Christina Graf, BSc

RF Pulse Design for Chemical Exchange Saturation Transfer by Optimal Control

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

O.Univ.-Prof. Karl Kunisch
Institut für Mathematik und Wissenschaftliches Rechnen, KFU Graz

Dr.rer.nat. Armin Rund
Institut für Mathematik und Wissenschaftliches Rechnen, KFU Graz

Dipl.-Ing. Christoph Aigner
Institut für Medizintechnik, TU Graz

Graz, April 2018

Danke Mama und Papa
Danke Philipp

Danke Patrick

Danke Armin und Christoph

Danke Prof. Kunisch und Prof. Stollberger

Danke IMT: Andi, Andi, Christian, Felix, Markus, Martin, Matthias, Oliver, Stefan
Und natürlich Pilvi the Dog

Danke Christoph, Kathi, Raphael, Richy

ABSTRACT

This thesis deals with RF pulse optimization for chemical exchange saturation transfer by optimal control.
Chemical exchange effects, which can arise from off-resonant RF irradiation, are described by the Bloch-McConnell equation. An analytical and two numerical solvers based on operator splitting are introduced and compared. A fast forward approximate numerical solution is possible with high accuracy due to second order convergence with respect to the time step together with exact solutions of the subsystems. Due to the special form of the Bloch-McConnell equation, part of said numerical solution can be described using fast rotation matrices. The accuracy of the numerical solution method is shown in various numerical experiments.
For the optimization of RF pulses, a reasonable cost functional is modeled and analyzed regarding various aspects. Including physical limitations, an optimal control problem is derived. For every suitable RF pulse, existence and uniqueness of a solution of the Bloch-McConnell equation are shown. The corresponding adjoint equation and linearized state equation are introduced and a connection between them is established. The Fréchet differentiability of the *control-to-state-operator* is derived. Finally, for the optimal control problem, first order necessary conditions for optimality are derived. These form the basis for numerical optimization.
Numerical experiments are performed with a two pool model, where magnetization exchange arises between a water proton pool and a solute proton pool. Optimization runs are performed regarding various aspects. Different functionals are compared. Since standard tracking turns out to be infeasible, a second dynamical system is added where the chemical exchange is turned off. A combined L^∞−tracking of both dynamical systems is proposed. It is reformulated with inequality constraints and solved approximately by penalization techniques. Since sparsity of the RF pulse is important for the application, numerical optimization results are presented both for given sparsity and sparse control. In postoptimal simulation studies, the stability of input parameters, especially of the magnetization exchange rates, is investigated. Another two pool problem is analyzed, which allows for validation by phantom experiments on a 3T MR−hardware.

> *There is increasing interest in chemical exchange saturation transfer contrast, as several metabolites have been reported to be detectable by this approach in vivo, including amide protons in proteins, creatine, glutamate and glucose.* [24]

ZUSAMMENFASSUNG

Diese Arbeit befasst sich mit der RF Puls Optimierung für Sättigungsübertragung durch chemischen Austausch mittels optimaler Steuerung.
Auswirkungen des chemischen Austausches auf die Magnetisierung der Protonen im Wasser, welcher durch off-resonante RF Anregung entstehen kann, werden durch die Bloch-McConnell Gleichung beschrieben. Ein analytischer und zwei numerische Löser, basierend auf einem Operatorsplitting, werden eingeführt und verglichen. Quadratisches konvergentes symmetrisches Splitting in Verbindung mit exakter Lösung der Subsysteme erlaubt eine genaue und effiziente numerische Lösung der Vorwärtsgleichung. Wegen der speziellen Form der Bloch-McConnell Gleichung kann ein Teil der numerischen Lösung günstig mit Rotationsmatrizen beschrieben werden. Die Genauigkeit der numerischen Löser wird durch verschiedene numerische Beispiele untermauert.
Für die Optimierung eines RF Pulses wird ein geeignetes Kostenfunktional modelliert und unter verschiedenen Aspekten analysiert. Das Optimalsteuerungsproblem wird aufgestellt unter Berücksichtigung der technischen Beschränkungen der MR−Hardware. Für jeden geeigneten RF Puls wird die Existenz und Eindeutigkeit einer Lösung der Bloch-McConnell Gleichung gezeigt. Die zugehörige adjungierte Gleichung und die linearisierte Zustandsgleichung werden eingeführt und es wird eine Verbindung zwischen diesen aufgebaut. Die Fréchet Differenzierbarkeit des *Steuerungs-Zustands-Operator* wird gezeigt. Alles zusammen erlaubt schließlich die notwendigen Optimalitätsbedingungen erster Ordnung herzuleiten. Diese bilden die Basis für numerische Optimierung.
Numerische Experimente konzentrieren sich zunächst auf ein Beispiel mit zwei Pools von Protonen, von Wasser und von einem darin gelösten Stoff. Verschiedene Optimierungsstudien werden präsentiert. Zuerst wird das Kostenfunktional näher untersucht. Ein sinnvoller Sollzustand wird hergeleitet. Da ein standard Tracking sich als nicht zielführend herausstellt, wird ein zweites dynamisches System hinzugefügt, wobei der chemische Austausch auf Null gesetzt wird. Ein kombiniertes L^∞−Tracking beider dynamischen Systeme ermöglicht das Design guter RF Pulse für CEST. Nach Umformulierung mittels Ungleichungen kann es approximativ über Penalisierung gelöst werden. Numerische Optimierungsergebnisse werden präsentiert, auch mit Fokus auf dünn besetzte Steuerungen. Im Anschluss an die Optimierung wird die Robustheit des optimierten RF Pulses bezüglich der Input Parameter, im besonderen der Austauschraten, untersucht. Ein zweites zwei Pool Modell wird untersucht, welches speziell für Phantom Messungen an einer 3T MR−Hardware definiert wurde.

> *Es besteht erhöhtes Interesse in Kontrasten durch chemischen Austausch und Sättigungsübertragung, da einige Stoffwechselvorgänge durch diesen Ansatz in vivo sichtbar werden, eingeschlossen Amid Protonen in Proteinen, Kreatin, Glutamat und Glucose.* [24]

CONTENTS

1 ANALYTICAL AND NUMERICAL SOLUTION OF THE BLOCH AND BLOCH–MCCONNELL EQUATIONS

For the Bloch and Bloch-McConnell equations three solvers - an analytical one and two fast numerical ones - are introduced. The analytical solvers are based on eigenvalues and eigenvectors and yield exact solutions at the cost of an increased runtime. However, in case of RF pulse optimization, the Bloch and the Bloch-McConnell equations need to be simulated frequently. Therefore, two fast numerical solvers are derived for the Bloch equation. One is based on an asymmetric operator splitting and is first order accurate, while the other one is based on a symmetric operator splitting and is second order accurate. The theory based on the asymmetric solver is already used in the well known solver of Brian Hargreaves, see [2]. Additional literature can be found in [8] and [9]. Afterwards, the numerical solvers are extended to the Bloch-McConnell equation.

1.1 Bloch's equation

The temporal evolution of the magnetization vector M in external fields can be modeled by the Bloch equation. In 1946, the physicist Felix Bloch stated the assumption, that individual nuclei in a sample can be described by a single vector M, called the net magnetization. Joseph Lamor investigated, that an external field B produces a "twisting force" on M resulting in its precession at frequency γB.
Bloch recognized that the signal quickly decays to zero, where the decay time is based on the studied material. Said signal decay arises from the interaction of the individual spins with each other and the environment. Therefore, M returns to its initial position parallel to the static magnetic field B with magnitude M_0 by releasing energy to its environment. This process is called relaxation. The relaxation times T_1 and T_2 reflect the regrowth of longitudinal magnetization M_z on the one hand, and the decay of transverse magnetization (M_x, M_y) on the other hand.
During returning to the equilibrium after a 90° pulse, the magnetization vector M performs as

$$\begin{aligned} M_x(t) &= M_0 e^{-t/T_2} \cos(\omega t), \\ M_y(t) &= M_0 e^{-t/T_2} \sin(\omega t), \\ M_z(t) &= M_0 (1 - e^{-t/T_1}), \end{aligned}$$

which describes a spiraling precession around B with frequency $\omega = \gamma B$. ω is called Lamor frequency and γ gyromagnetic ratio. Out of it, Bloch described this phenomena by the Bloch equation

$$\begin{aligned} \frac{dM_x(t)}{dt} &= \gamma (M(t) \times B(t))_x - \frac{M_x(t)}{T_2}, \\ \frac{dM_y(t)}{dt} &= \gamma (M(t) \times B(t))_y - \frac{M_y(t)}{T_2}, \\ \frac{dM_z(t)}{dt} &= \gamma (M(t) \times B(t))_z - \frac{M_z(t) - M_0}{T_1}. \end{aligned}$$

This is a macroscopic description which is a sufficient description for a large number of spins. They describe the motion of the sum of all nuclear magnetic moment in the sample. A microscopic description is governed by laws of quantum mechanics.

The full time dependent Bloch equation with relaxation in the rotating frame is therefore given as

$$\begin{cases} \frac{dM}{dt}(t,z) = \begin{pmatrix} -\frac{1}{T_2} & \gamma G_s(t)z & -\gamma B_{1,y}(t) \\ -\gamma G_s(t)z & -\frac{1}{T_2} & \gamma B_{1,x}(t) \\ \gamma B_{1,y}(t) & -\gamma B_{1,x}(t) & -\frac{1}{T_1} \end{pmatrix} M(t,z) + \begin{pmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{pmatrix}, \\ M(0,z) = M^0(z), \end{cases} \tag{1}$$

where $M(t,z) = (M_x(t,z), M_y(t,z), M_z(t,z))^T$ is the magnetization vector, γ the gyromagnetic ratio, $M^0 \in \mathbb{R}^3$ the initial magnetization, M_0 the equilibrium magnetization and T_1, T_2 are the relaxation times. $B_1(t) = (B_{1,x}(t), B_{1,y}(t))^T$ is the radio frequency (RF) pulse and $G_s(t)$ is the slice-selective gradient, see [11].
For fixed position z, (1) is a linear ordinary differential equation system with non-constant coefficients. Therefore, a full analytical solution is not possible in general, e.g. [15]. It is possible for the special case of *free induction decay*, where the RF pulse B_1 is zero. Furthermore, for piecewise constant RF pulse B_1 and gradient G_s, (1) is per time step a linear system with constant coefficients. Therefore, an analytic solution with eigenvalues and eigenvectors is possible, e.g. [15]. Hence, the time interval $[0,T]$ is uniformly discretized with step length τ using N steps

$$0 = t_0 < t_1 < \cdots < t_N = T$$

leading to piecewise constant matrices $A_1, \cdots A_N$, where

$$A_n = \begin{pmatrix} -\frac{1}{T_2} & \gamma G_s(t_{n-1})z & -\gamma B_{1,y}(t_{n-1}) \\ -\gamma G_s(t_{n-1})z & -\frac{1}{T_2} & \gamma B_{1,x}(t_{n-1}) \\ \gamma B_{1,y}(t_{n-1}) & -\gamma B_{1,x}(t_{n-1}) & -\frac{1}{T_1} \end{pmatrix}.$$

Now the piecewise constant Bloch equation in each time step $[t_{n-1}, t_n]$ reads

$$\begin{cases} \frac{dM}{dt} &= A_n \cdot M + \tilde{b} \quad \text{on} \quad (0,\tau), \\ M(0,z) &= M^{n-1}(z), \end{cases} \tag{2}$$

where M^{n-1} is given from the previous time step with M^0 being the initial condition. We set $M^n(z) = M(\tau, z)$.

1.1.1 Analytical solution based on eigenvalues and eigenvectors

Below an analytical solution for the piecewise constant Bloch equation (2) is designed based on eigenvalues and eigenvectors. Calculating the characteristic polynomial of the system matrix A_n leads to the problem

$$\chi(\lambda) = \lambda^3 + p\lambda^2 + q\lambda + r \stackrel{!}{=} 0 \tag{3}$$

for some coefficients $p, q, r \in \mathbb{R}$. The three roots of (3) are λ_1, λ_2 and λ_3, where $\lambda_i \in \mathbb{C}/\mathbb{R}$ is a root if and only if the complex conjugate $\overline{\lambda_i}$ is a root. Now the solution behavior of (2) can be characterized using the properties of geometric and algebraic multiplicity of the eigenvalues. For doing so, the definition of generalized eigenvectors shall be introduced.

Definition 1. [15] $v \in \mathbb{C}^n \backslash \{0\}$ is called *generalized eigenvector* of order $l \in \mathbb{N}$ corresponding to the eigenvalue λ of the matrix $A \in \mathbb{C}^{n\times n}$, if

$$(A - \lambda I)^l v = 0 \quad \text{and} \quad (A - \lambda I)^{l-1} v \neq 0,$$

where I is the identity matrix in $\mathbb{C}^{n\times n}$.

With this definition, the general solution of the homogenous equation

$$\frac{dM_h}{dt} = A_n M_h \tag{4}$$

is now given for different cases.
Case 1: λ_1 has algebraic multiplicity 3. Therefore $\lambda_1 = \lambda_2 = \lambda_3 \in \mathbb{R}$.

Case 1a: λ_1 has geometric multiplicity 1. Let v_1 be the eigenvector corresponding to λ_1. The corresponding generalized eigenvectors of order 2 and 3 are u_2 and u_3. Now the general solution of the homogenous equation (4) is given as

$$\begin{aligned} M_h(t) = {} & c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_1 t}(u_2 + t(A_n - \lambda_1 I)u_2) \\ & + c_3 e^{\lambda_1 t}(u_3 + t(A_n - \lambda_1 I)u_3 + \frac{t^2}{2}(A_n - \lambda_1 I)^2 u_3). \end{aligned}$$

with constants c_1, c_2, c_3 in $\mathbb{R}$.

Case 1b: λ_1 has geometric multiplicity 2. Let v_1 and v_2 the corresponding linearly independent eigenvectors. Let u_2 be the generalized eigenvector of order 2. Then the homogenous solution is

$$M_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_1 t} v_2 + c_3 e^{\lambda_1 t}(u_2 + t(A_n - \lambda_1 I)u_2).$$

Case 1c: λ_1 has geometric multiplicity 3. Therefore we can choose three linearly independent eigenvectors v_1, v_2 and v_3 and the homogenous solution is given as

$$M_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_1 t} v_2 + c_3 e^{\lambda_1 t} v_3.$$

Case 2: λ_1 has algebraic multiplicity 1 and $\lambda_2 \neq \lambda_1$ has algebraic multiplicity 2, i.e. $\lambda_2 = \lambda_3 \in \mathbb{R}$.

Case 2a: λ_2 has geometric multiplicity 1. Let v_1 the be eigenvector corresponding to λ_1 and v_2 the one corresponding to λ_2. Let the generalized eigenvector of order 2 associated with λ_2 be u_2. Now the homogenous solution is

$$M_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + c_3 e^{\lambda_2 t}(u_2 + t(A_n - \lambda_2 I)u_2).$$

Case 2b: λ_1 has geometric multiplicity 1 with eigenvector v_1 and λ_2 has geometric multiplicity 2 with linearly independent eigenvectors v_2 and v_3 which leads to

$$M_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + c_3 e^{\lambda_2 t} v_3.$$

Case 3: Three different eigenvalues λ_1, λ_2 and λ_3. Therefore, the geometric multiplicity of each eigenvalue is 1.

Case 3a: The eigenvalues are real, which means that the corresponding eigenvectors v_1, v_2 and v_3 are real as well and the homogenous solution is

$$M_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + c_3 e^{\lambda_3 t} v_3.$$

Case 3b: λ_1 is real, but λ_2 is not, i.e. $\lambda_{2,3} = \alpha \pm i\beta$ with corresponding eigenvectors v_1 and $v_{2,3} = a \pm ib$. Here the homogenous solution is

$$\begin{aligned} M_h(t) = c_1 v_1 e^{\lambda_1 t} &+ c_2 e^{\alpha t} \left(\cos(\beta t) \cdot a - \sin(\beta t) \cdot b\right) \\ &+ c_3 e^{\alpha t} \left(\sin(\beta t) \cdot a + \cos(\beta t) \cdot b\right). \end{aligned}$$

In all cases, the particular solution $M_p(t)$ of (2) is given as $M_p(t) = -A_n^{-1}\tilde{b}$ and therefore the full solution reads

$$M(t) = M_h(t) + M_p(t).$$

The constants c_1, c_2 and c_3 shall be calculated out of the initial condition

$$M(0) \;=\; M^{n-1}.$$

1.1.2 Numerical solution based on operator splitting

The idea of operator splitting methods is to decompose the problem into simpler subproblems and solve them individually using specific solvers. In this thesis, an asymmetric and a symmetric operator splitting scheme are used. Therefore, the Bloch matrix A_n in each time step and for fixed position z is decomposed additively into

$$A_n = R_n + C,$$

where

$$R_n = \begin{pmatrix} 0 & \gamma G_s(t_{n-1})z & -\gamma B_{1,y}(t_{n-1}) \\ -\gamma G_s(t_{n-1})z & 0 & \gamma B_{1,x}(t_{n-1}) \\ \gamma B_{1,y}(t_{n-1}) & -\gamma B_{1,x}(t_{n-1}) & 0 \end{pmatrix}$$

describes the rotation and

$$C = \begin{pmatrix} -\frac{1}{T_2} & & \\ & -\frac{1}{T_2} & \\ & & -\frac{1}{T_1} \end{pmatrix}$$

the relaxation effects. Note that C does not depend on the time t. Now an asymmetric operator splitting scheme is defined as

$$\frac{dM_{rot}}{dt} = R_n \cdot M_{rot}, \qquad M_{rot}(0) = M^{n-1}, \tag{5a}$$

$$\frac{dM_{rel}}{dt} = C \cdot M_{rel} + \tilde{b}, \qquad M_{rel}(0) = M_{rot}(\tau), \tag{5b}$$

where M^n is defined as

$$M^n = M_{rel}(\tau).$$

In contrast, a symmetric operator splitting scheme is defined as

$$\frac{dM_{rot,1}}{dt} = R_n \cdot M_{rot,1}, \qquad M_{rot,1}(0) = M^{n-1}, \tag{6a}$$

$$\frac{dM_{rel}}{dt} = C \cdot M_{rel} + \tilde{b}, \qquad M_{rel}(0) = M_{rot,1}\left(\frac{1}{2}\tau\right), \tag{6b}$$

$$\frac{dM_{rot,2}}{dt} = R_n \cdot M_{rot,2}, \qquad M_{rot,2}(0) = M_{rel}(\tau), \tag{6c}$$

where M^n is given as

$$M^n = M_{rot,2}(\tau).$$

The three rotation parts (5a), (6a) and (6c) can be solved easily using rotation matrices $D_{n,asy}$ and $D_{n,sy}$, see [7] and [5], so that

$$\begin{aligned} M_{rot} &= D_{n,asy} \cdot M_{rot}(0), \\ M_{rot,1} &= D_{n,sy} \cdot M_{rot,1}(0), \\ M_{rot,2} &= D_{n,sy} \cdot M_{rot,2}(0). \end{aligned}$$

The rotation matrices D underly the following definition.

Definition 2 (Rotation Matrices). [5] The rotation by an angle α about an axis in the mathematical positive sense in the direction of $n = (n_1, n_2, n_3)$ with $|n| = 1$ is given as

$$D(\alpha, n) = \begin{pmatrix} n_1^2(1-\cos\alpha)+\cos\alpha & n_1 n_2(1-\cos\alpha)-n_3\sin\alpha & n_1 n_3(1-\cos\alpha)+n_2\sin\alpha \\ n_1 n_2(1-\cos\alpha)+n_3\sin\alpha & n_2^2(1-\cos\alpha)+\cos\alpha & n_2^3(1-\cos\alpha)-n_1\sin\alpha \\ n_1 n_3(1-\cos\alpha)-n_2\sin\alpha & n_2 n_3(1-\cos\alpha)+n_1\sin\alpha & n_3^2(1-\cos\alpha)+\cos\alpha \end{pmatrix}.$$

Since C is a diagonal matrix, equations (5b) and (6b) can be solved as

$$M_{rel} = \exp(C\tau)(M_{rel}(0) - \tilde{b}) + \tilde{b}$$

where

$$\exp(C\tau) = \begin{pmatrix} \exp(-\frac{1}{T_2}\tau) & & \\ & \exp(-\frac{1}{T_2}\tau) & \\ & & \exp(-\frac{1}{T_1}\tau) \end{pmatrix}.$$

1.2 Bloch-McConnell's equation

Suppose you have two proton pools A and B. For each pool, magnetization can be described individually via the Bloch equation, each of them having individual relaxation times, equilibrium magnetization and initial magnetization. Through various physical processes, the individual Bloch equations are coupled and the two pools can exchange magnetization with rates $k_{A,B}$ and $k_{B,A}$, respectively, see Figure 1. This exchange can be modeled by adding exchange rates to the Bloch equation.

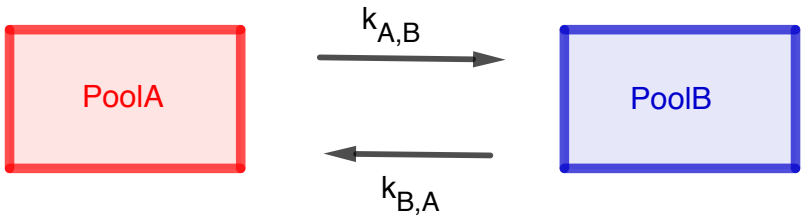

Figure 1: Magnetization exchange between pool A and pool B.

Expanding the Bloch equation (1) with magnetization transfer terms for a water proton pool w and two solute proton pools s_1 and s_2 yields the Bloch-McConnell equation

$$\begin{cases} \frac{dM}{dt}(t,z) = A(t,z)\cdot M(t,z) + \tilde{b}, \\ M(0,z) = M^0(z). \end{cases} \tag{7}$$

Therein the system matrix $A(t,z)$ is given as a block matrix

$$A(t,z) = \begin{pmatrix} A_{11}(t,z) & k_{s_1,w}I & k_{s_2,w}I \\ k_{w,s_1}I & A_{22}(t,z) & k_{s_2,s_1}I \\ k_{w,s_2}I & k_{s_1,s_2}I & A_{33}(t,z) \end{pmatrix} \tag{8}$$

where $I \in \mathbb{R}^{3\times 3}$ is the identity matrix. The submatrices $A_{11}(t,z)$, $A_{22}(t,z)$ and $A_{33}(t,z)$ are given as

$$A_{11}(t,z) = \begin{pmatrix} -\frac{1}{T_{2,w}} - k_{w,s_1} - k_{w,s_2} & \Delta\omega_w(z) & -\gamma B_{1,y}(t) \\ -\Delta\omega_w(z) & -\frac{1}{T_{2,w}} - k_{w,s_1} - k_{w,s_2} & \gamma B_{1,x}(t) \\ \gamma B_{1,y}(t) & -\gamma B_{1,x}(t) & -\frac{1}{T_{1,w}} - k_{w,s_1} - k_{w,s_2} \end{pmatrix},$$

$$A_{22}(t,z) = \begin{pmatrix} -\frac{1}{T_{2,s_1}} - k_{s_1,w} - k_{s_1,s_2} & \Delta\omega_{s_1}(z) & -\gamma B_{1,y}(t) \\ -\Delta\omega_{s_1}(z) & -\frac{1}{T_{2,s_1}} - k_{s_1,w} - k_{s_1,s_2} & \gamma B_{1,x}(t) \\ \gamma B_{1,y}(t) & -\gamma B_{1,x}(t) & -\frac{1}{T_{1,s_1}} - k_{s_1,w} - k_{s_1,s_2} \end{pmatrix},$$

$$A_{33}(t,z) = \begin{pmatrix} -\frac{1}{T_{2,s_2}} - k_{s_2,w} - k_{s_2,s_1} & \Delta\omega_{s_2}(z) & -\gamma B_{1,y}(t) \\ -\Delta\omega_{s_2}(z) & -\frac{1}{T_{2,s_2}} - k_{s_2,w} - k_{s_2,s_1} & \gamma B_{1,x}(t) \\ \gamma B_{1,y}(t) & -\gamma B_{1,x}(t) & -\frac{1}{T_{1,s_2}} - k_{s_2,w} - k_{s_2,s_1} \end{pmatrix}.$$

The inhomogeneity reads

$$\tilde{b} = \left(0,\ 0,\ \frac{M_{0,w}}{T_{1,w}},\ 0,\ 0,\ \frac{M_{0,s_1}}{T_{1,s_1}},\ 0,\ 0,\ \frac{M_{0,s_2}}{T_{1,s_2}}\right)^T. \tag{9}$$

The Bloch-McConnell equation (7) describes the magnetization of three proton pools during RF irradiation.

$$M = (M_{x,w}, M_{y,w}, M_{z,w}, M_{x,s_1}, M_{y,s_1}, M_{z,s_1}, M_{x,s_2}, M_{y,s_2}, M_{z,s_2})^T$$

is the vector describing the magnetization of the water proton pool w and the solute proton pools s_1 and s_2, respectively, γ the gyromagnetic ratio, $M^0 \in \mathbb{R}^9$ the initial magnetization, $M_{0,w}$ the equilibrium magnetization corresponding to pool w and M_{0,s_i} the one corresponding to pool s_i, $i = 1, 2$.
$T_{1,w}$, T_{1,s_1} and T_{1,s_2} are the longitudinal relaxation times, $T_{2,w}$, T_{2,s_1} and T_{2,s_2} are the transversal relaxation times. $B_1(t) = (B_{1,x}(t), B_{1,y}(t))^T$ is the radio frequency (RF) pulse. The frequency offset of the RF irradiation relative to the Larmor frequency of the water proton pool w and the solute proton pools $s_i, i = 1, 2$, is described by $\Delta\omega_w(z)$ and $\Delta\omega_{s_i}(z)$ and can be computed through $z - \omega_w$ and $z - \omega_{s_i}$. $k_{i,j}$ are the magnetization transfer rates from pool i to pool j. The Bloch-McConnell equation for another number of solute proton pools is defined analogously. In the following, the theory and methods are derived for two solute proton pools s_1 and s_2. It is obviously valid also for only one solute proton pool by simply setting all the transfer rates corresponding to s_2 to zero.
Again, for fixed z, the time interval $[0, T]$ is uniformly discretized using a step length τ into

$$0 = t_0 < t_1 < \cdots < t_N = T$$

yielding piecewise constant matrices

$$A_1, \cdots A_N. \tag{10}$$

1.2.1 Analytical solution based on diagonalization

A calculation of the analytical solution for Bloch-McConnell's equation in time step $[t_{n-1}, t_n]$ would result in

$$M(t) = (M^{n-1} + A_n^{-1}\tilde{b})\exp(A_n t) - A_n^{-1}\tilde{b}.$$

The matrix exponential $\exp(A_n t)$ can be calculated analytically only if the system matrix A_n is diagonalizable, which is not the case here in general. A fast solution can be found in using *Matlab's* matrix exponential function *expm*, if A_n is not diagonalizable. However, this results in an approximate numerical solution which has been shown to yield large errors in some examples, especially for small relaxation times, see [12]. Therefore, a more robust numerical solution method is introduced below.

1.2.2 Numerical solution based on operator splitting

As in Section 1.1.2 the numerical solvers for the Bloch-McConnell equation are based on an asymmetric and a symmetric operator splitting scheme. Consider a splitting of the discretized Bloch-McConnell matrices (10) of the form

$$A_n = R_n + C$$

where

$$R_n = \begin{pmatrix} R_{11,n} & & \\ & R_{22,n} & \\ & & R_{33,n} \end{pmatrix} \tag{11}$$

with

$$R_{11,n} = \begin{pmatrix} 0 & \Delta\omega_w(z) & -\gamma B_{1,y}(t_{n-1}) \\ -\Delta\omega_w(z) & 0 & \gamma B_{1,x}(t_{n-1}) \\ \gamma B_{1,y}(t_{n-1}) & -\gamma B_{1,x}(t_{n-1}) & 0 \end{pmatrix},$$
$$R_{22,n} = \begin{pmatrix} 0 & \Delta\omega_{s_1}(z) & -\gamma B_{1,y}(t_{n-1}) \\ -\Delta\omega_{s_1}(z) & 0 & \gamma B_{1,x}(t_{n-1}) \\ \gamma B_{1,y}(t_{n-1}) & -\gamma B_{1,x}(t_{n-1}) & 0 \end{pmatrix},$$
$$R_{33,n} = \begin{pmatrix} 0 & \Delta\omega_{s_2}(z) & -\gamma B_{1,y}(t_{n-1}) \\ -\Delta\omega_{s_2}(z) & 0 & \gamma B_{1,x}(t_{n-1}) \\ \gamma B_{1,y}(t_{n-1}) & -\gamma B_{1,x}(t_{n-1}) & 0 \end{pmatrix}$$

describes the rotation of each pool independently. The other part

$$C = \begin{pmatrix} C_{11} & k_{s_1,w}I & k_{s_2,w}I \\ k_{w,s_1}I & C_{22} & k_{s_2,s_1}I \\ k_{w,s_2}I & k_{s_1,s_2}I & C_{33} \end{pmatrix} \tag{12}$$

with

$$C_{11} = \begin{pmatrix} -\frac{1}{T_{2,w}} - k_{w,s_1} - k_{w,s_2} & -\frac{1}{T_{2,w}} - k_{w,s_1} - k_{w,s_2} & \\ & & -\frac{1}{T_{1,w}} - k_{w,s_1} - k_{w,s_2} \end{pmatrix},$$
$$C_{22} = \begin{pmatrix} -\frac{1}{T_{2,s_1}} - k_{s_1,w} - k_{s_1,s_2} & -\frac{1}{T_{2,s_1}} - k_{s_1,w} - k_{s_1,s_2} & \\ & & -\frac{1}{T_{1,s_1}} - k_{s_1,a} - k_{s_1,s_2} \end{pmatrix},$$
$$C_{33} = \begin{pmatrix} -\frac{1}{T_{2,s_2}} - k_{s_2,w} - k_{s_2,s_1} & -\frac{1}{T_{2,s_2}} - k_{s_2,w} - k_{s_2,s_1} & \\ & & -\frac{1}{T_{1,s_2}} - k_{s_2,w} - k_{s_2,s_1} \end{pmatrix}$$

models the relaxation effects and the exchange. These matrices are constant in time. Similarly to chapter 1.1.2 an asymmetric operator splitting scheme is defined as

$$\frac{dM_{rot,1}}{dt} = R_n \cdot M_{rot}, \qquad M_{rot}(0) = M^{n-1}, \tag{13a}$$
$$\frac{dM_{rel}}{dt} = C \cdot M_{rel} + \tilde{b}, \qquad M_{rel}(0) = M_{rot}(\tau), \tag{13b}$$

where τ is the step length and M^n is given as

$$M^n = M_{rot}(\tau).$$

Again, the symmetric operator splitting scheme is defined as

$$\frac{dM_{rot,1}}{dt} = R_n \cdot M_{rot,1}, \qquad M_{rot,1}(0) = M^{n-1}, \tag{14a}$$
$$\frac{dM_{rel}}{dt} = C \cdot M_{rel} + \tilde{b}, \qquad M_{rel}(0) = M_{rot,1}(\frac{1}{2}\tau), \tag{14b}$$
$$\frac{dM_{rot,2}}{dt} = R_n \cdot M_{rot,2}, \qquad M_{rot,2}(0) = M_{rel}(\tau), \tag{14c}$$

where

$$M^n = M_{rot,2}(\tau).$$

Since R_n is a blockdiagonal matrix of Bloch rotation matrices, one can use rotation matrices $D_{n,asy}$ and $D_{n,sy}$ so that

$$\begin{aligned} M_{rot} &= D_{n,asy} \cdot M_{rot}(0), \\ M_{rot,1} &= D_{n,sy} \cdot M_{rot,1}(0), \\ M_{rot,2} &= D_{n,sy} \cdot M_{rot,2}(0). \end{aligned}$$

The calculation of (13b) and (14b) is done using eigenvalues and eigenvectors and the theory deduced in 1.1.1. The following Lemma shows that a reduction to three smaller independent eigenvalue problems is possible by posing separate problems for the x, y and z components.

Lemma 1 (Calculation of the eigenvalues). The eigenvalues of matrix (12) are the eigenvalues of the three submatrices

$$C_1 = C_2 = \begin{pmatrix} -\frac{1}{T_{2,w}} - k_{w,s_1} - k_{w,s_2} & k_{s_1,w} & k_{s_2,w} \\ k_{w,s_1} & -\frac{1}{T_{2,s_1}} - k_{s_1,w} - k_{s_1,s_2} & k_{s_2,s_1} \\ k_{w,s_2} & k_{s_1,s_2} & -\frac{1}{T_{2,s_2}} - k_{s_2,w} - k_{s_2,s_1} \end{pmatrix}$$

and

$$C_3 = \begin{pmatrix} -\frac{1}{T_{1,w}} - k_{w,s_1} - k_{w,s_2} & k_{s_1,w} & k_{s_2,w} \\ k_{w,s_1} & -\frac{1}{T_{1s_1}} - k_{s_1,w} - k_{s_1,s_2} & k_{s_2,s_1} \\ k_{w,s_2} & k_{s_1,s_2} & -\frac{1}{T_{1,s_2}} - k_{s_2,w} - k_{s_2,s_1} \end{pmatrix}$$

of C.

Proof. The idea of the proof is to transform the submatrix C into a blockdiagonal matrix using similarity transformations.
Preliminary Remark Let Q be a unitary matrix, i.e. $QQ^T = Q^TQ = I$ and assume that (λ, x) is an eigenpair of C. Let $w := Q^Tx$, then

$$Q^TCQw = Q^TCQQ^Tx = Q^TCx = Q^T\lambda x = \lambda Q^Tx = \lambda w$$

which means that (λ, x) is an eigenpair of C if and only if (λ, Q^Tx) is an eigenpair of Q^TCQ. With that, define a unitary matrix Q as

$$Q = \begin{pmatrix} e_1 & e_4 & e_7 & e_2 & e_5 & e_8 & e_3 & e_6 & e_9 \end{pmatrix},$$

where e_i denotes the $i-$th unit vector and note that

$$Q^T C Q = \begin{pmatrix} C_1 & & \\ & C_2 & \\ & & C_3 \end{pmatrix}.$$

Since the eigenvalues of a blockdiagonal matrix are the eigenvalues of the blocks, the assertion follows. □

Now (13b) and (14b) can be solved using the theory of chapter 1.1.1 for each block of $Q^T C Q$.

1.3 Numerical experiments

The error behavior of the asymmetric operator splitting solver, abbreviated with ASY, and the symmetric operator splitting solver, abbreviated with SY, is analyzed using an example with a water proton pool w at $0\,\mathrm{ppm}$ and a solute proton pool s at $3.5\,\mathrm{ppm}$. The input data used for this comparison coincides with the model problem which was used in [23], where the parameters originate from an amide proton system in white matter. Various tests with different relaxation times and zero or non-zero exchange rates are performed. For the comparisons, three different real RF pulses with a pulse length of 2.1s are used, see Figure 2. The first one is block pulse, named BL, the other two are smoother Gaussian pulses, named G1 and G2. The z-spectrum $\Omega = [-5, 5]\,\mathrm{ppm}$ is discretized using 101 steps, whereas the time interval is discretized using a step length of $5e-4s$ in the first experiments. Afterwards, a convergence analysis is done with step lengths starting at $1e-3s$ refining up to $1e-5s$. Finally, the run-time of the analytical solver and both operator splitting solvers is analyzed using a small scale and two large scale examples.
Every coordinate of the magnetization vector M is compared at the terminal time T against the analytical solver, abbreviated with ANA, using specific relative errors

$$\varepsilon_{2,x,w} = \frac{\|M_{a,x,w}(z,T) - M_{n,x,w}(z,T)\|_{L^2(\Omega)}}{\|M_{a,x,w}(z,T)\|_{L^2(\Omega)}},$$

$$\varepsilon_{\infty,x,w} = \frac{\sup\limits_{z\in\Omega} |M_{a,x,w}(z,T) - M_{n,x,w}(z,T)|}{\sup\limits_{z\in\Omega} |M_{a,x,w}(z,T)|} = \frac{\|M_{a,x,w}(z,T) - M_{n,x,k}(z,T)\|_{L^\infty(\Omega)}}{\|M_{a,x,w}(z,T)\|_{L^\infty(\Omega)}}.$$

Here $M_{a,x,w}$ denotes the $x-$coordinate of the water magnetization calculated using the analytic solver, whereas $M_{n,x,w}$ was calculated using one of the numerical solvers. The errors corresponding to the $y-$ and $z-$coordinate and the solute proton pool s are defined analogously. For the tests with relaxation, not only the analytic solver and the two operator splitting solvers are compared, but the numerical solver which was used in [24], abbreviated with E3, as well. The numerical experiments were performed on a MacBook Pro (2GHz Intel Core i5, 8 GB RAM).

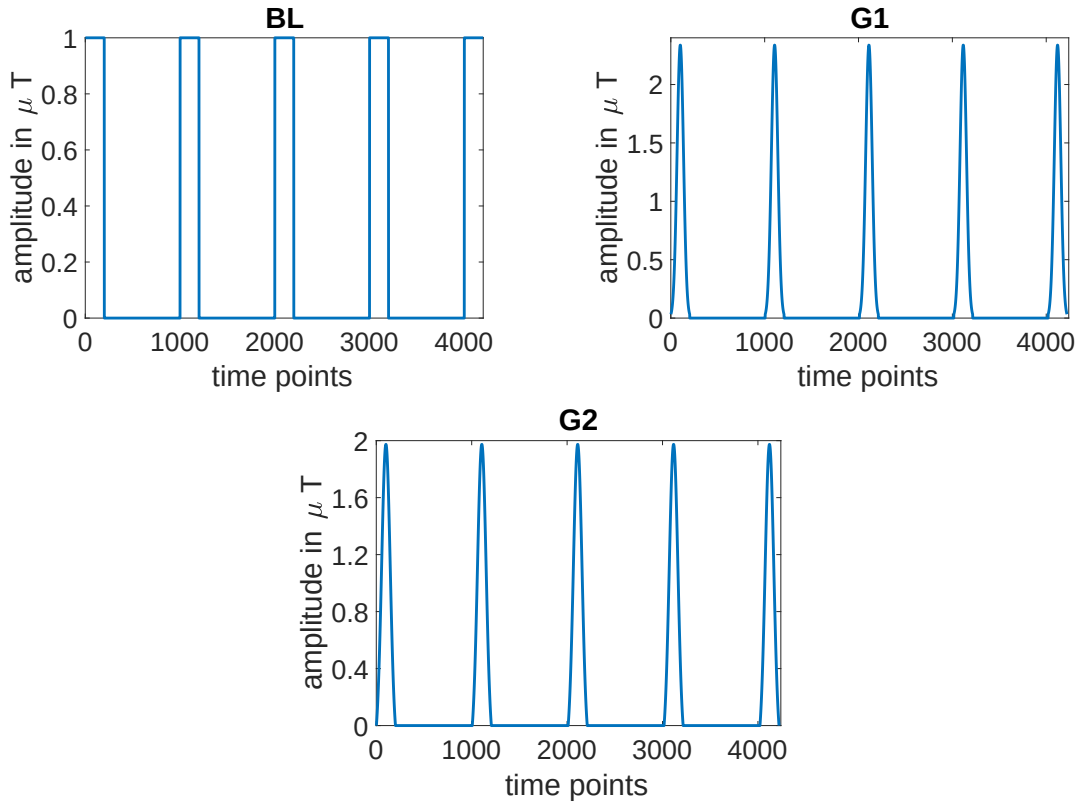

Figure 2: Initial RF pulses using a block type pulse train BL and two Gaussian type pulse trains G1 and G2.

1.3.1 *Comparison without relaxation and without exchange rates*

For the first comparison, no relaxation is assumed, which means setting both the longitudinal relaxation times and the transversal relaxation times to infinity. The exchange rates $k_{w,s}$ and $k_{s,w}$ are set to zero. Therefore, the resulting problem is to solve the Bloch equation two times individually without relaxation. Table 1 shows six error components in both error norms for the two different solvers SY and ASY, and for three different RF pulses BL, G1 and G2. It can be seen that all relative errors are nearly in machine precession. The excellent error behavior did not occur unexpectedly. Since the example is performed without relaxation and without transfer rates, both operator splitting solvers perform with rotation matrices only, which yield an exact solution in this case, as it was shown for the Bloch case we have here, see [7].

Table 1: Error study for block pulse BL and Gaussian pulses G1 and G2 with $T_1 = T_2 = \infty$, $k = 0$.

RF	Solver	Error	$\varepsilon_{i,x,w}$	$\varepsilon_{i,y,w}$	$\varepsilon_{i,z,w}$	$\varepsilon_{i,x,s}$	$\varepsilon_{i,y,s}$	$\varepsilon_{i,z,s}$
BL	SY	$i = 2$	$2.9e-13$	$3.2e-13$	$4.9e-13$	$3.7e-13$	$3.9e-13$	$5.2e-13$
		$i = \infty$	$2.8e-13$	$5.2e-13$	$4.2e-12$	$6.3e-13$	$4.4e-13$	$4.3e-12$
	ASY	$i = 2$	$2.7e-13$	$3.2e-13$	$4.9e-13$	$3.4e-13$	$3.9e-13$	$5.3e-13$
		$i = \infty$	$3.5e-13$	$6.4e-13$	$4.3e-12$	$4.9e-13$	$5.1e-13$	$4.3e-12$
G1	SY	$i = 2$	$1.4e-13$	$8.6e-13$	$4.2e-13$	$1.4e-13$	$9.2e-13$	$4.8e-13$
		$i = \infty$	$1.7e-13$	$1.1e-12$	$4.1e-12$	$1.7e-13$	$1.2e-12$	$4.1e-12$
	ASY	$i = 2$	$2.6e-13$	$8.7e-13$	$4.3e-13$	$2.1e-13$	$8.0e-13$	$4.8e-13$
		$i = \infty$	$2.4e-13$	$1.1e-12$	$4.1e-12$	$2.1e-13$	$1.0e-12$	$4.2e-12$
G2	SY	$i = 2$	$3.7e-13$	$4.2e-13$	$4.2e-13$	$3.5e-13$	$5.0e-13$	$4.7e-13$
		$i = \infty$	$3.7e-13$	$1.1e-12$	$4.1e-12$	$3.7e-13$	$8.5e-13$	$4.2e-12$
	ASY	$i = 2$	$4.7e-13$	$6.7e-13$	$4.3e-13$	$4.5e-13$	$6.1e-13$	$4.8e-13$
		$i = \infty$	$4.4e-13$	$1.1e-12$	$4.2e-12$	$4.7e-13$	$1.0e-12$	$4.2e-12$

1.3.2 Comparison without relaxation and with exchange rates

Both relaxation times for the water proton pool and the solute proton pool are set to infinity, but the exchange rates are assumed to be $k_{w,s} = 0.25\,\text{Hz}$ and $k_{s,w} = 25\,\text{Hz}$. Now, the numerical solution with operator splitting is not exact and

Table 2: Error study for block pulse BL and Gaussian pulses G1 and G2 with $T_1 = T_2 = \infty$, $k \neq 0$.

RF	Solver	Error	$\varepsilon_{i,x,w}$	$\varepsilon_{i,y,w}$	$\varepsilon_{i,z,w}$	$\varepsilon_{i,x,s}$	$\varepsilon_{i,y,s}$	$\varepsilon_{i,z,s}$
BL	SY	$i = 2$	$3.7e-4$	$3.8e-4$	$7.5e-5$	$9.5e-4$	$7.0e-3$	$6.1e-5$
		$i = \infty$	$4.2e-4$	$3.5e-4$	$2.8e-4$	$1.0e-3$	$2.5e-3$	$2.4e-4$
	ASY	$i = 2$	$3.8e-4$	$3.8e-4$	$7.8e-5$	$9.2e-3$	$6.2e-2$	$1.9e-3$
		$i = \infty$	$4.4e-4$	$3.9e-4$	$2.8e-4$	$7.2e-3$	$3.0e-2$	$5.2e-3$
G1	SY	$i = 2$	$1.9e-4$	$2.6e-4$	$1.9e-5$	$2.0e-3$	$2.3e-3$	$8.7e-6$
		$i = \infty$	$1.7e-4$	$2.3e-4$	$1.3e-4$	$2.1e-3$	$1.9e-3$	$6.4e-5$
	ASY	$i = 2$	$2.0e-4$	$2.7e-4$	$2.2e-5$	$1.8e-2$	$2.0e-2$	$9.9e-4$
		$i = \infty$	$1.8e-4$	$2.5e-4$	$1.4e-4$	$1.7e-2$	$1.7e-2$	$3.6e-3$
G2	SY	$i = 2$	$6.2e-4$	$1.2e-4$	$3.2e-5$	$2.3e-3$	$8.8e-4$	$7.9e-6$
		$i = \infty$	$4.6e-4$	$1.0e-4$	$2.2e-4$	$1.9e-3$	$8.6e-4$	$4.8e-5$
	ASY	$i = 2$	$6.3e-4$	$1.4e-4$	$3.3e-5$	$1.0e-2$	$2.1e-2$	$1.1e-3$
		$i = \infty$	$5.1e-4$	$1.5e-4$	$2.2e-4$	$8.7e-3$	$1.8e-2$	$3.9e-3$

the splitting errors can be observed in Table 2. It can be seen that the symmetric operator splitting solver *SY* shows in general smaller errors through all examples than the asymmetric operator splitting solver *ASY*, especially for the solute. The

error of the $z-$coordinate of both the water pool and the solute proton pool is observed to be slightly smaller than the one of the others.

1.3.3 Comparison with relaxation and without exchange rates

The longitudinal relaxation times are set to $T_{1,w} = T_{1,s} = 1.048s$ and the transversal relaxation times to $T_{2,w} = 0.069s$ and $T_{2,s} = 0.015s$ with zero exchange. The solver E3 of [24] can be included in the comparison.

Table 3: Error study for block pulse BL and Gaussian pulses G1 and G2 with $T_1 \neq \infty, T_2 \neq \infty$, $k = 0$.

RF	Solver	Error	$\varepsilon_{i,x,w}$	$\varepsilon_{i,y,w}$	$\varepsilon_{i,z,w}$	$\varepsilon_{i,x,s}$	$\varepsilon_{i,y,s}$	$\varepsilon_{i,z,s}$
BL	SY	$i=2$	$4.5e-5$	$1.0e-3$	$4.7e-6$	$4.6e-5$	$4.1e-2$	$1.5e-5$
		$i=\infty$	$4.3e-5$	$3.5e-4$	$8.7e-6$	$4.7e-5$	$1.8e-2$	$2.3e-5$
	ASY	$i=2$	$3.6e-3$	$4.1e-3$	$1.1e-4$	$1.7e-2$	$8.4e-2$	$1.3e-4$
		$i=\infty$	$3.6e-3$	$3.7e-3$	$2.5e-4$	$1.7e-2$	$3.2e-2$	$2.4e-4$
	E3	$i=2$	$1.5e-1$	$5.2e-1$	$1.3e-2$	$2.6e-1$	$2.7e-1$	$1.1e-2$
		$i=\infty$	$1.5e-1$	$1.8e-1$	$1.7e-2$	$2.6e-1$	$2.5e-1$	$1.3e-2$
G1	SY	$i=2$	$2.6e-5$	$6.9e-5$	$6.8e-6$	$1.6e-4$	$2.7e-3$	$2.6e-5$
		$i=\infty$	$2.2e-5$	$1.8e-5$	$1.2e-5$	$2.0e-4$	$4.8e-4$	$3.5e-5$
	ASY	$i=2$	$3.6e-3$	$3.6e-3$	$9.4e-5$	$1.7e-2$	$1.7e-2$	$1.6e-4$
		$i=\infty$	$3.6e-3$	$3.6e-3$	$2.3e-4$	$1.7e-2$	$1.7e-2$	$2.4e-4$
	E3	$i=2$	$6.1e-2$	$5.9e-2$	$2.6e-1$	$2.6e-1$	$2.6e-1$	$1.1e-2$
		$i=\infty$	$6.0e-2$	$5.8e-2$	$1.3e-2$	$2.6e-1$	$2.6e-1$	$1.3e-2$
G2	SY	$i=2$	$2.7e-5$	$3.1e-5$	$5.1e-6$	$4.7e-4$	$5.1e-4$	$2.3e-5$
		$i=\infty$	$1.7e-5$	$2.1e-5$	$9.6e-6$	$1.4e-4$	$3.2e-4$	$3.3e-5$
	ASY	$i=2$	$3.6e-2$	$3.6e-2$	$9.1e-5$	$1.7e-2$	$1.7e-2$	$1.5e-4$
		$i=\infty$	$3.6e-2$	$3.6e-2$	$2.4e-4$	$1.7e-2$	$1.7e-2$	$2.4e-4$
	E3	$i=2$	$5.9e-2$	$6.0e-2$	$1.2e-2$	$2.6e-1$	$2.6e-1$	$1.1e-2$
		$i=\infty$	$6.0e-2$	$6.0e-2$	$1.3e-2$	$2.6e-1$	$2.6e-1$	$1.3e-2$

In contrast to the case without relaxation, the symmetric operator splitting solver SY shows to have significantly improved accuracy compared to the asymmetric operator splitting solver ASY through all RF pulses, see Table 3. Both splitting schemes outperform E3 in accuracy. Figure 3 shows the absolute pointwise difference of the $z-$spectra of the different numerical solvers to the analytical solver. It can be seen, that the numerical solver E3 performs fine around the water proton pool at 0 ppm, it's difference to the analytical solution increases outside the neighborhood of the water proton pool, where the used RF pulse was the Gaussian G1.

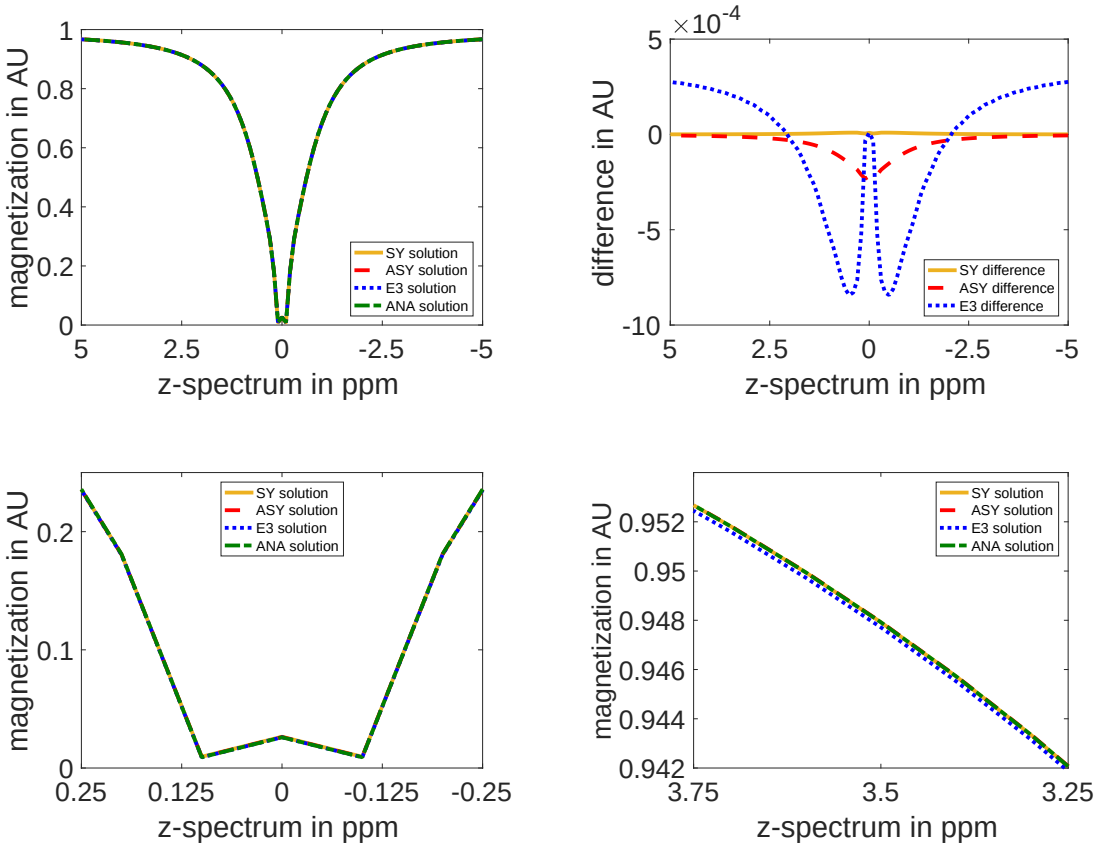

Figure 3: z−magnetization of water pool with $T_1 \neq \infty, T_2 \neq \infty$, $k = 0$ with G1. Full view, difference between analytical solution and numerical solution and two zoom pictures.

1.3.4 *Comparison with relaxation and with exchange rates*

This last comparison works with typical longitudinal relaxation times $T_{1,w} = T_{1,s} = 1.048s$ and the transversal relaxation times $T_{2,w} = 0.069s$ and $T_{2,s} = 0.015s$. The exchange rates are set to $k_{w,s} = 0.25\,\text{Hz}$ and $k_{s,w} = 25\,\text{Hz}$, which are the values from [24]. The symmetric operator splitting solver SY has the best error performance, the asymmetric operator splitting solver ASY is just between E3 and SY, see Table 4. SY shows a reduced L^∞−error compared to ASY of up to factor 1000 and of up to factor 2500 compared to E3. Again, the z−spectra in full view, zoom and difference view are depicted in Figure 4, 5 and 6 for the RF pulses under investigation. The highest absolute errors are observed in Figure 4 in the difference plot, where E3 yields large oscillating error. Obviously, the nonsmooth RF pulse, more precise the block pulse train BL, leads to increased numerical error here. The behavior among the two Gaussian pulses G1 and G2 is about the same, see Figure 5 and 6.

Table 4: Error study for block pulse BL and Gaussian pulses G1 and G2 with $T_1 \neq \infty, T_2 \neq \infty, k \neq 0$.

RF	Solver	Error	$\varepsilon_{i,x,w}$	$\varepsilon_{i,y,w}$	$\varepsilon_{i,z,w}$	$\varepsilon_{i,x,s}$	$\varepsilon_{i,y,s}$	$\varepsilon_{i,z,s}$
BL	SY	$i=2$	$4.4e-5$	$1.1e-3$	$5.1e-6$	$2.2e-4$	$6.0e-3$	$1.9e-5$
		$i=\infty$	$4.5e-5$	$3.7e-4$	$1.0e-5$	$4.4e-4$	$3.4e-3$	$3.2e-5$
	ASY	$i=2$	$3.7e-3$	$4.2e-3$	$1.1e-4$	$2.4e-2$	$3.1e-2$	$2.5e-3$
		$i=\infty$	$3.7e-3$	$3.8e-3$	$2.6e-4$	$2.3e-2$	$2.6e-2$	$5.7e-3$
	E_3	$i=2$	$1.5e-1$	$5.2e-1$	$1.3e-2$	$3.6e-1$	$3.6e-1$	$4.1e-2$
		$i=\infty$	$1.4e-1$	$1.8e-1$	$3.4e-1$	$3.4e-1$	$3.6e-1$	$9.6e-2$
G1	SY	$i=2$	$2.3e-5$	$7.1e-5$	$7.5e-6$	$1.3e-3$	$1.9e-3$	$2.1e-5$
		$i=\infty$	$2.7e-5$	$2.2e-5$	$1.2e-5$	$2.1e-3$	$1.4e-3$	$3.3e-5$
	ASY	$i=2$	$3.7e-3$	$3.7e-3$	$9.4e-5$	$2.6e-2$	$2.6e-2$	$1.6e-3$
		$i=\infty$	$3.7e-3$	$3.7e-3$	$2.3e-4$	$2.2e-2$	$2.3e-2$	$3.7e-3$
	E_3	$i=2$	$6.2e-2$	$6.0e-2$	$1.2e-2$	$3.9e-1$	$3.9e-1$	$2.4e-2$
		$i=\infty$	$6.1e-2$	$5.9e-2$	$1.3e-2$	$3.4e-1$	$3.5e-1$	$5.8e-2$
G2	SY	$i=2$	$2.9e-5$	$3.4e-5$	$5.9e-6$	$1.2e-3$	$1.4e-3$	$2.1e-5$
		$i=\infty$	$2.0e-5$	$2.8e-5$	$9.7e-6$	$1.8e-3$	$1.3e-3$	$3.2e-5$
	ASY	$i=2$	$3.7e-3$	$3.7e-3$	$9.2e-5$	$2.5e-2$	$2.5e-2$	$1.7e-3$
		$i=\infty$	$3.7e-3$	$3.7e-3$	$2.4e-2$	$2.2e-2$	$4.0e-3$	$4.0e-3$
	E_3	$i=2$	$6.1e-2$	$6.1e-2$	$1.2e-2$	$3.8e-1$	$3.8e-1$	$2.6e-2$
		$i=\infty$	$6.1e-2$	$6.1e-2$	$1.3e-2$	$3.5e-1$	$3.5e-1$	$6.5e-2$

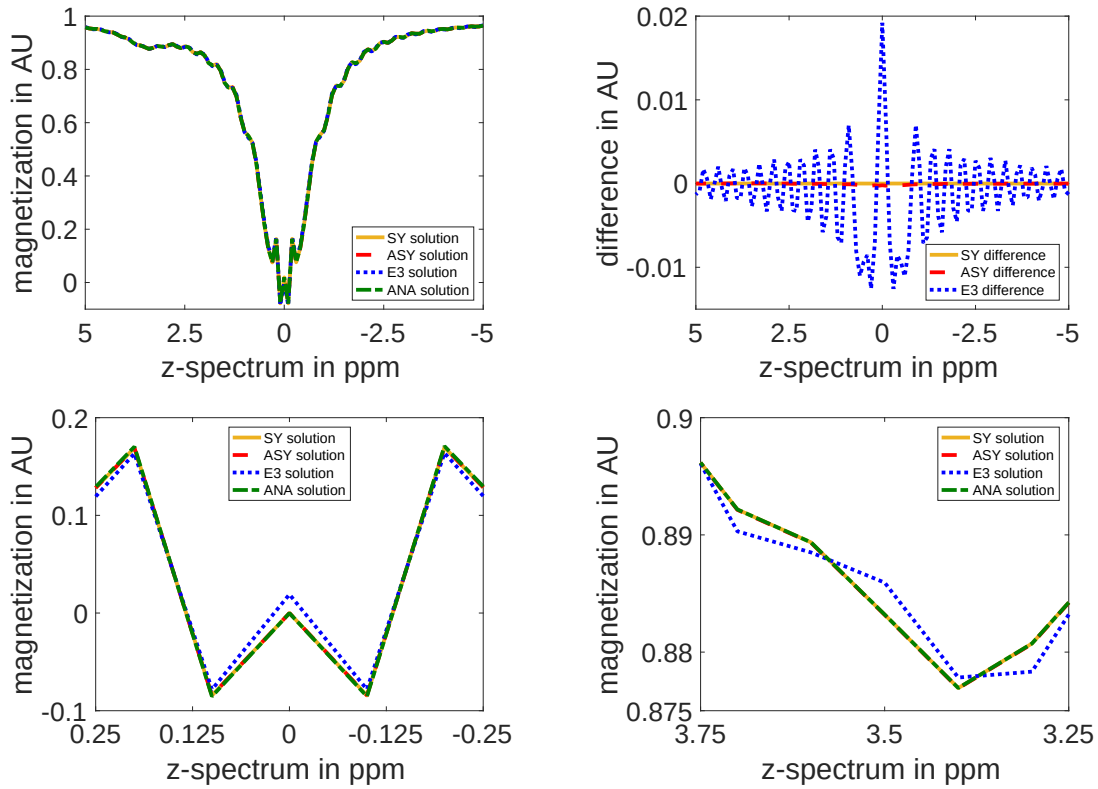

Figure 4: z−magnetization of water pool calculated with BL, where $T_1 \neq \infty$, $T_2 \neq \infty$, $k \neq 0$. Full view, difference between analytical and numerical solution and two zoom pictures.

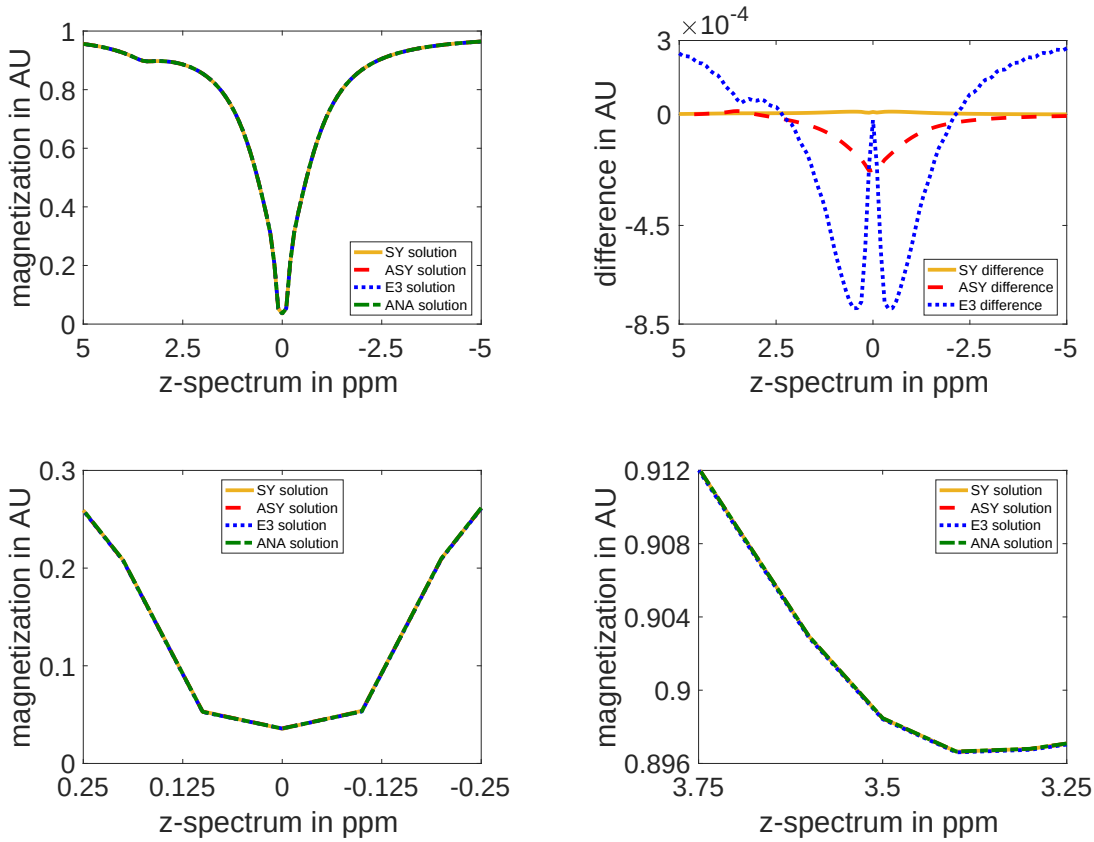

Figure 5: z—magnetization of water pool calculated with G2, where $T_1 \neq \infty, T_2 \neq \infty$, $k \neq 0$. Full view, difference between analytical and numerical solution and two zoom pictures.

1.3.5 *Comparison with various time discretizations*

In this series of tests, the block pulse BL is analyzed with relaxation and with exchange rates for time discretizations with different resolutions. The specific norm used for this comparison is the relative L^2—norm over all coordinates of the magnetization evaluated at the terminal time T, defined as

$$\varepsilon_2 = \frac{\|M_a(z,T) - M_n(z,T)\|_{L^2(\Omega)}}{\|M_a(z,T)\|_{L^2(\Omega)}},$$

where M_a is the magnetization which is calculated with the analytical solver and M_n the magnetization out of one of the numerical solvers, namely the symmetric and the asymmetric operator splitting solver. For a further investigation, the experimental order of convergence EOC is calculated, which is defined as

$$\mathrm{EOC}(\tau,\tau') = \frac{\log\frac{\varepsilon_2(\tau)}{\varepsilon_2(\tau')}}{\log\frac{\tau}{\tau'}}.$$

This approach for a different length of time steps τ results in errors printed in Table 5. From theory one expects a quadratic order of convergence for the symmetric

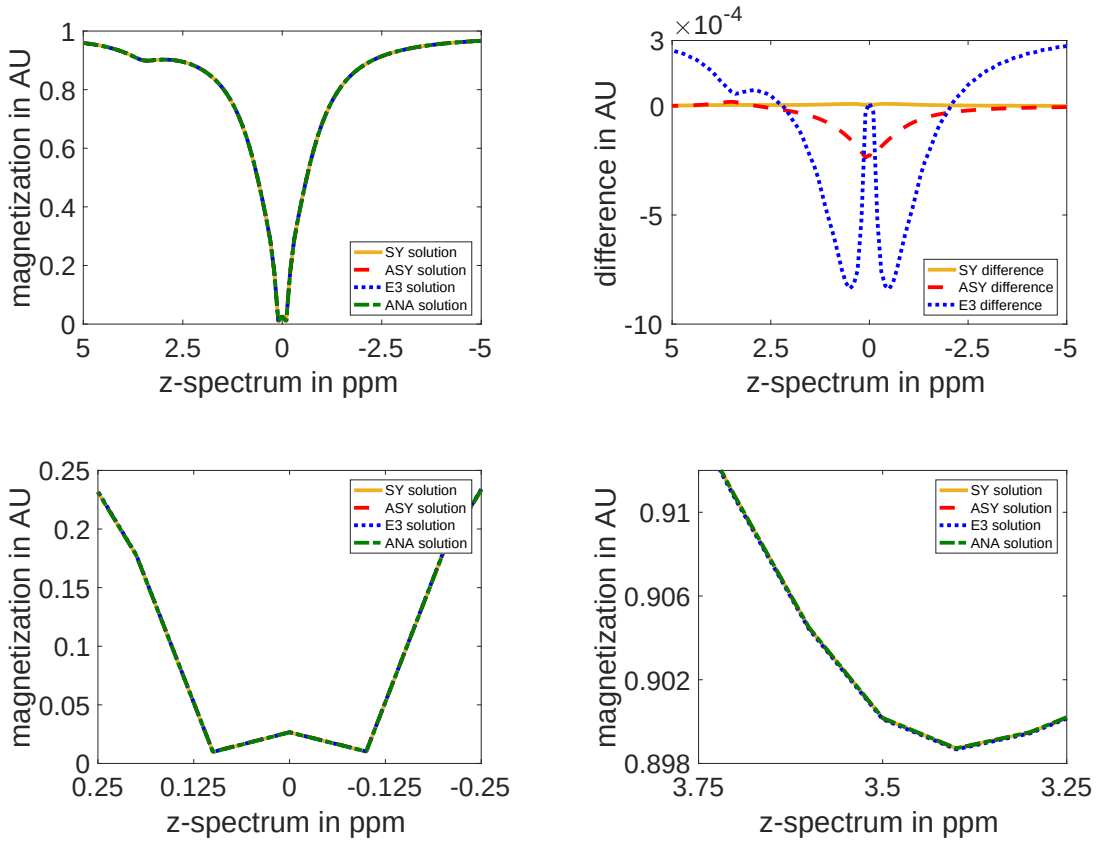

Figure 6: z—magnetization of water pool calculated with G1, where $T_1 \neq \infty, T_2 \neq \infty$, $k \neq 0$. Full view, difference between analytical and numerical solution and two zoom pictures.

operator splitting solver *SY* and a linear order of convergence for the asymmetric operator splitting solver *ASY*. The numerical results in Table 5 verify both orders over a wide range of step sizes.

Table 5: Errors for different time discretizations.

	SY		ASY	
τ	ε_2	EOC	ε_2	EOC
$1e-3$	$2.2e-4$		$1.2e-3$	
$5e-4$	$4.4e-5$	2.18	$5.9e-4$	1.02
$1e-4$	$1.7e-6$	2.02	$1.2e-4$	0.99
$5e-5$	$4.2e-7$	2.02	$5.9e-5$	1.02
$1e-5$	$1.7e-8$	1.99	$1.2e-5$	0.99
$5e-6$	$4.2e-9$	2.02	$5.9e-6$	1.02

1.3.6 Runtime

Since runtime is a vital point of RF pulse optimization, a study is done for the analytic, symmetric and asymmetric operator splitting solvers. For both operator splitting solvers, a vectorized and a non-vectorized implementation is considered.

The RF pulse used for this comparison is the block pulse BL. Relaxation is assumed with $T_{1,w} = T_{1,s} = 1.048s$, $T_{2,w} = 0.069s$ and $T_{2,s} = 0.015s$. The exchange rates are set to $k_{w,s} = 0.25\,\text{Hz}$ and $k_{s,w} = 25\,\text{Hz}$, [24]. At first, a small scale example is executed using $4.2e3$ time points and a time step length of $\tau = 5e-4s$. The z-spectrum $\Omega = [-5, 5]\,\text{ppm}$ is discretized using a step length of 0.1 into 101 parts. For the second example the time discretization is refined to $4.2e5$ time points and a time step length of $\tau = 5e-6s$. For the third example only the z-spectrum is refined instead using a step length of 0.01 into 1001 steps.

Table 6: Runtime in seconds for a different number of points in space N_x and time N_t.

				not vectorized		vectorized	
Example	N_x	N_t	ANA	SY	ASY	SY	ASY
Small scale	101	4.2e3	27.5	2.1	1.9	0.34	0.24
Large scale 1	101	4.2e5	~ 3000	225	218	100	91
Large scale 2	1001	4.2e3	~ 300	20	18	1.82	1.77

Increasing the number of points in time N_t yields a linear increase of runtime of the analytical solver ANA and both non-vectorized numerical solvers SY and ASY. Comparing the runtimes of the small scale example and the first large scale example for the vectorized implementations, see Table 6, a greater than linear increase in runtime can be investigated. However, when we compare the runtime for a different number of points in space N_x, the vectorized solvers SY and ASY show a less than linear increase, and not a linear increase as for the non-vectorized implementations and the analytical solver ANA.
SY is only slightly slower than ASY, and both are much faster than working with the analytical solution. Especially the vectorized implementations of the operator splitting solvers are quite fast and would be a good option for optimization.

Conclusion

For the Bloch equation and the Bloch-McConnell equation, three solvers were derived, an analytical one and two numerical ones. The analytical one allows to investigate the error properties of any numerical solver, but is not suitable for optimization due to an increased runtime.
In contrast, both numerical solvers SY and ASY are quite fast, whereby the asymmetric operator splitting solver is slightly faster. However, the results show that SY prevails with a quadratic order of convergence as well as reduced numerical error.
Therefore, the symmetric operator splitting solver SY is the preferred forward solver for optimal control and is used in the following chapters.

2 OPTIMAL CONTROL MODELING AND THEORY

The following problem, known as the *Problem of the Rocket Car*, see [22] demonstrates the goal of optimal control: Optimizing a certain behavior, where the governing equations are typically differential equations.
Suppose you are driving a car with maximal acceleration 1 and maximal negative acceleration -1, i.e. $-1 \leqslant u(t) \leqslant 1$, where $u(t)$ is the thrust of the car at time t. Your goal is to drive in shortest time from point A to point B. $y(t) \in \mathbb{R}$ denotes the position of the car at time t, m is the mass of the car. $y_0, y_T \in \mathbb{R}$ are the known and desired positions of the car according to point A and point B. Furthermore, the car should be at rest at $t = 0$ and $t = T$ and reach the desired position in minimum time.

$$
\begin{aligned}
&\min_{u \in U_{ad}} J = T, \\
&\text{subject to} \quad m\frac{d^2y}{dt^2}(t) = u(t) \quad \text{in } (0,T), \\
&\qquad\qquad y(0) = y_0, \qquad \frac{dy}{dt}(0) = 0, \qquad y(T) = y_T, \qquad \frac{dy}{dt}(T) = 0, \\
&\qquad\qquad U_{ad} = \{u \in L^2(0,T) \,\big|\, |u(t)| \leqslant 1 \text{ for a. a. } t \in [0,T]\}.
\end{aligned}
$$

This depicts a standard optimal control problem. J is called the objective functional, u is called the control variable and describes the essential degrees of freedom in the optimization. In contrast, y is given as solution of the differential equation. It is called the state variable. Accordingly, $|u(t)| \leqslant 1$ is usually referred to as control constraints. U_{ad} is called the admissible set of controls.

Now, a reasonable *cost functional* J for chemical exchange saturation transfer is introduced, an *admissible set* U_{ad} is constructed, which models technical limitations of the MR–scanner and as *constraint*, the Bloch-McConnell equation is chosen.

2.1 The optimal control problem for CEST RF pulse design

2.1.1 *Derivation of the objective function*

Preliminary Remark: The modeling of the objective function is demonstrated for a water proton pool w at 0ppm and a solute proton pool s at 3.5ppm with exchange rates $k_{s,w} = 25\text{Hz}$ and $k_{w,s} = 0.25\text{Hz}$. The external magnetic field is chosen to be $B_0 = 3\text{T}$ and the gyromagnetic ratio γ is set to $\gamma = 267.513\text{MHz} \cdot \text{T}^{-1}$. Those values correspond to [24]. However, the modeling approach is not restricted to two pool models and can be generalized comfortably to models with more than two pools. It is requested, that the peak of the solute proton pool in the z-spectrum is thin, but deep in order to separate it from the water proton pool and prevent a spill-over effect, which means that the peaks of the water proton pool and the solute proton pool merge. An example of a good z-spectrum is given in Figure 7. It was created using a 0.9s continuous wave RF pulse. However, such a long pulse cannot be implemented in the scanner due to duty cycle constraints of the RF

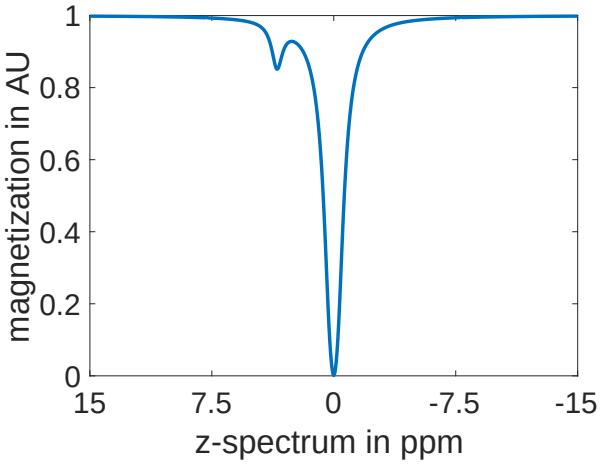

Figure 7: Desired z-spectrum.

amplifiers, which means that the RF pulse cannot be turned on for such a long time. Therefore, an optimization goal is to find a RF pulse, which rebuilds the $z-$spectrum in Figure 7 as good as possible while fulfilling all scanner limitations. In particular we require the RF pulse to be sparse in time, i.e. it should be zero for some points in time.

Let $M_{z,w}$ be the $z-$magnetization of the water proton pool, $M_{z,w,des}$ the desired $z-$ magnetization of the water proton pool as in Figure 7, where the exchange rates are included. A component J_1 of the full cost functional J reads

$$J_1 = \frac{\beta}{q_1} \int_\Omega \left| M_{z,w}(T,s) - M_{z,w,des}(s) \right|^{q_1} ds,$$

where $\beta > 0$ is a weighting parameter and $q_1 \in \mathbb{N}$ specifies the norm used to measure the distance to the target. J_1 should be evaluated at the terminal time T. This approach faces some difficulties, though. Numerical experiments have shown that with this objective it can not be guaranteed, that the RF pulse created rebuilds the CEST$-$peak through magnetization transfer and not through saturation of the RF pulse itself, see Section 3.1.2.
As remedy, not only the magnetization calculated with exchange rates should be tracked, but also the one without as well to guarantee that the desired $z-$spectrum is generated by CEST$-$effect only, and not by direct saturation of the water magnetization around the CEST peak. Let $\widetilde{M_{z,w}}$ be the $z-$magnetization of the water proton pool calculated without exchange rates and $\widetilde{M_{z,w,des}}$ the desired $z-$magnetization of the water proton pool without exchange rates. Now another part J_2 of the full cost functional J is given as

$$J_2 = \frac{\beta_0}{q_2} \int_\Omega \left| \widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s) \right|^{q_2} ds,$$

where $\beta_0 > 0$ and $q_2 \in \mathbb{N}$.
The RF pulse $B_1 = (u_1, u_2)$ is given in polar coordinates, i.e.

$$u_1 = r\cos(\varphi), \quad u_2 = r\sin(\varphi).$$

An advantage of this definition is that the control constraints transform into standard box constraints, see (15). Next, the RF pule undergoes some limitations. First, the pulse power shall be not too large, so

$$J_3 = \frac{\alpha}{2}\int_0^T r(t)^2 dt,$$

where $\alpha > 0$. Second, the RF pulse is desired to be zero for some points in time in order to fulfill the duty cycle constraints of the RF amplifiers. One method of ensuring this is given by sparse control [10]

$$J_4 = \eta \int_0^T |r(t)|\, dt,$$

where $\eta > 0$. Another possible solution of this task is to set $r \equiv 0$ on a given fixed subset of the time interval. Both variants are presented in the numerical experiments below. For the following theory part we set $\eta = 0$. Now the full objective function J reads

$$\begin{aligned} J(r,\varphi) = \frac{\alpha}{2}\int_0^T r(t)^2 dt + \frac{\beta}{q_1}\int_\Omega \left|M_{z,w}(T,s) - M_{z,w,des}(s)\right|^{q_1} ds \\ + \frac{\beta_0}{q_2}\int_\Omega \left|\widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s)\right|^{q_2} ds. \end{aligned}$$

2.1.2 Modeling of the constraints

Due to scanner limitations and physical limits, the RF amplitude is restricted with a maximum amplitude r_{max}

$$\sqrt{u_1^2 + u_2^2} \leqslant r_{max}. \tag{15}$$

Clearly, a transformation of the pulse into polar coordinates r and φ is useful since it transforms this constraint into box constraints

$$0 \leqslant r \leqslant r_{max}, \quad 0 \leqslant \varphi \leqslant 2\pi.$$

Therefore, the admissible set U_{ad} is given as

$$\begin{aligned} U_{ad} = \{&r \in L^\infty(0,T),\ \varphi \in L^\infty(0,T) \mid \\ &0 \leqslant r \leqslant r_{max},\ 0 \leqslant \varphi \leqslant 2\pi \text{ for almost all } t \in (0,T)\}. \end{aligned}$$

Note that as we only consider finite RF pulse duration times, $L^\infty(0,T) \hookrightarrow L^2(0,T)$. Therefore, a $f \in L^\infty(0,T)$ fulfills not only $\|f\|_{L^\infty(0,T)} < \infty$, but also $\|f\|_{L^2(0,T)} < \infty$. Finally, the Bloch-McConnell equation from Section 1.2 has to be fulfilled. Altogether, the optimal control problem can now be defined.

Definition 1 (Optimal control problem for CEST RF pulse design). Let $U = [L^\infty(0,T)]^2$ be the control space, $(r,\varphi) \in U$ the control,

$$M = (M_{x,w}, M_{y,w}, M_{z,w}, M_{x,s}, M_{y,s}, M_{z,s})^T \in [H^1(0,T)]^6$$

the state with exchange, and analogously $\widetilde{M}$ the state without exchange. Then define the objective function $J : U \times H^1(0,T) \times H^1(0,T) \to \mathbb{R}$,

$$J(r,\varphi,M_{z,w},\widetilde{M_{z,w}}) = \frac{\alpha}{2}\int_0^T r(t)^2 dt + \frac{\beta}{q_1}\int_\Omega \left|M_{z,w}(T,s) - M_{z,w,des}(s)\right|^{q_1} ds + \frac{\beta_0}{q_2}\int_\Omega \left|\widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s)\right|^{q_2} ds.$$

Let

$$U_{ad} = \{r \in L^\infty(0,T),\ \varphi \in L^\infty(0,T) \mid 0 \leqslant r \leqslant r_{max},\ 0 \leqslant \varphi \leqslant 2\pi \text{ for almost all } t \in (0,T)\}.$$

Then the *optimal control problem for RF pulse design for CEST imaging* is given as

$$\min_{(r,\varphi)\in U_{ad}} J(r,\varphi,M_{z,w},\widetilde{M_{z,w}}) \quad \text{s.t} \quad \begin{cases} \frac{dM}{dt} = A\cdot M + \tilde{b}, & \frac{d\widetilde{M}}{dt} = \widetilde{A}\cdot\widetilde{M} + \tilde{b}, \\ M(0,z) = M^0(z), & \widetilde{M}(0,z) = M^0(z), \quad z \in \Omega \end{cases} \tag{16}$$

with $A, \tilde{b}$ given by the Bloch-McConnell equation, see (7). $\widetilde{A}$ is A with exchange rates replaced by zero.

2.2 Existence and uniqueness of solutions of the Bloch-McConnell equation

Before having a closer look on the solution of the optimal control problem (16), existence and uniqueness of solutions of the forward Bloch-McConnell equation (7) are shown. For simplicity of representation, this section is written for two pools. Due to the pointwise nature of the Bloch-McConnell equation in spatial direction, it suffices to discuss most of the theory for a fixed spatial position.
We equip the space $[L^2(0,T)]^6$ with the usual norm

$$\|M\|^2_{[L^2(0,T)]^6} = \int_0^T |M|_2^2\, dt,$$

where $|\cdot|_2$ is the euclidean norm in $\mathbb{R}^6$. The norms in $[L^2(0,T)]^2$, $[L^\infty(0,T)]^2$ and $[H^1(0,T)]^6$ are defined analogously.

Theorem 1 (Existence and uniqueness of solutions of the state equation)**.**
Let $M^0 \in \mathbb{R}^6$, $\tilde{b} \in [L^\infty(0,T)]^6$. Let A be defined as in (8), but with $\omega = \omega_0 \in \Omega$ fixed and for 2 pools. The full Bloch-McConnell equation only dependent on the time

$$\begin{cases} \frac{dM}{dt}(t) = A\cdot M(t) + \tilde{b}, \\ M(0) = M^0, \end{cases} \tag{17}$$

admits for every RF pulse $B_1 \in [L^\infty(0,T)]^2$ a unique solution $M(t) \in [H^1(0,T)]^6$.

Before proving this essential theorem, we note that $\|A(t)\|_\infty$ is bounded, since all coefficients therein are bounded due to $B_1 \in [L^\infty(0,T)]^2$, so

$$\sup_{t\in(0,T)} \|A(t)\|_\infty < \infty.$$

It futher holds that $\|\tilde{b}\|_\infty \leqslant \infty$. Therefore, assuming $M \in [L^2(0,T)]^6$, from (17) it can be inferred that $\frac{dM}{dt} \in [L^2(0,T)]^6$, too. Hence, $M \in [H^1(0,T)]^6$ immediately follows. However, we first show the existence $M \in [\mathcal{C}([0,T])]^6$ with Banach's fixed point theorem.

Theorem 2 (Banach's fixed point theorem). *[18]*
Let (X,d) *be a complete metric space,* $M \subset X$ *non-empty and closed. Let* $\phi : M \to M$ *be a contraction with contraction constant smaller than* 1*. Then,* ϕ *admits exactly one fixed point.*

Based on that, Theorem 1 can now be proven.

Proof. $H^1(0,T)$ is densly embedded in $\mathcal{C}([0,T])$ and $(\mathcal{C}([0,T]), \|\cdot\|_\infty)$ is complete. In order to simplify the proof of existence, find solutions in $[\mathcal{C}([0,T])]^6$ in the first place and extend them to $[H^1(0,T)]^6$ afterwards. From (7) it follows that

$$M(\tau) = \underbrace{\int_0^\tau \left(A(t)M(t) + \tilde{b}\right) dt + M^0}_{=:\phi(M)(\tau)} \quad \forall\, \tau \in [0,\tau_1], \tag{18}$$

with $\phi : [\mathcal{C}([0,\tau_1])]^6 \to [\mathcal{C}([0,\tau_1])]^6$. Therefore, (18) yields a fixed point problem $M = \phi(M)$ on $[0,\tau_1]$. The fixed point theorem of Banach is applicable, since $(\mathcal{C}([0,\tau_1]), \|\cdot\|_\infty)^6$ is complete. What we have to show is, that ϕ is a contraction with contraction number K smaller than 1, i.e.

$$\|\phi(M_1) - \phi(M_2)\|_{[\mathcal{C}([0,\tau_1])]^6} \leqslant K \cdot \|M_1 - M_2\|_{[\mathcal{C}([0,\tau_1])]^6}, \quad K < 1.$$

In the first place, the contraction is shown in a small interval $[0,\tau_1]$. Afterwards, it is extended to $[0,T]$. For $[0,\tau_1]$ we get

$$\begin{aligned}
\|\phi(M_1) - \phi(M_2)\|_{[\mathcal{C}([0,\tau_1])]^6} &= \max_{i=1,\cdots 6} \sup_{\tau\in[0,\tau_1]} \left| [\int_0^\tau A(t)(M_1(t) - M_2(t))dt]_i \right| \\
&\leqslant \max_{i=1,\cdots 6} \sup_{\tau\in[0,\tau_1]} \int_0^\tau \underbrace{\|A(t)\|_\infty}_{\leqslant c} |[M_1(t) - M_2(t)]_i|\, dt \\
&\leqslant \left(\max_{i=1,\cdots 6} \sup_{\tau\in[0,\tau_1]} |[M_1 - M_2]_i| \right) \underbrace{c\tau_1}_{=:K}.
\end{aligned}$$

Therein, we would like to get $K < 1$. Note that $c < \infty$ for any fixed time horizon. Therefore, choose the endpoint τ_1 of the first interval as $\tau_1 < \frac{1}{c}$. It holds $\tau_1 > 0$ since $c > 0$. Since $\|A(t)\|_\infty \leqslant c$ holds for all times, we can repeat the step with the

same step length τ_1, such that for every interval $[\tau_1 n, \tau_1(n+1)]$, ϕ is a contraction. As the interval $[0, T]$ is finite, there exists a finite $N \in \mathbb{N}$, such that

$$[0, T] \subset \bigcup_{n=0}^{N} [n\tau_1, (n+1)\tau_1].$$

So ϕ has a unique fixed point in every time interval, yielding a unique solution M of the Bloch-McConnell equation. The continuity of the decomposed solution at the interval transitions is given by construction. Therefore, the full solution, denoted again by M, belongs to $[\mathcal{C}([0, T])]^6$. But as $[\mathcal{C}([0, T])]^6$ is continuously embedded in $[L^2(0, T)]^6$, M belongs to $[L^2(0, T)]^6$ as well. In the end, $M \in [H^1(0, T)]^6$ as noted above. □

Now we know that the Bloch-McConnell equation (17) has a unique solution for a given RF pulse B_1.

2.3 Auxiliary results for adjoint operators

We introduce the adjoint equation pointwise in space for $\omega = \omega_0$

$$\begin{cases} -\frac{dp}{dt} - A^T p = a_Q \quad \text{in } (0, T), \\ p(T) = a_\Omega, \end{cases} \tag{19}$$

where $a_Q \in [L^\infty(0, T)]^6$ and $a_\Omega \in \mathbb{R}^6$. Note that the matrix A is the Bloch-McConnell matrix (8). Further note that the equation is backward-in-time. The following Lemma guarantees that (19) has a solution in $[H^1(0, T)]^6$, too.

Lemma 2 (Solution of the adjoint equation). Under the prerequisites of Theorem 1, the adjoint equation (19) admits a unique solution $p \in [H^1(0, T)]^6$.

Proof. Introduce the inverse time $\tau \in (0, T)$, $\tau = T - t$, and define

$$\begin{aligned} \widetilde{p}(\tau) &= p(T - \tau), \\ \widetilde{A}(\tau) &= A(T - \tau), \\ \widetilde{a}_Q(\tau) &= a_Q(T - \tau). \end{aligned}$$

Due to $\widetilde{p}_\tau = \frac{d\widetilde{p}(\tau)}{d\tau} = \frac{dp(T-\tau)}{d\tau} = -p'(T - \tau)$, (19) is equivalent to the forward differential equation

$$\begin{cases} \widetilde{p}_\tau - \widetilde{A}(\tau)^T \widetilde{p}(\tau) = \widetilde{a}_Q(\tau) \quad \text{for} \quad \tau \in (0, T), \\ \widetilde{p}(0) = \tilde{a}_\Omega. \end{cases} \tag{20}$$

For this equation, Theorem 1 guarantees a unique solution $\widetilde{p} \in [H^1(0, T)]^6$. Retransforming the time again yields the solution $p \in [H^1(0, T)]^6$ of (19). □

For ease of representation, the full Bloch-McConnell matrix A is split into a constant part A_c and two pulse-dependent parts A_{u_1} and A_{u_2} such that

$$A = A_c + A_{u_1} u_1 + A_{u_2} u_2,$$

where we denote the RF pulse as $B_1 = u = (u_1, u_2)$. Then the directional derivative (see Definition 3 below), of A in direction $\partial u = (\partial u_1, \partial u_2)$ is

$$\partial A = A_{u_1} \partial u_1 + A_{u_2} \partial u_2.$$

In the following, we introduce the linearized state equation and draw a connection to the adjoint equation.

Theorem 3 (Auxiliary result)**.** *Let $\partial M \in [H^1(0,T)]^6$ be a solution of*

$$\begin{cases} \partial M_t - \partial A M - A \partial M = 0, \\ \partial M(0) = 0. \end{cases} \tag{21}$$

(21) is called linearized state equation in direction $(\partial u, \partial M)$. Then there holds for all $a_\Omega \in \mathbb{R}^6$, $a_Q \in [L^\infty(0,T)]^6$ and corresponding $p \in [H^1(0,T)]^6$ defined by (19)

$$a_\Omega^T \partial M(T) + \int_0^T a_Q^T \partial M dt = \int_0^T p^T \partial A M dt.$$

Proof. A variational formulation for ∂M with a testfunction p yields

$$\int_0^T p^T \partial M_t dt - \int_0^T p^T \partial A M dt - \int_0^T p^T A \partial M dt = 0,$$
$$\partial M(0) = 0.$$

On the other hand, a variational formulation for p with a testfunction ∂M results in

$$-\int_0^T p_t^T \partial M dt - \int_0^T [A^T p]^T \partial M dt = \int_0^T a_Q^T \partial M dt,$$
$$p(T) = a_\Omega.$$

Partial integration of the first variational formulation leads to

$$p^T \partial M \Big|_0^T - \int_0^T p_t^T \partial M dt - \int_0^T p^T \partial A M dt - \int_0^T p^T A \partial M dt$$
$$= a_\Omega^T \partial M(T) - \int_0^T p_t^T \partial M dt - \int_0^T p^T \partial A M dt - \int_0^T p^T A \partial M dt = 0.$$

With inserting the corresponding terms of the second variational formulation, it follows that

$$a_\Omega^T \partial M(T) + \int_0^T a_Q^T \partial M dt - \int_0^T p^T \partial A M dt = 0.$$

□

Later on, we use this result to obtain a simpler description of the first order optimality conditions.

2.4 First order necessary conditions

Numerical optimization methods are usually based on first order necessary conditions. For the derivation of such optimality conditions, we first need to become acquainted with some generalization of differentiation. Therefore, the concepts of Gâteaux and Fréchet differentiation are introduced.
Let X and Y be Banach spaces, $U \subset X$ open and $F : U \to Y$.

Definition 3 (First variation). [22] If for $u \in U$ and $h \in U$ the limit

$$\delta F(u, h) := \lim_{t \to 0} \frac{1}{t}(F(u + th) - F(u))$$

exists in Y, then it is called directional derivative of F at u in direction h. If this limits exists for all $h \in U$, then the mapping $h \mapsto \delta F(u, h)$ is called first variation of F at u.

Here we don not require linearity of the mapping. On the basis of above definition, we introduce the following.

Definition 4 (Gâteaux differentiability). [22] If the first variation $\delta F(u, h)$ exists at u and is linear, i.e. a linear continuous operator $A : X \to Y$ exists, such that

$$\delta F(u, h) = Ah$$

for all $h \in U$, then F is called Gâteaux differentiable and A Gâteaux derivative of F at u. We write $A = F'(u)$.

Definition 5 (Fréchet differentiability). [22] F is called Fréchet differentiable at $u \in U$, if there exist an operator $A \in \mathcal{L}(X, Y)$ and a mapping $r(u, \cdot) : X \to Y$ with the following properties. For all $h \in X$ with $u + h \in U$ it holds

$$F(u + h) = F(u) + Ah + r(u, h)$$

and for r it holds that

$$\frac{\|r(u, h)\|_Y}{\|h\|_X} \to 0 \quad \text{for} \quad \|h\|_X \to 0.$$

A is called Fréchet derivative of F at u. We write $A = F'(u)$.

The definition of Fréchet differentiability is equivalent to

$$\frac{\|F(u + h) - F(u) - Ah\|_Y}{\|h\|_X} \to 0 \quad \text{for} \quad \|h\|_X \to 0.$$

If F is Fréchet differentiable, it is Gâteaux differentiable as well and the differentials coincide. The next Lemma is the last piece of groundwork we need before we can derive optimality conditions.

Lemma 3 (Chain rule). [22] Let X, Y and Z be Banach spaces, $U \subset X$, $V \subset Y$ open and $F : U \to V$ as well as $G : V \to Z$ at $u \in U$ and $F(u) \in V$ Fréchet differentiable mappings. Then $E = G \circ F$ defined by $E(u) = G(F(u))$ is Fréchet differentiable at u and

$$E'(u) = G'(F(u))F'(u).$$

The following Theorem is the basis for the derivation of optimality conditions.

Theorem 4. *[22] Let X be a real Banach space, $U \subset X$ a convex set and f a real Gâteaux differentiable functional on an open subset of X, which encloses U. $\overline{u} \in U$ shall be a solution to*

$$\min_{u \in U} f(u),$$

so

$$f(\overline{u}) \leqslant f(u), \quad \forall u \in U.$$

Then the following variational inequality holds:

$$f'(\overline{u})(u - \overline{u}) \geqslant 0 \quad \forall u \in U. \tag{22}$$

On the other hand, let $\overline{u} \in U$ be a solution to the variational inequality (22) and f additionally convex, then $\overline{u}$ is a global solution of $\min_{u \in U} f(u)$.

Now, a line to the optimal control problem (16) is drawn. First, the solution operator is defined. Afterwards, it is shown that this operator is Fréchet differentiable, which allows the use of the chain rule. Last, Theorem 4 is applied.

Remark 1. Due to Theorem 1, there exists for every control $u \in [L^\infty(0,T)]^2$ a solution $M \in [H^1(0,T)]^6$ of (7). Therefore, define the solution operator G as

$$G : \begin{cases} [L^\infty(0,T)]^2 & \to [H^1(0,T)]^6, \\ u & \mapsto M. \end{cases}$$

Notation 1. In the following, $A(u)$ denotes the Bloch-McConnell matrix (8) which contains the RF pulse u. $M(u)$ is the magnetization satisfying the Bloch-McConnell equation (7) with the coefficient matrix $A(u)$.

Theorem 5. *The operator G is Fréchet differentiable.*

Before starting with proving Theorem 5, we repeat the *linearized state equation* in a point (M, u) in direction ∂u

$$\begin{cases} \partial \dot{M} = A(u)\partial M + \partial A(\partial u)M(u), \\ \partial M(0) = 0, \end{cases} \tag{23}$$

where $\partial A(\partial u)$ is the directional derivative of A with respect to u in direction ∂u. For $\partial u = u - \tilde{u}$ it holds $\partial A(\partial u) = A(u) - A(\tilde{u})$.

Lemma 4. The solution ∂M_φ of the linearized state equation (23) fulfills

$$\|\partial M(\varphi)\|_{[H^1(0,T)]^6} \leqslant c\|u - \tilde{u}\|_{[L^2(0,T)]^2}$$

with $c > 0$. Here u and $\tilde{u}$ are in $[L^\infty(0,T)]^2$, $\varphi = u - \tilde{u}$.

Before proving Lemma 4, we repeat the Lemma of Gronwall in its integral form.

Theorem 6 (Gronwall's Lemma). *[1]*
Let ζ, φ and χ be real, continuous functions, $\chi(t) \geqslant 0$ for all $t \in [a, b]$. If

$$\zeta(t) \leqslant \varphi(t) + \int_a^t \chi(s)\zeta(s)\,ds, \quad t \in [a, b]$$

then

$$\zeta(t) \leqslant \varphi(t) + \int_a^t \chi(s)\varphi(s) \exp\left(\int_s^t \chi(u)\,du\right) ds, \quad t \in [a, b].$$

Proof of Lemma 4. Step 1.
Testing the linearized state equation (23) with $\partial M(\varphi)$ itself and integrating yields for $t \in [0, T]$

$$\begin{aligned}\int_0^t \partial M(\varphi)^T \frac{d\partial M(\varphi)}{dt} ds &= \int_0^t \partial M(\varphi)^T A(u)\partial M(\varphi) ds \\ &\quad + \int_0^t \partial M(\varphi)^T [A(u) - A(\widetilde{u})] M(u) ds.\end{aligned}$$

Note that

$$\int_0^t \partial M(\varphi)^T \frac{d\partial M(\varphi)}{dt} ds = \frac{1}{2} \int_0^t \frac{d}{dt} |\partial M(\varphi)|_2^2 \, ds = \frac{1}{2} |\partial M(\varphi)(t)|_2^2 .$$

$|\cdot|_2$ is the euclidean vector norm in $\mathbb{R}^6$. With a generic constant $c > 0$ it follows that

$$\begin{aligned}\frac{1}{2} |\partial M(\varphi)(t)|_2^2 &= \int_0^t \partial M(\varphi)^T A(u)\partial M(\varphi) ds + \int_0^t \partial M(\varphi)^T (A(u) - A(\widetilde{u})) M(u) ds \\ &\leqslant \int_0^t \left| \partial M(\varphi)^T A(u)\partial M(\varphi) \right|_2 ds + \int_0^t \partial M(\varphi)^T (A(u) - A(\widetilde{u})) M(u) ds \\ &\leqslant \int_0^t |\partial M(\varphi)|_2^2 \, \|A(u)\|_2 ds + \int_0^t \partial M(\varphi)^T (A(u) - A(\widetilde{u})) M(u) ds \\ &\leqslant c \underbrace{\|A(u)\|_{L^\infty(0,t;\mathbb{R}^6)}}_{\leqslant c} \int_0^t |\partial M(\varphi)|_2^2 \, ds + \int_0^t \partial M(\varphi)^T (A(u) - A(\widetilde{u})) M(u) ds.\end{aligned}$$

Note that as $u \in [L^\infty(0, T)]^2$ it holds that $\|A(u)\|_{L^\infty(0,t;\mathbb{R}^6)} \leqslant c$.

Using Young's inequality [6] with $p = q = 2$ yields,

$$\int_0^t \partial M(\varphi)^T [A(u) - A(\widetilde{u})] M(u) ds \leqslant \int_0^t |\partial M(\varphi)|_2 \, |(A(u) - A(\widetilde{u}))M(u)|_2 \, ds$$
$$\leqslant \int_0^t \left[\frac{|\partial M(\varphi)|_2^2}{2} + \frac{|(A(u) - A(\widetilde{u}))M(u)|_2^2}{2} \right] ds.$$

Altogether we obtain

$$|\partial M(\varphi)(t)|_2^2 \leqslant c \int_0^t |\partial M(\varphi)(s)|_2^2 \, ds + \int_0^t |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, ds. \tag{24}$$

Step 2. The Lemma of Gronwall (Theorem) is now applied to (24), so

$$\frac{1}{2} |\partial M(\varphi)(t)|_2^2 \leqslant \int_0^t |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, ds \tag{25}$$
$$+ c \int_0^t \int_0^s |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, d\tau \underbrace{\exp\left(\int_s^t c \, dr \right)}_{\leqslant c} ds$$
$$\leqslant \int_0^t |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, ds + c \int_0^t \underbrace{\int_0^s |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, d\tau}_{\leqslant \int_0^t |(A(u) - A(\widetilde{u}))M(u)|_2^2 d\tau \text{ as } s \leqslant t} ds$$
$$\leqslant c \int_0^t |(A(u) - A(\widetilde{u}))M(u)|_2^2 \, ds \tag{26}$$

with a generic constant $c > 0$.

Step 3. Equation (26) is further estimated. Therefore define

$$N_1 = N_2 = \begin{pmatrix} 0 & 0 & (u - \widetilde{u})_1 \\ 0 & 0 & (u - \widetilde{u})_2 \\ -(u - \widetilde{u})_1 & -(u - \widetilde{u})_2 & 0 \end{pmatrix}$$

and

$$N := A(u) - A(\widetilde{u}) = \begin{pmatrix} N_1 & \\ & N_2 \end{pmatrix}.$$

The Frobenius norm $\|A\|_F = \left(\sum_i \sum_j |A_{ij}|^2 \right)^{\frac{1}{2}}$ is compatible with the L^2–norm and it can be easily evaluated that

$$\|N\|_F^2 = 4 \, |(u - \widetilde{u})_1|^2 + 4 \, |(u - \widetilde{u})_2|^2 = 4 \, |u - \widetilde{u}|_2^2 .$$

Therefore,

$$\begin{aligned}\|(A(u)-A(\widetilde{u}))M(u)\|^2_{[L^2(0,T)]^6} =& \|NM(u)\|^2_{[L^2(0,T)]^6} = \int_0^T |NM(u)|_2^2\,dt\\ &\leqslant \int_0^T \|N\|_F^2\,|M(u)|_2^2\,dt \leqslant \sup_{t\in(0,T)} |M(u)(t)|_2^2 \int_0^T \|N\|_F^2 dt\\ &= \underbrace{\sup_{t\in(0,T)} |M(u)(t)|_2^2}_{\leqslant c} \int_0^T 4\,|u-\tilde{u}|_2^2\,dt \leqslant c\|u-\widetilde{u}\|^2_{[L^2(0,T)]^2}.\end{aligned}$$

So

$$\|\partial M(\varphi)(t)\|^2_{[L^2(0,T)]^2} \leqslant c\|u-\widetilde{u}\|^2_{[L^2(0,T)]^2}.$$

Step 4.
The estimation for $\frac{d}{dt}\partial M(\varphi)$ is derived from the following.

$$\begin{aligned}\|\frac{d}{dt}\partial M(\varphi)\|^2_{[L^2(0,T)]^6} &= \|A(u)\partial M(\varphi) + [A(u)-A(\widetilde{u})]M(u)\|^2_{[L^2(0,T)]^6}\\ &\leqslant \|A(u)\partial M(\varphi)\|^2_{[L^2(0,T)]^6} + \|[A(u)-A(\widetilde{u})]M(u)\|^2_{[L^2(0,T)]^6}\\ &= \int_0^T |A(u)\partial M(\varphi)|_2^2\,dt + \underbrace{\|[A(u)-A(\widetilde{u})]M(u)\|^2_{[L^2(0,T)]^6}}_{\text{Step 3}}.\end{aligned}$$

We investigate that

$$\begin{aligned}\int_0^T |A(u)\partial M(\varphi)|_2^2\,dt &\leqslant \int_0^T \|A(u)\|_F^2\,|\partial M(\varphi)|_2^2\,dt\\ &\leqslant \sup_{t\in(0,T)} \|A(u)(t)\|_F^2 \|\partial M(\varphi)\|^2_{[L^2(0,T)]^6} \overset{\text{Step 3}}{\leqslant} c\|u-\widetilde{u}\|^2_{[L^2(0,T)]^2}.\end{aligned}$$

The matrix $A(u)$ is not only bounded in $\|\cdot\|_\infty$, but in $\|\cdot\|_F$ as well, which allowed the last inequality. Therefore,

$$\|\frac{d}{dt}\partial M(\varphi)\|^2_{[L^2(0,T)]^6} \leqslant c\|u-\widetilde{u}\|^2_{[L^2(0,T)]^2}.$$

Step 5. Finally, everything is put together yielding

$$\|\partial M(\varphi)\|^2_{[H^1(0,T)]^6} = \|\partial M(\varphi)\|^2_{[L^2(0,T)]^6} + \|\frac{d}{dt}\partial M(\varphi)\|^2_{[L^2(0,T)]^6} \leqslant c\|u-\widetilde{u}\|^2_{[L^2(0,T)]^2}.$$

□

Now Theorem 5 can be proven.

Proof of Theorem 5. Let $u \in U$ and $U \subset [L^\infty(0,T)]^2$ a suitable open neighborhood. Let $\widetilde{u} \in U$ with $\|\widetilde{u} - u\|_{[L^2(0,T)]^2} \leqslant \delta$, $\varphi := \widetilde{u} - u$ and $\delta > 0$. Let $\partial M(\varphi)$ be the solution to the linearized state equation

$$\begin{cases} \partial \dot{M}(\varphi) = A(u)\partial M(\varphi) + \partial A(\partial u)M(u), \\ \partial M(\varphi)(0) = 0, \end{cases} \tag{27}$$

where $\partial A(\partial u) = A(\widetilde{u}) - A(u)$.
The solution $\partial M(\varphi)$ exists in $[H^1(0,T)]^6$ due to Theorem 1. Then

$$\begin{aligned} \frac{d}{dt}(\underbrace{M(\widetilde{u}) - M(u) - \partial M(\varphi)}_{=:S}) &= A(\widetilde{u})M(\widetilde{u}) - A(u)M(u) \\ &\quad - A(u)\partial M(\varphi) - (A(\widetilde{u}) - A(u))M(u) \\ &= A(\widetilde{u})(\underbrace{M(\widetilde{u}) - M(u) - \partial M(\varphi)}_{=:S}) \\ &\quad + \underbrace{(A(\widetilde{u}) - A(u))\partial M(\varphi)}_{=:b}. \end{aligned}$$

Therefore we consider the problem

$$\begin{cases} \frac{dS}{dt} = A(\widetilde{u})S + b, \\ S(0) = 0. \end{cases} \tag{28}$$

Now test (28) with S itself and integrate

$$\int_0^t S^T \frac{d}{dt} S ds = \int_0^t S^T A(\widetilde{u}) S ds + \int_0^t S^T b ds.$$

Note that

$$\int_0^t S^T \frac{d}{dt} S ds = \frac{1}{2} \int_0^t \frac{d}{dt} |S|_2^2 \, ds = \frac{1}{2} |S(t)|_2^2,$$

where we used the euclidean vector norm. With a generic constant $c > 0$ it follows

$$\begin{aligned} \frac{1}{2} |S(t)|_2^2 &\leqslant c \|A(\widetilde{u})\|^2_{L^\infty(0,t;\mathbb{R}^6} \int_0^t |S|_2^2 \, ds + \int_0^t S^T b ds \\ &\leqslant c \int_0^t |S|_2^2 \, ds + \int_0^t |S|_2 \, |b|_2 \, ds \\ &\leqslant c \int_0^t |S|_2^2 \, ds + \int_0^t \left(\frac{|S|_2^2}{2} + \frac{|b|_2^2}{2} \right) ds, \end{aligned}$$

where in the last line Young's inequality [6] was applied. Therefore

$$|S(t)|_2^2 \leqslant c \int_0^t |S|_2^2 \, ds + \int_0^t |b|_2^2 \, ds.$$

The Lemma of Gronwall yields

$$\begin{aligned}
|S(t)|_2^2 &\leqslant \int_0^t |b(s)|_2^2 \, ds + \int_0^t c \int_0^s |b(v)|_2^2 \, dv \underbrace{\exp\left(\int_s^t c \, dv\right)}_{\leqslant c} ds \\
&\leqslant \int_0^t |b(s)|_2^2 \, ds + c \int_0^t \underbrace{\int_0^s |b(v)|_2^2 \, dv}_{\leqslant \int_0^t |b(v)|_2^2 dv \text{ as } s \leqslant t} ds \\
&\leqslant c \int_0^t |b(s)|_2^2 \, ds.
\end{aligned}$$

Now we resubstitute and get

$$\begin{aligned}
|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)|_2^2 &\leqslant c \int_0^t |(A(\widetilde{u}) - A(u))\partial M(\varphi)|_2^2 \, ds \\
&\leqslant c \int_0^t \|A(\widetilde{u}) - A(u)\|_F^2 \, |\partial M(\varphi)|_2^2 \, ds \\
&\leqslant c \int_0^t 4 \, |u - \widetilde{u}|_2^2 \, |\partial M(\varphi)|_2^2 \, .
\end{aligned}$$

Those steps were done analogously to Step 3, Proof of Lemma 4. We further investigate that

$$\begin{aligned}
c \int_0^t |u - \widetilde{u}|_2^2 \, |\partial M(\varphi)|_2^2 &\leqslant c \|\partial M(\varphi)\|_{[L^\infty(0,T)]^6}^2 \|u - \widetilde{u}\|_{[L^2(0,T)]^2}^2 \\
&\leqslant c \|\partial M(\varphi)\|_{[H^1(0,T)]^6}^2 \|u - \widetilde{u}\|_{[L^2(0,T)]^2}^2.
\end{aligned}$$

The last inequality is a consequence of Morrey's Theorem [3]. With the result of Lemma 4 it finally follows that

$$|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)|_2^2 \leqslant c \|\widetilde{u} - u\|_{[L^2(0,T)]^2}^4.$$

Putting everything together yields

$$
\begin{aligned}
&\|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)\|^2_{[H^1(0,T)]^6} \\
&= \|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)\|^2_{[L^2(0,T)]^6} \\
&+ \|\frac{d}{dt}(M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t))\|^2_{[L^2(0,T)]^6} \\
&\leqslant c\|\widetilde{u} - u\|^4_{[L^2(0,T)]^2} + \|A(\widetilde{u})(M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t))\|^2_{[L^2(0,T)]^6} \\
&+ \|(A(\widetilde{u}) - A(u))\partial M(\varphi)\|^2_{[L^2(0,T)]^6} \\
&\leqslant c\|\widetilde{u} - u\|^4_{[L^2(0,T)]^2} + \underbrace{\sup_{t\in(0,T)} \|A(\widetilde{u})(t)\|_F^2}_{\leqslant c} \|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)\|^2_{[L^2(0,T)]^6} \\
&+ \underbrace{\|\partial M(\varphi)\|^2_{[L^\infty(0,T)]^6}}_{\leqslant \|\partial M(\varphi)\|^2_{c[H^1(0,T)]^6}} \|\widetilde{u} - u\|^2_{[L^2(0,T)]^2}.
\end{aligned}
$$

With the previous step and Lemma 4 it follows that

$$\|M(\widetilde{u})(t) - M(u)(t) - \partial M(\varphi)(t)\|^2_{[H^1(0,T)]^6} \leqslant c\|\widetilde{u} - u\|^4_{[L^2(0,T)]^2}.$$

In conclusion,

$$\|G(\widetilde{u}) - G(u) - G'(u)(\widetilde{u} - u)\|_{[H^1(0,T)]^6} \leqslant c\|\widetilde{u} - u\|^2_{[L^2(0,T)]^6} \leqslant c\|\widetilde{u} - u\|^2_{[L^\infty(0,T)]^6},$$

due to $L^\infty(0,T) \hookrightarrow L^2(0,T)$. As $\|\widetilde{u} - u\|^2_{[L^\infty(0,T)]^6} \leqslant \delta$,

$$\|G(\widetilde{u}) - G(u) - G'(u)(\widetilde{u} - u)\|_{[H^1(0,T)]^6} \leqslant c\delta^2.$$

This implies the claimed Fréchet differentiability. □

Now, the first order necessary conditions for (16) can be derived. Let p be the solution to the adjoint equation (19) with

$$a_Q = 0,\ a_\Omega = (0, 0, \beta(M_{z,w}(T,s) - M_{z,w,des}(s))^{q_1 - 1}, 0, 0, 0)^T.$$

Let analogously $\widetilde{p}$ be the solution to the adjoint equation (19) with

$$\widetilde{a_Q} = 0,\ \widetilde{a_\Omega} = (0, 0, \beta_0(\widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s))^{q_2 - 1}, 0, 0, 0)^T$$

and zero exchange rates. Then, Theorem 3 yields

$$p^T(T)\partial M(T) = \int_0^T p^T \partial A M dt,$$

$$\widetilde{p}^T(T)\partial \widetilde{M}(T) = \int_0^T \widetilde{p}^T \partial \widetilde{A}\widetilde{M} dt.$$

Therein, $\partial\widetilde{M}$ is the tangent state defined with zero exchange rates. We rewrite the L^2- regularization of the RF pulse in Cartesian coordinates as

$$\frac{\alpha}{2}\int_0^T r(t)^2 dt = \frac{\alpha}{2}\int_0^T u_1(t)^2 dt + \frac{\alpha}{2}\int_0^T u_2(t)^2 dt.$$

Using the chain rule Lemma 3, differentiation of the cost functional

$$\begin{aligned} J(u_1, u_2, M_{z,w}, \widetilde{M_{z,w}}) = \frac{\alpha}{2}\int_0^T u_1(t)^2 dt + \frac{\alpha}{2}\int_0^T u_2(t)^2 dt \\ + \frac{\beta}{q_1}\int_\Omega \left|M_{z,w}(T,s) - M_{z,w,des}(s)\right|^{q_1} ds \\ + \frac{\beta_0}{q_2}\int_\Omega \left|\widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s)\right|^{q_2} ds \end{aligned}$$

in direction $h = (\partial u_1, \partial u_2)$ yields

$$\begin{aligned} J'(u_1, u_2, M_{z,w}, \widetilde{M_{z,w}})h &= \alpha\int_0^T u_1 \partial u_1 dt + \alpha\int_0^T u_2 \partial u_2 dt \\ &+ \beta\int_\Omega \left|M_{z,w}(T,s) - M_{z,w,des}(s)\right|^{q_1-1} \partial M ds \\ &+ \beta_0\int_\Omega \left|\widetilde{M_{z,w}}(T,s) - \widetilde{M_{z,w,des}}(s)\right|^{q_2-1} \partial\widetilde{M} ds \\ &= \alpha\int_0^T u_1 \partial u_1 dt + \alpha\int_0^T u_2 \partial u_2 dt \\ &+ \int_\Omega\int_0^T p^T \partial A M dt ds + \int_\Omega\int_0^T \widetilde{p}^T \partial\widetilde{A}\widetilde{M} dt ds, \end{aligned}$$

where in the last line the auxiliary result 3 was used. Now we define as before $\partial A = A_{u_1}\partial u_1 + A_{u_2}\partial u_2$ and $\partial\widetilde{A} = A_{u_1}\partial u_1 + A_{u_2}\partial u_2$ and note, that ∂A and $\partial\widetilde{A}$ are equal.

Using that, we further investigate

$$
\begin{aligned}
J'(u_1, u_2, M_{z,w}, \widetilde{M_{z,w}})h &= \alpha \int_0^T u_1 \partial u_1 \, dt + \alpha \int_0^T u_2 \partial u_2 dt \\
&+ \int_\Omega \int_0^T p^T (A_{u_1} \partial u_1 + A_{u_2} \partial_{u_2}) M dt ds + \int_\Omega \int_0^T \widetilde{p}^T (A_{u_1} \partial u_1 + A_{u_2} \partial_{u_2}) \widetilde{M} dt ds \\
&= \int_0^T \left(\left(\alpha u_1 + \int_\Omega p^T A_{u_1} M ds + \int_\Omega \widetilde{p}^T A_{u_1} \widetilde{M} ds \right) \partial u_1 \right. \\
&+ \left. \left(\alpha u_2 + \int_\Omega p^T A_{u_2} M ds + \int_\Omega \widetilde{p}^T A_{u_2} \widetilde{M} ds \right) \partial u_2 \right) dt.
\end{aligned}
$$

Therefore, we can identify the reduced gradient $g(u_1, u_2)$ of J as

$$
g(u_1, u_2) = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} \alpha u_1 + \int_\Omega p^T A_{u_1} M ds + \int_\Omega \widetilde{p}^T A_{u_1} \widetilde{M} ds \\ \alpha u_2 + \int_\Omega p^T A_{u_2} M ds + \int_\Omega \widetilde{p}^T A_{u_2} \widetilde{M} ds \end{pmatrix}. \tag{29}
$$

Now we assume the existence of a minimizer $(\overline{u}_1, \overline{u}_2)$. Then, according to Theorem 4, it must satisfy

$$
g(\overline{u}_1, \overline{u}_2) \begin{pmatrix} u_1 - \overline{u}_1 \\ u_2 - \overline{u}_2 \end{pmatrix} \geqslant 0 \quad \forall \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in U_{ad}.
$$

In practice, we have to start with transforming the RF pulse (r, φ) into cartesian coordinates (u_1, u_2). Then the gradient should be calculated as above. Afterwards, retransformation of the RF pulse into polar coordinates yields the reduced gradient in polar coordinates

$$
\hat{g}(r, \varphi) = \begin{pmatrix} g_1 \cos(\varphi) + g_2 \sin(\varphi) \\ -g_1 r \sin(\varphi) + g_2 r \cos(\varphi) \end{pmatrix}.
$$

3 NUMERICAL OPTIMIZATION AND EXPERIMENTS

In the preceding chapter, the fundament for numerical optimization was laid. An optimal control model was designed and analyzed. First order optimality conditions were derived, that form the basis for numerical optimization. Therefore, this chapter can now proceed with the numerical solution of different application problems. First, a sophisticated desired state is derived for a 2 pool model from literature. The cost functional is further analyzed and adapted compared to Section 2.1 and different initializations for optimization are compared. Optimization runs using the optimization method of [17] are performed, see Section 3.1.3. The results are discussed. Then, after the optimization, the robustness of the optimized RF pulse with respect to relaxation times, exchange rates and frequency offsets is investigated in numerical experiments. Afterwards, the optimization model and method are applied to a second 2 pool model with values that are directly related to an experiment on an MR scanner, and allow for phantom measurements.

3.1 Optimization of a 2 pool model

The input data used for this comparison coincides with the model problem which was used in [24], where the parameters originate from an amide proton system in white matter. Due to the specifications of the MR scanner (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany), the external magnetic field was chosen to be $B_0 = 3T$. The longitudinal relaxation times are $T_{1,w} = T_{1,s} = 1.048s$ and the transversal relaxations times are $T_{2,w} = 0.069s$ and $T_{2,s} = 0.015s$. Magnetization exchange is modeled by the frequency offset of the RF irradiation $\omega_w = 0ppm$ and $\omega_s = 3.5ppm$ and exchange rates $k_{ws} = 0.25Hz$ and $k_{sw} = 25Hz$. The normalized equilibrium magnetizations are set to $M_{0,w} = 1$ and $M_{0,s} = 0.01$. The $z-$spectrum $\Omega = [-15, 15]ppm$ is discretized into 601 parts using a step length of $0.05ppm$. The duration of the RF$-$pulse is $0.9s$.

3.1.1 Determination of a desired state

As already mentioned in Section 2.1, the peak of the solute proton pool in the $z-$spectrum is required to be thin, but deep. A good candidate RF pulse for generating such a $z-$ spectrum is a continuous wave RF pulse with a long duration, here $0.9s$. What needs to be calibrated is the amplitude of this pulse. Due to limitations of the RF amplifiers, it is generally restricted with $2.5\mu T$. Therefore, several pulses varying in amplitude are compared. The comparison was done using a step length of $\Delta t = 1e-4s$.
Figure 8 shows several $z-$spectra calculated with different values of B_1. While the yellow curve with $B_1 = 0.25\mu T$ shows a very thin and distinct CEST$-$peak in the $z-$spectrum (left picture), the larger B_1-amplitude of $0.58\mu T$ prevails clearly in the difference between $z-$magnetization of the water proton pool with and without exchange M_{diff}. Therefore, the latter value is chosen as amplitude for generation

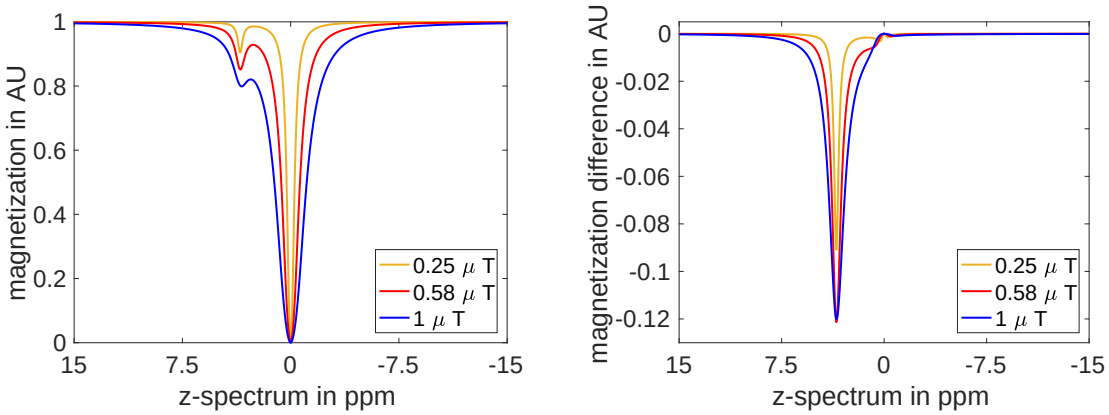

Figure 8: Desired z-spectrum and magnetization difference $M_{diff} = M_{z,w} - \widetilde{M_{z,w}}$ for different values of B_1, where M_{diff} arises from the difference of the z−magnetization of the water proton pool w with and without exchange rates.

of the desired state. This value was experimentally determined and corresponds with the value derived in [21].

Time discretization

For model validity, and also for accurate numerical solutions with the symmetric operator splitting solver, see Section 1.3.5, a minimum sampling rate (corresponding to a maximum time step size) has to be prescribed and met. In particular, the *Nyquist-Shannon sampling theorem* must be regarded for practical application of results.

Theorem 7 (Nyquist-Shannon sampling theorem). *[4] If a function* $x(t)$ *contains no frequencies higher than* B *hertz, it is completely determined by giving its ordinates at a series of points spaced* $1/(2B)$ *seconds apart.*

Reformulating this theorem for our problem yields a minimum sampling rate of

$$\Delta t < \frac{1}{2 \max_{zSpec} B_0 \frac{\gamma}{2\pi}}$$

to guarantee good reconstruction, where $\max_{zSpec} = |\Omega|$. For our proposed example, this leads to

$$\Delta t < \frac{1}{2 \cdot 30 \cdot 3 \frac{267.5129 \cdot 10^6}{2\pi}} s \approx 1.30 \cdot 10^{-4} s.$$

However, such a short sampling rate results in time-consuming optimization runs. Therefore, we use a sampling rate of $\Delta t = 1e-3s$ for analyzing the cost functional, where some moderate irregularities within the z-spectrum can be expected due to the Nyquist Theorem 7. After the deduction of the right functional and optimization parameters, the sampling rate is reduced to $\Delta t = 1e-4s$ and optimal RF pulses for practical application are generated.

3.1.2 Investigation of different cost functionals

In the following, we take a closer look on the cost functional and explain its form. Three different cost functionals are investigated. First, only a simpler version of the full functional J, which only uses tracking of the $z-$magnetization of the water proton pool and L^2-regularization of the RF pulse is considered. Second, tracking is applied to the difference between $z-$magnetization of the water proton pool with and without exchange M_{diff}. Third, the full cost functional J, which was derived in Section 2.1 is utilized in a slightly modified form. It is shown, that this choice fulfills the needs of Chemical Exchange Saturation Transfer problems best. For the first two cases a block-type pulse train is chosen as initial for optimization with an amplitude of $B_1 = 0.58\mu T$. The terminal time is set to $0.9s$ with a step length of $\Delta t = 1e-3s$ yielding 900 time discretization points.

Tracking of water magnetization with exchange

We consider the cost functional J as

$$J_1(r,\varphi) = \frac{\alpha}{2}\int_0^T r(t)^2 dt + \frac{\beta}{q_1}\int_\Omega \left|M_{z,w}(T,s) - M_{z,w,des}(s)\right|^{q_1} ds.$$

Typical regularization parameters are chosen as $\alpha = 10^{-2}$ and $\beta = 10^2$. Optimization with $q_1 = 2$ yields an RF pulse, which results in the $z-$spectra in Figure 9.

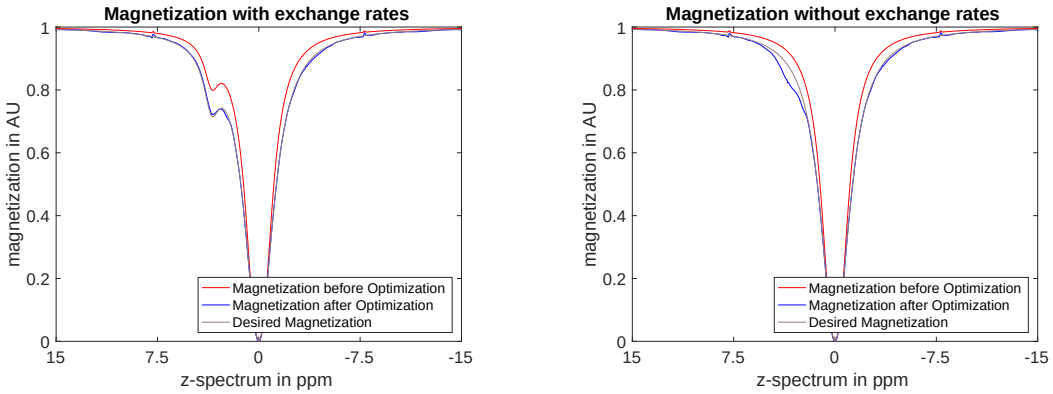

Figure 9: $z-$spectra using J_1 calculated with and without exchange rates.

In the left plot of Figure 9, the optimized $z-$magnetization of the water proton pool w calculated with exchange rates seems to be excellent, the desired magnetization was rebuild almost perfectly. However, the $z-$magnetization calculated without exchange, see the right plot, is perturbed around the solute proton pool at $3.5ppm$. This behavior is undesired, the water magnetization without exchange should not deviate from its desired magnetization. Otherwise, it can not be guaranteed, that the CEST$-$peak created with exchange stems from magnetization exchange and not from saturation.
To overcome this issue, a modified cost functional is presented below, which uses tracking of the difference of the $z-$magnetization of the water proton pool calculated with and without magnetization exchange in order to prevent this undesired direct saturation.

Tracking of a difference

For the next two objectives, the dynamical system is doubled in its size. Additionally to the Bloch-McConnell equation, also the system with zero exchange rates (corresponding to the Bloch equation per pool) is included. The difference of the $z-$magnetization of the water proton pool, which was calculated with exchange rates and the one without exchange rates, named M_{diff}, is then defined as

$$M_{diff}(T,s) = M_{z,w}(T,s) - \widetilde{M_{w,z}}(T,s), \quad s \in \Omega.$$

In Section 3.1.1 it is mentioned, why this value is important for Chemical Exchange. Intuitively, one would assume that tracking of the M_{diff} towards a desired difference $M_{diff,des}$ is a reasonable task for optimization and prevents the problems, which occurred with the previous cost functional. Therefore, define a cost functional J_2 as

$$J_2(r,\varphi) = \frac{\alpha}{2}\int_0^T r(t)^2 dt + \frac{\delta}{q_2}\int_\Omega \left|M_{diff}(T,s) - M_{diff,des}(s)\right|^{q_2} ds.$$

Again, L^2-regularization is included. The norm is defined with $q_2 = 2$, the weighting parameters are set to $\alpha = 10^{-2}$ and $\delta = 10^2$. Then optimization results in the $z-$spectra in Figure 10.

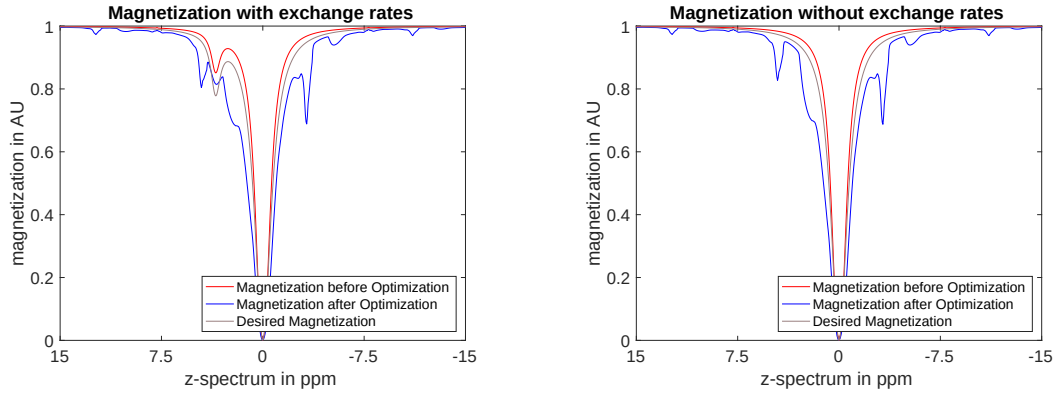

Figure 10: $z-$spectra using J_2 calculated with and without exchange rates.

What we can conclude is that all the instabilities occur in both the $z-$spectra with and without exchange simultaneously, since the cost functional J_2 does not penalize a gap between the desired magnetizations, but only between the differences, which cancel away here. Therefore, tracking of M_{diff} alone is also not a suitable optimization task.

In contrast, the cost functional J introduced in Section 2.1 overcomes these problems and that is exactly why it was proposed.

Best approximation in L^∞*-norm*

Due to the preceding results, compared to Section 2.1, a slightly modified cost functional J is introduced, which includes L^∞−tracking of the water proton pool, as this approach allows to set taylored boundaries for the solution. Also L^2−regularization of the RF pulse is included. This approach will overcome the issues that occurred in the previous cost functionals. On the one hand, L^∞−tracking of the desired z−spectrum with exchange will guarantee a visible CEST−peak after optimization. On the other hand, L^∞−tracking of the z−spectrum without exchange as well will prevent unwanted magnetization saturation around the CEST−peak. Also, the presented tracking of M_{diff} is implicitly included in this approach, but without allowing to rebuild the instabilities which we faced above.

Due to [22], Section 6.1.2, we know that the problem of pure minimization

$$\min_{u\in U}\|y-y_d\|_{L^\infty(\Omega_1)}+\|y-y_d\|_{L^\infty(\Omega_2)}$$

is equivalent to the problem with pointwise state constraints

$$\begin{aligned}\min_{u\in U,\varepsilon_1,\varepsilon_2\geqslant 0}\ & \varepsilon_1+\varepsilon_2,\\ \text{s.t.}\quad & \|y-y_d\|_{L^\infty(\Omega_1)}\leqslant\varepsilon_1,\\ & \|y-y_d\|_{L^\infty(\Omega_2)}\leqslant\varepsilon_2.\end{aligned}$$

In this spirit, we define the optimal control problem below and the solution theory of Section 2 can be applied.

$$\begin{aligned}&\min_{(r,\varphi)\in U_{ad},\varepsilon_1\geqslant 0} J(r,\varphi)=\varepsilon_1+\frac{\alpha}{2}\int_0^T r(t)^2dt,\\ &\text{s.t.}\begin{cases}\frac{dM}{dt}=AM+\tilde{b}, & \frac{d\widetilde{M}}{dt}=\widetilde{A}\widetilde{M}+\tilde{b},\\ M(0,s)=M^0(s), & \widetilde{M}(0,s)=M^0(s),\quad s\in\Omega,\end{cases}\\ &\|M_w(T)-M_{w,des}\|_{L^\infty(\Omega_C)}\leqslant\varepsilon_1,\\ &\|M_w(T)-M_{w,des}\|_{L^\infty(\Omega_0)}\leqslant\varepsilon_2,\\ &\|\widetilde{M}_w(T)-\widetilde{M_{w,des}}\|_{L^\infty(\Omega_C)}\leqslant\varepsilon_3,\\ &\|\widetilde{M}_w(T)-\widetilde{M_{w,des}}\|_{L^\infty(\Omega_0)}\leqslant\varepsilon_4,\end{aligned}\tag{30}$$

where ε_2, ε_3 and ε_4 are given error parameters. Therein, the main objective is to approach the CEST−peak in an environment Ω_C as close as possible in the L^∞−norm, i.e. the most important task is to reduce ε_1. Simultaneously, certain error bounds ε_2 for the z−spectrum with exchange within $\Omega_0=\Omega\backslash\Omega_C$, as well as ε_3, ε_4 for the z−spectrum without exchange are prescribed.

Inspired by [17], the state constraints are treated with a L^{q_i}—penalization, where $q_i \in \mathbb{N}$, $i = 1, \cdots 4$, and the q_i are even. With parameters μ_i, $i = 1, \cdots 4$, the optimal control problem (30) is modified to

$$\begin{aligned}
&\min_{(r,\varphi)\in U_{ad}} J_{pen}(r,\varphi) = J(r,\varphi) \\
&+ \frac{\mu_1}{q_1} \int_{\Omega_C} \left(\frac{M_w(T) - M_{w,des}}{\varepsilon_1} \right)^{q_1} ds + \frac{\mu_2}{q_2} \int_{\Omega_0} \left(\frac{M_w(T) - M_{w,des}}{\varepsilon_2} \right)^{q_2} ds \\
&+ \frac{\mu_3}{q_3} \int_{\Omega_C} \left(\frac{\widetilde{M_w}(T) - \widetilde{M_{w,des}}}{\varepsilon_3} \right)^{q_3} ds + \frac{\mu_4}{q_4} \int_{\Omega_0} \left(\frac{\widetilde{M_w}(T) - \widetilde{M_{w,des}}}{\varepsilon_4} \right)^{q_4} ds, \qquad (31) \\
&\text{s.t.} \begin{cases} \frac{dM}{dt} = AM + \tilde{b}, & \frac{d\widetilde{M}}{dt} = \widetilde{A}\widetilde{M} + \tilde{b}, \\ M(0,s) = M^0(s), & \widetilde{M}(0,s) = M^0(s), \quad s \in \Omega. \end{cases}
\end{aligned}$$

3.1.3 About the optimization method

The penalized optimal control problem (31) has a similar form to the one in [17]. Therefore, the optimization method and software of [17] are used for all optimization within this thesis. A short description of the underlying methods is given here. For the extension of the software to sparse control, see [14]. The optimization method is a matrix-free second-order method with exact discrete derivative using adjoint calculus based on equation (29). The core of the optimization method is a combination of semismooth Newton method for the box control constraints (as well as the sparsity) with a quasi-Newton method for the smooth parts given by the tracking term. The method is analyzed in [14]. For robustness, the method is embedded into a trust-region framework build on Steihaug-cg [20]. The semismooth Newton method is based on the reformulation of the first-order necessary condition with Robinson's normal map, confer [16]. In this work the proximal map of the box constraints, and of sparsity with box constraints are applied here. As quasi-Newton method, the BFGS method is used, in compact limited-memory implementation and with a limit of 100. The objective parameters μ are calibrated automatically with the techniques from [17].

3.1.4 Reduction of the error band

In the cost functional (31), four different error bounds ε_i, $i = 1, \cdots 4$ for reproducing the desired magnetization are defined. The task is to minimize ε_1. Algorithmically, this is done in an outer homotopy loop, so that all ε_i, $i = 1, \cdots 4$, are fixed in the optimization. This allows to apply the optimization method of [17]. As soon as the optimization method finds an admissible point, ε_1 is reduced and the optimization is continued. Within this framework, it is also possible to reduce the other parameters ε_2, ε_3 and ε_4. Therefore, we compare two different strategies, a uniform reduction of those values $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4$, with a version, where only the error bounds around the CEST—peak with and without exchange are reduced.

Due to the duty-cycle constraint in the RF amplifiers, we aim at RF pulses that are sparse in time. Two different scenarios will be investigated. First, the sparsity is given and fixed. Later, the sparsity will be optimized by sparse control. The fixed sparsity is prescribed as

$$\begin{aligned} 0 \leqslant r(t) \leqslant r_{max} \quad \text{for} \quad t \in T_{on}, \\ r(t) = 0 \quad \text{for} \quad t \in T_{off}, \end{aligned}$$

with $I = T_{on} \cup T_{off}$, $T_{on} \cap T_{off} = \emptyset$ and $I = [0, T]$. Here we would like the RF to be 100ms on and 100ms off alternating. As initialization for optimization, a 0.9s Gaussian pulse with a mean amplitude of 0.58μT and $r_{max} = 2.5\mu T$ is used which fulfills the mentioned box constraints, see Figure 11. The step length is chosen to be $\Delta t = 1e-3s$, while the $z-$spectrum $\Omega = [-15, 15]ppm$ is discretized with a step size of 0.05ppm. The optimization is done with on the one hand fixed values

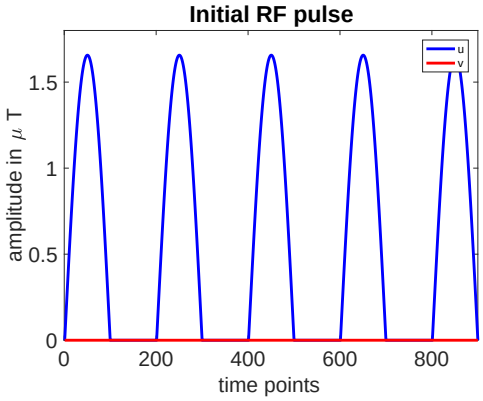

Figure 11: Initial RF pulse for optimization using a Gaussian type pulse train, where $u = r \cdot \cos \varphi$ and $v = r \cdot \sin \varphi$.

of q_i, and on the other hand with a homotopy loop for increasing q_i around the optimization.

UNIFORM REDUCTION The error bounds are initially set to $\varepsilon_i = 0.03$, $i = 1, \cdots 4$ and they are uniformly reduced by 1%, with regularization parameters $\mu_i = 10$, and $i = 1, \cdots 4$. This study of optimization is done with the L^2- norm for each boundary part, so $q_i = 2$, $i = 1, \cdots 4$. We require a uniform reduction of all bounds and a maximum iteration number of 10000.

NON-UNIFORM REDUCTION The parameters ε_i, μ_i and q_i are initialized with the same values as above and a maximum iteration number of 10000. But we require a non-uniform reduction, tightening only the error bounds around the CEST−peak specified by ε_1 and ε_3, by reducing stepwise by 1%.

In Table 7 we see the optimization results, where DE denotes the outcome using the desired RF pulse, IN the initial and UN and $NON-UN$ are the two optimized pulses corresponding to the description above. The optimization runs indicated by H, SC and SCL will be discussed in Section 3.1.5 and 3.1.6. Here, the value CEST describes the difference of the water magnetization in the CEST−peak with

Table 7: Optimization results from different studies. The rows show the desired RF pulse DE, the initial IN and the optimized UN, NON − UN, H and SC with its finer version SCL. The columns depict the CEST−peak with its increase, the duty cycle, the SAR and the error parameters.

RF	CEST	increase in %	duty cycle	SAR	ε_1	ε_2	ε_3	ε_4
DE	0.123		100%	0.3028				
IN	0.079		55.56%	0.6792				
UN	0.0893	23.4%	55.56%	0.2934	0.0197	0.0197	0.0197	0.0197
NON − UN	0.0897	24.3%	55.56%	0.2169	0.0083	0.0300	0.0083	0.0300
H	0.0900	25.0%	55.56%	0.2799	0.0158	0.0158	0.0158	0.0158
SC	0.0954	37.3%	56.22%	0.2394	0.0130	0.0130	0.0130	0.0130
SCL	0.0930	31.8%	57.72%	0.2259	0.0142	0.0142	0.0142	0.0142

Table 8: Optimization results from different studies. The rows show the initial RF pulse IN and the optimized UN, NON − UN, H and SC with its finer version SCL. The columns depict the relative $L^\infty-$ and L^2-errors to the desired magnetization DE.

RF	$\varepsilon_{\infty,1}$	$\varepsilon_{\infty,2}$	$\varepsilon_{\infty,3}$	$\varepsilon_{\infty,4}$	$\varepsilon_{2,1}$	$\varepsilon_{2,2}$	$\varepsilon_{2,3}$	$\varepsilon_{2,4}$
IN	$6.0e-2$	$1.6e-1$	$5.0e-2$	$1.6e-2$	$3.3e-2$	$4.8e-2$	$3.2e-2$	$4.7e-2$
UN	$5.1e-3$	$2.0e-2$	$1.3e-2$	$2.0e-2$	$6.3e-3$	$3.3e-3$	$4.3e-3$	$3.4e-3$
NON − UN	$4.9e-3$	$3.0e-2$	$1.3e-2$	$3.1e-2$	$6.0e-3$	$5.1e-3$	$4.2e-3$	$5.2e-3$
H	$8.9e-3$	$1.6e-2$	$1.6e-2$	$1.6e-2$	$6.6e-3$	$4.8e-3$	$6.8e-3$	$5.1e-3$
SC	$9.4e-3$	$1.3e-2$	$1.3e-2$	$1.3e-2$	$7.0e-3$	$3.9e-3$	$7.5e-3$	$4.2e-3$
SCL	$5.3e-3$	$1.3e-2$	$1.4e-2$	$1.4e-2$	$5.8e-3$	$3.0e-3$	$5.7e-3$	$3.3e-3$

and without exchange at the CEST−peak, here at 3.5ppm. Compared to the initial value, the CEST−peak increases 23.4% for the uniform reduction UN and 24.3% for the non-uniform reduction NON − UN. The value obtained by the desired magnetization DE is unreachable, which was expected since it was created with a pulse that does not fulfill the enforced sparsity pattern.
The duty cycle depicts the number of points in time, where the RF pulse is turned on. As we applied the same fixed sparsity pattern for the initial RF pulse IN and the two optimized pulses UN and NON − UN, this value is obviously the same for all of them. The next column monitors the global RF pulse power

$$\mathrm{SAR}(r) = \int_0^T r(t)^2 dt,$$

where $r(t)$ is the amplitude of the pulse during a specific point in time. Due to application reasons, the SAR is wished to be limited, therefore it was also included in the cost functional, see Section 2.1. The SAR of the initial pulse IN was in both optimization runs UN and NON − UN reduced to less than a half that value, and now admits a lower SAR than the desired RF pulse DE.

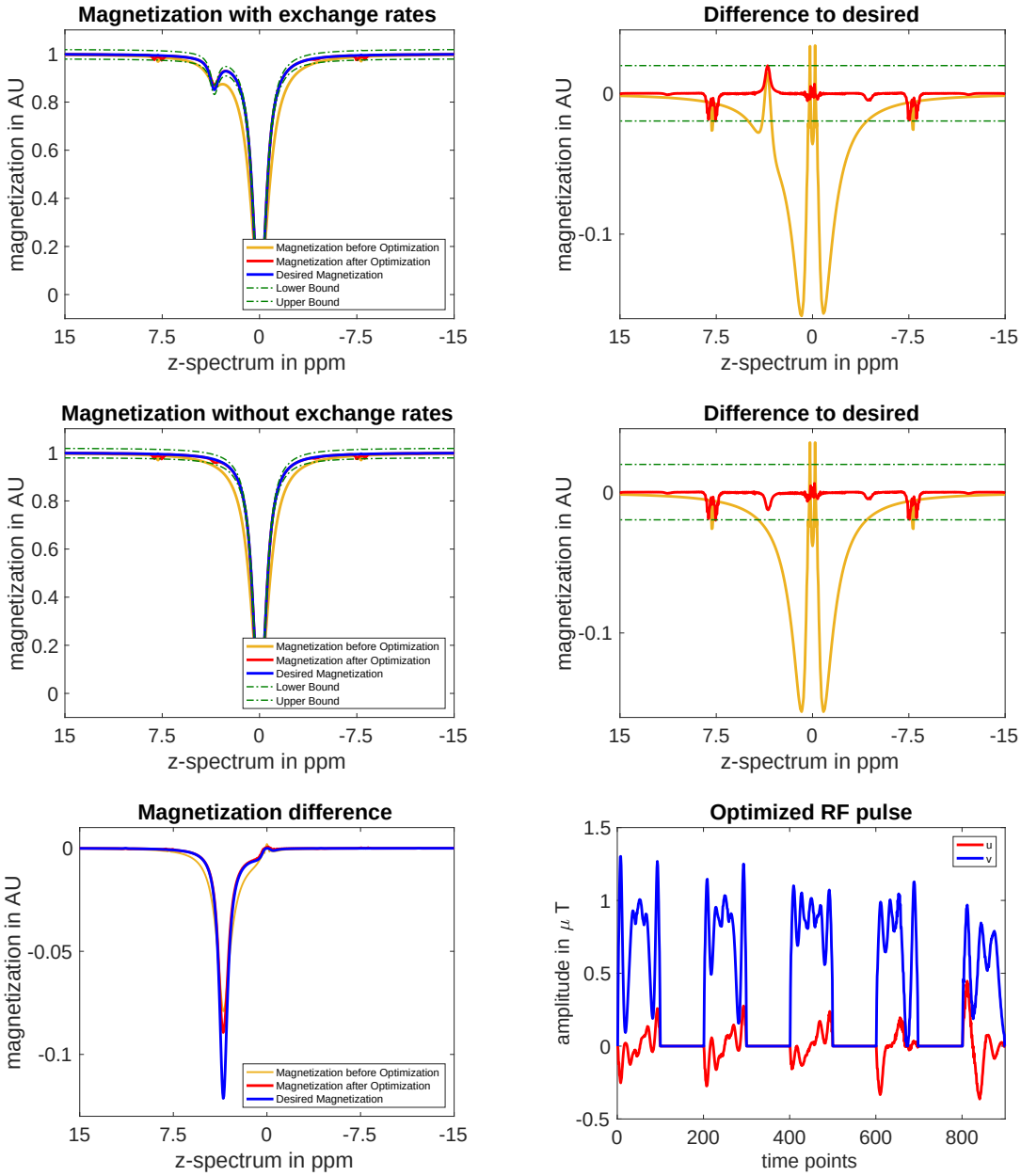

Figure 12: Uniform reduction of the bounds resulting in $\varepsilon = 0.0197$ with a maximum iteration number of 10000 and $q_i = 2$, $i = 1, \cdots 4$. Plots of $z-$magnetization (left) and the difference to the desired magnetization (right) of the water proton pool with (row one) and without (row two) exchange. Row three shows the difference between magnetization with and without exchange and the optimized RF pulse, where $u = r \cdot \cos\varphi$, $v = r \cdot \sin\varphi$.

The values $\varepsilon_{\infty,i}$, $i = 1, \cdots 4$ and $\varepsilon_{2,i}$, $i = 1, \cdots 4$ in Table 8 define the relative maximum error and the L^2-error of the optimized and initial magnetizations compared to the magnetization DE in all four tracking regions. Now we see that both optimized pulses $NON-UN$ and UN dramatically reduce those values, where the one with a uniform reduction UN has a better overall error, while the non-uniform one $NON-UN$ is slightly better in the regions around the CEST$-$peak, but not so good outside.
Due to the better overall error and only a slightly worse CEST$-$peak, the following optimization runs should consider a uniform reduction of the error parameters.
Figure 12 shows the graphical results of the optimization run with the resulting RF

pulse UN. Both in the magnetization with and without exchange rates, we see that the optimized magnetization scratches the lower bound at 7.5ppm and -7.5ppm, respectively. This phenomena is a result of the Nyquist-Theorem 7, as the sampling rate was chosen too small. We expect to overcome this instabilities when using a smaller time step length Δt. Moreover, the magnetization without exchange rates still shows some direct saturation of the water proton pool at 3.5ppm. A further drive of optimization should reduce this undesired behavior. The optimized RF pulse has left its initial Gaussian shape to a slightly oscillating one.

3.1.5 *Homotopy loop*

Next, optimization with a fixed set of values $q_i, i = 1, \cdots 4$ is compared with a version, where these experiments are increased in a homotopy loop around the optimization in order to approach the $L^\infty-$tracking. The expected outcome of this comparison is, that the homotopy loop can drive optimization further towards the desired state (in an $L^\infty-$sense).

OPTIMIZATION WITH FIXED EXPONENTS First, optimization of above problem is done without a homotopy loop, i.e. without increasing the values q_i. We set $q_i = 2$ and $\mu_i = 2$, $i = 1 \cdots 4$ and $\varepsilon_i = 0.03, i = 1, \cdots 4$ in the first place. This case corresponds to the uniform reduction with optimized RF pulse UN from above.

OPTIMIZATION WITH INCREASING EXPONENTS We perform optimization for fixed exponents to some maximal iteration number. Then the exponents q_i are increased and optimization is started again, where this time the initial RF is the optimized RF pulse of the last optimization, see Algorithm 1.

Algorithm 1 Homotopy loop

Data: Initial RF pulse (r_0, φ_0), initial parameters q_i, allowed maximum number of iterations $maxIter$, number of homotopy runs n
Result: Optimized RF pulse (r, φ)
for *i=1 to number homotopy runs* n **do**
    optimize yielding RF pulse (r, φ)
    increase q_i
    set $(r_0, \varphi_0) = (r, \varphi)$

In particular, optimization is here done with increasing the q_i from 2 to 10 in steps of 2, allowing a maximum number of iterations 2000 for every q_i. Again, all μ_i are set to 10 and $\varepsilon_1 = \cdots = \varepsilon_4 = 0.03$. As before we require a uniform reduction of the bound. During optimization, the ε_i were decreased to $\varepsilon_i = 0.0158$, $i = 1, \cdots 4$. The μ_i are recalibrated within the optimization method of [17] automatically. Table 7 shows that the optimized RF pulse H has an increasing CEST$-$peak of 25%, where the SAR as about the same as for the other optimized pulses. The errors towards the desired state DE are about the same as for UN, see Table 8.

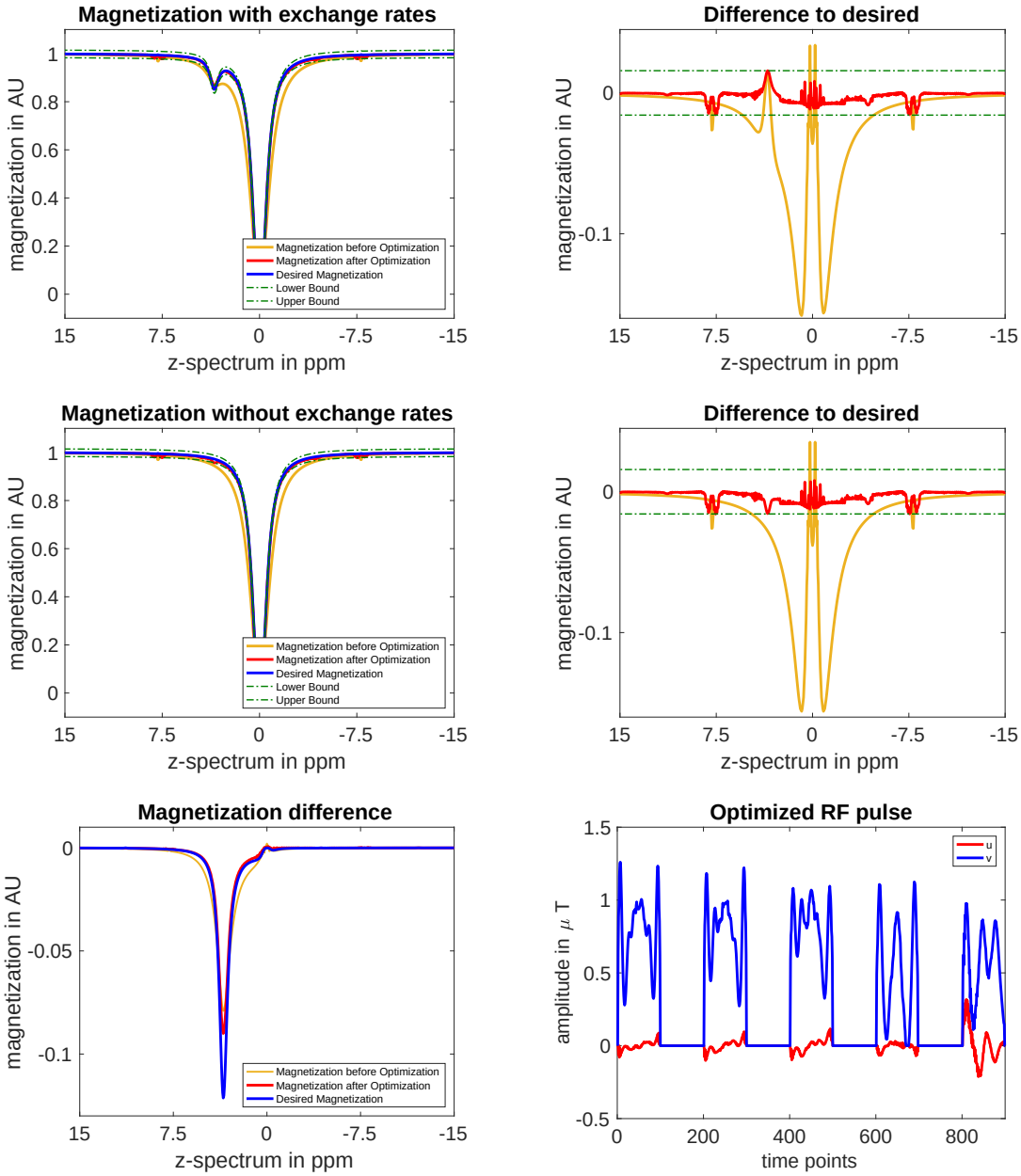

Figure 13: Uniform reduction of the bounds resulting in $\varepsilon = 0.0158$, where $q_i = 2$ to 10 with a maximum iteration number of 2000. Plots of $z-$magnetization (left) and the difference to the desired magnetization (right) of the water proton pool with (row one) and without (row two) exchange. Row three shows the difference between magnetization with and without exchange and the optimized RF pulse, where $u = r \cdot \cos\varphi$, $v = r \cdot \sin\varphi$.

In Figure 13 we see that compared to the run without homotopy, the optimized RF pulse has a similar shape. But magnetization with and without exchange is now significantly closer to the desired one which is a consequence of the smaller value of ε_1.

Comparing these two approaches, performing with a homotopy loop prevails here. The effort for both optimization tasks was the same, as we have chosen the total number of iterations being the same, but the optimized RF pulse with homotopy loop yields a better $z-$spectrum both with and without exchange than the one without homotopy.

3.1.6 Sparsity

In this section, motivated by the fact, that the RF amplifiers underly a certain duty cycle constraint, a comparison regarding two different methods for obtaining sparsity is done. The first one allows the control only to work on a certain set of time points T_{on}, whereas on T_{off} it is forced to be zero. The second one includes L^1−regularization of the RF amplitude in the cost functional, the points in time where the RF should be zero are not prescribed here.

FIXED SPARSITY Analogously to Section 3.1.2, we define a region T_{on} of the time set and a counterpart T_{off} such that $[0,T] = T_{on} \cup T_{off}$. Both controls u and v are allowed to work on T_{on}, but are forced to be zero on T_{off}. The initial Gaussian pulse, see Figure 11 fulfills this requirements with $T_{on} = [0,100] \cup [200,300] \cdots \cup [800,900]$ and $T_{off} = [0,T] \backslash T_{on}$, where $T = 900ms$. All regularization parameters correspond to Section 3.1.4. This case was already analyzed within the RF pulse UN, yielding a defined sparsity of 44.44%.

SPARSITY CONTROL It is well known, that sparsity of the optimal control can be achieved by adding a L^1−cost of the optimal control to the objective, confer [19]. This task is performed using a two step approach.
Step 1. The cost functional J is enhanced with additional L^1−regularization, so

$$J_{sparse} = J + \eta \int_0^T |r(t)| \, dt,$$

with a regularization parameter $\eta \geqslant 0$, see Section 2.1. During this part of optimization it is detected, where the RF pulse is switched on and off, i.e. certain sets T_{on} and T_{off} can be defined. The exponents are fixed to q_i, $i = 1, \cdots 4$, the ε_i are fixed to 0.03. In contrast to the case of fixed sparsity, optimization is started with the desired RF pulse as initial pulse. Therefore, the duty cycle in the beginning is 100% and the error of the initial towards the desired is zero in all regions of interest. The regularization parameters were set in a way, that a similar sparsity to the case with fixed sparsity was obtained, see Table 7.
The implementation is based on an extension of [17] to sparsity by using Robinson's normal map, [16].
Step 2. Optimization with the original cost functional J is done analogously to the case of fixed sparsity with the sets T_{on} and T_{off}, which were examined in Step 1. As initial RF pulse, the optimized one of Step 1 is used. Table 7 shows that the CEST−peak was increased by 37.3%, which is by far the best of all results. The SAR is in the order of the others. Also the ε−bounds are reduced far more than the others with uniform reduction. In Table 8 we see that the the errors around all regions of interest are in the order of the others. The interesting form of the optimized RF pulse can be seen in Figure 14.

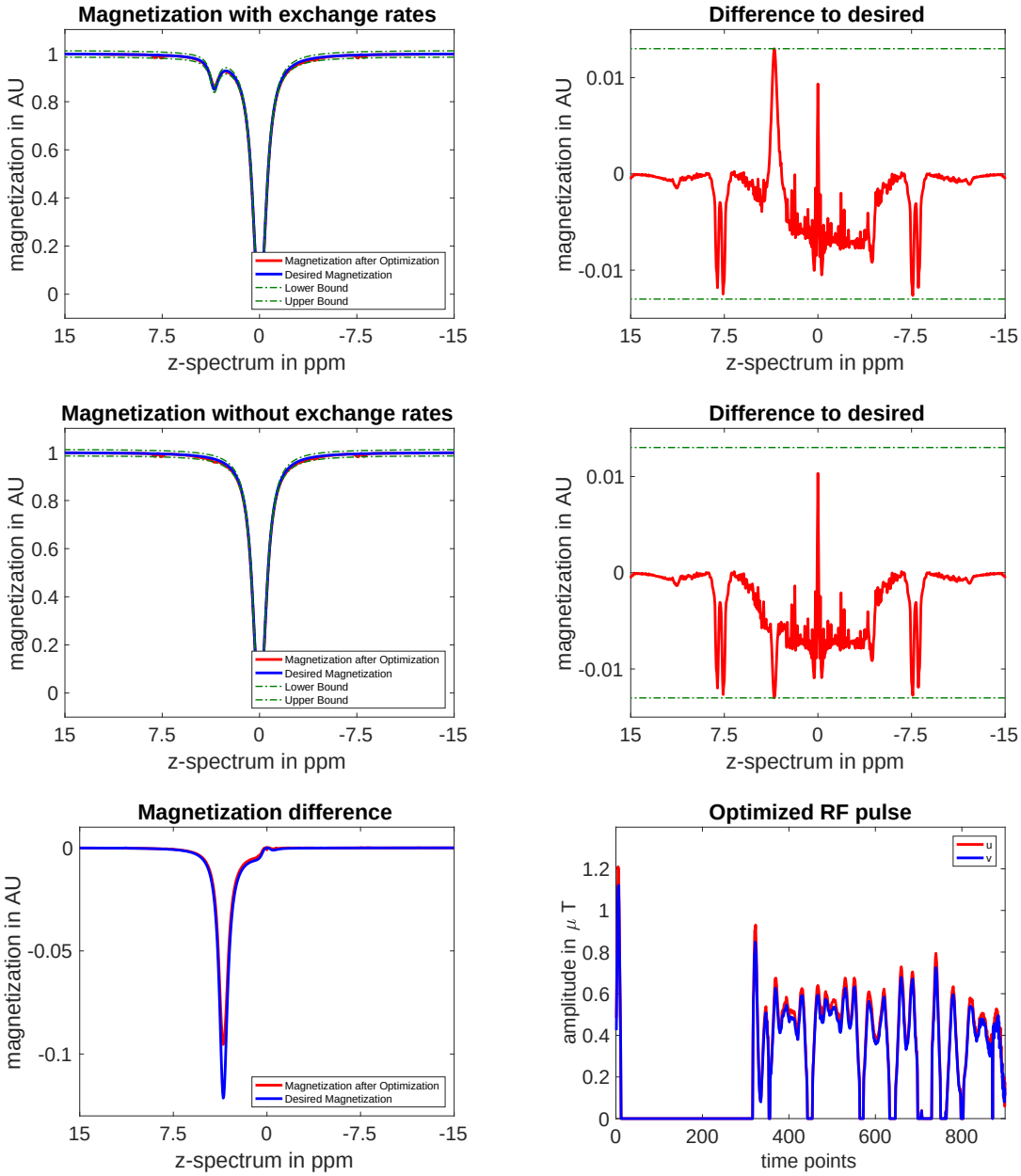

Figure 14: Two step approach including sparse control resulting in $\varepsilon = 0.013$. Plots of z—magnetization (left) and the difference to the desired magnetization (right) of the water proton pool with (row one) and without (row two) exchange. Row three shows the difference between magnetization with and without exchange and the optimized RF pulse, where $u = r \cdot \cos\varphi$, $v = r \cdot \sin\varphi$.

We compared different method of performing optimization. First, a uniform reduction of the error bounds was identified to be more stable compared to the non-uniform one. Second, the homotopy loop around the q_i was presented as method to drive optimization further towards the desired state. In the end, we introduced L^1—sparsity and showed that it is possible to also optimize the sparsity pattern itself, together with the RF pulse. These first runs with sparsity optimization yield promising results.

3.1.7 *Optimized* RF *pulse*

Now we turn to the optimization with fine time step size $\Delta t = 1e-4s$. As this value is a correct discretization choice by means of the Nyquist Theorem 7, we expect that the instabilities around 7.5 and $-7.5ppm$ vanish. This behavior occurred indeed, see Figure 15.
The optimized sparse RF pulse is depicted in the lower right plot. Similarly to the case with the larger step length, the optimized RF pulse is after a first amplitude zero for quite a long time and then gains an oscillating behavior with some additional zero parts. The optimized $z-$spectrum stays always inside the box around the desired spectrum and is active at the CEST peak, see the upper right plot. Other results are summarized in Table 7, line SCL. There it can be seen that the optimized error bound ends up slightly larger than in the coarse discretization SC. The fine solution is slightly sparser and exhibits a smaller SAR value. The increase in CEST$-$peak is 31.8%. The relative errors are depicted in Table 8, line SCL. Here, lower values than in the coarse case SC are observed.

Robustness

In this last series of comparison, robustness of the optimized RF pulse with respect to slight changes of relaxation times, exchange rates and frequency offsets is examined. Again, as desired state, the output of a 0.9s continuous wave pulse is used. The values ε_∞ and ε_2 denote the relative $L^\infty-$ and L^2-errors of the optimized $z-$magnetization of the water proton pool with CSL towards the desired one DE with exchange. In contrast, $\widetilde{\varepsilon}_\infty$ and $\widetilde{\varepsilon}_2$ denote the relative errors without exchange rates.

Table 9: Robustness regarding a change of relaxation times $T_{1,w}$, $T_{1,s}$ and $T_{2,w}$, $T_{2,s}$.

Example	RF	CEST	ε_∞	$\widetilde{\varepsilon}_\infty$	ε_2	$\widetilde{\varepsilon}_2$
1	DE	0.123				
	SCL	0.093	$1.3e-2$	$1.4e-2$	$3.5e-3$	$3.7e-3$
2	DE	0.125				
	SCL	0.095	$1.2e-2$	$1.4e-2$	$3.8e-3$	$3.8e-3$
3	DE	0.122				
	SCL	0.094	$1.9e-2$	$1.9e-2$	$4.4e-3$	$4.7e-3$
4	DE	0.126				
	SCL	0.096	$1.9e-2$	$2.0e-2$	$4.7e-3$	$4.9e-3$
5	DE	0.116				
	SCL	0.089	$8.6e-3$	$1.4e-2$	$3.0e-3$	$3.2e-3$

First, we analyze the behavior regarding a change of relaxation times T_1, and T_2, see Table 9.

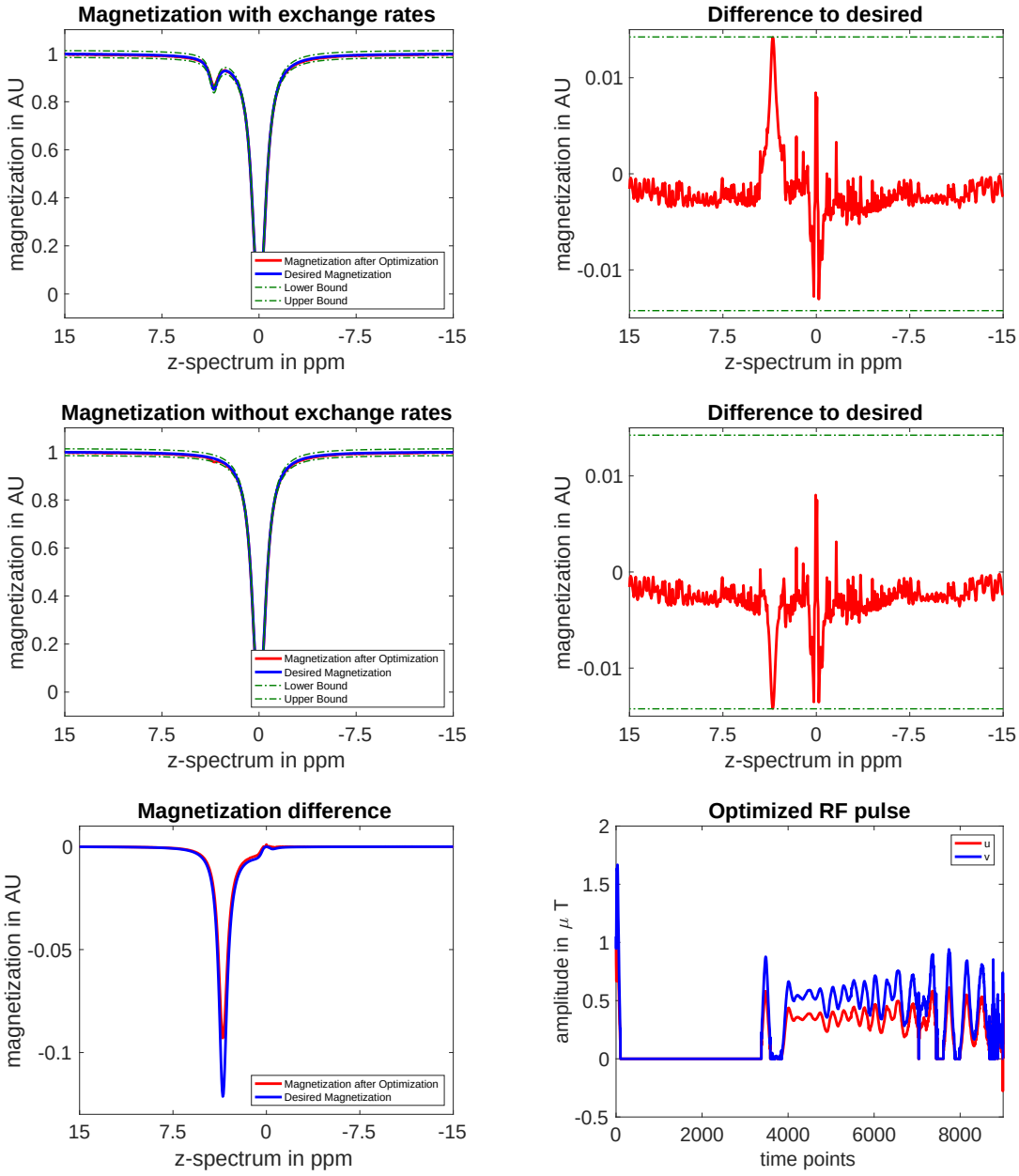

Figure 15: Two step approach including sparse control resulting in $\varepsilon = 0.0142$. Time step length of $\Delta t = 1e-4s$. Plots of $z-$magnetization (left) and the difference to the desired magnetization (right) of the water proton pool with (row one) and without (row two) exchange. Row three shows the difference between magnetization with and without exchange and the optimized RF pulse, where $u = r \cdot \cos\varphi$, $v = r \cdot \sin\varphi$.

The first example describes the original parameter set. In the second respective the third example, the longitudinal respective the transverse relaxation times are increased by 10%. During the fourth example, all relaxation times are increased by 10%. In contrast, the relaxation times are reduced by 10% in the fifth example. In Table 9 wee see that CEST and the four errors change only slightly with the

changed relaxation times. Note that the relative errors are always related to the continuous waver pulse using the corresponding changed relaxation times.

Next, the behavior among changes of magnetization exchange rates is analyzed. Again, the first example uses the initial parameter set to have a certain reference. During the second and the third one, the exchange rates k_{ws} and k_{sw} are increased by 10% and 20%, while in the fourth, they are decreased by 10%. In the fifth one, only the exchange rate from solute to water k_{sw} is increased by 30%. The sixth one reduced k_{ws} by 50% and increases k_{sw} by 50%.

Table 10: Robustness regarding a change of magnetization exchange rates k_{ws} and k_{sw}.

Example	RF	CEST	ε_∞	$\widetilde{\varepsilon}_\infty$	ε_2	$\widetilde{\varepsilon}_2$
1	DE	0.123				
	SCL	0.093	$1.3e-2$	$1.4e-2$	$3.5e-3$	$3.7e-3$
2	DE	0.131				
	SCL	0.100	$1.3e-2$	$1.4e-2$	$3.7e-3$	$3.7e-3$
3	DE	0.140				
	SCL	0.106	$1.3e-2$	$1.4e-2$	$3.9e-3$	$3.7e-3$
4	DE	0.111				
	SCL	0.086	$1.3e-2$	$1.4e-2$	$3.4e-3$	$3.7e-3$
5	DE	0.117				
	SCL	0.088	$1.3e-2$	$1.4e-2$	$3.5e-3$	$3.7e-3$
6	DE	0.055				
	SCL	0.040	$1.3e-2$	$1.4e-2$	$3.2e-3$	$3.7e-3$

In Table 10 we see that although a change of exchange rates has a significant influence on the CEST−peak, the behavior of the optimized RF pulse compared to the desired one is about the same. The values specified by $\widetilde{\varepsilon}_\infty$ and $\widetilde{\varepsilon}_2$ are in all examples the same, as those are calculated without exchange rates.

Table 11: Robustness regarding a change of the frequency offset of the solute ω_s.

Example	RF	CEST	ε_∞	$\widetilde{\varepsilon}_\infty$	ε_2	$\widetilde{\varepsilon}_2$
1	DE	0.123				
	SCL	0.093	$1.3e-2$	$1.4e-2$	$3.5e-3$	$3.7e-3$
2	DE	0.122				
	SCL	0.095	$1.3e-2$	$1.4e-2$	$4.1e-3$	$3.7e-3$
3	DE	0.123				
	SCL	0.096	$1.3e-2$	$1.4e-2$	$4.5e-3$	$3.7e-3$
4	DE	0.120				
	SCL	0.093	$1.3e-2$	$1.4e-2$	$4.1e-3$	$3.7e-3$
5	DE	0.117				
	SCL	0.091	$1.3e-2$	$1.4e-2$	$4.4e-3$	$3.7e-3$

Last, robustness of the CEST effect with respect to change in the frequency offset of the solute proton pool ω_s is regarded. For the second and the third example, it is set to 4ppm and 4.5ppm, while for the fourth and fifth they are reduced to 3 and 2.5ppm. As can be seen in Table 11, both the performance of the continuous wave pulse DE and the optimized pulse SCL are nearly not influenced by altered frequency offsets.

3.2 Optimization of 2 pool problem for phantom measurements

Another two pool model, which was designed to be measured on the MR scanner is presented and optimized here. The parameters originate from creatine in muscle and correspond to [13]. Due to the specification of the MR−scanner, $B_0 = 3T$. The longitudinal relaxation times are set to $T_{1,w} = 1.3s$ and $T_{1,s} = 0.5s$, where the transversal relaxation times are $T_{2,w} = 0.18s$ and $T_{2,s} = 0.011s$. The frequency offsets are described with $\omega_w = 0ppm$ and $\omega_s = 1.9ppm$, where the exchange rates between the water proton pool and the solute proton pool are specified by $k_{ws} = 0.702Hz$ and $k_{sw} = 389Hz$. As normalized initial magnetization, $M_{0,w} = 1$ and $M_{0,s} = 0.0018$ are used. We require a pulse duration of 0.9s, where corresponding to Nyquist's Theorem 7, a time step length of $\Delta t = 1e-4s$ is applied. The z−spectrum $\Omega = [-15, 15]ppm$ is discretized into 601 steps using a step length of 0.05. As desired state for optimization, the magnetization calculated with a 0.9s continuous wave RF pulse with an amplitude of $0.58\mu T$ was used.
Optimization was done analogously to Section 3.1.7, i.e. a uniform reduction of the error bands, a homotopy loop and sparsity via L^1−regularization were used.

Table 12: Optimization results from different studies. The rows show the desired RF pulse DE and the optimized OP. The columns depict the CEST−peak, the duty cycle, the SAR and the error parameters.

RF	CEST	duty cycle	SAR	ε_1	ε_2	ε_3	ε_4
DE	0.048	100%	0.3028				
OP	0.029	33.6%	0.1796	0.0102	0.0102	0.0102	0.0102

Table 13: Optimization results for the second two pool model with the optimized OP. Optimization results from different studies. The row shows the optimized RF pulse OP. The columns depict the relative L^∞− and L^2−errors to the desired magnetization DE.

RF	$\varepsilon_{\infty,1}$	$\varepsilon_{\infty,2}$	$\varepsilon_{\infty,3}$	$\varepsilon_{\infty,4}$	$\varepsilon_{2,1}$	$\varepsilon_{2,2}$	$\varepsilon_{2,3}$	$\varepsilon_{2,4}$
OP	$3.7e-3$	$5.9e-3$	$1.0e-2$	$9.5e-3$	$5.2e-3$	$1.5e-3$	$5.1e-3$	$1.5e-3$

With a duty cycle of only 33.6%, the error of the ε−bands could be reduced to $\varepsilon_i = 0.0102$, see Table 12. Additionally, the SAR was reduced by about 41% compared to the continuous wave pulse. The desired state is reached closely, the relative errors are depicted in Table 13. Figure 16 shows the visualized results, especially the interesting shape of the optimized RF pulse.

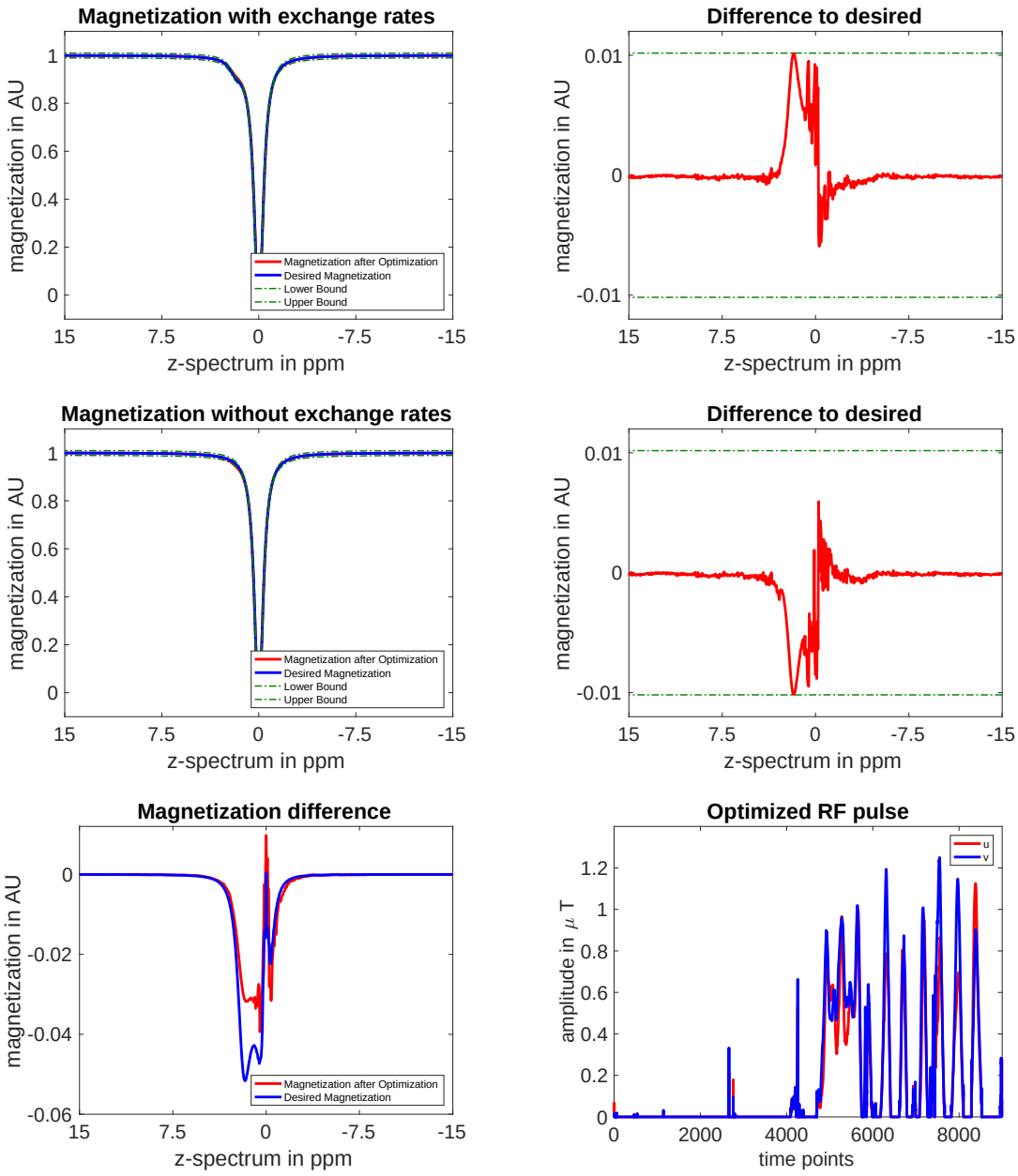

Figure 16: Two step approach including sparse control resulting in $\varepsilon = 0.0102$. Time step length of $\Delta t = 1e-4s$. Plots of $z-$magnetization (left) and the difference to the desired magnetization (right) of the water proton pool with (row one) and without (row two) exchange. Row three shows the difference between magnetization with and without exchange and the optimized RF pulse, where $u = r \cdot \cos\varphi$, $v = r \cdot \sin\varphi$.

CONCLUSION AND OUTLOOK

A numerical forward solver for the Bloch-McConnell equation was presented which is second order accurate and provides solutions in short time. The optimal control problem for Chemical Exchange Saturation Transfer was modeled and analyzed. First order necessary conditions were derived. The next section covered comprehensive numerical optimization studies for a selected 2 pool model, including numerical experiments with different objective functions, sparse control, and robustness studies. Afterwards, optimization was done for a second two pool model, yielding a RF pulse which will be implemented on the scanner for performing phantom measurements.

In the future, the sparse control for CEST should be further investigated, especially the joint optimization of sparsity and $L^\infty-$tracking in one step is expected to further increase the quality of the solutions. Furthermore, weighting or penalization terms should be added in order to more evenly distribute the sparsity pattern. Besides, the investigation of other homotopy strategies for the ε_i and q_i might additionally improve the CEST effect of the optimized pulses. However, the state and the adjoint solvers should be parallelized first using C++ with OpenMP, or using graphic cards. Besides, also RF pulse design for models with three or more pools needs to be investigated for improving the practical realization of CEST imaging.

REFERENCES

[1] Gronwall's lemma. https://www.math.uni-bielefeld.de/~rkruse/files/gronwall.pdf.

[2] Hargreaves solver. http://mrsrl.stanford.edu/~brian/bloch/.

[3] Morrey's theorem. http://page.math.tu-berlin.de/~baerwolf/num_pde_ss10/vortrag_sobolev.einbettung.pdf.

[4] Nyquist-shannon sampling theorem. https://en.wikipedia.org/wiki/Nyquist?Shannon_sampling_theorem.

[5] Rotation matrices. https://en.wikipedia.org/wiki/Rotation_matrix.

[6] Young's inequality. https://en.wikipedia.org/wiki/Youngś_inequality_for_products.

[7] Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. *Handbook of MRI Pulse Sequences*. Elsevier Academic Press, 2004.

[8] J. Bittoun, J. Taquin, and M. Sauzade. *A computer algorithm for the simulation of any Nuclear Magnetic Resonance (NMR) imaging method*, volume 2. 1984.

[9] Zhipeng Cao, Sukhoon Oh, Christopher T. Sica, John M. McGarrity, Timothy Horan, Wei Luo, and Christopher M. Collins. *Bloch-based MRI system simulator considering realistic electromagnetic fields for calculation of signal, noise, and specific absorption rate*, volume 72.

[10] Massimo Fornasier. *Learning and Sparse Control of Multiagent Systems*. 2010.

[11] Nishimura D. G. *Principles of Magnetic Resonance Imaging*. 1996.

[12] Christina Graf. *On the Numerical Solution of the Bloch Equation with Relaxation*. Report des Technomathematik Praktikums, 2017.

[13] Feliks Kogan, Mohammad Haris, Anup Singh, Kejia Cai, Catherine Debrosse, Ravi Prakash Reddy Nanga, Hari Hariharan, and Ravinder Reddy. *Method for high-resolution imaging of creatine in vivo using chemical exchange saturation transfer*, volume 71.

[14] Florian Kruse and Armin Rund. *A hybrid semismooth quasi-Newton method*. 2018.

[15] Kurt Meyberg and Peter Vachenauer. *Höhere Mathematik 2*. Springer Verlag, 1997.

[16] Konstantin Pieper. *Finite element discretization and efficient numerical solution of elliptic and parabolic sparse control problems*. PhD Dissertation, Technische Universität München, April 2015.

[17] Armin Rund, Christoph Stefan Aigner, Karl Kunisch, and Rudolf Stollberger. *Magnetic Resonance RF pulse design by optimal control with physical constraints*. 2017.

[18] Mohamed Khamsi Saleh Almezel, Qamrul Hasan Ansari. *Topics in Fixed Point Theory*. Springer Verlag, 2014.

[19] Georg Stadler. *Elliptic optimal control problems with* L^1*—control cost and applications for the placement of control devices*. 2009.

[20] Trond Steihaug. *The Conjugate Gradient Method and Trust Regions in Large Scale Optimization*, volume 20. 1983.

[21] Phillip Zhe Sun, Enfeng Wang, Jerry S. Cheung, Xiaoan Zhang, Thomas Brenner, and A. Gregory Sorensen. *Simulation and Optimization of Pulsed Radio Frequency Irradiation Scheme for Chemical Exchange Saturation Transfer (CEST) MRI - Demonstration of pH-Weighted Pulsed-Amide Proton CEST MRI in an Animal Model of Acute Cerebral Ischemia*. 2011.

[22] Fredi Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg und Teubner, 2009.

[23] Moritz Zaiss and Peter Bachert. *Exchange-dependent relaxation in the rotating frame for slow and intermediate exchange - modeling off-resonant spin-lock and chemical exchange saturation transfer*. 2012.

[24] Moritz Zaiss, Zhongliang Zu, Junzhong Xu, Patrick Schuenke, Daniel F. Gochberg, John C. Gore, Mark E. Ladd, and Peter Bachert. *A combined analytical solution for chemical exchange saturation transfer and semi-solid magnetization transfer*. 2014.

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature