On the Application of Compressed Sensing to Magnetic Resonance Imaging

Dissertation zur Erlangung

des naturwissenschaftlichen Doktorgrades

der Julius-Maximilians-Universität Würzburg

vorgelegt von

André Fischer

aus Rothenburg ob der Tauber

Würzburg, 2011

Eingereicht am: 30.11.2011
bei der Fakultät für Physik und Astronomie

1. Gutachter: Prof. Dr. Peter M. Jakob

2. Gutachter: Prof. Dr. Herbert Köstler
der Dissertation

1. Prüfer: Prof. Dr. Peter M. Jakob

2. Prüfer: Prof. Dr. Herbert Köstler

3. Prüfer: Prof. Dr. Haye Hinrichsen
im Promotionskolloquium

Tag des Promotionskolloquiums: 20.07.2012

Doktorurkunde ausgehändigt am: ____________________________

“The more I know, the less I understand
All the things I thought I knew, I’m learning again
I’ve been tryin’ to get down to the heart of the matter
But my will gets weak
And my thoughts seem to scatter“

Excerpt from: “The Heart of the Matter“ (1990)
From the album: “The End of the Innocence“

Author: Don Henley (22.07.1947)*
American Singer and Songwriter

Contents

Introduction

Magnetic Resonance Imaging (MRI) is now an established modality in clinical diagnostics as well as in medical, biological, chemical, and physical research. It is a non-invasive technique leaving the patient/volunteer/sample unharmed. Furthermore, it carries no radiation burden unlike Computed Tomography (CT). Additionally, radioactive markers are unnecessary unlike Positron Emission Tomography (PET) or Single Photon Emission Computed Tomography (SPECT). However, the relatively slow imaging compared to methods such as CT has always been a drawback of MRI. In the early days of MRI, a typical imaging experiment took several minutes. This was unsuitable for wide clinical application because patients must lie absolutely still to obtain artifact-free images. Obviously, in many cases this is an unattainable requirement.

With the discovery of fast imaging sequences such as Fast Low Angle Shot (FLASH [1]) and Rapid Acquisition with Relaxation Enhancement (RARE [2]), the drawback of slow imaging speed could be overcome. FLASH is a gradient echo sequence that benefits from a significantly lower excitation flip angle than the 90° pulses usually used in spin echo sequences. Thus, the long latency times during which the longitudinal magnetization relaxes and rebuilds can be discarded. Furthermore, the T_R times can be subsequently reduced from the typically used $T_R \approx 5 \cdot T_1$ (several 100 ms to s) to a couple of ms. This allowed an enormous acceleration of the imaging process, enabling the acquisition of a fully sampled image within several hundred ms instead of minutes. However, this benefit comes at the expense of a lower Signal-to-Noise-Ratio (SNR) since only a fraction of the longitudinal magnetization is excited.
The RARE sequence is a spin echo technique. Instead of acquiring one echo per T_R, an echo train consisting of multiple, differently phase encoded echoes is recorded. Because of this characteristic, the RARE technique is also known as Turbo Spin Echo (TSE) or Fast Spin Echo (FSE). RARE leads to an intrinsic T_2 weighting of each individual echo that must be taken into consideration to reduce imaging artifacts. However, the SNR in the RARE images is comparable to conventional spin echo images if the first echo is reordered to the k-space center. This echo is then responsible for the overall SNR and contrast.

Although the advance in imaging speed opened new possibilities for clinical and research use, further acceleration of the data acquisition process remained important. Further acceleration would allow real-time imaging; detection of fast dynamic processes such as cardiac motion; and further reduction of artifact susceptibility due to patient motion. Because of the MR scanner hardware and safety margins for the Specific Absorption Rate (SAR) set by the World Health Organization (WHO), restrictions are placed on the potential acceleration of the imaging sequences. Faster switching of the gradient fields would be technically possible, but would induce currents in the nervous system of patients leading to undesirable side effects such as nervous stimulations. Applying more radiofrequency radiation for the excitation process is not permitted by the SAR limiting value. Thus, the data acquisition can be further accelerated only by reducing the amount of sampled data. The missing data must therefore be appropriately reconstructed by a suitable scheme.

In 1990, the NMR phased array was introduced by Roemer [3]. Instead of using a single

receiver coil, an array of multiple receivers is used to detect the MR signal. Every array element has an individual coil sensitivity profile. By properly arranging the coil array so that the coil sensitivities sufficiently vary, an intrinsic spatial encoding can be achieved by the array. This was successfully exploited by the Simultaneous Acquisition of Spatial Harmonics (SMASH [4]) technique. It was the first successful parallel imaging method and is k-space based. In SMASH, only every n^{th} phase encoding step is acquired during data acquisition, which consequently reduces the measurement time by a factor of n. This introduces aliasing in image space due to the reduction of the Field-Of-View (FOV) by a factor of n. From a fully sampled reference scan and the measured coil sensitivity profiles, weighting factors are derived and used to calculate additional spatial harmonics. Those harmonics shift the acquired k-space lines by $\Delta k, 2 \cdot \Delta k, ..., (n-1) \cdot \Delta k$. Thereby, the missing lines in k-space are reconstructed and an unaliased image can be obtained. The method, however, originally suffered from remaining aliasing artifacts due to errors in the weighting factors. SMASH was further optimized by the self-calibrating AUTO-SMASH [5] and VD-AUTO-SMASH [6] approaches.
Sensitivity Encoding (SENSE [7]), an image-space based formulation of the parallel imaging problem, overcame the drawbacks of SMASH. Similar to SMASH, the coil sensitivities must be explicitly measured for use in the reconstruction process. With SENSE, this reconstruction is a robust and SNR-optimized matrix inversion that is used to unfold the aliased images.
A general implementation and consequent advancement of AUTO-SMASH and VD-AUTO-SMASH was Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA [8]). With GRAPPA, the missing lines in k-space are reconstructed by a weighted sum of measured k-space lines of all coils. The required weights can be determined from a small, fully sampled region in the k-space center, the so-called Autocalibrating Signal (ACS) lines first proposed by the AUTO-SMASH approach. In contrast to SENSE, AUTO-SMASH and VD-AUTO-SMASH, this allowed reconstruction of unaliased images for every individual coil. Unlike SENSE or SMASH, no coil sensitivities need to be determined since they are implicitly included in the GRAPPA weights.
GRAPPA and SENSE are the most commonly used parallel imaging techniques. Besides the two mentioned methods, there exist a variety of other techniques such as PILS [9], SPACE RIP [10], Generalized SMASH [11], and mSENSE [12].

The intrinsic spatial encoding of NMR phased arrays that is exploited in parallel imaging allows the accurate reconstruction of undersampled data when the acceleration factor does not exceed the number of coils in the array. This is because each individual coil “sees“ the whole object that is weighted with the respective coil sensitivity profile from a different view. Thus, the object information is encoded in each coil, which introduces, if fully acquired, redundancy in the acquired data. However, a further reduction of the acquired data is only possible if something about the object to be imaged is known *a priori*. This so-called “prior knowledge“ can be obtained by methods such as acquiring a training dataset closely related to the desired undersampled dataset. An example for this would be k-t BLAST or k-t-SENSE [13]. For these methods, a low spatial resolution training dataset of the dynamic event is acquired in the desired temporal resolution and used as prior knowledge. This knowledge is then used in the data reconstruction process of the undersampled high spatial resolution dataset.

A new way to exploit prior knowledge to accelerate the data acquisition process was developed in the signal processing community. The researchers investigated so-called “sparse“ signals. These signals can be represented by only a small fraction of signal coefficients (in any arbitrary mathematical basis). Their number is significantly smaller than the number of all signal coefficients. A surprising result led to the discovery of Compressed Sensing (CS, also termed Compressive Sampling, Compressive Sensing, or Compressed Sampling). Under certain conditions, sparse signals can be recovered from undersampled data that violate the Nyquist-Shannon criterion. The pioneer works in this field were published by Emmanuel Candès and David Donoho [14, 15]. In contrast to parallel imaging, the reconstruction process does not require coil information. In fact, the intrinsic redundancy in the k-space data of a sparse signal is exploited. Each k-space point is a superposition of the intensities of all image pixels, with each pixel weighted by the respective Fourier coefficient. Thus, each k-space point carries information about the complete object. The Fourier transform is a (possibly) large linear system of equations. In order to exactly solve the linear system of equations, all Fourier coefficients must be acquired to obtain a unique solution. However, if the signal is sparse in image space, most pixel values are zero (in the ideal noise-free case). If which of the pixels in the image are zero is known, fewer equations of the linear system would be sufficient to obtain the unique solution by matrix inversion. However, which of the pixels in the image are zero is unknown. If it is known *a priori* that the desired image is sparse, the image can be recovered by minimizing a cost function that models the sparsity of the image (in any arbitrary mathematical basis). The minimization is iteratively achieved by appropriate algorithms. This means a drastic change from reconstruction techniques based purely on physical information toward signal processing techniques that make *a priori* assumptions of the desired image and then model it accordingly. This was uncommon in the MR community before 2007.

The rise of CS triggered great research efforts in this new field. In a short period of time, contributions to the ISMRM Annual Meetings regarding CS exploded. The amount of presentations at the annual congress exclusively dealing with CS demonstrate the enormous hopes and research activities currently in this field. This thesis discusses and explains the possibilities and limitations of CS as well as several MR applications that promise to benefit from CS. These applications are not restricted to clinically relevant applications as CS can also be helpful for topics in basic clinical research. Furthermore, an approach to combine CS with parallel imaging is introduced.

Outline of this thesis

After introducing the basic principles of MR imaging relevant to this thesis, the concept of CS is explained. A special emphasis is laid on the three important requirements for CS. These are sparsity, sampling pattern, and reconstruction. Two different algorithms were considered for reconstructing the undersampled sparse data with the CS concept. In this thesis, the Nonlinear Conjugate Gradient based technique with a relaxed data consistency constraint as suggested by Lustig et al. [16] is termed Relaxed DC method. An alternative represents the Gradient or Steepest Descent algorithm with strict data consistency and is, therefore, termed the Strict DC method. Chapter 3 presents simulations illustrating which of these two reconstruction algorithms is best suited to recover undersampled sparse MR datasets. The results lead to the decision for the Strict DC method as reconstruction technique in this thesis.

After laying the basics and deciding on the optimal reconstruction algorithm for CS, different applications and extensions of CS are demonstrated. Chapter 4 shows how CS benefits spectroscopic ^{19}F imaging at 7 T, allowing a significant reduction of measurement times during *in vivo* experiments. Furthermore, it allows highly resolved spectroscopic 3D imaging in acceptable measurement times for *in vivo* applications.

Chapter 5 introduces an extension of the Strict DC method called CS-CC (CS on Combined Coils), which allows efficient processing of sparse undersampled multi-coil data. It takes advantage of a concept named “Joint Sparsity“, which exploits the fact that all channels of a coil array detect the same sparse object weighted with the coil sensitivity profiles. The practical use of this new algorithm is demonstrated in dynamic radial cardiac imaging. Accurate reconstructions of cardiac motion in free breathing without ECG triggering were obtained for high undersampling factors.

An Iterative GRAPPA algorithm is introduced in Chapter 6 that can recover undersampled data from arbitrary (Non-Cartesian) trajectories and works solely in the Cartesian plane. This characteristic makes the proposed Iterative GRAPPA computationally more efficient than SPIRiT [17]. Iterative GRAPPA was developed in a preceding step to combine parallel imaging with CS. Optimal parameters for Iterative GRAPPA (e.g. number of iterations, GRAPPA kernel size) will be determined in phantom experiments and verified by retrospectively undersampling and reconstructing a radial cardiac cine dataset.

The synergistic combination of the coil-by-coil Strict DC CS method and Iterative GRAPPA called CS-GRAPPA is presented in Chapter 7. CS-GRAPPA allows accurate reconstruction of undersampled data from even higher acceleration factors than each individual method. It is a formulation equivalent to L_1-SPIRiT [17–19] but computationally more efficient. Additionally, a comparison with CS-CC is given. Interestingly, exploiting joint sparsity in CS-CC is slightly more efficient than the proposed CS-GRAPPA, a hybrid of parallel imaging and CS.

The last chapter of this thesis concludes the findings presented in this dissertation. Future applications expected to benefit from CS are discussed and possible synergistic combinations with other existing MR methodologies for accelerated imaging are also contemplated.

1 Basic Principles of MRI

This chapter explains the physical basics of magnetic resonance imaging (MRI) relevant to this work. In MRI, the quantity $\vec{B}$ is usually referred to as the "magnetic field" even though the correct physical term is "magnetic flux density". Both quantities are connected by

$$\vec{B} = \mu_R \cdot \mu_0 \cdot \vec{H} \qquad \mu_o = 4\pi \cdot 10^{-7} \frac{\mathrm{Vs}}{\mathrm{Am}} \tag{1.1}$$

where μ_R is the dimensionless relative permeability of the sample present in the magnetic field $\vec{H}$ and μ_0 is the permeability of the vacuum. The dimensions of both quantities are given as $\left[\vec{B}\right] = \frac{\mathrm{Vs}}{\mathrm{m}^2} = \mathrm{T}$ and $\left[\vec{H}\right] = \frac{\mathrm{A}}{\mathrm{m}}$.

It is recognized that the nomenclature usually used in the MR community is physically incorrect. However, this thesis will follow the well-established terms of the community by referring to $\vec{B}$ as the magnetic field.

In this dissertation, only ^{1}H and ^{19}F MRI experiments were performed. Both nuclei exhibit the spin $s = 1/2$. Therefore, the following discussion of MR basics is focused on spin 1/2 particles.

1.1 Magnetic resonance

The phenomenon of magnetic resonance was discovered by Bloch and Purcell in 1946 [20, 21]. They observed that a nucleus with a non-vanishing nuclear spin precesses around the direction of a static external magnetic field $\vec{B}_0$. A spin 1/2 particle can either align parallel to or anti-parallel to $\vec{B}_0$. Thus, due to the Zeeman splitting two energy levels emerge that are separated by

$$\Delta E = \hbar\omega_0 \tag{1.2}$$

In this equation, the so-called Larmor frequency can be identified as

$$\omega_0 = 2\pi\gamma B_0 \tag{1.3}$$

where γ is the gyromagnetic ratio given in the meaningful units MHz T^{-1}. The gyromagnetic ratio is unique for each contemplated nucleus. In this work, ^{1}H and ^{19}F were the nuclei of interest with $\gamma_{^1\mathrm{H}} = 42.58$ MHz T^{-1} and $\gamma_{^{19}\mathrm{F}} = 40.04$ MHz T^{-1}. Equation 1.3 reveals the important relationship between the resonance frequency and the applied external field strength $|\vec{B}_0|$.

The spins precess around the direction of the external magnetic field because of the quantum mechanical nature of nuclear spins. The Eigenvalue of the spin operator $\hat{S}$ is given by

$$|\hat{S}| = \sqrt{s\,(s+1)} \cdot \hbar = \sqrt{\frac{3}{4}}\hbar \tag{1.4}$$

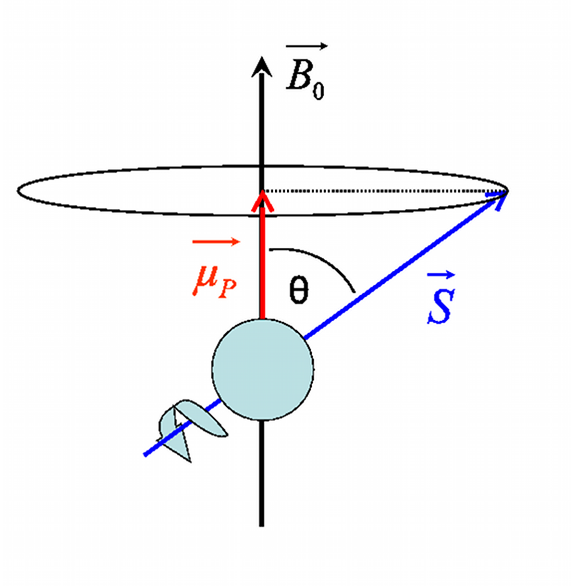

Figure 1: *Semi-classical illustration of a spin in an external magnetic field $\vec{B}_0$. $\vec{\mu}_p$ is hereby the magnetic moment of a particle created by the z-component of the spin $\vec{S}$.*

The only observable component of $\hat{S}$ is the z-component $\vec{S}_z$, which results in a magnetic moment of the observed particle,

$$\vec{\mu}_p = 2\pi\gamma\vec{S}_z \tag{1.5}$$

This magnetic moment $\vec{\mu}_p$ aligns either parallel to or anti-parallel to the main magnetic field. This corresponds to the magnetic quantum numbers $m = \pm 1/2$, where the favorable state in terms of energy is the $m = +1/2$ (parallel aligned) state. A semi-classical illustration is given in Figure 1. As seen, the spin $\vec{S}$ is incompletely aligned along the magnetic field direction $\vec{B}_0$. The angle Θ (in this semi-classical illustration) between the magnetic field direction and the spin direction can be calculated by

$$\Theta = \arccos\frac{m}{|\hat{S}|} = 54.74^\circ\ (m = +1/2)\ \text{ and } 125.26^\circ\ (m = -1/2) \tag{1.6}$$

Thus, it can be semi-classically explained why the external magnetic field $\vec{B}_0$ creates a torque on the spin $\vec{S}$ which results in the precession.
The quantitative behavior of a macroscopic number of spins in an external magnetic field can be calculated using the Bloch equations

$$\frac{d}{dt}\begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} = 2\pi\gamma \cdot \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} \times \begin{pmatrix} B_{0,x} \\ B_{0,y} \\ B_{0,z} \end{pmatrix} - \begin{pmatrix} \frac{M_x}{T_2} \\ \frac{M_y}{T_2} \\ \frac{M_z - M_0}{T_1} \end{pmatrix} \tag{1.7}$$

M_0 denotes the equilibrium magnetization that occurs by placing a sample in a magnetic field $\vec{B}_0$ without disturbing the spin ensemble. The vector on the far right side explicitly considers relaxation effects due to T_1 (spin-lattice interaction) and T_2 (spin-spin interaction). By convention, the direction of $\vec{B}_0$ is the z-direction, i.e. $B_{0,x} = B_{0,y} = 0$ T. This leads to the following simplifications:

$$\frac{dM_x}{dt} = 2\pi\gamma \cdot B_0 \cdot M_y - \frac{M_x}{T_2} \tag{1.8}$$

$$\frac{dM_y}{dt} = -2\pi\gamma\!\!\!\!- \cdot B_0 \cdot M_x - \frac{M_y}{T_2} \tag{1.9}$$

$$\frac{dM_z}{dt} = \frac{M_0 - M_z}{T_1} \tag{1.10}$$

Equation 1.10 can be solved by simply separating the variables and integrating. This results in:

$$M_z(t) = M_z(t=0) \cdot e^{-\frac{t}{T_1}} + M_0 \left(1 - e^{-\frac{t}{T_1}}\right) \tag{1.11}$$

This equation models the so-called longitudinal relaxation. The characteristic time constant describing this process is T_1. If spins are excited by radiofrequency (rf) radiation with the Larmor frequency ω_0, then M_z is no longer equal to M_0 because spins were excited from the lower energy state (m = +1/2, parallel) to the higher state (m = -1/2, anti-parallel) and tend to relax toward the energetically lowest state. This is caused by the interaction between the spin and the environment (the lattice) leading to the term “spin-lattice relaxation“. Spins that have been excited are then aligned anti-parallel to the main magnetic field direction and relax to the energetically lower parallel-aligned state. After infinite time, the thermodynamical equilibrium is again reached.

The x- and y-components can be solved, as with Equation 1.10, by introducing the transversal magnetization:

$$M_\perp(t) \equiv M_x(t) + i \cdot M_y(t) \tag{1.12}$$

This immediately leads to

$$\frac{dM_\perp(t)}{dt} = \frac{dM_x(t)}{dt} + i \cdot \frac{dM_y(t)}{dt} = \left(-i \cdot 2\pi\gamma\!\!\!\!- \cdot B_0 - \frac{1}{T_2}\right) M_\perp(t) \tag{1.13}$$

This differential equation is once more solvable after separating the variables and integrating:

$$M_\perp(t) = M_\perp(t=0) \cdot e^{-i2\pi\gamma\!\!\!\!- B_0 t - \frac{t}{T_2}} = M_\perp(t=0) \cdot e^{-i\omega_0 t - \frac{t}{T_2}} \tag{1.14}$$

Usually, the precession of the spins is observed in the rotating reference frame with frequency ω_0, resulting in

$$M_\perp(t) = M_\perp(t=0) \cdot e^{-\frac{t}{T_2}} \tag{1.15}$$

This equation models the decay of transversal magnetization in a perfectly homogeneous magnetic field. The decay is a result of locally varying magnetic fields caused by the randomly moving spins (the so-called spin-spin interaction). Therefore, this dephasing of transversal magnetization is termed spin-spin relaxation. Unfortunately, the main magnetic field of a MR scanner is not perfectly homogeneous. Thus, spins at different positions in the scanner exhibit (slightly) different resonance frequencies due to these B_0 inhomogeneities. Furthermore, in addition to the B_0 imperfections, all external effects causing local inhomogeneities in the magnetic field result in an accelerated decay of transversal magnetization compared to pure T_2 signal decay. This can be modeled by:

$$T_2^* = \left(\frac{1}{T_2} + \frac{1}{T_2'}\right)^{-1} \tag{1.16}$$

T_2^*, instead of T_2, is thus the actually observed relaxation time that should be considered in Equation 1.7. T_2' models the loss of magnetization due to inhomogeneities induced by all possible sources of external fields. This decay is reversible since these inhomogeneities are constant over time.

1.2 Signal intensity

In Equation 1.7, the equilibrium magnetization M_0 is introduced. M_0 is directly proportional to the observed signal intensity S:

$$S \propto M_0 \tag{1.17}$$

Thermodynamic considerations using the Boltzmann distribution lead to a quantitative expression for M_0:

$$M_0 = \rho_0 2\pi\gamma\hbar \cdot \frac{\sum_{m=-s}^{s} m \cdot e^{m\cdot\frac{\hbar\omega_0}{kT}}}{\sum_{m=-s}^{s} e^{m\cdot\frac{\hbar\omega_0}{kT}}} \tag{1.18}$$

where k is the Boltzmann constant ($k = 1.3807{\cdot}10^{-23}$ J K^{-1}), T is the absolute temperature in K, s is the spin quantum number (1/2 for the purposes of this thesis), and ρ_0 is the physical spin density (number per volume).

All experiments in this thesis were performed at approximately $T = 300$ K and at high magnetic field strengths ($|\vec{B}| \geq 1.5$ T). This allows approximation $\hbar\omega_0 \ll kT$ and leads, after Taylor expansion of Equation 1.18 and using $s = 1/2$, to :

$$M_0 \simeq \rho_0 \frac{4\pi^2\gamma^2\hbar^2}{4kT} B_0 \qquad (\hbar\omega_0 \ll kT) \tag{1.19}$$

Since $M_0 \propto \gamma^2$, this equation demonstrates that the choice of the nucleus directly influences the expected signal level . Thus, nuclei with a high gyromagnetic ratio and a high ρ_0 are preferable in terms of signal intensity.
At $T = 300$ K and $B_0 = 1.5$ T, only 5 out of 1.000.000 nuclei are aligned parallel to the main magnetic field. All others, however, fluctuate randomly due to the thermal energy. Nonetheless, the large number of nuclei in the sample volume makes tomographic images possible. These images exhibit a signal intensity directly proportional to the physical spin density ρ_0 (obeying Equations 1.17 to 1.19):

$$S \propto M_0 \propto \rho_0$$

Variations in the spin density are one of the mechanisms that create contrast in the final MR image. Other possible mechanisms introducing different contrasts are T_1, T_2, T_2^*, and diffusion.

1.3 Spatial encoding and Fourier transform

Equation 1.3 states that the precession frequency is proportional to $|\vec{B}_0|$. This leads to the idea of spatial encoding by varying the local magnetic field strength by an additional spatially varying field. One simple way to achieve this spatial encoding is to superimpose a linearly varying magnetic field, a so-called gradient field, to the existing main magnetic field. Consider this situation in 1D:

$$B_x(x,t) = B_0 + x \cdot G_x(t) \qquad G_x(t) \equiv \frac{\partial B_x}{\partial x} \tag{1.20}$$

Please note that the direction of the gradient field G_x and the main magnetic field B_0 is the z-direction (by convention). Only the amplitude of the field varies along the x-direction. When switched on, the gradient field must not vary. However, the explicit time dependency present in Equation 1.20 accounts for the possibility of switching the gradient on and off. Following Equation 1.3 leads immediately to:

$$\begin{aligned} \omega_x(x,t) &\equiv \omega_0 + \omega_G(x,t) \\ &= \omega_0 + 2\pi\bar{\gamma} x G_x(t) \end{aligned} \tag{1.21}$$

Thus, the Larmor frequency varies linearly with the spatial position. This can be exploited by considering the 1D imaging equation given in the rotating reference frame with frequency ω_0:

$$S = \int \rho(x) e^{i\phi_G(x,t)} dx = \int \rho(x) e^{-i\int_0^t \omega_G(x,t^{'})dt^{'}} dx = \int \rho(x) e^{-i2\pi\bar{\gamma} x \int_0^t G_x(t^{'})dt^{'}} dx \tag{1.22}$$

The spatial frequency k can be introduced by defining

$$k(t) = \bar{\gamma} \int_0^t G(t^{'})dt^{'} \tag{1.23}$$

Inserting this relation into Equation 1.22 results in the compact representation

$$S(k) = \int \rho(x) e^{-i2\pi k_x x} dx \tag{1.24}$$

This equation reveals that the image space x and the space of spatial frequencies k are connected via the Fourier transform. Therefore, the effective spin density ρ can be obtained by the inverse Fourier transform of the k-space signal $S(k)$:

$$\rho(x) = \int S(k) e^{+i2\pi k_x x} dk \tag{1.25}$$

The concept of spatial encoding using linear gradient fields can naturally be extended to all three spatial dimensions:

$$\begin{aligned} S(k_x,k_y,k_z) &= \int\int\int \rho(x,y,z) e^{-i2\pi\bar{\gamma}\int_0^t \left(xG_x(t^{'})+yG_y(t^{'})+zG_z(t^{'})\right)dt^{'}} dxdydz \\ &= \int\int\int \rho(x,y,z) e^{-i2\pi(k_x x + k_y y + k_z z)} dxdydz \end{aligned} \tag{1.26}$$

Most MR imaging experiments are 2D experiments, i.e. no phase-encoding in the z-direction is applied ($G_z = 0$ T m^{-1}). In this case, Equation 1.26 shows that the 2D imaging case is a projection of the excited 3D volume along the z-direction. First consider the resulting 2D k-space:

$$S(k_x,k_y) = \int\int\int \rho(x,y,z) e^{-i2\pi(k_x x + k_y y)} dxdydz \tag{1.27}$$

After inverse Fourier transform and using Equation 1.27, the effective 2D spin density results in:

$$\rho(x,y) = \int\int S(k_x,k_y) e^{+i2\pi(k_x x + k_y y)} dk_x dk_y = \int \rho(x,y,z) dz \tag{1.28}$$

The Fourier transform is a continuous function. However, MR data are discretized, motivating the use of the discrete Fourier transform. Discretization of the continuous 1D Fourier transform changes Equations 1.24 and 1.25 into:

$$s(p\Delta k) = \Delta x \sum_{q=-n}^{n-1} \hat{\rho}(q\Delta x) e^{-\frac{i\pi pq}{n}} \tag{1.29}$$

and

$$\hat{\rho}(q\Delta x) = \Delta k \sum_{p=-n}^{n-1} s(p\Delta k) e^{\frac{i\pi pq}{n}} \tag{1.30}$$

with $n \in \mathbb{N}$ and $p, q \in \mathbb{Z}$. In both equations, $\Delta x = \frac{L}{2n}$ is the pixel size if $2n$ pixels are desired. Furthermore, $\Delta k = \frac{1}{L}$ is the step size in k-space determined by the desired Field-Of-View (FOV) L. Because of the discretization, only an approximate effective spin density $\hat{\rho}\,(q\Delta x)$ can be reconstructed.

The claim for $\Delta k = \frac{1}{L}$ is known as the Nyquist-Shannon theorem. This theorem states that a signal with a maximum frequency f_{max} and a minimum frequency f_{min} must be acquired with a sampling frequency $f_{sample} \geq 2(f_{max} - f_{min})$ for unambiguous detection. Otherwise, since more than one frequency is capable of explaining the measured data (see Figure 2), so-called aliasing effects occur.

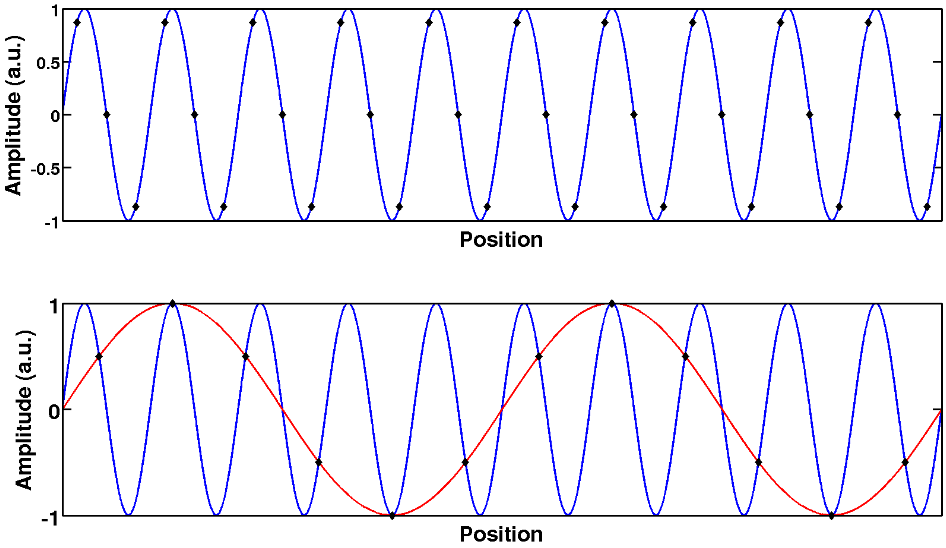

Figure 2: *The upper diagram shows a sinusoid sampled according to the Nyquist-Shannon theorem. Thus, only the displayed sinus fits the data points (black diamonds). In the lower diagram, the sampling theorem is violated, leading to undersampled data. Therefore, both the blue sinus and the red sinus with lower frequency fits the data. The red curve is the difference frequency between the sampling frequency and the desired blue frequency.*

After laying the foundational basics, the process of exciting and encoding a typical MRI experiment can be explained. First, a slice position within the 3D sample volume must be chosen. Therefore, a slice selection gradient is switched on and a radiofrequency pulse with the resonance frequency of the desired slice position excites the spins (Figure 3). The bandwidth of the rf pulse determines the slice profile. Subsequently, the spatial encoding within the slice is achieved by two perpendicular gradient fields. These gradients are the readout gradient and the phase encoding gradient. In the case of 3D imaging, the whole sample volume is excited and then encoded by all three gradients. In addition to the readout gradient, two phase encoding gradients now exist. The readout gradient implies a frequency encoding along the readout direction, while the two phase encoding directions alter the phase of the spins perpendicular to the frequency encoding axis. Thus, every position in space is eventually encoded. However, the value of the phase that is imposed by the phase encoding (all other phase variations are neglected) varies between 0 and 2π. This could lead to foldover artifacts after Fourier transform due to the ambiguity of the phase if $\Delta k \geq \frac{1}{L}$.

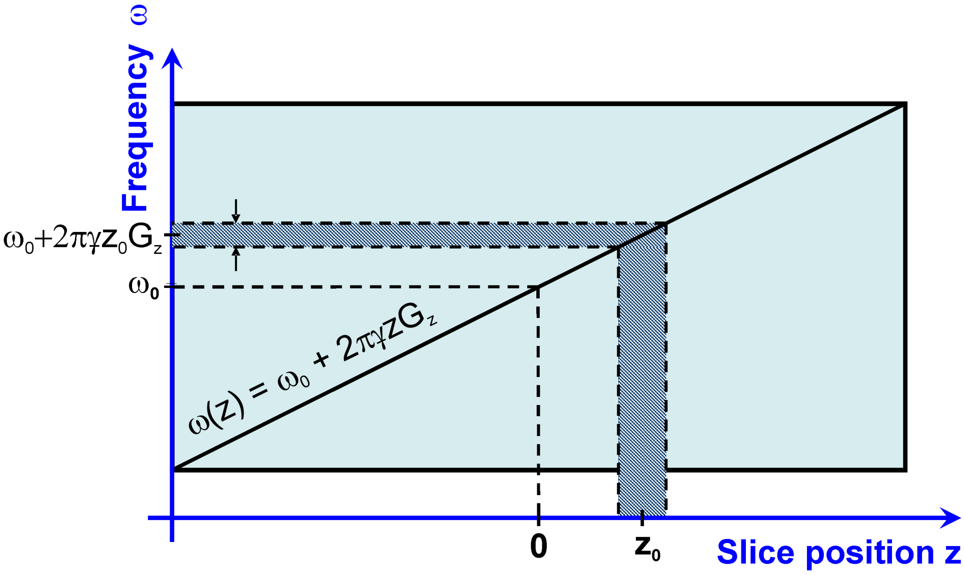

Figure 3: *Principle of slice selection: A slice selection gradient $G_{SS} = G_z$ (arbitrary choice) is used to encode the z-direction. Next, an rf-pulse with the right center frequency $\omega_0 + 2\pi\gamma z_0 G_z$ and the fitting bandwidth $\Delta\omega$ for the desired slice thickness is applied. The spins within this frequency bandwidth are then excited by the rf-pulse.*

2 Compressed Sensing

The theory of Compressed Sensing (CS; also known as Compressive Sensing or Compressed/Compressive Sampling) was introduced in 2004 by Candès and Donoho [22, 23] and later published in 2006 [14, 15, 24]. Its origin lies in the field of information theory, where research efforts were dealing with the problem of how to reconstruct an object from incomplete frequency samples (i.e. Fourier coefficients). If it is *a priori* known that the signal which to be reconstructed is sparse, then Candès and Donoho could prove that, with overwhelming probability, the signal can be exactly reconstructed from randomly measured Fourier coefficients. This methodology, however, is not limited to Fourier imaging (e.g., MRI) as discussed in this chapter.

CS is a sophisticated and advanced mathematical theory. However, for the purposes of this work, it is not necessary to review all mathematical proofs, lemmas and characteristics of CS. Therefore, this section will summarize the most important findings of the CS framework and how to exploit them in MRI. Readers interested in the mathematical publications introducing CS are thus referred to [14, 15] as well as the websites of Emmanuel Candès (http://www-stat.stanford.edu/∼candes/) and David Donoho (http://www-stat.stanford.edu/∼donoho/).

The successful application of CS requires the consideration of three essential aspects:

1. **Sparsity**: A sparse representation of the desired signal is crucial for the success of CS. A signal is sparse if a domain exists where the dataset is characterized by a number of elements significantly lower than the number of all elements in the dataset. The sparser the description of the desired signal in any arbitrary mathematical basis, the fewer data are necessary for an accurate reconstruction.
2. **Sampling Pattern**: In conventionally undersampled data (as given e.g. in Cartesian parallel imaging [7, 8]), coherent foldover artifacts occur. Because CS cannot distinguish between a foldover artifact and the desired unfolded signal, a sampling strategy that creates incoherent artifacts in case of undersampling is necessary.
3. **Reconstruction**: An accurate reconstruction scheme is crucial for "decompressing" the "compressed" signal. For this decompression, the reconstruction algorithm solves an optimization problem that models the expected image.

Each of these issues is explained in greater detail with a specific focus on MRI in the following three sections.

2.1 Sparsity

Sparsity is the most important foundation of CS. The sparsity of the desired signal allows the amount of acquired data to be reduced while preserving the ability to accurately reconstruct the data. The assumption of sparsity is later included in the data reconstruction process as prior knowledge.

Consider a noise-free ideal signal s composed of N elements, where m elements are non-zero.

This signal is referred to as “sparse“ if m $\ll$ N. In mathematical terms, the l_0-norm of the signal is much smaller than the total number of signal elements N:

$$\|s\|_0 = \sum_{i=1}^{N} |s_i|^0 = m \tag{2.1}$$

if $0^0 := 0$ is defined. Thus, the l_0-norm gives 1 where $s_i \neq 0$ and 0 otherwise. The general definition of a norm is:

$$\|u\|_p = \sqrt[p]{\sum_i |u_i|^p} \tag{2.2}$$

where u is an arbitrary vector with elements $u_i \in \mathbb{C}$ and $p \in \mathbb{R}_+$.
Importantly, no strict borderline dividing sparse from non-sparse signals exists. To illustrate the phenomenon discussed above, Figure 4 shows a sparse signal.

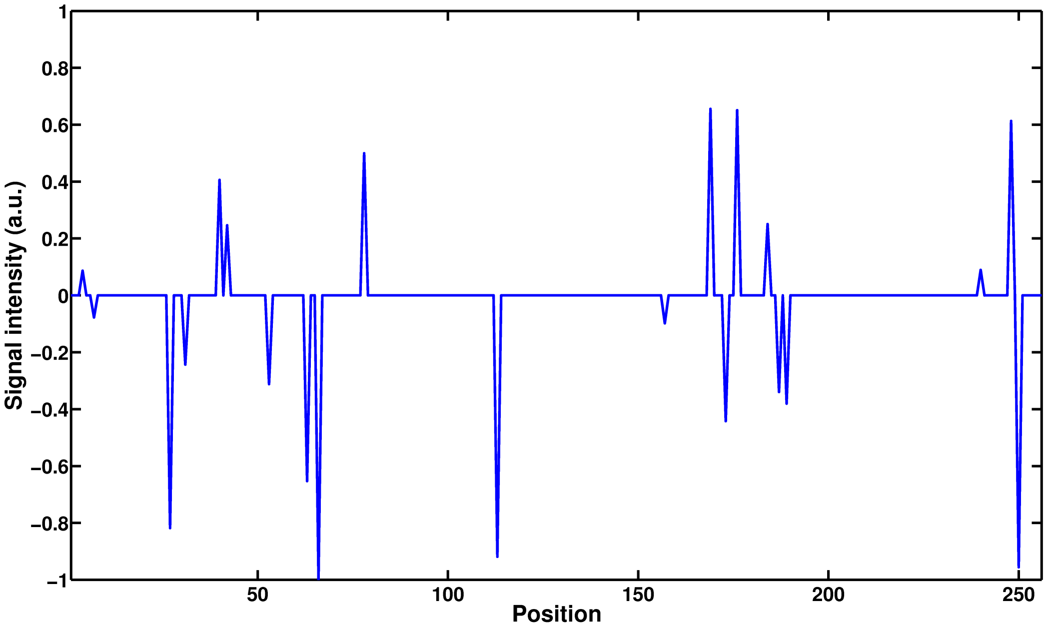

Figure 4: *A sparse 1D signal of length 256. The maximum signal intensity is normalized to 1. Only 21 elements are non-zero* ($\|s_{256}\|_0 = 21$)*.*

In the context of MR, a very simple and popular example for sparse data is contrast-enhanced angiography (Figure 5). If noise is neglected and the background tissue intensity is sufficiently low, everything else can be regarded as zero except for the desired signal (the blood vessels) located in defined regions. In the case of angiography, the sparse domain is the image domain and is, therefore, very intuitive.

As mentioned above, the signal can be sparse in any arbitrary mathematical basis. It is obvious that most MR examinations do not result in datasets that are sparse in the image domain. However, many MR datasets can be sparsified by transformation into the wavelet [16] or the Discrete Gradient domain [16, 26, 27]. The wavelet transformation is commonly used in image compression techniques (e.g. JPEG2000). When the wavelet transform coefficients are sorted by their amplitude and a substantial percentage of the smaller wavelet coefficients are discarded, the image remains very well represented after back-transformation into the image domain. This has been extensively discussed in the Sparse MRI paper and in the dissertation of Michael Lustig [16, 27].

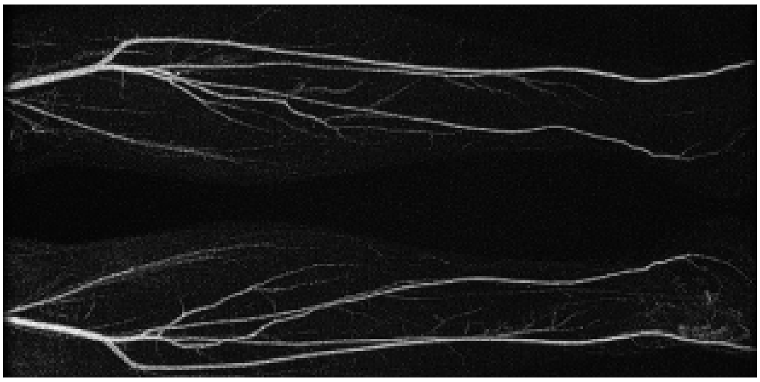

Figure 5: *Typical angiography result of the lower human extremities. This image was used with permission and appeared in [25].*

When sparsifying by Discrete Gradient transformation, the difference between neighboring pixels in all spatial dimensions is calculated. This can also be referred to as forward differencing:

$$\nabla_{DG} x = \begin{pmatrix} x_{i,j,k} - x_{i-1,j,k} \\ x_{i,j,k} - x_{i,j-1,k} \\ x_{i,j,k} - x_{i,j,k-1} \end{pmatrix} \tag{2.3}$$

This leads to a new image representation in which transform coefficients with high intensities occur at sharp transitions between adjacent pixel values. Consequently, adjacent pixels with the same intensity value lead to values of zero after Discrete Gradient transformation. Thus, this can provide effective sparsification of smooth images consisting of sharp transitions (edges) between (spatially extended) piecewise constant image intensities. As illustration, the Shepp-Logan phantom in Figure 6 can be very effectively sparsified using Discrete Gradient transformation.

In addition to image space, wavelet space and Discrete Gradient space, any other sparsifying transform can be utilized. In many cases, finding an appropriate sparsifying transform can be challenging. In this dissertation, the data are either sparse in the image domain or were sparsified by Discrete Gradients.

2.1.1 Dependency on SNR

As pointed out by Candès and Donoho [15, 24], it is important to note that sparse signals only exist in a noise-free setup. In the presence of noise, all elements have non-vanishing intensities. Sparsity, however, can be approximately defined when the background noise is significantly lower than the intensity of signal-containing elements. Thus, it could be stated that a signal is sparse when the number of signal elements (i.e., elements with significantly larger intensity than the background noise) is small compared to the number of all the elements in the dataset. Therefore, sparsity always depends on the SNR. The higher the SNR, the better the noise-free case can be approximated (Figure 7).

Figure 6: *Left: The standard noise-free Shepp-Logan phantom that is not sparse in the image domain; right: The Shepp-Logan phantom after Discrete Gradient transform. Only the edges of the object are visible, leading to a very sparse description of the phantom.*

2.2 Incoherence and sampling strategy

As pointed out by Candès [14, 24] and Donoho [28], the incoherence of the basis in which data are acquired and the sparse domain (where data are subject to reconstruction) are particularly important for CS. In the context of MR, the data acquisition domain is the k-space. Incoherence thus means that a sparse signal (e.g., the image) does not have a sparse representation in the other domain (e.g., k-space). For MRI, this claim is perfectly fulfilled since one pixel in image space is encoded by all k-space values and vice versa. Incoherence is important for two reasons:

- If data are acquired in k-space, then even an undersampled k-space will bear information about all non-zero pixels of the sparse signal.
- With incoherent undersampling, the artifacts resulting from the incomplete information will also be incoherent and appear noise-like. For example, incoherent undersampling can be accomplished by randomly acquiring Fourier coefficients. However, coherent undersampling leads to coherent foldover artifacts that can be observed if, for example, every second phase encoding line in k-space is discarded. These artifacts are well-known from parallel imaging [7, 8]. More generally formulated, incoherent undersampling artifacts can be created by any sampling pattern in k-space that leads to a point spread function with a sharp main peak and low sidelobes. Thus, besides pseudo-random sampling, Non-Cartesian trajectories such as radial and rosette are well suited in the CS context. In Cartesian sampling, improved pseudo-random sampling patterns have been reported [18, 19, 29]. In this dissertation, pseudo-random sampling schemes and radial trajectories are employed.

In conclusion, using a sampling pattern that introduces incoherent noise-like undersampling artifacts in MRI allows reconstruction of sparse MR signals. This work focuses on signals that are sparse either in the image domain or in the Discrete Gradient domain. Because MR data are acquired in k-space and transformed into the image domain, the claim for

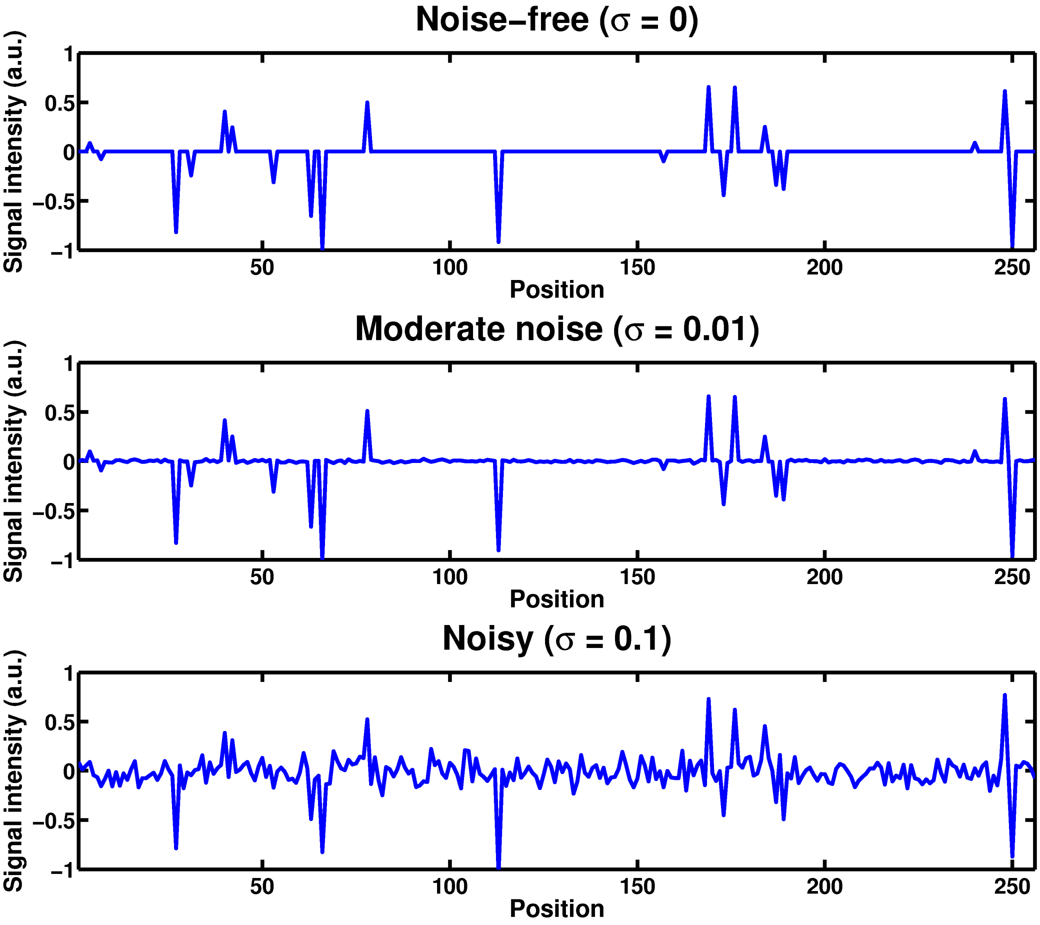

Figure 7: *Illustration of the dependency of sparsity on SNR. Sparse signals only exist in the noise-free case as shown in the upper plot. The same signal (maximum intensity normalized to 1) is shown as displayed in Figure 4. As soon as noise is introduced (middle), all signal elements have signal intensities > 0. However, the original sparse signal can still be observed and describes as "approximately sparse". If the signal is severely corrupted by noise (bottom), the underlying sparse signal is no longer perceivable. The displayed noisy signal is not even approximately sparse.*

incoherence between the data acquisition domain and the sparse domain is given. This is also true for the Discrete Gradient domain because ∇_{DG} is a transform that stays close to the image domain.

2.3 Reconstruction

The CS reconstruction process is an iterative algorithm which solves an optimization problem. In mathematical notation, the problem can be expressed as a so-called cost-function

$$\min \|\Psi x\|_p \qquad \text{such that} \quad \|F_u x - y\|_2 \leq \mu \tag{2.4}$$

where x is the image, y are the undersampled k-space data, Ψ is a linear sparsifying transform, and μ is chosen such that the desired error becomes smaller than the noise level. F_u is the undersampled Fourier transform that gives Fourier coefficients only at positions where actual data are acquired. F_u can be implemented as follows:

1. FFT of the image $x \rightarrow$ k-space k
2. Create logical matrix of the sampling pattern SP: $\mathrm{SP}_i = 1$ where data were collected and $\mathrm{SP}_i = 0$ otherwise
3. Pointwise multiplication of k with SP

This can be summarized by

$$F_u = SP \odot k \tag{2.5}$$

where $\odot$ denotes point-wise multiplication of the matrices or vectors.
In this work, Ψ is either the identity matrix $\mathbb{1}$ or ∇_{DG}. By using $\mathbb{1}$, the sparse domain is the image domain itself. The Discrete Gradient transform is well suited to sparsify images with sharp pixel intensity transitions and piecewise constant pixel intensity areas (e.g., the Shepp-Logan phantom).
Equation 2.4 means that the undersampled image x can be reconstructed by minimizing the l_p-norm of the transform coefficients Ψx. The *a priori* sparsity assumption is explicitly incorporated in the term $\|\Psi x\|_p$. The second part, $\|y - F_u x\|_2 \leq \mu$, enforces data consistency. This constraint is important because the minimum of every minimization will eventually be zero. Furthermore, each iteration step alters the k-space data, which results in successively increasing reconstruction deviations from the measured data. Both effects are effectively addressed by the data consistency constraint.
Ideally, the minimization of the l_0-norm would be desirable:

$$\min \|\Psi x\|_0 \qquad \text{such that} \quad \|F_u x - y\|_2 \leq \mu \tag{2.6}$$

However, the minimization of the l_0-norm is non-deterministic polynomial-time hard, or short NP hard [30, 31]. It is non-deterministic since no information about the position or amplitude of the non-zero pixels contributes to the l_0-norm; this means that all possible combinations of pixel positions and amplitudes must be tested, leading to a combinatorial problem of immense complexity. “Polynomial time“ means that the problem can be solved in a time proportional to the power-scaled size of the problem. For a signal of length N, the polynomial time T in dependency of the problem size N would be $T(N) = O(N^c)$ for a constant $c \in \mathbb{R}_+$ and $c \neq \infty$. In principle, the problem is solvable, although the time required to find a solution would be very long. Thus, instead of minimizing the very complex l_0-norm cost function, the easier problem of l_p-minimization is addressed.
Minimizing the l_p-norm is simpler since, unlike the l_0-case, the information about the amplitudes is preserved. By choosing a starting point close to the desired solution, the convergence toward that desired solution should be guaranteed. Usually, the so-called initial guess for the reconstruction is the inverse Fourier transform of the undersampled k-space data. It is optimal in the least squares sense: $\|F_u x - y\|_2^2 = 0$. For all CS reconstructions presented in this work, choosing the optimal l_2-solution of the undersampled data as an initial guess resulted in meaningful reconstructions.

The choice of p leads to the distinction between “convex“ ($p \geq 1$) and “nonconvex“ ($p < 1$) CS. Usually, the convex l_1-norm is addressed to reconstruct undersampled sparse data [14, 15]. Donoho proved that in most cases, the solution with the minimum l_1-norm is also the sparsest solution [31]. For this reason, $p > 1$ was not investigated in this thesis. Thus, the optimal solution for an ideal sparse signal can be obtained by solving a much simpler problem than minimizing l_0. Minimizing the l_1-norm of the Discrete Gradient transform coefficients is also called “Total Variation Minimization“, since $\|\nabla_{DG} x\|_1$ is known as the Total Variation of the image [26]. Recently, Chartrand [32] demonstrated that using a nonconvex norm $p < 1$ leads to results equivalent to l_1 minimization. Moreover, it allows the exact reconstruction of a sparse signal (for ideal noise-free data) from even less sampled data than in the convex case $p = 1$. Please note that for $p < 1$, the resulting function does not define a norm in the strict mathematical sense. Nonetheless, following up authors like David Donoho and Rick Chartrand [15, 32], throughout this thesis norms with $p < 1$ will still be referred to as a norm.

2.3.1 Formulation with relaxed data consistency: Relaxed DC method

Consider a dataset exhibiting sparsity in an arbitrary domain. Equation 2.4 can be rewritten to explicitly incorporate the data consistency constraint. This is the so-called Lagrangian form:

$$\min(\underbrace{\|F_u x - y\|_2^2}_{\substack{\text{Data}\\ \text{consistency}}} + \lambda \cdot \underbrace{\|\Psi x\|_p}_{\text{Sparsity}}) \tag{2.7}$$

The first term is the data consistency term, while the second term models the sparsity of the expected image. A new parameter, λ, occurs in this equation and weights data consistency against sparsity. Such parameters are called “regularization parameters“ and are critical for the success of the reconstruction [33]. The accurate determination of regularization parameters for CS optimization functions is often empirically achieved by “trial and error“. For example, Equation 2.7 is first solved for various values of λ before the “best“ λ is chosen. The term “best“ hereby refers to that regularization parameter for which

$$\|F_u x - y\|_2^2 \leq \mu \tag{2.8}$$

In the presence of noise, μ is related to the noise level. If no fully sampled reference is available for comparison, it might be problematic to determine the optimal λ. However, the typical purpose of CS is to accelerate the data acquisition process in time-critical MR experiments. Therefore, no fully sampled reference will be available. Consequently, the subjective view of the user has critical influence on the reconstruction result since no objective measure is known that can reliably distinguish a “good“ from a “bad“ reconstruction result. This is certainly a drawback of the formulation in Equation 2.7 since the outcome of the data reconstruction must not depend on the user. The use of the relaxed data consistency constraint in Equation 2.7 motivated the term **Relaxed Data Consistency** (Relaxed DC) method, which will be used throughout this thesis.

2.3.2 Formulation with strict data consistency: Strict DC method

By looking at 2.7, the question might be asked: Is it possible to reconstruct the data without regularization parameters? In order to achieve this, one of the two terms in 2.7 must be forced to be zero. It is clear that forcing the sparsity constraint $\|\Psi x\|_p \overset{!}{=} 0$ is unreasonable since this term models the expected image. However, the data consistency term can be removed by simply reinserting the actually sampled k-space data at the respective positions in the reconstructed k-space. This leads to $\|F_u x - y\|_2^2 = 0$. Thus, only the sparsity term remains, which allows the arbitrary but reasonable choice $\lambda = 1$. Therefore, the final optimization function becomes

$$\min\|\Psi x\|_p \qquad \text{such that} \quad y = F_u x \tag{2.9}$$

The difference between Equations 2.4 and 2.9 is that strict data consistency, $\|F_u x - y\|_2^2 = 0$, is enforced. This reduces the computational load compared to Equation 2.7 since the testing

of a variety of regularization parameters can be omitted. Furthermore, no user interaction is necessary. Because of the strict data consistency, this method will be referred to throughout this thesis as **Strict Data Consistency** (Strict DC) method.
In the ideal noise-free case, strict data consistency is well-suited since CS enables an exact reconstruction of the undersampled data. However, as soon as noise is introduced, noisy imperfect data are regarded as “perfect“ data since they are reinserted after each iteration step. It is obvious that this will lead to artifacts in the final reconstructed image. The simulations performed in this thesis were, besides the determination of interdependencies between several investigated parameters, designed to investigate if these artifact are tolerable in dependence on the noise level.

2.3.3 Optimization algorithms

In this dissertation, two different iterative optimization algorithms were implemented and used to reconstruct retrospectively undersampled datasets. The Nonlinear Conjugate Gradient algorithm was used to reconstruct only retrospectively undersampled data since a regularization parameter is present in the CS cost function. Retrospectively undersampled data were also reconstructed using the Gradient Descent algorithm, which minimizes Equation 2.9. Furthermore, prospectively undersampled datasets were exclusively reconstructed using the Gradient Descent method.

Nonlinear Conjugate Gradient

One way to solve Equation 2.7 is to use a Nonlinear Conjugate Gradient algorithm. Conjugate Gradient algorithms are converged to the global minimum in n iterations if the problem to solve has n dimensions. More details and a proof for the convergence of the Conjugate Gradient method can be found in [34]. This Nonlinear Conjugate Gradient method was used to solve Equation 2.7 in the first CS publication with emphasis on MRI [16]. In this paper, the algorithm is explained in detail for the convex case $p = 1$ and has been implemented for this thesis accordingly which involves the calculation of the derivative of Equation 2.7. Let Equation 2.7 be abbreviated as $f(x)$ and the derivative operation with respect to x^* (* denoting the complex conjugate) as ∇_{x^*} (Wirtinger Calculus, see Appendix A), then, using Equation 2.2, the k^{th} component of the derivative of $f(x)$ becomes (obeying the Einstein Summation Convention):

$$\begin{aligned}
\left[\nabla_{x^*} f(x)\right]_k &= \left[\nabla_{x^*}\left(\|F_u x - y\|_2^2 + \lambda \cdot \|\Psi x\|_1\right)\right]_k \\
&= \left[\nabla_{x^*}\left(\textstyle\sum_i |\left(F_u x - y\right)_i|^2 + \lambda \cdot \sum_i |\left(\Psi x\right)_i|\right)\right]_k \\
&= \frac{\partial}{\partial x_k^*}\Bigg(\textstyle\sum_i \left(x_j^* F^*_{u\,j,i} - y_i^*\right)\left(F_{u\,i,m} x_m - y_i\right) + \\
&\qquad + \lambda \cdot \textstyle\sum_i \sqrt{\left(x_j^* \Psi^*_{j,i}\right)\left(\Psi_{i,m} x_m\right)}\Bigg)
\end{aligned} \tag{2.10}$$

Please note that the derivative is a vector with the same number of components as the image vector x. The calculation of the l_2-norm of the data consistency term and the l_1-norm of the sparsity transform coefficients involves the absolute value function. This function is not smooth around zero and, thus, cannot be derivated at this point. An approximation of the absolute value function solves this problem (Figure 8):

$$|x| = \sqrt{x^* x} \approx \sqrt{x^* x + \epsilon^2} \tag{2.11}$$

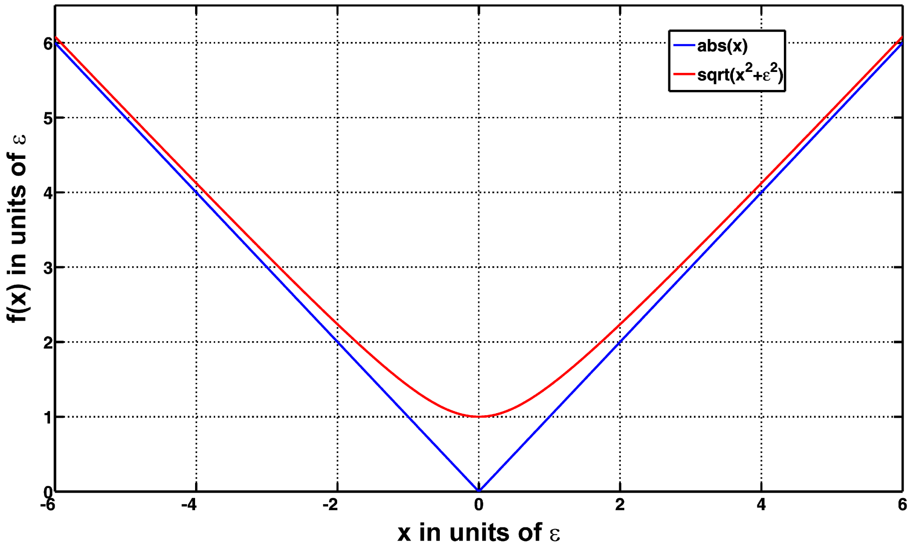

Figure 8: *Comparison of* $|x|$ *and the approximation* $|x| \approx \sqrt{x^*x+\epsilon^2}$*. The approximation smooths the discontinuity around zero with the minimum value being* ϵ*. Thus, the approximation can be derivated around zero. Please note that the axes are in units of* ϵ*. The approximation asymptotically approaches the real absolute value function with increasing distance from the origin.*

with ϵ being a smoothing factor. In the implementation of the Nonlinear Conjugate Gradient used in this thesis, ϵ was fixed to 10^{-4}.
The derivative using Equation 2.11 becomes (for details see Appendix A.1):

$$\begin{aligned}
[\nabla_{x^*} f(x)]_k &\approx \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \left\{ \left(x_j^* F_{u\,j,i}^* - y_i^* \right) \left(F_{u\,i,m} x_m - y_i \right) + \epsilon^2 \right\} + \\
&\quad + \lambda \cdot \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \sqrt{ \left(x_j^* \Psi_{j,i}^* \right) \left(\Psi_{i,m} x_m \right) + \epsilon^2 } \\
&= \left[F_u^* \left(F_u x - y \right) \right]_k + \tfrac{\lambda}{2} \cdot \left[\Psi^* \left(\frac{\Psi x}{W_\Psi} \right) \right]_k
\end{aligned} \tag{2.12}$$

Ψ^* is hereby the inverse transform of Ψ and W_Ψ is a vector with the components

$$W_{\Psi i} = \sqrt{ (\Psi x)_i^* \, (\Psi x)_i + \epsilon^2 } \approx |\, (\Psi x)_i \,| \tag{2.13}$$

The factor $\frac{1}{2}$ can be absorbed in the regularization parameter $\lambda : \frac{\lambda}{2} \mapsto \lambda$. All factors that have been neglected are scaling factors or powers. They can be absorbed in the step size (see Appendix of [16]), which is determined by backtracking line search [35] analogous to [16].
For $\Psi = \mathbb{1}$, Equation 2.12 results in

$$[\nabla_{x^*} f(x)]_k = \left[F_u^* \left(F_u x - y \right) \right]_k + \lambda \cdot \left[\frac{x}{W_{\mathbb{1}}} \right]_k \tag{2.14}$$

On the other hand, if $\Psi = \nabla_{DG}$, the derivative becomes

$$[\nabla_{x^*} f(x)]_k = \left[F_u^* \left(F_u x - y \right) \right]_k + \lambda \cdot \left[\nabla_{DG}^* \left(\frac{\nabla_{DG} x}{W_{\nabla_{DG}}} \right) \right]_k \tag{2.15}$$

where ∇_{DG}^* is defined as

$$\nabla_{DG}^* x = \begin{pmatrix} x_{i-1,j,k} - x_{i,j,k} \\ x_{i,j-1,k} - x_{i,j,k} \\ x_{i,j,k-1} - x_{i,j,k} \end{pmatrix} \tag{2.16}$$

and is the backward differencing operation (see the definition of the forward differencing operator in Equation 2.3).

It turned out empirically that decreasing the regularization parameter λ every 30^{th} iteration by a factor of 0.5 accelerated the convergence and enabled more successful reconstructions than using a static λ.

In conclusion, a Nonlinear Conjugate Gradient algorithm for the convex formulation ($p = 1$) of Equation 2.7 was implemented according to [16]. In contrast to the description of the algorithm in [16], the applied implementation decreased the regularization factor λ every 30^{th} iteration by a factor of 0.5.

Gradient Descent or Steepest Descent

The simplest and most intuitive way to minimize a function is to calculate the derivative and step along the direction of steepest descent. In principle, this method is also applicable to Equation 2.7; however, for regularized CS, only the Nonlinear Conjugate Gradient algorithm was used. The convergence speed of the Steepest Descent method depends on the conditioning of the linear system and is explained in greater detail in [34]. Consider Equation 2.9 for an arbitrary p and calculate the derivative using Equation 2.2 (obeying the Einstein Summation Convention and using Wirtinger Calculus, see Appendix A):

$$\begin{aligned} \left[\nabla_{x^*} f(x)\right]_k &= \left[\nabla_{x^*}\left(\|\Psi x\|_p\right)\right]_k \\ &= \frac{\partial}{\partial x_k^*}\left(\sum_i \left\{\sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,j} x_j\right)}\right\}^p\right)^{\frac{1}{p}} \end{aligned}$$

The situation here is similar to that in Equation 2.10 and can be treated accordingly. Again, the derivative using 2.11, neglecting the scaling power $\frac{1}{p}$, and dividing by p becomes (details in Appendix A.2):

$$\begin{aligned} \left[\nabla_{x^*} f(x)\right]_k &\approx \frac{1}{p} \cdot \frac{\partial}{\partial x_k^*} \sum_i \left\{\sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2}\right\}^p \\ &= \frac{1}{p} \cdot \frac{p}{2} \cdot \sum_i \left\{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2\right\}^{\frac{p}{2}-1} \cdot \left(\delta_{k,j} \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) \\ &= \frac{1}{2} \cdot \left[\Psi^* \left(W_\Psi^{p-2} \Psi x\right)\right]_k \end{aligned} \tag{2.17}$$

W_Ψ is defined as given in Equation 2.13. W_Ψ^{p-2} means an elementwise raising to the power $p-2$.

Similar to the last section, if $\Psi = \mathbb{1}$, this equation becomes

$$\left[\nabla_{x^*} f(x)\right]_k = \left[W_{\mathbb{1}}^{p-2} x\right]_k \tag{2.18}$$

and, respectively, for $\Psi = \nabla_{DG}$

$$\left[\nabla_{x^*} f(x)\right]_k = \left[\nabla_{DG}^* \left(W_{\nabla_{DG}}^{p-2} \nabla_{DG} x\right)\right]_k \tag{2.19}$$

In order to update the image x, a step size t must be determined. This step size t implicitly incorporates the factor 2 by which Equation 2.17 is divided. After the determination of the step size, the derivative of the cost function must be multiplied by that step size and subtracted from the current x. The step size can be determined by any suitable line search method, such as backtracking line search or exact line search. Both methods are described and explained in [35]. The step size can also be manually adjusted. Repeatedly determining the derivative, a step size, and the update of the current x, leads to an iterative algorithm outlined in [32]. The exact implementation used in this thesis can be formulated in pseudo code:

```
Input
    x: Image
    y: Undersampled k-space data
    $\Psi$: Sparsifying transform
    FFT: Fast Fourier Transform (image space → k-space)
    IFFT: Inverse FFT (k-space → image space)
    Approximate norms using Equation 2.11

Initialize
    $x_0$ = IFFT(y), normalize $x_0$ to 1, n = 1, $\epsilon = 1$

Iterations
    while $\epsilon > 10^{-4}$
        $d = \Psi^* \left( W_\Psi^{p-2} \Psi x_{n-1} \right)$        , $W_\Psi$ calculated from $x_{n-1}$
        Determine step size $t$ such that $\| \Psi \left( x_{n-1} - t \cdot d \right) \|_p = min$
            (e.g. in MATLAB® using fminbnd, an exact line search)
        $x_n = x_{n-1} - t \cdot d$
        $y_n$ = FFT($x_n$)
        Reinsert acquired k-space points at respective positions
            in $y_n \rightarrow y_{n,SDC}$ (Strict Data Consistency)
        $x_n$ = IFFT($y_{n,SDC}$)
        if n is multiple of 30
            $\epsilon = \epsilon \cdot 0.5$
        end
        n = n + 1
    end
```

Please note that in contrast to the Nonlinear Conjugate Gradient algorithm, ϵ is not a fixed value but is repeatedly decreased. The stopping criterion “$\epsilon > 10^{-4}$“, is derived in Section 3.2 and turned out to be a robust and reliable choice. In the presented implementation, this leads to 420 iterations for each reconstruction.

Relation to Iterative Soft Thresholding

The Gradient or Steepest Descent method used to solve Equation 2.9 in the Strict DC method can be viewed from a different, more intuitive perspective. This perspective involves thresholding algorithms described in the following paragraph.

In a recent publication from 2007 [36], it was demonstrated that minimizing the l_1-norm in an iterative reconstruction process is equivalent to iterative thresholding. Iterative thresholding techniques are very simple and intuitive algorithms widely used in signal processing (see Figure 9). Two different approaches exist: In *hard thresholding*, data smaller than a predefined threshold value τ are set to zero. Thus, all data remain unchanged except data smaller than τ. This can be abbreviated as:

Hard thresholding

$$|f_i| \leq \tau \Rightarrow f_i = 0 \quad \text{for } f_i \in \mathbb{C} \tag{2.20}$$

In *soft thresholding*, the threshold value τ is subtracted from all data points and data smaller than τ before subtraction are set to zero. Thus, all data points are affected by the thresholding operation:

Soft thresholding

$$\begin{aligned} |f_i| > \tau \quad &\Rightarrow \quad f_i = f_i - \tau \\ |f_i| \leq \tau \quad &\Rightarrow \quad f_i = 0 \qquad\qquad \text{for } f_i \in \mathbb{C} \end{aligned}$$

A closer look at Equations 2.17 to 2.19 for $p = 1$ reveals that the derivative is only a normalization to 1 of the transform sparsity coefficients Ψx. Since all transform sparsity coefficients are normalized to 1, the multiplication with a step size leads to a global threshold for all pixels. Therefore, the following subtraction from a signal is equivalent to soft thresholding. This equivalence holds for the convex norm $p = 1$.

However, in this thesis, a great emphasis lies on nonconvex CS where $p < 1$. For nonconvex norms, Equations 2.17 to 2.19 are no longer a simple normalization, but a normalization to a power-scaled version of the sparsity coefficients. The thresholding value is hereby the step size t multiplied with the value of the derivative at the respective pixel. Please note that this leads to different thresholds for each individual pixel instead of a global threshold (as given in soft or hard thresholding). However, there is still a close relationship between soft thresholding and l_p-minimization.

The observed equivalence between thresholding and Steepest Descent is not relevant for this work. Rather, it is a different, more intuitive view on what CS minimization actually does.

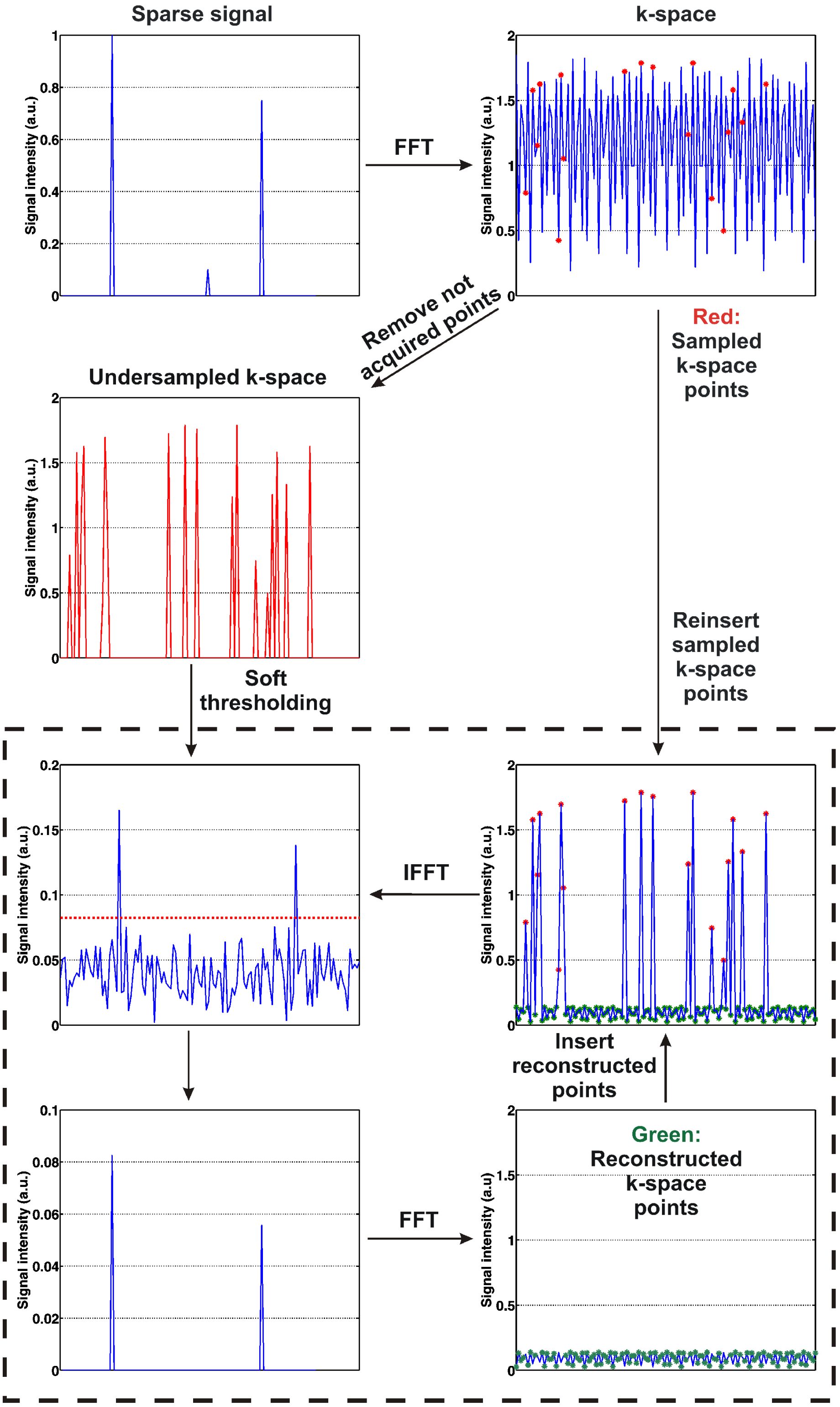

Figure 9: *Schematic diagram of the iterative soft thresholding method. The undersampled, zerofilled k-space data act as starting point. After inverse Fourier transform, a threshold is selected. All values below this threshold are set to zero, from all remaining points the threshold is subtracted. The result is Fourier transformed, leading to a complete k-space. The acquired (red asterisks) and reconstructed (green asterisks) data are combined to a new complete k-space. After that, the iteration loop restarts. Since there is a close relation between soft thresholding and l_p-minimization, this chart is an intuitive approach to explaining how CS works*

3 Simulations

In Chapter 2, two different CS reconstruction algorithms were introduced: The Nonlinear Conjugate Gradient method used in [16] and the Gradient or Steepest Descent approach used in [32]. The first method solves the cost function with a relaxed data consistency constraint (Relaxed DC method, Equation 2.7). The latter approach minimizes the cost function using strict data consistency (Strict DC method, Equation 2.9).

In the first section of this chapter, the different characteristics of these two approaches regarding data consistency are investigated and compared. It was not determined which algorithm is faster since the speed of an algorithm depends on multiple variables. These variables include ability of the user to program efficient code, performance of the software that executes the code, the machine speed, etc.. Furthermore, because both algorithms can appropriately optimize, it was deemed unnecessary to determine which algorithm would be better suited for optimizing the respective cost function. As will be demonstrated in this chapter, however, the determination of a fitting regularization parameter was found to be cumbersome. Thus, simulation results that altered exemplary parameters influencing the regularization parameter λ in the Relaxed DC method are presented and discussed. These results demonstrate why the Strict DC method is preferable to the Relaxed DC method.

The second section investigates the parameter ϵ in the Strict DC method that might be critical to the reconstruction process. The parameter ϵ was introduced in Equation 2.11 to approximate the norm function. This allows the derivation of this equation. The section demonstrates that ϵ can be chosen so that the Strict DC method works for all investigated applications.

In this chapter, a pseudo-random sampling pattern with a densely sampled inner k-space was used in most simulations. More sophisticated methods to derive random sampling patterns leading to incoherent artifacts exist (e.g. Poisson-disc [18, 19, 27] or variable density [16, 27, 29, 37]). However, as recently shown [38], the differences between these patterns are of minor importance. The purpose of this chapter is to clarify differences between the Relaxed DC method and the Strict DC method and not to derive optimal sampling strategies.

All presented algorithms and simulations were implemented and conducted using MATLAB® (The MathWorks, Inc., Natick/MA, USA).

Finally, it should be noted that it is possible to solve Equation 2.9 using the Nonlinear Conjugate Gradient algorithm. However, this possibility was not investigated in order to remain close and comparable to the original presentations of the algorithms (Relaxed DC method in [16, 27], Strict DC method in [32]).

3.1 Comparison of the Relaxed DC and Strict DC method

3.1.1 Motivation

The first publication describing the potential of CS in the context of MRI [16], used a Nonlinear Conjugate Gradient algorithm to optimize the cost function presented in Equation 2.7:

$$\min\left(\|F_u x - y\|_2^2 + \lambda \cdot \|\Psi x\|_p\right)$$

In accordance with [16], the implemented version of the Relaxed DC method uses $p = 1$ for the sparsity constraint (last term). It is obvious that a relaxed data consistency constraint is reasonable in the presence of noise since the data are only as accurate as the noise level. However, this leads to the introduction of the regularization parameter λ. This parameter must be determined by "trial and error" which is often termed "empirically". This was already mentioned in Section 2.3.1. It is obvious that this can be a time consuming procedure even when assuming that an optimal determination of λ is possible without an available reference.

Later in 2008, Rick Chartrand presented the nonconvex CS method. This discarded the first term in Equation 2.7 by simply considering it as zero [32]:

$$\min\|\Psi x\|_p \qquad \text{such that} \quad y = F_u x$$

Thus, the choice of the regularization parameter λ is arbitrary and simply set to 1. This circumvents the time consuming determination of λ and therefore eliminates a source of user influence on the reconstruction quality. Nevertheless, regarding noisy data as "perfect" data can certainly be a source of reconstruction artifacts.

Taking all these considerations into account, some questions occur, for example

- if it is acceptable to treat imperfect, noisy data as perfect (Strict DC method) or
- if it is preferable to optimize the regularization parameter λ and tolerate the resulting user influence on the reconstruction quality (Relaxed DC method).

Simulations were thus performed to critique the advantages and disadvantages of the Relaxed DC and Strict DC method. Based on the simulations outcomes, a decision on which method is preferable was made.

All simulations were performed on the Shepp-Logan phantom. The maximum signal intensity of this phantom was normalized to 1. For the quantification of the reconstructions, the root mean square error (RMSE) metric was used. The RMSE is defined as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_i (R_i - O_i)^2}{\sum_i O_i^2}} \tag{3.1}$$

where R_i represents the i-th pixel in the reconstructed phantom and O_i represents the i-th pixel in the noise-free reference phantom.

The Shepp-Logan phantom is sparsified by the Discrete Gradient transform as introduced in Equation 2.3. Therefore, the reconstruction optimized Equations 2.7 and 2.9 using $\Psi = \nabla_{DG}$.

In all reconstructions with the Relaxed DC method, the regularization parameter λ has to be chosen. In this work, this parameter was chosen by selecting values from uniformly distributed pseudo-random numbers of a certain interval. The strategy was chosen for two reasons. First, if no information about λ is available (which is normally the case), the parameter must be empirically determined. In this process, an arbitrary value for λ must be selected in the beginning. Thus, this situation is mimicked by choosing random values for the regularization parameter. Second, if uniform spacing on the interval of regularization parameters is chosen, grid artifacts can occur. “Grid artifacts“ in this context are wrong observations/conclusions resulting from a disadvantageously spaced grid. For example, the optimal regularization parameter λ might be 0.023, but the spacing was such that only the values 0.022 and 0.024 were investigated. Choosing random values reduces the probability to miss such trends.

Orientation for RMSE values

All reconstructions in this chapter were quantified using the RMSE metric described in Equation 3.1. To illustrate the meaning of the RMSE values, reconstruction results for different RMSE values are shown in Figure 10. As can be seen, the reconstruction with RMSE $= 0.0533$ can be regarded as successful while all other reconstructions exhibit different degrees of degradation. Thus, reconstruction with a RMSE ≤ 0.05 were assumed to be successful in all following sections of this chapter. However, please note that this somewhat arbitrary value was used in only this chapter to distinguish successful from failed reconstructions.

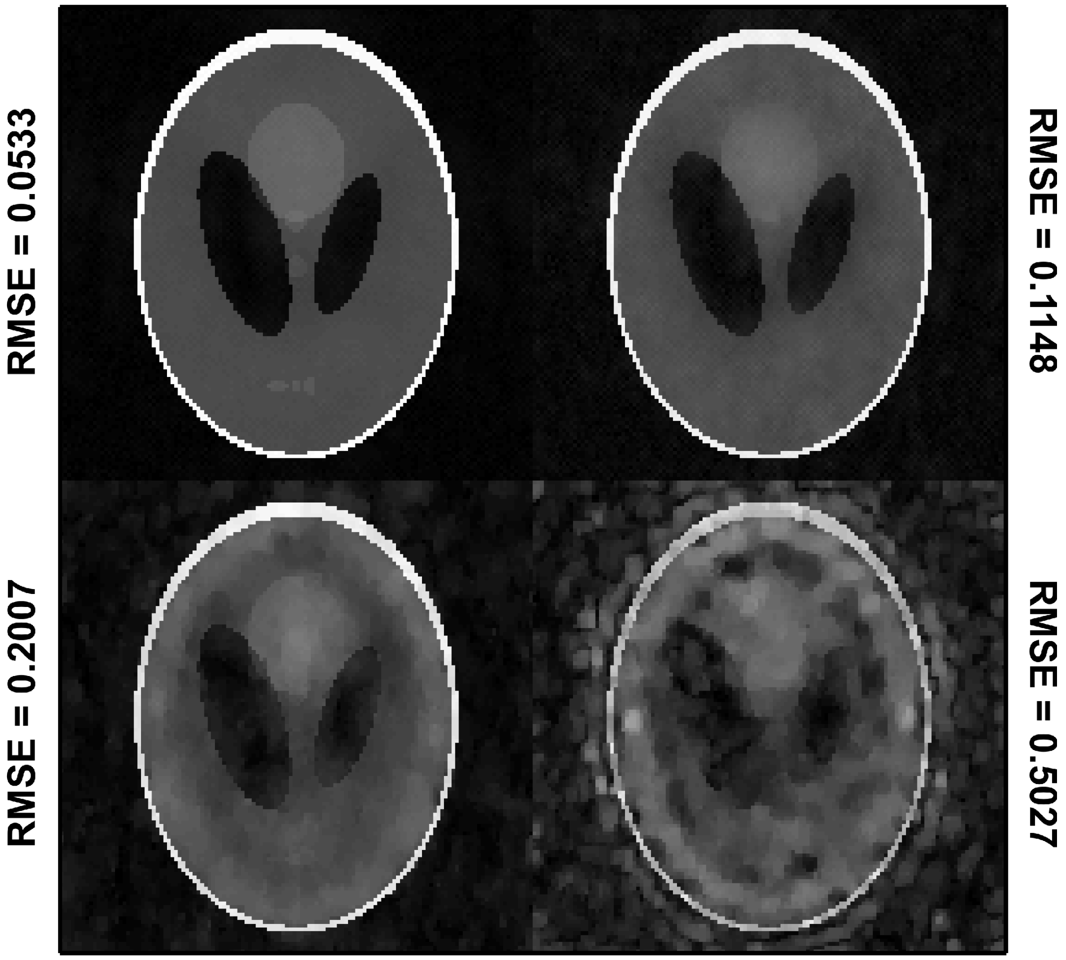

Figure 10: *Exemplary reconstruction results to illustrate the meaning of certain RMSE values.*

3.1.2 Relaxed DC method: Dependency of the regularization parameter λ on the acceleration factor and the sampling pattern

Several parameters could influence the choice of the regularization parameter λ in Equation 2.7. However, in this chapter, only two exemplary important parameters are investigated:

- the acceleration factor
- the sampling pattern

Furthermore, as briefly mentioned in the Discussion section 3.1.2, other parameters such as noise or the object resolution could also influence the regularization parameter λ.
To compare the dependency of the Relaxed DC method on one particular parameter (such as the acceleration factor) to the dependency of the Strict DC method on the same parameter, additional reconstructions with the Strict DC method were performed.

Methods

Dependency on the acceleration factor

To reveal the dependency of the regularization parameter on the applied acceleration factor, simulations testing four acceleration factors ($af = \{2, 4, 6, 8\}$) on the 128x128 noise-free Shepp-Logan phantom were performed. For each reconstruction, a pseudo-random sampling pattern was created with a randomly undersampled outer k-space and a fully sampled k-space center (9x9 k-space points corresponding to approximately 0.5 % of the total k-space). This sampling pattern was then used for all reconstructions with the chosen acceleration factor. An exemplary sampling pattern can be observed in Figure 11. For each acceleration factor, 100 random regularization parameters were drawn from uniformly distributed random numbers between 0 and $\{10, 1, 0.1, 0.01\}$. These parameters were used to reconstruct the undersampled Shepp-Logan phantom using the Relaxed DC method. This mimics the situation where no or only limited *a priori* knowledge about the choice of the regularization parameter λ is available. The results of these 4 (af) $\cdot$ 4 (λ range) $\cdot$ 100 (chosen λs) = 1600 reconstructions were again quantified with the RMSE metric.
For the purpose of comparison, for each acceleration factor, the undersampled data were also reconstructed using the Strict DC method. Five norms ($p = \{0.25, 0.50, 0.75, 0.90, 1.00\}$) were tested to determine an optimal value for the (non-)convex norm p. This resulted in 20 additional reconstructions that were quantified using the RMSE and compared to the outcome of the Relaxed DC method.

Optimal range for the regularization parameter λ

The correct range from which the regularization parameter should be chosen must be determined. “Optimal“ hereby means optimal in terms of the RMSE. Therefore, the mean RMSE from all 100 reconstructions per af and λ-range (see paragraph above) was calculated. The results were evaluated with respect to the optimal λ-range for all investigated acceleration factors.

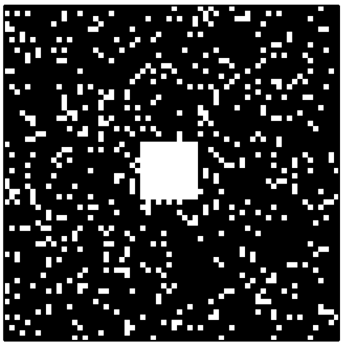

Figure 11: *Exemplary pseudo-random sampling pattern. The outer k-space is randomly undersampled according to the acceleration factor (here:* af = *8) while the inner k-space is densely sampled.*

Dependency on the sampling pattern

As mentioned above, the sampling pattern itself might influence the proper choice of the regularization parameter λ. This was studied on the 128x128 noise-free Shepp-Logan phantom. Thus, 20 random regularization parameters from uniformly distributed random numbers between 0 and 0.1 were drawn. A fixed acceleration factor of 6 was used to investigate the influence of the sampling pattern. For each chosen parameter, 20 pseudo-random sampling patterns (all with the same acceleration factor of 6) were created. For each sampling pattern, reconstructions were performed with the Relaxed DC method and quantified using the RMSE.
Furthermore, to investigate the influence of different sampling patterns with the same acceleration factor on the Strict DC method, 20 sampling patterns were created and used to reconstruct the Shepp-Logan phantom for five p-values (0.25, 0.50, 0.75, 0.90, 1.00). Again, the RMSE of each reconstruction was calculated and compared to the outcome of the Relaxed DC method.
In this thesis, in addition to Cartesian data, radially acquired undersampled MR data were CS reconstructed after gridding to the Cartesian grid using GRAPPA Operator Gridding (GROG, [39, 40]). GROG leads to a pseudo-radial distribution of gridded k-space points in the Cartesian plane. To study whether or not these post-gridding pseudo-radial data influence the choice of the regularization parameter, simulations on the noise-free 128x128 Shepp-Logan phantom were performed. A pseudo-radial sampling pattern was created consisting of 20 pseudo-radial projections with a Cartesian acceleration factor of approximately 6. The pseudo-radial sampling pattern was compared to a pseudo-random Cartesian sampling pattern with densely sampled inner k-space (9x9 k-space points corresponding to 0.5 % of the total k-space). 100 start values for the regularization parameter λ were chosen from uniformly distributed pseudo-random numbers between 0 and 0.1. For each start value, the data were undersampled using both sampling patterns and CS reconstructed with the Relaxed DC method. The results were quantified using the RMSE metric.

Results

Optimal range for the regularization parameter λ

The results displayed in Figure 12 demonstrate that the optimal range from which the

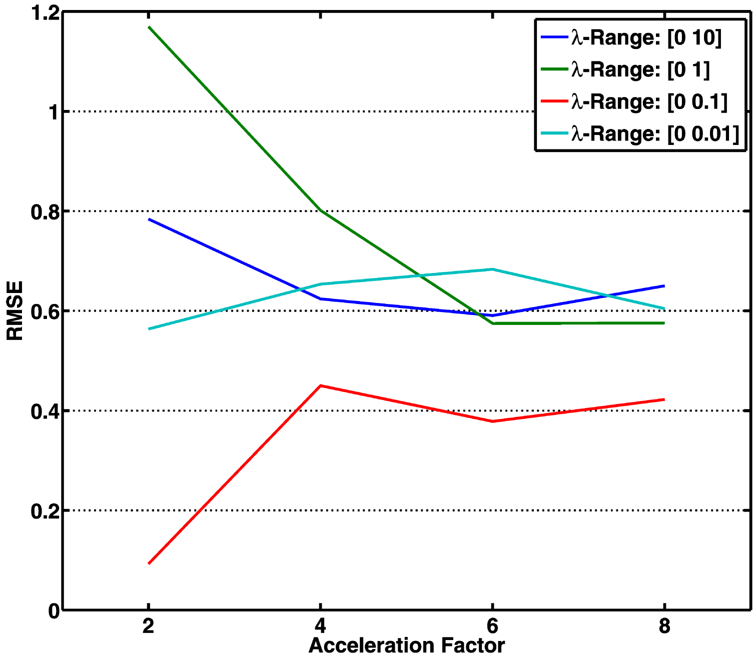

Figure 12: *Mean RMSE values for the reconstruction results of the noise-free Shepp-Logan phantom. These were dependent on the regularization parameter range and the acceleration factor. It can be clearly seen that the average RMSE was minimal for λ's in the range* [0 0.1] *for all investigated acceleration factors.*

start value for λ should be chosen is [0 0.1]. For all investigated acceleration factors, the reconstructions using a start value for λ from the range [0 0.1] offer the lowest mean RMSE values.

Dependency on the acceleration factor

Figure 13 reveals that the optimal choice for the start value of λ is heavily dependent on the acceleration factor. All reconstructions denoted by the blue line were obtained using the same start value for $\lambda = 0.0169$. This value was chosen because it was optimal (in the presented simulations) for $af = 2$. While a reconstruction for $af = 2$ leads to a successful result, the outcome for all other acceleration factors is highly erroneous. The red line shows the RMSE values of the best reconstructions obtained by the simulations. As can be seen, the reconstruction quality in terms of RMSE was significantly better and, especially for high acceleration factors, the optimal values for λ differed tremendously from 0.0169. However, even the optimal parameters obtained from the simulations failed to achieve successful reconstructions (RMSE ≤ 0.05) for $af \geq 6$.

Table 1 shows that the Relaxed DC method resulted in most cases in significantly higher RMSE values than the Strict DC method (Table 2). For $af = 2$, the Relaxed DC method resulted in a very low RMSE for noise-free data. However, the average RMSE of the reconstructions with the Relaxed DC method was significantly higher than the best RMSE value obtained with the Relaxed DC method and the corresponding RMSE values of the Strict DC method. Furthermore, in Figure 14, it can be seen that the optimal norm in terms of RMSE was nonconvex ($p \leq 0.50$). This was true for all investigated acceleration factors. By choosing $p \leq 0.50$, the Strict DC method achieved successful reconstructions (RMSE ≤ 0.05) for all acceleration factors.

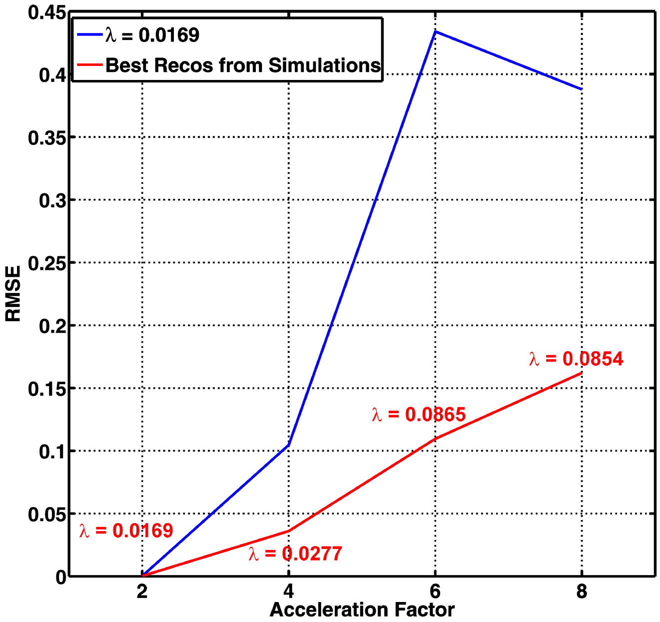

Figure 13: *Reconstructions of the noise-free Shepp-Logan phantom for different acceleration factors with the same start value $\lambda = 0.0169$ (blue line). As seen, the RMSE shows significant dependency on the acceleration factor. The RMSE values of the best reconstructions from the simulations (red line) are significantly lower.*

	min	mean	max
$af = 2$	0.0002	0.0924	1.4931
$af = 4$	0.0360	0.4499	1.7384
$af = 6$	0.1096	0.3782	2.0143
$af = 8$	0.1620	0.4223	0.9831

Table 1: *The results of the reconstructions obtained with the Relaxed DC method. For each acceleration factor, the minimum, maximum, and mean RMSE values are given. As can be seen, the range of values is very large.*

	$\boldsymbol{af = 2}$	$\boldsymbol{af = 4}$	$\boldsymbol{af = 6}$	$\boldsymbol{af = 8}$
Relaxed DC	0.0002	0.0360	0.1096	0.1620
Strict DC	0.0102	0.0136	0.0221	0.0416

Table 2: *Comparison of the RMSE values obtained with the Strict DC method for* p = *0.50 with the best RMSE values obtained from the Relaxed DC method. With the exception of* af = *2, the Strict DC method leads to significantly lower RMSE values than the Relaxed DC method.*

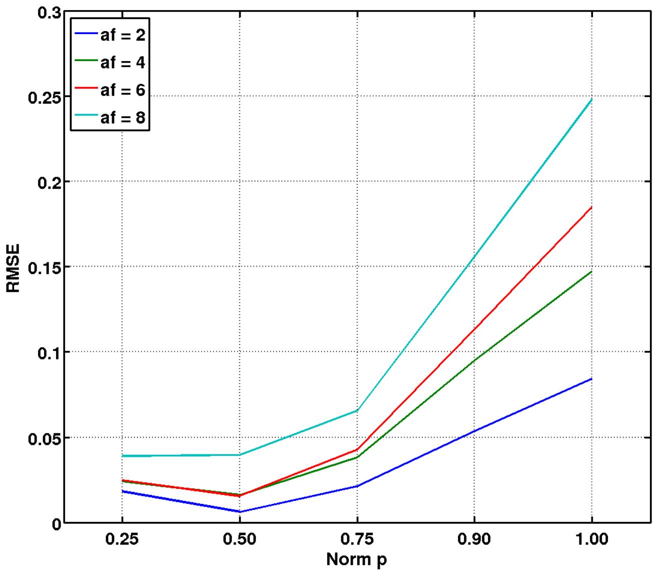

Figure 14: *Results from reconstructions with the Strict DC method. The RMSE for the investigated acceleration factors is plotted against the applied norm* p*. It can be seen that the optimal norm in terms of RMSE is nonconvex (*$\mathrm{p} \leq 0.50$*). For all acceleration factors, choosing* $\mathrm{p} \leq 0.50$ *led to successful reconstructions (RMSE* $\leq$ *0.05).*

Dependency on the sampling pattern

The influence of the sampling pattern on the reconstruction quality was studied. The specific focus was thus to determine if the pattern using the Relaxed DC method has critical influence on the choice of the regularization parameter λ. The results are displayed in Figure 15.

In the upper row, the RMSE values of the reconstructions are sorted and plotted. The dashed red line indicates the RMSE threshold. Below 0.05 the reconstruction was considered successful. As can be seen, all reconstructions with the Relaxed DC method exceeded this threshold. Therefore, no successful reconstruction could be obtained. Most of the reconstructions resulted in a RMSE > 0.25, which indicates erroneous reconstruction outcomes. However, for the Strict DC method, 60 out of the 100 reconstructions resulted in a RMSE of approximately 0.05 or lower. The remaining 40 reconstructions did not exceed a RMSE value of 0.25; hence, the Strict DC method converged in most cases to a significantly better solution than the Relaxed DC method. The “plateaus“ in the RMSE plot for the Strict DC method can be assigned to the different norms: The lowest plateau belongs to $p = 0.25$ and 0.50, the next plateau to $p = 0.75$, etc.. This clarifies the benefit of using nonconvex norms in the Strict DC method.

In the lower row of Figure 15, the RMSE values are plotted in histograms. Again, the dashed red line indicates the RMSE threshold of 0.05. The Relaxed DC method offers a broad spectrum of RMSE values with a maximum of approximately 0.20, while the Strict DC method exhibits RMSE values in a small range (RMSE < 0.20). These histograms demonstrate the preferability (in terms of RMSE) of using the Strict DC method.

In Figure 16, the dependency of λ on the sampling pattern becomes even more obvious. This figure shows the same RMSE values displayed in Figure 15 but in a different presentation.

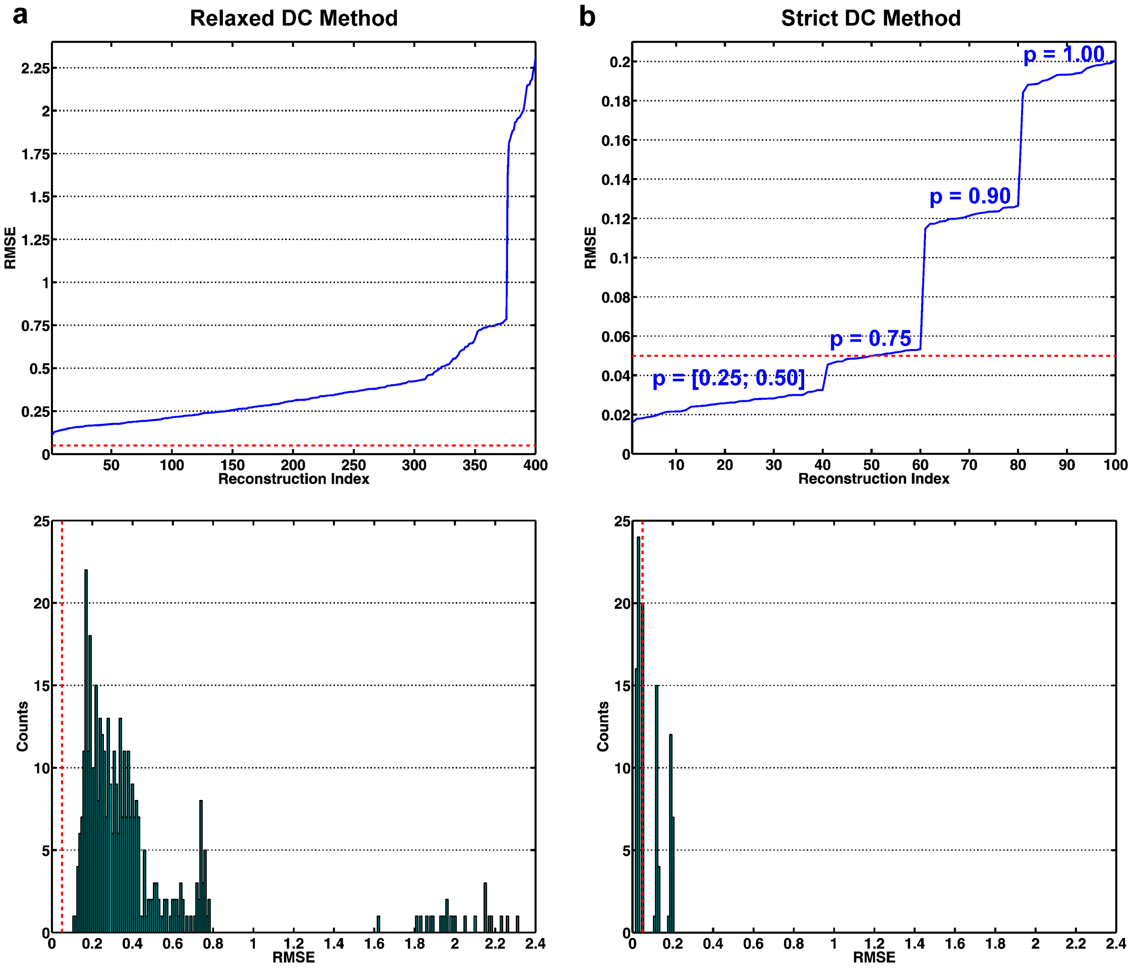

Figure 15: *Investigation outcome of the regularization parameter λ dependency on the sampling pattern. (a) The Relaxed DC method fails to successfully reconstruct the Shepp-Logan phantom in all simulations. A reconstruction was considered successful when RMSE $<$ 0.05. This threshold was indicated by the dashed red lines in the diagrams. The upper diagram shows that the majority of the CS reconstructions with the Relaxed DC method resulted in RMSE values larger than 0.25. This is further clarified in the lower diagram where the RMSE values are displayed in a histogram. A peak occurs around RMSE = 0.20. (b) The Strict DC method resulted in successful CS reconstructions for* $\mathrm{p} \leq 0.75$*. The plateaus in the graph of the upper diagram can be attributed to the investigated values for* p*. The histogram in the lower diagram further clarifies that the Strict DC method leads in general to lower RMSE values than the Relaxed DC method.*

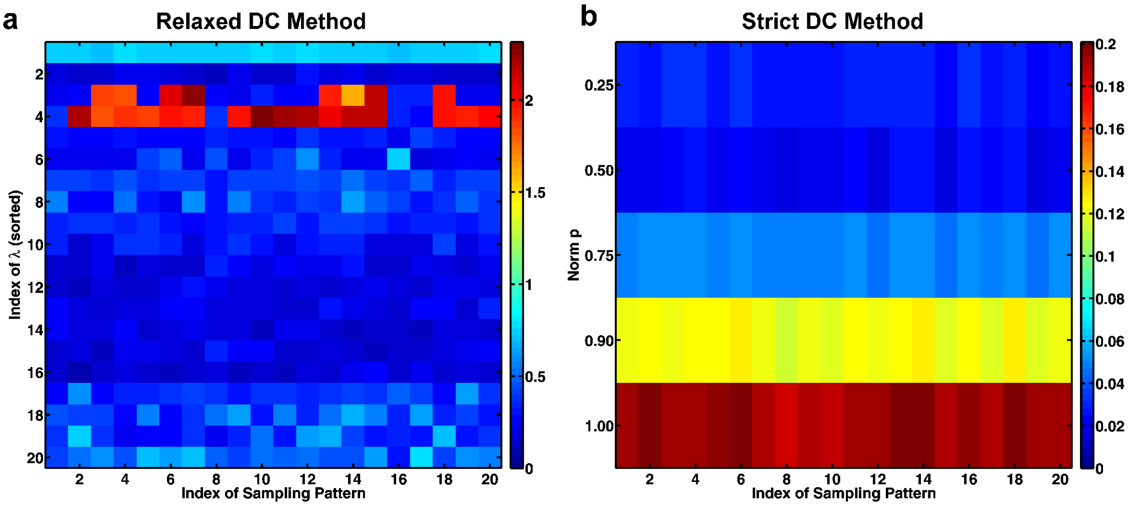

Figure 16: *Investigation results of the regularization parameter λ dependency on the sampling pattern. (a) RMSE values for the Relaxed DC method. As seen, the sampling pattern has significant influence on the reconstruction quality for a given λ. Each index of λ represents a start value drawn from uniformly distributed random numbers between 0 and 0.1. The λs are sorted (ascending order from top to bottom). Each index of the sampling pattern denotes a particular pseudo-random pattern with the same acceleration factor ($\mathrm{af} = 6$) and a densely sampled 9x9 k-space center. (b) The RMSE values of the Strict DC method exhibit only minor variations for a given norm p. The sampling pattern shows a negligible influence on the reconstruction quality. The x-axis is equivalent to that of the Relaxed DC method. Please note the different scaling of the RMSE values in the two images.*

Norm p	**Mean RMSE**	**RMSE Range**
0.25	0.0288	[0.0245 0.0325]
0.50	0.0220	[0.0159 0.0268]
0.75	0.0499	[0.0454 0.0533]
0.90	0.1215	[0.1148 0.1265]
1.00	0.1936	[0.1842 0.2008]

Table 3: *The mean RMSE values and the range of these values with the Strict DC method for the same $\mathrm{af} = 6$ but different pseudo-random sampling patterns. As seen, the nonconvex norm $\mathrm{p} = 0.50$ was the best choice in terms of RMSE. The range of the RMSE values offers only minor variations around the mean value.*

For the Relaxed DC method, Figure 16a exhibits significant variations in the RMSE values with a given λ start value for different sampling patterns. The start values of λ have been sorted in ascending order from top to bottom to better visualize possible trends. It can be seen that at λ indices 3 and 4 ($\lambda = 0.0255$ and 0.0260), significantly higher RMSE values were obtained. However, at index 5 ($\lambda = 0.0278$), the RMSE values returned to lower values.

These strong variations cannot be observed in Figure 16b for the Strict DC method. Here, the reconstruction quality in terms of RMSE was significantly dependent on the chosen norm p, but not on the sampling pattern. For each individual value for p, the CS reconstruction using the Strict DC method resulted in an equivalent reconstruction quality. Again, a nonconvex norm was beneficial. In this particular case, as shown in Figures 14 and 15, $p = 0.50$ was optimal in terms of RMSE. The mean RMSE values and the range of the RMSE values are given in Table 3.

Only 90 out of 400 reconstructions with the Relaxed DC method resulted in RMSE values that were lower than the maximum RMSE value for the Strict DC method (0.2008). The lowest RMSE value obtained with the Relaxed DC method was 0.1081.

The best and worst reconstruction results from these simulation data are shown in Figure 17.

The regularization parameter λ must be re-adjusted when a pseudo-radial sampling pattern is chosen instead of a pseudo-random. Figure 18 shows that, for the same start value for λ (i.e., same x-axis index in Figure 18), a pseudo-radial sampling pattern generally led to lower RMSE values than the pseudo-random Cartesian pattern. The spikes in the RMSE course were also less prominent in the pseudo-radial reconstructions. Figure 19 shows the best reconstruction results for both investigated sampling patterns. While the pseudo-random Cartesian pattern led, in the best case, to a corrupted reconstruction, the pseudo-radial pattern resulted in a nearly successful reconstruction.

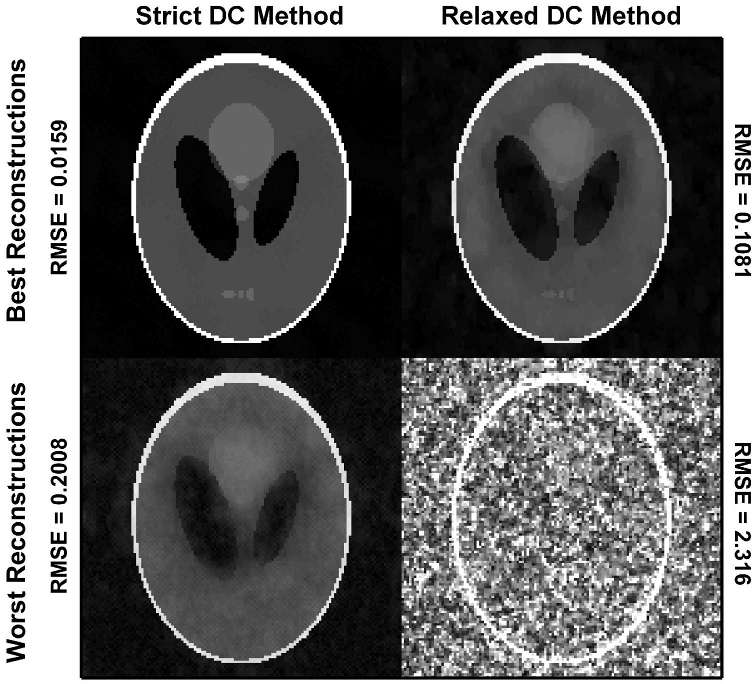

Figure 17: *Comparison of the best and the worst reconstruction results obtained with the Strict and the Relaxed DC method.*

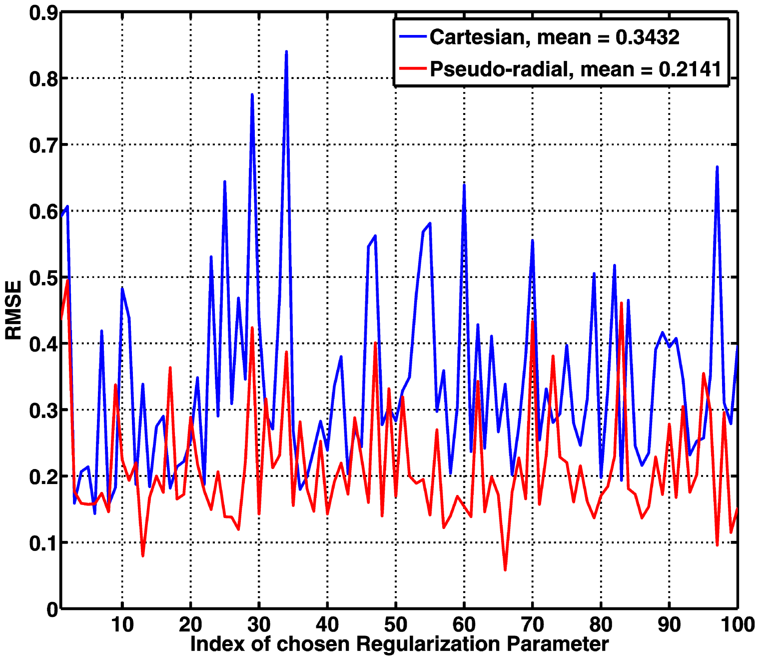

Figure 18: *Results of the comparison between a pseudo-random Cartesian sampling pattern and a pseudo-radial pattern. Each index on the x-axis corresponds to one particular start value for λ. Each start value was used for a reconstruction with the Cartesian and the radial sampling patterns. As seen, the radial pattern generally led to lower RMSE values. Furthermore, the results for both sampling patterns significantly differed in parts. Thus, the optimal start value for the regularization parameter changed when a pseudo-radial pattern was chosen instead of a pseudo-random Cartesian pattern.*

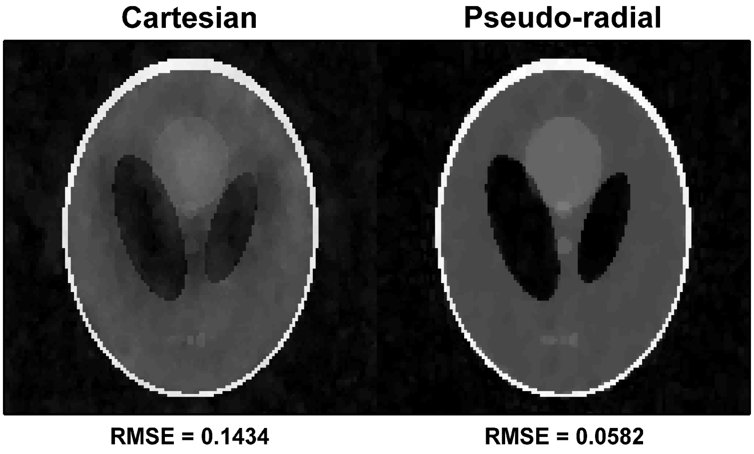

Figure 19: *The best reconstruction results for both sampling patterns. As seen, the pseudo-radial sampling pattern led to a nearly successful reconstruction.*

Discussion

Dependency on the acceleration factor

The strong dependency of λ on the applied acceleration factor makes it cumbersome to use the Relaxed DC method. The first term in Equation 2.7 $\|F_u x - y\|_2^2$ is solely based on the difference between the undersampled acquired data and the reconstructed data at the same k-space locations. Thus, this term differs for each individual acceleration factor after the first iteration. Even though the sparsity constraint ($\|\Psi x\|_p$, second term in Equation 2.7) is the same, the weighting between these two terms must be altered. This is achieved by re-adjusting the regularization parameter λ.

Dependency on the sampling pattern

According to Figure 16, the choice of λ is significantly affected by the sampling pattern. Even though the same acceleration factor and the same type of sampling pattern (pseudo-random with densely sampled inner k-space) was chosen, the RMSE values differed tremendously according to the start value of λ. This can be explained by the different point spread functions that result from different sampling patterns. These point spread function variations lead to different incoherent artifact patterns in the initial guess for the CS reconstruction. The initial guess has direct influence on the sparsity term $\|\Psi x\|_p$ in Equation 2.7, since different incoherent undersampling artifacts lead to different values for the sparsity term in the first iteration. Thus, this should influence the choice of the regularization parameter, which is reflected by the data. The Strict DC method is not significantly affected by the exact choice of the sampling pattern. Again, this clarifies why it is more convenient and beneficial to use a CS algorithm without a regularization parameter.

Furthermore, the correct choice of the regularization parameter is affected when a pseudo-radial sampling pattern is used instead of a pseudo-random Cartesian sampling pattern. This makes it difficult to apply the Relaxed DC method to undersampled MR data. In the Relaxed DC method, the correct choice of λ for a particular undersampling factor in one particular sampling pattern cannot be conveyed to a different sampling pattern (e.g., from Cartesian to radial). The reason can again be found in the different point spread functions previously mentioned. The radial reconstructions were generally of better quality than the Cartesian reconstructions. This is a result of the artifact nature of (pseudo-)radial trajectories where the resulting incoherent artifacts lead to a less corrupted initial guess than with a pseudo-random Cartesian sampling pattern (Figure 20).

Dependency on noise and resolution

In this chapter, the possible dependency on noise and resolution was not further investigated. However, both parameters should influence the choice of the regularization parameters. This shall be argued by considering Equation 2.7.

The possible influence of noise results from altering both the data consistency term $\|F_u x - y\|_2^2$ and the sparsity term $\|\Psi x\|_p$. Even if the same sampling patterns are used, a noise-free dataset offers different values for both terms of Equation 2.7 than a noisy dataset. The quantitative change depends on the noise level. Therefore, a readjustment of the regularization parameter λ is very likely. However, for low noise levels, this influence should be less pronounced.

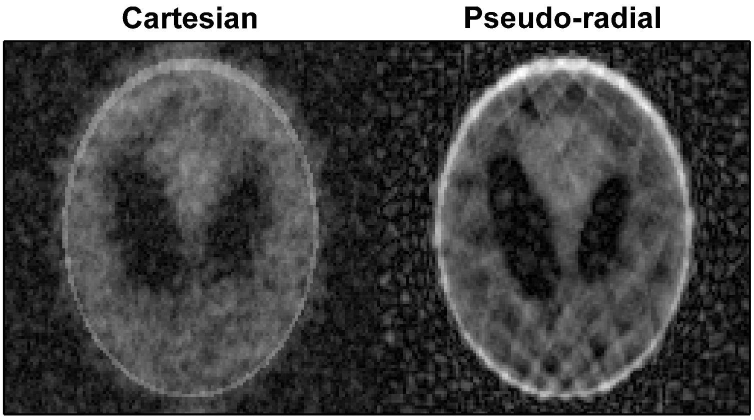

Figure 20: *Artifact nature in the undersampled Shepp-Logan phantom for an acceleration factor of approximately 6 (20 pseudo-radial projections). The pseudo-radial trajectory exhibits a better representation of the phantom, more of the ellipses can be recognized.*

Altering the resolution can also require a readjustment of the regularization parameter. This possible influence can be explained by the different phantom sparsity levels for different resolutions. This is clarified in Table 4, where the sparsity of the noise-free Shepp-Logan phantom is given after transformation into the Discrete Gradient domain (see also Figure 6). By changing the resolution, the sparsity is directly affected. Since the sparsity term $\|\Psi x\|_p$ changes, this leads to different solutions for Equation 2.7 depending on the resolution. This change usually directly affects the proper choice of λ.
The Strict DC method is expected to lead to successful reconstructions independent of the resolution. In fact, the sparsity term $\|\Psi x\|_p$ is also affected. However, since only this term is minimized in Equation 2.9 and no regularization parameters need to be adjusted, the Strict DC method will minimize the sparsity term and, thus, converge to the desired solution. Since no parameter must be readjusted if the resolution of the object changes, the Strict DC method is, therefore, more applicable than the Relaxed DC method. However, it should be kept in mind that assuming “imperfect“ noisy data as perfect through the strict data consistency constraint is a sure artifact source. These artifacts will be investigated in greater detail in the following chapters.

Resolution	**# Non-zero elements**	**Sparsity**
128x128	1080	6.59 %
192x192	1630	4.42 %
256x256	2183	3.33 %

Table 4: *Effect of the resolution on the sparsity of the Shepp-Logan phantom in the Discrete Gradient domain. The sparsity reduces as the resolution increases. The sparsity is calculated as the number of non-zero elements per total pixel number.*

3.1.3 Conclusion

It could be demonstrated that, with the Relaxed DC method, the choice of the regularization parameter λ exhibits significant dependency on the acceleration factor and the applied sampling pattern. Furthermore, it was argued that the noise level and the resolution should

also affect the choice of λ. The dependency of λ on the acceleration factor and the sampling pattern required the time-consuming empirical determination of a correct regularization parameter. Therefore, the Relaxed DC method will likely be difficult to use. For this reason, a detailed investigation of the dependency of λ on noise or the resolution was discarded.
In the presented simulations, determination of only one regularization parameter was necessary. If further sparsity constraints were introduced into Equation 2.7, more regularization parameters occurred. It is obvious that the computational complexity of determining two independent regularization parameters is higher than for only one parameter. Therefore, this would very likely be even more time-consuming and cumbersome.
The Strict DC method resulted in reproducible reconstruction quality independent of the acceleration factor and the applied sampling pattern.
The simulations were performed on the Shepp-Logan phantom, which can be sparsified by Discrete Gradients. However, similar results can be expected for data sparse in a different domain (e.g., image space).
The Strict DC method was chosen as the CS reconstruction algorithm since determination of a regularization parameter was unnecessary. However, the interdependencies between the chosen norm, the noise level, and the acceleration factor should be investigated. This will be done in each individual chapter to create a stringent, logical line of argumentation from simulation results to *in vivo* measurements.

3.2 Influence of ϵ on the reconstruction quality of the Strict DC method

3.2.1 Motivation

In Equation 2.11, ϵ was introduced as a smoothing parameter that allows derivation of the norm. As described by Lustig et al. [16], this parameter was set to a fixed value $\epsilon = 10^{-4}$ in the Nonlinear Conjugate Gradient algorithm. However, in the Gradient Descent algorithm, ϵ was decreased until a value ϵ_{end} was reached at which the reconstruction terminated. It might be thus argued that instead of a regularization parameter such as that given in the Relaxed DC method, another parameter has to be appropriately adjusted to obtain an accurate reconstruction. This section investigates the influence of ϵ_{end} on the reconstruction quality of the Strict DC method and whether or not it is problematic to correctly adjust the terminating value ϵ_{end}.

3.2.2 Methods

The simulations were performed on the noise-free 256x256 Shepp-Logan phantom. A fixed acceleration factor $af = 6$ and a fixed nonconvex norm $p = 0.75$ were chosen for all reconstructions. The start value for ϵ was always set to 1 while the end values were varied: $\epsilon_{end} = \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$. The choice of the start value of ϵ as 1 always holds if the signal intensity in the image to be reconstructed is normalized to 1, or to a value on the order of magnitude 1. This can be achieved by normalizing the initial guess (the undersampled image) to 1. To exclude the influence of the sampling pattern, 20 different pseudo-random sampling patterns with $af = 6$ and a densely sampled inner k-space (17x17 k-space points corresponding to 0.4 % of the total k-space) were created. For each sampling pattern, the reconstructions with the Strict DC method for the different values for ϵ_{end}

were performed. The resulting quality of the reconstructions was again quantified by the RMSE metric (Equation 3.1).

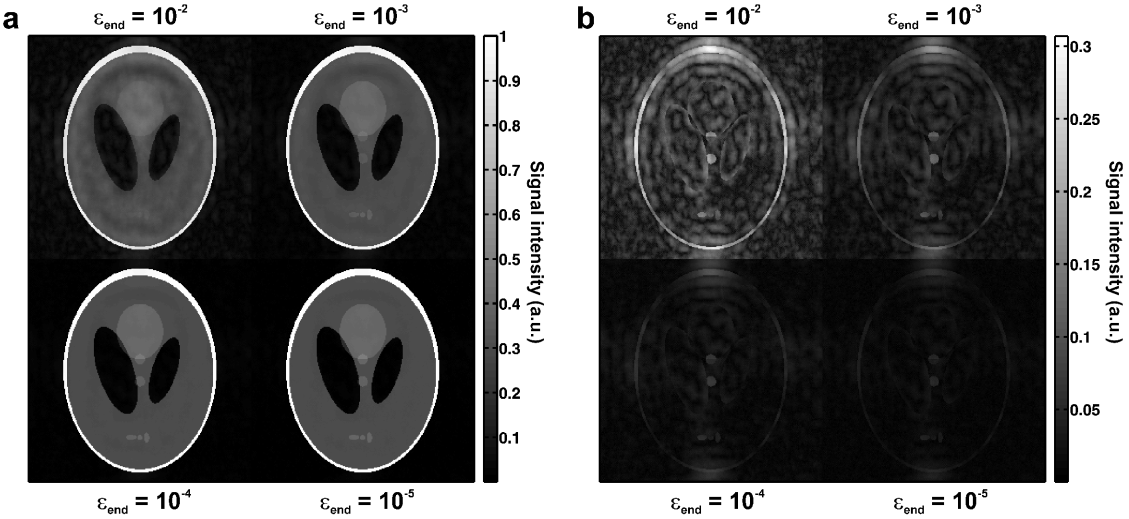

Figure 21: *Influence of ϵ_{end} on the reconstruction quality. (a) The recovered Shepp-Logan phantoms for different values of ϵ_{end}. The related RMSE values can be found in Table 5. Using $\epsilon_{end} \leq 10^{-4}$ led to successful reconstructions. (b) The differences between the reconstructions and the noise-free reference phantom demonstrate that there is only minor improvement using $\epsilon_{end} = 10^{-5}$ instead of 10^{-4}.*

3.2.3 Results

Figure 21 shows exemplary reconstruction results for different values of ϵ_{end}. In Figure 21a, the recovered Shepp-Logan phantom is presented. As confirmed by the difference images between the reconstructions and the reference in Figure 21b, the reconstruction for $\epsilon_{end} \leq 10^{-4}$ was successful. The benefit of using $\epsilon_{end} = 10^{-5}$ over $\epsilon_{end} = 10^{-4}$ was only minor. However, compared to the results using $\epsilon_{end} = 10^{-4}$ (Table 5), 90 additional iterations were necessary for $\epsilon_{end} = 10^{-5}$ to reduce the mean RMSE value by a factor of 2. The range of RMSE values offered only minor variations around the mean value (Table 5).

ϵ_{end}	**# Iterations**	**Mean RMSE**	**RMSE Range**
10^{-2}	210	0.1509	[0.1451 0.1590]
10^{-3}	300	0.0616	[0.0581 0.0657]
10^{-4}	420	0.0240	[0.0219 0.0271]
10^{-5}	510	0.0144	[0.0133 0.0161]

Table 5: *The effect of ϵ_{end} on the Strict DC method reconstruction quality for an acceleration factor af = 6 and a fixed nonconvex norm p = 0.75. As seen, the RMSE decreased when ϵ_{end} was decreased. When the RMSE of the reconstruction fell below the threshold RMSE $\leq$ 0.05, the reconstruction was assumed to be successful. Thus, $\epsilon_{end} = 10^{-4}$ was sufficient to achieve successful reconstructions using the Strict DC method. No standard deviation is given since the distribution of RMSE values cannot be analytically given. The difference in iterations between neighboring ϵ_{end}s is unequal because ϵ was reduced by a factor of 2 every 30^{th} iteration.*

3.2.4 Discussion

The influence of ϵ_{end} on the reconstruction quality was investigated in this section. An appropriate value for ϵ_{end} has to be found to achieve successful reconstructions. As previously mentioned, reconstructions exhibiting a RMSE ≤ 0.05 can be viewed as successful. This was realized for $\epsilon_{end} \leq 10^{-4}$. However, the differences in the visible reconstruction quality (compare Figure 21a) for $\epsilon_{end} = 10^{-4}$ and 10^{-5} were negligible. This can also be seen in the difference images in Figure 21b, where only minor improvements for $\epsilon_{end} = 10^{-5}$ over 10^{-4} could be observed. Thus, decreasing ϵ to 10^{-4} for noise-free data was sufficient.

For noisy data, $\epsilon_{end} = 10^{-4}$ could also be an appropriate choice. By considering Equation 2.11, it becomes clear that ϵ measures the accuracy of the approximated norm against the "real" norm. Assuming that each pixel value x is as accurate as ϵ results in an absolute value of:

$$|x \pm \epsilon| = \sqrt{(x \pm \epsilon)^2} = \sqrt{x^2 \pm 2x\epsilon + \epsilon^2} \tag{3.2}$$

It is important to remember that the image intensity is normalized to 1 (i.e., $x \leq 1$ and $x \geq 0$).

Assuming a SNR of 100, each pixel has an uncertainty of 10^{-2}. If $x = 10^{-2}$ and $\epsilon = 10^{-4}$, this leads to:

$$\sqrt{x^2 \pm 2x\epsilon + \epsilon^2} \approx x \tag{3.3}$$

The same holds for the other extremum $x = 1$. Therefore, using $\epsilon_{end} = 10^{-4}$ should lead to sufficiently accurate reconstructions for MR datasets.

3.2.5 Conclusion

The presented results and the discussion demonstrate that using $\epsilon_{end} = 10^{-4}$ allows reliable, reproducible, and accurate reconstructions for all SNR levels if the signal intensity in the initial guess is normalized to 1. The choice of ϵ_{end} affects the desired accuracy of the reconstruction. Appropriately adjusting ϵ_{end} is unproblematic since it must only be guaranteed that ϵ_{end} is sufficiently small to achieve the desired reconstruction accuracy. Usually, this accuracy is limited by the noise corrupting the data. Nonetheless, as seen in previous sections of this chapter, reconstructions of noisy data using $\epsilon_{end} = 10^{-4}$ led to successful results.

3.3 Conclusion of the simulation results

The purpose of this chapter was to determine which CS method is better suited to reconstructing MR datasets. With the Relaxed DC method, the proper regularization parameter λ choice depended on the acceleration factor and the applied sampling pattern. Furthermore, it was argued that the noise level and the resolution of the desired object should also influence the choice of the regularization parameter. Because the regularization parameter must be adjusted for each individual application, the Relaxed DC method can thus be considered cumbersome and time consuming. This introduces a major user dependent influence on the reconstruction quality since the outcome of the CS reconstruction using the Relaxed DC method depends on the choice of the regularization parameter. It is impossible to tell if the chosen regularization parameter is optimal if no reference image is available.

The simulations revealed that the Strict DC method leads to reliable, stable, and reproducible results without requiring regularization parameter adjustment. Therefore, user interaction is unnecessary during the reconstruction process. It was demonstrated that the only remaining parameter, ϵ_{end}, can be appropriately chosen to 10^{-4} if the image intensity in the initial guess is normalized to 1. This results in a “Black Box“ algorithm that reliably leads to the desired reconstruction. The drawback to the Strict DC method is that it treats “imperfect“ noisy data as perfect, which is a source of reconstruction artifacts. These artifacts will be further investigated in the following chapters. Furthermore, the interdependencies of the Strict DC method between noise level, acceleration factor, and norm will be examined in each individual chapter to allow a stringent, logical line of argumentation for each demonstrated application.

All presented results were derived on the Shepp-Logan phantom. However, when a phantom with a comparable sparsity in any other sparsity domain would be used, similar results can be expected. Specific simulations will be presented for each investigated application in the following chapters.

It was demonstrated that the regularization parameter λ is dependent on factors such as the sampling pattern and the acceleration factor. However, the presented dependency is only valid for the Shepp-Logan phantom. When an object with a different sparsity in another basis than the Shepp-Logan phantom is investigated, the dependency might be different.

It should be mentioned that it is possible to solve Equation 2.9 using the Nonlinear Conjugate Gradient algorithm. However, this should not significantly alter the results of the presented simulations for the Strict DC method since Gradient Descent and Conjugate Gradient should lead (in the case of sufficient iterations and the same choice for the (final) approximation parameter ϵ in Equation 2.11) to the same result.

In conclusion, the Strict DC method was chosen as the preferred CS reconstruction method. It will be used throughout the remaining chapters in this thesis. The implementation used is, if not differently specified, given in Section 2.3.3 in pseudo code.

4 Accelerated ^{19}F Chemical Shift Imaging at 7 T

4.1 Motivation

The MR community has regained interest in ^{19}F MRI during the last few years. The low natural abundance of ^{19}F in living tissue leads to a negligible ^{19}F background signal [41]. This allows unambiguous detection of many exogenously administered ^{19}F markers exhibiting unique spectral signals. These advantages, along with its gyromagnetic ratio comparable to ^{1}H, make fluorine a suitable marker for molecular imaging [42], cell tracking [41, 43], and other biological and medical applications [44–46]. The identification of different ^{19}F markers can be achieved using chemical shift imaging (CSI) [42, 47, 48], the specific properties of which will be discussed in the next section.
Certain limitations, however, must be considered if ^{19}F CSI is used. In some applications, ^{19}F imaging suffers from a low concentration of markers at the region of interest, resulting in a low signal-to-noise ratio (SNR). Furthermore, spectral information acquired in an acceptable *in vivo* measurement time comes at the expense of spectral or spatial resolution. Additionally, when the spatial distribution of markers is unknown, such as with ^{19}F tracking of labeled cells, 3D imaging is preferable. Moreover, the combination of spectroscopic and 3D imaging can lead to unacceptably long measurement times for *in vivo* experiments. Therefore, strategies to accelerate 3D ^{19}F CSI experiments are of special interest and have been proposed. These include echo-shifting methods [49], wavelet encoding [50], and Non-Cartesian trajectories [51].
CS is a novel strategy to accelerate spectroscopic experiments. The possible reduction in scan time using CS with spectroscopic imaging has been investigated in hyperpolarized ^{13}C CSI experiments [52–54]. In these papers, the sparsity in the spectral dimension was exploited using an echo-planar flyback method. The current chapter investigates the potential of CS using fully phase-encoded ^{19}F CSI. Because its signal distribution is often spatially sparse, ^{19}F is a nucleus well-suited for CS.

To study the potential of CS to accurately reconstruct undersampled ^{19}F CSI data, a proof-of-principle investigation using 2D CSI datasets was performed. For this, retrospectively undersampled phantom and *in vivo* mouse datasets were reconstructed. The 2D CSI experiments will be presented in the next section of this chapter.

Following these initial 2D demonstrations of principal feasibility, the proposed method was extended to 3D. Since with some applications ^{19}F images suffer from low SNR, the noise level and its effect on reconstruction quality were of special focus in this work. To address this issue, simulations were performed with various combinations of noise levels and reconstruction parameters. The simulation results were confirmed by a real ^{19}F signal distribution that was obtained from a 7 T *ex vivo* mouse dataset. To validate the proposed method *in vivo*, undersampled and fully sampled datasets of a photothrombotic (PT) mouse model were acquired at 7 T. The undersampled *in vivo* data were reconstructed using optimized parameters obtained from the results of the simulations and *ex vivo* experiments.

4.2 Biological background

Conventional proton-based MRI exams only disclose non-specific tissue alterations (e.g., damage, or edema formation). However, the particular nature of these tissue alterations can be investigated in a more differentiated way by the targeted application of contrast agents or markers. In this context, ^{19}F markers proved superior to iron-oxide particle-based contrast agents ((U)SPIOs: (**U**ltrasmall) **S**uper**P**aramagnetic **I**ron-**O**xide particles) since they provide a unique *in vivo* signal. The (U)SPIOs lead to a locally faster T_2 decay, resulting in dark spots in a T_2-weighted MR image. This effect enables *in vivo* assessment of cellular infiltration into the nervous system of animal models [55] and patients with stroke and multiple sclerosis [56]. However, the (U)SPIO-induced image hypointensity has been shown as ambiguous to hemorrhages or passive diffusion of the (ultra)small iron-oxide particles through the defective blood-brain barrier [56]. Therefore, markers containing ^{19}F are advantageous because they can be unambiguously detected and provide high specificity. It has been recently demonstrated [46] that an intravenous injection of perfluorocarbons (PFC) leads to a significant uptake by macrophages. Thus, areas infiltrated by macrophages can be visualized by ^{19}F MRI. This chapter investigates a model of brain ischemia in mice at 7 T. The model is called photothrombosis (PT) and represents a focal lesion model of the central nervous system characterized by thrombotic occlusion of small cortical vessels due to endothelial damage (endothelium: thin layer of cells lining the interior surface of blood vessels) and subsequent macrophage inflammation [57]. These PT lesions can be visualized using (U)SPIOs in T_2-weighted MR images [58, 59] and by ^{19}F imaging [46, 47]. This chapter focuses on ^{19}F imaging of PT lesions since ^{19}F MRI allows unambiguous detection of the utilized marker. Furthermore, as already outlined in the introduction, ^{19}F signal distribution is sparse due to the negligible background signal which is well-suited for CS. The hypointensities caused by (U)SPIOs, however, cannot be easily transformed to a sparse domain. Therefore, it was unknown how to exploit them in the CS framework.

4.3 Chemical Shift Imaging

This section discusses the physical basics of CSI and presents the sequence utilized throughout this chapter.

4.3.1 Chemical Shift

Several mechanisms influencing the Larmor frequency were mentioned in Chapter 1:

- The Larmor frequency is primarily dependent upon the magnetic field strength $|\vec{B}_0|$ that the spins observe. This also includes inhomogeneities in the main magnetic field that lead to an accelerated decay of transversal magnetization. The resulting characteristic time constant T_2^*, can be modeled using Equation 1.16, where B_0 inhomogeneities are explicitly included.
- Furthermore, the Larmor frequency depends on the nucleus species which is investigated. This influence is represented by the gyromagnetic ratio γ in Equation 1.3 .
- The spins exhibit a magnetic moment, which results from the non-vanishing nuclear spin. The magnetic moment is described by Equation 1.5. This means that the spin

is itself a small magnet, creating its own magnetic field. Due to diffusion, spins can approach each other very closely. This leads to a shift of the local Larmor frequency since the static magnetic field $\vec{B}_0$ is superimposed by the weaker magnetic field of the passing spin. Hence, this spin-spin interaction eventually results in the T_2 decay of the transversal magnetization.

Besides these mechanisms, the immediate chemical environment also influences the Larmor frequency. The electron hulls of neighboring atoms and molecules can shield or anti-shield the main magnetic field $\vec{B}_0$. This influence can be modeled by the so-called chemical shift σ, which leads to a shift of the locally observed magnetic field strength:

$$B_{s,j} = (1 - \sigma_j)\, B_0 \tag{4.1}$$

Here, $B_{s,j}$ is the shifted magnetic field strength caused by the chemical species j, while σ_j is the dimensionless chemical shift of this species.
It can be immediately seen that this shift in the field strength alters the resonance frequency:

$$\Delta\omega\,(\sigma_j) = 2\pi\bar{\gamma}\,(B_0 - B_{s,j}) = -2\pi\bar{\gamma}\sigma_j B_0 = -2\pi\bar{\gamma}\sigma_j B_0 = -\sigma_j\omega_0 \tag{4.2}$$

Using this relation, a chemical shift axis can be defined by calculating $\sigma_j = -\frac{\Delta\omega(\sigma_j)}{\omega_0}$ with respect to a reference nucleus with the Larmor frequency ω_0. Typically, the chemical shift is given in ppm (parts per million). In the current chapter, this reference nucleus is a ^{19}F species of a certain fluorinous substance. The chemical shift axis adds one dimension in the dataset to the existing spatial dimensions.

4.3.2 Sequence for accelerated CSI

In chemical shift imaging (CSI), frequency information must be acquired in addition to the spatial information. Therefore, at a certain point in k-space the free induction decay (FID) must be sampled. In the CSI implementation for this work, all gradients used to spatially encode the sample volume were phase encoding gradients. The accumulated phase due to the chemical shift can be expressed as

$$\phi_j(t) = -\sigma_j 2\pi\bar{\gamma} B_0 t = -\sigma_j 2\pi f_0 t = -\sigma_j\omega_0 t \tag{4.3}$$

where f_0 is the Larmor frequency in Hz. Using Equation 4.3 and observing 1.26 results in an analytical description of the 3D CSI signal:

$$S(k_x, k_y, k_z, t) = \sum_s \int\int\int \rho(x, y, z, \sigma_s) e^{-i2\pi(k_x x + k_y y + k_z z - \sigma_s f_0 t)} dx dy dz \tag{4.4}$$

The sum, therefore, runs over the contemplated nucleus species, which are chemically shifted.
Switching on all required phase encoding gradients leads to a certain point in k-space. The gradients are then switched off, the FID is sampled and stored, and the next k-space point is subsequently chosen. This is schematically shown in the sequence diagram in Figure 22. The measurement time T_{meas} for a CSI experiment is given by:

$$T_{meas} = N_x N_y N_z T_R \tag{4.5}$$

where N_x, N_y and N_z are the number of voxels along the x-, y- and z-direction and T_R is the repetition time of the MR sequence. Especially in 3D, this can lead to extraordinary long

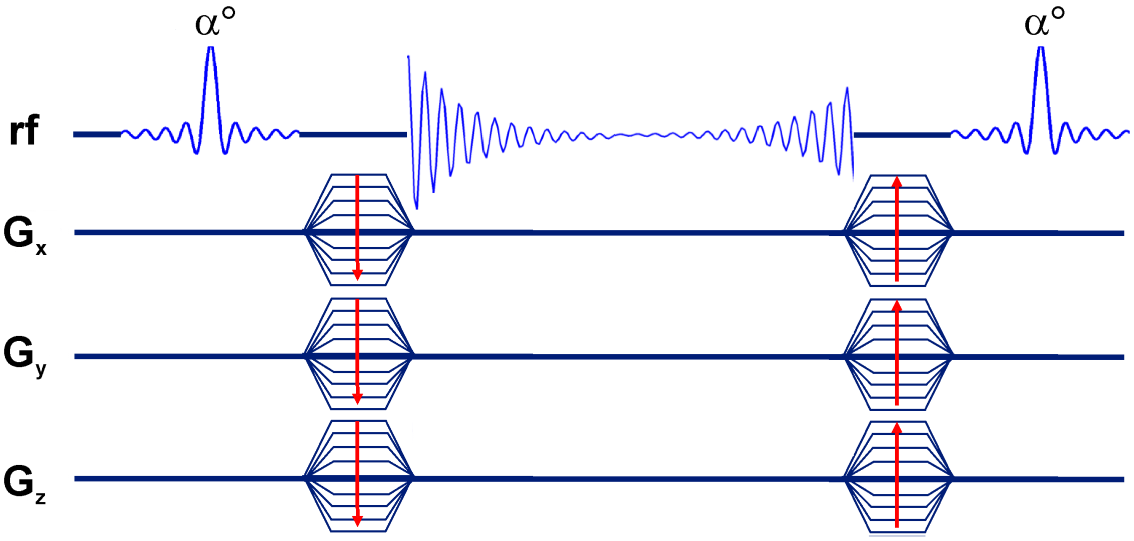

Figure 22: *Schematic sequence diagram of the SSFP CSI sequence used in this dissertation for the 2D and 3D* ^{19}F *experiments. Details on the imaging parameters are given in the respective section.*

measurement times. Thus, either the spatial resolution is reduced or 2D CSI is performed. The latter is only reasonable if it is known *a priori* where the expected signal will appear. Otherwise, 3D imaging is always preferable.

For the purposes of this work, a Steady State Free Precession (SSFP) CSI sequence was used with either two or three spatial dimensions (Figure 22). A detailed description of the imaging parameters is given in the respective section below. The signal level for a SSFP sequence can be estimated as (if $T_R \ll T_1, T_2$):

$$S_{ss} \propto \frac{M_0 \sin\alpha}{1 + T_1/T_2 - (T_1/T_2 - 1)\cos\alpha} \tag{4.6}$$

where S_{ss} is the steady state signal intensity, M_0 is the equilibrium magnetization and α is the flip angle of the SSFP sequence.

4.4 Compressed Sensing aspects

This section discusses how the basic requirements for CS (see also Chapter 2) are met by the CSI experiments.

4.4.1 Sparsity

In the context of this chapter, the sparse domain is the image space because only regions containing ^{19}F markers provide signal. Therefore, no transformations or any other data preparation steps must be undertaken prior to CS reconstruction ($\Psi = \mathbb{1}$, see also Section 2.3).

4.4.2 Sampling Pattern

This work focuses on purely phase-encoded 2D and 3D CSI. Thus, a pseudo-random 2D or 3D sampling scheme was chosen for all undersampled experiments. It has been previously

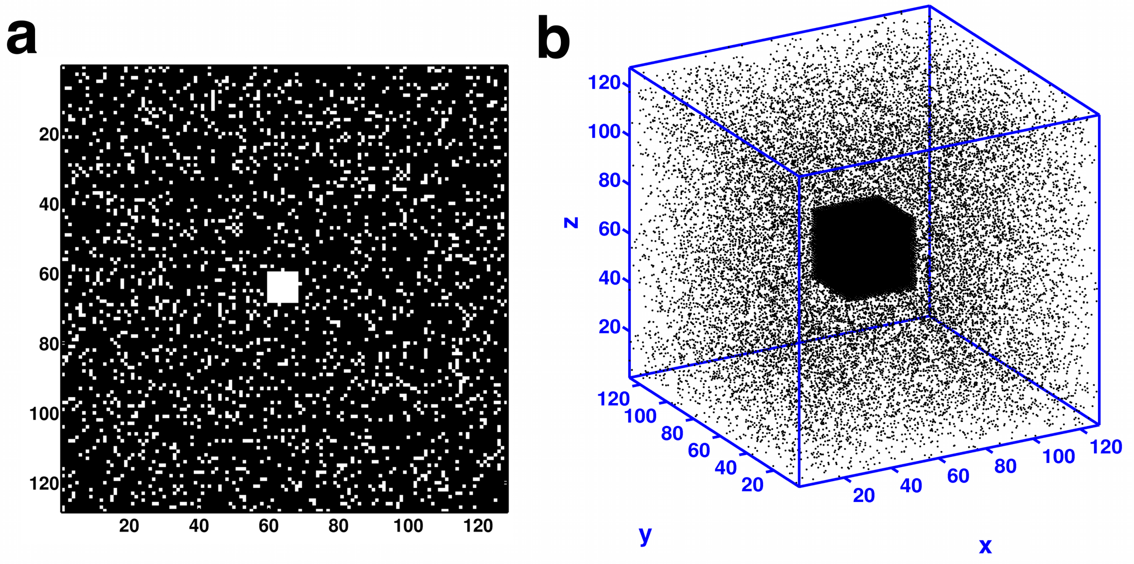

Figure 23: *Sampling patterns for 2D and 3D CSI as utilized in this chapter. (a) The displayed acceleration factor in this 2D pattern was 8. This pattern can be directly applied for accelerated CSI. (b) For illustration purposes, the acceleration factor in the displayed 3D sampling pattern was set to 100 and the fully sampled central k-space region was exaggerated. The actually used patterns exhibited acceleration factors of 2 to 10 and the fully sampled central k-space region was significantly smaller.*

demonstrated that such a pattern is well-suited for CS [16, 60, 61] due to the incoherent artifacts introduced in cases of undersampling. The k-space center was densely sampled to obtain a low resolution image of the object. This provided a good initial guess for the CS algorithm and, thus, accelerated the algorithm convergence. Sampling patterns for accelerated CSI experiment are shown in Figure 23. Details about the specific utilized sampling patterns are given in the respective sections.

4.4.3 Reconstruction

All datasets presented in this chapter were reconstructed using the Strict DC method with $\Psi = \mathbb{1}$ as described in Section 2.3.2. This CS method was chosen because it was unnecessary to adjust regularization parameters. This prevented subjective influences on the reconstruction quality due to the choice of such parameters.

4.5 2D ^{19}F CSI

4.5.1 Methods

In order to give a proof of principle that ^{19}F CSI can benefit from CS, two 2D CSI experiments were performed:

- A phantom of six tubes filled with perfluoro-15-crown-5-ether (PF15C) emulsion (VS580H, Celsense Inc., Pittsburgh, PA, USA) was imaged using a CSI sequence described in Section 4.3.2 with the following imaging parameters: Bruker BioSpec 7 T (Bruker Biospin GmbH, Rheinstetten, Germany), matrix size 64x64, FOV 40x40 mm^2, $T_R = 17$ ms, receive bandwidth 38.55 Hz per spectral point, 1 and 6 averages. The dataset was retrospectively undersampled by various acceleration factors ($af =$ 2, 4, 6, 8) using a sampling pattern with a densely sampled central k-space region (11x11 k-space points accounting for 2.95 % of the total k-space, see also Figure 23a). The acceleration factor used in this work is defined as the ratio of all data points required for full Fourier encoding N_{full} to the acquired data points N_{acq}:

  $$af = \frac{N_{full}}{N_{acq}}$$

  Thus, the acceleration factor is equivalent to the reduction factor in scan time compared to a fully sampled k-space. The undersampled data were then CS reconstructed point-by-point along the spectral dimension using $p = 0.75$. Thus, only spatial sparsity was exploited.

- Animal experiments were performed in accordance with institutional guidelines and were approved by state authorities. Focal cerebral ischemia was induced in an adult C57/BL6-mouse by photothrombosis (PT) of cortical microvessels under inhalation anesthesia with enflurane in a 2:1 nitrogen/oxygen atmosphere, as previously described [58, 62]. A fiber optic bundle of a cold light source was stereotactically centered 2 mm posterior and 2.4 mm lateral from Bregma exposed via a dorsal midline skin incision. A sterile-filtered rose Bengal solution (0.2 ml) was given intraperitoneally and the brain was illuminated for 20 min. Afterwards, the skin was sutured and the mouse was allowed to recover. This procedure resulted in cone-shaped cortical infarction without clinically overt symptoms. For MRI measurements, the mouse was anesthetized using 1.5 % isofluran in a 2 l/min oxygen atmosphere.

  For the 2D ^{19}F CSI experiments, 125 μl of a commercial VS580H emulsion was injected into the tail vein of the mouse directly after PT. A CSI experiment was performed 72 hours after administration of the ^{19}F marker. The sequence used is described in Section 4.3.2 with the following parameters: Bruker BioSpec 7 T, matrix size 48x48, FOV 30x30 mm^2, $T_R = 24$ ms, receive bandwidth 25.00 Hz per spectral point, 65 averages.

  The *in vivo* data were retrospectively undersampled by various acceleration factors ($af =$ 2, 4, 6, 8) using a similar sampling pattern as described above (11x11 k-space points accounting for 5.25 % of the total k-space). The undersampled data were then CS reconstructed point-by-point along the spectral dimension using $p = 0.75$. Again, only the spatial sparsity of the dataset was exploited.

Both phantom and *in vivo* experiments were fully sampled and retrospectively undersampled to mimic accelerated imaging.

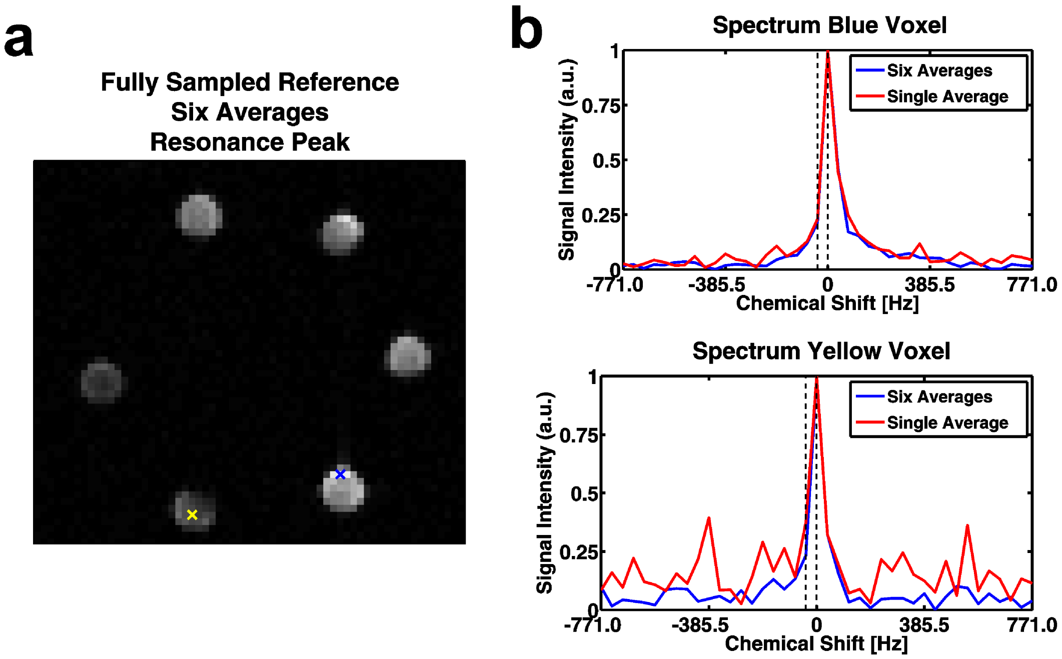

Figure 24: *(a) Reference image from the fully sampled six times averaged dataset to illustrate the position of the chosen voxels for the spectra of (b). The blue voxel is in a high SNR region while the yellow one lies in a low SNR area. (b) Comparison of the spectra of the single-averaged and six times averaged dataset. As seen in the upper plot of the spectra from the blue voxel, there is no significant difference visible besides the improved SNR of the six times averaged dataset. However, in the lower plot from the yellow voxel, the SNR difference results in significant differences of the spectra. The vertical lines illustrate the spectral positions of the images in Figures 25 (on resonance) and 26 (one spectral point shifted from the main resonance).*

4.5.2 Results

Phantom experiments

Figures 25 and 26 clarify how averaging benefits the CS reconstruction quality. CS reconstructions of two adjacent spectral points are shown. Figure 24 illustrates the position of the spectral points within the spectra. On the main resonance (dashed vertical line at 0 Hz), the signal intensity in both chosen voxels (indicated by the blue and yellow cross in Figure 24a) is significantly greater than the average noise level. However, as soon as a spectral point adjacent to the main resonance is considered (left dashed line), the six-times-averaged dataset exhibits improved SNR compared to the single averaged dataset. The spectra of the blue voxel (upper diagram of Figure 24b) exhibits sufficient SNR for both averaging factors. This is no longer the case for the signal in the yellow voxel (lower diagram of Figure 24b). With the single-averaged dataset, the signal is comparable to the noise level while sufficient SNR can be observed for the six times averaged dataset. Thus, the CS reconstruction is expected to improve for the six times averaged dataset compared to the single averaged dataset. In Figures 25 and 26, this is demonstrated in greater detail.

Figure 25 compares at the VS580H marker resonance frequency the fully sampled six times averaged reference to the undersampled and CS reconstructed chemical shift images. At the resonance frequency of the ^{19}F marker substance, sensitivity differences can be observed between the single-averaged and the six times averaged reference dataset (as seen in Figure 25a). However, the information content is the same, all tubes can be clearly seen, and all

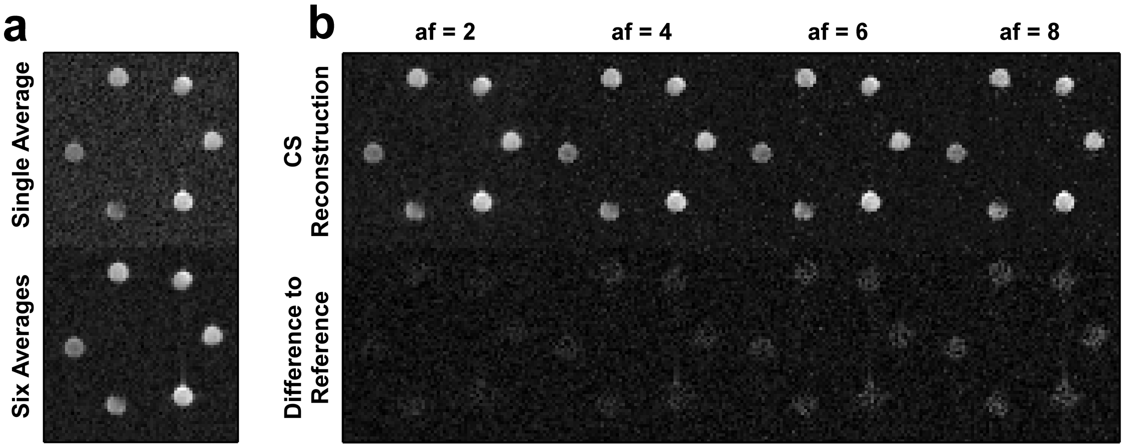

Figure 25: *CS reconstruction of the phantom tubes on the VS580H marker resonance peak. (a) No significant differences in sensitivity can be observed between the single averaged and the six times averaged reference. (b) All CS reconstructions from the six times averaged dataset resulted in successful images exhibiting no visible information loss compared to the reference.*

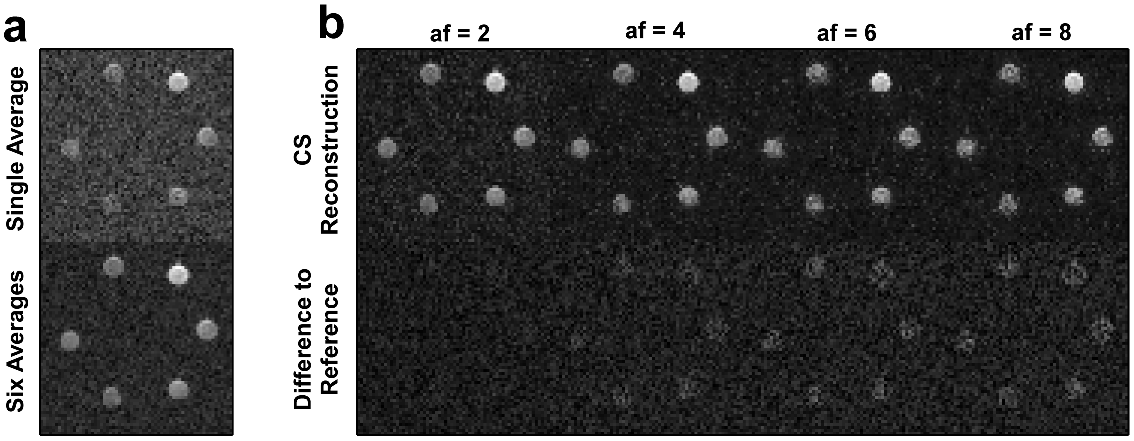

Figure 26: *CS reconstruction of the phantom tubes shifted one spectral point from the VS580H marker resonance line. (a) Significant differences in sensitivity can be observed between the single averaged and the six times averaged reference. (b) All CS reconstructions from the six times averaged dataset resulted in successful images which exhibit no visible information loss compared to the reference.*

exhibit sufficient SNR as described in the first paragraph.

Figure 25b displays the reconstruction results. For the investigated range of acceleration factors, all retrospectively undersampled six times averaged datasets could be recovered by CS without visible information loss. The apparent SNR level is similar to the six times averaged reference in Figure 25a. However, as could be expected, the deviations between the CS reconstruction and the fully sampled reference increase with higher acceleration.

If the same reconstructions are shifted one spectral point from the VS580H marker resonance frequency, SNR differences in the fully sampled images become visible (Figure 26a). The six times averaged dataset exhibits significantly higher SNR than the single averaged reference. Thus, due to the better definition of the signal bearing tubes in the six times averaged data, the CS reconstructions at this spectral point were again successful and show all six tubes (Figure 26b). However, the SNR in the reference data is lower at this spectral point than

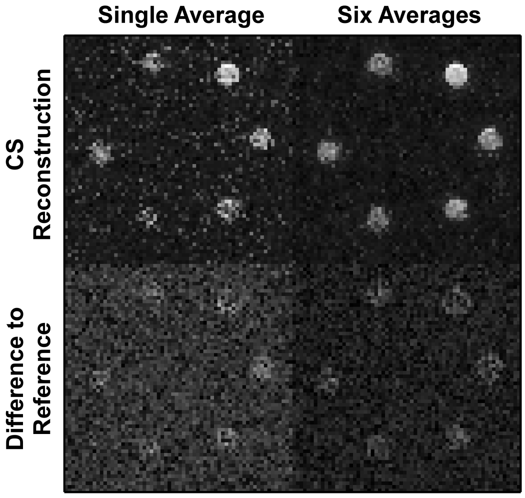

Figure 27: *Comparison of retrospectively undersampled (*af = *6.6) and CS reconstructed images from single averaged and six times averaged data. It can be clearly seen that averaging undersampled data significantly improves the CS reconstruction quality. Moreover, the undersampled and six times averaged dataset would have been acquired in approximately the same measurement time as a fully sampled single averaged reference.*

at the main resonance frequency, leading to CS reconstructions with lower apparent SNR.

Furthermore, a comparison was drawn between a CS reconstruction of the single averaged and the six times averaged dataset. Again, the chosen spectral point was shifted from the main resonance frequency by one spectral point (compare Figure 24). Both datasets were undersampled by a factor of 6.6; thus, the six times averaged and factor 6.6 accelerated dataset would have been acquired in approximately the same time as a fully sampled single averaged dataset. Thus, either the measurement time can be reduced by a factor of 6.6 (while acquiring every phase encoding step only once) or by spending the same amount of time reducing the acquired data but averaging every phase encoding step six times. The results are displayed in Figure 27. It can be clearly seen that a measurement time reduction is unreasonable for this setup. The CS reconstructions of the single averaged undersampled dataset is artifact corrupted and suffers from an inherently low SNR. Due to the low SNR, the dataset is no longer sparse and, thus, CS reconstruction fails. However, by averaging the undersampled data, the CS reconstructed tubes are significantly better defined.

In vivo experiments

Figure 28 exhibits the successful application of CS to undersampled *in vivo* data. Accurate reconstructions could be obtained for all investigated acceleration factors (Figure 28a). Again, as already seen in the phantom experiments, the deviations between the fully sampled reference and the CS reconstruction increase with higher acceleration. Nonetheless, the biologically relevant information is preserved and the reconstructed signal can be correlated to the area of the stroke (Figure 28b).

4.5.3 Discussion

This initial proof of principle demonstrated that 2D ^{19}F CSI can benefit from CS. There are two main reasons why ^{19}F CSI datasets are well-suited for CS: First, the low natural

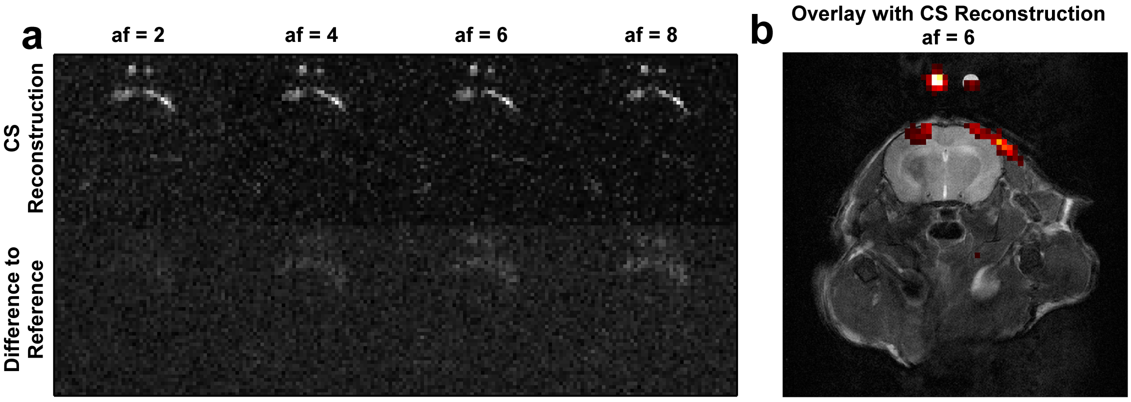

Figure 28: *Results from the 2D* in vivo *experiments. (a) A successful and reproducible reconstruction could be obtained for all investigated acceleration factors. The biologically relevant information is preserved for all CS reconstructions. (b) The biological information successfully recovered by CS can be correlated to the infarcted area underneath the skull of the mouse.*

abundance of ^{19}F in biological tissue leads to a negligible background signal. Thus, signal in the datasets can only be observed in regions where exogenous ^{19}F marker has been introduced. This leads to a very sparse signal distribution in the image space, which is crucial for the success of CS as pointed out in Chapter 2. Second, purely phase-encoded CSI allows application of every desired k-space sampling pattern. Therefore, a pseudo-random sampling pattern with a densely sampled k-space center was chosen. Such a pattern is beneficial for CS since it introduces incoherent undersampling artifacts that can be removed using CS.

This study demonstrates that CS offers the possibility to drastically reduce the number of phase-encoding steps in ^{19}F CSI experiments while preserving the spectral information. The last claim will be further investigated in Section 4.6 in an experiment with two different ^{19}F markers.
The gain in measurement time can be used in multiple ways.

- First, it can be invested into averaging the undersampled k-space data. This emphasizes the ^{19}F signal against the noise level, leading to better CS reconstructions (see Figure 27). To understand why averaging helps to improve the visual image reconstruction quality, consider a dataset with a low SNR where certain structures are below the noise level and cannot be observed. Consider a second dataset with a high SNR that can be obtained by averaging the low SNR dataset a certain number of times until the invisible structures can again be observed. If the high SNR dataset is retrospectively undersampled and CS reconstructed, the reconstruction will exhibit an apparent SNR similar to the fully sampled high SNR dataset. Thus, all structures will be recovered which can be observed in the fully sampled reference. If the low SNR dataset is undersampled in the same manner, the reconstruction will exhibit an apparent SNR which is comparably low as in the fully sampled reference. Thus, the structures that were lost below the noise level remain lost.
  Consider now the same situation for an undersampled dataset (not retrospectively undersampled). For the CS reconstruction process, it does not matter if data are removed from a high SNR dataset or if low SNR undersampled data are averaged several times. Both strategies should lead to similar results on similar datasets. This qualitative argument makes clear that averaging data helps improve the CS reconstruction quality. In the limiting case, the data could be averaged until the

(almost) noise-free case is reached, which would be optimal for CS.

- Second, if sufficient signal intensity is given, the measurement time can be reduced by accelerating the data acquisition process without further averaging or with few averages. This would allow for more experiments in the same measurement time, increasing the amount of information that can be collected per unit time. Therefore, the measurement time can be more efficiently exploited.
- Third, the gain in measurement time might be used to increase the resolution of the CSI dataset. Even though this immediately leads to prolonged measurement times, the ability to undersample this higher resolved data might keep the measurement time in reasonable limits for *in vivo* imaging.

The chosen value for $p = 0.75$ was empirically derived. It proved stable, accurate and reproducible throughout all of the presented reconstructions in this section. However, no detailed investigations regarding norm, noise, and acceleration have been performed. This will be discussed in the next section on 3D ^{19}F CSI (Section 4.6).

In conclusion, the presented proof of principle on the application of CS on 2D ^{19}F CSI datasets resulted in successful reconstructions of retrospectively undersampled datasets. The biological information in the *in vivo* experiments was preserved and allowed correlation of the reconstructed ^{19}F signal intensity to the area of PT-induced stroke. These encouraging results motivated the investigation of the proposed method on 3D ^{19}F CSI imaging which is topic of the following section.

4.6 3D ^{19}F CSI

The 2D ^{19}F experiments presented in the last section encouraged to transfer the idea of accelerating spatially sparse ^{19}F CSI experiments by CS to 3D. In 3D, the measurement time is naturally prolonged due to the additional phase encoding direction. Nonetheless, as pointed out in the motivation to this chapter, 3D imaging is preferable since the exact marker distribution is not *a priori* known. Therefore, CS offers a possibility to reduce the measurement time while preserving the biologically and spectrally relevant information.

The following experiments investigated the quality of CS reconstructed spatially sparse 3D data. First, simulations were performed on a numerical 3D phantom to determine the influence of noise, acceleration factor, and norm on reconstruction quality. Second, a measured *ex vivo* dataset was retrospectively undersampled and reconstructed to verify the simulation results. Two ^{19}F markers with different chemical shifts were used in these *ex vivo* experiments to investigate whether or not CS reconstruction preserves the ability of CSI to distinguish between these markers. Third, optimized parameters were derived from these simulations and experiments and used to reconstruct undersampled *in vivo* photothrombotic (PT) mouse datasets. In contrast to the *ex vivo* experiments, only one marker was applied. It is important to note that, for each sampled k-space point in the *ex* and *in vivo* experiments a complete spectrum was acquired. Therefore, no CS reconstruction exploiting spectral sparsity was performed.

4.6.1 Methods

Phantom simulations

In Section 4.5, no detailed investigation on the interdependencies of norm, noise, and acceleration factor was performed. This is caught up now in this section. In order to study the behavior of the utilized CS technique, simulations were performed to investigate these interdependencies of the algorithm. These simulations aimed at optimizing the reconstruction parameters for later experimental realization. Therefore, a numerical mouse phantom was created using MATLAB®. The phantom dimensions were 90x90x180 spatial points with a spatial sparsity of 6 % (i.e. 6 % of all pixels were non-zero). Spectral information was not simulated since different spectral points on one particular resonance line differ primarily in SNR. Thus, the spectral points on the slopes of the resonance line offer lower SNR than the peak. Therefore, simulating different noise levels is equivalent to examining different spectral points.
The phantom provided signal from various simulated structures known to accumulate ^{19}F markers such as the liver, spleen, and bone marrow [28]. Additionally, signal reflecting stroke structures in the brain were integrated to match *ex* and *in vivo* PT mouse models. The maximum signal intensity was normalized to 1. To mimic the biological situation, the signal level variations of the different structures throughout the phantom covered a range of several magnitudes. For example, the simplified liver exhibited approximately 20 times more signal than the stroke structures in the brain. This was in accordance with the *ex vivo* measurements in this study. Pseudo-random Gaussian white noise was separately added to the real and imaginary parts of the dataset. The standard deviation of the Gaussian noise distribution was set to $\sigma = (0; 0.0001; 0.0005; 0.001; 0.0025; 0.005; 0.01)$. Furthermore, the value of the norm was varied by $p = (0.25, 0.50, 0.75, 0.90, 1.00)$. The norm $p = 1.00$ is the so-called “limiting convex case“, which is often used in CS reconstructions [16, 29]. It has been shown that solving the minimization problem with $p = 1.00$ leads, in most cases, to the sparsest solution [31]. Therefore, $p > 1$ was not further investigated. The acceleration factor was the third parameter studied. The following undersampling values were investigated: 2, 4, 6, 8, 10, 12. The sampling pattern consisted of a densely sampled k-space center (7x7x11 cuboid accounting for 0.04 % of the total k-space) and a randomly undersampled outer k-space.
The reconstruction quality of the results was quantified using the RMSE metric as defined in Section 3.1.

Ex vivo experiments

The *ex vivo* experiments were performed to verify the simulation results with a real ^{19}F signal distribution. Two ^{19}F markers with different chemical shifts were used to examine whether or not CS reconstruction preserves the ability of CSI to distinguish between these markers.

Animal preparation

Ex vivo animal experiments were performed in accordance with institutional guidelines and approved by Bavarian state authorities. As previously described [62], focal cerebral ischemia was induced in one adult C57/BL6-mouse by PT of cortical microvessels under inhalation anesthesia with enflurane in a 2:1 nitrogen/oxygen atmosphere. A cold light source was stereotactically centered on the intact skull 2 mm posterior and 2.4 mm right from Bregma. After intraperitoneal administration of sterile-filtered Rose Bengal solution (0.2 ml), the brain was illuminated for 20 min. Immediately after illumination, 250 μl of an emulsion containing 30 % v/v perfluoro-15-crown-5-ether (VS580H, Celsense, Inc., Pittsburgh, PA, USA) was applied intravenously. The skin was afterwards sutured and the mouse was allowed to recover. The same procedure was repeated on the left hemisphere eight days later, leading to a second cortical infarction. Immediately following illumination, 250 μl of a second emulsion containing 30 % v/v perfluoro-polyethelene-oxide (VS1000H, Celsense, Inc., Pittsburgh, PA, USA) was applied intravenously.

14 days after induction of the first PT, the mouse was sacrificed by CO_2 narcosis and fixed in 4 % paraformaldehyde. The fixed animal served as model for the *ex vivo* experiments.

Regarding the emulsions, the PFC core compound of the VS580H emulsion is single resonant and the PFC core compound of the VS1000H emulsion can also be considered single resonant. Due to the chemical shift, the separation of both PFC resonances is approximately 0.8 ppm, which corresponds to approximately 230 Hz at 7 T.

Magnetic Resonance Imaging

The MR measurements were performed on a 7 T Bruker Biospec system (Bruker BioSpin GmbH, Rheinstetten, Germany) using a home-built surface coil with an inner diameter of 20 mm. The coil was adjustable to both the ^{1}H resonance frequency of 300.3 MHz and the ^{19}F resonance frequency of 282.4 MHz.

For use as an anatomical reference, an axial 3D ^{1}H turbo spin echo (TSE) scan of the animal's head was performed ($T_{E,eff}/T_R$: 53.6 ms/1000 ms; inter-echo time: 6.7 ms; turbo factor: 16; FOV: 30x40x30 mm^3; matrix: 150x200x60; NA: 1). After ^{1}H imaging, the coil was set to the ^{19}F frequency. The same geometry was used for both the ^{1}H image and the fully phase-encoded 3D ^{19}F steady-state free precession CSI (SSFP-CSI) [63] experiments (pulse shape: hermite; pulse bandwidth: 5400 Hz; T_{ACQ}/T_R: 10.7 ms/14 ms; spectral points: 128; FOV: 30x40x30 mm^3; matrix: 74x100x60).

The frequency of the non-selective pulse was centered between the resonance frequencies of both administered PFC emulsions. The resonance frequencies were derived from a non-spatially resolved spectrum.

Five fully sampled ^{19}F datasets were obtained and each experiment lasted 1 h 45 min. Thus, the total measurement time was 8 h 45 min. To improve the point spread function (PSF), each dataset was retrospectively weighted based on a modified Bartlett-Hanning window provided in MATLAB®. A five times averaged dataset was created and retrospectively undersampled with an acceleration factor of 8. Additionally, one of the fully sampled datasets was undersampled by an acceleration factor of $\frac{8}{5}$ and reconstructed using the same parameters previously described. Moreover, a zerofilled low resolved dataset from $\frac{5}{8}$ of the total k-space was generated. This allowed comparison of opposite sampling strategies with the same reduced measurement time. In the same context, a single averaged dataset was CS reconstructed from $af = 8$ ($p = 0.75$) and compared to a zerofilled low resolved dataset from $\frac{1}{8}$ of the total k-space data. An overview of the different investigated combinations of averaging and undersampling is given in Table 6.

For all CS-reconstructed *ex vivo* experiments, the densely sampled inner k-space was a cuboid consisting of 7x7x11 k-space points. These points accounted for approximately 0.12% of the total k-space. For these datasets, 30 spectral points covering the ^{19}F signal peaks were CS reconstructed for different values of p (0.25; 0.50; 0.75; 0.90; 1.00).

Reconstruction method	CS	CS	Zerofilled	CS	Zerofilled
Acceleration factor *af*	8	$\frac{8}{5}$	$\frac{8}{5}$	8	8
Number of averages NA	5	1	1	1	1
Possible measurement time (min)	65	65	65	13	13
Figure	31,33,34	31	31	32	32

Table 6: *Overview of the different investigated combinations of averaging and undersampling as utilized in the retrospectively undersampled* in vivo *experiments.*

In vivo experiments

Undersampled *in vivo* mouse datasets were acquired and reconstructed to demonstrate the *in vivo* applicability of the proposed method. For the reconstruction, optimized parameters were derived from the previously described simulations and *ex vivo* experiments.

Animal preparation

Animal experimentation was performed in accordance with institutional guidelines and approved by Bavarian state authorities. *In vivo* scanning was performed on one mouse. As described in the previous section, focal cerebral ischemia by PT was induced only on the right hemisphere. Immediately after illumination, 500 μl of a customized 10 % v/v perfluoro-15-crown-5-ether emulsion was intravenously applied. The preparation of this emulsion is described in detail by Floegel et al. [46]. For *in vivo* scanning, the mouse was anesthetized with 1.5 % isoflurane in a 2 L/min oxygen atmosphere.

Magnetic Resonance Imaging

MR measurements were performed on a 7 T Bruker Biospec system using a home-built, double resonant birdcage coil with an inner diameter of 40 mm [64].
For use as an anatomical reference, an axial 3D ^{1}H turbo spin echo (TSE) whole body scan was performed ($T_{E,eff}/T_R$: 44.8 ms/1000 ms; inter-echo time: 5.6 ms; turbo factor: 16; FOV: 30x30x70 mm^3; matrix: 192x192x70; NA: 1). After ^{1}H imaging, the coil was set to the ^{19}F frequency and multiple 3D ^{19}F SSFP-CSI experiments were performed. The same geometry was used for both the ^{1}H image and the 3D ^{19}F SSFP-CSI experiments (pulse shape: hermite; pulse bandwidth: 5400 Hz; T_{ACQ}/T_R: 10.1 ms/13.6 ms; spectral points: 64; matrix: 48x48x70; NA: 1). A fully sampled dataset was acquired in a total measurement time of 37 min. Additionally, eight identically undersampled datasets were acquired in only 4 min 50 s each. All scans had an approximate acceleration factor of 8. To improve the PSF, all datasets were retrospectively weighted based on a modified Bartlett-Hanning window provided in MATLAB®. The eight identical datasets allowed reconstruction of undersampled datasets with 1-8 k-space averages. The densely sampled central k-space was a 16x16x22 cuboid, which accounted for 3.50 % of the total k-space. The *in vivo* mouse datasets were reconstructed at 15 spectral points covering the ^{19}F signal peak. To show the influence of the SNR, the eight times accelerated datasets were reconstructed for $p = 0.75$ and multiple averaging factors (1, 2, 4, 8).

4.6.2 Results

Phantom simulations

The effect of different noise levels and acceleration factors on the reconstruction quality is shown in Figure 29. To visualize the influence of the acceleration factor, the RMSE was plotted against the reconstruction norm for four noise levels in the subplots of Figure 29a. The diagrams are almost identical, indicating that, for the investigated range of parameters, the acceleration factor had only a minor influence on the reconstruction quality. This indication is also seen in Figure 29b.

The effect of different noise levels on the reconstruction quality is the focus of the subplots in Figure 29b. Similar to Figure 29a, the RMSE is displayed as dependent on the reconstruction norm. In each subplot, the effect is visualized for three acceleration factor values. As seen, increasing the noise level led to a shift of the minimum RMSE from lower ($p = 0.25$) to higher p-values ($p = 1.00$). This tendency can also be observed in Figure 29a by comparing the different noise levels of each subplot.

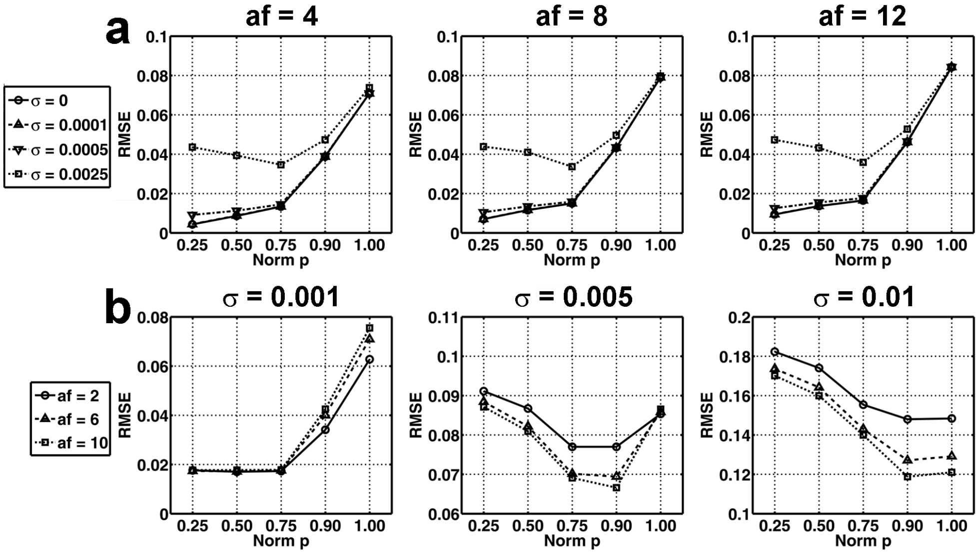

Figure 29: *Results of the simulations. (a) The RMSE for multiple noise levels in dependence on the norm* p *is shown for different acceleration factors (*af $= 4,8$ *and 12). Please note that the minimum RMSE value shifts from* p $= 0.25$ *to* p $= 0.75$ *for higher noise levels. Furthermore, the curves for* $\sigma = 0.0001$ *are positioned almost exactly on top of the noise-free curve (*$\sigma = 0$*). (b) The RMSE for multiple acceleration factors in dependence on the norm* p *is shown for different noise levels (*$\sigma = 0.001$*, 0.005 and 0.01). Please note that the y-axis is scaled differently for the plots and that the displayed noise levels are up to a factor of 100 higher than in (a). In addition, the qualitative appearance of the curves changes with increasing noise level. The minimum RMSE value shifts from* p $= 0.25$ *to* p $= 1$*. To prevent the presentation of redundant information, different combinations of noise levels and acceleration factors are displayed in (a) and (b).*

Figure 30 shows that a nonconvex norm led to a faster convergence of the minimization algorithm and a lower remaining background intensity. This becomes obvious when comparing the convex ($p = 1.00$) with the nonconvex ($p = 0.25$ and 0.75) reconstructions. For the nonconvex norms, more details were recovered for the given number of iterations than for the convex norm. Furthermore, the background intensity was significantly lower. When moderate noise was added (Figure 30, $\sigma = 0.001$), the results were almost identical to the noise-free case.

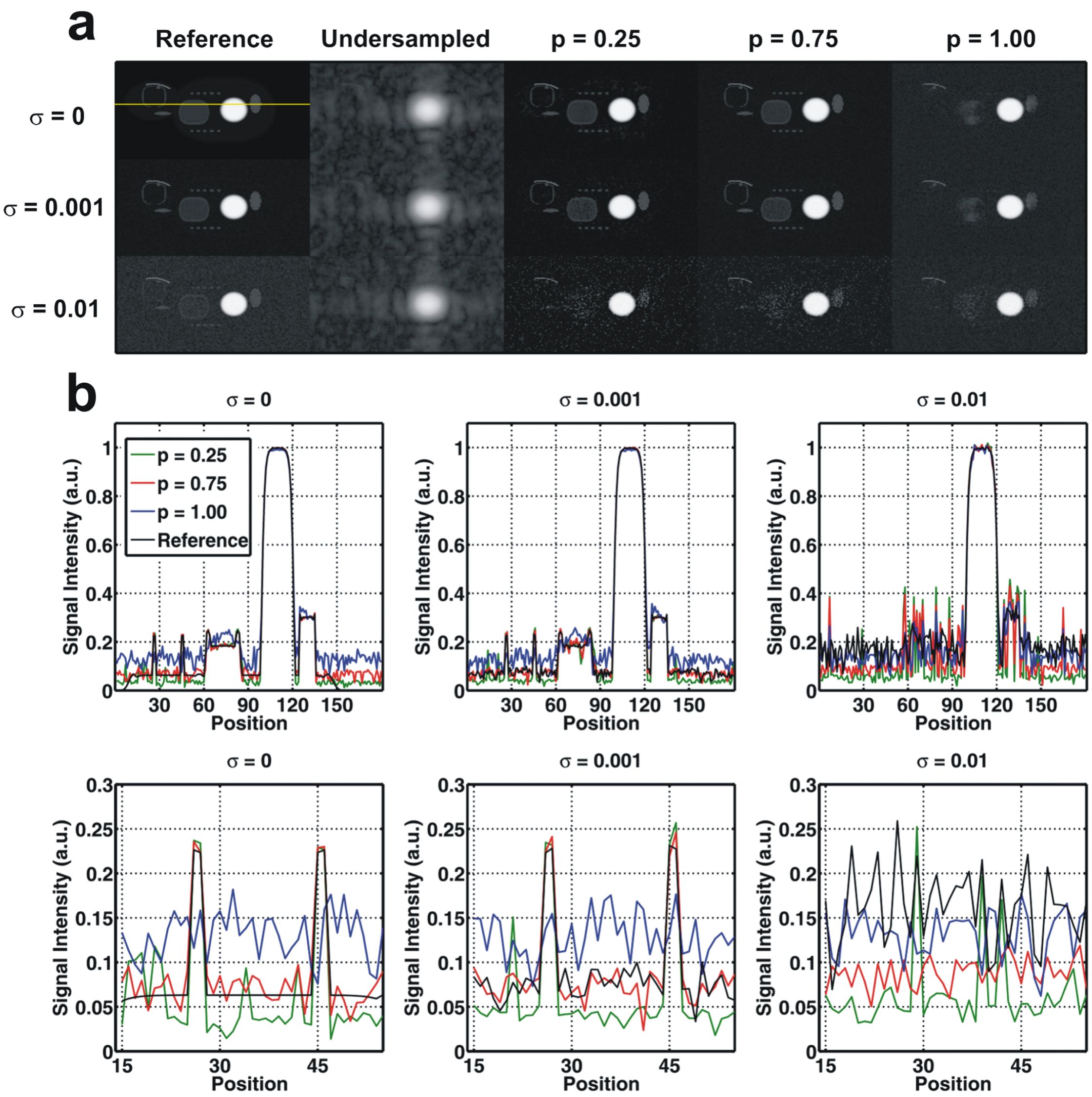

Figure 30: *(a) A representative slice of the numerical mouse phantom dataset is shown with results for an exemplary acceleration factor of 10. The influence of the chosen norm* p *on the visual reconstruction quality for multiple noise levels is displayed. As seen, the limiting convex case* p *= 1 fails to recover some of the fine structures. Please note that the increasing background intensity in the CS reconstructions depends on the norm and the noise level. The yellow line in the* $\sigma = 0$ *reference image indicates the position of the plotted profile shown in (b). For better visibility, the images are scaled by power 0.4. (b) A single 1D profile from the numerical mouse phantom dataset is shown. It can be seen that the smaller the norm* p *is, the lower the remaining overall background intensity. Spike artifacts, however, occur with increasing noise. These can be clearly seen in the lower row for* $\sigma = 0.001$ *and* p *= 0.25 (e.g., position 21) and for* $\sigma = 0.01$ *and* p *= 0.25 (e.g., position 29). Please note that the lower row is a magnified view of positions 15-55 of the upper row. The SNR of the displayed structure in the lower row is infinity (*$\sigma = 0$*), 24 (*$\sigma = 0.001$*) and 2.4 (*$\sigma = 0.01$*). The plots are scaled by power 0.4.*

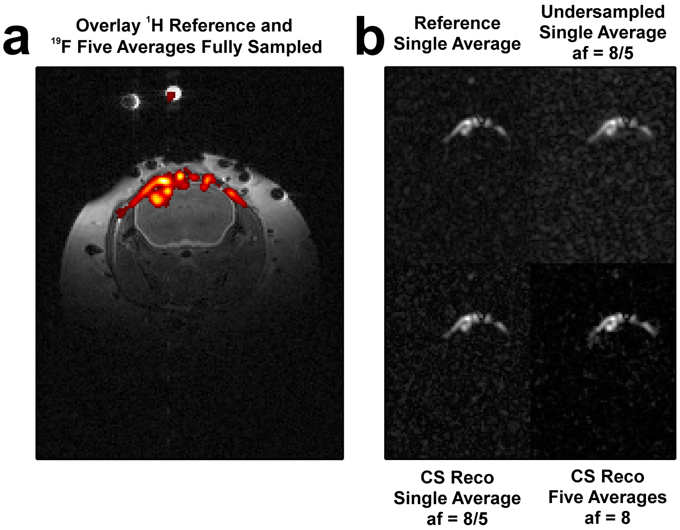

Figure 31: *(a) Anatomical reference overlaid with a fully sampled ^{19}F signal of a five times averaged dataset. ^{19}F signal can be observed in the ring-like lesion area and supracranially at the site of the skin incision. (b) Comparison of different undersampling schemes leading to the same measurement time of 65 min. Both CS reconstructions were obtained using* $\mathrm{p} = 0.75$*. Please note the reduced blurring and artifact level in both CS reconstructions compared to the undersampled image.*

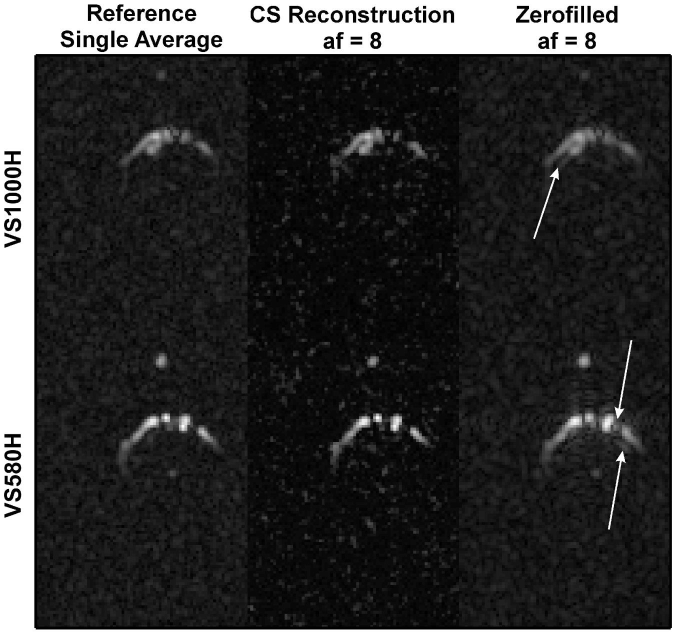

Figure 32: *Comparison of a low resolution dataset ($\frac{1}{8}$ of the whole data and zerofilled to full matrix size) and a CS reconstruction (*$\mathrm{p} = 0.75$*;* $\mathrm{af} = 8$*). The reference dataset was not averaged. All images show the same exemplary slice. In the upper row, the quality of the zerofilled dataset is comparable to the reference and the CS reconstruction. However, in addition to blurring, a ringing artifact is observed in slice direction (perpendicular to image plane, arrow). In the bottom row, blurring and ringing artifacts are observed in the zerofilled dataset. Again, the arrows point to ringing artifacts in slice direction. The CS image, however, offers an accurate reconstruction exhibiting typical spike artifacts.*

It is important to remember the influence of the noise level on CS reconstructions. Low signal components can no longer be recovered when the overall noise level is high (e.g., $\sigma = 0.01$). Additionally, spike artifacts occur dependent on the norm and the noise level (Figure 30a). A smaller norm will lead to more pronounced spike artifacts. This can, especially in the case of high noise levels, significantly corrupt the reconstruction quality.
In Figure 30b, a close-up of a 1D profile of the numerical mouse phantom is displayed. For noise-free data, only the two nonconvex reconstructions ($p = 0.25$ and 0.75) were able to adequately recover the profile. Furthermore, the nonconvex norms provided a lower overall background intensity for the same number of iterations than the limiting convex $p = 1.00$. This held for noise-levels varying over several orders of magnitude as seen in the two profiles with $\sigma = 0.001$ and 0.01. The magnified details (Figure 30b, lower row) of positions 15-55 in the three corresponding profile plots (Figure 30b, upper row) show that the displayed low intensity structures could not be recovered from noise-free and low-noise datasets using $p = 1.00$. For high noise levels (e.g., $\sigma = 0.01$), these structures were below the reference noise level.

Retrospectively undersampled *ex vivo* MRI measurements

Figure 31a displays the ^{1}H reference scan and the fully sampled five times averaged ^{19}F reference overlay, allowing anatomical correlation. Thus, ^{19}F signal can be observed in the lesion area and supracranially at the site of the skin incision. Figure 31b compares a zerofilled dataset from the inner $\frac{5}{8}$ of the single averaged k-space to CS reconstructions of single averaged data with $af = \frac{8}{5}$ and five times averaged data with $af = 8$. All images provide the relevant biological information. No significant difference between the single averaged zerofilled data and the fully sampled single averaged reference can be seen in the figures.
Differences, however, can be observed between the CS reconstructions as a result of averaging and different acceleration factors. These differences become apparent when observing the reduced background intensity and stronger contrast of the five times averaged and CS reconstructed dataset. Spike artifacts, however, appear more pronounced in the five times averaged and CS reconstructed dataset. Nevertheless, the CS reconstruction of the single averaged dataset with $af = \frac{8}{5}$ shows no significant differences from the reference and zerofilled images in Figure 31b.
In order to demonstrate that the proposed method allowed effective acceleration of a factor of eight, a comparison is shown in Figure 32 of a zerofilled low resolved dataset from $\frac{1}{8}$ of the k-space and a CS reconstruction ($p = 0.75$, $af = 8$). All data were obtained from a single averaged dataset. In all images, the same exemplary slice is displayed at different spectral points. In the upper row, the slice is shown at the resonance peak of VS1000H. The zerofilled dataset offers a similar image quality as the reference dataset and the CS reconstruction. However, the CS reconstruction exhibits a suppressed background. In the zerofilled image, besides blurring due to the reduced resolution, a ringing artifact in the slice direction (perpendicular to displayed image plane, arrow) is observed. At the VS580H peak, more prominent blurring and ringing artifacts can be observed in the zerofilled image. Furthermore, similar to the VS1000H zerofilled image, a ringing artifact in the slice direction (arrow) is observed. The CS reconstruction at the VS580H resonance peak shows a highly accurate reconstruction without blurring or ringing artifacts. Spike artifacts, however, can be observed.

To illustrate the influence of the norm, CS reconstructions for different p-values of the five

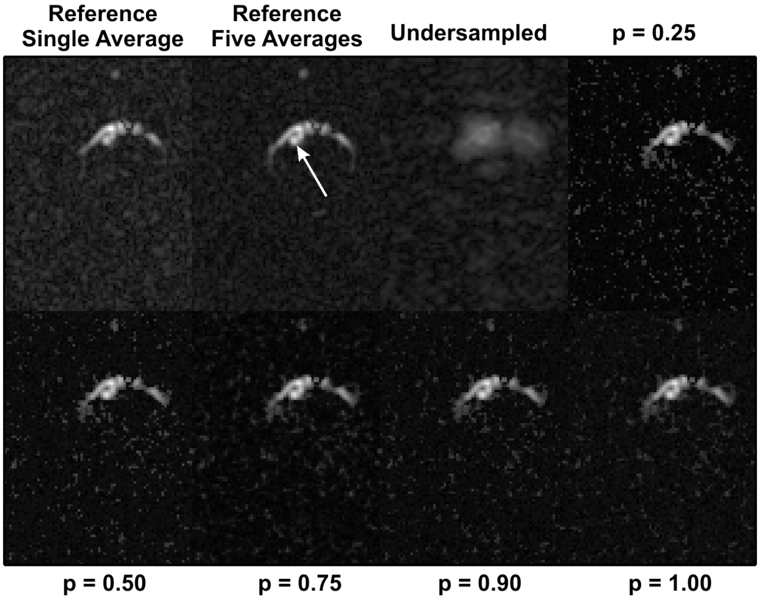

Figure 33: *Results of the retrospectively undersampled* ex vivo *mouse experiments (*af = *8). A representative slice from the brain is shown. Reconstruction results for the different values of the norm* p *are compared to a single average and a five times averaged reference. All norms recovered the relevant information. The arrow marks the area where the VS1000H signal can be correlated with the infarcted area. Additional ^{19}F signal can be observed supra cranially at the location of skin incision.*

times averaged and eight times undersampled dataset are displayed in Figure 33. This figure clearly shows that all CS reconstructed datasets provided superior image quality compared to the randomly undersampled dataset. As already seen in the simulations, all norms showed a similar influence on reconstruction quality. CS reconstructions with low p-values tended to exhibit more prominent spike artifacts while the quality of the reconstructed ^{19}F signals remained similar for all norms. All CS reconstructed images, however, provided at least the same information content as the single averaged, fully sampled dataset. Furthermore, the signal components in the CS reconstructions were better defined. Additionally, Figure 33 demonstrates that the proposed method is able to reproduce ^{19}F signals correlated to a photothrombosis stroke from the accelerated data. Thus, in this case no information relevant to biological/clinical research would be lost.

Figure 34 demonstrates that different ^{19}F markers could be distinguished in the CS reconstruction. Figure 34a compares the five times averaged reference to the eight times undersampled and CS reconstructed datasets at two different spectral points. These points correspond to the peak intensities of VS1000H and VS580H. In Figure 34a and b, no significant deviations between the reference images and the reconstructions can be observed. Spectra of representative voxels from both stroke structures (Figure 34a, arrows) are shown for the reference and the CS reconstructions. In addition to the peak intensities, the resonance line shapes of both markers were also sufficiently recovered. As previously mentioned, the spatial data for each spectral point were independently reconstructed.

Undersampled *in vivo* MRI measurements

In Figure 35a, the fully sampled, single averaged reference is compared to the reconstructions of the undersampled datasets with different averaging factors. A representative slice in the

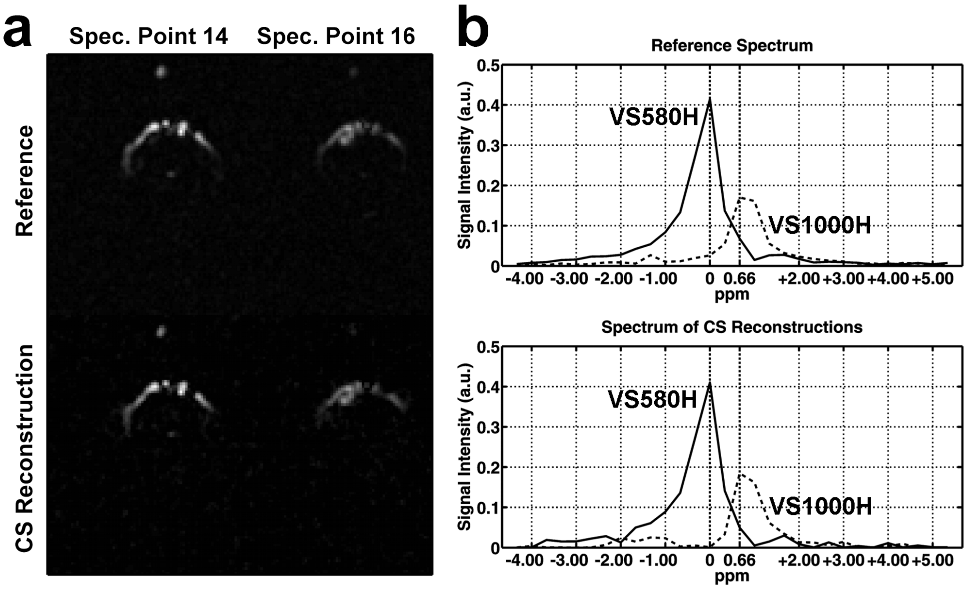

Figure 34: *Influence of the CS reconstruction on the spectral information. (a) Images of the peak intensities of the resonance lines of VS580H (spectral line 14) and VS1000H (spectral line 16) are shown. The same slice from the 3D dataset for both resonance lines is shown. The solid arrows point to voxels in a stroke labeled with VS580H and the dashed arrows point to voxels in a stroke labeled with VS1000H. (b) Comparison of the reference spectrum with the spectrum of the CS reconstructions. The spectra (solid or dashed) were obtained from areas marked by the corresponding (solid or dashed) arrows in (a). As can be seen, the essential spectral information is the same for the reference and the CS reconstructions. Thus, the relevant biological information was preserved in the CS reconstructions. Please keep in mind that each spectral point was independently reconstructed.*

mouse brain is shown, exhibiting ^{19}F signal correlated to the area of the stroke (Figure 35b). It can be seen that the spike artifacts in the CS reconstructed images became less pronounced with increased averaging. Importantly, the stroke-related structure could be clearly extracted for all averaging factors. Thus, it was possible to obtain the relevant biological information in less than 5 min ($af = 8$). Furthermore, the blurring of the undersampled dataset was completely eliminated, preserving the spatial resolution of the fully sampled data. The fully sampled, single average dataset, however, lasted 37 min.

4.6.3 Discussion

This chapter demonstrated the ability of CS to accurately and reproducibly reconstruct spatially sparse undersampled 3D ^{19}F CSI datasets. There were four main findings. First, CS can be used to shorten the measurement time in the presence of sufficient signal intensity. Second, the gain in measurement time due to undersampling provides more flexibility in ^{19}F MR experiments. Third, the optimal norm for CS reconstruction depends on the noise level. A good compromise, however, is $p = 0.75$ for all investigated noise levels. Fourth, CS does not affect the ability of CSI to distinguish between different markers. Thus, the proposed method can be a valuable tool for multiple biological and medical applications.

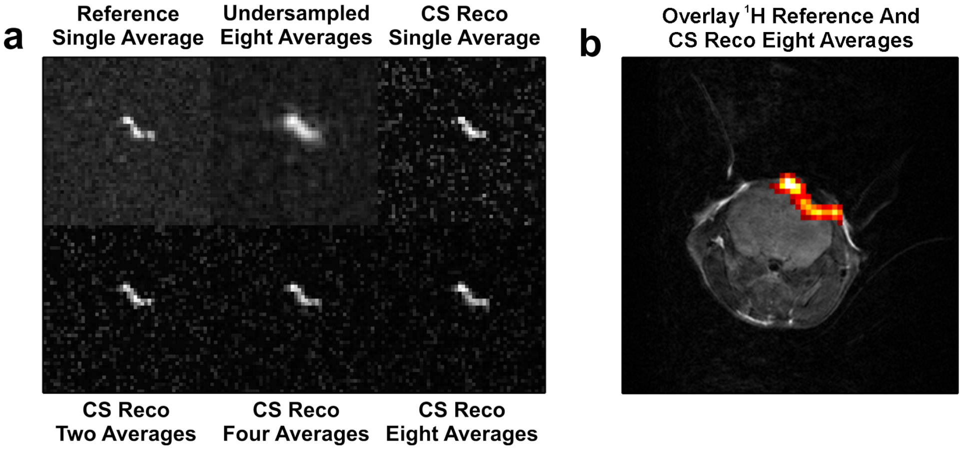

Figure 35: *Results of the undersampled* in vivo *experiments. The acceleration factor was 8. (a) A representative slice of the data is shown. The benefit of averaging on the visible CS reconstruction quality is clear. Furthermore, the relevant biological information could be recovered for all averaging factors. Thus, this information was acquired in reduced measurement time (approx. 5 min instead of approx. 37 min) by up to a factor of 8. (b) Overlay of an* in vivo *^1H mouse reference dataset and the CS reconstruction of an 8 times undersampled, 8 times averaged dataset. The peak SNR of the shown biological structure in the single average reference scan was approximately 10.*

Flexibility due to undersampling

The ability to undersample spatially sparse spectroscopic ^{19}F datasets offers new flexibility for experiments. One possibility is a reduction in measurement time given sufficient signal intensity. Thus, in a fraction of the measurement time needed for a fully sampled dataset, enough k-space data can be acquired to reconstruct a full image. If the SNR in a dataset is low, additional averaging of the undersampled data can be applied to improve the CS reconstruction quality (Figures 33 and 35). However, in case of low SNR, it could be more advantageous to collect additional data points instead of averaging. Furthermore, CS reconstruction of undersampled data allows utilizing time-consuming strategies that would otherwise exceed acceptable *in vivo* measurement times. Examples include using a higher spectral resolution, acquisition-weighted imaging requiring averaging of k-space data [65], or increased temporal resolution for dynamic imaging. A detailed investigation of sampling strategies will be a topic of future studies.

Influence of the norm p

A paper by Chartrand [32] shows that a nonconvex norm could be more beneficial in the noise-free case than using the convex norm $p = 1.00$. The present study thus required a detailed investigation of the norm influence. As previously mentioned, a nonconvex norm $p < 1$ leads to exact reconstructions from even fewer sampled k-space points than the convex norm $p = 1.00$ [32]. Since, however, noise is always present in physical measurements such as MR, this theory was investigated for noisy data (Figures 29 and 30).

The noise-free case is well approximated for low noise levels. Additionally, the profiles of the numerical mouse phantom show that each norm accurately recovered the components with significantly higher intensity than the noise level (upper row, Figure 30b). However, the fine structure magnified in the lower row of Figure 30b shows that for the given number of iterations, the low signal details were not recovered from the convex norm $p = 1.00$. This confirmed the faster convergence of a nonconvex CS algorithm for data with moderate to no noise. It also explains why using a nonconvex instead of a convex norm is beneficial. Spike artifacts, however, appear in the presence of noise due to the strict data consistency constraint of the applied algorithm. These artifacts are more pronounced the smaller the norm. This is because a decreasing norm also decreases the remaining background intensity of the nonconvex reconstructions. Thus, the spikes are much more pronounced for lower p-values and easier to recognize (Figure 30b, lower row, $\sigma = 0.001$ and 0.01). This effect degrades the visual reconstruction quality for the $p < 1$ norms, making the $p = 1.00$ images appear less corrupted (Figure 30a, $\sigma = 0.01$). In most cases, since spikes rarely emerge in clusters and thus resemble "salt-and-pepper" noise, they can be clearly distinguished from the relevant signal components.

In summary, for low to moderate noise levels, a nonconvex norm can be beneficial in terms of accelerated CS algorithm convergence. In the case of clearly visible spike artifacts, however, it must be decided in each case whether or not these artifacts are tolerable. This is because they can corrupt the reconstructed data. However, a preliminary study has been recently published offering a possibility to reduce spike artifacts [66]. For the *in vivo* experiments in this study, the nonconvex norm $p = 0.75$ was chosen as a good compromise. Further explanation for this choice is given in the noise section.

Influence of the acceleration factor

The simulations, reflected by the RMSE, suggest that the acceleration factor is of minor importance (Figure 29). The RMSE, however, can be a misleading metric to measure reconstruction quality in cases of high noise levels. This is discussed in detail in the following section. Since the exact sparsity of the underlying data is unknown in most experiments, the use of relatively low acceleration factors up to 8 is recommended. Based on these findings, the *in vivo* datasets were undersampled by a factor of 8.

Influence of noise and effects on the RMSE metric

As already implied, noise is the dominating factor influencing reconstruction quality. The underlying assumption of sparsity, which is crucial for CS, can be properly defined only when used with noise-free data. However, as soon as an infinitesimally low noise level is added, sparsity can only be given as an approximate. Thus, CS reconstruction fails when applied on data with a high noise level. As shown in recent mathematical [67] and MR-related papers [16, 52–54, 60], however, CS reconstructions from noisy data are possible. While the influence of noise on CS reconstruction quality has already been investigated [68], the present study examines in detail the influence of noise on CS reconstruction.

The results demonstrated that by increasing the noise level, the minimum RMSE value shifts from nonconvex norms to the convex norm $p = 1.00$ (Figure 29b). This implies that the norm providing the best reconstruction quality is dependent on the noise level. A good compromise is $p = 0.75$ since it performs well for all investigated noise levels. Using $p = 0.75$, sufficiently accurate CS reconstructions of undersampled *in vivo* ^{19}F CSI datasets were obtained with an SNR as low as 10 in the fully sampled reference (Figure 35). Besides the investigated animal model, other ^{19}F applications, such as inflammation models [69, 70], provide sufficient SNR for CS reconstruction. In these works, SNR values larger than 20 were achieved, which are, based on the results of the current study, sufficient for proper CS reconstruction.

Since a noise-free reference was available, the simulations were quantified using the RMSE.

The RMSE metric is widely used and merges deviations between reconstructions and a noise-free reference image into a single value. However, even when the deviation between a completely sampled noisy dataset and a noise-free reference are quantified using this metric, the RMSE never equals zero due to noise. This becomes especially problematic in the presence of high noise levels. For noise levels $\sigma = 0.005$ and 0.01 (Figure 29b), the RMSE increases with decreasing acceleration factors. This is a misleading tendency that is a result of the increasing suppression of background intensity with higher acceleration factors (Figure 36a). If the noise level is low, this effect is less pronounced (Figure 36b). Furthermore, in contrast to human visual perception, no structural information is taken into account using the RMSE metric. Two images with the same RMSE value may have different types of image errors (e.g., artifacts or noise). Since the RMSE might to lead to false conclusions in cases of high noise, an additional visual inspection of the reconstructed images must be performed. Which image metric is best suited to quantify CS reconstruction quality should be further investigated.

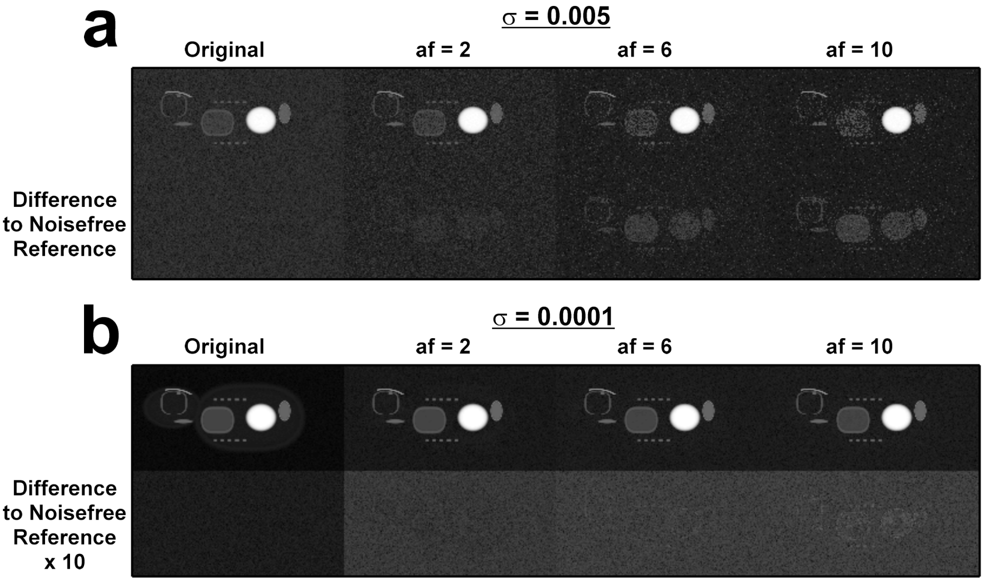

Figure 36: *Illustration of the RMSE metric problematic. (a) In the presence of high noise, CS reconstructions (*$p = 0.75$*) for different acceleration factors are displayed. Although increasingly more non-recovered structures occur in the difference images, the overall difference intensities decrease the higher the acceleration factor. This is due to the increasing suppression of remaining background intensity. Thus, lower RMSE values are obtained for higher acceleration factors (Figure 29). (b) CS reconstructions (*$p = 0.75$*) for different acceleration factors are displayed in the presence of low noise. The difference intensities are comparable regardless of the acceleration factor. However, the higher the acceleration factor, the more non-recovered signal components are revealed in the difference images. Thus, higher RMSE values are obtained for higher acceleration factors (Figure 29).*

Influence of averaging

As mentioned above, CS is an exact framework only for noise-free data. This implies that the SNR in the undersampled data should be as high as possible in order to approximate the noise-free case as close as possible. In the phantom simulations, the effect of higher SNR can be observed by comparing the quality of reconstructions obtained for different noise levels but the same acceleration factor and norm (Figures 29 and 30). For the *ex* and *in vivo* data, the SNR was improved by averaging the acquired k-space points. Although it was possible to successfully reconstruct the single averaged *ex* and *in vivo* datasets from $af = 8$, the effect of different SNR levels on the reconstruction quality was also studied. The better-defined signal components in Figures 31 and 33 resulted from the improved SNR of the acquired k-space points. In the CS reconstruction of the undersampled ($af = 8$) five times averaged dataset (Figure 31), the low background intensity is due to the higher acceleration factor and the five averages. High undersampling leaves more degrees of freedom for the CS reconstruction, which results in increased suppression of background intensity (Figure 36a). Higher SNR allows greater distinction between signal and background, resulting in higher contrast in the CS reconstruction.

In the presence of low SNR (single average), however, low acceleration factors (CS reconstruction with $af = \frac{8}{5}$ in Figure 31) can be advantageous. Since most k-space data are collected using low acceleration factors, few degrees of freedom are left to the CS reconstruction. Therefore, no visible spike artifacts appear. In this particular case, zerofilling a low resolved dataset from $\frac{5}{8}$ of all k-space points (Figure 31, zerofilled) resulted in nearly the same image quality as the fully sampled reference. This is because the resolution of the zerofilled dataset in 3D imaging is still approx. 85 % (third root of $\frac{5}{8}$) of the fully sampled reference.

Figure 35 demonstrates the CS reconstruction improvement in the *in vivo* experiments due to averaging. Additionally, these experiments underline the relation of the spike artifacts to the noise level. As the number of averages increases, the SNR in the undersampled data is improved. As can be seen in the phantom simulations, higher SNR in undersampled data leads to significantly lower RMSE values (e.g., Figures 29 and 30). Hence, averaging reduces the intensity of the spikes and improves the visual reconstruction quality. Based on the presented results, no conclusion can be drawn as to which of the presented sampling strategies (zerofilling from $\frac{5}{8}$ of the total k-space, single average; $af = 8$, five averages; $af = \frac{8}{5}$, single average) is best. Furthermore, the utilized CS reconstruction is a nonlinear algorithm and, as described in the next section, it is unknown how to determine the SNR in a CS reconstruction. These issues should be further investigated and were beyond the scope of this study.

Remarks on SNR in CS reconstructions

Since the utilized CS algorithm is nonlinear, the Gaussian noise distribution in the data will not be reproduced in the CS reconstruction. The result is a non-Gaussian noise distribution where the underlying statistics are unknown. In Figure 30, the noise level in the reference is always higher than the remaining background intensity of the CS reconstructions. Furthermore, as seen in the simulations, the appearance of spike artifacts is more pronounced the smaller the norm and the higher the noise level (Figure 30). These results indicate spike artifacts are noise related. Moreover, as previously discussed, the background intensity depends on the norm and on the noise level. Thus, the remaining background intensity and spike artifacts in the CS reconstructions cannot be viewed as an actual noise level even though they appear to be noise-like.

Remarks on the CS algorithm

CS allows reconstruction of undersampled datasets when it is not possible to obtain fully sampled data. CS algorithms often require regularization parameters [16, 29]. However, these parameters can be hard to obtain without prior knowledge and might significantly influence the reconstructed image [33]. The CS algorithm used in this work enforces strict data consistency and optimizes only one minimization term [32, 71]. Therefore, no regularization is necessary in order to reconstruct undersampled datasets.

Remarks on the sampling pattern

The exact distribution of ^{19}F marker substance *in vivo* was unknown *a priori*. This is the case in the majority of applications. Therefore, a pseudo-random pattern was chosen as the k-space sampling strategy. This strategy can be used independent of measurement parameters (e.g., matrix size) as long as the expected data are sparse. The densely sampled k-space center size must also be adapted to the particular matrix size. As demonstrated in this paper, sampling a central k-space region that is approximately 1 % of the total k-space is sufficient. However, if the signal distribution is known *a priori*, the sampling strategy can be adopted to the specific situation. For example, k-space points could be acquired with the highest expected SNR [72, 73].

General comments

Zerofilling is a common strategy to interpolate low resolved datasets to larger matrix sizes. For low acceleration factors, this method can provide acceptable results since 3D resolution decreases only with the third root of the acceleration factor. However, in the case of $af = 8$, the resolution is already reduced by a factor of 2. As seen in Figure 32, blurring and ringing artifacts occur in the zerofilled dataset and hamper the image quality. In such cases, CS is preferable and, for lower afs, CS provides image quality at least comparable to zerofilling. The applied CS algorithm enforces strict data consistency as suggested in [32]. This means that the acquired data repeatedly reinserted in the CS reconstructed k-spaces are assumed to be perfect. This is logical for ideal, noise-free data. However, in the presence of noise, noise-corrupted data are assumed to be optimal. This is a source of artifacts in the reconstructions. These errors manifest themselves as spike artifacts in the final images.

The ability to distinguish between compounds with different spectral signatures is the key feature of CSI. As shown in Figure 34, different PFC markers can be clearly distinguished after CS reconstruction of ^{19}F CSI. Furthermore, spectral points are well reconstructed in cases of high signal. The line shapes differ noticeably for low signal components since the CS reconstruction fails (Figure 34b). This issue is closely related to the quantification of ^{19}F markers from CS reconstructed data. If the line shapes are correctly reconstructed, quantification is possible (Figure 34). However, because CS reconstructed line shapes differ noticeably from the real spectrum for low signal components, quantification in such instances is likely to fail. Nevertheless, quantification with low signal is always challenging.

The free choice of a sampling pattern for the undersampled k-space is one advantage of using fully phase-encoded CSI. In this study, a densely sampled k-space center surrounded by a pseudo-random outer k-space was chosen. As mentioned in the methods section, this pattern is known to be advantageous for CS. However, other trajectories that introduce incoherent artifacts in cases of undersampling can be easily employed in fully phase-encoded CSI. Examples include variable density [29] and (pseudo-)radial [74] methods.

4.7 Conclusion

The presented results demonstrate that ^{19}F CSI can benefit from CS. The proposed method yielded highly accurate and reproducible reconstructions. In simulations, the interdependency of the acceleration factor, the noise level and the reconstruction norm was studied. The simulation results were then confirmed using *ex vivo* experiments, through which the proposed method was tested on a realistic signal distribution. In a final step, undersampled *in vivo* 3D CSI mouse datasets were successfully reconstructed. This chapter showed that CS can lead to significantly reduced measurement times and more flexibility in the data acquisition process. Additional topics such as the quantification of SNR and marker content are of special interest and must be further investigated. The question of how to best exploit the flexibility provided by undersampling should also be addressed in future studies.

In conclusion, CS has the potential to significantly improve and extend the applicability of *in vivo* ^{19}F MR imaging. Additionally, the proposed CS reconstruction method is not limited to the presented ^{19}F nucleus or the utilized spectroscopic imaging sequence.

5 Radial Dynamic Cardiac Imaging

5.1 Motivation

Dynamic cardiac imaging with high spatial and temporal resolution is a demanding task in clinical MRI. Capturing the cardiac motion is of great interest as information about the cardiac anatomy, function, and myocardial perfusion can help to diagnose diseases. Cine imaging of the beating heart is currently an important method to obtain functional and anatomical information. In cine imaging, both ECG triggering and imaging in breath-hold are mandatory to prevent motion artifacts. However, prospective segmented acquisition using ECG triggering is only effective if successive heartbeats are sufficiently similar during the data acquisition process. Thus, patients suffering from cardiac arrhythmia can only be examined at the cost of reduced image quality. Furthermore, acquiring data over multiple R-R intervals leads to an averaged representation of the cardiac cycle. Several methods using retrospective gating techniques have been applied to avoid using ECG gating and imaging in breath-hold. One approach is self-gating, where a navigator signal that is not encoded by any gradient is acquired. This is equivalent to sampling the k-space center, which is the average of the complete image space information of the given dataset. This signal can be separately acquired either before or after the data readout, as with Cartesian imaging [75–77]. Alternatively, it can be simply selected out of Non-Cartesian trajectories such as radial [78, 79], where the central k-space point is acquired during each readout. Retrospective gating is performed by extracting the cardiac and respiratory motion state from the navigator signal. Based on thresholding, data are accepted or rejected, ideally leading to images free of motion artifacts.

Another approach has been presented by Kellman et al. [80], where real-time data with high spatial and low temporal resolution are acquired. These data are first reconstructed using SENSE [7], where auto-calibration data are obtained using TSENSE [81]. From these data, an image-based navigator signal [82] is derived that allows correction for cardiac and respiratory motion. The originally acquired data are then reordered to obtain motion artifact-free cine data with high temporal and spatial resolution.

Similar to prospective gating using ECG triggering, all retrospective gating techniques depend on sufficiently similar cardiac cycles during data acquisition. In patients with cardiac arrhythmia, high quality images can be obtained by rejecting the respective data from arrhythmic cycles. However, acquisition of a large amount of data is necessary to prevent possible undersampling after retrospective gating, leading to prolonged measurement times. Additionally, in patients with arrhythmia, the (potentially) interesting information about the abnormal cardiac motion is rejected. In this context, real-time data acquisition strategies are highly preferable.

In real-time imaging, a compromise between spatial and temporal resolution has to be found. A powerful and widespread method for the reconstruction of undersampled data with high spatial and temporal resolution is k-t BLAST or its multi-coil variant k-t SENSE [13]. k-t SENSE exploits the spatiotemporal correlations the algorithm gains from a training dataset with low spatial and the desired high temporal resolution. After that, an undersampled dataset with high spatial and high temporal resolution is acquired. Given that the dynamic event in the training data sufficiently reflects the motion of the undersampled dataset, the missing data can be reconstructed, leading to a complete dataset with high spatial and high temporal resolution. However, if the monitored events in the training stage are non-periodic, k-t SENSE can lead to results with compromised image quality.

Recently, a number of publications demonstrated real-time imaging with high temporal and spatial resolution using the radial trajectory [83–86]. It is well-known that the radial trajectory introduces relatively benign artifacts in cases of undersampling. In the mentioned works, a nonlinear inverse reconstruction was used to achieve a temporal resolution as high as 20 ms by including temporal filtering [85, 86]. Thereby, it was even possible to visualize turbulent flow patterns introduced by stirring in a cylindrical beaker [86].

In this chapter, a real-time data acquisition scheme in combination with CS is proposed. The radial trajectory was chosen due to its relatively benign undersampling artifacts. It has been repeatedly demonstrated that radial imaging can be employed in connection with CS [68, 71, 74, 87–89]. In contrast to an earlier publication that described accelerated dynamic imaging using CS [89], the sparse dynamic differences between a temporally averaged composite image and the desired timeframes were reconstructed using the Strict DC method (see Section 2.3.3) that was extended to incorporate joint sparsity. The principle of joint sparsity was introduced earlier [90] and has found wide-spread application in the MR community (e.g., [91, 92]). In this work, the joint sparsity of the receiver coils was exploited by adaptively combining the coils [93] prior to CS reconstruction. After the CS step, the data were re-distributed to the individual coils.

The presented method is computationally efficient since all reconstruction steps are performed on a Cartesian grid after gridding the undersampled radial data using GRAPPA Operator Gridding (GROG, [39, 40]). To optimize the parameters of the utilized CS algorithm, simulations on a numerically generated phantom were performed. The results from these simulations were transferred to a retrospectively undersampled radial cine dataset. The outcome of these reconstructions confirmed the findings of the simulations. In a last step, the applicability of the proposed method was tested on an undersampled radial real-time dataset. The proposed joint sparsity CS algorithm was found to be well-suited to the reconstruction of the sparse differences between a composite image and the desired timeframe. Thus, it was possible to obtain real-time images with high spatial and high temporal resolution using this method.

5.2 Methods

5.2.1 Compressed Sensing

CS is a sophisticated technique used to precisely reconstruct undersampled sparse datasets and was described in detail in Chapter 2.

In this chapter, the undersampled data were recovered using an algorithm that solves the reconstruction problem:

$$\min \|\nabla_{DG} x\|_p \tag{5.1}$$

such that strict data consistency holds. The Discrete Gradient operator ∇_{DG} calculates the differences between adjacent pixels as already shown in Equation 2.3. It can be interpreted as a transform from image space to the Discrete Gradient domain. The l_1-norm of Equation 5.1 is usually referred to as the Total Variation of the image [26].

The Total Variation $\nabla_{DG} x$ was calculated according to Equation 2.3. The result was then stacked into a vector allowing calculation of the norm. The $\|.\|_p$-norm ($p \in \mathbb{R}$) is defined according to Equation 2.2. The norm is “nonconvex“ if $p < 1$, otherwise it is referred to as “convex“. Theoretically, with noise-free data, a nonconvex norm $p < 1$ leads with overwhelming probability to exact reconstructions from even fewer sampled k-space points than the convex case $p = 1.00$ [32]. The minimization of Equation 5.1 was achieved using a Steepest Descent algorithm similar to the Strict DC method described in Sections 2.3.2 and 2.3.3. However, as explained below, the Strict DC method was extended to exploit the joint sparsity of the receiver array. The concept of joint sparsity is explained in the following section and the reconstruction algorithm with pseudo-code in Section 5.2.3.

Equation 5.1 models the expected image under the assumption that there exist few “edges“ (i.e., spatially localized sharp transitions between high and low signal intensity) and many areas with the same image intensity. Even though the dynamic differences were also sparse in the image domain, minimizing Equation 5.1 led to significantly more accurate results than utilizing image space sparsity. This is further clarified in Section 5.2.3.

5.2.2 Joint sparsity with common sparse support

Joint sparsity is based upon sparse multi-receiver signals. This immediately shows the relevance to MRI since many MR exams are performed using coil arrays with multiple receivers. Duarte described two joint sparsity models in his conference paper [90]; however, only the model he termed JSM-2 applies to MRI. This model will be explained in more detail.

In MRI, the response of the spins after excitation in the sample/volunteer/patient in the magnet is often received by an array of multiple receiver coils. These coils have sensitivity profiles that depend on the geometry of the individual coils. Nonetheless, all receiver coils in the array detect the same object. The object exhibits a certain sparsity in the combined image from all receiver coils. However, due to the coil sensitivity profiles and the noise, this sparsity might be altered in some of the coils since sparsity depends on the SNR as described in Section 2.1.1 and visualized in Figure 7. If the CS reconstruction is applied to each individual receiver coil, this might lead to erroneous reconstruction results (e.g., smoothing of desired structures that exhibit low SNR in certain coils). Nonetheless, all receivers should exhibit the same sparsity, i.e. they share a common sparse support. In the context of MR, this means that all receiver coils “see“ the same object.

The next section describes how this knowledge can be exploited and used to improve CS reconstructions from multi-receiver measurements.

5.2.3 The CS-CC algorithm

Performing the CS step on an SNR-optimized combined image of the coil array would be desirable. thereby, CS would be sensitive to the more homogeneous SNR in the combined image and no longer to the SNR in the individual receivers. However, to obey strict data consistency (i.e., reinserting acquired k-space points at the respective k-space positions of the individual coil signals), the CS-processed combination image must be re-distributed to the individual receiver coils. The adaptive combination method proposed by Walsh et al. [93] offers a possibility to obtain SNR-optimized combination images. Hence, this method was used to determine weights w_{comb} that can be used to combine the individual coil images. The complex conjugate weights w^*_{comb} can be used to re-distribute the combined image to the individual coils after performing the CS reconstruction step. The minimization of Equation 5.1 was achieved using the Strict DC method with the coil combination mentioned above. This algorithm will be termed CS-CC (Compressed Sensing on Combined Coils) throughout this chapter and is presented in pseudo-code below.

```
Input
    x: Image
    y: Undersampled k-space data
    $\Psi$: Sparsifying transform
    FFT: Fast Fourier Transform (image space → k-space)
    IFFT: Inverse FFT (k-space → image space)
    Approximate norms using Equation 2.11
    Determine $w_{comb}$ and $w^*_{comb}$ using an adaptive combination
        algorithm [93]

Initialize
    $x_0$ = IFFT(y), normalize $x_0$ to 1, n = 1, $\epsilon = 1$

Iterations
    while $\epsilon > 10^{-4}$
        Combine multi-coil dataset using $w_{comb}$
            → Single-coil dataset $x_{n-1}$
        $d = \Psi^* \left(W_\Psi^{p-2} \Psi x_{n-1}\right)$        , $W_\Psi$ calculated from $x_{n-1}$
        Determine step size $t$ such that $\|\Psi\left(x_{n-1} - t \cdot d\right)\|_p = min$
            (e.g. in MATLAB® using fminbnd, an exact line search)
        $x_n = x_{n-1} - t \cdot d$
        Multiply single coil dataset with $w^*_{comb}$ → Multi-coil dataset $x_n$
        Reinsert acquired k-space points at respective positions
            (strict data consistency)
        if n is multiple of 30
            $\epsilon = \epsilon \cdot 0.5$
        end
        n = n + 1
    end
```

W_Ψ is hereby defined as in Equation 2.13. W_Ψ^{p-2} means an elementwise raising to the power $p - 2$. In the CS-CC algorithm, ∇_{DG} served as sparsifying transform Ψ. The reason for not using the identity matrix (i.e., sparsity in the image domain) was that ∇_{DG} leads to smoother reconstruction results while preserving sharp transitions (edges). The CS-CC algorithm and all other data processing steps were implemented in MATLAB® (The MathWorks, Inc., Natick/MA, USA).

5.2.4 Sparsification and timeframe reconstruction

Figure 37 describes how the data for an individual timeframe can be effectively sparsified to enable CS reconstruction. First, a composite image is calculated. This can be performed by collecting an additional calibration dataset where several fully sampled radial k-spaces are acquired. The fully sampled k-space data exhibits temporal blurring due to the relatively long acquisition time for each k-space compared to the timescale of the cardiac motion. All

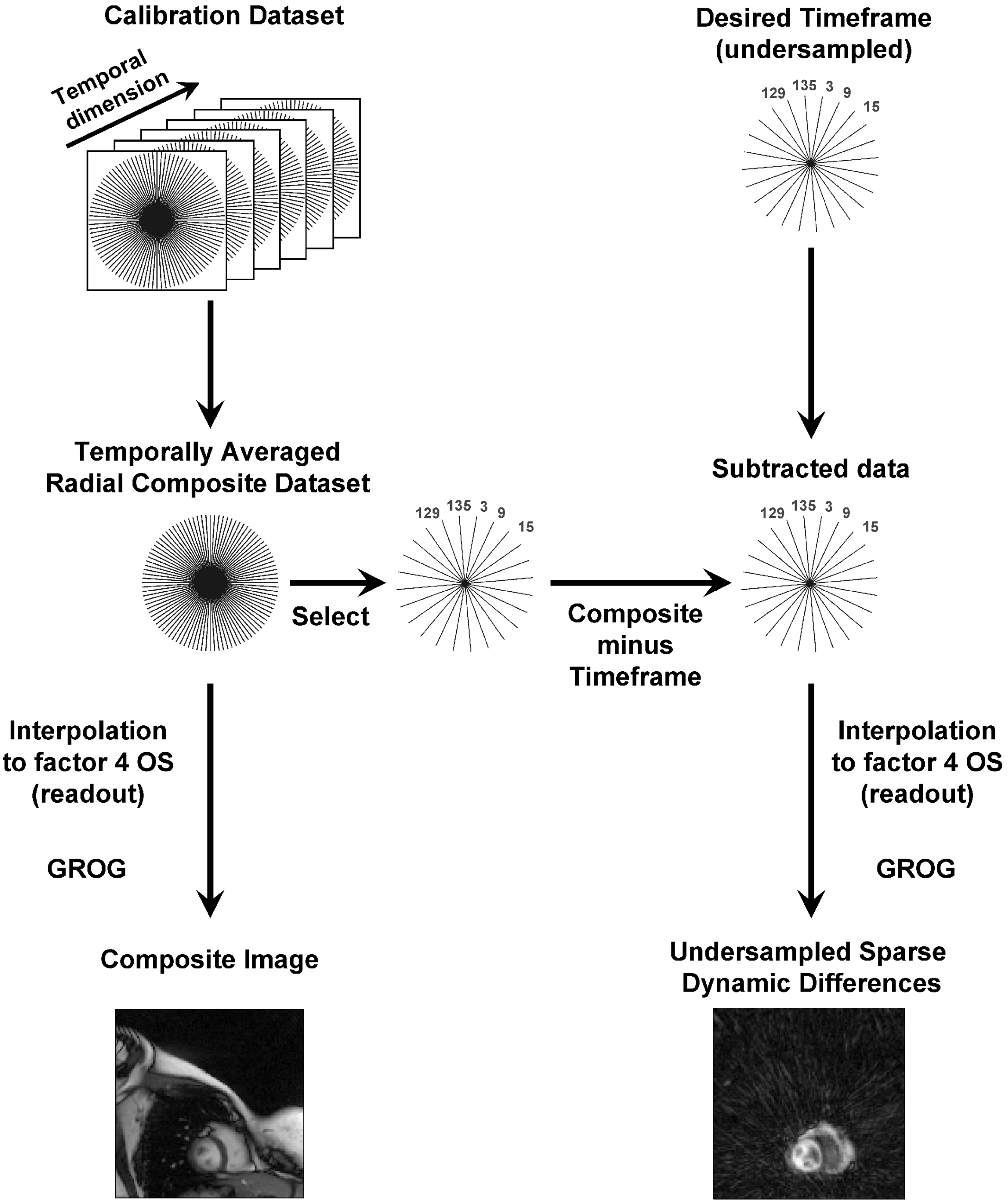

Figure 37: *Method used to sparsify the data in this chapter. The undersampled timeframe data are subtracted from the respective projections of a composite radial dataset. After interpolating and gridding using GROG [39, 40], the obtained undersampled Cartesian data are subject to CS reconstruction. Details can be found in the text.*

fully sampled radial k-spaces are averaged leading to a radial composite dataset. Another possibility would be to acquire interleaved undersampled radial data where all projections are acquired several times. These data can then be averaged, leading to a fully sampled composite radial dataset. Usually, the radial data are oversampled by a factor of 2 along the readout direction to prevent foldover artifacts. These oversampled radial data can be interpolated to an oversampling factor of 4 along the readout direction. This procedure is suggested in [39] to avoid SNR losses in the final gridded image. The interpolated radial composite data are then gridded using GROG [39, 40], leading to the desired Cartesian composite image.

The undersampled data from the individual timeframes are sparsified by subtracting the undersampled projections from the respective projections of the radial composite data. After this subtraction, these data are interpolated and gridded using GROG analogous to the procedure described for the composite image. The result is a sparse undersampled Cartesian dataset that can be reconstructed with the proposed CS-CC algorithm. The respective timeframe image can be obtained from the CS reconstructed dynamic differences by subtraction from the composite image (see Figure 38). Throughout this chapter, for all timeframes of the respective datasets, the same azimuthally equally-spaced projections were used for subtraction and subsequent CS reconstruction.

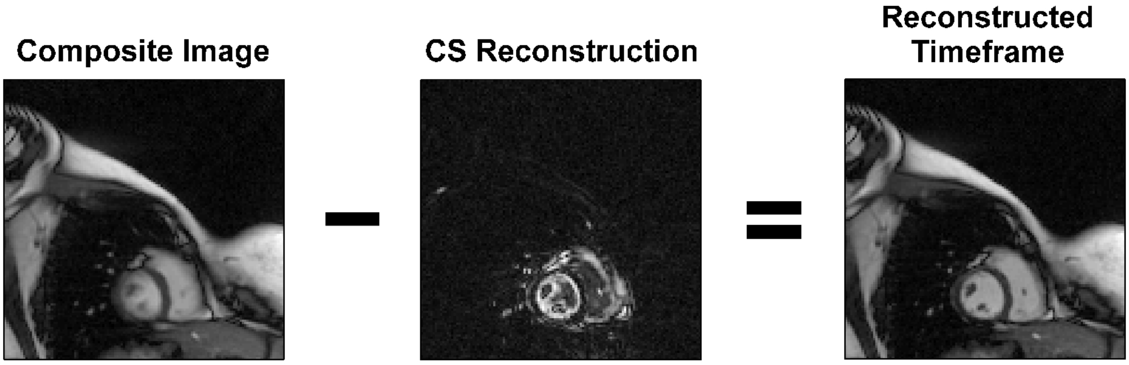

Figure 38: *The desired timeframe image can be obtained by subtracting the CS reconstructed sparse dynamic differences from the composite image.*

5.2.5 Simulations

All simulations were performed using MATLAB®. A numerical dynamic five-coil phantom (matrix size 128x128) mimicking the cardiac cycle over 20 timeframes (see Figure 39) was used to investigate the dependency of the proposed method on noise, acceleration factor, and the norm p. To this end, the 20 timeframes of the Cartesian multi-coil phantom dataset (maximum intensity of the dataset was normalized to 1) were interpolated onto a radial grid with 256 readout-points and 144 projections per frame. These radial data were then further processed according to Figure 37. The sparse differences for all timeframes were calculated from the same azimuthally equally-spaced projections (angular spacing being dependent on the applied acceleration factor). The coil information of the final Cartesian composite image was used to calculate weights to combine (w_{comb}) and de-combine (w^*_{comb}) the individual channels using an adaptive coil combination procedure [93].

The dependency of the reconstruction quality on noise, acceleration factor, and the applied norm p in the CS-CC method were tested. Therefore, pseudo-random Gaussian noise of different standard deviations ($\sigma = \{0; 0.10; 0.25; 0.50; 0.75; 1.00; 2.50; 5.00\}$) were added to

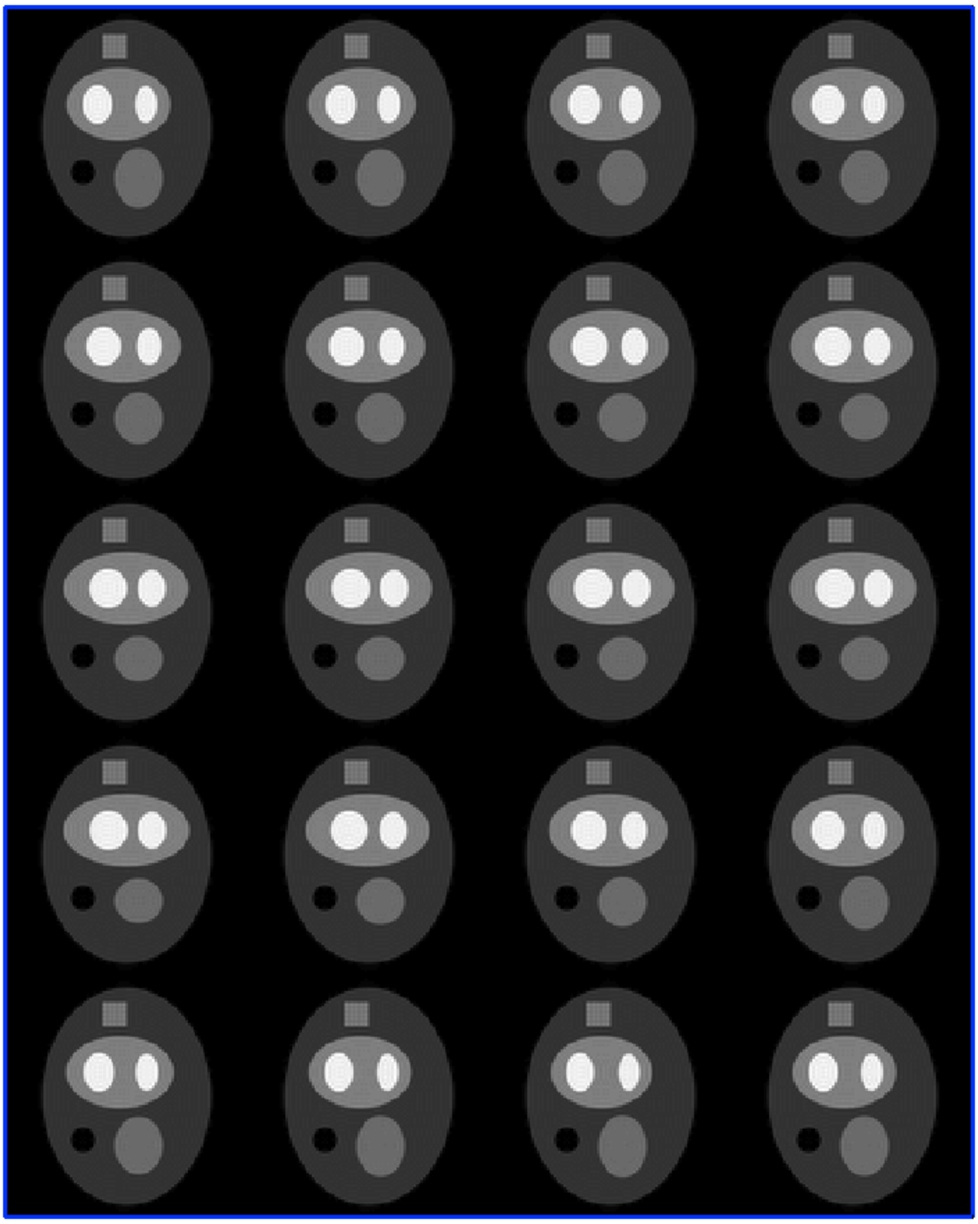

Figure 39: *The dynamic phantom utilized in the simulations in this chapter. The timeframes are depicted from the upper left to the lower right corner. The large ellipsis with the two white smaller ellipses exhibits a contraction. A timecourse is shown in Figure 42.*

the real and imaginary part of the radial data prior to interpolating the readout direction to an oversampling factor of 4. The phantom was reconstructed from the following numbers of projections: 72, 48, 36, 24, 18, 16. The norm p was varied as follows: 0.25, 0.50, 0.75, 0.90, 1.00. The case $p = 1.00$ is the "limiting convex case" and is often used in CS reconstructions (e.g., [16, 29, 68]) as it has been demonstrated that in most cases an l_1-minimization leads to the sparsest solution [31]. However, Chartrand demonstrated that using a l_p-norm with $p < 1$ leads to equivalent results [32].
For convolution gridding, **N**on-**U**niform **F**ast **F**ourier **T**ransform (NUFFT) [94] was used with a Kaiser-Bessel window of width 6. All data were gridded to a 128x128 matrix. The reconstruction quality was quantified using the RMSE metric as defined in Equation 3.1.

5.2.6 Retrospectively undersampled radial cardiac cine dataset

The results from the simulations were transferred and tested on a retrospectively undersampled radial cine dataset of a short-axis view of the heart of a healthy volunteer. Imaging parameters: Siemens Espree 1.5 T clinical scanner (Siemens Healthcare, Erlangen, Germany), 15 channels, radial SSFP, $\alpha = 70°$, $T_E = 1.91$ ms, $T_R = 3.82$ ms, FOV = 280x280 mm^2, 144 projections, 256 readout points, 14 timeframes. The data were acquired in breath-hold and with ECG triggering. The composite image was obtained by averaging the complete cine dataset as depicted in Figure 37. Again, all timeframes consist of the same azimuthally equally spaced projections which were used for subtraction from the radial composite dataset. The coil information of the final Cartesian composite image was used to calculate the weights to combine (w_{comb}) and de-combine (w^*_{comb}) the individual channels using the adaptive coil combination procedure [93]. The subsequent CS reconstruction was performed with the nonconvex norm $p = 0.75$ based on the results of the simulations. The following numbers of projections were tested for the accelerated dataset: 48, 24, and 16. The data were reconstructed using CS and a profile across the heart was selected. This profile was then displayed for each timeframe, depicting the timecourse of the cardiac motion that was additionally compared to a convolution gridded dataset of the non-subtracted projections.
For convolution gridding, NUFFT [94] was used with a Kaiser-Bessel window of width 6. The data were gridded to a 128x128 matrix. For GROG, the data were gridded to a 192x192 matrix to prevent foldover artifacts.

5.2.7 Undersampled radial cardiac real-time dataset

The potential of the CS-CC method was tested on radially acquired dynamic cardiac data from a healthy volunteer. First, a calibration dataset was acquired by sampling 144 projections 20 times with a linear increasing angle. The following imaging parameters were used: Siemens Espree 1.5 T clinical scanner (Siemens Healthcare, Erlangen, Germany), 18 channels, radial SSFP, $\alpha = 70°$, $T_E = 1.38$ ms, $T_R = 2.76$ ms, FOV = 300x300 mm^2, 144 projections/k-space, 256 readout points. The data were acquired in free breathing without cardiac gating. A composite radial dataset was calculated by averaging the fully-sampled radial datasets. Due to the long data acquisition time at this stage, information about all cardiac cycle phases appears in the composite dataset. This radial composite data was further processed as depicted in Figure 37. The coil information of the final Cartesian composite image was used to calculate the weights to combine (w_{comb}) and de-combine (w^*_{comb}) the individual channels using the adaptive coil combination procedure [93]. In the second stage, an undersampled radial real-time dynamic cardiac dataset was acquired by sampling 24 azimuthally equally spaced projections a total of 150 times, corresponding to an angular undersampling factor of 6 with respect to the calibration dataset (144 projections). The angular undersampling factor is more than 8 with respect to the Nyquist-Shannon theorem, which dictates 200 projections. Again, the data were acquired in free-breathing and without cardiac gating. Besides the number of projections, the same imaging parameters as above were used. Each timeframe had a temporal resolution of $24{\cdot}T_R \approx 66$ ms and was reconstructed using the CS-CC method with the nonconvex norm $p = 0.75$ based on the results of the simulations.

Analogous to the retrospectively undersampled cine dataset, a profile across the heart was selected and the resulting CS reconstructed timecourse of the cardiac motion was visually compared to the non-CS reconstructed data, a convolution gridded dataset of the non-subtracted projections, and a GROG-gridded dataset of the non-subtracted projections. For convolution gridding, NUFFT [94] was used with a Kaiser-Bessel window of width 6. The data were thereby gridded to a 128x128 matrix. For GROG, the data were gridded to a 256x256 matrix to prevent foldover artifacts.

5.3 Results

5.3.1 Simulations

Figure 40 shows the dependency of the norm p on the noise level and the acceleration factor. For noise-free data as well as low and moderate noise levels (data not shown), minima in the RMSE courses at values of $p < 1.00$ (nonconvex norms) can be observed. In most cases, $p = 0.75$ was the optimal norm in terms of RMSE. An example of this is shown in Figure 40a for noise-free data. However, as the data became increasingly corrupted by noise, the minimum shifted to $p = 1.00$ as demonstrated in Figure 40b. As seen, this was true for all investigated numbers of projections. In Figure 40b, the order of the RMSE courses is reversed, i.e. higher acceleration resulted in lower RMSE values compared to low acceleration. This finding is counterintuitive since higher acceleration should lead to high RMSE values while low acceleration should lead to low RMSE values. An explanation for this observation is provided in the Discussion section.

In Figures 40c and d, 24 projections were used for reconstruction while the noise level was varied. The findings of Figures 40a and b (existence of a minimum in the RMSE courses, shift of minimum toward larger norms) were confirmed. The shift of the optimal norm is highlighted in Figure 40d by the black circles which indicate the position of the minimum in the RMSE courses. It can be seen that this minimum shifts from $p = 0.75$ ($\sigma = 0.75$) via $p = 0.90$ ($\sigma = 1.00$) to $p = 1.00$ ($\sigma = 2.50$ and 5.00).

To illustrate the meaning of the RMSE plots in Figure 40, an example image from the dynamic dataset as well as timecourses of dynamic datasets are shown in Figures 41 and 42. In Figure 41, the CS-CC reconstruction (24 projections, $\sigma = 1.00$, $p = 0.75$) offers an accurate depiction of the phantom without visible artifacts. The convolution gridded data exhibited lower apparent contrast and streaking artifacts. The profile line, which is marked in the Reference image of Figure 41, resulted in the timecourses shown in Figure 42. As seen, the CS reconstruction (Figure 42b) led to an accurate representation of the timecourse with low artifact level and without blurring. The convolution gridded data (Figure 42c) led to a correct reconstruction of the timecourse. However, undersampling artifacts hampered image quality.

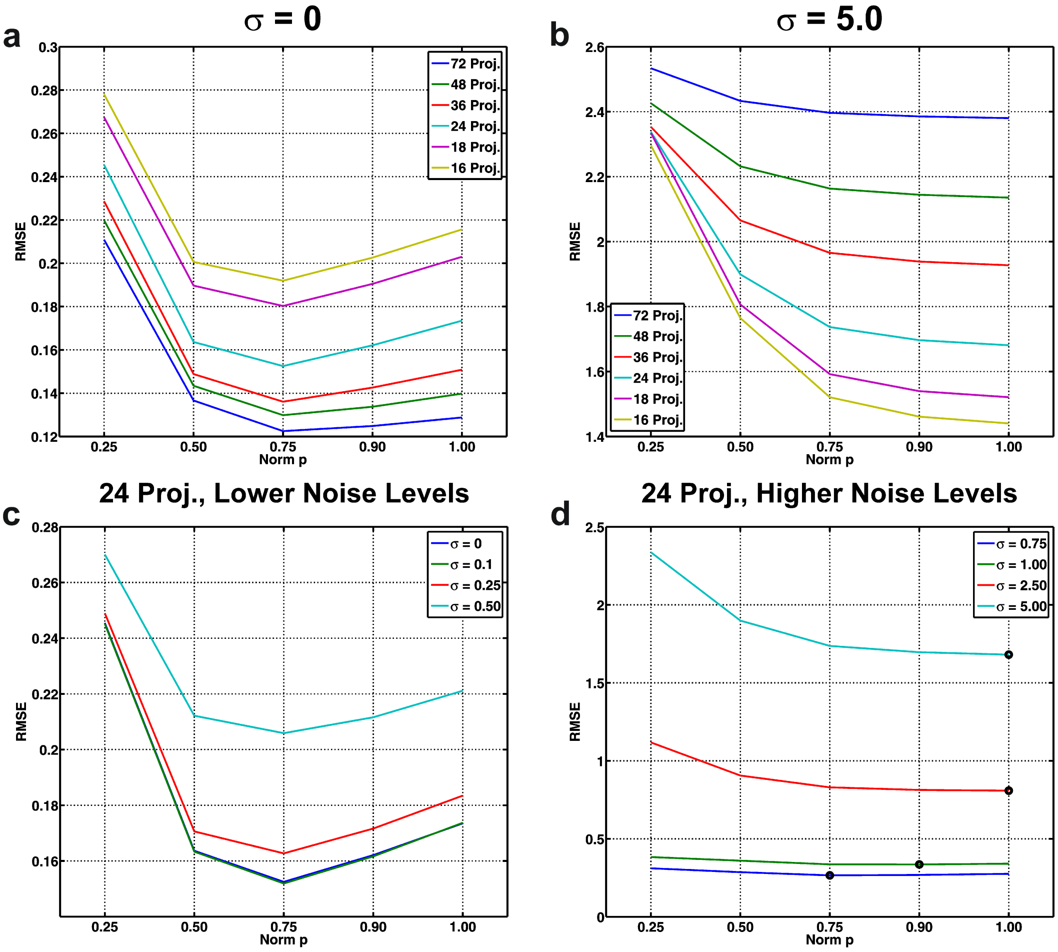

Figure 40: *Resulting exemplary RMSE plots from the simulations. (a) For all investigated numbers of projections, the RMSE increased with fewer projections used to reconstruct the noise-free data. Furthermore, a minimum in the RMSE course indicating an optimal norm for reconstruction can be found. Here, the optimal norm in terms of RMSE is* $\mathrm{p} = 0.75$. *(b) For noisy data* ($\sigma = 5.0$)*, the minima of the RMSE courses appear at* $\mathrm{p} = 1.00$. *Please note the reversed order of the courses compared to (a). (c) For the lower investigated noise levels* ($\sigma \leq 0.50$)*, the results from 24 projections indicate an optimal nonconvex norm* $\mathrm{p} = 0.75$. *(d) When the noise was further increased* ($\sigma \geq 0.75$)*, the minima of the RMSE courses (black circles) shifted from nonconvex norms to the limiting convex case* $\mathrm{p} = 1.00$.

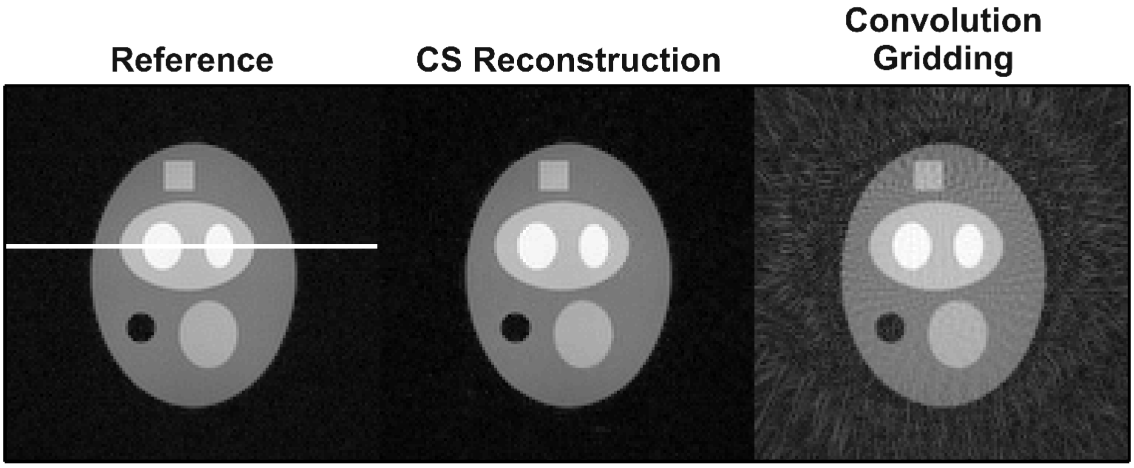

Figure 41: *Results obtained from 24 projections, noise level* $\sigma = 1.00$, *and* $\mathrm{p} = 0.75$. *These images display an example timeframe from the dynamic dataset. The profile line which led to the timecourses in Figure 42 is displayed in the reference image.*

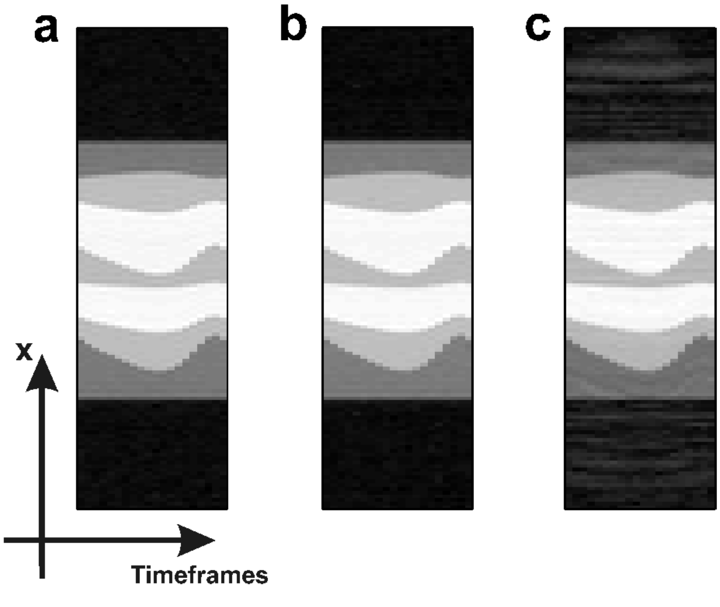

Figure 42: *Resulting timecourses for the profile line depicted in Figure 41 from 24 projections, noise level* $\sigma = 1.00$, *and* $\mathrm{p} = 0.75$. *(a) Reference; (b) CS reconstruction; (c) Convolution gridding of non-subtracted projections.*

5.3.2 Retrospectively undersampled radial cardiac cine dataset

To demonstrate the quality of the CS reconstructed images, two example reconstructions from 24 projections are presented in Figure 43. In the upper row of Figure 43, a timeframe from mid-systole is displayed. The cardiac phase is correctly represented in the convolution gridded image. However, prominent streaking artifacts as well as reduced apparent contrast hamper image quality. The CS reconstruction offers an accurate representation of the systolic motion state without visible artifacts. The same holds for the reconstructed data from end-diastole (Figure 43, bottom row).

The chosen profile line that is marked in Figure 43, Reference, is used to illustrate the quality of the reconstructed timecourse of the cardiac motion. Figure 44 presents the resulting timecourses for the different acceleration factors. For reconstruction from 48 projections, the cardiac motion in the CS reconstruction (Figure 44b, upper row) and the convolution gridded data (Figure 44c, upper row) is equally well represented. However, the convolution gridded data offer a lower apparent contrast than the CS reconstruction. The improved apparent contrast in the CS reconstruction compared to the convolution gridded data becomes even more prominent when fewer projections are used for reconstruction (Figure 44, middle and bottom row). The timecourse in the CS reconstruction remained accurate as seen by comparison with the reference.

5.3.3 Undersampled radial cardiac real-time dataset

In Figure 45, example timeframes from the real-time dataset are displayed. The CS reconstructions offer high apparent contrast and few remaining undersampling artifacts while the convolution gridded images exhibit remaining streaking artifacts and a reduced apparent contrast. The representation of the cardiac state is equivalent for the CS reconstruction and the convolution gridded images. In addition, no significant influence of the respiratory motion on the accuracy of the CS reconstructed images could be observed in the real-time dataset.

The timecourse of the profile whose position is depicted in Figure 45 (systole, CS Reconstruction) is displayed in Figure 46. By comparing the CS reconstruction with the convolution gridded data, it can be seen that the temporal fidelity of the cardiac motion is equivalently reflected. However, the CS reconstructed data exhibit fewer visible artifacts and an improved apparent contrast compared to the convolution gridded dataset.

Another important aspect for real-time imaging, which can be appreciated in these images, is the tolerance against respiratory motion. This motion is clearly seen as a sinusoidal movement of the heart in Figure 46. It can be seen that the respiratory motion does not affect the accuracy of the CS reconstruction.

5.4 Discussion

This work investigated the benefit of using the CS-CC method with dynamic radial cardiac datasets. Three main results were found in this study. First, the CS-CC method allowed accurate reconstructions from undersampled radial cardiac real-time datasets. Second, the optimal norm p for the CS reconstruction depended on the noise level. A good compromise for all investigated noise levels was $p = 0.75$. Third, even though the real-time dataset was acquired in free breathing, the CS-CC algorithm led to accurate reconstructions. This

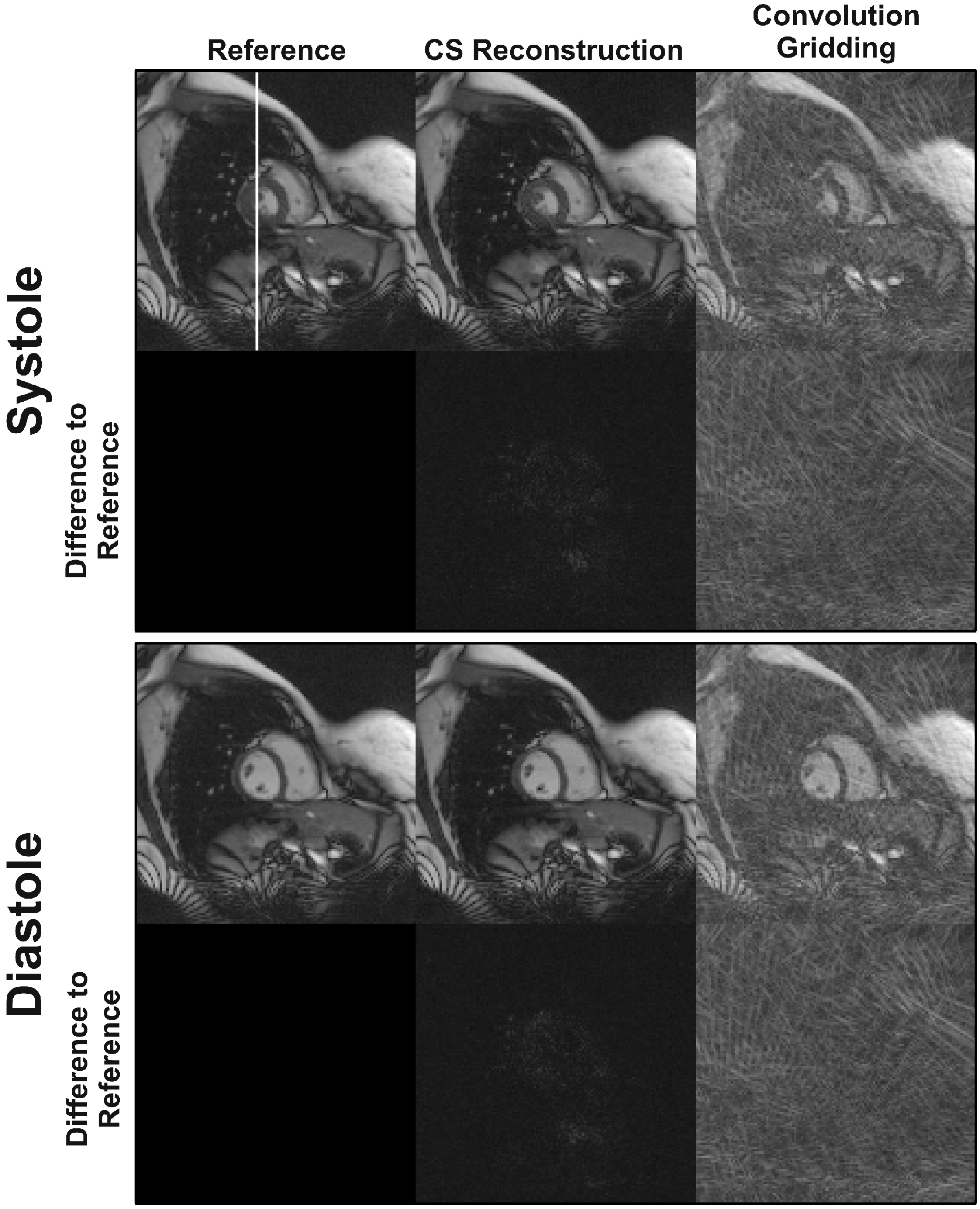

Figure 43: *Exemplary results obtained from 24 projections. Timeframes from mid-systole and end-diastole are displayed. The difference images for the CS reconstructions were obtained by subtraction from the GROG-gridded reference image (first row). Those for the undersampled convolution gridded data were calculated by subtraction from a fully sampled convolution gridded image. The profile line that led to the timecourses in Figure 44 is displayed in the systolic reference image.*

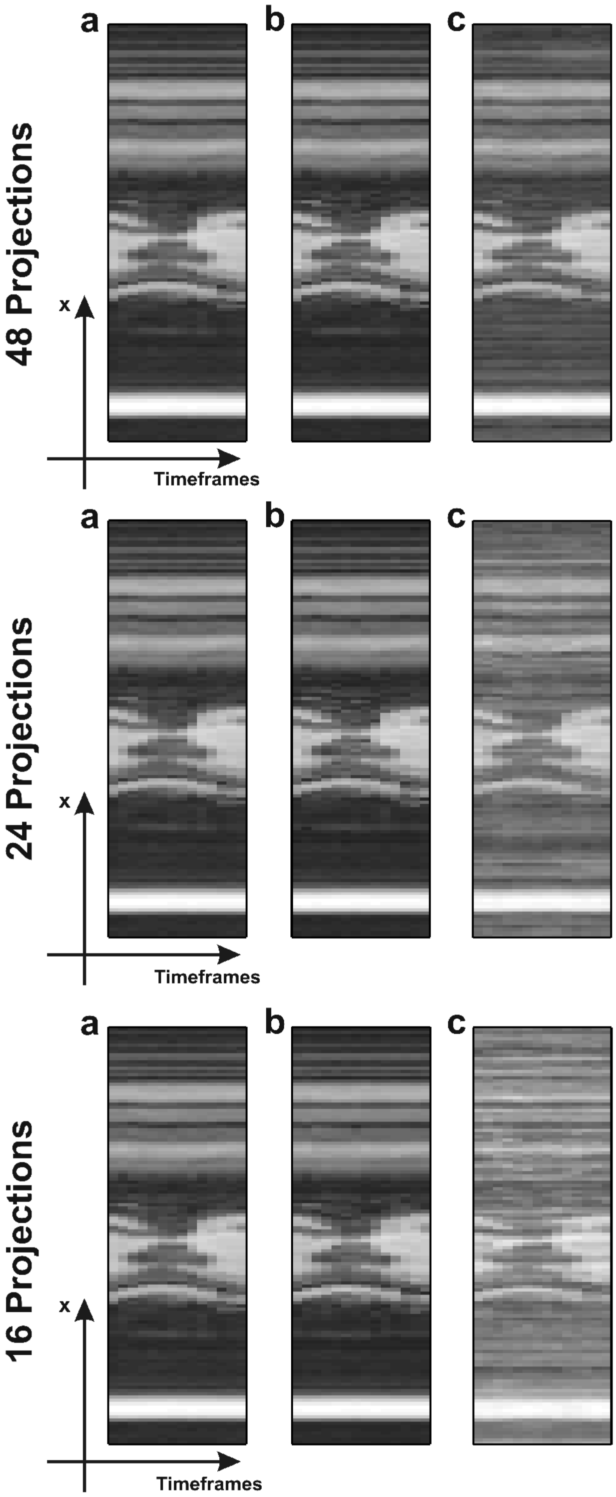

Figure 44: *Resulting timecourses for the profile line depicted in Figure 43 from 48, 24, and 16 projections. (a) Reference; (b) CS reconstruction; (c) Convolution gridding of non-subtracted projections.*

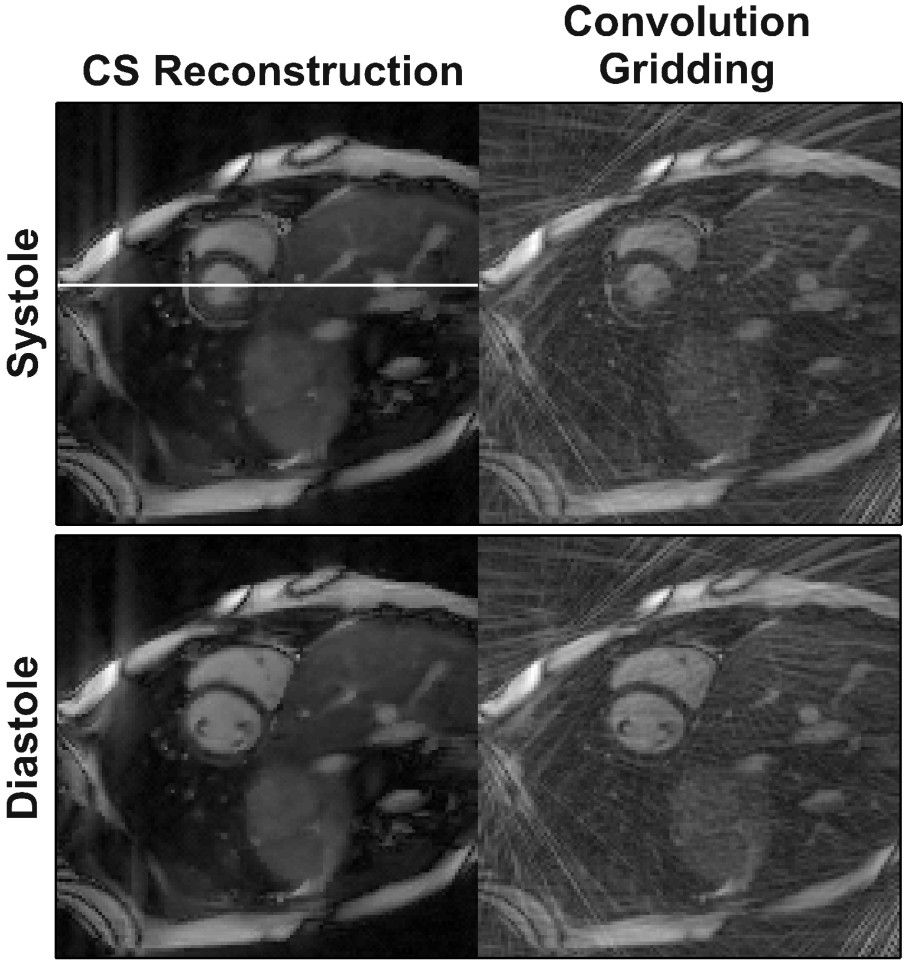

Figure 45: *Exemplary results from timeframes in mid-systole and end-diastole of the undersampled real-time dataset. While the cardiac phase display is similarly accurate in both the CS reconstruction and the convolution gridded image, undersampling artifacts are eliminated and the contrast is improved in the CS reconstruction. The profile line leading to the timecourses in Figure 46 is displayed in the systolic CS reconstructed image.*

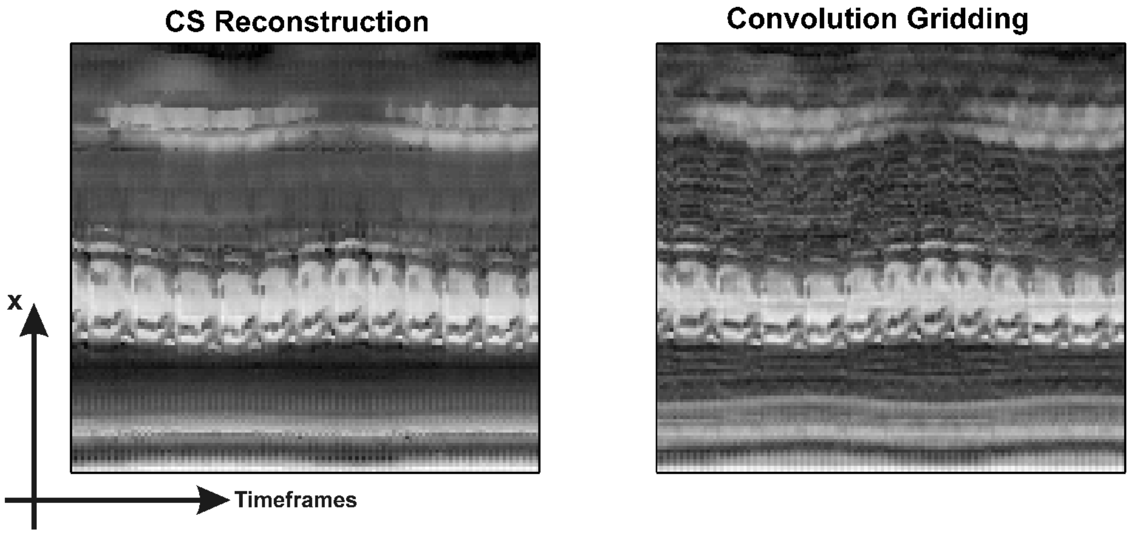

Figure 46: *Resulting timecourses for the profile line depicted in Figure 45. The results were obtained from 24 projections. As seen, the CS reconstruction and the convolution gridded data exhibit equivalent temporal fidelity. However, in the CS reconstruction, less visible undersampling artifacts than in the convolution gridded data can be observed. The breathing of the volunteer can be clearly seen by the sinusoidal motion of the heart.*

indicates that, even when the sparsity in the difference datasets is reduced due to the respiratory motion, this can be tolerated in the CS reconstruction.

5.4.1 Choice of the norm p

In the simulations, it could be demonstrated that the noise level critically influences the choice of the reconstruction norm p in terms of RMSE (see Figure 40). No such impact on the reconstruction quality could be found by reducing the number of projections. The simulation results indicate that $p = 0.75$ is a good compromise for the norm. However, if the minima in the RMSE courses did not occur at $p = 0.75$, applying this norm did not lead to significantly worse results in terms of RMSE than the optimal norm. Thus, $p = 0.75$ performed well for all investigated noise levels and was therefore chosen to reconstruct the retrospectively undersampled cine dataset as well as the real-time dataset.

A similar study to investigate the dependency of the CS method in [32] on noise, acceleration factor, and norm was presented in Chapter 4. It was also published in the Journal of Magnetic Resonance [95]. The CS algorithm in that study utilized image space sparsity in contrast to the Discrete Gradient transform ∇_{DG} used in the current chapter. Nonetheless, similar results were derived leading to the choice $p = 0.75$. However, the CS-CC algorithm exploits joint sparsity while the method in Chapter 4 and [95] is for use with single-receiver data. Both studies lead to equivalent results, which might indicate that using $p = 0.75$ is a good compromise for many sparsifying transforms. This claim should be further investigated.

5.4.2 Remarks on the RMSE metric

In Chapter 4, it was already argued that the RMSE metric is problematic in the presence of noise. In this connection it was demonstrated that the RMSE can lead to counterintuitive results. Such a case can be observed in Figure 40b, where the RMSE courses are in reversed order compared to Figure 40a. This leads to the observation that, with increasing acceleration factor, the reconstruction quality in terms of RMSE improves. However, the explanation for this behavior is that by using fewer projections for reconstruction, more degrees of freedom are left for the CS reconstruction algorithm. When using the Discrete Gradient operator ∇_{DG} in Equation 5.1, this leads to increased smoothing of noisy areas in the image. Thus, since the RMSE is calculated using a noise-free reference, reconstruction results with increased smoothing lead to lower RMSE values than results with a less degree of smoothing. Hence, as already concluded in Chapter 4, the RMSE should be considered critically in the presence of noise. Which image metric is best suited to quantify CS reconstruction quality should be further investigated. In any case, an additional visual inspection of the resulting images is strongly recommended.

5.4.3 Advantage of the joint sparsity approach

The easiest way to transfer the Gradient Descent method in [32] to multi-receiver datasets is to separately CS reconstruct each channel. However, in the presence of noise, it has been shown that spike artifacts can occur (see Chapter 4 and [95]). Furthermore, as already mentioned, the SNR in the individual channels decreases with increasing distance from the receiver coil. Thus, structures that can be recognized in one channel may offer low SNR or even become lost in the receiver noise of other channels. If CS is applied to such a coil, then the low SNR structures can be smoothed out using ∇_{DG} in Equation 5.1. Hence, in the sum-of-squares image of all CS reconstructed channels, blurring can be observed, which compromises the resulting image quality. This blurring in combination with possible spike artifacts at different positions in the individual channels leads to a severe impairment of the final image. In the CS-CC method that exploits the joint sparsity of the coil array, the CS step is performed on an SNR-optimized image, that was adaptively combined from all channels [93]. Thus, all structures with sufficient SNR will be accurately recovered so long as the acceleration factor is not too high. Furthermore, possible spike artifacts occur at only one distinct position in the combination image instead of at different positions in each channel. Hence, the resulting CS-CC reconstructed images offer an improved image quality compared to CS reconstructions of each separate channel. It is worth mentioning that in the case of very high SNR (approx. $\mathrm{SNR} \geq 50$), the differences between coil-by-coil CS and CS-CC diminish and vanish in the limit for noise-free data, especially if ∇_{DG} is used in Equation 5.1 as in the present study.

5.4.4 Remarks on the strict data consistency constraint

The CS-CC algorithm enforces strict data consistency as suggested in [32]. This means that the acquired data repeatedly reinserted in the CS reconstructed k-spaces are assumed to be perfect. This is logical for ideal, noise-free data. However, in the presence of noise, this assumption is violated which can lead to reconstruction artifacts. As demonstrated in this chapter, Chapter 4, and [95], these artifacts manifest in forms such as spike artifacts. Many previously proposed CS algorithms optimize a weighted sum of a relaxed data consistency constraint and a sparsity constraint in the cost function (e.g., [16, 29], see also Equation 2.7 in Section 2.3.1). The weighting factor is a regularization parameter that critically influences the reconstruction quality and is usually determined empirically (i.e., by “trial and error“, as mentioned in Sections 2.3.1 and 3.1). Solving a problem where a certain deviation is allowed between the acquired k-space data and the reconstructed data points (i.e., relaxed instead of strict data consistency) is reasonable for noisy data. However, without a reference image, as is the case in real-time imaging, it is always questionable if the applied regularization parameter is really the optimal choice. Since the CS-CC resulted in accurate reconstructions, enforcing strict data consistency was preferred to the time-consuming, user-influenced, and possibly suboptimal choice of a regularization parameter in a relaxed data consistency algorithm. Without the need to determine a regularization parameter, the CS-CC algorithm can be used as a “Black-Box-Algorithm“ to reconstruct sparse undersampled datasets.

5.4.5 General remarks

In contrast to recently published studies using a nonlinear inverse reconstruction [83–86], no sliding window reconstruction or temporal filtering of the reconstructed data was applied. However, the presented work can be combined with these windowing and filtering techniques. Furthermore, combining CS-CC with parallel imaging could allow accurate reconstructions from even fewer projections than used in this chapter. Additionally, incorporating sliding window, temporal filtering and parallel imaging, could reduce the temporal resolution to even less than the 66 ms given in this study.

It should be noted that the afore mentioned nonlinear inverse reconstruction does not include an explicit sparsity constraint. Rather, it exploits the coil sensitivity maps and, thus, is more related to parallel imaging (in particular SENSE [7]) than to CS.

5.5 Conclusion

In this chapter, a joint sparsity CS algorithm that enforces strict data consistency for radial dynamic cardiac imaging was introduced. With this method, the undersampled radial data is gridded using GROG, which results in undersampled pseudo-radial Cartesian data. Hence, the CS step can be performed in the Cartesian plane. Strict data consistency is guaranteed by reinserting, after each CS reconstruction step, the gridded undersampled radial data at the respective positions in k-space. Thus, de-gridding of the gridded data is unnecessary, making the proposed method computationally efficient. The individual coils are combined and de-combined by weight sets derived from an adaptive coil combination method. Thereby, the CS step can be performed on a SNR-optimized combination image while strict data consistency in each individual coil data can be enforced.

In simulations, the interdependencies of noise, acceleration factor, and the norm of the cost function on the reconstruction quality were investigated. Optimized parameters were derived, transferred, and tested on a retrospectively undersampled radial cardiac cine dataset. Finally, to demonstrate the applicability of the proposed method on real-time datasets, successful reconstructions were obtained from a radial real-time dataset with 24 projections per timeframe in free-breathing and without cardiac gating. A temporal resolution of 66 ms could be achieved. Thus, the proposed sparsification scheme in combination with the CS-CC method allowed reconstructing undersampled dataset with high spatial and high temporal resolution. In principle, the proposed technique allows the reconstruction of undersampled datasets from patients with cardiac arrhythmia. Thus, more detailed information on cardiac anatomy or function about arrhythmic diseases is feasible with MRI without the need for long breath-holds or ECG triggering. Furthermore, the presented technique may prove advantageous for other dynamic applications such as vocal tract imaging in speech production research, swallowing, or knee and shoulder imaging during motion.

6 Iterative GRAPPA

This chapter introduces a computationally efficient version of Iterative GRAPPA. Because it works solely in the Cartesian plane, gridding and de-gridding are unnecessary. The potentials and limitations of this method were investigated by phantom simulations. Additionally, the practical applicability was demonstrated on a retrospectively undersampled radial cardiac cine dataset.

This chapter covers the groundwork necessary to develop one possible combination of parallel imaging and CS. This combination termed CS-GRAPPA is described in detail in Chapter 7.

6.1 Motivation

In conventional Cartesian parallel imaging, the k-space is regularly undersampled. For example, only every second, third, etc., line in k-space is acquired. This reduces the FOV which leads to coherent foldover artifacts due to the periodicity of the Fourier transform. These coherent foldover artifacts can be unfolded using conventional Cartesian parallel imaging reconstruction methods such as SENSE [7] or GRAPPA [8]. In CS, sampling patterns need to be applied that lead to incoherent artifacts after Fourier transform of the undersampled k-space. These undersampling patterns are unsuitable for reconstruction using conventional Cartesian parallel imaging methods.

One solution is to apply a GRAPPA kernel by which the central k-space point is reconstructed from the surrounding k-space points. This strategy works even when k-space is irregularly undersampled (e.g., pseudo-random or GROG-gridded [39, 40] undersampled Non-Cartesian data). This idea was already published in 2005 [96]. In that abstract, gridding and de-gridding steps were required when the data were acquired using Non-Cartesian trajectories. However, in this thesis, all Non-Cartesian data were gridded using GROG prior to Iterative GRAPPA. Thus, all further reconstruction steps could be performed solely in the Cartesian plane, which is computationally more efficient than repeatedly applying gridding and de-gridding operations.

6.2 Theory of GRAPPA

In 2002, GRAPPA was introduced by Griswold et al. [8]. GRAPPA emerged from SMASH [4] and further improvements of SMASH called AUTO-SMASH [5] and VD-AUTO-SMASH [6]. The idea common to all of these techniques is that the intrinsic sensitivity variations of the coil array provide additional spatial information that can be exploited to reconstruct undersampled k-space data. Consider the 1D signal equation for the signal received in coil L:

$$S_L(k_y) = \int \rho(y) \cdot C_L(y) \cdot e^{ik_y y} dy \tag{6.1}$$

where $S_L(k_y)$ is the signal of coil L at k-space location k_y, $\rho(y)$ is the spin density at spatial position y, and $C_L(y)$ is the sensitivity of coil L at spatial position y.
Similarly, the signal at the k-space position $k_y + m\Delta k_y$ ($m \in \mathbb{N}$) is given by:

$$S_L(k_y + m\Delta k_y) = \int \rho(y) \cdot C_L(y) \cdot e^{i(k_y + m\Delta k_y)y} dy \tag{6.2}$$

Thus, the equations are equivalent except for the spatial harmonic $e^{im\Delta k_y y}$. GRAPPA aims to mimic this spatial harmonic by linearly combining the sensitivity variations of all coils of the array:

$$C_L(y) \cdot e^{im\Delta k_y y} \approx \sum_{K=1}^{N_C} w_{K,L,m} \cdot C_K(y) \tag{6.3}$$

The indices K and L run from 1 to N_C (number of coils in the array). This equation looks very similar to the SMASH equation; however, in SMASH, spatial harmonics are mimicked appropriately by combining the coil sensitivities. Here, missing spatial harmonics in each individual channel are created.
In conventional Cartesian parallel imaging, whole phase encoding lines in the k_y-directions are skipped. If every second line in k-space is missing, then m takes only the value 1. In general, if e.g. every R-th line is acquired, then m runs from 1 to R-1.
Equation 6.3 can be inserted into Equation 6.2 which yields:

$$\begin{aligned} S_L(k_y + m\Delta k_y) &= \int \rho(y) \cdot C_L(y) \cdot e^{im\Delta k_y y} \cdot e^{ik_y y} dy \\ &\approx \int \rho(y) \sum_{K=1}^{N_C} w_{K,L,m} \cdot C_K(y) \cdot e^{ik_y y} dy \\ &\approx \sum_{K=1}^{N_C} w_{K,L,m} \cdot S_K(k_y) \end{aligned} \tag{6.4}$$

In SMASH-like techniques (SMASH [4], AUTO-SMASH [5], VD-AUTO-SMASH [6]), data from the acquired lines in all coils at k-space position k_y are fitted to the missing line at k-space position $k_y + m\Delta k_y$ (compare Fig. 1 in [8]). However, in contrast to Equation 6.4, a combined image similar to SENSE is obtained instead of individual coil images. In GRAPPA, k-space data from more than one acquired line per coil is used to fit every missing line in each individual coil. By using more than one acquired line per coil for fitting, GRAPPA incorporates more information into each reconstructed line, resulting in a substantially improved fit:

$$S_L(k_y + (m+R)\Delta k_y) = \sum_{K=1}^{N_C} \sum_{b=0}^{N_b - 1} w_{L,b,K,m} S_K(k_y + bR\Delta k_y) \tag{6.5}$$

where N_b is the number of acquired k-space lines used to fit each missing line $k_y + mR\Delta k_y$. Additionally, b is an index that runs over these acquired lines and R is the reduction factor of the undersampled dataset. For example, if R $= 2$ and $N_b = 4$, acquired data from the four lines at k-space positions $k_y, k_y + 2\Delta k_y, k_y + 4\Delta k_y$ and $k_y + 6\Delta k_y$ are used to recover the line at $k_y + 3\Delta k_y$.
Equation 6.5 can be compactly expressed as matrix equation:

$$\hat{T} = \hat{w} \cdot \hat{S} \qquad \Longleftrightarrow \qquad \hat{w} = \hat{T} \cdot \text{pinv}(\hat{S}) \tag{6.6}$$

where $\hat{T}$ is the matrix representing the missing target points and $\hat{S}$ is the matrix containing the acquired source points. Equations 6.5 and 6.6 both state that a linear combination of acquired source points from all coils approximates the missing target points in one particular coil. The principle of GRAPPA is shown in Figure 47. The matrix $\hat{w}$ is called the GRAPPA weight set. It can be obtained by acquiring a fully-sampled calibration dataset in which source and target points are known. Alternatively, the central k-space of an undersampled Cartesian dataset can be fully-sampled. This is called the autocalibrating signal (ACS) and was introduced in the AUTO-SMASH technique [5] and is often used in conventional GRAPPA reconstructions [8].

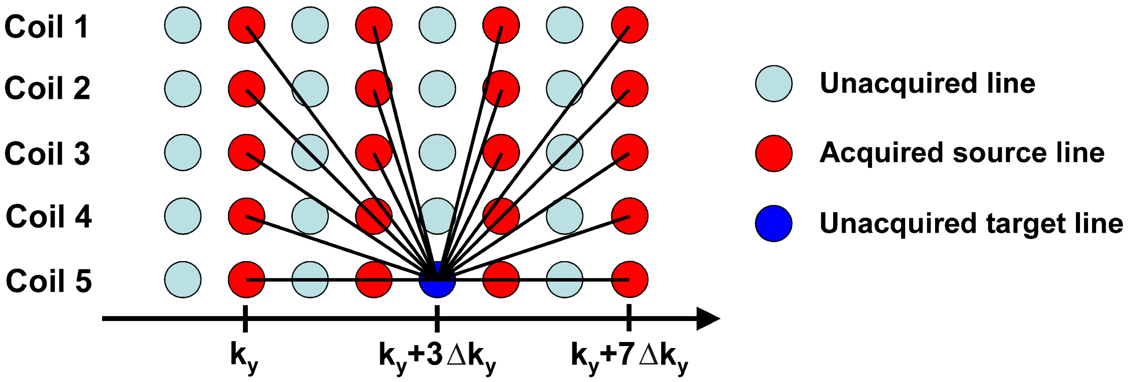

Figure 47: *Basic principle of GRAPPA: several acquired lines (red source lines) from all coils are fitted to each missing line (blue target line) in each individual channel. In the depicted example, R = 2,* $N_b = 4$ *and m = 1. The missing line in coil 1 at k-space position* $k_y + 3\Delta k_y$ *is obtained by a weighted sum of the acquired lines from all five coils at k-space positions* $k_y, k_y + 2\Delta k_y, k_y + 4\Delta k_y$ *and* $k_y + 6\Delta k_y$*. More details can be found in the text. This figure is inspired by Fig. 2 in the original GRAPPA paper [8].*

The GRAPPA weight set can be calculated by taking the Moore-Penrose pseudo inverse of the source points $\hat{S}$, which is defined as:

$$\mathrm{pinv}(\hat{A}) = (\hat{A}^H\hat{A})^{-1}\hat{A}^H \tag{6.7}$$

where $\hat{A}^H$ is the conjugate transpose of the matrix $\hat{A}$. Using a regularized version of the pseudo inverse helps guarantee a stable inversion despite insufficient coil sensitivity variations or insufficient calibration data:

$$\mathrm{pinv}(\hat{A}) = (\hat{A}^H\hat{A} + \mathbb{1}\cdot(\mu \cdot L_{max})^2)^{-1}\hat{A}^H \tag{6.8}$$

This kind of regularization is called Tikhonov regularization. In this equation, μ is a regularization parameter and L_{max} is the largest singular value of $\hat{A}$ that can be obtained by a singular value decomposition. The product $\mu \cdot L_{max}$ is squared since the matrix product $\hat{A}^H\hat{A}$ has squared entries. From a physicist's point of view, the physical units of $\hat{A}^H\hat{A}$ must correspond to those of $\mu \cdot L_{max}$. By normalizing the regularization to the maximum singular value of $\hat{A}$, the regularization parameter μ becomes independent from $\hat{A}$.
Without the additional regularization, the pseudo-inverse might lead to very large values due to small elements in the matrix $\hat{A}$.

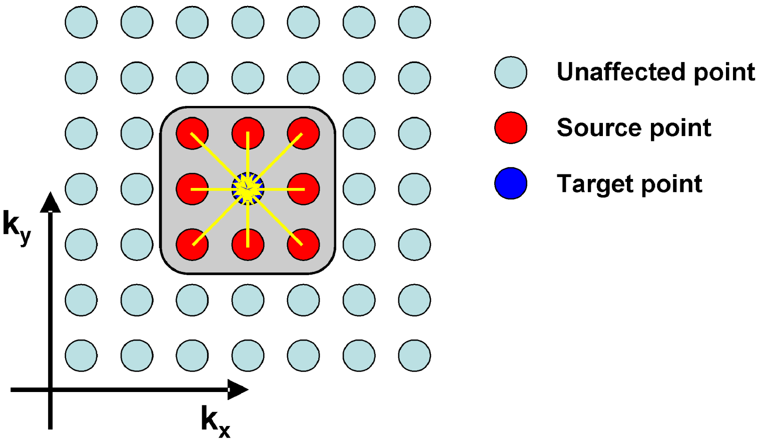

Figure 48: *The basic principle of Iterative GRAPPA as proposed in this thesis. The light blue circles denote k-space points that do not contribute to the GRAPPA reconstruction in this example. An exemplary 3x3 kernel is displayed in which the central target point is reconstructed by the surrounding 8 source points. For the sake of simplicity and clarity of depiction, the coil dimension has been discarded. The GRAPPA kernel calibration and GRAPPA reconstruction was performed according to [8].*

6.3 Iterative GRAPPA

Combining parallel imaging with CS would be desirable as the potential of both methods to accelerate the data acquisition process could be exploited. In CS, the data have to be irregularly sampled (e.g., randomly or Non-Cartesian) to obtain incoherent undersampling artifacts which can be eliminated during the CS reconstruction process. In this context, the idea for the presented implementation of Iterative GRAPPA arose from the need to reconstruct irregularly undersampled k-space data with a parallel imaging method. In contrast to conventional Cartesian parallel imaging from regularly undersampled k-space data, Iterative GRAPPA involves an iteration loop. Figure 48 clarifies the basic idea. A GRAPPA kernel is determined by a fully-sampled calibration dataset or additional ACS lines. This kernel is then applied to the undersampled data. The kernel used in the presented implementation of Iterative GRAPPA is a square one (e.g., 3x3 or 5x5 k-space points). The central point of this kernel is the target point for reconstruction and the surrounding k-space points are the source points. Throughout this thesis, the kernels used for Iterative GRAPPA reconstructions are termed according to their square dimensions in k-space (e.g., 3x3 or 5x5). The kernel runs over the whole k-space and reconstructs every single point at both sampled and non-sampled k-space locations. After that, the acquired data points are reinserted at their respective positions. By considering an undersampled GROG-gridded radial sampling pattern, it becomes immediately clear that reconstructing all k-space points in only one iteration is impossible. In Figure 49, the 3x3 GRAPPA kernel is at a position in outer k-space between two gridded radial projections. As seen, all source points are zero (i.e., not acquired). Thus, the target point is also zero. In other words, nothing has happened. The kernel, however, is close to one projection. After one iteration, some of the source points in the displayed kernel in Figure 49 would therefore no longer be zero. Thus, after a second iteration, the target point in the displayed kernel would also be $\neq 0$. With each iteration, more and more remote k-space points are reconstructed. However, the reconstruction remains imperfect since only a few source points contributed to the target point. Therefore, several iterations have to be performed in which all k-space

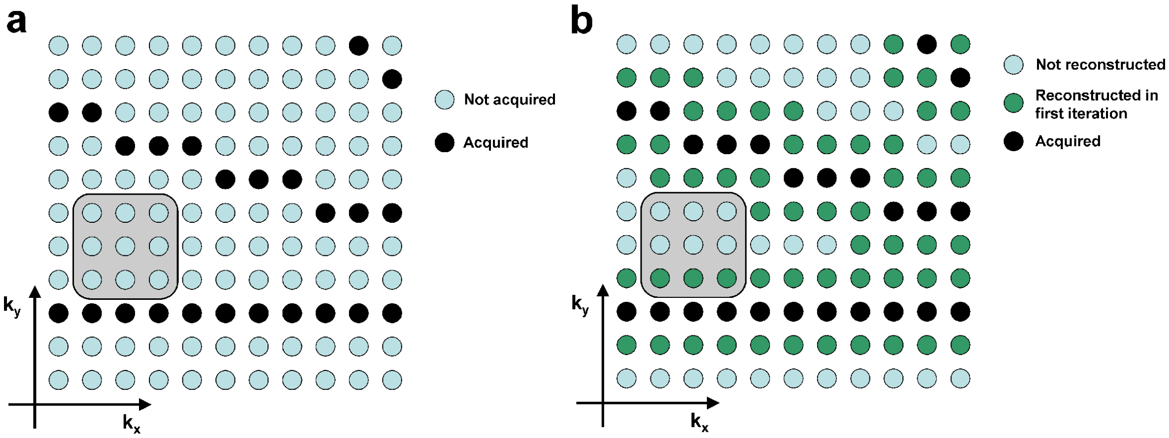

Figure 49: *Iterative GRAPPA is based on running a GRAPPA kernel (similar to the one displayed in Figure 48) multiple times over the whole k-space. In this example, the acquired data are pseudo-radial (e.g., after GROG-gridding). (a) In the first iteration, the GRAPPA kernel can end up in a region where no k-space points were acquired. In such cases, the target point will also be zero after GRAPPA reconstruction. (b) After the second iteration, k-space points in the immediate vicinity of the acquired pseudo-radial data are no longer zero. They are not, however, perfectly reconstructed due to the acquisition of only a few source points. Nonetheless, by repeating the GRAPPA procedure multiple times, increasing numbers of unacquired k-space points are more accurately reconstructed. The iterative process asymptotically converges to a solution of which the accuracy depends on the kernel size, the acceleration factor, the noise level and, of course, the sensitivity variations of the receiver array.*

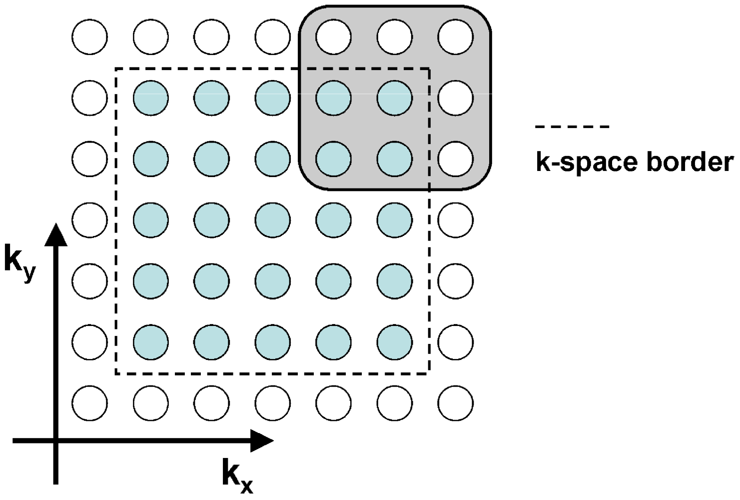

Figure 50: *If the GRAPPA kernel is used to reconstruct the outermost points in k-space, some of the source points of the GRAPPA kernel protrude the boundaries of k-space. Thus, points containing zeros are added at the k-space borders so the GRAPPA kernel can function.*

points are reconstructed using the respective GRAPPA kernel. After each iteration, the original acquired data are reinserted at the respective positions in k-space. The behavior of Iterative GRAPPA at k-space borders is explained in Figure 50.
The idea for Iterative GRAPPA was presented several years ago by Lustig et al. [96]. It is worth noting that the presented concept of Iterative GRAPPA was conceived independently from the abstract of Lustig et al.. Only after the presented method was already implemented, it was noticed that there had been an abstract on this topic before. The main difference to the Lustig abstract is that Iterative GRAPPA, as proposed in this dissertation, works completely on Cartesian data. Thus, gridding and de-gridding are unnecessary. Undersampled Non-Cartesian data can be effectively gridded to the Cartesian plane using GROG, leading to undersampled gridded Cartesian data. Please note that this is not possible using the standard convolution gridding [97]. By discarding the gridding and de-gridding steps necessary in [96], the proposed method becomes much more computationally efficient.

6.4 Simulations and experiments

This section first explores simulations performed to investigate and optimize the Iterative GRAPPA algorithm (e.g., the utilized kernel size). Additionally, it investigates the possibility to retrospectively undersample and reconstruct a radial cardiac cine dataset using optimized parameters for Iterative GRAPPA. The outcome conclusively demonstrates the applicability of Iterative GRAPPA with *in vivo* MR datasets.

6.4.1 Simulations

All simulations were performed in MATLAB® on the Shepp-Logan phantom. A 96x96 phantom was embedded into a 128x128 matrix of zeros and the maximum signal intensity was normalized to 1. Furthermore, this phantom was multiplied with the coil sensitivities of an eight-element one-ring head coil array. The sensitivities were derived using an analytic integration of the Biot-Savart equation. For the quantification of the reconstructions, the RMSE metric introduced and defined in Section 3.1 was used.

Similar to the CS simulations chapter (Chapter 3), there are a number of possible influences on the reconstruction quality that should be examined. For Iterative GRAPPA, these influences are:

- noise level
- acceleration factor
- sampling pattern
- size of the GRAPPA kernel
- choice of the regularization parameter μ in Moore-Penrose pseudo inverse

Additionally, to obtain accurate reconstructions, an appropriate number of iterations must be performed.
If not otherwise specified, the regularization parameter in the Moore-Penrose pseudo inverse was set to $\mu = 0.001$.

Iterative GRAPPA was developed in view of a possible combination with CS (see Chapter 7). Therefore, two different sampling patterns were investigated: A pseudo-random Cartesian

pattern and a radial sampling pattern. Both patterns have been shown to be beneficial for CS and, therefore, the performance of Iterative GRAPPA was tested on both patterns.
In the case of the pseudo-random Cartesian pattern, the outer k-space was undersampled according to the acceleration factor while the inner k-space (9x9 k-space points corresponding to 0.5% of the total k-space, see Figure 11) was densely sampled to guarantee convergence to the desired solution. The GRAPPA kernels were calibrated using the inner 51x51 points of the non-undersampled k-space in the (noisy) reference datasets. The choice of 51x51 was not optimized; however, it was arbitrarily selected in order to have enough calibration data to obtain accurate GRAPPA weights.
The radial acquisition scheme was mimicked by interpolating the Cartesian phantom data onto a radial grid with 240 projections and 256 readout points. Gaussian white noise with different standard deviations was added to each individual coil of the radial dataset. The (noisy) data were then undersampled so a certain number of projections could be used for reconstruction. The noisy and undersampled radial data were interpolated (by linear interpolation) along the readout direction to an oversampling factor of four, leading to 512 readout points. This procedure was suggested in [39] to reduce undesired artifacts after gridding noisy data and to improve the overall SNR. Next, the data were gridded using GROG [39, 40] onto a 128x128 Cartesian grid. The obtained undersampled Cartesian k-space was subject to reconstruction using Iterative GRAPPA. The GRAPPA kernels were calibrated similar to conventional GRAPPA kernels in regularly undersampled datasets. Subsequently, the central 51x51 k-space points of the gridded data from all 210 (as used in the subsection "Influence of the number of iterations") or 240 projections (subsections "Influence of the kernel size" and "Influence of the regularization parameter μ in the Moore-Penrose pseudo inverse") of the respective (noisy) radial dataset were used for calibration.

Influence of the number of iterations

Simulations were performed on the Shepp-Logan phantom with three different noise levels (Cartesian: $\sigma = \{0, 0.001, 0.01\}$, radial: $\sigma = \{0, 0.1, 1\}$) which were thus added to each individual coil. The noise for the Cartesian data was added in image space to each individual channel, while for the radial data it was added to the radial k-spaces of each coil. The noise levels were arbitrarily chosen to obtain noise-free, medium-noisy and noisy data. A direct comparison of the results from both sampling patterns was not intended.
Three different acceleration factors were tested on each of the sampling patterns (af = 2, 4, 8 for the random Cartesian; 70, 30, and 15 projections for the radial pattern). This led to comparable numbers of utilized k-space points for the CS reconstruction (see Table 7). The undersampled data were reconstructed using a 5x5 GRAPPA kernel with 10, 15, 20, 25, and 30 iterations. In the end, this led to 45 different reconstructions per sampling pattern.

Influence of the kernel size

To study the influence of the kernel size on the reconstruction quality, simulations on the Shepp-Logan phantom were conducted. The simulations were performed using a pseudo-random Cartesian and a radial sampling pattern. The reconstructions using the pseudo-random Cartesian pattern were performed with several noise levels ($\sigma = \{0, 0.001, 0.0025, 0.0050, 0.0075, 0.010\}$), which were then added to each individual coil. The following acceleration factors were studied: $af = \{2, 3, ..., 9, 10\}$.
For the radial simulations, different noise levels were tested ($\sigma = \{0, 0.1, 0.5, 1.0, 2.5, 5.0\}$)

Pseudo-random Cartesian		**Radial**	
af	**k-space points**	**Projections**	**k-space points**
1	16384	**210**	12939
2	8201	**70**	8170
4	4152	**30**	3920
8	2108	**15**	2051

Table 7: *Utilized number of k-space points for the respective acceleration factors and projections. In the first row, the values for the fully sampled data are given. The number of k-space points per coil is displayed. As seen, the numbers are comparable across the two different sampling patterns.*

for the following numbers of projections: 120, 80, 50, 48, 40, 30, 24, 15, 10. In Table 8, the respective numbers of utilized k-space points and the corresponding Cartesian acceleration factors are given for the gridded radial sampling patterns.

All undersampled datasets were reconstructed using four different GRAPPA kernels: 3x3, 5x5, 7x7, and 9x9. The number of iterations was fixed to 20 since empirical testing revealed 20 iterations led to accurate reconstructions (see previous subsection and the results in Section 6.5.1).

Projections	**k-space points**	***af***	**Projections**	**k-space points**	***af***
120	11459	1.5	**30**	3938	4.2
80	8832	1.9	**24**	3051	5.4
50	7033	2.3	**15**	2043	8.0
48	5786	2.8	**10**	1395	11.7
40	4934	3.3			

Table 8: *Utilized number of k-space points per individual coil for the utilized numbers of projections with the radial sampling pattern. The corresponding Cartesian acceleration factor is also given.*

Influence of the regularization parameter μ in the Moore-Penrose pseudo inverse

The Tikhonov regularization parameter μ was introduced in Equation 6.8. The influence of this parameter on the reconstruction quality was studied for various noise levels (Cartesian: $\sigma = \{0, 0.001, 0.0025, 0.0050, 0.0075, 0.010\}$, radial: $\sigma = \{0, 0.1, 0.5, 1.0, 2.5, 5.0\}$) which were then added to each individual coil. A fixed acceleration factor $af = 3$ was used for the Cartesian pattern, while 40 projections were used for the radial pattern. The parameter μ was varied as follows: $\mu = \{0.1, 0.01, 0.001, 0.0001\}$. All reconstructions were performed using 20 iterations and a 5x5 kernel.

6.4.2 Radial cardiac cine dataset

A radial cardiac cine dataset of a healthy volunteer was acquired on a 1.5 T clinical MR scanner (Siemens Avanto, Siemens Healthcare, Erlangen, Germany) using a 32-channel coil array (Rapid Biomedical, Rimpar, Germany). The following imaging parameters were applied: 2D radial bSSFP sequence, $\alpha = 72^\circ$, $T_R = 47.36$ ms, $T_E = 1.48$ ms, FOV 300x300 mm^2. The dataset contained 21 timeframes with 224 projections and 192 readout points, oversampled by factor of 2 in the readout direction (384 points). As suggested in [39], the data were interpolated along the readout direction to an oversampling factor of four, leading to 768 readout points. This fully sampled interpolated dataset was gridded using GROG [39, 40] to a 192x192 matrix and used as reference. The non-interpolated radial dataset was then retrospectively undersampled. The applied sampling scheme mimicked an interleaved acquisition scheme with the following numbers of projections: 56, 32, 28, and 14. The undersampled projections were also gridded using GROG [39, 40] after interpolation along the readout direction to 768 points. The regularization parameter μ was varied as follows: $\mu = \{0.5, 0.1, 0.05, 0.01\}$. All timeframes were reconstructed using 20 iterations and a 5x5 kernel. A composite Cartesian k-space was obtained by summing the radial cine dataset along the timeframe dimension and gridding the dataset using GROG. This composite k-space was used to calibrate the GRAPPA kernel. Similar to the simulations, the central 51x51 k-space points of the composite k-space were used for calibration.

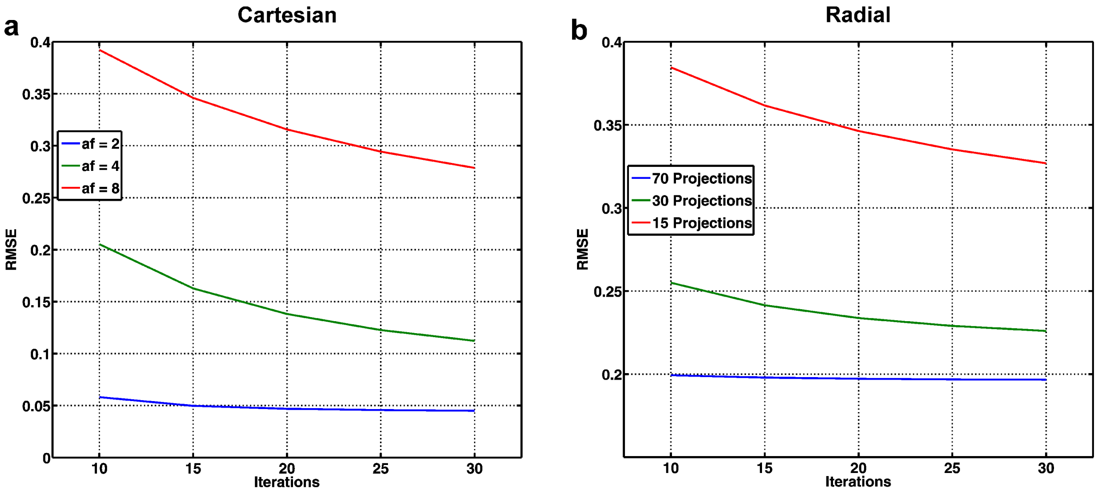

Figure 51: *RMSE plots for the pseudo-random Cartesian (a) and radial (b) sampling patterns. The noise level of each individual coil was* $\sigma = 0.001$ *(Cartesian) or* $\sigma = 0.1$ *(radial). The RMSE decreased with increasing number of iterations and trended towards an asymptotic limit.*

6.5 Results

6.5.1 Simulations

Influence of the number of iterations

In Figure 51, the RMSE is displayed for noisy data corrupted with $\sigma = 0.001$ (Cartesian) and $\sigma = 0.1$ (radial) in each individual coil. The results obtained with a pseudo-random Cartesian sampling pattern are shown in Figure 51a, while those obtained with the radial patterns are shown in Figure 51b. For both diagrams, the RMSE decreased with increasing number of iterations. The behavior of the curves in both diagrams indicates an asymptotical limit for each acceleration factor toward which the reconstruction quality in terms of RMSE

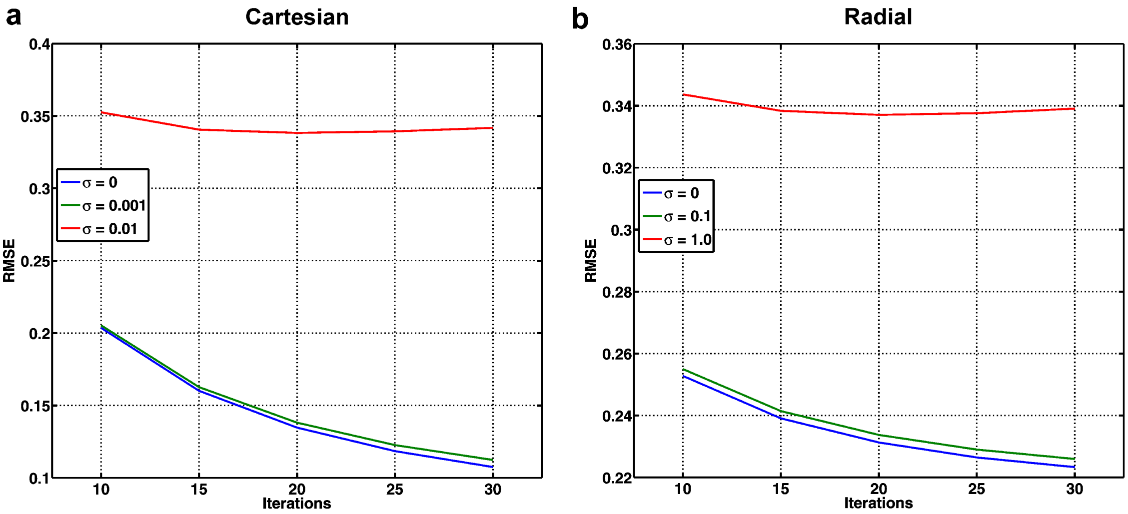

Figure 52: *RMSE plots for the pseudo-random Cartesian (a) and radial (b) sampling patterns. The acceleration factor for the Cartesian pattern was* af $= 4$*, 30 projections were used for the radial pattern. For both patterns and the two lower noise levels, the RMSE decreased with increasing number of iterations. For the high noise levels (*$\sigma = 0.01$ *for Cartesian,* $\sigma = 1.0$ *for radial), the RMSE ran through a minimum for 20 iterations.*

trends.

Figure 52 shows the RMSE in dependency on the noise level for a given acceleration factor (Cartesian: $af = 4$; radial: 30 projections). Similar to Figure 51a, for the two lower noise levels, the RMSE decreased with increasing number of iterations for both sampling patterns. However, this situation changed for high noise levels, in which the RMSE ran through a minimum at 20 iterations. Nonetheless, the difference between the optimal RMSE for 20 iterations and the non-optimal RMSE for other numbers of iterations is barely visible in the reconstructed images (data not shown).

In Figures 53 to 55, exemplary reconstruction results are shown to underline the meaning of the RMSE plots in Figure 51. The noise level in each coil in these figures was $\sigma = 0.001$ for the Cartesian pattern and $\sigma = 0.1$ for the radial pattern.

In Figure 53, the reconstruction quality for the radial reconstructions from 30 projections almost stagnated after 10 iterations and the benefit of using more iterations is barely visible. Close examination revealed that using more iterations led to a reduction of the "halo" around the outer ellipse of the Shepp-Logan phantom. Furthermore, the smaller ellipses in the inner phantom were better defined.

In Figure 54, results from the pseudo-random Cartesian dataset with $af = 2$ are displayed. In this example, the phantom is accurately recovered for all investigated numbers of iterations. Similar to the findings in Figure 53, the visible image quality almost stagnated when more than 10 iterations were applied. Nonetheless, more iterations reduced the (barely visible) remaining artifacts. This effect was clearer when the acceleration factor was increased.

Figure 55 exhibits results for data undersampled by a factor of 4 using the pseudo-random Cartesian pattern. Even though the phantom was not accurately recovered for all investigated iteration numbers, it can be seen that using more than 10 iterations led to increasingly improved results with reduced artifact level and a smoother phantom appearance. The benefit of using more than 20 iterations, however, is barely noticeable.

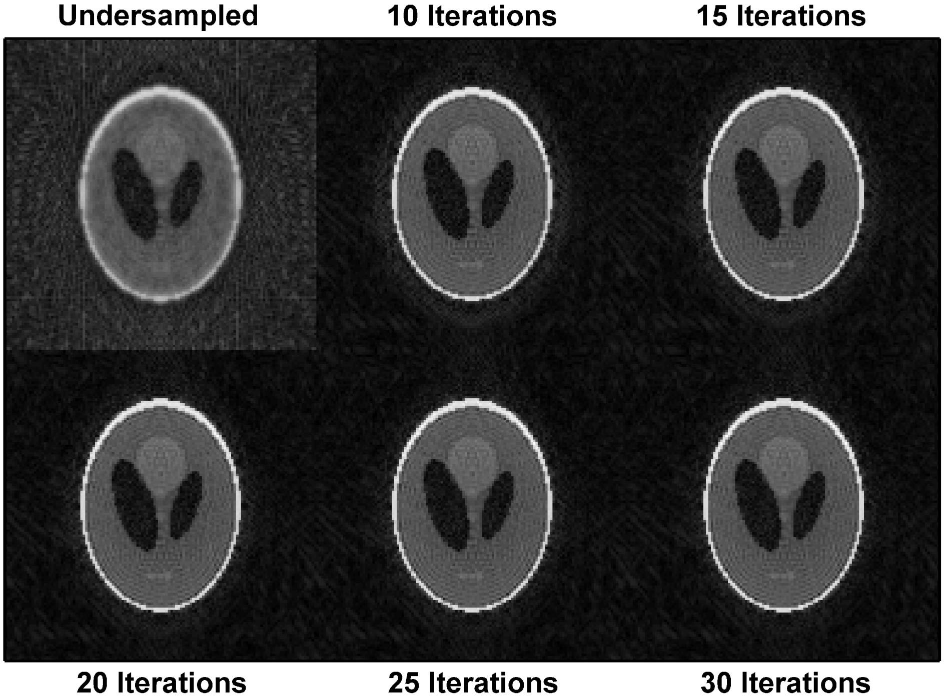

Figure 53: *Reconstruction results for the radial sampling pattern with 30 projections. The noise level in each individual coil of the phantom was $\sigma = 0.1$. As seen, the visible reconstruction quality barely increased when more than 20 iterations were applied (compare Figure 51b).*

Influence of the kernel size

Analogous to the previous section, two different sampling patterns were investigated. The results for the pseudo-random Cartesian sampling pattern are presented in Figures 56 to 58 and the results for the radial sampling pattern in Figures 59 to 61.

Figures 56 and 57 demonstrate that the RMSE decreased with increasing kernel size for all investigated noise levels and acceleration factors. However, the biggest gain in reconstruction quality was obtained from using a 5x5 instead of a 3x3 kernel. The improvement by using even larger kernels was only moderate except if high acceleration factors were used ($af \geq 6$). The visual improvement from using larger kernel sizes than 5x5 was only minor as seen in Figure 58. Furthermore, even for a moderate acceleration factor of 3 within 20 iterations, the reconstruction did not converge to the optimal solution.

The results of the simulations using the radial sampling pattern (Figures 59 to 61) confirmed the findings of the pseudo-random Cartesian simulations. However, as seen in Figure 61, even though the number of k-space points utilized for the Iterative GRAPPA reconstruction was comparable to that of a pseudo-random Cartesian acceleration factor of 3 (see Table 8), the visual reconstruction results were significantly better. Nonetheless, undersampling artifacts remained visible in the background.

Influence of the regularization parameter μ in the Moore-Penrose pseudo inverse

Figures 62 to 64 demonstrate the influence of μ on the reconstruction quality. For the calculated coil map, the optimal value for the regularization parameter in terms of RMSE was $\mu = 0.01$. This was true for all investigated noise levels. As seen, the radial reconstruction

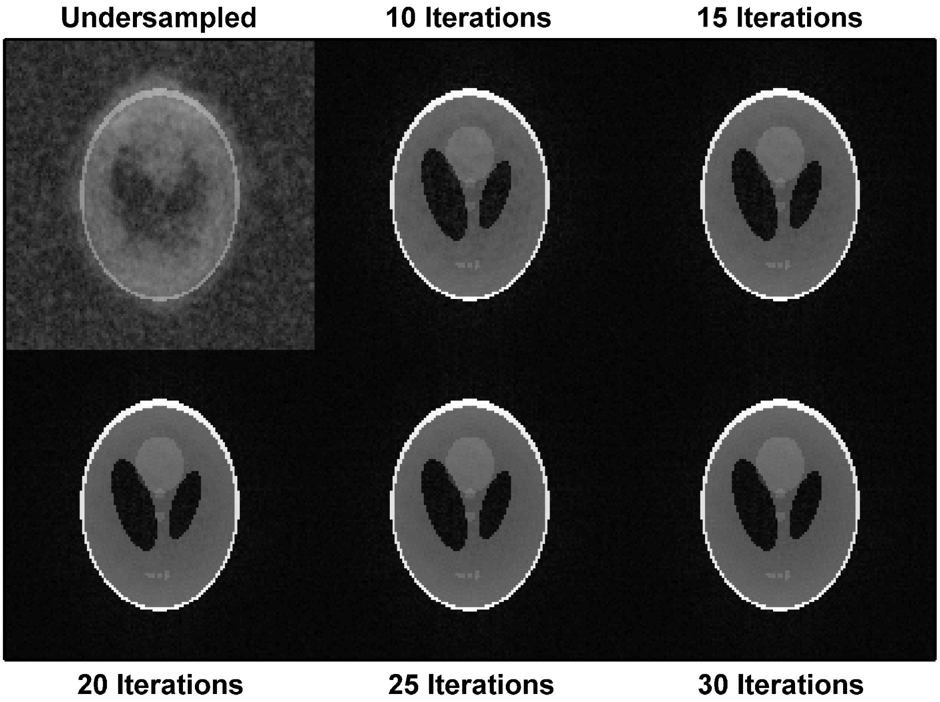

Figure 54: *Reconstruction results for the pseudo-random Cartesian sampling pattern with* af $= 2$. *The noise level in each individual coil of the phantom was* $\sigma = 0.001$. *Similar to Figure 53, the visible reconstruction quality stagnated if more than 10 iterations were applied (compare Figure 51a).*

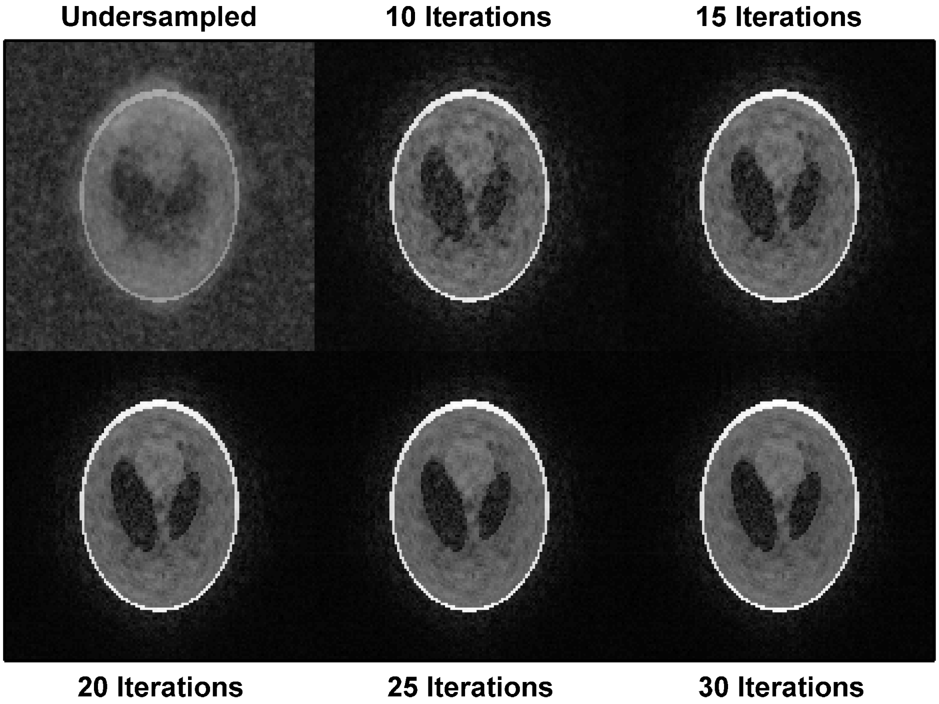

Figure 55: *Reconstruction results for the pseudo-random Cartesian sampling pattern with* af $= 4$. *The noise level in each individual coil of the phantom was* $\sigma = 0.001$. *The improvement in visible reconstruction quality after using more than 20 iterations is barely noticeable.*

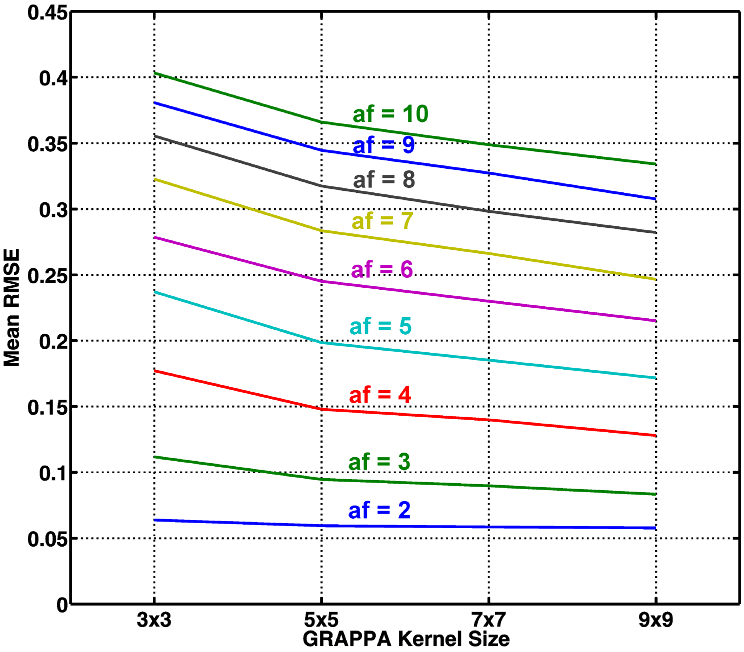

Figure 56: *Influence of the GRAPPA kernel size on the reconstruction quality in terms of RMSE. The results of the pseudo-random Cartesian sampling pattern with a noise level of* $\sigma = 0.0025$ *added to each individual coil are displayed. For all investigated acceleration factors, a larger kernel led to a decrease of the RMSE. Thus, a larger kernel led to better reconstruction quality in terms of RMSE for a given number of iterations (here: 20).*

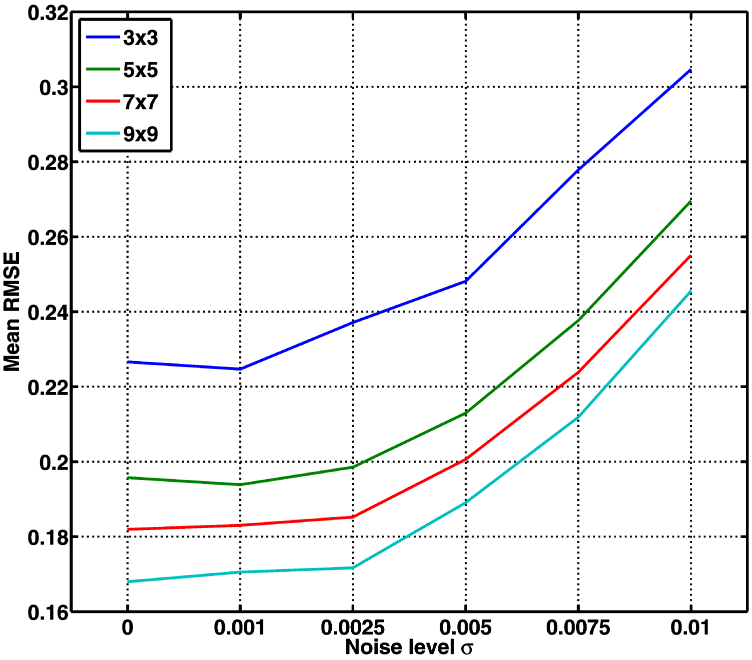

Figure 57: *For all investigated noise levels, the larger the GRAPPA kernel, the better the reconstruction quality in terms of RMSE. The displayed exemplary results were reconstructed for* af *= 5 using the pseudo-random Cartesian sampling pattern.*

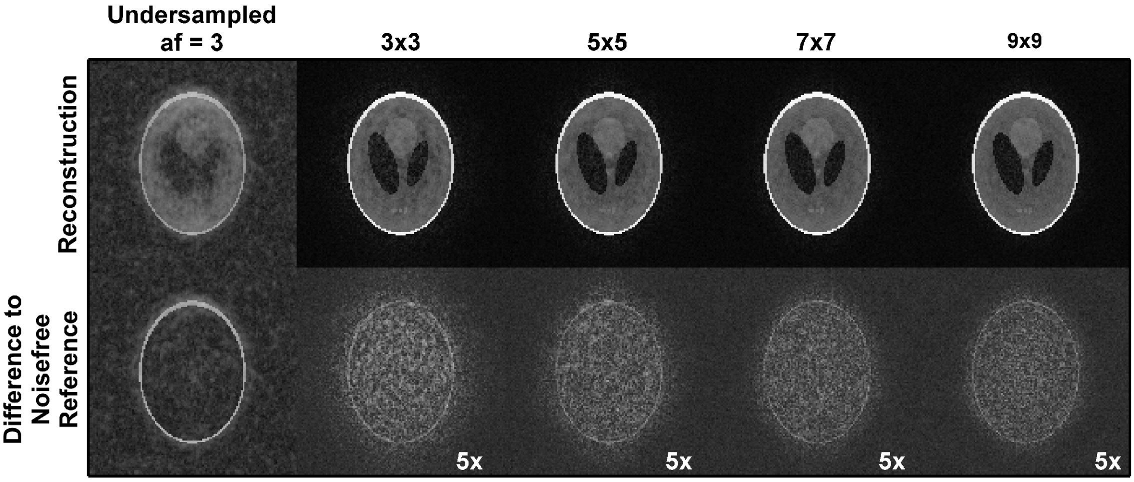

Figure 58: *Effect of different kernel sizes on the visible image quality for the pseudo-random Cartesian sampling pattern (*af = *3). Results for a noise level of* $\sigma = 0.0025$ *added to each individual coil are displayed. The effect of the GRAPPA reconstruction procedure is apparent. As seen, the quality improved significantly by moving from a 3x3 kernel to a 5x5 kernel. However, using an even larger kernel barely improved the visible image quality even though the RMSE decreased (see Figure 56 for* af = *3). All kernel sizes did not lead to the desired solution within 20 iterations. The difference images for the reconstructions from different kernel sizes are scaled by a factor of 5.*

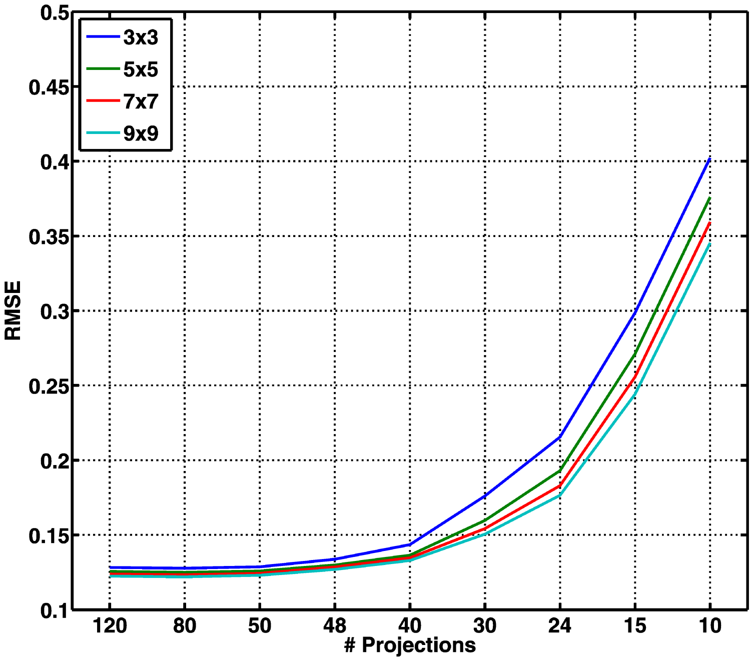

Figure 59: *Effect of the utilized number of projections and the GRAPPA kernel size on the reconstruction quality. Results from the radial sampling pattern for a noise level of* $\sigma = 0.5$ *added to each individual coil are displayed. For all investigated projection numbers, the RMSE remains lowest for the largest kernel. This confirms the findings of Figure 56.*

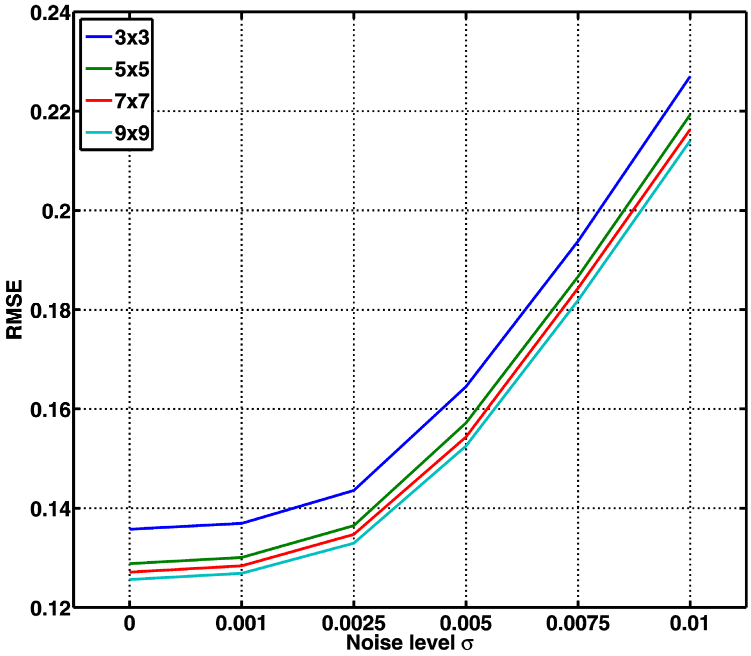

Figure 60: *For all investigated noise levels, the larger the GRAPPA kernel, the better the reconstruction quality in terms of RMSE. The displayed exemplary results were reconstructed from 40 projections using the radial sampling pattern. This figure confirms the results of Figure 57.*

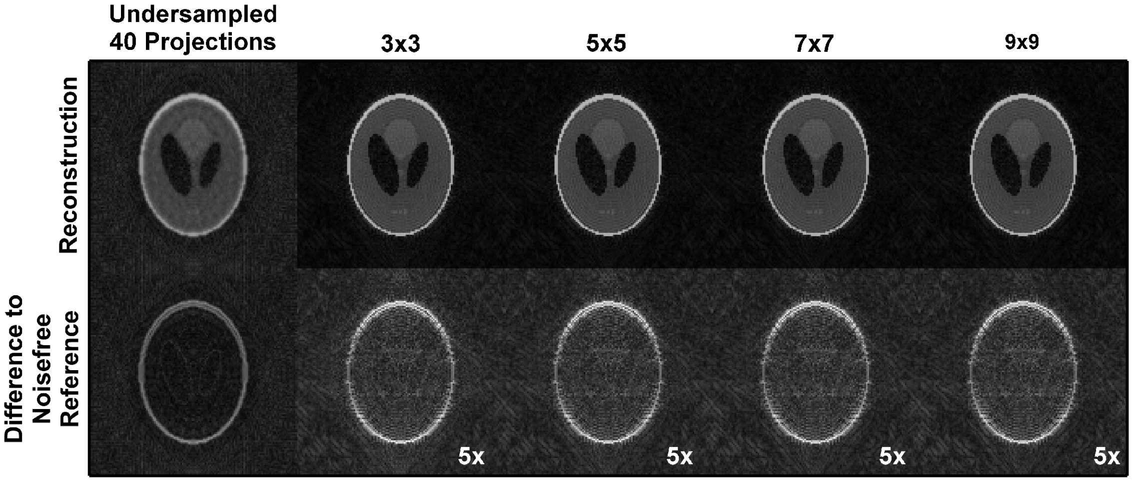

Figure 61: *Effect of different kernel sizes on the visible image quality for the radial sampling pattern (40 projections). The noise level in each individual coil was $\sigma = 0.1$. Again, the apparent effect of the GRAPPA reconstruction procedure is visible. As seen, the quality improved by moving from a 3x3 kernel to a 5x5 kernel. However, using an even larger kernel barely improved the visible image quality even though the RMSE decreased (see Figure 59). Similar to the previous two figures, the results from the Cartesian simulations are confirmed (Figure 58).*

for noise-free data and the smallest regularization parameter ($\mu = 0.0001$) diverged. In the reconstructed image, this is reflected at certain positions in the FOV as “hot spots“, which accumulate more and more image energy. In the end, the object itself can no longer be recognized (data not shown; however, the effect can be seen in Figure 67 for $\mu = 0.01$). This behavior cannot be observed in the Cartesian reconstructions.

Figures 63 and 64 clarify the meaning of different values for μ in terms of visible reconstruction quality. If the chosen value for μ was too large ($\mu \geq 0.1$), the Iterative GRAPPA reconstruction appeared blurred. The residual differences to the noise-free reference image were at a minimum for $\mu = 0.01$, while they slightly increased if $\mu < 0.01$. However, the effect of $\mu < 0.01$ on the visible image quality was marginal.

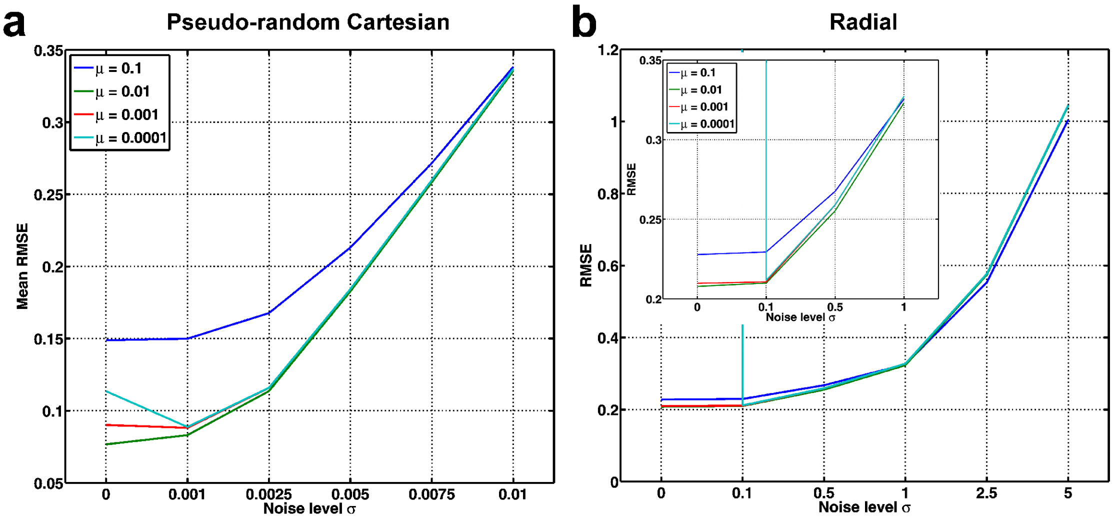

Figure 62: *Influence of the regularization parameter μ in the Moore-Penrose pseudo inverse (see Equation 6.8). The results for both investigated sampling patterns were similar, indicating that $\mu = 0.01$ is optimal for the given coil sensitivity map. (a) Results of the pseudo-random Cartesian sampling pattern for af = 3. On the y-axis, the mean RMSE values of the 30 reconstructions (per af and μ) is plotted. (b) Results of the radial sampling pattern for 40 projections corresponding to a Cartesian af = 3.3 (see Table 8). The embedded plot shows the interval from σ = 0 to 1 on the x-axis of the larger plot in a narrower RMSE scale. This was to clarify that $\mu = 0.01$ is, analogous to the pseudo-random Cartesian results, the optimal choice.*

6.5.2 Radial cardiac cine dataset

The Iterative GRAPPA reconstructions of the radial cardiac dataset led to successful results when the regularization parameter μ was properly chosen. In Figures 66 and 67, the influence of μ on the reconstruction quality can be seen. The position of the image pixels displayed in the x-t diagrams in Figure 66 is marked in Figure 65. When μ was properly adjusted, in this case $\mu = \{0.1, 0.05\}$, the dynamic information could be accurately recovered. This was because each individual timeframe was accurately reconstructed through Iterative GRAPPA (see Figure 66). When the value of μ was chosen too small ($\mu \leq 0.05$), the algorithm diverged and created “hot spots“ in the image that led to unacceptable image quality (see Figure 67, $\mu = 0.01$). However, if the regularization parameter was set too high ($\mu \geq 0.5$), Iterative GRAPPA barely improved the visual image quality compared to the undersampled and Fourier transformed image. When the undersampling was too high, Iterative GRAPPA failed to remove all undersampling artifacts. This can be best seen in

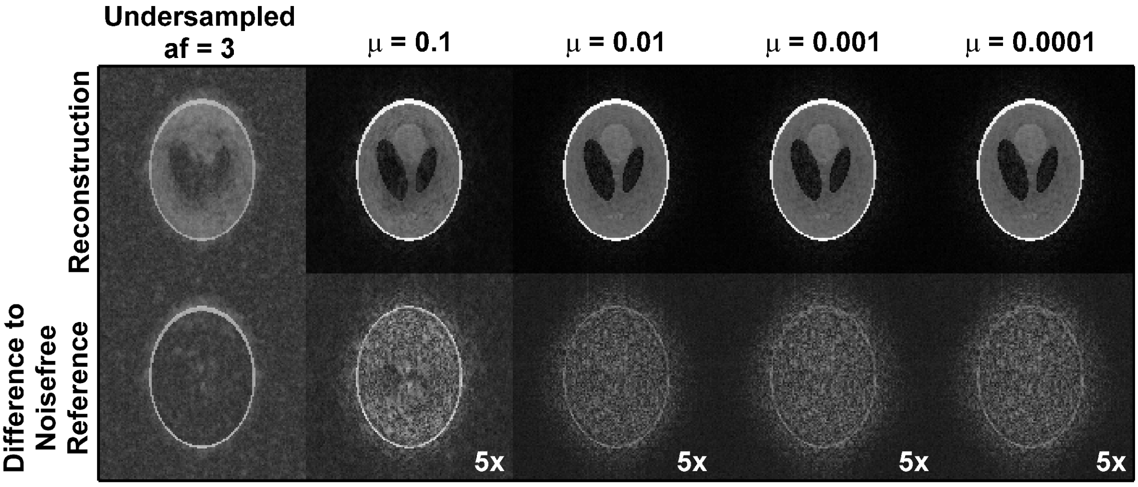

Figure 63: *Exemplary reconstruction results for the pseudo-random sampling pattern for* af = 3 *with a noise level* $\sigma = 0.001$ *added to each individual coil. Independent of the choice of* μ*, Iterative GRAPPA significantly improved the image quality compared to the undersampled and Fourier transformed image (first column). However, a properly adjusted regularization parameter* μ *led to improved reconstruction quality and, thus, less difference to the noise-free reference. For the given coil sensitivity map, choosing* $\mu = 0.01$ *led to the best results as already suggested in Figure 62. The difference images for the reconstructions from the investigated values of* μ *are scaled by a factor of 5.*

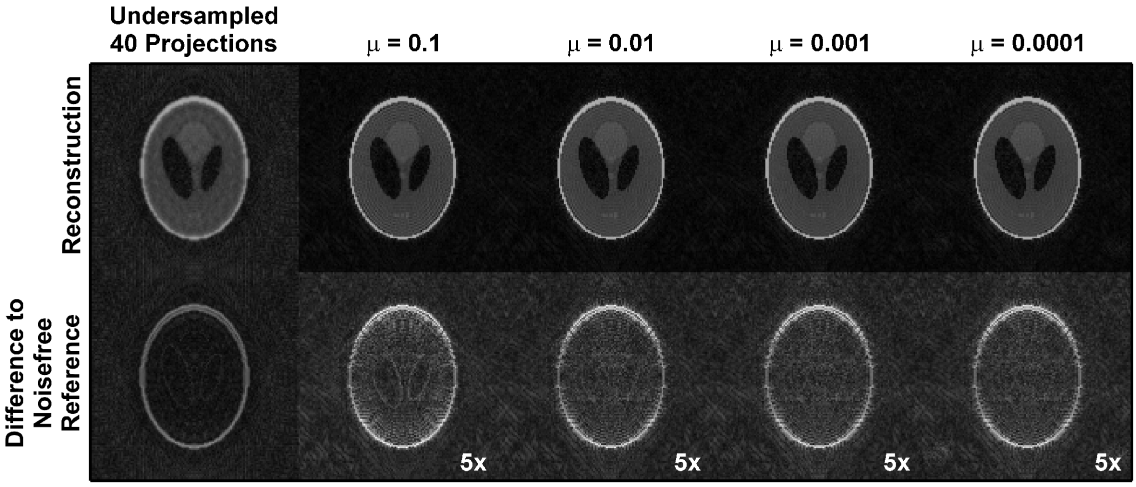

Figure 64: *Exemplary reconstruction results for the radial sampling pattern for 40 projections. The noise level in each coil was* $\sigma = 0.1$*. Again, the difference images for the reconstructions from the investigated values of* μ *are scaled by a factor of 5. The findings of Figure 63 are confirmed.*

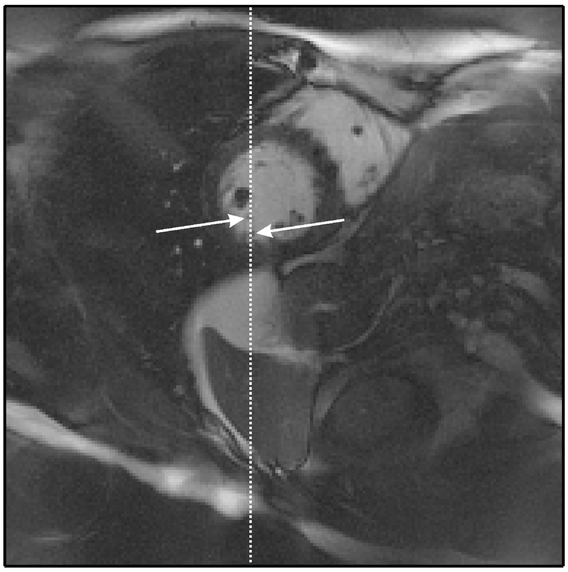

Figure 65: *Position of the image pixels displayed in Figure 66 as x-t diagrams. The arrows point to the position of two papillary muscles and correspond to the arrows in Figure 66.*

Figure 67, where the artifacts visible in the undersampled image are also noticeable in the Iterative GRAPPA reconstructions. Furthermore, the remaining undersampling artifacts can also be seen in the upper section of the x-t diagrams in Figure 66 for reconstructions obtained from 32 or fewer projections. By comparing the reconstructions for $\mu = 0.5$ and $\mu = \{0.1, 0.05\}$, it can be seen that the definition of the myocardium-ventricle border in Figure 67 was significantly improved for the lower values of μ. Noise-like reconstruction artifacts, however, remained visible in the region of the heart.

6.6 Discussion

6.6.1 How many iterations are necessary?

In a first step to investigate the characteristics of the proposed implementation of Iterative GRAPPA, simulations were performed. These were meant to help determine the necessary number of iterations for an accurate reconstruction. Two different sampling pattern were studied: A pseudo-random Cartesian and a radial sampling pattern. Both patterns can be beneficial when Iterative GRAPPA is combined with CS, the topic of which is covered in the following chapter of this thesis. However, the pseudo-random Cartesian pattern will be difficult to implement on scanners in most situations. Nevertheless, as shown in Chapter 4, when single point acquisition techniques such as CSI are applied, pseudo-random patterns are possible. Furthermore, in 3D imaging, the two phase-encoding directions can be undersampled using such a pattern. Radially acquired and undersampled multi-coil data can be gridded using GROG [39, 40] and then further processed in the Cartesian plane using Iterative GRAPPA. This is not possible when conventional convolution gridding is used since this technique distributes intensity to all Cartesian k-space points.

Both sampling patterns behaved similarly when the number of iterations was increased. For the two low noise levels, the RMSE of the reconstructed images decreased. However, although the RMSE decreased, the visible image quality did not significantly improve after 20 iterations (see also Figures 51 to 55). Moreover, in the presence of high noise, 20

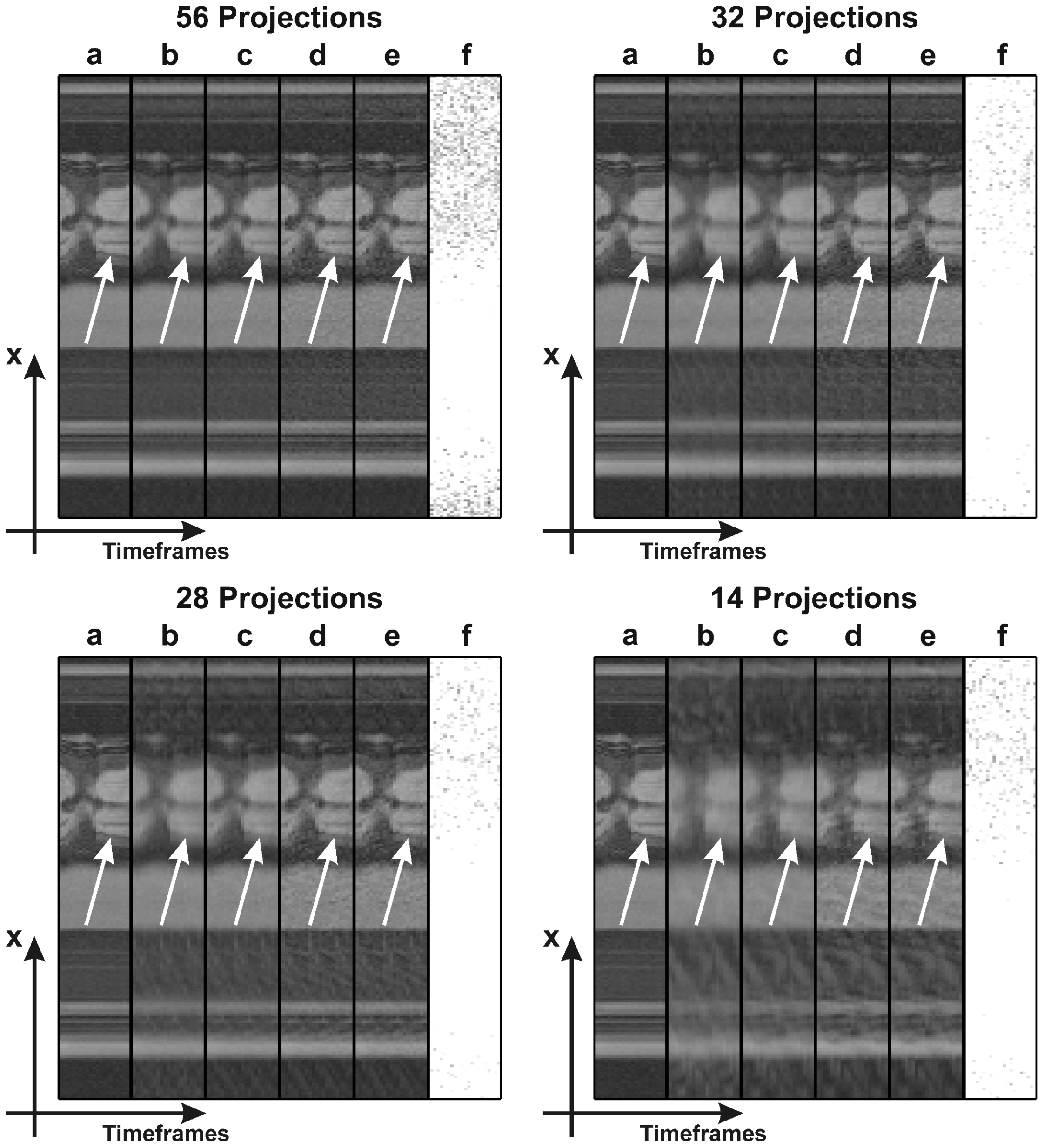

Figure 66: *Exemplary results from the radial cardiac cine dataset. The image pixels along the dotted line in Figure 65 are plotted for all timeframes. As seen for all shown data, when the chosen value for μ was too large ($\mu \geq 0.5$), the improvement compared to the undersampled dataset was marginal. However, when the parameter was chosen too small ($\mu \leq 0.01$), the Iterative GRAPPA algorithm diverged. A properly adjusted μ thus leads to accurate reconstructions. The arrows point to the position of the papillary muscles (two thin black lines). These can only be recognized for a properly adjusted μ. When the undersampling factor was too high ($\leq$ 32 projections), artifacts from the undersampling could be seen in the final image (upper part of the x-t diagrams). Legend: (a) Reference, (b) Undersampled, (c) $\mu = 0.5$, (d) $\mu = 0.1$, (e) $\mu = 0.05$, (f) $\mu = 0.01$.*

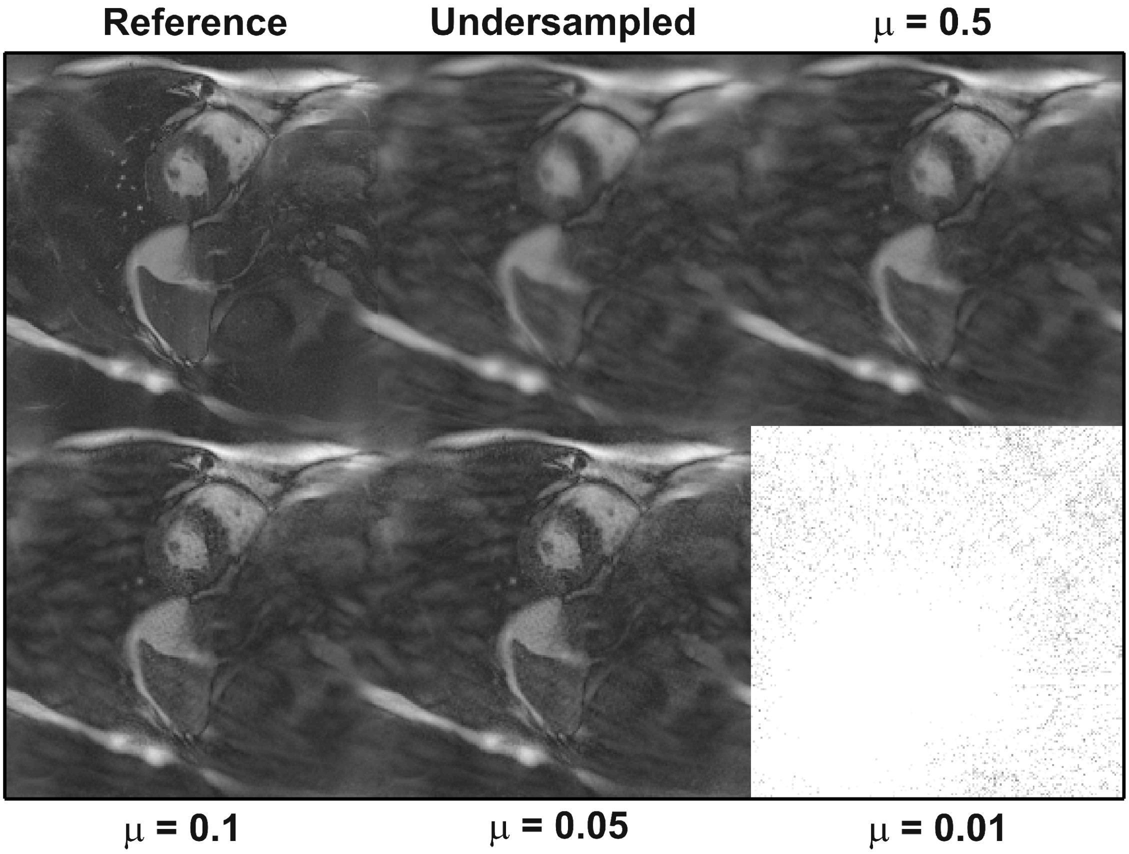

Figure 67: *Reconstruction results from an exemplary timeframe in mid-systole of the cardiac cine dataset. 28 projections were used for reconstruction. Again, as implied by Figure 66, only a properly adjusted* μ *led to accurate reconstructions (*$\mu = \{0.1, 0.05\}$*). However, noise-like reconstruction artifacts could be observed in the region of the heart. When the value of* μ *was chosen too small (*μ *= 0.01), the Iterative GRAPPA reconstruction diverged.*

iterations was the optimum in terms of RMSE. Based on these findings, the number of iterations was set to 20 for all further Iterative GRAPPA reconstructions in this thesis.

6.6.2 Influence of the kernel size

The GRAPPA kernel size determines how many surrounding k-space points contribute to the target point (see Figure 48). Therefore, the size of the kernel influences the reconstruction quality. This was reflected in the results displayed in Figures 56 to 61. However, the larger the kernel, the larger the computational burden. Thus, the necessity of a large kernel should be carefully considered. Using a 5x5 instead of a 3x3 kernel led to major improvements, all further Iterative GRAPPA reconstructions were performed using a 5x5 kernel. A kernel larger than 5x5 would have further improved the reconstruction quality in terms of RMSE. However, this benefit was barely reflected in the images (see Figures 58 and 61).

In both investigated sampling patterns there were regions in k-space where the density of sampled points was low. This led to “empty“ voids in k-space, as previously mentioned and pictured in Figure 49a. The larger the gaps, the worse the accuracy of the reconstructed k-space points in the void. In the final image, this was reflected as an insufficient removal of the incoherent artifacts introduced by the sampling pattern (Figures 58 and 61). However, the remaining artifact nature was different for each sampling pattern. The pseudo-random Cartesian pattern exhibited reduced but noticeable incoherent artifacts similar to those in the undersampled image. The radial reconstructions offered a “halo“ around the outer ellipse of the Shepp-Logan phantom. The intensity of this halo was reduced by using a larger kernel. Additionally, especially for higher undersampling factors, streaking artifacts typical for radial data acquisition could be visible (see Figure 61).

6.6.3 The role of the regularization parameter

The regularization parameter μ was introduced in Equation 6.8 and explained in greater detail in Section 6.2. Depending on the coil sensitivity variations, this parameter must be adjusted to enable a stable pseudo inverse. In the phantom simulations, an analytically derived coil sensitivity map was used. This map offered sufficient sensitivity variations and enabled μ values as low as 0.0001 for the Cartesian simulations. However, for noise-free data, the radial simulations diverged when $\mu = 0.0001$. Similarly, the radial cine dataset did not allow setting the regularization parameter lower than $\mu = 0.05$ because Iterative GRAPPA diverged in this case. However, as seen in the phantom simulations, when the chosen value of μ was not sufficiently small, the reconstruction results appeared blurry (see Figures 64, 66, and 67). From the results shown, the conclusion can be drawn that choosing $\mu \leq 0.05$ leads to accurate reconstructions without blurring.

6.6.4 Divergence of Iterative GRAPPA reconstructions

In Figures 62b and 67 ($\mu = 0.01$), diverging Iterative GRAPPA reconstructions could be identified. This behavior occurred when the regularization parameter μ was chosen too small. In the simulations, no divergence could be observed in the Cartesian reconstructions although the same coil map was used as that for the radial simulations. This, in connection with the fact that divergence also occurred in the radial cardiac dataset reconstruction, indicates that this behavior might be related to the radial acquisition scheme. However, further investigations regarding this topic should be undertaken.

6.6.5 Considerations regarding the g-factor

Noise-like reconstruction artifacts could be observed in Iterative GRAPPA reconstructions, especially in Figure 67. These artifacts might be related to a g-factor noise enhancement. Since, in contrast to conventional parallel imaging, the foldover artifacts due to undersampling were incoherently distributed in the FOV, the g-factor enhancement should have also been distributed all over the FOV. In the reconstructed images in Figure 67, these artifacts occurred mainly in the region around the heart. Thus, whether or not these noise-like artifacts are related to a g-factor noise enhancement should be further investigated.

6.7 Conclusion

A computationally efficient implementation of Iterative GRAPPA to reconstruct undersampled multi-coil MR data has been described. All reconstruction steps were performed on a Cartesian grid. In the case of undersampled Non-Cartesian data, GROG [39, 40] could be used to shift the data to the Cartesian grid. During the Iterative GRAPPA procedure, de-gridding onto the Non-Cartesian trajectory was unnecessary which decreased the computational burden. After each iteration step, the (gridded) acquired k-space points were reinserted at the respective positions to guarantee strict data consistency.

Radial sampling patterns led to better reconstructions than pseudo-random Cartesian patterns. This was due to the beneficial artifact nature of the radial trajectory. However, the higher the acceleration was, the larger the empty voids in k-space were. These voids were difficult to reconstruct and led to remaining artifacts in the reconstructions. Furthermore, noise-like reconstruction artifacts due to a possible g-factor enhancement could be observed.

The regularization parameter μ in the Moore-Penrose pseudo inverse must be appropriately adjusted to allow accurate reconstructions. In the shown experiments, choosing $\mu \leq 0.05$ resulted in accurate reconstructions for the phantom simulations as well as the cardiac dataset reconstruction.

The simulation results led to the choice of a 5x5 kernel and 20 iterations for the reconstruction process. These parameters were maintained for all further Iterative GRAPPA reconstructions in this thesis.

The proposed Iterative GRAPPA technique can be extended to 3D in a straight-forward way. However, this would lead to significantly increased reconstruction times.

7 CS-GRAPPA

The Iterative GRAPPA algorithm was introduced in the previous chapter of this thesis. It can be effectively combined with the Strict Data Consistency (DC) CS method to achieve higher acceleration factors than possible using each method individually. This was demonstrated in phantom simulations, which also determined the limitations of this Parallel Imaging-CS hybrid termed CS-GRAPPA.

7.1 Motivation

The potential of CS has already been demonstrated, investigated, and discussed in this thesis. Instead of reconstructing missing data points by incorporating intrinsically given physical information (as in Parallel Imaging), CS recovers data according to a model cost function based on prior knowledge about the desired object. As outlined in Chapter 2, it is known *a priori* that the object can be sparsely represented in an arbitrary mathematical basis. Parallel imaging, however, exploits the sensitivity profiles of the coil array that provide additional spatial information. In the reconstruction process, this inherent spatial encoding can be utilized. However, it would be desirable to incorporate the coil sensitivity information in the CS reconstruction algorithm to access the potential of parallel imaging and CS simultaneously. When combining these methods, however, the data must not only be in accordance with the measured k-space data points, but also with the coil sensitivity profiles. Through this, accurate reconstructions from even higher undersampling factors than possible with each individual method could be obtained.

This chapter describes a straight-forward extension of the Strict DC CS method by introducing an additional Iterative GRAPPA reconstruction step. Iterative GRAPPA was explained in the previous chapter and the Strict DC method was introduced in Sections 2.3.2 and 2.3.3. In CS-GRAPPA, the Strict DC method is applied coil-by-coil, followed by an Iterative GRAPPA step. This procedure is repeated several times, leading to an iterative algorithm. Empirical observation revealed that when the Iterative GRAPPA step is performed first and followed by a coil-by-coil CS step the CS-GRAPPA approach led to equivalent results. This is topic is revisited in Section 7.5.7.

Chapter 5 demonstrated that performing the CS step on a combined image of all coils in the receiver array (CS-CC) leads to significantly improved image quality compared to coil-by-coil CS. This is due to exploitation of the joint sparsity of the receiver coils. However, the idea of combining CS and parallel imaging that is described in this chapter was published earlier in 2009 in two conference abstracts [98, 99]. At that time, it was not yet known that CS on Combined Coils (CS-CC, developed in early 2010) is superior to coil-by-coil CS.

This leads immediately to the idea of combining CS-CC with Iterative GRAPPA. Unfortunately, as discussed in Section 7.5.5, this combination, utilizing the same idea outlined for CS-GRAPPA, failed to result in improved reconstructions compared to pure CS-CC.

The current chapter further demonstrates that CS-CC outperforms CS-GRAPPA in terms of reconstruction quality. Therefore, the main intention of this chapter is to present the idea behind CS-GRAPPA and to provide a possible initial point to develop a Parallel Imaging-CS hybrid.

7.2 CS-GRAPPA

CS-GRAPPA is a straight-forward extension of the Strict DC method for CS. To guarantee consistency with not only the acquired k-space data points but also the inherent coil sensitivity profiles, an additional Iterative GRAPPA step is introduced after the CS step. The resulting algorithm can be presented analogous to Section 2.3.3 in pseudo-code on the following page.

In contrast to pure Iterative GRAPPA, more than 20 iterations were performed. In fact, an Iterative GRAPPA step followed each CS step leading to an equal number of CS and Iterative GRAPPA steps. In the presented implementation, 420 iterations were performed. As demonstrated in the previous chapter, Iterative GRAPPA is not merely a coil sensitivity constraint. It reconstructs missing data points and, thus, contributes to an improved image reconstruction quality. The Iterative GRAPPA reconstructed data points are recovered by exploiting the inherent coil sensitivity variations while CS reconstructs the data according to a model cost function. In the CS step, deviations from the coil sensitivity profiles might occur in the reconstructed data points, which can subsequently be balanced and improved by the Iterative GRAPPA step.

```
Input
    x: Image
    y: Undersampled k-space data
    $\Psi$: Sparsifying transform
    FFT: Fast Fourier Transform (image space → k-space)
    IFFT: Inverse FFT (k-space → image space)
    Approximate norms using Equation 2.11
    Calibrate 5x5 GRAPPA kernel $G_w$ on fully sampled calibration dataset
        or ACS lines

Initialize
    $x_0$ = IFFT(y), normalize $x_0$ to 1, n = 1, $\epsilon = 1$

Iterations
    while $\epsilon > 10^{-4}$
------------------------------------------------------------------------
                    CS Step (Coil-by-Coil)

        for L = 1:$N_C$          $N_C$: Number of coils
            $d_L = \Psi^* \left( W_\Psi^{p-2} \Psi x_{n-1,L} \right)$      , $W_\Psi$ calculated from $x_{n-1,L}$
            Determine step size $t_L$
                such that $\| \Psi \left( x_{n-1,L} - t_L \cdot d_L \right) \|_p = min$
                (e.g. in MATLAB® using fminbnd, an exact line search)
            $x_{CS,L} = x_{n-1,L} - t_L \cdot d_L$
            Reinsert acquired k-space points at respective positions
                (strict data consistency)
        end
========================================================================
                    Iterative GRAPPA Step

        $x_n = \text{IFFT} \left( G_w \odot \text{FFT} \left( x_{CS} \right) \right)$          $\odot$: Convolution operation
        Reinsert acquired k-space points at respective positions
            (strict data consistency)
------------------------------------------------------------------------
        if n is multiple of 30
            $\epsilon = \epsilon \cdot 0.5$
        end
        n = n + 1
    end
```

W_Ψ is hereby defined as in Equation 2.13. W_Ψ^{p-2} means an elementwise raising to the power $p-2$.

7.3 Methods

To study the potential of CS-GRAPPA, simulations in MATLAB® on the Shepp-Logan phantom were performed. For this, a 192x192 Shepp-Logan phantom was embedded into a 256x256 matrix of zeros. This phantom was multiplied with a twelve-element one-ring head coil array. The sensitivities were derived using an analytic integration of the Biot-Savart equation. For the quantification of the reconstructions, the RMSE metric introduced and defined in Section 3.1 was used. Additionally, a visual inspection of the image quality was performed.

Four different reconstruction schemes were tested to investigate the potential of CS-GRAPPA:

- Coil-by-coil CS (CS-CbC)
- CS on combined coils according to Chapter 5 (CS-CC)
- Iterative GRAPPA
- CS-GRAPPA

A radial acquisition scheme was simulated. Therefore, the k-space of the phantom was interpolated onto a radial grid with 300 projections and 512 readout points along each projection. Gaussian white noise was added to each individual coil of the radial dataset with the following standard deviations: $\sigma = \{0, 0.1, 0.5, 1, 2.5, 5\}$. The (noisy) data were undersampled so that the following number of projections were used for reconstruction: 150, 100, 50, 30, 20, and 15. Before gridding, the data were interpolated (by linear interpolation) to an oversampling factor of 4 along the readout direction, leading to 1024 readout points. This procedure was suggested in [39] to reduce undesired artifacts after gridding noisy data. The undersampled and interpolated data were gridded using GROG [39, 40] onto a 256x256 matrix and then reconstructed. Table 9 gives the number of Cartesian k-space points per coil after GROG-gridding. Additionally supplied are the respective Cartesian acceleration factors that could be obtained by dividing the number of all Cartesian k-space points by the number of the Cartesian k-space points after GROG-gridding.

Projections	**k-space points**	***af***	**Projections**	**k-space points**	***af***
150	35439	1.8	**30**	8729	7.5
100	25695	2.6	**20**	5601	11.7
50	14070	4.7	**15**	4470	14.7

Table 9: *Number of k-space points used per individual coil after GROG-gridding for the corresponding number of projections used with the radial sampling pattern. Additionally provided are the corresponding Cartesian acceleration factors.*

For all CS methods including CS-GRAPPA, the standard CS implementation described in Section 2.3.3 was used with the nonconvex norm $p = 0.75$. The sparsifying transform Ψ was the Discrete Gradient transform ∇_{DG} as given in Equation 2.3. Iterative GRAPPA reconstructions were performed according to the findings of the previous chapter using a 5x5 kernel and 20 iterations. CS-GRAPPA also utilized a 5x5 kernel. All GRAPPA kernels were calibrated using the 51x51 central points of the gridded k-space from all 300 projections of the (noisy) datasets. The regularization parameter $\mu = 0.005$ in the Moore-Penrose pseudo inverse was chosen for the noise-free data, otherwise $\mu = 0.001$. This inconsistent choice for μ was necessary because the Iterative GRAPPA and CS-GRAPPA reconstruction of the noise-free data tended to diverge. Such behavior was also described in the last chapter.

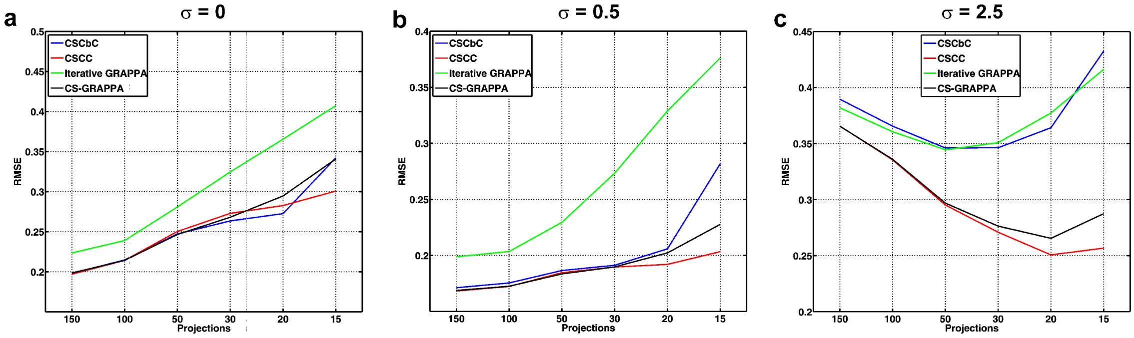

Figure 68: *Comparison of the four investigated reconstruction schemes in terms of RMSE. (a) With noise-free data, all techniques involving CS resulted in similar RMSE values. Iterative GRAPPA offered significantly higher RMSE values. (b) For moderately noisy data and high undersampling factors, it became clear that CS-CC and CS-GRAPPA are superior than CS-CbC and Iterative GRAPPA in terms of RMSE. CS-CC exhibited the best reconstruction quality in terms of RMSE. (c) For heavily noise-corrupted data, CS-CbC and Iterative GRAPPA resulted in significantly higher RMSE values than CS-CC and CS-GRAPPA. Again, CS-CC exhibited the best reconstruction quality in terms of RMSE. Furthermore, in contrast to (a) and (b), a minimum in the RMSE values can be observed for all four techniques.*

7.4 Results

Figure 68 shows RMSE plots of all four investigated reconstruction schemes for three exemplary noise levels. In Figure 68a, it can be seen that for noise-free data, all techniques involving CS exhibited similar RMSE values. Iterative GRAPPA resulted in significantly worse RMSE values. This was also true for Figure 68b; however, for higher undersampling factors (i.e., fewer projections used for reconstruction), the CS-CC method was superior to CS-GRAPPA and CS-CbC. Nonetheless, CS-GRAPPA led to lower RMSE values than CS-CbC. When data were heavily corrupted by noise (Figure 68c), CS-CbC and Iterative GRAPPA resulted in high RMSE values. CS-CC and CS-GRAPPA were significantly better techniques in terms of RMSE than the other methods, with CS-CC being superior to CS-GRAPPA at high undersampling factors. While, for noise-free and moderately noisy data (Figure 68a and b), the RMSE values increased with decreasing number of projections utilized for reconstruction, this changed for data with high noise levels. In those situations, the RMSE decreased for higher undersampling, reached a minimum and then increased again. For CS-CbC and Iterative GRAPPA, this minimum was observed for 50 and 30 projections while for CS-CC and CS-GRAPPA it occurred at 20 projections.

Figures 69 and 70 are meant to clarify the RMSE plots of the visual reconstruction quality. In Figure 69, reconstructions from moderately noisy data ($\sigma = 0.5$) are shown. As seen, the differences between the reconstructions and the noise-free Cartesian reference phantom were minimal using the CS-CC technique, followed by CS-GRAPPA. CS-CbC offered left residual intensity at the edges of the smaller ellipses in the inner Shepp-Logan phantom. Iterative GRAPPA showed a noise-like enhancement of signal intensity in the difference image, which was homogeneously distributed over the phantom. In all reconstructions, streaking artifacts from the radial undersampling could still be observed. Nonetheless,

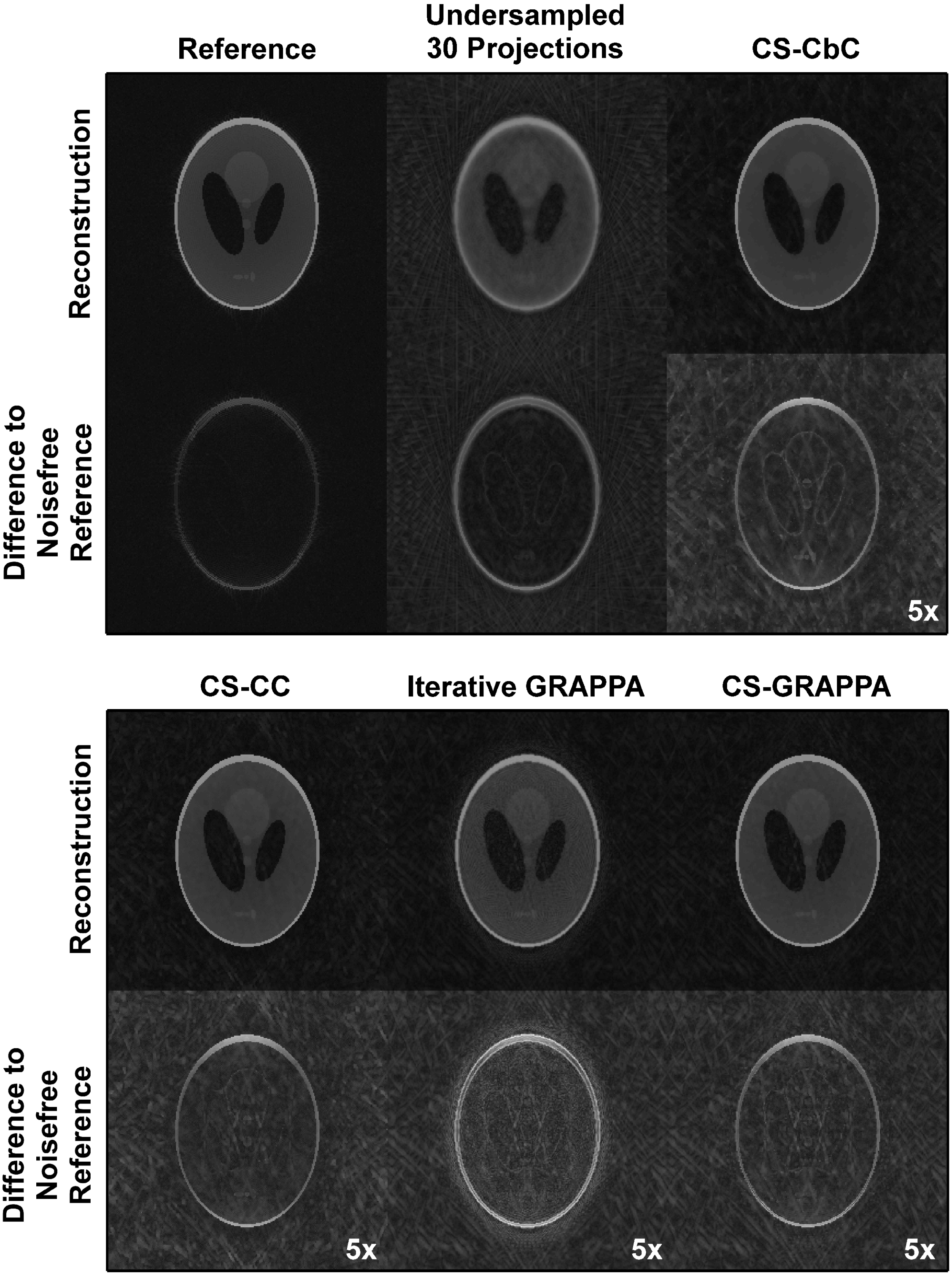

Figure 69: *Exemplary reconstruction results from noisy data with* $\sigma = 0.5$ *and 30 projections. The difference images of CS-CbC, CS-CC, Iterative GRAPPA, and CS-GRAPPA are scaled by a factor of 5. The residual image intensity in the difference images for CS-CbC was higher than for CS-CC. Iterative GRAPPA exhibited a noise-like enhancement of signal intensity in the region of the phantom that could no longer be observed in the CS-GRAPPA reconstruction. However, fewer streaking artifacts in the CS-CC image led to a smoother appearance of the phantom than seen with the CS-GRAPPA reconstruction.*

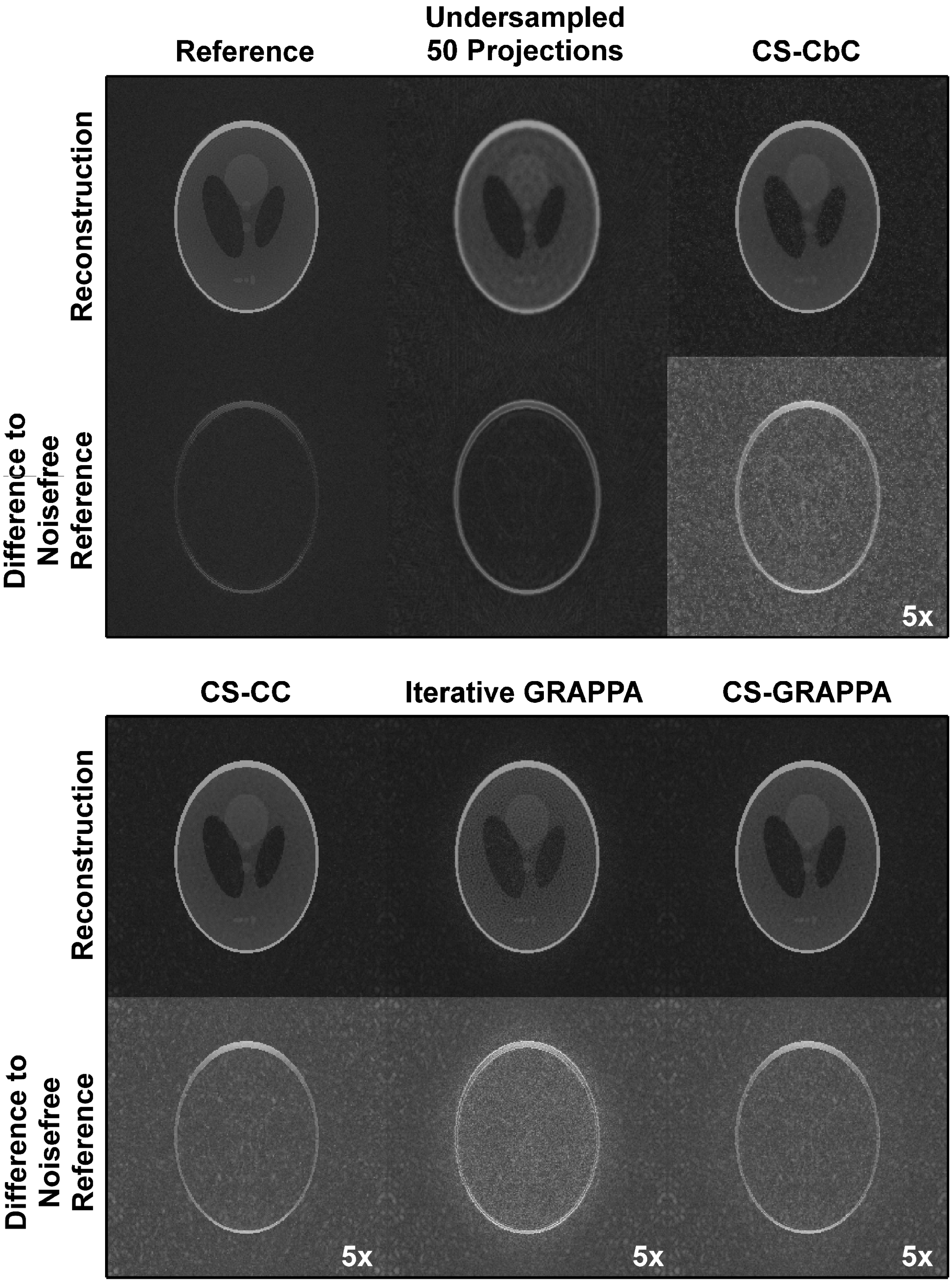

Figure 70: *Exemplary reconstruction results from heavily noise-corrupted data with* $\sigma = 2.5$ *and 50 projections. The difference images of CS-CbC, CS-CC, Iterative GRAPPA, and CS-GRAPPA are scaled by a factor of 5. The CS-CbC image exhibited spike artifacts visible throughout the FOV. CS-CC, Iterative GRAPPA, and CS-GRAPPA resulted in accurate reconstructions in which all features of the Shepp-Logan phantom could be recognized. However, Iterative GRAPPA exhibited a noise-like signal enhancement at the region of the phantom. This enhancement was significantly reduced in the CS-GRAPPA reconstruction. Here, the CS-CC and CS-GRAPPA image quality was equivalent and the reconstructions appeared less noisy than the Iterative GRAPPA image.*

only CS-CC and CS-GRAPPA allowed a proper distinction of the three small ellipses in the lower part of the Shepp-Logan phantom. While the visual reconstruction quality of CS-CC and CS-GRAPPA was very similar at first glance, a closer look revealed more streaking artifacts in the CS-GRAPPA reconstruction. Thus, the phantom in the CS-CC reconstruction appeared smoother than in the CS-GRAPPA reconstruction.
When the phantom had greater noise corruption (Figure 70, $\sigma = 2.5$), the CS-CbC technique exhibited spike artifacts that appeared all over the FOV. Furthermore, the CS-CbC reconstructed image from 50 projections looked blurry. CS-CC, Iterative GRAPPA, and CS-GRAPPA resulted in accurate reconstructions without blurring, where all features of the phantom could be recognized. However, similar to Figure 69, Iterative GRAPPA left a noise-like enhancement of signal intensity at the position of the phantom. This enhancement was significantly reduced in CS-GRAPPA. Nonetheless, a slight noise-like enhancement could still be observed. CS-CC again showed minimal residual signal intensity in the difference images. However, the visual reconstruction quality of CS-CC and CS-GRAPPA was equivalent.

7.5 Discussion

7.5.1 Comments on the difference images

The difference images in Figures 69 and 70 were computed between the undersampled GROG-gridded and differently reconstructed data, and the noise-free Cartesian reference phantom. In the reconstructed radial data, several sources can contribute to possible deviations from the original Cartesian phantom:

- The Cartesian k-space data must be interpolated onto a radial grid. Here, erroneous radial k-space points can be calculated which lead to differences in the final image after gridding even when the gridding procedure is optimal.
- The Non-Cartesian data in this chapter are gridded using GROG. GROG uses a GRAPPA-like weight set calibrated from the (noisy) radial data. Therefore, weight uncertainties are introduced by the interpolation procedure used to obtain radial data from the Cartesian reference. Furthermore, Gaussian noise also leads to errors in the weights calibration. Thus, the weights can be incorrect and lead to an imperfect shifting of the radial k-space points to the desired Cartesian k-space position. This, in turn, leads to increased deviations between the Cartesian reference phantom and the GROG-gridded, reconstructed image.
- The different investigated reconstruction schemes lead to imperfect recovery of missing k-space data. This is due to the first two aspects discussed above as well as the reconstruction algorithm itself. All techniques involving a CS step must introduce errors in the presence of noise, since the concept for CS is based on ideal noise-free data. All techniques requiring an Iterative GRAPPA step need to calibrate a GRAPPA kernel from gridded (noisy) radial data. Thus, the data are erroneous because of the first two aspects mentioned as well as noise. Therefore, the GRAPPA kernel is imperfect and leads to a g-factor noise enhancement and, hence, to increased differences between the final image and the noise-free Cartesian reference phantom.

The discussed problems can explain why differences in Figures 69 and 70 between the noise-free Cartesian reference phantom and even the fully sampled GROG-gridded reference occur.

7.5.2 Synergistic aspects of CS-GRAPPA

This chapter presents a hybrid technique of a CS and a parallel imaging method termed CS-GRAPPA. Figures 68 to 70 clearly demonstrate that this combination has synergistic effects, meaning that the hybrid CS-GRAPPA can accurately recover data from higher acceleration factors than possible using each method individually. It is important to note that, in contrast to the CS-CC method introduced in Chapter 5, the CS steps works on a coil-by-coil basis. The reason why CS-CC was not combined with Iterative GRAPPA is discussed in Subsection 7.5.5.

CS exploits the data sparsity in any arbitrary mathematical basis while parallel imaging recovers missing data points by combining coil sensitivity variations to mimic the missing spatial harmonics. CS-GRAPPA effectively reduces the noise enhancement of the Iterative GRAPPA step. This step is most likely related to a g-factor enhancement and was discussed in Section 6.6.5. Furthermore, spike artifacts that can occur in CS reconstructions when using strict data consistency are significantly reduced by the Iterative GRAPPA step. Thus, an overall improved image quality can be obtained using CS-GRAPPA.

7.5.3 On the different approaches for higher acceleration in CS-CC and CS-GRAPPA

In this subsection, the characteristics of CS-CC are first described and discussed. Next, CS-GRAPPA and the differences to CS-CC are discussed.

As mentioned in Chapter 5, CS-CC combines the individual coils by determining weights using an adaptive combination strategy [93]. Then, the CS step is performed on the combined SNR-optimized image and the CS-processed image is again distributed to the individual coils. This is done by applying the complex conjugate weights used to combine the images. The strict data consistency constraint can then be fulfilled by reinserting the acquired k-space data points at the respective positions in the k-space of the CS-processed image. This procedure implicitly exploits the joint sparsity of the data in all receiver channels as was already discussed in Chapter 5.

Using the aforementioned adaptive combination method leads to a SNR-optimized “single coil“ image. CS is SNR-sensitive, as demonstrated in detail in several chapters throughout this thesis. In each individual coil, the SNR decreases according to the coil sensitivity profile. The numerically generated 12-channel coil-map used for the simulations in this chapter exhibits a decreasing SNR in each coil toward the center of the FOV. The utilized sparsity transform for the Shepp-Logan phantom is the Discrete Gradient transform, which tends to smooth structures with low SNR or CNR. Thus, in each coil, structures toward the center of the FOV (i.e., the middle of the Shepp-Logan phantom) are increasingly smoothed with increasing noise level or undersampling factor. This can be clearly seen in Figure 70 for the CS-CbC method where the phantom appears blurred in the center.

Exploiting the joint sparsity of all receiver channels by adaptively combining the individual coil images has two advantages: First, the CS step is performed on a “single coil“ image with high SNR throughout the FOV, thereby reducing the smoothing due to the CS model. Second, after re-distributing the CS-processed image to each coil, all channels still offer the same information content of the object, without structure loss in certain coils due to insufficient SNR. Hence, no or reduced blurring should be visible as compared to the CS-CbC method. This behavior is confirmed in Figure 70, where the CS-CC technique does not exhibit noticeable blurring in the center of the phantom.

CS-GRAPPA exploits the sparsity given in each individual coil and the inherent spatial encoding of the receiver sensitivity profiles. While the coil-by-coil CS step suffers from the problems mentioned in the preceding paragraph, the parallel imaging step works on the whole dataset and reconstructs missing k-space data points by mimicking the missing spatial harmonics necessary to shift the source points to the target point. Since data are reconstructed based on given physical information, the negative effects of coil-by-coil CS are balanced by the Iterative GRAPPA step. This significantly improves the image quality compared to CS-CbC and can be seen in all presented data except for noise-free data. In the ideal noise-free case, no structures of the Shepp-Logan phantom will be lost in noise. This is true for all receivers, leading to results equivalent to the CS-CC method.
Interestingly, exploiting joint sparsity or combining coil-by-coil CS with parallel imaging leads to similar results with slight advantages for the CS-CC method. Therefore, a combination of joint sparsity CS and parallel imaging would be desirable (see also Subsection 7.5.5).

7.5.4 Relation to L_1-SPIRiT and ESPIRiT

L_1-SPIRiT [17–19] was introduced in 2009 as an alternative approach to combining an Iterative GRAPPA technique with CS. At the ISMRM 18^{th} Annual Meeting in 2010 in Stockholm, a computationally more efficient implementation of L_1-SPIRiT, termed ESPIRiT [92], was presented. First, the main characteristics of L_1-SPIRiT will be explained. Next, the extension to ESPIRiT will be described. Finally, the differences between these methods and CS-GRAPPA will be compiled and analyzed.

L_1-SPIRiT is, in principle, the same formulation as CS-GRAPPA. It combines a CS step with an Iterative GRAPPA step to effectively and synergistically combine CS and parallel imaging. The Iterative GRAPPA procedure was introduced as Iterative Self-consistent Parallel Imaging Reconstruction (SPIRiT, [17]). SPIRiT iteratively solves the following equation:

$$\min(\underbrace{\|Dx - y\|_2^2}_{\substack{\textit{Data}\\ \textit{consistency}}} + \lambda \underbrace{\|(G - \mathbb{1})x\|_2^2}_{\substack{\textit{Iterative}\\ \textit{GRAPPA}}}) \tag{7.1}$$

In this equation, D is a linear operator that relates the acquired (Cartesian or Non-Cartesian) data y to the reconstructed Cartesian image x. Therefore, it can involve gridding and re-gridding steps when used with Non-Cartesian data or the Fourier transform. The operator G is a GRAPPA kernel analogous to the kernels used in CS-GRAPPA (e.g., of size 5x5). This kernel is repeatedly convolved with the k-space data until all missing data points have been reconstructed. Thus, G is in fact created by a series of convolution operations and a Fourier transform to obtain Cartesian k-space data from the image x. λ is a regularization parameter.
The first term of Equation 7.1 is a data consistency term. The second terms states that, if G is applied to x and x is the correct solution, then subtracting $\mathbb{1}x$ from Gx must equal zero. Otherwise, this second term does not vanish. The operator G works solely on Cartesian data x.
The SPIRiT equation (Equation 7.1) can be solved by Conjugate Gradients or by Projections onto Convex Sets (POCS). As pointed out in [17], the POCS algorithm is applicable to Cartesian data and can be implemented so that, instead of the relaxed data consistency constraint in Equation 7.1, a strict data consistency holds (i.e. $\|Dx - y\|_2^2 \overset{!}{=} 0 \Leftrightarrow y = Dx$). As a further advantage, the determination of the regularization parameter λ is not required.

However, this chapter deals with Non-Cartesian radial data. Therefore, according to [17], the Conjugate Gradient algorithm is preferred.

CS can be implemented in a straight-forward way into the SPIRiT formulation, resulting in the following equation for L_1-SPIRiT:

$$\min(\underbrace{\|Dx-y\|_2^2}_{\substack{\textit{Data}\\\textit{consistency}}}+\lambda\underbrace{\|(G-\mathbb{1})x\|_2^2}_{\substack{\textit{Iterative}\\\textit{GRAPPA}}}+\rho\underbrace{\|\nabla_{DG}x\|_1}_{\textit{Sparsity}}) \tag{7.2}$$

Here, the l_1 penalty $\|\nabla_{DG}x\|_1$ relevant for this chapter is incorporated. However, any other regularization term better suited for the specific data can be used. l_1 penalties are well-known in the context of CS. Therefore, L_1-SPIRiT can be seen as a hybrid of parallel imaging and CS. The CS constraint in L_1-SPIRiT is also formulated for Cartesian data x. Thus, if Non-Cartesian data appear, they must only be dealt with in the data consistency constraint (first term of Equations 7.1 and 7.2).
Equation 7.2 can, similarly to SPIRiT, be solved either by a Conjugate Gradient algorithm or POCS.

ESPIRiT was presented for Cartesian 3D volumetric data and combines the individual coil images into one “single coil“ image prior to the CS step. After the CS step, the data is again re-distributed to the individual coils within a strict data consistency. In other words, the originally acquired data are reinserted at their respective positions.

There are various differences between CS-GRAPPA, L_1-SPIRiT and ESPIRiT. The most obvious difference is that two regularization parameters need to be appropriately adjusted in order to obtain optimal results with L_1-SPIRiT. As shown in Chapter 3, this can be cumbersome and time-consuming. Furthermore, it is generally questionable, whether or not the optimal parameters can be derived without an available reference image. In CS-GRAPPA, no regularization parameters must be determined. In ESPIRiT, the parallel imaging step is performed before entering the iteration loop. Since ESPIRiT obeys a strict data consistency constraint analogous to CS-GRAPPA, no regularization parameters must be determined.
Furthermore, due to the gridding and de-gridding operations necessary for L_1-SPIRiT and ESPIRiT, the computational burden in CS-GRAPPA is much lower and should allow for faster reconstructions than SPIRiT-based methods. CS-GRAPPA operates completely in the Cartesian plane after gridding undersampled Non-Cartesian data using GROG.
ESPIRiT exploits the joint sparsity along the coil direction as well as the sensitivity variations of the coil array. Thereby, it can be viewed as a more advanced hybrid of CS and parallel imaging. CS-GRAPPA and L_1-SPIRiT both perform the CS steps on a coil-by-coil basis.
In the SPIRiT paper [17], no comparison of L_1-SPIRiT to a joint sparsity CS algorithm was performed. The same held for the ESPIRiT abstract [92]. Therefore, it cannot be stated if L_1-SPIRiT or ESPIRiT are superior to pure joint sparsity CS. Furthermore, it is unknown if ESPIRiT is indeed a technique combining joint sparsity CS and parallel imaging. These aspects should be further investigated.

7.5.5 Combination of CS-CC and Iterative GRAPPA

The Iterative GRAPPA step can also be inserted after a CS-CC step. However, combining the CS-CC method with the Iterative GRAPPA step in the CS-GRAPPA scheme did not improve the reconstruction quality. Instead, it led to images with a degraded image quality compared to the pure CS-CC result (data not shown). Thereby, it was not possible to effectively simultaneously exploit joint sparsity and coil sensitivity variations.

7.5.6 Considerations regarding the g-factor

Similar to Section 6.6.5 in the Iterative GRAPPA chapter, a g-factor noise enhancement is to be expected in CS-GRAPPA reconstructions since parallel imaging is involved. By comparing the reconstructions and difference images of CS-CC, Iterative GRAPPA and CS-GRAPPA in Figure 70, a clear noise-like enhancement in the Iterative GRAPPA images can be observed. Similarly, a slight noise-like enhancement is noticeable in the CS-GRAPPA difference image. The phantom in the CS-GRAPPA reconstructed image appears noisier than the CS-CC reconstructed phantom. These observations might indicate a g-factor enhancement that is less pronounced in the CS-GRAPPA images than in the Iterative GRAPPA images. The reduced enhancement of the CS-GRAPPA images is caused by the smoothing characteristics of the utilized CS cost function. However, as with Section 6.6.5, further investigation on whether or not the observed enhancements are caused by the g-factor is necessary.

7.5.7 Sequence of the CS step and the Iterative GRAPPA step

In the presented implementation of CS-GRAPPA, the CS step was performed prior to the Iterative GRAPPA step. However, it was empirically observed that reversing this sequence led to equivalent results (data not shown). Further combinations are also imaginable. Examples include:

- Several CS steps followed by a single Iterative GRAPPA step
- Several Iterative GRAPPA steps followed by a single CS step
- Pure CS reconstruction according to either Chapter 5 or in a coil-by-coil fashion, followed by an Iterative GRAPPA reconstruction according to Chapter 6
- Reversed sequence of the last item

The first two items lead to arbitrarily many combinations. The goal of CS-GRAPPA was to develop a simple combination of parallel imaging and CS. Therefore, the method of alternating single steps of CS and Iterative GRAPPA was chosen. Thus, it cannot be stated if there exist more beneficial combinations. In sporadic attempts , in which the first two items of the above list were investigated, no improvement could be observed by either performing more CS/Iterative GRAPPA steps prior to each Iterative GRAPPA/CS step (data not shown).

The last two items pose the additional problem of deciding which k-space points should be kept for the Iterative GRAPPA/CS reconstruction that follows the initial CS/Iterative GRAPPA reconstruction. This problem requires deriving a quality criterion for deciding (without available reference) which k-space points are sufficiently accurately reconstructed to treat as actually “sampled“ in the second reconstruction (either Iterative GRAPPA of CS). This was not further investigated in this thesis.

7.6 Conclusion

This chapter introduced a straight-forward extension of the Strict DC CS method in which an Iterative GRAPPA reconstruction step was implemented after each CS step. This significantly improved coil-by-coil CS reconstruction since the Iterative GRAPPA step introduces a new consistency constraint on the reconstructed data. The data must not only be strictly consistent with the sampled k-space data, but also with the inherent coil sensitivity profiles. By exploiting the inherent spatial encoding of the coil array, missing data points can be iteratively reconstructed using parallel imaging. Noise-like signal enhancements observed in Iterative GRAPPA are significantly reduced by the CS step, while inconsistencies with the coil sensitivity profiles of the CS reconstructed data are balanced by the parallel imaging step. Thus, CS and parallel imaging are combined in an efficient, synergistic manner.

The CS method using a combined coil image was introduced in Chapter 5. This method proved to be slightly superior to CS-GRAPPA. While the visible reconstruction quality of CS-CC and CS-GRAPPA was usually comparable, the RMSE values were lower for the CS-CC reconstructions. Thus, exploiting joint sparsity in CS-CC appears to be more efficient than the proposed CS-GRAPPA method that is a hybrid of CS and parallel imaging. Nonetheless, CS-GRAPPA allows higher acceleration than possible using each individual method. Even though it appears less effective than CS-CC, it marks a possible starting point in developing a more advanced method combining CS and parallel imaging. Exploiting joint sparsity as well as the coil sensitivity variations to further push the limits for achievable acceleration in multi-coil MR imaging is highly recommended.

Summary and Conclusion

This thesis investigated the potential of Compressed Sensing (CS) applied to Magnetic Resonance Imaging (MRI). CS is a novel image reconstruction method that emerged from the field of information theory. The framework of CS was first published in technical reports in 2004 by Candès and Donoho [22, 23]. Two years later, the theory of CS was published in a conference abstract and two papers in “IEEE Transactions on Information Theory“ [14, 15, 24].

Candès and Donoho proved that it is possible, with overwhelming probability, to reconstruct a noise-free sparse signal from incomplete frequency samples (e.g., Fourier coefficients) [22, 23]. Hereby, it is assumed *a priori* that the desired signal for reconstruction is sparse. A signal is considered “sparse“ when the number of non-zero elements is significantly smaller than the number of all elements. Sparsity is the most important foundation of CS. When an ideal noise-free signal with few non-zero elements is given, it should be understandably possible to obtain the relevant information from fewer Fourier coefficients than dictated by the Nyquist-Shannon criterion.

As already pointed out, the theory of CS is based on noise-free sparse signals. As soon as noise is introduced, no exact sparsity can be specified since all elements have signal intensities that are non-zero. However, with the addition of little or moderate noise, an approximate sparsity that can be exploited using the CS framework will still be given. The ability to reconstruct noisy undersampled sparse MRI data using CS has been extensively demonstrated [16, 27, 29, 52–54, 60, 68, 74, 87, 95, 100, 101].

The theory of CS is not limited to Fourier imaging. However, the claim for incoherence between the sparse domain and the domain where data are acquired clarifies why MRI is a natural application for CS. Incoherence in this context means that the sparse signal does not have a sparse representation in the domain where data are acquired and vice versa. In MRI, the data acquisition domain is the Fourier domain (k-space). The sparse domain in MRI can be image space itself. Examples for an image-space sparse signal in MRI are contrast-enhanced angiography or ^{19}F datasets. Image space and k-space are incoherent since a complete correct representation of an MR image requires a k-space that is sampled according to the Nyquist-Shannon criterion. Because of the given incoherence between image space and k-space, MRI was early considered a possible application for CS [24] and a first detailed investigation of CS in MR soon followed [16].

Although most MR datasets are not sparse in image space, they can be efficiently sparsified by a sparsifying transform. It has been shown that the wavelet [16, 27] and the Discrete Gradient [16, 26, 27] transform are well-suited in this context. In this thesis, the data are either sparse in the image domain, after Discrete Gradient transformation, or after subtraction of a temporally averaged dataset from the data to be reconstructed (dynamic imaging).

For the successful application of CS in MRI, three important aspects must be considered:

1. **Sparsity**: A sparse representation of the desired signal is crucial for the success of CS.
2. **Sampling Pattern**: A sampling strategy that creates incoherent artifacts in case of

undersampling is necessary. In this thesis, either pseudo-random Cartesian sampling patterns with densely sampled inner k-space are utilized or a Non-Cartesian radial trajectory.

3. **Reconstruction**: An accurate reconstruction scheme must be found to "decompress" the "compressed" signal. For decompression, the reconstruction algorithm iteratively solves an optimization problem that models the expected image. Two different algorithms were tested in this dissertation: A Nonlinear Conjugate Gradient algorithm with a **Relaxed** **D**ata **C**onsistency constraint, which is referred to as the **Relaxed DC** method and a Gradient or Steepest Descent algorithm with a **Strict** **D**ata **C**onsistency constraint, which is referred to as the **Strict DC** method.

The aim of this thesis was to identify possible applications of CS to MRI. Hereby, a CS reconstruction scheme should be found that requires the least amount of user interference. In cases where actually undersampled data without a reference are recovered, user interaction in the reconstruction process introduces an unwanted subjective user dependent influence on the reconstruction result.

In Chapter 3, simulations on the Shepp-Logan phantom were performed to determine the characteristics of the Relaxed DC and Strict DC method. The regularization parameter, which occurs in the Relaxed DC method, was of special importance since the correct adjustment of this parameter has critical influence on the reconstruction quality. Hence, whether or not changing relevant imaging parameters requires the readjustment of the regularization parameter was tested. The investigated imaging parameters were the acceleration factor and the applied sampling pattern. Results indicated that the regularization parameter is sensitive to both parameters and, thus, requires a time-consuming empirical determination of the correct regularization parameter. Therefore, a detailed study of the dependencies of the regularization parameter on factors such as noise or the resolution was discarded.
It could be shown that the Strict DC method led to reliable, stable, and reproducible results. The reconstructions with the Strict DC method mostly resulted in significantly lower RMSE values than those obtained with the Relaxed DC method. The main reason for not considering the Relaxed DC method to reconstruct undersampled sparse data was the cumbersome and time-consuming adjustment of the regularization parameter. Thus, the Strict DC method was chosen to serve as the CS reconstruction scheme for the remainder of this thesis. However, a parameter that controls the accuracy of the reconstruction in the Strict DC method has to be adjusted and was, therefore, investigated. It could be demonstrated that, if the signal intensity in the initial guess of the Strict DC method is normalized to 1, this parameter can be appropriately selected. Thus, the Strict DC method can be used as a "Black Box algorithm" to accurately and reliably reconstruct undersampled MR datasets. All CS reconstructions in the remaining chapters of this thesis were either obtained with the Strict DC method or a joint sparsity extension of the Strict DC method called CS-CC (see Chapter 5).

In Chapter 4, the benefit of using CS to accelerate the data acquisition process or improve the apparent SNR in 2D and 3D ^{19}F Chemical Shift Imaging (CSI) was demonstrated. Datasets of exogenously applied ^{19}F marker substances are sparse in the image domain due to the negligible natural abundance of ^{19}F in living tissue. Since fully phase-encoded CSI was used to acquire the data, a pseudo-random Cartesian sampling pattern with densely sampled inner k-space was applied to undersample the k-space data. The spectral dimension was not undersampled, leading to fully sampled spectra at each acquired k-space point. The corresponding image spaces to each individual spectral point were CS reconstructed using the Strict DC method. First, simulations were performed on thresholded, fully sampled ^{19}F data and a numerically generated mouse phantom. The interdependencies between noise, the acceleration factor, and the norm on the reconstruction quality of the minimization problem were investigated. From these results, optimized parameters for the CS algorithm were derived and tested on undersampled *ex vivo* and *in vivo* datasets. A comparison of the reconstructed *ex vivo* images and an additionally acquired fully sampled reference dataset again proved the stability, reliability, and reproducibility of the Strict DC method. It was further demonstrated that CS can accelerate the data acquisition process up to a factor of eight in the presence of sufficient SNR. This allows the acquisition of spatially high resolved 3D ^{19}F CSI datasets on timescales that can be tolerated with *in vivo* measurements. When the SNR in the undersampled data is low, the image quality in the resulting images can be improved by averaging the undersampled data several times and then performing the CS reconstruction. This leads to reconstructions with improved apparent SNR. An exact SNR quantification is unfeasible at the moment since how to analytically describe the noise distribution after the non-linear CS reconstruction is unknown.

The Strict DC method hitherto used on single-receiver data was extended to a multi-receiver application in Chapter 5. Here, the potential of CS in dynamic radial cardiac imaging was investigated. To efficiently and optimally perform the CS reconstruction step in the presence of noise on multi-channel data, the Compressed Sensing on Combined Coils (CS-CC) method was developed. It involves a coil combination step to perform the CS step on a SNR-optimized, combined image of all channels. The strict data consistency was enforced as follows: After each CS step, the CS processed “single-receiver“ data were re-distributed to the individual channels of the coil array; then, the gridded acquired data were reinserted at the respective k-space positions. The method is computationally efficient since all data processing steps after gridding the radial data with GROG [39, 40] could be performed in the Cartesian plane. Thus, a computationally expensive de-gridding step was unnecessary. Since the dynamic data are not sparse in the image domain, the data were sparsified by subtracting the radial projections of the desired timeframe from a temporally averaged composite dataset. This led to sparse undersampled dynamic differences in the image domain that can be efficiently recovered by the CS-CC algorithm. Simulations revealed the interdependencies of the CS-CC method on noise, acceleration factor, and norm. From the outcome of the simulations, optimized parameters for the CS-CC algorithm were derived, transferred, and tested on a retrospectively undersampled radial cardiac cine set. Finally, the applicability of the proposed method was demonstrated on an undersampled radial cardiac real-time dataset. Accurate reconstructions from as few as 24 projections per timeframe could be obtained, resulting in a temporal resolution of 66 ms. The method is, in principle, capable of reconstructing undersampled datasets from patients with cardiac arrhythmia. Thereby, it might offer new insights in cardiac anatomy and function in this disease.

A computationally efficient implementation of Iterative GRAPPA was presented in Chapter 6. The development on the proposed Iterative GRAPPA implementation followed the need for a parallel imaging reconstruction strategy that can reconstruct missing data points in irregularly undersampled Cartesian datasets. It was also a prior step for combining parallel imaging and CS (CS-GRAPPA). In Cartesian parallel imaging, the phase-encoding direction is regularly undersampled, i.e, only every second, third, ... k-space line is sampled. When Non-Cartesian data are gridded using GROG, an irregularly undersampled k-space is obtained that cannot be recovered using conventional parallel imaging strategies such as SENSE [7] or GRAPPA [8]. In Iterative GRAPPA, missing data points are reconstructed by a weighted sum of the surrounding k-space points. This is achieved by applying a square GRAPPA kernel (e.g., 5x5), where the central point is the target point. Subsequently, the originally sampled and gridded data are reinserted to preserve strict data consistency. This procedure must be repeated several times since large gaps can occur in the undersampled gridded Cartesian k-space. By repeatedly (i.e., iteratively) running a square GRAPPA kernel over the undersampled k-space data, the missing points in these gaps can be iteratively reconstructed. The original concept of Iterative GRAPPA was presented in 2007 at the ISMRM [96] and incorporated gridding and de-gridding steps to switch between the Cartesian and the Non-Cartesian k-space. This is computationally demanding and can be avoided by gridding undersampled Non-Cartesian data using GROG. Hence, using the proposed implementation with GROG, all iterative reconstruction steps can be performed in the Cartesian plane. In simulations on the Shepp-Logan phantom, the dependencies of the method on the noise level, the acceleration factor, the sampling pattern, the size of the GRAPPA kernel, and a regularization parameter in the Moore-Penrose pseudo inverse were investigated. Results showed that undersampled radial data lead to better reconstructions than pseudo-randomly undersampled Cartesian data. Furthermore, the efficiency of the Iterative GRAPPA algorithm could be optimized by using a 5x5 kernel, which led to accurate reconstructions in 20 iterations. The regularization parameter in the Moore-Penrose pseudo inverse can be appropriately adjusted to a value that leads to accurate reconstructions for all investigated situations. As expected, the reconstruction quality decreased with increasing noise level and acceleration factor. These findings could be confirmed by a successful reconstruction from a retrospectively undersampled radial cardiac cine dataset.

Finally, Chapter 7 highlighted a straight-forward possibility to synergistically combine the Strict DC method with Iterative GRAPPA. This hybrid algorithm of coil-by-coil CS and parallel imaging is termed, CS-GRAPPA. It involves an Iterative GRAPPA reconstruction step after each coil-by-coil CS step. This introduces a new consistency constraint on the reconstructed data: Besides strict data consistency, the data must also be consistent with the inherent coil sensitivity profiles. Furthermore, as demonstrated in the previous chapter, Iterative GRAPPA itself recovers data from the intrinsic encoding properties of the coil array. Thereby, errors from the CS reconstruction step are balanced, while noise amplification due to the parallel imaging step is effectively reduced by the CS step. Simulations were performed on the Shepp-Logan phantom to investigate the potential of CS-GRAPPA. It was of special importance to determine whether or not CS-GRAPPA allows higher acceleration than coil-by-coil CS or Iterative GRAPPA can separately achieve. Furthermore, a comparison to the joint sparsity CS-CC method was performed, which was introduced in Chapter 5. The results demonstrated that CS-GRAPPA allows recovery of higher undersampled data than possible using either individual method. Furthermore, noise amplification of Iterative GRAPPA and spike artifacts from coil-by-coil CS were significantly reduced. However, the CS-CC method resulted in slightly superior image quality in terms of RMSE, while the visible image quality was comparable to CS-GRAPPA. Thus, exploiting the joint sparsity of the receiver coils in the CS-CC method seems to be more beneficial than combining parallel imaging with a coil-by-coil CS approach. Nonetheless, CS-GRAPPA marks a possible starting point to develop a more advanced combination of CS that exploits joint sparsity and parallel imaging.

Perspectives

Chapter 4 demonstrated that CS is beneficial for single-receiver data to which parallel imaging methods cannot be applied. Therefore, CS is an interesting tool to accelerate the acquisition of sparse data on high field NMR systems. On these systems, often only (home-build) single-receiver coils are available. Similar to ^{19}F imaging, other x-nuclei such as ^{13}C or ^{23}Na lead to sparse signal distributions due to negligible amounts of these isotopes in living tissue. However, the gyromagnetic ratio of ^{19}F ($\gamma_{^{19}\mathrm{F}} = 40.04$ MHz T^{-1}) is the second-largest following protons ($\gamma_{^{1}\mathrm{H}} = 42.58$ MHz T^{-1}). Since the equilibrium magnetization and, therefore, the receiver signal is proportional to γ^2 (see Equation 1.19), the SNR in other x-nuclei imaging experiments is expected to be significantly lower than with ^{1}H or ^{19}F (equal spin density assumed). In this context, the gyromagnetic ratio of ^{13}C is $\gamma_{^{13}\mathrm{C}} = 10.71$ MHz T^{-1} and that of ^{23}Na is $\gamma_{^{23}\mathrm{Na}} = 11.26$ MHz T^{-1}. Furthermore, similar to ^{19}F imaging, the spin density of other x-nuclei is very low, resulting in even lower SNR. However, as shown in Chapter 4, it might be possible to undersample the x-nuclei dataset and then average the undersampled data several times, thereby increasing the apparent SNR in the reconstructed images.

In this context, the artifact reduction in CS reconstructions of low SNR data using the Strict DC method is of great importance. Hereby, a first promising idea has been published [66].

One possibility to significantly improve the SNR is hyperpolarization. First applications of CS to spectroscopic imaging of hyperpolarized ^{13}C have already been demonstrated [52–54]. Another nucleus well-suited for hyperpolarization is ^{3}He. First successful CS reconstructions of accelerated hyperpolarized ^{3}He lung imaging have also been reported [102–106]. Hence, hyperpolarized x-nuclei are a promising application for CS and should be further investigated.

Using conventional proton MRI, Dental MRI could be an interesting application for CS. In Chapter 5 of Olga Tymofiyeva's PhD thesis [107], an overview of the potential of CS for Dental MRI is given. A more detailed investigation of this topic can be found in Sven Vaegler's Diploma thesis [100]. In both works, it could be demonstrated that CS can effectively accelerate Dental MRI acquisitions by a factor of 2 while preserving the required accuracy in the tooth surface reconstruction. However, undersampled datasets were acquired using a modified Spin Echo sequence. An accelerated Turbo Spin Echo sequence would be desirable and should be further developed.

There are some more general problems related to CS. One important question is whether or not it is possible to derive, analytically or reliably in an automated way, the regularization parameters necessary in many CS algorithms, without unwanted subjective user influence. In this thesis, an approach without regularization parameter was chosen and, therefore, this dissertation did not deal with that problem.

Another important problem is the quest for the optimal sparsifying transform for a specific application. In this thesis, the data were either sparse in the image domain, in the Discrete Gradient domain, or after subtraction from a reference dataset. However, when the data do not fit the assumed sparsity model, the reconstruction will be erroneous. For some applications there exist model-based CS approaches [101, 108, 109]. An interesting tool in this context might be a Principle Component Analysis.

Great efforts are concentrated on combining parallel imaging and CS. As already demonstrated in Chapter 7, a synergistic combination of both methods allows an accurate reconstruction of higher undersampled data than achievable by each individual method. Several publications have already dealt with this topic [110–112]. The goal should be to combine joint sparsity CS with parallel imaging in a robust, reliable, and user-independent way to efficiently exploit the sparsity of the dataset as well as the implicit spatial encoding of the utilized coil array.

Zusammenfassung

In der vorliegenden Arbeit wurde untersucht, welches Potential die Anwendung von Compressed Sensing (CS) in der Magnetresonanztomographie (MRT) hat. CS ist eine neue Bildrekonstruktionsmethode welche ihre Ursprünge in der Informationstheorie hat. Das Grundgerüst für CS wurde zuerst in zwei technischen Berichten von Candès und Donoho aus dem Jahr 2004 [22, 23] vorgestellt. Zwei Jahre später wurde die CS-Theorie in einem Konferenzbeitrag und zwei wissenschaftlichen Artikeln im Fachmagazin „IEEE Transactions on Information Theory“ veröffentlicht [14, 15, 24].

Candés und Donoho zeigten, dass es mit überwältigender Wahrscheinlichkeit möglich ist, ein rauschfreies dünn besetztes (engl.: sparse) Signal aus unvollständig vorliegender Frequenzinformation zu rekonstruieren. Da es keine gute deutsche Übersetzung für das englische Wort „sparse“ gibt, wird im weiteren Fortlauf der Zusammenfassung dieses Wort sowie das zugehörige Nomen „sparsity“ der englischen Sprache entlehnt. Für die Rekonstruktion eines Signals aus unvollständiger Frequenzinformation mittels CS ist eine wichtige *A-priori*-Annahme, dass das gewünschte Signal, welches rekonstruiert werden soll, sparse sein soll. Man spricht von sparsen Signalen, falls die Anzahl der Elemente mit Intensität größer Null signifikant kleiner als die Anzahl aller Elemente ist. Sparsity ist das wichtigste Fundament, auf dem CS basiert. Betrachtet man ein ideales rauschfreies Signal mit wenigen signaltragenden Elementen, dann leuchtet es ein, dass sich die relevante Information aus weniger Fourier-Koeffizienten als durch das Nyquist-Shannon-Theorem vorgegeben bestimmen lassen sollte.

Wie bereits erwähnt, basiert die CS-Theorie auf rauschfreien, sparsen Signalen. Sobald Rauschen auftritt, kann keine exakte Sparsity mehr bestimmt werden, da alle Elemente Signalintensitäten größer Null haben. Falls jedoch nur wenig oder moderates Rauschen hinzugefügt wird ist immer noch näherungsweise eine Sparsity gegeben, die mit Hilfe von CS ausgenutzt werden kann. Es wurde bereits ausführlich gezeigt, dass es möglich ist, verrauschte, unterabgetastete, sparse MR-Daten mittels CS zu rekonstruieren [16, 27, 29, 52–54, 60, 68, 74, 87, 95, 100, 101].

Die CS-Theorie ist nicht auf Fourier-Bildgebung beschränkt. Die notwendige Inkohärenz zwischen der sparsen Domäne und der Datenakquisitionsdomäne offenbart jedoch, warum die MRT eine natürliche Anwendung für CS ist. In diesem Zusammenhang bedeutet Inkohärenz, dass das sparse Signal in der Datenakquisitionsdomäne nicht sparse repräsentiert werden kann und anders herum. In der MRT ist die Datenaufnahmedomäne der Fourier-Raum (k-Raum). Die sparse Domäne kann der Bildraum selbst sein. Beispiele für sparse Signale im Bildraum sind die kontrastmittelgestützte Angiographie oder ^{19}F-Datensätze. Bildraum und Fourier-Raum sind inkohärent, da eine artefaktfreie Darstellung eines MR-Bildes einen k-Raum benötigt, der nach dem Nyquist-Shannon-Theorem abgetastet worden ist. Wegen der gegebenen Inkohärenz zwischen Bild- und k-Raum wurde die MRT schon früh als mögliche Anwendung für CS vorgeschlagen [24], eine erste detaillierte Untersuchung von CS in der MRT folgte kurz darauf [16].

Die meisten MR-Datensätze sind nichtsdestotrotz nicht-sparse im Bildraum, können allerdings durch eine sog. Sparsifizierungstransformation effektiv sparsifiziert werden. Als besonders geeignet haben sich in diesem Zusammenhang die Wavelet-Transformation [16, 27]

und die Diskrete-Gradienten-Transformation [16, 26, 27] erwiesen. In der vorliegenden Arbeit sind daher die Daten entweder im Bildraum sparse, nach einer Diskreten-Gradienten-Transformation oder nachdem bei dynamischen Daten ein zeitlich gemittelter Datensatz von den zu rekonstruierenden Daten abgezogen worden ist.
Drei wichtige Aspekte müssen berücksichtigt werden, um CS erfolgreich in der MRT anwenden zu können:

1. **Sparsity**: Eine sparse Darstellung des gewünschten Signals ist zwingend notwendig für eine erfolgreiche Anwendung von CS.
2. **Abtastschema**: Eine Abtaststrategie ist notwendig, die zu inkohärenten Artefakten bei Unterabtastung führt. In dieser Arbeit wurde entweder ein pseudo-zufälliges kartesisches Abtastschema mit vollständig aufgenommenem innerem k-Raum oder die nicht-kartesische radiale Trajektorie verwendet.
3. **Rekonstruktion**: Ein akkurates Rekonstruktionsschema muss gefunden werden, um die „komprimierten“ (engl.: compressed) Signale zu „dekomprimieren“ (engl.: decompress). Der Rekonstruktionsalgorithmus löst zur Dekomprimierung ein Optimierungsproblem, welches das zu erwartendende Bild modelliert. Zwei unterschiedliche Algorithmen wurden in dieser Dissertation getestet: Ein nicht-linearer Algorithmus, der die Methode der konjugierten Gradienten mit gelockerter Datenkonsistenzbedingung (**Relaxed** **D**ata **C**onsistency) anwendet. In dieser Arbeit wird dieses Verfahren als **Relaxed-DC**-Methode bezeichnet. Der zweite Algorithmus ist ein Gradienten-Abstieg- oder Steilster-Abstieg-Verfahren mit strikter Datenkonsistenzbedingung (**Strict** **D**ata **C**onsistency). Diese Methode wird im weiteren Fortlauf der Arbeit als **Strict-DC**-Methode bezeichnet.

Das Ziel dieser Arbeit war es, mögliche Anwendungen für CS in der MRT zu identifizieren. Hierzu sollte ein Rekonstruktionsschema gefunden werden, welches das Eingreifen des Experimentators in den Rekonstruktionsprozess auf ein Minimum beschränkt, da dieses zu einem ungewünschten subjektiven, benutzerabhängigen Einfluss auf das Rekonstruktionsergebnis führt in Fällen, in denen tatsächlich unterabgetastete Daten ohne ein Referenzbild rekonstruiert werden.

Zu diesem Zweck wurden in Kapitel 3 Simulationen mit dem Shepp-Logan-Phantom durchgeführt, um die Charakteristiken der Relaxed-DC- und Strict-DC-Methode zu bestimmen. Der Regularisierungsparameter, welcher in der Relaxed-DC-Methode auftritt, war hierbei von besonderer Wichtigkeit, denn die korrekte Anpassung dieses Parameters hat kritischen Einfluss auf die Rekonstruktionsqualität. Daher wurde getestet, ob das Variieren von relevanten Bildgebungsparametern eine Neuanpassung des Regularisierungsparameters erfordert. Die untersuchten Bildgebungsparameter waren der Beschleunigungsfaktor sowie das angewendete Abtastschema. Es zeigte sich, dass der Regularisierungsparameter von beiden Größen abhängt und daher eine zeitraubende empirische Bestimmung eines korrekten Regularisierungsparameters notwendig ist. Aus diesem Grund wurde eine weitere detaillierte Untersuchung der Abhängigkeiten des Regularisierungsparameters vom Rauschen oder der Auflösung verworfen.
Es konnte gezeigt werden, dass die Strict-DC-Methode zu verlässlichen, stabilen und reproduzierbaren Resultaten führt. Die Rekonstruktionen mit der Strict-DC-Methode resultierten in den meisten Fällen in signifikant kleineren RMSE-Werten als diejenigen, die mit der Relaxed-DC-Methode erzielt wurden. Der wichtigste Grund, warum die Relaxed-DC-Methode nicht weiter zur Rekonstruktion von unterabgetasteten, sparsen Daten in Betracht gezogen worden ist, war die bereits angesprochene aufwendige und zeitraubende Bestimmung des Regularisierungsparameters. Daher wurde die Strict-DC-Methode für

den Rest dieser Arbeit als CS-Rekonstruktionsschema gewählt. Es wurde allerdings ein weitere Parameter untersucht, der die Genauigkeit der Rekonstruktion in der Strict-DC-Methode kontrolliert und angepasst werden musste. Es konnte demonstriert werden, dass der Parameter geeignet gewählt werden kann, wenn die Signalintensität des Startbildes (Initial Guess) in der Strict-DC-Methode auf 1 normiert wird. Daher kann die Strict-DC-Methode als „Black-Box"-Algorithmus benutzt werden, um akkurat und verlässlich unterabgetastete MR-Datensätze rekonstruieren zu können. Alle CS-Rekonstruktionen in den übrigen Kapiteln dieser Arbeit wurden entweder mit der Strict-DC-Methode oder einer Joint-Sparsity-Erweiterung der Strict-DC-Methode erzielt (siehe Kapitel 5).

In Kapitel 4 wurde gezeigt, dass CS entweder zur Beschleunigung des Datenakquisitionsprozesses oder zur Erhöhung des apparenten SNRs in 2D oder 3D Chemical Shift Imaging (CSI) vorteilhaft eingesetzt werden kann. Datensätze von exogen verabreichten ^{19}F-Markersubstanzen sind sparse im Bildraum auf Grund der vernachlässigbaren Menge an natürlich vorkommendem ^{19}F in lebendem Gewebe. Es wurde ein pseudo-zufälliges kartesisches Abtastschema mit vollständig aufgenommenem innerem k-Raum angewendet, da eine voll-phasenkodierte CSI-Sequenz zur Datenakquisition eingesetzt wurde. Die spektrale Dimension wurde hierbei nicht unterabgetastet, vollständige Spektren wurden an jedem aufgenommenen k-Raum-Punkt erhalten. Die entsprechenden Bildräume wurden an jedem einzelnen spektralen Punkt mit der Strict-DC-Methode CS-rekonstruiert. Zunächst wurden Simulationen mit schwellenwert-begrenzten ^{19}F-Daten und mit einem numerisch generierten Maus-Phantom durchgeführt. An diesen simulierten Datensätzen wurden die Abhängigkeiten des CS-Algorithmus vom Rauschen, vom Beschleunigungsfaktor sowie von der Norm des Minimierungsproblems untersucht. Aus den hieraus erhaltenen Ergebnissen wurden optimierte Parameter für den CS-Algorithmus abgeleitet und an unterabgetasteten *ex-vivo*- und *in-vivo*-Datensätzen getestet. Ein Vergleich zwischen den rekonstruierten *ex-vivo*-Bildern und einem zusätzlich akquirierten voll-aufgenommenen Referenzdatensatz bewies erneut die Stabilität, Verlässlichkeit und Reproduzierbarkeit der Strict-DC-Methode. Es konnte demonstriert werden, dass CS den Datenakquisitionsprozess bis zu einem Faktor Acht bei ausreichendem SNR beschleunigen kann. Dies erlaubt die Aufnahme eines räumlich hochaufgelösten 3D ^{19}F-CSI-Datensatzes auf Zeitskalen, die bei *in-vivo*-Messungen tolerierbar sind. Falls das SNR im unterabgetasteten Datensatz niedrig ist, kann die Bildqualität in den resultierenden Bildern durch mehrmaliges Mitteln der unterabgetasteten Daten und anschließende CS-Rekonstruktion verbessert werden. Dies führt zu Rekonstruktionen mit verbessertem apparenten SNR. Eine exakte SNR-Quantifizierung ist zur Zeit nicht möglich, da es unbekannt ist, wie die Rauschverteilung nach der nicht-linearen CS-Rekonstruktion analytisch beschrieben werden kann.

Die Strict-DC-Methode, die bislang auf Einzel-Empfänger-Daten angewendet wurde, wurde in Kapitel 5 auf Multi-Empfänger-Anwendungen erweitert. Hierbei wurde das Potential von CS in der dynamischen radialen Herzbildgebung untersucht. Um den CS-Schritt effizient und optimal auf verrauschten Multi-Empfänger-Daten ausführen zu können, wurde „Compressed Sensing On Combined Coils" (CS-CC, dt.: Compressed Sensing auf kombinierten Spulen) entwickelt. Diese Technik beinhaltet einen Schritt zur Spulenkombination, um den CS-Schritt auf einem SNR-optimierten Kombinationsbild aller Einzelspulen durchführen zu können. Strikte Datenkonsistenz wurde wiefolgt gewährleistet: Nach jedem CS-Schritt wurden die CS-prozessierten, spulenkombinierten Daten entkombiniert, d.h. auf die Einzelspulen zurückverteilt. Im Anschluss wurden die gegriddeten akquirierten Daten an den entsprechenden k-Raum-Positionen wieder eingesetzt. Die Methode ist mit Blick auf den Rechenaufwand effizient, da alle Datenprozessierungsschritte nach Gridden der radialen Daten mit GROG [39, 40] in der kartesischen Ebene ausgeführt werden konnten.

Dadurch war kein rechnerisch aufwendiger Degridding-Schritt notwendig. Da die dynamischen Daten in der Bild-Domäne nicht sparse sind, wurden die Daten durch Abziehen der radialen Projektionen des gewünschten Timeframes von einem zeitlich gemittelten Datensatz sparsifiziert. Dies führte zu sparsen unterabgetasteten dynamischen Differenzen in der Bild-Domäne, welche effizient mit dem CS-CC-Algorithmus wiederhergestellt werden konnten. Simulationen zeigten die Abhängigkeiten der CS-CC-Methde von Rauschen, Beschleunigungsfaktor und Norm auf. Auf Basis der Simulationen konnten optimierte Parameter für den CS-CC-Algorithmus abgeleitet, transferiert und an einem retrospektiv unterabgetasteten radialen Herz-Cine-Datensatz getestet werden. Abschließend wurde die Anwendbarkeit der vorgeschlagenen Methode an einem unterabgetasteten radialen Echtzeit-Herz-Datensatz gezeigt. Akkurate Rekonstruktionen von 24 Projektionen pro Timeframe mit einer Zeitauflösung von 66 ms konnten erzielt werden. Im Prinzip ist die Methode in der Lage, unterabgetastete Datensätze von Patienten mit Herzrhythmusstörungen zu rekonstruieren und damit eventuell neue Einblicke in die Anatomie und Funktion des Herzens bei dieser Erkrankung zu erlauben.

Eine vom Rechenaufwand her effiziente Implementierung von Iterativem GRAPPA wurde in Kapitel 6 vorgestellt. Die Entwicklung der vorgeschlagenen Implementierung von Iterativem GRAPPA folgte aus der Notwendigkeit einer Parallelen Bildgebungsstrategie, die fehlende Datenpunkte auf irregulär unterabgetasteten kartesischen Daten rekonstruieren kann. Desweiteren war es ein Schritt, der einer späteren Kombination von Iterativem GRAPPA und CS namens CS-GRAPPA vorausging. In der kartesischen Parallelen Bildgebung wird die Phasenkodierrichtung in regelmäßiger Weise unterabgetastet, d.h. es wird nur jede zweite, dritte, ... k-Raum-Zeile aufgenommen. Wenn nicht-kartesische Daten mit GROG auf ein regelmäßiges kartesisches Gitter verschoben werden, erhält man einen irregulär unterabgetasteten k-Raum, der mit konventionellen Methoden der Parallelen Bildgebung wie SENSE [7] oder GRAPPA [8] nicht wiederhergestellt werden kann. Bei Iterativem GRAPPA werden fehlende Datenpunkte durch eine gewichtete Summe der umgebenden k-Raum-Punkte rekonstruiert. Dies wird durch die Anwendung eines quadratischen GRAPPA-Kerns (z.B. 5x5) erreicht, bei dem der zentrale Punkt der Zielpunkt ist. Danach werden die tatsächlich aufgenommenen und mit GROG auf ein kartesisches Gitter verschobenen Daten wieder eingesetzt, um die strikte Datenkonsistenz zu gewährleisten. Dieses Vorgehen muss mehrmals wiederholt werden, da im unterabgetasteten kartesischen k-Raum größere Lücken auftreten können. Durch wiederholtes (d.h. iteratives) Verschieben eines quadratischen GRAPPA-Kerns über die unterabgetasteten k-Raum-Daten können diese Lücken iterativ rekonstruiert werden. Das ursprüngliche Konzept von Iterativem GRAPPA wurde 2007 auf der ISMRM präsentiert [96] und beinhaltet Schritte, um die Daten auf ein kartesisches Gitter verschieben zu können (engl.: gridding) und von diesem wieder auf das nicht-kartesische (engl.: de-gridding). Dies ist vom Rechenaufwand her anspruchsvoll und kann durch das Gridden der nicht-kartesischen Daten mittels GROG umgangen werden. Daher können in der vorgeschlagenen Implementierung alle iterativen Rekonstruktionsschritte in der kartesischen Ebene durchgeführt werden. In Simulationen am Shepp-Logan-Phantom wurden die Abhängigkeiten der Methode vom Rauschlevel, dem Beschleunigungsfaktor, der Größe des GRAPPA-Kerns sowie eines Regularisierungsparameters in der Moore-Penrose-Pseudoinversen untersucht. Es zeigte sich, dass unterabgetastete radiale Daten zu besseren Rekonstruktionsergebnissen führen als pseudo-zufällig unterabgetastete kartesische Daten. Desweiteren konnte die Effizienz des Iterativen-GRAPPA-Algorithmus optimiert werden, indem ein 5x5-Kern eingesetzt wurde, welcher zu akkuraten Rekonstruktionen innerhalb von 20 Iterationen führte. Der Regularisierungsparameter in der Moore-Penrose-Pseudoinversen kann auf einen geeigneten Wert justiert werden, welcher zu akkuraten Rekonstruktionen in allen untersuchten Situationen führte. Erwartungsgemäß verschlechterte sich die Rekonstruktionsqualität, je höher der Rauschlevel und je größer der Beschleunigungsfaktor gewählt wurde. Diese Erkenntnisse konnten an Hand erfolgreicher Rekonstruktionen an einem retrospektiv unterabgetasteten radialen Herzdatensatz bestätigt werden.

Das abschließende Kapitel 7 beleuchtet eine naheliegende Möglichkeit, um die Strict-DC-Methode mit Iterativem GRAPPA synergetisch zu kombinieren. Dieser Hybrid-Algorithmus aus CS und Paralleler Bildgebung wird CS-GRAPPA genannt. Der Algorithmus beinhaltet einen Rekonstruktionsschritt mit Iterativem GRAPPA nach jedem einzelnen spulenweisen CS-Schritt. Dadurch wird ein neues Konsistenz-Kriterium eingeführt: Neben der strikten Datenkonsistenz müssen die Daten auch mit den inhärenten Spulensensitivitätsprofilen konsistent sein. Darüber hinaus kann, wie bereits im vorherigen Kapitel gezeigt, Iteratives GRAPPA selbst Daten durch die intrinsischen Kodiereigenschaften der Spulenanordnung wieder herstellen. Hierdurch werden Fehler aus dem CS-Rekonstruktionsschritt ausbalanciert während die Rauscherhöhung aus dem Parallelen Bildgebungsschritt durch den CS-Schritt effektiv reduziert wird. Es wurden Simulationen am Shepp-Logan-Phantom durchgeführt, um das Potential von CS-GRAPPA zu untersuchen. Hierbei war insbesondere von Interesse, ob CS-GRAPPA höhere Beschleunigungsfaktoren als die beiden Einzelmethoden CS bzw. Iteratives GRAPPA erlauben würde. Desweiteren wurde CS-GRAPPA mit der CS-CC-Methode verglichen, die in Kapitel 5 eingeführt wurde. Die Resultate zeigten, dass CS-GRAPPA die Rekonstruktion höher unterabgetasteter Daten erlaubt als die beiden Methoden für sich genommen. Desweiteren konnte die Rauscherhöhung durch Iteratives GRAPPA ebenso reduziert werden wie die Spike-Artefakte, die durch CS bedingt sind. Die CS-CC-Methode resultierte jedoch in einer geringfügig besseren Bildqualität im Sinne des RMSE als CS-GRAPPA, während die augenscheinliche Bildqualität vergleichbar zu CS-GRAPPA war. Daher scheint das Ausnutzen der Joint Sparsity der Empfängerspulen durch die CS-CC-Methode erfolgreicher zu sein als eine Kombination von Paralleler Bildgebung mit einer spulenweisen CS-Technik. Nichtsdestotrotz markiert CS-GRAPPA einen möglichen Startpunkt, um eine fortgeschrittenere Kombination aus CS, welches auch die Joint Sparsity ausnutzt, und Paralleler Bildgebung zu entwickeln.

Ausblick

Im Kapitel 4 wurde gezeigt, dass CS vorteilhaft für Einzelempfänger-Daten ist bei denen Parallele Bildgebungsmethoden nicht angewendet werden können. Aus diesem Grund ist CS eine interessante Möglichkeit, um die Aufnahme von sparsen Daten bei Hochfeld-NMR-Systemen zu beschleunigen. Bei diesen Systemen sind zumeist nur (selbstgebaute) Einzelempfänger-Spulen verfügbar. Ähnlich wie in der ^{19}F-Bildgebung führen auch andere X-Kerne wie ^{13}C oder ^{23}Na zu sparsen Signalverteilungen auf Grund der vernachlässigbaren Mengen dieser Isotope im lebenden Gewebe. Allerdings ist das gyromagnetische Verhältnis von ^{19}F mit $\gamma_{^{19}\mathrm{F}} = 40.04$ MHz T^{-1} das zweitgrößte nach dem für Protonen $\gamma_{^{1}\mathrm{H}} = 42.58$ MHz T^{-1}. Da die Gleichgewichtsmagnetisierung und damit auch das Empfängersignal proportional zu γ^2 (siehe Gleichung 1.19) ist, wird das SNR anderer X-Kern-Bildgebungsexperimente (bei gleicher Spindichte) signifikant kleiner als das von ^{1}H oder ^{19}F sein. An dieser Stelle seien die gyromagnetischen Verhältnisse von ^{13}C und ^{23}Na angegeben: $\gamma_{^{13}\mathrm{C}} = 10.71$ MHz T^{-1} bzw. $\gamma_{^{23}\mathrm{Na}} = 11.26$ MHz T^{-1}. Ähnlich wie bei ^{19}F-Bildgebung ist die Spindichte anderer X-Kerne sehr niedrig, was wiederum zu niedrigerem SNR führt. Wie in Kapitel 4 gezeigt wurde, könnte es jedoch möglich sein, die X-Kern-Datensätze unterabzutasten und anschließend mehrmals zu mitteln, um das

apparente SNR in den rekonstruierten Bildern zu erhöhen.
In diesem Zusammenhang ist die Reduktion von Artefakten in CS-Rekonstruktionen mit der Strict-DC-Methode aus Niedrig-SNR-Datensätzen von großer Wichtigkeit. Eine vielversprechende erste Idee wurde hierzu bereits publiziert [66].
Hyperpolarisation ist eine Möglichkeit, um das SNR signifikant zu verbessern. Erste Anwendungen von CS in spektroskopischer Bildgebung von hyperpolarisiertem ^{13}C wurden bereits vorgestellt [52–54]. ^{3}He ist ein anderer Kern, der gut geeignet wäre zur Hyperpolarisation; erste erfolgreiche CS-Rekonstruktionen von beschleunigten, hyperpolarsierten ^{3}He-Lungendaten wurden berichtet [102–106]. Hyperpolarisierte X-Kerne sind demzufolge eine vielversprechende Anwendung für CS und sollten weiter untersucht werden.
Eine interessante Anwendung für CS in konventioneller Protonen-MR könnte die Dental-MRT sein. In Kapitel 5 von Olga Tymofiyevas Doktorarbeit [107] wurde ein Überblick über das Potential von CS in der Dental-MRT gegeben. Eine detailliertere Untersuchung dieses Themas kann in Sven Vaeglers Diplomarbeit nachgelesen werden [100]. In beiden Arbeiten konnte demonstriert werden, dass CS Dental-MRT-Aufnahmen effektiv um einen Faktor 2 beschleunigen kann, wobei die notwendige Genauigkeit in der Rekonstruktion der Zahnoberflächen gewahrt bleibt. Die unterabgetasteten Datensätze wurden allerdings mit einer modifizierten Spinecho-Sequenz aufgenommen. Ein beschleunigtes Turbo-Spinecho wäre wünschenswert, weshalb weiter daran gearbeitet werden sollte.

Es gibt noch eine Reihe allgemeinerer Probleme im Zusammenhang mit CS.
Regularisierungsparameter treten in vielen CS-Algorithmen auf. Eine wichtige Frage ist, ob sie sich analytisch oder verlässlich auf automatische Art und Weise bestimmen lassen, ohne den ungewünschten Einfluss des Benutzers. In dieser Arbeit wurde ein Ansatz ohne Regularisierungsparameter gewählt, daher wurde diese Fragestellung in der vorliegenden Dissertation nicht untersucht.
Ein anderes wichtiges Problem ist die Suche nach einer optimalen Sparsifizierungstransformation für eine bestimmte Anwendung. In dieser Arbeit waren die Daten entweder sparse im Ortsraum, in der Diskreten-Gradienten-Domäne oder nach Subtraktion von einem Referenzdatensatz. Falls die Daten jedoch nicht zum angenommen Sparsity-Modell passen, wird die Rekonstruktion fehlerbehaftet sein. Für einige Anwendungen existieren modellbasierte Ansätze [101, 108, 109]. Ein interessantes Werkzeug in diesem Zusammenhang könnte die Principle Component Analysis sein.
Große Anstrengungen werden unternommen, um Parallele Bildgebung und CS miteinander zu kombinieren. Eine synergetische Kombination beider Methoden sollte es erlauben, stärker unterabgetastete Daten akkurat zu rekonstruieren als dies mit den jeweiligen einzelnen Methoden erreichbar wäre. Dies wurde auch in Kapitel 7 demonstriert. Einige Publikationen haben sich mit diesem Thema bereits beschäftigt [110–112]. Das Ziel sollte sein, Joint-Sparsity-CS mit Paralleler Bildgebung in einer robusten, verlässlichen und benutzer-unabhängigen Art und Weise zu kombinieren, um die Sparsity des Datensatzes ebenso effizient ausnutzen zu können wie die implizite räumliche Kodierungseigenschaft des verwendeten Spulen-Arrays.

APPENDIX

A Derivatives of the CS Cost Functions

In Chapter 2, Section 2.3.3, the cost functions are given which lead to the reconstruction of undersampled sparse datasets exploiting the CS theory. The optimization procedure requires the derivative of the respective cost functions. Since the detailed derivatives are not necessary to the understanding and implementation of the respective CS algorithms, they were discarded in Section 2.3.3. The explicit detailed derivatives will be given in this appendix. Hereby, the Einstein Summation Convention holds for the i-th component of a vector: $v_i = \hat{M}_{i,k} x_k = \sum_k \hat{M}_{i,k} x_k$ for Matrix $\hat{M}$ and vectors v and x.
All cost functions mentioned in this thesis are nonholomorphic since complex conjugate variables (complex conjugation indicated by *) appear. Therefore, the derivatives are obtained using the so-called "Wirtinger calculus" which requires for the optimization of a real-valued nonholomorphic function [113–115]:

$$\frac{\partial f}{\partial z^*} = \left(\frac{\partial f}{\partial z}\right)^* \overset{!}{=} 0 \tag{A.1}$$

with $f(z, z^*) \in \mathbb{R}, z \in \mathbb{C}$. The norms appearing in the cost functions (Equations 2.7 and 2.9) are real-valued nonholomorphic functions since the absolute value of a complex number is calculated according to $|z| = \sqrt{z^* z}$.
More details on Wirtinger calculus can be found in [113–116].

A.1 Derivative of Equation 2.7

The initial point for the derivative of Equation 2.7 is Equation 2.10 (obeying Equation A.1):

$$\begin{aligned}
[\nabla_{x^*} f(x)]_k &= \left[\nabla_{x^*} \left(\|F_u x - y\|_2^2 + \lambda \cdot \|\Psi x\|_1\right)\right]_k \\
&= \left[\nabla_{x^*} \left(\sum_i |\,(F_u x - y)_i\,|^2 + \lambda \cdot \sum_i |\,(\Psi x)_i\,|\right)\right]_k \\
&= \frac{\partial}{\partial x_k^*} \Bigg(\sum_i \left(x_j^* F^*_{u\,j,i} - y_i^*\right) \left(F_{u\,i,m} x_m - y_i\right) + \\
&\quad + \lambda \cdot \sum_i \sqrt{\left(x_j^* \Psi^*_{j,i}\right) \left(\Psi_{i,m} x_m\right)} \Bigg)
\end{aligned} \tag{A.2}$$

This is the kth component of the derivative. Please note that the derivative is a vector and has the same number of components as the image vector x.
As already pointed out in Section 2.3.3, the absolute value function is not smooth around zero, and, thus, cannot be derivated at this point. An approximation of the absolute value function solves this problem:

$$|x| = \sqrt{x^* x} \approx \sqrt{x^* x + \epsilon^2} \tag{A.3}$$

with x^* being the complex conjugate of x and ϵ being a smoothing factor.
The derivative using Equation 2.11 becomes:

$$\begin{aligned}
\left[\nabla_{x^*} f(x)\right]_k &= \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \left(x_j^* F_{u\,j,i}^* - y_i^*\right)\left(F_{u\,i,m} x_m - y_i\right) + \\
&\quad + \lambda \cdot \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right)} \\
&\approx \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \left\{\left(x_j^* F_{u\,j,i}^* - y_i^*\right)\left(F_{u\,i,m} x_m - y_i\right) + \epsilon^2\right\} + \\
&\quad + \lambda \cdot \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2} \\
&= \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \left\{\left(x_j^* F_{u\,j,i}^* - y_i^*\right)\left(F_{u\,i,m} x_m - y_i\right) + \epsilon^2\right\} + \lambda \cdot \tfrac{\partial}{\partial x_k^*} \textstyle\sum_i \sqrt{a_i}
\end{aligned} \tag{A.4}$$

Using the abbreviation $a_i = \left(\Psi x\right)_i^* \left(\Psi x\right)_i + \epsilon^2 \approx |\left(\Psi x\right)_i|^2$ for the radicants of the second term, the derivative can be further evaluated:

$$\begin{aligned}
\left[\nabla_{x^*} f(x)\right]_k &= \textstyle\sum_i \left(\delta_{k,j} F_{u\,j,i}^*\right)\left(F_{u\,i,m} x_m - y_i\right) + \\
&\quad + \lambda \cdot \textstyle\sum_i \frac{1}{2\sqrt{a_i}} \left(\delta_{k,j} \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) \\
&= \textstyle\sum_i F_{u\,k,i}^* \left(F_{u\,i,m} x_m - y_i\right) + \\
&\quad + \lambda \cdot \textstyle\sum_i \frac{1}{2\sqrt{a_i}} \Psi_{k,i}^* \left(\Psi_{i,m} x_m\right) \\
&= \left[F_u^* \left(F_u x - y\right)\right]_k + \tfrac{\lambda}{2} \cdot \textstyle\sum_i \cdot \frac{1}{\sqrt{a_i}} \Psi_{k,i}^* \left(\Psi_{i,m} x_m\right) \\
&= \left[F_u^* \left(F_u x - y\right)\right]_k + \tfrac{\lambda}{2} \cdot \textstyle\sum_i \cdot \Psi_{k,i}^* \underbrace{\frac{1}{\sqrt{a_i}} \Psi_{i,m} x_m}_{\approx \frac{\left(\Psi x\right)_i}{|\left(\Psi x\right)_i|}} \\
&= \left[F_u^* \left(F_u x - y\right)\right]_k + \tfrac{\lambda}{2} \cdot \left[\Psi^* \left(\frac{\Psi x}{W_\Psi}\right)\right]_k
\end{aligned} \tag{A.5}$$

where W_Ψ is a vector with the components $W_{\Psi i} = \sqrt{a_i} = \sqrt{\left(\Psi x\right)_i^* \left(\Psi x\right)_i + \epsilon^2} \approx |\left(\Psi x\right)_i|$ (see also Equation 2.13).
This is the detailed derivative which led to Equation 2.12. Ψ^* is hereby the inverse transform of Ψ and the factor $\frac{1}{2}$ can be absorbed in the regularization parameter $\lambda : \frac{\lambda}{2} \mapsto \lambda$.

A.2 Derivative of Equation 2.9

Consider now Equation 2.9 for an arbitrary p and calculate the derivative (again obeying Equation A.1):

$$\begin{aligned}
\left[\nabla_{x^*} f(x)\right]_k &= \left[\nabla_{x^*} \left(\|\Psi x\|_p\right)\right]_k \\
&= \tfrac{\partial}{\partial x_k^*} \left(\sum_i \left\{\sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,j} x_j\right)}\right\}^p\right)^{\frac{1}{p}}
\end{aligned} \tag{A.6}$$

Again, this is the kth component of the derivative. Please note that the derivative is a vector and has the same number of components as the image vector x.
In the following, the power $\frac{1}{p}$ is neglected. The derivative of this outer function would yield the power $\frac{1-p}{p}$ which just scales the vector of the derivative.
The derivative using the norm approximation given above and dividing by p becomes:

$$\begin{aligned}
\left[\nabla_{x^*} f(x)\right]_k &\approx \tfrac{1}{p} \cdot \tfrac{\partial}{\partial x_k^*} \sum_i \left\{\sqrt{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2}\right\}^p \\
&= \tfrac{1}{p} \cdot \tfrac{p}{2} \cdot \sum_i \left\{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2\right\}^{\frac{p}{2}-1} \cdot \left(\delta_{k,j} \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) \\
&= \tfrac{1}{2} \cdot \sum_i \Psi_{k,i}^* \underbrace{\left\{\left(x_j^* \Psi_{j,i}^*\right)\left(\Psi_{i,m} x_m\right) + \epsilon^2\right\}^{\frac{p}{2}-1}}_{\approx |(\Psi x)_i|^{p-2}} \cdot \left(\Psi_{i,m} x_m\right) \\
&= \tfrac{1}{2} \cdot \left[\Psi^* \left(W_\Psi^{p-2} \Psi x\right)\right]_k
\end{aligned} \tag{A.7}$$

This is equivalent to Equation 2.17. W_Ψ is hereby defined as in the previous section and in Equation 2.13. W_Ψ^{p-2} means that each element of W_Ψ is taken to that power.

Bibliography

[1] A. Haase, J. Frahm, D. Matthaei, W. Hänicke, and K. D. Merboldt. FLASH imaging. rapid NMR imaging using low flip-angle pulses. *J Magn Reson*, 67:258–266, 1986.

[2] J. Hennig, A. Nauerth, and H. Friedburg. RARE imaging: a fast imaging method for clinical MR. *Magn Reson Med*, 3(6):823–833, Dec 1986.

[3] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller. The NMR phased array. *Magn Reson Med*, 16(2):192–225, Nov 1990.

[4] D. K. Sodickson and W. J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn Reson Med*, 38(4):591–603, Oct 1997.

[5] P. M. Jakob, M. A. Griswold, R. R. Edelman, and D. K. Sodickson. AUTO-SMASH: a self-calibrating technique for SMASH imaging. *Magn Reson Mater Phy*, 7(1):42–54, Nov 1998.

[6] R. M. Heidemann, M. A. Griswold, A. Haase, and P. M. Jakob. VD-AUTO-SMASH imaging. *Magn Reson Med*, 45(6):1066–1074, Jun 2001.

[7] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magn Reson Med*, 42(5):952–962, Nov 1999.

[8] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med*, 47(6):1202–1210, Jun 2002.

[9] M. A. Griswold, P. M. Jakob, M. Nittka, J. W. Goldfarb, and A. Haase. Partially parallel imaging with localized sensitivities (PILS). *Magn Reson Med*, 44(4):602–609, Oct 2000.

[10] W. E. Kyriakos, L. P. Panych, D. F. Kacher, C. F. Westin, S. M. Bao, R. V. Mulkern, and F. A. Jolesz. Sensitivity profiles from an array of coils for encoding and reconstruction in parallel (SPACE RIP). *Magn Reson Med*, 44(2):301–308, Aug 2000.

[11] M. Bydder, D. J. Larkman, and J. V. Hajnal. Generalized SMASH imaging. *Magn Reson Med*, 47(1):160–170, Jan 2002.

[12] J. Wang, T. Kluge, M. Nittka, V. Jellus, B. Kühn, and B. Kiefer. Parallel acquisition techniques with modified SENSE reconstruction mSENSE. In *Proceedings of the First International Workshop on Parallel MRI, Würzburg, Germany*, page 89, 2001.

[13] J. Tsao, P. Boesiger, and K. P. Pruessmann. k-t BLAST and k-t SENSE: dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magn Reson Med*, 50(5):1031–1042, Nov 2003.

[14] Emmanuel Candès. Compressive sampling. In *Proceedings of the International Congress of Mathematicians, Madrid, Spain*, 2006.

[15] David Donoho. Compressed sensing. *IEEE Trans Inf Theory*, 52:1289–1306, 2006.

[16] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med*, 58(6):1182–1195, Dec 2007.

[17] M. Lustig and J. M. Pauly. SPIRiT: Iterative Self-consistent Parallel Imaging Reconstruction From Arbitrary k-Space. *Magn Reson Med*, 64:457–471, 2010.

[18] M. Lustig, M. Alley, S. Vasanawala, D. L. Donoho, and J. M. Pauly. L_1-SPIRiT: Autocalibrating Parallel Imaging Compressed Sensing. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 379, 2009.

[19] M. Lustig, M. Alley, S. Vasanawala, D. L. Donoho, and J. M. Pauly. Autocalibrating Parallel Imaging Compressed Sensing using L_1-SPIRiT. In *ISMRM Workshop on Data Processing and Image Reconstruction, Sedona/AZ, USA*, 2009.

[20] Felix Bloch. Nuclear induction. *Physical Review*, 70:460–469, 1946.

[21] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, 69:37–38, 1946.

[22] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. Technical report, California Institute of Technology, http://www-stat.stanford.edu/∼candes/papers/ExactRecovery.pdf, 2004.

[23] David Donoho. Compressed sensing. Technical report, Stanford University, http://www-stat.stanford.edu/∼donoho/Reports/2004/CompressedSensing091604.pdf, 2004.

[24] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory*, 52:489–509, 2006.

[25] J. Du, F. R. Korosec, Y. Wu, T. M. Grist, and C. A. Mistretta. Whole-body MR angiography using variable density sampling and dual-injection bolus-chase acquisition. *Magn Reson Imaging*, 26(2):181–187, Feb 2008.

[26] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[27] Michael Lustig. *Sparse MRI.* PhD thesis, Stanford University, http://www.eecs.berkeley.edu/∼mlustig/mlustigThesis.pdf, 2008.

[28] D. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans Inf Theory*, 47:2845–2862, 2001.

[29] T. Cukur, M. Lustig, and D. G. Nishimura. Improving non-contrast-enhanced steady-state free precession angiography with compressed sensing. *Magn Reson Med*, 61(5):1122–1131, May 2009.

[30] Balas Kausik Natarajan. Sparse Approximate Solutions to Linear Systems. *SIAM J Comput*, 24(2):227–234, 1995.

[31] David Donoho. For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. *Commun Pure Appl Math*, 59:797–829, 2006.

[32] Rick Chartrand. Exact reconstruction of sparse signals via nonconvex compressed sensing. *IEEE Signal Proc Let*, 14:707–710, 2007.

[33] T. Grotz, B. Zahneisen, A. Ella, M. Zaitsev, and J. Hennig. Fast functional brain imaging using constrained reconstruction based on regularization using arbitrary projections. *Magn Reson Med*, 62(2):394–405, Aug 2009.

[34] Jonathan Richard Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh/PA, USA, http://www.cs.cmu.edu/∼quake-papers/painless-conjugate-gradient.pdf, 1994.

[35] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[36] A. S. Stern, D. L. Donoho, and J. C. Hoch. NMR data processing using iterative thresholding and minimum l_1-norm reconstruction. *J Magn Reson*, 188(2):295–300, Oct 2007.

[37] J. Yerly, M. L. Lauzon, H. S. Chen, and R. Frayne. A simulation-based analysis of the potential of compressed sensing for accelerating passive MR catheter visualization in endovascular therapy. *Magn Reson Med*, 63(2):473–483, Feb 2010.

[38] M. Usman and P. G. Batchelor. Optimized Sampling Patterns for Practical Compressed MRI. In *8^{th} International Conference on Sampling Theory and Applications (SAMPTA), Marseille, France*, page 156, 2009.

[39] N. Seiberlich, F. A. Breuer, M. Blaimer, K. Barkauskas, P. M. Jakob, and M. A. Griswold. Non-Cartesian data reconstruction using GRAPPA operator gridding (GROG). *Magn Reson Med*, 58(6):1257–1265, Dec 2007.

[40] N. Seiberlich, F. Breuer, M. Blaimer, P. M. Jakob, and M. A. Griswold. Self-calibrating GRAPPA operator gridding for radial and spiral trajectories. *Magn Reson Med*, 59(4):930–935, Apr 2008.

[41] E. T. Ahrens, R. Flores, H. Xu, and P. A. Morel. *In vivo* imaging platform for tracking immunotherapeutic cells. *Nat Biotechnol*, 23(8):983–987, Aug 2005.

[42] V. D. Kodibagkar, J. Yu, L. Liu, H. P. Hetherington, and R. P. Mason. Imaging beta-galactosidase activity using ^{19}F chemical shift imaging of lacz gene-reporter molecule 2-fluoro-4-nitrophenol-beta-d-galactopyranoside. *Magn Reson Imaging*, 24(7):959–962, Sep 2006.

[43] K. C. Partlow, J. Chen, J. A. Brant, A. M. Neubauer, T. E. Meyerrose, M. H. Creer, J. A. Nolta, S. D. Caruthers, G. M. Lanza, and S. A. Wickline. ^{19}F magnetic resonance imaging for stem/progenitor cell tracking with multiple unique perfluorocarbon nanobeacons. *FASEB J*, 21(8):1647–1654, Jun 2007.

[44] G. Bringmann, K. Wolf, M. Meininger, M. Rokitta, and A. Haase. *In vivo* ^{19}F NMR chemical-shift imaging of Ancistrocladus species. *Protoplasma*, 218(3-4):134–143, 2001.

[45] C. W. Li and O. Gonen. Simultaneous 3D NMR spectroscopy of fluorine and phosphorus in human liver during 5-fluorouracil chemotherapy. *Magn Reson Med*, 35(6):841–847, Jun 1996.

[46] U. Flögel, Z. Ding, H. Hardung, S. Jander, G. Reichmann, C. Jacoby, R. Schubert, and J. Schrader. *In vivo* monitoring of inflammation after cardiac and cerebral ischemia by fluorine magnetic resonance imaging. *Circulation*, 118(2):140–148, Jul 2008.

[47] T. C. Basse-Lüsebrink, G. Ladewig, T. Kampf, G. Melkus, D. Haddad, W. R. Bauer, and P. M. Jakob. Multi-color ^{19}F CSI: Simultaneous detection of differently labeled cells *in vivo*. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 806, 2009.

[48] R. Lamerichs, C. Kammen, M. Yildirim, and K. Nicolay. *In vivo* ultra-fast spectroscopic imaging of ^{19}F containing contrast agents. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 616, 2009.

[49] Axel Haase. Snapshot FLASH MRI. Applications to T_1, T_2, and chemical-shift imaging. *Magn Reson Med*, 13(1):77–89, Jan 1990.

[50] H. Serrai and L. Senhadji. Acquisition time reduction in magnetic resonance spectroscopic imaging using discrete wavelet encoding. *J Magn Reson*, 177(1):22–30, Nov 2005.

[51] Y. S. Levin, D. Mayer, Y.-F. Yen, R. E. Hurd, and D. M. Spielman. Optimization of fast spiral chemical shift imaging using least squares reconstruction: application for hyperpolarized ^{13}C metabolic imaging. *Magn Reson Med*, 58(2):245–252, Aug 2007.

[52] S. Hu, M. Lustig, A. P. Chen, J. Crane, A. Kerr, D. A. C. Kelley, R. Hurd, J. Kurhanewicz, S. J. Nelson, J. M. Pauly, and D. B. Vigneron. Compressed sensing for resolution enhancement of hyperpolarized ^{13}C flyback 3D-MRSI. *J Magn Reson*, 192(2):258–264, Jun 2008.

[53] S. Hu, A. Balakrishnan M. Lustig, P. E. Z. Larson, R. Bok, J. Kurhanewicz, S. J. Nelson, J. M. Pauly, A. Goga, and D. B. Vigneron. High acceleration 3D compressed sensing hyperpolarized ^{13}C MRSI of a transgenic mouse model of liver cancer. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 131, 2009.

[54] P. E. Z. Larson, S. Hu, M. Lustig, A. B. Kerr, S. J. Nelson, J. Kurhanewicz, J. M. Pauly, and D. B. Vigneron. 3D dynamic MRSI for hyperpolarized ^{13}C with compressed sensing and multiband excitation pulses. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 257, 2009.

[55] J. W. M. Bulte and D. L. Kraitchman. Iron oxide MR contrast agents for molecular and cellular imaging. *NMR Biomed*, 17(7):484–499, Nov 2004.

[56] G. Stoll and M. Bendszus. Imaging of inflammation in the peripheral and central nervous system by magnetic resonance imaging. *Neuroscience*, 158(3):1151–1160, Feb 2009.

[57] B. D. Watson, W. D. Dietrich, R. Busto, M. S. Wachtel, and M. D. Ginsberg. Induction of reproducible brain infarction by photochemically initiated thrombosis. *Ann Neurol*, 17(5):497–504, May 1985.

[58] C. Kleinschnitz, M. Bendszus, M. Frank, L. Solymosi, K. V. Toyka, and G. Stoll. *In vivo* monitoring of macrophage infiltration in experimental ischemic brain lesions by magnetic resonance imaging. *J Cereb Blood Flow Metab*, 23(11):1356–1361, Nov 2003.

[59] C. Kleinschnitz, A. Schütz, I. Nölte, T. Horn, M. Frank, L. Solymosi, G. Stoll, and M. Bendszus. *In vivo* detection of developing vessel occlusion in photothrombotic ischemic brain lesions in the rat by iron particle enhanced MRI. *J Cereb Blood Flow Metab*, 25(11):1548–1555, Nov 2005.

[60] A. Fischer, T. C. Basse-Lüsebrink, T. Kampf, G. Ladewig, M. Blaimer, F. Breuer, G. Stoll, W. R. Bauer, and P. M. Jakob. Improved sensitivity in ^{19}F cellular imaging using nonconvex compressed sensing. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 3154, 2009.

[61] Rick Chartrand. Nonconvex compressed sensing and reconstruction of gradient-sparse images: Random vs. tomographic Fourier sampling. In *Proceedings of the IEEE International Conference on Image Processing, San Diego/CA, USA*, page 2793, 2008.

[62] M. Schroeter, S. Jander, and G. Stoll. Non-invasive induction of focal cerebral ischemia in mice by photothrombosis of cortical microvessels: characterization of inflammatory responses. *J Neurosci Methods*, 117(1):43–49, May 2002.

[63] O. Speck, K. Scheffler, and J. Hennig. Fast ^{31}P chemical shift imaging using SSFP methods. *Magn Reson Med*, 48(4):633–639, Oct 2002.

[64] P. M. Joseph and D. Lu. A technique for double resonant operation of birdcage imaging coils. *IEEE Trans Med Imaging*, 8(3):286–294, 1989.

[65] R. Pohmann and M. von Kienlin. Accurate phosphorus metabolite images of the human heart by 3D acquisition-weighted CSI. *Magn Reson Med*, 45(5):817–826, May 2001.

[66] T. C. Basse-Lüsebrink, T. Kampf, A. Fischer, G. Ladewig, G. Stoll, and P. M. Jakob. Spike artifact reduction in nonconvex Compressed Sensing. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 4886, 2010.

[67] R. Saab, R. Chartrand, and O. Yilmaz. Stable sparse approximations via nonconvex optimization. In *Proceedings of the 33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas/NV, USA*, pages 3885–3888, 2008.

[68] U. Gamper, P. Boesiger, and S. Kozerke. Compressed sensing in dynamic MRI. *Magn Reson Med*, 59(2):365–373, Feb 2008.

[69] V. Sturm, T. Hertlein, T. C. Basse-Lüsebrink, D. Haddad, K. Ohlsen, and P. M. Jakob. *In vivo* monitoring of bacterial infections using high-field MR microscopy. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 215, 2010.

[70] Y. Ye, T. C. Basse-Lüsebrink, P. Arias, K. Hu, T. Kampf, V. Kocoski, X. Helluy, P. M. Jakob, K.-H. Hiller, R. Jahns, and W. R. Bauer. MR cell tracking in reperfused myocardial infarction with microvascular obstruction and haemorrhage: fluorine-19 MR could be a better solution. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 1870, 2010.

[71] A. Fischer, F. Breuer, M. Blaimer, N. Seiberlich, and P. M. Jakob. Introduction of a nonconvex compressed sensing algorithm for MR imaging. In *Proceedings of the 16th Annual Meeting of the ISMRM, Toronto/ON, Canada*, page 1487, 2008.

[72] P. Parasoglou, A. J. Sederman, J. Rasburn, H. Powell, and M. L. Johns. Optimal k-space sampling for single point imaging of transient systems. *J Magn Reson*, 194(1):99–107, Sep 2008.

[73] P. Parasoglou, D. Malioutov, A. J. Sederman, J. Rasburn, H. Powell, L. F. Gladden, A. Blake, and M. L. Johns. Quantitative single point imaging with compressed sensing. *J Magn Reson*, 201(1):72–80, Nov 2009.

[74] K. T. Block, M. Uecker, and J. Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med*, 57(6):1086–1098, Jun 2007.

[75] M. Buehrer, J. Curcic, P. Boesiger, and S. Kozerke. Prospective self-gating for simultaneous compensation of cardiac and respiratory motion. *Magn Reson Med*, 60(3):683–690, Sep 2008.

[76] R. Manka, M. Buehrer, P. Boesiger, E. Fleck, and S. Kozerke. Performance of simultaneous cardiac-respiratory self-gated three-dimensional MR imaging of the heart: initial experience. *Radiology*, 255(3):909–916, Jun 2010.

[77] S. Weick, P. Ehses, M. Blaimer, F. A. Breuer, and P. M. Jakob. DC gated high resolution 3D MRI of the human lung under free breathing. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 2525, 2010.

[78] B. Hiba, N. Richard, M. Janier, and P. Croisille. Cardiac and respiratory double self-gated cine MRI in the mouse at 7 T. *Magn Reson Med*, 55(3):506–513, Mar 2006.

[79] B. Hiba, N. Richard, H. Thibault, and M. Janier. Cardiac and respiratory self-gated cine MRI in the mouse: comparison between radial and rectilinear techniques at 7T. *Magn Reson Med*, 58(4):745–753, Oct 2007.

[80] P. Kellman, C. Chefd'hotel, C. H. Lorenz, C. Mancini, A. E. Arai, and E. R. McVeigh. High spatial and temporal resolution cardiac cine MRI from retrospective reconstruction of data acquired in real time using motion correction and resorting. *Magn Reson Med*, 62(6):1557–1564, Dec 2009.

[81] P. Kellman, F. H. Epstein, and E. R. McVeigh. Adaptive sensitivity encoding incorporating temporal filtering (TSENSE). *Magn Reson Med*, 45(5):846–852, May 2001.

[82] P. Kellman, C. Chefd'hotel, C. H. Lorenz, C. Mancini, A. E. Arai, and E. R. McVeigh. Fully automatic, retrospective enhancement of real-time acquired cardiac cine MR images using image-based navigators and respiratory motion-corrected averaging. *Magn Reson Med*, 59(4):771–778, Apr 2008.

[83] S. Zhang, K. T. Block, and J. Frahm. Magnetic resonance imaging in real time: advances using radial FLASH. *J Magn Reson Imaging*, 31(1):101–109, Jan 2010.

[84] M. Uecker, S. Zhang, and J. Frahm. Nonlinear inverse reconstruction for real-time MRI of the human heart using undersampled radial FLASH. *Magn Reson Med*, 63(6):1456–1462, Jun 2010.

[85] S. Zhang, M. Uecker, D. Voit, K.-D. Merboldt, and J. Frahm. Real-time cardiovascular magnetic resonance at high temporal resolution: radial FLASH with nonlinear inverse reconstruction. *J Cardiovasc Magn Reson*, 12:39, 2010.

[86] M. Uecker, S. Zhang, D. Voit, A. Karaus, K.-D. Merboldt, and J. Frahm. Real-time MRI at a resolution of 20 ms. *NMR Biomed*, 23(8):986–994, Oct 2010.

[87] A. Fischer, F. A. Breuer, M. Blaimer, N. Seiberlich, and P. M. Jakob. Accelerated Dynamic Imaging by Reconstructing Sparse Differences using Compressed Sensing. In *Proceedings of the 16th Annual Meeting of the ISMRM, Toronto/ON, Canada*, page 341, 2008.

[88] M. Doneva, H. Eggers, J. Rahmer, P. Börnert, and A. Mertins. Highly Undersampled 3D Golden Ratio Radial Imaging with Iterative Reconstruction. In *Proceedings of the 16th Annual Meeting of the ISMRM, Toronto/ON, Canada*, page 336, 2008.

[89] A. Samsonov, Y. Jung, A. L. Alexander, W. F. Block, and A. S. Field. MRI Compressed Sensing via Sparsifying Images. In *Proceedings of the 16th Annual Meeting of the ISMRM, Toronto/ON, Canada*, page 342, 2008.

[90] M. F. Duarte, S. Sarvotham, M. B. Wakin, D. Baron, and R. G. Baraniuk. Joint Sparsity Models for Distributed Compressed Sensing. In *Online Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations (SPARS)*, Rennes, France, 2005.

[91] R. Otazo and D. K. Sodickson. Distributed Compressed Sensing for Accelerated MRI. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 378, 2009.

[92] P. Lai, M. Lustig, A. C. Brau, S. Vasanawala, P. J. Beatty, and M. Alley. Efficient L$_1$-SPIRiT Reconstruction (ESPIRiT) for Highly Accelerated 3D Volumetric MRI with Parallel Imaging and Compressed Sensing. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 345, 2010.

[93] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin. Adaptive reconstruction of phased array MR imagery. *Magn Reson Med*, 43(5):682–690, May 2000.

[94] Jeffrey A. Fessler. http://www.eecs.umich.edu/∼fessler/code/.

[95] T. Kampf, A. Fischer, T. C. Basse-Lüsebrink, G. Ladewig, F. Breuer, G. Stoll, P. M. Jakob, and W. R. Bauer. Application of Compressed Sensing to *in vivo* 3D ^{19}F CSI. *J Magn Reson*, 207:262–273, 2010.

[96] M. Lustig and J. M. Pauly. Iterative GRAPPA: A General Solution for the GRAPPA Reconstruction from Arbitrary k-Space Sampling. In *Proceedings of the 15th Annual Meeting of the ISMRM, Berlin, Germany*, page 333, 2007.

[97] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans Med Imaging*, 10(3):473–478, 1991.

[98] A. Fischer, N. Seiberlich, M. Blaimer, P. M. Jakob, F. A. Breuer, and M. A. Griswold. A Combination of Nonconvex Compressed Sensing and GRAPPA (CS-GRAPPA). In *ISMRM Workshop on Data Sampling and Image Reconstruction, Sedona/AZ, USA*, 2009.

[99] A. Fischer, N. Seiberlich, M. Blaimer, P. M. Jakob, F. A. Breuer, and M. A. Griswold. A Combination of Nonconvex Compressed Sensing and GRAPPA (CS-GRAPPA). In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 2813, 2009.

[100] Sven Vaegler. *Schnelle MR-Bildgebung der Zahnoberflächen mit Compressed Sensing (CS) (german)*. Diploma thesis, University of Wuerzburg, 2010.

[101] M. Doneva, P. Börnert, H. Eggers, C. Stehning, J. Sénégas, and A. Mertins. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magn Reson Med*, 64:1114–1120, 2010.

[102] L. V. Søgaard, T. Dorniok, F. Hengstenberg, S. Karpuk, J. Vestbo, P. Åkeson, and P. Magnusson. Compressed Sensing for hyperpolarized ^{3}He 3D ADC measurements. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 7, 2009.

[103] S. Ajraoui, K. J. Lee, J. Parra-Robles, M. H. Deppe, S. R. Parnell, and J. M. Wild. Compressed Sensing: Applications in Hyperpolarised ^{3}He lung MRI. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 2175, 2009.

[104] S. Ajraoui, K. J. Lee, M. H. Deppe, S. R. Parnell, J. Parra-Robles, and J. M. Wild. Compressed sensing in hyperpolarized ^{3}He lung MRI. *Magn Reson Med*, 63(4):1059–1069, Apr 2010.

[105] S. Ajraoui, J. Parra-Robles, H. Marshall, M. H. Deppe, S. R. Parnell, and J. M. Wild. Simultaneous acquisition of ^{3}He ventilation images, ADC, T_2^* and B_1 maps in a single scan with Compressed Sensing. In *Proceedings of the 17th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 201, 2010.

[106] H. Marshall, S. Ajraoui, and J. M. Wild. Golden Angle Radial Imaging for Improved Visualisation of Initial Stages of Inhalation in Dynamic ^{3}He Lung MRI. In *Proceedings of the 18th Annual Meeting of the ISMRM, Stockholm, Sweden*, page 2536, 2010.

[107] Olga Tymofiyeva. *Magnetic Resonance Imaging in Dental Medicine.* PhD thesis, University of Wuerzburg, 2010.

[108] M. Doneva, C. Stehning, P. Börnert, H. Eggers, and A. Mertins. Fast relaxation parameter mapping from undersampled data. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 385, 2009.

[109] M. Doneva, J. Sénégas, P. Börnert, H. Eggers, and A. Mertins. Accelerated MR parameter mapping using compressed sensing with model-based sparsifying transform. In *Proceedings of the 17th Annual Meeting of the ISMRM, Honolulu/HI, USA*, page 2812, 2009.

[110] D. Liang, B. Liu, J. Wang, and L. Ying. Accelerating SENSE using compressed sensing. *Magn Reson Med*, 62(6):1574–1584, Dec 2009.

[111] R. Otazo, D. Kim, L. Axel, and D. K. Sodickson. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magn Reson Med*, 64:767–776, Jun 2010.

[112] T. Wech, M. Gutberlet, A. Greiser, D. Stäb, C. O. Ritter, M. Beer, D. Hahn, and H. Köstler. High-Resolution Functional Cardiac MR Imaging using Density-Weighted Real-Time Acquisition and a Combination of Compressed Sensing and Parallel Imaging for Image Reconstruction. *Röfo*, 182(8):676–681, Aug 2010.

[113] D. H. Brandwood. A complex gradient operator and its application in adaptive array theory. *IEE Proceedings F Communications, Radar and Signal Processing*, 130(1):11–16, 1983.

[114] Ken Kreutz-Delgado. The Complex Gradient Operator and the CR-Calculus. *arxiv.org,* http://arxiv.org/abs/0906.4835v1, June 2009.

[115] Don Johnson. Optimization theory. *http://cnx.org/content/m11240/1.4/?format=pdf*, Mar 2008.

[116] Wilhelm Wirtinger. Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen. *Math Ann*, 97:357–376, 1927.

Publications related to this thesis

Papers

T. Kampf*, **A. Fischer***, T. C. Basse-Lüsebrink*, G. Ladewig, F. A. Breuer, G. Stoll, P. M. Jakob, W. R. Bauer:
Application of Compressed Sensing to *in vivo* 3D ^{19}F CSI
J Magn Reson, V. 207, pp. 262-273 (2010)

*: Authors contributed equally to this publication

Conference Talks

A. Fischer, F. A. Breuer, M. Blaimer, N. Seiberlich, P. Ehses, P. M. Jakob:
Beschleunigte radiale dynamische Bildgebung mit Hilfe von Compressed Sensing
10. Jahrestagung der Deutschen Sektion der ISMRM e.V., Würzburg, Germany, Abstract O6 (2007)

A. Fischer, F. A. Breuer, M. Blaimer, N. Seiberlich, P. M. Jakob:
Accelerated Dynamic Imaging by Reconstructing Sparse Differences using Compressed Sensing
16th Annual Meeting of the ISMRM, Toronto/ON, Canada, Abstract 341 (2008)

A. Fischer, N. Seiberlich, M. Blaimer, P. M. Jakob, F. A. Breuer, M. A. Griswold:
A Combination of Nonconvex Compressed Sensing and GRAPPA (CS-GRAPPA)
ISMRM Workshop on Data Sampling and Image Reconstruction, Sedona/AZ, USA (2009)

Electronic Posters

A. Fischer, M. A. Griswold, H. Jeong, T. J. Caroll, F. A. Breuer, P. M. Jakob:
Compressed Sensing in Dynamic Contrast-Enhanced Angiography
25th Annual Meeting of the ESMRMB, Valencia, Spain, Abstract 768 (2008)

A. Fischer, T. C. Basse-Lüsebrink, T. Kampf, G. Ladewig, M. Blaimer, F. A. Breuer, G. Stoll, W. R. Bauer, P. M. Jakob:
Improved Sensitivity in ^{19}F Cellular Imaging using Nonconvex Compressed Sensing
17th Annual Meeting of the ISMRM, Honolulu/HI, USA, Abstract 3377 (2009)

A. Fischer, N. Seiberlich, P. Ehses, M. A. Griswold, P. M. Jakob, F. A. Breuer:
Golden Angle radial cardiac imaging without ECG gating using nonconvex Compressed Sensing
19th Annual Meeting of the ISMRM, Montréal/QC, Canada, Abstract 4370 (2011)

Traditional Posters

A. Fischer, F. A. Breuer, M. Blaimer, N. Seiberlich, P. M. Jakob:
Introduction of a Nonconvex Compressed Sensing Algorithm to MR Imaging
16th Annual Meeting of the ISMRM, Toronto/ON, Canada, Abstract 1487 (2008)

M. Blaimer, **A. Fischer**, P. Ehses, N. Seiberlich, M. A. Griswold, P. M. Jakob, F. A. Breuer:
Simplified Iterative GRAPPA for fast and robust parallel MRI with arbitrary trajectories
17th Annual Meeting of the ISMRM, Honolulu/HI, USA, Abstract 2713 (2009)

A. Fischer, N. Seiberlich, M. Blaimer, P. M. Jakob, F. A. Breuer, M. A. Griswold:
A Combination of Nonconvex Compressed Sensing and GRAPPA (CS-GRAPPA)
17th Annual Meeting of the ISMRM, Honolulu/HI, USA, Abstract 2813 (2009)

T. C. Basse-Lüsebrink, T. Kampf, **A. Fischer**, G. Ladewig, G. Stoll, P. M. Jakob:
Spike artifact reduction in nonconvex Compressed Sensing
18th Annual Meeting of the ISMRM, Stockholm, Sweden, Abstract 4886 (2010)

T. C. Basse-Lüsebrink, **A. Fischer**, T. Kampf, V. Sturm, G. Ladewig, G. Stoll, P. M. Jakob:
^{19}F-Compressed-Sensing-CISS: Elimination of banding artifacts in ^{19}F bSSFP MRI/CSI without sacrificing time
18th Annual Meeting of the ISMRM, Stockholm, Sweden, Abstract 4888 (2010)

T. C. Basse-Lüsebrink, J. Beck, T. Kampf, **A. Fischer**, G. Stoll, P. M. Jakob:
Application of Compressed Sensing to ^{19}F turbo spin echo chemical shift imaging
19th Annual Meeting of the ISMRM, Montréal/QC, Canada, Abstract 1502 (2011)

F. A. Breuer, **A. Fischer**, N. Seiberlich, P. Ehses, M. Blaimer, D. Neumann, P. M. Jakob, M. A. Griswold:
Improved Compressed Sensing Reconstruction in Dynamic Contrast Enhanced MR Angiography by Means of Principal Component Analysis (PCA)
19th Annual Meeting of the ISMRM, Montréal/QC, Canada, Abstract 4379 (2011)

Further publications

Papers

E. D. Pracht, **A. Fischer**, J. F. T. Arnold, M. Kotas, M. Flentje, P. M. Jakob:
Single-shot quantitative perfusion imaging of the human lung
Magn Reson Med, V. 56, pp. 1347-1351 (2006)

A. Fischer, E. D. Pracht, J. F. T. Arnold, M. Kotas, M. Flentje, P. M. Jakob:
Assessment of pulmonary perfusion in a single shot using SEEPAGE
J Magn Reson Imaging, V. 27, pp. 63-70 (2008)

T. C. Basse-Lüsebrink, T. Kampf, **A. Fischer**, V. J. F. Sturm, D. Neumann, H. Köstler, D. Hahn, G. Stoll, P. M. Jakob:
SAR-reduced Spin-Echo-Based Bloch-Siegert B_1^+ Mapping: BS-SE-BURST
Magn Reson Med, V. 68, pp. 529-536 (2012)

T. Wech, D. Stäb, J. C. Budich, **A. Fischer**, J. Tran-Gia, D. Hahn, H. Köstler:
Resolution evaluation of MR images reconstructed by iterative thresholding algorithms for compressed sensing
Med Phys, V. 39, pp. 4328-4338 (2012)

Conference Talks

A. Fischer, E. D. Pracht, P. M. Jakob:
Single-shot perfusion-weighted imaging of the human lung
23rd Annual Meeting of the ESMRMB, Warsaw, Poland, Abstract 731 (2006)

A. Fischer, S. Weick, C. O. Ritter, D. Hahn, H. Köstler:
High resolution self-gated morphological and functional imaging of the human lung at 1.5 and 3.0 T
28th Annual Meeting of the ESMRMB, Leipzig, Germany, Abstract 66 (2011)

T. Wech, D. Stäb, **A. Fischer**, D. Hahn, H. Köstler:
On the Quality Evaluation for Images Reconstructed by Compressed Sensing
19th Annual Meeting of the ISMRM, Montréal/QC, Canada, Abstract 73 (2011)

Electronic Posters

A. Fischer, C. O. Ritter, M. Hagemeister, M. Oechsner, D. Hahn, H. Köstler:
Comparison of contrast agents for quantitative lung MRI at 1.5 T with respect to signal enhancement and dose
28th Annual Meeting of the ESMRMB, Leipzig, Germany, Abstract 454 (2011)

A. Gotschy, U. C. Hölscher, T. C. Basse-Lüsebrink, **A. Fischer**, M. Choli, T. Kampf, V. Sturm, D. Neumann, V. Herold, H. Köstler, D. Hahn, G. Stoll, W. R. Bauer, P. M. Jakob:
Rapid and Low SAR B1-Mapping Using a BURST-Based Bloch-Siegert-Shift Sequence
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 3365 (2012)

D. Stäb, **A. Fischer**, C. O. Ritter, D. Hahn, H. Köstler:
Truncation Artifacts and Their Impact on Morphological and Quantitative Lung Imaging
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 3965 (2012)

A. Fischer, C. O. Ritter, D. Hahn, H. Köstler:
Automatic Segmentation of Lung Parenchyma Using Fuzzy Clustering
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 3980 (2012)

A. Fischer, M. Gutberlet, J. Vogel-Claussen, C. O. Ritter, D. Hahn, H. Köstler:
Detection of functional lung defects using DC signal gated ^{1}H imaging
29th Annual Meeting of the ESMRMB, Lisbon, Portugal, Abstract 699 (2012)

Traditional Posters

A. Fischer, E. D. Pracht, P. M. Jakob:
Quantification of pulmonary perfusion using SEEPAGE
14th Annual Meeting of the ISMRM, Seattle/WA, USA, Abstract 1298 (2006)

K. Dara, J. J. Derakhshan, **A. Fischer**, S. R. Yutzy, N. Seiberlich, J. L. Duerk, M. A. Griswold:
True-SEEPAGE: A Tool for Evaluating Renal Perfusion and Function
17th Annual Meeting of the ISMRM, Honolulu/HI, USA, Abstract 2039 (2009)

A. Fischer, S. Weick, C. O. Ritter, D. Hahn, H. Köstler:
High Resolution Morphology, Ventilation, and Perfusion of the Human Lung by ^{1}H Imaging at 3.0 T
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 1339 (2012)

A. Fischer, C. O. Ritter, D. Hahn, H. Köstler:
Quantification of Pulmonary Perfusion: Comparison of DCE-MRI and SEEPAGE at 3.0 T
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 1342 (2012)

T. Wech, D. Stäb, **A. Fischer**, D. Hahn, H. Köstler:
Local Temporal Point Spread Function for CS Reconstructions Exploiting x-f-Sparsity
20th Annual Meeting of the ISMRM, Melbourne/VIC, Australia, Abstract 1342 (2012)