpha)\rho(x, d) + \alpha\rho(q, r) \tag{4}$$

where α controls the contributions of the correlations in image and parameter space. The selected atoms are now determined by

$$\mathcal{M}_n = \{d_c \in D, r_c \in R : f(x_n, d_c, \tilde{q}_n, r_c, \alpha) > \tau_C\}. \tag{5}$$

In the first iteration $\alpha = 0$ as we have no information on the map $\hat{Q}$ for our subject. In a second iteration we increase α, adding weight to the similarities in parameter space and compute Eq. 5 again to find a new set of dictionary atoms. The final version of the maps is given by a modified version of Eq. 3:

$$\hat{q}_{n,m} = \frac{\sum_{p \in \Omega_n} \sum_c f(x_p, d_c, \tilde{q}_n, r_c, \alpha) r_{c,pm}}{P \sum_c f(x_p, d_c, \tilde{q}_n, r_c, \alpha)}. \tag{6}$$

This procedure is essentially a 3D patch-match over a V-dimensional image space and M-dimensional parameter space, where the matching patches are combined by their weighted correlation to create a final result.

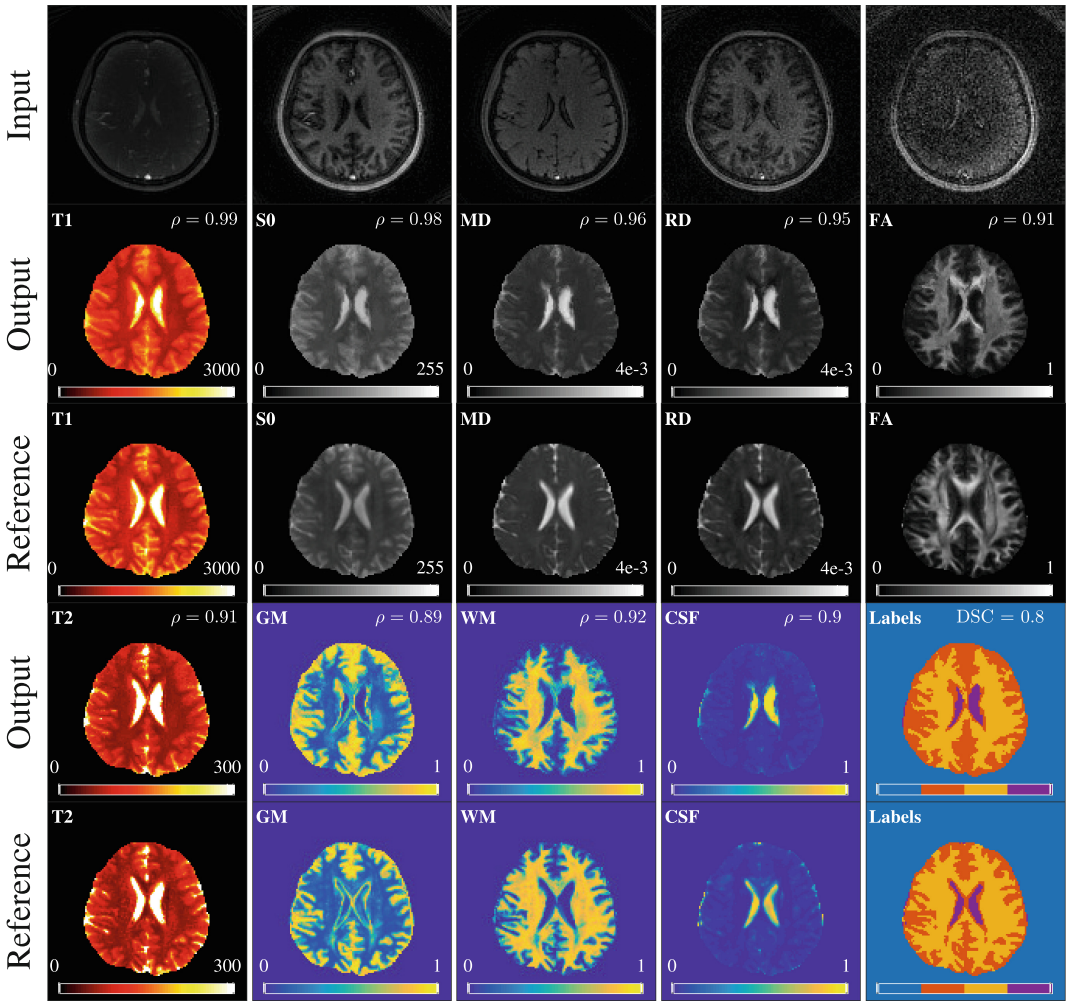

Fig. 2. Exemplary results of one test subject with $P = 3 \times 3 \times 3$. The upper row displays the first five singular images; while the second and fourth row show the output for different parametric maps and the correlation to the reference image, displayed in the third and fifth row, respectively. Additionally, the last column in rows four and five shows labels obtained from selecting the tissue class with highest probability and the dice similarity coefficient (DSC) from the output labels to the reference. The bar underneath represents, from left to right, background, GM, WM, and CSF; and the DSC was computed from the GM, WM, and CSF labels. T1 and T2 scale is displayed in ms; S0 is qualitatively scaled to 255 arbitrary units; MD and RD are in mm^2/s; FA, GM, WM, and CSF are fractional values between zero and one.

2.3 Data Acquisition and Pre-processing

We acquired data from six volunteers with a FISP-MRF pulse sequence [4] on a 3T GE HDx MRI system (GE Medical Systems, Milwaukee, WI) using an eight channel receive only head RF coil. After an initial inversion, a train of $T = 1024$ radio-frequency pulses with varying flip angles and repetition times following a Perlin noise pattern [4] was applied (see Fig. 1). We use one interleave of a zero-moment compensated variable density spiral trajectory per repetition, requiring

32 interleaves to sample a 22×22 cm field of view (FOV) with 1.7 mm isotropic resolution. We acquired 10 slices per subject with a scan time of 13.47 seconds per slice, performed a gridding reconstruction onto a 128×128 Cartesian grid, projected the data into SVD space, and truncated it to generate $V = 10$ singular images. The choice of $V = 10$ was motivated by the energy ratio, as this was the lowest rank approximation which still yielded an energy ratio of 1.0 [8]. The singular images were matched to a MRF dictionary comprising of T1 values ranging from 100 to 6,000 ms; and T2 values ranging from 20 ms to 3,000 ms.

In addition, we scanned each volunteer with a diffusion weighted imaging (DWI) protocol comprising of 30 directions in one shell with $b = 1000$ s/mm^2. The FOV, resolution, and acquired slices were the same as with MRF-FISP, resulting in a 15 min scan. We applied FSL processing to correct for spatial distortions derived from EPI readouts, skull strip, estimate the diffusion tensor and its derived metrics MD, RD, and FA; and used the non-diffusion weighted image S0 to compute probability maps of three tissue types (GM, WM, CSF) using [11]. Finally, we applied registration across all subjects to create the database.

3 Experiments and Results

For every subject, we performed a leave-one-out cross validation, wherein the dictionary was constructed from five subjects and the remaining subject was used as a test case. Following the procedure described in Sect. 2.2, we created a database of nine parametric maps (T1,T2,S0,MD,RD,FA,GM,WM,CSF) and compared the estimated metrics to the reference by their correlation.

We explored the influence of the window size W_n, the number of entries C, and the α on the estimated maps. We found correlations increased with diminishing returns as W_n increased, while adding more entries yielded smoother maps. Correlations were higher after a second iteration of data-driven updates with $\alpha > 0$, irrespective of the value of α. Nonetheless, variations of these parameters didn't have a significant effect on the overall results. To investigate the impact of using spatial information, we repeated the experiment for spatial patch sizes of $P = 1 \times 1 \times 1$, $3 \times 3 \times 3$, and $5 \times 5 \times 5$. For these experiments we used $W_n = 11 \times 11 \times 11$, $C = 5$, $\alpha = 0.5$, and two iterations.

3.1 Results

The reference T1 and T2 maps were estimated from a FISP-MRF sequence with a temporal dictionary, while we used a spatiotemporal dictionary with varying spatial patches. Estimated T1 and T2 maps were consistent with the reference, with increasing spatial smoothness for larger spatial patches. This also lead to a decrease in correlation to the reference, most notably in T2 estimation (see Fig. 3a–b), which could also be attributed noisier T2 estimates. In future experiments we will rely on standard relaxation mapping for reference comparison.

The synthetic S0 and diffusion metrics MD, RD, and FA show spatial coherence, achieving correlation values over 0.90 with respect to a standard DWI

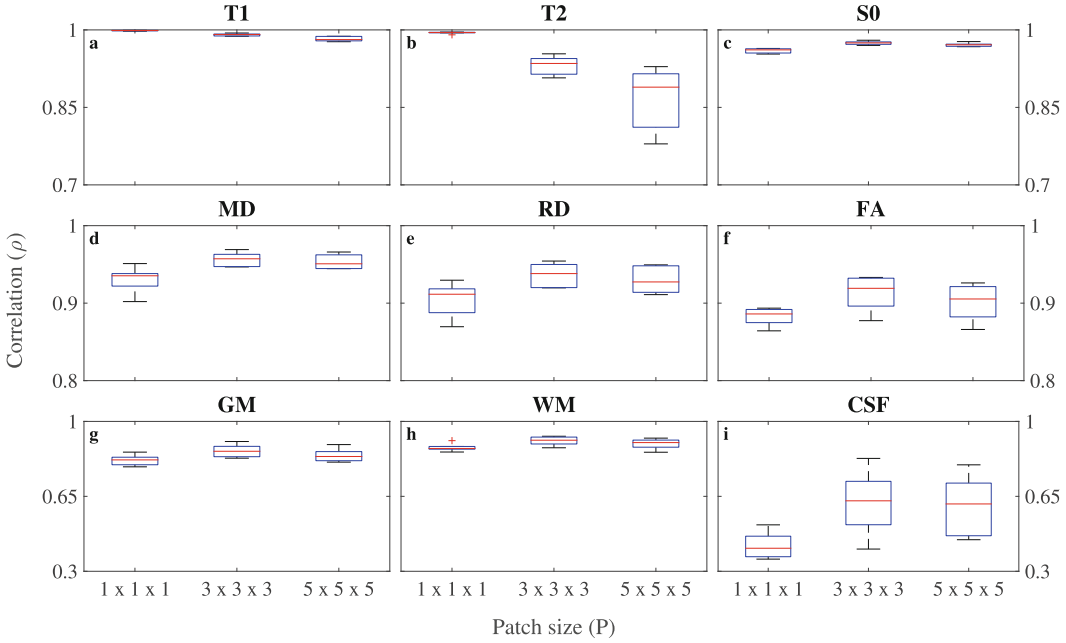

Fig. 3. Correlation as a function of spatial patches for all subjects. **a**–**b**, T1 and T2 parameter mapping. **c**–**f**, Synthesis of S0 and diffusion metrics. **g**–**i**, Tissue labeling.

acquisition (Fig. 2). Similar to [10], we found that FA maps were generally the least correlated to the reference. This is due to the fact that diffusion encoding in DWI acts as a proxy for underlying tissue anisotropy, whereas the measured fingerprints are not diffusion sensitive, failing to exactly recover directionality present in FA. In fact, the higher the directionality encoded in a given modality, the lower the correlation to the reference ($\overline{\rho_{S0}} > \overline{\rho_{MD}} > \overline{\rho_{RD}} > \overline{\rho_{FA}}$). Furthermore, for all cases in modality synthesis, incorporating spatial information generated increased consistency and higher correlated results (Fig. 3c–f).

Figure 2 shows the visual similarity between tissue probability maps obtained directly as an output from matching and those computed with [11] and the labels obtained by selecting the class with the highest probability. As with modality synthesis, anatomical labels improved when spatial information was taken into account (Fig. 3g–i). Particularly in CSF, incorporation of spatial information eliminated false positives, yielding better quality maps. On the other hand, thresholding of probability maps lead to an overestimation of GM labels, notably at tissue boundaries. Labeling at tissue boundaries could benefit from higher resolution scans and a multi-channel reference segmentation.

4 Discussion

This work proposes to replace a simulated temporal MRF dictionary with a spatiotemporal dictionary that can be learnt from data, increasing the efficiency of relaxation parameter mapping, and enabling the novel applications of modality synthesis and anatomical labeling. In terms of methodology, we borrow concepts

such as the search window and parameter space regularization from the image segmentation and synthesis literature [1,3,10], but change the input to a V-dimensional image space and the output to an M-dimensional parameter space, making it applicable to MRF. Moreover, our framework is valid for any MR sequence, provided signal evolutions can be computed from the training data.

Results indicate that it is possible to use MRF to simultaneously map T1 and T2 parameters, synthesize modalities, and classify tissues with high consistency with respect to established methods. While our method allows us to circumvent post-processing for diffusion metric estimation and tissue segmentation, it is important to note that changes in synthetic diffusion maps can only be propagated from the information available in the database. Therefore, creating the dictionary from pathology and exploring advanced learning techniques capable of capturing these changes is the subject of future work.

Acknowledgments. With the support of the Technische Universität München Institute for Advanced Study, funded by the German Excellence Initiative and the European Commission under Grant Agreement Number 605162.

References

1. Giraud, R., Ta, V.T., Papadakis, N., et al.: An optimized PatchMatch for multi-scale and multi-feature label fusion. NeuroImage **124**, 770–782 (2016)
2. Gómez, P.A., Ulas, C., Sperl, J.I., et al.: Learning a spatiotemporal dictionary for magnetic resonance fingerprinting with compressed sensing. In: MICCAI Patch-MI Workshop, vol. 9467, pp. 112–119 (2015)
3. Iglesias, J.E., Konukoglu, E., Zikic, D., et al.: Is synthesizing MRI contrast useful for inter-modality analysis? MICCAI **18**(9), 1199–1216 (2013)
4. Jiang, Y., Ma, D., Seiberlich, N., et al.: MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout. MRM **74**(6), 1621–1631 (2014)
5. Jog, A., Carass, A., Roy, S., et al.: MR image synthesis by contrast learning on neighborhood ensembles. Med. Image Anal. **24**(1), 63–76 (2015)
6. Konukoglu, E., van der Kouwe, A., Sabuncu, M.R., Fischl, B.: Example-based restoration of high-resolution magnetic resonance image acquisitions. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 131–138. Springer, Heidelberg (2013)
7. Ma, D., Gulani, V., Seiberlich, N., et al.: Magnetic resonance fingerprinting. Nature **495**, 187–192 (2013)
8. McGivney, D., Pierre, E., Ma, D., et al.: SVD compression for magnetic resonance fingerprinting in the time domain. IEEE TMI **0062**, 1–13 (2014)
9. Roy, S., Carass, A., Prince, J.L.: Magnetic resonance image example based contrast synthesis. IEEE TMI **32**(12), 2348–2363 (2013)
10. Ye, D.H., Zikic, D., Glocker, B., et al.: Modality propagation: coherent synthesis of subject-specifc scans with data-driven regularization. MICCAI **8149**, 606–613 (2013)
11. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE TMI **20**, 45–57 (2001)

5.2 RELEVANT PEER-REVIEWED ABSTRACTS

The abstracts contained in this section were presented at the ISMRM annual meeting. The work titled 3D Magnetic Resonance Fingerprinting with a Clustered Spatiotemporal Dictionary proposes an unsupervised learning technique to build a clustered dictionary for MRF, this method boosts efficiency and is robust to undersampling and shorter acquisitions. In this regards, Accelerated Parameter Mapping with Compressed Sensing: an Alternative to MR Fingerprinting take this idea further revisiting the requirements of MRF and removing dispensable elements to reach an optimized technique: QTI.

5.2.1 *3D Magnetic Resonance Fingerprinting with a Clustered Spatiotemporal Dictionary*

Peer-reviewed Conference Abstract

Authors: PA. Gómez, G. Buonincontri, **M. Molina-Romero**, C. Ulas, JI. Sperl, MI. Menzel, BH. Menze

In: *Proc Intl Soc Mag Reson Med* (2016) [26]

Abstract: We present a method for creating a spatiotemporal dictionary for Magnetic Resonance Fingerprinting (MRF). Our technique is based on the clustering of multi-parametric spatial kernels from training data and the posterior simulation of a temporal fingerprint for each voxel in every cluster. We show that the parametric maps estimated with a clustered dictionary agree with maps estimated with a full dictionary, and are also robust to undersampling and shorter sequences, leading to increased efficiency in parameter mapping with MRF.

Contribution of thesis author: Discussion of algorithmic implementation and experimental design, manuscript revision and editing.

3D Magnetic Resonance Fingerprinting with a Clustered Spatiotemporal Dictionary

Pedro A Gómez[1,2], Guido Bounincontri[3], Miguel Molina-Romero[1,2], Cagdas Ulas[1,2], Jonathan I Sperl[2], Marion I Menzel[2], Bjoern H Menze[1]

[1]Computer Science, Technische Universität München, Munich, Germany
[2]GE Global Research, Munich, Germany
[3]INFN Pisa, Pisa, Italy

Abstract. We present a method for creating a spatiotemporal dictionary for magnetic resonance fingerprinting (MRF). Our technique is based on the clustering of multi-parametric spatial kernels from training data and the posterior simulation of a temporal fingerprint for each voxel in every cluster. We show that the parametric maps estimated with a clustered dictionary agree with maps estimated with a full dictionary, and are also robust to undersampling and shorter sequences, leading to increased efficiency in parameter mapping with MRF.

1 Purpose

Magnetic resonance fingerprinting (MRF) allows for the simultaneous quantification of multiple tissue properties via the matching of acquired signals to a precomputed dictionary, created by sampling a wide range of the parameter space [4]. As the parameters of interest increase, so does the dictionary size, leading to long reconstruction times. One possibility for overcoming this limitation is to use a clustered dictionary with both spatial and temporal information [2]. This work aims at increasing MRF efficiency by using a clustered spatiotemporal dictionary and incorporating it into a MRF pipeline that includes B1 mapping and a view-sharing (VS) anti-aliasing strategy [1].

2 Methods

We tested our approach using 3D MRF data of a Lister-hooded adult rat brain adult acquired with a Bruker BioSpec 47/40 system (Bruker Inc., Ettlingen, Germany) [1]. The sequence was based on SSFP-MRF [3]with Cartesian sampling, $T = 1000$ shots, and 0.5 mm isotropic resolution. A dictionary $\mathbf{D} \in \mathbb{C}^{L \times T}$ was simulated using extended phase graphs with the following ranges: T1 from 100ms to 3,000ms in 20ms steps; T2 from 20ms to 100ms in 5ms steps and from 100 to 500ms in 10ms steps; and B1 as a flip angle factor from 50% to 150% in 1% steps, resulting in a dictionary of size 840522×1000. The acquired data was matched to the dictionary to create a reference dataset.

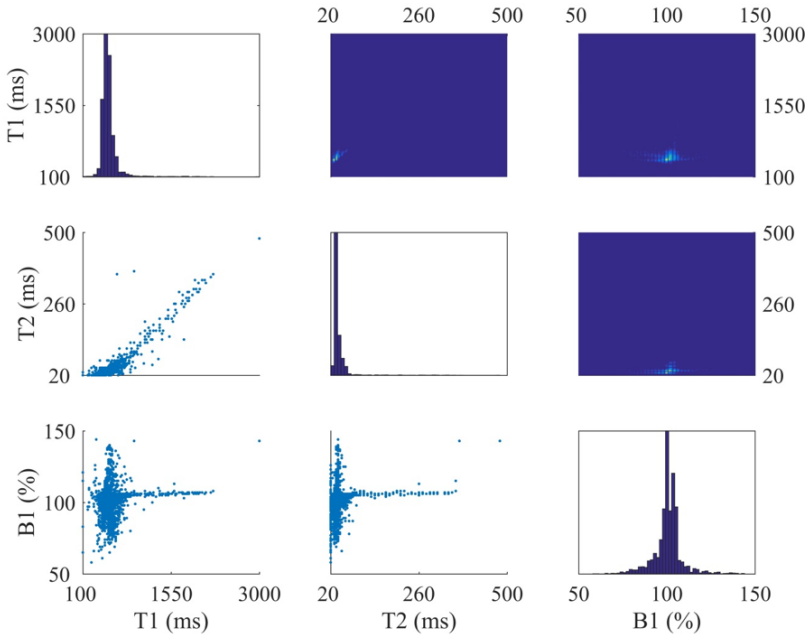

Fig. 1. Parameter distribution obtained from selected slices in the left hemisphere used as a training dataset. The upper triangle displays density plots, the diagonal histograms, and the lower triangle scatter plots. Note that parameters approach a Gaussian distribution and are densely scattered within a specific range.

Exploiting symmetry of the brain, the reference dataset was divided along the medial longitudinal fissure, separating the left and right hemisphere. The estimated parametric T1, T2 and B1 maps of the left hemisphere (see Fig. 1) were used to create spatiotemporal dictionaries of different sizes by first clustering multi-parametric (T1,T2,B1) spatial kernels using k-means and subsequently simulating the temporal signal of every voxel in each cluster. The right hemisphere of the reference dataset was then matched to dictionaries with spatial kernel sizes of $P = 1 \times 1 \times 1$ (clustered only), $P = 3 \times 3 \times 3$ and $P = 5 \times 5 \times 5$ (see Fig. 2).

We hypothesize that a dictionary that contains only feasible parameter combinations and spatial information should enable acceleration in both space and time. We test this by samplingless k-space points using a Gaussian mask in the phase encode directions with different acceleration factors (Figs. 3-4), and by reducing the sequence length (Fig. 4). Undersampled datasets were reconstructed with the original dictionary template matching (TM) [4]and with our VS approach, and compared to the reference dataset by their similarity index (SSIM) [5]. Furthermore, we study the amount of clusters required to accurately capture the entire spatio-parametric variability in our dataset by evaluating the mean

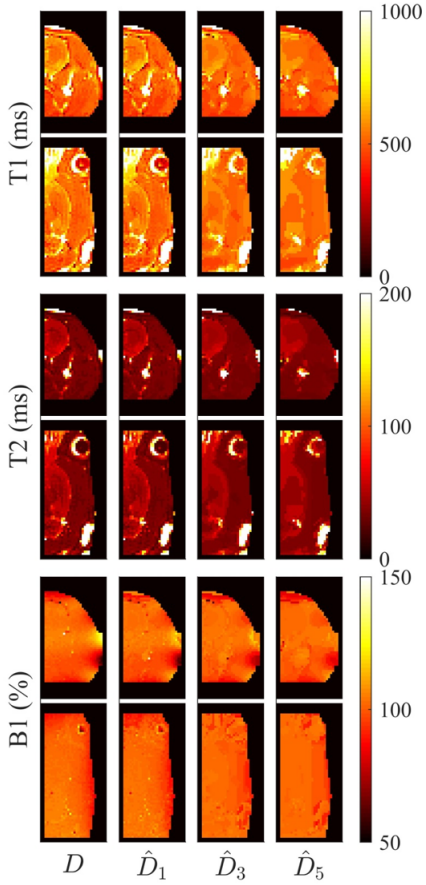

Fig. 2. Comparison of the estimated T1, T2 and B1 parametric maps from the fully sampled dataset with a temporal dictionary $\mathbf{D} \in \mathbb{C}^{L\times T}$ and three clustered dictionaries $\hat{\mathbf{D}}_{\sqrt[3]{P}} \in \mathbb{C}^{K\times TP}$ with $K = 300$, $T = 1000$, and $P = 1 \times 1 \times 1$, $3 \times 3 \times 3$, and $5 \times 5 \times 5$. Spatial smoothing obtained with and is achieved by averaging all contributing patches to a given voxel.

square error (MSE) of the training and testing data for different spatial kernels (Fig. 5).

3 Results

Figure 1 shows how the estimated parameters approximate a Gaussian distribution, and are scattered in a restricted range within the parameter space. Hence, using dictionaries trained from this distribution yields parametric maps that agree with maps estimated using the full dictionary (see Fig. 2). Figure 3 com-

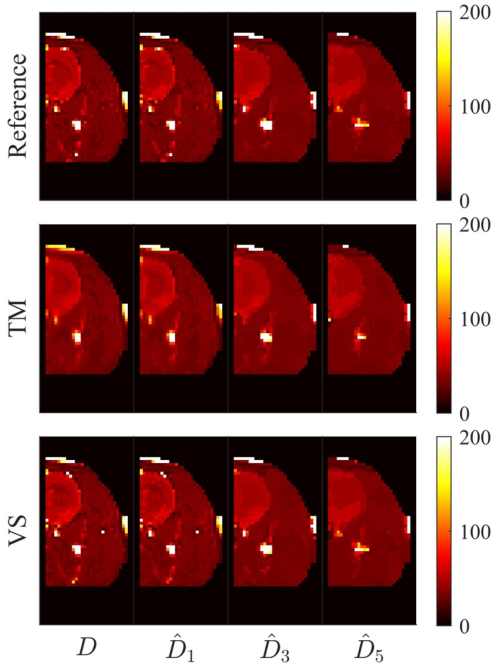

Fig. 3. Estimated T2 parametric maps from fully sampled reference data and data undersampled with an acceleration factor R=5 (20% of k-space) for two different reconstruction methods: template matching (TM) and view-sharing (VS). The clustered dictionaries $\hat{\mathbf{D}}_{\sqrt[3]{P}} \in \mathbb{C}^{K\times TP}$ consisted of $K = 300$, $T = 1000$, and $P = 1 \times 1 \times 1$, $3 \times 3 \times 3$, and $5 \times 5 \times 5$.

pares the reconstructed maps with 20% sampling of k-space, where $\mathbf{D}$ and $\hat{\mathbf{D}}_1$ combined with VS are the most similar to the reference dataset. Figure 4 shows smaller variation of the clustered dictionaries with undersampling, though having less similarity to the reference dataset in fully sampled cases. Fig. 5 evidences how the training error decreases for more clusters in all cases, while the testing error only decreases continuously for $\hat{\mathbf{D}}_1$.

4 Discussion

We use spatiotemporal dictionaries of different spatial kernel sizes with $K = 300$ clusters (0.036% of the original dictionary size) and obtain comparable parametric maps (see Fig. 2). Furthermore, Figs. 3-4 show that clustered dictionaries, especially if they contain spatial information, are more robust to undersampling and shorter sequences. Conversely, the spatial smoothing achieved with larger spatial kernels along with the constant testing errors for increasing clusters in Fig. 5 indicate that the training data does not accurately represent the testing

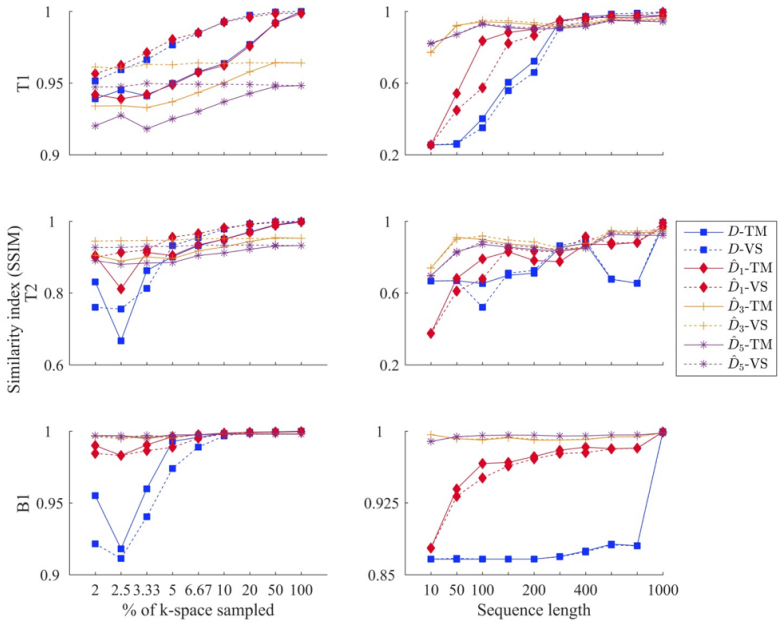

Fig. 4. Acceleration in space and time. The left column displays the SSIM for each of the dictionaries and two reconstruction methods: template matching (TM) and view-sharing (VS) for different levels of k-space sampling. The right column shows the estimated SSIM for increasing sequence length and an acceleration factor R=5 (20% of k-space).

data for kernel sizes larger than $P = 3 \times 3 \times 3$. In fact, the amount of training observations required and the corresponding size of the dictionary in terms of space, time, and clusters, leads to two important discussion points: 1) using clustering enables higher acceleration, at the expense of disregarding parameter combinations that are not present in the training set (e.g. pathology); and 2) adding spatial information increases the dimensionality of the dictionary, requiring approaches that can effectively deal with matching in high dimensional spaces.

5 Conclusions

We propose a method to create clustered MRF dictionaries and show the added benefit of combining it with a view-sharing strategy to enable both accelerated acquisitions by undersampling, and accelerated reconstructions through dictionary compression. Further investigation of data-driven approaches could pave the way towards tissue and disease specific dictionaries in clinical settings.

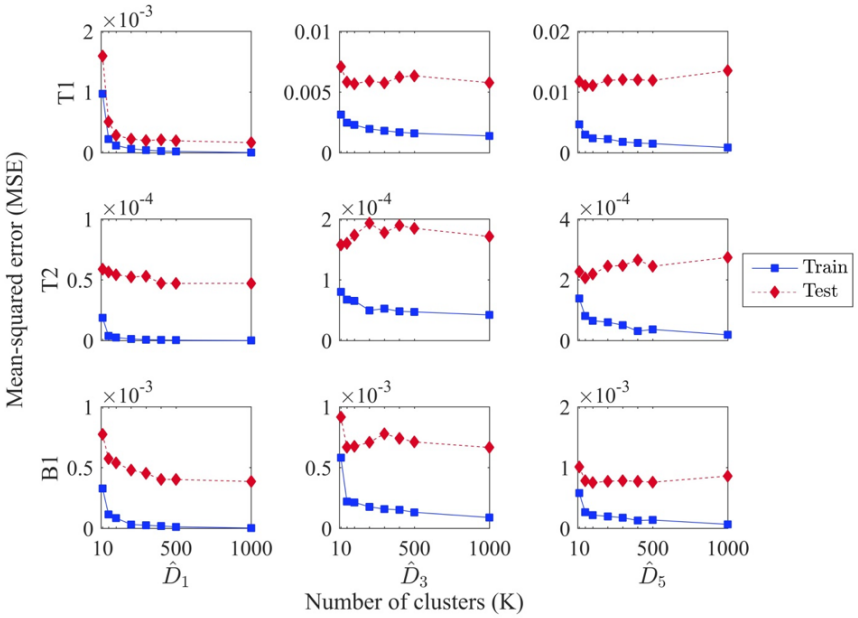

Fig. 5. Training and testing error for different cluster sizes K from the fully sampled reference dataset. For $\hat{\mathbf{D}}_1$ both the training and testing error reduce with an increasing number of clusters, while testing errors for $\hat{\mathbf{D}}_3$ and $\hat{\mathbf{D}}_5$ do not change significantly with increasing clusters.

References

1. Buonincontri, G., Sawiak, S.: Three-dimensional MR fingerprinting with simultaneous B1 estimation. Magnetic Resonance in Medicine 00, 1–9 (2015)
2. Gómez, P.A., Ulas, C., Sperl, J.I., Sprenger, T., Molina-Romero, M., Menzel, M.I., Menze, B.H.: Learning a spatiotemporal dictionary for magnetic resonance fingerprinting with compressed sensing. MICCAI Patch-MI Workshop 9467, 112–119 (2015)
3. Jiang, Y., Ma, D., Seiberlich, N., Gulani, V., Griswold, M.A.: MR Fingerprinting Using Fast Imaging with Steady State Precession (FISP) with Spiral Readout. MRM (2014)
4. Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A.: Magnetic resonance fingerprinting. Nature 495, 187–192 (2013)
5. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 13, 600–612 (2004)

5.2.2 *Accelerated Parameter Mapping with Compressed Sensing: an Alternative to MR Fingerprinting*

Peer-reviewed Conference Abstract

Authors: PA. Gómez, G. Buonincontri, **M. Molina-Romero**, JI. Sperl, MI. Menzel, BH. Menze

In: *Proc Intl Soc Mag Reson Med* (2017) [27]

Abstract: We introduce a method for MRI parameter mapping based on three concepts: 1) an inversion recovery, variable flip angle acquisition strategy designed for speed, signal, and contrast; 2) a compressed sensing reconstruction which exploits spatiotemporal correlations through low rank regularization; and 3) a model-based optimization to simultaneously estimate proton density, T_1, and T_2 values from the acquired measurements. Compared to Magnetic Resonance Fingerprinting (MRF), the proposed method achieves a five-fold acceleration in acquisition time, reconstructs an unaliased series of images, and does not rely on dictionary matching for parameter estimation.

Contribution of thesis author: Discussion of algorithmic implementation and experimental design, manuscript revision and editing.

Accelerated Parameter Mapping with Compressed Sensing: an Alternative to MR Fingerprinting

Pedro A Gómez[1,2], Guido Bounincontri[3], Miguel Molina-Romero[1,2], Jonathan I Sperl[2], Marion I Menzel[2], Bjoern H Menze[1]

[1]Computer Science, Technische Universität München, Munich, Germany
[2]GE Global Research, Munich, Germany
[3]INFN Pisa, Pisa, Italy

Abstract. We introduce a method for MR parameter mapping based on three concepts: 1) an inversion recovery, variable flip angle acquisition strategy designed for speed, signal, and contrast; 2) a compressed sensing reconstruction which exploits spatiotemporal correlations through low rank regularization; and 3) a model-based optimization to simultaneously estimate proton density, T1, and T2 values from the acquired measurements. Compared to MR Fingerprinting, the proposed method achieves a five-fold acceleration in acquisition time, reconstructs an unaliased series of images, and does not rely on dictionary matching for parameter estimation.

1 Purpose

MR fingerprinting (MRF) [12] has recently gained attention due to its ability to simultaneously estimate multiple parametric maps within clinically feasible scan times. MRF is based on three main ingredients: a pseudorandom acquisition, the reconstruction of aliased measurements, and the matching of these measurements to a precomputed dictionary. While conceptually appealing, MRF suffers from multiple methodological shortcomings. The purpose of this work is to present an alternative method to parameter mapping that addresses these limitations by optimizing the acquisition, reconstructing unaliased measurements, and fitting the measurements in a model-based optimization not subject to the discretization of the dictionary.

2 Methods

In an MRI experiment the observed signal $y(t)$ can be described by the combination of a spatial function with a temporal signal evolution:

$$y(t) = \int_{\mathbf{r}} \rho(\mathbf{r}) f_t(\mathbf{r}) e^{-2\pi \mathbf{k}(t)\cdot\mathbf{r}} d\mathbf{r}; \tag{1}$$

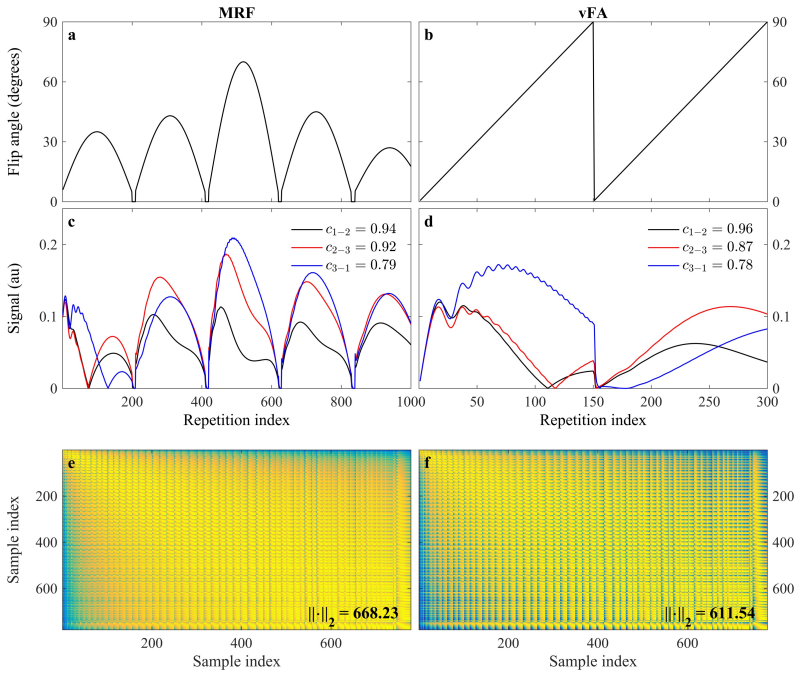

Fig. 1. Flip angle trajectories and correlation between observations. **a-b**, Flip angle trajectory for MRF and variable flip angles (vFA). **c-d**, Temporal signal evolution of three exemplary observations. The legend displays the correlation c between each of them. **e-f**, Correlation matrix for all observations and the norm of the matrix; where the observations are less correlated in vFA.

where $\rho(\mathbf{r})$ is the spatial distribution of the spin density at position $\mathbf{r}$, $\mathbf{k}(t)$ is the k-space trajectory, and $f_t(\mathbf{r})$ is the temporal signal, given by the recursion:

$$f_t(\mathbf{r}) = f_{t-1}(\mathbf{r}) g\left(\boldsymbol{\theta}_a(t); \boldsymbol{\theta}_b(\mathbf{r})\right). \tag{2}$$

The temporal signal $f_t(\mathbf{r})$ at time t is determined by the signal value at the previous time point $f_{t-1}(\mathbf{r})$ modulated by $g(\cdot)$, a function of two different parameter sets: the temporally varying acquisition parameters $\boldsymbol{\theta}_a(t)$, e.g. flip angle $\alpha(t)$ and repetition time $T_R(t)$; and the spatially dependent biological parameters of interest $\boldsymbol{\theta}_b(\mathbf{r})$, such as $T_1(\mathbf{r})$ and $T_2(\mathbf{r})$. In MRF, the temporal signals are denoted fingerprints, where the method aims at creating unique signals for different spatial locations through pseudorandom variations of the acquisition parameters $\boldsymbol{\theta}_a(t)$. Alternatively, we chose $\boldsymbol{\theta}_a(t)$ to satisfy three criteria: speed, signal, and contrast. We increased the speed by minimizing $T_R(t)$, and optimized $\alpha(t)$ for signal and contrast using a training dataset $\mathbf{x} \in \mathbb{C}^{L \times T}$ with L observations and T time points; wherein we experimentally attempted to increase both the orthogonality between observations, and the norm within observations [3,4,8].

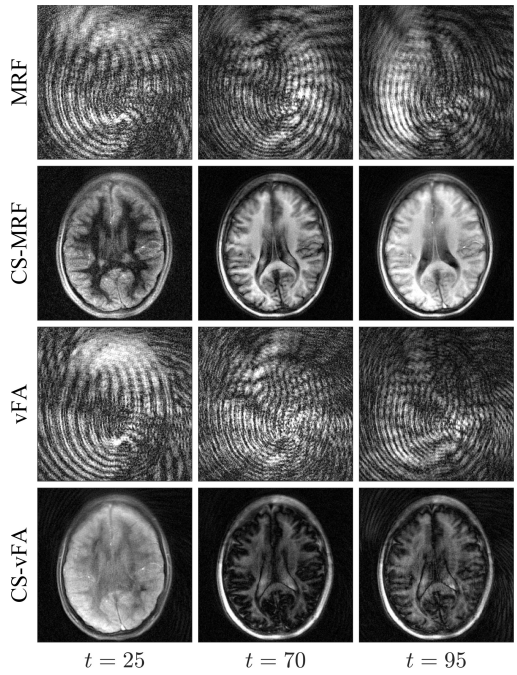

Fig. 2. Reconstructed images at different repetition indexes t. While the nuFFT operation reconstructs a series of aliased images, the proposed CS reconstruction removes the aliasing, allowing for an easier visualization of the temporal dynamics of the signal.

Whereas the original MRF reconstructs aliased images from the measurements, recent work has shown that the acquired data can also be reconstructed in an iterative framework [1,2,5,7,13,17,18]. Based on these ideas, we implemented a compressed sensing (CS) [11] reconstruction that constrains the temporal signal evolution to a low dimensional subspace [14,17], and regularizes the image series by promoting local low rank of spatiotemporal image patches [14,15]. Finally, once we reconstruct an unaliased image series, we propose to replace the matching to a simulated dictionary with an optimization based on least-squares curve fitting for the simultaneous estimationof $\rho(\mathbf{r})$, $T_1(\mathbf{r})$, and $T_2(\mathbf{r})$.

We acquired a single slice from a healthy volunteer based on the FISP implementation of MRF [10] on a GE HDx MRI system (GE Medical Systems, Milwaukee, WI), with an eight channel receive only head RF coil. After an initial inversion, a train of $T = 1000$ pulses with varying flip angles and repetition times was applied ($T_{acq} = 13.15$ s per slice). In addition, we acquired a train of $T = 300$ variable flip angles (vFA) with $T_R = 8$ ms ($T_{acq} = 2.42$ s per slice). For both acquisitions, we used a zero-moment compensated variable density spiral designed with 22.5×22.5 cm FOV, 256×256 matrix size, 1 mm in-plane resolution, 5 mm slice thickness, and golden angle rotations between every in-

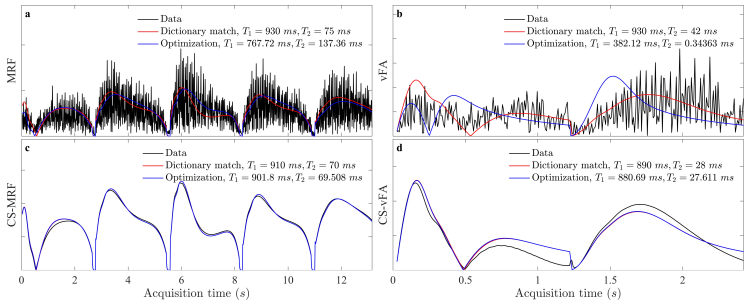

Fig. 3. Signal evolution of a single observation with dictionary matching and model-based optimization. **a-b**, The high levels of aliasing in the acquisitions lead to different parameter estimates between matching and optimization. **c-d**, Matching and optimization results are similar when the evolution is unaliased, where the optimization is not constrained to the discretization of the dictionary. The T_2 values obtained from CS-vFA present an underestimation with respect to the T_2 in MRF and CS-MRF.

terleave. Each acquisition was reconstructed using the nuFFT operator [6] and with the proposed CS method, and parameter maps were subsequently estimated with both dictionary matching and model-based optimization. We simulated the dictionary for a varying range of T_1 and T_2 values using the EPG formalism [9,16].

3 Results

We found that two linear ramps yielded T_1 and T_2 sensitivity while reducing the cost (see Fig. 1). Compared to MRF, the proposed strategy reduces the number of repetitions (Fig. 1a-b), while increasing the orthogonality of the signal evolutions between training observations (Fig. 1c-d). This acquisition, coupled with the proposed CS reconstruction, allows for the recovery of a series of unaliased images (Fig. 2), which in turn facilitate a model-based optimization for parameter mapping (Figs. 3,4).

4 Discussion

When the measurements are aliased, the optimization is subject to local minima due to high noise levels of the temporal signals, resulting in biased parameter estimates. On the other hand, as soon as the image series is unaliased, the optimization converges, yielding parametric maps that are not subject to the discretization of the dictionary. vFA maps display an underestimation of T_2 values, indicating that the vFA strategy could still benefit from better T_2 encoding using variable repetition or echo times.

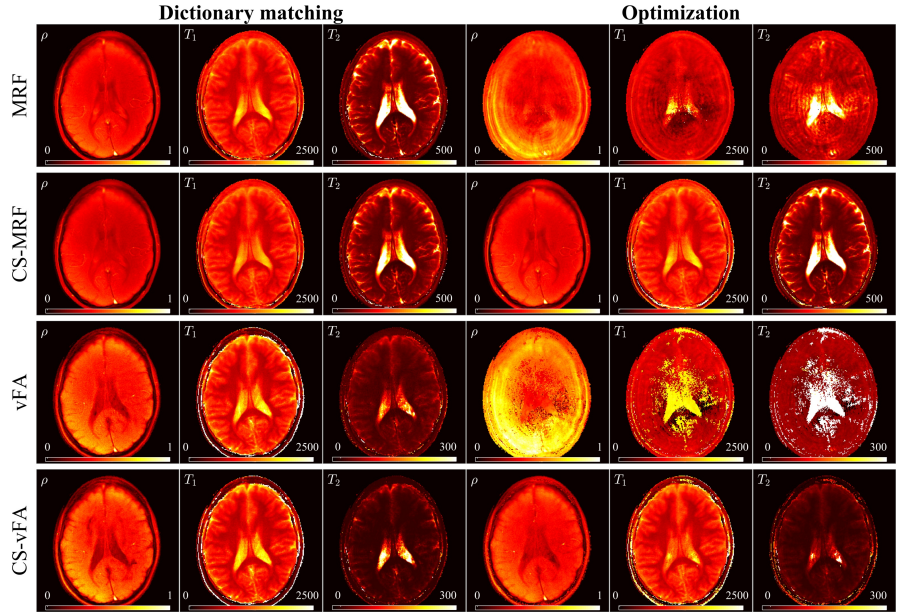

Fig. 4. Estimated parametric maps with dictionary matching and model-based optimization. Dictionary matching yields consistent results with noisy and unaliased signals, while the optimization only converges when fitting unaliased signals (rows 2 and 4, right). When fitting noisy signals, the optimization converges to local minima, leading to an incorrect estimation of parametric maps (rows 1 and 3, right). On the other hand, the optimization is not subject to dictionary discretization, allowing for a wider range of parameters in the solution space. T_2 maps in vFA are underestimated in comparison to MRF.

5 Conclusions

We demonstrated an alternative to MRF based on variable flip angles, a compressed sensing reconstruction, and a model-based optimization. Our proposal reduces the acquisition time by a factor of five, reconstructs unaliased temporal signals, and, notably, does not rely on dictionary matching for parameter estimation.

6 Acknowledgments

With the support of the TUM Institute for Advanced Study, funded by the German Excellence Initiative and the European Commission under Grant Agreement Number 605162.

References

1. Assländer, J., Cloos, M.A., Knoll, F., Sodickson, D.K., Hennig, J., Lattanzi, R.: Low Rank Alternating Direction Method of Multipliers Reconstruction for MR Fingerprinting. arXiv eprint p. arXiv:1608.06974 (2016)
2. Cline, C.C., Chen, X., Mailhe, B., Wang, Q., Nadar, M.: AIR-MRF: Accelerated iterative reconstruction for magnetic resonance fingerprinting. Proc Intl Soc Mag Reson Med (2016)
3. Cohen, O.: Magnetic Resonance Fingerprinting Trajectory Optimization. Proc Intl Soc Mag Reson Med 1, 1 (2014)
4. Cohen, O., Sarracanie, M., Rosen, M.S., Ackerman, J.L.: In Vivo Optimized Fast MR Fingerprinting in the Human Brain. Proc Intl Soc Mag Reson Med (2016)
5. Davies, M., Puy, G., Vandergheynst, P., Wiaux, Y.: A Compressed Sensing Framework for Magnetic Resonance Fingerprinting. SIAM Journal on Imaging Sciences 7(4), 2623–2656 (dec 2014)
6. Fessler, J.A., Sutton, B.P.: Nonuniform Fast Fourier Transforms Using Min-Max Interpolation. IEEE Trans Signal Processing 51(2), 560–574 (2003)
7. Gómez, P.A., Ulas, C., Sperl, J.I., Sprenger, T., Molina-Romero, M., Menzel, M.I., Menze, B.H.: Learning a spatiotemporal dictionary for magnetic resonance fingerprinting with compressed sensing. MICCAI Patch-MI Workshop 9467, 112–119 (2015)
8. Hamilton, J.I., Wright, K.L., Jiang, Y., Hernandez-Garcia, L., Ma, D., Griswold, M., Seiberlich, N.: Pulse Sequence Optimization for Improved MRF Scan Efficiency. Proc Intl Soc Mag Reson Med 23, 3386 (2015)
9. Hennig, J.: Echoes—how to generate, recognize, use or avoid them in MR-imaging sequences. Part I: Fundamental and not so fundamental properties of spin echoes. Concepts in Magnetic Resonance 3(3), 125–143 (jul 1991)
10. Jiang, Y., Ma, D., Seiberlich, N., Gulani, V., Griswold, M.A.: MR Fingerprinting Using Fast Imaging with Steady State Precession (FISP) with Spiral Readout. MRM (2014)
11. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Society of Magnetic Resonance in Medicine 58, 1182–1195 (2007)
12. Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A.: Magnetic resonance fingerprinting. Nature 495, 187–192 (2013)
13. Pierre, E.Y., Ma, D., Chen, Y., Badve, C., Griswold, M.A.: Multiscale reconstruction for MR fingerprinting. Magnetic Resonance in Medicine 2492, 2481–2492 (2015)
14. Tamir, J.I., Uecker, M., Chen, W., Lai, P., Alley, M.T., Vasanawala, S.S., Lustig, M.: T 2 shuffling: Sharp, multicontrast, volumetric fast spin-echo imaging. Magnetic Resonance in Medicine (2016)
15. Trzasko, J., Manduca, A.: Local versus Global Low-Rank Promotion in Dynamic MRI Series Reconstruction. Proc Intl Soc Mag Reson Med 24(7), 4371 (2011)
16. Weigel, M.: Extended phase graphs: Dephasing, RF pulses, and echoes - pure and simple. Journal of Magnetic Resonance Imaging (2014)
17. Zhao, B., Setsompop, K., Gagoski, B., Ye, H., Adalsteinsson, E., Grant, P.E., Wald, L.L.: A Model-Based Approach to Accelerated Magnetic Resonance Fingerprinting Time Series Reconstruction. Proc Intl Soc Mag Reson Med (2016)
18. Zhao, B., Setsompop, K., Ye, H., Cauley, S., Wald, L.L.: Maximum Likelihood Reconstruction for Magnetic Resonance Fingerprinting. IEEE Trans Med Imaging 0(0), 0 (2016)

Part III

DISCUSSION AND CONCLUSION

6 DISCUSSION OF THE PRESENTED METHODS

The core of this thesis aims to understand tissue composition beyond the resolution limit of MRI. To this end, three projects were completed focusing on dMRI and multiparametric mapping, using advanced analysis techniques like BSS, deep learning, and Bayesian inference. This dissertation is composed of three publications plus one under review, five abstracts, and one master thesis.

The formulation of BSS as a multicompartment tissue microstructure problem with multi-echo dMRI (section 3.1 and chapter 4) data yields two main advantages: first, it sets the basis for compartmental signal disentangling without geometrical assumptions [23] or ill-posed ILT; and second, it opens a new way for simultaneous quantification of volumetry and relaxometry of the tissue water compartments. Solutions based on SCA [50], NSC [51], and NMF [52] were explored with simulations, phantoms, and in vivo repeatably and reproducibility studies for more than twenty volunteers. NMF was proved to be a robust solver after the introduction of physically informed constraints for two compartments, producing free-water corrected T_2 maps equivalent to the reference method and paving the way for FWE.

The BSS project shed light into the second contribution of this dissertation. The classical two-compartments fitting of the DTI model is ill-posed and required regularization [57] of specific acquisition [31]. This thesis introduces a deep learning approach that estimates the volume fraction of the tissue from the evolution of the diffusion signal (section 3.2 and chapter 4) [53]. This transforms the two-compartments problem into a single compartment, shifting the perspective of the problem and eliminating the need of regularization. The clinical images from twenty glioblastoma patients were processed with this method, yielding more relevant information from the available data.

This thesis contributed to the multiparametric mapping techniques in the conceptual discussion, signal processing, and implementation (chapter 5). A Bayesian inference framework implemented on ΠΛU [29] was adapted to work on dynamic MRI signals inspired by MRF. As part of the co-supervision of a master thesis this framework was extended to model up to three microstructural compartments, filling the gap between multiparametric mapping and tissue microstructure characterization (appendix A.2.1). The results were consistent with literature values of myelin, parenchyma, and CSF.

7 OUTLOOK AND FUTURE WORK

This thesis pushes the state-of-the-art in tissue microstrucutre characterization in three ways. First, the bridge created between BSS theory and dMRI allows for a new manner of approaching the compartmental nature of brain tissue. This enables for minimum modeling and no ILT. With BSS not only the volumetry and relaxometry compartmental properties are estimated, also the diffusion signals from each component are disentangled and can be studied separately. This feature is exclusive of BSS paving the way for a deeper understanding. Second, there is a shift in perspective for the FWE issue. This dissertation introduces a deep learning pattern matching approach that substitutes previous regularization techniques, reaching an improved performance and independence of the diffusion protocol. Finally, we extended the work on multiparameter mapping to account for several sub-voxel components, opening a new way for ultra-fast myelin fraction estimation and g-ratio quantification in neuroimaging [46].

The ill-posed nature of NMF is well reported in literature. Here, the presented BSS is based on a ALS version that uses physical constraints to regularize the matrix factorization achieving feasible solutions. Nevertheless, ALS is an unsupervised machine learning algorithm that works on data patterns. As such, it is also subject of being learned by a deep learning method, merging two of the contributions of this thesis into a more powerful BSS approach. In that sense, the stability shown by this algorithms will be expected to improve the convergence issues for more than two compartments, extending its reach to myelin diffusion, volume and relaxometry estimation.

A major limitation found with the BSS multi-echo analysis lays in the restriction of the minimum achievable TE. This value falls near 60 ms in clinical scanners, hindering the observation of myelin. State-of-the-art research hardware [68] and preclinical scanners allow reaching TE values below 40 ms, opening a small window for the third compartment. Moreover, diffusion acquisition techniques with several echoes integrate multi-echo data into one acquisition, preventing the burden caused by protocol repetition.

Multiparametric mapping of tissue microstructure is highly relevant due to its efficiency in time-information terms. This work addresses the compartmental disentanglement using data acquired and reconstructed accounting only for one compartment. This is a sub-optimal approach and further research to extend the impact of this method is necessary.

Part IV

APPENDIX

A

MANUSCRIPT UNDER REVIEW AND MASTER THESIS

A.1 MANUSCRIPT UNDER REVIEW

The manuscript contained in this appendix, Deep learning with synthetic diffusion MRI data for free-water elimination in glioblastoma cases, is the continuation of the work presented in Deep learning with synthetic data for free water elimination in diffusion MRI, and was submitted to the MICCAI international conference. The clinical relevance of the deep learning FWE method is shown in 20 patients affected by gliablastoma. This techniques output more relevant information from the clinical data, including edema and tumor infiltration delineation.

A.1.1 *Deep learning with synthetic diffusion MRI data for free-water elimination in glioblastoma cases*

Peer-reviewed Conference Paper

Authors: M. Molina-Romero, B. Wiestler, PA. Gómez, MI. Menzel, BH. Menze

In: *Magnetic Resonance in Medicine*, doi: 10.1002/mrm.27181 (2018), *in press*, p. 10 [52].

Abstract: Glioblastoma is the most common and aggressive brain tumor. In clinical practice, diffusion MRI (dMRI) enables tumor infiltration assessment, tumor recurrence prognosis, and identification of white- matter tracks close to the resection volume. However, the vasogenic edema (free-water) surrounding the tumor causes partial volume contamination, which induces a bias in the estimates of the diffusion properties and limits the clinical utility of dMRI. We introduce a voxel-based deep learning method to map and correct free-water partial volume contamination in dMRI. Our model learns from synthetically generated data a non-parametric forward model that maps free-water partial volume contamination to volume fractions. This is independent of the diffusion protocol and can be used retrospectively. We show its benefits in glioblastoma cases: first, a gain of statistical power; second, quantification of free-water and tissue volume fractions; and third, correction of free-water contaminated diffusion metrics. Free-water elimination yields more relevant information from the available data.

Contribution of thesis author: Model development and implementation, experimental analysis, manuscript preparation and editing.

Deep learning with synthetic diffusion MRI data for free-water elimination in glioblastoma cases

Miguel Molina-Romero[1,2][0000−0001−8054−0426], Benedikt Wiestler[3][0000−0002−2963−7772], Pedro A. Gómez[1,2][0000−0003−3709−3557], Marion I. Menzel[2][0000−0003−0087−9134], and Bjoern H. Menze[1][0000−0003−4136−5690]

[1] Computer Science, Technischen Universität München, Munich, Germany
[2] GE Healthcare Global Research Organization, Munich, Germany
[3] Department of Neuroradiology, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany.

Abstract. Glioblastoma is the most common and aggressive brain tumor. In clinical practice, diffusion MRI (dMRI) enables tumor infiltration assessment, tumor recurrence prognosis, and identification of white-matter tracks close to the resection volume. However, the vasogenic edema (free-water) surrounding the tumor causes partial volume contamination, which induces a bias in the estimates of the diffusion properties and limits the clinical utility of dMRI.
We introduce a voxel-based deep learning method to map and correct free-water partial volume contamination in dMRI. Our model learns from synthetically generated data a non-parametric forward model that maps free-water partial volume contamination to volume fractions. This is independent of the diffusion protocol and can be used retrospectively. We show its benefits in glioblastoma cases: first, a gain of statistical power; second, quantification of free-water and tissue volume fractions; and third, correction of free-water contaminated diffusion metrics. Free-water elimination yields more relevant information from the available data.

Keywords: Glioblastoma, Brain Tumor, DTI, Deep learning, Fractional Anisotropy, Free-water elimination, Data harmonization.

1 Introduction

Glioblastomas are the most common primary brain tumor. Ninety percent of these are IDH-wild-type and have a dismal prognosis, with a 5-year survival rate of less than 10%. Diffusion magnetic resonance imaging (dMRI) and the diffusion tensor model (DTI) are used clinically for surgical planing. DTI yields quantitative estimates of the tissue diffusivity, e.g. mean diffusivity (MD), and fractional anisotropy (FA), an index of tissue microstructure organization. Previous research focused on FA as indicator of tumor grade, tumor cellularity, tumor infiltration and edema assessment [1], and tumor recurrence [2]. However, results are controversial due to reproducibility issues derived from differences in methology, image acquisition, or post-processing [1].

II

Data harmonization in dMRI is attracting attention to overcome these problems [3]. One way to reduce uncontrolled variability is to eliminate the free-water signal [4]. Free-water elimination (FWE) uses a two-compartments tissue model composed by tissue (or parenchyma) and free-water [5]. Fitting the diffusion tensor in a two-compartments model is an ill-posed problem that has been solved using spatial regularization [6] or optimized acquisition protocols [7] [8]. Harmonization of image resolution and diffusion directionality can be achieved by image quality transfer (IQT) with non-linear learninig algorithms, such as convolutional neural networks [9] or random forest [10]. IQT offers a new dimension in dMRI, enabling learning complex diffusion model on high quality data, to then transfer information captured in signal patterns to low quality data. However, this approach is limited by the availability of rich datasets.

We propose an new method for free-water elimination based on an artificial neural network (ANN), trained with synthetically generated data, that is independent of the number of diffusion shells (b-values) and can be applied retrospectively to any dMRI data. Instead of regularizing the FWE ill-posed inverse problem, we teach a non-parametric forward model to learn the mapping between partial volume contamination and free-water volume fraction from synthetic data. Besides, unlike IQT, our approach works only in the diffusion dimension, enabling an important simplification of the ANN model. We further show the advantages of FWE in glioblastoma cases: 1) a gain of statistical power through data harmonization, 2) complementary information of the tissue microstructure composition, and 3) better assessment of edema, tumor, and tumor infiltrated areas.

2 Methods

Diffusion signal modeling: Following previous work on free-water elimination we modeled the diffusion signal of a single voxel, along the diffusion directions $(b, \boldsymbol{g})$, as the contribution of tissue and free-water compartments:

$$S(TE, b, \boldsymbol{g}) = S_0 \left(\hat{f}_t e^{\frac{-TE}{T_{2_t}}} S_t(b, \boldsymbol{g}) + \hat{f}_{fw} e^{\frac{-TE}{T_{2_{fw}}}} S_{fw}(b, \boldsymbol{g}) \right), \tag{1}$$

where b and $\boldsymbol{g}$ summarize the gradient effects; S_0 is a scaling factor proportional to the proton density; $\hat{f}_t$, $\hat{f}_{fw}$, T_{2_t}, and $T_{2_{fw}}$ are the volume fraction and T_2 values of tissue and free-water respectively. Since $T_{2_t} < T_{2_{fw}}$, measurements at different echo-times (TE) yield distinct contributions of tissue and free-water. Thus, disentangling the volume fractions from the T_2 effects requires measurements at least two different TEs [8]. However, in clinical routine only one TE is acquired simplifying Eq.1:

$$S(b, \boldsymbol{g}) = S_0 \left(f_t S_t(b, \boldsymbol{g}) + f_{fw} S_{fw}(b, \boldsymbol{g}) \right), \tag{2}$$

where the T_{2_i} and $\hat{f}_i$ effects are integrated in f_i, inducing a positive bias towards the new free-water volume fraction (f_{fw}) as TE increases. The volume

III

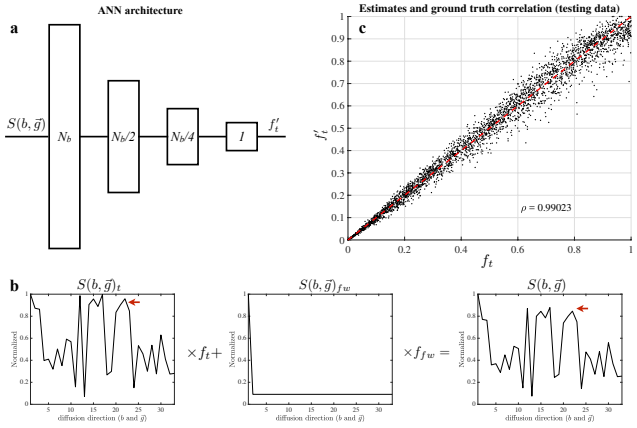

Fig. 1. Artificial neural network model. The ANN architecture (a). Training data was synthetically generated following Eq. 2 (b). Free-water partial volume contamination effects were visible (b, red arrows). We ran a correlation analysis for 3000 randomly generated samples (c) reaching a factor of 0.99.

fraction indexes are ratios relative to the signal contribution of each compartment and thus $f_{fw} + f_t = 1$.

Synthetic singal generation: The diffusion properties of free-water at body temperature are well characterized [5], presenting isotropic behavior and a diffusion coefficient $D_{fw} = 3 \times 10^{-3}$ mm^2/s, and $S_{fw}(b, \boldsymbol{g}) = e^{-bD_{fw}}$. On the other hand, S_t is unknown since it depends on the tissue microstructure organization and the orientation of the diffusion gradients, $\boldsymbol{g}$. Thus, we modeled its behavior with a random variable, $S_t \in \mathbb{R}^{N_b}$, following an uniform distribution, $U(0, 1)$, where N_b is the number of measured b-values, including non-diffusion-weighted volumes. Furthermore, the tissue volume fraction, $f_t \in [0, 1]$, varies from voxel to voxel depending on the amount of free-water infiltrated within the tissue, thus, we also represented it as a random uniform variable, $U(0, 1)$. Based on the prior knowledge of D_{fw}, the models of S_t and f_t, and knowing the acquisition b-values, it is possible to generate unlimited synthetic diffusion signals, $S(b, \boldsymbol{g})$, containing free-water partial volume effects (Eq. 2 and Fig. 1b).

ANN architecture and training: We designed a regression fully connected ANN capable of estimating the tissue volume fraction, f_t, directly from the diffusion signal $S(b, \boldsymbol{g})$ (Eq. 2). The input layer contained as many units as the number of acquired b-values (including non-diffusion-weighted volumes), N_b, and a single output unit yielding the estimate of f_t. The ANN had a pyramidal architecture with two hidden layers with $N_b/2$ and $N_b/4$ respectively (Fig. 1a). To train the ANN we used 20000 synthetic signals generated as explained above (70% training, 15% validation, and 15% testing). Convergence for $N_b = 33$ (Fig. 1) was reached after nine epochs and 4.7 seconds in a consumers laptop (Apple

IV

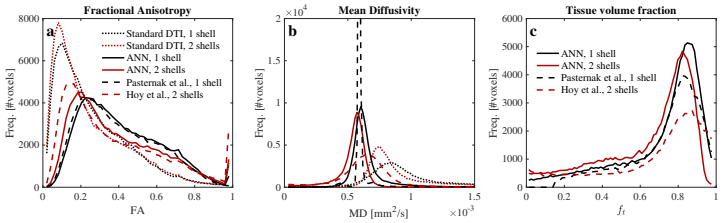

Fig. 2. Comparison of FWE with ANN and state of the art methods for two diffusion protocols. The correction effects on FA of ANN for one shell were equivalent to Pasternak's et al. algorithm (a), while ANN MD estimates were less over-regularized (b). Tissue volume fraction estimates were also in agreement although Pasternak's method failed to estimate $f_t < 0.18$ (c). The ANN for two shells and Hoy's et al. algorithm were in good agreement for FA (a) but larger differences were visible for MD and f_t (b and c).

MacBook Pro, Intel Core i5, 8GB RAM; MATLAB, Mathworks, Natwick, MA). The training process depended on the protocol prescribed b-values. Thus, we trained four networks to match the DWI data used in the experiments below.

3 Experiments and Results

Comparison with state of the art: Methods from Pasternak et al. [6] and Hoy et al. [7] are the state of the art for one and two shell acquisitions. For comparison we measured data from a volunteer in a GE 3T MR750w (GE Healthcare, Milwaukee, WI). The protocol comprised first, one diffusion weighted imaging (DWI) acquisition for 30 diffusion direction (b=500 s/mm^2) and two b=0 volume (N_b=32); and second a DWI for two shells (b=500 and 1000 s/mm^2) with 30 diffusion directions for each shell and four non-diffusion-weighted volumes (N_b=64). The data was processed with a pipeline including steps for: 1) head motion and eddy current corrections (FSL eddy); 2) denoising based on random matrix theory [11]; and 3) free-water elimination based on ANN, Pasternak's, and Hoy's methods (Fig. 2). The ANN results were comparable to those of the state of the art methods.

Data harmonization: The T_2 effects described in Eq.1 and Eq. 2 were investigated using data from a volunteer scanned in the same scanner as before. The protocol consisted of a DWI acquisition for 30 direction (b=1000 s/mm^2) and one non-diffusion-weighted volume (N_b = 31). This was repeated for seven equispaced TE = 74.9 – 134.9 ms. The data was processed as described before, but only ANN FWE was computed. For comparison two processing lines were created with and without ANN based FWE. Both were fitted with robust DTI (RESTORE) [12] to extract diffusion metrics.

The multi-echo diffusion data showed an increase of free-water (Fig. 3c) and its effects (Fig. 3a, larger low FA peak) with the TE, which agrees with the

V

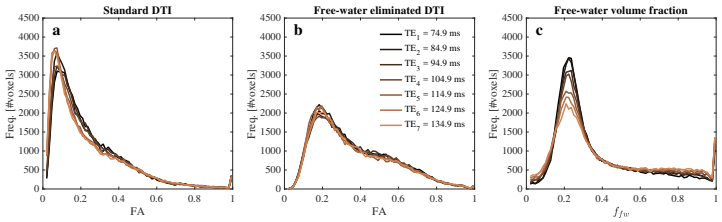

Fig. 3. Histogram comparison at several TEs. FA from standard DTI (a) showed larger variability across TEs for FA < 0.4, than the free-water corrected (b). The tissue volume fraction estimates are influenced by T_2 decay (Eq. 1) and thus sensitive to TE (c). The free-water elimination step has a TE harmonization effect on the diffusion signal (b), shifting the TE variability into the volume fraction estimate (c).

two-compartments tissue model (Eq. 1) accounting for T_2 effects, simplified in Eq. 2. Multi-center studies are often carried out in data acquired with heterogeneous protocols. The prescribed TEs are a source of variability that mostly depend on the gradient strength of the scanner and the image resolution. Data harmonization is important to remove uncontrolled variability and achieve a good statistical power. Free-water elimination plays a double role. First, it accounts for T_2 effects in the f_i, shifting this variability from the diffusion metrics to the volume fraction estimates (Fig. 3). And second, it removes the "diffusion isotropic noise" from the signal showing the actual tissue anisotropy, eliminating the variability induced by the presence of free-water.

Glioblastoma analysis: Data from 25 patients affected by glioblastoma (IDH wild-type, WHO 2016 classification) were provided by the Department of Neuroradiology, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany. All patients are part of a prospective glioma database, approved by the local ethics committee, and gave written informed consent. They were scanned in a 3T whole-body scanner (Achieva, Philips Medical Systems, Best, The Netherlands). The protocol included DWI for 32 directions (b=800 s/mm^2) and one non-diffusion-weighted volume ($N_b = 33$). Furthermore, T_2 turbo spin echo (T2w), T_2-FLAIR, and T_1 contrast enhanced (CE-T1w) were acquired. The DWI data was processed as described before, and T2w, CE-T1w, and FLAIR volumes were registered to the DWI space (Fig. 4).

Tissue and free-water volume fraction estimates: The free-water and tissue maps computed with ANN (Fig. 4d and h) showed tumor and edema areas that were in agreement with well established methods: T2w, CE-T1w, and FLAIR (Fig. 4e, f, and g). The estimation of tissue and free-water volume fraction maps is an important feature of FWE. They provide complementary information of the underlying tissue organization. Cytotoxic edema and necrosis areas that are not distinguishable can be better identified with the knowledge of the amount of tissue in the voxel (Fig. 4d, e, f, g, and h, light-blue arrows). Furthermore, in clinical routine tumor delineation is based on CE-T1w hyper-intensities (Fig. 4f

VI

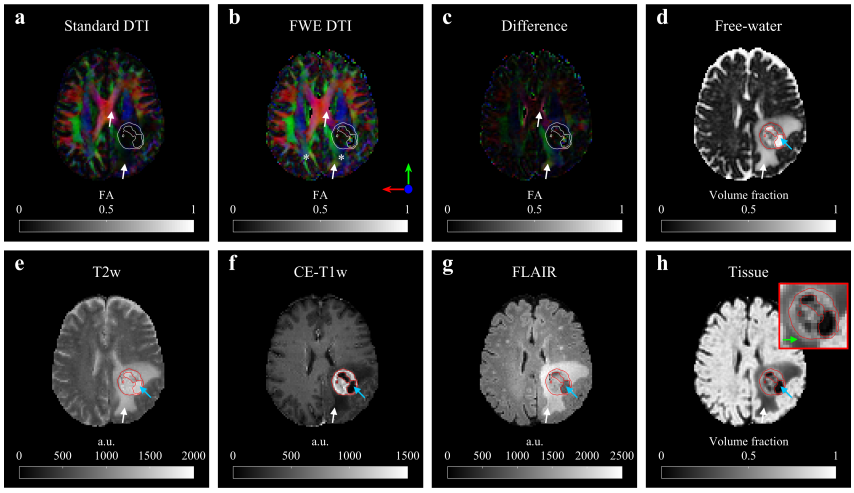

Fig. 4. Comparison of metrics and contrasts for one patient. Fractional anisotropy from standard DTI (a) showed a dimmed region corresponding to edema (white arrows). This area was recovered after free-water elimination (b and c). We observed alterations of the white matter integrity compared to the NAWM (b, white asterisks). Free-water (d) and tissue (h) volume fraction maps were estimated by the ANN. Edema regions were well defined and distinguishable from the tumor (red contour). Pools compatible with cytotoxic edema were observable inside the tumor (light-blue arrows). These findings agreed with the observations based on T2w (e), CE-T1w (f), and FLAIR (g). Extended tumor infiltration was derived from the comparison of CE-T1w and tissue volume fraction map (h, gree arrow).

red contour). When this is compared with tissue volume fraction maps (Fig. 4h zoomed area), we observed an increased tumor region of up to four millimeters that is compatible with tumor infiltration. This agrees with the radiated area after resection (Fig. 4h, green arrow).

Fractional anisotropy recovery: The comparison between standard and FWE FA maps (Fig. 4a and b) exhibited a recovery of the anisotropic information in the edema region and around the ventricles (Fig. 4c, white arrows). The elimination of the isotropic compartment from the diffusion signal leads to a recovery of the tissue anisotropy captured by the protocol, leading to an enhancement of the FA maps, especially in areas with large partial volume contamination like edema and the border of the ventricles (Fig. 4a, b, and c, white arrows). The correction of the FA maps provides new information of the tissue microstructure integrity hidden by the edema (Fig. 4b asterisks).

Edema and infiltration unmixing: To assess the impact of FWE in edema, tumor, and tumor infiltrated areas, regions of interests (ROI) were defined by a neuroradiologist for each subject using FLAIR, T2w, CE-T1w, f_t, and FWE

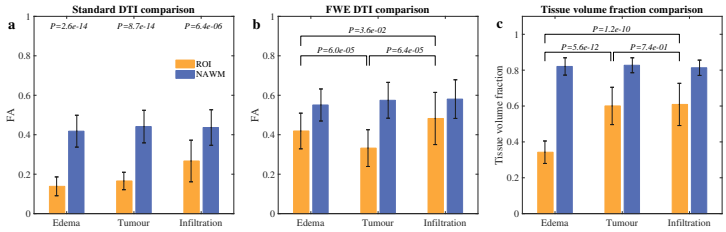

Fig. 5. Edema, tumor, and tumor infiltrated regions differentiation. The FA values for standard DTI (a) were statistically different ($\alpha = 0.01$) between NAWM areas. After free-water elimination (b) FA values in edema and tumor infiltrated areas were significantly larger than those for tumor (family-wise error = 0.03), indicating a better organized tissue microstructure. Complementary, the tissue volume fractions in tumor and tumor infiltrated regions were statistically larger than in edema areas, suggesting that the loss in FA in edema was mostly due to free-water infiltration.

FA maps(Fig. 4). To minimize the influence of outliers we used the median for each ROI, and ran t-test comparison for FA and tissue volume fraction values across subjects for the three type of ROIs (Fig. 5). For reference, extra ROIs were drawn in normal appearing white mater (NAWM) mostly contralateral.

Comparison of NAWM with tumor and edema areas for the 25 patients showed a statistically significant difference ($\alpha = 0.01$) in FA for standard DTI (Fig. 5a). Tumor infiltration and edema are the driving factors behind the loss in FA. After free-water elimination we compared FWE FA (Fig. 5b) and tissue volume fraction (Fig. 5c) in edema, tumor and tumor infiltrated areas (family-wise error = 0.03). Significantly larger FA was found in edema and tumor infiltrated areas compared to tumor regions. While, tissue volume fraction in edema was statistically lower than in tumor infiltrated and tumor sections.

We hypothesize that the combination of FWE FA and tissue volume fraction yields a better understanding of tissue microstructure integrity (Fig. 5). High FA and low tissue volume fraction might indicate well organized microstructure infiltrated by free-water (vasogenic edema). Low FA and high tissue volume fraction is compatible with the unstructured cellularity found in tumors. Finally, areas with high FA and tissue volume fraction might be caused by tumor infiltration in highly structured white matter bundles.

Discussion

The presented ANN model combines simplicity with high accuracy. We introduce for the first time an ANN design capable of learning partial volume effect features from synthetic data, and reach comparable results with the state of the art methods but at least 55-fold faster. Besides, our approach, unlike [6] is voxel-based, avoiding blurring artifacts induced by the use of patch based

regularization, and can be applied to any diffusion protocol beyond DTI. The robustness of our approach was shown with 25 patients.

The addition of free-water elimination in processing pipeline extracts more information from the data, which has a potential benefit for glioblastoma patients in three aspects: diagnosis, surgical planning, and guided radiotherapy. The diagnosis potentially improves with the quantification of the free-water volume fraction map, yielding the severity of the vasogenic edema. The precision of the surgical planning benefits from the corrected FA maps in edema areas providing a better definition of white matter fiber bundles and limiting the resection of healthy tissue. Finally, guided radiotherapy profits from identification of tumor infiltrated areas.

Acknowledgments

The authors want to thank Dr. Ofer Pasternak for his support in the comparison of the methods. This work was supported by the TUM Institute of Advanced Study, funded by the German Excellence Initiative, and the European Commission (Grant Agreement Number 605162).

References

1. A.S. Field. Diffusion imaging in brain tumors. In D.K. Jones, editor, *Diffusion MRI*, chapter 33, pages 547–563. Oxford University Press, 2010.
2. S. Bette et al. Local Fractional Anisotropy Is Reduced in Areas with Tumor Recurrence in Glioblastoma. *Radiology*, 283(2):499–507, 2017.
3. Z. Eaton-Rosen et al. Beyond the Resolution Limit: Diffusion Parameter Estimation in Partial Volume. *MICCAI*, pages 605–612, 2016.
4. C. Metzler-Baddeley et al. How and how not to correct for CSF-contamination in diffusion MRI. *Neuroimage*, 59(2):1394–1403, 2012.
5. C. Pierpaoli and D.K. Jones. Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model. In *ISMRM, Kyoto*, page 1215, 2004.
6. O. Pasternak et al. Free Water Elimination and Mapping from Diffusion MRI. *Magn. Reson. Med.*, 730:717–730, 2009.
7. A.R. Hoy et al. Optimization of a Free Water Elimination Two-Compartment Model for Diffusion Tensor Imaging. *Neuroimage*, (103):323–333, 2014.
8. M. Molina-Romero et al. Theory, validation and application of blind source separation to diffusion MRI for tissue characterisation and partial volume correction. In *ISMRM, Honolulu*, page 3462, 2017.
9. R. Tanno et al. Bayesian image quality transfer with CNNs: Exploring uncertainty in dMRI super-resolution. volume 10433 LNCS, pages 611–619, 2017.
10. D.C. Alexander et al. Image quality transfer and applications in diffusion MRI. *Neuroimage*, 152(March):283–298, 2017.
11. J. Veraart et al. Denoising of diffusion MRI using random matrix theory. *Neuroimage*, 142:394–406, nov 2016.
12. L. Chang et al. Restore: Robust estimation of tensors by outlier rejection. *Magnetic Resonance in Medicine*, 53(5):1088–1095, 2005.

A.2 RELEVANT CO-SUPERVISED MASTER THESIS

As part of this dissertation, the master thesis with title Brain Microstructure Quantification from Transient-State Magnetic Resonance Imaging was co-directed. This expands the work from chapter 5 into a tissue microstructure framework. Brain sub-voxel information accounting for meylin water, intra/extra-cellular water, and CSF are extracted from the evolution of the QTI signals. To that end, a Bayesian inference technique based on TMCMC was incorporated to the data processing. In this section, a summary of the master thesis is reported.

A.2.1 *Brain Microstructure Quantification from Transient-State Magnetic Resonance Imaging*

Master Thesis

Authors: MP. Orihuela Martinez-Costa, **M. Molina-Romero**, PA. Gómez, BH. Menze.

Abstract: Diffusion Magnetic Resonance Imaging (dMRI) is a powerful and non-invasive imaging technique that has become a key tool in Neuroscience and is continuously contributing to the expansion of our understanding of the human brain. One of the next goals in MRI research is to achieve quantification of tissue properties within clinically feasible scanning times. Quantitative maps can be more representative of underlying changes at the cellular level than the standard qualitative imaging. Bringing such methods to the clinical practice would allow objective comparison between examinations in follow-up studies and a better assessment of disease in the brain. This thesis aims to develop a new method to quantify brain microstructure using an ultrafast transient-state based image acquisition approach.

Multicomponent relaxometry is a quantitative technique for brain microstructure imaging that considers that the signal from each voxel in an image encodes the contribution from different water relaxing pools also called microstructural compartments. In this thesis, three multicomponent relaxometry models have been implemented: one-compartment, two-compartments and three-compartments, where the parameters to be quantified are the relaxation times and the volume fraction of each tissue compartment.

Bayesian uncertainty quantification and propagation has been used to estimate the microstructure parameters of each model. By using TMCMC sampling in a Bayesian estimation framework, it is possible to sample from a prior parameter distribution and iteratively converge to the posterior distribution of each microstructure parameter. Addressing the problem from a Bayesian approach enables quantification and uncertainty estimation of each of the model parameters which lead to quantitative maps of the microstructure parameters of interest. Three experiments have been performed to validate and test the microstructure models proposed: simulation experiments, phantom acquisition and one in vivo acquisition, obtaining quantitative maps of the relaxation times of each tissue compartment as well as the volume fraction of each compartment.

Additionally, the evidence of each model was calculated and used to define which one of the three models represents best the data at each voxel location.

A.2.1.1 *Introduction*

The aim of this work is to use the partial volume effects (section 1.1.4) induced in the QTI signal (section 1.1.3 and chapter 5) to gain microstructural information of the measured tissue.

A.2.1.2 *Methods*

The tissue modeling and quantification framework are described in section 3.3. Here, experimental design for simulations, phantom, and in vivo examination are described for one, two, and three compartments. All the measurements were performed in a 3T MR750w scanner (GE Healthcare, Milwaukee, WI).

SIMULATIONS Simulated data were generated using EPG [70, 71] for each compartment. To that end, T_{1_i}, T_{2_i}, and σ^2 values were defined. Gaussian noise was added to the signals for an SNR = 80 dB. The compartmental parameters were estimated with ΠΛU.

PHANTOM EXPERIMENTS A single slice of the Eurospin T05 phantom [44] was acquired and reconstructed with QTI. The signal from different tubes were artificially mixed for defined weights. These mixtures were then processed with ΠΛU for disentangling.

IN VIVO EXPERIMENTS One slice was acquired from a healthy volunteer and reconstructed with QTI. The protocol included 500 repetions for constant TR/TE = 8/2 ms, and linearly increasing $\alpha \in [1 - 70]$ degrees. The temporal evolution of the signal in each voxel was analyzed with ΠΛU for one, two, and three microstructure compartments. It was expected that the model evidence for one compartment was larger in the ventricles, composed mostly for CSF; for two compartments in Gray Matter (GM); and for three compartments in White Matter (WM) due to the influence of myelin.

A.2.1.3 *Results and Discussion*

The reader can find the results ilustrated with figures in the original document of the master thesis [55].

This work proves with simulation, phantom, and in vivo experiments that is possible to extract compartmental information from the signal evolution of QTI due to the partial volume effects induced in the signal. This approach introduces an ultra-fast method for tissue microstructure characterization.

The results show that the three compartments model represents better the signal evolution. This agrees with literature findings where tissue is composed by myelin, intra/extra-cellular water, and CSF [47, 72]. With this model, volume fractions of this components were anatomically meaningful, although biased for myelin. The relaxometry pa-

rameters for the three compartments were reliable in the regions where the volume fraction of the corresponding compartments were significant.

This work faced several limitations. First, the acquisition and reconstruction of QTI are design for a single compartment, introducing a bias in the volume fraction estimates. Second, the Gaussian assumption of noise might not be accurate since the Low Rank (LR) regularization affects the distribution of noise. And third, Π4U is computationally expensive, running from 1 to 2 days for a single slice. Further work is necessary to overcome this limitations and validate this approach in healthy subjects as well as patients.

BIBLIOGRAPHY

[1] Daniel C Alexander, Penny L Hubbard, Matt G Hall, Elizabeth A Moore, Maurice Ptito, Geoff J M Parker, and Tim B Dyrby. "Orientationally invariant indices of axon diameter and density from diffusion MRI." In: *Neuroimage* 52 (2010), pp. 1374–1389.

[2] Y Assaf and PJ Basser. "Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain." In: *Neuroimage* 27 (2005), pp. 48–58.

[3] Y Assaf, T Blumenfeld-Katzir, Y Yovel, and Peter J Basser. "AxCaliber: a method for measuring axon diameter distribution from diffusion MRI." In: *Magn. Reson. Med.* 59 (2008), pp. 1347–1354.

[4] P J Basser, J Mattiello, and D LeBihan. "MR diffusion tensor spectroscopy and imaging." In: *Biophys. J.* 66.1 (1994), pp. 259–267.

[5] Dan Benjamini and Peter J. Basser. "Use of marginal distributions constrained optimization (MADCO) for accelerated 2D MRI relaxometry and diffusometry." In: *J. Magn. Reson.* 271 (2016), pp. 40–45.

[6] Michael W. Berry, Murray Browne, Amy N. Langville, V. Paul Pauca, and Robert J. Plemmons. "Algorithms and applications for approximate nonnegative matrix factorization." In: *Comput. Stat. Data Anal.* 52.1 (2007), pp. 155–173.

[7] Wolfgang Betz, Iason Papaioannou, and Daniel Straub. "Transitional Markov Chain Monte Carlo : Observations and Improvements." In: *Journal of Engineering Mechanics* 142.5 (2016), pp. 1–10. ISSN: 0733-9399. DOI: 10.1061/(ASCE)EM.1943-7889.0001066..

[8] Sampada Bhave, Sajan Goud Lingala, Casey P. Johnson, Vincent A. Magnotta, and Mathews Jacob. "Accelerated whole-brain multi-parameter mapping using blind compressed sensing." In: *Magn. Reson. Med.* 75.3 (2016), pp. 1175–1186. ISSN: 15222594. DOI: 10.1002/mrm.25722.

[9] Denis Le Bihan and E. Breton. "Imagerie de diffusion in-vivo par résonance magnétique nucléaire." In: *Comptes-Rendus l'Académie des Sci.* 93.5 (1985), pp. 27–34. URL: https://hal.archives-ouvertes.fr/hal-00350090.

[10] F. Bloch. "Nuclear induction." In: *Physical Review* 70 (1946), pp. 460–474. ISSN: 0031899X. DOI: 10.1103/PhysRev.70.460.

[11] Pau Bofill and Michael Zibulevsky. "Underdetermined blind source separation using sparse representations." In: *Signal Processing* 81.11 (2001), pp. 2353–2362.

[12] Thomas F. Budinger, Stephen E. Derenzo, Grant T. Gullberg, William L. Greenberg, and Ronald H. Huesman. "Emission computer assisted tomography with single-photon and positron annihilation photon emitters." In: *J. Comput. Assist. Tomogr.* 1.1 (1977), pp. 131–145. ISSN: 03638715. DOI: 10.1097/00004728-197701000-00015.

[13] Paul T Callaghan, Christoph H Arns, Petrik Galvosas, Mark W Hunter, Ying Qiao, and Kate E Washburn. "Recent Fourier and Laplace perspectives for multidimensional NMR in porous media." In: *Magn. Reson. Imaging* 25 (2007), pp. 441–444.

[14] H. Y. Carr and E. M. Purcell. "Effects of diffusion on free precession in nuclear magnetic resonance experiments." In: *Phys. Rev.* 94.3 (1954), pp. 630–638. ISSN: 0031899X. DOI: 10.1103/PhysRev.94.630. arXiv: arXiv:1011.1669v3.

[15] H Carr. "Steady-State Free Precession in Nuclear Magnetic Resonance." In: *Phys. Rev.* 112.5 (1958), pp. 1693–1701. ISSN: 0031-899X. DOI: 10.1103/PhysRev.112.1693. URL: http://link.aps.org/doi/10.1103/PhysRev.112.1693{\%}5Cnpapers2://publication/doi/10.1103/PhysRev.112.1693.

[16] Jinaye Ching and Yi-Chu Chen. "Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating , Model Class Selection, and Model Averaging." In: *Journal of Engineering Mechanics* 7.133 (2007), pp. 816–832. DOI: 10.1061/(ASCE)0733-9399(2007)133.

[17] R. Damadian. "Tumor Detection by Nuclear Magnetic Resonance." In: *Science (80-.).* 171.3976 (1971), pp. 1151–1153. ISSN: 0036-8075. DOI: 10.1126/science.171.3976.1151. arXiv: 1151-1153. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.171.3976.1151.

[18] Sean C L Deoni, Lucy Matthews, and Shannon 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." In: *Magnetic Resonance in Medicine* 70.1 (2013), pp. 147–154. ISSN: 07403194. DOI: 10.1002/mrm.24429. arXiv: NIHMS150003.

[19] Sean C L Deoni, Brian K Rutt, Tarunya Arun, Carlo Pierpaoli, and Derek K Jones. "Gleaning multicomponent T_1 and T_2 information from steady-state imaging data." In: *Magn. Reson. Med.* 60 (2008), pp. 1372–1387.

[20] Mark D Does and John C Gore. "Compartmental study of diffusion and relaxation measured in vivo in normal and ischemic rat brain and trigeminal nerve." In: *Magn. Reson. Med.* 43.6 (2000), pp. 837–44.

[21] Mariya Doneva, Peter Börnert, Holger Eggers, Christian Stehning, Julien Sénégas, and Alfred Mertins. "Compressed sensing reconstruction for magnetic resonance parameter mapping." In: *Magn. Reson. Med.* 64.4 (2010), pp. 1114–1120. ISSN: 07403194. DOI: 10.1002/mrm.22483.

[22] A. Einstein. *Investigations on the Theory of the Brownian Movement*. 1905. DOI: 10.1002/andp.19053220806. arXiv: 9811186 [arXiv:cond-mat].

[23] Uran Ferizi, Torben Schneider, Eleftheria Panagiotaki, Gemma Nedjati-Gilani, Hui Zhang, Claudia A M Wheeler-Kingshott, and Daniel C Alexander. "A ranking of diffusion MRI compartment models with in vivo human brain data." In: *Magn. Reson. Med.* 72.6 (2014), pp. 1785–92.

[24] A. N. Garroway, P. K. Grannell, and P. Mansfield. *Image formation in NMR by a selective irradiative process*. 1974. DOI: 10.1088/0022-3719/7/24/006.

[25] Pedro A. Gómez, Cagdas Ulas, Jonathan I. Sperl, Tim Sprenger, Miguel Molina-Romero, Marion I. Menzel, and Bjoern H. Menze. "Learning a Spatiotemporal Dictionary for Magnetic Resonance Fingerprinting with Compressed Sensing." In: *Patch-Based Techniques in Medical Imaging*. Cham: Springer International Publishing, 2015, pp. 112–119.

[26] Pedro A Gómez, Miguel Molina-Romero, Guido Bounincontri, Jonathan I Sperl, Jones Derek K, Marion I Menzel, and Bjoern H Menze. "Simultaneous Parameter Mapping, Modality Synthesis, and Anatomical Labeling of the Brain with MR Fingerprinting." In: *MICCAI* (2016).

[27] Pedro A. Gómez, Guido Buonincontri, Miguel Molina-Romero, Jonathan I Sperl, Marion I Menzel, and Bjoern H Menze. *Accelerated parameter mapping with compressed sensing: an alternative to MR Fingerprinting*. 2017.

[28] A. Haase, J. Frahm, D. Matthaei, W. Hanicke, and K. D. Merboldt. "FLASH imaging. Rapid NMR imaging using low flip-angle pulses." In: *J. Magn. Reson.* 67.2 (1986), pp. 258–266. ISSN: 00222364. DOI: 10.1016/0022-2364(86)90433-6. URL: http://linkinghub.elsevier.com/retrieve/pii/0022236486904336.

[29] P.E. Hadjidoukas, P. Angelikopoulos, C. Papadimitriou, and P. Koumoutsakos. "Π4U: A high performance computing framework for Bayesian uncertainty quantification of complex models." In: *J. Comput. Phys.* 284 (2015), pp. 1–21. ISSN: 00219991.

DOI: 10.1016/j.jcp.2014.12.006. URL: http://www.sciencedirect.com/science/article/pii/S0021999114008134.

[30] G. N. Hounsfield. "Computerized transverse axial scanning (tomography): I. Description of system." In: *Br. J. Radiol.* 46.552 (1973), pp. 1016–1022. ISSN: 00071285. DOI: 10.1259/0007-1285-46-552-1016.

[31] Andrew R Hoy, Cheng Guan Koay, Steven R Kecskemeti, and Andrew L Alexander. "Optimization of a Free Water Elimination Two-Compartment Model for Diffusion Tensor Imaging." In: *Neuroimage* 103 (2014), pp. 323–333. DOI: 10.1016/j.neuroimage.2014.09.053.

[32] P.O. Hoyer. *Non-negative sparse coding*. 2002. DOI: 10.1109/NNSP.2002.1030067. URL: http://ieeexplore.ieee.org/document/1030067/.

[33] Junzhou Huang, Chen Chen, and Leon Axel. "Fast Multi-contrast MRI Reconstruction." In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*. Ed. by Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Mori. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 281–288.

[34] Aapo Hyvarinen and Erkki Oja. "Independent Component Analysis : A Tutorial." In: *Neural Networks* 1 (1999), pp. 1–30.

[35] Ian Jolliffe. *Principal component analysis*. New York: Springer Verlag, 2002.

[36] D.K. Jones. *Diffusion MRI*. Oxford University Press, 2010. ISBN: 9780199708703.

[37] Derek K. Jones and Mara Cercignani. *Twenty-five pitfalls in the analysis of diffusion MRI data*. 2010. DOI: 10.1002/nbm.1543.

[38] Jörg Kärger. "Der Einfluss der Zweibereichdiffusion auf die Spinechodaempfung unter Beruecksichtigung der Relaxation bei Messungen mit der Methode der gepulsten Feldgradienten." In: *Annalen der Physik* 7.27 (1971), pp. 107–109.

[39] I Kay and R M Henkelman. "Practical Implementation and Optimization of One-Shot TI Imaging." In: 424 (1991), pp. 414–424. URL: http://onlinelibrary.wiley.com/store/10.1002/mrm.1910220249/asset/1910220249{_}ftp.pdf?v=1{\&}t=j08pdf7y{\&}s=5bf2d02a030b3f4ec698509a3c0ff7f03f63cf58{\&}systemMessage=Wiley+Online+Library+will+be+unavailable+on+Saturday+25th+March+from+07{\%}253A00+GMT+{\%}252F+03{\%}253A00.

[40] Daeun Kim, Eamon K Doyle, Jessica L Wisnowski, Joong Hee Kim, and Justin P Haldar. "Diffusion-Relaxation Correlation Spectroscopic Imaging: A Multidimensional Approach for Probing Microstructure." In: *Magn. Reson. Med.* (2017). DOI: 10.1002/mrm.26629.

[41] L L Latour, K Svoboda, P P Mitra, and C H Sotak. "Time-dependent diffusion of water in a biological model system." In: *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994), pp. 1229–1233.

[42] P. C. Lauterbur. "Image formation by induced local interactions. Examples Employing Nuclear Magnetic Resonance." In: *Nature* 242 (1973), pp. 190–191. ISSN: 0028-0836. DOI: 10.1038/242190a0.

[43] Dd Lee and Hs Seung. "Algorithms for non-negative matrix factorization." In: *Adv. Neural Inf. Process. Syst.* 1. 2001, pp. 556–562. ISBN: 9781424418206. DOI: 10.1109/IJCNN.2008.4634046. eprint: 0408058v1 (arXiv:cs). URL: http://papers.nips.cc/paper/1861-algorithms-for-non-negative-matrix-factorization.

[44] RA Lerski and JD De Certaines. "II. Performance assessment and quality control in MRI by Eurospin test objects and protocols." In: *Magnetic resonance imaging* 11.6 (1993), pp. 817–833.

[45] Dan Ma, Vikas Gulani, Nicole Seiberlich, Kecheng Liu, Jeffrey L Sunshine, Jeffrey L Duerk, and Mark A Griswold. "Magnetic resonance fingerprinting." In: *Nature* 495.7440 (2013), pp. 187–92. ISSN: 1476-4687. DOI: 10.1038/nature11971. arXiv: NIHMS150003. URL: http://www.nature.com/nature/journal/v495/n7440/full/nature11971.html?WT.ec{_}id=NATURE-20130314.

[46] Alex L. MacKay and Cornelia Laule. "Magnetic Resonance of Myelin Water: Anin vivo Marker for Myelin." In: *Brain Plasticity* January (2016), pp. 1–21.

[47] Alex L MacKay, Cornelia Laule, Irene Vavasour, Thorarin Bjarnason, Shannon Kolind, and Burkhard Mädler. "Insights into brain microstructure from the T_2 distribution." In: *Magn. Reson. Med.* 24.4 (2006), pp. 515–25.

[48] *Magnetic Resonance Imaging*. Wiley-Blackwell, 2014. ISBN: 9781118633953. DOI: 10.1002/9781118633953. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118633953. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118633953.

[49] P. Mansfield. "Multi-planar image formation using NMR spin echoes." In: *J. Phys. C Solid State Phys.* 10.3 (1977). ISSN: 00223719. DOI: 10.1088/0022-3719/10/3/004.

[50] M. Molina-Romero, Pedro A Gómez, Jonathan I Sperl, Derek K Jones, Marion I Menzel, and Bjoern H Menze. "Tissue microstructure characterisation through relaxometry and diffusion MRI using sparse component analysis." In: *ISMRM Workshop on Breaking the Barriers of Diffusion MRI, Lisbon*. 2016, p. 17.

[51] M. Molina-Romero, Pedro A Gómez, Jonathan I Sperl, Andrew J Stewart, Derek K Jones, Marion I Menzel, and Bjoern H Menze. "Theory, validation and application of blind source separation to diffusion MRI for tissue characterisation and partial volume correction." In: *Proc Intl Soc Mag Reson Med*. 2017, p. 3462.

[52] Migue Molina-Romero, Pedro A Gómez, Jonathan I Sperl, Michael Czisch, Philipp G Sämann, Marion I Menzel, and Bjoern H Menze. "A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization." In: *Magnetic Resonance in Medicine* 0.0 (2018), pp. 1–18. DOI: 10.1002/mrm.27181.

[53] Miguel Molina-Romero, Pedro A Gómez, Shadi Albarqouni, Jonathan I Sperl, Marion I Menzel, and Bjoern H Menze. "Deep learning with synthetic data for free water elimination in diffusion MRI." In: *Proc Intl Soc Mag Reson Med*. 2018.

[54] Markus Nilsson, Jimmy Lätt, Danielle van Westen, Sara Brockstedt, Samo Lasič, Freddy Ståhlberg, and Daniel Topgaard. "Noninvasive mapping of water diffusional exchange in the human brain using filter-exchange imaging." In: *Magn. Reson. Med.* 69.6 (2013), pp. 1573–81. ISSN: 1522-2594. DOI: 10.1002/mrm.24395. URL: http://www.ncbi.nlm.nih.gov/pubmed/22837019.

[55] Paula Orihuela Martinez-Costa, Miguel Molina-Romero, Pedro A Gómez, and Bjoern H Menze. "Brain Microstructure Quantification from Transient-State Magnetic Resonance Imaging." MA thesis. Germany: Technical University of Munich, 2017.

[56] P Paatero and U Tapper. "Positive Matrix Factorization - A Nonnegative Factor Model With Optimal Utilization of Error-Estimates of Data Values." In: *Environmetrics* 5.2 (1994), pp. 111–126. ISSN: 11804009. DOI: 10.1002/env.3170050203.

[57] Ofer Pasternak, Nir Sochen, Yaniv Gur, Nathan Intrator, and Yaniv Assaf. "Free Water Elimination and Mapping from Diffusion MRI." In: *Magn. Reson. Med.* 62 (2009), pp. 717–730. DOI: 10.1002/mrm.22055.

[58] S Peled, D G Cory, S A Raymond, D A Kirschner, and F A Jolesz. "Water diffusion, T_2, and compartmentation in frog sciatic nerve." In: *Magn. Reson. Med.* 42.5 (1999), pp. 911–8. ISSN: 0740-3194.

[59] C. Pierpaoli and D.K. Jones. "Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model." In: *ISMRM, Kyoto*. 2004, p. 1215.

[60] E M Purcell, H C Torrey, and R B Pound. "Resonance Absorption by Nuclear Magnetic Moments in a Solid." In: *Phys Rev* (1946), pp. 37–38. ISSN: 0031-899X. DOI: 10.1103/PhysRev.69.37.

[61] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. "A New Method of Measuring Nuclear Magnetic Moment." In: *Physical Review* 53 (Feb. 1938), pp. 318–318. DOI: 10.1103/PhysRev.53.318.

[62] Saiprasad Ravishankar and Yoram Bresler. "l_0 Sparsifying Transform Learning With Efficient Optimal Updates and Convergence Guarantees." In: *IEEE Trans. Signal Process.* 63.9 (2015), pp. 2389–2404. ISSN: 1053587X. DOI: 10.1109/TSP.2015.2405503. arXiv: arXiv:1501.02859v1.

[63] Greg J Stanisz, Aaron Szafer, Graham A Wright, and R Mark Henkelman. "An Analytical Model of Restricted Diffusion in Bovine Optic Nerve." In: *Magn. Reson. Med.* 37 (1997), pp. 103–111.

[64] E. O. Stejskal and J. E. Tanner. "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient." In: *J. Chem. Phys.* 42.1 (1965), pp. 288–292.

[65] a Szafer, J Zhong, and J C Gore. "Theoretical model for water diffusion in tissues." In: *Magn. Reson. Med.* 33.5 (1995), pp. 697–712.

[66] Michel M. Ter-Pogossian, Michael E. Phelps, Edward J. Hoffman, and Nizar A. Mullani. "A Positron-Emission Transaxial Tomograph for Nuclear Imaging (PETT)." In: *Radiology* 114.1 (1975), pp. 89–98. ISSN: 0033-8419. DOI: 10.1148/114.1.89. URL: http://pubs.rsna.org/doi/10.1148/114.1.89.

[67] H. C. Torrey. "Bloch equations with diffusion terms." In: *Phys. Rev.* 104.3 (1956), pp. 563–565. ISSN: 0031899X. DOI: 10.1103/PhysRev.104.563.

[68] D. C. Van Essen et al. *The Human Connectome Project: A data acquisition perspective*. 2012. DOI: 10.1016/j.neuroimage.2012.02.018. arXiv: NIHMS150003.

[69] Julia V. Velikina, Andrew L. Alexander, and Alexey Samsonov. "Accelerating MR parameter mapping using sparsity-promoting regularization in parametric dimension." In: *Magn. Reson. Med.* 70.5 (2013), pp. 1263–1273. ISSN: 07403194. DOI: 10.1002/mrm.24577. arXiv: NIHMS150003.

[70] M. Weigel, S. Schwenk, V. G. Kiselev, K. Scheffler, and J. Hennig. "Extended phase graphs with anisotropic diffusion." In: *Journal of Magnetic Resonance* 205.2 (2010), pp. 276–285. ISSN: 10907807. DOI: 10.1016/j.jmr.2010.05.011. URL: http://dx.doi.org/10.1016/j.jmr.2010.05.011.

[71] Matthias Weigel. "Extended phase graphs: Dephasing, RF pulses, and echoes - Pure and simple." In: *Journal of Magnetic Resonance Imaging* 41.2 (2015), pp. 266–295. ISSN: 15222586. DOI: 10.1002/jmri.24619.

[72] Kenneth P Whittall, Alex L Mackay, Douglas A Graeb, Robert A Nugent, David K B Li, and Donald W Paty. "In vivo measurement of T_2 distributions and water contents in normal human brain." In: *Magn. Reson. Med.* 37.1 (1997), pp. 34–43.

[73] Xianchuan. Yu, Dan Hu, and Xu Jindong. *Blind source separation: theory and applications*. Science Press, 2014, p. 416.

[74] H Zhang, T Schneider, Claudia A M Wheeler-Kingshott, and Daniel C Alexander. "NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain." In: *Neuroimage* 61 (2012), pp. 1000–1016.

[75] Li Zhao, Xue Feng, and Craig H. Meyer. "Direct and accelerated parameter mapping using the unscented Kalman filter." In: *Magn. Reson. Med.* 75.5 (2016), pp. 1989–1999. ISSN: 15222594. DOI: 10.1002/mrm.25796.