

# Nonlinear Reconstruction Methods for Parallel Magnetic Resonance Imaging

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen  
Doktorgrades

“Doctor rerum naturalium”

der Georg-August-Universität Göttingen

vorgelegt von

Martin Uecker  
aus Würzburg

Göttingen 2009

angefertigt in der

Biomedizinischen NMR Forschungs GmbH  
am Max-Planck-Institut für biophysikalische Chemie

unter Betreuung durch das

Institut für Numerische und Angewandte Mathematik  
der Georg-August-Universität Göttingen

D7

Referent: Prof. Dr. T. Hohage

Korreferent: Prof. Dr. J. Frahm

Tag der mündlichen Prüfung: 15.7.2009

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Magnetic Resonance Imaging</b>	<b>8</b>
2.1	Quantum Physics of the Nuclear Spin . . . . .	8
2.2	Relaxation Effects . . . . .	11
2.3	Signal Types . . . . .	12
2.4	Spatial Encoding . . . . .	15
2.5	The Mathematics of Image Reconstruction . . . . .	19
2.5.1	Discretization . . . . .	20
2.5.2	Fast Fourier Transform Algorithms . . . . .	21
2.6	Summary . . . . .	22
<b>3</b>	<b>Parallel Imaging</b>	<b>23</b>
3.1	Introduction . . . . .	23
3.2	Phased-Array Coils . . . . .	23
3.2.1	Whitening . . . . .	24
3.2.2	Array compression . . . . .	25
3.3	Undersampling of k-space . . . . .	26
3.4	Image Reconstruction . . . . .	28
3.4.1	Discretization . . . . .	28
3.4.2	Parallel Imaging as Linear Inverse Problem . . . . .	30
3.5	Calibration of the Coil Sensitivities . . . . .	31
3.6	Algorithms . . . . .	34
3.6.1	SENSE . . . . .	34
3.6.2	Conjugate Gradient Algorithm . . . . .	36
3.6.3	SMASH, AUTO-SMASH, GRAPPA . . . . .	38
3.7	Summary . . . . .	42

<b>4</b>	<b>MRI System</b>	<b>43</b>
4.1	Magnet and Gradient System . . . . .	43
4.2	Radio Frequency Coils . . . . .	44
4.3	Computer System and Software . . . . .	45
<b>5</b>	<b>Joint Estimation of Image Content and Coil Sensitivities</b>	<b>46</b>
5.1	Introduction . . . . .	46
5.2	Parallel Imaging as Nonlinear Inverse Problem . . . . .	47
5.3	Algorithm . . . . .	47
5.3.1	Iteratively Regularized Gauss Newton Method . . . . .	47
5.3.2	Regularization of the Coil Sensitivities . . . . .	51
5.3.3	Choice of Parameters . . . . .	52
5.3.4	Postprocessing . . . . .	53
5.3.5	Computational Speed . . . . .	53
5.4	Experiments . . . . .	54
5.4.1	Methods . . . . .	54
5.4.2	Results . . . . .	58
5.5	Extensions . . . . .	59
5.5.1	Partial Fourier Imaging . . . . .	59
5.5.2	Reconstruction with Reduced Field of View . . . . .	63
5.5.3	Non-Cartesian Trajectories . . . . .	65
5.6	Discussion . . . . .	66
5.7	Summary . . . . .	67
<b>6</b>	<b>Segmented Diffusion Imaging</b>	<b>68</b>
6.1	Introduction . . . . .	68
6.2	Theory . . . . .	69
6.2.1	Diffusion Tensor Imaging . . . . .	69
6.2.2	Segmented Diffusion Imaging . . . . .	71
6.2.3	Reconstruction Algorithm . . . . .	73
6.3	Materials and Methods . . . . .	75
6.4	Results . . . . .	76
6.5	Discussion . . . . .	81
6.6	Summary . . . . .	84
<b>7</b>	<b>Image Regularization</b>	<b>85</b>
7.1	Tikhonov Regularization . . . . .	85
7.2	Choice of the Regularization Parameter . . . . .	87



<i>CONTENTS</i>	5
7.3 $L^1$ Regularization . . . . .	87
7.3.1 Nonlinear Transformation into a Quadratic Penalty . . . . .	89
7.3.2 Total Variation . . . . .	90
7.4 Examples . . . . .	91
7.4.1 Parallel Imaging . . . . .	92
7.4.2 Compressed Sensing . . . . .	92
7.5 Summary . . . . .	92
<b>8 Summary and Outlook</b>	<b>94</b>
8.1 Summary . . . . .	94
8.2 Future Work . . . . .	96
<b>Properties of the Fourier Transform</b>	<b>98</b>
<b>Abbreviations</b>	<b>100</b>
<b>References</b>	<b>102</b>
<b>Curriculum Vitae</b>	<b>113</b>
<b>List of Publications</b>	<b>114</b>
<b>Acknowledgements</b>	<b>117</b>

# 1

## Introduction

Magnetic resonance imaging (MRI) is a non-invasive method for cross-sectional imaging of humans and animals with a wide range of applications in both clinical radiology and biomedical research. A special emphasis is on the basic and clinical neurosciences. Advanced techniques comprise functional MRI (fMRI) of brain activation and diffusion tensor imaging (DTI) of white matter fiber architecture. Because such applications depend on much more information than usually required for the calculation of a structural image, the development of faster acquisition techniques is an ongoing research topic of utmost importance.

The relatively low acquisition speed of MRI is caused by the use of a point-by-point scanning scheme in Fourier space. This is in contrast to optical imaging methods which are inherently parallel and typically acquire whole images at the same time. A recent development in MRI is the adaptation of a *parallel imaging* concept which makes use of multiple receive coils to acquire data in parallel. As we will see, the reconstruction of such data requires the inversion of an ill-posed system and a very accurate calibration of the sensitivities from the individual receive coils. The main contribution of this thesis will be the development of a new algorithm that solves the *combined* coil calibration and image estimation. The idea is based on the formulation of autocalibrated parallel imaging in MRI as a nonlinear inverse problem. This reconstruction problem is then solved with a Newton-based regularization method. While conventional algorithms such as generalized autocalibrating partially parallel acquisitions (GRAPPA) and autocalibrating variants of sensitivity encoding (SENSE) are often limited by a miscalibration of the coil sensitivities, the new algorithm is only limited by the noise amplification that arises due to the bad

conditioning of the reconstruction problem.

Building on this algorithm, and in addition to its direct application for improved parallel MRI, this thesis also contributes a new reconstruction technique for segmented diffusion-weighted MRI. Here, conventional MRI reconstructions fail due to the presence of uncontrollable phase variations in the segmented data sets. These errors are caused by unavoidable brain pulsations and amplified by the strong diffusion gradients. They lead to unacceptable motion artifacts in images that represent a simple combination of the data from different segments. For this reason, it is proposed to determine high-resolution phase maps for each individual segment by parallel imaging, while taking advantage of the improved coil sensitivity calibration by nonlinear inversion.

A general limitation of parallel imaging is due to the bad conditioning of the inverse problem. Because quadratic regularization techniques are unable to distinguish between signal and noise, a trade-off between artifact power and noise has to be made. New non-quadratic regularization techniques derived from the  $L^1$  norm are able to suppress the noise much better. This thesis therefore presents a new idea about strategies that integrate such techniques into algorithms that are based on the iteratively regularized Gauss-Newton method. The development of these new regularization techniques naturally links to the emerging theory of *compressed sensing*. Here, the combination of non-quadratic regularization techniques and an appropriate encoding of the information allows for a sparse representation of the measurement object to be inferred from a reduced set of measurements, even without the additional information from multiple receive coils.

In summary, this work explores and extends the concept of parallel imaging in MRI. Despite being an established technique in clinical practice, parallel MRI acquisitions can only fully be exploited when based on novel mathematical algorithms. This thesis introduces some basic solutions.

# 2

## Magnetic Resonance Imaging

Atomic nuclei consist of protons and neutrons which are fermionic particles of spin  $1/2$ . Spin is angular momentum related to an internal degree of freedom. While for some atomic nuclei the total sum of the spin of all components is zero, some nuclei have residual spin such as  $^1H$ ,  $^{13}C$ ,  $^{15}N$ ,  $^{19}F$ , and  $^{31}P$ . The most important nucleus for MRI is the nucleus of the hydrogen atom  $^1H$  which consists of a single proton with spin  $1/2$ . The nuclear spin creates a magnetic moment  $\boldsymbol{\mu}$  connected to the spin  $\boldsymbol{S}$  by the gyromagnetic ratio  $\gamma$  according to

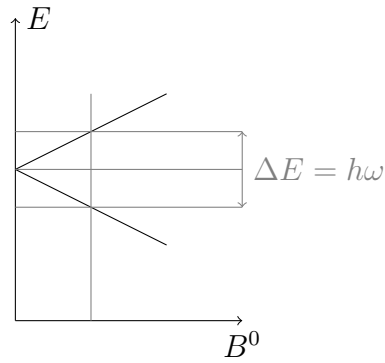
$$\boldsymbol{\mu} = \gamma \boldsymbol{S} \ .$$

The quantity measured in nuclear magnetic resonance (NMR) experiments is the electric current induced in a nearby coil by the rotating magnetic moment of a large ensemble of excited spins. The NMR effect in condensed matter was discovered by Purcell and Bloch in 1946 [87, 10].

### 2.1 Quantum Physics of the Nuclear Spin

The quantum mechanical description of a single isolated proton spin is given by a two-states system. Mathematically, the state of such a system is described by a normed vector  $\psi$  in a two-dimensional complex Hilbert space  $\mathbb{C}^2$ . Quantum mechanical observables are given by self-adjoint operators defined on this space. The expectation value of an observable  $A$  in a given state  $\psi$  is

$$\langle A \rangle = \langle \psi, A\psi \rangle \ .$$



**Figure 2.1:** *Splitting of the energy levels for a proton in an external magnetic field.*

The observables related to the three components of the spin of a spin 1/2 particle are the Pauli matrices

$$\sigma_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

The time derivative of the expectation value of an observable  $A$  can be computed with the commutator  $[\cdot, \cdot]$  with the Hamilton operator  $H$ :

$$\frac{d}{dt} \langle A \rangle = \frac{i}{\hbar} \langle [H, A] \rangle$$

The Hamilton operator for a particle with 1/2 spin contains an interaction term which describes the coupling of the magnetic field with the magnetic moment of the spin:

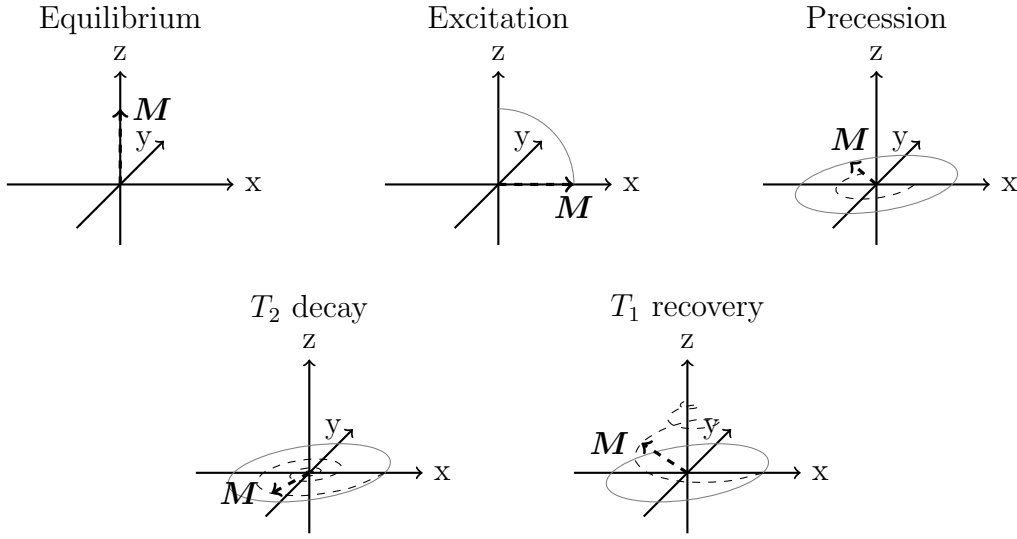
$$H = -\boldsymbol{\mu} \cdot \mathbf{B}$$

With the Hamilton operator the dynamical behaviour of the expectation values of the spin observables can be derived for arbitrary time-dependent magnetic fields (see for example [45]):

$$\frac{d}{dt} \langle \boldsymbol{\mu} \rangle = \gamma \langle \boldsymbol{\mu} \rangle \times \mathbf{B} \quad (2.1)$$

In NMR experiments the magnetic field consists of a static component  $\mathbf{B}^0$  which lies - by convention - parallel to the z-axis of the coordinate system. Due to the coupling of the nuclear magnetic spin with the magnetic field, the two energy levels split up with increasing field strength as described by the two discrete eigenvalues  $E_{\pm} = \pm \gamma \frac{\hbar}{2} B^0$  of the Hamilton operator. In thermodynamical equilibrium the occupation number of both energy levels is given according to the Boltzmann equation

$$\frac{N_+}{N_-} = e^{-\frac{\Delta E}{k_B T}} .$$



**Figure 2.2:** Pulsed MR experiment: In equilibrium, the magnetic moments align themselves along the static magnetic field  $B^0$ . After excitation with a RF pulse the spins are tilted into the  $xy$ -plane and precess with the Larmor frequency.

Here,  $k_B$  is the Boltzmann constant and  $T$  the temperature. A slightly larger occupation of the lower energy level causes an equilibrium magnetization

$$M_{eq} = \rho \frac{\gamma^2 \hbar^2}{4kT} B^0$$

parallel to the static  $B^0$  field ( $\rho$  the spin density). Associated to the energy difference is a characteristic resonance frequency, the *Larmor frequency*  $\omega_0 = \gamma B^0$ . In pulsed NMR experiments, a radio frequency (RF) field  $\mathbf{B}^1$  with frequency  $\omega_0$  is used to excite the spins. The RF field vector lies in the  $xy$ -plane, perpendicular to the  $B^0$  field:

$$\mathbf{B}(t) = \mathbf{B}^0 + \mathbf{B}^1(t) = \begin{pmatrix} 0 \\ 0 \\ B_z^0 \end{pmatrix} + \begin{pmatrix} B^1(t) \sin(\omega t) \\ B^1(t) \cos(\omega t) \\ 0 \end{pmatrix}$$

During excitation, the magnetization vector is tilted towards the  $xy$ -plane where it precesses with  $\omega_0$ . The *flip angle* is proportional to the integral over the envelope  $B^1(t)$  of the pulse (see [71]). A pulse which rotates the magnetization vector by a certain flip angle  $\alpha$  will be called an  $\alpha$ -pulse (e.g., a  $90^\circ$ -pulse) in the following.

## 2.2 Relaxation Effects

In 1946, the dynamical equations have been extended by Bloch with two phenomenological terms [10]. The *Bloch equation*

$$\frac{d}{dt}\mathbf{M} = \gamma\mathbf{M} \times \mathbf{B} + \begin{pmatrix} -\frac{1}{T_2}M_x \\ -\frac{1}{T_2}M_y \\ \frac{M_{eq}-M_z}{T_1} \end{pmatrix} \quad (2.2)$$

describes the relaxation toward thermodynamical equilibrium. The constant  $T_1$  is the spin-lattice relaxation time which describes the relaxation of the longitudinal magnetization caused by energy exchange with the surrounding environment. The constant  $T_2$  is called spin-spin relaxation time and describes the loss of transversal magnetization. Because this decrease in transversal magnetization is not only caused by the exchange of energy with the environment but also by energy exchange between spins,  $T_2$  is smaller than  $T_1$ .

To simplify the analysis of the experiments, the two components of the transversal magnetization are usually combined to one single complex-valued quantity

$$M_{\perp} = M_x(t) + iM_y(t) \quad .$$

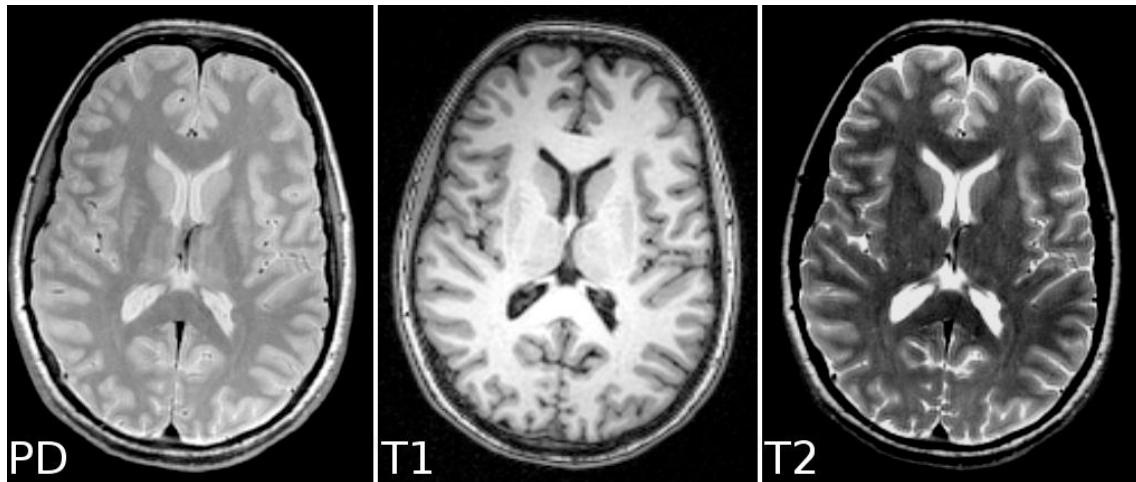
Note that for  $B_x = B_y = 0$  equation (2.2) implies that  $M_{\perp}$  satisfies the differential equation  $\frac{d}{dt}M_{\perp} = -(i\gamma B_z(t) + \frac{1}{T_2})M_{\perp}(t)$  with explicit solution:

$$M_{\perp}(t) = M_{\perp}(0)e^{-t/T_2 - i\gamma \int_0^t B_z(t') dt'} \quad (2.3)$$

$$M_z(t) = M_{eq} + (M_z(0) - M_{eq})e^{-t/T_1} \quad (2.4)$$

In the absence of the  $\mathbf{B}^1$  field the solution of the Bloch equations (2.2) is given by an exponential decline of the rotating transversal magnetization with  $T_2$  time and an exponential return of the longitudinal magnetization to the equilibrium value  $M_{eq}$  with  $T_1$ .

The basic MR experiment, which goes back to Hahn in 1950 [43], is the following: The sample is placed in a strong static magnetic field. Excitation with a RF field near the resonance frequency  $\omega_0$  of the system will rotate the spins into the  $xy$ -plane. The angle of this rotation depends on the integral of the modulating function of the applied RF pulse. After applying a  $90^\circ$  pulse, the magnetic moment lies in the  $xy$ -plane and precesses with the Larmor frequency. This rotating magnetic moment induces an electric current in nearby coils, which constitutes the basic signal in MR spectroscopy and imaging. The exponential loss of transversal magnetization



**Figure 2.3:** MR images with different contrast. From left to right: proton density,  $T_1$ -weighted, and  $T_2$ -weighted images.

corresponds to a rapid decline of the received signal. Only after the longitudinal magnetization has recovered, the experiment can be repeated.

The influence of the relaxation constants can be controlled by the parameters of the NMR (or MRI) experiment. The signal generated depends not only on the amount of proton spins in a given volume, but also on the time past after excitation. In this way, the  $T_2$  relaxation directly modulates the generated signal. Because often a repeated series of single excitations is used to create the signals, also the  $T_1$  relaxation time can influence the signal strength. If the repetition time (TR) is too short, the magnetization does not reach the full equilibrium value before the next excitation, and the signal depends on the  $T_1$  relaxation processes. The influence of the relaxation processes on MRI images is demonstrated in Figure 2.3.

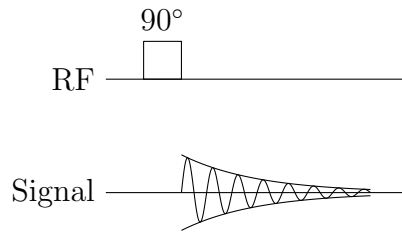
## 2.3 Signal Types

The basic techniques for generating a useful signal (an echo) will be described next.

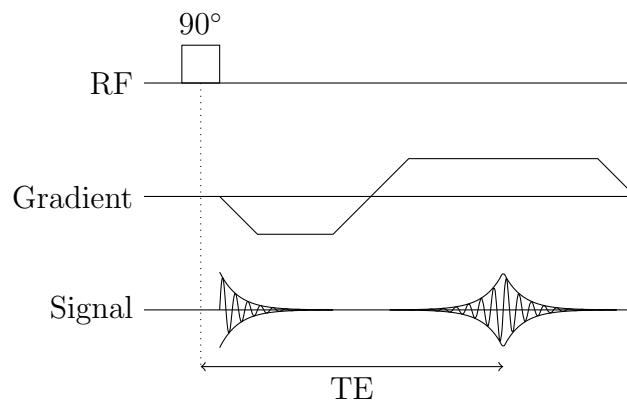
**Free Induction Decay** After excitation with a RF pulse the excited spins send out a signal with Larmor frequency. Because local field inhomogeneities contribute to the dephasing of the spins, the signal decays exponentially with an effective spin-spin relaxation time ( $T_2^*$ ), which is somewhat smaller than  $T_2$ .

**Gradient Echo** After excitation, the spin are dephased by an additional field gradient. Due to this field gradient the signal decays rapidly. For a gradient-recalled

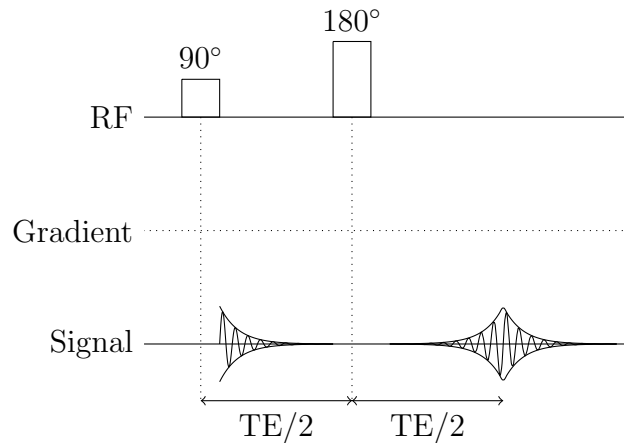




**Figure 2.4:** *Free induction decay: The signal send out from the excited spins decays according to  $T_2$  relaxation.*



**Figure 2.5:** *Gradient echo: After dephasing of the spins with a gradient, an echo can be created by rephasing with a gradient of opposite polarity.*

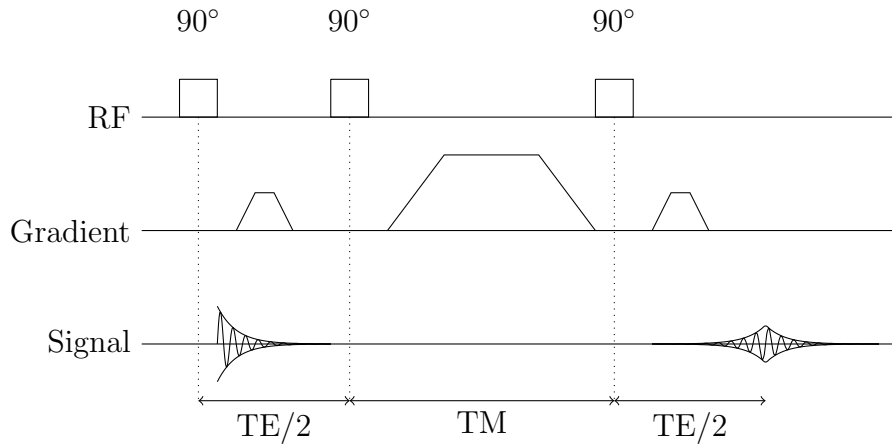


**Figure 2.6:** *Spin echo: Dephased spins are rephased with a  $180^\circ$  pulse which rotates all spins by  $180^\circ$  around the  $x$ -axis. The phase of the transversal magnetization is then exactly inverted.*

echo a gradient of opposite polarity is switched on which leads to a rephasing of the spins and the creation of an echo (see Figure 2.5). Because the spins are only partially rephased due to the combined effect of field inhomogeneities and  $T_2$  relaxation, the signal strength depends on the echo time (TE) according to  $\exp(-TE/T_2^*)$ . The gradient echo is the base of the fast low angle shot (FLASH) imaging sequence [33, 35, 41].

**Spin Echo** The spin echo was described first in 1950 by Hahn [44]. Similar to a gradient echo, dephased spins are refocused to obtain an echo. In contrast to gradient echoes, spin echoes are created by inverting the transversal magnetization vector with an additional  $180^\circ$  RF pulse. In this way, the accumulated phase differences between different spins subject to different magnetic fields are inverted. After the same time the spins are again exactly in phase and create a so-called spin echo. For the spin echo phase differences created by the gradient as well as by local field distortions are compensated. Therefore, the signal amplitude depends on the  $T_2$  relaxation time and not on  $T_2^*$ .

**Stimulated Echo** The stimulated echo was described by Hahn in 1950 [44]. The stimulated echo is created with the use of three RF pulses. The crucial property of the stimulated echo is the fact that the phase state of the excited spins is frozen after the second pulse and restored after the third. This is done by rotating the transversal magnetization into the longitudinal orientation with the second pulse.



**Figure 2.7:** *Stimulated echoes are created by three  $90^\circ$  pulses. After the second pulse the prepared transversal phase state is rotated into the longitudinal direction. In this state it is unaffected by  $T_2$  relaxation and off-resonance effects.*

Here, it is not affected by  $T_2$  relaxation and off-resonance effects anymore. Instead, it is subject to the much slower  $T_1$  attenuation. Fast imaging is possible by replacing the third  $90^\circ$  pulse with a series of small pulses, splitting the prepared magnetization into series of smaller echoes. Spectroscopy and imaging with stimulated echoes was first described by Frahm et al. [35, 34].

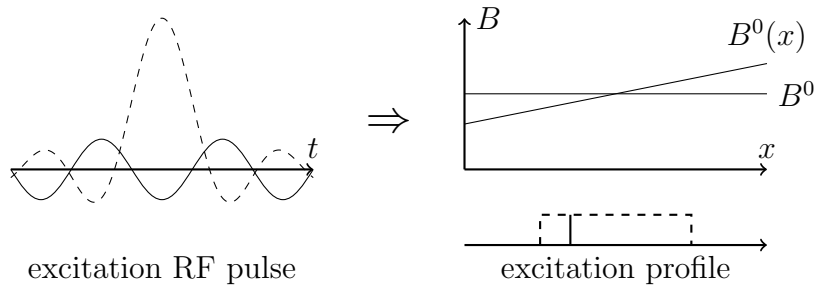
## 2.4 Spatial Encoding

To discern the signals from spins located at different positions in the sample some kind of spatial encoding has to be used. There are two basic principles which are commonly used in MRI: slice selection, where only a single slice is excited, and Fourier encoding, which can be used to encode the signal of the excited spins. The two techniques are complementary: In 2D imaging, a slice is selected and the remaining two dimensions are Fourier encoded, while in 3D imaging only Fourier encoding is used for all three dimensions.

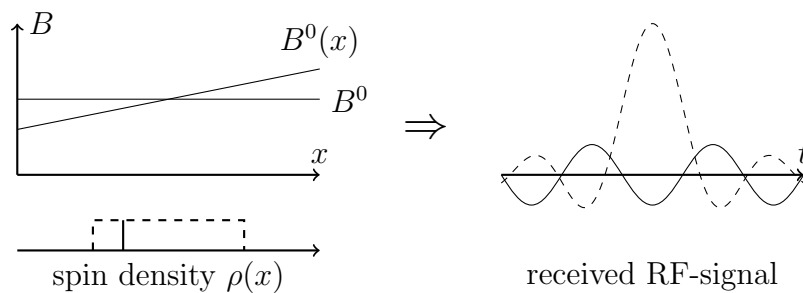
To excite only a slice of the sample an additional field gradient  $\mathbf{G}$  is switched on during the excitation pulse, giving a constant magnetic field which varies in space according to

$$\mathbf{B}^0(x) = \mathbf{B}_0^0 + \mathbf{G} \cdot x \quad .$$

In this way, the resonance frequency of the spins varies linearly along the gradient direction  $\omega(x) = \gamma B^0(x)$ . The application of a pulse which can be decomposed



**Figure 2.8:** *Slice selection: A slice selection gradient leads to a linearly varying resonance frequency of the spins. A sinc pulse then ideally excites a slice with rectangular profile.*



**Figure 2.9:** *Fourier encoding: A readout gradient leads to a linearly varying resonance frequency of the spins. The signal is related to the Fourier transform of the spatial profile of the transversal magnetization.*

into many frequencies, will in good approximation only excite the spins with the corresponding resonance frequencies (see [71]). For example, the application of a *sinc* pulse where the Fourier transform consists of a continuous block of frequencies between a lower and a higher limit excites only the corresponding region of spins located between a respective lower and higher position along the gradient direction. For practical reasons the pulse is cut off to a finite time which leads to an imperfect slice profile. After selectively exciting a thin slice of spins, only this slice creates a signal that needs to be encoded during the experiment.

For *frequency encoding* a gradient is turned on during the acquisition of the signal. Again, the resonance frequency depends on the position of the spin. Ignoring relaxation effects, the received signal is proportional to a superposition of these different resonance frequencies. The position of a spin is now encoded in the frequency of the received signal. Because only one dimension can be encoded in this simple way, frequency encoding has to be generalized by controlling the phase of the transversal magnetization by more complex gradient switching schemes. In the

following, the gradient may also vary in time:

$$\mathbf{B}^0(t, \mathbf{x}) = \mathbf{B}_0^0 + \mathbf{G}(t) \cdot \mathbf{x} \quad (2.5)$$

In this general situation, the phase of the transversal magnetization can be described by a k-space formalism, which will be derived in the following. According to (2.3), the phase depends on the time integral over the gradients:

$$M_{\perp}(t, \mathbf{x}) = M_{\perp}(0, \mathbf{x}) e^{-i(\omega_0 t + \gamma \mathbf{x} \cdot \int_0^t dt \mathbf{G}(t'))} \quad \text{with} \quad \omega_0 = \gamma B_0^0$$

To describe the time evolution of the gradient induced spatial phase variations, the k-space trajectory  $\mathbf{k}(t)$  is defined as

$$\mathbf{k}(t) := \frac{\gamma}{2\pi} \int_0^t d\tau \mathbf{G}(\tau) \quad .$$

The real-valued signal created in a large coil surrounding the sample (oriented with its symmetry axis parallel to the x-axis) can be calculated by integrating the magnetization over the complete volume:

$$\Re \int d\mathbf{x} M_{\perp}(\mathbf{x}, t) = \Re \int d\mathbf{x} M_{\perp}(\mathbf{x}, 0) e^{-2\pi i \mathbf{k}(t) \cdot \mathbf{x}} e^{-i\omega_0 t}$$

By quadrature demodulation the high frequency phase term  $e^{-i\omega_0 t}$  is removed, and real and imaginary parts of the remaining expression can be determined:

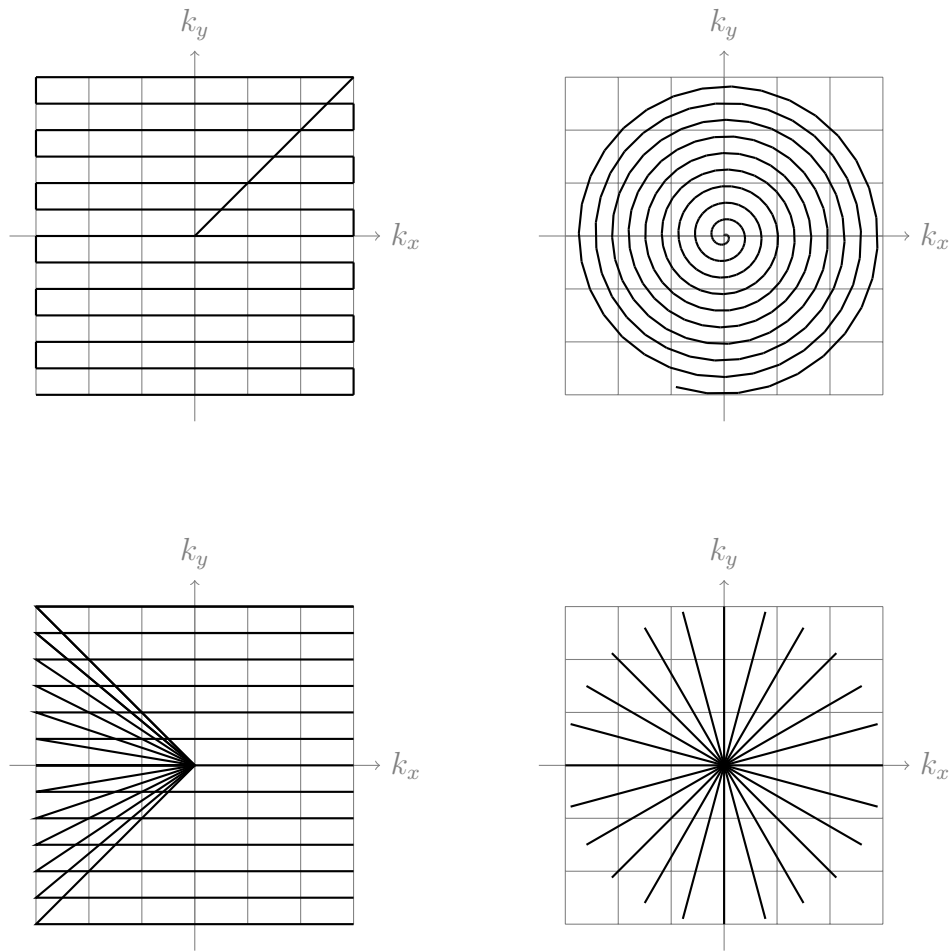
$$s(t) = \int d\mathbf{x} M_{\perp}(\mathbf{x}, 0) e^{-2\pi i \mathbf{k}(t) \cdot \mathbf{x}}$$

Assuming that the initial magnetization is directly proportional to the spin density  $\rho$ , the signal is then proportional to the Fourier transform of  $\rho$ , sampled on a k-space trajectory  $\mathbf{k}(t)$ :

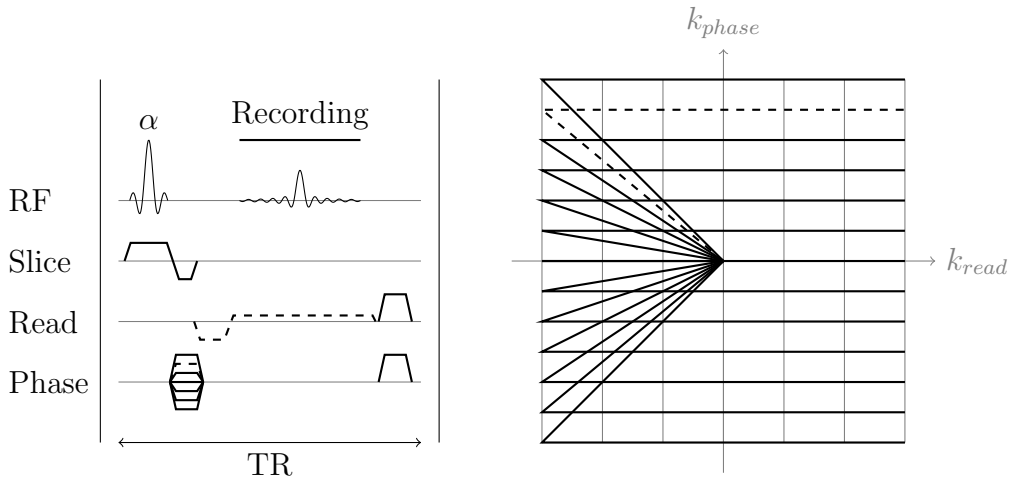
$$s(t) \propto \int_V d\mathbf{x} \rho(\mathbf{x}) e^{-i2\pi \mathbf{x} \cdot \mathbf{k}(t)}$$

Slight inhomogeneities in the  $\mathbf{B}^0$ -field as well as various other effects lead to derivations from this ideal signal equation.

In principle nearly arbitrary k-space trajectories can be used. The most important ones are shown in Figure 2.10. In practice, the use of a long trajectory is problematic because the  $T_2$  relaxation and off-resonance effects caused by field inhomogeneities lead to blurring and phase variations in the reconstructed images. Nevertheless, techniques such as echo planar imaging (EPI) [29] and spirals are often used when imaging speed is important. In the case of EPI, off-resonance effects lead to distortions primarily in the direction perpendicular to the long line elements



**Figure 2.10:** Typical  $k$ -space trajectories in MRI (clock-wise from upper left): echo-planar, spiral, radial, and Cartesian encoding schemes.



**Figure 2.11:** (Left) Timing diagram of a generic FLASH MRI sequence comprising the switching of a slice-selection, phase-encoding, and frequency-encoding gradient as well as spoiler gradients and the RF pulse; (right) corresponding sampling trajectory.

of the trajectory. The effects of spiral trajectories are reviewed in [12]. To avoid these problems, most image acquisition techniques use new magnetization for each k-space line. Residual magnetization is then often dephased (spoiled) by the use of a gradient after the acquisition of a line, so that its signal does not disturb the acquisition of the following lines. Hence, the use of line-by-line scanning with a Cartesian sampling scheme avoids most effects due to field inhomogeneities and relaxation and therefore allows for a simple reconstruction with a Fourier transform (FT). In such trajectories, the direction of the k-space lines is called read direction, while the directions perpendicular to the lines are called phase-encoding directions. Still, for very fast acquisition radial trajectories are attractive because they combine some of the advantages of line-by-line scanning with better undersampling behaviour and motion robustness. Radial trajectories are discussed in [11].

## 2.5 The Mathematics of Image Reconstruction

The image reconstruction problem in MRI can be stated as: Find a function  $\rho \in L^2(\mathbf{R}^d)$  ( $d \in \{2, 3\}$  for 2D or 3D imaging, respectively) with support in a given compact region  $\Omega \subset \mathbf{R}^d$  (the field of view (FOV)) and a Fourier transform which matches the measured data  $y$ :

$$y = P_{\mathbf{k}} \mathcal{F} \rho \quad \text{supp } \rho \subset \Omega \quad (2.6)$$

Here,  $\mathcal{F}$  denotes the Fourier transform and  $P_{\mathbf{k}}$  the projection defined by the restriction onto the measured k-space samples. Because the trajectory  $\mathbf{k}(t)$  is again sampled on discrete time points, the Fourier transform of  $\rho$  is only known on a discrete set of k-space samples, typically on a finite area around the origin of a Cartesian grid. Because the solution to this reconstruction problem is not unique, it is common practice to choose the one with minimal  $L^2$ -norm. In general, this is justified by the assumption that the missing higher frequencies are small. Basic properties of the Fourier transform are discussed in the appendix.

### 2.5.1 Discretization

Most MRI acquisitions sample the k-space in a finite area on a Cartesian grid described by

$$\Gamma_{\mathbf{N}} = \{ \mathbf{n} \in \mathbb{Z}^d : N_1/2 \leq n_1 < N_1/2, \dots, N_d/2 \leq n_d < N_d/2 \} \quad .$$

For simplicity, the size of the dimensions of the grid  $\mathbf{N} = (N_1, \dots, N_d)$  are all assumed to be even. The FOV is assumed to be a quadratic region  $Q_R = (-R/2, R/2)^d$  of size  $R$ . Let  $\rho$  be the real continuous object, then the ideal (noiseless) sample values are given by

$$y_{\mathbf{n}} = (\mathcal{F}\rho) \left( \frac{\mathbf{n}}{R} \right) \quad \text{for } \mathbf{n} \in \Gamma_{\mathbf{N}} \quad .$$

Let  $\rho_{per}$  denote the  $R$ -multiperiodic function  $\rho_{per}(\mathbf{x}) := \sum_{\mathbf{n} \in \mathbb{Z}^d} \rho(\mathbf{x} - R\mathbf{n})$ . Given the standard orthonormal basis  $\phi_{\mathbf{n}}(\mathbf{x}) = R^{-d} e^{\frac{2\pi i}{R} \mathbf{n} \cdot \mathbf{x}}$  of  $L^2(Q_R)$  it can be expanded into a Fourier series

$$\rho_{per} = \sum_{\mathbf{n} \in \mathbb{Z}^d} \hat{\rho}_{per}(\mathbf{n}) \phi_{\mathbf{n}} \quad \text{with} \quad \hat{\rho}_{per}(\mathbf{n}) = \langle \rho_{per}, \phi_{\mathbf{n}} \rangle_{L^2(Q_R)} \quad .$$

The Fourier coefficients of low order are given directly in terms of the measured data by

$$\hat{\rho}_{per}(\mathbf{n}) = R^{-d} y_{\mathbf{n}} \quad \text{for } \mathbf{n} \in \Gamma_{\mathbf{N}} \quad .$$

With these coefficients the orthogonal projection

$$P_{\mathbf{N}} \rho_{per} := \sum_{\mathbf{n} \in \Gamma_{\mathbf{N}}} \langle \rho_{per}, \phi_{\mathbf{n}} \rangle \phi_{\mathbf{n}}$$

can be defined.  $P_{\mathbf{N}} \rho_{per}$  is the best  $L^2$  approximation in span  $\{ \phi_{\mathbf{n}} : \mathbf{n} \in \Gamma_{\mathbf{N}} \} \subset L^2(Q_R)$  and is the desired solution of (2.6) with minimal  $L^2$  norm in  $L^2(Q_R)$ . Given



the assumption  $\text{supp } \rho \subset Q_R$ , the function  $\rho_{per}$  coincides with  $\rho$  on  $Q_R$ , and the truncation error is then given by the norm of the missing high frequencies

$$\|\rho_{per} - P_{\mathbf{N}}\rho_{per}\|_{L^2(Q_R)}^2 = \sum_{\mathbf{n} \in \mathbb{Z}^d \setminus \Gamma_{\mathbf{N}}} |\hat{\rho}_{per}(\mathbf{n})|^2 .$$

Should the assumption be violated, then  $\rho_{per}$  and its approximation  $P_{\mathbf{N}}\rho_{per}$  contain aliasing artifacts. This fact is related to the *Nyquist-Shannon sampling theorem* [95]. A discrete Fourier transform (DFT) yields the values of  $P_{\mathbf{N}}\rho_{per}$  at the nodal points  $\{\frac{R}{N}\mathbf{n} : \mathbf{n} \in \Gamma_{\mathbf{N}}\}$ , which are presented to the operator as the reconstructed image. A fast computation of the DFT is possible with fast Fourier transform (FFT) algorithms, as discussed in the next section.

## 2.5.2 Fast Fourier Transform Algorithms

A DFT amounts to the evaluation of the sum

$$FT_k^N \{f_n\}_{n=0, \dots, N-1} := \hat{f}_k = \sum_{n=0}^{N-1} e^{i\frac{2\pi}{N}kn} f_n \quad \text{for } k = 0, \dots, N-1 . \quad (2.7)$$

Because a direct evaluation would be quite expensive, the use of fast algorithms to calculate the DFT is required for MRI. Such fast algorithms are called FFTs and reduce the complexity of a DFT of size  $N$  from  $\mathcal{O}(N^2)$  for a direct evaluation to only  $\mathcal{O}(N \log N)$  multiplications. The best known FFT is the Cooley-Tukey [25] algorithm. It decomposes a DFT of size  $N = N_1 N_2$  into smaller DFT of sizes  $N_1$  and  $N_2$ . With the decomposition of the indices  $k = k_2 + k_1 N_2$  and  $n = n_2 N_1 + n_1$  and the shortcut  $\xi_N$  for an  $N$ -th root of the unit of highest order, i.e.  $(\xi_N)^N = 1$  and  $(\xi_N)^k \neq 1$  for  $k \in \{1, \dots, N-1\}$ , the following simple algebraic relation can be proved:

$$\begin{aligned} (\xi_N)^{kn} &= (\xi_{N_1 N_2})^{(k_2 + k_1 N_2)(n_2 N_1 + n_1)} \\ &= (\xi_{N_1 N_2})^{N_2 k_1 n_1} (\xi_{N_1 N_2})^{k_2 n_1} (\xi_{N_1 N_2})^{N_1 k_2 n_2} (\xi_{N_1 N_2})^{N_1 N_2 k_1 n_2} \\ &= (\xi_{N_1})^{k_1 n_1} (\xi_{N_1 N_2})^{k_2 n_1} (\xi_{N_2})^{k_2 n_2} \end{aligned}$$

In the last step the rules  $\xi_{AB}^A = \xi_B$  and  $\xi_A^A = 1$  have been used. Using this relation the derivation of a divide and conquer algorithm for the Fourier transform is

straightforward:

$$\begin{aligned}
FT_k^N \{f_n\}_{n=0, \dots, N-1} &= \sum_{n=0}^{N-1} (\xi_N)^{kn} f_n \\
&= \sum_{n_1=0}^{N_1-1} (\xi_{N_1})^{k_1 n_1} (\xi_N)^{k_2 n_1} \sum_{n_2=0}^{N_2-1} (\xi_{N_2})^{k_2 n_2} f_{n_1+n_2 N_1} \\
&= FT_{k_1}^{N_1} \left\{ (\xi_N)^{k_2 n_1} FT_{k_2}^{N_2} \{f_{n_1+n_2 N_1}\}_{n_2=0, \dots, N_2-1} \right\}_{n_1=0, \dots, N_1-1}
\end{aligned}$$

This recursive application reduces all DFTs to prime-sized DFTs. Because a DFT of size two is trivial, efficient computation for all powers of two is directly possible. A DFT can be re-expressed as a convolution, a fact that can be used to implement a prime-sized DFT with the help of an FFT of a different size [16, 89]. In this way, efficient algorithms for all  $N$  can be constructed. Higher dimensional DFTs can be decomposed into lower dimensional transforms in various ways.

## 2.6 Summary

The dynamical behaviour of the magnetic moment of a proton spin in external magnetic fields can be derived from a quantum mechanical description. In a strong static magnetic field  $\mathbf{B}^0$  the magnetization of an ensemble of spins acquires an equilibrium magnetization parallel to the  $\mathbf{B}^0$  field. Associated with the static magnetic field is a characteristic resonance frequency, the Larmor frequency  $\omega = \gamma B^0$ . By exciting spins with a RF pulse  $\mathbf{B}^1$  at the resonance frequency the spins are tilted towards the plane perpendicular to the direction of the  $B^0$  field where they start to precess, again with frequency  $\omega_0$ . The signal measured in MRI is related to the current in a coil which is induced by this rotating transversal magnetization, expressed as a complex-valued quantity. The return of the magnetization to its equilibrium value is described by the Bloch equation which phenomenologically describes the important  $T_1$  and  $T_2$  relaxation effects. By including time-varying field gradients to the static  $B^0$  field it is possible to manipulate the position-dependent phase state of the rotating transversal magnetization. This can be described with the k-space formalism, and is exploited in imaging experiments to acquire discrete samples of the Fourier transform of the spatial density distribution of the proton spins. The image reconstruction problem can be formulated in a continuous setting. In the common case where the discrete samples are given on a finite rectangular area of a Cartesian grid, the minimum norm solution can be efficiently calculated with the help of FFT algorithms.

# 3

## Parallel Imaging

### 3.1 Introduction

A drawback of MRI is the long acquisition time. Parallel imaging is a general technique to accelerate MRI by the simultaneous use of multiple receive coils. Roughly at the same time with the introduction of phased-array coils [91] parallel imaging was first conceived [57, 61, 88, 66, 65, 22]. Clinical applications appeared only much later with the introduction of the algorithms SMASH [99] and SENSE [86].

In parallel imaging, MRI acquisitions from multiple receive coils may be exploited for encoding part of the spatial information of an object by the spatially varying coil sensitivities. When used in conjunction with conventional phase-encoding by magnetic field gradients, coverage of the k-space for image reconstruction may become undersampled along a suitable phase-encoding dimension which in turn corresponds to a reduction of overall scan time.

### 3.2 Phased-Array Coils

Phased-array coils consist of many small coils which are electromagnetically decoupled as far as possible. They were originally introduced into MRI to combine the advantage of higher signal-to-noise ratio (SNR) obtained with small surface coils with the large FOV of volume coils [91]. The MRI signal obtained for multiple receive coils is given by

$$s_j(t) = \int d\mathbf{x} \rho(\mathbf{x}) c_j(\mathbf{x}) e^{-2\pi i \mathbf{k}(t) \mathbf{x}} + n(t) \quad j = 1, \dots, N \quad . \quad (3.1)$$

Here  $\rho$  denotes the proton density and  $c_j$  the complex-valued spatial sensitivity profiles of the individual receive coils.  $\mathbf{k}(t)$  is the chosen k-space trajectory. The signal  $s_j$  is further disturbed by noise  $n$ . In the fully sampled case the data from each individual channel can be reconstructed independently by Fourier transformation. Figure 3.1 shows the respective images  $\hat{s}_j$  for a water phantom.

To obtain a single reconstructed image, the data from all coils have to be combined. Assuming independent and identically distributed (i.i.d.) Gaussian white noise, the best unbiased estimate for the image is given by [91, 105, 18]

$$\rho_{est} = \frac{1}{\sum_j |c_j|^2} \sum_j c_j^* \hat{s}_j \quad (3.2)$$

with the complex conjugate  $c_j^*$  of the coil sensitivities and the image  $\hat{s}_j$  for each individual channel. Because this estimator requires the knowledge of the sensitivity profiles  $c_j$ , a root of sum of squares (RSS) reconstruction is often used instead. Here, a final image is reconstructed by calculating magnitude images with a point-wise root of the sum of squares

$$\rho_{RSS} = \sqrt{\sum_j |\hat{s}_j|^2} . \quad (3.3)$$

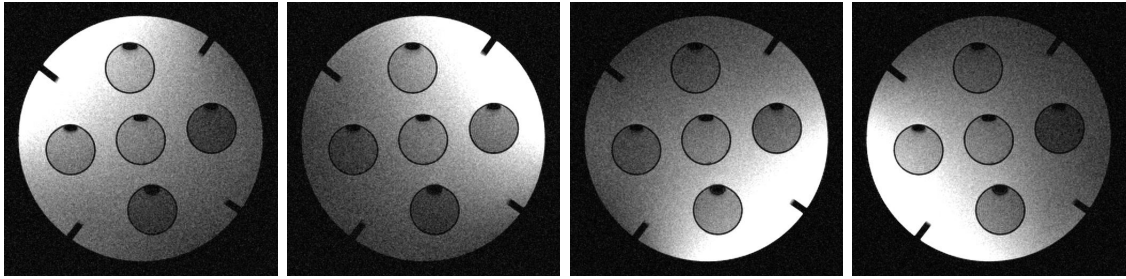
Apart from the fact that the image is modulated by the root of the sum of squares of the sensitivities of all receive coils, this can be considered an approximation of the optimal formula (3.2), where the sensitivity at a certain position is estimated from the corresponding signal  $c_j^* \approx \hat{s}_j^*$  itself. Because this approximation is valid only for locations with high signal, the RSS reconstruction introduces some bias and provides a lower SNR by a factor of  $\sqrt{N}$  in areas of low signal intensity [40, 24, 68].

### 3.2.1 Whitening

The assumption of white noise holds true only approximatively for real MRI coil arrays: The noise of different receive coils is correlated (often due to residual electromagnetic coupling) and of different variance. Given the corresponding statistical parameters, which can be estimated from a noise calibration scan, the reconstruction formulas and algorithms can be adapted for optimal results. Instead of modifying the algorithm, the data can be "whitened" in a pre-processing step which will be described in the following.

In a first step, the noise covariance matrix has to be calculated from noise  $n_i$  of each channel  $i$ , obtained for example during a calibration scan:

$$C_{ij} = \langle n_i - \bar{n}_i, n_j - \bar{n}_j \rangle$$



**Figure 3.1:** *Four individual images each calculated from the signal of its respective receive coil exhibit a different spatial sensitivity profile.*

In practice, the mean value  $\bar{n}$  can be assumed to be zero and its subtraction can be omitted. An eigen decomposition of the covariance matrix  $C$  consists of an unitary transformation  $U$  into a basis of eigenvectors and a diagonal matrix  $\Sigma$  of ordered eigenvalues:

$$C = U\Sigma U^H \quad \Sigma = (\sigma_{max}, \dots, \sigma_{min})$$

With this data, it is possible to transform the channels to synthetic channels

$$s'_i(t) = \sum_j (\Sigma^{-1/2} U^H)_{ij} s_j(t)$$

with uncorrelated noise of equal strength. In general, this transformation has to be taken into account in further processing step by modifying the signal equation accordingly. For MRI, this can be done simply by calibrating the sensitivity profiles from the transformed data.

When using coil elements which are accurately tuned to be decoupled and to have similar noise variance, this whitening step can be omitted. Because this technique can be applied to transform the noise statistics into noise with the identity as covariance matrix for all other cases, such statistics will be assumed in the rest of the thesis.

### 3.2.2 Array compression

Computational requirements increase linearly (in the case of the GRAPPA algorithm even quadratically) with the number of channels. For this reason, so-called array compression techniques have been developed [17]. They reduce the number of channels without significant loss of image quality. Similar to the whitening technique, this technique is based on an eigenvalue decomposition of a covariance matrix. Here, instead of the noise, the signal from different channels is decorrelated.

The covariance matrix is constructed from the data  $y_i$  as

$$C_{ij} = \langle y_i - \bar{y}_i, y_j - \bar{y}_j \rangle .$$

This matrix has size  $N \times N$  where  $N$  is the number of channels. Again, the covariance matrix has an eigen decomposition  $C = U\Sigma U^H$ . To decorrelate the acquired signal, the channels are then transformed according to

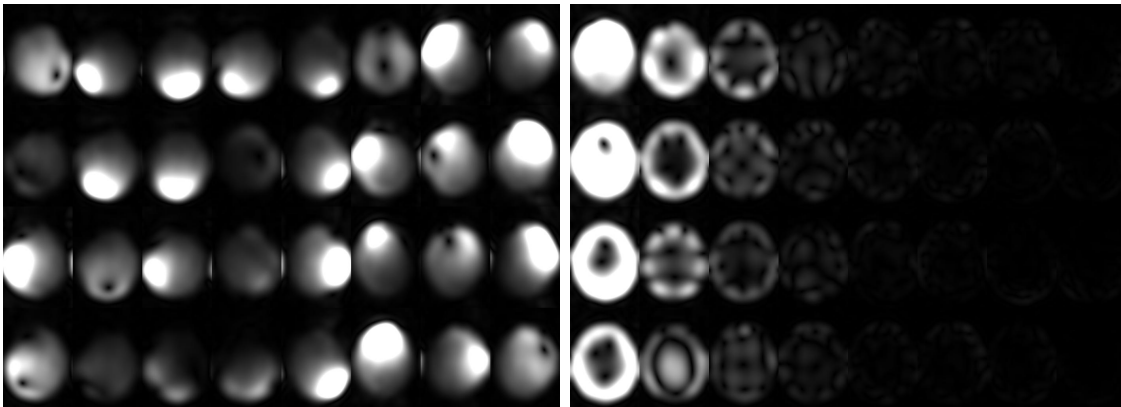
$$s'_i(t) = \sum_{j=1}^N U_{ij}^H s_j(t) .$$

It should be noted, that this transformation is an unitary transformation acting point-wise on k-space. Thus, when all transformed channels are reconstructed according to (3.2) with equally transformed sensitivities or with an RSS reconstruction (3.3), then the image reconstructed from the transformed channels is identical to the optimal image reconstructed from the original data.

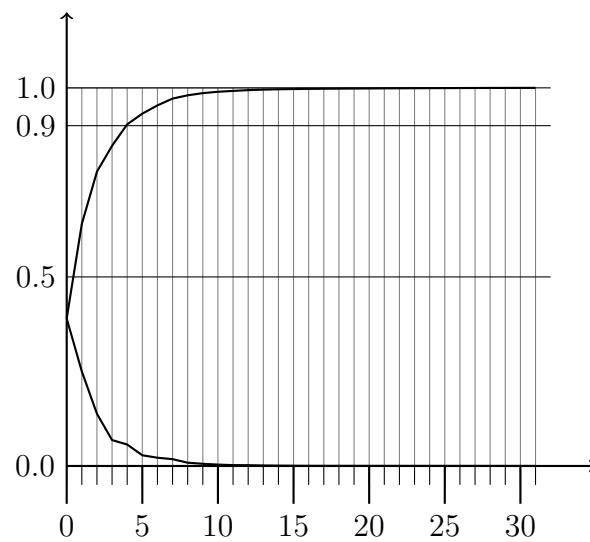
Sensitivities of the original and transformed channels are shown in Figure 3.2. Most energy is now concentrated in the first channels according to the value of their respective eigenvalue. In other words, the data is split into components, which are ordered according to their importance. Because of this property, the transformation is called principal components analysis (PCA). Computation time can be saved by using only the first most important channels and simply discarding the rest. To decide how many channels can be omitted, a possible strategy is to set an energy cutoff, which quantifies the fraction of total signal energy which must remain. Energy (corresponding to the eigenvalues) and cumulative energy for the synthetic channels are shown in Figure 3.3.

### 3.3 Undersampling of k-space

The main idea behind parallel imaging is, that the Fourier encoding can partially be replaced by the spatial information contained in the receive coil sensitivities. Hence, the MRI acquisition is undersampled by skipping some of the costly phase encoding steps, which directly translates into saved measurement time. If only every  $N$ -th line is measured, the measurement is accelerated by a factor of  $N$ , known as reduction or *acceleration factor*. In case of 2D imaging, the undersampling is employed in the direction of the phase encoding direction, while in 3D imaging both phase encoding directions can be used. The effect of regular undersampling on the individual coil images can be explained by the convolution theorem (8.1): Regular undersampling can be understood as multiplication with a Dirac comb function in k-space. In



**Figure 3.2:** *Left: Magnitude images of the sensitivities for all elements of the 32-channel head coil. Right: Principal components of all channels ordered from left to right and top to bottom.*



**Figure 3.3:** *Sorted eigenvalues of the covariance matrix used in array compression in descending order (lower graph) and cumulative energy content of the eigenvalues (higher graph).*

image space, this multiplication translates to a convolution with a comb function of reciprocal width (see appendix: Equation (8.5)), and an analogous result holds true in the discrete periodic setting. This aliasing effect is demonstrated in Figure 3.4.

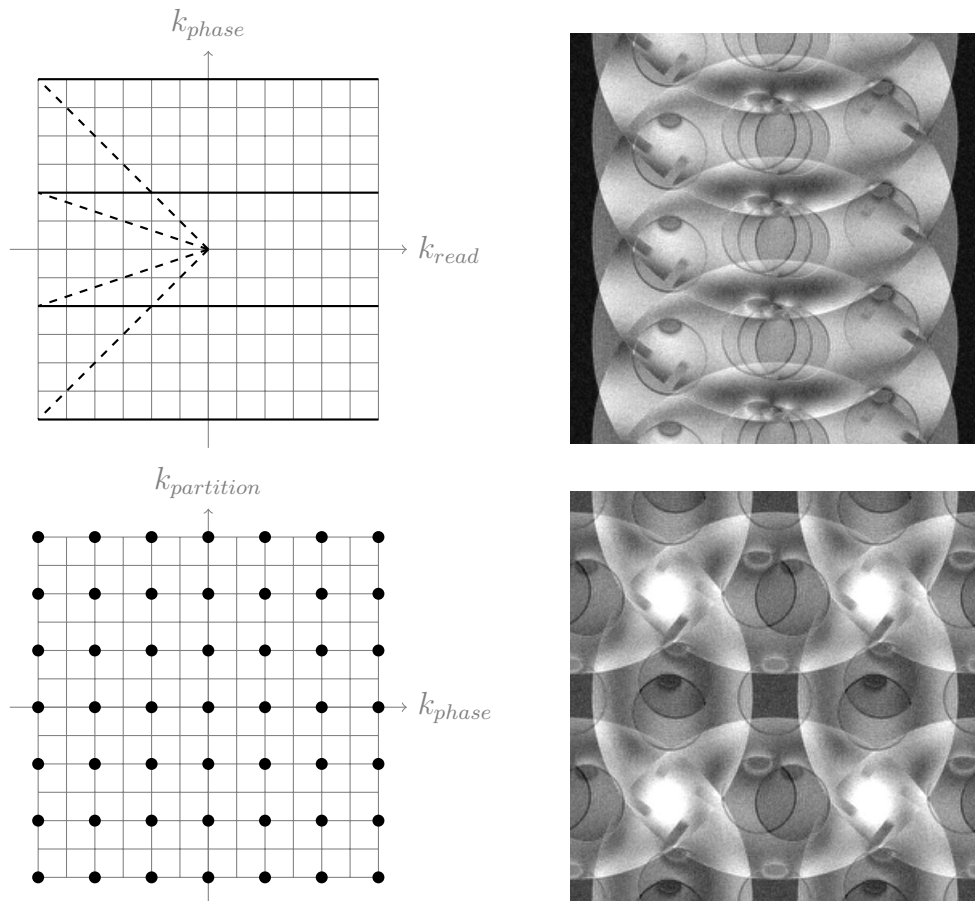
## 3.4 Image Reconstruction

To reconstruct an image from the undersampled multi-channel data, the signal equation (3.1) has to be solved for the unknown image  $\rho$ . If the coil sensitivity profiles are known, the signal equation represents a linear system, which can be discretized and solved numerically [53]. Existing direct methods to solve this system either utilize the decoupling of the equation in image space for regular sampling patterns like SENSE [88, 86, 85, 67], or approximate a sparse inverse in k-space such as simultaneous acquisition of spherical harmonics (SMASH) [99, 19] and its successors. Unfortunately, the parallel imaging reconstruction problem becomes increasingly ill-conditioned for large acceleration (or undersampling) factors. As a consequence, the inversion of the system leads to the amplification of noise that contributes to the data. To counter this effect the inversion has to be regularized [53]. Because the fundamental issues of parallel imaging can be understood best from the mathematical formulation as a linear inverse problem, it will shortly be introduced in the following, after discussing the discretization of the problem. Then the existing algorithms will be discussed.

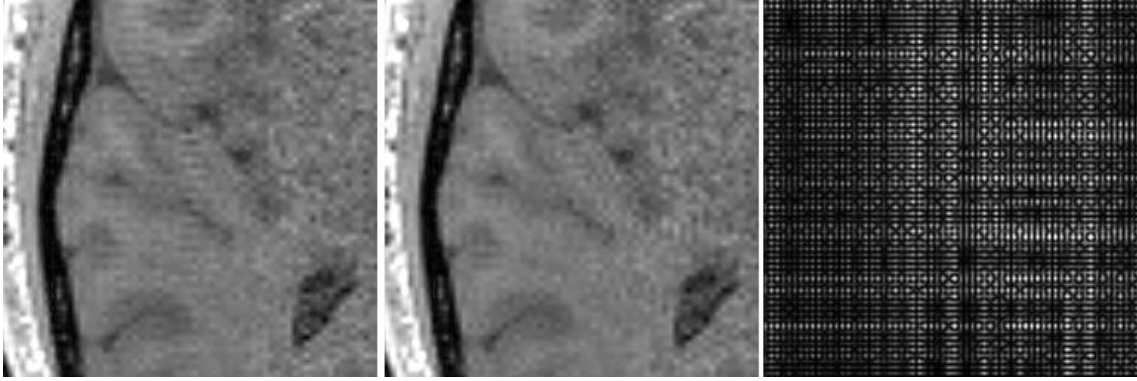
### 3.4.1 Discretization

Most of the time, the discretization scheme of Chapter 2.5.1 is implicitly assumed. Still, it is useful to reconsider this for the parallel imaging signal equation (3.1). Again, the function  $\rho$  as well as the coil sensitivities  $c_j$  can be assumed to have compact support. By periodic extension, this allows the k-space to be discretized on a grid (see Chapter 2.5). The situation is more complicated for the necessary cutoff in k-space. Here, a natural restriction of support is given for the data  $s$ , while the support of the image  $\rho$  as well as sensitivities  $c_j$  is a priori unbounded. Because the multiplication of  $\rho$  with the sensitivities  $c_j$  corresponds to a convolution of two functions of unbounded support in k-space (which has to be evaluated in a bounded region), at least one of these two functions has to be truncated in any numerical implementation. Given a cutoff for the sensitivities, the maximum frequency of the image, which can be shifted into the support of the data, can be computed. It is the sum of the maximum measured frequency and the highest frequency in





**Figure 3.4:** Undersampling in  $k$ -space (left) corresponds to aliasing in the image domain (right): (Top) In a 2D sequence, there is one phase encoding direction. In this example, this direction is undersampled by a factor of four, which leads to aliasing in the corresponding direction in the image domain. (Bottom) In 3D imaging, it is advantageous to split the acceleration factor to both phase encoding directions. The image on the right-hand side represents a section perpendicular to the read direction of the reconstructed 3D volume.



**Figure 3.5:** *Parallel image reconstruction using cyclic convolution (left) compared to normal convolution (middle) with the coil sensitivities. The periodic boundary conditions in k-space related to cyclic convolution lead to numerical errors. A difference image is shown on the right.*

the truncated k-space representation of the sensitivities. Because higher frequencies cannot possibly be determined from the data, they can be set to zero. After this implicit frequency cutoff for the image, the convolution can now numerically be computed by an FFT.

Most image-domain algorithms for parallel imaging ignore this issue and simply multiply sensitivities and object function and apply an FFT afterwards. This multiplication in the image-domain corresponds to cyclic convolution in k-space, which introduces some numerical noise at the k-space border. Although very small, the effect can sometimes lead to a visual degradation of image quality, as can be seen in Figure 3.5.

### 3.4.2 Parallel Imaging as Linear Inverse Problem

In the following, a linear inverse problem is considered, which is notated as a forward problem:

$$\mathbf{y} = A\mathbf{x} + \mathbf{n}$$

Here,  $\mathbf{x}$  is the unknown image,  $\mathbf{y}$  is the data,  $A$  the forward operator and  $\mathbf{n}$  the noise. The matrix  $A$  is composed of three components

$$A = P_{\mathbf{k}}\mathcal{F}C .$$

Here,  $C$  denotes the multiplication of the image with the coil sensitivities,  $\mathcal{F}$  is the Fourier transform, and  $P_{\mathbf{k}}$  the projection onto the trajectory. In the context of

parallel imaging, this problem is in general over-determined. This will be assumed in the following. A solution is therefore calculated in the least-squares sense:

$$\bar{x} = \operatorname{argmin}_x \|A\mathbf{x} - \mathbf{y}\|_2^2$$

Assuming that  $\ker(A) = \{0\}$ , a direct formula for this solution is given by the Moore-Penrose pseudo inverse [80, 83]

$$\bar{x} = A^\diamond \mathbf{y} \quad A^\diamond = (A^H A)^{-1} A^H \quad .$$

In absence of systematic errors, this solution is the sum of the true solution and a term corresponding of the reconstructed noise:

$$\bar{x} = \mathbf{x} + A^\diamond \mathbf{n}$$

Unfortunately, in the case of bad conditioning of the linear system, this noise term can become very large. The noise amplification can be reduced by introducing a small multiple of the identity matrix as a damping (or regularization) term into the inversion:

$$\bar{x}_\alpha = A_\alpha^\diamond \mathbf{y} \quad A_\alpha^\diamond = (A^H A + \alpha I)^{-1} A^H$$

Formulated as minimization problem, the regularized solution is then

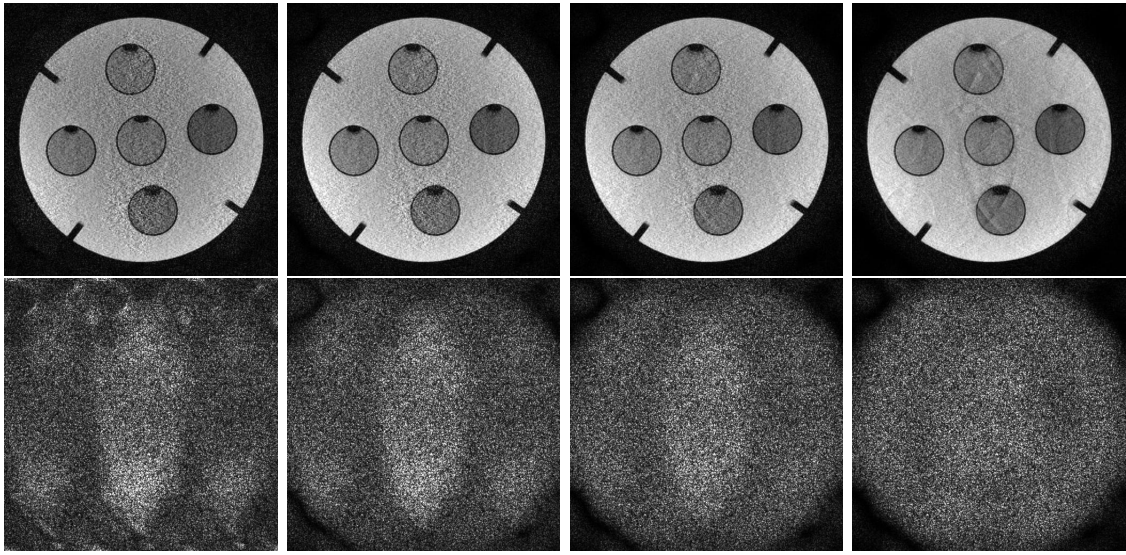
$$\bar{x}_\alpha = \operatorname{argmin}_x \|A\mathbf{x} - \mathbf{y}\|_2^2 + \alpha \|\mathbf{x}\|_2^2 \quad .$$

An example of the influence of this regularization parameter on parallel imaging reconstruction is given in Figure 3.6. A low value of  $\alpha$  leads to a noisy reconstruction, while a high value causes reconstruction artifacts.

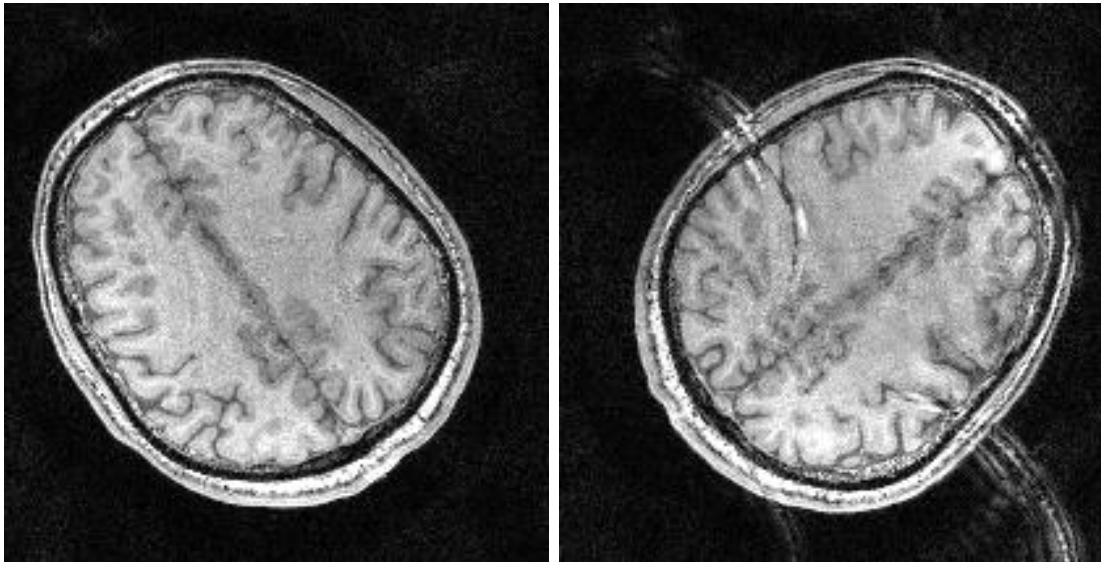
The identity matrix is the simplest choice for the regularization term, which corresponds to a penalty of the  $L^2$  norm of the image. A closer look at different regularization techniques is taken in Chapter 7.

## 3.5 Calibration of the Coil Sensitivities

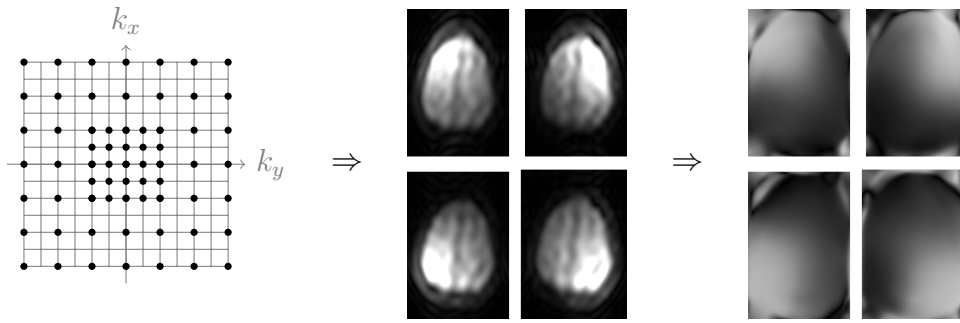
To take full advantage of parallel imaging techniques the information that necessarily needs to be derived from the sensitivities of the different receive coils has to be known with sufficiently high accuracy. Unfortunately, however, the receive sensitivities depend on the dielectric properties of the object under investigation and reflect even small object movements.



**Figure 3.6:** *Reconstruction artifacts: When choosing the regularization parameter, high noise amplification (left, weak regularization) has to be balanced against residual under-sampling artifacts (right, strong regularization).*



**Figure 3.7:** *Effects of coil sensitivity miscalibration: (Left) The sensitivities have been calibrated and allow for a reasonable reconstruction. (Right) Using the same coil sensitivities after moving the head to a new position leads to serious reconstruction artifacts.*



**Figure 3.8:** *Auto-calibration: From a fully sampled center low resolution images can be calculated. Division by an RSS image and post-processing yields approximate coil sensitivities.*

Coil sensitivities can be obtained with a pre-scan. Here, complete images  $\rho_i$  are acquired for each channel. When data from the body coil with a very homogeneous sensitivity profile  $c_{bc} \approx \text{const}$  is available, coil sensitivities can in principle be calculated up to a constant factor by division:

$$\frac{\hat{s}_j(\mathbf{x})}{\hat{s}_{bc}(\mathbf{x})} = \frac{c_j(\mathbf{x})\rho(\mathbf{x})}{c_{bc}(\mathbf{x})\rho(\mathbf{x})} \propto c_j(\mathbf{x})$$

In practice, a support mask  $\{\mathbf{x} \text{ with } |\rho(\mathbf{x})| > \epsilon\}$  has to be calculated, to exclude regions without signal. Also, the result has to be smoothed and extrapolated to a slightly larger region of support. When data from a body coil is not available, the sensitivities can be calculated relative to the RSS reconstruction (3.3).

While this approach yields good coil sensitivities for phantom studies, a major problem with this approach is the fact that movements of the subject can lead to inconsistencies between the calibrated coil sensitivities and the actual measurement. Especially when part of the subject moves into regions where no sensitivities could be determined during calibration, reconstructed images are affected by severe artifacts (see Fig. 3.7). But even without these problems a pre-scan is an additional time-consuming step during an examination, which must be repeated after each patient repositioning, and therefore is a major practical hurdle.

To avoid these problems, autocalibrating methods have been developed which determine the required information from a fully sampled block of reference lines in the center of k-space. Because the reference lines are usually acquired exactly at the same time as the actual object-defining lines in k-space, all aforementioned miscalibration problems are completely avoided. For methods where explicit coil sensitivities are required, the technique proceeds similar to the conventional calculation of the sensitivities from data acquired with a pre-scan (see Figure 3.8).

Methods which do not need explicit sensitivity maps will be discussed later. In all these techniques, the measurement time spent for the acquisition of the reference lines has to be balanced against truncation artifacts caused by the limited size of the fully sampled k-space center. A considerably improved method for autocalibrated parallel imaging will be presented in Chapter 5.

## 3.6 Algorithms

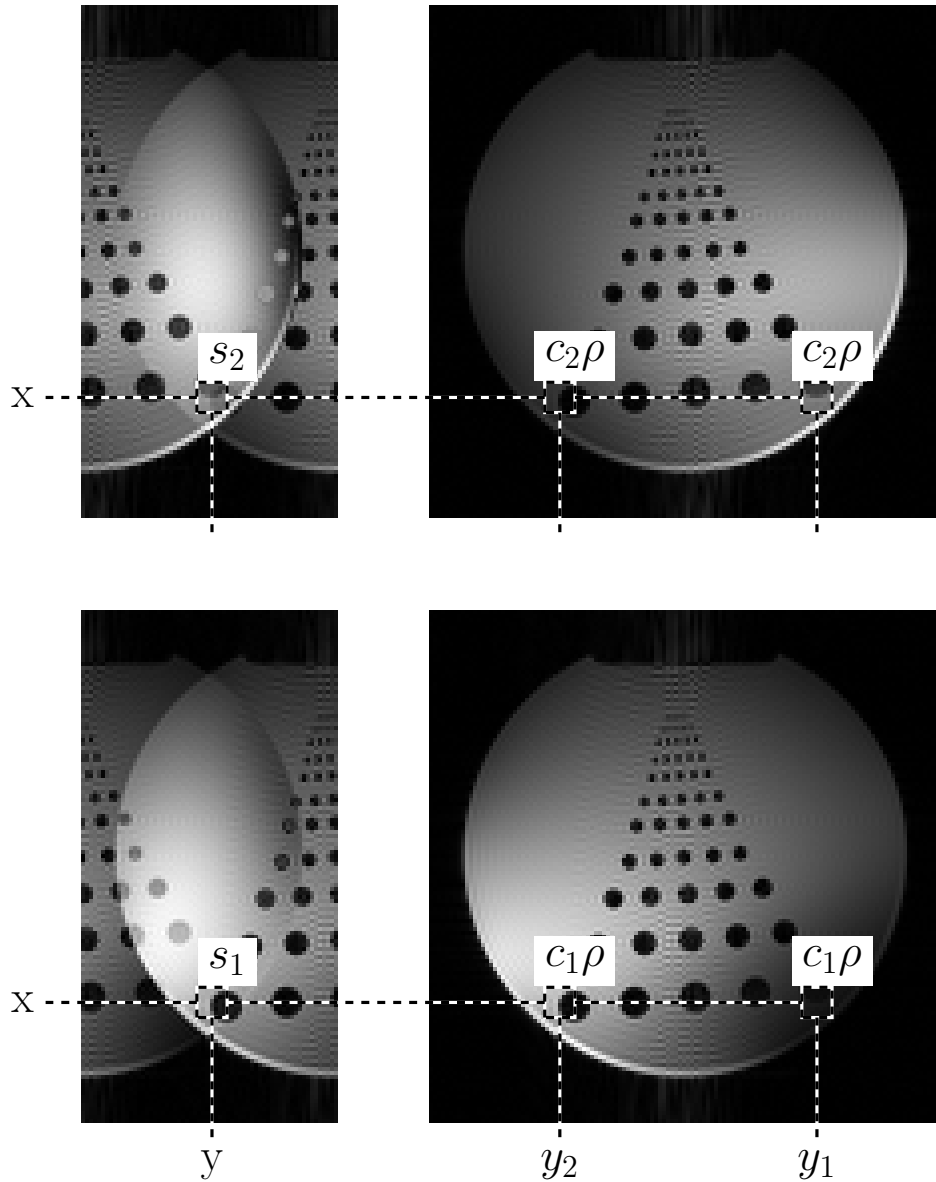
Most commercially available algorithms are currently based on direct matrix inversion methods, which use special techniques to calculate a sparse inverse for the linear system. There are many different variants, based on two lines of development: Methods formulated in the image domain were originally pioneered by Ra and Rim [88]. For practical computation, they exploit the decoupling of the equations in image space for regular sampling patterns. In [86], this algorithm was analyzed and extended to improve SNR in the case of correlated noise. Also, the term SENSE for this kind algorithms was introduced in this work. The other line of algorithms is based on a sparse approximation of the inverse in the Fourier domain. Starting with the SMASH algorithm [99] which gave the first in vivo demonstration of parallel imaging, this line was then developed in multiple steps to the GRAPPA algorithm, which is currently one of the most commonly used algorithms for parallel imaging. Beside these direct methods, iterative algorithms provide a generic alternative which overcome many limitations of the currently used direct techniques.

### 3.6.1 SENSE

The fast SENSE algorithm for regular Cartesian sampling, already conceived in [88], is based on the decoupling of the signal equation in image space. Reconstruction of the undersampled data for each channel  $j = 1, \dots, N_{coils}$  with an FFT leads to aliased images  $\hat{s}_j$ . For an acceleration factor of  $R$  in each point  $(x, y_1)$  of these aliased images exactly  $R$  equally spaced points  $(x, y_1), \dots, (x, y_R)$  in the given FOV are folded on top of each other: The linear system of equations decouples after Fourier transformation to a large number of small independent linear equations

$$\begin{pmatrix} \hat{s}_1(x, y_1) \\ \vdots \\ \hat{s}_{N_{coils}}(x, y_1) \end{pmatrix} = \begin{pmatrix} c_1(x, y_1) & \cdots & c_1(x, y_R) \\ \vdots & & \vdots \\ c_{N_{coils}}(x, y_1) & \cdots & c_{N_{coils}}(x, y_R) \end{pmatrix} \cdot \begin{pmatrix} \rho(x, y_1) \\ \vdots \\ \rho(x, y_R) \end{pmatrix}.$$

A graphical illustration is given in Figure 3.9 for a reduction factor of  $R = 2$ . For each of these equations, each corresponding to a set of aliased points, a (reg-



**Figure 3.9:** Decoupling of the linear system of equations for regular sampling patterns in image space. Only a number of points equal to the acceleration factor are aliased on top of each other. For each set of these points, the equations can be solved independently.

ularized) pseudo inverse can be calculated. While the calculation of a direct inverse of the complete system for  $N_{pixels}$  image pixels and  $N_{coils}$  channels would be prohibitively expensive in computation time and storage (with a matrix size of  $N_{pixels} \times (N_{pixels}/R) \times N_{coils}$ ), the calculation of the inverse for  $N_{pixels}/R$  equations of size  $R \times N_{coils}$  is cheap. This sparse inverse can be stored in  $N_{pixels} \times N_{coils}$  variables.

To obtain optimal SNR in the case of correlated noise, a whitening technique can be used, or the inversion can be adapted as described in [86]. At least for higher acceleration factors, regularization terms should be included. The extension of this algorithm to 3D imaging is straightforward [107].

The restriction to regular undersampling patterns is removed in an algorithm known as SPACE-RIP [67]. Here, the equations are decoupled only along the (fully sampled) frequency-encoded direction into  $N_{phase}$  equations by Fourier transformation of the data along this axis. Each individual equation is again solved by applying the pseudo inverse. The size of the individual equations increases from  $R \times N_{coils}$  to  $N_{phase} \times N_{coils}$ , which requires somewhat more computation time as compared to SENSE.

### 3.6.2 Conjugate Gradient Algorithm

The conjugate gradient algorithm can be used to iteratively solve linear inverse systems, which are too large to be solved efficiently with a direct matrix inversion [51]. In the context of parallel imaging, iterative algorithms, mostly based on the conjugate gradient algorithms, present a generic alternative to the established direct algorithms. Such algorithms are often referred to as conjugate gradient SENSE (CG-SENSE). While in the past iterative algorithms have been used only rarely due to their large computational demand, continuous progress in the development of computer hardware render them viable.

For a Hermitian and positive definite matrix  $A$ , the conjugate gradient algorithm calculates in each iteration an approximate solution to the equation  $A\mathbf{x} = \mathbf{y}$ , which minimizes the distance to the exact solution  $\|\mathbf{x}_n - \mathbf{x}_*\|_A$  in a so-called Krylov subspace. This distance is measured with a norm  $\|\mathbf{x}\|_A := \sqrt{\mathbf{x}^H A \mathbf{x}}$ . The Krylov subspace is increased by one dimension in each iteration step  $n$ . The subspaces are constructed by the repeated application of the symmetric system matrix to the initial data vector  $\mathbf{y}$ :

$$K_{A,\mathbf{y}}^n = \text{span} \{A^n \mathbf{y}, A^{n-1} \mathbf{y}, \dots, A \mathbf{y}, \mathbf{y}\}$$

This construction yields a good approximate solution already in a Krylov subspace



of small dimension, and, as a consequence, after only a small number of iterations. Typically, the algorithm is stopped, when the residuum  $\|A\mathbf{x} - \mathbf{y}\|_2$  becomes smaller than a given accuracy  $\epsilon$ . For more information about the conjugate gradient algorithm see [46].

Because the system is not symmetric for parallel imaging and additionally needs to be regularized, the algorithm is applied to the regularized normal equation

$$(A^H A + \alpha I)\mathbf{x} = A^H \mathbf{y} .$$

The algorithm then converges to the desired solution  $\bar{x}_\alpha = A_\alpha^\diamond \mathbf{y}$ .

### Extension to Non-Cartesian Trajectories

For non-Cartesian trajectories, Fourier transform and projection  $P_{\mathbf{k}}$  onto the measured data space cannot be implemented with a DFT anymore. For a direct reconstruction of non-Cartesian data, a technique called *gridding* is used to approximate an inverse of  $P_{\mathbf{k}}\mathcal{F}$  with interpolation techniques. In the context of iterative reconstruction techniques, only the forward operator, which maps from image to k-space, and its adjoint have to be implemented, which avoids some steps of the gridding technique. To evaluate the forward operator, the Fourier transformation is first calculated on a Cartesian grid with a DFT and then interpolated to the desired sampled points. According to the Whittaker-Shannon interpolation formula, the exact values at the sample positions can be obtained with a *sinc* interpolation. For practical reasons, this convolution has to be approximated by using some finite convolution kernel (typically a *Kaiser-Bessel-function*) and a roll-off correction in the image domain.

While this procedure is fast compared to a direct computation of the Fourier transform, it is still a major computational burden. In iterative reconstructions it is possible to completely avoid this interpolation step during the iteration [104, 30]. Instead, the data is interpolated only once at the beginning of the iteration. The conjugate gradient algorithm is applied to the normal equation:

$$\begin{aligned} A^H \mathbf{y} &= A^H A \mathbf{x} \\ &= C^H \mathcal{F}^{-1} P_{\mathbf{k}} P_{\mathbf{k}} \mathcal{F} C \mathbf{x} \\ &= C^H \underbrace{\mathcal{F}^{-1} P_{\mathbf{k}} \mathcal{F}}_{\text{convolution}} C \mathbf{x} \end{aligned}$$

The left-hand side of this equation, the part which is only evaluated once, is approximated by using an interpolation technique for the adjoint  $A^H$  as detailed above,

while the right-hand side, which needs to be evaluated in each iteration step, can be implemented exactly without any approximation and without the use of expensive operations. The right side contains a convolution  $[\star]$  with the point spread function (PSF) of the trajectory:

$$\mathcal{F}^{-1}P_{\mathbf{k}}\mathcal{F} = [\star]_{\text{PSF}}$$

Due to the compact support (FOV) of the image  $\mathbf{x}$  a corresponding projection  $P_{\text{FOV}}$  may be inserted into the equation:

$$P_{\text{FOV}}\mathcal{F}^{-1}P_{\mathbf{k}}\mathcal{F}P_{\text{FOV}} = P_{\text{FOV}}[\star]_{\text{PSF}}P_{\text{FOV}}$$

Hence, the convolution kernel (the PSF) may be truncated to region of twice the size and then implemented with FFTs and multiplication of the Fourier transform  $\hat{T}_{\mathbf{k},2\times\text{FOV}}$  of the truncated PSF:

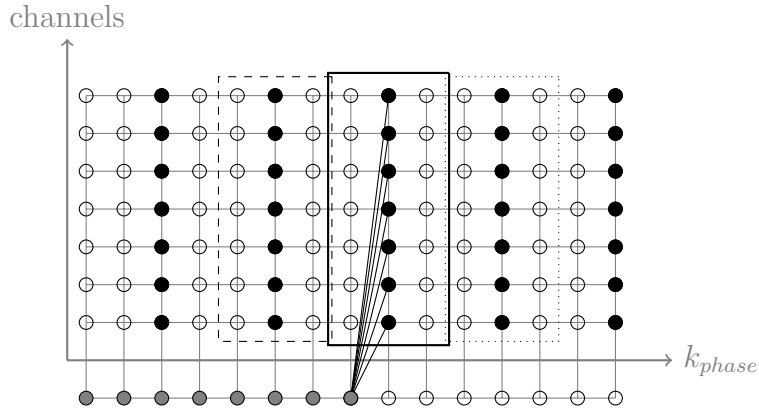
$$P_{\text{FOV}}[\star]_{\text{PSF}}P_{\text{FOV}} = P_{\text{FOV}}\mathcal{F}^{-1}\hat{T}_{\mathbf{k},2\times\text{FOV}}\mathcal{F}P_{\text{FOV}}$$

Here, the diagonal operator  $T_{\mathbf{k},2\times\text{FOV}}$  is a smeared version of the projection  $P_{\mathbf{k}}$ . Because this convolution is implemented by two FFTs and point-wise multiplication, it is not only much faster than any k-space interpolation technique, but also much simpler to implement than gridding techniques. This is especially helpful for an implementation on a graphical processing unit (GPU), which are difficult to program.

### 3.6.3 SMASH, AUTO-SMASH, GRAPPA

A second line in the evolution of methods for parallel imaging are algorithms which try to complete the missing data in k-space before reconstructing the final image with an FFT. These algorithms are all based on the *k-space locality principle* [112], which postulates that points which are nearby in k-space are strongly correlated, while this correlation decreases rapidly with the distance of the points. This is also assumed to be the case between k-space samples from different coils. The principle finds its justification in the limited FOV which allows for an exact *sinc* interpolation in k-space and the extreme smoothness of the coil sensitivities. Because the modulation of the spin density with the sensitivities corresponds to a convolution with the Fourier transform of the sensitivities in k-space, smooth sensitivities relate to convolution with a sharp function.

Because image reconstruction with known sensitivities is linear, the k-space locality principle expresses missing samples approximately as a linear combination of nearby k-space samples from all coils.



**Figure 3.10:** *SMASH algorithm: Each k-space sample of the image is assembled with a linear combination of the nearest k-space samples from each channel.*

The first algorithm of this type was the SMASH algorithm [98]. Here, harmonic functions of low order are approximated by a suitable linear combination of the coil sensitivity functions:

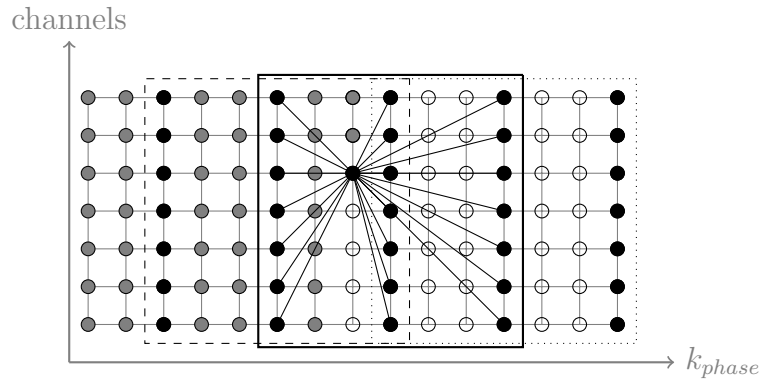
$$e^{i2\pi kx} \approx \sum_{l=1}^N w_{lk} c_l(x) \quad k = -\frac{R-1}{2} \Delta k, \dots, \frac{R-1}{2} \Delta k$$

K-space samples are then reconstructed by a linear combination from the nearest measured k-space samples from all channels:

$$\hat{\rho}(k) = \sum_{l=1}^N w_{lk} s_l(k)$$

This process is shown in Figure 3.10. Because the use of only the nearest k-space samples leads to a bad approximation which is only reasonable for very special coils in specific geometric settings, the algorithm was soon generalized to include larger k-space areas into the linear combination [19].

Further step-wise improvements have been the extension to autocalibration with variable density trajectories. Here, the weights are not determined by a fit of the coil sensitivities, but instead by a direct fit of some measured signal lines against one [58] or more [48] reference lines. Because inaccurate calibration might lead to cancellation artifacts, coil-by-coil reconstruction [77] was introduced. Here, the missing data points are recovered for each channel, and then combined with the actually measured data. The completed data set can then be reconstructed by an FFT for each channel, followed by an RSS reconstruction (3.3). Thus, no signal energy is lost in a linear combination of the channels. All these advancements were finally integrated into the GRAPPA algorithm [37] (see Figure 3.11), where the



**Figure 3.11:** *GRAPPA: In GRAPPA, all missing data from all coils are reconstructed from neighboring  $k$ -space data.*

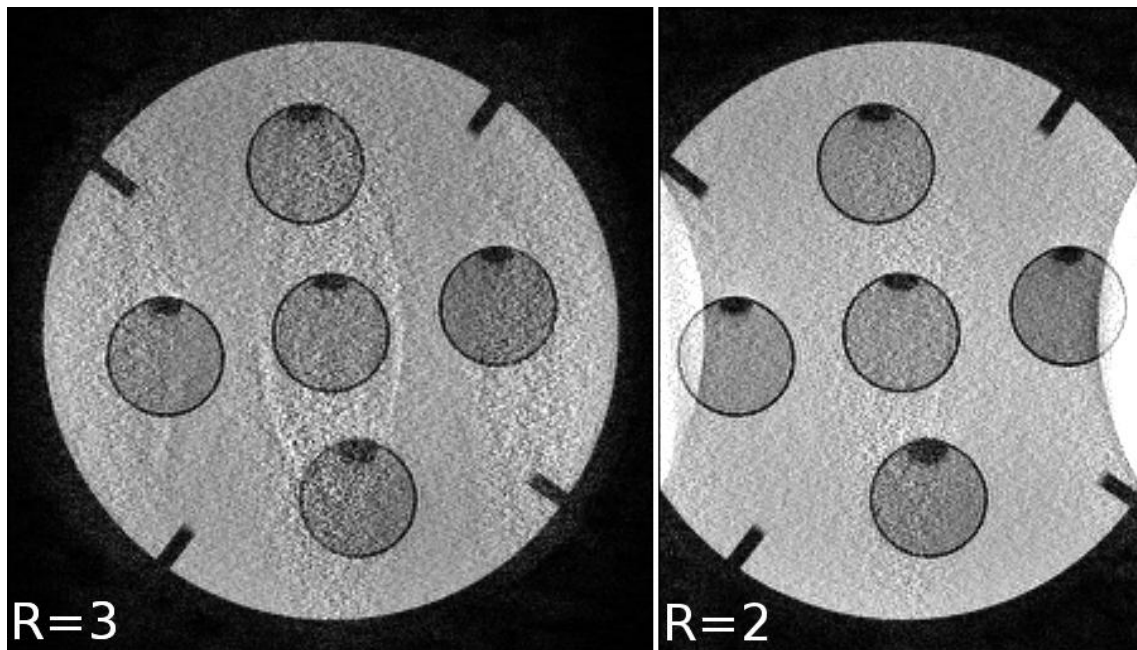
combination of coil-by-coil reconstruction with a variable density sampling scheme has the additional advantage, that the reference lines can be directly incorporated into the reconstruction, thus increasing the SNR. Further developments include the extension to 2D acceleration in 3D imaging [9] and to dynamic imaging [56]. Various extensions and similar algorithms based on the  $k$ -space locality principle have been proposed for non-Cartesian imaging [112, 49, 2, 50, 73, 94].

### Field of View Limitations

In MRI without parallel imaging, it is a common practice to reduce the FOV as much as possible to save phase-encoding steps. Making the FOV slightly smaller than the actual object causes aliasing artifacts which appear on the border of the object (see Chapter 2.5.1). Whenever one is only interested in the center part of the image, this aliasing artifacts can be tolerated. For this reason, this technique is often used to reduce the measurement time in MRI.

When used in combination with autocalibrating parallel imaging, this has the effect that the normally fully sampled block of reference lines becomes undersampled. Many autocalibration techniques which are based on the direct estimation of the sensitivities from the  $k$ -space center rely on sensitivities with aliasing artifacts and reconstruct images with aliasing in the center of the image. This typically renders them useless. An interesting property of the GRAPPA algorithm is its ability to reconstruct images which have residual aliasing only at the border of the image, exactly as the images which can be obtained without parallel imaging.

The reason why this is possible is not yet fully understood. In [39] it is described as an experimental fact, but no real explanation is given. In [8] it is stated that



**Figure 3.12:** *GRAPPA reconstruction: (left) full FOV with an acceleration factor  $R = 3$  (right) reduced FOV with an acceleration factor  $R = 2$ . Both measurements have a similar acquisition times. Although the acquisition with a reduced FOV has aliasing artifacts at the border of the image, the quality in the center of the image is better than in the image acquired with full FOV but higher acceleration factor.*

the GRAPPA algorithm calculates a low-resolution convolution kernel to unfold the image. The proposed explanation is that due to the redundancy of multiple receive coils, there are low-resolution reconstruction kernels even when the coil sensitivities are not smooth.

### 3.7 Summary

Reconstruction algorithms for parallel imaging are based on the inversion of a linear system of equations. In the case of autocalibrated parallel imaging, the reconstruction process is split into two separate linear steps: A linear calibration step and a linear reconstruction step. For autocalibrated SENSE, the linear calibration step consists in the determination of the coil sensitivities from the fully sampled k-space center, followed by the inversion of a decoupled system of equations in image space. For GRAPPA, the calibration step consists in the determination of a sparse approximation to the inverse system matrix. In the reconstruction step this approximate inverse is applied to the data to calculate the solution. In their original formulation, both techniques rely on regular sampling patterns, although various studies generalized the algorithms to irregular and non-Cartesian sampling schemes. A generic alternative to these direct algorithms is the use of iterative techniques as these allow for arbitrary sampling patterns, although at the expense of a prolonged computation time. For a large number of channels, the computation time can be reduced for all algorithms with the use of the array compression technique.

# 4

## MRI System

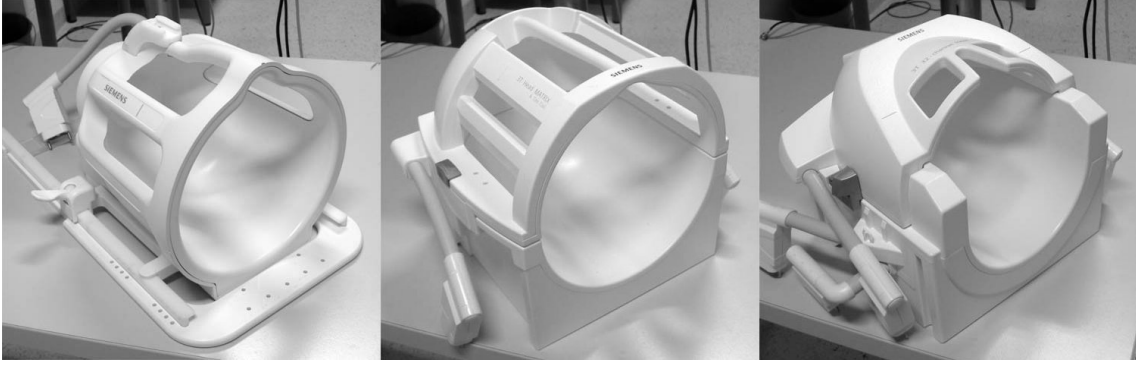
In this chapter the MRI system which has been used for all experiments is described. The discussion covers properties of the superconducting magnet, gradient system, RF coils, computer system, and software framework.

### 4.1 Magnet and Gradient System

The main components of a MRI system are a strong magnet, the gradient system, and RF coils. The magnet creates the static  $B^0$  field, while the gradient system creates additional gradient fields, which can be changed during the experiments. RF coils are used to transmit the excitation pulses, as well as to receive the measurement signal. The MRI system used in this work is a Siemens TIM Trio whole body human



**Figure 4.1:** Whole-body MRI system: (left) with and (right) without housing.



**Figure 4.2:** 8-channel, 12-channel, and 32-channel head coil arrays.

scanner (Siemens AG, Erlangen, Germany), which contains a liquid helium cooled super-conducting magnet. The bore has a diameter of 60 cm offering a FOV of 50 cm in each direction. The main magnetic field is  $B_0 = 2.89$  T. At this field strength, the Larmor frequency is  $\omega/2\pi = 123.2$  MHz. The gradient system has a maximum gradient strength of  $38 \text{ mT m}^{-1}$  per axis and a maximum slew rate of  $170 \text{ mT m}^{-1} \text{ ms}^{-1}$ . The system provides 32 receive channels, which are digitized at 10 MHz with 24 bit analog-to-digital converters. Quadrature demodulation, low pass filtering, and downsampling yield the real and imaginary part of the MRI signal.

## 4.2 Radio Frequency Coils

Built into the structure of the magnet is the so-called body coil, which is used as a transmit coil for the excitation of the spins. It can also be used as a receive coil, but provides only one single output channel and has a relatively low signal due to its distance from the patient. For this reason, various receive coils are available, which are often specialized for different body parts. In this work, 8, 12, and 32-channel head coils have been used, with most experiments conducted with the 12-channel head coil. This coil has additional electronics, which transform the received signals in hardware. The signals from four sets of three coil elements ( $L, M, R$ ) are independently combined in hardware, according to the following unitary transformation:

$$\begin{pmatrix} P \\ S \\ T \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{-i}{\sqrt{2}} & \frac{-1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & \frac{-1}{2} \end{pmatrix} \begin{pmatrix} L \\ M \\ R \end{pmatrix}$$

This transform is adapted to the properties of the coil so that most signal energy is concentrated into the four primary modes ( $P$ ). The receive hardware has two



operating modes: circularly polarized (CP)-mode and triple mode. In triple mode the signals from all 12 channels are available for image reconstruction, while in CP-mode, only the four primary modes ( $P$ ) are used. This mode can be used to save computation time during image reconstruction, and is a very simple and hard-coded alternative to the array compression technique described in Chapter 3.2.2.

### 4.3 Computer System and Software

The MRI system is equipped with three computers, the host, where the user interface is running, a real-time computer system, which is used to control the image acquisition processes, and an image reconstruction computer. The latter receives the digitized data and reconstructs the images, which are then send back to the host system to be presented to the operator. The image reconstruction system also stores the raw data of all acquired data for some limited time. From here, raw data can be acquired and transferred to a different computer for off-line processing.

For off-line image reconstruction two PowerEdge 2900 computer systems (Dell Inc., Round Rock, USA) running the Ubuntu Linux operating system were available. The smaller system is equipped with two dual core Intel (Intel Corporation, Santa Clara, USA) Xeon u5060 CPUs at 3.20 GHz and with 4 GB RAM. The bigger system has two quad core Intel Xeon (E5345) CPUs at 2.33 GHz and 8 GB RAM. The systems are connected with a dedicated 1 Gigabit ethernet connection to the scanner network as well as to each other.

Image reconstruction was done with programs written in the  $C$  programming language (ISO/IEC 9899:1999) and making use of the POSIX application programmer interface (API) (ISO/IEC 9945) provided by the Ubuntu Linux operating system. The programs were compiled with the GNU compiler collection (GCC)<sup>1</sup>. The FFTW<sup>2</sup> library provided a fast implementation of FFT algorithms. The program collection includes a tool to extract the raw measurement data from the file obtained from the scanner, pre-processing tools for whitening and array compression, and implementations of various reconstruction algorithms. An image viewer (written with the GTK library<sup>3</sup>) was used to visualize the reconstructed images. The algorithm presented in the next chapter is also available as Matlab (or octave) code and can be downloaded from the internet.<sup>4</sup>

---

<sup>1</sup><http://www.gcc.org>

<sup>2</sup><http://www.fftw.org>

<sup>3</sup><http://www.gtk.org>

<sup>4</sup><http://www.biomednmr.mpg.de>

# 5

## Joint Estimation of Image Content and Coil Sensitivities

### 5.1 Introduction

The use of parallel imaging for scan time reduction in MRI faces problems with image degradation when using GRAPPA or SENSE for high acceleration factors. While an inherent loss of SNR in parallel MRI is inevitable due to the reduced measurement time, the sensitivity to image artifacts that result from severe undersampling can be ameliorated by alternative reconstruction methods. Here, an algorithm based on a Newton-type method with appropriate regularization terms is demonstrated to improve the performance of autocalibrating parallel MRI – mainly due to a better estimation of the coil sensitivity profiles. The approach yields images with considerably reduced artifacts for high acceleration factors and/or a low number of reference lines forming the fully sampled k-space center.

The common reconstruction methods for autocalibrated parallel imaging are based on a sequential approach: the determination of the information about the coil sensitivities from the reference lines is followed by the reconstruction of an image by a *linear* process. As will be explained later such two-step techniques make only suboptimal use of the available data. With the help of an alternating minimization method, Ying and Sheng [113] recently proposed to improve this situation by iteratively optimizing both the coil sensitivities and the image content until a joint solution is found. Extending these ideas, the purpose of this work is to show how a regularized *nonlinear* inversion technique based on a Newton-type method with appropriate regularization terms provides a generic and convenient framework

for solving this problem in the context of MRI reconstruction.

## 5.2 Parallel Imaging as Nonlinear Inverse Problem

In general, a determination of coil sensitivities from only the center of k-space does not take advantage of all available information. Although the information about a smooth coil profile is mostly localized in the k-space center, the measured data represents the convolution of the coil profiles with the object function which shifts information from the center of k-space to its outer parts. Because even small errors in the sensitivity profiles lead to residual aliasing artifacts in the resulting image, an optimal reconstruction should exploit all available k-space data rather than only a small part in its center. This can be accomplished by a nonlinear inversion technique.

In order to solve for the object function and the sensitivity profiles at the same time, the method proposed here applies a regularized nonlinear inversion to the reconstruction problem. The basic MRI signal equation is understood as a nonlinear operator equation with an operator  $F$  which maps the proton density, the image content, and the coil sensitivity profiles to the measured data

$$F(x) = y \quad \text{with} \quad x = \begin{pmatrix} \rho \\ c_1 \\ \vdots \\ c_N \end{pmatrix} .$$

This equation is solved by a Newton-type algorithm, more precisely the iteratively regularized Gauss-Newton method (IRGNM), see [31, 5, 59] as general references. In the following it is assumed that the operator equation is given in a discretized form where all functions are represented by vectors of point values on a rectangular grid.

## 5.3 Algorithm

### 5.3.1 Iteratively Regularized Gauss Newton Method

The general idea of the Newton algorithm is as follows. A linearization of the equation around an initial guess  $x_n$  yields

$$F(x_n + dx) \approx DF(x_n)dx + F(x_n) .$$

Here  $DF(x_n)$  denotes the Fréchet derivative (or Jacobian) of  $F$  at the point  $x_n$ . This linearized equation is then solved for the update  $dx$

$$DF(x_n)dx + F(x_n) = y \quad .$$

Under certain conditions iterative updates of  $x_{n+1} = x_n + dx$  converge to a solution.

To calculate an approximate solution to the linearized problem the conjugate gradient algorithm can be used. This algorithm needs to be applied to a symmetric matrix, a condition which can be ensured by multiplying both sides with the adjoint of  $DF(x_n)$

$$DF(x_n)^H DF(x_n)dx = DF(x_n)^H (y - F(x_n)) \quad .$$

Due to the bad conditioning of the linearized equations the inversion must be regularized. Adding a positive definite regularization matrix (e.g., the identity) to the matrix  $DF(x_n)^H DF(x_n)$  yields the well-known Levenberg-Marquardt algorithm [70, 76]

$$(DF(x_n)^H DF(x_n) + \alpha_n I) dx = DF(x_n)^H (y - F(x_n)) \quad .$$

The regularization parameter  $\alpha_n$  is reduced in each step. For large values of  $\alpha_n$  the algorithm is related to the gradient descent algorithm

$$(\alpha_n I) dx = DF(x_n)^H (y - F(x_n)) \quad .$$

For low values the algorithm represents the classic Gauss-Newton method

$$DF(x_n)^H DF(x_n)dx = DF(x_n)^H (y - F(x_n)) \quad .$$

The Levenberg-Marquardt algorithm takes advantage of both these ideas by using the more robust gradient descent algorithm at the beginning of the iterative process (far from the solution) and the faster Gauss-Newton algorithm at the end (near the solution). Its properties can be understood in a reformulation as a sequence of quadratic optimization problems where the update  $dx$  corresponds to the unique minimizer of the functional

$$\|DF(x_n)dx - (y - F(x_n))\|^2 + \alpha_n \|dx\|^2 \quad .$$

Although each individual update is regularized to suppress noise, it is still possible that small residual noise in each update accumulates in the final solution. A more stable algorithm is given by changing the update rule to minimize

$$\|DF(x_n)dx - (y - F(x_n))\|^2 + \alpha_n \|x_n + dx - x_0\|^2 \quad .$$

In this case the regularization no longer applies to the update itself but to the result of the update with respect to the initial guess. A suitably revised algorithm, the IRGNM, then relates to an update rule

$$dx = (DF(x_n)^H DF(x_n) + \alpha_n I)^{-1} (DF(x_n)^H (y - F(x_n)) + \alpha_n (x_n - x_0)) \quad .$$

The regularization parameters are always chosen to be of the form  $\alpha_n = \alpha_0 q^n$  with the same  $q \in (0, 1)$ , usually  $q = \frac{2}{3}$ .

To apply this IRGNM algorithm to parallel MRI reconstructions, implementations of the operator, its derivative, and the adjoint of the derivative are needed. The operator is given by

$$F : x \mapsto \begin{pmatrix} P_{\mathbf{k}} \mathcal{F}\{c_1 \cdot \rho\} \\ \vdots \\ P_{\mathbf{k}} \mathcal{F}\{c_N \cdot \rho\} \end{pmatrix} \quad \text{with} \quad x = \begin{pmatrix} \rho \\ c_1 \\ \vdots \\ c_N \end{pmatrix},$$

where  $\mathcal{F}$  is the (multidimensional) Fourier transform and  $P$  is the orthogonal projection onto the trajectory which – for Cartesian sampling – is a diagonal matrix with ones at the positions of the measured data and zeros elsewhere. The derivative of the operator can easily be calculated by using the linearity of the Fourier transform and the product rule of derivatives

$$DF(x) \begin{pmatrix} d\rho \\ dc_1 \\ \vdots \\ dc_N \end{pmatrix} = \begin{pmatrix} P_{\mathbf{k}} \mathcal{F}\{\rho \cdot dc_1 + d\rho \cdot c_1\} \\ \vdots \\ P_{\mathbf{k}} \mathcal{F}\{\rho \cdot dc_N + d\rho \cdot c_N\} \end{pmatrix} \quad .$$

Due to the unitarity of  $\mathcal{F}$  the adjoint is then given by

$$DF^H(x) \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^N c_j^* \cdot \mathcal{F}^{-1}\{P_{\mathbf{k}} y_j\} \\ \rho^* \cdot \mathcal{F}^{-1}\{P_{\mathbf{k}} y_1\} \\ \vdots \\ \rho^* \cdot \mathcal{F}^{-1}\{P_{\mathbf{k}} y_N\} \end{pmatrix}$$

with the star denoting pointwise complex conjugation.

Once the dependency on the coil sensitivities is removed, that is the coil sensitivities are treated as constants, the resulting operator becomes linear while the parts corresponding to the update of the coil sensitivities in the derivative and its adjoint vanish. As a consequence, the algorithm effectively transforms into a conjugate

gradient (CG) variant of SENSE, as commonly used for non-Cartesian MRI [85]. When started with suitable coil sensitivities this modified algorithm produces solutions similar to those obtainable from a CG-SENSE algorithm with decreasing regularization.

## Convergence

For Newton-type algorithms convergence is not automatically guaranteed and usually convergence results assume an initial guess close to the solution. In practice, this seems to be no problem for the proposed algorithm. For example, in situations where GRAPPA produces a reasonable image, the nonlinear inversion algorithm decreases the residual in each iteration yielding a good solution after only 10 – 25 iterations – even for very low numbers of reference lines. In all these cases, the initial guess was always chosen as described below. In extreme situations, that is without any reference lines or for acceleration factors greater than the number of receive channels, the algorithm is still able to decrease the residual but without generating reasonable images. Very slow convergence in the sense that the norm of the residual decreases very slowly is only observed when the image object is completely empty and the algorithm tries to fit only noise. This case could be dealt with by calculating the amount of energy in the data before starting the iteration.

The theoretical treatment of convergence is complicated by the fact, that parallel imaging is increasingly worse conditioned for higher acceleration factors. This holds true even in the linear case and for perfectly known coil sensitivities. The exact solution of the system is very far from the desired solution, because of the amplified noise. Every parallel imaging algorithm should therefore contain some form of regularization. For the proposed algorithm this is accomplished by terminating the iteration long before convergence to an undesirable solution is reached. In fact, it would be advantageous to replace the assessment of convergence by a different mathematical criterion. It might quantify the ability of the algorithm in conjunction with its automatic stopping criterion (e.g., the discrepancy principle) to produce a series of solutions for input vectors with decreasing noise that converges to the exact solution of the noiseless case. While various formal theorems of this kind have been proved for this algorithm [4, 54, 59], it was not yet attempted to verify their assumptions in this specific case.

### 5.3.2 Regularization of the Coil Sensitivities

A direct application of the described IRGNM algorithm to the problem of auto-calibrated parallel imaging would yield an unrealistic solution to the MRI signal equation. Basically, the coil sensitivities would include a substantial portion of the object information and the object part would not be unfolded correctly. The reason is that the equation is highly underdetermined – even in the fully sampled case. This can be seen by the fact that multiplying the object part of a solution  $(\rho, c_j)$  by some arbitrary complex function  $g$  and dividing the coil sensitivities by the same function gives another solution  $(\rho \cdot g, c_j/g)$ .

The problem may be overcome by adding a priori knowledge about the object and the coil sensitivities. For example, while the object may contain edges, the coil sensitivities are generally rather smooth. This can be ensured by a smoothness enforcing norm for the coil profiles that is used during the quadratic optimization problems solved in each Newton step. An appropriate choice is the norm given by

$$\|f\|_l = \|(I - s\Delta)^{l/2}f\| \quad (5.1)$$

for some chosen index  $l$  and scaling parameter  $s$ . The 2D Laplacian is given by  $\Delta = \partial_x^2 + \partial_y^2$ . This norm can easily be calculated by weighting the standard  $L^2$  norm in Fourier space by an additional term  $(1 + s\|\mathbf{k}\|^2)^{l/2}$ . It penalizes high frequencies where the penalty is a polynomial with degree  $l$  in the distance to the k-space center. To achieve the desired regularization, the operator and the representation of the coil profiles are transformed with a preconditioning matrix  $W$  which contains a Fourier transform of the coil part multiplied by a diagonal weighting matrix

$$\begin{pmatrix} \rho \\ \hat{c}_1 \\ \vdots \\ \hat{c}_N \end{pmatrix} = \begin{pmatrix} I & & & \\ & (1 + s\|\mathbf{k}\|^2)^{l/2}\mathcal{F} & & \\ & & \ddots & \\ & & & (1 + s\|\mathbf{k}\|^2)^{l/2}\mathcal{F} \end{pmatrix} \begin{pmatrix} \rho \\ c_1 \\ \vdots \\ c_N \end{pmatrix} .$$

The IRGNM algorithm is then applied to the following transformed but equivalent system of equations

$$\begin{aligned} \hat{x} &= W^{-1}x \\ G\hat{x} &= FW\hat{x} = y \quad . \end{aligned}$$

Using these definitions the new minimization problem – to be solved in each Newton step – is given by

$$\|DG(\hat{x}_n)d\hat{x} - (y - G(\hat{x}_n))\|^2 + \alpha_n\|\hat{x}_n + d\hat{x} - \hat{x}_0\|^2 \quad .$$

A closer look at the regularization term reveals the effect of the above modification (assuming  $x_0 = \mathbf{0}$  for simplicity)

$$\begin{aligned}
\|\hat{x} - \hat{x}_0\|^2 &= \|W^{-1}x - W^{-1}x_0\|^2 \\
&= \|\rho\|^2 + \sum_{j=1}^N \|\hat{c}_j\|^2 \quad \text{using } x_0 = \mathbf{0} \\
&= \|\rho\|^2 + \sum_{j=1}^N \|(1 + s\|\mathbf{k}\|^2)^{l/2} \mathcal{F}c_j\|^2 \\
&= \|\rho\|^2 + \sum_{j=1}^N \|c_j\|_l^2 .
\end{aligned}$$

The  $L^2$  norm used in the regularization of the coil part of the transformed system therefore corresponds to the norm (5.1) in the original space, penalizing high frequencies and enforcing the smoothness of the coil sensitivities.

### 5.3.3 Choice of Parameters

The parameters for the proposed nonlinear inversion algorithm were chosen as follows: As initial guess  $x_0$  the object function  $\rho$  was set to constant 1 and the coil profiles  $c_1, \dots, c_N$  to zero. The data vector was normalized in the  $L^2$ -norm to 100. The parameters in the regularization term of the coil sensitivities were set to  $l = 16$  and  $s = 220 \text{ mm}^2$ . The first regularization parameter was chosen as  $\alpha_0 = 1$  and reduced by a factor  $q = 2/3$  in each Newton step. To our experience the final results are not sensitive to the choice of any of these parameters. The only critical value is the regularization parameter  $\alpha_{n_*}$  at the stopping index  $n_*$ , and hence the stopping criterion for the Newton method. The choice of  $n_*$  is a trade-off between small noise (small  $n_*$ ) and small undersampling artifact (large  $n_*$ ).

Although the choice of the stopping index has been studied intensively in the literature on inverse problems, and a number of methods such as Morozov's discrepancy principle [31] or Lepskii's balancing principle [6] are available, no satisfactory solution for the parallel MRI problem has been described to the best of our knowledge. Part of the reason is that for images the  $L^2$  error does not accurately describe the visual quality of a reconstruction. For the data presented here, the stopping index was therefore chosen by visual inspection. Although it may be fixed for repeated measurements with the same MRI protocol, the automatic determination of a robust stopping criterion is certainly a most desirable feature, but outside the scope of this work. Further comments on the choice of the regularization parameter can be found in Chapter 7.2.



It should be noted that the choice of a regularization parameter is not unique to the present algorithm but a general problem common to all parallel imaging methods. Even some of the GRAPPA or SENSE images shown in this work could probably be improved by manually optimizing the regularization parameters (chosen by the vendor). Nevertheless, a simple adjustment of the GRAPPA and SENSE regularization would not be enough to generate images equivalent to those obtained by nonlinear inversion. For example, pronounced qualitative differences are demonstrated for reconstructions with a low number of reference lines. As demonstrated later in Figure 5.6 respective GRAPPA reconstructions exhibited both high noise and residual ghosting artifacts, whereas nonlinear inversion resulted in very much improved images. Because the overall regularization determines the trade-off between noise and artifacts, the advantage of the proposed algorithm relative to GRAPPA cannot be explained by a better choice of regularization parameters.

### 5.3.4 Postprocessing

As mentioned before, the system of equations is underdetermined even in the fully sampled case. Although this ambiguity may be removed by the choice of the regularization terms, the result is not necessarily identical to a typical RSS reconstruction. This difference manifests itself as slight changes in the large-scale intensity distribution. By multiplying the resulting image with the RSS of the estimated coil sensitivities and dividing the coil sensitivities by the same quantity, this difference can be removed in a simple step

$$\rho_{final} = \rho \cdot \sqrt{\sum_{j=1}^N |c_j|^2} \ .$$

While not strictly necessary, this postprocessing is useful when comparing images or coil sensitivities to images reconstructed with different regularization parameters, other parallel imaging algorithms, or from fully sampled data sets.

### 5.3.5 Computational Speed

In the present implementation the reconstruction of a single  $256 \times 256$  image by nonlinear inversion takes less than a minute on a modern processor. While this is too slow for applications to a complete 3D MRI data set, significant acceleration can already be expected by simple measures such as a multithreaded implementation utilizing multiple processor cores and switching from double to single precision floating

point numbers. Further possibilities include preconditioning techniques (e.g., the method presented in [55]), which would reduce the number of conjugate gradient iterations, and adaptive discretization in particular of the coil profiles, which would reduce the cost of the first Newton steps by an order of magnitude.

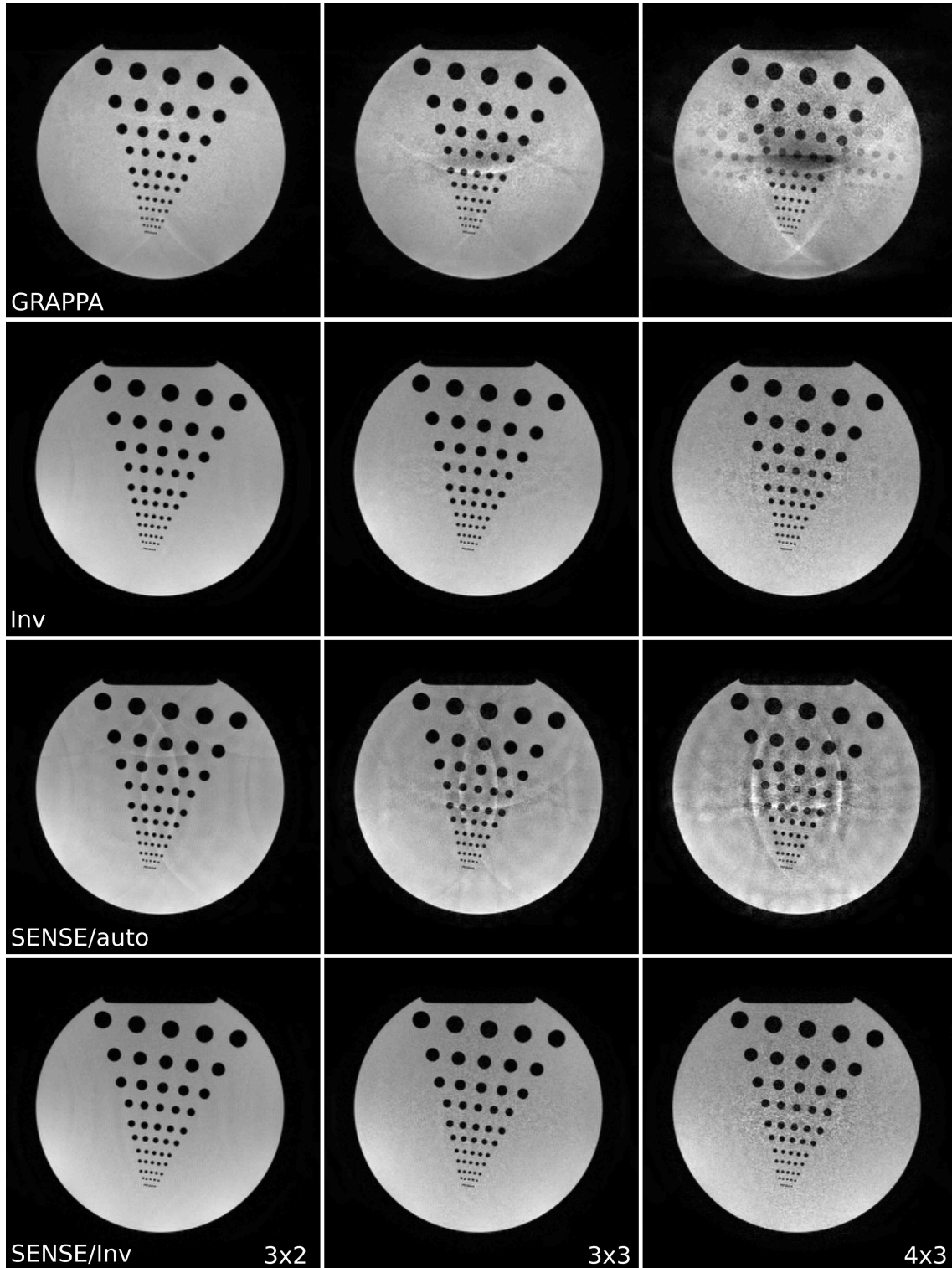
## 5.4 Experiments

### 5.4.1 Methods

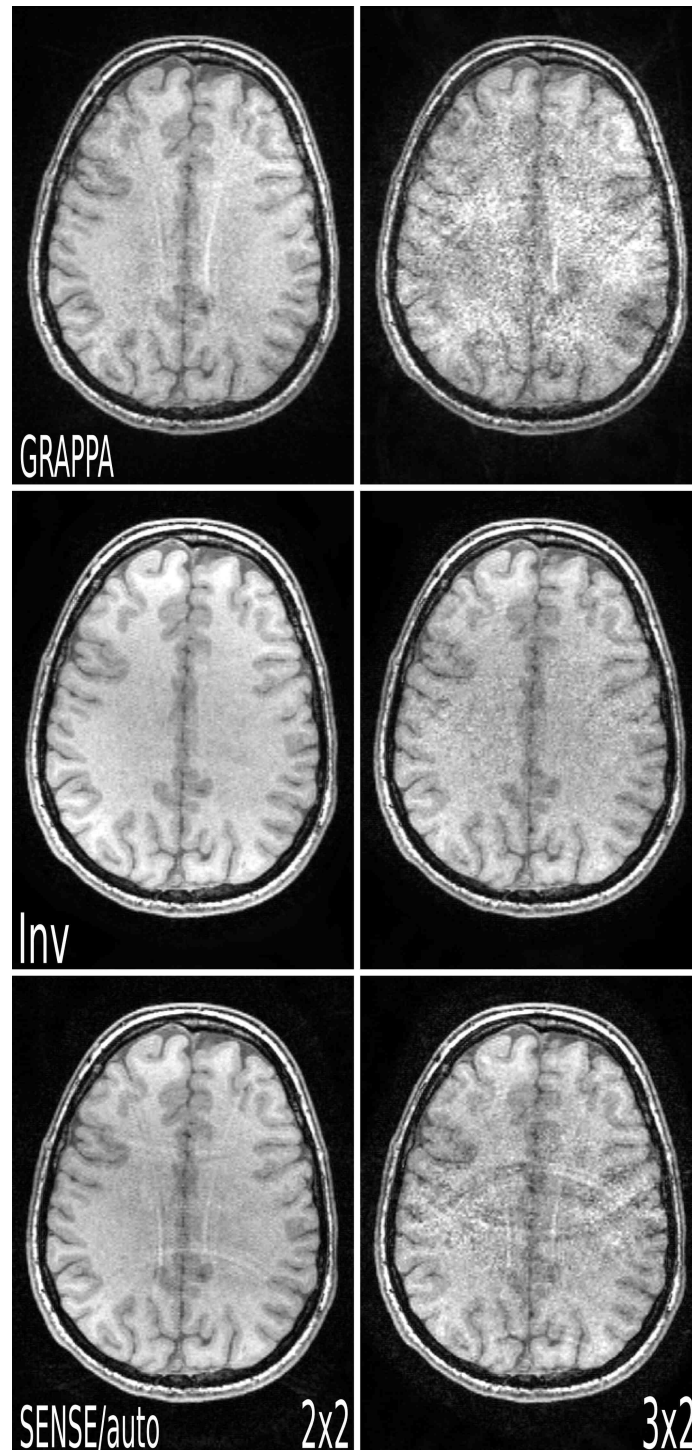
Experimental examples were obtained for a water phantom and the brains of healthy volunteers. Written informed consent was obtained from all subjects prior to the examination.

The data was reconstructed with the proposed algorithm, and results were compared to the GRAPPA algorithm implemented on the MRI system (software version: VB13) and to an autocalibrated version of SENSE. As confirmed by personal communication the GRAPPA implementation is based on the algorithm presented in [37] but contains unpublished proprietary modifications. It may still serve as a valid reference because it is widely distributed and its characteristics are documented in numerous studies about parallel imaging methods and applications. The SENSE algorithm used here relies on the autocalibration technique from [78] – including an apodization with the Kaiser window ( $\beta = 4$ ) – and employs an iterative (CG) SENSE reconstruction [85] for image reconstruction. This generic SENSE version was chosen, because it allows for the inclusion of reference lines in the reconstruction – similar to GRAPPA and the proposed nonlinear inversion algorithm.

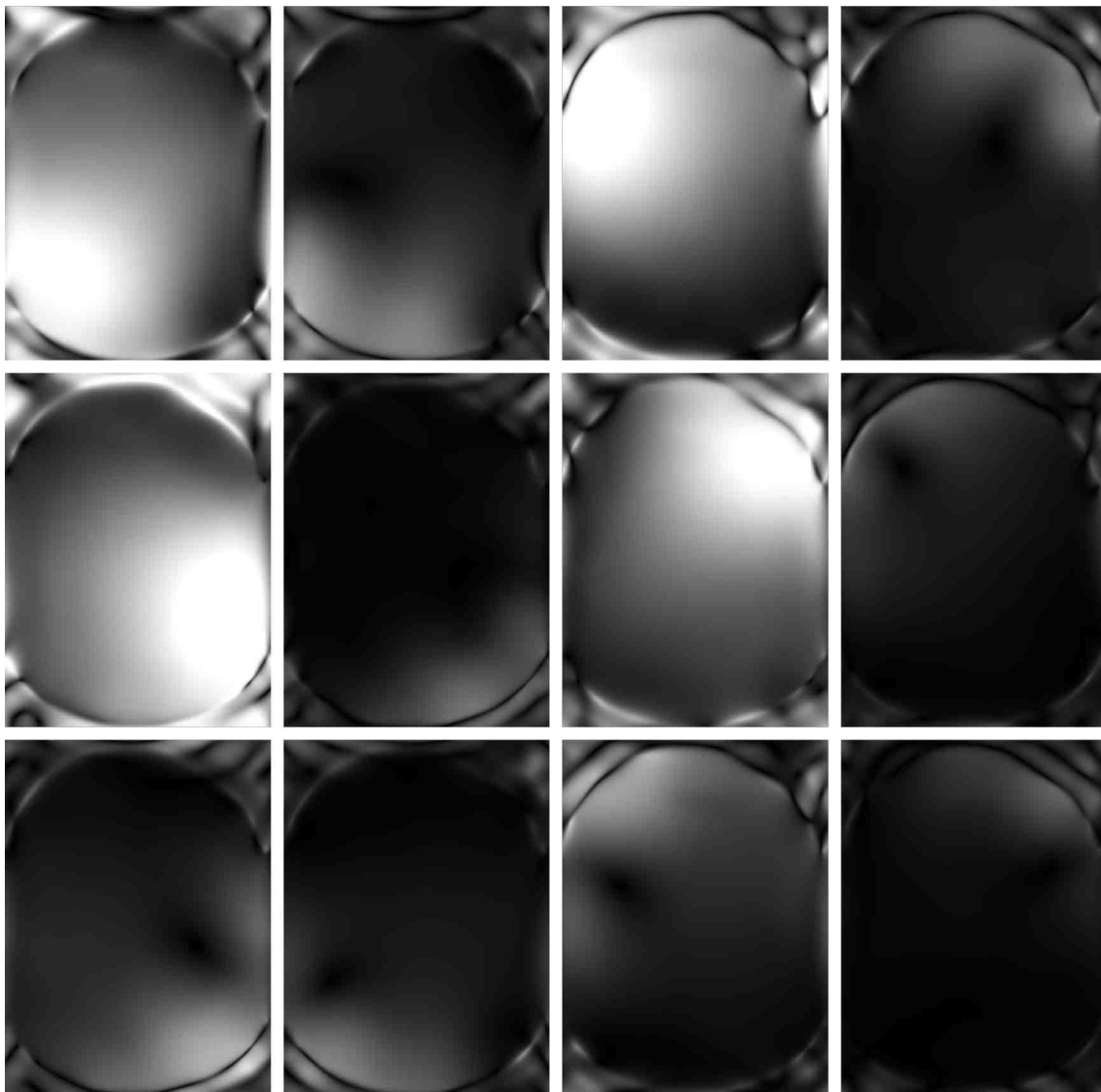
Images were acquired with the use of a 3D RF-spoiled FLASH sequence (TR/TE = 10.6/4.2 ms, flip angle = 17°) with an isotropic spatial resolution of 1 mm. Parallel imaging was performed with variable acceleration (undersampling) factors in all phase-encoding directions and with variable numbers of reference lines. To improve the SNR of the 3D gradient-echo images of the human brain, the acquisitions employed two accumulations. For demonstration purposes individual partitions were selected from the 3D data orthogonal to the readout direction, that is after 1D Fourier transformation along the frequency-encoding axis. These 2D partitions were extracted and subsequently reconstructed using the proposed algorithm.



**Figure 5.1:** *3D FLASH MRI partitions of a phantom reconstructed with GRAPPA, auto-calibrated SENSE (SENSE/auto), nonlinear inversion (Inv) and SENSE with coil sensitivities taken from nonlinear inversion (SENSE/Inv) for 2D acceleration factors of  $6 = 3 \times 2$ ,  $9 = 3 \times 3$ , and  $12 = 4 \times 3$  using  $24 \times 24$  reference lines and 16, 19, and 21 iterations of the nonlinear inversion algorithm, respectively.*



**Figure 5.2:** 3D FLASH MRI partitions of the human brain reconstructed with GRAPPA, autocalibrated SENSE and nonlinear inversion (Inv) for 2D acceleration factors of  $4 = 2 \times 2$  and  $6 = 3 \times 2$  using  $16 \times 16$  reference lines using 14 and 18 iterations, respectively.



**Figure 5.3:** *Coil sensitivities corresponding to the Inv reconstruction (2D acceleration factor 4, 14 iterations) shown in Fig. 5.2.*

### 5.4.2 Results

Figure 5.1 shows 3D MRI partitions of a water phantom obtained for 2D acceleration factors of  $6 = 3 \times 2$ ,  $9 = 3 \times 3$ , and  $12 = 4 \times 3$  using  $24 \times 24$  reference lines. All acquisitions employed the 12 channels of the head coil and the images were reconstructed with GRAPPA, autocalibrated SENSE and nonlinear inversion, respectively. Visual inspection reveals that GRAPPA and autocalibrated SENSE lead to ghosting artifacts for a 2D acceleration factor of 6 that get increasingly worse for higher degrees of undersampling. In contrast, the proposed nonlinear inversion method is able to reconstruct almost artifact-free images up to the theoretical maximum of 12, although at this level the quality in central areas of the image begins to deteriorate. The SENSE reconstructions in the bottom part of Figure 5.1 demonstrate that the high quality of the nonlinear reconstructions is primarily due to an improved estimation of the sensitivity profiles. SENSE reconstructions achieve a similar quality if the coil sensitivities are not estimated from the reference lines but taken from the calculations of the nonlinear inversion algorithm.

The above findings are confirmed for in vivo conditions. GRAPPA, autocalibrated SENSE and nonlinear inversion reconstructions for 3D MRI partitions of the human brain are compared in Fig. 5.2 for 2D acceleration factors of  $4 = 2 \times 2$  and  $6 = 3 \times 2$  using  $16 \times 16$  reference lines. Again, GRAPPA and autocalibrated SENSE images yield ghosting artifacts and considerably more noise than reconstructions by nonlinear inversion (using 14 and 18 iterations). For the nonlinear inversion reconstruction with 2D acceleration factor  $2 \times 2$ , Figs. 5.3 to 5.5 detail the extreme smoothness of the estimated coil sensitivities as well as the influence of the total number of iterations and the number of iterative updates of the coil information, respectively. The progressive reduction of the undersampling artifacts by increasing the number of iterations from 11 to 14 is demonstrated in Fig. 5.4. Complementary, Fig. 5.5 supports the notion that the improved reconstructions by nonlinear inversion are largely due to the improved estimation of the coil sensitivities. When limiting the update of the coil information to 8, 10, 12 or 14 iterations, the remaining iterations of the image information (up to 14) refer to a linear inversion similar to CG-SENSE. The comparison of respective reconstructions in Figure 5.5 clearly indicates that an early termination of the coil update leaves the rest of the iterations with slightly incorrect coil sensitivities which in turn give rise to residual ghosting artifacts in the final image. It should be noted that because these artifacts are related to errors in the coil sensitivities and not due to the regularization, they cannot be eliminated by adding further iterations without including an update of

the coil information.

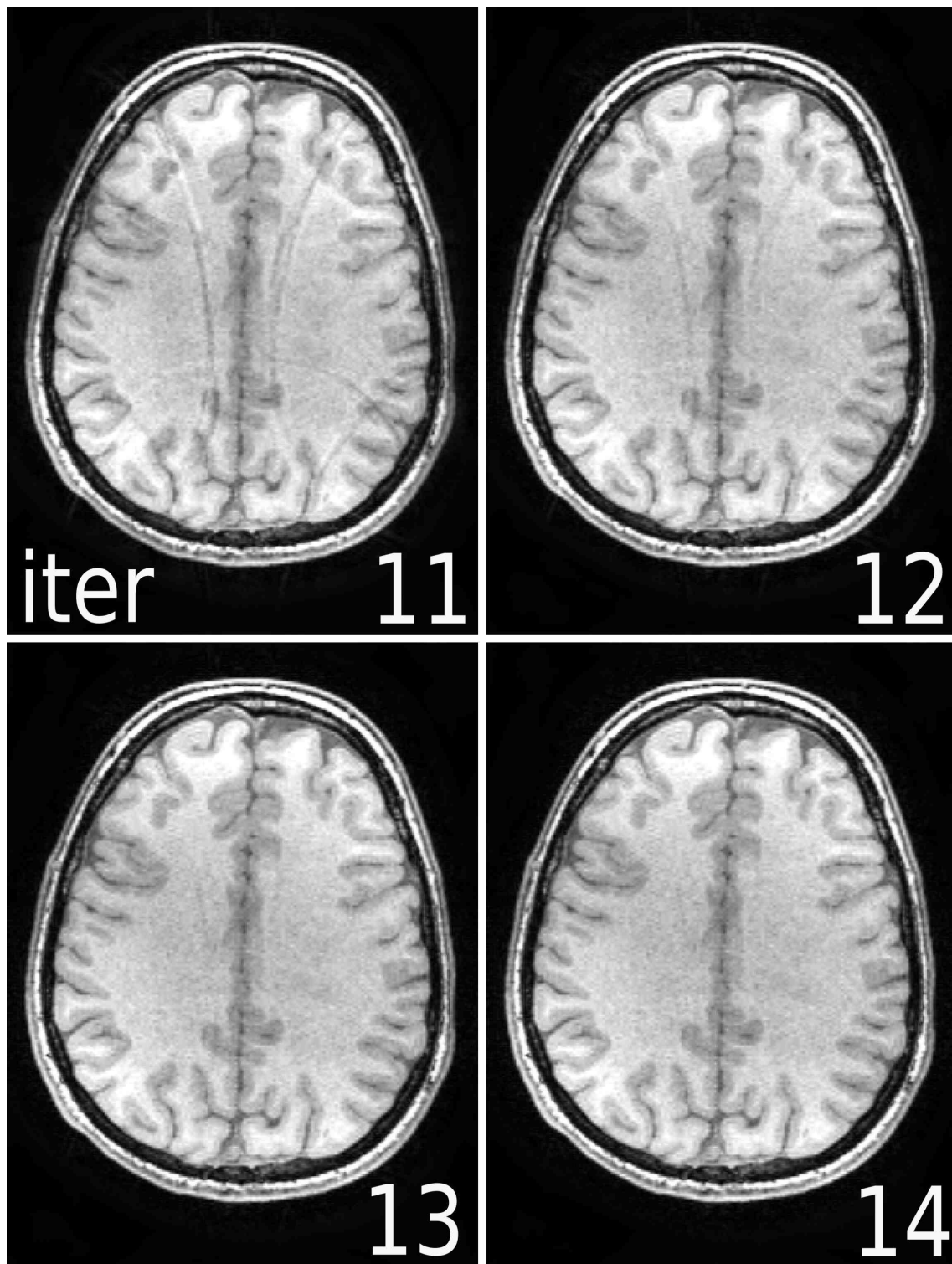
The relevance of the number of reference lines is examined closer in Figure 5.6 which compares 3D FLASH MRI partitions of the human brain reconstructed with GRAPPA, autocalibrated SENSE and nonlinear inversion (2D acceleration factor 4) for  $24 \times 24$ ,  $16 \times 16$ , and  $8 \times 8$  reference lines. While reconstructions by nonlinear inversion exhibit no ghosting artifacts and only a moderate increase of central noise for the lowest number of reference lines, GRAPPA and SENSE reconstructions are much more sensitive to accurate estimations of the coil information from a sufficiently high number of reference lines. Severe undersampling artifacts occur for both  $16 \times 16$  and  $8 \times 8$  reference lines.

## 5.5 Extensions

### 5.5.1 Partial Fourier Imaging

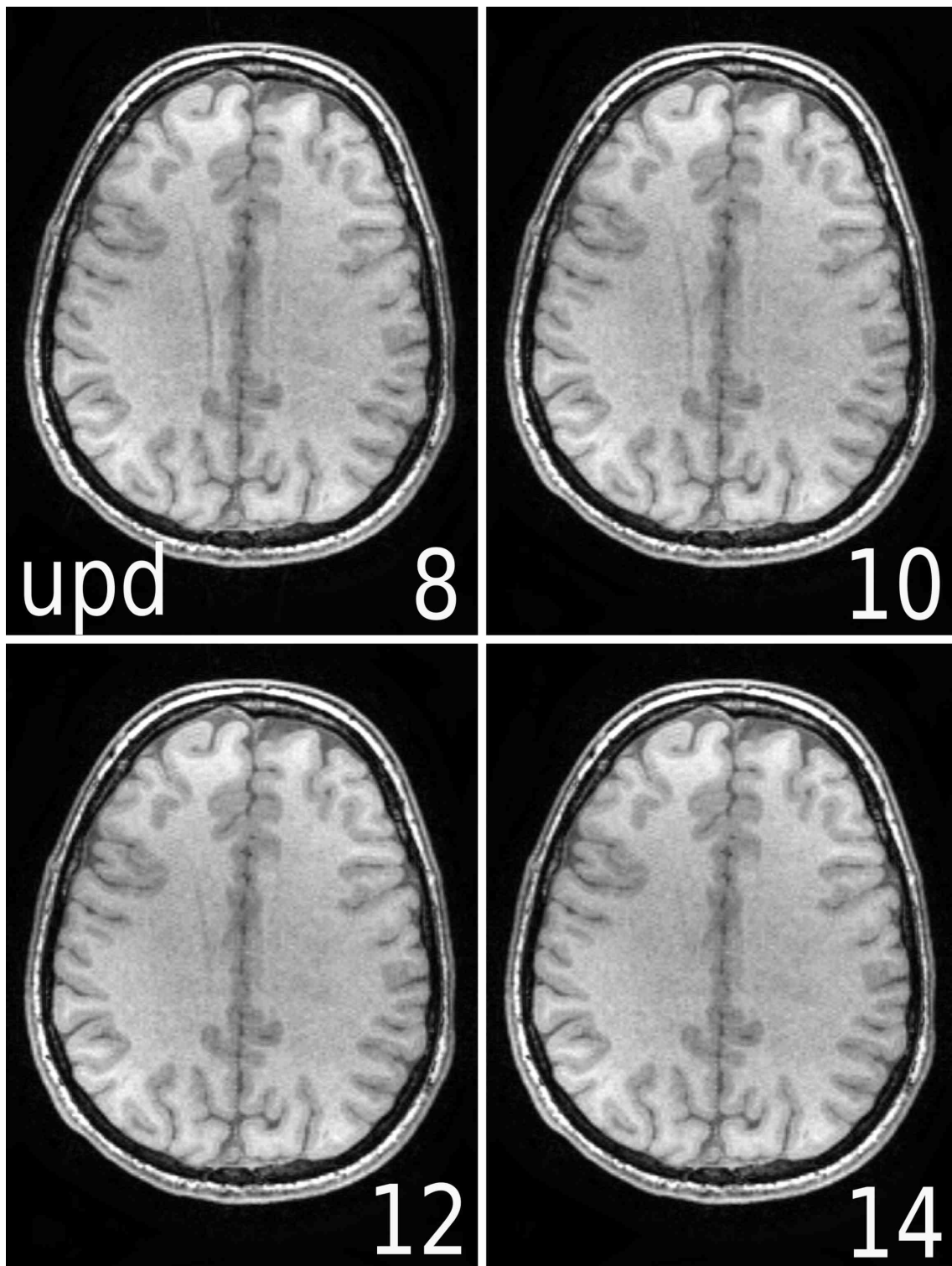
Partial Fourier imaging exploits the fact, that the fundamental measurement quantity, the spin density, is real. Under ideal circumstances, only half of the k-space must be measured, whereas the other half could be completed by complex conjugation. Unfortunately, various kinds of systematic errors, especially off-resonance effects lead to phase variations in the magnetization, even after taking into account the complex-valued sensitivity maps of the receive coils. For this reason, partial Fourier techniques typically estimate a low-resolution phase map from the fully sampled center and use this map in the reconstruction algorithm to compensate the phase variations. Of course, this only works, if the phase variations are rather smooth and can be obtained from the k-space center. Whenever the fundamental assumption of a low-varying phase is violated and a high-resolution phase map can not be determined with other methods (e.g., with a previously obtained field map in the absence of other phase effects), data from partial Fourier acquisitions can not properly be reconstructed.

The integration of partial Fourier acquisition with parallel imaging sometimes realizes higher acceleration factors than each method alone [110, 20]. For the proposed algorithm, this combination turns out to be almost trivial: Simply by constraining the image function to be real, the (low-resolution) phase variations are automatically moved to the sensitivity profiles. An example is shown in Figure 5.7, where only 5/8 of the full k-space was acquired. According to the convolution theorem, a direct reconstruction of a zero-filled k-space yields images which are blurred and affected by ringing. In comparison, the reconstruction obtained by nonlinear in-

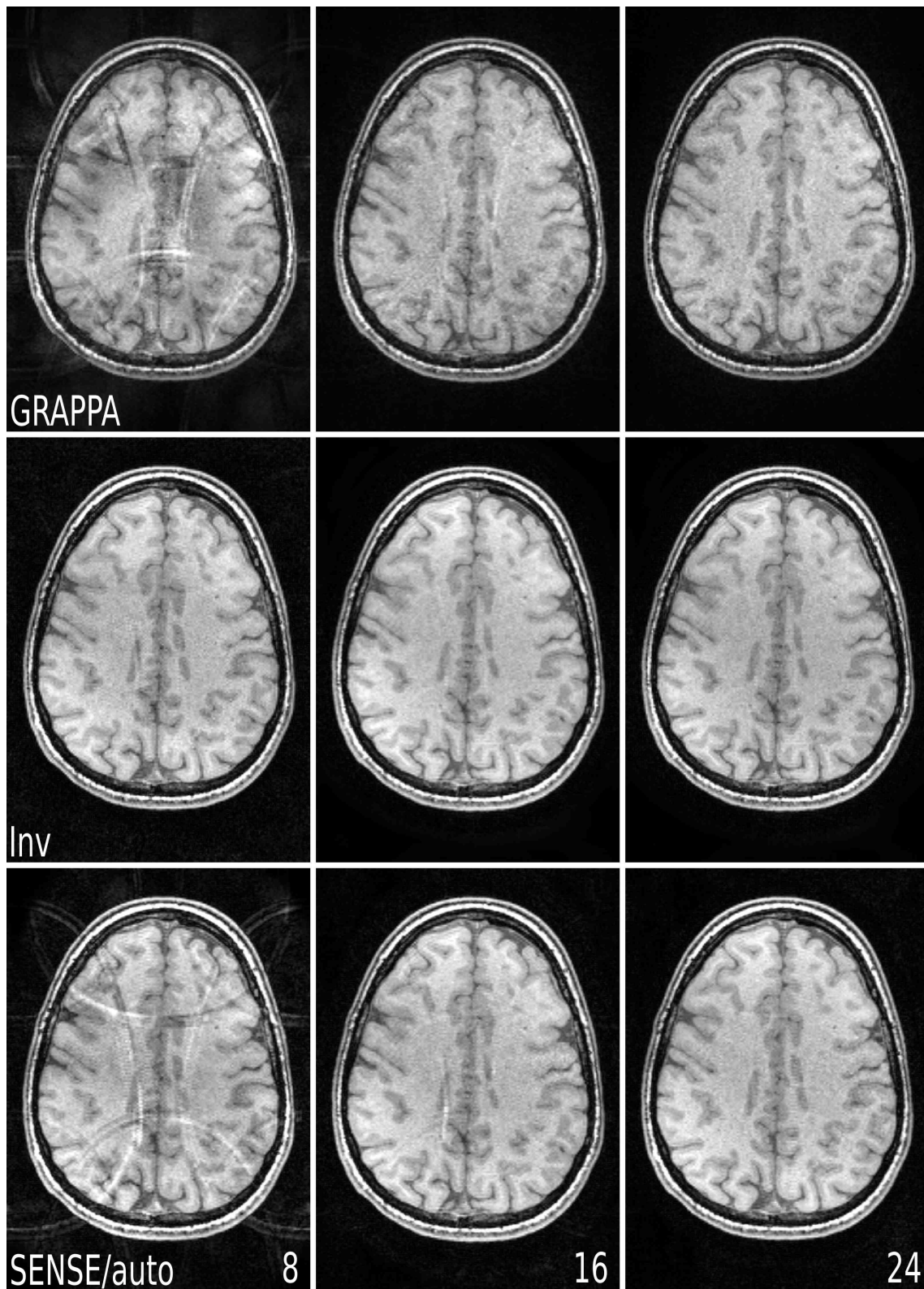


**Figure 5.4:** Influence of the number of iterations (11 to 14) on the Inv reconstruction (2D acceleration factor 4) shown in Fig. 5.2.

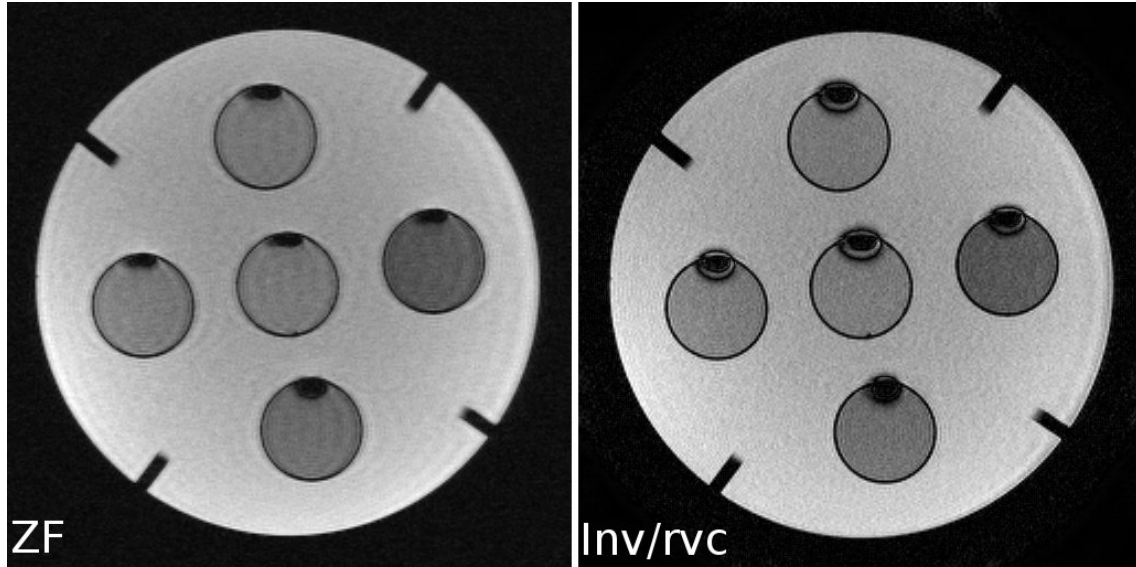




**Figure 5.5:** Influence of the number of coil updates (8 to 14) on the Inv reconstruction (2D acceleration factor 4, 14 iterations) shown in Fig. 5.2.



**Figure 5.6:** 3D FLASH MRI partitions of the human brain reconstructed with GRAPPA, autocalibrated SENSE and nonlinear inversion (Inv) for a 2D acceleration factor of 4 using  $24 \times 24$ ,  $16 \times 16$ , and  $8 \times 8$  reference lines with 14 and 22 ( $8 \times 8$ ) iterations.



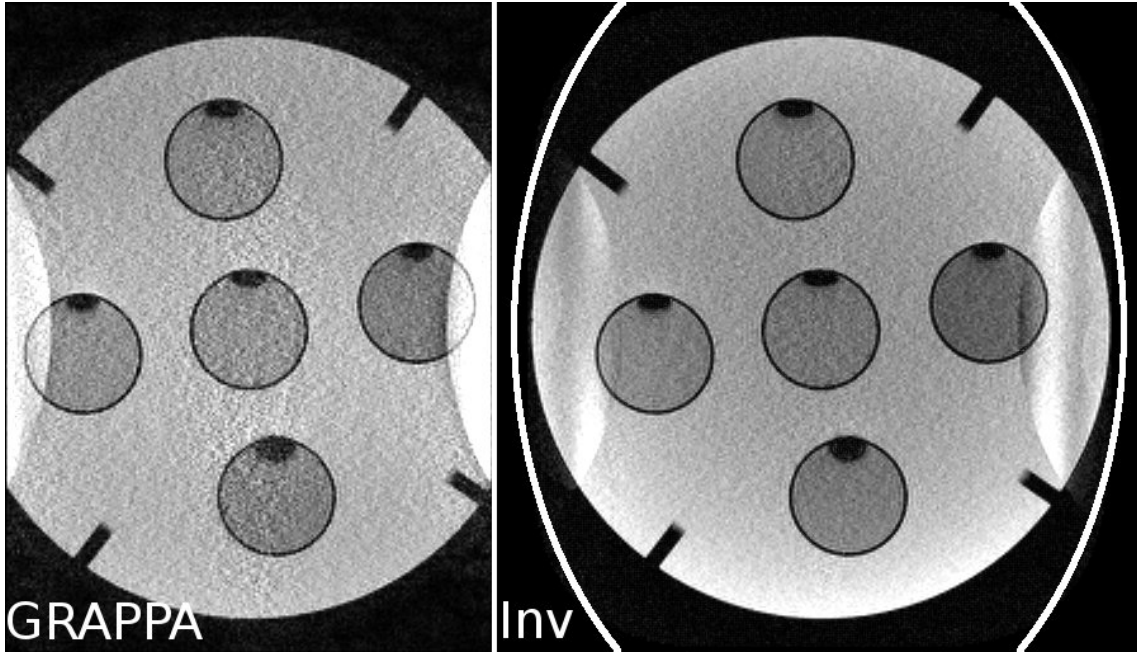
**Figure 5.7:** *Partial Fourier acquisition (left) Zero-filling vs a reconstruction with regularized nonlinear inversion and real-value constraint. The zero-filled  $k$ -space translates into a blurred image. (Right) This blurring is avoided in the phase-constrained reconstruction. Here, phase-cancellation effects are visible at the border of the air bubbles.*

version with real-value constraint is free from these problems. This example also demonstrates the occurrence of phase cancellation artifacts at the border between water and air bubbles in the phantom, where the magnetic field is disturbed by severe susceptibility changes.

### 5.5.2 Reconstruction with Reduced Field of View

As discussed in Chapter 3.6.3, one of the advantages of the GRAPPA algorithm is the possibility to reconstruct images with a FOV, which was made smaller in phase direction than the size of the measurement object. This situation, which basically corresponds to the use of a fractional acceleration factor slightly higher than the nominal one, turns out to be difficult to handle for autocalibration techniques of image-based reconstruction algorithms. Here, the low-resolution image which can be calculated from the fully sampled center already contains aliasing artifacts, which makes the determination of the necessary coil sensitivities impossible. Because this situation is of practical importance, it is shown how to adapt the regularized nonlinear inversion algorithm to this case.

First, it should be noted, that for algorithms which operate in the image domain it is possible to constrain the support of the object in arbitrary ways by including an



**Figure 5.8:** *Reconstruction with reduced FOV. (Left) GRAPPA reconstruction (Right) Reconstruction with nonlinear inversion (Inv). The size of the mask is indicated.*

appropriately chosen mask in the forward operator. This mask reduces the admissible support of the object during reconstruction to arbitrary regions inside of the FOV used for the measurement. Second, when the FOV as well as the acceleration factor are changed by the same factor (e.g., doubled), but the mask constrains the support of the object to the original value, then the reconstruction problem remains basically unchanged. The only difference are the reference lines in the k-space center, which are now closer together, yielding an aliasing-free determination of the coil sensitivities. Starting from here, it is possible to increase the support of the object to regions larger than the original FOV, also realizing fractional acceleration factors. For the nonlinear inversion algorithm, it turns out, that the higher sampling density of the reference lines is often not required. When the mask is only slightly extended over the original FOV and includes the complete object, the algorithm is able to reconstruct images, which are free from artifacts in the relevant center part. A comparison with a GRAPPA reconstruction is shown in Figure 5.8 for data acquired with an acceleration factor of 2, 4 channels, and 24 reference lines.

While in contrast to GRAPPA, this currently requires manual intervention to determine the true size of the measurement object, in principle it is possible to transfer the superior reconstruction quality of the nonlinear algorithm even to this situation. Another slight disadvantage is the need to double the matrix size in the

reconstruction. As described in the following section, this doubling is also the main step needed to allow for the reconstruction of non-Cartesian data. This opens up an even nicer solution for this imaging scenario: With the ability to reconstruct non-Cartesian data, there is no reason any more to require the fully sampled center to be on the same grid as the rest of the data. Here, the distance of the measured  $k$ -space points in the center should be chosen small enough as required by the Nyquist theorem, while the outer part can then be sampled with a pattern which can be freely chosen to match the given speed and quality requirements. This would result in reconstructions that are completely free from artifacts and thus allow arbitrary fractional acceleration factors.

### 5.5.3 Non-Cartesian Trajectories

An extension to non-Cartesian trajectories can be achieved with the inclusion of an interpolation step [62]. An even more promising approach, however, results from a technique which is based on the convolution of the PSF. This is similar to the convolution-based SENSE algorithm described earlier and will be described in the following.

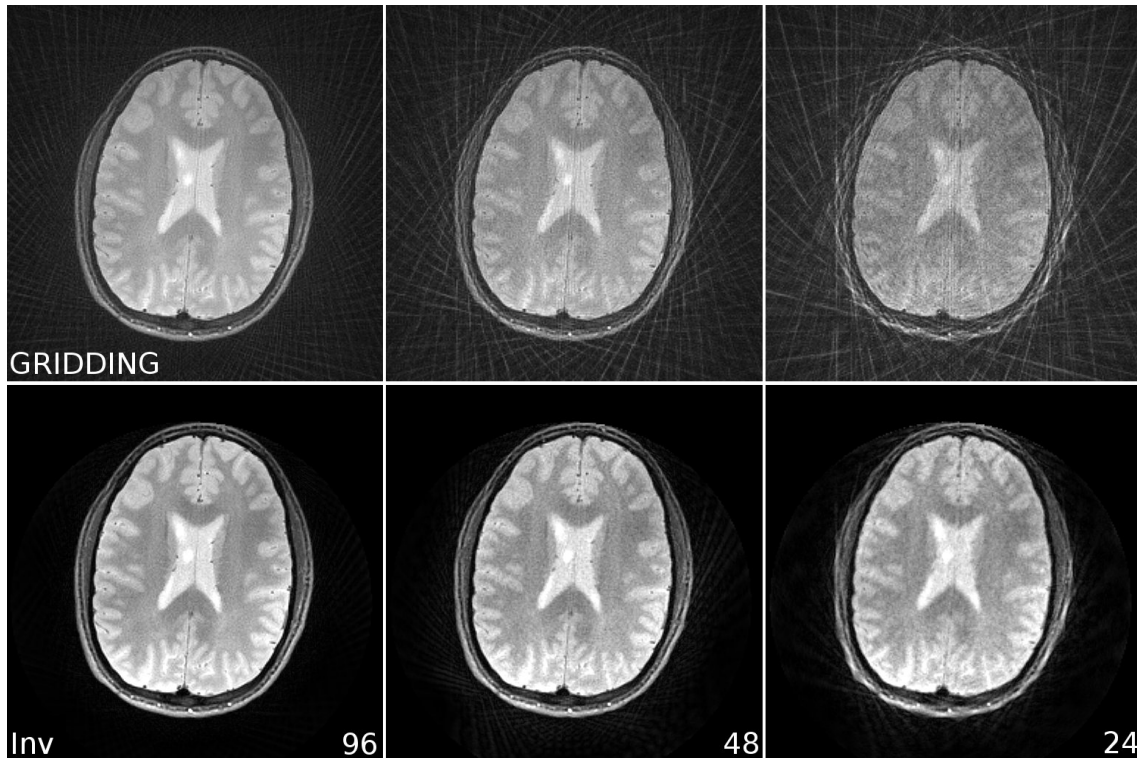
Again, starting from a continuous description, the forward operator  $F$  can be decomposed into a (nonlinear) operator  $C$ , which contains the multiplication of the object with the sensitivities, a projection onto the FOV, the Fourier transformation  $\mathcal{F}$ , and a projection  $P_{\mathbf{k}}$  onto the trajectory:

$$F = P_{\mathbf{k}}\mathcal{F}P_{FOV}C$$

Inserting this operator into the update rule of the IRGNM and making use of idempotence and self-adjointness of an orthogonal projection ( $PP = P$  and  $P^H = P$ ) and the fact that the data  $y$  is already given on the trajectory  $P_{\mathbf{k}}y = y$  yields

$$\begin{aligned} dx = & (DC(x_n)^H P_{FOV}\mathcal{F}^{-1}P_{\mathbf{k}}\mathcal{F}P_{FOV}DC(x_n) + \alpha_n I)^{-1} \\ & (DC(x_n)^H P_{FOV}\mathcal{F}^{-1}(y - P_{\mathbf{k}}\mathcal{F}P_{FOV} \circ C(x_n)) + \alpha_n(x_n - x_0)) \quad . \end{aligned}$$

As in the case of the convolution-based SENSE algorithm, both terms  $\mathcal{F}^{-1}P_{\mathbf{k}}\mathcal{F}$  can be understood as a convolution with the PSF. Because this convolution is to be calculated on an area with compact support, it can, after discretization, efficiently be implemented with the use of an FFT algorithm. In a preparation step, the vector  $y$  is discretized by interpolating the data onto a grid and the convolution kernel is calculated.



**Figure 5.9:** *Reconstruction from data acquired with a radial spin-echo sequence. 96, 48 and 24 spokes: (top) gridding and (bottom) Inv reconstruction*

Figure 5.9 shows images calculated from data of a radial spin-echo sequence in comparison to a gridding reconstruction. Here, the data was acquired with 96, 48, and 24 spokes and using the 12 channel head coil. The gridded images are increasingly affected by strong undersampling (streaking) artifacts for lower number of spokes, which may be avoided by the nonlinear inversion algorithm.

## 5.6 Discussion

This work introduces a new reconstruction method for autocalibrated parallel imaging which is based on regularized nonlinear inversion. The approach allows for a simultaneous calculation of the unknown coil sensitivity maps and the unknown spin density of the object using all available data. At least for the experimental conditions examined here, that is 2D acceleration factors of up to 12, reference lines as few as  $8 \times 8$ , and conventional 2D and 3D MRI sequences, the proposed strategy yields images with visually reduced artifacts compared to the two-step approaches GRAPPA and autocalibrated SENSE. These methods first estimate information about the coil sensitivities (which in the case of GRAPPA is encoded in the recon-

struction weights) from only a part of the data and then solve a linear equation where the coil information remains fixed.

The necessity to improve the estimation of the coil sensitivity profiles in parallel image reconstructions has also been recognized by others. As mentioned before, it has recently been proposed to exploit the bilinear structure of the MRI signal equation to solve the system of equations for the coil profiles and object functions in an alternating way [113]. In comparison to the alternating minimization scheme, an advantage of the Newton methods used here is a greater flexibility for incorporating additional nonlinear constraints and regularization terms. Such options will be even more important for higher acceleration factors that suffer from increased noise amplification. A Newton-type method has also been used for parallel MRI in [7], but the algorithm presented there computes the full Jacobian  $DF(x_n)$  and uses a QR decomposition to solve the regularized Newton equations. Therefore, the time and memory complexity is much higher compared to the method presented here. Moreover, the use of a small number of basis functions, which ensures the smoothness of the coil profiles in [113] and [7] seems problematic because it limits the accuracy of the profile reconstructions as opposed to the use of a regularization term.

## 5.7 Summary

The formulation of the MRI reconstruction for multiple receive coils as a nonlinear inversion problem and its application to autocalibrated parallel imaging has been demonstrated to markedly improve the achievable image quality for high acceleration factors when compared to conventional methods.

In addition to the basic algorithm, which is published in [103], three additional extensions have been described: The use of a real-value constraint in combination with a partial Fourier acquisition scheme, an extension to situations with a reduced FOV, and, finally, the extension to non-Cartesian trajectories. Here, the improved sensitivities allow for reconstructions of radially sampled data with an image quality approaching that of the algorithm described in [13, 11], but without the use of the – sometimes undesirable – total variation (TV) penalty.

# 6

## Segmented Diffusion Imaging

### 6.1 Introduction

Diffusion-weighted (DW) MRI is a technique with important clinical and scientific applications. For example, it is a highly relevant tool in the diagnosis of acute stroke, as the degree of water diffusion in affected brain regions is reduced. Because it further provides directional information about the diffusion process, which in turn can be used to track nerve fibers in the living brain, it emerges as an increasingly used method to study the structural integrity and fiber architecture of white matter.

In this chapter, a reconstruction technique for segmented DW MRI is presented. In MRI segmented k-space acquisitions are often attractive because they reduce the length of an acquired echo train. For example, in EPI a shorter train of gradient echoes translates into reduced susceptibility artifacts, while in rapid stimulated echo acquisition mode (STEAM) MRI fewer stimulated echoes yield higher flip angles and an improved SNR as recently demonstrated for black-blood cardiac MRI [60]. Unfortunately, however, segmented multi-shot acquisitions are not easily applicable to DW MRI. This is because of the occurrence of spatially varying (nonlinear) phase variations which are caused by unavoidable non-rigid brain pulsations and patient movements during the action of the self-compensating diffusion-encoding gradients. This phase information differs for each individual segment and therefore prohibits a direct image reconstruction by a simple combination of all k-space data.

Current attempts to reconstruct motion-affected segmented k-space data make use of low-resolution phase maps that are either obtained by a two-dimensional (2D) navigator acquisition [79, 74, 3] or extracted from the fully sampled center of a self-



calibrating k-space trajectory [84, 72, 82]. As a common feature, these techniques rely on the assumption that the motion-associated phase varies only smoothly, so that it can be accurately described by low spatial frequencies. In DW MRI of the human brain this condition is not generally fulfilled. A robust and artifact-free image reconstruction from segmented k-space acquisitions therefore requires high-resolution phase maps. These experimental approaches, however, face shortcomings: For DW EPI, a large 2D navigator may be compromised by susceptibility problems which limit the accuracy of the resulting phase maps, while for DW STEAM MRI, the need for multiple navigator acquisitions would eliminate the SNR advantage of a shortened echo train.

Here, an alternative solution is presented which is based on an iterative inverse image reconstruction technique that exploits the information from the entire k-space as well as from multiple receive coils. The new method allows for the reconstruction of DW images from segmented k-space acquisitions without motion-induced artifacts and without the need for additional navigator acquisitions. It provides a more general and robust solution than a recent proposal by Skare and colleagues [97] which relies on a standard non-iterative SENSE or GRAPPA reconstruction and a homogeneously sampled EPI trajectory.

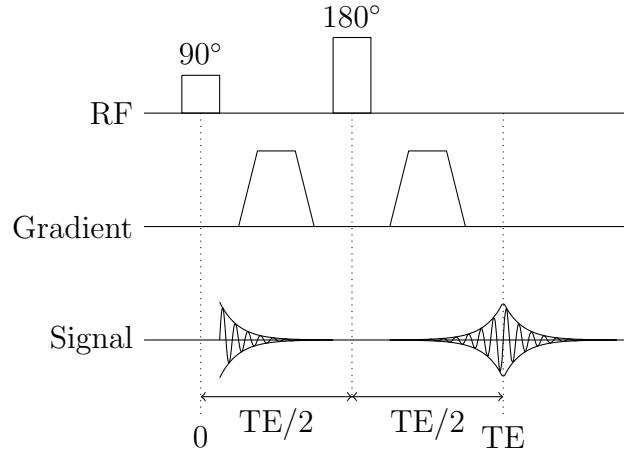
## 6.2 Theory

### 6.2.1 Diffusion Tensor Imaging

Diffusion is random molecular motion. Often, diffusion processes are obstructed in various directions. In the presence of magnetic field gradients moving spins acquire an additional phase relative to stationary spins. For diffusion processes this movement is random, causing a loss of coherence and corresponding signal loss. To describe the influence of the diffusion process onto the MRI experiment, the Bloch equations have been extended to the Bloch-Torrey equations [102]

$$\frac{\partial}{\partial t} \mathbf{M}(\mathbf{x}, t) = \gamma \mathbf{M}(\mathbf{x}, t) \times \mathbf{B}(\mathbf{x}, t) + \begin{pmatrix} \frac{-M_x(\mathbf{x}, t)}{T_2} \\ \frac{-M_y(\mathbf{x}, t)}{T_2} \\ \frac{M_z(\mathbf{x}, 0) - M_z(\mathbf{x}, t)}{T_1} \end{pmatrix} + \text{div } D \text{ grad } \mathbf{M}(\mathbf{x}, t) .$$

In this equation a diffusion term based on a symmetric positive definite tensor  $D$  as a simple model for the directional nature of the diffusion process was included. Note, that divergence and gradient operations act component-wise on the magnetization. In the absence of a  $\mathbf{B}^1$  field and for  $\mathbf{B}^0$  as in (2.5), the evolution of the transversal



**Figure 6.1:** *Stejskal-Tanner-Sequence: Strong gradients are used to create a diffusion-weighted spin echo. The two gradients left and right of the 180° pulse exactly compensate each other. Thus, stationary spins are unaffected, while moving spins acquire some phase. Due to the random Brownian movement of the diffusion process, this phase offset leads to an incoherent dephasing of the spin and in consequence to a reduced signal strength. In contrast, joint movement of many spins leads to coherent dephasing, which is visible as phase distortions in the reconstructed image.*

magnetization is then described by

$$\frac{\partial}{\partial t} M_{\perp}(\mathbf{x}, t) = - \left( \frac{1}{T_2} + i(\omega_0 + \gamma \mathbf{x} \cdot \mathbf{G}(t)) \right) M_{\perp}(\mathbf{x}, t) + \text{div } D \text{ grad } M_{\perp}(\mathbf{x}, t) .$$

The effect of diffusion in the presence of field gradients on the transversal magnetization can be demonstrated with the simple ansatz

$$M_{\perp}(\mathbf{x}, t) = m(t) e^{-\left(\frac{1}{T_2} + i(\omega_0 + 2\pi \mathbf{k}(t) \cdot \mathbf{x})\right)} .$$

It yields an ordinary differential equation

$$\frac{dm(t)}{dt} = -4\pi^2 m(t) \mathbf{k}(t)^T D \mathbf{k}(t)$$

for  $m(t)$ . The solution of this equation is given by

$$\begin{aligned} m(t) &= m(0) e^{-4\pi^2 \int_0^t dt' \mathbf{k}(t')^T D \mathbf{k}(t')} \\ &= m(0) e^{-\text{Tr}(DB(t))} \quad \text{with} \quad B(t) = 4\pi^2 \int_0^t dt' \mathbf{k}(t') \mathbf{k}(t')^T . \end{aligned}$$

Whenever the direction of the field gradient  $\mathbf{G}(t)$  is constant, the effect on the measurement can be specified by its direction  $\mathbf{g}$  and the so called *b-value*  $b = \text{Tr}B(\text{TE})$  at the time TE of the acquired echo.

This effect is used in the Stejskal-Tanner sequence [100, 101] (see Figure 6.1) to create a diffusion-weighted echo. Extending the sequence to a complete imaging sequence allows to acquire diffusion-weighted images. Here, the obtained signal is weighted by an additional diffusion term:

$$s(t) = \int d\mathbf{x} \underbrace{\rho(\mathbf{x}) e^{-b\mathbf{g}^T D(\mathbf{x})\mathbf{g}}}_{\rho_{b,\mathbf{g}}(\mathbf{x})} e^{-2\pi i\mathbf{k}(t)\mathbf{x}}$$

The direction  $\mathbf{g}$  of the diffusion gradient and the b-value can be experimentally controlled. After the acquisition of diffusion-weighted images  $\rho_{b,\mathbf{g}}(\mathbf{x})$  for various directions, the diffusion tensor can be estimated for each voxel.

The diffusion tensor can then be analyzed. A measure of the isotropic degree of diffusion is given by the ADC value

$$ADC = \frac{1}{3} \text{tr}(D) .$$

Various derived quantities can be calculated from the eigenvalues  $v_1, v_2, v_3$  of the tensors: A useful quantity for practical purposes is the fractional anisotropy (FA)

$$FA = \sqrt{\frac{3}{2} \frac{(v_1 - ADC)^2 + (v_2 - ADC)^2 + (v_3 - ADC)^2}{v_1^2 + v_2^2 + v_3^2}} ,$$

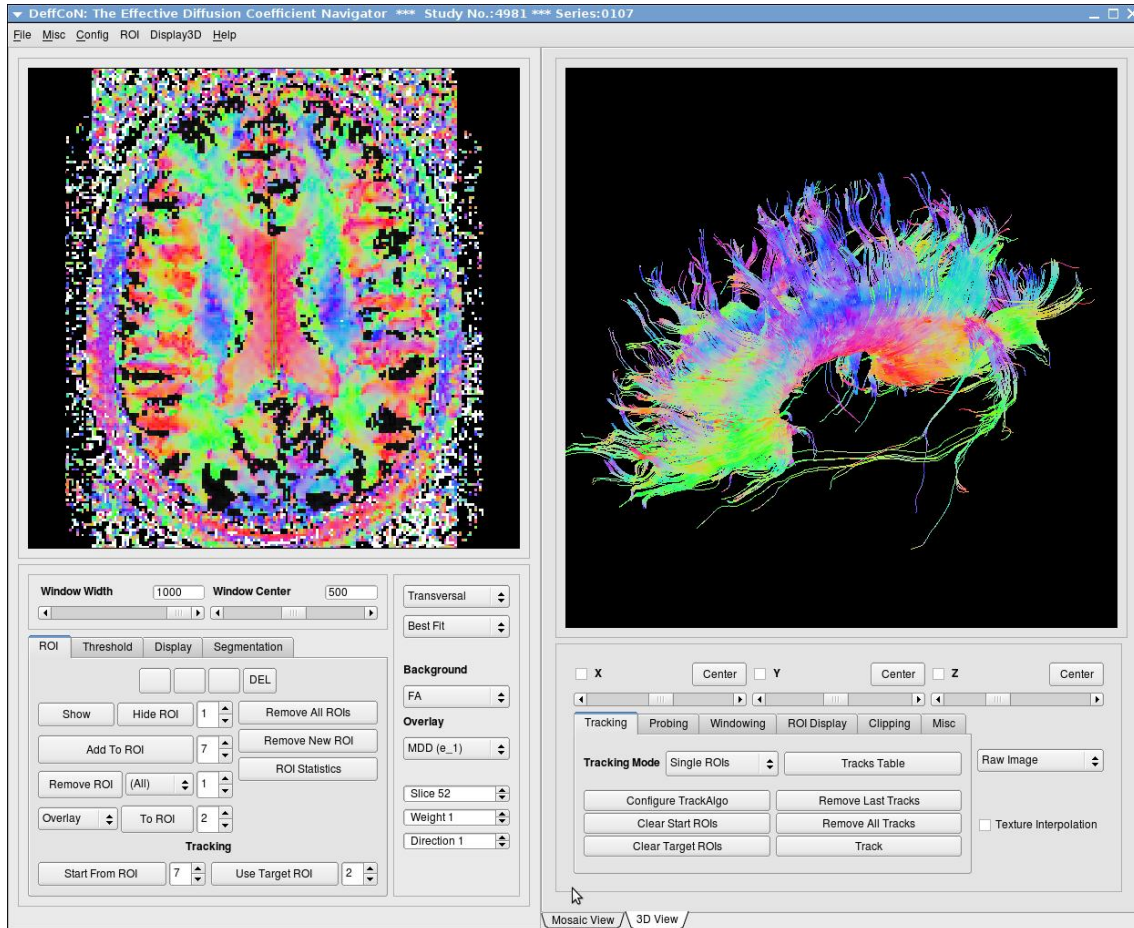
which quantifies the anisotropy of the diffusion tensor. Where FA values are high, it makes sense to study the main diffusion direction, which is given by the eigen vector corresponding to the largest eigenvalue. In the white matter of the brain, the direction of this eigen vector is assumed to point into the direction of nerve fibers. This is exploited to analyze fiber connections in the brain (see Fig. 6.2). A more detailed introduction to DTI can be found in [64].

### 6.2.2 Segmented Diffusion Imaging

In DW MRI as in many other imaging scenarios phase effects have to be taken into account by including non-trivial phase maps into the signal equation

$$s_{lj}(t) = \int d\mathbf{x} e^{-2\pi i\mathbf{k}_l(t)\mathbf{x}} \rho_{b,\mathbf{g}}(\mathbf{x}) e^{ip_l(\mathbf{x})} c_j(\mathbf{x}) \quad j = 1, \dots, N; \quad l = 1, \dots, M.$$

The index describes the segment  $l$  of a multi-shot acquisition with  $M$  segments. Because in DW MRI the phase information is not consistent between different segments, the situation requires individual phase maps  $e^{ip_l(\mathbf{x})}$  for each segment  $l$ . In most current approaches, phase maps and coil sensitivity maps are obtained with the use of a 2D navigator acquisition or from additional autocalibration lines in the



**Figure 6.2:** Fiber tracking with the Deffcon program [64]. Diffusion tensors are calculated for each voxel in a 3D data set. Fiber connections in the white matter can be visualized by integrating the vector fields given by the first eigen vector (the main diffusion direction).

k-space center. Given those maps, the resulting reconstruction problem is linear and can be solved with iterative methods such as conjugate gradient-based versions of SENSE.

In Chapter 5 a new algorithm for autocalibrated parallel imaging was presented, which combines the estimation of the coil sensitivities  $c_j$  and the calculation of the image  $\rho$  into a nonlinear reconstruction problem. The advantage of this nonlinear inversion technique, namely a better estimation of both the image and coil sensitivities, allows for higher reduction factors and fewer autocalibration lines than linear inversion techniques. In a similar way, nonlinear inversion might be helpful to improve the phase maps in multi-shot DW MRI as the approach would include the entire available k-space data, instead of relying on a small number of navigator echoes. In fact, after only minor adjustments of the nonlinear inversion algorithm presented earlier, its application to DW MRI lends itself as a generic alternative to conventional reconstruction methods. Despite some advances, however, the algorithm did not yield completely satisfactory results for data sets with pronounced high-frequency phase variations. For this reason, this work presents the development of a new multi-step algorithm, which robustly reconstructs even severely motion-disturbed data.

### 6.2.3 Reconstruction Algorithm

By combining the coil sensitivities  $c_j$  and the motion-associated phase maps  $e^{ip_l}$  into  $N \times M$  generalized maps  $cp_{lj}(\mathbf{x}) = e^{ip_l(\mathbf{x})} \cdot c_j(\mathbf{x})$ , the mathematical problem becomes identical to that of autocalibrated parallel imaging. In fact, when including a low number of reference lines in each segment of a multi-shot acquisition, the regularized nonlinear inversion algorithm may be directly applied for image reconstruction. The implicit combination of the coil sensitivity maps and motion-associated phase maps, however, is physically unmotivated as the occurrence of high spatial frequencies in the phase maps violates the assumption of spatial smoothness for the coil profiles. It nevertheless turns out that a proper choice of the weighting matrix  $W$  offers better image quality than obtainable by a simple navigator-based approach.

An improved solution may be achieved with the use of a 2-step procedure. First, the non-DW images are reconstructed by parallel imaging using nonlinear inversion. This step also yields respective coil sensitivity maps which in a subsequent step are exploited to calculate individual images for each segment by iterative linear inversion. These images have the same motion robustness as any other single-shot acquisition, so that a final image may be calculated by averaging the magnitude images from all segments. This 2-step method mimics the ideas developed for diffusion-

weighted EPI by Skare et al [97], but differs in the use of the nonlinear inversion algorithm to obtain optimum coil sensitivities in the first step. In general, however, the ill-conditioned parallel image reconstructions for individual segments lead to a high noise amplification which affects the SNR of the final combined image.

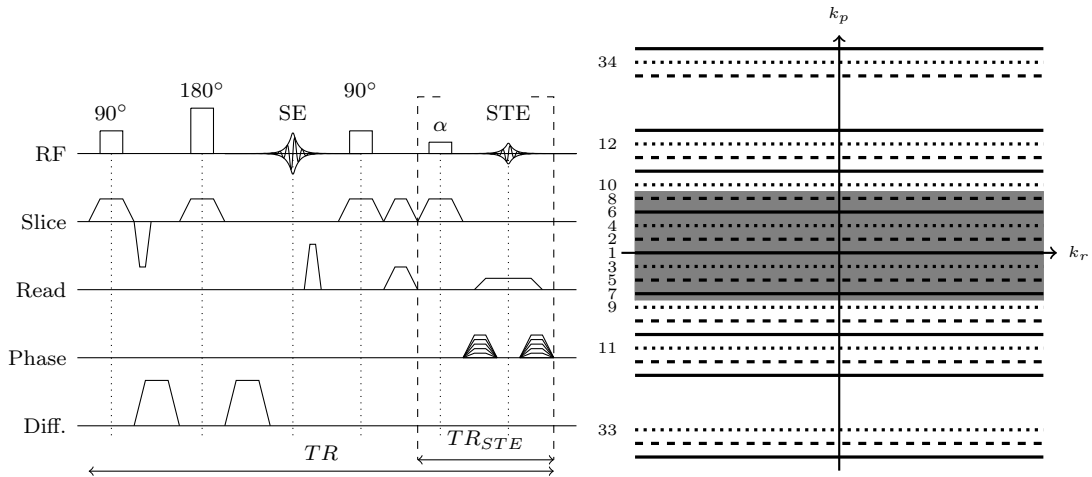
For this reason, the above 2-step method was complemented by another parallel image reconstruction process based on iterative linear inversion. This third step replaces the simple averaging of images from all segments by a true reconstruction of the final image and therefore takes advantage of all data from all segments and coils. The required phase maps for each segment are obtained from the images calculated in the second step and possess the same resolution as the final image. The algorithm retains the motion robustness of the 2-step method, but avoids its inherent noise amplification. Thus, the proposed 3-step method comprises three consecutive calculations:

(1) In a first step, a non-DW image, that is an image without diffusion encoding but otherwise identical acquisition parameters, is reconstructed from the data of all segments by regularized nonlinear inversion as described above. In this step all phase issues are ignored. This is possible because the relevant brain motions do not give rise to pronounced phase variations in the absence of diffusion-encoding gradients. The algorithm simultaneously recovers a high-quality image and sensitivity profile for each receive coil.

(2) In a second step, complex-valued DW images are reconstructed separately for each segment (and all diffusion directions) by linear inversion, that is with the use of an iterative conjugate-gradient version of the SENSE algorithm. This calculation takes advantage of the coil sensitivities  $c_j$  determined in the first step.

(3) In a third step, real-valued DW images are reconstructed from the data of all segments, again by iterative SENSE. This reconstruction uses the coil sensitivities  $c_j$  from the first step and the phase maps  $e^{ip}$  that are available from the images of each segment in the second step. This final step is similar to navigator-based reconstructions, but differs in its use of motion-associated phase maps with much higher spatial resolution.

Taken together, the first step serves to calculate the coil sensitivity maps while ignoring phase problems, the second step estimates high-resolution phase maps that represent motion-associated phase variations (for each individual segment), and the third step calculates real-valued images (combining the data of all segments) by using the previously determined coil sensitivities and phase maps.



**Figure 6.3:** (a) Schematic diagram of a segmented multi-shot STEAM DTI sequence comprising an initial spin-echo (SE) diffusion module and a high-speed STEAM MRI sequence. The sequence generates differently phase-encoded stimulated echoes (STE) by repeating the final readout interval with variable low-flip angle RF pulses ( $\alpha$ , repetition time  $TR_{STE}$ ). The acquisition of multislice diffusion-weighted images is repeated for multiple segments and different diffusion-encoding gradient directions (repetition time  $TR$ ). (b) Coverage of (segmented)  $k$ -space for Cartesian encoding with centric reordering. The example refers to the acquisition of 84 Fourier lines with the use of three segments and 8 reference lines. For details see text.

## 6.3 Materials and Methods

Experiments were performed using either the 8-channel or 32-channel phased-array head coil. Apart from preliminary studies of water phantoms, applications involved DTI of the brain of young healthy adults. Written informed consent was obtained from all subjects prior to the examination.

Segmented DW MRI was based on a rapid STEAM MRI sequence [33, 81] without cardiac gating. As shown in Figure 6.3a the first  $90^\circ$  pulse of the STEAM sequence is replaced by a spin-echo (SE) diffusion module  $90^\circ - DW - 180^\circ - DW - SE$ , while the final acquisition part is repeated in order to generate  $n$  stimulated echoes (STE) – corresponding to  $n$  Fourier lines – for each segment. This readout interval employs RF pulses with variable flip angles to ensure similar signal strengths for each STE. The flip angles may be iteratively calculated according to

$$\alpha_{i-1} = \arctan(\sin\alpha_i) \cdot e^{-TR/T_1}$$

with  $\alpha_n = 90^\circ$  and  $T_1$  the spin-lattice relaxation time. For studies of the human brain  $T_1$  was chosen to be 800 ms for white matter at 3 T.

DTI was performed at 2 mm isotropic spatial resolution using one non-DW image and 24 DW images with  $b$  values of  $1000 \text{ s mm}^{-2}$  along different directions. A total of 51 transverse-to-coronal 2 mm thick sections (orientation along the anterior to posterior commissure) covered the brain with a rectangular  $168 \times 192 \text{ mm}^2$  FOV and a matrix of  $84 \times 96$  complex data points (corresponding to a full Fourier acquisition). For display purposes individual DW images were mildly processed by adaptive filtering taking into account the local intensity distribution and continuation of structures (software supplied by the manufacturer).

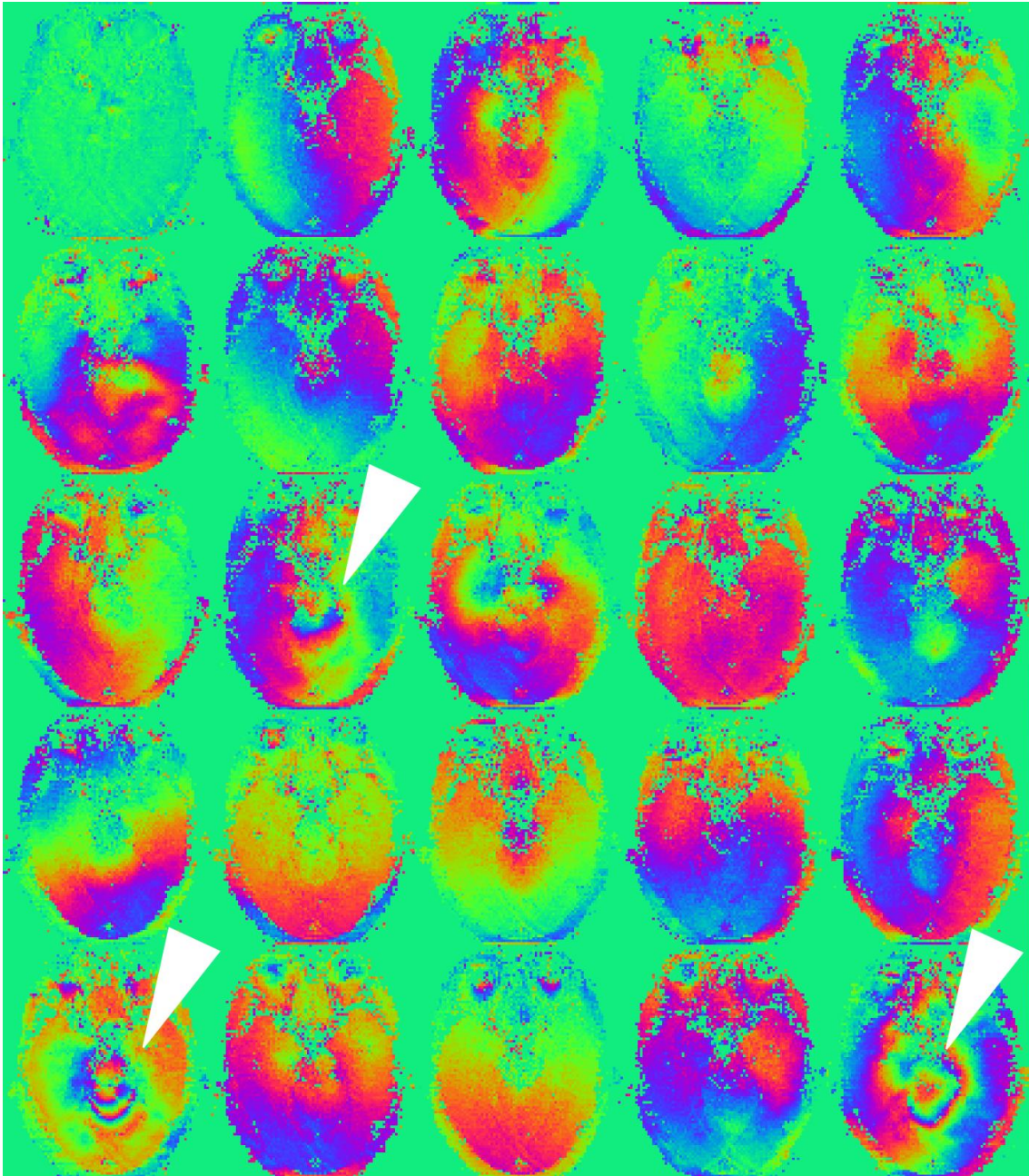
The coverage of k-space by segmented acquisitions with central reference lines is illustrated in Figure 6.3b. The example refers to the case of three segments and 8 reference lines. For the above mentioned image matrix three segments reduce the number of k-space lines (stimulated echoes) per segment from 84 to 33 or 34. While a simple division would lead to 28 lines that homogeneously cover the entire k-space, the use of 8 reference lines adds 5 or 6 more lines not counting those lines that are already included in the original 28 lines.

For the sequence with three segments and 8 reference lines the repetition time per segment was 15.3 s for 51 sections. The corresponding measuring time was 19 min for a DTI data set with 24 diffusion directions. Studies with four segments and 16 reference lines yielded 33 lines per segment, a repetition time of 15.0 s, and a measuring time of 25 min. The total reconstruction time on a computer with two quad-core central processing units was 7 min for the 3-segment data set with 8 channels and 30 min for the 4-segment data set with 32 channels.

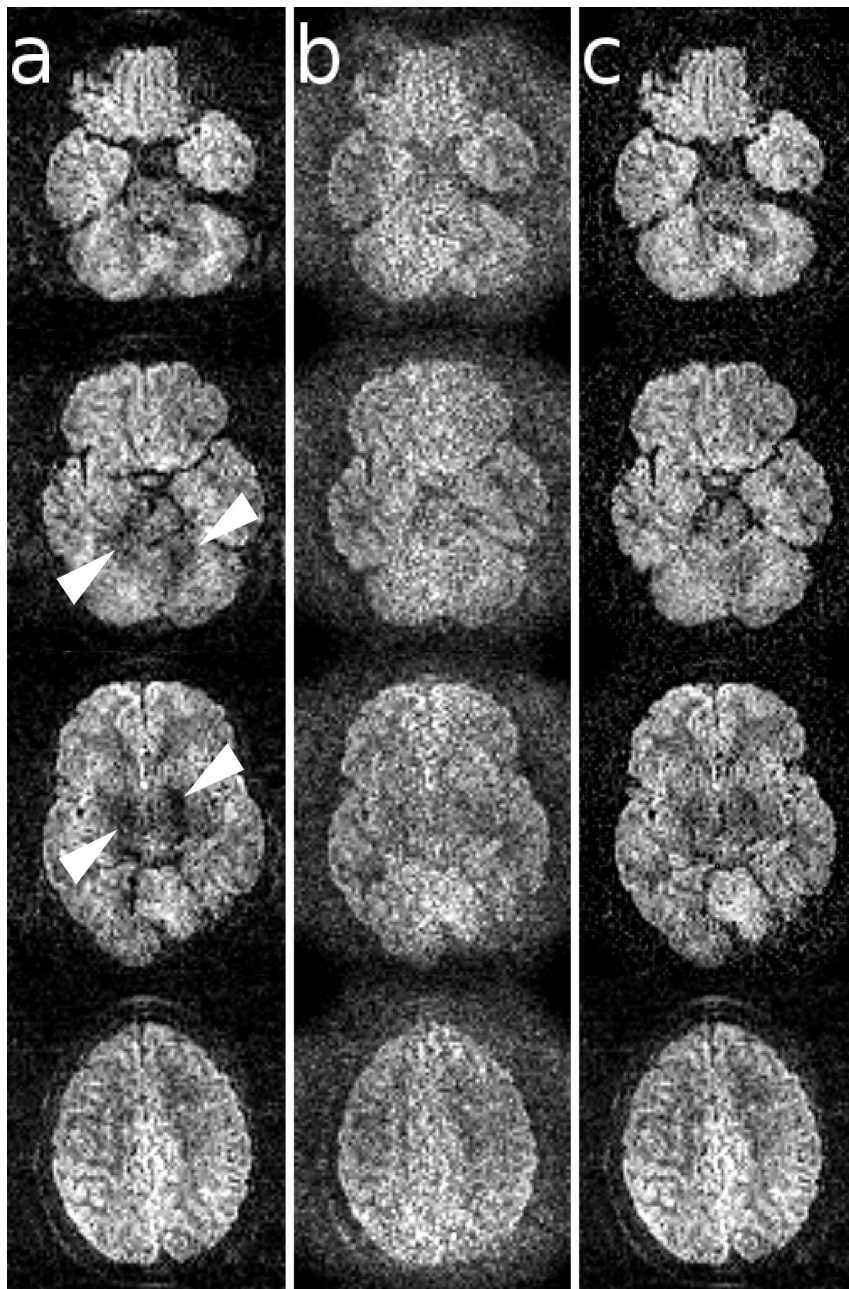
## 6.4 Results

Figure 6.4 shows motion-associated phase maps for a selected section and one segment of a 4-segment DW MRI data set of the human brain. These maps were estimated according to the second step of the proposed reconstruction algorithm taking into account the coil sensitivity maps obtained in the first step. The phase maps refer to the non-DW image (upper left corner) and 24 DW images. The color code facilitates the recognition of major phase changes. The absence of any visible phase variations in the non-DW image confirms the assumption of a constant phase in the first step, that is for the determination of the coil sensitivities from the non-DW acquisitions by regularized nonlinear inversion. While most DW images lead to maps with only moderate phase variations, some images are affected by phase

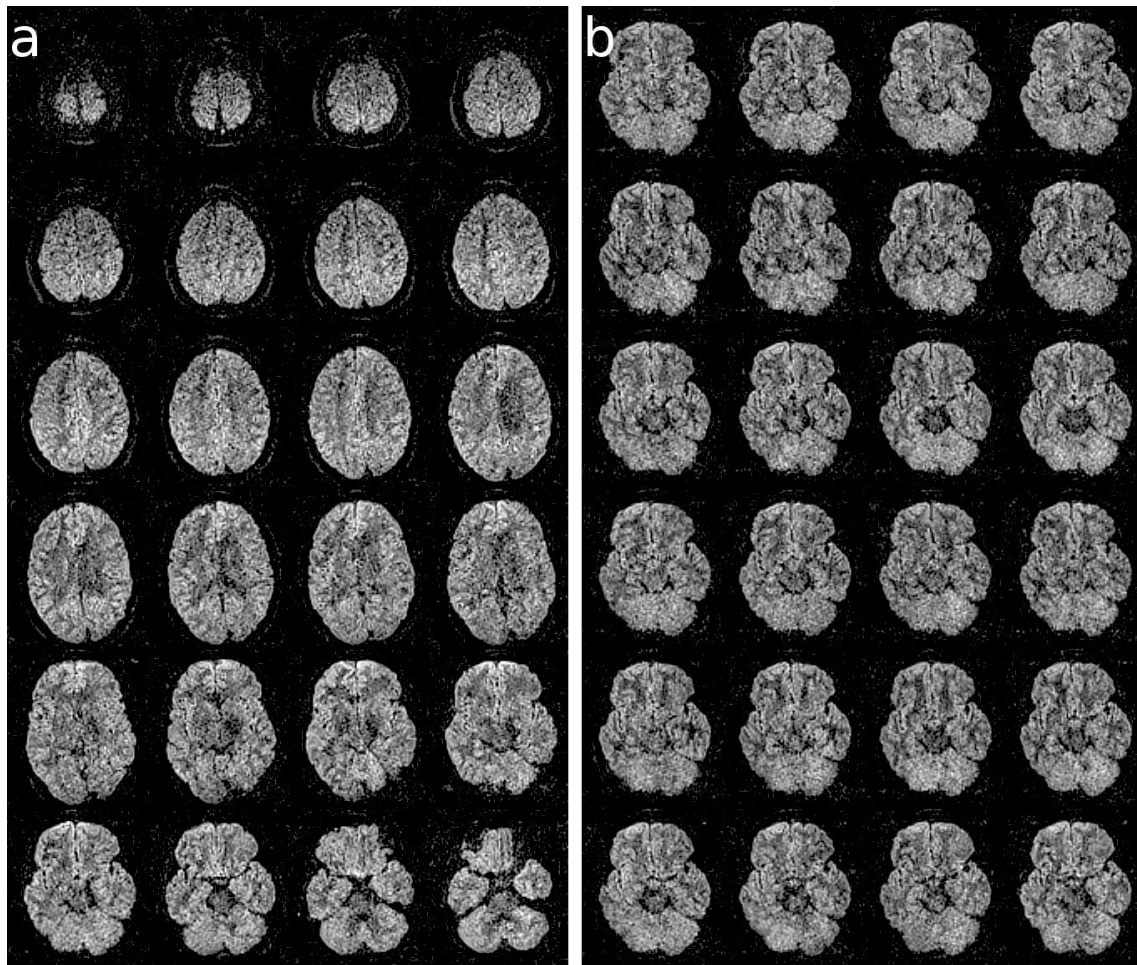




**Figure 6.4:** Motion-associated phase maps from a 4-segment DW MRI data set of the human brain (single section). The images were obtained during application of the second step of the proposed 3-step method. They represent a non-DW image (upper left corner) and all DW images ( $b = 1000 \text{ s mm}^{-2}$ , 24 directions) for one selected segment. The color encodes phase values between  $0$  and  $2\pi$ . Some images are affected by marked phase changes with high spatial frequencies (arrows).

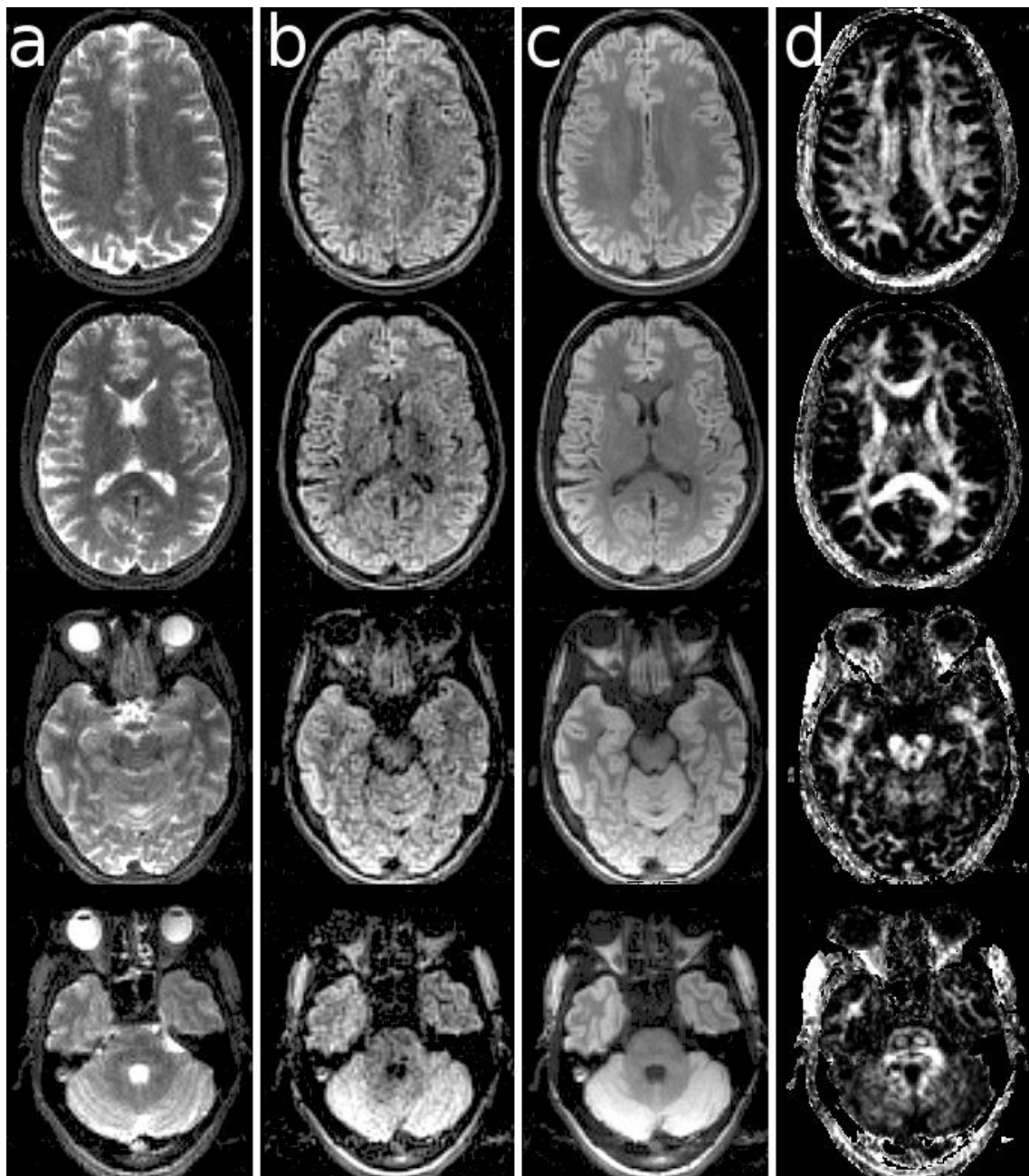


**Figure 6.5:** *DW images ( $b = 1000 \text{ s mm}^{-2}$ , four sections) reconstructed from a 3-segment DW MRI data set of the human brain with the use of (a) a direct application of the nonlinear inversion algorithm, (b) a 2-step approach averaging the magnitude images reconstructed for individual segments, and (c) the proposed 3-step method. The motion-induced signal void obtained for the first approach (arrows) is avoided by the more complex strategies, while the 3-step method further reduces the sensitivity of the 2-step approach to noise.*



**Figure 6.6:** (a) Selected 24 of 51 multislice DW images for a single diffusion direction and (b) all 24 DW images of a single section ( $b = 1000 \text{ s mm}^{-2}$ ) reconstructed from a 3-segment DW MRI data set of the human brain with the proposed 3-step method.





**Figure 6.7:** (a) Non-DW images, (b) DW images for a single diffusion direction ( $b = 1000 \text{ s mm}^{-2}$ ), (c) isotropically DW images, and (d) maps of the fractional anisotropy reconstructed from a 4-segment DW MRI data set of the human brain with the proposed 3-step method (four sections).

changes with high spatial frequencies (arrows). In particular, this applies to the vicinity of the brain stem where brain pulsations are most pronounced.

Figure 6.5 compares the performance of three different reconstruction techniques for four selected sections of a 3-segment acquisition. The results shown in Figure 6.5a represent a direct application of the nonlinear inversion algorithm. It yields the final DW image in only one reconstruction step, but merges the information from the coil sensitivities and phase variations into a single map. In this case, the images still suffer from residual motion artifacts (some signal void, arrows). Such problems are avoided by the 2- and 3-step methods shown in Figures 6.5b and 6.5c, respectively. However, because the 2-step method averages the noise-affected magnitude images from the individual segments, the final DW images (Fig. 6.5b) exhibit a lower SNR than the images obtained by the proposed 3-step method (Fig. 6.5c) in spite of an identical regularization.

The general performance of the new method is demonstrated in Figures 6.6 and 6.7 for a 3- and 4-segment acquisition, respectively. Figure 6.6 summarizes 24 of 51 DW images of a multislice data set for a single diffusion direction (Fig. 6.6a) as well as all 24 DW images of a single section (Fig. 6.6b). Finally, using the proposed algorithm, Figure 6.7 shows non-DW images, DW images for one diffusion direction, isotropically DW images, and maps of the fractional anisotropy for four selected sections.

## 6.5 Discussion

It is common knowledge that segmented multi-shot DW MRI is affected by motion-associated phase differences that preclude the reconstruction of artifact-free images by a straightforward combination of respective k-space segments. Here it is demonstrated that the typical phase variations in DW MRI of the human brain that are due to cardiac-induced brain pulsations and residual subject movements cannot be represented by low-resolution k-space data as usually obtained from (external) navigator acquisitions or (internal) autocalibration or reference lines. Instead, a substantial fraction of DW images presents with phase variations that have to be characterized at full spatial resolution using the entire available k-space data.

In a first attempt, the direct application of a regularized nonlinear inversion algorithm turned out to be suboptimal as it forces smooth (low-resolution) coil sensitivities and irregular (high-resolution) motion-associated phase maps into a single

“reference” map that complements the image of the true object. An improvement was achieved by a 2-step procedure which first reconstructs the non-DW images by nonlinear inversion and then exploits the respective coil sensitivities to obtain DW images for all segments with the use of an iterative SENSE-like algorithm. However, when simply averaging the magnitude images of the individual segments to calculate the desired DW image [97], then the enhanced noise sensitivity of the ill-conditioned reconstruction problem for the undersampled segments affects the SNR of the final image. The least motion sensitivity and best SNR was obtained by adding another reconstruction by iterative linear inversion that exploits both the coil sensitivities from the first step and the motion-associated phase maps that may be extracted from the reconstructions of the second step. Because this third step uses the complementary k-space data from all segments, it is much better conditioned than the reconstructions from individual segments in the second step. The corresponding absence of any detectable noise amplification during the final reconstruction leads to a better SNR than obtainable by the 2-step method.

In comparison with other reconstruction techniques for segmented motion-affected data sets, the main advantage of the proposed 3-step method is the consideration of phase disturbances with high spatial frequencies. With respect to the work by Skare et al[97] the approach further benefits from the high-quality coil sensitivities obtained by nonlinear inversion in the first step, the more flexible conjugate gradient-based version of the SENSE algorithm in the second step (see below), and the avoidance of any noise amplification in the third true reconstruction step.

The current implementation employed a Cartesian encoding scheme where in each segment of both the non-DW and the DW acquisitions the center of k-space is fully sampled by 8 or 16 lines. It should be noted, however, that the reference lines for the DW images are not necessarily required for the proposed 3-step algorithm. Their use was originally motivated to allow for a fair comparison with the direct use of the nonlinear inversion technique, while later trials with reference lines and iterative SENSE-like reconstructions in the second and third step helped to improve the SNR. As far as non-DW images are concerned, the nonlinear inversion technique is much less sensitive to a low number of central reference lines than conventional approaches (see Chapter 5).

In addition, also other specific aspects of the actual work such as the use of a full Fourier acquisition or even the choice of a Cartesian encoding scheme pose no general restrictions for the 3-step method. Preliminary trials of both partial Fourier schemes and radial encoding schemes for the same segmented DW STEAM MRI

sequence proved to be successful.

In order to be applicable the proposed method has to meet only two experimental conditions: (i) the non-DW images must be reconstructable by nonlinear inversion, and (ii) each undersampled segment must contain enough data to allow for a reasonable reconstruction by parallel imaging, for example iterative SENSE. The first condition simply refers to the fact that the joint k-space of the non-DW images from all segments constitutes a fully sampled center. It is needed to properly estimate the coil sensitivities by regularized nonlinear inversion. The second condition requires the k-space lines in each segment of the DW images to be sampled in a sufficiently interleaved and dense manner to enable adequate reconstructions by parallel imaging. Apart from these requirements, however, the algorithm imposes no further constraints on the trajectory.

An interesting question arises for the dependence of the final image on the accuracy or noise of the estimated phase maps. For the experimental parameters chosen in this work, no difficulty was observed. In general, however, this may become a concern for very high reduction factors, that is a large number of segments and a correspondingly ill-conditioned linear system. A possible solution may be to constrain the second reconstruction step by an even stronger regularization. Of note, the choice of regularization is not critical in the first and third step. Another possible source of reconstruction error may be due to alterations of the coil sensitivities that could occur due to severe macroscopic motions during the acquisition process. If this does not lead to a total corruption of the data set but emerges as a tractable problem, then a possible remedy may be obtained by using a regularized nonlinear inversion algorithm also in the second step. This would allow for a re-estimation of the coil sensitivities for each individual segment.

Finally, a most elegant algorithm to tackle the reconstruction problem in segmented DW MRI would be a regularized nonlinear inversion method which simultaneously treats coil sensitivities, motion-associated phase maps, and an object image as three independent unknowns. Such a method should achieve the same quality and robustness as the 3-step algorithm proposed here, but not suffer from the need to first reconstruct images from individual segments. Unfortunately, preliminary trials required a good initial guess for the phase maps which so far renders the approach useless. Nevertheless, foreseeable improvements of the algorithm are likely to alter the situation in future.

## 6.6 Summary

This work presents a new inverse reconstruction method for segmented multi-shot DW MRI which is based on the concepts of parallel imaging. Experimental applications deal with DW STEAM MRI of the human brain using 3 or 4 segments. The algorithm first determines separate coil sensitivity and motion-associated phase maps for each segment by taking advantage of the entire k-space data from multiple receive coils. Subsequently, these maps are used to reconstruct object images without motion artifacts – again by parallel imaging based on the data from all segments. In contrast to existing approaches, the proposed method provides robust solutions without compromised SNR even in cases where the phase variations are characterized by high spatial frequencies.



# 7

## Image Regularization

As evident from the PCA of the multi-channel MRI data, most of the energy is concentrated in very few principal components (see Chapter 3.2.2). To reach high acceleration factor the information contained in the principal components of low energy needs to be exploited. Hence, even with a large number of channels and perfectly known coil sensitivities, parallel imaging is limited by the high noise amplification due to the bad conditioning of the system. In parallel imaging, regularization allows the noise to be reduced, but only at the cost of residual undersampling artifacts. While Tikhonov regularization is commonly used, more sophisticated regularization techniques might be able to suppress the noise much better, mitigating this situation.

In improved regularization, additional information about the mathematical structure of the image complements the weakly encoded information in the data. As it turns out, this additional information even allows for a complete replacement of the missing information under certain conditions. This is the topic of the new research area called *compressed sensing* [21, 27]. After the discussion of the conventional quadratic regularization, this chapter explores a novel idea how to extend algorithms based on the IRGNM to include  $L^1$  and TV regularization terms.

### 7.1 Tikhonov Regularization

With the use of a regularization term preference for solution with certain properties can be expressed. The most commonly used technique is quadratic or *Tikhonov regularization*

$$\|A\mathbf{x} - \mathbf{y}\|_2^2 + \alpha\|\Gamma\mathbf{x}\|_2^2 \ .$$

The simplest choice for  $\Gamma$  is the unit matrix, which corresponds to a penalty in the  $L^2$  norm. The solution is then given by

$$\mathbf{x}_\alpha = A_\alpha^\diamond \mathbf{y} \quad A_\alpha^\diamond = (A^H A + \alpha^2 \Gamma^H \Gamma)^{-1} A^H .$$

To understand the effect of the regularization, it is helpful to consider the singular value decomposition (SVD) of the system matrix

$$A = U \Sigma V^H \quad \text{with} \quad \Sigma = \text{diag}(\sigma_1, \dots, \sigma_n) .$$

The matrix is decomposed into two unitary matrices  $U$  and  $V$  and a (in general non-quadratic) diagonal matrix consisting of the singular values  $\sigma_1, \dots, \sigma_n$  which will be assumed to be ordered in magnitude. The pseudo inverse is then given by

$$A^\diamond = V \Sigma^{-1} U^H .$$

Corresponding to small singular values there is a subspace in the domain, which is encoded very weakly by the data. In the calculation of the solution by application of the pseudo inverse this subspace is multiplied by reciprocals of the respective singular values causing a strong increase of noise. A quantitative description of this problem can be given by the condition number, which is defined as the relation between the largest and smallest eigenvalue:

$$\text{cond}(A) = \frac{\sigma_{\max}}{\sigma_{\min}}$$

Given a singular value decomposition of  $A = U \Sigma V^H$ , the solution can be expressed as

$$A_\alpha^\diamond = V \Sigma^\diamond U^H \quad \text{with} \quad \Sigma_i^\diamond = \frac{\sigma_i}{\sigma_i^2 + \alpha^2} .$$

For large singular values  $\sigma \gg \alpha$  the regularization term  $\alpha$  has a small effect, while for small singular values  $\sigma \ll \alpha$  the explosion of the inverse is avoided.

Instead of choosing  $\Gamma = I$ , another common choice is the use of a derivative operator  $\Gamma = D$ , which has a smoothing effect along the corresponding dimension. Although it is commonly implemented with finite differences in image space, it is a special case of a more general class, which can be expressed as diagonal weighting  $W$  in the Fourier domain:

$$\Gamma = W \mathcal{F}$$

Here, the regularization penalty is invariant to arbitrary shifts in the image domain, which is a natural assumption in digital signal processing and imaging.

## 7.2 Choice of the Regularization Parameter

Many techniques have been proposed, how to best choose the regularization parameter  $\alpha$ . The most well known are the discrepancy principle, L-curve methods and cross-validation methods. For example, the discrepancy principle chooses  $\alpha$  such that the norm of the residual is equal to a constant times the norm of the data error. In fact, all these methods try to optimize the regularization parameter in order to minimize the reconstruction error in the  $L^2$  norm. In the case of linear inverse problems, the expectation value of the error splits into two components

$$\mathbf{E} \left\{ \left\| (A^H A - \alpha^2 I)^{-1} A^H y - x \right\|_2^2 \right\} = \left\| \left( (A^H A - \alpha^2 I)^{-1} A^H A - I \right) x \right\|_2^2 + \mathbf{E} \left\{ \left\| (A^H A - \alpha^2 I)^{-1} A^H n \right\|_2^2 \right\} .$$

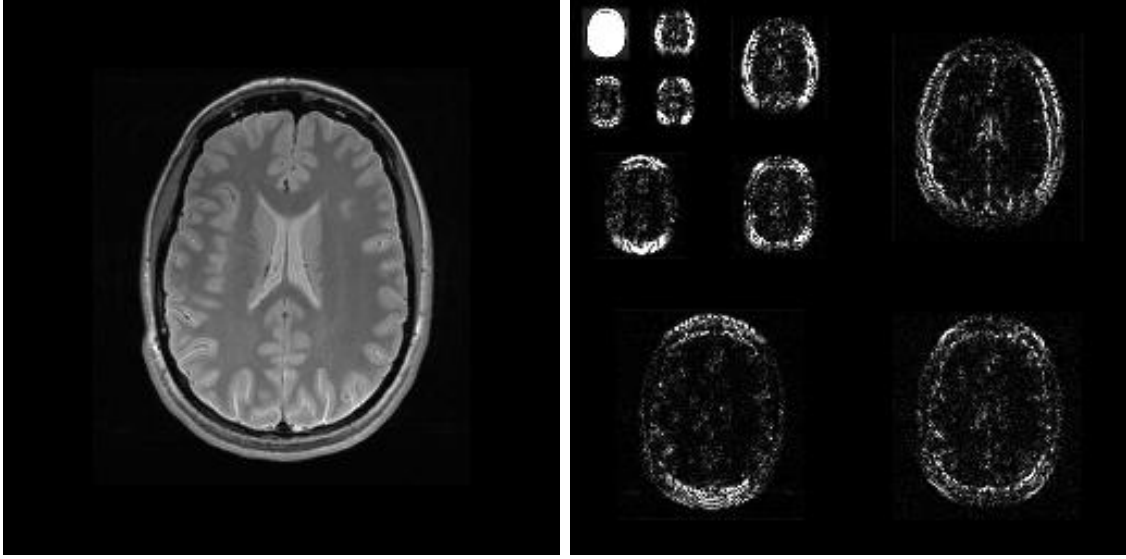
The first term is caused by the suppression of legitimate signal during the reconstruction which increases with stronger regularization, while the second error term quantifies the reconstructed noise which decreases with regularization. In the total error both terms are weighted equally.

Unfortunately, in the context of parallel imaging, minimizing this sum often does not produce the visually best reconstruction. Although this depends on the exact purpose, aliasing artifacts are usually considered worse than noise, because they might lead to an erroneous interpretation of the image.

A heuristic criterion which seems to be useful in practice compares some mean energy of k-space samples at actually measured positions to the k-space mean energy at predicted positions and chooses the value, where both quantities are equal. Compared to criteria which are based on the  $L^2$  norm of error, this criterion chooses a smaller regularization parameter, but still prevents that the noise dominates the solution. Finally, it should be noted, that with many irregular sampling patterns or non-Cartesian trajectories, the visual appearance of the undersampling artifacts is completely different and often looks much more like noise. This might justify the use of a conventional criterion, for example the discrepancy principle.

## 7.3 $L^1$ Regularization

$L^2$  regularization by its quadratic nature penalizes large values much stronger than small values. In case of sparse images, where a transformed image vector  $\Gamma \mathbf{x}$  has only very few non-vanishing entries, it makes sense to replace the  $L^2$  regularization



**Figure 7.1:** *MR image (left) and corresponding multi-scale wavelet transformation (right). For many images, a wavelet transformation is able to compress most energy into only few wavelet coefficients, providing a sparse representation of the image.*

term with a regularization in the  $L^1$  norm.

$$\operatorname{argmin}_{\mathbf{x}} \{ \|A\mathbf{x} - \mathbf{y}\|_2^2 + \|\Gamma\mathbf{x}\|_1 \} \quad (7.1)$$

The few largest values carry most of the image information, while the noise is distributed equally over all pixels. In the  $L^1$  norm the small values are penalized much more, effectively suppressing the noise. This allows to find a sparse solution, which consists only of the few non-vanishing coefficients which represent the real object.

Because most images are not sparse, a sparsity transformation  $\Gamma$  must be applied for this technique to be effective. For example, after a wavelet transform a typical (medical) image has only few non-vanishing coefficients (see Fig. 7.1): a fact, which is also exploited in some noise filter and image compression algorithms (e.g., the JPEG2000 compression standard).

In the following, an algorithm for integration of  $L^1$  regularization into algorithms based on regularized nonlinear inversion with the IRGNM is developed. The algorithm is based on the simple idea that  $L^1$  and  $L^2$  norm of a function can be related by point-wise squaring, according to

$$\|f\|_2^2 = \int d\mathbf{x} |f(\mathbf{x})|^2 = \int d\mathbf{x} |f(\mathbf{x})^2| = \|f^2\|_1 \quad .$$

Point-wise squaring translates the quadratic regularization of a transformed system to the desired  $L^1$  regularization for the original problem. An exact generalized

formulation of this idea can be found in the next section. A first mathematical analysis can be found in [108].

### 7.3.1 Nonlinear Transformation into a Quadratic Penalty

Because in the following the nonlinear transformation is needed also for multi-variate functions  $f(\mathbf{x}) \in \mathbf{R}^N$ , it is formulated for this general case. A suitable transformation should establish the following correspondence

$$\int d\mathbf{x} \sum_{j=1}^N f_j(\mathbf{x})^2 = \int d\mathbf{x} \sqrt{\sum_{j=1}^N \hat{f}_j(\mathbf{x})^2} .$$

A point-wise transformation that maps the components  $f_j \mapsto \hat{f}_j$  correctly is given by

$$T_\epsilon : (x_1, \dots, x_N) \mapsto (\hat{x}_1, \dots, \hat{x}_N) = (x_1, \dots, x_N) \sqrt{\epsilon^2 + \sum_{j=1}^N x_j^2} .$$

Here, a small parameter  $\epsilon > 0$  was introduced which will make the inverse of the transformation differentiable at the origin. From rotational invariance in the components it is clear that the inverse of  $T$  must have the following form:

$$Q_\epsilon : (\hat{x}_1, \dots, \hat{x}_N) \mapsto \frac{(\hat{x}_1, \dots, \hat{x}_N)}{\sqrt[4]{\hat{\epsilon}^2 + \sum_{i=1}^N \hat{x}_i^2}}$$

For  $\epsilon = 0$ , calculation yields the identity  $T_0 \circ Q_0 = Id$  on  $R^N \setminus (0, \dots, 0)$  with  $\hat{\epsilon} = 0$ . To determine the function  $\hat{\epsilon}$  in the general case, the transformation is expressed as part of a transform with dimension  $N + 1$  and  $\epsilon = 0$ :

$$Q_0(\hat{\epsilon}, \hat{x}_1, \dots, \hat{x}_N) = \left( \frac{\hat{\epsilon}}{\sqrt[4]{\hat{\epsilon}^2 + \sum_{j=1}^N \hat{x}_j^2}} \right) \oplus Q_\epsilon(\hat{x}_1, \dots, \hat{x}_N)$$

With  $Q_0 \circ T_0 = I$  the first component yields a quadratic equation for  $\hat{\epsilon}^2$  in terms of  $\epsilon$  and  $\hat{x}_1, \dots, \hat{x}_N$ :

$$\epsilon = \frac{\hat{\epsilon}}{\sqrt[4]{\hat{\epsilon}^2 + \sum_{j=1}^N \hat{x}_j^2}} \quad \Rightarrow \quad \hat{\epsilon} = \sqrt{\frac{\epsilon^4 + \epsilon^2 \sqrt{\epsilon^2 + 4 \sum_{j=1}^N \hat{x}_j^2}}{2}}$$

### 7.3.2 Total Variation

The TV semi-norm of a continuously differentiable function  $f$  is given by

$$TV(f) = \int \|\nabla f\|_2 \ .$$

A TV penalty suppresses noise very well and can be used to obtain a denoised image from an original  $f_0$  by the minimization of the functional [92]

$$\operatorname{argmin}_f \{ \|f - f_0\|_2^2 + \alpha TV(f) \} \ .$$

In one dimension, the total variation norm is just the  $L^1$  norm of the derivative, which motivates the following algorithm. Following the idea presented above, a nonlinear inverse problem  $F(x) = y$  with TV regularization should be transformed into one with quadratic regularization. For simplicity,  $x$  will be assumed to be a real or complex function defined on  $\mathbb{R}^2$  (e.g., an image). As a first step to get an explicit representation of the needed partial derivatives of  $x$  additional equations have to be added:

$$\underbrace{\begin{pmatrix} F & 0 & 0 \\ \partial_1 & -I & 0 \\ \partial_2 & 0 & -I \end{pmatrix}}_G \begin{pmatrix} x \\ \partial_1 x \\ \partial_2 x \end{pmatrix} = \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix}$$

The components representing the partial derivatives are now transformed with  $Q_\epsilon$  which yields new variables

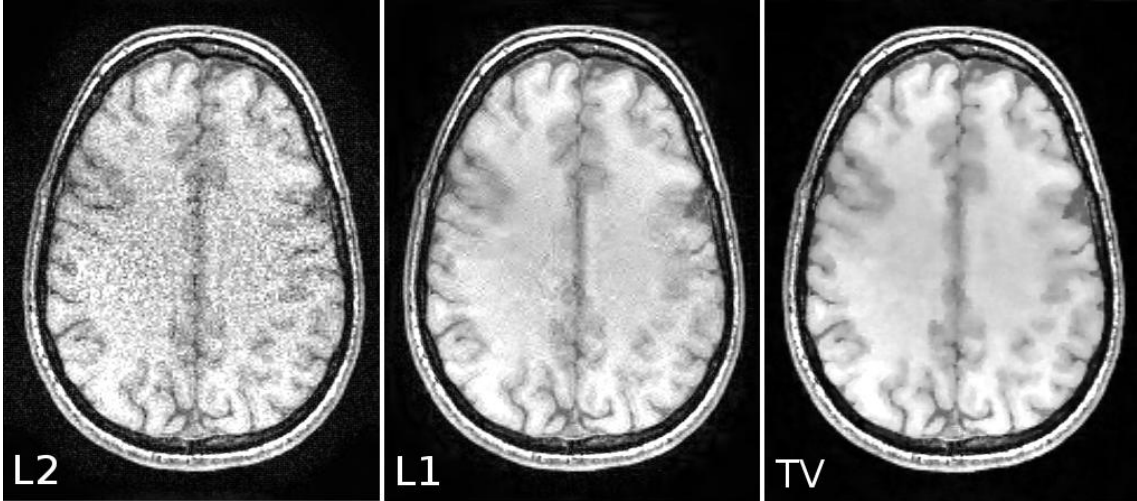
$$\begin{aligned} z_0 &= x \\ z_1 &= (Q_\epsilon(\partial_1 x, \partial_2 x))_1 \\ z_2 &= (Q_\epsilon(\partial_1 x, \partial_2 x))_2 \ . \end{aligned}$$

The final transformed problem is then given by

$$\hat{G}z = \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} \quad \text{with} \quad \hat{G}z = G \begin{pmatrix} x \\ (T_\epsilon(z_1, z_2))_1 \\ (T_\epsilon(z_1, z_2))_2 \end{pmatrix} \ .$$

This nonlinear system of equations can be solved with the IRGNM (see Chapter 5). In each iteration, the updated estimate for the solution is the minimizer of a linear least-squares problem

$$\|D\hat{G}^H \Delta z_n - (y - \hat{G}(z_{n-1}))\|_2^2 + \alpha_n \|z_n\|_2^2 \ .$$



**Figure 7.2:** *Parallel imaging reconstruction (3D FLASH, 2D acceleration factor  $R = 3 \times 2$ , 12 channels and  $24 \times 24$  reference lines) with different regularization techniques: Conventional  $L^2$  regularization,  $L^1$  regularization in a wavelet basis, and TV regularization.*

Reverting the transformation for the regularization term reveals the TV semi-norm

$$\|x\|_2^2 + \int z_1^2 + z_2^2 = \|x\|_2^2 + \int \sqrt{(\partial_1 x)^2 + (\partial_2 x)^2 + \epsilon^2} \ .$$

The constant  $\epsilon > 0$  must be included to be able to initialize the variables  $z_{1,2}$  to zero, which would otherwise be “stuck” at zero due to the vanishing derivative. To reduce its influence, it might be decreased in each iteration step. The influence of the  $L^2$  term  $\|x\|_2^2$  can be reduced by including additional weights in the transformed system, which have been left out for simplicity.

## 7.4 Examples

As proof-of-principal, two examples are presented in this chapter. They have been computed with the IRGNM applied to a suitable transformed system. The regularization parameter was decreased only by a small factor 1.1 in each iteration step. Unfortunately, this leads to a long reconstruction time of several minutes, but is necessary to prevent instability of the reconstruction process in the current version of the algorithm.

### 7.4.1 Parallel Imaging

Figure 7.2 shows a single section of a full 3D MRI brain scan, accelerated with parallel imaging by a factor of 6. The image was reconstructed with the nonlinear regularized inversion algorithm (Chapter 5) which was extended with the new regularization scheme. As wavelet transform biorthogonal Cohen-Daubechies-Feauveau (CDF) 9/7 wavelets [23] have been used. The  $L^2$  regularized image is affected by a high amount of noise, while the  $L^1$  wavelet and TV regularized images are able to suppress most of the noise.

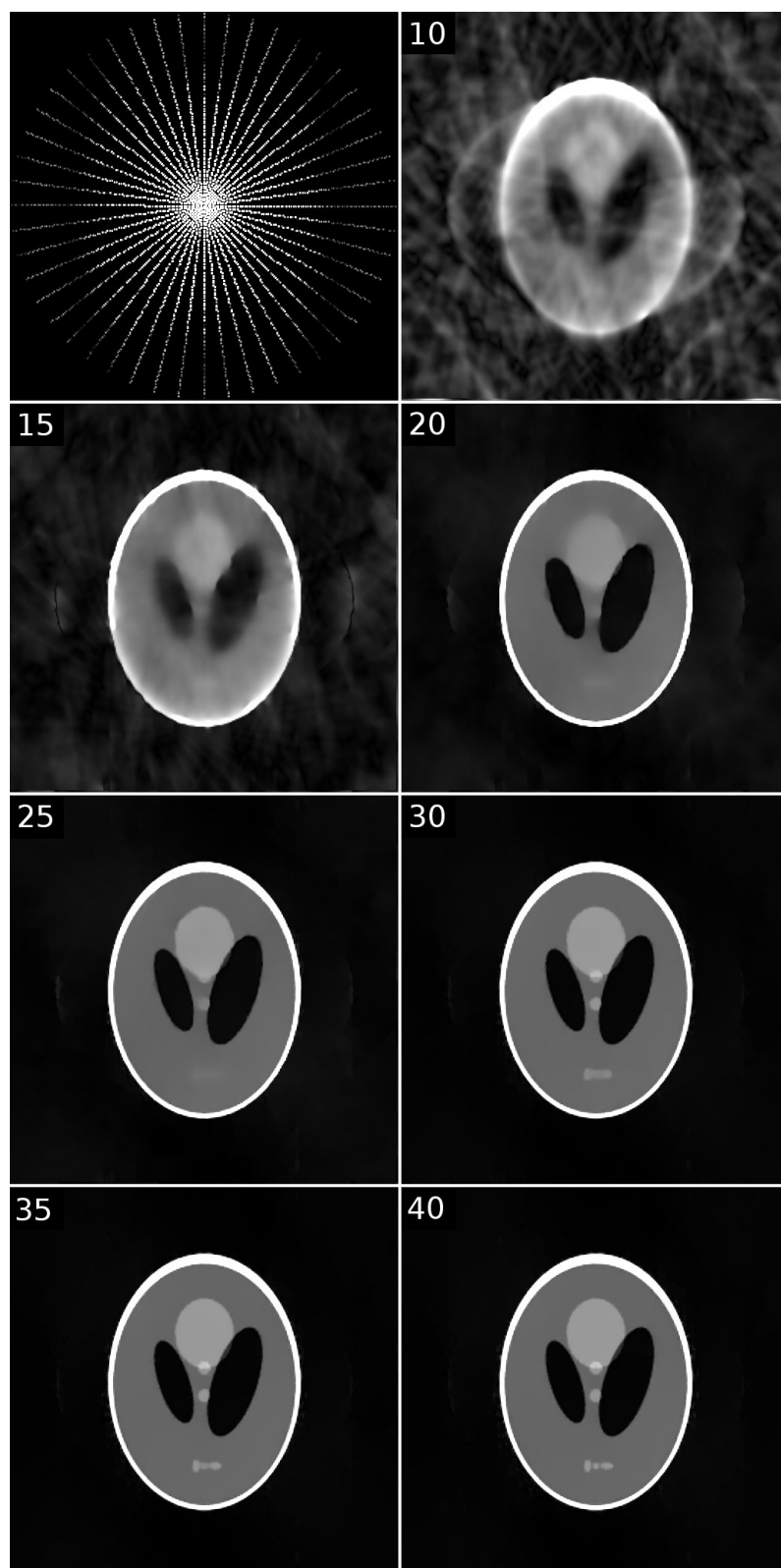
### 7.4.2 Compressed Sensing

While the use of the TV semi-norm for denoising is well known [92], it is also related to a recent new area of research called compressed sensing. The first application of the compressed sensing idea to MRI was given in [111]. A combination with parallel imaging for radially sampled data was described in [13]. A general description of compressed sensing for MRI can be found in [75]. Under suitable conditions an object with a sparse representation can be recovered exactly from undersampled data, for example randomly sampled Fourier data. Although theoretical results currently restrict to simpler cases, one well-known example is the exact recovery of the Shepp-Logan head phantom [96] from its radially sampled Fourier data with the use of a TV constraint. Here, the algorithm is applied to a linear system  $Ax = y$ , where  $A$  is a Fourier transform followed by a projection onto the radially sampled data. Figure 7.3 shows the reconstructions of the Shepp-Logan phantom from 24 spokes with the use of the algorithm presented here.

## 7.5 Summary

The development of better regularization techniques is critical to exploit the full potential of parallel imaging. Without efficient noise suppression higher acceleration factors can not be used. After discussion of Tikhonov regularization, a new idea to use  $L^1$ -based regularization terms with the IRGNM was discussed. While the algorithm presented in this chapter needs further work to increase its speed and stability, the combination of the IRGNM with a nonlinear transform to achieve non-quadratic regularization shows promise as a general framework for image reconstruction. This was demonstrated with two examples from parallel imaging and compressed sensing.





**Figure 7.3:** *TV-constrained reconstruction of the Shepp-Logan phantom as a function of the number of iterations from a data set with 24 radial spokes.*

# 8

## Summary and Outlook

### 8.1 Summary

In this thesis, the reconstruction problem of parallel imaging was investigated for the case of MRI with multiple receive coils. Starting from conventional descriptions of parallel imaging as a linear inverse problem, GRAPPA and SENSE algorithms and their respective variants are discussed as the most commonly used algorithms in commercially available MRI systems. As a more recent development, (CG-SENSE) was evaluated as an iterative algorithm that is based on the conjugate gradient algorithm. Iterative algorithms provide the highest flexibility, because they can easily be extended to irregular sampling patterns and non-Cartesian trajectories. For the non-Cartesian case an efficient implementation based on a convolution could be developed as an alternative to the standard interpolation techniques. Although this implementation has already been described in the literature, it does not seem to be well known. In the course of these initial analyses the main causes of residual image artifacts with existing parallel imaging approaches could be identified as problems that arise from a miscalibration of the receive coil sensitivities and the bad conditioning of the linear problem.

The problem of inaccurate coil sensitivities is addressed in Chapter 5. Based on the formulation as a nonlinear problem a novel algorithm was developed that considerably improves the estimation of the coil sensitivities in autocalibrated parallel imaging. It employs IRGNM with appropriate regularization terms and uses the conjugate gradient algorithm to invert the Jacobian in each Newton step. The method allows for a reconstruction of artifact-free images for higher acceleration fac-

tors and with less reference lines than existing approaches. Its better performance was experimentally demonstrated for MRI of a water phantom and a human brain.

In addition to the basic algorithm, the method was extended to allow for partial Fourier imaging, imaging scenarios with a reduced field of view, and non-Cartesian trajectories. In the latter case, the extension is based on the efficient convolution implementation known from CG-SENSE and therefore lends itself to an easy implementation on a GPU.

The development of a new reconstruction technique for segmented DW MRI is outlined in Chapter 6. In general, both rapid acquisition techniques for diffusion-weighted imaging, STEAM and EPI, benefit from a shorter echo train length. Apart from other strategies, a most efficient shortening can be achieved with a segmented acquisition, which splits the measurement of all k-space data into a multiple measurements. The major problem in segmented DW MRI, however, is the occurrence of motion-induced phase variations in each data segment, which precludes a direct combination of respective k-space lines. The new reconstruction proposed here splits the problem into three steps. First, the non-DW images and the coil sensitivities are jointly estimated with the algorithm from Chapter 5. Subsequently, the data from each segment are separately reconstructed with an iterative linear inversion to yield high-resolution estimates of the phase variations in each segment. This information is used in a third step to combine the data from all segments. In comparison to existing approaches, the new method allows for reconstructions with pronounced motion robustness and without any compromise in SNR.

Finally, a general limitation of parallel imaging is addressed which is caused by the bad conditioning of the inverse problem. It requires a trade-off between the degree of residual undersampling artifacts (strong regularization) and increased noise (weak regularization). Better regularization techniques are expected to mitigate this situation. Chapter 7 presents a novel idea which explores the use of either a  $L^1$  regularization or TV regularization by a nonlinear transformation of the image. While these algorithms need further work to improve their speed and stability, they are able to achieve a substantial noise suppression even for high acceleration (data reduction) factors. These regularization techniques bridge the gap to the new field of compressed sensing, which was demonstrated by a simple numerical example.

## 8.2 Future Work

A major drawback of advanced iterative techniques for image reconstruction is the increased computation time. While this is the main reason why such methods have been disregarded in clinical settings, the situation may change in the near future. Firstly, the progressional development of computer technology together with the parallelization of the algorithm is expected to yield reasonable reconstruction times. For example, for segmented DW MRI the proposed three-step technique needs to iteratively calculate more than 1000 images for 3 to 4 segments and up to 32 receive coils. This can be accomplished with a parallelized implementation in only 7 to 30 min on a modern 8-core machine. Hence, implementation on a massively parallel GPU will further speed up the reconstruction time considerably.

Secondly, apart from parallelized implementations, the possibility of algorithmic improvements should be investigated, as can be demonstrated by the application of the array compression technique discussed in Chapter 3.2.2. For the example mentioned above, this technique allows for an acceleration by a factor of two yielding 3 to 15 min. And thirdly, preconditioning techniques are known to speed up the conjugate gradient algorithm for many applications. Preconditioners are often constructed from approximative inverse solutions which, for example, may be obtained from the known direct methods for parallel imaging. For the IRGNM, it is possible to build a preconditioner dynamically during the iteration [55], a technique which is attractive because it does not need domain-specific knowledge.

Because noise amplification is the limiting factor in parallel imaging, a better regularization seems to be the most important next step. The integration of a non-quadratic regularization into the regularized nonlinear inversion with the IRGNM would provide a more powerful mathematical framework for many reconstruction problems. Pertinent applications may range from parallel imaging to compressed sensing and model-based reconstruction techniques (see below). Further work is required to improve the speed and robustness of the ideas presented in Chapter 7.

Another remaining issue that needs to be addressed is the robustness of the reconstruction. Especially in dynamic applications, it does not seem possible to fully avoid randomly occurring non-stochastic errors in the data. Thus, when the reconstruction by parallel imaging relies on only few measured data, the effect of corrupted data on the result may be fatal. For this reason, such data must be reliably identified and removed from the reconstruction. A possible solution could result from the use of robust estimators as known from stochastics. For example, the incorporation of a  $L^1$  norm in the data fit term is known to lessen the impact

of outliers. Such strategies could be exploited in image reconstruction to ameliorate the impact of corrupted data.

Finally, model-based reconstruction techniques emerge as another interesting area which can be solved from the perspective of a nonlinear inverse problem. For example, by extending the MRI signal equation with the relaxation terms, the estimation of relaxation maps can be accelerated as shown in [15] for radial trajectories. A similar idea was described in [1] for DTI. Here, instead of first calculating images for all diffusion directions and then estimating the tensor for each voxel, the complete tensor maps are directly calculated from the data. Such approaches rely on accurate field maps, and for this reason would be a logical extension to the algorithm described in Chapter 6.

In summary, advanced reconstruction techniques are necessary to fully exploit the potential of parallel acquisition techniques in MRI. The results of this thesis demonstrate that the formulation of the reconstruction as a nonlinear inverse problem adequately addresses many of the fundamental challenges. Based on experimental demonstrations, the proposed methods not only outperform currently used algorithms, but should provide a good mathematical framework for many other reconstruction problems in MRI.

# Appendix

## Definition of the Fourier Transform

In the space  $L^2(\mathbb{R})$  of equivalence classes of complex square integrable functions on  $\mathbb{R}$ , the Fourier transformation can be defined on the dense subset of  $L^1$ -integrable functions with the integral

$$\mathcal{F}f(k) = \int_{-\infty}^{\infty} dx f(x) e^{-2\pi i k x} .$$

Since

$$\langle f, g \rangle_{L^2(\mathbb{R})} = \langle \mathcal{F}f, \mathcal{F}g \rangle_{L^2(\mathbb{R})} ,$$

this formula defines a linear and isometric map from the dense subset  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$  to  $L^2(\mathbb{R})$ . For this reason, it has a unique extension to a unitary operator on  $L^2(\mathbb{R})$ . On the subset of  $L^1$ -integrable functions, a formula for the inverse is given by

$$\mathcal{F}^{-1}f(x) = \int_{-\infty}^{\infty} dk \hat{f}(k) e^{2\pi i k x} .$$

Many practical problems need to be formulated in a more general setting than given by the  $L^2$ -space. The Fourier transform can be extended to the space  $\mathcal{S}'$  of tempered distributions, which is the dual space of the Schwartz space  $\mathcal{S}$  of rapidly decreasing functions. Since it can be shown that  $\mathcal{F}(\mathcal{S}) = \mathcal{S}$ , the Fourier transform can be extended to  $\mathcal{S}'$  by the formula

$$\langle \mathcal{F}f, \phi \rangle = \langle f, \mathcal{F}\phi \rangle \quad f \in \mathcal{S}', \phi \in \mathcal{S} .$$

Here,  $\langle \cdot, \cdot \rangle$  denotes the canonical bilinear dual pairing of  $\mathcal{S}'$  and  $\mathcal{S}$ . The Fourier transform can be generalized to higher dimensions, starting from the integral

$$\mathcal{F}^d f(\mathbf{k}) = \int_{\mathbb{R}^d} d\mathbf{x} f(\mathbf{x}) e^{-2\pi i \mathbf{k} \cdot \mathbf{x}} .$$

$$\mathcal{F} \left\{ e^{-(\cdot)^2} \right\} (k) = \sqrt{\pi} e^{-\pi^2 k^2} \quad (8.2)$$

$$\mathcal{F} \text{rect}_{[-1/2, 1/2]}(k) = \frac{\sin(\pi k)}{\pi k} \quad (8.3)$$

$$\mathcal{F} \left\{ e^{2\pi i a(\cdot)} \right\} (k) = \delta(k - a) \quad (8.4)$$

$$\mathcal{F} \text{comb}_a(k) = \frac{1}{a} \text{comb}_{1/a}(k) \quad (8.5)$$

**Figure 8.1:** The Fourier transform of some important functions and distributions.

## Properties of the Fourier Transform

Some of its properties will be discussed shortly: According to the *convolution theorem*, the Fourier transform translates convolution to point-wise multiplication:

$$\mathcal{F}(f \star g) = (\mathcal{F}f) \cdot (\mathcal{F}g) \quad (8.1)$$

A shift is translated into a (linear) phase factor of the transformed function:

$$\mathcal{F} \{ f(\cdot - \Delta x) \} = e^{-2\pi i \Delta x k} \mathcal{F}f$$

Another important property relates to the behavior under scaling with a constant  $\lambda$ :

$$\mathcal{F} \{ f(\lambda \cdot) \} (k) = \frac{1}{|\lambda|} \mathcal{F}f\left(\frac{k}{\lambda}\right)$$

Building on the generalization to distributions, many processes arising in the digitalization of continuous problems may be described. An important example is sampling on a regular grid (see Chapter 2.5.1), which can be formulated as multiplication by a comb function, which is defined as sum of delta distributions shifted to equidistant positions)

$$\text{comb}_a = \sum_{l=-\infty}^{\infty} \delta_{la} \ .$$

By the Poisson summation formula

$$\sum_{n=-\infty}^{\infty} f(t + nT) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \hat{f}\left(\frac{k}{T}\right) e^{2\pi i \frac{k}{T} t} \ ,$$

its Fourier transform is again a comb function. A list of important Fourier transforms of various functions and distributions can be found in Figure 8.1.

# Abbreviations

**CG-SENSE** conjugate gradient SENSE

**CP** circularly polarized

**CT** computed tomography

**DFT** discrete Fourier transform

**DTI** diffusion tensor imaging

**DW** diffusion-weighted

**EPI** echo planar imaging

**FA** fractional anisotropy

**fMRI** functional MRI

**GCC** GNU compiler collection

**GPU** graphical processing unit

**GRAPPA** generalized autocalibrating partially parallel acquisitions

**IRGNM** iteratively regularized Gauss-Newton method

**JSENSE** joint SENSE

**FFT** fast Fourier transform

**FID** free induction decay

**FLASH** fast low angle shot

**FOV** field of view

**FT** Fourier transform



- MR** magnetic resonance
- MRI** magnetic resonance imaging
- NMR** nuclear magnetic resonance
- PCA** principal components analysis
- PF** partial Fourier
- PSF** point spread function
- RF** radio frequency
- RSS** root of sum of squares
- SENSE** sensitivity encoding
- SMASH** simultaneous acquisition of spherical harmonics
- SNR** signal-to-noise ratio
- STEAM** stimulated echo acquisition mode
- SVD** singular value decomposition
- $T_1$  spin-lattice relaxation time
- $T_2$  spin-spin relaxation time
- $T_2^*$  effective spin-spin relaxation time
- TE** echo time
- TR** repetition time
- TV** total variation

# References

- [1] AKSOY, M., LIU, C., MOSELEY, M., AND BAMMER, R. Single-step nonlinear diffusion tensor estimation in the presence of microscopic and macroscopic motion. *Magnetic Resonance in Medicine* 59, 5 (2008), 1138–1150.
- [2] ARUNACHALAM, A., SAMSONOV, A., AND BLOCK, W. Self-calibrated GRAPPA method for 2D and 3D radial data. *Magnetic Resonance in Medicine* 57, 5 (2007), 931–938.
- [3] ATKINSON, D., COUNSELL, S., HAJNAL, J., BATCHELOR, P., HILL, D., AND LARKMAN, D. Nonlinear phase correction of navigated multi-coil diffusion images. *Magnetic Resonance in Medicine* 56, 5 (2006), 1135–1139.
- [4] BAKUSHINSKY, A. The problem of convergence of the iteratively regularized Gauss-Newton method. *Computational Mathematics and Mathematical Physics* 32 (1992), 1354–1359.
- [5] BAKUSHINSKY, A., AND KOKURIN, M. *Iterative methods for approximate solution of inverse problems*. Kluwer Academic Publishers, 2004.
- [6] BAUER, F., AND HOHAGE, T. A Lepskij-type stopping rule for regularized Newton methods. *Inverse Problems* 21, 6 (2005), 1975–1991.
- [7] BAUER, F., AND KANNENGIESSER, S. An alternative approach to the image reconstruction for parallel data acquisition in MRI. *Mathematical Methods in the Applied Sciences* 30, 12 (2007), 1437–1451.
- [8] BEATTY, P., AND BRAU, A. Understanding the GRAPPA paradox. In *Proceedings of the ISMRM 14th Annual Meeting, Seattle, USA* (2006), p. 2467.
- [9] BLAIMER, M., BREUER, F., MUELLER, M., SEIBERLICH, N., EBEL, D., HEIDEMANN, R., GRISWOLD, M., AND JAKOB, P. 2D-GRAPPA-operator for faster 3D parallel MRI. *Magnetic Resonance in Medicine* 56, 6 (2006), 1359–1364.

- [10] BLOCH, F., HANSEN, W., AND PACKARD, M. Nuclear induction. *Physical Review* 70, 7-8 (1946), 460–474.
- [11] BLOCK, K. *Advanced Methods for Radial Data Sampling in Magnetic Resonance Imaging*. Doctoral Thesis, Universität Göttingen, 2008.
- [12] BLOCK, K., AND FRAHM, J. Spiral imaging: a critical appraisal. *Journal of Magnetic Resonance Imaging* 21, 6 (2005), 657–668.
- [13] BLOCK, K., UECKER, M., AND FRAHM, J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine* 57, 6 (2007), 1086–1098.
- [14] BLOCK, K., UECKER, M., AND FRAHM, J. Suppression of MRI Truncation Artifacts Using Total Variation Constrained Data Extrapolation. *International Journal of Biomedical Imaging 2008* (2008), 184123.
- [15] BLOCK, K., UECKER, M., AND FRAHM, J. Model-Based Iterative Reconstruction for Radial Fast Spin-Echo MRI. *IEEE Transactions on Medical Imaging* (accepted 2009).
- [16] BLUESTEIN, L. A linear filtering approach to the computation of discrete Fourier transform. *IEEE Transactions on Audio and Electroacoustics* 18, 4 (1970), 451–455.
- [17] BUEHRER, M., PRUESSMANN, K., BOESIGER, P., AND KOZERKE, S. Array compression for MRI with large coil arrays. *Magnetic Resonance in Medicine* 57, 6 (2007), 1131–1139.
- [18] BYDDER, M., LARKMAN, D., AND HAJNAL, J. Combination of signals from array coils using image-based estimation of coil sensitivity profiles. *Magnetic Resonance in Medicine* 47, 3 (2002), 539–548.
- [19] BYDDER, M., LARKMAN, D., AND HAJNAL, J. Generalized SMASH imaging. *Magnetic Resonance in Medicine* 47, 1 (2002), 160–170.
- [20] BYDDER, M., AND ROBSON, M. Partial Fourier partially parallel imaging. *Magnetic Resonance in Medicine* 53, 6 (2005), 1393–1401.
- [21] CANDÉS, E., ROMBERG, J., AND TAO, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52, 2 (2006), 489–509.

- [22] CARLSON, J., AND MINEMURA, T. Imaging time reduction through multiple receiver coil data acquisition and image reconstruction. *Magnetic Resonance in Medicine* 29, 5 (1993), 681–687.
- [23] COHEN, A., DAUBECHIES, I., AND FEAUVEAU, J. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* 45, 5 (1992), 485–560.
- [24] CONSTANTINIDES, C., ATALAR, E., AND MCVEIGH, E. Signal-to-noise measurements in magnitude images from NMR phased arrays. *Magnetic Resonance in Medicine* 38, 5 (1997), 852–857.
- [25] COOLEY, J., AND TUKEY, J. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation* (1965), 297–301.
- [26] DAMADIAN, R. Tumor detection by nuclear magnetic resonance. *Science* 171, 3976 (1971), 1151–1153.
- [27] DONOHO, D. Compressed sensing. *IEEE Transactions on Information Theory* 52, 4 (2006), 1289–1306.
- [28] DONOHO, D. For most large underdetermined systems of linear equations the minimal  $l_1$ -norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics* 59, 6 (2006), 797–829.
- [29] EDELMAN, R., WIELOPOLSKI, P., AND SCHMITT, F. Echo-planar MR imaging. *Radiology* 192, 3 (1994), 600.
- [30] EGGERS, H., BOERNERT, P., AND BOESIGER, P. Comparison of gridding- and convolution-based iterative reconstruction algorithms for sensitivity-encoded non-Cartesian acquisitions. In *Proceedings of the ISMRM 10th Annual Meeting, Honolulu, USA* (2002), p. 743.
- [31] ENGL, H., HANKE, M., AND NEUBAUER, A. *Regularization of inverse problems*. Kluwer Academic Publishers, 1996.
- [32] FINSTERBUSCH, J., FRAHM, J., AND KOCH, M. Gradient and stimulated echo (GRASTE) imaging. *Magnetic Resonance in Medicine* 55, 2 (2006), 455–459.
- [33] FRAHM, J., HAASE, A., MATTHAEI, D., MERBOLDT, K., AND HÄNICKE, W. Rapid NMR imaging using stimulated echoes. *Journal of Magnetic Resonance* 65, 1 (1985), 130–135.

- [34] FRAHM, J., MERBOLDT, K., AND HÄNICKE, W. Localized proton spectroscopy using stimulated echoes. *Journal of Magnetic Resonance* 72, 3 (1987), 502–508.
- [35] FRAHM, J., MERBOLDT, K., HÄNICKE, W., AND HAASE, A. Stimulated echo imaging. *Journal of Magnetic Resonance* 64, 1 (1985), 81–93.
- [36] GRISWOLD, M., BREUER, F., BLAIMER, M., KANNENGIESSER, S., HEIDEMANN, R., MUELLER, M., NITTKA, M., JELLUS, V., KIEFER, B., AND JAKOB, P. Autocalibrated coil sensitivity estimation for parallel imaging. *NMR in Biomedicine* 19, 3 (2006), 316–324.
- [37] GRISWOLD, M., JAKOB, P., HEIDEMANN, R., NITTKA, M., JELLUS, V., WANG, J., KIEFER, B., AND HAASE, A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine* 47, 6 (2002), 1202–1210.
- [38] GRISWOLD, M., JAKOB, P., NITTKA, M., GOLDFARB, J., AND HAASE, A. Partially parallel imaging with localized sensitivities (PILS). *Magnetic Resonance in Medicine* 44, 4 (2000), 602–609.
- [39] GRISWOLD, M., KANNENGIESSER, S., HEIDEMANN, R., WANG, J., AND JAKOB, P. Field-of-view limitations in parallel imaging. *Magnetic Resonance in Medicine* 52, 5 (2004), 1118–1126.
- [40] GUDBJARTSSON, H., AND PATZ, S. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine* 34, 6 (1995), 910.
- [41] HAASE, A., FRAHM, J., MATTHAEI, D., HÄNICKE, W., AND MERBOLDT, K. FLASH imaging rapid NMR imaging using low flip-angle pulses. *Journal of Magnetic Resonance* 67, 2 (1986), 258–266.
- [42] HAGER, W., AND ZHANG, H. A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on Optimization* 16 (2005), 170.
- [43] HAHN, E. Nuclear induction due to free Larmor precession. *Physical Review* 77, 2 (1950), 297–298.
- [44] HAHN, E. Spin echoes. *Physical Review* 80 (1950), 580.
- [45] HAKEN, H., AND WOLF, H. *Atom- und Quantenphysik*. Springer, 2003.

- [46] HANKE, M. *Conjugate gradient type methods for ill-posed problems*. CRC Press, 1995.
- [47] HEBERLEIN, K., AND HU, X. Kriging and GRAPPA: A new perspective on parallel imaging reconstruction. In *Proceedings of the ISMRM 14th Annual Meeting, Seattle, USA (2006)*, p. 2465.
- [48] HEIDEMANN, R., GRISWOLD, M., HAASE, A., AND JAKOB, P. VD-auto-SMASH imaging. *Magnetic Resonance in Medicine* 45, 6 (2001), 1066–1074.
- [49] HEIDEMANN, R., GRISWOLD, M., SEIBERLICH, N., KRUGER, G., KAN-NENGISSER, S., KIEFER, B., WIGGINS, G., WALD, L., AND JAKOB, P. Direct parallel image reconstructions for spiral trajectories using GRAPPA. *Magnetic Resonance in Medicine* 56, 2 (2006), 317–326.
- [50] HEIDEMANN, R., GRISWOLD, M., SEIBERLICH, N., NITTKA, M., KAN-NENGISSER, S., KIEFER, B., AND JAKOB, P. Fast method for 1D non-cartesian parallel imaging using GRAPPA. *Magnetic Resonance in Medicine* 57, 6 (2007), 1037–1046.
- [51] HESTENES, M., AND STIEFEL, E. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards* 49, 6 (1952), 409–436.
- [52] HOGE, W., BROOKS, D., MADORE, B., AND KYRIAKOS, W. On the regularization of SENSE and Space-RIP in parallel MR imaging. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004* (2004), pp. 241–244.
- [53] HOGE, W., BROOKS, D., MADORE, B., AND KYRIAKOS, W. A tour of accelerated parallel MR imaging from a linear systems perspective. *Concepts in Magnetic Resonance Part A* (2005), 13–37.
- [54] HOHAGE, T. Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems* 13, 5 (1997), 1279–1299.
- [55] HOHAGE, T. On the numerical solution of a three-dimensional inverse medium scattering problem. *Inverse Problems* 17, 6 (2001), 1743–1764.

- [56] HUANG, F., AKAO, J., VIJAYAKUMAR, S., DUENSING, G., AND LIMKEMAN, M. k-t GRAPPA: A k-space implementation for dynamic MRI with high reduction factor. *Magnetic Resonance in Medicine* 54, 5 (2005), 1172–1184.
- [57] HUTCHINSON, M., AND RAFF, U. Fast MRI data acquisition using multiple detectors. *Magnetic Resonance in Medicine* 6, 1 (1988), 87–91.
- [58] JAKOB, P., GRISOWLD, M., EDELMAN, R., AND SODICKSON, D. AUTO-SMASH: a self-calibrating technique for SMASH imaging. *Magnetic Resonance Materials in Physics, Biology and Medicine* 7, 1 (1998), 42–54.
- [59] KALTENBACHER, B., NEUBAUER, A., AND SCHERZER, O. *Iterative regularization methods for nonlinear ill-posed problems*. de Gruyter, 2008.
- [60] KARAS, A., MERBOLDT, K., GRAESSNER, J., AND FRAHM, J. Black-blood imaging of the human heart using rapid stimulated echo acquisition mode (STEAM) MRI. *Journal of Magnetic Resonance Imaging* 26, 6 (2007), 1666–1671.
- [61] KELTON, J., MAGIN, R., AND WRIGHT, S. An algorithm for rapid image acquisition using multiple receiver coils. In *Proceedings of the SMRM 8th Annual Meeting, Amsterdam, Netherlands* (1989), p. 1172.
- [62] KNOLL, F., CLASON, C., UECKER, M., AND STOLLENBERGER, R. Improved Reconstruction in Non-Cartesian Parallel Imaging by Regularized Non-linear Inversion. In *Proceedings of the ISMRM 17th Annual Meeting, Honolulu, USA* (2009), p. 2721.
- [63] KUMAR, A., WELTI, D., AND ERNST, R. Imaging of macroscopic objects by NMR Fourier zeugmatography. *Naturwissenschaften* 62, 1 (1975), 34–34.
- [64] KÜNTZEL, M. *Parallele Datenakquisition zur Beschleunigung Diffusionsgewichteter Kernspintomographie mit Stimulierten Echos*. Doctoral Thesis, Universität Göttingen, 2006.
- [65] KWIAT, D., AND EINAV, S. Preliminary experimental evaluation of an inverse source imaging procedure using a decoupled coil detector array in magnetic resonance imaging. *Medical Engineering and Physics* 17, 4 (1995), 257–263.
- [66] KWIAT, D., EINAV, S., AND NAVON, G. A decoupled coil detector array for fast image acquisition in magnetic resonance imaging. *Medical Physics* 18 (1991), 251.

- [67] KYRIAKOS, W., PANYCH, L., KACHER, D., WESTIN, C., BAO, S., MULKERN, R., AND JOLESZ, F. Sensitivity profiles from an array of coils for encoding and reconstruction in parallel (SPACE RIP). *Magnetic Resonance in Medicine* 44, 2 (2000), 301–308.
- [68] LARSSON, E., ERDOGMUS, D., YAN, R., PRINCIPE, J., AND FITZSIMMONS, J. SNR-optimality of sum-of-squares reconstruction for phased-array magnetic resonance imaging. *Journal of Magnetic Resonance* 163, 1 (2003), 121–123.
- [69] LAUTERBUR, P. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature* 242, 5394 (1973), 190–191.
- [70] LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. *The Quarterly of Applied Mathematics* 2, 2 (1944), 164–168.
- [71] LIANG, Z., AND LAUTERBURG, P. *Principles of Magnetic Resonance Imaging*. IEEE Press Series on Biomedical Engineering, 2000.
- [72] LIU, C., BAMMER, R., KIM, D., AND MOSELEY, M. Self-navigated interleaved spiral (SNAILS): application to high-resolution diffusion tensor imaging. *Magnetic Resonance in Medicine* 52, 6 (2004), 1388–1396.
- [73] LIU, C., BAMMER, R., AND MOSELEY, M. Parallel imaging reconstruction for arbitrary trajectories using k-space sparse matrices (kSPA). *Magnetic Resonance in Medicine* 58, 6 (2007), 1171–1181.
- [74] LIU, C., MOSELEY, M., AND BAMMER, R. Simultaneous phase correction and SENSE reconstruction for navigated multi-shot DWI with non-Cartesian k-space sampling. *Magnetic Resonance in Medicine* 54, 6 (2005), 1412–1422.
- [75] LUSTIG, M., DONOHO, D., AND PAULY, J. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine* 58, 6 (2007), 1182–1195.
- [76] MARQUARDT, D. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* (1963), 431–441.
- [77] MCKENZIE, C., OHLIGER, M., YEH, E., PRICE, M., AND SODICKSON, D. Coil-by-coil image reconstruction with SMASH. *Magnetic Resonance in Medicine* 46, 3 (2001), 619–623.



- [78] MCKENZIE, C., YEH, E., OHLIGER, M., PRICE, M., AND SODICKSON, D. Self-calibrating parallel imaging with automatic coil sensitivity extraction. *Magnetic Resonance in Medicine* 47, 3 (2002), 529–538.
- [79] MILLER, K., AND PAULY, J. Nonlinear phase correction for navigated diffusion imaging. *Magnetic Resonance in Medicine* 50, 2 (2003), 343–353.
- [80] MOORE, E. On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society* 26 (1920), 394–395.
- [81] NOLTE, U., FINSTERBUSCH, J., AND FRAHM, J. Rapid isotropic diffusion mapping without susceptibility artifacts: whole brain studies using diffusion-weighted single-shot STEAM MR imaging. *Magnetic Resonance in Medicine* 44, 5 (2000), 731–736.
- [82] NUNES, R., JEZZARD, P., BEHRENS, T., AND CLARE, S. Self-navigated multishot echo-planar pulse sequence for high-resolution diffusion-weighted imaging. *Magnetic Resonance in Medicine* 53, 6 (2005), 1474–1478.
- [83] PENROSE, R. A generalized inverse for matrices. In *Proceedings of the Cambridge Philosophical Society* (1955), vol. 51, pp. 406–413.
- [84] PIPE, J. Multishot diffusion weighted FSE with PROPELLER. *Magnetic Resonance in Medicine* 47, 1 (2002), 42–52.
- [85] PRUESSMANN, K., WEIGER, M., BORNERT, P., AND BOESIGER, P. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magnetic Resonance in Medicine* 46, 4 (2001), 638–651.
- [86] PRUESSMANN, K., WEIGER, M., SCHEIDEGGER, M., AND BOESIGER, P. SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine* 42, 5 (1999), 952–962.
- [87] PURCELL, E., TORREY, H., AND POUND, R. Resonance absorption by nuclear magnetic moments in a solid. *Physical Review* 69, 1-2 (1946), 37–38.
- [88] RA, J., AND RIM, C. Fast imaging using subencoding data sets from multiple detectors. *Magnetic Resonance in Medicine* 30, 1 (1993), 142–145.
- [89] RADER, C. Discrete Fourier transforms when the number of data samples is prime. *Proceedings of the IEEE* 56, 6 (1968), 1107–1108.

- [90] RIESEBERG, S., MERBOLDT, K., KÜNTZEL, M., AND FRAHM, J. Diffusion tensor imaging using partial Fourier STEAM MRI with projection onto convex subsets reconstruction. *Magnetic Resonance in Medicine* 54, 2 (2005), 486–490.
- [91] ROEMER, P., EDELSTEIN, W., HAYES, C., SOUZA, S., AND MUELLER, O. The NMR phased array. *Magnetic Resonance in Medicine* 16, 2 (1990), 192–225.
- [92] RUDIN, L., OSHER, S., AND FATEMI, E. Nonlinear total variation based noise removal algorithms. *Physica D* 60, 3 (1992), 259–268.
- [93] SAMSONOV, A., BLOCK, W., ARUNACHALAM, A., AND FIELD, A. Advances in locally constrained k-space-based parallel MRI. *Magnetic Resonance in Medicine* 55, 2 (2006), 431–438.
- [94] SEIBERLICH, N., BREUER, F., BLAIMER, M., BARKAUSKAS, K., JAKOB, P., AND GRISWOLD, M. Non-Cartesian data reconstruction using GRAPPA operator gridding (GROG). *Magnetic Resonance in Medicine* 58, 6 (2007), 1257–1265.
- [95] SHANNON, C. Communication in the presence of noise. *Proceedings of the IRE* 37, 1 (1949), 10–21.
- [96] SHEPP, L., AND LOGAN, B. The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science* 21, 1 (1974), 21–43.
- [97] SKARE, S., NEWBOULD, R., CLAYTON, D., ALBERS, G., NAGLE, S., AND BAMMER, R. Clinical multishot DW-EPI through parallel imaging with considerations of susceptibility, motion, and noise. *Magnetic Resonance in Medicine* 57, 5 (2007), 881–890.
- [98] SODICKSON, D., GRISWOLD, M., JAKOB, P., EDELMAN, R., AND MANNING, W. Signal-to-noise ratio and signal-to-noise efficiency in SMASH imaging. *Magnetic Resonance in Medicine* 41, 5 (1999), 1009–1022.
- [99] SODICKSON, D., AND MANNING, W. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magnetic Resonance in Medicine* 38, 4 (1997), 591–603.

- [100] STEJSKAL, E., AND TANNER, J. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics* 42 (1965), 288.
- [101] TANNER, J. Self diffusion of water in frog muscle. *Biophysical Journal* 28, 1 (1979), 107–116.
- [102] TORREY, H. Bloch equations with diffusion terms. *Physical Review* 104, 3 (1956), 563–565.
- [103] UECKER, M., HOHAGE, T., BLOCK, K., AND FRAHM, J. Image reconstruction by regularized nonlinear inversion — Joint estimation of coil sensitivities and image content. *Magnetic Resonance in Medicine* 60, 3 (2008), 674–682.
- [104] WAJER, F., AND PRUESSMANN, K. Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories. In *Proceedings of the ISMRM 9th Annual Meeting, Glasgow, Scotland* (2001), p. 767.
- [105] WALSH, D., GMITRO, A., AND MARCELLIN, M. Adaptive reconstruction of phased array MRI imagery. *Magnetic Resonance in Medicine* 43, 5 (2000), 682–690.
- [106] WANG, Z., AND FERNANDEZ-SEARA, M. 2D partially parallel imaging with k-space surrounding neighbors-based data reconstruction. *Magnetic Resonance in Medicine* 56, 6 (2006), 1389–1396.
- [107] WEIGER, M., PRUESSMANN, K., AND BOESIGER, P. 2D SENSE for faster 3D MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine* 14, 1 (2002), 10–19.
- [108] WERNER, F. *Ein neuer numerischer Ansatz zur  $L^p$ -Regularisierung*. Diploma Thesis, Universität Göttingen, 2008.
- [109] WHITTAKER, J. The Fourier theory of the cardinal function. *Proceedings of the Edinburgh Mathematical Society* 1, 03 (2009), 169–176.
- [110] WILLIG-ONWUACHI, J., YEH, E., GRANT, A., OHLIGER, M., MCKENZIE, C., AND SODICKSON, D. Phase-constrained parallel MR image reconstruction. *Journal of Magnetic Resonance* 176, 2 (2005), 187–198.
- [111] XIANG, Q. Accelerating MRI by skipped phase encoding and edge deghosting (SPEED). *Magnetic Resonance in Medicine* 53, 5 (2005), 1112–1117.

- [112] YEH, E., MCKENZIE, C., OHLIGER, M., AND SODICKSON, D. Parallel magnetic resonance imaging with adaptive radius in k-space (PARS): Constrained image reconstruction using k-space locality in radiofrequency coil encoded data. *Magnetic Resonance in Medicine* 53, 6 (2005), 1383–1392.
- [113] YING, L., AND SHENG, J. Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). *Magnetic Resonance in Medicine* 57, 6 (2007), 1196–1202.

# Curriculum Vitae

## Personal Data

Name: Martin Uecker  
Date and Place of Birth: 1979/01/31 in Würzburg, Germany  
Nationality: German

## Education

1985-1989 Elementary school (Johannes-Kepler-Gundschule, Würzburg)  
1989-1998 Gymnasium (Wirsberg-Gymnasium, Würzburg)  
Abitur (1998/06/26)

## Community Service

1998-1999 Community Service (Zivildienst) at the  
Department of Radiotherapy of the  
Medical Center of the University of Würzburg

## Studies

1999-2005 Studies of Physics and Mathematics  
at the Georg-August-Universität Göttingen  
Interim exam in Mathematics (2001/05/20)  
Interim exam in Physics (2001/07/12)  
Diploma in Physics (2005/07/08)

Since 2005 PhD School of Mathematical Sciences (University of Göttingen),  
PhD Student at the Biomedizinische NMR Forschungs GmbH  
am Max-Planck-Institut für biophysikalische Chemie in Göttingen

Since 2008 Associated member of the graduate school  
„Identifikation in mathematischen Modellen:  
Synergie stochastischer und numerischer Methoden“  
at the Faculty of Mathematics of the University of Göttingen

# List of Publications

## Journal Publications

- Kai Tobias Block, Martin Uecker and Jens Frahm, Undersampled Radial MRI with Multiple Coils. Iterative Image Reconstruction Using a Total Variation Constraint, *Magnetic Resonance in Medicine*, 57(6):1086-1098, 2007.
- Martin Uecker, Thorsten Hohage, Kai Tobias Block and Jens Frahm, Image Reconstruction by Regularized Nonlinear Inversion - Joint Estimation of Coil Sensitivities and Image Content, *Magnetic Resonance in Medicine*, 60(3):674-82, 2008.
- Kai Tobias Block, Martin Uecker and Jens Frahm, Suppression of MRI Truncation Artifacts Using Total Variation Constrained Data Extrapolation, *International Journal of Biomedical Imaging*, 2008:184123, 2008.
- Kai Tobias Block, Martin Uecker and Jens Frahm, Model-based Iterative Reconstruction for Radial Fast Spin-Echo MRI, *IEEE Transactions on Medical Imaging*, in press, 2009.
- Martin Uecker, Alexander Karaus and Jens Frahm, Inverse Reconstruction Method for Segmented Multi-shot Diffusion-weighted MRI With Multiple Coils, *Magnetic Resonance in Medicine*, in press, 2009.

## Conference Contributions

- Kai Tobias Block, Martin Uecker and Jens Frahm, Iterative Image Reconstruction of Undersampled Radial MRI Data from Multiple Coils, *ISMRM Workshop on Non-Cartesian MRI Sedona 2007*, Workshop Syllabus:Poster 41.
- Kai Tobias Block, Martin Uecker and Jens Frahm, Undersampled Radial MRI with Multiple Coils. Iterative Image Reconstruction Using a Total Variation

- Constraint, Joint Annual Meeting ISMRM-ESMRMB Berlin 2007, In Proc. Intl. Soc. Mag. Reson. Med. 15 (2007):145.
- Martin Uecker, Kai Tobias Block and Jens Frahm, Parallel Imaging as a Non-linear Inversion Problem — Improved Reconstructions, Joint Annual Meeting ISMRM-ESMRMB Berlin 2007, In Proc. Intl. Soc. Mag. Reson. Med. 15 (2007):1740.
  - Kai Tobias Block, Martin Uecker and Jens Frahm, Iterative Reconstruction for R2 Mapping Based on Radial Fast Spin-Echo MRI, Annual Meeting ISMRM Toronto 2008, In Proc. Intl. Soc. Mag. Reson. Med. 16 (2008):1432.
  - Kai Tobias Block, Martin Uecker and Jens Frahm, Suppression of MRI Truncation Artifacts Using Total Variation Constrained Data Extrapolation, Annual Meeting ISMRM Toronto 2008, In Proc. Intl. Soc. Mag. Reson. Med. 16 (2008):1477.
  - Martin Uecker, Kai Tobias Block and Jens Frahm, Non-linear Inversion with L1-Wavelet Regularization — Application to Autocalibrated Parallel Imaging, Annual Meeting ISMRM Toronto 2008, In Proc. Intl. Soc. Mag. Reson. Med. 16 (2008):1479.
  - Martin Uecker, Kai Tobias Block and Jens Frahm, An Automatic Stopping Criterion for Iterative MRI Reconstructions, Annual Meeting ISMRM Toronto 2008, In Proc. Intl. Soc. Mag. Reson. Med. 16 (2008):1499.
  - Martin Uecker, Simultane Schätzung von Spulensensitivitäten und Bildinhalt in der parallel MRT — Bildrekonstruktion mit einer nichtlinearen Regularisierungsmethode, 11. Jahrestagung der Deutschen Sektion der ISMRM e.V. Frankfurt 2008
  - Martin Uecker, Alexander Karaus and Jens Frahm, Improved Image Reconstruction for Multi-Shot Multi-Coil Diffusion Imaging, Annual Meeting ISMRM Honolulu 2009, In Proc. Intl. Soc. Mag. Reson. Med. 17 (2009):2842
  - Florian Knoll, Christian Clason, Martin Uecker, Rudolf Stollenberger, Improved Reconstruction in Non-Cartesian Parallel Imaging by Regularized Non-linear Inversion, Annual Meeting ISMRM Honolulu 2009, In Proc. Intl. Soc. Mag. Reson. Med. 17 (2009):2721

## Other Publications

- Martin Uecker, Lokale Gleichgewichtszustände des elektromagnetischen Feldes (Diploma thesis), Institut für Theoretische Physik, Universität Göttingen 2006
- Martin Uecker, Reconstruction from Undersampled Multichannel Data in Parallel Magnetic Resonance Imaging, MPIbpc News, Max-Planck-Institut für biophysikalische Chemie, Issue 6, 2008



# Acknowledgements

First of all, I have to thank Prof. Dr. Jens Frahm, head of the Biomedizinische NMR Forschungs GmbH am Max-Planck-Institut für biophysikalische Chemie. The present work would not have been possible without his continuous support.

I am very grateful to Prof. Dr. Thorsten Hohage from the Institut für Numerische und Angewandte Mathematik of the Georg-August-Universität Göttingen for the supervision of this thesis, his interest in my work, and all the time he spent discussing mathematical questions with me.

Next, I want to thank Dr. Tobias Block for leading my way into the field of Magnetic Resonance Imaging and for a decade-long friendship.

I want to thank Mr. Alexander Karaus, who shared a room with me during the last years and always helped to keep my spirits up. He also contributed directly to this thesis by providing me with the data used in Chapter 6.

Further, I would like to thank all members of the BiomedNMR for creating a pleasant working atmosphere.

I am deeply indebted to my parents, who encouraged my scientific interest from the early on and continuously supported me during my studies.

Finally, I want to thank my girlfriend Anna and our daughter Maja Antonina for their endurance, company, and love.

