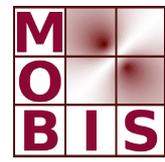
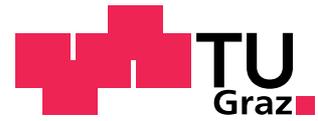




SpezialForschungsBereich F 32



Karl-Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz



Efficient high-resolution RF pulse design applied to simultaneous multi-slice excitation

C. S. Aigner C. Clason A. Rund
R. Stollberger

SFB-Report No. 2015-001

November 2015

A-8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)



SFB sponsors:

- **Austrian Science Fund (FWF)**
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**



Efficient high-resolution RF pulse design applied to simultaneous multi-slice excitation

Christoph Stefan Aigner* Christian Clason† Armin Rund‡
Rudolf Stollberger*

November 23, 2015

RF pulse design via optimal control is typically based on gradient and quasi-Newton approaches and therefore suffers from slow convergence. We present a flexible and highly efficient method that uses exact second-order information within a globally convergent trust-region CG-Newton method to yield an improved convergence rate. The approach is applied to the design of RF pulses for single- and simultaneous multi-slice (SMS) excitation and validated using phantom and in-vivo experiments on a 3 T scanner using a modified gradient echo sequence.

Keywords: pulse design, optimal control, second-order methods, simultaneous multi-slice excitation

1. Introduction

For many applications in MRI there is still demand for the optimization of selective RF excitation, e.g., for simultaneous multi-slice excitation [30, 49], UTE imaging [8], or velocity selective excitation [15]. To achieve a well-defined slice profile at high field strength while meeting B_1 peak amplitude limitations is a challenge for RF pulse design and becomes especially critical for quantitative methods.

Correspondingly, many approaches for general pulse design have been proposed in the literature. RF pulses with low flip angles can be designed using the small tip angle simplification [38], which makes use of an approximation of the Bloch equation to compute a pulse via the Fourier

*Institute of Medical Engineering, Graz University of Technology, Kronesgasse 5, and BioTechMed Graz, 8010 Graz, Austria (christoph.aigner@tugraz.at)

†Faculty of Mathematics, University of Duisburg-Essen, 45117 Essen, Germany

‡Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria

transform of the desired slice profile. However, this simplification breaks down for large flip angles. The resulting excitation error for large flip angle pulses can be reduced by applying the Shinnar–Le Roux (SLR) technique [37] or optimization methods, e.g., simulated annealing, evolutionary approaches or optimal control [12, 26, 29, 45, 51, 55, 56]. The SLR method is based on the hard pulse approximation and a transformation of the excitation problem, allowing to solve the excitation problem recursively by applying fast filter design algorithms such as the Parks–McClellan algorithm [37]. Originally, this approach only covered special pulses such as 90° and 180° excitation or refocusing, but Lee [31] generalized this approach to arbitrary flip angles with an exact parameter relation. Despite its limitations due to neglected relaxation terms and sensitivity to B_1 inhomogeneities, it found widespread use (see, e.g., [3, 20, 31, 32]) and is considered to be the gold standard for large tip angle pulse design. An alternative approach is based on optimizing a suitable functional; see, e.g., [7, 19, 37, 50, 56]. In particular, optimal control (OC) approaches involve the solution of the Bloch equation describing the evolution of the magnetization vector in an exterior magnetic field [5, 12, 21, 27, 29, 46, 51, 56]. They often lead to better excitation profiles due to a more accurate design model and are increasingly used in MRI, for instance, to perform multidimensional and multichannel RF design [21, 56], robust 2D spatial selective pulses [51] and saturation contrast [29]. In addition, arbitrary flip angles and target slice profiles, as well as inclusion of additional physical effects such as, e.g., relaxation can be handled. However, so far OC approaches are limited by the computational effort and require a proper modeling of the objective. In particular, standard gradient-based approaches suffer from slow convergence, imposing significant limitations on the accuracy of the obtained slice profiles. On the other hand, Newton methods show a locally quadratic convergence, but require second-order information which in general is expensive to compute [1]. Approximating the Hessian using finite differences causes loss of quadratic convergence due to the lack of exact second-order information and typically requires significantly more iterations. Superlinear convergence can be obtained using quasi-Newton methods based on exact gradients [14], although their performance can be sensitive to implementation details. The purpose of this work is to demonstrate that for the OC approach to pulse design, it is in fact possible to use exact second-order information while avoiding the need of computing the full Hessian, yielding a highly efficient numerical method for the optimal control of the full time-dependent Bloch equation. In contrast to [1] (which uses black-box optimization method and symbolically calculated Hessians based on an effective-matrix approximation of the Bloch equation), we propose a matrix-free Newton–Krylov method [28] using first- and second-order derivatives based on the adjoint calculus [23] together with a trust-region globalization [48]; for details we refer to Section 2.2. Recently, similar matrix-free Newton–Krylov approaches with line search globalization were presented for optimal control of quantum systems in the context of NMR pulse sequence design [10, 11]. In comparison, the proposed trust-region framework significantly reduces the computational effort, particularly for the initial steps far away from the optimum. The effectiveness of the proposed method is demonstrated for the design of pulses for single and simultaneous multi-slice excitation (SMS).

SMS excitation is increasingly used to accelerate imaging experiments [17, 30, 36, 49]. Conventional design approaches, based on a superposition of phase-shifted sub-pulses [33] or sinusoidal modulation [30], typically result in a linear scaling of the B_1 peak amplitude, a quadratic peak

power and a linear increase in the overall RF power [2, 35]. The required maximal B_1 peak amplitude of conventional multi-slice pulses therefore easily exceeds the transmit voltage of the RF amplifier. In this case, clipping will occur, while rescaling will decrease and limit the maximal flip angle of such a pulse. On the other hand, restrictions of the specific absorption rate limit the total (integrated) B_1 power and therefore the maximal number of slices as well as the pulse duration and flip angle. The increase of B_1 power can be addressed by the Power Independent of Number of Slices (PINS) technique [35], which was extended to the kT-PINS method [44] to account for B_1 inhomogeneities. This approach leads to a nearly slice-independent power requirement, but the periodicity of the resulting excitation restricts the slice orientation and positioning. Furthermore, the slice profile accuracy is reduced [17], and a limited ratio between slice thickness and slice distance may further restrict possible applications. The combination of PINS with regular multi-band pulses was shown to reduce the overall RF power by up to 50 % (MultiPINS [16]) and was applied to refocusing pulses in a multi-band RARE sequence with 13 slices [18].

A different way to reduce the maximum B_1 amplitude is to increase the pulse length; however, this stretching increases the minimal echo and repetition times and decreases the RF bandwidth, thus reducing the slice profile accuracy [2]. Applying variable rate selective excitation [13, 41] avoids this problem but leads to an increased sensitivity to slice profile degradations at off-resonance frequencies. In addition, they require specific sequence alterations, e.g., variable slice gradients or gradient blips. Instead of using the same phase for all sub-slices, the peak power can be reduced by changing the uniform phase schedule to a different phase for each individual slice [53]. Alternative approaches [43, 57] using phase-matched excitation and refocusing pairs show that a nonlinear phase pattern can be corrected by a subsequent refocusing pulse. Another way to reduce the power deposition and SAR of SMS pulses is to combine them with parallel transmission [25]. This allows to capitalize transmit sensitivities in the pulse design and leads to a more uniform excitation with an increased power efficiency [39, 54]. Recently, Guerin et al. [22] demonstrated that it is possible to explicitly control both global and local SAR as well as the peak power using a spokes-SMS-pTx pulse design.

The focus of this work, however, is on single channel imaging, where we apply our OC-based pulse design for efficient SMS pulse optimization using a direct description of the desired magnetization pattern. Its flexible formulation allows a trade-off between the slice profile accuracy and the required pulse power and is well suited for the reduction of power and amplitude requirements of such pulses, even for a large number of slices or large flip angles or in presence of relaxation. The efficient implementation of the proposed method allows to optimize for SMS pulses with a high spatial resolution to achieve accurate excitation profiles. The RF pulses are designed to achieve a uniform effective echo time and phase for each slice and use a constant slice-selective gradient, allowing to insert the RF pulse into existing sequences and opening up a wide range of applications.

2. Theory

This section is concerned with the description of the optimal control approach to RF pulse design as well as of the proposed numerical solution approach.

2.1. Optimal control framework

Our OC approach is based on the full time-dependent Bloch equation, which describes the temporal evolution of the ensemble magnetization vector $M(t) = (M_x(t), M_y(t), M_z(t))^T$ due to a transient external magnetic field $B(t)$ as the solution of the ordinary differential equation (ODE)

$$\begin{cases} \dot{M}(t) = \gamma B(t) \times M(t) + R(M(t)), & t > 0, \\ M(0) = M^0, \end{cases}$$

where γ is the gyromagnetic ratio, M^0 is the initial magnetization and

$$R(M(t)) = (-M_x(t)/T_2, -M_y(t)/T_2, -(M_z(t) - M_0)/T_1)^T$$

denotes the relaxation term with relaxation times T_1, T_2 and the equilibrium magnetization M_0 . To encode spatial information in MR imaging, the external magnetic field B (and thus the magnetization vector) depends on the slice direction z , hence the Bloch equation can be considered as a parametrized family of three-dimensional ODEs. In the on-resonance case and ignoring spatial field inhomogeneities, the Bloch equation can be expressed in the rotating frame as

$$(1) \quad \begin{cases} \dot{M}(t; z) = A(u(t); z)M(t; z) + b(z), & t > 0, \\ M(0; z) = M^0(z), \end{cases}$$

where the control $u(t) = (u_x(t), u_y(t))$ describes the RF pulse,

$$A(u; z) = \begin{pmatrix} -\frac{1}{T_2} & \gamma G_z(t)z & \gamma u_y(t)B_1 \\ -\gamma G_z(t)z & -\frac{1}{T_2} & \gamma u_x(t)B_1 \\ -\gamma u_y(t)B_1 & -\gamma u_x(t)B_1 & -\frac{1}{T_1} \end{pmatrix}, \quad b(z) = \begin{pmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{pmatrix},$$

and G_z is the slice-selective gradient; see, e. g., [34, Chapter 6.1].

The OC approach consists in computing for given initial magnetization $M^0(z)$ the RF pulse $u(t)$, $t \in [0, T_u]$, that minimizes the squared error at read-out time $T > T_u$ between the corresponding solution $M(T; z)$ of (1) and a prescribed slice profile $M_d(z)$ for all $z \in [-a, a]$ together with a quadratic cost term modeling the SAR of the pulse, i.e., solving

$$(2) \quad \min_{(u, M) \text{ satisfying (1)}} J(M, u) = \frac{1}{2} \int_{-a}^a |M(T; z) - M_d(z)|_2^2 dz + \frac{\alpha}{2} \int_0^{T_u} |u(t)|_2^2 dt.$$

The parameter $\alpha > 0$ controls the trade-off between the competing goals of slice profile attainment and SAR reduction.

2.2. Adjoint approach

The standard gradient method for solving (2) consists of computing for given u^k the gradient $g(u^k)$ of $j(u) := J(M(u), u)$ and setting $u^{k+1} = u^k - s^k g(u^k)$ for some suitable step length s^k . The gradient can be calculated efficiently using the adjoint method, which in this case yields

$$(3) \quad \begin{aligned} g(u^k)(t) &= \alpha u(t) + \gamma B_1 \left(\int_{-a}^a M_z(t; z) P_y(t; z) - M_y(t; z) P_z(t; z) dz \right) \\ &= \alpha u(t) + \left(\int_{-a}^a M(t; z) A_1 P(t; z) dz \right), \quad 0 \leq t \leq T_u, \end{aligned}$$

where M is the solution to (1) for $u = u^k$ and $0 < t \leq T$, P is the solution to the adjoint (backward in time) equation

$$\begin{cases} -\dot{P}(t; z) = A(u(t); z)^T P(t; z), & 0 \leq t < T, \\ P(T; z) = M(T; z) - M_d(z), \end{cases}$$

and for the sake of brevity, we have set

$$A_1 := \gamma B_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 := \gamma B_1 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

However, this method requires a line search to converge and usually suffers from slow convergence close to a minimizer. This is not the case for Newton's method (which is a second-order method and converges locally quadratically), where one additionally computes the Hessian $H(u^k)$ of j at u^k , solves for δu in

$$(4) \quad H(u^k) \delta u = -g(u^k),$$

and sets $u^{k+1} = u^k + \delta u$. While the full Hessian $H(u^k)$ is very expensive to compute in practice, solving (4) using a Krylov method such as conjugate gradients (CG) only requires computing the Hessian action $H(u^k)h$ for a given direction h per iteration; see, e.g., [28]. The crucial observation in our approach is that the adjoint method allows computing this action exactly (e.g., without employing finite difference approximations) and without knowledge of the full Hessian. Since Krylov methods usually converge within very few iterations, this so-called "matrix-free" approach amounts to significant computational savings. To derive a procedure for computing the Hessian action $H(u^k)h$ for a given direction h directly, we start by differentiating (3) with respect to u in direction h and applying the product rule. This yields

$$(5) \quad [H(u^k)h](t) = \alpha h(t) + \left(\int_{-a}^a \delta M(t; z) A_1 P(t; z) + M(t; z) A_1 \delta P(t; z) dz \right), \quad 0 \leq t \leq T_u,$$

where δM – corresponding to the directional derivative of M with respect to u – is given by the solution of the linearized state equation

$$(6) \quad \begin{cases} \delta \dot{M}(t; z) = A(u^k; z) \delta M(t; z) + A'(h)M, & 0 < t \leq T, \\ \delta M(0; z) = (0, 0, 0)^T, \end{cases}$$

with

$$A'(h) = \gamma B_1 \begin{pmatrix} 0 & 0 & h_y(t) \\ 0 & 0 & h_x(t) \\ -h_y(t) & -h_x(t) & 0 \end{pmatrix},$$

and δP – corresponding to the directional derivative of P with respect to u – is the solution of the linearized adjoint equation

$$(7) \quad \begin{cases} -\delta \dot{P}(t; z) = A(u^k; z)^T \delta P(t; z) + A'(h)^T P, & 0 \leq t < T, \\ \delta P(T; z) = \delta M(T; z). \end{cases}$$

This characterization can be derived using formal Lagrangian calculus and rigorously justified using the implicit function theorem; see, e.g., [24, Chapter 1.6]. Since (5) can be computed by solving the two ODEs (6) and (7), the cost of computing a single Hessian action is comparable to that of a gradient evaluation; cf. (3). This has already been observed in the context of seismic imaging [40], meteorology [52], and optimal control of partial differential equations [23], but has received little attention so far in the context of optimal control of ODEs.

One difficulty is that the Bloch equation (1) is bilinear since it involves the product of the unknowns u and M . Hence, the optimal control problem (2) is not convex and the Hessian $H(u)$ is not necessarily positive definite (or even invertible), thus precluding a direct application of the CG-Newton method. We therefore embed Newton's method into the trust-region framework of Steihaug [48], where a breakdown of the CG method is handled by a trust-region step and the trust region radius is continually adapted. This allows global convergence (i.e., for any starting point) to a local minimizer as well as transition to fast quadratic convergence; see [48]. As an added advantage, computational time is saved since the CG method is usually not fully resolved far away from the optimum. The full algorithm is given in Appendix A.

2.3. Discretization

For the numerical computation of optimal controls, both the Bloch equation (1) and the optimal control problem in (2) need to be discretized. Here, the time interval $[0, T]$ is replaced by a time grid $0 = t_0 < \dots < t_N = T$ with time steps $\Delta t_m := t_m - t_{m-1}$, chosen such that $t_{N_u} = T_u < T$ for some $N_u < N$. The domain $[-a, a]$ is replaced by a spatial grid $-a = z_1 < \dots < z_Z = a$ with grid sizes $\Delta z_m := z_m - z_{m-1}$. We note that for each z_i , the corresponding ODEs can be solved independently and in parallel. The Bloch equation is discretized using a Crank-Nicolson method, where the state M is discretized as continuous piecewise linear functions with values $M_m := M(t_m)$, and the controls u are treated as piecewise constant functions, i.e., $u = \sum_{m=1}^{N_u} u_m \chi_{(t_{m-1}, t_m]}(t)$, where $\chi_{(a, b]}$ is the characteristic function of the half-open interval $(a, b]$.

For the efficient computation of optimal controls, it is crucial that both the gradient and the Hessian action are computed in a manner consistent with the chosen discretization. This implies that the adjoint state P has to be discretized as piecewise constant using an appropriate time-stepping scheme [4], and that the linearized state δM and the linearized adjoint state δP have to be

discretized in the same way as the state and adjoint state, respectively. Furthermore, the conjugate gradient method has to be implemented using the scaled inner product $\langle u, v \rangle := \sum_{m=1}^{N_u} \Delta t_m u_m v_m$ and the corresponding induced norm $\|u\|^2 := \langle u, u \rangle$. For completeness, the resulting schemes and discrete derivatives are given in Appendix B.

3. Methods

This section describes the computational implementation of the proposed pulse design and the experimental protocol for its validation.

3.1. Pulse design

The OC approach described in Section 2 is implemented in MATLAB (The MathWorks, Inc., Natick, USA) using the Parallel Toolbox for parallel solution of the (linearized) Bloch and adjoint equation for different values of z_i . In the spirit of reproducible research, the code used to generate the results in this paper can be downloaded from <https://github.com/chaigner/rfcontrol/releases/v1.2>.

The initial magnetization vector is set to equilibrium, i.e., $M^0(z) = M_0(0, 0, 1)^T$. The slice-selective gradient $G_z(t)$ is extracted out of a standard Cartesian GRE sequence simulation and consists of a trapezoidal shape of length 2.56 ms that is followed by a re-phasing part of length 0.92 ms to correct the phase dispersion using the maximal slew rate; i.e., $T_u = 2.56$ ms and $T = 3.48$ ms with a temporal resolution of $\Delta t = 5 \mu\text{s}$ for the single-slice excitation (see dashed line in Figure 2a) and $T_u = 10.24$ ms and $T = 13.92$ ms with a temporal resolution of $\Delta t = 20 \mu\text{s}$ for the SMS excitation (see dashed line in Figure 4a). This corresponds in both cases to $N = 697$ uniform time steps for the time interval $[0, T]$ and $N_u = 512$ time steps for the control interval $[0, T_u]$. For the spatial computational domain, $a = 0.5$ m is chosen to consider typical scanner dimensions; the domain $[-a, a]$ is discretized using $Z = 5001$ equidistant points to achieve a homogeneous spatial resolution of $\Delta z = 0.2$ mm.

For the desired magnetization vector, we consider three examples:

Single-slice excitation To validate the design procedure, we compute an optimized pulse for a single slice of a given thickness Δ_w and a flip angle of 90° , i.e., we set

$$\tilde{M}_d(z) = \begin{cases} (0, \sin(90^\circ), \cos(90^\circ))^T & \text{if } |z| < \Delta_w/2, \\ (0, 0, 1)^T & \text{else,} \end{cases}$$

as visualized in Figure 1a. To reduce Gibbs ringing, this vector is filtered before the optimization with a Gaussian kernel with a full width at half maximum of 1.6 mm. For comparison, an SLR pulse [31, 37] with an identical temporal resolution and pulse duration is designed to the same specification (slice width, flip angle, full width at half maximum) using the Parks–McClellan

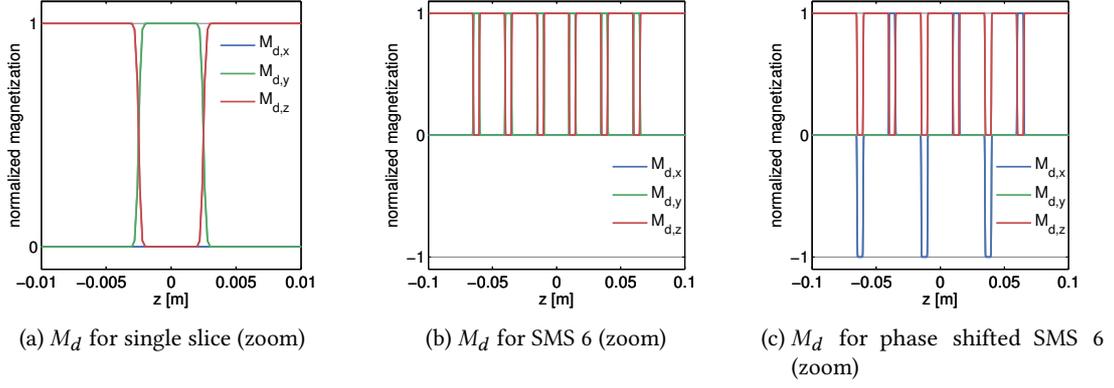


Figure 1: Desired magnetization for single- (a) and multi-slice (b,c) pulse design

(PM) algorithm [37] with a 1% in- and out-of slice ripple as usual [37] and a bandwidth of 2.35 kHz. To achieve a fully refocused magnetization, the refocusing area of the slice selective gradient for the SLR pulse is increased by 4.1 percent compared to the OC pulse (Figure 2a).

SMS excitation: phantom RF pulses for the simultaneous excitation of two, four, and six equidistant rectangular slices with a flip angle of 90° are computed, i.e., we set

$$\tilde{M}_d(z) = \begin{cases} (0, \sin(90^\circ), \cos(90^\circ))^T & \text{if } z \text{ in slice,} \\ (0, 0, 1)^T & \text{if } z \text{ out of slice,} \end{cases}$$

and apply Gauss filtering; see Figure 1b for the case of six slices.

Since PINS pulses are not suitable for axial or axial-oblique slice preference as they generate a periodic slice pattern extending outside the field of interest [49], the optimized pulses are compared with conventional SMS pulses obtained using superposed phase-shifted sinc-based excitation pulses, again for the same slice width, flip angle and full width at half maximum.

SMS excitation: in-vivo Since multi-slice in-vivo imaging using slice-GRAPPA starts to suffer from g-factor problems for more than three slices, we modify the above-described SMS pulses using a CAIPIRINHA-based excitation pattern [6], which alternates two different pulses to achieve phase-shifted magnetization vectors in order to increase the spatial distance of aliased voxels. Here, the first vector and pulse are identical to those designed for the phantom. For odd slice numbers, the second vector is modified by adding a phase term of π to every second slice of the desired magnetization, i.e.,

$$\tilde{M}_d(z) = \begin{cases} (0, \pm \sin(90^\circ), \cos(90^\circ))^T & \text{if } z \text{ in odd/even slice,} \\ (0, 0, 1)^T & \text{if } z \text{ out of slice} \end{cases}$$

(before filtering). For even slice numbers, the transverse pattern has to be further shifted by $\frac{\pi}{2}$, i.e.,

$$\tilde{M}_d(z) = \begin{cases} (\pm \sin(90^\circ), 0, \cos(90^\circ))^T & \text{if } z \text{ in even/odd slice,} \\ (0, 0, 1)^T & \text{if } z \text{ out of slice,} \end{cases}$$

see Figure 1c for the case of six slices. The additional phase shift is balanced before reconstruction by subtracting a phase of $\frac{\pi}{2}$ from every second phase-encoding line of the measured k-space data. Since typical relaxation times in the human brain are at least an order of magnitude bigger than the pulse duration, relaxation effects are neglected in the optimization.

The starting point for the optimization is chosen in all cases as $u^0 = [0, \dots, 0]$. The control cost parameter is fixed at $\alpha = 10^{-4}$ for both the single-slice and the multi-slice optimization. The parameters in Algorithm 1 are set to $\text{tol}_N = 10^{-9}$, $\text{maxit}_N = 5$, $\text{tol}_C = 10^{-6}$, $\text{maxit}_C = 50$, $\rho_0 = 1$, $\rho_{\max} = 2$, $q = 2$, $\sigma_1 = 0.03$, $\sigma_2 = 0.25$, $\sigma_3 = 0.7$.

All calculations are performed on a workstation with a four-core 64 bit processor with 3.1 GHz (Intel i5-3350P) and 16 GB of RAM.

3.2. Experimental validation

Fully sampled experimental data for a phantom and a healthy volunteer were acquired on a 3 T MR scanner (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany) using the built-in body coil to transmit the RF pulse. The MR signals were received using a body coil for the phantom experiments and a 32-channel head coil for the in-vivo experiments. A standard Cartesian GRE sequence was modified to import and apply external RF pulses. By changing the read-out gradient from the frequency-axis to the slice direction, the excited slice can be measured and visualized. The single-slice excitation was measured using a water filled sphere with a diameter of 170 mm. To acquire a high resolution in z-direction, we used a matrix size of 512×384 with a FOV of $250 \text{ mm} \times 187 \text{ mm}$ and a bandwidth of 390 Hz. The echo time was $T_E = 5 \text{ ms}$ and the repetition time $T_R = 2000 \text{ ms}$ to get fully relaxed magnetization before the next excitation. The SMS phantom experiments were performed using a homogeneous cylinder phantom with diameter of 140 mm, length of 400 mm, and relaxation times $T_1 = 102 \text{ ms}$, $T_2 = 81 \text{ ms}$, and $T_2^* = 70 \text{ ms}$. The sequence parameters were $T_E = 10 \text{ ms}$, $T_R = 1000 \text{ ms}$, bandwidth 390 Hz, matrix size 512×288 , and a field of view of $250 \text{ mm} \times 141 \text{ mm}$.

To verify the in-vivo applicability, human brain images of a healthy volunteer were acquired using the above described GRE sequence modified to include the optimized CAIPIRINHA-based pulses. The sequence parameters were set to $T_E = 10 \text{ ms}$, $T_R = 4000 \text{ ms}$, bandwidth 390 Hz, matrix size 192×120 and FOV $300 \text{ mm} \times 187 \text{ mm}$. After acquisition, the k-space data of the individual slices were separated using an offline slice-GRAPPA (32 coils, kernel size of 4×4) reconstruction [9, 42]. The reference scans used in the slice-GRAPPA reconstruction were performed with the same sequence using an optimized single-slice pulse (not shown here). To decrease the scanning time, we acquired 25 k-space lines (1/5 of the full dataset) around

the k-space center for each reference scan. After this separation, a conventional Cartesian reconstruction was performed individually for each slice.

4. Results

Single-slice excitation Figure 2 shows the results of the design of an RF pulse for the excitation of a single slice of width $\Delta_w = 5$ mm; see Figure 1a. The computed pulse (after 4 Newton iterations and a total number of 28 CG steps taking 989 s on the above-mentioned workstation is shown in Figure 2a. (To indicate the sequence timing, the slice-selective gradient G_z – although not part of the optimization – is shown dashed.) It can be seen that $u_x(t)$ is similar, but not identical, to a standard sinc shape, and that $u_y(t)$ is close to zero, which is expected due to the symmetry of the prescribed slice profile. Figure 2b contains a detail of the corresponding transverse magnetization $M_{xy}(T) = (M_x(T)^2 + M_y(T)^2)^{1/2}$ obtained from the numerical solution of the Bloch equation, which is confirmed by experimental phantom measurements in Figure 2c,d. Both simulation and measurement show an excitation with a steep transition between the in- and out-of-slice regions and a homogeneous flip angle distribution across the target slice.

Figure 3 compares the optimized (OC) pulse with a standard SLR pulse by showing details of the corresponding simulated magnetizations (Figure 3a for OC and Figure 3b for SLR; in both cases the targeted ideal magnetization is shown dashed). It can be seen that the in-slice magnetization of the optimized pulse has oscillations of higher frequency but of much smaller amplitude than that of the SLR pulse. This becomes especially visible when comparing the resulting in-slice phases (Figure 3c).

This is achieved by allowing higher ripples close to the slice while decreasing the amplitude monotonically away from the slice. (Note that only a small central segment of this region is shown in the figures.) This leads to the total root mean squared error (RMSE) and the mean absolute error (MAE) with the ideal rectangular magnetization pattern (Figure 3d) matching the full width at half maximum of both pulses being smaller for the OC pulse (1.46×10^{-2} and 1.10×10^{-4} , respectively) compared to the SLR pulse (1.62×10^{-2} and 2.27×10^{-4} , with an equal power demand for both pulses).

SMS excitation: phantom Figure 4 shows the results of the design of RF pulses for simultaneous excitation of two, four and six equidistant slices with a separation of 25 mm and a thickness $\Delta_w = 5$ mm; see Figure 1b. The computational effort in all cases is similar to that in the single-slice case. The corresponding computed pulses are shown in Figures 4a–c. A graphical analysis shows that instead of higher amplitudes, the optimization distributes the total RF power (which increases with the number of slices) more uniformly over the pulse length. A central section of the corresponding optimized slice profiles are given in Figures 4d–f. It can be seen that all slices have a sharp profile which does not deteriorate as the number of slices increases (although it decreases slightly farther from the center and the bandwidth is slightly reduced). These results are validated by the experimental phantom measurements using the computed

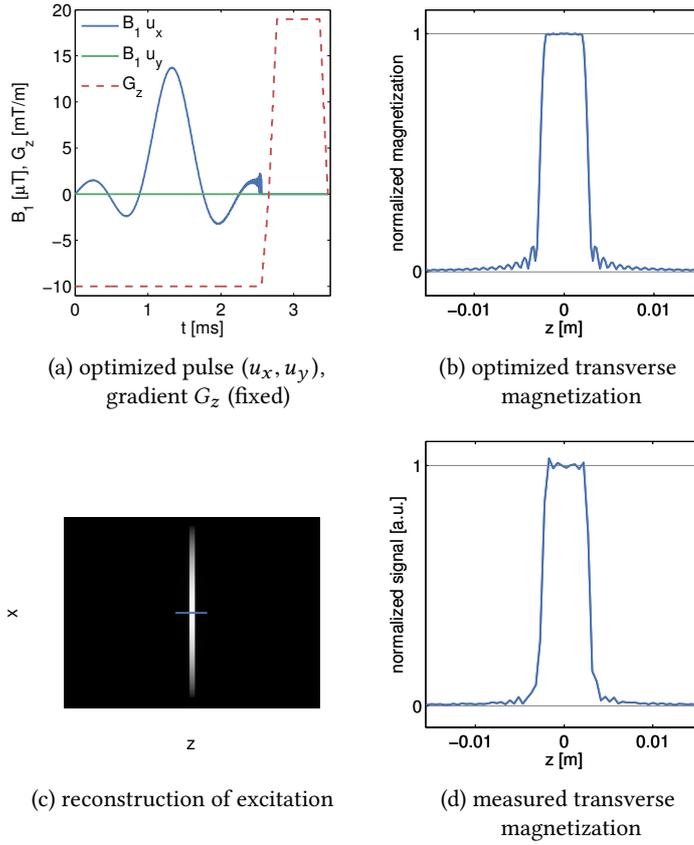


Figure 2: Optimized pulse and slice profile for single-slice excitation

pulses: Figures 4g–i show the reconstructed excitation inside the phantom, while Figures 4j–l show the measured slice profiles along a cut parallel to the x -axis in the center of the previous images.

A quantitative comparison of SLR and OC-based SMS pulses from one to six simultaneous slices is given in Table 1, which shows both the power requirement of the computed pulses, both in total B_1 energy

$$\|B_{1,x}\|_2^2 = \int_0^T |B_1 u_x(t)|^2 dt$$

and in peak B_1 amplitude

$$\|B_{1,x}\|_\infty = \max_{t \in [0, T]} |B_1 u_x(t)|,$$

as well as the mean absolute error (MAE) with respect to the ideal (unfiltered) slice profiles for the in-slice and the out-of-slice regions. While both methods lead to a linear increase of the total energy with the number of slices, the peak amplitude increases more slowly for the OC

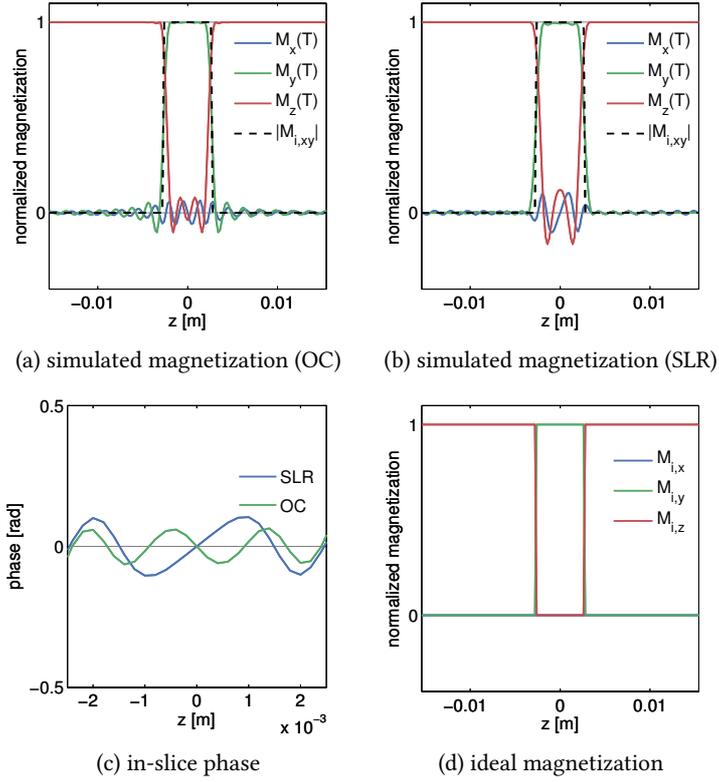


Figure 3: Comparison of SLR and OC pulse

pulses than for the conventional pulses. Furthermore, we remark that the peak B_1 amplitude for four, five and six slices remain similar. Regarding the corresponding slice profiles, the OC pulses lead to a significantly lower MAE in both the in-slice and out-of-slice regions compared to the SLR pulses. Visual inspection of Figure 4d–f shows that this is due to the fact that the out-of-slice ripples are concentrated around the in-slice regions while quickly decaying away from them.

Finally, we illustrate the influence of the regularization parameter α in Table 2, where the root of mean square error (RMSE), the total B_1 energy as well as the B_1 peak of the OC SMS 6 pulses is shown for different values of the control cost parameter α . As can be seen, a bigger α leads to an increase in the error between desired and controlled magnetization while both the total B_1 power and the peak B_1 amplitude are reduced, although these effects amount to less than 20 percent over a range of parameters spanning two orders of magnitude. This demonstrates that the results presented here are robust with respect to the choice of the control cost parameter.

SMS excitation: in-vivo The CAIPIRINHA-based modifications to the SMS pulse design (see Figure 1c) are illustrated in Figure 5 (showing the case of five slices for the sake of variation).

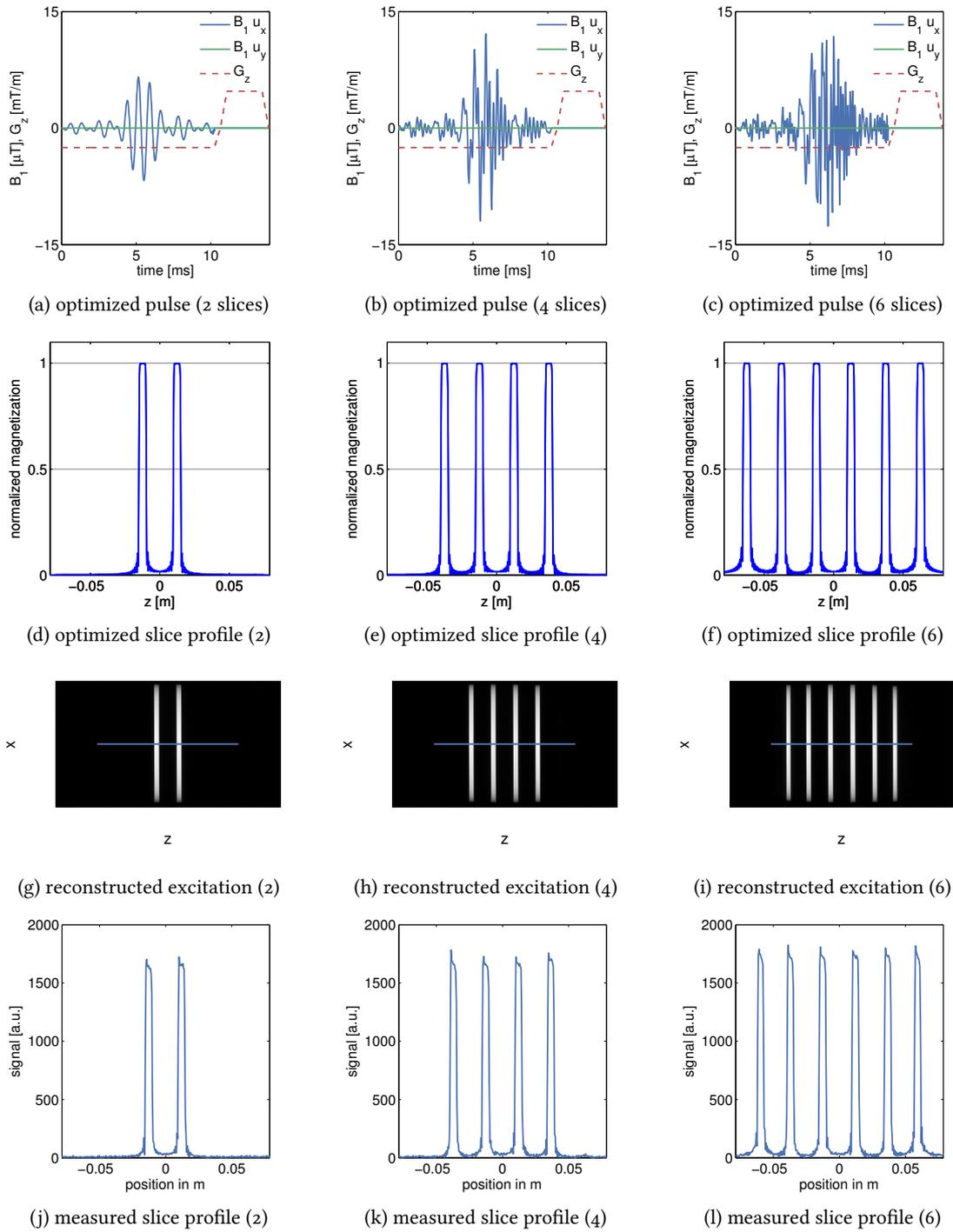


Figure 4: Optimized pulses and slice profiles for SMS excitation (phantom)

Table 1: Comparison of B_1 power and the mean absolute error (MAE) of the transverse magnetization after excitation for conventional and OC based SMS pulses

slices	$\ B_{1,x}\ _2^2$ [a.u.]		$\ B_{1,x}\ _\infty$ [μ T]		MAE in-slice [a.u.]		MAE out-of-slice [a.u.]	
	conv	OC	conv	OC	conv	OC	conv	OC
1	19.5	19.5	3.5	3.49	0.062	0.052	0.0039	0.0014
2	38.9	38.1	7.0	6.78	0.060	0.052	0.0040	0.0018
3	58.4	57.2	10.5	10.02	0.054	0.053	0.0039	0.0030
4	77.9	76.3	14.0	12.13	0.065	0.045	0.0086	0.0031
5	97.3	95.5	17.5	11.38	0.059	0.053	0.0078	0.0051
6	116.8	113.9	21.0	12.63	0.068	0.053	0.0075	0.0067

Table 2: Comparison of RMSE, B_1 power and B_1 peak for different values of α

α [a.u.]	RMSE [a.u.]	$\ B_{1,x}\ _2^2$ [a.u.]	$\ B_{1,x}\ _\infty$ [μ T]
1×10^{-5}	2.374×10^{-2}	117.0	12.75
5×10^{-5}	2.375×10^{-2}	115.1	12.71
1×10^{-4}	2.377×10^{-2}	113.9	12.62
5×10^{-4}	2.437×10^{-2}	106.7	12.14
1×10^{-3}	2.591×10^{-2}	98.9	11.63

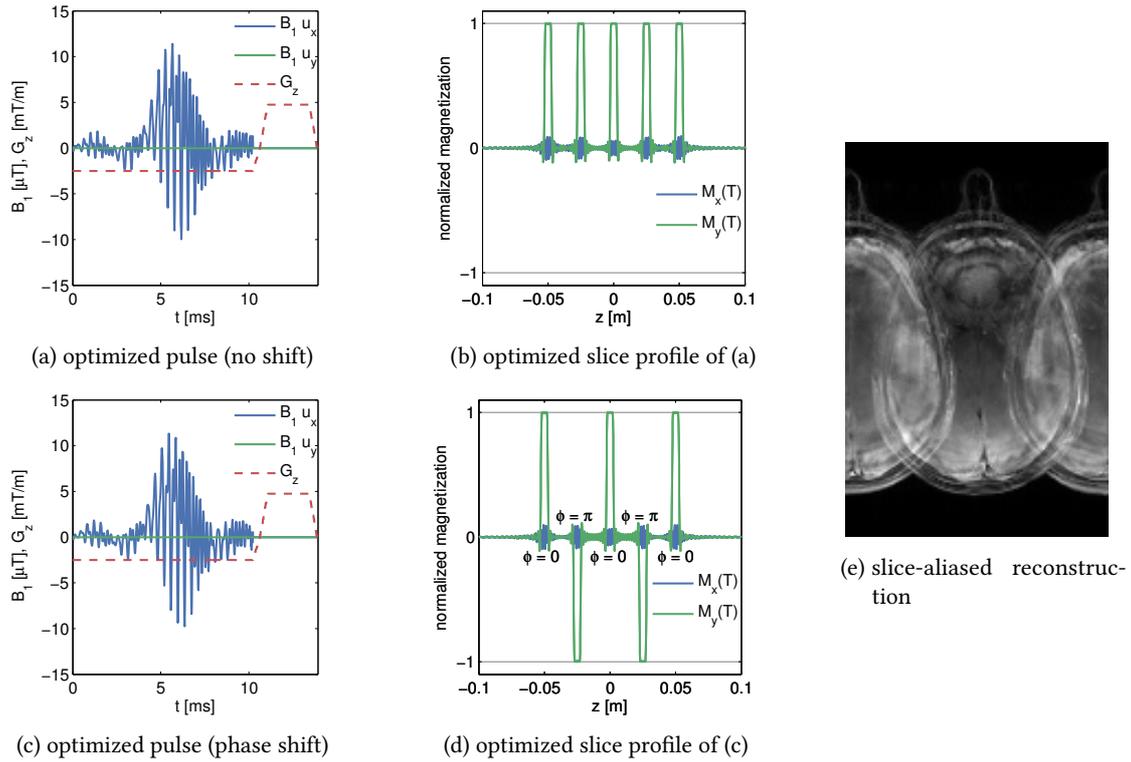


Figure 5: Optimized pulses, slice profiles and slice-aliased Cartesian reconstruction for CAIPIRINHA-based SMS excitation pattern (five slices)

Figure 5a shows the unmodified pulse, which differs in structure from the cases with an even number of slices in, e.g., Figure 4c due to the different symmetry of the slice profile (see Figure 5b). On the other hand, the pulse is very similar to the modified pulse for the alternating phase shift; see Figure 5c for the computed pulse and Figure 5d for the resulting slice profile. For illustration, a slice-aliased reconstruction of the acquired in-vivo data using this pulse sequence is shown in Figure 5e.

Figure 6 shows the image reconstruction using optimized RF pulses for simultaneous excitation of two, four and six slices with the same slice separation and thickness as above. As can be seen clearly in the first column, all three pulses lead to the desired excitation pattern in-vivo as well. The remaining columns show the slice-GRAPPA reconstructions, which illustrate that the excitation is uniform across the field of view.

5. Discussion

Our optimization approach is related to the basic ideas presented by Conolly et al. [12]. In the context of MRI, the implementation of this principle was also carried out by other groups using

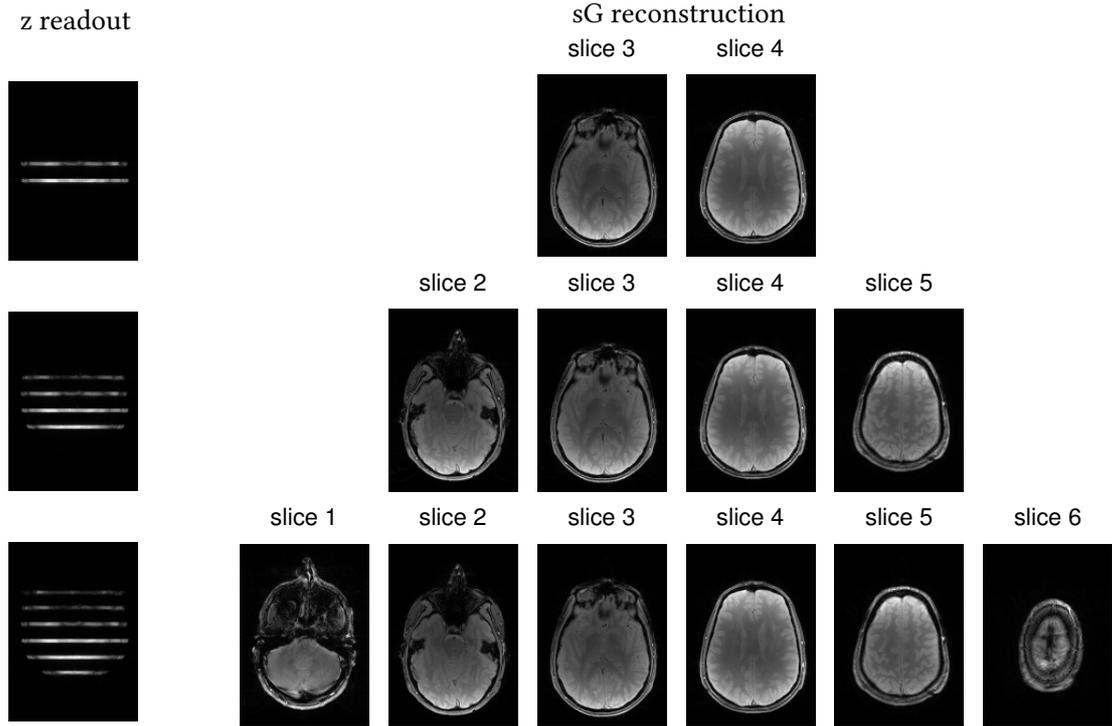


Figure 6: Slice-GRAPPA reconstruction of in-vivo data using CAIPIRINHA-based SMS excitation pattern for two (top), four (middle) and six (bottom) slices (left: conventional reconstruction showing the collapsed data in slice-encoding direction; right: reconstruction of GRAPPA-separated slices)

gradient [21,56] and quasi-Newton [51] methods. However, these methods do not make full use of second-order information and therefore achieve at best superlinear convergence. In contrast, our Newton method makes use of exact second derivatives and is therefore quadratically convergent. In particular, the main contribution of our work is the efficient computation of exact Hessian actions using the adjoint approach and its implementation in a matrix-free trust-region CG–Newton method. The use of exact derivatives speeds up convergence of the CG method, while the trust-region framework guarantees global convergence and terminates the CG method early especially at the beginning of the optimization. Both techniques save CG steps and therefore computations of Hessian actions, allowing the use of second-order information with limited computational effort and memory requirements. Since computing a Hessian action incurs the same computational cost as a gradient evaluation (i.e., the solution of two ODEs; compare (3) with (5)), we were able to compute a minimizer, e.g., for the single-slice example, with a computational effort corresponding to 32 gradient evaluations (4 for the right-hand side in each Newton iteration and 28 for the Hessian action in each CG iteration). This is less than the same number of iterations of a gradient or quasi-Newton method with line search (required in this case for global convergence), demonstrating the efficiency of the proposed approach. Therefore, our method can be used to compute RF pulses with a high temporal resolution, allowing the

design of pulses for a desired magnetization on a very fine spatial scale, in particular for the excitation of a sharp slice profile.

Furthermore, the proposed algorithm does not require an educated initial guess for global convergence (to a local minimizer, which might depend on the initial guess if more than one exists) and allows for pulse optimization in non-standard situations where no analytic RF pulse exists (e.g., for large flip angles). Compared to design methods using a simplification or approximation of the Bloch equation [37, 38], our OC based approach is capable of including relaxation terms. However, for standard in-vivo imaging applications of the human head, the relevant relaxation times are very long compared to the RF pulse duration. Thus, in our examples the influence of relaxation during excitation on the designed pulses is insignificant and has been neglected in the optimization process (although the inclusion may be indicated for other applications). The presented direct design approach allows to specify the desired magnetization in x-, y- and z-direction independently for every point in the field of view. This spatial independence of each control point allows to directly apply parallel computing to speed up the optimization process. While real-time optimization was not the aim of this work, a proof-of-concept implementation of the proposed approach on a GPU system (CUDA, double precision, GeForce GTX 550 Ti with 192 cores and 1024 MB of RAM) shows an average speedup of 135 (e.g., 6.8 s instead of 989 s for the single-slice example) while yielding identical results, thus making patient-specific design feasible as well as making the gap between OC and SLR pulse design nearly negligible. This allows efficient and fast generation of accurate slice profiles – important for minimal slice gaps, optimal contrast and low systematic errors in quantitative imaging – for arbitrary flip angles and even for specialized pulses such as refocusing or inversion.

In particular, our approach can be used to design pulses for the simultaneous excitation of multiple slices, which increases the temporal efficiency of advanced imaging techniques such as diffusion tensor imaging, functional imaging or dynamic scans. In these contexts, SMS excitation is successfully used to reduce the total imaging time [17, 30, 36, 49]; however, the peak B_1 amplitude of conventional SMS pulses is one of the main restrictions of applying SMS imaging to high-field systems [49]. The performed studies show that compared to conventional SMS design, the presented procedure yields pulses with a reduced B_1 peak amplitude (e.g., 40 % reduction for six simultaneous slices). Depending on the desired temporal resolution, the bandwidth and the slice profiles of the outer slices are slightly changed, which results in a decreased B_1 peak amplitude. It could be shown that the peak B_1 amplitude does not increase linearly with the number of slices, while the power requirement per slice remains constant and the overall power consumption is comparable to that of conventional pulses. To further reduce the SAR it is necessary to either change the excitation velocity using a time-varying slice selective gradient [13], or to extend the pulse design to parallel transmit [22, 25, 39, 54]. Furthermore, our OC-based pulses produce sharp slice profiles with a lower mean absolute error compared to the used PM-based SLR pulse, both in- and out-of-slice, at the cost of slightly larger out-of-slice ripples close to the in-slice regions. Of course, the ripple behavior of the SLR pulse can be balanced with the transition steepness by using different digital filter design methods (i.e. PM for minimizing the maximum ripple or a least squares linear-phase FIR filter for minimizing integrated squared error). The OC ripple amplitude close to the transition band can be further controlled by using offset-dependent weights as demonstrated by Skinner et al. [47]. In addition,

the computational complexity of OC methods is significantly higher than for direct or linearized methods. This implies that OC-based pulse design is advantageous in situations where high in-slice contrast and low B_1 peak amplitude are important, while SLR pulses should be used when minimal near-slice excitation and computational effort are crucial.

The presented OC approach is able to avoid some possible disadvantages of previously proposed design methods for SMS excitation. In particular, the OC design method prescribes each slice with the same uniform echo-time and phase in comparison to time-shifted [2], phase relaxation [53] and nonlinear phase design techniques [43, 57]. On the other hand, some of their features such as different echo times [2] or a non-uniform phase pattern [43, 53, 57] (e.g., for spin echo experiments) can be incorporated in our approach to further reduce the B_1 peak amplitude. It also should be possible to combine the OC design method with other techniques analogous to MultiPINS [16, 18] that combine PINS with conventional multiband pulses for a further reduction of SAR. Finally, the phantom and in-vivo experiments demonstrate that it is possible to simply replace standard pulses by optimized pulses in existing imaging sequences, and that the proposed method is therefore well suited for application in a wide range of imaging situations in MRI.

6. Conclusions

This paper demonstrates a novel general-purpose implementation of RF pulse optimization based on the full time-dependent Bloch equation and a highly efficient second-order optimization technique assuring global convergence to a local minimizer, which allows large-scale optimization with flexible problem-specific constraints. The power and applicability of this technique was demonstrated for SMS, where a reduced B_1 peak amplitude allows exciting a higher number of simultaneous slices or achieving a higher flip angle. Phantom and in-vivo measurements (on a 3 T scanner) verified these findings for optimized single- and multi-slice pulses. Even for a large number of simultaneously acquired slices, the reconstructed images show good image quality and thus the applicability of the optimized RF pulses for practical imaging applications. While the computational requirements for optimal control approaches are of course significantly greater than for, e.g., SLR-based approaches, a proof-of-concept GPU implementation indicates that this gap can be sufficiently narrowed to make patient-specific design feasible.

Due to the flexibility of the optimal control formulation and the efficiency of our optimization strategy, it is possible to consider field inhomogeneities (B_1 , B_0), design complex RF pulses for parallel transmit, or to extend the framework to include pointwise constraints due to hardware limits such as peak B_1 amplitude and slew rate.

Acknowledgments

This work is funded and supported by the Austrian Science Fund (FWF) in the context of project "SFB F3209-18" (Mathematical Optimization and Applications in Biomedical Sciences). Support from BioTechMed Graz and NAWI Graz is gratefully acknowledged. We would like to thank

Markus Böderler from the Graz University of Technology for the CUDA implementation of our algorithm.

A. Trust-region algorithm

Algorithm 1: Trust-region CG-Newton algorithm

Input: Trust region parameters $\text{tol}_N, \text{maxit}_N, \text{tol}_C, \text{maxit}_C, \rho_0, \rho_{\max}, q > 1,$

$0 < \sigma_1 < \sigma_2 < \sigma_3 < 1$

Output: Control u

```

1 Set  $u^0 \equiv 0, k = 0, g \equiv 1, \rho = \rho_0$  // initialization
2 while  $\|g\| > \text{tol}_N$  and  $k < \text{maxit}_N$  do // TR-Newton step
3   Compute gradient  $g(u^k)$ 
4   Set  $p^0 = r^0 = -g(u^k), \delta u = 0, i = 0$ 
5   while  $\|r^i\| > \text{tol}_C \|r^0\|$  and  $i < \text{maxit}_C$  do // TR-CG step
6     Compute  $H(u^k)p^i$ 
7     if  $\langle p^i, H(u^k)p^i \rangle < \varepsilon$  then // negative curvature: CG fails
8       Compute  $\max\{\tau : \|\delta u + \tau p^i\| \leq \rho\}$  // go to boundary of trust region
9       Set  $\delta u = \delta u + \tau p^i$ ; break
10    Compute  $\alpha = \|r^i\| / \langle p^i, H(u^k)p^i \rangle$ 
11    if  $\|\delta u + \alpha p^i\| \geq \rho$  then // step too large: model not trusted
12      Compute  $\max\{\tau : \|\delta u + \tau p^i\| \leq \rho\}$  // go to boundary of trust region
13      Set  $\delta u = \delta u + \tau p^i$ ; break
14    Set  $r^{i+1} = r^i - \alpha H(u^k)p^i$ 
15    Set  $p^{i+1} = r^{i+1} + \|r^{i+1}\|^2 / \|r^i\|^2 p^i$ 
16    Set  $\delta u = \delta u + \alpha p^i, i = i + 1$ 
17  Compute  $\delta J_a = J(u^k) - J(u^k + \delta u)$  // actual function decrease
18  Compute  $\delta J_m = -\frac{1}{2} \langle \delta u, H(u^k) \delta u \rangle - \langle \delta u, g(u^k) \rangle$  // predicted function decrease
19  if  $\delta J_a > \varepsilon$  and  $\delta J_a > \sigma_1 \delta J_m$  then // sufficient decrease
20    Set  $u^{k+1} = u^k + \delta u$  // accept step
21  if  $\delta J_a > \varepsilon$  and  $|\delta J_a / \delta J_m - 1| \leq 1 - \sigma_3$  then // step accepted, model good
22    Set  $\rho = \min\{q\rho, \rho_{\max}\}$  // increase radius
23  else if  $\delta J_a \leq \varepsilon$  then // step rejected, no decrease
24    Set  $\rho = \rho/q$  // decrease radius
25  else if  $\delta J_a < \sigma_2 \delta J_m$  then // model bad
26    Set  $\rho = \rho/q$  // decrease radius

```

B. Discretization

Cost functional:

$$J(M, u) = \frac{1}{2} \sum_{i=1}^Z \Delta z_i |M_{N,i} - M_d(z_i)|_2^2 + \frac{\alpha}{2} \sum_{m=1}^N \Delta t_m |u_m|_2^2$$

Bloch equation for all $i = 1, \dots, Z$:

$$\begin{aligned} \left[I - \frac{\Delta t_m}{2} A(u_m; z_i) \right] M_{m,i} &= \left[I + \frac{\Delta t_m}{2} A(u_m; z_i) \right] M_{m-1,i} + \Delta t_m b, \quad m = 1, \dots, N \\ M_{0,i} &= M^0(z_i) \end{aligned}$$

Adjoint equation for all $i = 1, \dots, Z$:

$$\begin{aligned} \left[I - \frac{\Delta t_m}{2} A(u_m; z_i)^T \right] P_{m,i} &= \left[I + \frac{\Delta t_{m+1}}{2} A(u_{m+1}; z_i)^T \right] P_{m+1,i}, \quad m = 1, \dots, N-1 \\ \left[I - \frac{\Delta t_N}{2} A(u_N; z_i)^T \right] P_{N,i} &= M_{N,i} - M_d(z_i) \end{aligned}$$

Discrete gradient for all $m = 1, \dots, N_u$: $\bar{M}_m := \frac{1}{2}(M_m + M_{m-1})$,

$$g_m = \alpha u_m + \gamma B_1 \left(\sum_{i=1}^Z \Delta z_i \begin{pmatrix} P_{m,i}^T A_1 \bar{M}_{m,i} \\ P_{m,i}^T A_2 \bar{M}_{m,i} \end{pmatrix} \right)$$

Linearized state equation for all $i = 1, \dots, Z$:

$$\begin{aligned} \left[I - \frac{\Delta t_m}{2} A(u_m; z_i) \right] \delta M_{m,i} &= \left[I + \frac{\Delta t_m}{2} A(u_m; z_i) \right] \delta M_{m-1,i} + \Delta t_m A'(\delta u_m) \bar{M}_{m,i}, \quad m = 1, \dots, N \\ \delta M_{0,i} &= 0 \end{aligned}$$

Linearized adjoint equation for all $i = 1, \dots, Z$:

$$\begin{aligned} \left[I - \frac{\Delta t_m}{2} A(u_m; z_i)^T \right] \delta P_{m,i} &= \left[I + \frac{\Delta t_{m+1}}{2} A(u_{m+1}; z_i)^T \right] \delta P_{m+1,i} + \frac{\Delta t_m}{2} A'(\delta u_m)^T P_{m,i} \\ &\quad + \frac{\Delta t_{m+1}}{2} A'(\delta u_{m+1})^T P_{m+1,i}, \quad m = 1, \dots, N-1 \\ \left[I - \frac{\Delta t_N}{2} A(u_N; z_i)^T \right] \delta P_{N,i} &= \delta M_{N,i} + \frac{\Delta t_N}{2} A'(\delta u_N)^T P_{N,i} \end{aligned}$$

Discrete Hessian action for all $m = 1, \dots, N_u$: $\delta \bar{M}_m := \frac{1}{2}(\delta M_m + \delta M_{m-1})$

$$[H(u)h]_m = \alpha h_m + \gamma B_1 \left(\sum_{i=1}^Z \Delta z_i \begin{pmatrix} \delta P_{m,i}^T A_1 \bar{M}_{m,i} + P_{m,i}^T A_1 \delta \bar{M}_{m,i} \\ \delta P_{m,i}^T A_2 \bar{M}_{m,i} + P_{m,i}^T A_2 \delta \bar{M}_{m,i} \end{pmatrix} \right)$$

References

- [1] C. K. Anand, A. D. Bain, A. T. Curtis, and Z. Nie. Designing optimal universal pulses using second-order, large-scale, non-linear optimization. *Journal of Magnetic Resonance*, 219:61–74, 2012. doi:[doi:10.1016/j.jmr.2012.04.004](https://doi.org/10.1016/j.jmr.2012.04.004).
- [2] E. J. Auerbach, J. Xu, E. Yacoub, S. Moeller, and K. Uğurbil. Multiband accelerated spin-echo echo planar imaging with reduced peak RF power using time-shifted RF pulses. *Magnetic Resonance in Medicine*, 69(5):1261–1267, 2013. doi:[10.1002/mrm.24719](https://doi.org/10.1002/mrm.24719).
- [3] P. Balchandani, J. Pauly, and D. Spielman. Designing adiabatic radio frequency pulses using the Shinnar–Le Roux algorithm. *Magnetic Resonance in Medicine*, 64(3):843–851, 2010. doi:[10.1002/mrm.22473](https://doi.org/10.1002/mrm.22473).
- [4] R. Becker, D. Meidner, and B. Vexler. Efficient numerical solution of parabolic optimization problems by finite element methods. *Optimization Methods and Software*, 22(5):813–833, 2007. doi:[10.1080/10556780701228532](https://doi.org/10.1080/10556780701228532).
- [5] B. Bonnard, M. Claeys, O. Cots, and P. Martinon. Geometric and numerical methods in the contrast imaging problem in nuclear magnetic resonance. *Acta Applicandae Mathematicae*, online first, 2014. doi:[10.1007/s10440-014-9947-3](https://doi.org/10.1007/s10440-014-9947-3).
- [6] F. A. Breuer, M. Blaimer, M. F. Mueller, N. Seiberlich, R. M. Heidemann, M. A. Griswold, and P. M. Jakob. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). *Magnetic Resonance in Medicine*, 55(3):549–556, 2006. doi:[10.1002/mrm.20787](https://doi.org/10.1002/mrm.20787).
- [7] M. H. Buonocore. RF pulse design using the inverse scattering transform. *Magnetic Resonance in Medicine*, 29(4):470–477, 1993. doi:[10.1002/mrm.1910290408](https://doi.org/10.1002/mrm.1910290408).
- [8] G. M. Bydder, G. D. Fullerton, and I. R. Young. *MRI of tissues with short T2s or T2*s*, volume 1. Wiley, 2012.
- [9] S. F. Cauley, J. R. Polimeni, H. Bhat, L. L. Wald, and K. Setsompop. Interslice leakage artifact reduction technique for simultaneous multislice acquisitions. *Magnetic Resonance in Medicine*, 72(1):93–102, 2014. doi:[10.1002/mrm.24898](https://doi.org/10.1002/mrm.24898).
- [10] G. Ciaramella and A. Borzi. SKRYN: A fast semismooth-Krylov–Newton method for controlling Ising spin systems. *Computer Physics Communications*, 190:213–223, 2015. doi:[10.1016/j.cpc.2015.01.006](https://doi.org/10.1016/j.cpc.2015.01.006).
- [11] G. Ciaramella, A. Borzi, G. Dirr, and D. Wachsmuth. Newton methods for the optimal control of closed quantum spin systems. *SIAM Journal on Scientific Computing*, 37:A319–A346, 2015. doi:[10.1137/140966988](https://doi.org/10.1137/140966988).
- [12] S. Conolly, D. Nishimura, and A. Macovski. Optimal control solutions to the magnetic resonance selective excitation problem. *Medical Imaging, IEEE Transactions on*, MI-5(2):106–115, 1986. doi:[10.1109/TMI.1986.4307754](https://doi.org/10.1109/TMI.1986.4307754).

- [13] S. Conolly, D. Nishimura, A. Macovski, and G. Glover. Variable-rate selective excitation. *Journal of Magnetic Resonance*, 78:440–458, 1988. doi:10.1016/0022-2364(88)90131-X.
- [14] P. de Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov. Second order gradient ascent pulse engineering. *Journal of Magnetic Resonance*, 212(2):412–417, 2011. doi:10.1016/j.jmr.2011.07.023.
- [15] L. de Rochefort, X. Maître, J. Bittoun, and E. Durand. Velocity-selective RF pulses in MRI. *Magnetic Resonance in Medicine*, 55(1):171–176, 2006. doi:10.1002/mrm.20751.
- [16] C. Eichner, L. L. Wald, and K. Setsompop. A low power radiofrequency pulse for simultaneous multislice excitation and refocusing. *Magnetic Resonance in Medicine*, 72(4):949–953, 2014. doi:10.1002/mrm.25389.
- [17] D. A. Feinberg and K. Setsompop. Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of Magnetic Resonance*, 229:90–100, 2013. doi:10.1016/j.jmr.2013.02.002.
- [18] B. A. Gagoski, B. Bilgic, C. Eichner, H. Bhat, P. E. Grant, L. L. Wald, and K. Setsompop. Rare/turbo spin echo imaging with simultaneous multislice wave-caipi. *Magnetic Resonance in Medicine*, 73(3):929–938, 2015. doi:10.1002/mrm.25615.
- [19] J. D. Gezelter and R. Freeman. Use of neural networks to design shaped radiofrequency pulses. *Journal of Magnetic Resonance (1969)*, 90(2):397–404, 1990. doi:10.1016/0022-2364(90)90149-4.
- [20] W. A. Grissom, G. C. McKinnon, and M. W. Vogel. Nonuniform and multidimensional Shinnar–Le Roux RF pulse design method. *Magnetic Resonance in Medicine*, 68(3):690–702, 2012. doi:10.1002/mrm.23269.
- [21] W. A. Grissom, D. Xu, A. B. Kerr, J. A. Fessler, and D. C. Noll. Fast large-tip-angle multi-dimensional and parallel RF pulse design in MRI. *Medical Imaging, IEEE Transactions on*, 28(10):1548–1559, 2009. doi:10.1109/TMI.2009.2020064.
- [22] B. Guérin, K. Setsompop, H. Ye, B. A. Poser, A. V. Stenger, and L. L. Wald. Design of parallel transmission pulses for simultaneous multislice with explicit control for peak power and local specific absorption rate. *Magnetic Resonance in Medicine*, 73(5):1946–1953, 2014. doi:10.1002/mrm.25325.
- [23] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM Journal on Control and Optimization*, 40(3):925–946, 2001. doi:10.1137/S0363012999361810.
- [24] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009. doi:10.1007/978-1-4020-8839-1.
- [25] U. Katscher, H. Eggers, I. Graesslin, G. Mens, and P. Boernert. 3D RF shimming using multi-frequency excitation. In *Proc. ISMRM 16*, page 1311, 2008.

- [26] A. A. Khalifa, A.-B. M. Youssef, and Y. M. Kadah. Optimal design of RF pulses with arbitrary profiles in magnetic resonance imaging. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 3, pages 2296–2299, 2001. doi:10.1109/IEMBS.2001.1017234.
- [27] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser. Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. *Journal of Magnetic Resonance*, 172(2):296–305, 2005. doi:10.1016/j.jmr.2004.11.004.
- [28] D. A. Knoll and D. E. Keyes. Jacobian-free Newton–Krylov methods: a survey of approaches and applications. *Journal of Computational Physics*, 193(2):357–397, 2004. doi:10.1016/j.jcp.2003.08.010.
- [29] M. Lapert, Y. Zhang, M. A. Janich, S. J. Glaser, and D. Sugny. Exploring the physical limits of saturation contrast in magnetic resonance imaging. *Scientific Reports*, 2, 2012. doi:10.1038/srep00589.
- [30] D. J. Larkman, J. V. Hajnal, A. H. Herlihy, G. A. Coutts, I. R. Young, and G. Ehnholm. Use of multicoil arrays for separation of signal from multiple slices simultaneously excited. *Journal of Magnetic Resonance Imaging*, 13(2):313–317, 2001. doi:10.1002/1522-2586(200102)13:2<313::AID-JMRI1045>3.0.CO;2-W.
- [31] K. J. Lee. General parameter relations for the Shinnar–Le Roux pulse design algorithm. *Journal of Magnetic Resonance*, 186(2):252–258, 2007. doi:10.1016/j.jmr.2007.03.010.
- [32] C. Ma and Z.-P. Liang. Design of multidimensional Shinnar–Le Roux radiofrequency pulses. *Magnetic Resonance in Medicine*, 73(2):633–645, 2015. doi:10.1002/mrm.25179.
- [33] S. Müller. Multifrequency selective RF pulses for multislice MR imaging. *Magnetic Resonance in Medicine*, 6(3):364–371, 1988. doi:10.1002/mrm.1910060315.
- [34] D. G. Nishimura. *Principles of Magnetic Resonance Imaging*. Stanford University, 1996.
- [35] D. G. Norris, P. J. Koopmans, R. Boyacioğlu, and M. Barth. Power independent of number of slices (PINS) radiofrequency pulses for low-power simultaneous multislice excitation. *Magnetic Resonance in Medicine*, 66(5):1234–1240, 2011. doi:10.1002/mrm.23152.
- [36] R. Nunes, J. V. Hajnal, X. Golay, and D. J. Larkman. Simultaneous slice excitation and reconstruction for single shot EPI. In *Proc. ISMRM 14*, page 293, 2006.
- [37] J. Pauly, P. Le Roux, D. Nishimura, and A. Macovski. Parameter relations for the Shinnar–Le Roux selective excitation pulse design algorithm. *Medical Imaging, IEEE Transactions on*, 10(1):53–65, 1991. doi:10.1109/42.75611.
- [38] J. Pauly, D. Nishimura, and A. Macovski. A k-space analysis of small-tip-angle excitation. *Journal of Magnetic Resonance (1969)*, 213(2):544–557, 1989. doi:10.1016/0022-2364(89)90265-5.

- [39] B. A. Poser, R. J. Anderson, B. Guérin, K. Setsompop, W. Deng, A. Mareyam, P. Serano, L. L. Wald, and V. A. Stenger. Simultaneous multislice excitation by parallel transmission. *Magnetic Resonance in Medicine*, 71(4):1416–1427, 2014. doi:10.1002/mrm.24791.
- [40] F. Santosa and W. W. Symes. Computation of the Hessian for least-squares solutions of inverse problems of reflection seismology. *Inverse Problems*, 4(1):211, 1988. doi:10.1088/0266-5611/4/1/017.
- [41] K. Setsompop, J. Cohen-Adad, B. A. Gagoski, T. Raij, A. Yendiki, B. Keil, V. J. Wedeen, and L. L. Wald. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *NeuroImage*, 63(1):569–580, 2012. doi:10.1016/j.neuroimage.2012.06.033.
- [42] K. Setsompop, B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen, and L. L. Wald. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic Resonance in Medicine*, 67(5):1210–1224, 2012. doi:10.1002/mrm.23097.
- [43] A. Sharma, M. Lustig, and W. A. Grissom. Root-flipped multiband refocusing pulses. *Magnetic Resonance in Medicine*, Early View, 2015. doi:10.1002/mrm.25629.
- [44] A. Sharma, R. O’Halloran, E. Aboussouan, A. Tu Van, J. Maclaren, M. Aksoy, V. A. Stenger, R. Bammer, and W. A. Grissom. kT-PINS RF pulses for low-power field inhomogeneity-compensated multislice excitation. In *Proc. ISMRM 21*, page 44, 2013.
- [45] J. Shen. Delayed-focus pulses optimized using simulated annealing. *Journal of Magnetic Resonance*, 149(2):234–238, 2001. doi:10.1006/jmre.2001.2306.
- [46] T. E. Skinner, N. I. Gershenzon, M. Nimbalkar, W. Bermel, B. Luy, and S. J. Glaser. New strategies for designing robust universal rotation pulses: Application to broadband refocusing at low power. *Journal of Magnetic Resonance*, 216:78–87, 2012. doi:10.1016/j.jmr.2012.01.005.
- [47] T. E. Skinner, N. I. Gershenzon, M. Nimbalkar, and S. J. Glaser. Optimal control design of band-selective excitation pulses that accommodate relaxation and rf inhomogeneity. *Journal of Magnetic Resonance*, 217:53–60, 2012. doi:10.1016/j.jmr.2012.02.007.
- [48] T. Steihaug. The conjugate gradient method and trust regions in large scale optimization. *SIAM Journal on Numerical Analysis*, 20(3):626–637, 1983. doi:10.1137/0720042.
- [49] K. Ugurbil, J. Xu, E. J. Auerbach, S. Moeller, A. T. Vu, J. M. Duarte-Carvajalino, C. Lenglet, X. Wu, S. Schmitter, P. F. Van de Moortele, J. Strupp, G. Sapiro, F. De Martino, D. Wang, N. Harel, M. Garwood, L. Chen, D. A. Feinberg, S. M. Smith, K. L. Miller, S. N. Sotiropoulos, S. Jbabdi, J. L. R. Andersson, T. E. J. Behrens, M. F. Glasser, D. C. Van Essen, and E. Yacoub. Pushing spatial and temporal resolution for functional and diffusion MRI in the human connectome project. *NeuroImage*, 80:80–104, 2013. doi:10.1016/j.neuroimage.2013.05.012.

- [50] J. L. Ulloa, M. Guarini, A. Guesalaga, and P. Irarrazaval. Chebyshev series for designing RF pulses employing an optimal control approach. *Medical Imaging, IEEE Transactions on*, 23(11):1445–1452, 2004. doi:10.1109/TMI.2004.835602.
- [51] M. S. Vinding, I. I. Maximov, Z. Tošner, and N. C. Nielsen. Fast numerical design of spatial-selective RF pulses in MRI using Krotov and quasi-Newton based optimal control methods. *Journal of Chemical Physics*, 137(5):054203, 2012. doi:10.1063/1.4739755.
- [52] Z. Wang, I. Navon, F. Le Dimet, and X. Zou. The second order adjoint analysis: theory and applications. *Meteorology and Atmospheric Physics*, 50(1–3):3–20, 1992. doi:10.1007/BF01025501.
- [53] E. Wong. Optimized phase schedules for minimizing peak RF power in simultaneous multi-slice RF excitation pulses. In *Proc. ISMRM 20*, page 2209, 2012.
- [54] X. Wu, S. Schmitter, E. J. Auerbach, S. Moeller, K. Uğurbil, and P.-F. Van de Moortele. Simultaneous multislice multiband parallel radiofrequency excitation with independent slice-specific transmit B1 homogenization. *Magnetic Resonance in Medicine*, 70(3):630–638, 2013. doi:10.1002/mrm.24828.
- [55] X.-L. Wu, P. Xu, and R. Freeman. Delayed-focus pulses for magnetic resonance imaging: An evolutionary approach. *Magnetic Resonance in Medicine*, 20(1):165–170, 1991. doi:10.1002/mrm.1910200118.
- [56] D. Xu, K. F. King, Y. Zhu, G. C. McKinnon, and Z.-P. Liang. Designing multichannel, multi-dimensional, arbitrary flip angle RF pulses using an optimal control approach. *Magnetic Resonance in Medicine*, 59(3):547–560, 2008. doi:10.1002/mrm.21485.
- [57] K. Zhu, A. B. Kerr, and J. M. Pauly. Nonlinear-phase multiband 90° – 180° RF pair with reduced peak power. In *Proc. ISMRM 22*, page 1440, 2014.