Advances in MRI-guided radiotherapy

Acquisition, analysis and evaluation methods

Rahbek, Sofie

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

Sofie Rahbek

ADVANCES IN MRI-GUIDED RADIOTHERAPY: ACQUISITION, ANALYSIS AND EVALUATION METHODS

PhD Thesis, June 2022

DTU Health Tech
Department of Health Technology

Region of Southern Denmark
OUH Odense University Hospital

Advances in MRI-guided radiotherapy: Acquisition, analysis and evaluation methods

PhD dissertation by:
Sofie Rahbek

Main Supervisor:
Lars G. Hanson, Associate Professor, DTU Health Technology

Co-supervisor:
Faisal Mahmood, Associate Professor, University of Southern Denmark, Department of Clinical Research
Kristoffer H. Madsen, Associate Professor, DTU Applied Mathematics and Computer Science

DTU Health Technology
Magnetic Resonance Section
Technical University of Denmark
Ørsteds Plads 349, 1st floor
2800 Kgs. Lyngby
Denmark
Tel: +45 4525 3518

Project period:	August 2018 – June 2022
Class:	Public
Edition:	1st Edition
Remarks:	This report is submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering at the Technical University of Denmark.

ABSTRACT

In radiotherapy the main goal is to kill tumor cells with ionizing radiation, typically high-energy photons, while sparing the surrounding healthy tissue. The precision of the radiation dose delivery is crucial for the efficacy of radiotherapy and imaging is utilized for tumor and normal tissue delineation and for guidance at the moment of treatment, leading to the concept *image-guided radiotherapy* (IGRT). The most recent advancement here is the hybrid MRI – linear accelerator system (MR-linac), which enables high soft-tissue contrast images for adapting the dose plan to the anatomy-of-the-day.

The MR-linac also makes it feasible to obtain daily advanced MRI which may potentially be used for revealing functional features of the tumor informative of local radio-sensitivities in the tumor and predictive of the overall treatment outcome. Utilizing MRI for such purposes requires high image quality and advanced data processing that involves analysis of multi-contrast measurements, such as multi-echo T_2-weighted measurements, to extract characteristics of the tumor tissue. A model-based method is most commonly used, but the resulting quantitative maps are often prone to partial volume effects and may be biased or uninformative if the chosen model does not properly fit the data. Here, *monotonous slope non-negative matrix factorization* (msNMF) is proposed as a novel data-driven method. This extended version of the NMF decomposes the data under monotony constraints that fit many types of MRI data. A demonstration of the method showed its ability to extract interpretable components related to the underlying tissue micro-structure, and applications was also exemplified by estimation of edema water fractions in spinal cord white matter.

MR-linacs enable longitudinal imaging data series, which requires dedicated handling of the covarying time-resolved measurements to investigate the tumor dynamics during the course of fractionated radiotherapy. This work also proposes a prediction framework as a tool to search for biomarkers using longitudinal MRI data. The framework relies on an initial data-driven decomposition and includes fitting over time to capture therapy-induced tumor changes that may be predictive of the outcome. Its feasibility was demonstrated using example datasets and the msNMF for decomposition, and results indicated a value of early T_2-relaxation changes for predicting tumor response.

Diffusion-weighted imaging (DWI) is an interesting multi-contrast technique due to the lower diffusivity of many tumor types that is altered when the ionizing radiation induces micro-structural changes. Unfortunately, standard DWI suffers from geometric distortions, incompatible with radiotherapy purposes which includes both target delineation and response prediction. To address this, the current work includes optimization of the single-shot split acquisition for fast spin-echo (SPLICE) sequence. Contrary to standard diffusion-weighted echo-planar imaging, the SPLICE sequence is based on a fast spin-echo readout and results in geometrically robust images, but with a relatively poor voxel shape due to signal modulation during readout. The

suggested optimization method maximizes the signal-to-noise ratio (SNR) for a controlled point-spread-function by varying the refocusing flip angles. A clear SNR gain, which also improved the accuracy of apparent diffusion coefficient (ADC) estimates, was seen for a healthy subject brain.

In summary, the acquisition and analysis strategies developed during this research project may provide directions for future radiotherapy studies and can advance the usage of MRI for both treatment planning and evaluation.

RESUMÉ

I strålebehandling er hovedmålet at dræbe tumorceller med ioniserende stråling (typisk højenergi fotoner), og samtidigt sikre at det omgivende raske væv ikke påvirkes unødigt. Præcisionen af dosisplanlægningen er afgørende for effektiviteten af strålebehandlingen, og billeddiagnostik bruges til at optegne tumor og normalt væv, samt til vejledning under selve behandlingen, hvilket fører til konceptet *image-guided radiotherapy* (IGRT). Det seneste fremskridt her er et hybridt system bestående af en MR-skanner og en lineær accelerator (MR-linac), som kan optage billeder med høj kontrast i blødt væv med henblik på at adaptere dosisplanen til patientens anatomi på dagen.

MR-linac systemer gør det også muligt at opnå avancerede MR-skanninger på daglig basis, som kan anvendes til at måle egenskaber ved tumorvævet, der kan være informative for den lokale strålefølsomhed i tumoren og muligvis bruges til at forudsige det samlede behandlingsresultat. Anvendelse af MRI til sådanne formål kræver høj billedkvalitet og avanceret databehandling, der involverer analyse af multikontrast målinger såsom diffusionsvægtet MRI, til at udtrække tumorvævets karakteristika. En modelbaseret metode er mest almindeligt anvendt, men de resulterende kvantitative billeder er ofte begrænsede af det faktum at billedvoxels kan indeholde information fra flere vævstyper (eng.: *partial volume effects*) og kan være misvisende, hvis den valgte model ikke passer tilstrækkeligt til målingerne. Her foreslås *monotonous slope non-negative matrix factorization* (msNMF) som en ny, datadrevet metode. Denne udvidelse af standard NMF dekomponerer data under hensyn til at signalkomponenter og deres afledte skal være monotone, hvilket er tilfældet for mange typer af MR-målinger. En demonstration af metoden viste dens evne til at udtrække realistiske og tolkelige komponenter relateret til den underliggende vævsmikrostruktur og kunne f.eks. bruges til at estimere den relative mængde ødem i hvid substans i rygmarven.

Under et fraktioneret radioterapiforløb muliggør MR-linacen optagelse af longitudinelle dataserier, hvilket kræver dedikeret databehandling til håndtering af de tidsopløste, men kovarierende målinger for at undersøge tumordynamikken i løbet af behandlingen. I dette projekt er der udviklet en fremgangmåde til prædiktionsanalyse i den forbindelse, og den foreslås som et værktøj til at søge efter biomarkører vha. longitudinelle MR-data. Analysen bygger på en indledende datadrevet dekomposition, og inkluderer et fit henover tid for at identificere terapi-inducerede tumorændringer, som kan være prædiktive for behandlingsresultatet. Analysemetoden blev demonstreret i to datasæt, hvor den nyudviklede msNMF blev brugt til dekompositionen. Resultater indikerede at tidlige T_2-relaksationsændringer kan være prædiktive for tumorresponset.

Diffusionsvægtet MRI er en interessant multi-kontrast teknik fordi der er en lav diffusivitet i mange tumortyper, og denne diffusivitet ændres, når den ioniserende stråling inducerer mikrostrukturelle ændringer. Desværre kan billeder fra standard diffusionsvægtet MRI (udlæst med *echo planar imaging*) være udsat for geometrisk forvrængning, hvilket gør dem uegnede til

formål indenfor strålebehandling, der omfatter præcis tumoroptegning og evaluering af respons. For at adressere denne problematik, blev der i projektet optimeret en enkelt-excitation split-ekko sekvens egnet til diffusionvægtet MRI (SPLICE). SPLICE sekvensen er baseret på en hurtig spin-ekko udlæsning og resulterer i geometrisk robuste billeder. Dog kan signal-modulation under udlæsningen af data medføre at SPLICE billederne fremstår slørede. Den foreslåede optimeringsmetode maksimerer signal-støj-forholdet (SNR) for en kontrolleret punktspredningsfunktion ved at variere refokuseringsvinklerne i ekkotoget, og forbedrer derved billedkvaliteten. En testskanning af en rask hjerne viste tydelige SNR forbedringer, hvilket medførte en mere nøjagtig bestemmelse af diffusionskoefficienter.

Sammenfattende kan optage- og analysestrategierne udviklet i løbet af dette forskningsprojekt forsyne fremtidige strålebehandlingsstudier med retningslinjer og værktøjer, der kan forbedre og udvikle brugen af MRI til både terapiplanlægning og -evaluering.

PREFACE

This thesis has been written as partial fulfillment of the requirements for obtaining the degree of Doctor of Philosophy in Engineering at the Technical University of Denmark.

The thesis presents research conducted from August 2018 to May 2022 with affiliation to the Magnetic Resonance Section at the Department of Health Technology at the Technical University of Denmark, Kongens Lyngby, Denmark, and to the Department of Oncology, Odense University Hospital, Odense, Denmark. Two months in 2020 were used for an external research stay at University Medical Center Utrecht (UMCU), Utrecht, The Netherlands. COVID-19 shortened the stay from a planned period of five months, but collaborations with the site continued throughout the rest of the research period.

Other research collaborations include the Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Denmark; the Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark; and the Copenhagen University Hospital Herlev-Gentofte, Denmark.

The research was collaboratively funded by the Technical University of Denmark and the Danish Cancer Society, grant number R167-A10637-17-S2. The external stay was economically supported by a grant awarded by Knud Højgaards Fond (19-02-3076).

Sofie Rahbek

CONTRIBUTIONS

JOURNAL PAPERS

- Rahbek, Sofie; Madsen, Kristoffer H.; Lundell, Henrik; Mahmood, Faisal; Hanson, Lars G. **Data-driven separation of MRI signal components for tissue characterization**. *Published in Journal of Magnetic Resonance, 2021.*
- Rahbek, Sofie; Mahmood, Faisal; Tomaszewski, Michal R.; Hanson, Lars G; Madsen, Kristoffer H. **Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI**. *Manuscript submitted.*
- Rahbek, Sofie; Schakel, Tim; Mahmood, Faisal; Madsen, Kristoffer H.; Philippens, Marielle E.P.; Hanson, Lars G. **Optimized flip angle schemes for the diffusion-weighted SPLICE sequence.** *Manuscript under revision for publication, 2022*

CONFERENCE ABSTRACTS

- Rahbek, Sofie; Madsen, Kristoffer H.; Lundell, Henrik; Mahmood, Faisal; Hanson, Lars G. **Data-driven separation of MRI signal components for tissue characterization**. *Presented at ISMRM & SMRT Annual Meeting & Exhibition, 2021*
- Rahbek, Sofie; Mahmood, Faisal; Madsen, Kristoffer H; Hanson, Lars G. **Decomposition-based framework for prediction of radiotherapy response from longitudinal DW-MRI data.** *Presented at Virtual 8th MR in RT Symposium, 2021*
- Rahbek, Sofie; Schakel, Tim; Mahmood, Faisal; Madsen, Kristoffer H.; Philippens, Marielle E.P.; Hanson, Lars G. **Flip-angle optimization for the diffusion-weighted SPLICE sequence for applications in brain imaging.** *Presented at ISMRM & SMRT Annual Meeting & Exhibition, 2021*

ACKNOWLEDGMENTS

I would first and foremost like to thank my three supervisors for their priceless collaboration and contribution throughout this research project. With their different expertise they have complemented each other well and given me a complete palette of helpful advise and guidance. My main supervisor, Associate professor Lars G. Hanson, has contributed with his profound knowledge of MR methodology. With his MRI and research experience he was always able to raise relevant perspectives that increased the quality of the research remarkably. My second supervisor, Associate professor Faisal Mahmood, has contributed with important knowledge within radiation oncology. He has a finger on the pulse of the newest trends and developments within the field and has ensured a clinical perspective when things became theoretically heavy. My third supervisor, Associate professor Kristoffer H. Madsen, contributed with his extensive knowledge of computational modelling and analysis. If you are in doubt of any method or concept within this area, he is the one to go to. I am really grateful that he joined the project as the research turned out to involve computer science to a higher degree. My weekly meetings with Lars, Faisal and Kristoffer have included fruitful discussions, academic sparring, and high-class joking, and I always left the meetings motivated.

Then, I would like to thank Associate professor Marielle Philippens and Postdoctoral researcher Tim Schakel who acted as "supervisors" during my external stay at University Medical Center Utrecht (UMCU). Though the external stay only lasted two months due to COVID-19, I gained a lot from the visit, and it was inspiring to work together with two skilled researches as Tim and Marielle. It was also inspiring to work in a research environment at a hospital, where facilities made fast experimental testing possible. I was lucky to be placed in a group of other young researchers, which were all very welcoming and socially engaged, so a thanks to them for making my visit even better. I am glad that it was possible to continue the collaboration with Tim and Marielle after the research stay and I really appreciate that Tim took his time to record data, share software, and guide us during data acquisition at Odense University hospital.

Other researchers who deserves a big thanks for their collaboration and contribution with key points, discussions and sharing of data are: Senior researcher Henrik Lundell, Postdoctoral researcher Yi He and Postdoctoral researcher Vanessa Wiggerman from Danish Research Centre for Magnetic Resonance (DRCMR); Professor Mark D. Does and Assistant Professor Kevin D. Harkins from Vanderbilt University Medical Center; and Dr. Michal R. Tomaszewski and Professor Robert Gillies from Moffitt Cancer center.

I would also like thank my local coworkers, the entire group at the Magnetic Resonance Section at DTU Health Technology. They have all contributed in creating an inspiring research environment and a nice social atmosphere with silly discussions over lunch and funny social events such as a curling tournament and laser tag. A special thanks goes to my office mate, Postdoctoral researcher Rie Beck Olin, who has always taken her time to help me when I needed

advise or sparring.

Finally, I would like to thank my family for always supporting me in my work and for listening to my nerdy talk on MRI and to my complaints when something in the project was difficult or not going my way.

TABLE OF CONTENTS

LIST OF FIGURES

NOMENCLATURE

ADC	apparent diffusion coefficient
ANLS	alternating non-negative least squares
AUC	area under the curve
CPMG	Carr-Purcel-Meiboom-Gill
CT	computed tomography
CV	cross validation
DCE	dynamic contrast enhanced
DTI	diffusion tensor imaging
DW	diffusion-weighted
DWI	diffusion-weighted imaging
EPG	extended phase graph
EPI	echo-planar imaging
ESP	echo spacing
ETL	echo train length
EWF	edema water fraction
HALS	hierarchical alternating least squares
IGRT	image-guided radiotherapy
IMRT	intensity modulated radiotherapy
IVIM	intra-voxel incoherent motion
JEMRIS	Jülich extensible MRI simulator
linac	linear accelerator
MLR	multiple linear regression
MR	magnetic resonance
MRI	magnetic resonance imaging

msNMF	monotonous slope non-negative matrix factorization
NMF	non-negative matrix factorization
NNLS	non-negative least squares
PCA	principal component analysis
PET	positron emission tomography
PSF	point spread function
RARE	rapid acquisition with relaxation enhancement
RF	radio frequency
RMSE	root mean squared error
ROI	region of interest
SE	spin echo
SNR	signal-to-noise ratio
SoS	sum-of-squares
SPLICE	split acquisition for fast spin-echo
STE	stimulated echo
TE	echo time
TR	repetition time

1 INTRODUCTION

Radiation therapy is a standard of care for treating malignancies, and the primary goal is to irradiate and eliminate tumor cells while minimizing healthy tissue toxicity. In modern external beam radiotherapy, this is achieved through advanced computer-based treatment planning which allows highly conformal radiation dose distributions calculated directly on a 3D CT scan of the patient. In the treatment phase, computed tomography (CT) based imaging systems integrated into the linear accelerator (linac) are used to verify patient positioning and do adjustments if required, prior to delivery of the beam. This is known as *image-guided radiotherapy* (IGRT) [1, 2].

A potentially important improvement in IGRT has been introduced with *MR-linacs*, hybrid magnetic resonance imaging – linear accelerator systems [3, 4]. Magnetic resonance imaging (MRI) not only delivers anatomical images with high soft-tissue contrast for localization of tumor and adjacent organs, but also includes advanced multi-contrast imaging techniques for assessment of functional and micro-structural characteristics of the tissue. Increased knowledge of the heterogeneous tumor biology supports the concept of *biological target volumes* [5] which accounts for variation in radiation sensitivity across the tumor. This information can potentially be used in the treatment planning phase, for *biologically-guided* dose escalation to tumor sub-regions, using standard MR scanners, and for daily biological adaptation in the treatment phase, using MR-linacs [5]. Additionally, revealing tissue characteristics non-invasively increases the possibility to find prognostic and predictive biomarkers. The gross tumor volume change, which is the current standard of radiotherapy response evaluation [6], happens at a much longer timescale (weeks or months) than the cellular reaction to ionizing radiation which occurs already within minutes/hours [7]. An early prediction of a tumor's response to therapy increases the chance to timely adjust the therapy plan and also to avoid unnecessary irradiation. MRI-guided radiotherapy is therefore considered important in the development of personalized oncology [8, 9].

Ionizing radiation from radiotherapy induces cell death lowering the cell density of the irradiated region[10]. As this alters the local microscopic water mobility, diffusion-weighted imaging (DWI) is considered a candidate for probing tumor micro-structure changes that correlate with the radiation response. Many investigators have studied DWI-derived parameters for their potential as biomarkers of outcome [9, 11]. Although several of them showed some capacity to stratify tumor response, some findings are contradictory [12]. Challenges that potentially cause the ambiguous results are related to both data analysis and data quality, which are compromised for diffusion-weighted measurements compared to anatomical MRI. Low signal-to-noise ratio (SNR) data in combination with geometrical distortion are typically encountered. Moreover, model-based analysis which typically constitutes standard DWI processing, may result in biased or misleading estimates to characterize the diffusion in a voxel, and robust quantification can be difficult.

Apart from DWI, many other MRI techniques may provide useful information about the tumor biology and structure, including measurements of perfusion, relaxation, metabolism, and

hypoxia [13]. As for DWI processing, the analysis strategies most often consist of fitting the multi-contrast MRI measurements to a predefined model [14, 15]. Some models contain multiple parameters in the attempt to reproduce the measured signal in a voxel more accurately, but the demand for high data quality increases with the model complexity [16], and in any case, the sensitivity to noise and model imperfections is inevitable. Thus, there is reason to challenge model-based strategies for analyzing MRI measurements and explore data-driven methods as an alternative for discovering relevant tissue features.

The clinical introduction of MR-linacs paves the way for a large increase in the number of longitudinal MRI studies with frequent or even daily measurements during the course of therapy. Therefore, another aspect of analyses strategies for multi-contrast MRI is the handling of multiple time-resolved measurements. Among the few current longitudinal MRI tumor studies, most of them treat each measurement independently in their analysis and ignore the potentially interesting temporal variation and trend of the dynamic tumor biology. There is thus a need for new strategies to analyze time-resolved MRI data that utilizes the covariance between measurements. Preferably, an analysis framework should be applicable for a broad range of MRI parameters since the search for strong imaging biomarkers for radiotherapy purposes is still ongoing.

The applicability of DWI, and of MRI in general, for delineation and mapping of tumor heterogeneity depends on the image quality; a high resolution and the best possible SNR is desired given the fast imaging technique required to avoid effects of movement. Additionally, artifacts should be minimized, especially those compromising the geometrical precision of the images. A flawed geometry can be detrimental, especially in conformal radiotherapy [17]. Inconveniently, diffusion-weighted MRI scans are commonly obtained with echo planar imaging (EPI), a sequence that is prone to susceptibility artifacts manifesting as geometrical distortions [18]. Among alternatives, rapid acquisition with relaxation enhancement (RARE) sequences also exist for DWI. This is a fast imaging technique which is less prone to the mentioned type of artifact. However, RARE images are typically blurred due to signal modulation from T_2-weighted relaxation during readout. If the image quality, in terms of SNR and voxel-shape, of a RARE-based diffusion-weighted sequence could be improved, it may become a relevant replacement for standard EPI-based DWI.

1.1 OBJECTIVES

Overall, MRI-guided radiotherapy is under development. Especially, the utilization of biological information from functional MRI techniques need to be explored as it holds potential for higher degree of individualized radiotherapy for patients. This project is focused at both data acquisition and analysis, and specifically the thesis describes:

1. how to decompose multi-contrast MRI data for tissue characterization
2. how to analyze longitudinal MRI data and utilize the dynamic tumor information for prediction of outcome

3. how to optimize diffusion-weighted data acquisition such that images are improved for radiotherapy purposes

1.2 THESIS OVERVIEW

The dissertation first provides background knowledge of topics relevant to the research. This includes a section on radiotherapy, a section on diffusion-weighted MRI and a section on data analysis, given in Chapter 2. The description of radiotherapy is focused on the current standard methodologies and on on-going developments. The description of diffusion-weighted MRI includes details on sequence designs, simulation tools, and data processing. The description of data analysis contains a brief review of model-based analysis, data-driven analysis and supervised learning, and highlights methods utilized during the project. The following three chapters present the three sub-projects allocated to each of the above-mentioned objectives, respectively. Each of the chapters is based on a manuscript (placed in Appendix) that should be read in connection with the chapter, and each is divided into two sections: an introduction to the specific study, which describes the rationale, and a brief assessment of the methods and findings reported in the manuscript. Although only one of the manuscripts has been published to date, the three manuscripts are throughout the dissertation referred to as Paper I, Paper II, and Paper III, respectively. Subsequently, a discussion of the clinical perspectives and possible implications for future radiotherapy studies, is given in Chapter 6. There is a focus on how the three initiatives presented in Chapter 3-5 can influence and guide the data acquisition and analysis of new studies for improved MRI-guided radiotherapy. Lastly, Chapter 7 provides a summary conclusion and an outlook.

2 BACKGROUND

2.1 RADIOTHERAPY

The use of radiation therapy for treating cancer patients was initiated in 1951, and today it is a cornerstone of modern cancer therapy, functioning both as curative and palliative treatment. Standard external beam radiotherapy consists of irradiation with high energy X-rays (in the order of MeV) which inflicts both ionization of the DNA molecules directly and through oxidative stress, which can result in cell death [1]. External beam radiation therapy is delivered by a linear particle accelerator (linac), constructed to allow irradiation of the target from many angles for high degree of dose conformity.

A main challenge in radiotherapy is to deliver a sufficient radiation dose to obtain tumor control while sparing surrounding critical organs. The therapeutic window describes the separation between the probability of tumor control and compromising normal tissue for a given dose [19]. Obviously, it is desired to increase this window as much as possible, and fractionated therapy regimes are used for this. By doing so, radiobiological differences between the tumor and normal tissue are exploited. In brief, the elevated proliferation (fast mitotic cycle) of tumor cells compared to normal cells renders tumor cells more sensitive to ionising radiation. Delivering the full dose in fractions gives time for normal tissue to repair in the interval between the fractionated doses, while tumor tissue suffers more damage [1, 7]. The radiation dose and fractionation scheme depend on tumor type, location etc, but typically radiation is delivered in daily fractions of 2-3 Gy, as this is considered a reasonable compromise between sparing of normal tissue and tumor control, within a feasible overall treatment period (2-5 weeks).

Overall, the radiotherapy workflow consists of two phases, the planning phase and the treatment delivery phase. In the preparation phase, the patient is immobilized and CT scanned, covering the relevant anatomy. Relevant anatomical structures are delineated on the CT scan and the radiation dose distribution is calculated. In the delivery phase, the patient is prepared using the immobilization devices and scanned to verify the position before irradiation.

Both *intensity-modulated radiotherapy* (IMRT) and *image-guided radiotherapy* (IGRT) are important technical developments to enable *precision radiotherapy* [20]. IMRT is a conformal radiotherapy technique, meaning that the dose fall-off is very steep between the target and the surrounding tissue. This is possible with a multileaf collimator which is part of modern linacs. A multileaf collimator consists of a dynamic grid typically made of tungsten leaves which regulate and form the beam intensity at multiple discrete gantry angles or continuously while the gantry moves around the patient. The beam delivery is computer controlled and programmed according to an inverse planning algorithm. In inverse planning, an optimization algorithm calculates the best possible dose plan according to a list of desired dose prescriptions for the target tissue structures, and dose constraints for organs at risk. [21].

While IMRT facilitates a conformal dose distribution, IGRT minimizes the uncertainties of the radiation delivery and allows for reducion of planning margins and potentially an increase of tumor dosage. In modern radiotherapy, IGRT is used for daily online patient alignment in the treatment room to match the position of the target and organs at risk as closely as possible to the reference treatment plan created before start of the treatment course. Modern linacs are combined with an X-ray based imaging system, most commonly a cone beam CT (CBCT), which helps in positioning the patient and in detection of anatomical changes from planning scans. CBCT contrast relies on electron density differences and visualises lesions nicely in the lung, for example, but lacks soft-tissue contrast in most other sites where electron density differences are smaller [22]. CBCT also has the potential to be used for an online adaptive workflow where the dose plan is reoptimized for the current anatomy and tumor position, a concept known as *adaptive radiotherapy* [23]. A recent and arguable the most promising development in the context of adaptive radiotherapy is the development of a linac combined with an MRI system (MR-linac) [3, 4]. With MR-linacs, it is possible to obtain images with much greater soft-tissue contrast compared to cone beam CT, to visualize the anatomy-of-the-day [24]. Additionally, scanning during treatment delivery makes it possible to monitor the anatomy during delivery for surveillance and gated irradiation. The implementation of MR-linacs has also made it logistically feasible to obtain multiple MRI scans for advanced evaluation of the therapy (without prolonging the treatment time and inconvenience for the patient) [19].

In the planning phase, advanced imaging, including CT, MRI, and positron emission tomography (PET), can be used for accurate target delineation. With the possibility to deliver conformal and non-uniform dose patterns (via IMRT), it becomes even more important to obtain the exact location of the tumor and surrounding organs. Currently, target delineation implies defining the *gross tumor volume*, which is the area of known tumor infiltration. From here, margins are added to cover areas of microscopic disease (*clinical target volume*) and to account for delivery uncertainties (*planning target volume*) [25]. However, on-going research attempts to utilize knowledge of the tumor heterogeneity obtained from biological or functional imaging techniques to define a *biological target volume* and guide a more advanced dose distribution [5]. The aim is to identify sub-volumes of the tumor with different radio-sensitivities or resistances and then prescribe a non-uniform dose distribution according to this (known as *dose painting*) to obtain higher treatment efficacy [26]. Several imaging techniques may have a role in biological target volume delineation, though no single imaging modality can reflect all underlying mechanisms of radio-resistance. It is an area of on-going research, and examples of techniques which may be relevant are: fluorodeoxyglucose (FDG-)PET that reveals metabolic activity and hypoxia response, dynamic contrast enhanced (DCE-)MRI that reflects tumor permeability and perfusion, and diffusion-weighted MRI that measures local water mobility [20, 27].

2.2 DIFFUSION-WEIGHTED MAGNETIC RESONANCE IMAGING

Diffusion-weighted MRI (DWI) was introduced in 1985 [28], and is basically an imaging technique with an increased sensitivity to the Brownian motion of water molecules, which is diffusion.

Einstein described this stochastic, thermal motion of particles by the Einstein equation [29], $r = \sqrt{2 \cdot D \cdot t}$, where r is the 1D root mean squared displacement, t is the time for movement, and D is the diffusion coefficient of a homogeneous medium.

Considering transversal magnetization, the random movement of water molecules in inhomogeneous fields affects the phase dispersion and thus the attenuation of the proton MRI signal. While the Einstein equation describes the free, unrestricted displacement of molecules, the movement of water in the tissue is affected by the micro-structure and e.g. restricted by cell-membranes. The degree of signal loss caused by water movement can therefore be revealing of the underlying tissue structure, which makes diffusion-weighted MRI interesting.

2.2.1 *Diffusion-weighted sequences*

An MRI sequence becomes sensitive to diffusion by introduction of a diffusion-weighting gradient pulse pair consisting of two strong gradient lobes with the same net magnitude area. In spin-echo sequences the two gradient lobes are identical, but placed before and after a refocusing RF pulse, respectively. In gradient echo sequences, the two lobes have opposite sign but are otherwise identical [30]. The delay between the gradient pair leaves time for water movement, and only static spins experiencing both gradients in the same position will be fully rephased to create the strongest possible echo. This means that when water molecules diffuse to other areas, the total signal, S, is decreased, as illustrated in Figure 2.1 and expressed by Equation 2.1 derived by Stejskal and Tanner [31]. S_0 is the signal when no diffusion-encoding is present. As the equation reveals, the signal attenuation due to diffusion depends on D, the diffusion coefficient of the medium, and properties of the diffusion-weighting gradients given by the so-called b-value (b) [18].

$$S = S_0 \cdot e^{-b \cdot D} \tag{2.1}$$

The b-value depends on the diffusion-weighting gradient's waveform. Typically, a rectangular gradient pair (trapezoidal lobes with negligible ramp-up time) is used, and in such case, the b-value is given by the following expression:

$$b = (\gamma \cdot G \cdot \delta)^2 \cdot (\Delta - \delta/3) \tag{2.2}$$

As Figure 2.1 also indicates, G is the gradient strength, δ is the duration of the gradient, and Δ is the time between the two diffusion-weighting gradients, i.e. the time over which the sequence is diffusion sensitive (*diffusion time*). The expression in the last parenthesis is the so-called *effective diffusion time*, which also accounts for diffusion taking place from the moment the gradient pulse is applied and phase-accumulation is initiated. The expression in the first parenthesis is also known as q, and essentially being the area of the diffusion gradient (ignoring the gyromagnetic constant, γ), this factor defines the spin displacement resolution [32]. By measuring the signal at several gradient configurations, i.e. different b-values, information about the diffusion and thus the tissue structure can be revealed.

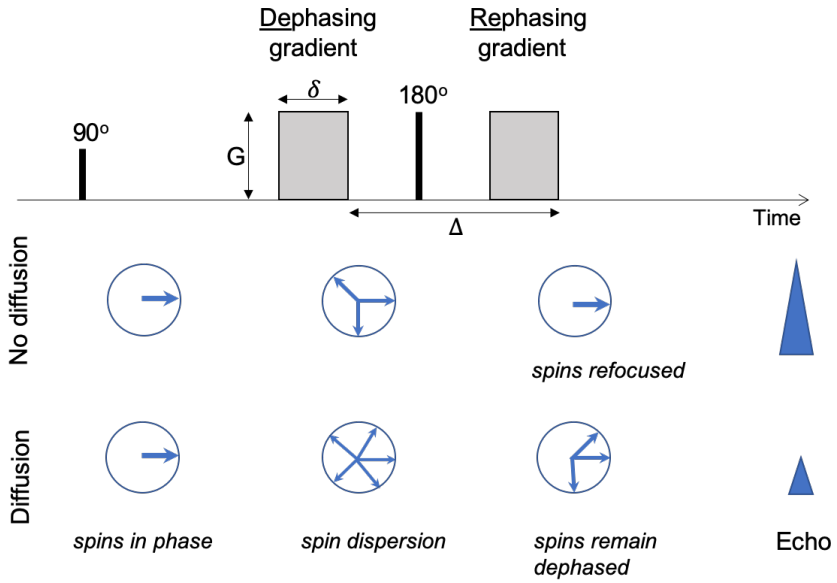

Figure 2.1: Excitation followed by diffusion-weighting gradients (gray blocks) and the effect on "spins" (blue arrows) for a case with and without diffusion. The first gradient lobe creates a dispersion of the spins (dephasing). The 180° pulse between the gradients inverts the dispersion created by the second gradient, and static spins will return to being in phase (refocus). However, when diffusion is present, the movement of spins causes a phase dispersion, which cannot be completely rephased by the second gradient. The result is a reduced echo compared to static conditions. Variables controlling the degree of diffusion-weighting are specified: the gradient strength (G), the gradient duration (δ), and the time delay between gradients (Δ). The figure is inspired by Patterson *et al.* [33]

As with all other gradients, diffusion-weighting gradients can be applied along all/each of the three physical directions, x, y and z, in the laboratory frame of reference. The direction of the resulting gradient vector field determines the direction of diffusion that the signal will be sensitive to. This means that the diffusion-weighted signal varies with the spatial orientation of the patient. In a DWI sequence, the diffusion-weighted signal is typically obtained for three orthogonal directions, and the geometric mean of these is calculated, leaving one image which is independent of orientation [18].

The diffusion-weighted echo-planar imaging sequence

A diffusion-weighted sequence can be divided into a *diffusion-preparation module*, containing the diffusion-endocing gradients, and a *readout module* for the image acquisition. In standard DWI, an echo-planar imaging (EPI) sequence is used as readout module [34]. In this sequence, the full k-space is traversed after a single excitation (single-shot EPI) or a few excitations (multi-shot EPI). Gradients in the phase-encoding (PE) and frequency-encoding (FE) directions are designed to create a specific sampling pattern covering the k-space. Figure 2.2 shows such a typical DWI sequence: a spin-echo diffusion-preparation module followed by a single-shot EPI with a Cartesian

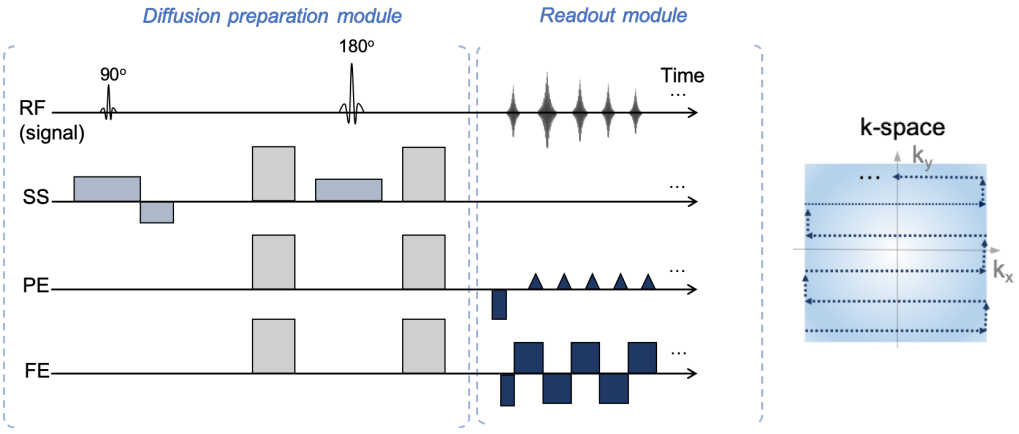

Figure 2.2: Schematic presentation of the DW-EPI sequence. The gray blocks are diffusion-encoding gradients, the light blue blocks are slice-selective (SS) gradients, and the dark blue blocks are the imaging gradients in the phase-encoding (PE) and frequency-encoding (FE) direction. The sampling trajectory for these gradients is shown in k-space.

sampling pattern. The advantage of using a single-shot EPI sequence for diffusion-weighted imaging is the rapid traversal through k-space (high acquisition speed), which is necessary to minimize the sensitivity to (intra-shot) bulk motion. On the other hand, a range of artifacts accompany the usage of single-shot EPI. One is Nyquist ghosting, which may originate from several sources of system imperfections or physical phenomenons causing signal modulation or displacement of k-space data, such as eddy currents. Another is image distortions and signal loss from off-resonance effects, e.g. magnetic susceptibility variations. The high sensitivity to this is associated with the relatively long readout period after each excitation, leaving room for accumulation of phase errors for off-resonance spins [18]. The effects of signal loss and geometrical distortions are prominent at areas with large susceptibility variations. For some imaging purposes this is not a big problem, but for applications within radiotherapy for example (including target delineation), it is highly problematic if the geometry is not trustworthy [35]. EPI is not the only type of fast imaging technique which can be used for DWI, and DWI sequences introduced to the clinic also include *fast spin-echo* (FSE) also known as a *rapid-acquisition with relaxation enhancement* (RARE) sequences. However, combining diffusion-weighted gradients with a RARE readout module poses difficulties. The following sections describe the RARE sequence and how to use it for DWI.

The RARE sequence and CPMG conditions

In an FSE or RARE sequence a single excitation pulse is followed by a set of refocusing radio-frequency (RF) pulses to produce a train of echoes [36]. Each RF pulse will create an echo which is distinctively spatially encoded, such that multiple or all lines of k-space can be sampled for one excitation. The number of refocusing RF pulses is known as the echo train length (ETL), and this parameter controls the scan-time reduction compared to a conventional spin-echo sequence, where only one k-space line is sampled per excitation.

As in conventional spin-echo sequences, the flip angle of the RF pulses should be 180° to maximize the spin-echo signal. However, in reality B1-field inhomogeneity and slice-profile imperfections will result in flip angle deviations from 180°, at least for part of the image volume. Furthermore, there may be advantages to intentionally reducing the flip angle from 180°. One aspect is the RF power deposition indicated by the specific absorption rate (SAR) which is necessary to take into account when using a large ETL. Another aspect is SNR-optimal acquisition weighting, a concept of controlling the flip angles for a desired signal weighting during k-space readout. This means that the RARE sequence in general consists of a set of refocusing RF pulses with a flip angle <180°. In any case, full refocusing will never be obtained due to the mentioned scanner imperfections, and instead an RF pulse can be considered as having three different effects on the magnetization: a 180°-effect that refocus transverse magnetization, a 90°-effect that tip magnetization into an orthogonal direction, and a 0°-effect that leave magnetization unaffected. The size of each effect or magnetization "component" depends on the actual flip angle. For a repeated set of RF pulses, as in multi-echo pulse experiments such as RARE, many magnetization components, also referred to as *coherence pathways*, are created. When the phase dispersion of a coherence pathway crosses zero an echo appears (more details on this are given in subsection 2.2.3). However, the general phase inconsistency between the different coherence pathways can result in signal cancellation. To obtain a high signal and avoid destructive interference of echoes, a set of conditions is used to ensure that echoes with the same phase coincide at desired time-points and that the rest of the echoes are eliminated [18, 37].

These conditions are called Carr-Purcell-Meiboom-Gill (CPMG) conditions, and consist of specific timing and phase relations for the RF pulse train [38–40]. First of all, the refocusing pulses must be 90° out of phase with respect to the excitation pulse. Secondly, the spacing between consecutive refocusing pulses must be the same throughout the train, and this must be twice the interval between the excitation pulse and the first refocusing pulse. These relations can be written as: $[90^\circ_x$ - τ - θ_y - 2τ - θ_y - 2τ - ...] where τ is the time interval, θ is the refocusing flip angle between 0-180° and subscripts specify along which axis the pulse is given in the rotating frame of reference. Lastly, the phase accumulation of a spin must be the same between any two consecutive RF pulses. This condition is satisfied using crusher gradient pairs of same area and phase-rewinding gradients prior to the next RF pulse. When the mentioned conditions are met, the primary spin-echoes and stimulated echoes will both occur exactly at the midpoint between two RF pulses, while the free induction decays (FIDs) and secondary spin-echoes will be eliminated [18]. Figure 2.3 shows a schematic presentation of a RARE pulse sequence with the mentioned CPMG conditions satisfied, though crusher gradients are left out for simplicity.
Combining a RARE readout with a diffusion-preparation module, however, is not trivial. The strong diffusion-encoding gradient pair creates a high sensitivity to any motion (including bulk motion), which means that minimal unavoidable movement will create a random phase of the different coherence pathways, and the CPMG conditions are no longer obeyed. This means that the echoes may sometimes interfere destructively and other times constructively creating severe artifacts. Strategies have been developed to handle this problem among which a relevant technique

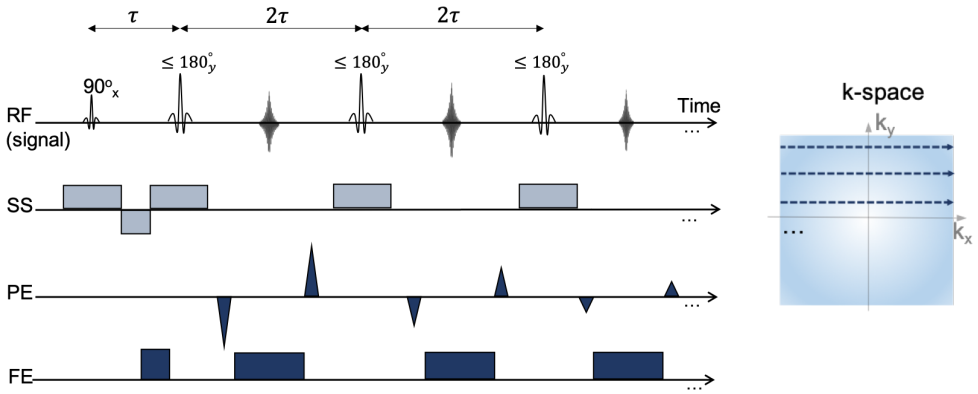

Figure 2.3: Schematic presentation of a RARE sequence. The light blue blocks are the slice-selective gradients, and the dark blue blocks are imaging gradients in the PE and FE direction. The timing and phase relations required to satisfy the CPMG conditions are indicated. The corresponding k-space is shown to the right. Crusher gradients are left out for a simplistic representation, but phase-rewinder gradients can be seen in the PE direction prior to each RF pulse.

is the split acquisition of fast spin-echo signals for diffusion imaging (SPLICE) sequence [41].

2.2.2 *The SPLICE sequence*

The SPLICE sequence [42] is adapted to handle the problem of random phase of the different coherence pathways generated by a RARE module. It includes a "split-echo" design consisting of a prolonged/imbalanced readout gradient which creates a separation of the signal components with different phases. The phase will be random, but it will be the same within two "families" of echoes created in the echo train, including both stimulated echoes and spin-echoes. If the readout gradient is prolonged sufficiently (e.g., with an area four times the size of the pre-phaser gradient), the two echo families (E1 and E2) will be well separated, allowing two individual k-spaces to be sampled simultaneously. The final diffusion-weighted image is then generated by reconstructing each set of k-space data and summing the two magnitude images for an improved SNR. As the split-echo acquisition do not require any phase relations of the signal components, the phase of the RF pulses can be arbitrarily chosen [18, 42]. Figure 2.4 illustrates the SPLICE sequence with a stimulated echo diffusion preparation, as originally presented in [42]. The more common spin-echo preparation consisting of a 180°-pulse between the two strong gradients has been replaced by two 90°-pulses leading to a stimulated first echo. The phase of the magnetization is stored along the longitudinal axis between the second and third pulse, which can be an advantage as this allows a long separation of the two diffusion-encoding gradients to create a high b-value without incurring T_2 or T_2^* signal decay. On the other hand, the maximal amplitude of a stimulated echo is only half the magnitude of a spin-echo (the remaining magnetization is dephased), so it may also be convenient to use a spin-echo diffusion-preparation [18, 42].

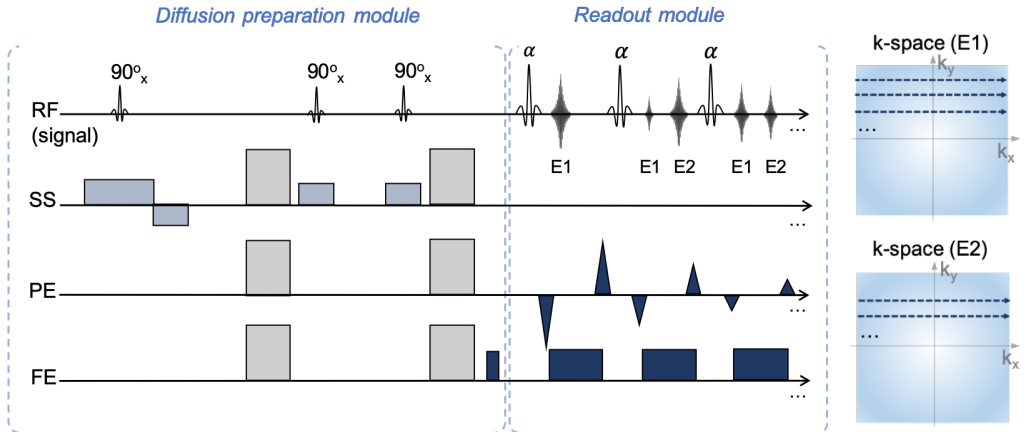

Figure 2.4: Schematic presentation of the SPLICE sequence. The gray blocks are the diffusion-encoding gradient pairs, the light blue blocks are the slice-selective gradients, and the dark blue blocks are the imaging gradients. The gradients in the FE direction are imbalanced, as the pre-phaser (gradient prior to the readout module) is much less than half the area of the following readout gradients. The two echo families appearing (E1 and E2) are illustrated on the RF line (top), and the corresponding k-spaces are illustrated to the right. Three 90° pulses along the same axis (specified in the subscript) are used for a stimulated echo diffusion preparation. α represents the flip angle (0-180°) of the refocusing pulses in the readout pulse train. No phase relation is required between the signal phase the the refocusing pulse phase, why no axis is specified in the subscript. Crusher gradients are left out for a simplistic representation.

The advantage of the SPLICE sequence is its robustness towards image distortions and thus its geometric accuracy compared to the more conventional DW-EPI sequence, but drawbacks of the SPLICE sequence also exist. For both types of rapid image acquisition, traversing through the entire k-space after a single excitation leaves time for T_2-relaxation during readout which creates a modulation of the signal in the Fourier domain. The SPLICE sequence is particularly prone to this effect as the necessary prolongation of the readout gradients increases the minimum possible echo spacing. The result is an increased image blurring compared to both conventional RARE and EPI images. Introducing acceleration methods such as parallel imaging to lower the ETL can reduce the problem of blurriness, but it may still be relevant to optimize the sequence to achieve a higher resolution without lowering the already limited SNR of DWI. The mentioned concept of acquisition weighting may be useful in such optimization.

Multi-shot DWI sequences, whether EPI or RARE based, results in images of better SNR and fewer artifacts, but they require acquisition of navigator echoes for phase correction [43]. This prolongs the scan time, and the phase correction is only partly effective if movement-induced phases vary at voxel-level. Therefore, single-shot sequences are most often used.

2.2.3 *Simulating multi-echo sequences*

Extended phase graphs

As mentioned above, for CPMG or other multi-pulse experiments, many coherence signal pathways are created and multiple echoes are produced. Every new RF pulse generates three new coherence pathways or "magnetization components" for each existing component, and the size of these depends on the flip angle. To keep track of the signal coherence pathways and their echo generation, Hennig [44] introduced an algorithm in 1988 depicting the magnetization response for a given set of RF pulses, gradients, relaxation parameters and timing conditions, namely the extended phase graph (EPG).

In EPG calculations the known orthogonal basis vectors $\mathbf{M}(\omega) = [\mathrm{M}_x(\omega), \mathrm{M}_y(\omega), \mathrm{M}_z(\omega)]^\top$ describing the complex magnetization for an ensemble of spins, are replaced by a new basis set, $\mathbf{F}(k)$, describing the phase state configurations of the spins. The variable k indicates levels of dephasing. The phase states, $\mathbf{F}(k)$=$[\mathrm{F}_k^+\ \mathrm{F}_k^-\ \mathrm{Z}_k]^\top$, are derived from the Fourier Transform of $\mathbf{M}(\omega)$, where $\theta \in [0, 2\pi]$ quantifies the angle of dephasing due to off-resonance induced by a gradient:

$$\mathrm{M}_x(\theta) \pm i\mathrm{M}_y(\theta) = \sum_{k=-\infty}^{\infty} \mathrm{F}_k^{\pm} e^{ik\theta} \quad \text{and} \quad \mathrm{M}_z(\theta) = \sum_{k=-\infty}^{\infty} \mathrm{Z}_k e^{ik\theta} \tag{2.3}$$

For a periodic pulse experiment (fixed TR and TE) the magnetization can simply be evaluated as a function of discrete states given by integer values of k, why a summation is used in Eq. 2.3. The F_k states are the transverse basis depicting a "phase twist" and when a component (ensemble of spins) refocuses from F_1^- to F_0 an echo is generated. The Z_k state is the longitudinal basis describing sinusoids, and here the spins are unaffected by new phase dispersion, i.e. magnetization is stored. Now to simulate a multi-echo experiment, three operators can be applied iteratively to the phase states. The effect of an RF pulse is applied through a rotation operator describing the rotation of spins for an RF pulse of a given flip angle, α, and phase, ϕ, (Eq. 2.4). A *hard pulse approximation* is used, i.e. the pulse is treated as an instantaneous pulse without any off-resonance effects, e.g. from variations across the pulse profile.

$$\mathbf{T}_\mathrm{k}(\alpha) = \begin{bmatrix} \cos^2(\alpha/2) & e^{2i\phi}\sin^2(\alpha/2) & -ie^{i\phi}\sin(\alpha) \\ e^{-2i\phi}\sin^2(\alpha/2) & \cos^2(\alpha/2) & ie^{-i\phi}\sin(\alpha) \\ -\frac{1}{2}e^{-i\phi}\sin(\alpha) & \frac{1}{2}e^{i\phi}\sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{2.4}$$

A shift operator is used to apply the dephasing effect occurring between an RF pulse and an echo, i.e. during a half ESP, and it basically shifts the spins of an F_k state to the successive state, $\mathbf{S}_\mathrm{k}(\mathrm{ESP}/2) : F_k \longrightarrow F_{k+1}$. Finally, the relaxation operator given in Eq. 2.5 is used to describe the T_1 and T_2 relaxation [44–47].

$$\mathbf{R}_\mathrm{k}(\mathrm{ESP}/2) = \begin{bmatrix} e^{-\mathrm{ESP}/(2\mathrm{T}_2)} & 0 & 0 \\ 0 & e^{-\mathrm{ESP}/(2\mathrm{T}_2)} & 0 \\ 0 & 0 & e^{-\mathrm{ESP}/(2\mathrm{T}_1)} \end{bmatrix} \tag{2.5}$$

The graphical representation of the EPG algorithm, an extended phase diagram, is helpful in understanding the evolution of a spin system and generation of different types of echoes during a multi-pulse experiment, though the effect of relaxation is left out. Figure 2.5 presents an extended phase diagram for the first four RF pulses of a CPMG experiment. To fulfill the CPMG conditions there is a phase difference of $\pi/2$ from the excitation pulse and the following refocusing pulses. The diagram presents the relevant signal pathways by color-coding and marks every zero-crossing, i.e. every echo not coinciding with an RF pulse. An example of a spin-echo (SE) and a stimulated echo (STE) is indicated as the first and second echo, respectively. Continuing the RF-pulse train, the echoes will be a result of several coinciding pathways and thus a mix of several types of echoes. The figure also presents an extended phase diagram for a SPLICE experiment. The characteristic prolonged gradient in a SPLICE readout causes a more rapid spin dephasing, indicated by the more steep slope of the signal pathways compared to the CPMG experiment, i.e. spins switch to higher phase states between two refocusing pulses. This results in the generation of two echoes between each RF pulse except after the first refocusing pulse. Again, the first echo will be a spin-echo. The second and third echo appearing after three RF pulses is a stimulated echo and secondary spin-echo (SE2), respectively. Considering a long echo train, it becomes complex to draw all signal pathways and visualize them in a diagram. Conveniently, the EPG algorithm, relying on the three mentioned operators, keeps track of all pathways and calculates the magnitude of each echo. It is thus a relatively simple tool for simulating the signal evolution throughout a multi-pulse experiment.

JEMRIS tool

The Jülich Extensible MRI Simulator (JEMRIS) is another tool for simulating MRI sequences, and opposite to the EPG algorithm it can include and deal with arbitrary RF waveforms, gradient waveforms, coil geometries and more, which allows for modeling of spin systems under the influence of important off-resonance effects. It is based on the Bloch equations, which formulated in cylindrical coordinates is well-suited for numerical implementations [48]. Together with the utilization of parallel programming, JEMRIS provides an optimized platform for the general solution of the Bloch equations of non-interacting spins. The C++ based software is open-source and includes a MATLAB based graphical user interface (GUI) consisting of a sequence development and sequence simulation environment. In the development environment small basic modules can be utilized to build a sequence of interest, and in the simulation environment a coil geometry, object and other effects such as chemical shift are chosen for a simulation of the built sequence. More details on JEMRIS can be found in Stöcker et al [48], and user-guide and software download are available at `https://www.jemris.org/`.

In this project both the EPG algorithm and JEMRIS tool are used for simulation of a SPLICE experiment. The JEMRIS simulation is utilized to validate that EPG simulations are adequate for the purpose of optimizing the flip angle scheme in a SPLICE sequence.

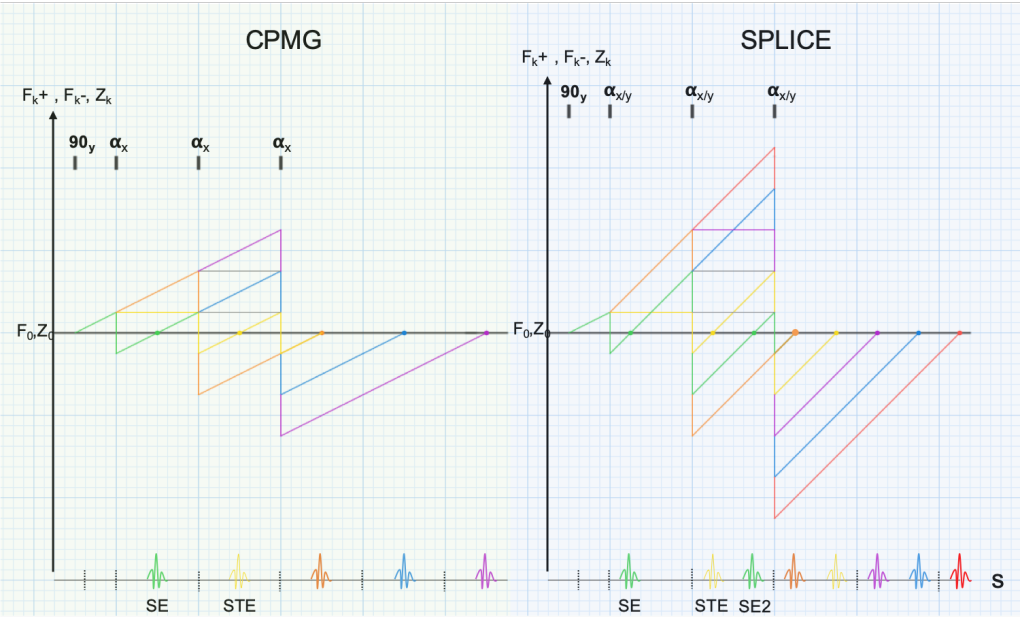

Figure 2.5: Extended phase diagram for a CPMG experiment (green background) and a SPLICE experiment (blue background), respectively. The different signal pathways and echoes are color-coded, though several coincide. The green pathway is an example of a spin-echo (SE): The first pulse acts as a 90° pulse, the second as a 180° pulse. The yellow pathway is an example of a stimulated echo (STE): the first pulse act as a 90° pulse, the second as a 90° pulse, turning the magnetization into the corresponding z-state (longitudinal), and the third as a 90° pulse, turning the magnetization transversal again. For the SPLICE experiment, a larger gradient area causes a faster dephasing and the spins switch to higher phase states between two pulses than for the regular CPMG experiment. This results in the appearance of two echoes after each refocusing pulse (except after the first). Again, the first and second echo is a spin-echo and stimulated echo, respectively. The third echo is an example of a secondary spin-echo (SE2): the third pulse acts as a 180° pulse on the green signal pathway and another spin-echo is generated.

2.2.4 *Analysis of diffusion-weighted MRI data*

A common way to handle and analyse diffusion-weighted MRI data is by calculation of the quantitative apparent diffusion coefficient (ADC) map. In the tissue, the diffusion will be restricted by micro-structural components such as cell membranes and will deviate from pure Gaussian diffusion described by the Einstein equation. A Gaussian-based diffusion model is still used, but the decay constant is expressed as an "apparent" diffusion coefficient to emphasize this deviation from a "true" diffusion coefficient [49, 16]. To estimate the ADC, multiple measurements with different b-value (at least two) are fitted to the mono-exponential Stejskal and Tanner model stated in Eq. 2.1, with D replaced with ADC. Rewritten and at logarithmic scale, the ADC simply becomes the slope of a linear relationship (Equation 2.6). Estimating the ADC in each voxel generates the ADC map.

$$\ln(S_i/S_0) = -b_i \cdot \mathrm{ADC} \tag{2.6}$$

Calculation of the ADC map is thus a relatively simple way to retrieve the local average diffusivity and it is utilized clinically and in research e.g. for brain examination.

However, there are also issues to be aware of. Though the ADC is quantitative it is a relative measure and depends on image quality and scan parameters such as the diffusion time and b-values [49, 50]. For example, for low b-values, the signal attenuation will mainly be due to perfusion of the water molecules, whereas for higher b-values the attenuation is more diffusion dependent, and the transition between these effects depends on the tissues vascularity [16, 11]. So for a large range of b-values, a bi-exponential model may explain the signal attenuation more accurately [14]. Such a model has been suggested, the intra-voxel incoherent motion (IVIM) model, but it requires acquisition of many b-values and a high quality of data to fit the four model parameters without large uncertainties [16]. For high b-values, the signal is much attenuated and may reach the noise floor. A positive bias due to noise for these measurements will lead to underestimated ADC values.

Using diffusion times of tens of milliseconds makes the DWI signal sensitive to tissue structures at micrometer scale. The main part of the signal attenuation is caused by water movement in the extracellular space and the intracellular diffusion is without significant impact, but micro-structural features such as cell density and the extracellular composition influence the diffusion [51]. However, these signal features are not resolved if voxel-based processing is combined with a relatively large voxel size (1-3 mm), which is the condition in DWI [52]. The voxel size is limited by the SNR which again is limited in single-shot sequences where a short TE is desired. In DWI, the bandwidth is particularly high because a fast k-space coverage is required to leave minimal time for relaxation. Thus, describing the diffusion-weighted signal by a mono-exponential model and estimating a single coefficient for a voxel is a simplistic representation of the underlying water movement, and though it can be informative, knowledge of the tissue structure at smaller scales is averaged away. The large voxel size also causes the signal to reflect diffusion in several macroscopic tissue types, both white and gray matter for example, and there is a partial volume effect to be aware of when interpreting ADC maps.

Besides from the IVIM model, other methods exist which may describe the diffusion-weighted signal attenuation more accurately. As mentioned, the ADC calculation is based on the assumption of a Gaussian, isotropic diffusion, though the diffusion in tissue typically is non-Gaussian and spherically asymmetric (anisotropic) due to restrictive structures. Kurtosis is the lowest-order symmetric deviation from Gaussian behavior, and a polynomial model can be used to derive a kurtosis-corrrected diffusion coefficient [49, 14, 53]. In regard to the anisotropic behaviour, diffusion tensor imaging (DTI) is used to calculate directional diffusion parameters expressed in a *diffusion tensor*. The tensor elements describe diffusion along each of the three spatial coordinates, x, y and z, defined in the laboratory reference frame, and the pairwise correlation between them. To calculate all these parameters, acquisition of at least six different diffusion gradient orientations

are needed, all non-colinear and non-coplanar [18].

2.3 DATA ANALYSIS

In the previous section, the most common ways of analyzing diffusion-weighted data were reviewed. This section reviews data analysis strategies more generally, though still with a perspective on multi-contrast MRI, where the main objective is to estimate and map functional and structural tissue properties. I here divide data analysis approaches into two broad branches, model-based and data-driven.

2.3.1 *Model-based analysis*

In model-based analysis, a mathematical, physical or biophysical model is defined a priori. It thus requires knowledge of the system based on a theoretical foundation or learned from previous experiences. The data is then fitted to this model by estimating the model parameters using a maximum likelihood approach, typically assuming a Gaussian noise distribution leading to the well-known least-squared-error fitting [54, 55].

Within MRI, modelling is used to quantify tissue parameters both at and below the nominal resolution leading to micro-structural mapping. The latter requires a decomposition of the signal into a weighted sum of latent factors, which often leads to the ill-posed nature of inverting a sum of functions. A specific example is the decomposition of multi-echo T_2-weighted data for relaxometry analysis. Typically, the signal is modelled as a weighted sum of multiple exponentially decaying signal components spanning the T_2 domain. This ill-conditioned linear system is then usually solved by implying a non-negativity constraint and solving for the model weights using the non-negative least squares (NNLS) algorithm [56] or variations hereof [57, 58].

If the presumed model accurately describes the underlying contributions to the observed signal, it is a straight-forward method for mapping tissue parameters independent from acquisition and system details, i.e for quantifying the MRI data. However, there are several pitfalls in regard to modelling of imaging data. First of all, the presumed model or the associated assumptions can be wrong or inadequate for describing the dominant physical effects. In such case, the estimated model parameters will be misleading. It is also a challenge to separate the desired biological signal variability from undesired variability stemming from e.g. acquisition artifacts or noise [59]. As mentioned for the ADC calculation, the impact of noise may result in biased estimates.

To reduce the impact of noise and the risk of overestimation in model-based analysis, prior information in the form of a prior distribution on the model parameters can be introduced. This is often based on a zero mean Gaussian distribution which leads to the widely applied L2-regularization, penalizing the sum of squared model weights. Using the NNLS approach for relaxometry implies voxel-wise fitting. The method is thus heavily affected by noise and regularization can advantageously be introduced. The L2-regularized NNLS algorithm is defined in Eq. 2.7 with $\mathbf{x}$ being the model weights, $\mathbf{A}$ the system matrix, $\mathbf{y}$ the measured signal, and λ the

regularization constant [55].

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_2^2 \quad \text{subject to} \quad \mathbf{x} \geq 0 \tag{2.7}$$

There are several examples of voxel-wise fitting, i.e. univariate modelling within MRI, but multivariate models also exist such as the component-resolved NMR (CORE) processing family for signal decomposition. An advantage of a multivariate method is the utilization of all available information simultaneously while estimating the model parameters.

2.3.2 *Data-driven analysis*

In data-driven analysis, the data structure and patterns are derived from the data itself. All data points are observed simultaneously to find the set of components representing the data the best. Although linearity of the system often is assumed, the strategy is considered unsupervised. The only necessary a priori knowledge (besides from such linearity assumption) is the rank of the underlying system, i.e. number of components. As for modelling, a Gaussian noise distribution and thus a least squares objective function is a typical choice for finding the representative components.

A very well-known data-driven analysis technique is the principal component analysis (PCA) [60]. In PCA, the data reconstruction consists of orthogonal rank-one components sequentially explaining the maximal amount of variance in the data. The method is guaranteed to find the lowest-error reconstruction for the chosen rank [61]. However, the assumptions of orthogonality and sequential sorting of components, are often not desirable in practice.

Non-negative matrix factorization

A technique which has been popular within image processing (and several other areas) is the non-negative matrix factorization (NMF) presented in Eq. 2.8 [62]. The data matrix consisting of n observations each of m elements, $\mathbf{X} \in \mathbb{R}^{m \times n}$, is decomposed into two factors; k non-negative basis functions, $\mathbf{W} \in \mathbb{R}_+^{m \times k}$, and the associated non-negative weights for the linear combination of basic functions to reconstruct the data, $\mathbf{H} \in \mathbb{R}_+^{k \times n}$. An error matrix, $\mathbf{E} \in \mathbb{R}^{m \times n}$, is included to account for noise and non-factorizable signals. So, instead of an orthogonality requirement as in PCA, there is a component-wise non-negativity requirement, which fits well to imaging intensities which by nature are non-negative [63].

$$\mathbf{X} = \mathbf{W}\mathbf{H} + \mathbf{E} \tag{2.8}$$

Several objective functions can be used such as the Euclidean distance and the Kullback-Leibler divergence, and the appropriate choice depends on the noise model. The most common choice is the Frobenius norm which implicitly assumes Gaussian i.d.d. residuals [64]. The optimization then becomes:

$$\min_{W,H} \|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 \quad \text{such that} \quad \mathbf{W} \geq 0 \quad \text{and} \quad \mathbf{H} \geq 0 \tag{2.9}$$

The problem stated is non-convex and to solve it, most NMF algorithms utilize that the sub-problem of optimizing just one factor while keeping the other fixed is convex. Each sub-problem is then equivalent to an NNLS problem, described above. This design is known as a two-block coordinate descent scheme, and thus consists of updating the solution by alternately optimizing the two factors, $\mathbf{W}$ and $\mathbf{H}$, respectively, in which the objective function decreases. There are several ways to define the update rules for estimating the two factors, and many variations of the NMF algorithms exist [63, 64]. One approach is the multiplicative update rules (MU), also seen as a rescaled gradient method [65], which was originally used for solving standard NNLS problems [66]. This method is simple to implement and works well on sparse matrices, but converges relatively slowly.

Another approach is the alternating non-negative least squares (ANLS) algorithms, which update the factors by alternately solving each of the constrained convex sub-problems utilizing information from the Hessian of the optimization problem:

$$\begin{aligned} \mathbf{W}^{l+1} &= \min_{W \geq 0} \|\mathbf{X} - \mathbf{W}\mathbf{H}^{l}\|_F^2 \\ \mathbf{H}^{l+1} &= \min_{H \geq 0} \|\mathbf{X}^T - \mathbf{H}^T[\mathbf{W}^{l+1}]\|_F^2 \end{aligned} \tag{2.10}$$

Without constraints, the exact solution to the problem can be obtained by deriving the stationary points from equating the gradient to zero. *With* constraints, which within variations of the NMF can be more than just the non-negativity constraint, it can be beneficial to reformulate the problem in quadratic programming form (only possible using the Gaussian noise model) [67, 68]. Equation 2.11 shows the general case of such reformulation:

$$\begin{aligned} &\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{R}\mathbf{x} - \mathbf{d}\|^2 \longrightarrow \min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{f}^{\mathrm{T}}\mathbf{x} \quad \text{subject to: } \mathbf{G}\mathbf{x} \leq \mathbf{a} \\ &\text{where } \mathbf{Q} = \mathbf{R}^{\mathrm{T}}\mathbf{R} \text{ and } \mathbf{f} = -\mathbf{R}^{\mathrm{T}}\mathbf{d} \end{aligned} \tag{2.11}$$

Besides from the inequality constraints described by $\mathbf{G}\mathbf{x} \leq \mathbf{a}$, the problem can instead or additionally be constrained by equality constraints, $\mathbf{G}_{\text{eq}}\mathbf{x} = \mathbf{a}_{\text{eq}}$, and lower and upper bounds: $LB \leq \mathbf{x} \leq UB$. Constraints can therefore easily be implemented, and several algorithms exist to solve the final dual optimization problem, e.g. the interior-point or the active set algorithm [69].

Lastly, the hierarchical alternating least squares (HALS) method should also be mentioned as one of the common strategies to solve the problem. The method utilizes that columns of $\mathbf{W}$ do not interact, and updates one column of $\mathbf{W}$ at a time, i.e the sub-problem is solved by an exact coordinate descent method. So similar to PCA, though without orthogonality, the method iteratively estimates rank-one components [70].

In general, the standard NMF problem has non-unique solutions. First of all, there is a permutation and scaling ambiguity which means that components can be permuted or scaled without changing the cost. However, more generally it applies that if $\mathbf{X} \approx \mathbf{W}_0\mathbf{H}_0$ is a solution, and $\tilde{\mathbf{W}} = \mathbf{W}_0\mathbf{D}$, $\tilde{\mathbf{H}} = \mathbf{D}^{-1}\mathbf{H}_0$ where $\mathbf{D}$ is a rotational matrix, then $\mathbf{X} \approx \tilde{\mathbf{W}}\tilde{\mathbf{H}}$ is also a solution. $\mathbf{D}$ can be a product of a permutation ($\mathbf{P}$) and scaling matrix ($\mathbf{S}$), $\mathbf{D}$=$\mathbf{PS}$, but its definition is broader than that,

and transformations beyond generalized permutations exist where the above algebra still applies. Overall, the uniqueness of the standard NMF is not guaranteed [69].

Similar to model-based analysis, a prior can be introduced, which here advantageously can reduce the number of non-unique solutions. A prior can be introduced either as a regularisation or as constraints where some solutions are simply not allowed. This has resulted in many variations of NMF, each with different priors. For further details on the NMF algorithms, the reader is referred to the book by Cichocki *et al*: Nonnegative Matrix and Tensor factorizations, 2009 [64].

2.3.3 *Supervised learning*

In supervised learning, data examples ("training data") are used to train a model that maps an input to an output. The strategy is called *supervised* because the output object (e.g. a "label") of the training data is known, such that a relationship to the input object (a set of variables) can be learned. The learned model is typically evaluated using "test data", where the model's predicted output is compared to the actual output, i.e. the generalization of the learned pattern is tested. Several supervised learning techniques exist, but overall they can be divided into classification and regression algorithms. A classification algorithm is used when the output object is a discrete variable, and a regression algorithm is used when the output object is a continuous variable.

An example of a regression algorithm is *multiple linear regression* (MLR) which can be used when a linear relationship between the independent variables ("regressor") and the dependent output variable is expected. In a high-dimensional space (many regressors), the relationship will often be linear. The MLR can for N observations be formulated in matrix form as:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \tag{2.12}$$

where $\mathbf{y}$ is the N output variables, $\mathbf{X}$ is an $[N \times M{+}1]$ matrix consisting of M regressors for each observation, and a one-column to include a bias (intercept) in the model. β consists of the $M{+}1$ model parameters ("weights"), which are learned via the model training, and ϵ is an error vector (residual). Most often, the sum of squared errors $\|\epsilon\|_2^2$ is minimized in the process of estimating β. The method can then also be seen as the ordinary least squares (OLS) regression extended to multiple explanatory variables [71, 55].

A model equivalent to linear regression but for binary classification problems is *logistic regression* which belongs to the larger class of models known as generalized linear models (GLMs). The logistic sigmoid function (Eq. 2.13) is used to turn the otherwise linear function of model features (Eq. 2.12) into a probability of a class, which then is mapped to a binary outcome. If the probability for "class A" is below 0.5, the outcome is "class B", otherwise the outcome is "class A".

$$\sigma(y) = \frac{1}{1 + e^{-y}} \tag{2.13}$$

Whereas the sum of squared errors is used as objective function to estimate the parameters in linear regression, the negative logarithm of the likelihood (also known as the *log loss*) is typically

used as objective function to estimate the logistic regression parameters. The expression for the log loss error function is given in Eq. 2.14 with y_n being the predicted outcome probability for observation n, and $t_n \in \{0, 1\}$ being the actual class label. β is the model parameters as defined in Eq. 2.12 [55, 72].

$$E(\beta) = -\sum_{n=1}^{N} t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n) \tag{2.14}$$

In supervised learning, the models are typically prone to overfitting, meaning that the estimated parameters fit to the training data but generalizes poorly. Especially using maximum likelihood, as in logistic regression, there is a high risk of overfitting if the data is linearly separable [55]. General solutions to avoid overfitting consist of adding a regularization term to the objective function and using cross-validation (CV). Regularization has also been mentioned in context with model-based and data-driven analysis and consists of adding a penalty term to the objective function. In L1-regularization, the penalty term is: $\lambda \sum_m |\beta_m|$ leading to a sparse solution, which in case of a large feature space often is desired. The appropriate penalty strength, i.e. the regularisation constant λ, depends on the data and needs to be estimated which CV is useful for. Instead of L1-regularisation, a sparsity in the feature domain can also be introduced directly by reducing the number of input features with a preceding feature selection process [55].

CV is a useful technique for estimating model prediction performance and generalization, especially when the amount of available data is limited. It is a resampling technique where the data is divided into subsets used for model training and testing, respectively. The subset used for testing is changed during the iterative resampling until all data have been used as "test data". This way, the method informs on both training and test performance and can reveal problems as overfitting and selection bias. The process is typically repeated several times with different data partitioning to reduce the effect of variability [73]. Typical metrics to evaluate the model performance is the *accuracy* or *area under the ROC curve* in classification cases, and the *root mean squared error* or *Pearson's correlation coefficient* in regression cases [55].

As stated above, CV is also used for model selection, i.e. for estimating a *hyperparameter* as the penalty term or number of input features, for example. If CV is introduced for model training already, an extra internal CV loop is needed for estimation of the hyperparameters. In this loop, a "grid" of discrete values are tested as hyperparameters, and the combination leading to the overall best performance is carried on to the outer CV loop concerning estimation of the model weights. This full process is referred to as *nested cross-validation.* A "grid search" (as it as called) is simple but relatively time consuming, especially if several hyperparameters are to be estimated. There are more advanced model selection strategies available, such as Bayesian optimization [74].

DATA-DRIVEN ANALYSIS FOR TISSUE CHARACTERIZATION

3

In the Background chapter, strategies for analysing diffusion-weighted MRI data and multi-contrast MRI data in general, have been reviewed. In this project, the specific aim of such analysis was to reveal information on the tissue microstructure of tumors as this can be prognostic for or predictive of the treatment response. To reveal tissue information at sub-voxel level, the ill-posed problem of decomposing the MRI signal into smaller compartment-specific components is met. While a model-based approach is restrictive in terms of dictating the signal behaviour in advance and may result in misleading or biased estimates if inaccurate, a model-free or data-driven approach such as the PCA is relatively unrestrictive and may result in components with an unrealistic signal behaviour being difficult to interpret and relate to tissue compartments. As a solution, we have developed a method which is somewhere between model-based and completely model-free approaches with a potential for tissue characterization.

In data-driven approaches, the data is decomposed into a lower-dimensional structure under some assumptions, which can leave a high flexibility for the decomposition or oppositely include several constraints that components must comply with. Our developed method therefore consists of a data-driven approach which have been formed by a set of constraints that fits to the assumption of the natural behavior of MRI signals, such that extracted signal components become more realistic and informative. By using a data-driven approach, we avoid having to forecast and model a complex tissue system and its associated biophysical processes (e.g., perfusion and water exchange). Additionally, the method identifies latent mixtures within each measurement, making it robust to partial volume effects.

The suggested solution is an extension of the NMF which in standard form only implies non-negativity of the latent factors. We imposed two additional constraints that force the signal components to be monotonous, and their first derivative (slope) to be monotonous as well. While the first part was originally introduced by Bhatt and Ayyer [68] in the form of the monotonous NMF, the constraint imposed on the first derivative has not been investigated earlier. It was added to limit the flexibility of the extracted components further, while allowing a signal behavior expected for many types of MRI measurements, which is a mono- or multi-exponential behavior or at least near exponential in many cases.

$$b_0 + \sum_{n=1}^{N} b_n \cdot e^{a_n x} \tag{3.1}$$

The group of mono- and multi-exponential functions described by Eq. 3.1 are monotonous as long as the exponents a_n have a common sign and the arbitrary number of non-zero coefficients b_n also share sign. In other words, the first derivative has no zeros in the case of $a_n < 0$ or $a_n > 0$ for all n, and $b_n \leq 0$ or $b_n \geq 0$ for all n. The derivatives of Eq. 3.1 fall within the same class of

functions, and are therefore also monotonous under these conditions. The question of interest is whether there are cases where signal components no longer can be described by a sum of exponentials of shared sign and thus deviate from the monotonicity constraints.

Consider a case of just two idealized tissue compartments (for simplification) which could represent inter-cellular and extra-cellular tissue, for example. The signal or magnetization "populations" of the compartments can then be described by the system:

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t} = \mathbf{A}\cdot\mathbf{y}(t), \quad \mathbf{A} = \begin{pmatrix} -R1-K & K \\ K & -R2-K \end{pmatrix} \tag{3.2}$$

where $\mathbf{y}(t)$ is the compartment populations at time t, and $\mathbf{A}$ is a system matrix holding the population decay rates, $R1$ and $R2$, and an exchange rate, K, between the two compartments (accounting for water exchange). This formalism applies to more than two compartments and is relevant for many cases of MRI measurements, e.g. the T_2-weighted signal with respect to TE or the diffusion-weighted signal with respect to the b-value, i.e. where the population decays towards zero. For T_1-relaxation, the relevant system may instead describe a change from equilibrium: $\mathrm{d}\mathbf{y}/\mathrm{d}t = \mathbf{A}\cdot(\mathbf{y}(t)-\mathbf{y}_0)$, but this can easily be rewritten to the simpler form in Eq. 3.2 by a parameter substitution. Factors affecting the populations in $\mathbf{y}$(t) is the decay and exchange rates, K, $R1$, and $R2$, and further the initial conditions, i.e. $\mathbf{y}(t=0)$. To check the monotonicity of the components in $\mathbf{y}(t)$, it is easiest to look at specific cases. In the simple case of no water exchange, $K=0$, $\mathbf{A}$ becomes a diagonal matrix and the two differential equations is decoupled. Hence, both populations simply decay towards zero and monotonicity is obeyed.

In the complex case, where exchange is present, it becomes more difficult to state something general for the populations in the physical compartments. For very slow water exchange compared to the signal relaxation times ($K \ll R1, R2$), the compartments can be approximated as non-interacting (i.e. no exchange). For very fast water exchange ($K \gg R1, R2$) between two compartments, the compartments can be approximated as one big water pool represented by just one monotonous signal component. For the more complex case of intermediate exchange (K in the order of $R1$ and $R2$), examples exist where monotonicity is violated for the populations in the physical compartments (shown below). However, since $\mathbf{A}$ is negative definite (symmetric with negative eigenvalues), a solution can conveniently be found from matrix diagonalization. This is equivalent to a rotation into another basis, and the resulting mono-exponential signal evolutions are then associated with "pseudo-compartments" instead of the physical ones. The coefficients describing these pseudo populations may be positive or negative, depending on the freely chosen sign of the eigenvectors. So, in any case it is possible to find a solution that only includes monotonous signals, though it may be for pseudo-compartments. An example is given below (Fig. 3.1), where one of the populations of the physical compartments violates the monotonicity conditions, while those of the pseudo-compartments comply with them. In the example, the two populations start with very different initial conditions: $\mathbf{y}(t=0) = [1, 0]$. This could simulate a situation were the signal is pre-nulled for a tissue, e.g. of edema in inversion recovery experiments. As the figure shows, the signal of the nulled compartment (blue, physical)

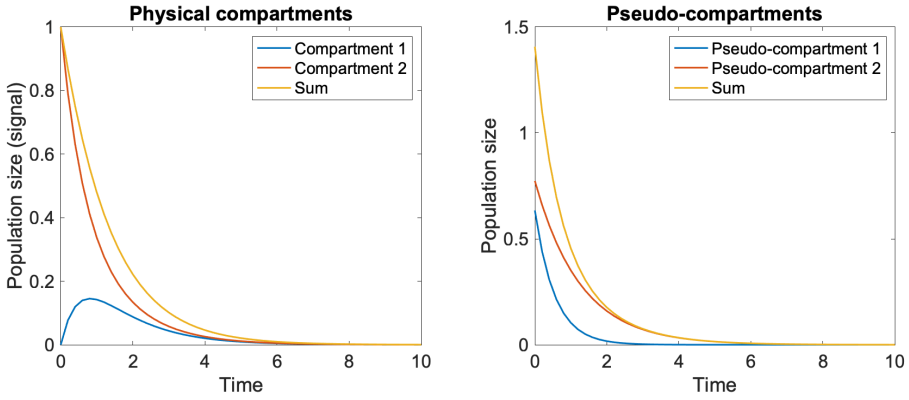

Figure 3.1: A two compartment system where $\mathbf{y_0} = [0, 1]$, $K = 0.5$, $R1 = 0.9$, and $R2 = 0.7$. The left plot show the signal/population size for the physical compartments, and the right plot show the corresponding populations for the pseudo-compartments obtained from diagonalization of the system matrix. The sum of the pseudo-compartment populations is not proportional to signal.

is not monotonous. (In relation to Eq. 3.1 the blue curve can be described as a linear combination of exponentials with both negative and positive coefficients, and violates the condition of shared sign coefficients). The corresponding populations in the pseudo-compartments (from matrix diagonalization) are both mono-exponential. However, as the figure shows, their total sum differs from the original total signal, and one must be aware that scaling of the individual populations is required to obtain the same sum.

Combining intermediate exchange with a pre-nulled compartment will always cause non-monotonic signal behaviour in a physical compartment. If the initial conditions instead are similar for the two compartments, e.g. $\mathbf{y_0} = [1, 1]$, there are many situations with intermediate exchange where monotony persists for the physical system. Figure 3.2 shows for which combinations of exchange and decay rates, monotony is applicable.

Overall, it is desired that the proposed data-driven solution reveals components of physical compartments as they are interpretable and representative of the tissue micro-structure. However, components of pseudo-compartments may also be informative and may well have predictive value. Putting aside situations of intermediate exchange there are truly many cases of MRI experiments where the assumptions are valid and signal components behave monotonously. It should be recalled that using a data-driven method with monotonicity constraints added is still less restrictive than using a mono/multi-exponential model.

Returning to the aim of this sub-project, it consisted of developing a decomposition method tailored to find realistic signal components from MRI measurements that are related to the underlying tissue structure. The developed extended NMF, named monotonous slope non-negative matrix factorization (msNMF), is suggested, and while it has now been reviewed in which situations the monotony constraints are suitable for signals of physical compartments, proving a relationship between the components and the underlying tissue structure is yet another task. When using an unsupervised technique, it is necessary to have e.g. histology data available for verification. The

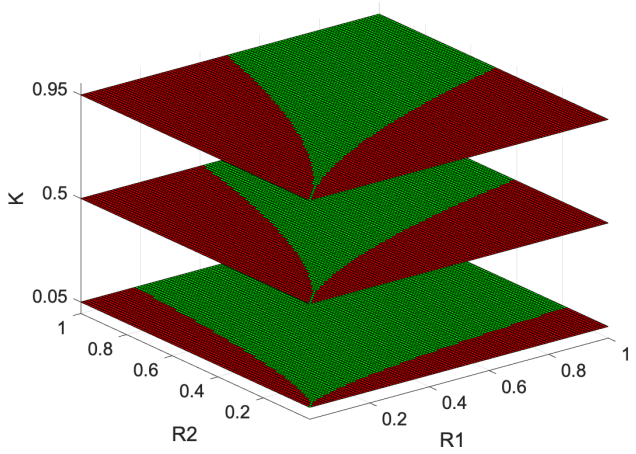

Figure 3.2: Combinations of K, R1 and R2, where monotony of the physical compartment populations is violated (red areas) or complied with (green areas). Three levels of exchange is presented for this overview: $K = [0.05, 0.5, 0.95]$. The situation is for equal initial conditions for the two compartments: $\mathbf{y_0} = [1, 1]$.

paper *Data-driven separation of MRI signal components for tissue characterization* demonstrates the msNMF using simulated data and two pre-clinical datasets of which one of them also included histological measurements. The paper is described in brief in the following section, and the full research article published in Journal of Magnetic Resonance (JMR) can be found in Appendix A.

3.1 PAPER I: DATA-DRIVEN SEPARATION OF MRI SIGNAL COMPONENTS FOR TISSUE CHARACTERIZATION

In this study, the developed msNMF was tested and demonstrated using data of ex vivo diffusion-weighted MRI measurements of a primate brain [75] and of in vivo multi-echo T_2-weighted "relaxometry" measurements of a rat spinal cord [76], respectively. In both cases, the total MRI signals were in compliance with the monotonicity constraints, and this was also expected for the partial signal components. The resulting multi-component data representation was evaluated using knowledge of local tissue microstructure and histology data. Before testing with pre-clinical data, the method was demonstrated using simulated data where the true data structure was known in advance. Simulated data was also appropriate for testing the method against standard and monotonous NMF and a typical model-based strategy using regularized NNLS.

3.1.1 *Methods*

The msNMF was implemented by adding constraints to the standard NMF, which decomposes the data matrix into a factor of *signal components* (**W**) and a factor holding the spatial distribution of these, i.e. *mixture maps* (**H**). In reference to the NMF algorithms described in subsection 2.3.2,

we used an ANLS method with the Frobenius norm as objective function (Gaussian noise model), reformulated as a quadratic programming problem for each mode. This made it straight-forward to implement the conditions implying monotony of the signals and their slopes as an inequality constraint for the optimization of **W**. Mathematical details of relevance are included in the paper. It is worth to mention that the objective function was extended to include a weight vector, such that it became possible to weight the individual observations with the initial signal amplitude. This was done to assign less value to measurements of low SNR (e.g. from voxels representing bone) in the optimization process.

3.1.2 *Results*

For the simulated data, the msNMF was able to reconstruct the underlying components with high accuracy, i.e. with high correlation to the true components, and it outperformed the standard and monotonous NMF (and the NNLS for high levels of Rician noise).

For the ex vivo diffusion-weighted data of a primate brain, the msNMF mixture maps (**H**) were compared for two datasets which differed by the diffusion time used during acquisition. The comparison was made by a simple subtraction between the two set of mixture maps, and the resulting map highlighted tissue areas of high cell-density (white matter, the cerebellum, and the visual cortex) with very convincing contrast. This indicates that extracted signal components are relevant for describing the underlying tissue micro-structure.

For the in vivo relaxometry data of rat spinal cords, the msNMF components were used to generate “quantitative” measures. One of the found signal components was a slowly decaying signal, consistent with edema water. The fraction of this, given by **H**, was used as an estimate of the edema water fraction (EWF) in the tissue. The estimated EWFs for all rats (which were divided into groups of different diets of an intra-myelinic edema-inducing toxin) were very similar to those estimated by the original publication of this data [76], and correlated well with histology (R^2 = 0.88). The high correlation with histological measures, strongly indicates that the msNMF signal components are related to the tissue structure.

DECOMPOSITION-BASED PREDICTION FRAMEWORK FOR ANALYSIS OF LONGITUDINAL MRI STUDIES

4

This part of the project is motivated by the introduction of the MR-linac systems, which truly creates the best conditions for tumor surveillance during the course of fractionated radiotherapy, and in that regard will increase the number of longitudinal MRI studies in the near future. An ultimate goal of analyzing data from such studies is to predict the effect of the treatment for a given tumor. Early prediction leaves time for treatment interventions. The problem is, that the current availability of longitudinal MRI studies for tumor surveillance is low, and among those available, most of them observes the time-resolved measurements independently in their analysis. A typical approach is to calculate the mean ADC value, or another quantitative metric, within the tumor ROI (region of interest) for each scan day and compare them to the baseline scan obtained before first treatment. Then a statistical analysis is used to test for group differences of the change from baseline between e.g. responders and non-responders [77–84]. Figure 4.1(a) is a graphical representation of such an approach. We have developed and explored an approach which differs in three main areas from the mentioned example.

The MR-linacs enables the acquisition of image data at each fraction, i.e. daily measurements, and it thus becomes less sensible to handle each measurement as independent. Our approach attempts to capture the early tumor tissue dynamics induced by the radiation by observing the variation or trend over multiple time-points instead of making pairwise comparisons with the baseline, which also quickly leads to a multiple comparisons problem.

Secondly, our developed framework is a *prediction* framework, which concludes with supervised learning to produce a model that predicts e.g. the tumor type or treatment outcome for an "unseen" observation. This deviates from the typical former approach (Fig. 4.1(a)) where group differences are presented based on known stratification. Proving statistical group differences is of course positive and informative, but unless a separation is shown as general, it does not guarantee that individual observations can be predicted. We directly test the predictability of extracted tumor features and include comparisons to empirical null distributions derived via random permutations tests to enable an unbiased performance evaluation of the prediction model.

Finally, the calculation of an ADC map or another quantitative metric is replaced with a data-driven decomposition. The benefits of data-driven analysis over model-based have been reviewed in the previous chapter, and in this context it is specifically advantageous that information is not lost to partial volume effects and that e.g. biased estimates are not passed on to the subsequent processing. Additionally, the interpretability and quantifiability of a metric such as the ADC becomes less relevant in combination with a subsequent prediction analysis, where the predictive value of the data is of most importance. A decomposition may consist of signal components corresponding to pseudo-compartments which might leave interpretation difficult, but it may

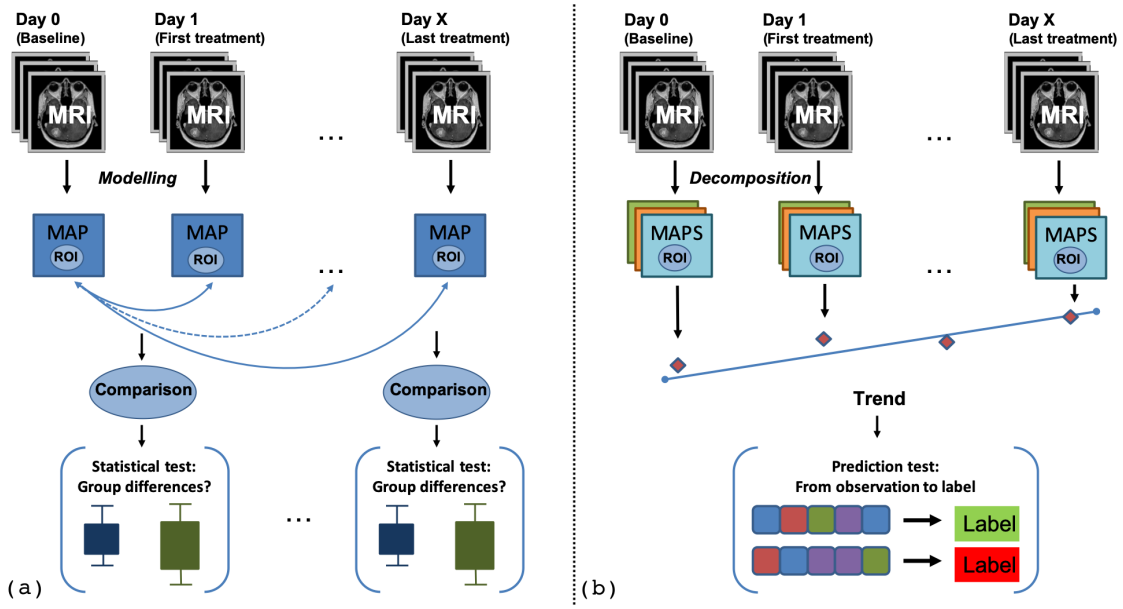

Figure 4.1: Analysis of longitudinal data. (a): A typical framework used in previous studies: a quantitative map is generated from the multi-constrast MRI data using a model. Then pair-wise comparisons between a measurement and the baseline is calculated, and finally statistical testing investigates differences between groups, e.g. responders and non-responders. (b): The suggested framework: the multi-constrast MRI data is decomposed into component maps. All measurements are used to capture a trend, a development over time, and this information is carried on to a prediction test where the group label is predicted for each observation.

still be predictive of the treatment outcome.

Figure 4.1(b) visualizes the principles of our developed framework juxtaposed to the graphical representation of earlier frameworks (Figure 4.1(a)). The discussed differences are emphasized. We aimed for developing a rather general framework, which is reflected in the graphical representation. No specific data type, decomposition technique, prediction analysis etc. are specified. Several types of MRI data are currently investigated for biomarkers predictive of tumor treatment response [8], and our framework is thus deliberately developed with a flexibility to choose different sub-techniques, e.g. different decomposition methods, resulting in a wide applicability.

In the manuscript *Decomposition-based framework for tumour classification and prediction of treatment response from longitudinal MRI*, presented in Appendix B, we demonstrate the developed framework using two previously published and very different datasets, and propose techniques for each of the analysis steps. Paper I demonstrated a value of the msNMF in terms of decomposing MR signals into informative tissue components, the msNMF is thus proposed as decomposition technique. Therefore, the following manuscript not only presents a prediction framework, but it also includes the first testing of the predictive capabilities of discriminative models based on msNMF-discovered tumor tissue components.

4.1 PAPER II: DECOMPOSITION-BASED FRAMEWORK FOR TUMOUR CLASSIFICATION AND PREDICTION OF TREATMENT RESPONSE FROM LONGITUDINAL MRI

In this paper, our proposed prediction framework designed for longitudinal MRI data is presented and demonstrated using two previously published datasets where subjects were scanned before, during and after the course of therapy. The first dataset consisted of multi-echo T_2-weighted scans of mice grafted with two different pancreatic tumor types [85]. It was used to test the framework for classification of the two tumor types, and further for prediction of the late tumor volume change after radiotherapy. The second dataset consisted of DW-MRI of human brain metastasis and the framework was used to classify the tumors into responders and non-responders [86].

4.1.1 *Methods*

After an initial data decomposition leading to msNMF component maps, histogram values (first order radiomics) for the target ROI were used to capture the tumor heterogeneity. The trend of these tumor features over time was obtained by fitting the time-resolved measurements to a linear regression curve. A Hubert loss function was used to increase robustness to outliers [87]. The fitted parameters (intercept and rate of change) were carried on to the prediction model, consisting of either a logistic regression (for binary outputs) or a multiple linear regression (for continuous outputs). The number of input features was reduced in a feature-selection process integrated in a nested cross-validation (CV) approach for model estimation. This means, overfitting was prevented by directly constraining the number of input features instead of implementing a sparse regularization term in the cost function. The model performance was accessed by comparing the prediction outcome to a null distribution generated from 5000 runs with permuted class labels.

4.1.2 *Results*

For the multi-echo T_2-weighted data, the two tumor types were more or less perfectly separated (AUC=0.99, P=0.0004), and the late tumor volume change (therapy response) was predicted with a significant correlation to the true values (correlation score=0.513, P=0.034). In both cases, the long-lived T_2-component in the tissue was an important feature to the prediction. For the DW-MRI data, the classification of metastases into responders and non-responders was encouraging although the accuracy did not reach statistical significance (AUC=0.74, P=0.065). The T_2-weighted part of the data (b=0 s/mm^2), which was entered explicitly as a feature in the model, appeared to carry the most valuable information, i.e. adding the diffusion-weighted data did not improve the outcome prediction in this case.

OPTIMIZATION OF THE DIFFUSION-WEIGHTED SPLICE SEQUENCE

5

Diffusion-weighted MRI is an important multi-contrast imaging technique within MRI-guided radiotherapy because mapping of the local water movement can reflect the micro-structure, specifically the local cell density. A high cell density is characteristic for tumors and a decrease in the tumor cell density is an expected response to radiation. DWI can thus provide useful information for both therapy planning and monitoring of the response during the course of therapy. In the Background Chapter two sequences for DWI were reviewed: The DW-EPI sequence, which is the current standard for DWI, and the SPLICE sequence, which combines a diffusion-weighted preparation with a RARE readout and overcomes the CPMG violation by utilizing the split echo acquisition principle. The problem of the EPI sequence is its relatively high sensitivity to field inhomogeinities, e.g. stemming from tissue susceptibility variations, which causes geometrical distortions in the images. This is highly inconvenient in the context of radiotherapy planning, which involves tumor delineation, where geometrical precision is crucial. The RARE readout of the SPLICE sequence makes it a geometrically accurate competitor to the DW-EPI sequence, but the SPLICE image quality is limited by its relatively low SNR and poor resolution.

Generally, SNR is limited for diffusion-weighted imaging because a fast imaging technique (low scan time) is needed to measure the diffusion-weighted signals, and because diffusion of spins costs signal. SNR of SPLICE is further limited by the prolonged readout gradient (i.e. prolonged interval between two refocusing pulses) increasing TE. Also in cases of preserved TE and scan duration, the split echo method reduces SNR compared to standard RARE: the division into two echos implies that only half the time is used on each of them. This increases the noise with a square root of two for each echo. However, averaging over two echoes will lower the noise with the same factor, i.e. the noise remains the same. The signal, on the other hand, is reduced because the average of two half-amplitude echoes results in a factor of two lower signal than one full-amplitude echo. As mentioned in subsection 2.2.2, the prolonged readout period in SPLICE also leaves more time for the relaxation-induced signal modulation taking place during readout, which leads to a poor voxel shape and thus blurred images. This part of the project is hence concerned with optimization of the single-shot SPLICE sequence to obtain better SNR and reduced blurring, as such improvements will make it a relevant candidate for DWI in MRI-guided radiotherapy. We have chosen not to focus on multi-shot sequence variations, such as the SPLICE PROPELLER [43], because they typically have a longer scan time and are more motion sensitive. Even though navigator echoes are used for phase corrections, there are local motion differences, and motion artifacts can still be a problem.

In general in MRI, there is a trade-off between the SNR and the resolution (voxel size), i.e. a higher resolution leads to a lower SNR for the same scan time. The SNR also generally depends on the voxel shape, or more intuitively on how the magnetization is spent in k-space during readout.

This is described via a signal modulation or k-space weighting function in the phase-encoding direction and corresponds to the point-spread-function (PSF) through the Fourier transform. The highest SNR is obtained when all magnetization is read out during its entire maximal lifetime T_2, i.e. when the total sum of the weighting function (area under curve) is maximized. However, such case may result in a poor PSF. An optimal PSF has a main peak with low full-width half-maximum and minimal side bands (as close to a delta function as possible). To accommodate both SNR and PSF requirements, we have employed an optimization approach where the SNR is maximized for a chosen PSF. It is utilized that varying the flip angles in the refocusing pulse train modulates the weighting function and thus the PSF. If the desired PSF can be obtained solely by appropriate flip angle variations, there is minimal SNR cost. The following paper presents the method and demonstrates the value of such an *acquisition weighting* approach for improving the image quality of the SPLICE sequence. Before we implemented the optimization strategy, we verified that EPG calculations were sufficient for simulating the k-space signal weighting. To this end, we made a thorough comparison to JEMRIS simulations and actual scans. The full manuscript is available in Appendix C and Supplementary Materials include the mentioned comparison.

5.1 Paper III: Optimized flip angle schemes for the diffusion-weighted SPLICE sequence

This study presents a method to optimize the variable flip angle scheme in a diffusion-weighted RARE-based sequence to improve image quality. Schemes of variable flip angles have been utilized before in optimization of RARE sequences, but to our knowledge not in combination with SPLICE or diffusion-weighting. When diffusion-weighting is aimed at, other weighting aspects such as the T_2-weighted contrast become less relevant, and the optimization can focus on maximizing the SNR for a given PSF without explicitly considering effects on the effective TE.

5.1.1 *Methods*

EPG calculations are used to estimate the acquisition weighting generated by a set of variable flip angles. A filter is calculated to compensate for the differences between this acquisition weighting and the desired k-space weighting (desired PSF) [88]. For normal linear image reconstruction and fixed PSF, the SNR only depends on the flip angle scheme via the filter's effect on noise, and the SNR is thus inversely proportional to the filter coefficients. The flip angles are optimized by maximizing SNR (minimizing filter coefficients) using an interior-point method.

5.1.2 *Results*

Scanning the brain of a healthy test subject with an optimized flip angle scheme and a reference scheme consisting of a constant flip angle (90°) showed a clear SNR improvement for the optimized scheme. The SNR gain was particularly evident for sequences with relatively long echo trains. ADC maps revealed that the SNR gain lowered the noise bias for high b-values resulting in more accurate ADC estimates.

6 DISCUSSION

A main goal of utilizing MRI in radiotherapy is to exploit its superior soft-tissue contrast for better visualisation of the target and relevant critical structures. This implies higher geometrical accuracy which can increase target control and/or reduce toxicity. Due to the versatility of MRI systems, biological information can also be measured and quantified with advanced imaging techniques. This holds the promise of biological guided radiotherapy, and fundamentally better individualization. One perspective is to have MRI biomarkers that potentially can identify which patients respond well and which patients respond poorly to the treatment [89]. The patient treatment may be adjusted based on this information, both during early or late phases of the radiotherapy course. The information can also be used as a decision-making tool to terminate the treatment midway, for example. The implementation of the MR-linacs paves the way for a large volume of new radiotherapy studies which can investigate how to optimally utilize MRI for that. Three initiatives have been presented for improving the usage of advanced MRI in radiotherapy studies, each assessing different aspects or limitations of biology-guided radiotherapy.

Specifically, the first sub-project presented a tool for analysing multi-contrast MRI data to obtain a tissue characterization, which is relevant in the search for MRI-based biomarkers and possibly also for biological target volume definition. The second sub-project presented a strategy to analyse longitudinal MRI measurements with the aim of finding prognostic or predictive biomarkers. The last sub-project presented a sequence design which improves the SPLICE sequence, and makes it a candidate as a geometrically accurate DWI technique.

With these novel tools at hand, how could a researcher plan a future MR-guided radiotherapy study? Considering this question will also answer how the developed tools can be effectuated and gain impact, which is the main motivation for me as a researcher. In the included papers, the technical feasibility of the methods is demonstrated and discussed, but I am aware that further steps and developments are needed for clinical translation, where practical limitations such as patient comfort, scan time, and computation time needs to be considered. The developed tools are currently ideal for use in designing studies that investigate the best imaging procedures and analyses for adaptive radiotherapy and response evaluation.

The metastasis study presented in Paper II and previously published by Mahmood *et al* [86, 90], is an example of a study which explores longitudinal DWI data for response evaluation by scanning the patients at each therapy fraction. It also represents a type of study which at the time of data collection (2013) was quite unique, but which in the near future will be very accessible. In the following I look into how my research can be used in the process of designing such a study today, i.e. can guide how to record and analyse data.

An appropriate starting point is to consider which sequences to use for data acquisition, which may depend on the tumor site and the specific application. In the metastasis study, DWI data was

used for tumor response evaluation and ROI delineation. More generally, DWI may enter in the course of radiotherapy as part of the planning phase, during the treatment phase for "biological" adaption of the dose plan "on the fly", and as part of the longitudinal tumor surveillance for response evaluation and prediction [35, 26]. In all cases geometrical distortions are undesired, but particularly for planning and adapting the treatment plan, geometrical accuracy is a high priority and the optimized SPLICE may be a better solution than the standard EPI-based sequence. Particularly, the optimized SPLICE is a useful sequence if the tumor site is near an air cavity, such as in head and neck cancer, where geometrical distortions in EPI can be prominent [17, 91]. Further testing against the "gold standard", the EPI-DWI, is a relevant step for implementing the SPLICE in the clinic more permanently. Currently, clinicians are used to inspect EPI images that have a particular contrast and image quality. This should not be overlooked when attempting to implement a new technique which may slightly alter the visual contrast. Potentially, a preliminary scan session including both an EPI-DWI and SPLICE sequence could be used to determine which images are preferred for visualizing the tissue diffusivities for a given site. A comparison to EPI-DWI could also be relevant for testing whether SPLICE data with optimized SNR is more useful in a prediction analysis.

When planning the DWI sequence (whether SPLICE- or EPI-based), choosing the b-values is an important consideration that should be made in light of the subsequent data processing. Effectuating the msNMF needs a range of contrast measurements for a given sequence type (the more the merrier), and in the case of DWI, this means obtaining multiple b-values ranging from low to relatively high values (e.g. 10 values ranging from 0-1000 s/mm^2). On the contrary, limitations on the scan time and gradient system needs to be considered and only the most necessary values should be acquired, thus avoiding too high b-values. The best choice of b-values can be difficult to know in advance as it depends on the tissue under investigation. This is also a challenge considering other post-processing methods, and the standard ADC calculation, which depends on the chosen b-values, may also benefit from optimising the b-value scheme [11].

A possible strategy to find an optimal b-value scheme is to use a preliminary scan session including many b-values together with the msNMF for extraction of informative components that can guide the study design. In more detail, a decomposition based on well-resolved signals can be considered the "ground truth" and be used in a "supervised" search of which b-values that are necessary for maintaining the right decomposition. Techniques within active machine learning such as uncertainty sampling and expected error reduction that can find the most valuable data-points [92], or simple methods such as *backward selection* are possible strategies for this [93]. Even though substantial new research dedicated to this is necessary to fully uncover the potential, the following figure exemplifies such an investigation using the diffusion-weighted data of a primate brain originally presented by Lundell *et al* [75]) and reused in Paper I. It consists of 12 b-values ranging from 242-4832 s/mm^2 and is therefore considered more suited for a demonstration of this aspect than the DWI data from the metastasis study, which consisted of eight non-equidistant b-values ranging from 0-800 s/mm^2.

To shortly recall the investigation from Paper I, a decomposition of the primate brain data

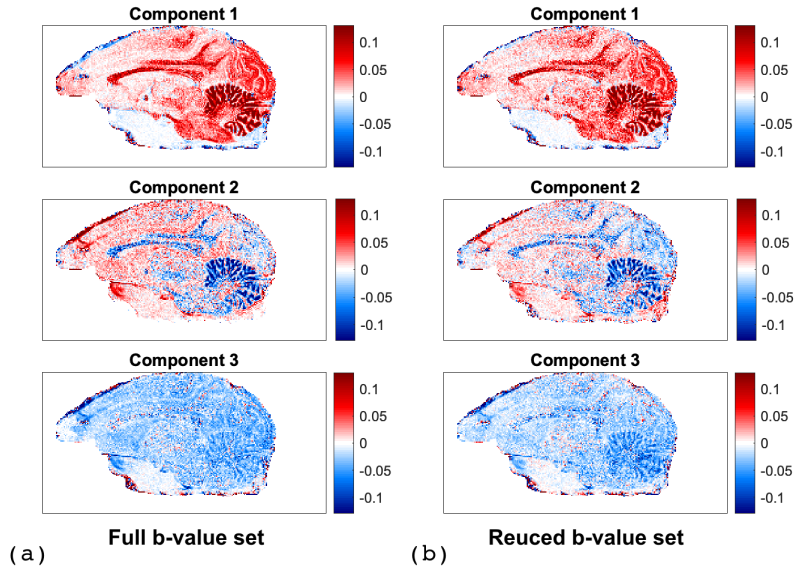

Figure 6.1: (a): The original difference map using the full b-value set. (b): The difference map when using the backwards eliminated b-value set consisting of three b-values: 242, 944, and 3681 s/mm^2. A maximum of RMSE=0.03 was used as threshold to stop the backward elimination. The resulting optimized b-value set gave a RMSE=0.027

revealed high cell-density areas by comparing component maps for two diffusion values via subtraction. These "difference" maps, one for each component, were considered the desired result in the b-value selection process, and the deviation from this (calculated as the root-mean squared error (RMSE)) was used as "cost" measure. Then a backward elimination strategy was used to repeatedly eliminate the b-value which resulted in the lowest cost. This continued until some RMSE threshold was reached. The result is presented in Figure 6.1 by showing the "difference" maps using the final set of b-values side by side with the original maps using all b-values. As can be seen, they are very similar, implying that the reduced set of three b-values appears adequate for the investigation of interest. The second highest b-value at 3681 s/mm^2 is included in this set, but is not clinically feasible to obtain in human DWI. However, the test demonstrates how to utilize the msNMF decomposition to guide sequence design details for similar scans/studies.

Returning to the hypothetical study planning, one must consider which multi-contrast MRI measurements to include in the protocol, which is especially important for exploiting the developed prediction framework to search for response biomarkers. While DWI is an obvious technique to include due to the biological link to cell density, the results of Paper II indicated that multi-echo T_2-weighted (relaxometry) data contains valuable information for a response prediction, possibly due to a link to tissue heterogeneity and necrosis. It could thus be interesting to include a multi-echo T_2-weighted sequence in the protocol. Several other multi-contrast techniques might also deserve to be explored (e.g. DCE-MRI), and though there is a limitation in terms of total scan time, the prediction framework is suited for many types of data and even for a combination of multiple data types.

In terms of the longitudinal measurement time schedule, the optimal case for the prediction framework is to obtain a baseline scan, a scan per day or per treatment fraction, and a follow-up scan, all of which are relatively easily accomplished with an MR-linac. There is a desire to achieve early response prediction (as soon as possible) in order to allow more time for therapy interventions. The importance of time-resolved measurements for capturing significant tumor dynamics can be studied in the same way as the demonstrated b-value selection. For example, one technique could be to run the prediction framework with fewer than the entire set of measurements and observe the prediction's uncertainty. The datasets included in Paper II was not ideal for such an investigation. The DWI metastases data was compromised by several missing measurements, and the mice dataset only included three to four measurements post radiation. In addition, the mice study imitated a stereotatic radiotherapy study, which was useful and relevant for framework testing, but which is not the main application here. The prediction framework is most of all intended for fractionated radiotherapy studies where the longitudinal study structure makes most sense. Overall, future studies and a prediction tool such as the developed framework are ideal for thoroughly investigating at which time-point a significant outcome prediction is possible, as well as which data features or combinations of data features are most predictive.

It has been shown through Paper I and Paper II that the msNMF can be used both as a stand-alone analysis for tissue characterization and in combination with further processing steps in a prediction framework. Using it for tissue characterization, it may be a beneficial tool in treatment planning, another aspect to consider when designing a radiotherapy study. As described in section 2.1, biological target volume and dose painting are hot topics within radiotherapy planning, and msNMF could be a relevant technique for segmentation of the tumor into compartments of different radio-sensitivities based on DWI data, for example. A research project has recently been initiated for improving tumor segmentation and delineation using biological information, and investigating the msNMF is part of the scope [94]. Using biological information for the initial treatment planning is one thing, another is to use it during treatment delivery to adapt the treatment plan. Currently, the treatment plan is adapted to the "anatomy-of-the-day" obtained with either cone-beam CT scans in standard linacs or with anatomical MRI scans in MR-linacs. In the future, it is desired to use biological information to effectuate "biological adaptive treatments" [9]. In such a setup, there is a high demand on the workflow and only a short time for analysis and decision making since a patient is waiting in the scanner. Thus, even though the msNMF may be relevant for tissue characterization and segmentation, practical conditions such as establishment of components in advance needs to be considered before implementing it into a workflow.

Furthermore, changing researchers' and clinicians' usage of standard model-based processing, which produces quantitatively interpretable measurements, to a less interpretable decomposition, may be challenging. If so, the msNMF could be utilized to generate priors for a subsequent model-based analysis or to separate noise from the data before further processing, reducing the risk of getting misleading or biased estimates. Paper I showed how the msNMF is relatively robust to noise and even includes the possibility of "modeling" unwanted signal behavior, e.g.

a noise-introduced bias, as a component. With the simulated data, it was shown that when sum-of-squares reconstruction is applied on coil-array data and the Rician noise distribution becomes distinct, adding a monotonous growing component to model the positive signal bias improved the decomposition.

Overall, the developed methods presented throughout this project has the potential for improving the usage of MRI within radiotherapy for better cancer management. Both the optimized DWI acquisition and the alternative data-driven processing suggested, aim at a better utilization of biological information from advanced MRI data to increase the opportunity to establish biological treatment planning/adaption and generate predictive biomarkers. The three research papers have demonstrated the technical feasibility of the proposed strategies, and perspectives on how to use them to form new and improved radiotherapy studies have now been discussed. All in all, steps have been taken towards a higher degree of personalized therapy and an improved treatment efficacy.

7 CONCLUSION

This dissertation presents three specific initiatives for advancement of MRI-guided radiotherapy.

A data-driven technique, the msNMF, has been developed for decomposition of multi-contrast MRI data to characterize the tissue. The method is insensitive to partial volume effects and relies on realistic assumptions on the MR signal behavior which facilitates an interpretable decomposition. A thorough demonstration using first simulated data and then two clinical datasets showed the potential of the method for identifying MR signal features specific to the underlying tissue structure. Hence, the method is a promising candidate for tumor tissue characterization which is relevant in treatment planning or adaption and for identifying both prognostic and predictive biomarkers.

A prediction framework has been established for analysis of repeated MRI measurements as a potential tool for finding biomarkers within the field of cancer treatment. The framework is suited for the increasing volume of longitudinal MRI studies prompted by the recent introduction of MR-linacs, and aims at using all time-resolved measurements simultaneously in order to capture dynamic tumor behaviour that may be predictive of the treatment outcome. The framework is based on data-driven decomposition of multi-contrast MRI data and was demonstrated using the newly developed msNMF for that. By showing prediction of the tumor treatment outcome using T_2-weighted multi-echo data as an example application, the merits of the msNMF for tumor tissue signal decomposition was also demonstrated.

Finally, the single-shot SPLICE sequence has been optimized to improve the quality of diffusion-weighted images for RT purposes. By optimizing the refocusing flip angle scheme in the RARE-based readout, it was possible to increase SNR for a controlled level of blurring. In combination with an increased geometric accuracy relative to EPI, this improved image quality makes the SPLICE more attractive for radiotherapy planning.

7.1 OUTLOOK

Future investigations of the msNMF may concern both its usage and technical advancements. It would be interesting to investigate it for tumor segmentation in radiotherapy planning, which has already been initiated in a different project, as mentioned in the discussion. The aim is to use the component maps in a segmentation analysis to find areas of the tumor with higher or lower radio-sensitivity. A big challenge here is to define the "truth" in the validation process.

The msNMF could also be tested more thoroughly for the application of denoising MRI data [95], e.g. for reducing the signal-dependent noise bias present for MRI magnitude data. Adding a growing component to "absorb" the noise bias was shown effective for simulated data. It could be interesting to test the strategy using real data and further investigate its effectiveness,

if the spatial distribution of such bias is non-uniform, i.e. if the method is upgraded to constrain mixture maps according to different tissue classes, for example.

Other future technical advancements include optimizing the msNMF in regard to computation time, which is relevant for clinical translation. Settings which affects the computation time and could be further considered are the choice of optimization algorithm, initialization strategy, stopping criteria and the size of the input data. Parameters such as the input data batch size and number of iterations could be tuned, for example.

For the prediction framework, an obvious next step is simply to obtain more experience with the method and apply it to larger datasets, preferably in a setting were the data volume is sufficient to reserve part of the data as a completely independent test set, perhaps even from another site. With improved data availability it will also be easier to test e.g. the importance of each of the time-resolved measurements as touched upon in the Discussion. In the search for predictive or prognostic biomarkers, it would be worthwhile to further investigate the link between T_2-weighted relaxation components and tumor necrosis suggested by the framework demonstration.

For the SPLICE optimization, next steps are first of all to demonstrate the image improvement for more datasets. Specifically, it is relevant to show an improvement for scans where the advantage of acceleration techniques is utilized to increase the nominal resolution. The current data example showed only a slight SNR improvement when acceleration techniques were used to reduce the ETL to 35. In the same context, scan time could also be considered as this is relevant to minimize for clinical translation. The recording of multiple b-values takes time, but redundancy in the data may be possible to utilize with more involved acceleration-reconstruction strategies.

Future initiatives could also investigate if the sequence design can be further improved. Currently, a prolonged gradient separates the two echo families and the two echos are simply summed without considering their individual contributions. However, the individual echoes have different intensities and maybe there are ways to utilize this during the readout. Norris *et al*, for example, have suggested a method for utilizing the echo of most signal, i.e. a selective parity approach [96], and showed a strategy to combine it with acceleration methods in diffusion weighted imaging [97]. The strategy seems compatible with implementation of the optimized variable flip angle schemes presented here.

Taken together, the advances presented in this thesis are promising for advancing future radiotherapy and can hopefully improve treatment for patients with cancer.

BIBLIOGRAPHY

[1] Michael C Joiner, Albert J Van der Kogel, and Gordon G Steel. *Basic Clinical Radiobiology, chapter 1 (2009).pdf*. CRC press, 2009.

[2] Uwe Oelfke. Image Guided Radiotherapy. In Yves Lemoigne and Alessandra Caner, editors, *Radiotherapy and Brachytherapy*, pages 113–125, Dordrecht, 2009. Springer Netherlands.

[3] B. W. Raaymakers, I. M. Jürgenliemk-Schulz, G. H. Bol, M. Glitzner, A. N.T.J. Kotte, B. Van Asselen, J. C.J. De Boer, J. J. Bluemink, S. L. Hackett, M. A. Moerland, S. J. Woodings, J. W.H. Wolthaus, H. M. Van Zijp, M. E.P. Philippens, R. Tijssen, J. G.M. Kok, E. N. De Groot-Van Breugel, I. Kiekebosch, L. T.C. Meijers, C. N. Nomden, G. G. Sikkes, P. A.H. Doornaert, W. S.C. Eppinga, N. Kasperts, L. G.W. Kerkmeijer, J. H.A. Tersteeg, K. J. Brown, B. Pais, P. Woodhead, and J. J.W. Lagendijk. First patients treated with a 1.5 T MRI-Linac: Clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. *Physics in Medicine and Biology*, 62(23):L41–L50, 2017.

[4] S Mutic, D Low, T Chmielewski, G Fought, M Hernandez, I Kawrakow, A Sharma, S Shvartsman, and J Dempsey. TU-H-BRA-08: The Design and Characteristics of a Novel Compact Linac-Based MRI Guided Radiation Therapy (MR-IGRT) System. *Medical Physics*, 2016.

[5] C. Clifton Ling, John Humm, Steven Larson, Howard Amols, Zvi Fuks, Steven Leibel, and Jason A. Koutcher. Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. *International Journal of Radiation Oncology Biology Physics*, 47(3):551–560, 2000.

[6] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2):228–247, 2009.

[7] Stephen J. McMahon and Kevin M. Prise. Mechanistic modelling of radiation responses. *Cancers*, 11(2), 2019.

[8] Vincent Grégoire, Matthias Guckenberger, Karin Haustermans, Jan J.W. Lagendijk, Cynthia Ménard, Richard Pötter, Ben J. Slotman, Kari Tanderup, Daniela Thorwarth, Marcel van Herk, and Daniel Zips. Image guidance in radiation therapy for better cure of cancer. *Molecular Oncology*, 14(7):1470–1491, 2020.

[9] Petra J. van Houdt, Yingli Yang, and Uulke A. van der Heide. Quantitative Magnetic Resonance Imaging for Biological Image-Guided Adaptive Radiotherapy. *Frontiers in Oncology*, 10(January):1–9, 2021.

[10] David Eriksson and Torgny Stigbrand. Radiation-induced cell death mechanisms. *Tumor Biology*, 31(4):363–372, 2010.

[11] Anwar R. Padhani, Guoying Liu, Dow Mu-Koh, Thomas L. Chenevert, Harriet C. Thoeny, Taro Takahara, Andrew Dzik-Jurasz, Brian D. Ross, Marc Van Cauteren, David Collins, Dima A. Hammoud, Gordon J.S. Rustin, Bachir Taouli, and Peter L. Choyke. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia*, 11(2):102–125, 2009.

[12] Serena Satta, Miriam Dolciami, Veronica Celli, Francesca Di Stadio, Giorgia Perniola, Innocenza Palaia, Angelina Pernazza, Carlo Della Rocca, Stefania Rizzo, Carlo Catalano, Silvia Capuani, and Lucia Manganaro. Quantitative diffusion and perfusion MRI in the evaluation of endometrial cancer: validation with histopathological parameters. *The British Journal of Radiology*, 94(1125):20210054, 2021.

[13] Concepción González Hernando, Laura Esteban, Teresa Cañas, Enrique Van Den Brule, and Miguel Pastrana. The role of magnetic resonance imaging in oncology. *Clinical and Translational Oncology*, 12(9):606–613, 2010.

[14] Nikolaos Kartalis, Georgios C. Manikis, Louiza Loizou, Nils Albiin, Frank G. Zöllner, Marco Del Chiaro, Kostas Marias, and Nikolaos Papanikolaou. Diffusion-weighted MR imaging of pancreatic cancer: A comparison of mono-exponential, bi-exponential and non-Gaussian kurtosis models. *European Journal of Radiology Open*, 3:79–85, 2016.

[15] Y. Wang, S. Tadimalla, R. Rai, J. Goodwin, S. Foster, G. Liney, L. Holloway, and A. Haworth. Quantitative MRI: Defining repeatability, reproducibility and accuracy for prostate cancer imaging biomarker development. *Magnetic Resonance Imaging*, 77:169–179, 2021.

[16] Sara Leibfarth, René M. Winter, Heidi Lyng, Daniel Zips, and Daniela Thorwarth. Potentials and challenges of diffusion-weighted magnetic resonance imaging in radiotherapy. *Clinical and Translational Radiation Oncology*, 13:29–37, 2018.

[17] Tim Schakel, Johannes M. Hoogduin, Chris H.J. Terhaard, and Marielle E.P. Philippens. Technical Note: Diffusion-weighted MRI with minimal distortion in head-and-neck radiotherapy using a turbo spin echo acquisition method: Diffusion-weighted. *Medical Physics*, 44(8):4188–4193, 2017.

[18] M.L; Bernstein, Kevin F King, and Xiaohong J Zhou. *Handbook of MRI Pulse Sequences*. Number 1. Elsevier, 2014.

[19] Moataz Reda, Alexander F. Bagley, Husam Y. Zaidan, and Wassana Yantasee. Augmenting the therapeutic window of radiotherapy: A perspective on molecularly targeted therapies and nanomaterials. *Radiotherapy and Oncology*, 150:225–235, 2020.

[20] Laura Beaton, Steve Bandula, Mark N. Gaze, and Ricky A. Sharma. How rapid advances in imaging are defining the future of precision radiation oncology. *British Journal of Cancer*, 120(8):779–790, 2019.

[21] David Jaffray, Patrick Kupelian, Toufik Djemil, and Roger M. Macklis. Review of image-guided radiation therapy. *Expert Review of Anticancer Therapy*, 7(1):89–103, 2007.

[22] Kavitha Srnivasan, Mohammad Mohammadi, and Justin Shepherd. Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: A review. *Polish Journal of Radiology*, 79:181–193, 2014.

[23] Sevimir Özdemir and Ömer Uzel. Adaptif radyoterapi. *Turk Onkoloji Dergisi*, 28(3):136–141, 2013.

[24] Bilgehan Sahin, Teuta Zoto Mustafayev, Gorkem Gungor, Gokhan Aydin, Bulent Yapici, Banu Atalar, and Enis Ozyar. First 500 Fractions Delivered with a Magnetic Resonance-guided Radiotherapy System: Initial Experience. *Cureus*, 11(12):1–10, 2019.

[25] Hans Georg Menzel. The international commission on radiation units and measurements. *Journal of the ICRU*, 10(1):1–106, 2010.

[26] Uulke A. van der Heide, Antonetta C. Houweling, Greetje Groenendaal, Regina G.H. Beets-Tan, and Philippe Lambin. Functional MRI for radiotherapy dose painting. *Magnetic Resonance Imaging*, 30(9):1216–1223, 2012.

[27] Piet Dirix, Vincent Vandecaveye, Frederik De Keyzer, Sigrid Stroobants, Robert Hermans, and Sandra Nuyts. Dose painting in radiotherapy for head and neck squamous cell carcinoma: Value of repeated functional imaging with 18F-FDG PET, 18F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. *Journal of Nuclear Medicine*, 50(7):1020–1027, 2009.

[28] Denis Le Bihan and Eric Breton. Imagerie de diffusion in-vivo par résonance magnetique nucléaire. *Comptes-Rendus de l'Académie des sciences*, 1985.

[29] J. Duclaux. Le mouvement brownien et la formule d'Einstein. *Journal de Physique et le Radium*, 1940.

[30] Andrew L. Alexander, Jay S. Tsuruda, and Dennis L. Parker. Elimination of eddy current artifacts in diffusion-weighted echo-planar images: The use of bipolar gradients. *Magnetic Resonance in Medicine*, 38(6):1016–1021, 1997.

[31] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1):288–292, 1965.

[32] P T Callaghan, C D Eccles, and Y Xia. {NMR} microscopy of dynamic displacements: k-space and q-space imaging. *Journal of Physics E: Scientific Instruments*, 21(8):820–822, aug 1988.

[33] Daniel M Patterson, Anwar R Padhani, and David J Collins. Technology Insight: water diffusion MRI—a potential new biomarker of response to cancer therapy. *Nature Clinical Practice Oncology*, 5(4):220–233, 2008.

[34] M. Poustchi-Amin, S. A. Mirowitz, J. J. Brown, R. C. McKinstry, and T. Li. Principles and applications of echo-planar imaging: a review for the general radiologist. *Radiographics : a review publication of the Radiological Society of North America, Inc*, 21(3):767–779, 2001.

[35] Tim Schakel, Boris Peltenburg, Jan Willem Dankbaar, Carlos E. Cardenas, Michalis Aristophanous, Chris H.J. Terhaard, Johannes M. Hoogduin, and Marielle E.P. Philippens. Evaluation of diffusion weighted imaging for tumor delineation in head-and-neck radiotherapy by comparison with automatically segmented 18F-fluorodeoxyglucose positron emission tomography. *Physics and Imaging in Radiation Oncology*, 5(December 2017):13–18, 2018.

[36] J. Hennig, A. Nauerth, and H. Friedburg. RARE imaging: A fast imaging method for clinical MR. *Magnetic Resonance in Medicine*, 3(6):823–833, 1986.

[37] Robert W. Brown, Yu-Chung N. Cheng, E. Mark Haacke, Michael R Thompson, and Ramesh Venkatesan. *Fast Imaging in the Steady State*, chapter 18, pages 447–510. John Wiley & Sons, Ltd, 2014.

[38] H. Y. Carr and E. M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94(3):630–638, 1954.

[39] S. Meiboom and D. Gill. Modified spin-echo method for measuring nuclear relaxation times. *The review of scientific instruments*, 29(8), 1958.

[40] E L Hahn. Spinechoes, 1950.

[41] Eric K Gibbons, Patrick Le Roux, Shreyas S. Vasanawala, John M. Pauly, and Adam B. Kerr. Body diffusion weighted imaging using non-CPMG fast spin echo. *IEEE Trans Med Imaging.*, 36(2):549–559, 2018.

[42] Fritz Schick. SPLICE: Sub-second diffusion-sensitive MR imaging using a modified fast spin-echo acquisition mode. *Magnetic Resonance in Medicine*, 38(4):638–644, 1997.

[43] Jie Deng, Reed A. Omary, and Andrew C. Larson. Multishot diffusion-weighted SPLICE PROPELLER MRI of the abdomen. *Magnetic Resonance in Medicine*, 59(5):947–953, 2008.

[44] J. Hennig. Multiecho imaging sequences with low refocusing flip angles. *Journal of Magnetic Resonance (1969)*, 78(3):397–407, 1988.

[45] Juergen Hennig, Matthias Weigel, and Klaus Scheffler. Calculation of Flip Angles for Echo Trains with Predefined Amplitudes with the Extended Phase Graph (EPG)-Algorithm: Principles and Applications to Hyperecho and TRAPS Sequences. *Magnetic Resonance in Medicine*, 51(1):68–80, 2004.

[46] Matthias Weigel. Extended phase graphs: Dephasing, RF pulses, and echoes - Pure and simple. *Journal of Magnetic Resonance Imaging*, 41(2):266–295, 2015.

[47] Alessandro Sbrizzi, Hans Hoogduin, Joseph V. Hajnal, Cornelis A.T. van den Berg, Peter R. Luijten, and Shaihan J. Malik. Optimal control design of turbo spin-echo sequences with applications to parallel-transmit systems. *Magnetic Resonance in Medicine*, 77(1):361–373, 2017.

[48] Tony Stöcker, Kaveh Vahedipour, Daniel Pflugfelder, and N. Jon Shah. High-performance computing MRI simulations. *Magnetic Resonance in Medicine*, 64(1):186–193, 2010.

[49] Denis Le Bihan and Mami Iima. Diffusion magnetic resonance imaging: What water tells us about biological tissues. *PLoS Biology*, 13(7):1–13, 2015.

[50] Sergey Morozov, Kristina Sergunova, Alexey Petraikin, Ekaterina Akhmad, Stanislav Kivasev, Dmitry Semenov, Ivan Blokhin, Igor Karpov, Anton Vladzymyrskyy, and Alexander Morozov. Diffusion processes modeling in magnetic resonance imaging. *Insights into Imaging*, 11(1), 2020.

[51] Francesca Fornasa. Diffusion-weighted magnetic resonance imaging: What makes water run fast or slow? *Journal of Clinical Imaging Science*, 1(1):1–7, 2011.

[52] Alard Roebroeck, Karla L. Miller, and Manisha Aggarwal. Ex vivo diffusion MRI of the human brain: Technical challenges and recent advances. *NMR in Biomedicine*, 32(4):1–14, 2019.

[53] Jens H. Jensen, Joseph A. Helpern, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005.

[54] Daniel Rueckert and Julia A. Schnabel. Model-Based and Data-Driven Strategies in Medical Image Computing. *Proceedings of the IEEE*, 108(1):110–124, 2020.

[55] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[56] Charles L. Lawson and Richard J. Hanson. *Solving Least Squares Problems*. 1995.

[57] Mark D. Does. Inferring brain tissue composition and microstructure via MR relaxometry. *NeuroImage*, 182(December 2017):136–148, 2018.

[58] V. Wiggermann, I. M. Vavasour, S. H. Kolind, A. L. MacKay, G. Helms, and A. Rauscher. Non-negative least squares computation for in vivo myelin mapping using simulated multi-echo spin-echo T2 decay data. *NMR in Biomedicine*, XXX(Xx):1–14, 2020.

[59] Dmitry S. Novikov, Valerij G. Kiselev, and Sune N. Jespersen. On modeling. *Magnetic Resonance in Medicine*, 79(6):3172–3193, 2018.

[60] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 1933.

[61] Philip M. Kim and Bruce Tidor. Subsystem identification through dimensionality reduction of large-scale gene expression data. *Genome Research*, 13(7):1706–1718, 2003.

[62] Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999.

[63] Nicolas Gillis. The Why and How of Nonnegative Matrix Factorization. *arXiv*, 7 Mar:1–25, 2014.

[64] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, and Shun-ichi Amari. *Nonnegative Matrix and Tensor Factorizations Applications to Exploratory Multiway Data Analysis and Blind Source Separation*. John Wiley and Sons Inc., Tokoy, 2009.

[65] Daniel D. Lee and H. Sebastian Seung. Algorithms for Non-negative Matrix Factorization. *Journal of Engineering Science and Technology Review*, 9(6):82–86, 2016.

[66] Margaret E Daube-Witherspoon and Gerd Muehllehner. An Iterative Image Space Reconstruction Algorthm Suitable for Volume ECT. *IEEE Transactions on Medical Imaging*, 5(2):61–66, 1986.

[67] Rafal Zdunek and Andrzej Cichocki. Nonnegative matrix factorization with quadratic programming. *Neurocomputing*, 71(10-12):2309–2320, 2008.

[68] Nirav Bhatt and Arun Ayyar. Monotonous (Semi-) Nonnegative Matrix Factorization. In *Proceedings of the Second ACM IKDD Conference on Data Sciences*, pages 92–97, 2015.

[69] Yu Xiong Wang and Yu Jin Zhang. Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering*, 25(6):1336–1353, 2013.

[70] Andrzej Cichocki, Rafal Zdunek, and Shun-ichi Amari. Hierarchical ALS Algorithms for Nonnegative Matrix and 3D Tensor Factorization. *Independent Component Analysis and Signal Separation*, pages 169–176, 2007.

[71] David Freedman. *Statistical Models: Theory and Practice*. 2005.

[72] Hoss Belyadi and Alireza Haghighat. *Supervised learning*. Number Ml. Elsevier Inc., 2021.

[73] Gavin C. Cawley and Nicola L.C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107, 2010.

[74] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando De Freitas. Taking the human out of the loop: A review of Bayesian optimization, 2016.

[75] H Lundell, M Nilsson, T B Dyrby, G. J.M. Parker, P. L.Hubbard Cristinacce, F. L. Zhou, D Topgaard, and S Lasič. Multidimensional diffusion MRI with spectrally modulated gradients reveals unprecedented microstructural detail. *Scientific Reports*, 9(1):1–12, 2019.

[76] Kevin D Harkins, William M Valentine, Daniel F Gochberg, and Mark D Does. In-vivo multi-exponential T 2 , magnetization transfer and quantitative histology in a rat model of intramyelinic edema. *NeuroImage: Clinical*, 2:810–817, 2013.

[77] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, and Rasmus H. Hansen. Repeated diffusion MRI reveals earliest time point for stratification of radiotherapy response in brain metastases. *Physics in Medicine and Biology*, 62(8):2990–3002, 2017.

[78] Ying Shi Sun, Yong Cui, Lei Tang, Li Ping Qi, Ning Wang, Xiao Yan Zhang, Kun Cao, and Xiao Peng Zhang. Early evaluation of cancer response by a new functional biomarker: Apparent diffusion coefficient. *American Journal of Roentgenology*, 197(1):23–29, 2011.

[79] Lan Wang, Lihong Liu, Chun Han, Shutang Liu, Hua Tian, Zhensheng Li, Xuejiao Ren, Gaofeng Shi, Qi Wang, and Guangda Wang. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. *Radiotherapy and Oncology*, 121(2):246–251, 2016.

[80] Gang Cai, Ye Xu, Ji Zhu, Wei Lie Gu, Shuai Zhang, Xue Jun Ma, San Jun Cai, and Zhen Zhang. Diffusion-weighted magnetic resonance imaging for predicting the response of rectal cancer to neoadjuvant concurrent chemoradiation. *World Journal of Gastroenterology*, 19(33):5520–5527, 2013.

[81] Patrick A. Hein, Christian Kremser, Werner Judmaier, Jürgen Griebel, Karl Peter Pfeiffer, Alfons Kreczy, Eugen B. Hug, Peter Lukas, and Alexander F. DeVries. Diffusion-weighted magnetic resonance imaging for monitoring diffusion changes in rectal carcinoma during combined, preoperative chemoradiation: Preliminary results of a prospective study. *European Journal of Radiology*, 45(3):214–222, 2003.

[82] Ying Liu, Haoran Sun, Renju Bai, and Zhaoxiang Ye. Time-window of early detection of response to concurrent chemoradiation in cervical cancer by using diffusion-weighted MR imaging: A pilot study. *Radiation Oncology*, 10(1):1–8, 2015.

[83] Marcel A. van Schie, Petra J. van Houdt, Ghazaleh Ghobadi, Floris J. Pos, Iris Walraven, Hans C.J. de Boer, Cornelis A.T. van den Berg, Robert Jan Smeenk, Linda G.W. Kerkmeijer, and Uulke A. van der Heide. Quantitative MRI Changes During Weekly Ultra-Hypofractionated Prostate Cancer Radiotherapy With Integrated Boost. *Frontiers in Oncology*, 9(December), 2019.

[84] Ramesh Paudyal, Jung Hun Oh, Nadeem Riaz, Praveen Venigalla, Jingao Li, Vaios Hatzoglou, Jonathan Leeman, David Aramburu Nunez, Yonggang Lu, Joseph O. Deasy, Nancy Lee, and Amita Shukla-Dave. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. *Journal of Magnetic Resonance Imaging*, 45(4):1013–1023, 2017.

[85] Michal R. Tomaszewski, William Dominguez-Viqueira, Antonio Ortiz, Yu Shi, James R. Costello, Heiko Enderling, Stephen A. Rosenberg, and Robert J. Gillies. Heterogeneity analysis of MRI T2 maps for measurement of early tumor response to radiotherapy. *NMR in Biomedicine*, (September):1–12, 2020.

[86] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, Giske F. Opheim, and Rasmus H. Hansen. The effect of region of interest strategies on apparent diffusion coefficient assessment in patients treated with palliative radiation therapy to brain metastases. *Acta Oncologica*, 54(9):1529–1534, 2015.

[87] Peter J. Huber. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.

[88] Li Zhao, Ching Di Chang, and David C. Alsop. Controlling T2 blurring in 3D RARE arterial spin labeling acquisition through optimal combination of variable flip angles and k-space filtering. *Magnetic Resonance in Medicine*, 80(4):1391–1401, 2018.

[89] Julianne M. Pollard, Zhifei Wen, Ramaswamy Sadagopan, Jihong Wang, and Geoffrey S. Ibbott. The future of image-guided radiotherapy will be MR guided. *British Journal of Radiology*, 90(1073), 2017.

[90] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, and Rasmus H. Hansen. Ultra-early apparent diffusion coefficient change indicates irradiation and predicts radiotherapy outcome in brain metastases. *Acta Oncologica*, 56(11):1651–1653, 2017.

[91] Tim Schakel, Johannes M. Hoogduin, Chris H.J. Terhaard, and Marielle E.P. Philippens. Diffusion weighted MRI in head-and-neck cancer: Geometrical accuracy. *Radiotherapy and Oncology*, 109(3):394–397, 2013.

[92] Burr Settles. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2012.

[93] Ian H. Witten, Eibe Frank, and Mark a. Hall. *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition.* Elsevier, 2011.

[94] Anne Louise Højmark Bisgaard, Carsten Brink, Maja Lynge Fransen, Tine Schytte, Claus P. Behrens, Ivan Vogelius, Henrik Dahl Nissen, and Faisal Mahmood. Robust extraction of biological information from diffusion-weighted magnetic resonance imaging during radiotherapy using semi-automatic delineation. *Physics and Imaging in Radiation Oncology*, 21(September 2021):146–152, 2022.

[95] Oliver J. Gurney-Champion, David J. Collins, Andreas Wetscherek, Mihaela Rata, Remy Klaassen, Hanneke W.M. Van Laarhoven, Kevin J. Harrington, Uwe Oelfke, and Matthew R. Orton. Principal component analysis for fast and model-free denoising of multi b-value diffusion-weighted MR images. *Physics in Medicine and Biology*, 64(10):ab1786, 2019.

[96] David G. Norris. Selective parity RARE imaging. *Magnetic Resonance in Medicine*, 58(4):643–649, 2007.

[97] Aidin Arbabi, Vitaly Khlebnikov, David G. Norris, and José P. Marques. Robust and motion-insensitive approach to diffusion-weighted half-fourier acquisition single-shot turbo spin-echo imaging. In *ISMRM*, page 1039, 2022.

Appendices

PAPER I

A

TITLE

Data-driven separation of MRI signal components for tissue characterization

AUTHORS

Rahbek, Sofie; Madsen, Kristoffer H.; Lundell, Henrik; Mahmood, Faisal; Hanson, Lars G.

JOURNAL

Journal of Magnetic Resonance

YEAR

2021

PUBLICATION HISTORY

Submission date: 8 March 2021
Acceptance date: 2 November 2021

Journal of Magnetic Resonance 333 (2021) 107103

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Data-driven separation of MRI signal components for tissue characterization ☆

Sofie Rahbek [a], Kristoffer H. Madsen [b,c], Henrik Lundell [b], Faisal Mahmood [d,e], Lars G. Hanson [a,b,*]

[a] Department of Health Technology, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
[b] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, 2650, Denmark
[c] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
[d] Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital, Odense C 5000, Denmark
[e] Department of Clinical Research, University of Southern Denmark, Odense 5000, Denmark

ARTICLE INFO

Article history:
Received 8 March 2021
Revised 14 October 2021
Accepted 2 November 2021
Available online 5 November 2021

Keywords:
Magnetic resonance imaging
non-negative matrix factorization
Data-driven decomposition
Monotonous slope
Tissue characterization
Diffusion
Relaxometry

ABSTRACT

Purpose: MRI can be utilized for quantitative characterization of tissue. To assess e.g. water fractions or diffusion coefficients for compartments in the brain, a decomposition of the signal is necessary. Imposing standard models carries the risk of estimating biased parameters if model assumptions are violated. This work introduces a data-driven multicomponent analysis, the monotonous slope non-negative matrix factorization (msNMF), tailored to extract data features expected in MR signals.
Methods: The msNMF was implemented by extending the standard NMF with monotonicity constraints on the signal profiles and their first derivatives. The method was validated using simulated data, and subsequently applied to both ex vivo DWI data and in vivo relaxometry data. Reproducibility of the method was tested using the latter.
Results: The msNMF recovered the multi-exponential signals in the simulated data and showed superiority to standard NMF (based on the *explained variance*, *area under the* ROC *curve*, and *coefficient of variation*). Diffusion components extracted from the DWI data reflected the cell density of the underlying tissue. The relaxometry analysis resulted in estimates of edema water fractions (EWF) highly correlated with published results, and demonstrated acceptable reproducibility.
Conclusion: The msNMF can robustly separate MR signals into components with relation to the underlying tissue composition, and may potentially be useful for e.g. tumor tissue characterization.

1. Introduction

Quantitative assessment of tissue properties has become possible with specific MRI sequences that enable extraction of parameters such as T_1 and edema water fraction. With a quantitative characterization of the tissue, it is possible not only to compare data across scans and subjects, but also to capture changes at sub-voxel scale [1,2]. This can potentially improve our understanding of pathophysiology and generate biomarkers for e.g. abnormality detection, disease staging or treatment response assessment. However, probing tissue parameters requires acquisition of high-dimensional datasets and appropriate parameterization. A well-known example is the calculation of the apparent diffusion coefficient (ADC) from a set of pulsed gradient spin-echo scans with varying diffusion weighting (b-value). The ADC describes the mean diffusivity of water in a voxel and is obtained by fitting the diffusion-weighted signal to a mono-exponential model, as expressed in the Stejskal-Tanner equation [3]. Important factors modulating the diffusivity are the volume of extracellular water with a less restricted mobility [4], the size of restricting domains such as cells, and further, the orientation of elongated restrictions which makes the diffusivity directionally dependent. However, a voxel can contain multiple tissue compartments (partial volume effect) or orientational dispersion of anisotropic domains, and the diffusion can be affected by water exchange between compartments [5,6]. Furthermore, biophysical processes such as flow and perfusion can cause extra signal decay [7,8]. Hence, tissues exhibit non-exponential signal decay if a broad range of b-values is used. Several more complex models have been suggested and used to improve fitting, e.g. the

☆ **Funding information** Danish Cancer Society, Grant Number: R167-A10637-17-S2; European Union's Horizon 2020 research and innovation programme, Grant agreement No 804746.
* Corresponding author at: Technical University of Denmark, Ørsteds Plads, Building 349, Kgs. Lyngby 2800, Denmark.
E-mail address: lghan@dtu.dk (L.G. Hanson).

https://doi.org/10.1016/j.jmr.2021.107103
1090-7807/

intravoxel incoherent motion model or water exchange models [6]. Imposing models, however, always carries the risk of estimating biased or uninformative parameters, because the model may be misleading, or too restricted or flexible considering the underlying signal variation [8,9]. Nevertheless, decomposition of the signal into multiple compartment-related components is desired for tissue characterization, which is an essential challenge that is repeatedly encountered in quantitative analysis of MRI data. Mapping parameters for relaxometry or magnetic resonance fingerprinting, for example, faces the ill-posed problem of decomposing a signal into a weighted sum of exponentials or other latent factors [1,7]. A class of methods designed to solve this ill-posed inversion, is regularized voxel-wise spectral fitting [10–12], which in the case of multi-exponential decay is also referred to as regularized inverse Laplace transforms [13]. A frequently used version is the non-negative least squares (NNLS) technique [14,15] which has proven useful, e.g. for estimation of myelin water fractions (MWF) [14,10,16]. The regularization is necessary to handle the inevitable large impact of noise in a voxel-wise fitting regime [17], but requires the choice of a penalty parameter. Though a cross-validated or data-driven estimation of this is possible [16], it may introduce a bias in quantitative estimates due to the well known bias/variance trade-off. Additionally, the method relies on an identification of spectral clusters or averaging across regions of interest (ROIs) requiring prior knowledge of the tissue homogeneity [17,18].

Another group of strategies avoiding these limitations and presumptive physiological models is blind source separation (BSS) techniques, including widely used unsupervised linear dimensionality reduction (LDR). Here, all measurements are analyzed simultaneously and represented as a product of two factors; a matrix of fundamental basis functions and a matrix of the associated spatial distributions [9,19,20]. An attractive feature of LDR techniques is their insensitivity to partial volume effects often hampering traditional signal modeling. What sets the different LDR techniques apart is the assumptions on the structure of the factors. The well-known principal components analysis yields orthogonality between the principal components [21–23], while the independent component analysis, often used for separation of functional MRI signals, ensures independence between them [19,24]. While these methods may result in low-error reconstructions of the data since the underlying singular value decomposition is guaranteed to produce the minimum error for a given dimensionality reduction [25], the components may be physically unrealistic and difficult to interpret. For medical image intensities, a natural assumption is component-wise non-negativity leading to the LDR technique, non-negative matrix factorization (NMF). This method is popular within a broad range of applications ranging from text mining to image processing because the non-negativity assumption promotes an attractive part-based representation with more realistic factors, although the general non-uniqueness of NMF solutions can render the interpretation of the estimated components difficult [19,26–28]. In general, NMF is non-convex and the estimated solution is often sensitive to initialization and small perturbations of the input data. In practice, a common strategy to tackle these issues is to use priors on the factors or regularization during optimization [19,27,29].

This work proposes a novel extension of the NMF by introducing additional constraints on the structure of the factors, which tailors the method for extraction of multi-exponential or similar signals, expected for many phenomena both within and beyond the field of MRI. This loosely defined class of functions is called *approximately multi-exponential* in the following. The extension is inspired by Bhatt and Ayyar who introduced the monotonous NMF [30], mNMF. They enforced monotony of NMF components, which in practise give results similar to NMF for a range of MR signals and is insufficient to enforce realistic signal behaviour (see Supplementary Material). Hence, we introduce a method, which constrains the problem further. It imposes two conditions on the non-negative basis functions: they should be monotonous and have a monotonous first derivative (slope). Thus, the proposed method is referred to as *monotonous slope non-negative matrix factorization* (msNMF). It is expected that the new constraint on the first derivative makes msNMF superior to the mNMF in regards to separation of mixed MRI signal components. Yet, the method leaves more freedom to the signal components compared to the spectral fitting techniques where the decomposition is within a pre-determined multi-exponential basis constraining also higher order derivatives. In more general terms, the additional constraints enforce prior expectations of MR signal decay and represents a balance between model-based analysis that may rely on inadequate model assumptions, and model-free analysis that may be too flexible to give interpretable results.

We introduce and demonstrate the msNMF for multi-component analysis of MRI data, and show examples where the estimated components are related to the underlying tissue structure, thus demonstrating the use of the method as a robust alternative to advanced modeling.

2. Methods

2.1. Non-negative matrix factorization

The NMF approximates a given data matrix, $\mathbf{X} \in \mathbb{R}^{m\times n}$, with the low rank matrix, $\mathbf{WH}$:

$$\mathbf{X} = \mathbf{WH} + \mathbf{E} \tag{1}$$

$\mathbf{W} \in \mathbb{R}_+^{m\times k}$ is a set of non-negative basis functions, $\mathbf{H} \in \mathbb{R}_+^{k\times n}$is the associated non-negative weights for the linear combination of basic functions to reconstruct the data, and $\mathbf{E} \in \mathbb{R}^{m\times n}$is an error matrix accounting for noise and non-factorizable signals. With the assumption of Gaussian i.i.d. residuals in $\mathbf{E}$, the Frobenius norm is used as objective function, $\|\mathbf{X} - \mathbf{WH}\|_F^2$. This is a non-convex optimization problem. However, using the two-block coordinate descent framework known as the alternating non-negative least squares (ANLS) algorithm, it can be formulated as two convex sub-problems:

$$\min_{\mathbf{W}\geqslant 0} \|\mathbf{X} - \mathbf{WH}\|_F^2 \quad \text{and} \quad \min_{\mathbf{H}\geqslant 0} \|\mathbf{X} - \mathbf{WH}\|_F^2 \tag{2}$$

The problem is solved by alternately re-estimating one of the two factors, $\mathbf{W}$ or $\mathbf{H}$, under a non-negativity constraint, while keeping the other fixed.

In this study demonstrating analysis of volumetric 4D MRI datasets with m measurements and n signal-carrying voxels, the k basis functions in $\mathbf{W}$ are referred to as *signal components*. The weights in $\mathbf{H}$, specifying how these are mixed in each voxel, are referred to as *mixture maps*.

2.2. Monotonous slope non-negative matrix factorization

The idea of introducing constraints that enforce monotonicity in the NMF emerges from expectations for noise-free MR signal behaviour. It can be shown, that the group of mono- and multi-exponential functions (described by: $b_0 + \sum_{n=1}^{N} b_n \cdot e^{a_n x}$) are monotonous, so the continuous derivative has no zeros, provided the exponents a_n have a common sign, and the arbitrary number of non-zero coefficients b_n also share sign. The offset b_0 has no impact. The derivatives fall within the same class of functions, and are therefore also monotonous. In unusual specific cases involving pronounced intermediate water exchange, mostly in

combination with inversion recovery, the signal no longer complies with the assumption of shared sign coefficients b_n, but normal multi-exponential signal decay and longitudinal magnetization recovery fulfills this requirement adopted in the following [31–33].

The two constrained least squares problems in (2) can easily be restated in dual form to allow optimization via standard quadratic programming. Eq. 3 shows the general case of such reformulation:

$$\min_{\mathbf{x}} \tfrac{1}{2}\|\mathbf{Rx}-\mathbf{d}\|^2 \longrightarrow \min_{\mathbf{x}} \tfrac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Px}+\mathbf{f}^{\mathrm{T}}\mathbf{x} \quad \text{subject to}: \ \mathbf{Gx} \leqslant \mathbf{a} \qquad \text{where } \mathbf{P}=\mathbf{R}^{\mathrm{T}}\mathbf{R} \text{ and } \mathbf{f}=-\mathbf{R}^{\mathrm{T}}\mathbf{d} \tag{3}$$

In the optimization problem at hand, $\mathbf{x}$ is a vectorization of either $\mathbf{W}$ or $\mathbf{H}$. The inequality constraint can be revised to include both the non-negativity and monotonicity constraints. This approach was inspired by the work of Bhatt and Ayyar [30] and results in the following reformulated optimization with the appropriate constraints added:

$$\text{Given } \mathbf{W}: \quad \min_{\mathbf{H}} \ \mathrm{vec}(\mathbf{H})^{\mathrm{T}}(\mathbf{I}_n \otimes \mathbf{W}^{\mathrm{T}}\mathbf{W})\mathrm{vec}(\mathbf{H}) - 2\mathrm{vec}(\mathbf{X}^{\mathrm{T}}\mathbf{W})^{\mathrm{T}}\mathrm{vec}(\mathbf{H})$$

$$\text{subject to}: \quad -\mathbf{I}_{nk}\mathrm{vec}(\mathbf{H}) \leqslant \mathbf{0}_{nk}$$

$$\text{Given } \mathbf{H}: \min_{\mathbf{W}} \ \mathrm{vec}(\mathbf{W})^{\mathrm{T}}(\mathbf{H}\mathbf{H}^{\mathrm{T}} \otimes \mathbf{I}_m)\mathrm{vec}(\mathbf{W}) - 2\mathrm{vec}(\mathbf{X}\mathbf{H}^{\mathrm{T}})^{\mathrm{T}}\mathrm{vec}(\mathbf{W})$$

$$\text{subject to}: \quad \begin{bmatrix} -\mathbf{I}_{mk} \\ \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathrm{vec}(\mathbf{W}) \leqslant \mathbf{0}_{mk}$$

vec() is a column vectorization of the matrix coefficients and $\otimes$ is the Kronecker tensor product. For optimization of $\mathbf{H}$ (Eq. 4), only the non-negativity constraint is included in the inequality. For optimization of $\mathbf{W}$ (Eq. 5), the new constraints are added in $\mathbf{A} \in \mathbb{R}^{(m-1)k\times mk}$ ensuring a monotonous behavior of the basis functions, and in $\mathbf{B} \in \mathbb{R}^{(m-2)k\times mk}$ ensuring monotonous behavior of their first derivatives. $\mathbf{A}$ and $\mathbf{B}$ can be designed such that the basis functions are either increasing or decreasing (Supplementary Material). The implementation enables mixing of basis functions to involve both decaying and growing signal components if desired. This may become relevant, e.g. for saturation recovery or if a Rician noise bias arises for small signal values.

The programming language MATLAB 2018b (MathWorks, Inc., Natick, Massachusetts, United States) was chosen for implementation, and the trust-region-reflective algorithm was used for the quadratic programming.

To facilitate straightforward adoption into existing processing pipelines the software is made publicly available, https://github.com/sofierahbek/msNMF.

2.3. Practical implementation

Initialization and stopping criterion A basic random initialization of $\mathbf{H}$ with coefficients drawn independently from a uniform distribution between 0 and 1 was used. For small datasets, all voxels were given as input to the iterative ANLS which continued until a stopping criterion was met, being either a maximum of 500 iterations or a relative change in factors below 10^{-8}. To reduce the unavoidable risk of ending in a local minimum due to the non-convex nature of the optimization problem, the process was carried out 10 times (each time with a different initialization), and the lowest cost solution was chosen. This is also known as a "multi-start" strategy [27,19]. For large datasets ($> 10^5$ observations), a stochastic gradient descent framework was utilized due to memory limitations which also served to reduce computation time. Data was divided into randomly selected batches, and the input batch was replaced during optimization for every T'th iteration. This continued until all data had been included (one full epoch) or until fulfilling a convergence criterion based on relative reduction in cost function (for the full data) averaged across the latest five results. Here, the threshold for the relative reduction of the cost function was set to 0.1%, the input batch size was chosen to $\approx 1\%$ of all observations and T was set to 30. We found that the optimization is robust to these choices and that they will primarily affect computation time rather than the obtained result [34].

Rank determination The resulting factorization of data is dependent on the number of components, i.e. the rank, k, which was chosen using prior knowledge or a trial-and-error approach, where the factorization and its residual are used as feedback for rank adjustments [19,27]. If anatomical structures were visible in the residual image, the rank was increased. Conversely, if two components were very similar, the rank was decreased.

Preprocessing steps A few preprocessing steps were implemented to reduce the impact of noise on the factorization. First, the data used as input to the msNMF was required to carry the highest signal in the first measurement (m = 1) to exclude irrelevant signal behaviour from vascular components, for example. Secondly, data was normalized to have initial value 1 before factorization. This removes the influence of the initial signal amplitudes during estimation of the signal components in $\mathbf{W}$. A subsequent projection of the original (unscaled) data onto these ($\mathbf{H} = (\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{X}$) leaves $\mathbf{H}$ with the amount of total signal distributed over the identified components. The introduced data normalization simultaneously amplifies the noise, which was compensated by weighting the ANLS with the initial signal amplitudes. An updated objective function is given in Supplementary Material.

2.4. Validation

Simulated data was used to validate the implementation (visualized in Supplementary Material, Figure S1). To imitate a set of realistic MR measurements, eight non-equidistant samples of three decaying signals, two mono-exponential and one bi-exponential function ($y_1 = e^{-0.01x}, y_2 = 0.6e^{-0.08x} + 0.4e^{-0.03x}, y_3 = e^{-0.15x}$), were generated and mixed across 64x64 voxels. The bi-exponential component was included to demonstrate the method's ability to detect realistic signals beyond the mono-exponential domain. Normally distributed complex noise ($\sigma = 0.06$) was added to the data. This left the final magnitude data with a relatively low SNR and a Rician noise distribution, which is more realistic than the assumption of Gaussian independent and identically distributed (i.i.d) residuals implied by the formulation of the cost function. Knowing the underlying data structure it was possible to quantitatively evaluate the result of the method. The *explained variance* (EV), with 1 indicating a perfect reconstruction, was used to compare the reconstructed curves in $\mathbf{W}$ with the true signals ($[y_1, y_2, y_3]$). The *area under the* ROC *curve* (AUC) was used to compare the spatial distributions in $\mathbf{H}$ with the true mixture maps. This measure explains to what extent the spatial pattern of the components is recognized, and is not affected by the absolute values of the loadings in $\mathbf{H}$. The performance of the msNMF was compared to that of the standard NMF. For a fair comparison, the exact same conditions for both factorizations were used except for the monotonicity constraints, i.e. data normalization and weights were also applied for the NMF which makes it more constrained than standard NMF. Additionally, a stability test was carried out to investigate how much the added constraints stabilize the solution across 50 random initializations. The *coefficient of variation* (CV) was used as

metric with a low percentage indicating high stability. The specified value is the average CV across all elements in **H** (leaving out background voxels).

2.5. Data

The application of the msNMF was demonstrated using two previously published MRI datasets: ex vivo DWI of a monkey brain acquired with 12 b-values at two different temporal scales [35], and in vivo multiple spin-echo imaging with 40 echo times for multi-exponential T_2 (MET_2) analysis of rat spinal cord [11]. The relatively large number of measurements in both datasets yielded well-characterized signal curves, which in both cases were expected to exhibit an approximately multi-exponential signal decay.

Ex vivo DWI data Data recorded and published by Lundell et al. [35] with multidimensional diffusion encoding revealed microstructural details in a post mortem brain of a Vervet monkey. The sample was perfusion fixed and placed in a buffer solution using procedures optimised for post mortem DWI [36], and data was recorded with a 4.7 T pre-clinical MRI system. We refer to the original publication for more details regarding the experimental setup and the data, which is available on reasonable request [35]. In short, the experiment probed the effect of anisotropy as well as time-dependent diffusion, depending on the size of restrictions. The latter was explored by comparing two diffusion acquisitions recorded with the same echo time TE = 68 ms and so-called "tuned" and "detuned" directional encoding gradient waveforms with an average diffusion time of $\sim$ 10 ms ("short diffusion time") and of $\sim$ 20 ms ("long diffusion time") respectively. Both acquisitions included the same 12 b-values in the range 242–4832 s/mm^2 in five slices of central brain. The difference in diffusion time is expected to influence the signal response to restricted diffusion on length scales $\sim$ 3–10 μm, whereas free diffusion will provide similar signal signatures [5]. All brain voxels from the two data sets were used as input for the msNMF and the stochastic gradient descent framework as described in Section 2.3 was used. This resulted in a computation time of 266 s using a workstation equipped with a 28-core system having 16 GB RAM. Signals were decomposed using an initial rank $k = 3$ due to expectations of at least a bi-component signal decay in the brain tissue and an additional component from the surrounding buffer solution. The resulting mixture maps (**H**) for the two data sets were normalized, leaving the coefficients as signal *fractions*, and compared through voxel-wise subtraction. The decomposition was evaluated using knowledge of cell densities in primate brain and findings by Lundell et al. [35].

In vivo MET_2 data Relaxometry data originally presented in [11] by Harkins et al. demonstrated a model for intra-myelinic edema (IME) quantification in rat spinal white matter evaluated by comparisons with histology. In this study, 24 rats were divided into three groups of eight, each receiving a different amount of the IME-inducing toxin hexachlorophene (HCP) in their diet: 0 ppm, 300 ppm and 600 ppm respectively. Forty T_2-weighted spin-echo images were recorded at 9.4 T using a fluid-attenuated inversion-recovery-prepared multiple spin-echo imaging sequence [37] (TI = 2 s, TR = 6 s). The first 32 echoes were collected with an echo spacing of 9 ms starting from TE = 7.4 ms. The last 8 were collected with an echo spacing of 50 ms. Recordings were obtained for a 1.5 mm thick axial slice (0.2x0.2 mm^2 in-plane resolution) of the cervical spinal cord, and histology data were obtained using light microscopy of 1 μm thick slices dissected from the same location for three rats from each group. Both MR images and histology data were analyzed for the water content of edema, myelin and intra/extra axonal space in four different white matter tracts: dorsal cortical spinal tract, funicilus gracilis, rubrospinal tract, and vestibulospinal tract. These regions-of-interest (ROIs) defined the voxels to include in the msNMF, and the input signals were decomposed into three decaying components, one for each of the mentioned compartments, i.e. the rank best suited for comparison was known in advance. **H** was normalized to provide the fraction of each compartment and to facilitate comparison to both histology values and imaging metrics (water fractions) estimated by Harkins et al. The water fractions were defined as in Mackay et al. [14], assuming proportionality between total signal amplitude and total water content. Signal from myelin lipids was considered to be negligible given its rapid T_2 decay (<1 ms) [38–40].

Reproducibility test As a novel method based on a non-convex problem and random initializations, it is sensible to test the reproducibility of the full framework, which was done using the in vivo data [11] and a split-half (NPAIRS) approach [41]: The full data was divided in two sets, each consisting of every second echo, i.e. half of the variables but all observations. The resulting **H** for each dataset was compared using Pearson correlation coefficient to calculate a *reproducibility* measure. The two factorizations were evaluated by projection of **H** onto the other unseen dataset and calculating the EV for each data reconstruction. The average defined a *prediction* measure. A *prediction*/*reproducibility* curve (pr-curve) showing the trade-off between prediction accuracy and pattern reproducibility for a range of ranks k was produced.

3. Results

3.1. Validation data

Fig. 1 shows the final factorization of the simulated data. The true exponential signals and true mixture maps are included for visual comparison, and the quantitative comparison revealed $EV_{\mathbf{W}} = 0.984$ and $AUC_{\mathbf{H}} = 0.999$. The three components are well separated but the spatial distributions seen in **H** are not flawlessly presenting homogeneous circles, especially the purple and yellow component contain traces of other signal components and a higher level of noise (confirmed by the indicated summary statistics). The lower mean value for the purple component illustrates that part of the signal has been incorrectly distributed to the other components, consistent with the "shadow" seen in the yellow mixture map. Please note that even for a hypothetical method perfectly identifying the true signals (dotted lines in Fig. 1a), the residual noise amplitude will be spatially inhomogeneous due to the structure and noise level of the data.

In comparison, the standard NMF extracted erroneous signal components and slightly worse spatial distributions (Supplementary Material, Figure S3). The evaluation metrics were correspondingly lower than those obtained with the msNMF ($EV_{\mathbf{W}} = 0.922$ and $AUC_{\mathbf{H}} = 0.989$). The mean CV for 50 initializations was 5.07 % for the NMF and 1.13 % for the msNMF, i.e. a higher stability of the solution was observed for the msNMF. A table in Supplementary Material shows the msNMF result for simulated data with coil arrays (8, 16 and 32 elements) and other SNR (σ) levels, included to demonstrate the dependence on noise characteristics.

3.2. Ex vivo DWI data

Example raw data are shown in Fig. 2a in the shape of a somewhat T_2-weighted image (TE = 68 ms, b = 242 s/mm^2). Trial-and-error testing of the rank led to a final choice of k = 3. A decrease to k = 2 resulted in a residual map with clear cerebral structures. Conversely, an increment to k = 4 resulted in two of the four components to be very similar for both factors, and **H** was distinctly contaminated with noise compared to the decomposition presented here using k = 3. The resulting three diffusion-related

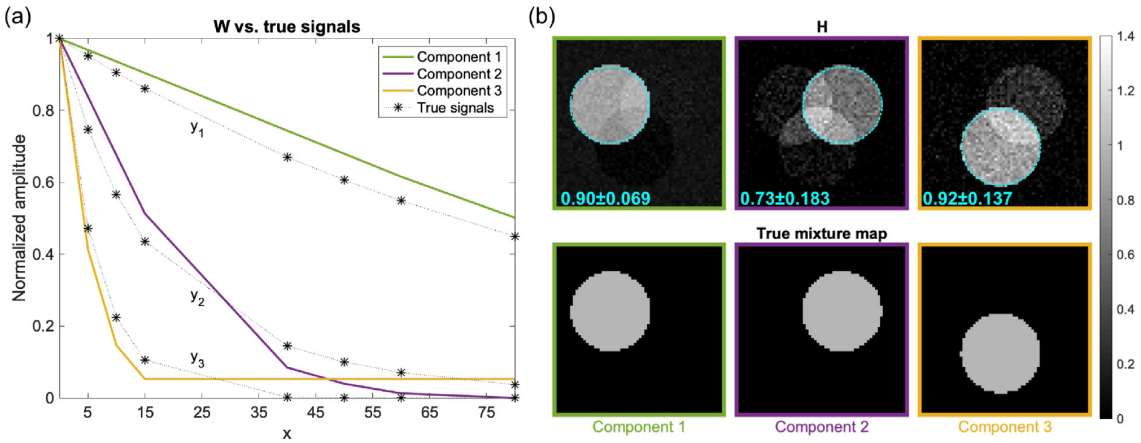

Fig. 1. True data compared to the msNMF result. (a) Estimated signal components, **W**, and true signals ($y_1 = e^{-0.01x}, y_2 = 0.6e^{-0.08x} + 0.4e^{-0.03x}, y_3 = e^{-0.15x}$). (b): Associated mixture maps, **H**, indicated by frame colors. The cyan numbers are mean±SD across the encircled voxel values. Similar illustrations obtained with standard NMF and the mNMF are provided in Supplementary Material.

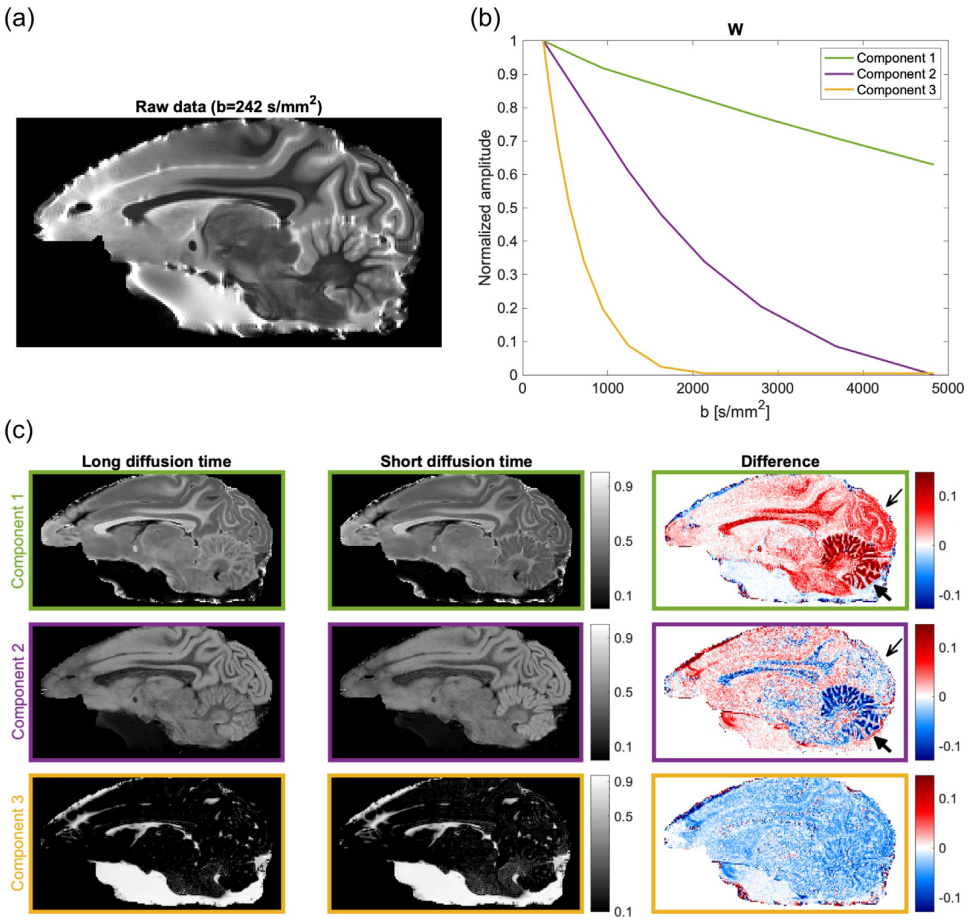

Fig. 2. (a): Measured image of the monkey brain for the lowest acquired b-value (242 s/mm^2) (b): The signal components, **W**, from the msNMF. (c): The associated normalized mixture maps, **H**, labelled by the frame colors. A logarithmic colorscale is used for the yellow component. Maps are given for both data sets and the right column shows the difference between the two ("Short diffusion time" subtracted from "Long diffusion time"). The thin and thick black arrows mark the conspicuous visual cortex and cerebellum, respectively.

components seen in Fig. 2b reveal long, short and intermediate signal lifetime (green, yellow and purple curve, respectively).

The associated normalized mixture maps in Fig. 2c reveal a relatively high fraction of the purple signal component in the cortex, creating a clear separation of grey matter from white matter, which mostly contains the green signal component. The signal from aqueous compartments (buffer solution surrounding the brain sample) is almost exclusively described by the yellow component. The difference between the two sets of maps, provided to capture the effect of time-dependent diffusion, is presented in the third column of images. Positive values (red colors) indicate that an increase in diffusion time results in a higher fraction of the respective component. This effect is observed for the long-lived signal component (green), especially in white matter including corpus callosum, visual cortex (thin arrow), and the cerebellar cortex (thick arrow), which exhibit high contrast. Simultaneously, the opposite effect (negative values) is seen for the intermediate decaying signal component (purple) in these areas. The cerebellum is particularly prominent (thick arrow), and superior to this, the visual cortex (thin arrow) clearly differs from the rest of the cerebral cortex, which otherwise exhibits positive changes for this signal component. The small uniform change across the brain seen for the short-lived signal (yellow) is in agreement with the expectation that a fast decaying liquid component is time independent.

3.3. In vivo MET_2 data

Fig. 3a and 3b show the result of the msNMF: the signal components and mixture maps, respectively. The mixture maps illustrate the distribution of the signals across the spinal cord for a rat from each group. Since only voxels from the white matter tracts have been included in the factorization, the gray matter signals may not be well represented. For all three components, the contrast between the butterfly-shaped gray matter and the surrounding white matter increases with increasing amount of HCP in the diet. Especially the abundance of the long-lived component (green) is increased in specific areas of the white matter for rats on HCP diet. The three signal components are thought to reflect relaxation of water within different compartments in the white matter. Fitting a mono-exponential decay to the curves resulted in relaxation times of 12 ms, 42 ms and 181 ms for the yellow, purple and green curves respectively. With expectations of IME to cause T_2 hyperintensity [11,42], the long-living signal component is consistent with IME. Thus, the fraction of the green component (given by the normalized **H**) is considered an estimate of the EWF.

Fig. 4b presents the mean EWF for the four ROIs visualized in Fig. 4a averaged across the eight rats in each group. The error bars indicate the inter-animal standard deviation. For comparison, Fig. 4c shows the equivalent result presented in [11]. The two bar plots are very similar considering the relative increase of EWF in the different ROIs for the HCP groups. Quantitatively, the paper reports slightly higher values than the method proposed here.

Fig. 5 compares the EWF estimates with the histology-based edema water content fraction (W_{ed}) for three rats from each group. The correlation metric for the linear regression together with the associated p-value reveal a high agreement with histology data (r^2 = 0.88, $P<0.001$). The trendline is closer to unity, the intra-ROI standard errors are lower, and the correlation metric is slightly higher in comparison to the corresponding analysis in the original paper [11]. For both approaches, non-zero intercepts remain in the comparisons with histology estimates.

Fig. 6 shows the result of the split-half test, the pr-curve. The plot reveals a high Pearson correlation (>0.8) for a rank below 5, and a high predictability for a rank above 1. This indicates an acceptable reproducibility of the framework for the chosen rank k = 3 with predictability and correlation both being high.

4. Discussion

4.1. Simulated data

The simulated data was included to validate the method before applying the msNMF to measured data where the true underlying structure is unknown. Three fundamental signals, only eight temporal samples, and a Rician noise distribution were selected to generate a clinically realistic MR dataset, and a non-trivial decomposition problem.

It was confirmed both by the visual comparisons (Fig. 1) and evaluation metrics that the msNMF provided a decomposition in high accordance with the true data, thus validating the implementation. The example, however, demonstrates both the ability and limitation of the method. The imperfect mixture maps observed for the purple and yellow component, in particular, are consistent with a lower integrated signal intensity for these components and a larger similarity between them. Both properties can make them more difficult to distinguish in the factorization. Comparisons with results from the standard NMF and mNMF [30] confirmed that adding constraints tailored to this type of data improves the result and increases the stability of the solution.

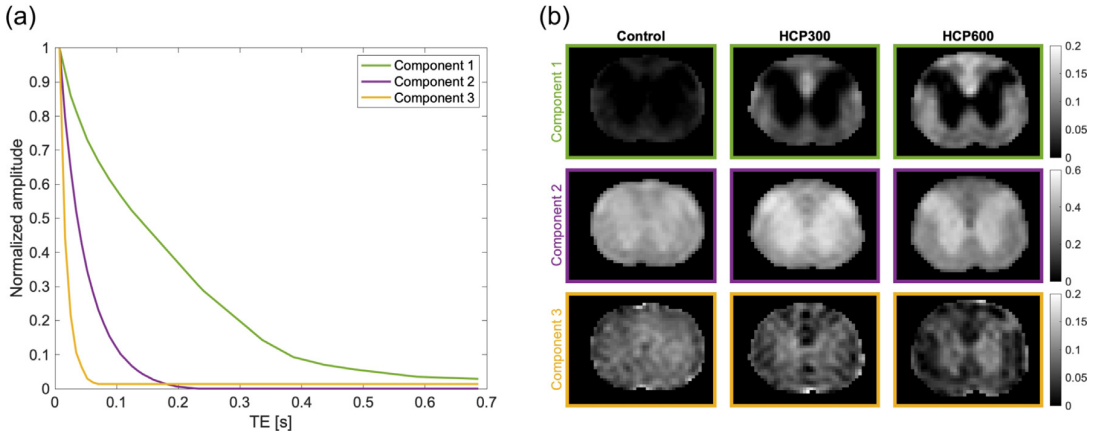

Fig. 3. Result of the msNMF for MET_2 analysis of rat spinal cord. (a): The signal components, **W**. (b): The associated mixture maps, **H**, indicated by the frame and label colors. Each column of images shows the maps for a specific rat (one from each group).

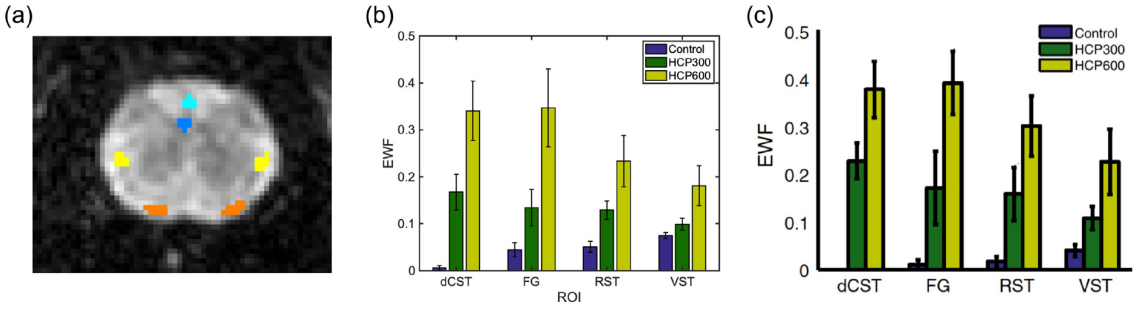

Fig. 4. EWF estimates for each of the spinal cord tracts (ROIs). (a): The ROI locations for an example rat; dorsal cortical spinal tract (blue), funicilus gracilis (cyan), rubrospinal tract (yellow), vestibulospinal tract (orange). (b): Estimates provided by the msNMF: the mean fraction of component 1 (green in Fig. 3) within a ROI averaged across rats. (c): Estimates reported in [11]. Error-bars indicate inter-animal sta.ndard deviation.

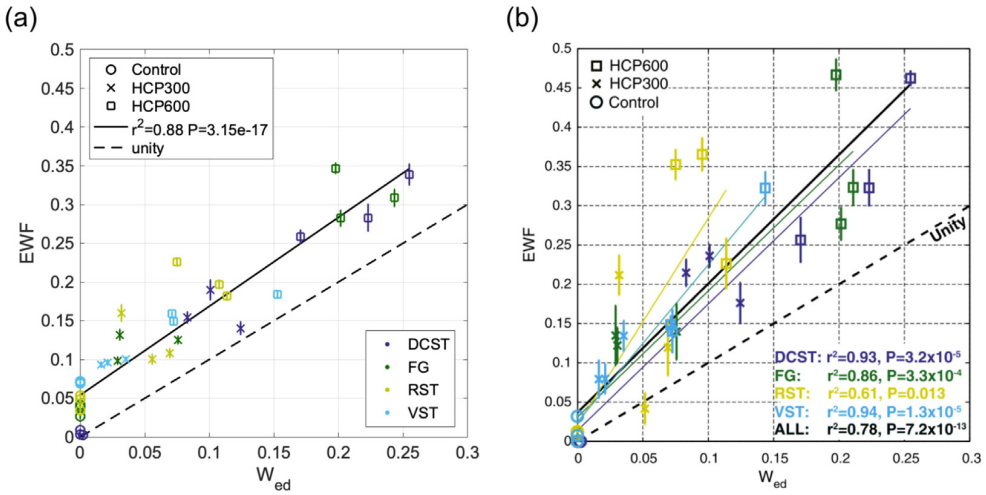

Fig. 5. (a): Scatter plot comparing EWF estimates from the msNMF with W_{ed} from histology data. Colors indicate the ROI, symbols indicate the rat group, and error bars indicate intra-ROI standard error. The black solid line shows a linear regression for all data merged together ($r^2 = 0.88$). (b): The equivalent scatter plot presented in the source paper [11].

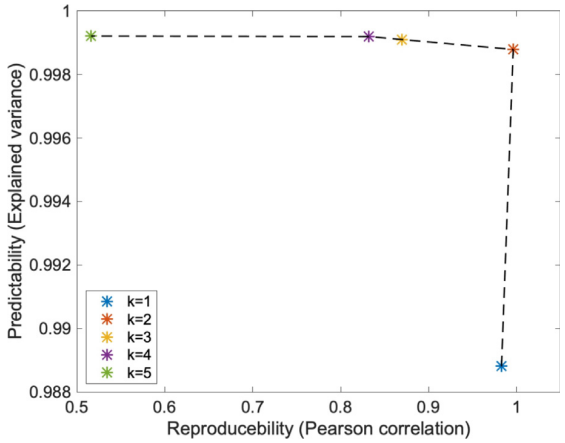

Fig. 6. Prediction/reproducibility curve for the msNMF of the MET_2 data set. The curve is constructed by varying the rank of the decomposition and calculating the Pearson correlation and explained variance using the split-half test.

The superiority to mNMF is particularly important, as this method, among several BSS techniques extended with priors or regularization in order to solve related problems [13,15,43–46], is most similar to our proposed method. We emphasize that the msNMF does not rely on specification of regularization parameters, nor requires information from several contrast dimensions, and we therefore consider it out of scope to compare msNMF to most alternative BSS methods imposing particular signal behaviour. The simple prior assumption of monotonicity of the individual signals and their first-order derivatives have to our knowledge not been utilized in a decomposition method before. These constraints intuitively fit many types of MR data.

A common problem of the Rician or noncentral chi distributed noise present in MRI magnitude data, is a positive bias (noise floor) for low SNR, and the effect is exacerbated using large coil arrays and conventional sum-of-squares coil combination [47,48]. In the validation example, this effect became evident for the yellow signal component (Fig. 1) for the added noise amplitude ($\sigma = 0.06$). A pragmatic strategy to reduce such signal-dependent noise bias is to estimate a single increasing component with monotonous decreasing slope in addition to the decreasing components. This may absorb a considerable fraction of the noise floor. To demonstrate such flexibility of the msNMF, the strategy was tested for coil array data in addition to the single element coil. The addition of a spatially uniform growing component was found to improve the estimates (Supplementary Material, Figure S4 and Table S1). The uniformity condition on the mixture map was used to avoid excessive "cancellation" of the decaying signal components, which may

occur if the noise floor is not reached for all signals and thus is difficult to identify. However, the presence of a positive bias depends on the level of signal in the given voxel, so the assumption of a constant bias across all voxels is not accurate in general. Even so, the proposed strategy can be advantageous to directly account for noise or other sources of bias as demonstrated with the simulated data. Also, it is possible to enforce more involved constraints for the mixture map such as a low frequency basis for each tissue class. For the animal data, the noise bias was not an issue so the described correction strategy was not applied.

The msNMF has proven effective for decomposition of the simulated data generated as a weighted mixture of three main signal components. However, signals of biological origin is often expected to have a continuous distribution of decay rates, e.g. a range of diffusivities from variation in cell sizes, centered around a mean decay rate. In the Supplementary Material (Figure S2), an extended example of the simulated data is presented to show how the method handles such data. For two of the main signal components, the decay rate was drawn from a normal distribution. As the result reveals, the mean signals were recognized, i.e. the decomposition outcome did not change significantly. This shows that msNMF provides an informative and compact representation of the data without necessarily describing all aspects of the underlying data.

As already mentioned, another typical quantitative analysis strategy is the NNLS approach. In spite of the necessity for regularization and averaging across ROIs, the NNLS approach is a relevant reference for the proposed strategy, and results of the method for two sets of simulated data are provided in the Supplementary Material (Figure S5). The presented results rely on post-processing involving k-means partitioning of the raw spectra into three clusters. Final results are comparable to those for the msNMF considering single-coil data, but are more affected by the Rician noise for the coil array data reconstructed using conventional sum-of-squares. Notably, our method decomposes mixed signals in the domain of the acquisition, e.g. the echo time domain, and is not relying on critical or equidistant sampling of the distribution of T_2 values, for example.

An alternative group of strategies worth mentioning is the CORE (component-resolved NMR) processing family, typically used for analysis of diffusion-ordered NMR spectroscopy (DOSY) data sets. Examples are CORE [49], GRECORD [50], OUTSCORE [51], SILT-DOSY [52], and one of the newest presented in the literature, the InSpect algorithm [18]. As for the BSS methods, these techniques have replaced the univariate processing (voxel-wise fitting) with a multivariate decomposition. However, they still require choosing an explicit MR signal model for the reconstruction [18]. As described in the Introduction, complex biological conditions make it difficult to formulate an accurate signal model a priori. Our introduction of msNMF is motivated by this and a direct comparison to methods from the CORE family is therefore not possible.

4.2. Ex vivo DWI data

The DWI dataset was used to investigate whether the extracted components were in fact related to the local diffusion and thus a "fingerprint" of the underlying tissue micro-structure. Being an ex vivo dataset with a broad range of b-values, measurements for two different temporal scales, and an accompanying thorough analysis published in [35], it was well-suited for testing the proposed method.

The comparison of the mixture maps differing by diffusion time (Fig. 2c, right column), uncovered areas of high cell density, especially visual cortex and cerebellum [53,54]. In more detail, the enhancement of the long-lived diffusion-weighted signal (green component) for increased diffusion time is associated with the degree of restrictions in the local tissue environment [35,55]. Therefore, the identified areas can be further categorized into very restricted environments (cerebellum) and intermediate restricted cell environments (visual cortex). This strong contrast across different areas of the cerebral and cerebellar cortex, revealed by the maps, have to our knowledge not been presented based on other MR contrast mechanisms before.

Thus, using two diffusion times and the msNMF for decomposition, the effect of time-dependent diffusion was convincingly demonstrated, exactly as Lundell et al. aimed at when setting up their experiment. Their analysis consisted of a voxel-wise calculation of the powder averaged signal difference at b = 4800 s/mm^2 normalized to the b = 0 signal. The analysis revealed similar time-dependent diffusion contrast, but was dependent on estimating the b = 0 signal from a 3rd order expansion of the signal in b and thus delivered a more noisy map than the one presented in Fig. 2c.

Through the use of non-conventional diffusion encodings with general gradient waveforms, different effects related to the signal attenuation in DWI (e.g. anisotropy, restrictions, exchange) can independently be probed and detected as direct features of the data or contrasts between measurements exploring single encoding dimensions [35,56]. Time-dependent diffusion is a promising marker of cell morphology and density in tumors [57]. This makes the combination of msNMF applied to novel DWI data a potentially powerful tool for characterization of tumor tissue that is generally difficult to model, and thus relevant for e.g. therapy response evaluation. However, to adapt the acquisition for in vivo measurements, fewer b-values (i.e. shorter scan time) and a more simple gradient waveform are desired. For this data, it may thus be interesting to investigate whether the separation into the three diffusion-related components can be obtained with fewer measurements, for example. Such analysis is beyond the scope of this paper but possible now that the components are characterized.

Consideration should also be given to what tissue information can be derived if only one data set was available, e.g if only the left column of mixture maps is used (long diffusion time). The three macrostructural brain tissues are easily distinguished. Gray matter is separated from white matter, and firm tissue from aqueous solutions. If the three decaying signal components each are fitted to a mono-exponential function, the resulting decay constants would be 0.1×10^{-3}, 0.6×10^{-3} and 2.3×10^{-3} mm^2/s respectively for the green, purple and yellow signals. These values are within a range of ADC values observed in the literature for white matter, gray matter and CSF respectively, though the ADC is dependent on acquisition method, handling of ex vivo material, temperature etc. [58–60]. The fitted values are thus only included here to relate the found components to a familiar characterization of diffusion. It is evident that especially the purple signal component deviates from an exponential decay as it reaches 0 for b = 4842 s/mm^2. This is an unrealistic behaviour indicating that all noise/left-over signal is assigned to the green component for the last measurement, possibly an effect of the implicit sparsity of the NMF [19]. The effect can potentially be reduced with signal sampling at even higher b-values to allow all individual signals to reach the noise floor, so they are easier to resolve.

A noticeable feature of the factorization is the alternative representation of the signal variation between the two tissue types, gray and white matter. A difference in the mixture of (mainly) two signal components discriminates the tissue types rather than a difference in a mono-exponential decay constant. This exemplifies that the analysis provides a compact representation of the data that may reveal relevant signal features without necessarily reflecting the underlying time-dependent diffusion phenomena [61], especially when data points are limited. The

decomposition-based representation may robustly reveal useful information about the local diffusion compared to a simplified signal model vulnerable to noise and partial volume effects. Here, it proved possible to capture cell density variation by means of the factorization, thereby demonstrating a connection between components and microstructure.

4.3. In vivo MET_2 data

To further demonstrate the versatility of the method, in vivo data was analyzed quantitatively to extract estimates of edema water content similar to those obtained by Harkins et al. in [11]. Simultaneously, the method was tested as an alternative solution for MET_2 analysis, which in general is challenged due to the ill-posed problem of inverting a sum of exponentials [31], but also highly desired in order to obtain stronger quantitative relationships between MR contrast and tissue microstructure [37,11,10].

The three extracted signal components (Fig. 3) are consistent with T_2-relaxation in the different compartments of white matter, i.e. IME (green component), intra/extra cellular space (purple component), and myelin water (yellow component). The visual contrast change observed in the mixture maps for HCP-affected rats affirmatively indicates that the green component can be interpreted as the "IME signal", and literature confirms a decay constant within the range 5–50 ms for water trapped in myelin sheaths [39].

The strategy used in Harkins et al. is similar to the NNLS approach tested for the simulated data, but assumes a Gaussian prior for the distribution of decay rates, and was found superior to the standard NNLS [62]. The T_2-distribution was fitted with up to three Log-Gaussian shaped peaks (components), and this was followed by a k-means algorithm to cluster the T_2-values into groups of short, intermediate and long decay. The resulting mean decay constants of the distributions were roughly 9 ms, 40 ms, and 150 ms, respectively, thus not far from the values obtained by a mono-exponential fit to the three signal components of the msNMF (12 ms, 44 ms, and 182 ms). The fitted decay rates are presented only to include a well-known characterization of the signal profiles and show that results are consistent with the decomposition reported by the source paper [11] (also confirmed by the bar-plots presented in Fig. 4).

Comparison to histology (Fig. 5a) is important for evaluation of the proposed factorization. The high agreement between EWF estimates and histology, revealed by an r^2 of 0.88 ($P< 0.001$), confirms that the decomposition is related to the underlying compartmentalization of the tissue, and demonstrates the method's ability to detect IME severity from T_2-weighted relaxometry measurements. However, similar to the results reported by Harkins et al., a small overestimation of EWF compared to W_{ed} was observed. In general, one must be critical to the absolute values, from both factorization (or any model) and histology. Explanations for the discrepancies were discussed in [11].

Although the split-half experiment revealed a high reproducibility for a decomposition of rank $k = 3$, and the CV metric for the simulation data revealed a relatively high stability of the method, a stable quantitative measure is in general not guaranteed. A variability of the absolute EWF estimates across multiple runs of the algorithm (as part of the initialization strategy) was indeed observed, however at a lower level than the subject-variability. This means that the relative difference of EWF between rat groups and ROIs was stable, and the correlation of r^2=0.88 with histology was robustly observed (SD = 0.0019 for 50 runs).

The rank of the factorization was predetermined to $k = 3$ to match the decomposition presented by Harkins et al. However, the pr-curve (Fig. 6) supports that two components may be sufficient for a representative reconstruction of the data. Reasons could be that the water exchange effectively merges several physical compartments and diminish the MWF [11], or simply that compartments co-vary so a linear combination of two components is sufficient to describe the relaxation within three different compartments. This also explains why estimation of the MWF may not be trustworthy and thus is not considered here, although it is sometimes a target for relaxometry measurements. Dependence on the choice of dimensionality of the sub-system is a common challenge for decomposition methods as there is no established procedure for rank determination, and metrics from heuristic approaches suggested in the literature are highly dependent on the data properties, making it difficult to know when a metric is suitable [27,63]. The pr-curve in Fig. 6 illustrates the problem, as for a range of rank values ($k = [2, 4]$) an acceptable compromise is seen between reproducibility and degree of data explanation. The suggested "trial-and-error" rank-adjustment guided by left-over anatomy in the residual image relies on a subjective inspection, which is a limitation of the method present when the underlying sub-system dimensionality is unknown. That being said, the rank is the only free parameter of the analysis (with or without the inclusion of a bias-compensating component). In the spectrum of techniques ranging from restrictive modeling to completely data-driven signal processing, the msNMF is in the unrestrictive end while relying on the existence of compartments and using this to enforce realistic signal behaviour, though possibly approximate, e.g. due to pronounced intermediate exchange or parameter variability within physical compartments. The technique used by Harkins et al. is also flexible, but still assumes a distribution of one to three Log-Gaussian shaped peaks of exponential time-constants.

Although assumptions of the signal behaviour are less specific than in model-based analysis, the monotonocity requirement must essentially be fulfilled. For clinical multi-echo spin-echo sequences at fields higher than 1.5 T, for example, transients will typically be present causing initial signal oscillations, which makes the msNMF inappropriate as analysis strategy. Also, RF inhomogeneity correction may be necessary to make signal decay rates independent of position in the coil. No such transients or inhomogeneity were observed in the MET_2 data, recorded with a preclinical scanner delivering a train of near-accurate 180° refocusing flip angles, why it was suited for demonstration of the msNMF.

Overall, the pursuit of micro-structural specificity, pioneered with the myelin water determination for multiple sclerosis patients, have throughout the last decade been in rapid development, and several strategies exist for analysis of relaxometry data [14,10,39]. The msNMF is not promoted as a replacement or advocated as a general new tool for MET_2 analysis. The two clinical data examples included in this paper are used to exemplify the characteristics and potential of the proposed analysis method, but not to promote the method for these applications in particular or in general. Many MR signals fall within the class of *approximately multi-exponential* signals and when prerequisites are met, the method offers a data-driven decomposition which may be useful either for direct analysis or to inspire modeling, e.g. let the detected data structure inform priors for a more specific parameterization. When a monotonous slope signal decay or increase is expected, this relatively simple and rapid analysis can give a compact, informative representation of data.

5. Conclusion

The msNMF robustly separated MR signals into identifiable components specific to the underlying structure, and produced quantitative results consistent with those published in the source papers using tailored analysis. The method is insensitive to partial-volume effects and requires only a choice of rank, and

optionally correction of noise biases. The msNMF is potentially applicable to a broad range of multi-dimensional MR data. The demonstrated sensitivity to cell density, for example, makes it a candidate for tumor tissue characterization needed for prognosis and treatment planning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Mark D. Does and Kevin D. Harkins for their cooperativeness and sharing of MET_2 data, and Nirav Bhatt and Arun Ayyar for providing their software for monotonous NMF. The study was funded by the Danish Cancer Society (grant R167-A10637-17-S2) and Henrik Lundell additionally received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 804746).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jmr.2021.107103.

References

[1] M.D. Does, J.L. Olesen, K.D. Harkins, T. Serradas-Duarte, D.F. Gochberg, S.N. Jespersen, N. Shemesh, Evaluation of principal component analysis image denoising on multi-exponential MRI relaxometry, Magnetic Resonance in Medicine 81 (2019) 3503–3514, https://doi.org/10.1002/mrm.27658.
[2] S.M. Meyers, I.M. Vavasour, B. Mädler, T. Harris, E. Fu, D.K. Li, A.L. Traboulsee, A. L. Mackay, C. Laule, Multicenter measurements of myelin water fraction and geometric mean T 2: Intra- and intersite reproducibility, Journal of Magnetic Resonance Imaging 38 (2013) 1445–1453, https://doi.org/10.1002/jmri.24106.
[3] G. Soujanya Chilla, C. Heng Tan, C. Xu, C. Loo Poh, Diffusion weighted magnetic resonance imaging and its recent trend-a survey, Quant Imaging Med Surg 5 (2015) 407–422, https://doi.org/10.3978/j.issn.2223-4292.2015.03.01, URL www.amepc.org/qims..
[4] F. Fornasa, Diffusion-weighted magnetic resonance imaging: What makes water run fast or slow?, Journal of Clinical Imaging Science 1 (2011) 1–7, https://doiorg/10.4103/2156-7514.81294.
[5] J.S. Nielsen, T.B. Dyrby, H. Lundell, Magnetic resonance temporal diffusion tensor spectroscopy of disordered anisotropic tissue, Scientific Reports 8 (2018), https://doi.org/10.1038/s41598-018-19475-y.
[6] A.R. Padhani, G. Liu, D. Mu-Koh, T.L. Chenevert, H.C. Thoeny, T. Takahara, A. Dzik-Jurasz, B.D. Ross, M. Van Cauteren, D. Collins, D.A. Hammoud, G.J. Rustin, B. Taouli, P.L. Choyke, Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations, Neoplasia 11 (2009) 102–125, https://doi.org/10.1593/neo.81328, http://linkinghub.elsevier.com/retrieve/pii/S1476558609800249.
[7] M. Nagtegaal, P. Koken, T. Amthor, M. Doneva, Fast multi-component analysis using a joint sparsity constraint for MR fingerprinting, Magnetic Resonance in Medicine 83 (2020) 521–534, https://doi.org/10.1002/mrm.27947.
[8] D.K. Jones, M. Cercignani, Twenty-five pitfalls in the analysis of diffusion MRI data, NMR in Biomedicine 23 (2010) 803–820, https://doi.org/10.1002/nbm.1543.
[9] E. Eyal, H. Degani, Model-based and model-free parametric analysis of Breast dynamic-contrast-enhanced MRI, NMR in Biomedicine 22 (2009) 40–53, https://doi.org/10.1002/nbm.1221.
[10] A.L. Mackay, C. Laule, Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin, Brain Plasticity 2 (2016) 71–91, https://doi.org/10.3233/bpl-160033.
[11] K.D. Harkins, W.M. Valentine, D.F. Gochberg, M.D. Does, In-vivo multi-exponential T 2, magnetization transfer and quantitative histology in a rat model of intramyelinic edema, NeuroImage: Clinical 2 (2013) 810–817, https://doi.org/10.1016/j.nicl.2013.06.007.
[12] M. Nagtegaal, P. Koken, T. Amthor, J. de Bresser, B. Mädler, F. Vos, M. Doneva, Myelin water imaging from multi-echo T2 MR relaxometry data using a joint sparsity constraint, NeuroImage 219 (2020), https://doi.org/10.1016/j.neuroimage.2020.117014.
[13] M. Molina-Romero, P.A. Gómez, J.I. Sperl, M. Czisch, P.G. Sämann, D.K. Jones, M. I. Menzel, B.H. Menze, A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization, Magnetic Resonance in Medicine 80 (2018) 2155–2172, https://doi.org/10.1002/mrm.27181.
[14] A. Mackay, K. Whittall, J. Adler, D. Li, D. Paty, D. Graeb, In vivo visualization of myelin water in brain by magnetic resonance, Magnetic Resonance in Medicine 31 (1994) 673–677, https://doi.org/10.1002/mrm.1910310614.
[15] D. Kim, J.P. Haldar, Nonnegative matrix factorization for tissue mixture modeling with noisy MR magnitude image sequences, Proceedings - International Symposium on Biomedical Imaging 2015-July (2015) 1028–1031, https://doi.org/10.1109/ISBI.2015.7164046.
[16] T.A. Bjarnason, C.R. McCreary, J.F. Dunn, J.R. Mitchell, Quantitative T2 analysis: The effects of noise, regularization, and multivoxel approaches, Magnetic Resonance in Medicine 63 (2010) 212–217, https://doi.org/10.1002/mrm.22173.
[17] V. Wiggermann, I.M. Vavasour, S.H. Kolind, A.L. MacKay, G. Helms, A. Rauscher, Non-negative least squares computation for in vivo myelin mapping using simulated multi-echo spin-echo T2 decay data, NMR in Biomedicine XXX (2020) 1–14, https://doi.org/10.1002/nbm.4277.
[18] P.J. Slator, J. Hutter, R.V. Marinescu, M. Palombo, L.H. Jackson, A. Ho, L.C. Chappell, M. Rutherford, J.V. Hajnal, D.C. Alexander, Data-Driven multi-Contrast spectral microstructure imaging with InSpect: INtegrated SPECTral component estimation and mapping, Medical Image Analysis 71 (2021) 102045, https://doi.org/10.1016/j.media.2021.102045.
[19] N. Gillis, The Why and How of Nonnegative Matrix Factorization, arXiv 7 Mar (2014) 1–25. http://arxiv.org/abs/1401.5226. arXiv:1401.5226.
[20] H. Benali, I. Buvat, F. Frouin, J.P. Bazin, J. Chabriais, R. Di Paola, Factor Analysis of Medical Image Sequences (FAMIS): Fundamental principles and applications, in: E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, B. Burtschy (Eds.), New Approaches in Classification and Data Analysis, Springer Berlin Heidelberg, Berlin, Heidelberg, 1994, pp. 619–627.
[21] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006, URL http://www.library.wisc.edu/selectedtocs/bg0137.pdf. doi:10.1641/B580519. arXiv:0-387-31073-8..
[22] D.J. Bartholomew, International Encyclopedia of Education, 3rd Edition, Chapter 18: Principal Components Analysis, 3rd ed., Elsevier Ltd, 2010. doi:10.1016/B978-0-08-044894-7.01358-0. arXiv:arXiv:1011.1669v3.
[23] M. Mudrova, A. Prochazka, Principal component analysis in image processing, in: Proceedings of the MATLAB Technical Computing Conference, 3, 2005, p. 4. http://dsp.vscht.cz/konference_matlab/matlab05/prispevky/mudrova/mudrova.pdf.
[24] H. Huang, J. Lu, J. Wu, Z. Ding, S. Chen, L. Duan, J. Cui, F. Chen, D. Kang, L. Qi, W. Qiu, S.W. Lee, S.J. Qiu, D. Shen, Y.F. Zang, H. Zhang, Tumor tissue detection using blood-oxygen-level-dependent functional MRI based on independent component analysis, Scientific Reports 8 (2018) 1–16, https://doi.org/10.1038/s41598-017-18453-0.
[25] P.M. Kim, B. Tidor, Subsystem identification through dimensionality reduction of large-scale gene expression data, Genome Research 13 (2003) 1706–1718, https://doi.org/10.1101/gr.903503.
[26] K. Huang, N.D. Sidiropoulos, A. Swami, Non-Negative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition, IEEE Transactions on Signal Processing 62 (2014) 211–224, https://doi.org/10.1109/TSP.2013.2285514.
[27] Y.X. Wang, Y.J. Zhang, Nonnegative matrix factorization: A comprehensive review, IEEE Transactions on Knowledge and Data Engineering 25 (2013) 1336–1353, https://doi.org/10.1109/TKDE.2012.51.
[28] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (1999) 788–791.
[29] H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, S.H. Jensen, Theorems on positive data: On the uniqueness of NMF, Computational Intelligence and Neuroscience 2008 (2008), https://doi.org/10.1155/2008/764206.
[30] N. Bhatt, A. Ayyar, Monotonous (Semi-) Nonnegative Matrix Factorization, in, in: Proceedings of the Second ACM IKDD Conference on Data Sciences, 2015, pp. 92–97, https://doi.org/10.1145/2732587.2732600.
[31] M.D. Does, Inferring brain tissue composition and microstructure via MR relaxometry, NeuroImage 182 (2018) 136–148, https://doi.org/10.1016/j.neuroimage.2017.12.087.
[32] J. Pfeuffer, U. Flo, W. Dreher, D. Leibfritz, Restricted diffusion and exchange of intracellular water: theoretical modelling and diffusion time dependence of 1H NMR measurements on perfused glial cells, NMR in Biomedicine 11 (1998) 19–31.
[33] S.C. Grant, D.L. Buckley, S. Gibbs, A.G. Webb, S.J. Blackband, MR microscopy of multicomponent diffusion in single neurons, Magnetic Resonance in Medicine 46 (2001) 1107–1112, https://doi.org/10.1002/mrm.1306.
[34] Y. Bengio, Practical recommendations for gradient-based training of deep architectures, in: Practical Recommendations for Gradient-Based Training of Deep Architectures, in: G. Montavon, G.B. Orr, K.R. Müller (Eds.), Neural Networks: Tricks of the Trade, Lecture Notes in Computer Science, vol 7700. Springer, volume 7700, Springer, Berlin Heidelberg, 2012, pp. 437–478, https://doi.org/10.1007/978-3-642-35289-8-26, arXiv:1206.5533.
[35] H. Lundell, M. Nilsson, T.B. Dyrby, G.J. Parker, P.L. Cristinacce, F.L. Zhou, D. Topgaard, S. Lasič, Multidimensional diffusion MRI with spectrally modulated gradients reveals unprecedented microstructural detail, Scientific Reports 9 (2019) 1–12, https://doi.org/10.1038/s41598-019-45235-7.
[36] T.B. Dyrby, W.F. Baaré, D.C. Alexander, J. Jelsing, E. Garde, L.V. Søgaard, An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets, Human Brain Mapping 32 (2011) 544–563, https://doi.org/10.1002/hbm.21043.

[37] K.D. Harkins, A.N. Dula, M.D. Does, Effect of intercompartmental water exchange on the apparent myelin water fraction in multiexponential T 2 measurements of rat spinal cord, Magnetic Resonance in Medicine 67 (2012) 793–800, https://doi.org/10.1002/mrm.23053.
[38] K.P. Whittall, A.L. MacKay, D.A. Graeb, R.A. Nugent, D.K. Li, D.W. Paty, In vivo measurement of T2 distributions and water contents in normal human brain, Magnetic Resonance in Medicine 37 (1997) 34–43, https://doi.org/10.1002/mrm.1910370107.
[39] P.S. Tofts, Quantitative MRI of the Brain: Measuring Changes Caused by Disease, John Wiley and Sons Inc., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2003.
[40] H.W. Fischer, P.A. Rinck, Y. van Haverbeke, R.N. Muller, Nuclear relaxation of human brain gray and white matter: Analysis of field dependence and implications for MRI, Magnetic Resonance in Medicine 16 (1990) 317–334, https://doi.org/10.1002/mrm.1910160212.
[41] P.M. Rasmussen, L.K. Hansen, K.H. Madsen, N.W. Churchill, S.C. Strother, Model sparsity and brain pattern interpretation of classification models in neuroimaging, Pattern Recognition 45 (2012) 2085–2100, https://doi.org/10.1016/j.patcog.2011.09.011.
[42] J.A. Cohen, R.S. Fisher, M.G. Brigell, R.G. Peyster, S. Gordon, The potential for Vigabatrin-induced intramyelinic edema in humans, Epilepsia 41 (2000) 148–157, https://doi.org/10.1111/j.1528-1157.2000.tb00134.x.
[43] J. Zhong, N. Didonato, P.G. Hatcher, Independent component analysis applied to diffusion-ordered spectroscopy: Separating nuclear magnetic resonance spectra of analytes in mixtures, Journal of Chemometrics 26 (2012) 150–157, https://doi.org/10.1002/cem.2423.
[44] B. Jeurissen, J.-D. Tournier, J. Sijbers, Tissue-type segmentation using non-negative matrix factorization of multi-shell diffusion-weighted MRI images, Proc. Intl. Soc. Mag. Reson. Med. 23 (2015) 0349.
[45] D. Christiaens, S. Sunaert, P. Suetens, F. Maes, Convexity-constrained and nonnegativity-constrained spherical factorization in diffusion-weighted imaging, NeuroImage 146 (2017) 507–517, https://doi.org/10.1016/j.neuroimage.2016.10.040.
[46] I. Toumi, B. Torrésani, S. Caldarelli, Effective processing of pulse field gradient NMR of mixtures by blind source separation, Analytical Chemistry 85 (2013) 11344–11351, https://doi.org/10.1021/ac402085x.
[47] N. Dikaios, S. Punwani, V. Hamy, P. Purpura, S. Rice, M. Forster, R. Mendes, S. Taylor, D. Atkinson, Noise estimation from averaged diffusion weighted images: Can unbiased quantitative decay parameters assist cancer evaluation?, Magnetic Resonance in Medicine 71 (2014) 2105–2117, https://doiorg/10.1002/mrm.24877.
[48] O. Dietrich, J.G. Raya, S.B. Reeder, M. Ingrisch, M.F. Reiser, S.O. Schoenberg, Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics, Magnetic Resonance Imaging 26 (2008) 754–762, https://doi.org/10.1016/j.mri.2008.02.001.
[49] P. Stilbs, K. Paulsen, Global least-squares analysis of large, correlated spectral data sets and application to chemical kinetics and time-resolved fluorescence, Review of Scientific Instruments 67 (1996) 4380–4386, https://doi.org/10.1063/1.1147539.
[50] P. Stilbs, Automated CORE, RECORD, and GRECORD processing of multi-component PGSE NMR diffusometry data, European Biophysics Journal 42 (2013) 25–32, https://doi.org/10.1007/s00249-012-0794-8.
[51] A.A. Colbourne, S. Meier, G.A. Morris, M. Nilsson, Unmixing the NMR spectra of similar species - Vive la différence, Chemical Communications 49 (2013) 10510–10512, https://doi.org/10.1039/c3cc46228e.
[52] B. Yuan, Y. Ding, G.M. Kamal, L. Shao, Z. Zhou, B. Jiang, P. Sun, X. Zhang, M. Liu, Reconstructing diffusion ordered NMR spectroscopy by simultaneous inversion of Laplace transform, Journal of Magnetic Resonance 278 (2017) 1–7, https://doi.org/10.1016/j.jmr.2017.03.004.
[53] C.E. Collins, D.C. Airey, N.A. Young, D.B. Leitch, J.H. Kaas, Neuron densities vary across and within cortical areas in primates, Proceedings of the National Academy of Sciences of the United States of America 107 (2010) 15927–15932. doi:10.1073/pnas.1010356107.
[54] J. Wen, M.S. Goyal, S.V. Astafiev, M.E. Raichle, D.A. Yablonskiy, Genetically defined cellular correlates of the baseline brain MRI signal, Proceedings of the National Academy of Sciences of the United States of America 115 (2018) E9727–E9736. doi:10.1073/pnas.1808121115.
[55] C. Najac, F. Branzoli, I. Ronen, J. Valette, Brain intracellular metabolites are freely diffusing along cell fibers in grey and white matter, as measured by diffusion-weighted MR spectroscopy in the human brain at 7 T, Brain Structure and Function 221 (2016) 1245–1254, https://doi.org/10.1007/s00429-014-0968-5.
[56] D. Topgaard, Multidimensional diffusion MRI, Journal of Magnetic Resonance 275 (2017) 98–113, https://doi.org/10.1016/j.jmr.2016.12.007.
[57] O. Reynaud, Time-dependent diffusion MRI in cancer: Tissue modeling and applications, Frontiers in Physics 5 (2017) 1–16, https://doi.org/10.3389/fphy.2017.00058.
[58] D. Xing, N.G. Papadakis, C.L.H. Huang, V.M. Lee, T.A. Carpenter, L.D. Hall, Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain, Magnetic Resonance Imaging 15 (1997) 771–784, https://doi.org/10.1016/S0730-725X(97)00037-4.
[59] S. Rane, T.Q. Duong, Comparison of In Vivo and Ex Vivo Diffusion Tensor Imaging in Rhesus Macaques at Short and Long Diffusion Times, The Open Neuroimag Journal 5 (2011) 172–178, https://doi.org/10.2174/1874440001105010172.
[60] A. Roebroeck, K.L. Miller, M. Aggarwal, Ex vivo diffusion MRI of the human brain: Technical challenges and recent advances, NMR in Biomedicine 32 (2019) 1–14, https://doi.org/10.1002/nbm.3941.
[61] D.S. Novikov, V.G. Kiselev, S.N. Jespersen, On modeling, Magnetic Resonance in Medicine 79 (2018) 3172–3193, https://doi.org/10.1002/mrm.27101.
[62] G.J. Stanisz, R.M. Henkelman, Diffusional anisotropy of T2 components in bovine optic nerve, Magnetic Resonance in Medicine 40 (1998) 405–410, https://doi.org/10.1002/mrm.1910400310.
[63] L. Muzzarelli, S. Weis, S.B. Eickhoff, K.R. Patil, Rank Selection in Non-negative Matrix Factorization: Systematic comparison and a new MAD metric, International Joint Conference on Neural Networks (IJCNN), volume July, IEEE, Budapest, Hungary, 2019, pp. 1–8, https://doi.org/10.1109/IJCNN.2019.8852146.

Supplementary Material

Data-driven separation of MRI signal components for tissue characterization

Sofie Rahbek, Kristoffer H. Madsen, Henrik Lundell, Faisal Mahmood and Lars G. Hanson

Lars G. Hanson
lghan@dtu.dk

The supplementary material includes:

- Supplementary text
- Supplementary tables (S1)
- Supplementary figures (S1-S5)

Supplementary text

A. Matrix definition

The definition of the matrices, **A** and **B**:

$$\mathbf{A} = \begin{bmatrix} \mathbf{D}_1 & & & \\ & \mathbf{D}_2 & & \\ & & \ddots & \\ & & & \mathbf{D}_k \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{E}_1 & & & \\ & \mathbf{E}_2 & & \\ & & \ddots & \\ & & & \mathbf{E}_k \end{bmatrix}$$

Where $\mathbf{D}_{i=1,2,\dots,k} \in \mathbb{R}^{(m-1)\times m}$ and $\mathbf{E}_{i=1,2,\dots,k} \in \mathbb{R}^{(m-2)\times m}$ are defined as the following Toeplitz matrices:

$$\mathbf{D}_i = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix}, \quad \mathbf{E}_i = \begin{bmatrix} \frac{-1}{\Delta t_1} & \frac{\Delta t_1+\Delta t_2}{\Delta t_1 \cdot \Delta t_2} & \frac{-1}{\Delta t_2} & & \\ & \frac{-1}{\Delta t_2} & \frac{\Delta t_2+\Delta t_3}{\Delta t_2 \cdot \Delta t_3} & \frac{-1}{\Delta t_3} & \\ & & \ddots & \ddots & \ddots \\ & & & \frac{-1}{\Delta t_{m-2}} & \frac{\Delta t_{m-2}+\Delta t_{m-1}}{\Delta t_{m-2} \cdot \Delta t_{m-1}} \quad \frac{-1}{\Delta t_{m-1}} \end{bmatrix}$$

$\mathbf{D}_i$ and $-\mathbf{D}_i$ are used for monotonically increasing and monotonically decreasing components, respectively. In the same way, $\mathbf{E}_i$ and $-\mathbf{E}_i$ are used for components with a monotonically increasing first derivative and a monotonically decreasing first derivative, respectively. The first derivative of the signal is approximated by the slope of the linear segment between two consecutive measurements (mean value theorem). $\Delta t_{j=1,2\dots,m-1}$ is the interval between two consecutive measurements (j and $j+1$) of the independent variable (time, b-value, etc.).

B. Objective function

Since the Frobenius norm of the residual is used as objective function, the overall problem can be stated as

$$J = \frac{1}{2}\sum_{ij}\left(X_{ij} - [WH]_{ij}\right)^2 = ||\,\mathbf{X} - \mathbf{WH}||_F^2$$

Including a vector, $\mathbf{u} \in \mathbb{R}^{1\times n}$, to allow individual weights of the n observations, the objective function becomes:

$$J_u = \frac{1}{2}\sum_{i=1}^{n} u_i(\mathbf{x}_i - \mathbf{W}\mathbf{h}_i)^2$$

Supplementary tables

For coil array data, coil combination using the complex sensitivity profiles as weights is preferred as it does not introduce a noise bias in the resulting phased image. However, sum-of-squares coil combination is common and the effect of a Rician bias on the decomposition is therefore explored using realistic coil sensitivity profiles generated with the JEMRIS simulator (open source software, http://www.jemris.org/). The bias becomes particularly severe for low SNR and many coil elements, as reflected by the results presented in the following table.

N_{coils}	SNR (σ)	$RMSE_W$		EV_W		ρ_H		AUC_H	
1	65.5 (0.02)	0.045	0.028	0.986	0.991	0.979	0.990	1.000	1.000
	21.8 (0.06)	0.059	0.063	0.961	0.954	0.938	0.943	0.999	1.000
	13.1 (0.10)	0.071	0.065	0.934	0.942	0.892	0.901	0.993	0.995
8	65.5 (0.02)	0.071	0.050	0.919	0.962	0.938	0.952	1.000	1.000
	21.8 (0.06)	0.186	0.092	0.327	0.915	0.758	0.767	0.934	0.867
	13.1 (0.10)	0.231	0.163	0.699	0.899	0.361	0.318	0.686	0.706
16	65.5 (0.02)	0.070	0.048	0.909	0.964	0.939	0.955	1.000	1.000
	21.8 (0.06)	0.206	0.070	0.118	0.940	0.773	0.799	0.938	0.951
	13.1 (0.10)	0.365	0.137	0.119	0.929	0.313	0.600	0.707	0.818
32	65.5 (0.02)	0.074	0.059	0.900	0.942	0.945	0.954	1.000	1.000
	21.8 (0.06)	0.207	0.133	0.071	0.595	0.796	0.824	0.944	0.958
	13.1 (0.10)	0.221	0.115	0.798	0.908	0.557	0.653	0.728	0.774

Table S1. The performance of the msNMF for a set of simulations with varying SNR and number of receive coils (N_{coils}), indicated in the two left-most columns. The SNR is calculated as the mean signal across the "object" for the first measurement (m=1) divided by σ, specifying the level of noise added. The following columns show evaluation metrics for the factorization: **W** is compared to true signals using the root mean squared error (RMSE) and the explained variance (EV), and **H** is compared to the binary true maps using the Pearson correlation coefficient (ρ) and the area under the receiver operating characteristics curve (AUC). The orange values indicate the results when one growing component has been added ($k = 4$, $c = 1$) to partially compensate Rician noise bias. The improvement is prominent for the combination of coil arrays and low SNR, especially considering reconstruction of the signal profiles (**W**) as also illustrated in Figure S4.

Supplementary figures

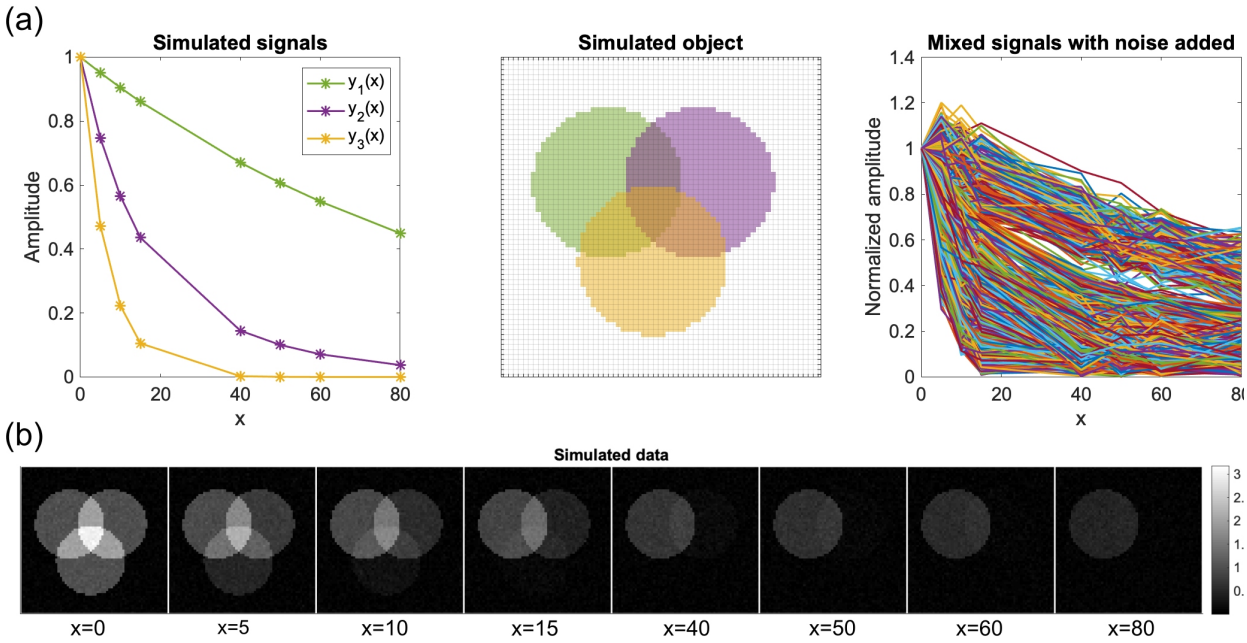

Figure S1. Simulated data ($\sigma = 0.06$, one uniform receive coil). (a): The three basic signals are seen to the left. In the middle their mixing pattern is illustrated. To the right, the final mixed signals scaled to start in 1 are seen. (b): The spatial presentation of the final simulated data for all eight measurements.

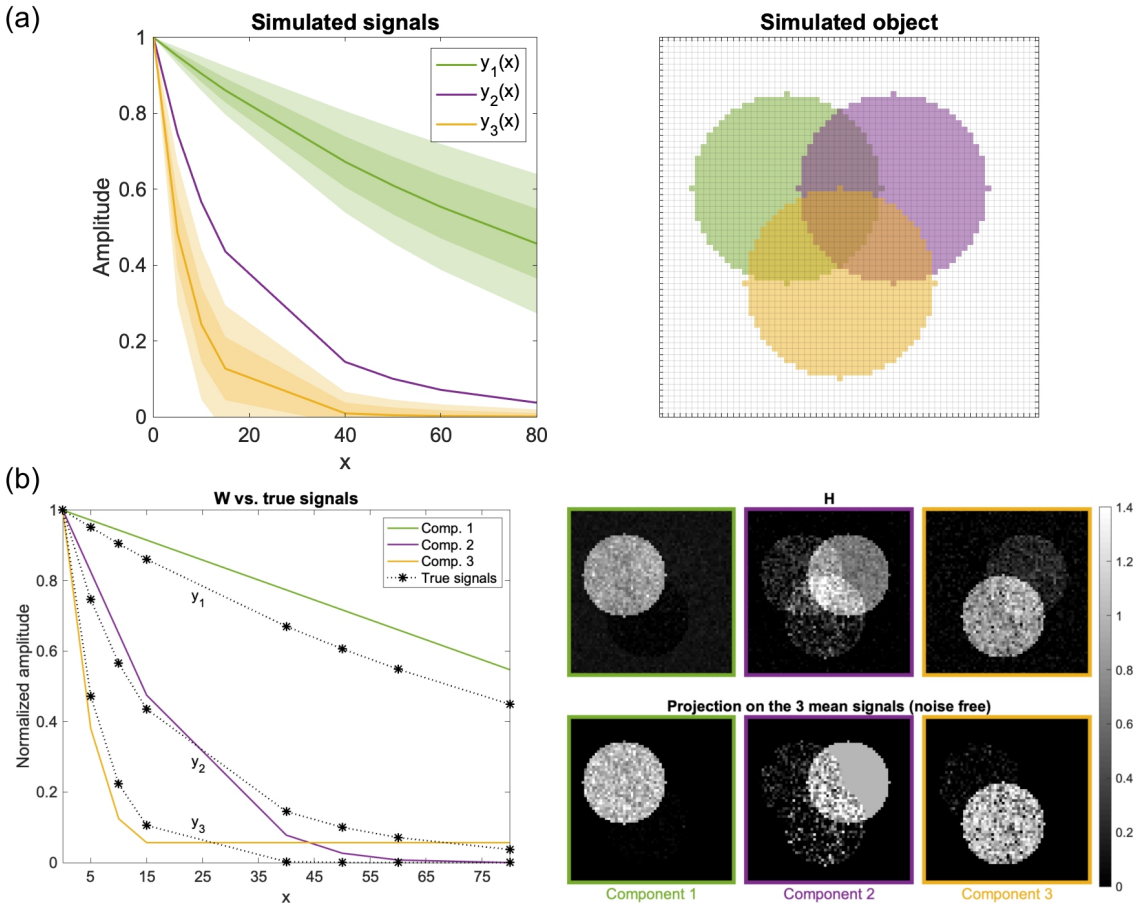

Figure S2. (a): Simulated data ($\sigma = 0.06$, one uniform receive coil) where the two mono-exponential signal components, $y_1(x)$ and $y_3(x)$, have decay rates (r) drawn from a normal distribution, $\mathcal{N}(\mu, \sigma^2)$. For $y_1(x)$, $r \sim \mathcal{N}(0.01, 0.0025^2)$ and for $y_3(x)$, $r \sim \mathcal{N}(0.15, 0.0375^2)$, i.e. the means correspond to the original decay rates and the standard deviations correspond to 25% of these. To the left, the signal components are visualized with $\pm\sigma$ and $\pm 2\sigma$ marked by the dark and light shaded areas, respectively. To the right, their mixing patterns are illustrated. (b): The msNMF result compared to the true decomposition. The spatial mixing map (**H**) is compared to the projection of the noise-free data on to the three mean signal components, representing the optimal 3-component decomposition in absence of noise.

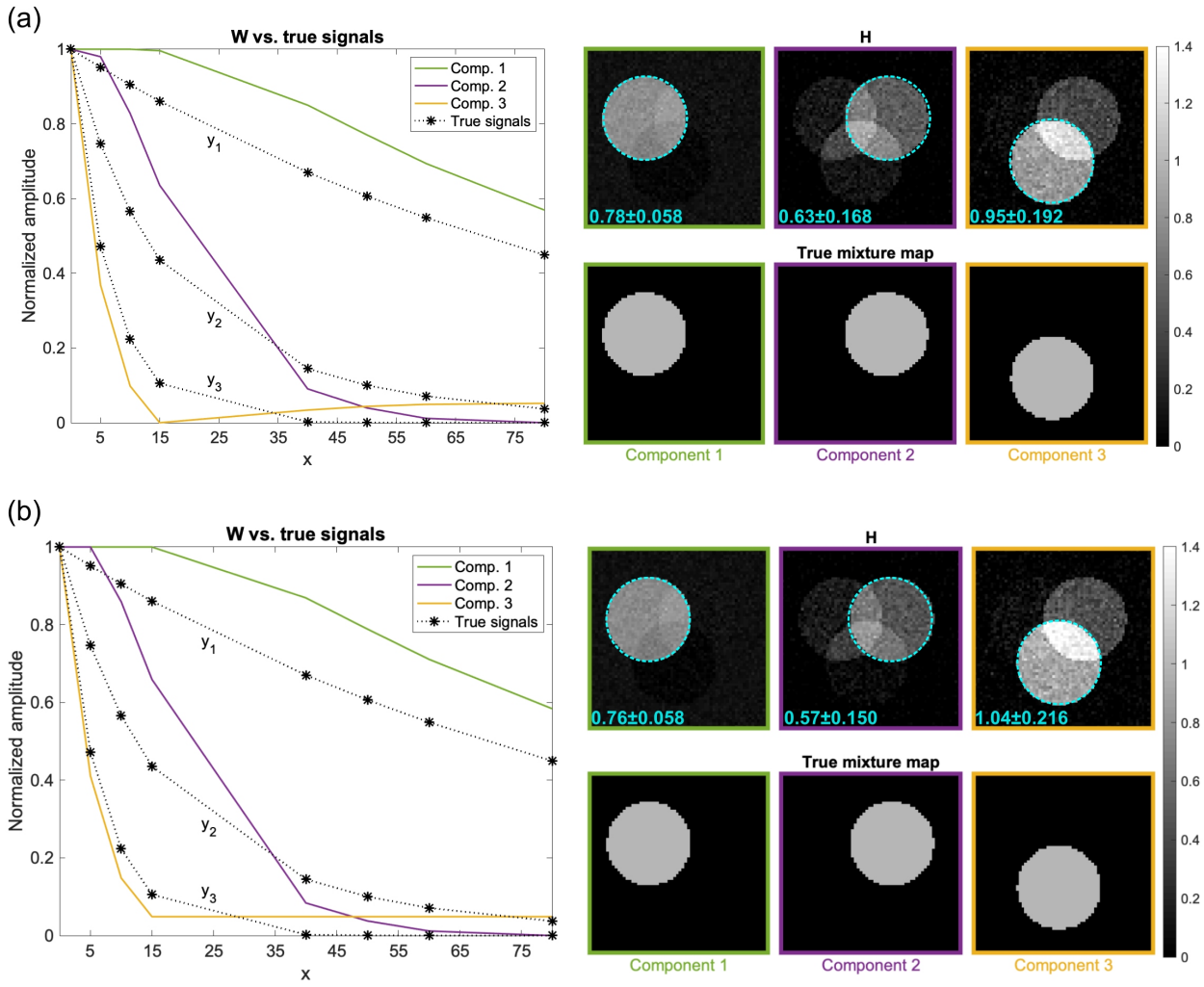

Figure S3. (a): The true decomposition compared to the result of the standard NMF, i.e. without any monotonicity constraints ($EV_{\mathbf{W}} = 0.922$ and $AUC_{\mathbf{H}} = 0.989$). (b): The true decomposition compared to the result of the monotonous NMF, i.e. with a monotonicity constraint only for the non-differentiated ($EV_{\mathbf{W}} = 0.926$ and $AUC_{\mathbf{H}} = 0.983$). The signal components are given to the left and the corresponding mixture maps to the right. The cyan numbers are mean±SD of the encircled voxel values. Comp. = component.

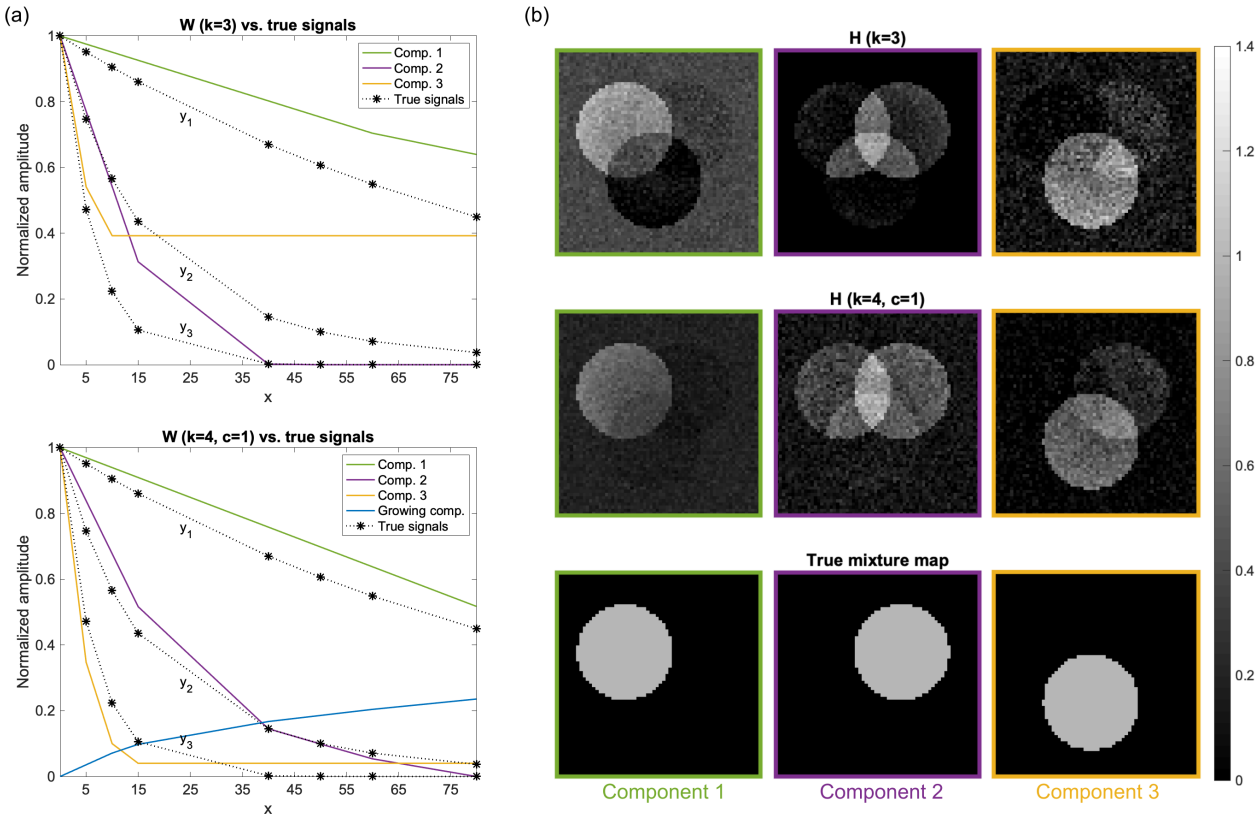

Figure S4. Results of the msNMF for simulated data with $\sigma = 0.06$ and 16 receive coils (sum-of-squares reconstruction) with realistic sensitivity profiles generated using the JEMRIS simulator (open source software, http://www.jemris.org/). Factorizations without ($k = 3$) and with ($k = 4$, $c = 1$) the addition of a spatially uniform growing component are compared. (a): The extracted signal components. (b): The mixture maps. The uniform map for the growing component is left out. The true components are included for comparisons. Comp. = component.

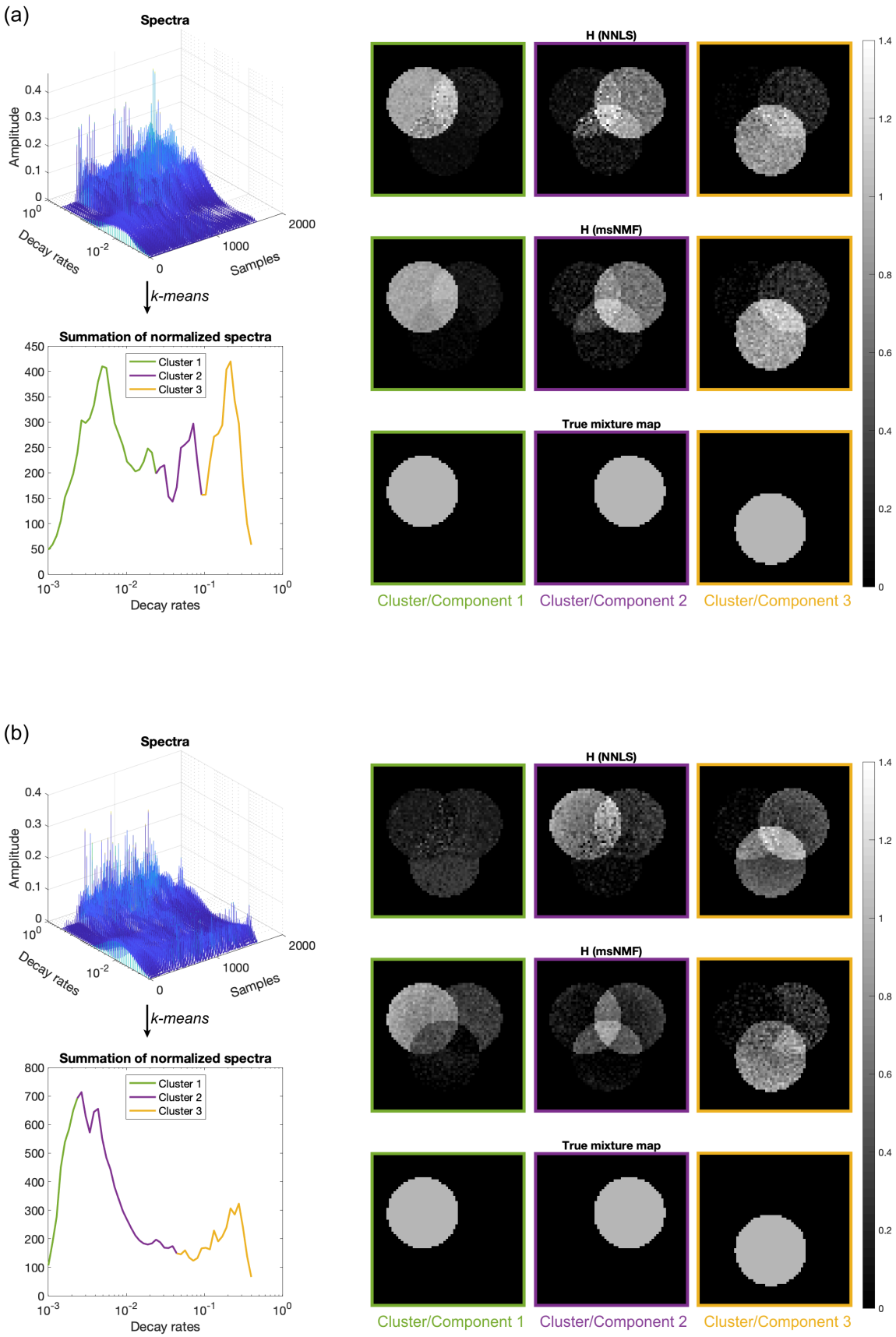

Figure S5. Results using regularized NNLS with cross-validated penalty parameter and log-spaced decay rates (software from the open-source AnalyzeNNLS, http://sourceforge.net/projects/analyzennls/). In each subfigure, the result is presented by the raw spectra, the summed spectrum with the partitioning of the decay rates using k-means clustering ($k = 3$) indicated, and finally the resulting "cluster maps" (**H**), compared to mixture maps of the msNMF and true maps (with the background masked away). (a): Results using simulated data with $\sigma = 0.06$ and a single receive coil. ρ and AUC for **H** (NNLS) compared to true maps are 0.915 and 0.996 respectively. (b): Results using simulated data with $\sigma = 0.06$ and a 16-element coil array (sum-of-squares reconstruction). ρ and AUC for **H** (NNLS) compared to true maps are 0.640 and 0.621 respectively.

PAPER II

B

TITLE

Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI

AUTHORS

Rahbek, Sofie; Mahmood, Faisal; Tomaszewski, Michal R.; Hanson, Lars G; Madsen, Kristoffer H.

YEAR

2022

PUBLICATION HISTORY

Manuscript ready for submission
Submission date: June, 2022

Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI

Sofie Rahbek[1], Faisal Mahmood[2,3], Michal R. Tomaszewski[4], Lars G. Hanson[1,5] and Kristoffer H. Madsen[5,6]

[1]Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark

[2]Department of Clinical Research, University of Southern Denmark, Odense, 5000, Denmark

[3]Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital, Odense C, 5000, Denmark

[4]Translation Imaging Department, Merck & Co, West Point, PA, USA

[5]Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, 2650, Denmark

[6]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark

E-mail: khma@dtu.dk

Abstract. *Objective* In the field of radiation oncology, the benefit of MRI goes beyond that of providing high soft-tissue contrast images for staging and treatment planning. With the recent clinical introduction of hybrid MRI linear accelerators (MR-Linacs) it has become feasible to map physiological parameters describing diffusion, perfusion, and relaxation during the entire course of radiotherapy, for example. However, advanced data analysis tools are required for extracting qualified prognostic and predictive imaging biomarkers from longitudinal MRI data. In this study, we propose a new prediction framework tailored to exploit temporal dynamics of tissue features from repeated measurements. We demonstrate the framework using a newly developed decomposition method for tumor characterization. *Approach* Two previously published MRI datasets with multiple measurements during and after radiotherapy, were used for development and testing: T_2-weighted multi-echo images obtained for two mouse models of pancreatic cancer, and diffusion-weighted images for patients with brain metastases. Initially, the data was decomposed using the novel monotonous slope non-negative matrix factorization (msNMF) tailored for MR data. The following processing consisted of a tumor heterogeneity assessment using descriptive statistical measures, robust linear modelling to capture temporal changes of these, and finally logistic regression analysis for stratification of tumors and volumetric outcome. *Main Results* The framework was able to classify the two pancreatic tumor types with an AUC of 0.999, $P < 0.001$ and predict the tumor volume change with a correlation coefficient of 0.513, $P = 0.034$. A classification of the human brain metastases into responders and non-responders resulted in an AUC of 0.74, $P = 0.065$.

Significance A general data processing framework for analyses of longitudinal MRI data has been developed and applications were demonstrated by classification of tumor type and prediction of radiotherapy response. Further, as part of the assessment, the merits of msNMF for tumor tissue decomposition were demonstrated.

Keywords: prediction framework, decomposition, longitudinal MRI, radiotherapy, tumor features, treatment response. Submitted to: *Phys. Med. Biol.*

1. Introduction

MRI has become a crucial modality within cancer management. With the ability to deliver high soft-tissue contrast, conventional MRI techniques are part of clinical routine to enable morphological characterization of tumors. Beyond that, advanced MRI techniques that allow quantitative analysis can characterize the local tumor microenvironment. This increases accuracy in tumor grading, subtype classification and the chance to offer an optimal therapy plan [1, 2]. An example is diffusion-weighted MRI (DWI), which is sensitive to the local cellular density, potentially important for identification of sub-regional tumor load and for early evaluation of a tumor's response to therapy [3]. This is due to the response to ionizing radiation at a cellular level occurring on a much shorter timescale (hours to days) than the volumetric response [4]. With an early response assessment there is time to adjust the therapy plan to increase treatment efficacy or reduce unnecessary radiation [5, 6, 7, 8].

The recent introduction of hybrid MRI linear accelerator systems (MR-Linacs)[9, 10] facilitates daily MRI-based adaption of the dose plan and offers opportunities for longitudinal MRI studies of therapy response within the normal clinical workflow. However, there is a need for more custom analysis strategies for the growing pool of longitudinal MRI data. Currently, only 16 previous human studies include more than two measurements during treatment [8], and as few as two datasets include daily measurements [11, 12]. An analysis pipeline customized for longitudinal datasets requires a method for investigating temporal tumor dynamics, but also a procedure for extracting quantitative information from multi-contrast MRI data (consisting of three spatial dimensions and one b-value dimension for diffusion-weighting, for example). Most previous studies have utilized a model-based parametrization of the data for quantification and e.g. showed correlations between the DWI-derived apparent diffusion coefficient (ADC) and treatment outcome, e.g. [13, 14], or between dynamic contrast-enhanced (DCE) MRI parameters and tumor subtype and prognosis, e.g. [15, 16]. Model-based analyses, although useful, carry the risk of introducing biased or misleading estimates if model assumptions are wrong, e.g due to partial-volume effects. Novikov et al [17] outlines pitfalls of modelling tissue microstructure, and Satta et al [18] highlights contradicting correlations found across studies that analyzed DWI- and DCE-derived parameters, demonstrating the problem and hence a need for alternatives such as data-driven analysis.

Among the few existing studies that included longitudinal measurements in their correlation analysis, most of them investigated the temporal changes only by pairwise comparisons [19, 20, 21, 22, 23, 24]. This results in extensive multiple comparisons problems for studies involving more than 2-3 measured time points. Additionally, data trends unfolding over several time-points will be difficult to identify, though they may be important in a therapy outcome prediction, for example. Advanced models such as neural networks has been used to discover patterns across time-resolved measurements

[25, 26], but these analyses were voxel-based and thus depended on an exact image registration between measurements. This may be problematic if morphological changes occur during the course of disease.

Therefore, we propose an analysis framework which avoids both model-based parametrization for quantitative mapping and the need for voxel correspondence over time [7]. The framework provides all processing steps necessary to analyze and exploit the dynamics across longitudinal measurements for a final prediction. As an alternative to model-based analysis, the framework utilizes a data-driven blind-source separation technique to decompose the MRI data. Though any decomposition technique in principle can be used, we suggest and explore the recently published monotonous slope non-negative matrix factorization (msNMF)[27]. The msNMF is specifically tailored for decomposition of MR signals and has proven useful for extracting realistic components informative of the underlying tissue structure.

The aim of this paper is to present a full prediction framework in which longitudinal MRI data is analyzed in order to extract information with potential prognostic or predictive value in cancer management, and to explore the msNMF as the initial decomposition method for tumor tissue characterization. We use two previously analysed and published data sets, T_2-weighted multi-echo data for pancreatic tumor type classification in mice, and DWI of human brain metastases for outcome prediction.

2. Methods

2.1. The prediction framework

An overview of the prediction framework is summarized in Figure 1a, and described below.

2.1.1. Data-driven decomposition As a first step, the data matrix was separated into latent signal components using msNMF, which is an extension of standard non-negative matrix factorization (NMF). It constrains the signal components and their slopes (first derivates) to be monotonous, thus enforcing near-exponential signal decays consistent with prior expectations for the relevant data types (e.g. DWI or relaxometry). A detailed description of the msNMF was published by Rahbek et al [27]. For brevity, only the resulting optimization problem is summarized here:

$$\mathcal{C} = \|\mathbf{X} - \mathbf{WH}\|_F^2 \tag{1}$$

The norm of the residual, $\mathcal{C}$, is minimized, where $\|\cdot\|_F^2$ denotes the Frobenius norm. $\mathbf{X} \in \mathbb{R}_+^{m\times n}$ is the data matrix for m temporal measurement points (e.g. echo times) and n sources (e.g. voxels). $\mathbf{WH}$ is the low-rank representation of the data with $\mathbf{W} \in \mathbb{R}_+^{m\times k}$ being the k fundamental signal profiles and $\mathbf{H} \in \mathbb{R}_+^{k\times n}$ the associated spatial distributions ("mixture maps"). $\mathcal{C}$ is minimized using an alternating non-negative least

squares (ANLS) algorithm, i.e. by optimizing one factor while keeping the other fixed. **H** is optimized under a non-negativity constraint, and **W** is optimized under both a non-negativity constraint and the additional monotonicity constraints. Strategies demonstrated by Rahbek *et al* [27] were followed regarding implementation including initialization, stopping criteria and rank determination.

2.1.2. Tumor heterogeneity assessment The decomposition output, k mixture maps describing the spatial distributions of different signal features across the volume, was analyzed for assessment of tumor features using a set of first-order radiomics. Five percentiles, (0.1, 0.25, 0.5, 0.75 and 0.9), were used to characterize the intensity distribution across the tumor volume, defined by a region-of-interest (ROI) for each component. ROI histogram values have previously been demonstrated as relevant descriptors of tumor structures [5, 7, 28, 29]. Additionally, they are robust to outliers and imperfect image registrations as opposed to voxel-based values and the ROI mean. The quartiles (0.25, 0.5, and 0.75) were supplemented with bins for extreme values, (0.1 and 0.9), to capture features of non-normal distributions while disregarding outliers, if such are present. The ROI delineation is explained in coming subsections dedicated to the individual data sets.

2.1.3. Temporal dynamics analysis The five component percentiles were calculated for all subjects and scan time-points. To capture a trend across time-resolved measurements, a linear fit was made for each percentile as a function of days after first treatment. A Huber loss function was used in the regression for the fit to reduce influence of potential outliers [30]. The baseline scan (obtained prior to treatment) was left out from the fit. Instead, the difference between the values at baseline and the first scan after treatment onset was calculated to include the immediate treatment response explicitly in the prediction model. The final prediction features then constituted the slope and intercept of the linear fit and this "difference to baseline" for each of the five percentiles generated for each of the components. An example is presented in Figure 2 using one of the observations in the DWI dataset. Lastly, each feature was normalized to the range [-1 1], such that all features were weighted equally in the prediction analysis.

2.1.4. Feature selection and prediction modelling The above processing resulted in a large number of features compared to the number of samples. To avoid overfitting and numerical instability [31, 32], the prediction analysis was thus carried out using an integrated feature selection process. Logistic regression was used as statistical model for prediction of classes and multiple (linear) regression for prediction of continuous outcomes. A nested cross-validation (CV) procedure was implemented for model training to avoid information leakage into the test data. In the outer CV loop, stratified K-fold cross-validation was used to divide data into a number of randomly stratified groups equal to the number of samples in the least frequent class for the classification analysis. For the regression analysis, leave-one-out cross-validation was used, resulting

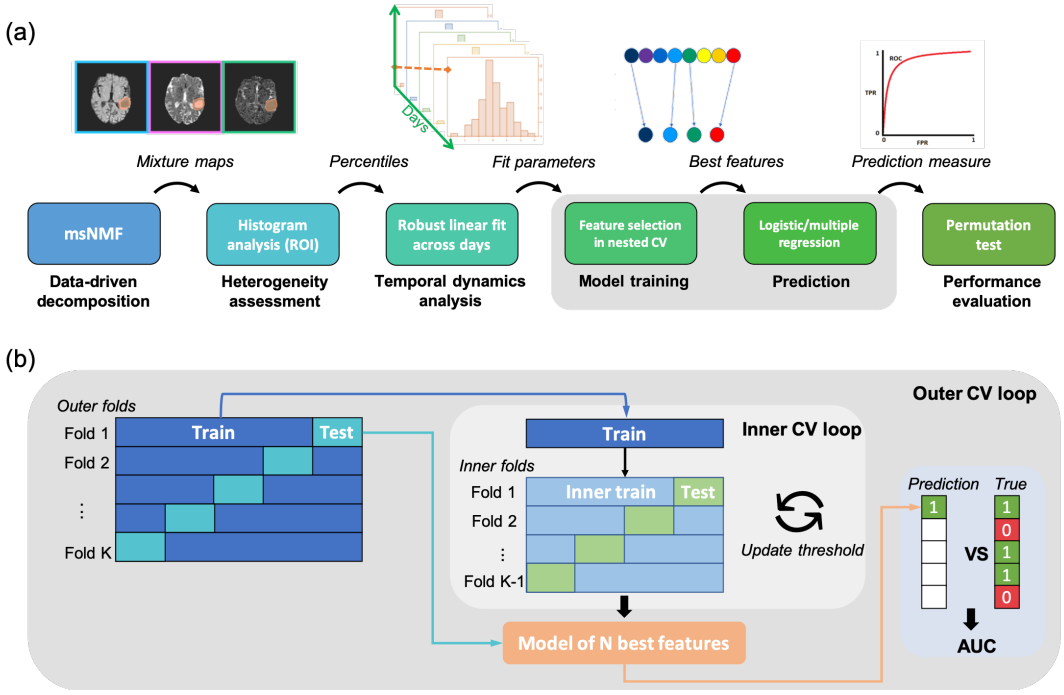

Figure 1. (a) The full prediction framework from the data decomposition to the final prediction. A nested cross-validation procedure was used for the feature selection and model training (gray box). This process is illustrated more detailed in (b).

in one sample per group. As the total number of samples were low, this setting was used to ensure that the maximum number of samples were available for training. The inner CV loop then served to determine which features were most relevant for prediction, and the input data was thus again divided into test and training data (Figure 1b). To this end, we relied on a greedy approach where the features were ranked according to test statistics from a univariate Student's t-test with pooled variance within the remaining groups. The performance estimated in the inner CV loop was used to determine the optimal threshold, and the model was retrained on the entire training set with the optimal threshold. As we did not use regularization in the current analysis, the maximum number of features was set to 10 to avoid overfitting and rank deficiency. The final generalization performance was evaluated in terms of the area under the receiver operating characteristic curve (AUC) or the Pearson correlation coefficient (ρ). The significance of the performance metrics was evaluated by a permutation test where an empirical null distribution was formed by repeating the entire analysis 5000 times with randomly permuted labels, which is a commonly used method for validating prediction performance [33]. A threshold of $\alpha = 0.05$ defined statistical significance.

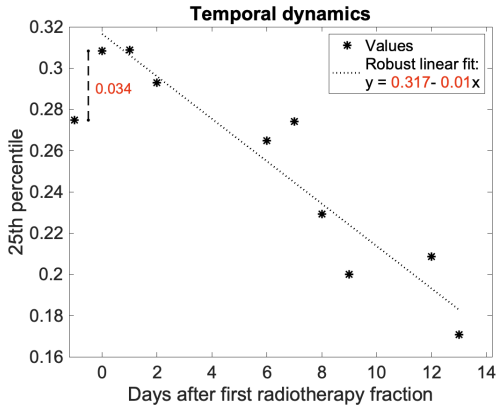

Figure 2. An example of the temporal fit using a Hubert loss function (dotted line). "Days after first treatment" is used as explanatory variable. The 25th percentile of component 1 for a random tumor is used as example data. The fitting parameters together with the difference to baseline (dashed line) are marked with a red font as these constitute the prediction features for this given case.

2.2. T_2-weighted multi-echo imaging of mice

This dataset was previously published by Tomaszewski et al [34] and consists of MR imaging of mice grafted subcutaneously with either slow growing human pancreatic adenocarcinoma BXPC3 (N=8) or more aggressive mouse pancreatic adenocarcinoma Panc02 tumors (N=10). A multi-slice multi-echo sequence with 32 echoes and 7 ms echo spacing was used. The sequence was part of a full MRI protocol applied two hours prior to irradiation and every 3 days after, using a 7 T animal scanner (Bruker horizontal bore, running Paravision 6.0.1). The final imaging was performed at day 9 and day 12, respectively, for the Panc02 and BXPC3 group. The radiotherapy consisted of a single dose of 10 Gy irradiation, matching the dose per fraction of a clinical stereotactic body radiation protocol. The events of imaging and radiation is schematically illustrated in Figure 3a.

For each mouse and scan session, the tumor region, i.e. the region-of-interest (ROI), was delineated on a high-resolution T_2-weighted anatomical scan included in the full protocol. The ROI voxels for all mice and all scan days were pooled to a single [$N_{echoes} \times N_{voxels}$] matrix for the msNMF analysis. The data was normalized to start at intensity 1 to mitigate effects of irrelevant signal behavior across the brain, e.g. stemming from receive coil inhomogeneities, which otherwise would influence the decomposition.

In the source paper [34], Tomaszewski et al calculated model-based T_2-values for each voxel in the tumor ROI and showed (statistical) differences between the two tumor types. Additionally, the paper revealed correlations between the T_2 inter-quantile range and the therapy response defined as the tumor volume change measured between the two last MRI scans. In this study, the data was thus ideal to test the proposed framework for both a classification of the two tumor types and a prediction of the therapy response.

For the therapy response prediction, the BXPC3 and Pan02 groups were pooled together due to their small sample sizes.

2.3. DWI of brain metastases

This clinical dataset was previously analyzed by Mahmood et al [11, 35, 36] and consists of MRI brain scans of cancer patients undergoing palliative radiotherapy. The data was recorded using a 1 T MR system (Panorama, Philips Healthcare, The Netherlands). In addition to T_1- and T_2-weighted imaging, the MRI-protocol included an echo-planar imaging (EPI) DWI sequence with eight b-values of [0, 50, 100, 150, 400, 500, 600, 800] s/mm^2, three orthogonal diffusion-weighting directions, an effective diffusion time ∼49 ms, and a prior spectral inversion module (SPIR) for fat-suppression. The radiotherapy consisted of 10 fractions of 3 Gy whole-brain irradiation delivered over two weeks. The MRI scan sessions took place prior to, during, and after the treatment period (Figure 3b).

For each patient, one to three brain metastases were delineated by a radiologist using b=800 s/mm^2 images, guided by the T_2-weighted data. All scans were aligned with the baseline scan using rigid registration (SPM12 Statistical Parametric Mapping software, version 7487 [37]) such that the baseline tumor delineation could define the ROI for all scans. However, ROIs were expanded using a morphological dilation of 3 voxels (5.4 mm). This was both to include sub-clinical disease, edema etc. and to increase robustness to imperfect image registrations or inaccurate tumor delineation. ROI voxels were pooled for all metastases, i.e. from all subjects and scan days, resulting in a [N_{voxels} × $N_{bvalues}$] data matrix for the msNMF, although input signals largely affected by noise were discarded. These voxels were defined as having maximum signal for a non-zero b-value. The data was normalized to start at an intensity of 1.

The full framework was tested for prediction of the local control evaluated at a follow-up scan obtained 2-3 months after last irradiation. The tumor volume change from baseline to follow-up, based on a radiologist's delineation on a high-resolution T_1-weighted scan, was used to divide the metastases into responders (more than 30 % volume shrinkage) and non-responders (less than 30 % volume shrinkage). A total of 31 metastases (23 responders and 8 non-responders) distributed in 15 patients were analysed.

3. Results

3.1. T_2-weighted multi-echo imaging of mice

3.1.1. Decomposition with msNMF An inspection of the decomposition of ranks two to four resulted in a decomposition into three components, corresponding to 99.4% explained data variance. Figure 4a shows the resulting signal components, all being relatively smooth, realistic signal decays. The associated mixture maps (**H**) are presented in Figure 4b for a representative mouse from each group (BXPC3 and Panc02)

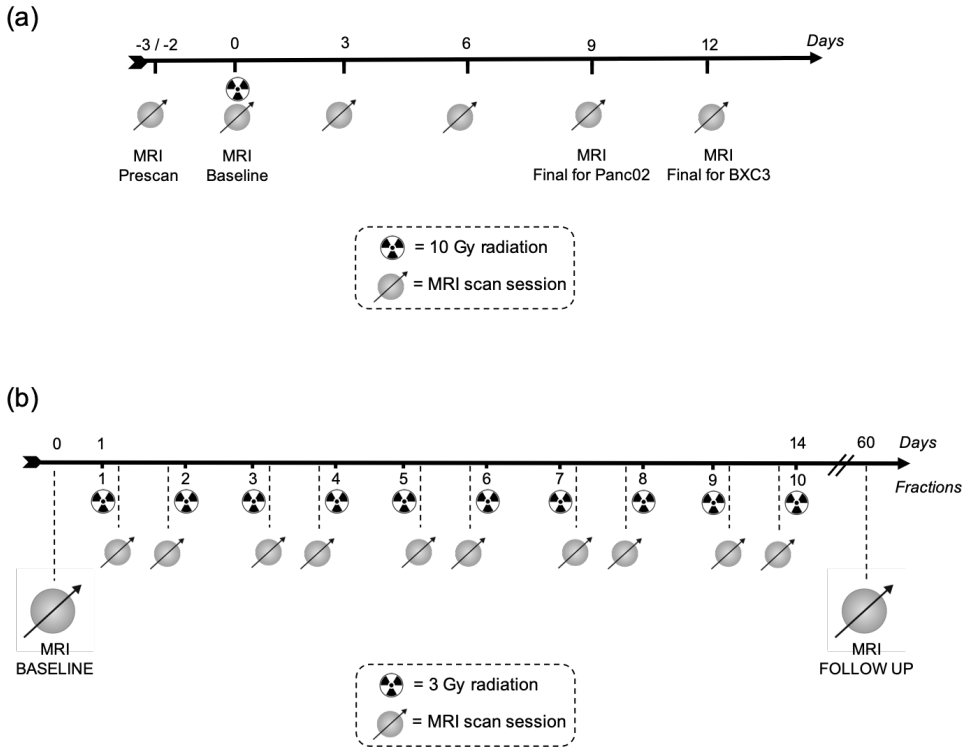

Figure 3. Timeline of the radiotherapy and MRI events, where the MRI baseline scan (last scan before radiation) defines "day 0". (a): The events for the mice pancreatic cancer study. (b): The events for the human brain metastases study. Notice the interleaved structure with an MRI scan right after and right before every second radiotherapy fraction*.

with the tumor ROIs shown in red. Even though both tumor types predominantly contain the intermediate decaying signal component (yellow), it is clear from the images that the two tumors differ in tissue heterogeneity. Looking at the mixture maps for the long-lived signal component (blue), for example, only the BXPC3 tumor contains a sub-area with relatively high intensity.

3.1.2. Classification of the tumor type Due to the randomization included in the K-fold CV procedure, running the prediction framework 50 times resulted in a mean AUC score of 0.999, i.e. an almost perfect classification of the two tumor types. Figure 5 shows this score relative to the null distribution of 5000 runs with permuted class labels. The corresponding p-value was equal to 0.0004, which confirms that the classifier has learned a significant class structure and could distinguish the two tumor types.

Except for a single feature related to the green signal component, the blue long-lived signal component was the only relevant predictor in a classification of the tumor type (Figure 5). Four percentiles are represented for this component distribution, and both changes in intercepts and slopes explain a significant part of the variance. The

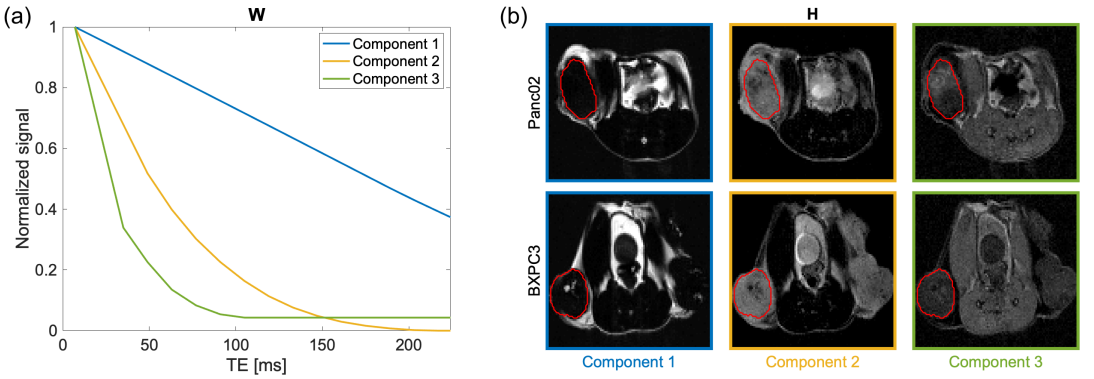

Figure 4. (a) The msNMF signal components (**W**) of mouse data. (b) The associated mixture maps (**H**) for two mice: one from each group. The red contour indicates the tumor ROI. The images are presented with a common, arbitrary intensity scale.

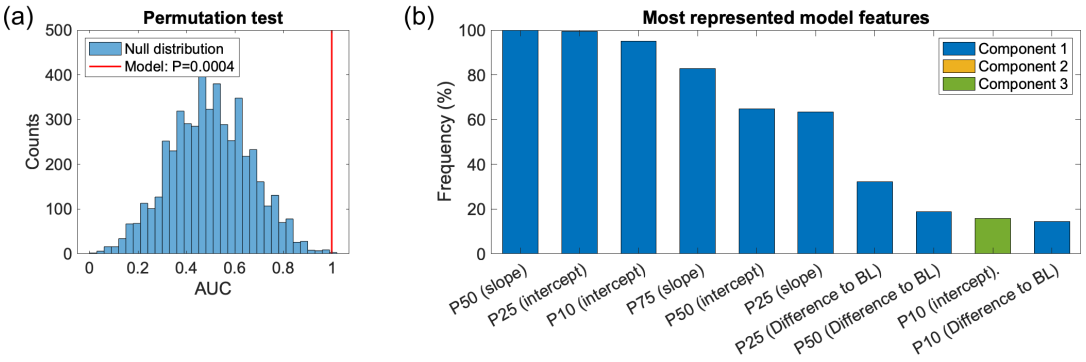

Figure 5. The result of the logistic regression for classification of tumor type. (a): The model AUC score (red line) compared to the null distribution generated from 5000 runs with permuted class labels (P=0.0004). (b): The most represented model features. The bar color indicates msNMF component, and PXX indicates percentile XX. The "slope" and "intercept" are from the temporal linear regression model. The "difference to BL" refers to the difference between the baseline scan and the subsequent scan at day 3.

latter indicates that the signal characteristics of the two tumor types differ with respect to both starting point and changes following radiation.

3.1.3. Prediction of volume change The two groups were pooled together to a total of 18 samples for the prediction of the tumor volume change, i.e. the growth between day 6 and 9 for Panc02 tumors and between day 9 and 12 for BXPC3 tumors. A correlation between the true and predicted values resulted in a correlation score $\rho = 0.513$ (Figure 6b), which compared to the null-distribution corresponded to a p-value of 0.034 (Figure 6a) reaching statistical significance. Again, features of the blue long-lived signal component were clearly of most importance to the prediction, especially the 90th percentile intercept and the 25th percentile slope (Figure 6c).

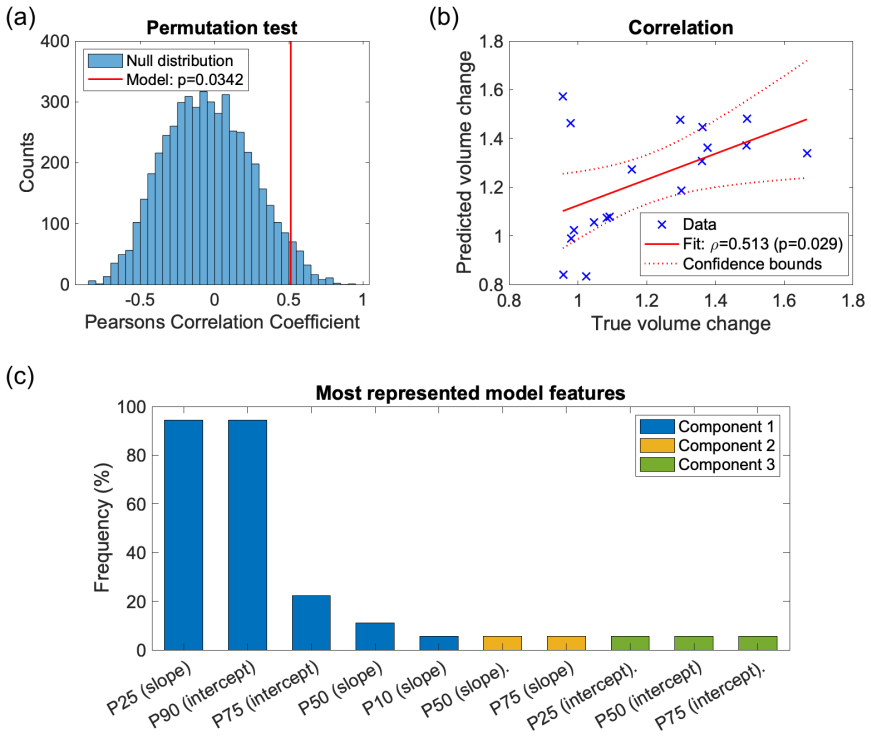

Figure 6. The result of the multiple linear regression for prediction of tumor volume change. (a): The model Pearson's correlation score (red line) compared to the null distribution generated from 5000 runs with permuted labels (P=0.034). (b): True versus predicted volume changes together with a linear fit presented with 95 % confidence bounds. (c): The most represented model features in the 18 CV runs.

3.2. DWI of brain metastases

3.2.1. Decomposition with msNMF Considering the mix of tumor and sub-clinical tissues included in the ROIs, it was expected that at least three signal components were distinguishable. The data was thus inspected and found to contain three to five components, and as Figure 7 reveals, a rank of four was chosen for the final decomposition. The four components were able to explain 99.1 % of the data variation, and their behavior differed from each other both as a function of b-value (Figure 7b) and spatially (Figure 7c). The close-ups in Figure 7c show that for this example, the main tumor region (central part of the magenta ROI) primarily contains the two intermediate decaying signal components (yellow and green) while a different signal composition is seen for the surrounding abnormal tissues. The very rapidly decaying signal component (red) shows no contrast in the presented slice except for a few small spots of high intensity which could well be blood vessels. The intensity may be affected by flow but the attenuation of blood signal with b-value will be strong in any case.

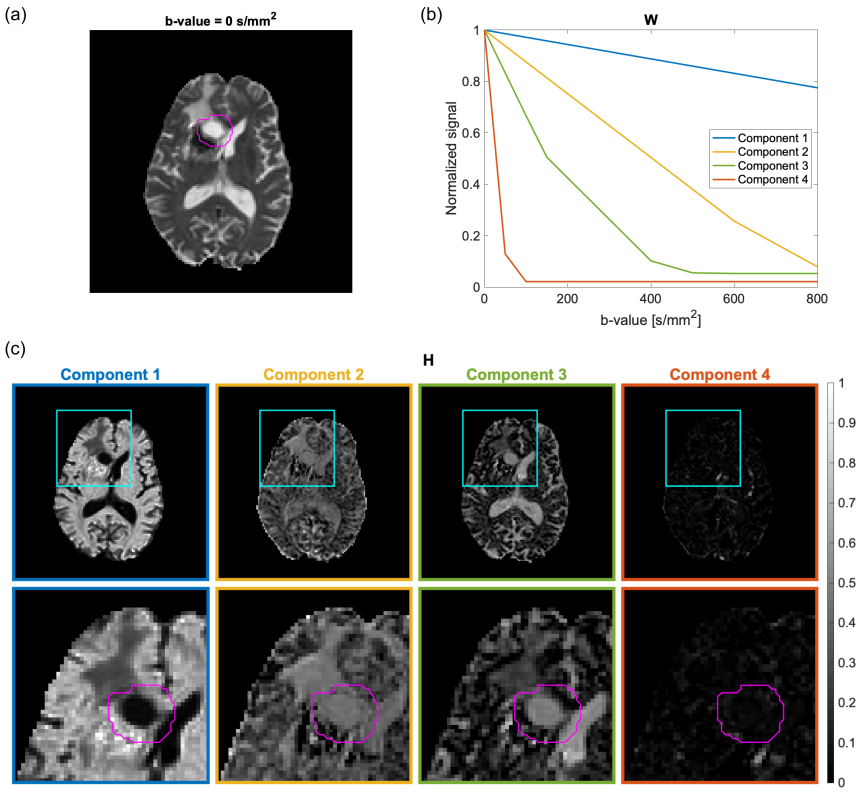

Figure 7. The msNMF components for an example patient with a brain metastasis, at scan day 1. (a): b-value=0 s/mm^2 image from the initial scan session. The magenta ROI is the result of dilating the radiologist's delineation. (b): The signal components (**W**). (c): The associated normalized mixture maps (**H**) indicated by frame colors. The cyan box surrounds the tumor area and marks the region magnified in the bottom row of images.

3.2.2. Prediction of therapy response For the classification of responders versus non-responders 50 prediction models were generated due to the randomization included in the K-fold CV procedure. This resulted in a mean AUC score of 0.74, which compared to the null distribution corresponded to a p-value of 0.065 (Figure 7). The relation between tumor features and therapy outcome was thus not strong enough for the prediction to be significant at an $\alpha = 0.05$ level. The most informative feature was the start median (P50 intercept) of the b=0 s/mm^2 data (Figure 6).

4. Discussion

The main purpose of this study was to present and demonstrate a full analysis framework utilizing longitudinal MRI measurements for prognosis and prediction of cancer treatment outcome. Our work was motivated by the recent release of MR-Linacs, which is expected to increase the amount of longitudinal MRI cancer studies in the coming years. Additionally, the potential problems and pitfalls connected with

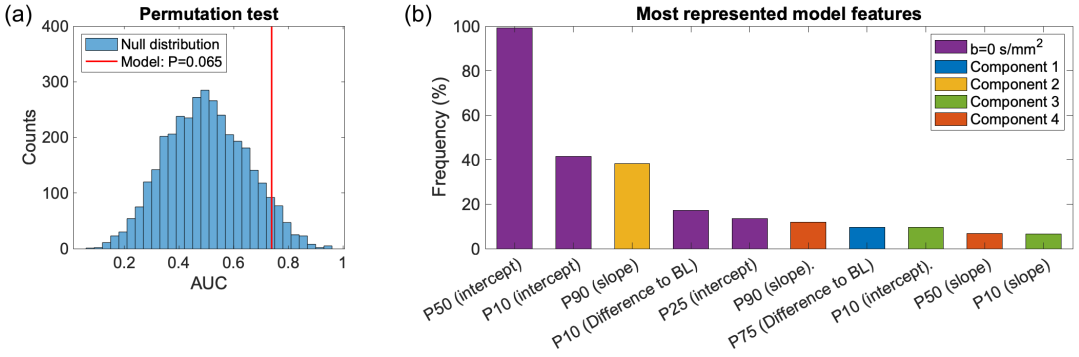

Figure 8. The result of the logistic regression for classification of metastasis into responders/non-responders. (a): The model AUC score (red line) compared to the null distribution generated from 5000 runs with permuted class labels (P=0.065). (b): The most represented model features. The bar color indicates signal/component, and PXX indicates percentile XX. The "slope" and "intercept" are from the temporal linear regression model. The "difference to BL" refers to the difference between the baseline scan and the subsequent scan (day 1).

model-based analysis pointed out by, e.g., Novikov et al [1] and Satta et al [2] motivated a framework that avoids model-based parametrization. To show its flexibility and range of applications, the framework was demonstrated using two datasets that varied both regarding MR contrast, number of contrast measurements, and number of time-resolved scans. Though results were statistically significant only for the T_2-weighted multi-echo dataset, the framework could handle both types of data.

4.1. T_2-weighted multi-echo imaging of mice

The preclinical data published by Tomaszewski et al[3] was ideal to use for framework testing. Firstly, it included a set of well-resolved signals (short echo spacing) measured multiple times during the treatment period. Secondly, the analyses of Tomaszewski et al revealed interesting findings regarding the predictive value of the T_2 inter-quantile range (reflecting tumor heterogeneity), which were supported by histology. Finally, the two pancreatic tumor models allowed to test the framework's ability to do a simpler tumor classification before testing for prediction of outcome.

The three signal components detected by the msNMF analysis clearly presents different relaxation features of the tumor tissue. In the raw data, decay curves contained signal transients from RF inhomogenety in the beginning of the echo train appearing as small oscillations, but the components are unaffected due to the msNMF constraints (Supplementary Figure S1). As discussed in Rahbek et al[27], the decomposition is tailored to give physically meaningful components. In some cases, however, the data may not be sufficiently informative to distinguish the true (underlying) signal components and different solutions may explain the data almost equally well. Nevertheless, the re-

sulting decompositions may well carry predictive value.

The prediction analysis was able to distinguish between the two tumor types with very high accuracy, and the successful stratification largely depended on tissue information involving the long-lived signal component. This agrees with results of the source paper [34] revealing that the BXPC3 tumor has a broader distribution of T_2-values including a longer T_2 relative to the Panc02 tumor, though direct comparisons are difficult since standard T_2 estimates are compromised by partial-volume effects. Satisfactorily, histology data confirmed a lower cellular density and more heterogeneous tissue for the BXPC3 tumor, consistent with the detected prominence of a long-lived fluid signal [38]. Although the convincing histological difference between the two tumor types allowed a robust classification to be expected, the results confirm that relevant tissue information is preserved throughout the comprehensive processing and used in the classification.

The analysis was also able to predict the tumor volume change with a significance of p=0.034 compared to the null distribution, a satisfactory prediction considering the low number of samples for both training and testing as well as the heterogeneity introduced by the use of two distinct tumor models. Again, the long-lived signal component was crucial, both the estimated starting point and the rate of change. One explanation could be that these values relate to the level of cell death and formation of necrosis (supported by the histological analysis), which naturally alters the local T2-weighted relaxation. Specifically, T_2-values decrease with increasing cell density [38]. While necrosis is an expected response to radiation, the level of necrosis before treatment can conversely also affect the response [39, 40]. Tomaszewski et al found necrosis to be reflected in a heterogeneity measure (the T_2 interquartile range) [34]. Our results point at the abundance of the long-lived signal component as a possible indicator of necrosis.

Due to limited data, it was necessary to pool the two tumor groups for the outcome prediction. However, underlying biological differences may make it difficult to find a general correlation between the data and the tumor volume response. For example, cell swelling, the extent of which depends on the tumor type, may occur during the initial phase of the therapy-induced apoptosis [7, 41]. If this or other radiobiological traits differ between Panc02 and BXPC3 tumors, it may explain why features of the immediate radiation response were not relevant to the prediction model. Tomaszewski et al supports this, as they reported that slowdown of tumor growth rate was detected already as early as day 3 for the Pan02 tumors but not until day 6 for the BXPC3 tumors. Nevertheless, the use of two groups in one model makes the solution more generalizable, even if the model performance is not as good. It is highly encouraging that a significant prediction of the tumor volume change was obtained despite of variations in tumor type.

The analysis of the mice relaxometry data has confirmed the potential of our framework and revealed that the signal component with a slow decay (long T_2) was

important in terms of successfully classifying tumor type and predicting tumor volume change. We cannot expect it to be a general finding, but the meaningful biological link to necrosis and high heterogeneity is interesting. Larger datasets are necessary to robustly identify signal trends and relate these to tissue features.

4.2. DWI of brain metastases

The DWI dataset previously presented by Mahmood et al [11] is rather unique as it contains many b-values and consists of scans acquired at each of the ten fractions during the radiotherapy course. These properties made the dataset suited for testing of the proposed framework, although the sample size is too small to expect reliable predictions. Additionally, the work by Mahmood et al [35, 36] showed differences between responders and non-responders with respect to the relative ADC change during therapy and thus indicated that DWI signals carry information useful for early stratification of treatment response, in consistency with other studies [8, 42, 20].

The prediction was close to the threshold of significance (AUC=0.74, p=0.065) indicating a possible relationship between the data and the response labels. The heterogeneity of the data could be one reason for the non-significance. The metastasis originated from six different primary tumor types (listed in Mahmood et al [36]), and excessive variation can make it difficult to establish a general relationship between the signal signature and the treatment outcome with limited data. The driving feature in the prediction originated from the b=0 s/mm^2 data. A prediction test that included only these T_2-weighted data led to a very similar result (AUC=0.74, p=0.052), indicating that local T_2-relaxation was relevant, while diffusion-weighted signal did not provide additional information for the response stratification in this case. It should be emphasized that demonstrating group differences of the ADC as in [35, 36] does not necessarily imply that a general pattern from input data to output label can be learned. Inter- and intra-tumor heterogeneity may dominate the variation in the diffusion-weighted data.

There is a notable limitation connected to the use of numerical bounds to separate and define non-responders from responders as was the case for this data. The RECIST guidelines[43] were followed except for using tumor volume instead of tumor diameter (as in [35, 36]). The dichotomization becomes particularly problematic here because several observations out of the small cohort (N=31) had a relative volume change close to the threshold. We also attempted prediction of the actual volume change using a regression model instead, but this resulted in a poor prediction outcome (analysis not shown). Despite the highlighted data challenges, it was demonstrated that the framework is well-suited for longitudinal study designs and relevant data types, which was a main objective.

A challenge not specific to this data, but more general for studies relying on tumor

features, involves the definition of the tumor ROI, for example which images (DWI, T_2-weighted,...) to use for delineation and which target volume (gross tumor volume, clinical target volume, viable tumor volume, etc.) to choose [44, 45, 46]. Mahmood et al showed how the ROI delineation strategy in conjunction with the signal fitting method affected the estimated ADC changes [11]. This is one of the reasons that our framework avoids model-based parameter estimation and is designed to be less sensitive to ROI variability. The unsupervised component analysis is not confounded by intra-voxel heterogeinity, and the ROI dilation and histogram-based assessment of tumor features reduce the significance of exact delineations. Nevertheless, there is still reason to be critical towards the choice of target and ROI dilation. We utilized the delineations from [35], i.e. high-intensity regions at b=800 s/mm^2 corresponding to "viable" tumor dilated by approximately 5 mm, and obtained reasonable results of the framework. These choices deserve investigation in studies with more data, and optimally an automatic segmentation process can eventually replace the manual delineation.

For the preclinical data example, the mice tumors, which are generated from subcutaneous cell injections into the hind leg, resulted in clearly delimited large tumor regions throughout the course of monitoring, why the ROI strategy was not an issue there.

An important aim of the proposed framework is the utilization of information from all scans simultaneously. Previous longitudinal studies assumed independence between the data points and treated them individually, although being related to the same patient. We argue it is better data handling to include the measurements' dependency, especially as there may be predictive value of the early changes for the later therapy outcome. Also, with the improved availability of MR-Linacs there is less reason to settle for a single scan time point. A robust linear fit has been proposed for this processing step because it is less sensitive to small variations (noise) in the data than a fit of higher order, which also introduces ambiguity of fitted coefficients when linear fitting is sufficient. Linear fitting is also more robust towards missing measurements and which time-points that are part of the longitudinal scan scheme. It is, however, possible to use more flexible modelling in the temporal dynamics analysis, if this can be trusted to generalize across observations without overfitting.

The proposed framework for analyzing longitudinal MRI datasets for tumor disease management was presented and demonstrated in a structured manner (Figure 1), highlighting all processing steps necessary to go from high-dimensional MR data to a final prediction including an unbiased performance evaluation. The cornerstones of this unique framework are the data-driven signal decomposition that avoids model-based parametrization, and the histogram-based tumor heterogeneity assessment, where tracking of individual voxels is unnecessary. Simultaneously, the framework is flexible and allows the user to replace some of the suggested methods, including the choice of data-driven decomposition, ROI definition, and temporal dynamics analysis. The decomposition strategy and its associated assumptions regarding the data structure

must be appropriate for the particular data. The unique constraints of the msNMF made it an ideal choice for the given data examples. The fact that the framework functioned as anticipated and even delivered convincing results for the mice data reinforces the choices made here, and proves msNMF to be a potentially relevant decomposition technique for identifying tumor-specific signal features. The two demonstration examples indicated a value of the sources of tissue contrast to stratify tumors and predict therapy response. These results show the potential of the method for detection of biomarkers sensitive to e.g. treatment-induced changes.

For research studies, it is in principle advantageous to use independently suggested analysis frameworks to avoid confirmation biases, but the analysis should also match the particular data at hand. The diverse data of the example studies demonstrated that the framework is sufficiently versatile, and thus a relevant candidate in many cases. Overall, there are no limitations to the type of input features in the prediction, which can also be a mix of e.g. imaging and biometric features.

5. Conclusion

Our study presented a new prediction framework developed for analysis of MRI data from repeated measurements and showed its potential as a tool to identify possible MRI biomarkers, specifically within the field of cancer treatment. In particular, the framework is potentially suited for the increasing volume of longitudinal MRI data already coming from clinical trials with the recent introduction of MR-Linacs in radiation therapy. Furthermore, the merits of the newly developed msNMF for tumor tissue signal decomposition was demonstrated as part of the assessment. Applications were exemplified by classification of tumor tissue and prediction of therapy outcome using T_2-weighted multi-echo data and DWI data, respectively, demonstrating the wide applicability of the framework.

Acknowledgements

The authors would like to thank Dr. Robert J. Gillies for discussions. The study was funded by the Danish Cancer Society (grant R167-A10637-17-S2).

6. References

[1] Mehrabian H, Detsky J, Soliman H, Sahgal A and Stanisz G J 2019 *Frontiers in Oncology* **9** 1–16 ISSN 2234943X

[2] Dhermain F G, Hau P, Lanfermann H, Jacobs A H and van den Bent M J 2010 *The Lancet Neurology* **9** 906–920 ISSN 14744422 URL http://dx.doi.org/10.1016/S1474-4422(10)70181-2

[3] Dappa E, Elger T, Hasenburg A, Düber C, Battista M J and Hötker A M 2017 *Insights into Imaging* **8** 471–481 ISSN 18694101

[4] McMahon S J and Prise K M 2019 *Cancers* **11** ISSN 20726694

[5] Beigi M, Kazerooni A F, Safari M, Alamolhoda M, Moghdam M S, Moghadam S, SalighehRad H and Ameri A 2018 *Radiologia Medica* **123** 36–43 ISSN 18266983

[6] wei Zhao D, jun Fan W, ling Meng L, rong Luo Y, Wei J, Liu K, Liu G, feng Li J, Zang X, Li M, xin Zhang X and Ma L 2021 *Cancer Imaging* **21** 1–12 ISSN 14707330

[7] Padhani A R, Liu G, Mu-Koh D, Chenevert T L, Thoeny H C, Takahara T, Dzik-Jurasz A, Ross B D, Van Cauteren M, Collins D, Hammoud D A, Rustin G J, Taouli B and Choyke P L 2009 *Neoplasia* **11** 102–125 ISSN 14765586 URL http://linkinghub.elsevier.com/retrieve/pii/S1476558609800249

[8] van Houdt P J, Yang Y and van der Heide U A 2021 *Frontiers in Oncology* **10** 1–9 ISSN 2234943X

[9] Raaymakers B W, Jürgenliemk-Schulz I M, Bol G H, Glitzner M, Kotte A N, Van Asselen B, De Boer J C, Bluemink J J, Hackett S L, Moerland M A, Woodings S J, Wolthaus J W, Van Zijp H M, Philippens M E, Tijssen R, Kok J G, De Groot-Van Breugel E N, Kiekebosch I, Meijers L T, Nomden C N, Sikkes G G, Doornaert P A, Eppinga W S, Kasperts N, Kerkmeijer L G, Tersteeg J H, Brown K J, Pais B, Woodhead P and Lagendijk J J 2017 *Physics in Medicine and Biology* **62** L41–L50 ISSN 13616560

[10] Mutic S, Low D, Chmielewski T, Fought G, Hernandez M, Kawrakow I, Sharma A, Shvartsman S and Dempsey J 2016 *Medical Physics* ISSN 1522-8541

[11] Mahmood F, Johannesen H H, Geertsen P, Opheim G F and Hansen R H 2015 *Acta Oncologica* **54** 1529–1534 ISSN 1651226X

[12] Bostel T, Dreher C, Wollschläger D, Mayer A, König F, Bickelhaupt S, Schlemmer H P, Huber P E, Sterzing F, Bäumer P, Debus J and Nicolay N H 2020 *Radiation Oncology* **15** 1–12 ISSN 1748717X

[13] Zhao B, Cao K, Li X T, Zhu H T and Sun Y S 2019 *BMC cancer* **19** 1115 ISSN 14712407

[14] Hatakenaka M, Nakamura K, Yabuuchi H, Shioyama Y, Matsuo Y, Ohnishi K, Sunami S, Kamitani T, Setoguchi T, Yoshiura T, Nakashima T, Nishikawa K and Honda H 2011 *International Journal of Radiation Oncology Biology Physics* **81** 339–345 ISSN 03603016

[15] Haldorsen I S, Grüner R, Husby J A, Magnussen I J, Werner H M, Salvesen Ø O, Bjørge L, Stefansson I, Akslen L A, Trovik J, Taxt T and Salvesen H B 2013 *European Radiology* **23** 2916–2925 ISSN 09387994

[16] Loncaster J A, Carrington B M, Sykes J R, Jones A P, Todd S M, Cooper R, Buckley D L, Davidson S E, Logue J P, Hunter R D and West C M 2002 *International journal of radiation oncology, biology, physics* **54** 759–767 ISSN 03603016

[17] Novikov D S, Kiselev V G and Jespersen S N 2018 *Magnetic Resonance in Medicine* **79** 3172–3193 ISSN 15222594

[18] Satta S, Dolciami M, Celli V, Di Stadio F, Perniola G, Palaia I, Pernazza A, Della Rocca C, Rizzo S, Catalano C, Capuani S and Manganaro L 2021 *The British Journal of Radiology* **94** 20210054 ISSN 0007-1285

[19] Paudyal R, Oh J H, Riaz N, Venigalla P, Li J, Hatzoglou V, Leeman J, Nunez D A, Lu Y, Deasy J O, Lee N and Shukla-Dave A 2017 *Journal of Magnetic Resonance Imaging* **45** 1013–1023 ISSN 15222586

[20] Foltz W D, Wu A, Chung P, Catton C, Bayley A, Milosevic M, Bristow R, Warde P, Simeonov A, Jaffray D A, Haider M A and Ménard C 2013 *Journal of Magnetic Resonance Imaging* **37**

909–916 ISSN 10531807
[21] Cai G, Xu Y, Zhu J, Gu W L, Zhang S, Ma X J, Cai S J and Zhang Z 2013 *World Journal of Gastroenterology* **19** 5520–5527 ISSN 22192840
[22] Wang L, Liu L, Han C, Liu S, Tian H, Li Z, Ren X, Shi G, Wang Q and Wang G 2016 *Radiotherapy and Oncology* **121** 246–251 ISSN 18790887 URL http://dx.doi.org/10.1016/j.radonc.2016.10.021
[23] Liu Y, Sun H, Bai R and Ye Z 2015 *Radiation Oncology* **10** 1–8 ISSN 1748717X URL http://dx.doi.org/10.1186/s13014-015-0493-6
[24] Hein P A, Kremser C, Judmaier W, Griebel J, Pfeiffer K P, Kreczy A, Hug E B, Lukas P and DeVries A F 2003 *European Journal of Radiology* **45** 214–222 ISSN 0720048X
[25] Wang C, Rimner A, Hu Y C, Tyagi N, Jiang J, Yorke E, Riyahi S, Mageras G, Deasy J O and Zhang P 2019 *Medical Physics* **46** 4699–4707 ISSN 00942405
[26] Nasiri F and Acosta-Tamayo O 2018 *arXiv* 1–14 (*Preprint* 1804.04590) URL http://arxiv.org/abs/1804.04590
[27] Rahbek S, Madsen K H, Lundell H, Mahmood F and Hanson L G 2021 *Journal of Magnetic Resonance* **333** 107103 ISSN 10907807 URL https://doi.org/10.1016/j.jmr.2021.107103
[28] Hötker A M, Lollert A, Mazaheri Y, Müller S, Schenk J P, Mildenberger P C, Akin O, Graf N and Staatz G 2020 *Abdominal Radiology* **45** 3202–3212 ISSN 23660058 URL https://doi.org/10.1007/s00261-020-02475-w
[29] Najafi M, Soltanian-Zadeh H, Jafari-Khouzani K, Scarpace L and Mikkelsen T 2012 *PLoS ONE* **7** 1–11 ISSN 19326203
[30] Alma Ö G 2011 *Int. J. Contemp. Math. Sciences* **6** 409–421 URL http://m-hikari.com/ijcms-2011/9-12-2011/almaIJCMS9-12-2011.pdf
[31] Hosmer D W and Lemeshow S 2000 *Applied Logistic Regression* 2nd ed (John Wiley and Sons Inc.)
[32] Bursac Z, Gauss C H, Williams D K and Hosmer D W 2008 *Source Code for Biology and Medicine* **3** 1–8 ISSN 17510473
[33] Ojala M and Garriga G C 2010 *Journal of Machine Learning Research* **11** 1833–1863 ISSN 15324435
[34] Tomaszewski M R, Dominguez-Viqueira W, Ortiz A, Shi Y, Costello J R, Enderling H, Rosenberg S A and Gillies R J 2020 *NMR in Biomedicine* 1–12 ISSN 10991492
[35] Mahmood F, Johannesen H H, Geertsen P and Hansen R H 2017 *Acta Oncologica* **56** 1651–1653 ISSN 1651226X URL https://doi.org/10.1080/0284186X.2017.1348627
[36] Mahmood F, Johannesen H H, Geertsen P and Hansen R H 2017 *Physics in Medicine and Biology* **62** 2990–3002 ISSN 13616560
[37] Penny W, Friston K, Ashburner J, Kiebel S and Nichols T 2007 *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Elsevier science) ISBN 9780123725608
[38] Saito S, Tsugeno M, Koto D, Mori Y, Yoshioka Y, Nohara S and Murase K 2012 *International Journal of Nanomedicine* **7** 5415–5421 ISSN 11769114
[39] Baskar R, Dai J, Wenlong N, Yeo R and Yeoh K W 2014 *Frontiers in Molecular Biosciences* **1** 1–9 ISSN 2296889X
[40] Eriksson D and Stigbrand T 2010 *Tumor Biology* **31** 363–372 ISSN 10104283
[41] Patterson D M, Padhani A R and Collins D J 2008 *Nature Clinical Practice Oncology* **5** 220–233 ISSN 1743-4262 URL https://doi.org/10.1038/ncponc1073
[42] Sun Y S, Cui Y, Tang L, Qi L P, Wang N, Zhang X Y, Cao K and Zhang X P 2011 *American Journal of Roentgenology* **197** 23–29 ISSN 0361803X
[43] Eisenhauer E A, Therasse P, Bogaerts J, Schwartz L H, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D and Verweij J 2009 *European Journal of Cancer* **45** 228–247 ISSN 09598049 URL http://dx.doi.org/10.1016/j.ejca.2008.10.026
[44] Mahmood F, Hjorth H, Geertsen P and Hvass R 2020 *Radiotherapy and Oncology* **144** 121–126

ISSN 0167-8140 URL https://doi.org/10.1016/j.radonc.2019.11.012

[45] Lambregts D M, Beets G L, Maas M, Curvo-Semedo L, Kessels A G, Thywissen T and Beets-Tan R G 2011 *European Radiology* **21** 2567–2574 ISSN 09387994

[46] Gity M, Moradi B, Arami R, Arabkheradmand A and Kazemi M A 2018 *Asian Pacific Journal of Cancer Prevention* **19** 2765–2770 ISSN 2476762X

Supplementary Materials

For the manuscript: *Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI*

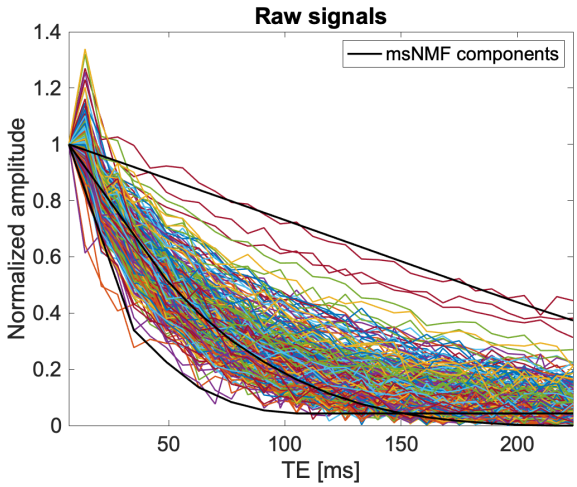

Figure S1: A sample of 1000 raw T_2-weighted multi-echo signals normalized to start at 1, together with the signal components from the msNMF (black curves). For the first few measurements, oscillations beyond the noise level are present. Such initial signal transients are expected due to RF inhomogeneity.

PAPER III

C

TITLE

Optimized flip angle schemes for the diffusion-weighted SPLICE sequence

AUTHORS

Rahbek, Sofie; Schakel, Tim; Mahmood, Faisal; Madsen, Kristoffer H.; Philippens, Marielle E.P.; Hanson, Lars G.

YEAR

2022

PUBLICATION HISTORY

Manuscript under revision for publication as a Technical Note
Submission date: 29 April, 2022
Review decision date: 21 May, 2022

Optimized flip angle schemes for the diffusion-weighted SPLICE sequence

Sofie Rahbek[a], Tim Schakel[b], Faisal Mahmood[c,d], Kristoffer H. Madsen[e,f], Marielle E.P. Philippens[b], Lars G. Hanson[a,e,*]

[a]*Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*
[b]*Department of Radiotherapy, University Medical Center Utrecht, Utrecht, Netherlands*
[c]*Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital, Odense C, 5000, Denmark*
[d]*Department of Clinical Research, University of Southern Denmark, Odense, 5000, Denmark*
[e]*Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, 2650, Denmark*
[f]*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*

Abstract

Purpose The diffusion-weighted SPLICE (split acquisition for fast spin-echo) sequence employs split-echo rapid acquisition with relaxation enhancement (RARE) readout to provide images almost free of geometric distortions. However, due to the varying T_2-weighting during k-space traversal, SPLICE suffers from blurring. This work presents a method for controlling the spatial point spread function (PSF) while optimizing the signal-to-noise ratio (SNR) achieved by optimizing the flip angles in the refocusing pulse train of SPLICE.

Methods An algorithm based on extended phase graph (EPG) simulations optimizes the flip angles by maximizing SNR for a flexibly chosen predefined target PSF that describes the desired k-space density weighting and spatial resolution. An optimized flip angle scheme and a corresponding post-processing correction filter which together achieve the target PSF was tested by healthy subject brain imaging using a clinical 1.5 T scanner.

Results Brain images showed a clear improvement over those obtained with a standard constant flip angle scheme. SNR was increased and ADC estimates were more accurate. The correction filter was relatively flat indicating a high benefit from acquisition weighting by flip angle control.

Conclusion The presented flexible method for optimizing SPLICE flip angle schemes offers improved MR image quality of geometrically accurate diffusion-weighted images that makes the sequence a strong candidate for radiotherapy planning or stereotactic surgery.

[⋆]**Funding information** Danish Cancer Society, Grant Number: R167-A10637-17-S2

[⋆]**Correspondence** Lars G. Hanson, Technical University of Denmark, Ørsteds Plads, Building 349, Kgs. Lyngby, 2800, Denmark.

Email address: lghan@dtu.dk (Lars G. Hanson)

Keywords: diffusion-weighted magnetic resonance imaging, SPLICE, sequence optimization, variable flip angle, SNR, point spread function

1. Introduction

Diffusion-weighted (DW) MRI reflects the micro-anatomy of tissues as it probes the local molecular mobility of water. This is particularly useful for cancer imaging as the high cellular density of tumour tissue restricts the water movement, which results in a high DW signal intensity of tumours compared to healthy tissue. Currently, the most common read-out used in clinical DW-MRI is single shot echo-planar imaging (EPI). Fast k-space traversal following a single excitation pulse ensures a low scan time minimizing the problem of motion artifacts. However, EPI images are prone to geometrical distortions due to the high sensitivity to static field inhomogeneities. Therefore, the benefit of DW-MRI contrast for tumour delineation in radiotherapy planning, for example, is compromised, especially when the region-of-interest is near an air cavity.[1]

Other DW-MRI sequences combine diffusion-encoding with a rapid acquisition with relaxation enhancement (RARE) readout.[2] In this combination, it is important to address potential violations of the Carr Purcell Meiboom Gill (CPMG) conditions that the RARE acquisition is performed under. The CPMG conditions dictate the timing and phase relations of the refocusing pulses so they form coinciding stimulated and spin echoes with the same phase. However, the diffusion-encoding in combination with subject motion essentially randomizes the initial phase of the echoes. This violates CPMG conditions, causing echoes to interfere destructively leading to severe artifacts.

Related to the work of Alsop[3] and Norris *et al*,[4] Fritz Schick introduced the Split Acquisition of Fast Spin Echo Signal for Diffusion Imaging (SPLICE) sequence,[5] which overcomes the CPMG condition violation. In SPLICE, a prolonged and imbalanced read-out gradient is used to split echoes with different phases into two families (Supporting Information Figure S1). Thus, two sets of echoes are collected, reconstructed, and combined as magnitude images. Other RARE-based DW-MRI sequences eliminate or suppress one of the echo families to overcome the CPMG violation, resulting in a lower SNR than for SPLICE.[6] The geometrical robustness of the RARE-based SPLICE makes it a compelling rival to EPI based DW-MRI especially when accurate anatomical delineations are critical.[7] However, the longer signal sampling period causes pronounced blurring due to T_2-decay during k-space traversal.

A strategy to reduce blurring and improve image quality is to modify the signal modulation by acquisition weighting[8] using a variable refocusing flip angle in the echo train. The extended phase graph (EPG) algorithm[9] calculates the magnetization response during a multi-pulse experiment and is useful for determining the resulting k-space weighting for a RARE or SPLICE readout. Earlier work[10,11,12,13] utilized the prospective EPG formulation

introduced by Hennig *et al*[14] to address the inverse problem of finding the flip angle scheme that generates a desired relaxation time contrast, and further used static pseudo steady states (PSS) to control the signal decay. This strategy resulted in improvements, e.g., low RF power deposition and reasonable PSFs for a given contrast, but it did not optimize the SNR directly for a given PSF. Therefore, we employ a method for optimizing the flip angle scheme with a main focus of improving the SNR for a chosen spatial resolution in the phase-encoding direction, expressed via the PSF (controlled blurring).[15] For diffusion-weighted sequences, a main focus is SNR. Hence, there is limited need to control the relaxation time contrast, which is a main difference from the earlier work referred to. Additionally, the method avoids the analytical solution of Hennig *et al*,[14] which may have an imaginary outcome and need iterative regulation of a set of flip angle control points. Instead, the method provides an optimized refocusing scheme and a corresponding filter for post-processing to compensate for remaining differences between the final signal modulation and the desired k-space weighting. The strategy also facilitates comparison between refocusing schemes since partial volume effects are inherently made similar. We validate and demonstrate the value of the method with simulations, phantom scans and brain scanning of a healthy volunteer. A preliminary report of the method was reported in.[16] The abstract included a single imaging example with linear phase ordering as well as an example with a center out trajectory which is not discussed in the current paper.

2. Methods

2.1. Optimization algorithm

A normalized target distribution, $T(k)$, is defined to describe the k-space density weighting in the phase encoding-direction corresponding to the desired spatial PSF through the Fourier transform. The sinc-shaped PSF for normal uniform weighting (constant $T(k)$) of central k-space, is not desirable due to pronounced spatial side-bands, for example. $T(k)$ is freely selectable and given as input to the optimization algorithm together with tissue relaxation parameters (T_2, T_1) and sequence specifications (echo spacing (ESP), echo train length (ETL)). The events of a SPLICE readout are simulated using EPG calculations, from which the final echo response, $I(k)$, is given by the signal magnitude sum of the two echo families (E1 and E2) for a spatial point source. Hence, $I(k)$ represents the k-space weighting introduced by the flip-angle scheme, and it depends on both acquisition and tissue parameters. A filter, $F(k)$, is calculated to compensate for the differences between $I(k)$ and $T(k)$ to ensure the desired PSF in the resulting images (Eq. 1). For normal linear image reconstruction and fixed PSF, the SNR only depends on the flip angle scheme via the filter's effect on noise (Eq. 2).

$$T(k) = I(k) \cdot F(k) \tag{1}$$

$$\mathrm{SNR} \propto \frac{1}{\sqrt{\sum_k \left(\frac{T(k)}{I(k)}\right)^2}} \tag{2}$$

The algorithm maximizes this SNR by updating the flip angle scheme using a nonlinear programming solver utilizing an interior-point method (MATLAB 2018b, MathWorks, Inc., Natick, Massachusetts, United States). The output is the optimized set of flip angles together with the final filter, preferably relatively flat as the echo response then already matches the desired signal weighting.[8] Strong filters are undesirable due to decreased SNR and the user must therefore choose a reasonable PSF (voxel size and shape as exemplified below) considering the application at hand. The algorithm makes it possible to compare choices theoretically, and it provides an optimal refocusing scheme for each. The filter ensures that the target PSF is met for the chosen tissue parameters, which enables direct SNR comparisons. Software with examples is available at `https://github.com/sofierahbek/flip-angle-optimization`.

A standard EPG framework is used, i.e. effects of finite-duration RF pulses are neglected.[9] We therefore investigated whether off-resonance effects across the slice-profile are negligible for a SPLICE sequence with a long echo train and a realistic RF pulse shape, using the open-source JEMRIS MRI simulator.[17] Additionally, JEMRIS was used to examine the build-up of signal that is not diffusion-encoded, a possible consequence of unwanted coherence pathways generated by the repeated refocusing pulses. Finally, the k-space signal weightings obtained with JEMRIS simulations, EPG simulations and phantom scans were compared to verify that the simulations are realistic and thus useful for optimization.

2.2. The choice of the target function

Inspired by Pohmann *et al*,[18] an example target function (T(k)) was chosen as a Hanning function defined as in Pohmann *et al*: $w(k_n) \propto 1 + \cos(\frac{2\pi \cdot k_n \cdot \Delta y}{\alpha})$, with Δy as the nominal spatial resolution and the constant α set to 1.5. The resulting PSF has both a small full width at half maximum (FWHM) and suppressed sidebands. Figure 1 illustrates this by comparing PSFs corresponding to Hanning, rectangular (Rect), and Gaussian functions, respectively. Additionally, the necessary compensation filters reveal that a Hanning target function also results in a relatively flat filter (yellow) indicating that the available magnetization is used efficiently.

2.3. Data acquisition and reconstruction

Healthy subject brain data was recorded using a 1.5 T MRI system (Ingenia, Philips Healthcare, Best, The Netherlands) and a multi-slice single-shot SPLICE sequence modified

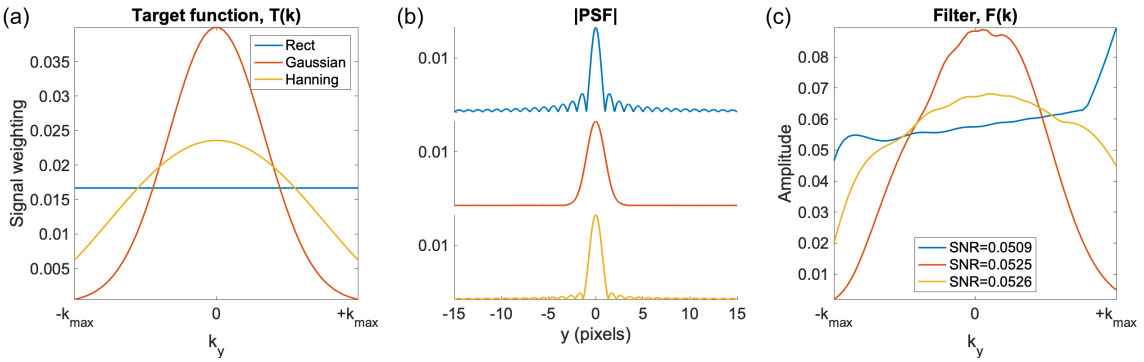

Figure 1: (a) Three different designs for the target function, all normalized so $\sum_k T(k) = 1$. (b) The corresponding PSF by Fourier transformation. (c) The filters needed to obtain the target functions after optimizing the FAs, for a linear k-space sampling order. The specified SNRs are calculated by Eq. 2.

to apply user-defined flip angles. A healthy subject volunteered for scanning, after informed consent. For all SPLICE scans, ESP=4.9 ms, in-plane nominal resolution 1.98×1.98 mm^2, slice thickness 5 mm, slice gap 5 mm, 29 slices, b-values [0, 800] s/mm^2, diffusion time 40 ms, and 3 discarded startup echoes. Measurements were conducted with a scheme of flip angles rapidly converging to 90 degrees [145, 90, 90,... 90] (default setting by the scanner vendor) and optimized schemes of variable flip angles.[19] For each case, a dataset was recorded with a fully sampled k-space, with a parallel imaging (PI) sampling scheme, and with a PI and partial Fourier (PF) imaging sampling scheme, respectively. A linear k-space sampling order, a PI acceleration factor of 2 and a PF factor of 0.6 was used. Table 1 shows timing parameters for each case. Raw k-space data was extracted from the scanner to be able to apply the filters calculated in the optimization procedure. The data was then reconstructed using MATLAB 2018b with help from the Berkeley Advanced Reconstruction Toolbox (BART)[20] and a Projection onto Convex Sets (POCS) algorithm.[21]

Table 1: SPLICE sequence parameters

Sampling	ETL	TE [ms]	TR [ms]	scan time [s]
Full	110	328	10525	295
PI	55	193	7892	221
PI + PF	34	86	6762	189

A multi-shot RARE sequence was included in the protocol to generate high-resolution reference images used for tissue segmentation. A white-matter (WM) mask was created using the Statistical Parametric Mapping toolbox, (SPM12, version 7487)[22] and used for calculation of ADC values.

The phantom experiments for validating consistency between simulations and real data are described in Supporting Information.

3. Results

The optimized flip angle scheme for the fully sampled data is presented together with the compensation filter in Figure 2. The filter is relatively flat compared to the filter needed for the reference flip angle scheme of repeated 90° pulses, shown for comparison. For the last eight echoes, the flip angle is 180°, indicating that a depletion of the longitudinal magnetization (so-called z-storage) is effectively reached.

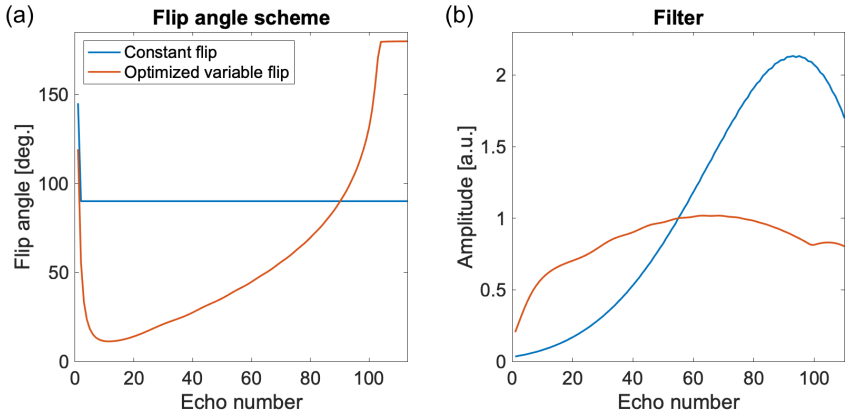

Figure 2: (a) Two flip angle schemes, the optimized (red) and a constant (blue). (b) The corresponding compensation filters are normalized to have value 1 in the k-space center.

In Figure 3 the SPLICE b=0 s/mm^2 images scaled to have equal background noise level, are presented for a comparison between the two flip angle schemes. The introduction of optimized flip angles has changed the signal intensity and contrast of the images, especially for the fully sampled data. The b=800 s/mm^2 data are presented in Supporting Information. The ratio images in the third column, showing the voxel-to-voxel relative signal change between the two datasets, make it clear that the SNR of brain matter has been improved using the optimized variable flip angles. The mean signal improvement across the WM segment is a factor of 2.00, 1.24, and 1.06 for the fully sampled, PI, and PI with PF data, respectively. Thus, the SNR gain increases for longer ETLs. Contrary to the brain matter, the signal of the CSF is slightly reduced.

ADC maps in Figure 4 show a generally higher ADC value in brain matter for the data recorded with the optimized variable flip angle scheme. For the WM segment, an average ADC value was estimated to 0.35×10^{-3} mm^2/s and 0.72×10^{-3} mm^2/s for the constant and variable flip angle data, respectively. The latter is comparable to reported ADC values.[23,24]

The results of the slice-profile investigation using the JEMRIS simulator are presented in Supporting Information. Initial transients occurred for the first three echoes, but otherwise the slice-profile appeared relatively stable and without severe imperfections (Supporting Information Figure S3). The build-up of steady-state signal was very low ($< 2.1\%$) during the

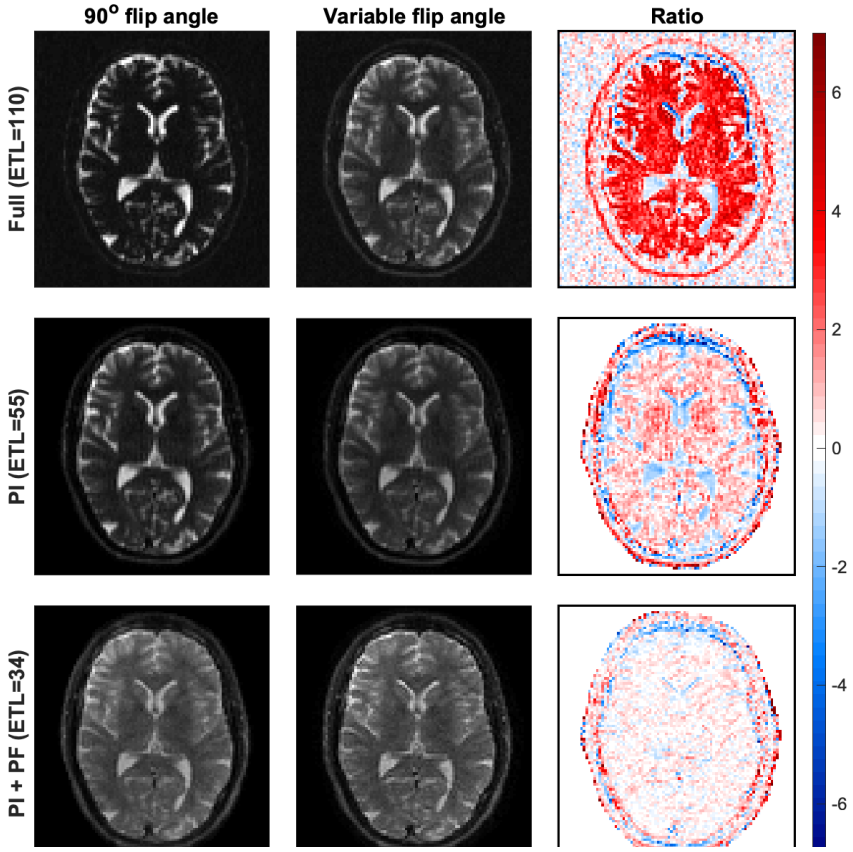

Figure 3: An example slice for the SPLICE b=0 s/mm^2 data scaled with the background noise level. The constant flip angle data, the variable flip angle data, and the ratio in dB between those are presented in the left, center, and right column, respectively. Red colors (positive values) represent a signal increase (SNR gain) for the variable flip angle data, and the opposite accounts for the blue colors (negative values). The fully sampled data is presented in the top row, and the undersampled data in the two bottom rows. Coil sensitivities were estimated only for areas within the subject, so the background is removed for undersampled data.

entire echo train compared to the desired attenuated diffusion-weighted signal (Supporting Information Figure S4). Additionally, Figure S4 shows that only small oscillations appeared in each echo family although a relatively large flip angle was used in the simulation (120°). This warrants that E1 and E2 are not considered individually, but only summed in the optimization algorithm. Finally, the expected consistency between simulations and MRI phantom scans was found (Supporting Information Figure S5).

4. Discussion

The suggested optimized variable flip angle scheme resembles designs presented in earlier work where signal stabilization during long echo trains was explicitly established.[12,25,26] This indicates that our solutions, which are optimal with respect to SNR for the given PSF, may

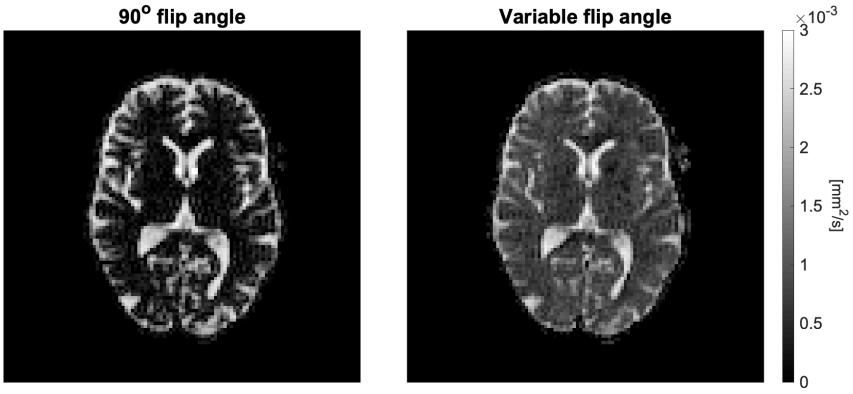

Figure 4: ADC maps for the fully sampled SPLICE data, same slice as visualized in Figure 3. The left and right map is for the constant and variable flip angle data, respectively.

also be near-optimal considering other image traits. Our method introduces flexibility in terms of choosing the target PSF, and though a filter may be needed to realize the chosen voxel shape, consideration of flip angle control points are not needed. Using the methods in,[25,11] for example, three angles must be chosen.

The target function suggested by Pohman et al, resulted in a relatively flat compensation filter which indicates a high benefit of the acquisition weighting obtained by varying the refocusing flip angle. As the filter is only fully effective for tissue with the parameters used in the optimization algorithm, strong filters may create substantial blurring or artifacts for other tissues. Relatively flat filters will increase the robustness towards tissue differences or transmit field inhomogeneities, though the latter are limited for volume coils. Figure S6 in Supporting Information documents the robustness by presenting the PSF for different tissues and a 10 % flip angle error. The PSF is relatively unchanged for all tissues using the optimized flip angle scheme, whereas the sub-optimal 90° flip angle scheme results in a PSF for CSF with small "ghost" peaks and a large FWHM compared to the other tissues.

The validation test presented in Supporting Information justified the use of simulations in the optimization algorithm. Even though JEMRIS simulations showed a better resemblance to actual MRI measurements, EPG simulations are computationally efficient, easy to implement and still provide reasonable estimates of the magnetization response. This is consistent with Weigel *et al*[27] who compared EPG simulations with clinical measurements and found less than 2 % difference. Full simulations (e.g. JEMRIS) or extended versions of EPG, which takes into account the slice profile,[28] are options, but our results do not show a need for more accuracy.

The SNR improvement seen in Figure 3 confirms that an optimized design of the flip

angle scheme is effective in practice, and further that the gain is increased for longer echo trains. Importantly, the use of optimized flip angle schemes involves no significant trade-offs. The lower gain for sequences with a relatively short ETL is expected, as the echo time is shorter and the undesired signal modulation caused by the T_2-weighted signal decay already is reduced by lowering the number of echoes needed for k-space traversal. Using an optimized flip angle design leaves room for increasing the ETL to e.g. obtain a higher spatial resolution or avoid undersampling strategies without a large SNR trade-off.

While earlier work optimized the flip angle scheme primarily for standard RARE sequences, we have implemented the optimized scheme for the diffusion-weighted SPLICE sequence and demonstrated its significance for ADC estimation (Figure 4). The SNR improvement has increased the accuracy of the ADC estimate over the reference scan due to the signal bias from low SNR for high b-values when constant flip angles are used. This strongly argues for adopting optimized flip angle designs when using RARE-based DWI sequences. A related positive effect of the variable flip angle design is a reduction of T_2 shine-through by the inherent shortening of the contrast-equivalent TE.[26] This effect was evident for the fully sampled b=800 s/mm^2 data in Supporting Information as the image contrast of the sub-optimal 90° flip angle sequence indicates a strong T_2 contrast compared to the optimized sequence. A major motivation for focusing on SPLICE is its potential for DWI when geometric distortion is particularly problematic such as for radiotherapy planning, where PSF sidebands in addition compromises quantitative analysis. To that end, the results presented in Figures 3 and 4 are promising. Despite the SPLICE focus, the presented strategy also applies to other single-shot multi-echo sequences. However, the focus on optimizing SNR without controlling relaxation time contrast is not ideal for most RARE applications. DWI is an example where the method can be very useful.

The method was demonstrated using a single example subject, but the findings are general since the optimization is independent from the diffusion-weighting, and since the motion-sensitivity of single-shot sequences is limited. The optimization was performed under an assumption of linear reconstruction, but non-linear methods such as compressed sensing are similarly expected to benefit from sampling schemes guided by the desired spatial resolution. The method is straight-forward to implement (code is provided) and directly applicable on scanners where user-defined flip angle schemes can be introduced.

5. Conclusion

Optimizing the refocusing flip angle scheme of the single-shot SPLICE sequence improves image quality, in particular with respect to SNR and spatial specificity. This directly benefits

DWI, but also reduces T_2 shine-through, ADC bias resulting from low SNR, and partial volume effects. Together with the improved geometric accuracy of SPLICE relative to EPI, this makes the diffusion-weighted SPLICE attractive for radiotherapy planning, for example. The algorithms are readily available, and the sequence and reconstruction changes needed are only flip angle adaptions and filtering.

Acknowledgements

The work was funded by the Danish Cancer Society (grant R167-A10637-17-S2).

References

[1] Schakel Tim, Hoogduin Johannes M., Terhaard Chris H.J., Philippens Marielle E.P.. Technical Note: Diffusion-weighted MRI with minimal distortion in head-and-neck radiotherapy using a turbo spin echo acquisition method: Diffusion-weighted *Medical Physics.* 2017;44:4188–4193.

[2] Hennig J., Nauerth A., Friedburg H.. RARE imaging: A fast imaging method for clinical MR *Magnetic Resonance in Medicine.* 1986;3:823–833.

[3] Alsop David C.. Phase insensitive preparation of single-shot RARE: Application to diffusion imaging in humans *Magnetic Resonance in Medicine.* 1997;38:527–533.

[4] Norris David G., Börnert Peter, Reese Torsten, Leibfritz Dieter. On the application of ultra-fast rare experiments *Magnetic Resonance in Medicine.* 1992;27:142–164.

[5] Schick Fritz. SPLICE: Sub-second diffusion-sensitive MR imaging using a modified fast spin-echo acquisition mode *Magnetic Resonance in Medicine.* 1997;38:638–644.

[6] Dietrich Olaf, Raya José G., Sommer Julia, Deimling Michael, Reiser Maximilian F., Baur-Melnyk Andrea. A comparative evaluation of a RARE-based single-shot pulse sequence for diffusion-weighted MRI of musculoskeletal soft-tissue tumors *European Radiology.* 2005;15:772–783.

[7] Schakel Tim, Peltenburg Boris, Dankbaar Jan Willem, et al. Evaluation of diffusion weighted imaging for tumor delineation in head-and-neck radiotherapy by comparison with automatically segmented 18F-fluorodeoxyglucose positron emission tomography *Physics and Imaging in Radiation Oncology.* 2018;5:13–18.

[8] Zeller Mario, Gutberlet Marcel, Stäb Daniel, et al. Density weighted turbo spin echo imaging *Journal of Magnetic Resonance Imaging.* 2013;37:965–973.

[9] Weigel Matthias. Extended phase graphs: Dephasing, RF pulses, and echoes - Pure and simple *Journal of Magnetic Resonance Imaging.* 2015;41:266–295.

[10] Busse Reed F.. Reduced RF Power Without Blurring: Correcting for Modulation of Refocusing Flip Angle in FSE Sequences *Magnetic Resonance in Medicine.* 2004;51:1031–1037.

[11] Loening Andreas M., Saranathan Manojkumar, Ruangwattanapaisarn Nichanan, Litwiller Daniel V., Shimakawa Ann, Vasanawala Shreyas S.. Increased speed and image quality in single-shot fast spin echo imaging via variable refocusing flip angles *Journal of Magnetic Resonance Imaging.* 2015;42:1747–1758.

[12] Busse Reed F., Hariharan Hari, Vu Anthony, Brittain Jean H.. Fast spin echo sequences with very long echo trains: Design of variable refocusing flip angle schedules and generation of clinical T2 contrast *Magnetic Resonance in Medicine.* 2006;55:1030–1037.

[13] Weigel Matthias, Hennig Jürgen. Diffusion sensitivity of turbo spin echo sequences *Magnetic Resonance in Medicine.* 2012;67:1528–1537.

[14] Hennig Juergen, Weigel Matthias, Scheffler Klaus. Calculation of Flip Angles for Echo Trains with Predefined Amplitudes with the Extended Phase Graph (EPG)-Algorithm: Principles and Applications to Hyperecho and TRAPS Sequences *Magnetic Resonance in Medicine.* 2004;51:68–80.

[15] Zhao Li, Chang Ching Di, Alsop David C.. Controlling T2 blurring in 3D RARE arterial spin labeling acquisition through optimal combination of variable flip angles and k-space filtering *Magnetic Resonance in Medicine.* 2018;80:1391–1401.

[16] Rahbek S., Schakel T., Mahmood F., Madsen K. H., Philippens M. E. P., Hanson L. G.. Flip-angle optimization for the diffusion-weighted SPLICE sequence for applications in brain imaging in *Abstract from 2021 ISMRM & SMRT Annual Meeting & Exhibition*:Abstract 1338, 2021.

[17] Stöcker Tony, Vahedipour Kaveh, Pflugfelder Daniel, Shah N. Jon. High-performance computing MRI simulations *Magnetic Resonance in Medicine.* 2010;64:186–193.

[18] Pohmann Rolf, Von Kienlin Markus. Accurate phosphorus metabolite images of the human heart by 3D acquisition-weighted CSI *Magnetic Resonance in Medicine.* 2001;45:817–826.

[19] Alsop David C. The Sensitivity of RARE low flip angle *Magnetic Resonance in Medicine.* 1997;37:176–184.

[20] Uecker Martin, Tamir Jonathan I, Ong Frank, Lustig Michael. The BART Toolbox for Computational Magnetic Resonance Imaging *Proceedings of the 23rd Annual Meeting of ISMRM, Toronto, Canada.* 2016:2486.

[21] Völker Michael. MRI Partial Fourier reconstruction with POCS (https://www.mathworks.com/matlabcentral/fileexchange/39350-mri-partial-fourier-reconstruction-with-pocs), MATLAB Central File Exchange. 2012.

[22] Penny William, Friston Karl, Ashburner John, Kiebel Stefan, Nichols Thomas. *Statistical Parametric Mapping: The Analysis of Functional Brain Images.* Elsevier science 2007.

[23] Oh Joonmi, Cha Soonmee, Aiken Ashley H., et al. Quantitative apparent diffusion coefficients and T2 relaxation times in characterizing contrast enhancing brain tumors and regions of peritumoral edema *Journal of Magnetic Resonance Imaging.* 2005;21:701–708.

[24] Sener R. N.. Diffusion MRI: Apparent diffusion coefficient (ADC) values in the normal brain and a classification of brain disorders based on ADC values *Computerized Medical Imaging and Graphics.* 2001;25:299–326.

[25] Busse Reed F., Brau Anja C.S., Vu Anthony, et al. Effects of refocusing flip angle modulation and view ordering in 3D fast spin echo *Magnetic Resonance in Medicine.* 2008;60:640–649.

[26] Mugler John P.. Optimized three-dimensional fast-spin-echo MRI *Journal of Magnetic Resonance Imaging.* 2014;39:745–767.

[27] Weigel Matthias, Hennig Juergen. Contrast behavior and relaxation effects of conventional and hyperecho-turbo spin echo sequences at 1.5 and 3 T1 *Magnetic Resonance in Medicine.* 2006;55:826–835.

[28] Guenthner Christian, Amthor Thomas, Doneva Mariya, Kozerke Sebastian. A unifying view on extended phase graphs and Bloch simulations for quantitative MRI *Scientific Reports.* 2021;11:1–19.

Supporting Information

SPLICE sequence diagram

Figure S1 illustrates the SPLICE sequence used in this study. It was originally published with a stimulated echo diffusion preparation in [1].

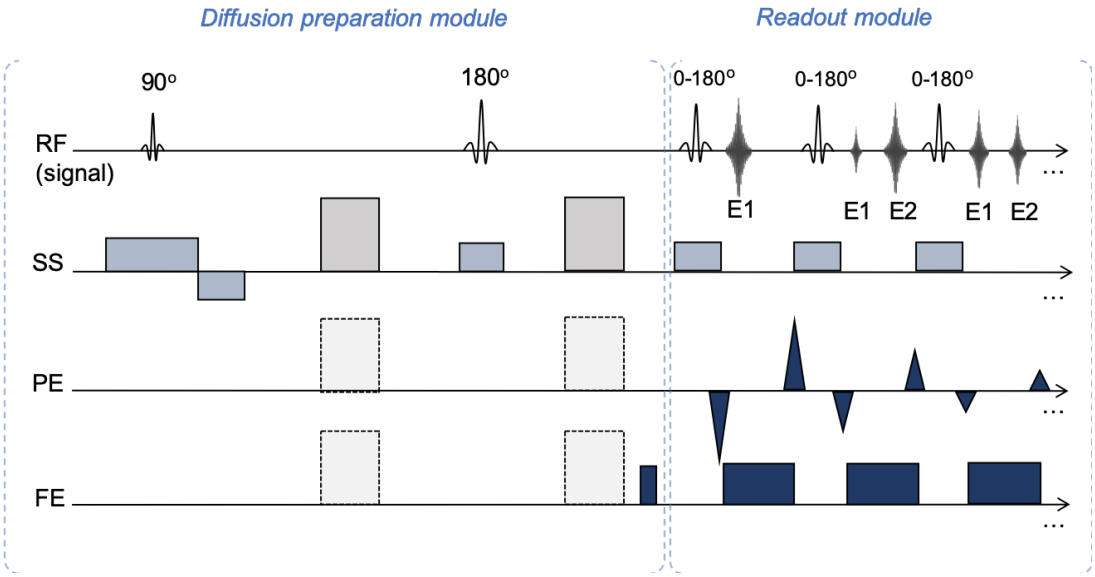

Figure S1: Schematic presentation of the SPLICE sequence, where SS, PE and FE are the slice-selection, phase-encoding and frequency-encoding directions, respectively. The gray blocks are diffusion-encoding gradient pairs (which can be applied in any direction), the light blue blocks are slice-selective gradients, and the dark blue blocks are imaging gradients. The gradients in the FE direction are imbalanced, as the pre-phaser (gradient prior to the read-out module) is much less than half the area of the following read-out gradients. The two families of echoes appearing (E1 and E2) are illustrated on the top line together with the RF pulses. The flip angle is indicated for each pulse. The split-echo mode implies that no phase relation for the refocusing pulses is required [1]. For simplicity, crusher gradients are not shown.

Diffusion-weighted images

Figure S2 presents an example diffusion-weighted SPLICE image for each of the acquisitions. The sum-of-squares combination of coil images results in a non-central chi signal distribution, and hence a bias for low SNR data.

Slice-profile investigation

The slice-profile, which depends on the refocusing pulse design (duration, shape, sidelobes, etc.) have imperfections that may accumulate throughout the RARE readout. Severe imperfections implying significant phase and flip angle variations across a slice could make the EPG calculations an inaccurate representation of the actual signal. The JEMRIS simulator has been used to simulate the standard SPLICE sequence as implemented by the scanner vendor, though leaving out the diffusion-weighting module. Relevant simulation settings are:

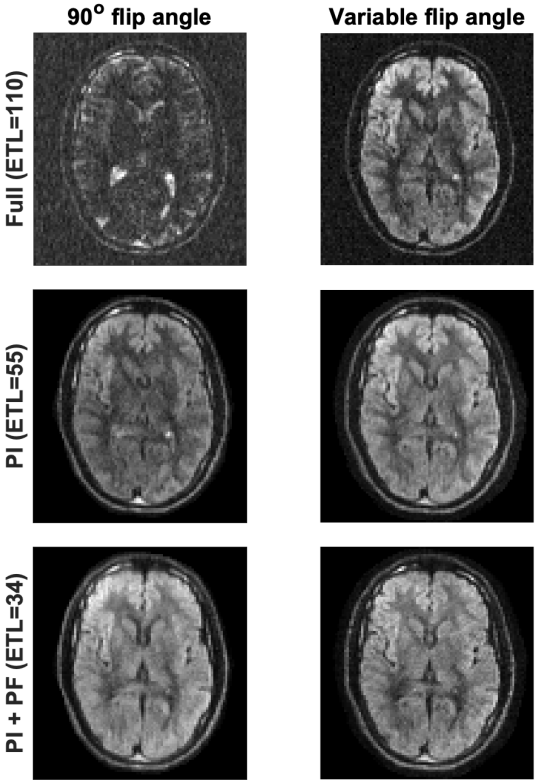

Figure S2: A single slice for the SPLICE b=800 s/mm^2 data. Images are presented with individual, arbitrary intensity scales. The fully sampled data is presented in the top row, and the undersampled data in the two bottom rows. Coil sensitivities were estimated only for areas within the subject, so the background is removed for undersampled data.

T_2=100 ms, T_1=1000 ms, ESP = 5.2 ms, ETL = 70, flip angles: [156, 127, 120, 120,... 120]°. Having two echo families, as is the case with the SPLICE sequence, complicates the slice profile investigation. Since absolute images from each echo family are combined in the reconstruction, a normal slice profile is not meaningful, except for the two echo families separately. When separated, the profiles provide little insight into the combined spatial variation of sensitivity, however. Hence, an effective slice profile is here defined to be the position-dependent sum of the two absolute transversal magnetizations, i.e. a measure of the sensitivity to magnetization in each position. The slice profiles for the individual echoes differ, as can be expected since they represent different coherence pathways, and hence tend to vary oppositely as a function of position across the slice: In one position along the slice selection direction, the stimulated echoes may contribute most signal, whereas the spin echoes may give most signal in another position (due to the flip angle variation along that direction). Conveniently, the sum sensitivities, and therefore the effective slice profile, tends to be smooth as seen in Figure S3. It is also seen to be well-behaved over echoes in the sense that the slice width remains relatively constant during the echo train, except for the initial echoes (indicated by red dots).

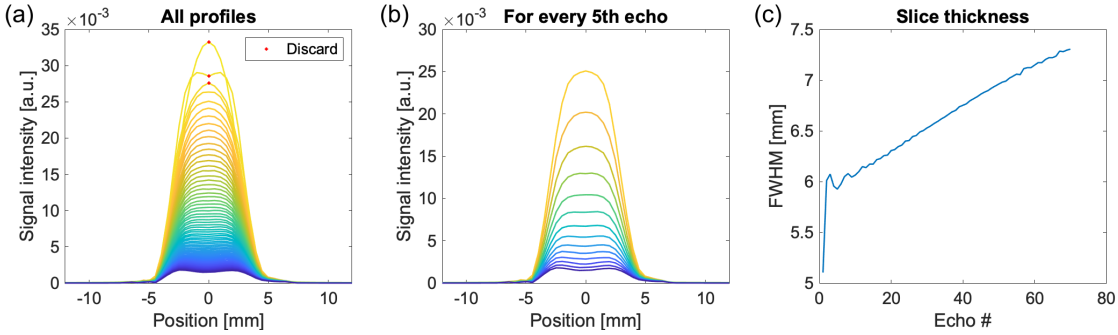

Figure S3: (a) All 70 slice-profiles and (b) every fifth of these. The slice thickness calculated as FWHM for all echoes (c). The first three echoes, denoted by red dots, are recommended discarded due to the initial instability.

This investigation was also performed for a flip angle scheme with more variation (not shown). Results were similar to the presented example to the extent that conclusions were the same. The JEMRIS simulations thus indicate that EPG calculations can be used as a good approximation of the signal in an flip angle optimization when the first three echoes are discarded.

Steady state signal investigation

A build-up of non-diffusion-weighted steady state signal may occur as a consequence of unwanted coherence signal pathways. This problem was investigated by comparing the signal from a JEMRIS simulation *with* a 90° excitation pulse with one *without* excitation ("no excitation"). Diffusion-weighting of the excited magnetization was simulated by multiplying the JEMRIS output with a weighting factor ($e^{-b \cdot \mathrm{ADC}}$). Figure S4 shows a thorough comparison, presenting both the slice-profiles for the two cases as well as signal integrated over the slice profile (echo signal). Throughout the echo train, the steady state signal is low relative to the diffusion-weighted SPLICE signal, corresponding to approximately 2.1 % after 110 echoes.

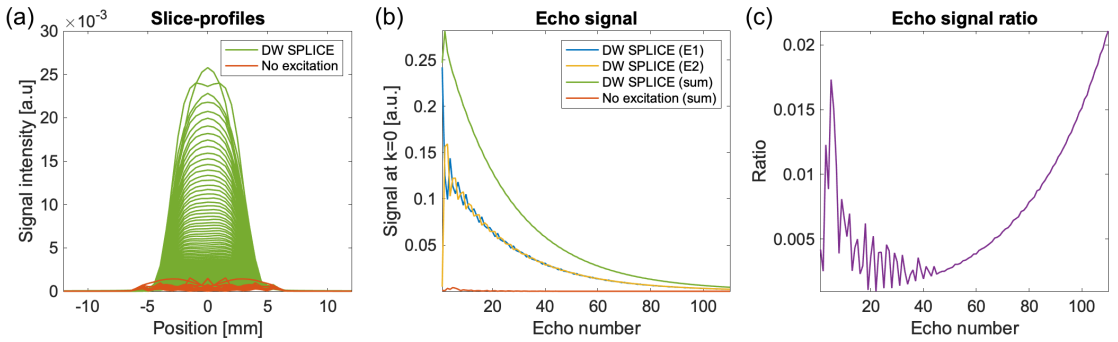

Figure S4: Comparisons of the DW SPLICE signal with the no-excitation signal. (a): Slice-profiles. (b): Echo signals. (c): The ratio between the echo signals, i.e. the red and green curve in (b). A b-value of 500 s/mm^2 and ADC value of 800×10^{-6} mm^2/s were used for the DW factor.

Figure S4 (b) also shows the echo amplitude for the two echo families (E1, E2) separately. Only very small oscillations are present even though a relatively large flip angle (120°) is used throughout the echo train. Theoretically, a repeated large flip angle will shift a large portion of the signal between the two echo families for each echo resulting in oscillations in the individual echo families. However, the imperfections of the slice-profile reduce this behaviour and the oscillations are dampened. A JEMRIS simulation has confirmed this effect why oscillations within each echo family have been ignored in the flip angle optimization.

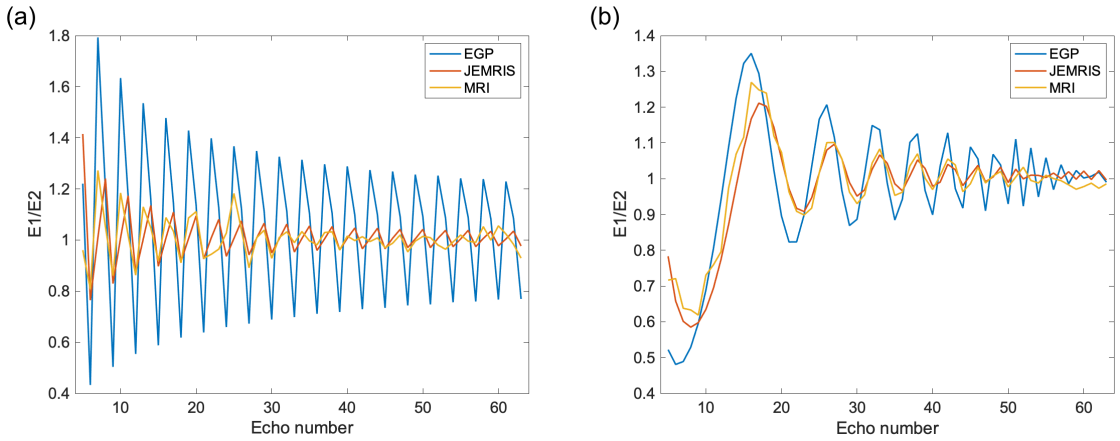

Figure S5: Comparison of the EPG simulation, the JEMRIS simulation, and the MRI phantom scan for a constant flip angle scheme (a) and a variable flip angle scheme (b). The curves represents the ratio between the raw k-space signals from the two echo families, E1 and E2. For the EPG and JEMRIS simulations, the k=0 signal is used, for the MRI phantom data, the square root of the signal power over the frequency encoding direction is used.

Validation of simulations

To validate that simulations are in agreement with actual MRI scans, the k-space weighting along the phase-encoding direction was compared for EPG simulations, JEMRIS simulations and MRI of a spherical phantom with a homogeneous center (PIQT phantom (Philips Healthcare, Best, The Netherlands), T_1=350 ms, T_2=320 ms) for two different flip angle schemes. For the SPLICE scan of the phantom, ESP=5.2 ms, in-plane nominal resolution 3.95x3.95 mm, slice thickness 5 mm, splice gap 5 mm, 20 slices, ETL=63, TE=226 ms, TR=4030 ms and scan time=64.5 s. Only b=0 s/mm^2 data was used, and from a single central slice. To eliminate the effect of the object, the ratio between the two echo families (E1/E2) was considered. The resulting oscillating pattern, stemming from a signal shift between echo families, was relatively consistent between simulations and phantom scans for both of the tested flip angle schemes (Figure S5). This consistency supports that simulations can be trusted in the optimization.

Robustness

The PSF for other tissues than the target is not controlled via the optimization method and may be affected by the compensation filter. Figure S6 shows that the FWHM and shape of the PSF remained relatively unchanged when using the variable flip angle scheme,

while a clear increase of the FWHM for the PSF of CSF is seen for the reference scheme of repeated 90° pulses. Additionally, two small "ghost" peaks appeared. These results remained effectively unchanged after introducing a flip angle error of 10 % and was thus robust towards realistic RF inhomogeneities.

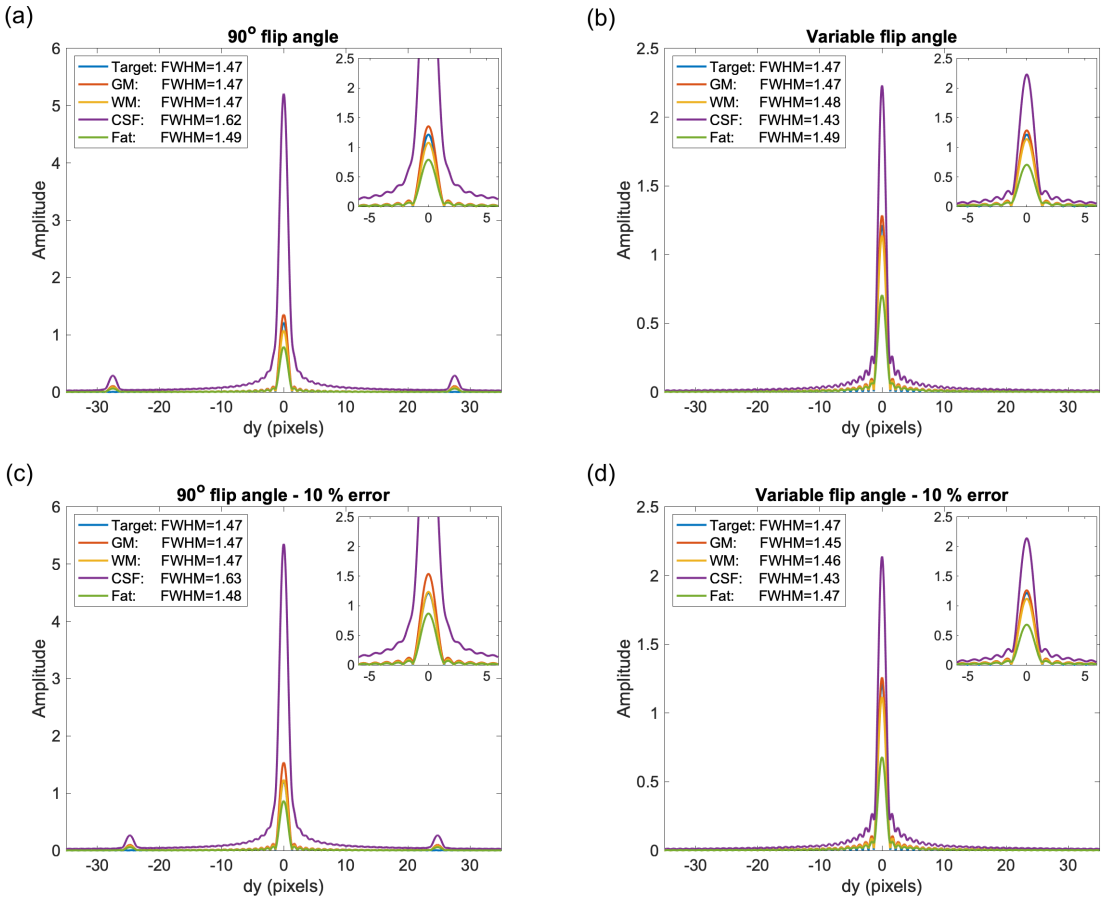

Figure S6: The PSFs for different tissues after applying the correction filters of Figure 2: target brain tissue (T_1=900 ms, T_2=95 ms), GM (T_1=1000 ms, T_2=100 ms), WM (T_1=800 ms, T_2=90 ms), CSF (T_1=2000 ms, T_2=250 ms), and fat (T_1=300 ms, T_2=85 ms). The PSFs are presented for the 90° flip angle scheme (a) and the optimized variable flip angle scheme (b), and again with an introduced RF error, corresponding to 10 % reduced flip angles, in (c) and (d). The FWHMs of each PSF are specified in the legends.

References

[1] F. Schick, SPLICE: Sub-second diffusion-sensitive MR imaging using a modified fast spin-echo acquisition mode, Magnetic Resonance in Medicine 38 (1997) 638–644. doi:10.1002/mrm.1910380418.