UNIVERSITÄT ZU LÜBECK

From the Institute for Signal Processing
of the University of Lübeck
Director: Prof. Dr.-Ing. Alfred Mertins

Model-based Image Reconstruction for Multi-shot Diffusion-weighted Imaging

Dissertation
for Fulfillment of
Requirements
for the Doctoral Degree
of the University of Lübeck

from the Department of Computer Sciences and Technical Engineering

Submitted by

Malte Riedel
from Diekholzen

Lübeck, 2021

First referee:	Prof. Dr.-Ing. Alfred Mertins
Second referee:	Prof. Dr. Martin Koch
Date of oral examination:	12.08.2021
Approved for printing.	Lübeck, 16.08.2021

Questions you cannot answer are usually far better for you than answers you cannot question.

- Yuval Noah Harari,
21 Lessons for the 21st Century

Contents

1

General Introduction

Magnetic resonance imaging (MRI) is a well-established non-invasive imaging modality for medical diagnostics and many research disciplines, which provides unique anatomical and functional insights into the living human body [1–3]. As a central characteristic, MRI provides various contrast mechanisms enabling unparalleled soft tissue differentiation. As such, diffusion-weighted imaging (DWI) [4–8] incorporates a special MRI contrast mechanism that allows to encode the three-dimensional diffusion process into the MRI signal, using the abundant water molecules in the human body to probe the tissue micro-structure far below the actual image resolution. The characterization of the diffusion process provides sensitive markers for stroke diagnostics and to differentiate the abnormal tissue cellularity of cancer [7]. The analysis of directional diffusion anisotropies can further be linked to the white matter connectivity allowing to study the physiological brain function as well as pathological alterations [4]. Besides the brain, the diffusion contrast also offers valuable tissue characterization for body anatomies like the liver or the heart, although body imaging usually suffers from faster signal decays and more complex motion profiles [4].

DWI sequences are particularly demanding due to the high gradient hardware requirements and its inherent sensitivity to patient motion including physiological and voluntary effects [8, 9]. Therefore, clinical applications have been dominated for many years by single-shot echo-planar imaging (EPI) acquisitions [8], which rapidly acquire one snapshot image per signal excitation [10]. Moreover, the image acquisitions commonly make use of parallel imaging techniques [11] like SENSE [12, 13] or GRAPPA [14] employing multiple receive coils with varying spatial sensitivity profiles.

The need for improved image resolution and signal-to-noise ratio (SNR) has been fostering multi-shot applications [8] which obtain the image information from multiple successively acquired data segments offering more flexible sampling [15, 16]. However, the motion sensitivity of DWI entails signal variations which spoil the correlations between the individual shot datasets and severely affect image quality. As a major source of shot variations in DWI, the displacements from tiny subject motion [17] and cardiac pulsations of the brain [18, 19] during the strong diffusion-sensitizing gradients lead to shot-specific phase accruals. Macroscopic motion of the patient is another inconvenient and costly [20] source of signal variations, which involves further secondary sampling-related artifacts [21, 22]. Hence, the signal variations must be appropriately taken into account to harvest the joint information from large multi-shot multi-coil datasets.

Model-based image reconstructions for MRI [23] move beyond the classical inverse Fourier transform, parameterizing the signal variations in tailored signal models. Unfortunately, the parameter identification is typically ill-posed and non-linear [24]. This challenging problem gives rise to a wide variety of methods for estimating the shot-specific phase maps [17, 18, 25–42] and macroscopic motion [43, 44] in multi-shot DWI including different strategies to sense the signal variations and solve or substitute the ill-posed problem. *Navigated* methods, for example, use additional signals in the MRI sequence for the parameter identification, while *self-navigated* methods deduce navigation signals from the imaging data itself. In addition, MRI has been revolutionized by compressed sensing [45] and deep learning [46, 47] which exploit the compressibility of data in appropriate domains to regularize the ill-posed problem. The multitude of models, navigation strategies and constraints constitutes a wide and active research field.

This thesis explores four new model-based image reconstructions [48–51] for motion-corrected multi-shot DWI in the human brain addressing relevant confounding factors for both clinical and research practice. Starting from pure shot phase corrections, the proposed methods explore additional 2D in-plane and fully 3D rigid motion modeling with different navigation strategies. Overall, the main contributions of this dissertation for multi-shot DWI include:

1. Deep learning support for self-navigated shot phase estimations in an alternating optimization [48] with conventional image reconstruction (Section 3.1),

2. A self-navigated algorithm [49] with shot phase and rigid in-plane motion correction enabling high segmentations (Section 3.2),

3. A navigated algorithm [50] with shot phase and rigid in-plane motion correction for fast and robust image reconstructions (Section 4.1),

4. A comparison of the navigated and self-navigated algorithms for shot phase and rigid in-plane motion correction (in Section 4.1),

5. A navigated algorithm [51] with shot phase and 3D rigid motion correction including simultaneous multi-slice sampling and multi-slice-to-volume registration in a full-volume image reconstruction (Section 4.2).

This work is separated into five main chapters. After the introduction, the second chapter provides a brief introduction to MR image formation, DWI and its multi-shot modeling. The third chapter covers the two self-navigated image reconstruction approaches, including the deep learning supported shot phase estimation and the rigid in-plane motion corrections. The fourth chapter contains the two navigated algorithms, presenting the rigid in-plane and the fully 3D rigid motion corrections. The final chapter puts the presented approaches together in a generalized context providing a joint discussion and conclusions. Note that the order of the chapters on self-navigated and navigated methods is unrelated to any technological preferences but rather attempts to group similar reconstruction concepts.

2

Theoretical Background

Diffusion-weighted imaging has been present in the MR research community for almost 40 years and is still a rapidly evolving research field with continuous innovations ranging from sequence design with advanced encoding technologies over intelligent reconstruction strategies to high-end diffusion modeling and post-processing. At the same time, DWI is a challenging imaging contrast due to the demanding hardware requirements and its inherent motion sensitivity. This work contributes to the technical developments of multi-shot DWI in the presence of shot phase variations and patient gross motion. This chapter provides a theoretical background of the image encoding and reconstruction which is required to provide a comprehensible description of the thesis contributions. The first section gives a brief introduction to magnetic resonance imaging and the state-of-the-art image formation principles that are relevant for DWI along with the corresponding reconstruction philosophies. The second section introduces diffusion-weighted imaging, describes the modeling framework for multi-shot DWI and reviews related work on this topic.

The theoretical background in this thesis is presented with a focus on multi-shot DWI. For a broader theoretical coverage, the reader is referred to textbooks on MRI [1–3], convex optimization [52, 53] and inverse problems [24]. For the topic of this dissertation, specialized articles on diffusion-weighted imaging [4–8], model-based reconstruction [23], motion correction in MRI [21, 22] and simultaneous multi-slice acquisition [54] are recommended.

2.1

MR Image Formation

An MRI sequence is built from radio-frequency excitation pulses, which produce a detectable transverse magnetization, and a spatial encoding process that allows to resolve the image. Today's MR image formation is mainly driven by two powerful signal encoding technologies: the Fourier encoding using gradient fields and the sensitivity encoding using localized coil sensitivity profiles. Although these technologies are mutually independent in the first place, their interplay has been fruitfully developed over the last decades and still represents an active research field. This section begins with some elementary MRI principles. Next, a selection of image formation techniques for the state-of-the-art of DWI is presented covering slice-selective excitations, the Fourier and the sensitivity encoding. In the end, simultaneous multi-slice imaging is introduced, which has led to considerable scan accelerations at minor SNR penalty for DWI.

Magnetic Resonance Imaging Principles

Magnetic resonance imaging in today's clinics predominantly examines the resonance of hydrogen nuclei (protons) and is based on three types of magnetic fields that are jointly applied to obtain non-invasive tomographic images [1]. First, a strong static magnetic field $\mathbf{B}_0$, by convention applied along the longitudinal z direction, induces a net magnetization in the object, whose equilibrium value is proportional to the present magnetic field and the spin density of the object. Moreover, the nuclear spins exhibit the magnetic resonance phenomenon at the *Larmor frequency* in the presence of a magnetic field $\mathbf{B}$:

$$\boldsymbol{\omega} = \gamma \mathbf{B}, \tag{2.1}$$

where γ is the gyromagnetic ratio, which is $\gamma' = \gamma/(2\pi) = 42.58$ MHz/T for protons. Second, a transmit radio-frequency (RF) field B_1^+, which is applied at the Larmor frequency in the transversal x-y-plane, excites the spin system and produces transverse magnetization components, which themselves precess at the Larmor frequency.

By Faraday's law of induction, the precessing magnetization is detectable by transversely aligned receiver coils, which sense complex-valued signals using phase-sensitive

acquisitions [1]. The return to the thermodynamic equilibrium of the spin system is governed by two relaxation processes. The longitudinal magnetization recovers with time constant T_1via energy dissipation into the local environment, whereas the transverse magnetization decays with the statistical loss of phase coherence described by the relaxation constant T_2. The Larmor precession and the relaxation processes jointly describe the basic MR spin dynamics in the Bloch equations [1].

As the third field, the gradient field **G** provides the fundamental spatial encoding technology for MRI [1]. The gradient coils are designed, one for each of the three orthogonal spatial directions, to produce a longitudinal (z) magnetic field component that varies linearly along the gradient axis. The linear field variation causes a frequency modulation according to Eq. 2.1, which translates to a linear phase accrual of the transverse magnetization characterized by the spatial frequency:

$$\mathbf{k}(t) = \gamma \int_0^t \mathbf{G}(\tau)\, d\tau. \tag{2.2}$$

The baseband signal equation describes the signal formation acquired in the volume of interest (VOI) after demodulation with the Larmor frequency $\boldsymbol{\omega}_0 = \gamma \mathbf{B}_0$ ignoring the relaxation terms [1]:

$$d(\mathbf{k}) = \int_{\mathrm{VOI}} \rho(\mathbf{r}) e^{-j\mathbf{kr}}\, d\mathbf{r}. \tag{2.3}$$

Hence, the linear spatial phase modulation induced by the gradient field effectively encodes the Fourier, or k-space, representation of the detectable transverse magnetization. The acquired MR data d is therefore the Fourier representation of the image object ρ. **k** and **r** are the k-space and spatial coordinates, respectively. The image is readily recovered by an inverse Fourier transform that is efficiently implemented in practice using the fast Fourier transform (FFT).

SLICE-SELECTIVE EXCITATION

In the thermal equilibrium, the object magnetization is aligned with the main magnetic field $\mathbf{B}_0$. Governed by the Bloch equations, the magnetization can be excited to obtain receivable transverse components using a RF magnetic field applied on-resonance in the transversal plane [1]. In the presence of the homogeneous $\mathbf{B}_0$ field, a resonant RF pulse equally excites the full volume and is called *non-selective*. Using an additional gradient field, the resonance frequency varies spatially and the excitation can thus be restricted to a specific region by RF pulses with dedicated frequency bands. By this *selective* excitation, the image encoding can be slice-specifically localized reducing the acquisition and reconstruction to a manageable two-dimensional problem. DWI sequences are commonly based on slice-selective excitations [5] and they are therefore considered in this work.

The resonance condition of an ideal slice with a rectangular profile in the presence of a gradient field corresponds to a rectangular radio-frequency band according to the linear Larmor relation in Eq. 2.1. The time-domain RF pulse for this single-band (SB) excitation has a sinc waveform $W_{\mathrm{SB}}(t) = \Delta f\, \mathrm{sinc}(\pi \Delta f\, t)$ with the frequency $\Delta f = \gamma' G_z \Delta z$ and the slice thickness Δz.

The complex-valued RF pulse $\mathrm{RF}_{\mathrm{SB}}(t)$ in the transverse plane can be constructed using the RF waveform W_{SB} and a phase modulation term [54]:

$$\mathrm{RF}_{\mathrm{SB}}(t) = W_{\mathrm{SB}}(t) \cdot e^{j\omega_l t}, \tag{2.4}$$

where ω_l accounts for the Larmor frequency offset caused by the spatial offset z_l of slice l from the iso-center via $\omega_l = \gamma G_z z_l$. The infinite sinc waveform $W_{\mathrm{SB}}(t)$ is replaced by filtered finite waveforms in practice.

FOURIER ENCODING

As briefly outlined before, the gradient system provides the basic encoding technology of MRI and offers a flexible sampling machinery for various applications. After some general sampling considerations, this subsection introduces the echo-planar imaging (EPI) concept for fast snap-shot imaging, which has become the workhorse of state-of-the-art DWI and functional MRI. Different undersampling strategies are covered for EPI and, finally, multi-shot acquisition techniques are introduced.

Sampling considerations

In MRI, the gradient fields are used to apply spatially linear phase modulations to the excited transverse magnetization. According to the signal equation in Eq. 2.3, a specific phase modulation corresponds to acquiring a spatial frequency weight at $\mathbf{k}$ of the image object's k-space representation. By playing out a sequence of gradients, the spatial frequency content can be successively encoded and sampled. The sampling trajectory describes the acquisition sequence covering the set of sampling positions K.

According to the Nyquist theorem, a signal can be unambiguously recovered if it is band-limited to half of the sampling frequency [1]. This condition relates the Cartesian image field of view FOV_i and its resolution Δr_i to the k-space sampling width W_i and its increment Δk_i by $\mathrm{FOV}_i = \Delta k_i^{-1}$ and $\Delta r_i = W_i^{-1}$ for the spatial encoding dimensions $i \in \{x, y, z\}$. An image is fully sampled if the k-space sampling in K satisfies these relations.

The discretized Fourier image encoding is described by the forward model relating the k-space data vector $\check{\mathbf{d}} \in \mathbb{C}^{N_k}$ with N_k samples to the image $\boldsymbol{\rho} \in \mathbb{C}^{N_p}$ with N_p pixels:

$$\check{\mathbf{d}} = \check{M}\check{F}\,\boldsymbol{\rho}, \tag{2.5}$$

involving the Fourier operator $\check{F}$ $(N_p \times N_p)$ and the sampling operator $\check{M}$ $(N_k \times N_p)$, which applies a mask for the N_k samples of the sampling trajectory K. In case of full Cartesian Nyquist sampling, $N_k = N_p$ and the sampling operator is the identity $\check{M} = I_{N_p}$. The image $\boldsymbol{\rho} = \check{F}^H\check{\mathbf{d}}$ is then efficiently recovered using the FFT.

Standard *spin warp imaging* [55] repeatedly performs three steps to sample the Cartesian k-space grid of an image. First, a *slice-selective* RF pulse produces transverse magnetization for one slice. By this restriction along the slice direction, the encoding is reduced to a 2D imaging task. In the second *phase encoding* step, a gradient pulse adjusts the spatial frequency weighting along one in-plane direction. In the third *frequency encoding* step, one line of k-space samples is acquired in the residual orthogonal in-plane direction under a constant readout gradient. The procedure is repeated line-by-line until the k-space grid is fully sampled. In the following, the coordinates x, y and z denote the readout, phase and slice encoding direction, respectively.

Single-shot echo-planar imaging

Spin warp imaging uses multiple excitations, also called shots, which jointly yield the required image information in k-space. Such multi-shot approaches are prone to shot-specific variations that spoil the overall data consistency and produce image artifacts. DWI sequences are naturally sensitive to motion and produce strong inter-shot variations [8]. Therefore, single-shot EPI [10], which rapidly acquires the k-space in one single shot, has become a clinical standard for DWI by avoiding the multi-shot variations.

Single-shot EPI is a fast image acquisition technique that samples the spatial frequencies of the image by a one-time k-space traversal in a zig-zag pattern [10]. Figure 2.1 shows a spin-echo sequence with EPI sampling and the associated k-space trajectory. Starting with the sequence in Fig. 2.1a, the slice-selective excitation (90°) and the refocusing (180°) RF pulse form a spin echo at the image echo time TE_{img}, around which the EPI acquisition is centered. The readout gradients are played out with alternating signs successively producing gradient echoes, while the gradients blips in the phase encoding direction k_y induce small Δk_y steps. The resulting single-shot EPI k-space trajectory is shown in Fig. 2.1b. Starting from the k-space center, the trajectory moves to the upper left by the dephasing gradients on the read and phase encoding channel. The sampling starts under the initial read train lobe alternately moving left and right in readout direction k_x parted by Δk_y blips, which shift the sampling line-by-line along k_y.

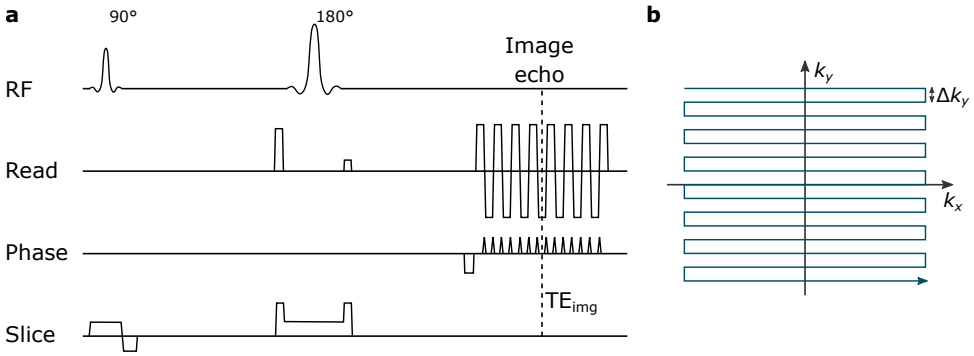


Figure 2.1: EPI sampling overview. (a) Spin-echo sequence example with EPI sampling. A slice-selective RF excitation pulse (90°) and a refocusing pulse form a spin-echo at the image echo time TE_{img}. The image echo is sampled by the EPI acquisition window centered around TE_{img}. (b) Single-shot EPI k-space trajectory.

Single-shot EPI offers fast acquisitions and remains relatively robust to off-resonance effects, but it suffers from several limiting factors [9]. The high acquisition speed at full k-space coverage avoids multi-shot variations and also reduces the impact of motion through the mere scan time reduction. Apart from ramp sampling, EPI is a Cartesian method, which allows for efficient reconstructions without further interpolation. In contrast, single-shot EPI resolution is limited by the signal decay through T_2^* blurring and the technical as well as physiological gradient limitations which impede sampling speed-ups [9]. Off-resonance effects from eddy currents, local field inhomogeneities and susceptibility variations manifest as geometric distortions in the low-bandwidth phase encoding direction [9]. In addition, EPI suffers from Nyquist ghosting [56], which results from inconsistencies between the odd and even phase encoding lines.

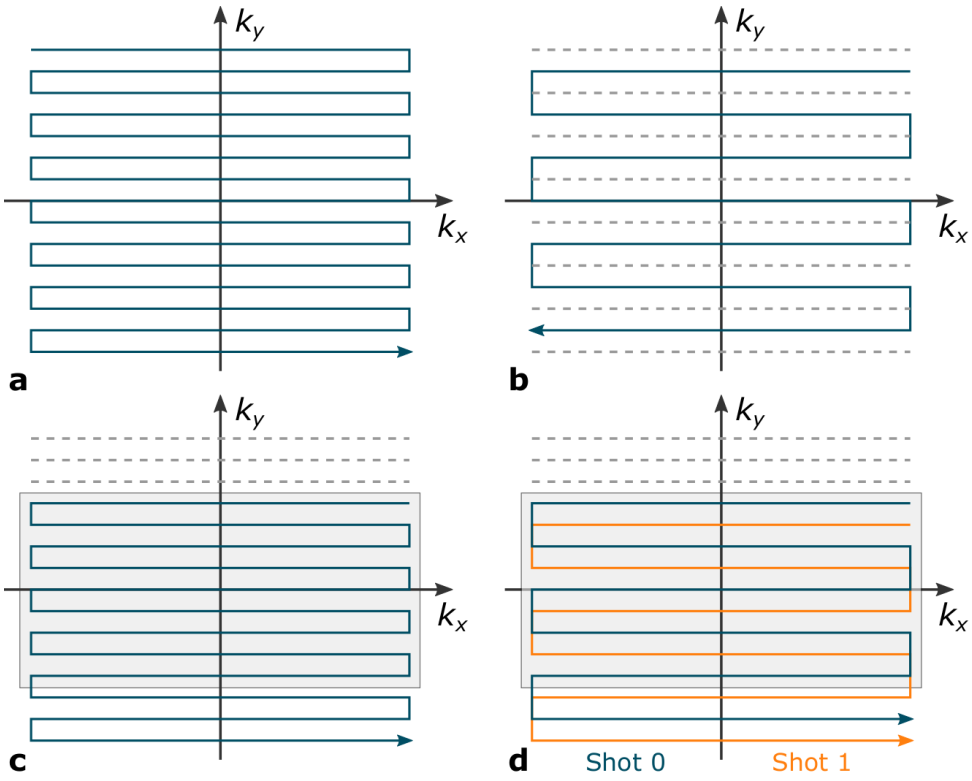


Figure 2.2: Overview of common EPI k-space trajectories at equal resolution. (a) Standard single-shot EPI. (b) Two-fold regular undersampling for single-shot EPI. The gray dotted lines indicate k-space positions that were not sampled. (c) Partial Fourier undersampling for single-shot EPI. (d) Multi-shot EPI with 2 interleaves and partial Fourier sampling. The individual interleaves are indicated by different colors. Note that each of the two interleaves is interpretable as an individual single-shot experiment. The gray boxes mark the symmetric partial Fourier sampling areas, which restrict the resolution of the object phase.

EPI acquisitions are frequently combined with undersampling and multi-shot strategies, which are visualized in Fig. 2.2. The standard single-shot EPI sampling is shown in Fig. 2.2a for comparison. Regular undersampling is a common technique to accelerate the trajectory and balance speed, resolution, artifacts and noise. Undersampled acquisitions violate the Nyquist criterion causing aliasing in the image space signal, which is described by its point spread function:

$$\text{PŠF} = \check{F}^H \check{M}^H \check{M} \check{F}. \tag{2.6}$$

The aliased image $\check{\boldsymbol{\rho}}$ is described by the PSF using Eq. 2.5 as $\check{\boldsymbol{\rho}} = \check{F}^H \check{M}^H \check{\mathbf{d}} = \text{PŠF}\, \boldsymbol{\rho}$. Parallel imaging techniques like SENSE [12], which are discussed in the next subsection, can resolve the aliasing from undersampled trajectories to a certain extent using the additional coil sensitivity encoding. By this, common sampling trajectories can be accelerated intentionally violating the Nyquist rate. Figure 2.2b shows the shortened single-shot EPI train for regular undersampling, which regularly skips k-space lines by taking larger steps $R\,\Delta k_y$ with reduction factor R ($R = 2$ for this example).

Another common scan acceleration for EPI is the partial Fourier acquisition, which exploits smoothness assumptions of the signal phase. Partial Fourier trajectories asymmetrically omit parts of the higher frequencies in the phase encoding direction k_y, as shown in Fig. 2.2c. According to the symmetry properties of the Fourier transform, the k-space signal d of a purely real-valued image ρ comprises Hermitian symmetry [3]:

$$d(\mathbf{k}) = d^{\star}(-\mathbf{k}). \tag{2.7}$$

It is therefore sufficient to acquire half of the k-space for real-valued images. MR images are usually not real-valued, but the phase is often assumed to be smooth. As a relaxation of the full Hermitian symmetry, partial Fourier acquisitions sample more than half of the k-space, usually 60-70%, and assume that the image phase is fully described by the symmetric area in the k-space center, which is shaded in gray in Fig. 2.2c.

There are several approaches to exploit the phase smoothness for the recovery of the full k-space signal [3]. One common algorithm, for example, uses projections onto convex sets (POCS) to iteratively enforce data consistency in k-space and phase smoothness in image-space. The data consistency step projects the image estimate on the measured k-space data and the phase smoothness step substitutes the image ρ by $|\rho| \cdot e^{j\theta}$ with the smooth phase map θ obtained from the fully sampled gray area. Besides the partial Fourier approach along k_y, the same philosophy can be used to asymmetrically reduce the sampling in readout direction k_x, which is called partial echo acquisition.

The regular undersampling and partial Fourier acceleration techniques have pushed the technical limitations of single-shot EPI and have therefore found widespread clinical acceptance. Nevertheless, a relaxation of the single-shot aspirations offers flexibility for EPI that can be used to mitigate artifacts and improve SNR or resolution.

Multi-shot echo-planar imaging

Multi-shot techniques repeat the single-shot sequence multiple times and sequentially acquire different k-space subsets per shot, which are also called segments or interleaves for EPI. Figure 2.2d shows an example with 2-fold segmentation in phase encoding direction [15, 16] including partial Fourier acquisition. Each individual segment can be interpreted as a single shot with regular undersampling. Without partial Fourier sampling, the two shots jointly yield the full k-space.

In this way, the sampling speed in the low-bandwidth k_y direction is increased, which results in less geometric distortions at the cost of fewer k-space coverage of the signal. Multi-shot scans lengthen the total acquisitions time and allow trading off SNR, artifacts, image resolution and hardware constraints [9]. On the other hand, segmented acquisitions are prone to inter-shot variations and require dedicated reconstruction models to combine the shot datasets. Another possible multi-shot approach is readout-segmented EPI [29], which performs segmentation in the readout direction k_x.

Besides EPI, non-Cartesian trajectories like spiral or radial sampling offer beneficial PSF patterns [13] and higher motion robustness [57]. Non-Cartesian Fourier encoding generally requires some sort of interpolation to translate the samples to a Cartesian image grid. The interpolation is commonly implemented by a gridding operation [58] that replaces the sampling operator M as for CG-SENSE [13]. The downsides of non-Cartesian acquisitions are the increased computational requirements due to the additional interpolation and the susceptibility to off-resonance-related blurring effects [57].

COIL SENSITIVITY ENCODING

The second major technology for MRI encoding is parallel imaging [11], which uses phased arrays for parallel signal reception. Instead of one huge homogeneous coil, multiple surface coils are arranged around the object. The localized and inhomogeneous coil sensitivity profiles pick up less patient-induced noise allowing for considerable SNR improvements [11]. The spatial encoding of dedicated coil designs provides valuable data redundancies among the coil signals, which can be leveraged to unravel the aliasing of undersampled data by methods like SENSE [12, 13] or GRAPPA [14]. Therefore, parallel imaging has found widespread adoption offering either speed, resolution or SNR gains.

Extended signal equation

The Fourier-based signal model in Eq. 2.3 is extended to incorporate the inhomogeneous sensitivity profile $c_\mathsf{c}(\mathbf{r})$ of receive coil c at location $\mathbf{r}$ [12]. The received data d_c of coil c integrates the Fourier components of the sensitivity-weighted image:

$$d_\mathsf{c}(\mathbf{k}) = \int_{\mathrm{VOI}} \rho(\mathbf{r})\, c_\mathsf{c}(\mathbf{r})\, e^{-j\mathbf{kr}}\, d\mathbf{r}. \tag{2.8}$$

A discrete realization of the sampling with known sensitivity and gradient encoding can be translated to a SENSE forward model, as visualized in Fig. 2.3, describing the data vector $\mathbf{d} \in \mathbb{C}^{N_c N_k}$ for N_c coils and N_k k-space samples [12]:

$$\mathbf{d} = M\, F\, C\, \boldsymbol{\rho}., \tag{2.9}$$

including the sensitivity operator C ($N_c N_p \times N_p$). The Fourier operator F ($N_c N_p \times N_c N_p$) and the sampling operator M ($N_c N_k \times N_c N_p$) are adapted with block-diagonal structure to the multi-coil setting. The model can be written in image-space using the PSF:

$$\tilde{\mathbf{d}} = F^H M^H \mathbf{d} = (\mathrm{PSF} \cdot C)\, \boldsymbol{\rho} = \tilde{C}\, \boldsymbol{\rho}, \tag{2.10}$$

where the PSF and the SENSE operator have been merged into $\tilde{C}$ ($N_c N_p \times N_p$). Regular Cartesian undersampling as presented in Fig. 2.2b results in a dirac-comb PSF in image-space, which produces small numbers of regularly spaced aliasing pixels $\boldsymbol{\rho}_\mathsf{r}$ and thereby disentangles the linear system $\tilde{C}$ into small problem subsets with matrix $\tilde{C}_\mathsf{r}$ [12]:

$$\tilde{\mathbf{d}}_\mathsf{r} = \tilde{C}_\mathsf{r}\, \boldsymbol{\rho}_\mathsf{r}. \tag{2.11}$$

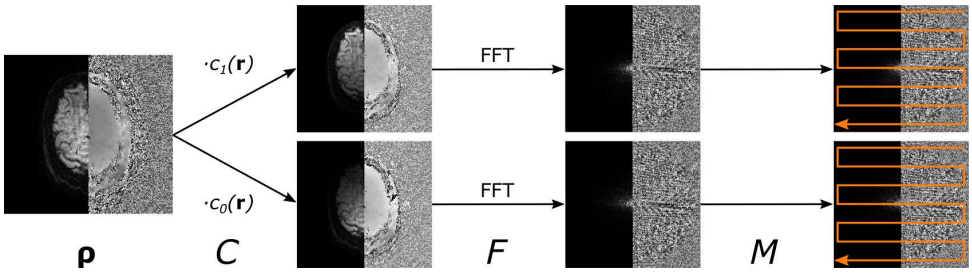


Figure 2.3: Visualization of the SENSE forward model with two coils. The image $\boldsymbol{\rho}$ is encoded by the sensitivity operator C, the Fourier operator F and the sampling operator M (trajectory indicated by orange arrows).

Image reconstruction

The image subproblems based on a weighted $\ell 2$-norm penalty are independently solved using the pseudo-inverse [12]:

$$\boldsymbol{\rho}_{\mathsf{r}} = (\tilde{C}_{\mathsf{r}}^{H}\,\Psi^{-1}\,\tilde{C}_{\mathsf{r}})^{-1}\,\tilde{C}_{\mathsf{r}}^{H}\,\Psi^{-1}\,\tilde{\mathbf{d}}_{\mathsf{r}}, \tag{2.12}$$

with the noise covariance matrix Ψ for SNR-optimal reconstruction, which is obtained in a noise pre-scan [11]. The SNR implications [12] can be summarized for the aliasing pixel a in the aliasing group r as:

$$\mathrm{SNR}_{\mathsf{r},\mathsf{a}}^{\mathrm{red}} = \frac{\mathrm{SNR}_{\mathsf{r},\mathsf{a}}^{\mathrm{full}}}{g_{\mathsf{r},\mathsf{a}}\sqrt{R}}. \tag{2.13}$$

The factor $\sqrt{R}$ in the denominator takes into account the reduced number of samples for accelerated acquisitions with reduction factor R. The g-factor $g_{\mathsf{r},\mathsf{a}}$ quantifies the local noise propagation properties depending on the coil geometry and the k-space trajectory-induced aliasing:

$$g_{\mathsf{r},\mathsf{a}} = \sqrt{(\tilde{C}_{\mathsf{r}}^{H}\,\Psi^{-1}\,\tilde{C}_{\mathsf{r}})_{\mathsf{a},\mathsf{a}}^{-1}\quad(\tilde{C}_{\mathsf{r}}^{H}\,\Psi^{-1}\,\tilde{C}_{\mathsf{r}})_{\mathsf{a},\mathsf{a}}} \geq 1. \tag{2.14}$$

The g-factor is greater or equal to one and represents an important aspect of coil designs as it expresses the capability to disentangle a pixel in a certain location from its aliases.

The coil sensitivity data can either be obtained from a pre-scan [12] or from calibration data [14, 59] within an image dataset. Sensitivity estimation from a pre-scan avoids changes to the imaging sequence and is acquired once in advance. Data-inherent calibration, on the other hand, is performed by sampling additional calibration data during the acquisition window. This technique might capture sensitivity changes due to, for example, motion or field drifts, but interferes with the sampling design.

In case of non-Cartesian sampling or motion-induced trajectory deviations, the effective PSF has a more complex shape in image-space and does not separate as for regular Cartesian undersampling. A pseudo-inverse solution for the full linear model in Eq. 2.9 is generally intractable. The $\ell 2$-norm data discrepancy is normally solved by methods like conjugate gradients (CG) [24] as for CG-SENSE [13]. As a preconditioner for the CG method, an intensity correction can be applied to normalize the pixel-wise sensitivities of all coils and enable faster convergence of the optimization. The intensity correction operator $\tilde{I}$ is constructed from the coil sensitivities by [13]:

$$(\,\tilde{I}\,)_{\mathsf{p},\mathsf{p}} = \frac{1}{\sqrt{\sum_{\mathsf{c}}|c_{\mathsf{c}}(\mathbf{r}_{\mathsf{p}})|^2}}, \tag{2.15}$$

for each pixel p. The diagonal operator is then integrated into the normal equations for the CG method with the forward model $A = MFC$ from Eq. 2.9 as:

$$(\tilde{I}\,A^{H}\,A\,\tilde{I})\,(\tilde{I}^{-1}\,\boldsymbol{\rho}) = \tilde{I}\,A^{H}\,\mathbf{d}, \tag{2.16}$$

The intensity corrected algorithm effectively solves for $(\tilde{I}^{-1}\,\boldsymbol{\rho})$ and the final image is thus recovered by: $\boldsymbol{\rho} = \tilde{I}\,(\tilde{I}^{-1}\,\boldsymbol{\rho})$. In practice, such algorithms involve some kind of regularization like weighted Tikhonov regularizations [24], smoothness-enforcing total variation (TV) regularizations [24] or sparsity-enforcing $\ell 1$-norm regularizations as used for compressed sensing [45].

Simultaneous Multi-slice Encoding

Simultaneous multi-slice (SMS) acquisitions excite multiple slices at the same time using specialized RF pulses. The signals of the slices are superimposed in the acquired data through the joint SMS reception. The simultaneous slice package must therefore be disentangled in a dedicated reconstruction. This subsection describes the design of SMS pulses, its sequence integration and SENSE-based reconstruction approaches.

The selective excitation of one slice through a single-band RF pulse was described in Eq. 2.4. To excite multiple slices simultaneously, the corresponding single-band pulses are added to a multi-band (MB) RF pulse [54]. The complex-valued MB pulse $\mathrm{RF}_{\mathrm{MB}}$ is commonly separated into the RF waveform W_{MB}, which describes the excitation waveform for each of the N_{MB} individual slices, and a slice-specific phase modulation term:

$$\mathrm{RF}_{\mathrm{MB}}(t) = W_{\mathrm{MB}}(t) \quad \cdot \sum_{l=0}^{N_{MB}-1} e^{j\omega_l t + j\varphi_l^{RF}}. \tag{2.17}$$

The circular frequency ω_l accounts for the spatial offset z_l of slice l from the iso-center via $\omega_l = \gamma G_z z_l$. The phase offsets φ_l^{RF} are slice-specifically tunable to reduce the peak RF power [60] of the total SMS RF pulse. As mentioned for the single-band case, the ideal sinc waveform is replaced by filtered finite waveforms in practice. The single-slice excitation represents the special case with $N_{MB} = 1$ and $\varphi_l^{RF} = 0$.

SMS is usually considered in conjunction with controlled aliasing in parallel imaging (CAIPI) schemes for improved g-factor properties. CAIPIRINHA [61, 62] uses dedicated RF phase modulations [54] to impose slice-specific linear phase modulations in phase encoding direction, which translate to different slice shifts in image space and improve the spatial coil encoding. This formalism has been adopted for EPI by blipped-CAIPI [63] using gradient blips in the slice encoding direction instead of RF encodings.

An extension to blipped-CAIPI SMS [64] requires two adaptions of the spin-echo sequence. First, the single-band RF pulses are replaced by multi-band pulses according to Eq. 2.17 with equally spaced slices. Second, CAIPI blips Δk_z are added to the slice encoding during the EPI. An adapted sequence is visualized in Fig. 2.4.

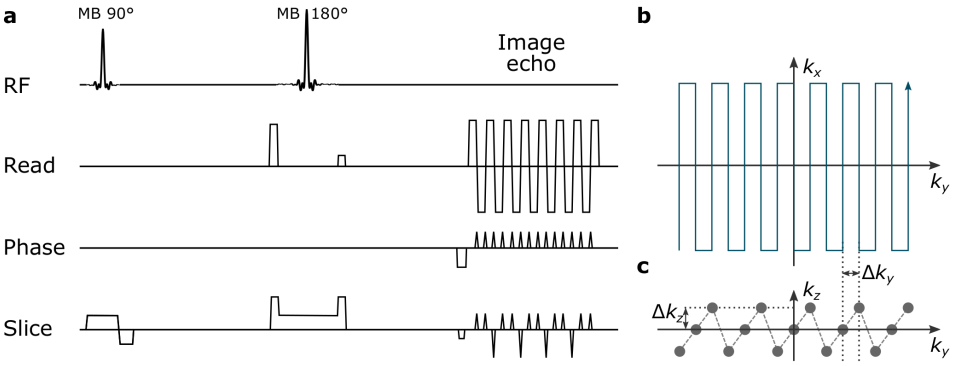


Figure 2.4: EPI sampling overview with SMS extension. (a) Spin-echo sequence with EPI and blipped-CAIPI SMS sampling. The RF pulses are changed to multi-band pulses adding phase modulated sinc-like waveforms to excite or refocus multiple slices. Slice gradient blips Δk_z in the image-echo EPI acquisitions produce individual slice shifts that significantly improve the coil encoding. The slice encoding is shown with a FOV/3 shift between adjacent slices. (b, c) k_x-k_y and k_y-k_z trajectories.

SMS-adapted SENSE reconstructions need to manage the additional encoding dimension, slice shifts and slice undersampling patterns [54]. The extra dimension can either be handled by implementing a 2D-SENSE approach [65] or by artificially concatenating the slice dimension in phase encoding direction (1D-SENSE). The slice shifts are integrated by adapting the PSF or simply shifting the coil sensitivity maps in the reconstruction. Zahneisen et al. [66, 67] embedded the SMS sampling (and undersampling) into a pseudo-3D k-space formalism, which connects blipped-CAIPI SMS to the SENSE signal equation (Eq. 2.8) and enables efficient SENSE algorithms.

SMS with blipped-CAIPI encoding has found widespread use in DWI, because it allows for a direct reduction of the number of excitations by the MB factor at only minor g-factor penalty [54]. In comparison, in-plane acceleration in the first place only reduces the trajectory length and an omission of excitations is penalized by a drastically reduced SNR. Drawbacks of SMS include an increased susceptibility to off-resonances, slice leakage artifacts arising from incomplete unaliasing in a SMS group, and potential slice-wise Nyquist ghosting effects.

2.2

DWI AND MULTI-SHOT MODELING

The diffusion contrast in MRI lends singular, non-invasive insights into the diffusion mechanisms of living subjects [4]. The contrast characteristics contain valuable information about the tissue micro-structures that are of particular interest for both clinical and research domains. DWI has been dominated for many years by single-shot acquisition techniques, which are robust but limited in resolution and corrupted by strong geometric distortions. Multi-shot DWI offers potentials to overcome these limitations, but the segmented acquisition, on the other hand, is by design prone to shot-wise signal variations, which have to be taken into account by appropriate modeling. This section starts with an introduction to diffusion-weighted imaging, followed by a description of the state-of-the-art modeling for the diffusion-related shot variations. Next, the developed models are combined to the multi-shot reconstruction problem. Finally, a literature review on related algorithms is provided.

DIFFUSION-WEIGHTED IMAGING

DWI uses strong diffusion-sensitizing gradients [6] to probe the Brownian motion of the spins within the tissue anatomies, obtaining the diffusion contrast with valuable structural information at the scale of 10 μm [4]. This remarkable method for micro-structural imaging has found widespread application in both neurological research and clinical practice. In neurology, DWI is used as a non-invasive technique to study in-vivo brain function and connectivity. At the same time, the diffusion contrast has become an indispensable part of the diagnostics for cancer, stroke and multiple sclerosis. A recent review of applications and technical developments was provided by Miller and Wu [8].

Stejskal-Tanner diffusion encoding

Diffusion-weighted images are normally acquired using the Stejskal-Tanner sequence [6]. The RF pulses and gradient encodings of such a sequence are visualized in Fig. 2.5a. Starting with the basic sequence, the slice-selective excitation (90°) and the first refocusing (180°) RF pulse form a spin echo at the echo time TE_{img}. Two identical diffusion gradients, placed before and after the 180° pulse, sensitize the sequence to tiny displace-

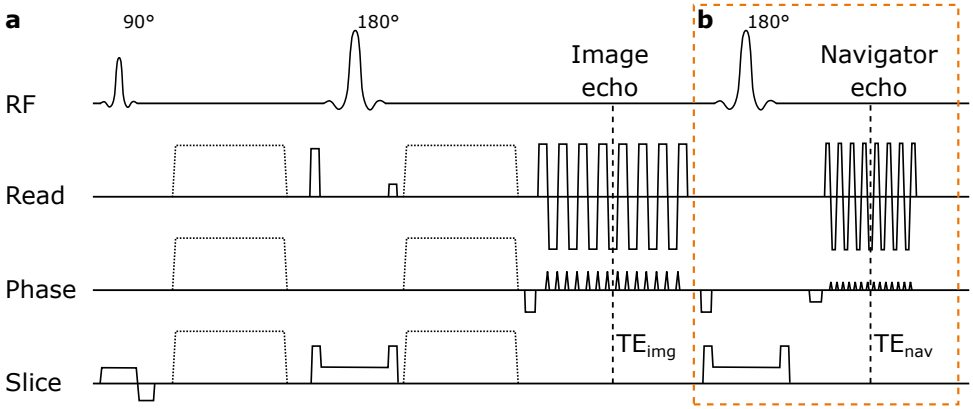


Figure 2.5: Stejskal-Tanner diffusion sequence with EPI sampling and an optional navigator echo. a: Basic diffusion-weighted spin-echo sequence. A slice-selective RF excitation pulse (90°) and a refocusing pulse (180°) form a spin-echo at the center of the image echo sampling window. The strong diffusion-sensitizing gradients (dotted lines) before and after the refocusing pulse introduce the diffusion-weighting of the signal. b: A second 180° pulse optionally refocuses the signal and forms another spin-echo for a low-resolution navigator acquisition. The sequence uses echo-planar imaging (EPI) for image and navigator sampling. Please note that the sequence supports segmented image echo acquisitions by an appropriately designed sampling.

ments of the magnetization along the gradient direction. The spatially linear phase accrual of the diffusion gradients cancels out exactly under ideal static conditions, while the phase reversal remains incomplete for protons that have moved along the gradient direction. Averaged over a voxel, the stochastic Brownian motion of the magnetization results in a phase dispersion leading to a signal attenuation that depends on the local diffusion characteristics [6]. The image sampling of the diffusion-weighted image data is centered at the echo time TE_{img}.

A common extension introduces a second refocusing pulse after the image echo to acquire a low-resolution navigator at echo time TE_{nav} [31] as shown in Fig. 2.5b. The additional signal provides valuable image information at a coarse resolution especially if the image echo is undersampled. Such navigated acquisitions have been investigated to overcome motion-related artifacts by model-based reconstructions [28, 31, 44, 50].

Diffusion tensor model

The Stejskal-Tanner equation characterizes the diffusion weighting of each voxel as an exponential decay governed by a three-dimensional apparent diffusion tensor and the diffusion encoding properties [6]. Let $\mathsf{Q} = \{1, ..., N_\mathsf{q}\}$ be the set of N_q diffusion-weighted experiments. The signal $\rho_\mathsf{q}(\mathbf{r})$ for a diffusion experiment $\mathsf{q} \in \mathsf{Q}$ is related to the non-diffusion weighted signal $\rho_0(\mathbf{r})$ by:

$$\rho_\mathsf{q}(\mathbf{r}) = \rho_0(\mathbf{r})\, e^{-b_\mathsf{q}\, \mathbf{g}_\mathsf{q}^T\, \mathbf{D}(\mathbf{r})\, \mathbf{g}_\mathsf{q}}. \tag{2.18}$$

For in-vivo experiments, the phase of ρ_q and ρ_0 are normally dropped to reduce the influence of physiological disturbances [6]. The apparent diffusion tensor $\mathbf{D}$ (3×3), measured in mm^2/s, is sampled in the direction of the unit diffusion gradient vector $\mathbf{g}_\mathsf{q}$ and is weighted by the diffusion encoding strength b_q, called b-value, with units s/mm^2.

The scalar diffusion encoding strength b_q and the unit diffusion direction vector $\mathbf{g}_\mathsf{q}$ span a three-dimensional diffusion-weighting space, which is called the *q-space* [8]. DWI scans normally use multiple diffusion encodings to characterize the 3D diffusion properties of each voxel. Joint samplings of both the spatial and diffusion dimensions are rather time-consuming, but the DWI contrast variations in the q-space are rather smooth. Models like the mono-exponential diffusion tensor model in Eq. 2.18 formulate the low-dimensional correlations through the symmetric tensors, which require only six voxel parameters to describe the q-space relations. Nevertheless, it should be noted that the tensor model is a relatively simple approximation to real-world diffusion characteristics and there are many extensions to this basic model using, for example, generalized tensor concepts or multiple tissue compartments within each individual voxel [6]. Clinical diffusion experiments are usually sampled in multiple q-space directions with one fixed b-value, which is called *single-shell* acquisition. In contrast, *multi-shell* scans acquire multiple shells and model relations among them [8].

Clinical diffusion-weighted imaging is often acquired in three orthogonal diffusion directions with a constant b-value b. Note that the spatial dependence of $\mathbf{r}$ is dropped here for the following pixel-specific measures. The three acquisitions probe the diagonal of the diffusion tensor $\mathbf{D}$. The *mean diffusivity* D_{avg} is a basic isotropic DWI measure connected to the trace of the apparent diffusion tensor $\mathbf{D}$:

$$D_{avg} = \text{trace}(\mathbf{D})/3. \tag{2.19}$$

For three orthogonal measurements, D_{avg} is just the average apparent diffusion coefficient (ADC) of the three directions. The geometric mean of the (magnitude) images is called *isotropic diffusion-weighted image* ρ_{iso} and represents an image with averaged attenuation of ρ_0 according to the mono-exponential DTI model in Eq. 2.18:

$$\rho_{iso} = \left(\prod_{\mathsf{q}\in\mathsf{Q}} |\rho_\mathsf{q}| \right)^{|\mathsf{Q}|^{-1}} = |\rho_0| \; e^{-b\,|\mathsf{Q}|^{-1} \sum_{\mathsf{q}\in\mathsf{Q}} \mathbf{g}_\mathsf{q}^T \,\mathbf{D}\, \mathbf{g}_\mathsf{q}}. \tag{2.20}$$

The isotropic DWI is connected to the mean diffusivity by $\rho_{iso} = |\rho_0| \; e^{-bD_{avg}}$ for three orthogonal measurements.

In diffusion tensor imaging (DTI), the symmetric diffusion tensor $\mathbf{D}$ with 6 unknown parameters per voxel is estimated from the DWIs in Q and at least one non-diffusion-weighted image from Q_0 by standard regularized regression [6]. The diffusion ellipsoids corresponding to the symmetric tensors are typically analyzed using an eigenvalue decomposition to derive measures of anisotropy, which are related to the local micro-anatomy within the tissue.

The fractional anisotropy (FA) is a common measure to quantify the normalized variance of the tensor eigenvalues λ_i for $i = \{1,2,3\}$:

$$\text{FA} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1-\bar{\lambda})^2 + (\lambda_2-\bar{\lambda})^2 + (\lambda_3-\bar{\lambda})^2}{\lambda_1^2+\lambda_2^2+\lambda_3^2}}, \tag{2.21}$$

with the mean eigenvalue $\bar{\lambda}$. The values range from zero to one. An FA measure of zero means isotropic diffusion, whereas one indicates completely anisotropic characteristics.

FA is commonly visualized in colored images, where the anisotropy measure controls the voxel's brightness and the color is adapted to the main diffusion direction, which is the eigenvector associated with the highest eigenvalue. The RGB color-code is red for left-right, green for anterior-posterior and blue for feet-head vector components.

This work is focused on single-shell DWI and DTI models for image reconstruction. The DTI evaluations in this work are performed using the *Dipy* library for python [68]. In the following explanations, the diffusion experiment subscript q is dropped, when the formulations are equally valid for all diffusion directions, as for general image encoding techniques.

MODELING OF SHOT VARIATIONS

Variations of the signal density over the scan time may be caused by magnetic field fluctuations or motion. The impact of the variations onto the signal acquisition depends on its effective interaction with the MR sequence and may lead to image artifacts or even irreversible signal loss. This work focuses on motion-related effects, although the general approach can be adopted for other signal variations. For further background on the multifaceted effects of motion on the MR signal acquisition, a review by Zaitsev et al. [22] is recommended.

Model-based reconstructions for multi-shot imaging require a shot-specific *signal model* describing the data formation and a *shot model* connecting the individual shot datasets. The shot-specific signal equation is based on the two established encoding engines of MRI: the Fourier encoding of the MR gradient system and the coil sensitivity encoding provided by phased arrays. As for single-shot imaging, it is commonly assumed that the violation of the encoding model by the signal variations remains negligible.

The signal equation from Eq. 2.8 can then be adopted for multi-shot imaging, relating the shot image ρ_s of shot s to the data $d_{\mathsf{s},\mathsf{c}}$ with the set of sampling positions K_s:

$$d_{\mathsf{s},\mathsf{c}}(\mathbf{k}) = \int_{\mathrm{VOI}} \rho_\mathsf{s}(\mathbf{r})\, c_\mathsf{c}(\mathbf{r})\, e^{-j\mathbf{kr}}\, d\mathbf{r}, \quad \mathbf{k} \in \mathsf{K}_\mathsf{s}. \tag{2.22}$$

A discrete realization of the sampling is represented by a linear encoding model with the shot data vector $\mathbf{d}_\mathsf{s} \in \mathbb{C}^{N_c N_k}$ and the shot image vector $\boldsymbol{\rho}_\mathsf{s} \in \mathbb{C}^{N_p}$:

$$\mathbf{d}_\mathsf{s} = M_\mathsf{s} F C \boldsymbol{\rho}_\mathsf{s}. \tag{2.23}$$

Next, the *shot model* describes how the shot images ρ_s for $\mathsf{s} \in \mathsf{S}$ are related to the underlying *joint image* ρ. S is the set of shots. The trivial approach, which ignores shot variations completely, corresponds to the identity mapping $\rho_\mathsf{s}(\mathbf{r}) = \rho(\mathbf{r})$ for all $\mathsf{s} \in \mathsf{S}$ and results in a standard SENSE encoding model *without* further shot-specific parameters.

In the following, two types of shot variations, which are important for multi-shot DWI and which are addressed in this work, are presented in detail: shot phase variations and inter-shot macroscopic motion. Figure 2.6 visualizes both effects for a single-slice multi-shot experiment. The datasets of the four shots are sampled with different EPI interleaves (Fig. 2.6b) and are subject to varying phase maps and macroscopic motion states (Figs. 2.6a and c). The trivial SENSE reconstruction, which is blind to the shot variations, results in severe ghosting and blurring as shown in Fig. 2.6d.

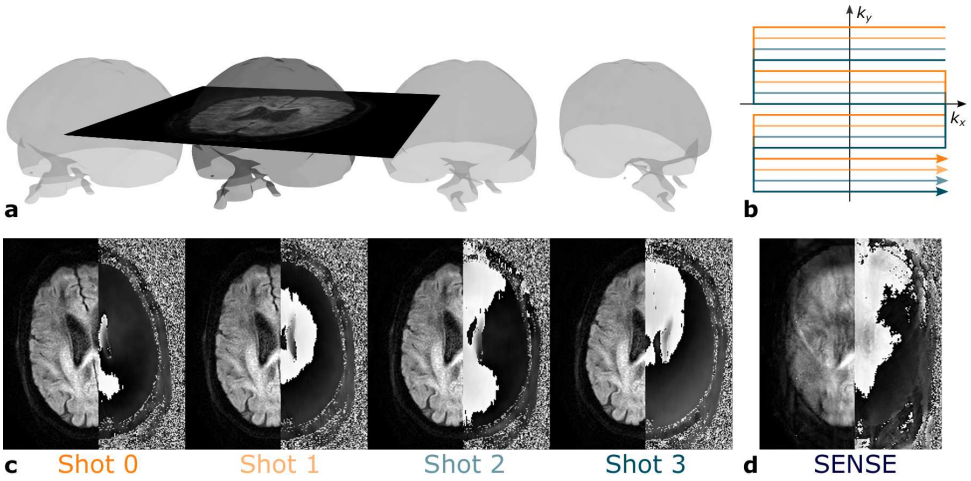


Figure 2.6: Example of the shot variations for a single-slice 4-shot DWI case. a: Visualization of a single-slice acquisition under 3D shot-to-shot bulk motion. b: Multi-shot sampling with 4 interleaves. c: Shot images estimated from motion-corrupted in-vivo data showing continuous in-plane rotations over the shots and significant shot phase variations. The images are separated by half into magnitude and phase (left/right) to emphasize and illustrate again the complex-valued nature of the MR signals. d: SENSE reconstruction of all shots ignoring the multi-shot variations. The shot phase variations cause ghosting artifacts, whereas the macroscopic motion leads to a geometric shot signal mismatch and blurring of the image structures.

Shot-to-Shot phase variations

Shot-to-shot phase variations arise mainly from tiny patient motion [17] and cardiac pulsation of the brain [18, 19] during the diffusion encoding. Such inevitable motion effects influence the phase accrual and lead to an incomplete shot-specific phase reversal of the strong diffusion-sensitizing gradients. General nonlinear shot phase variations can be incorporated into the model as shot phase maps ϕ_s with shot index s [18, 32]:

$$\rho_s(\mathbf{r}) = \rho(\mathbf{r})\, e^{j\phi_s(\mathbf{r})}, \tag{2.24}$$

which translates to a diagonal shot phase matrix $\Phi_s = \mathrm{diag}(e^{j\boldsymbol{\phi}_s})$ per shot s:

$$\boldsymbol{\rho}_s = \Phi_s\, \boldsymbol{\rho}. \tag{2.25}$$

Each shot phase operator Φ_s ($N_p \times N_p$) generally involves N_p real-valued phase parameters and, thus, tremendously increases the number of unknown variables. Smoothness assumptions [18, 28] are commonly imposed to reduce the number of effective phase variables balancing phase accuracy with SNR. Nevertheless, the assumption remains questionable as there are high-frequent phase components in MRI signals [34].

Macroscopic shot motion

Patient gross motion between the shots involves all kinds of geometric transformations happening to the image during the acquisition. The time between two shots of the same anatomy, which is related to the repetition time TR, is on the order of seconds and thus considerably longer than the sequence itself, which is about 100 ms. Therefore, macroscopic shot motion is commonly modeled as inter-shot motion with one motion state per shot and intra-shot motion during the fast acquisition is neglected. The inter-shot motion is described by a coordinate transformation $T_{\Omega,s}$ giving $\rho_s(\mathbf{r}) = \rho(T_{\Omega,s}(\mathbf{r}))$.

For DWI, rigid rotational components, described by the transformation T_R, additionally change the diffusion gradient direction $\mathbf{g}$ in the patient frame [43] by $T_R^{-1}(\mathbf{g})$. The diffusion tensor model in Eq. 2.18 can be used to correct for potential contrast variations, based on a diffusion tensor map. For a shot s, the joint image ρ is thus affected by the geometrical coordinate transformation $T_{\Omega,\mathsf{s}}$ and the diffusion contrast-reweighting according to the directional variation $\tilde{\mathbf{g}}_\mathsf{s} = T_{R,\mathsf{s}}^{-1}(\mathbf{g}) - \mathbf{g}$:

$$\rho_\mathsf{s}(\mathbf{r}) = \rho(T_{\Omega,\mathsf{s}}(\mathbf{r}))\; e^{-b\,\tilde{\mathbf{g}}_\mathsf{s}^T\,\mathbf{D}(T_{\Omega,\mathsf{s}}(\mathbf{r}))\,\tilde{\mathbf{g}}_\mathsf{s}}. \tag{2.26}$$

For discrete realizations, this can be included using matrix formulations [69]:

$$\boldsymbol{\rho}_\mathsf{s} = \Omega_\mathsf{s}\,\Upsilon_\mathsf{s}\,\boldsymbol{\rho}, \tag{2.27}$$

with the macroscopic motion operator Ω_s ($N_p \times N_p$) and the diffusion contrast-reweighting operator Υ_s ($N_p \times N_p$).

A numerical implementation for macroscopic inter-shot motion requires the definition of a motion model for the coordinate transformation $T_{\Omega,\mathsf{s}}$ and an interpolation scheme. The number of shot parameters depends on the selected motion model, ranging from 3 parameters per shot for 2D rigid in-plane motion, over 12 parameters for 3D affine motion up to pixel-wise deformation vectors for elastic transformations. Head motion is well captured by rigid motion models due to the fixation of the skull, whereas the motion profiles might be generally non-rigid for other anatomies, requiring affine or elastic transformations. Besides the motion transformation, the diffusion contrast correction requires a spatial map of symmetric tensors giving rise to six dimensions per voxel. The tensor map is not shot-specific but determines the effective contrast according to the shot-specific rotation parameters.

Joint shot model

The joint signal model for multi-shot DWI includes both the shot phase variations and macroscopic inter-shot motion:

$$\rho_\mathsf{s}(\mathbf{r}) = \rho(T_{\Omega,\mathsf{s}}(\mathbf{r}))\; e^{-b\,\tilde{\mathbf{g}}_\mathsf{s}^T\,\mathbf{D}(T_{\Omega,\mathsf{s}}(\mathbf{r}))\,\tilde{\mathbf{g}}_\mathsf{s}}\; e^{j\phi_\mathsf{s}(T_{\Omega,\mathsf{s}}(\mathbf{r}))}, \tag{2.28}$$

which relates to a linear shot model for a discrete realization:

$$\boldsymbol{\rho}_\mathsf{s} = \Phi_\mathsf{s}\,\Omega_\mathsf{s}\,\Upsilon_\mathsf{s}\,\boldsymbol{\rho}. \tag{2.29}$$

Hence, the shot model requires estimating shot-specific phase maps for the operator Φ_s, shot-specific motion parameters for Ω_s, and pixel-wise diffusion tensors along with the shot-specific rotation parameters for Υ_s.

Further signal variations

The model formulation is generally based on two assumptions. First, the *shot model* is assumed to sufficiently describe the relation of the shot data to one individual, consistent joint image. Second, the *signal encoding model* remains unaffected and valid, apart from the modeled effects. The two models thus respectively constitute consistency of the image object and the image encoding.

Besides shot phases and macroscopic motion, there are several other signal variations that can be encountered by dedicated model-based reconstructions [23]. Among those are B_0-field inhomogeneities, which affect the signal phase evolution especially for long readouts leading to geometric distortions in phase encoding direction for EPI or blurring for spiral sampling. The resulting image artifacts can be mitigated by dedicated acquisitions like read-out segmented EPI [29] or by integrating B_0-map estimates into the reconstruction [70, 71]. Other approaches address multiple chemical shift-encoded species [71], intra-shot rigid motion [72] or relaxation parameter estimation [73].

Regarding the signal encoding model, model-based reconstructions may include sampling trajectory deviations [74] or unknown coil sensitivities [75]. Nyquist ghosting from odd/even echo inconsistencies for EPI is normally corrected in advance based on a reference scan [8], but can also be addressed by a self-navigated reconstruction [76].

MULTI-SHOT RECONSTRUCTION PROBLEM

The signal encoding and shot models jointly describe the relation between the multi-shot data and the underlying image. This subsection formulates the multi-shot DWI reconstruction problem and categorizes strategies to perform the optimization.

Joint multi-shot DWI problem

The multi-shot problem with shot phase and macroscopic motion variations [43] is based on the data consistency of the shot-wise linear signal models in Eq. 2.23 and enforces similarity through the shot models in Eq. 2.29:

$$\begin{aligned} &\text{minimize} && \frac{1}{2}\sum_{\mathsf{s}\in\mathsf{S}'} \|M_\mathsf{s} F C\, \boldsymbol{\rho}_\mathsf{s} - \mathbf{d}_\mathsf{s}\|_2^2 \\ &\text{subject to} && \boldsymbol{\rho}_\mathsf{s} = \Phi_\mathsf{s}\,\Omega_\mathsf{s}\,\Upsilon_\mathsf{s}\,\boldsymbol{\rho}, \quad \mathsf{s}\in\mathsf{S}', \end{aligned} \tag{2.30}$$

where the joint image $\boldsymbol{\rho}$ and the shot parameters in Φ_s, Ω_s and Υ_s are the optimization variables. The adapted *set of included shots* S' allows to reject deteriorated shot datasets from the optimization remaining with $N'_{shots} = |\mathsf{S}'|$ included shots.

Substituting the shot models for $\boldsymbol{\rho}_\mathsf{s}$ into the signal model yields an unconstrained formulation that integrates the shot variations as part of the image encoding:

$$\underset{\boldsymbol{\rho},\Phi_\mathsf{s},\Omega_\mathsf{s},\Upsilon_\mathsf{s}}{\text{minimize}} \quad \frac{1}{2}\sum_{\mathsf{s}\in\mathsf{S}'} \|M_\mathsf{s} F C \Phi_\mathsf{s}\,\Omega_\mathsf{s}\,\Upsilon_\mathsf{s}\,\boldsymbol{\rho} - \mathbf{d}_\mathsf{s}\|_2^2. \tag{2.31}$$

For a compact notation, the shot forward models can be integrated into a multi-shot forward model with the stacked multi-shot multi-coil data vector $\grave{\mathbf{d}} \in \mathbb{C}^{N'_{shots} N_c N_k}$:

$$A\,\boldsymbol{\rho} = \grave{\mathbf{d}}, \tag{2.32}$$

and the forward operator $A = \grave{M}\grave{F}\grave{C}\grave{\Phi}\grave{\Omega}\grave{\Upsilon}\,P$ $(N'_{shots} N_c N_k \times N_p)$. The grave accents indicate multi-shot operators, which embed the shot operators into block diagonal matrices. The *joint shot phase operator* $\grave{\Phi}$, for example, contains the individual *shot phase operators* Φ_s with $\mathsf{s}\in\mathsf{S}'$ on the diagonal and zeros off the diagonal. The shot copy operator P $(N'_{shots} N_p \times N_p)$ vertically stacks N'_{shots} identity matrices I_{N_p} providing a copy of the joint image $\boldsymbol{\rho}$ for each shot in S'.

By this, the optimization problem can be compactly rewritten as:

$$\underset{\boldsymbol{\rho},\grave{\Phi},\grave{\Omega},\grave{\Upsilon}}{\text{minimize}} \qquad \frac{1}{2}\,\|\grave{M}\grave{F}\grave{C}\grave{\Phi}\grave{\Omega}\grave{\Upsilon}\,P\boldsymbol{\rho}-\grave{\mathbf{d}}\|_2^2. \tag{2.33}$$

Typically, a regularization term is added to the functional to include some form of prior knowledge about the image or shot parameters and thereby improve the problem conditioning [23]. Common choices are variations of the Tikhonov regularization, smoothness, sparsity and total variation constraints. Recently, deep learning methods have been studied as tools to learn useful representations of the data and support [47] or even replace [46] the reconstruction.

The joint optimization of the image and shot parameters is demanding due to the non-convexity of the problem and the large problem size. The metric landscape for macroscopic motion and phase parameters is subject to local minima and requires careful optimization [77]. Moreover, even relatively small datasets with 256 × 256 single-slice image size, 16 coils and 4 shots lead to forward models exceeding the size of 1 million × 65,000. Such matrices are too large to be densely stored in memory and require exploiting the sparse problem structure.

Strategies for solving the multi-shot DWI problem

If the shot parameters are known, Eq. 2.31 represents a convex but still large linear reconstruction problem for the joint image $\boldsymbol{\rho}$. For Cartesian sampling, the aliasing pattern reduces to small pixel groups, for which the pseudoinverse involved in the SENSE reconstruction (Eq. 2.12) is feasible. Nevertheless, macroscopic motion generally introduces deviations from the intended sampling and requires dedicated iterative algorithms to unfold the resulting aliasing pattern. Some gradient-based algorithms allow to perform the optimization requiring only matrix-vector products of an image estimate with the forward model and its adjoint [23]. The algorithms can then easily exploit diagonal or other sparse matrix structures and use the FFT. The convergence properties thereby largely depend on the condition number, which is the ratio of the largest and the smallest singular value of the forward model [24]. Common iterative algorithms are the conjugate gradient (CG) method [24] used in CG-SENSE [13] and proximal methods like the fast iterative shrinkage-thresholding algorithm (FISTA) [53] used for compressed sensing algorithms [45].

Shot parameter estimation strategies are often divided into extra-navigated, navigated, self-navigated, and navigator-free algorithms depending on the data that is used for this step [39, 40]. *Extra-navigated* techniques derive the parameters from an external device, such as an optical camera for motion tracking [21]. This strategy can be implemented with extra equipment independent of the MR scanner and does not need sequence changes. *Navigated* techniques use additional MR signals for this and, thus, interfere with the MR sequence but without requiring extra hardware. *Self-navigated* methods estimate the shot parameters from the image data itself exploiting data redundancies, which are provided by parallel imaging. *Navigator-free* techniques use problem reformulations without an explicit estimation of the shot parameters. The latter can, for example, be achieved by incorporating the shot models through low-rank constraints into the optimization [78]. By this, some non-convex problems can be converted into

equivalent convex formulations, but not every problem is expressible in this way. Self-navigated and navigator-free methods are considered *data-driven* and often involve demanding iterative algorithms. At the same time, these methods do not need extra equipment, scan time or sequence changes.

Image correction strategies are furthermore separated into prospective and retrospective methods. The latter estimate the modeled variations and perform the image correction as an off-line step after the acquisition. Conversely, prospective methods use the shot parameter estimates to update the sequence in real-time and directly acquire uncorrupted data. The immediate correction during the scan poses on-line demands for the shot parameter estimation and not every variation can be implemented prospectively on the scanner, i.e. bulk motion is representable by 3D MR gradient changes, while shot phases are hard to predict and almost uncorrectable by the scanner encoding.

MULTI-SHOT DWI ALGORITHMS

The variety of optimization algorithms, regularizations, navigation schemes, and encoding technologies has been continuously growing over the last decades along with the increasing performance of the available computers, pushing forward the role of model-based image reconstruction [23]. This subsection is based on the review of recent data-driven multi-shot DWI algorithms from the published article in Ref. [49], which is extended to extra-navigated and navigated methods here.

The literature provides a wide range of multi-shot DWI methods employing different strategies to solve the optimization. In general, all algorithms correct for phase variations from physiological motion due to its severe and unavoidable nature in multi-shot DWI. Conversely, macroscopic motion is mainly neglected. Table 2.1 gives an overview of key methods for shot phase and macroscopic motion modeling in multi-shot DWI.

Shot phase corrected algorithms

Shot phase corrections for DWI reach back almost 30 years and started with navigated methods using 1D navigator projections in readout direction [17]. Later both in-plane directions were addressed [25] to estimate zeroth and first order phase terms, which are mainly caused by small rigid body motion. The 1D approaches were outperformed by 2D navigators using low-resolution spiral [26] and EPI [27] scans. In 2003, Miller and Pauly extended the model to nonlinear phase maps from 2D navigators to address pulsatile brain motion from cardiac activity and to remove the time-consuming cardiac triggering [18]. Subsequent navigated techniques embedded the nonlinear phase variations into a multi-coil formulation [28], which yields the shot phase-corrupted multi-shot problem without macroscopic motion:

$$\underset{\boldsymbol{\rho},\Phi_s}{\text{minimize}} \qquad \frac{1}{2}\sum_{s\in S'} \|M_s F C \Phi_s \boldsymbol{\rho} - \mathbf{d}_s\|_2^2. \tag{2.34}$$

With navigated phase operators Φ_s, this optimization problem can be solved using CG. The resulting problem is often interpreted as a virtual sensitivity encoding $C\Phi_s$ with $N_{shots}N_c$ shot phase-adapted sensitivities $c_c(\mathbf{r})e^{j\phi_s(\mathbf{r})}$. In analogy to SENSE, efficient image-space algorithms were published for Cartesian trajectories that solve the reconstruction problem by matrix inversion for the reduced aliasing pixel groups according to the PSF [31].

Table 2.1: Chronological list of key multi-shot DWI methods for shot parameter and/or macroscopic motion estimation. k-space windows or kernels are indicated by a preceding 'k- '. Partial Fourier acquisitions for EPI are marked by 'PF '. Steady-state free precession (*SSFP*) and *spin warp* sampling acquire one k-space line per acquisition, rapid stimulated echo acquisition mode (*STEAM*) acquires multiple phase encoding lines per shot similar to EPI.

Algorithm	**Authors**	**Year**	**Navigation type**	**Alternating structure**	**Motion Correction**		**Sampling trajectory**	**SMS inclusion**
					Phase	**Gross**		
	Anderson and Gore [17]	1994	1D-nav.	-	0. & 1. order k_x	-	Spin warp	-
	Butts et al. [25]	1996	1D-nav.	-	0. & 1. order k_x-k_y	-	EPI	-
	Butts et al. [26]	1997	2D-nav.	-	0. & 1. order k_x-k_y	-	EPI	-
	Atkinson et al. [27]	2000	2D-nav.	-	0. & 1. order k_x-k_y	-	EPI	-
PROPELLER	Pipe et al. [32]	2002	self-nav.	-	k-Triangular	-	Rotated EPI	-
	Miller and Pauly [18]	2003	2D-nav.	-	k-Rectangular	-	SSFP	-
	Liu et al. [33]	2005	self-nav.	-	k-Hamming	-	Spiral	-
	Atkinson et al. [28]	2006	2D-nav.	-	Gaussian	-	EPI	-
	Uecker et al. [34]	2009	self-nav.	-	No filter	-	STEAM	-
IRIS	Jeong et al. [31]	2013	2D-nav.	-	k-Hamming	-	EPI + PF	-
MUSE	Chen et al. [35]	2013	self-nav.	-	TV	-	EPI + PF	-
SENSE+CG	Truong and Guidon [36]	2014	self-nav.	-	Median	-	Spiral	-
SF-MUSE	Zhang et al. [37]	2015	self-nav.	✓	Gaussian	-	EPI + PF	-
	Chang et al. [79]	2015	self-nav.	-	TV	-	EPI + PF	✓
POCSMUSE	Chu et al. [38]	2015	self-nav.	✓	k-Hann	-	EPI + PF	-
	Herbst et al. [80]	2015	extra-nav.	-	TV	3D Rigid	EPI	
POCS-ICE	Guo et al. [39]	2016	self-nav.	✓	k-Triangular	-	Spiral	-
AMUSE	Guhaniyogi et al. [43]	2016	self-nav.	-	TV	2D Rigid	EPI + PF	-
	Dai et al. [64]	2017	2D-nav.	-	Compact k-kernel	-	EPI + PF	✓
	Herbst et al. [81]	2017	extra-nav.	-	TV	3D Rigid	EPI + PF	✓
MUSSELS	Mani et al. [40]	2017	nav.-free	-	Low-rank	-	EPI + PF	-
	Dong et al. [44]	2018	2D-nav.	-	Compact k-kernel	2D Rigid	EPI	-
Shot-LLR	Hu et al. [41]	2018	nav.-free	-	Low-rank	-	EPI + PF	-
SMS-NEATR	Bilgic et al. [77]	2019	self-nav.	-	DL image prior	-	EPI	✓
SMS MUSSELS	Mani et al. [82]	2020	nav.-free	-	Low-rank	-	EPI + PF	✓
SPA-LLR	Hu et al. [42]	2020	nav.-free	-	Low-rank	-	EPI + PF	-

The first self-navigated approaches employed a two-step strategy, obtaining the shot phase maps first and then reconstructing the joint image. Initially, the approaches used the fully sampled central k-space to estimate the low-resolution phase variations for rotated EPIs in PROPELLER [32] and variable density spirals [33]. After that, the algorithms leveraged the data redundancy provided by parallel imaging to reconstruct high-resolution shot images from undersampled data and used them for the same two-step procedure, as done by Uecker et al. [34] and in MUSE [35] and SENSE+CG [36]. These algorithms with SENSE-based self-navigation are limited to relatively low segmentations, because the individual shot reconstructions become increasingly ill-conditioned for higher segmentation. The g-factor noise [12] thus propagates into the shot images deteriorating the phase estimates. Smoothness-enforcing phase filters are used to trade off SNR with artifacts from over-smoothing.

To address the g-factor problem for self-navigated algorithms, SF-MUSE [37], POCSMUSE [38] and POCS-ICE [39] use an *alternating strategy* to repeatedly estimate the shot phase variations and the joint image. POCS-ICE is presented in more detail here, because it serves as a reference method in this work. The algorithm uses projections onto convex sets (POCS) to alternately perform shot image and joint image updates. Initialized by zero, the algorithm performs a shot data projection through the SENSE-based forward model by inserting the measured shot data samples into the current shot image estimates. After eliminating the shot phase by a low-resolution triangular k-space window, the phase-corrected shots are averaged to an updated joint image. The joint image is combined with the shot phase estimates and projected back again to the shot data. The iterative feedback successively supports the shot estimation by an updated joint image enabling higher segmentations.

In contrast, the navigator-free algorithms MUSSELS [40] and Shot-LLR [41] derive compressible matrices with non-empty null spaces from the correlations of the phase-disturbed shot models and thereby achieve a convex formulation for the phase-corrupted multi-shot problem. SPA-LLR [42] is a locally low-rank approach that exploits the q-space compressibility of DTI using the correlations of the diffusion directions. These approaches are navigator-free, because the shot parameters are not explicitly included but enforced through iteratively posed low-rank matrix constraints.

As a new versatile technology for image processing, deep learning (DL) techniques are entering the field of MR image reconstruction during the recent years. A first approach for shot phase-corrected multi-shot DWI was proposed by Bilgic et al. [77], termed network acceleration estimation for tempered reconstruction (NEATR). This method applies a U-Net for image enhancement to a MUSSELS [40] reconstruction and uses the improved joint images for an explicit shot phase estimation by a projected gradient descent method [83]. Finally, a joint image is reconstructed using the updated shot phase estimates by a conventional multi-shot reconstruction method. Hence, this method starts from a navigator-free algorithm and performs deep learning-supported self-navigation for phase-corrected multi-shot DWI.

Joint shot phase and macroscopic motion corrected algorithms

A generalized matrix framework for macroscopic motion corrected reconstructions of multi-shot data was published by Batchelor et al. in 2005 [69] and combined with parallel imaging by Bammer et al. in 2007 [84]. The latter multi-coil approach reconstructs

shot images from undersampled data using SENSE (as for the phase corrections above), determines rigid motion parameters using image registration and performs a macroscopic motion-compensated multi-shot reconstruction.

Image registration algorithms optimize a similarity metric M_{sim} between a fixed image $\rho_{fix}(\mathbf{r})$ and a moving image $\rho_{mov}(T_\Omega(\mathbf{r}))$ to find the optimal motion parameters of the coordinate transformation T_Ω [85]:

$$\underset{\Omega}{\text{minimize}} \quad M_{sim}\big(\rho_{fix}(\mathbf{r}),\, \rho_{mov}(T_\Omega(\mathbf{r}))\big). \tag{2.35}$$

The metric, the interpolation strategy and the optimizer must be adequately designed for the application-specific registration task. The image registration metric is usually different from the data consistency functional employed for image reconstruction.

More recent (non-DWI) algorithms perform *alternating* optimizations of the data consistency functional by repeatedly estimating the joint image with fixed shot parameters and, next, the shot parameters with a fixed joint image [86, 87]. The functional equals Eq. 2.31 without the diffusion-related shot phase operator Φ_s and shot contrast operator Υ_s. The shot motion parameter update is performed by Newton's method, while the image reconstruction uses CG. Nevertheless, macroscopic motion estimation in the diffusion realm is additionally complicated by the low SNR, the diffusion-related phase and contrast variations, and the multi-contrast nature between diffusion- and T2-weighted images as well as amongst the diffusion directions themselves. Therefore, DWI requires specialized solutions to address the macroscopic motion problem.

For self-navigated multi-shot DWI, AMUSE [43] extends the two-step MUSE framework for shot phase-corrected reconstructions [35] by macroscopic motion and the diffusion contrast corrections. The algorithm reconstructs shot navigators using SENSE and obtains the shot-wise rigid in-plane motion parameters by rigid registration. AMUSE furthermore uses the rotation parameters and a diffusion tensor estimate from the shot navigators to equalize the shot contrasts when the diffusion-sensitization direction changes. Analogous to the two-step phase-corrected algorithms, the motion estimation is sensitive to noise propagation of the shot navigators, which limits its use to low segmentations. A navigated method for macroscopic motion and shot phase corrected multi-shot DWI was proposed by Dong et al. [44]. This method uses the low-resolution sampling of a second navigator echo to estimate the phase maps and the rigid in-plane shot motion parameters by image registration.

Extra-navigated methods are mostly based on optical camera tracking and implemented as prospective motion corrections to directly acquire uncorrupted data [21]. A multi-shot DWI approach with prospective motion correction was published by Herbst et al. [80], in which rigid motion parameters are provided by external tracking, while the shot phase variations are estimated by SENSE-based shot reconstructions using the MUSE approach. The on-line estimation also allows to adapt the diffusion-sensitizing gradients to the motion on the fly and thereby reduce shot contrast variations and signal loss due to incomplete gradient refocusing [88].

Data rejection strategies

The multi-shot reconstruction assumes that the encoding model remains unaffected by the adverse signal variations. However, strong motion during the sequence can result

in complete signal loss if, for example, the excited volume has fully left the volume that experiences the refocusing pulse. In this case, data rejection schemes are integrated to remove datasets that are compromised and would rather decrease image quality.

Several data rejection criteria can be found in the literature to reject degenerate shots from the reconstruction. Navigator energy-based criteria apply a threshold to the energy of each shot navigator with respect to a signal reference [26, 27]. Shots that have experienced significant signal loss are thus detected and can be removed. A similar approach measures the correlation of the shot navigators with respect to an image reference [32]. Other algorithms use first [30] or higher [37] moments of the k-space signal to detect relevant image-space phase effects for DWI that lead to k-space peak shifting or broadening effects. Instead of a complete rejection, some approaches prefer a data weighting by the criterion to at least partially use the SNR provided by the degenerate signals [28].

SMS integration for enhanced image encoding

Simultaneous multi-slice imaging with blipped-CAIPI offers an effective slice encoding strategy that achieves considerable scan accelerations at low g-factor penalty. SMS has been successfully combined with multi-shot DWI through a self-navigated approach by Chang et al. [79] and a navigated approach by Dai et al. [64]. The combined encoding allows to distribute the aliasing in more dimensions and, in this way, make efficient use of the 3D coil sensitivity profiles. The adjustment of the SMS factor and the number of shots offers more flexible sampling to trade off scan time, SNR, and geometric distortions. SMS MUSSELS, a MUSSELS extension, provides a navigator-free shot phase correction for multi-shot SMS data [82]. Herbst et al. have furthermore extended the prospective rigid motion correction for multi-shot DWI to SMS acquisitions using camera-based extra-navigation [81].

Convergence criteria for iterative algorithms

The iterative algorithms covered in this work require criteria to evaluate the convergence and to stop the algorithm for an appropriate tolerance τ. Commonly, a maximum number of iterations is set to avoid unmanageable reconstruction scenarios.

The residual norm criterion measures the normalized euclidean distance between the measured data $\mathbf{d}$ and the synthetic data $A\boldsymbol{\rho}$ from the image estimate $\boldsymbol{\rho}$ with the forward model A [13]:

$$\delta_r(\boldsymbol{\rho}) = \frac{\|A\boldsymbol{\rho} - \mathbf{d}\|_2}{\|\mathbf{d}\|_2}. \tag{2.36}$$

This criterion is often used to stop the conjugate gradients method.

A similar $\ell 2$-norm criterion measures the euclidean distance of the joint images from two subsequent iterations [39]:

$$\delta_s(\boldsymbol{\rho}^{(k)}, \boldsymbol{\rho}^{(k-1)}) = \frac{\|\boldsymbol{\rho}^{(k)} - \boldsymbol{\rho}^{(k-1)}\|_2^2}{\|\boldsymbol{\rho}^{(k-1)}\|_2^2}, \tag{2.37}$$

where k is the iteration number. Instead of an evaluation of the optimization functional, this criterion evaluates the progression of the image results. This criterion is connected to the normalized root-mean-square error $\text{nRMSE}(\boldsymbol{\rho}, \boldsymbol{\rho}_{ref}) = \sqrt{\delta_s(\boldsymbol{\rho}, \boldsymbol{\rho}_{ref})}$ for image quality evaluation between an image $\boldsymbol{\rho}$ and a reference image $\boldsymbol{\rho}_{ref}$.

3

SELF-NAVIGATED MULTI-SHOT DWI

Self-navigated approaches derive the shot parameters of the optimization problem from the imaging data itself. The MR sequence therefore does not require navigator acquisitions and can thus be shortened providing more efficient sampling. Instead of the navigator signal, self-navigated approaches rely on more expensive computational models. The joint optimization for image and shot parameters in multi-shot DWI is often of a non-convex nature and, thus, represents challenging reconstruction problems. The first section of this chapter investigates the potential of a proximal gradient method with deep learning-based image priors for the shot phase estimation. The second section includes macroscopic in-plane motion and presents an alternating image reconstruction framework with joint shot phase and 2D rigid in-plane motion correction.

3.1

Shot Phase Estimation with Deep Learning Support

In this work, a self-navigated iterative reconstruction algorithm is proposed for high-quality multi-shot DWI, which adopts a proximal gradient method to perform shot phase updates and includes a U-Net to improve the joint image magnitude prior for this step. This data-driven framework incorporates the deep learning module for an enhanced shot phase estimation, while keeping a conventional joint image production. The U-Net is trained on reliable navigator-based reference data to mitigate phase cancellation artifacts in joint multi-shot images. The proposed algorithm with and without U-Net support is compared to a self-navigated and a navigated reference algorithm. The U-Net approach effectively mitigates phase-related signal cancellation artifacts. The improved joint image priors effectively regularize the shot phase estimation, enabling self-navigated multi-shot diffusion-weighted echo-planar imaging for challenging and highly segmented datasets.

This section is based on a conference submission to the 2020 ISMRM & SMRT Virtual Conference & Exhibition. (Abstract number: 4367, 2020, Ref. [48])

INTRODUCTION

The shot phase variations in multi-shot DWI have been addressed by a wide variety of methods over the last decades, which are either based on navigator signals [17, 18, 25–31] or fully data-driven [32–41]. A detailed historical review is provided in Section 2.2. Data-driven approaches are usually proposed to improve the scan efficiency by sparing the navigation signal in the sequence, but, conversely, rely on more expensive computational models. Nevertheless, the data-driven algorithms still depend on the conditioning of the signal encoding model and break down for high segmentations or strongly corrupted datasets.

Deep learning techniques are currently investigated to push the image reconstruction limits for ill-conditioned datasets by introducing prior knowledge learned from reference data. Bilgic et al. [77] proposed a four-step approach, termed network estimated artifacts for tempered reconstruction (NEATR), to improve the shot phase estimation. After a SMS MUSSELS reconstruction [82], a U-Net performs image enhancement on the resulting image. The improved joint image is then used as an image prior for a shot phase estimation using a projected gradient descent method [83]. Finally, the joint image is reconstructed by a conventional multi-shot reconstruction method using the updated shot phase estimates. The neural network is thus trained for an image enhancement task to remove artifacts in SMS MUSSELS reconstructions. The final image reconstruction is performed without deep learning.

The following work presents the magnitude-regularized phase estimation (MAPE) framework with U-Net support that adopts the proximal gradient descent method for shot phase estimation with a fixed joint image [83]. In contrast to NEATR, the proposed method repeatedly integrates the neural network as a regularization tool into an alternating reconstruction that continuously provides enhanced image priors for the shot phase estimation. A prior MUSSELS-based reconstruction can thus be omitted. Moreover, the training data generation is based on the navigated IRIS method [31], which represents a robust clinically accepted reference for multi-shot DWI and offers time savings for the learning task. After the training process, the data-driven framework with U-Net support can be employed for un-navigated sequences.

MODEL-BASED IMAGE RECONSTRUCTION

The self-navigated MAPE algorithm optimizes the shot phase-corrected multi-shot problem in Eq. 2.34 with respect to the shot phase estimates $\hat{\boldsymbol{\phi}}_\mathsf{s} \in \mathbb{R}^{N_p}$ and the joint image estimate $\hat{\boldsymbol{\rho}} \in \mathbb{C}^{N_p}$:

$$(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}_\mathsf{s}) = \underset{\boldsymbol{\rho}, \boldsymbol{\phi}_\mathsf{s}}{\operatorname{argmin}} \quad \sum_{\mathsf{s} \in \mathsf{S}'} \| M_\mathsf{s} F C (e^{j\boldsymbol{\phi}_\mathsf{s}} \circ \boldsymbol{\rho}) - \mathbf{d}_\mathsf{s} \|_2^2 \; + \; \lambda_{\boldsymbol{\rho}} R_{\boldsymbol{\rho}}(\boldsymbol{\rho}) \; + \; \lambda_{\boldsymbol{\phi}} R_{\boldsymbol{\phi}}(\boldsymbol{\phi}_\mathsf{s}). \tag{3.1}$$

s is the shot index from the set of included shots S'. $\circ$ is the element-wise Hadamard product. The diagonal operator formulation $\Phi_\mathsf{s} = \operatorname{diag}(e^{j\boldsymbol{\phi}_\mathsf{s}})$ (in Eq. 2.25) is equivalent to the element-wise product $(e^{j\boldsymbol{\phi}_\mathsf{s}} \circ \boldsymbol{\rho})$, whereby the latter is used to obtain an explicit expression of the shot phase vectors $\boldsymbol{\phi}_\mathsf{s}$ here. The regularization terms contain the respective weight parameters $\lambda_{\boldsymbol{\phi}}$ and $\lambda_{\boldsymbol{\rho}}$ as well as the regularization functions $R_{\boldsymbol{\phi}} : \mathbb{R}^{N_p} \to \mathbb{R}^{N_p}$ and $R_{\boldsymbol{\rho}} : \mathbb{C}^{N_p} \to \mathbb{C}^{N_p}$ with N_p image pixels.

MAPE Algorithm

The proposed framework consists of three iteratively repeated steps and is visualized in Fig. 3.1. The algorithm is initialized with zero shot phase maps resulting in the identity operator $\Phi_s = I_{N_p}$. This corresponds to a phase-uninformed SENSE reconstruction of the joint image from the multi-shot data:

$$\hat{\boldsymbol{\rho}} = \underset{\boldsymbol{\rho}}{\operatorname{argmin}} \quad \sum_{s \in S'} \| M_s F C \, \boldsymbol{\rho} - \mathbf{d}_s \|_2^2 + \lambda_{\boldsymbol{\rho}} R_{\boldsymbol{\rho}}(\boldsymbol{\rho}), \tag{3.2}$$

A weighted $\ell 2$-norm regularization function was implemented as $R_{\boldsymbol{\rho}}(\boldsymbol{\rho}) = \| W \boldsymbol{\rho} \|_2^2$ with a weight matrix W ($N_p \times N_p$). A body coil magnitude image $\mathbf{w}$ from the SENSE reference scan was used for the weighting $(W)_{p,p} = w_p^{-1}$ per pixel p. With zero shot phase maps and the complementary sampling masks M_s, the problem is equivalent to a single-shot SENSE reconstruction [12] of the merged multi-shot k-space data as in Eq. 2.12. The resulting joint image estimate is prone to ghosting and signal cancellation artifacts due to the neglected shot phase variations.

In the first step of the algorithm, a neural network (NN) performs image enhancement on the joint image $\hat{\boldsymbol{\rho}}$ to provide an improved joint image magnitude prior $\tilde{\boldsymbol{\rho}} \in \mathbb{R}^{N_p}$ for the subsequent shot phase estimation. The NN, represented by $f_{NN} : \mathbb{C}^{N_p} \rightarrow \mathbb{R}^{N_p}$, is trained to mitigate shot phase-related artifacts and denoise the joint image magnitude:

$$\tilde{\boldsymbol{\rho}} = f_{NN}(\hat{\boldsymbol{\rho}}). \tag{3.3}$$

The NN can be bypassed to omit the deep learning module, which corresponds to a magnitude mapping $f_{NN}(\boldsymbol{\rho}) = \mathrm{abs}(\boldsymbol{\rho})$, where abs takes the element-wise absolute value here.

In the second step, the shot phases $\boldsymbol{\phi}_s \in \mathbb{R}^{N_p}$ are optimized as proposed by Ong et al. [83] using a proximal gradient descent method [53] of the phase part only, while the joint magnitude image is hold fix:

$$\hat{\boldsymbol{\phi}}_s = \underset{\boldsymbol{\phi}_s}{\operatorname{argmin}} \quad \| M_s F C (e^{j\boldsymbol{\phi}_s} \circ \tilde{\boldsymbol{\rho}}) - \mathbf{d}_s \|_2^2 + \lambda_{\boldsymbol{\phi}} R_{\boldsymbol{\phi}}(\boldsymbol{\phi}_s). \tag{3.4}$$

The phase regularization function $R_{\boldsymbol{\phi}} : \mathbb{R}^{N_p} \rightarrow \mathbb{R}^{N_p}$ is implemented as an $\ell 1$-norm minimization of the signal in the Daubechies 6 wavelet domain [83]. The update rule is independent for the individual shots and is therefore fully parallelizable.

The framework furthermore allows to perform the gradient updates with only the central k-space portion on a reduced resolution. This measure imposes phase smoothness [36] and tremendously decreases the computational demands due to the reduced number of pixels. To implement this, the k-space data is cropped to the required resolution and the coil sensitivities are adapted by linear interpolation. The image prior is cropped in k-space to the reduced resolution window and masked to the acquired samples for partial Fourier acquisition. The shot images $(e^{j\hat{\boldsymbol{\phi}}_s} \circ \tilde{\boldsymbol{\rho}})$ are finally upsampled using a triangular k-space window to recover the high-resolution shot phase maps $\hat{\boldsymbol{\phi}}_s$.

In the third step, the shot phase-corrected multi-shot problem is solved using the conventional IRIS formalism [31] with the updated shot phase maps in $\hat{\Phi}_s = \mathrm{diag}(e^{j\hat{\boldsymbol{\phi}}_s})$:

$$\hat{\boldsymbol{\rho}} = \underset{\boldsymbol{\rho}}{\operatorname{argmin}} \quad \sum_{s \in S'} \| M_s F C \hat{\Phi}_s \, \boldsymbol{\rho} - \mathbf{d}_s \|_2^2 + \lambda_{\boldsymbol{\rho}} R_{\boldsymbol{\rho}}(\boldsymbol{\rho}). \tag{3.5}$$

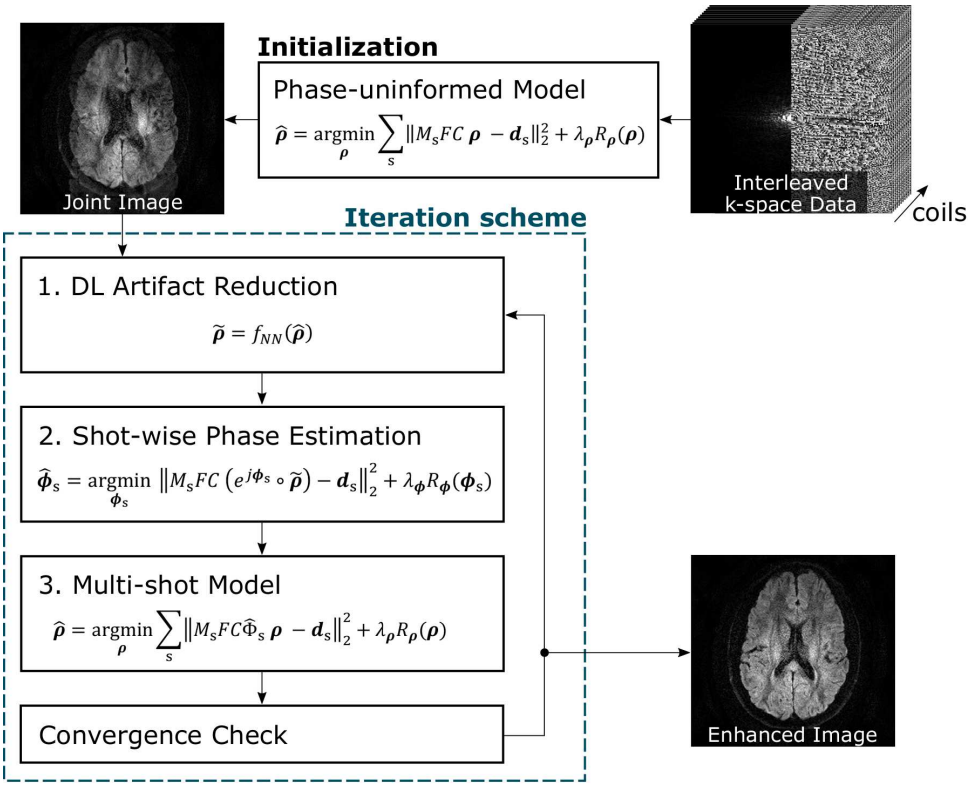


Figure 3.1: Magnitude regularized phase estimation (MAPE) framework for shot phase-corrected multi-shot DWI with deep learning support. The method is initialized by phase-uninformed SENSE of the joint multi-shot data. The resulting joint image contains severe artifacts due to the uncorrected shot phase variations. The iterative algorithm then repeatedly executes three steps until convergence. First, a U-Net performs artifact reduction of the joint image $\hat{\boldsymbol{\rho}}$ yielding the joint image magnitude prior $\tilde{\boldsymbol{\rho}}$. Second, the shot phase maps $\hat{\boldsymbol{\phi}}_s$ are updated by a proximal gradient descent with fixed joint image magnitude $\tilde{\boldsymbol{\rho}}$. Third, a conventional phase-informed multi-shot algorithm reconstructs an enhanced joint image $\hat{\boldsymbol{\rho}}$, which is fed back to the first step.

The NN thereby remains isolated from this joint image production step and enters only as an image prior for the shot phase estimation. Going on with the enhanced joint image $\hat{\boldsymbol{\rho}}$, the three steps are repeated in an iterative fashion until either the δ_s-criterion (Eq. 2.37), which is the euclidean distance of two subsequent joint images, falls below a tolerance τ_s or a maximum number of iterations is exceeded.

DEEP LEARNING IMPLEMENTATION

The deep learning-based artifact reduction was implemented through a U-Net [89]. The implementation is described in the following regarding the training data generation, the network architecture and the training process.

Training data was synthesized from in-vivo navigated multi-shot DWI data for a residual learning task [90] using the forward model with randomly scaled in-vivo phase variations. The data generation included the following seven steps:

1. *Reconstruct clean multi-shot images* by the navigated IRIS algorithm [31].
2. *Randomly sample image slices and DWI directions* from a uniform distribution.
3. *Prepare in-vivo shot phases* for the forward modeling: The shot navigator phases are unwrapped (using *restoration.unwrap_phase* from the *scipy* library) and scaled by a factor drawn from a Gaussian distribution with zero mean and variance of 0.5. Low-resolution object phases from the clean IRIS images are removed by a triangular k-space window of halved matrix size as done for POCS-ICE [39] to match the reconstruction conditions.
4. *Apply the forward model* using the random shot phases, the coil sensitivity maps and the original Fourier shot sampling.
5. *Add noise* with SNR of approx. 20 (manually matched to in-vivo conditions).
6. *Produce phase-corrupted input images* by performing phase-blind SENSE [12].
7. *Produce residual targets* by subtracting the input images from the ground truths.

A U-Net[89] was set up in PyTorch as visualized in Fig. 3.2. The network maps the input image of size (2, 240, 240) with magnitude and phase in the two channels to a magnitude output image of the size (1, 240, 240). The network is build from blocks, each concatenating a (3×3) 2D convolution layer, 2D batch normalization (BN) [91], rectified linear unit (ReLU) activation and a dropout layer (5%). The encoder path composes two of such blocks on each of five resolution levels, whereby the number of channels is doubled per level from 32 on full resolution to 512 channels on the coarsest level. The decoder path is structured in symmetric equivalence, except for the integration of the skip connections on each resolution level. Downsampling is performed by max pooling and upsamling by bilinear interpolation. For this regression task, the output layer is a 2D convolution layer (without BN and ReLU activation) combining the 32 channels to a single-channel output. In total, the network comprises 7,852,513 parameters.

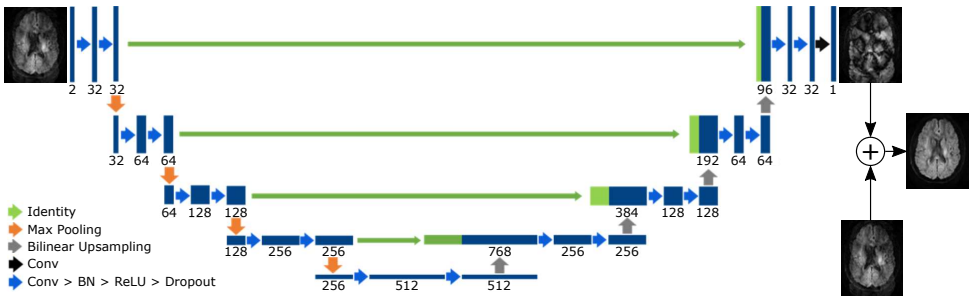


Figure 3.2: U-Net architecture in a residual learning setting for the MAPE framework. The network maps a complex input image of size (2, 240, 240), where the magnitude and phase are stacked in the two channels, to a magnitude output image of size (1, 240, 240). The U-Net consists of 5 resolution levels and applies two blocks of 2D convolutions (Conv), batch normalization (BN), ReLU activations and dropout layers per level. The number of channels is indicated below the blocks. Skip connections are marked by green arrows. The output layer is a single 2D convolution layer.

The U-Net was trained with standard training parameters and additional on-the-fly data augmentation. The ADAM [92] optimizer (beta1 = 0.5, beta2 = 0.999) was used for 100 epochs with a batch size of 30. The training was executed using $\ell 1$-loss. The learning rate was started with 10^{-5} amd reduced to $5 \cdot 10^{-6}$ and 10^{-6} after every 40 epochs. The training was performed on a GeForce RTX 2080 Ti. Further data augmentation during the training process involved random flipping of the training images along both in-plane axes. Moreover, the network should also perform well on uncorrupted datasets, because it is supposed to be used within a converging iterative process. Therefore, 25% of the corrupted input images were replaced by their clean counterparts to achieve a solid representation of uncorrupted images in the training set.

METHODS

IMAGE ACQUISITION

The aim of the deep learning supported MAPE method is to facilitate robust self-navigated multi-shot DWI without navigators. Nonetheless, additional navigation is used in this work to provide robust ground truth data for the learning task in the demanding diffusion realm. The navigated Stejskal-Tanner sequence, shown in Fig. 2.5, provides a navigator acquisition with moderate undersampling, which offers diffusion shot phase information at low g-factor penalty. The navigators can be omitted for the intended self-navigated usage after the network has been trained, allowing to shorten the sequence and thereby achieve more time-efficient volume coverage.

Diffusion-weighted brain data was acquired from 7 healthy volunteers. Informed consent was attained according to the rules of the institution. The data was acquired on a 3T Philips Ingenia Scanner using a 13-channel head coil (Philips, The Netherlands). DWI data was sampled with 4, 5 and 8 shots and DTI data with 5 shots. Important scan parameters are listed in Table 3.1. Furthermore, the acquisitions used spectral pre-saturation with inversion recovery (SPIR) for fat suppression [93], volume shimming and interleaved slice scan order. In contrast to standard DWI scans, no averaging was performed. Coil sensitivity maps were acquired in a gradient-echo prescan. EPI ramp samples were interpolated to the Cartesian positions by gridding [58] and an odd/even Nyquist correction was applied for the EPI data using an EPI reference scan [94].

ALGORITHM OVERVIEW

The MAPE framework is evaluated in two versions, one with and one without deep learning. The first version including the U-Net is explicitly termed *MAPE+U-NET* in the following evaluations. A second deep learning-free version that skips the U-Net with $f_{NN}(\boldsymbol{\rho}) = \text{abs}(\boldsymbol{\rho})$ is just called *MAPE*. The proposed methods are compared to the navigated IRIS method [31] and the self-navigated POCS-ICE algorithm [39].

The IRIS method is used as the gold standard for the network training and the evaluation, as it represents a robust reconstruction method of clinical importance. The POCS-ICE implementation generally follows the published descriptions, except for the changes mentioned in the following to improve the comparability of the methods. The four methods were used to reconstruct the DWI and DTI datasets of subjects that were not included in the training data.

Table 3.1: In-vivo measurement parameters for multi-shot DWI and DTI scans with varying segmentations.

Parameters	**DWI**			**DTI**
#shots (=segmentation factor)	4	5	8	5
Partial Fourier factor	0.63	0.62	0.63	0.62
Resolution (R × P × S) / mm^3	1 × 1 × 4	1 × 1 × 4	1 × 1 × 4	1 × 1 × 4
Image size (R × P)	232 × 228	232 × 225	232 × 216	232 × 225
Number of slices (S)	26	26	2	4
Slice gap / mm	0.4	0.4	0.4	8
Repetition time TR/ s	5	5	5	5
b-value / (s/mm^2)	1,000	1,000	1,000	1,000
#diffusion directions N_p	3	3	3	15
Image echo time TE$_{img}$ / ms	69	66	63	66
Navigator echo time TE$_{nav}$ / ms	141	126	104	126
Navigator reduction factor	1.62	2.00	2.00	2.00
Scan time [m:ss]	1:50	2:15	3:30	7:15

Implementation Details

The MAPE methods were executed with a maximum of 60 feedback iterations (60 joint image updates) and 5 shot phase updates per iteration. POCS-ICE was performed with a maximum of 300 iterations, so that the number of shot phase updates equals 300 for all iterative methods. For the 8-shot datasets, a maximum of 500 shot phase updates was set, implemented by increasing the number of feedback iterations for MAPE. The image difference criterion δ_s of two subsequent iteration in Eq. 2.37 was used with tolerance $\tau_s = 10^{-8}$. Coil compression [95] was applied for all self-navigated methods with a principal component analysis threshold of 99%, reducing the set of coils from 13 to 7. The regularization parameter for the MAPE methods and IRIS was set to $\lambda_{\boldsymbol{\rho}} = 10^{-2}$.

The shot phase estimation by the proximal gradient descent involves several implementation details. The smoothness-enforcing k-space window of the phase estimation was set to 5 mm for all self-navigated methods, which is equal to the navigator resolution and thus makes the navigated and self-navigated methods comparable. The resolution reduction of the MAPE methods was accordingly adapted and the regularization parameter was set to $\lambda_{\boldsymbol{\phi}} = 10^3$. Following Ref. [49], the forward model of the shot phase estimation was formulated in image space using the shot PSF according to Eq. 2.6. Moreover, a coil sensitivity mask was generated from the body coil reference image $\mathbf{w}$ by thresholding the absolute values at 10% of its maximum. Then, 10 binary closings and 5 binary dilations were performed on the mask using the *scipy* library.

The implementation of the proximal gradient descent method for the shot phase updates involves an appropriate step size strategy. Typically, the maximum eigenvalue of the forward model is a useful measure to set the step size of the proximal gradient descent [83], but it is hard to estimate for the present problem size. Therefore, the step size of the proximal gradient descent method was scaled once in the beginning to a maximum step of $\pi/12$ per shot. The proximal method was then implemented using a simple

step size strategy [96], which was implemented to reduce the step size by a factor of 0.1 if the functional value after the gradient step has increased.

Partial Fourier reconstructions were included by homodyne approaches [2]. The multi-shot SENSE reconstructions used by the navigated IRIS [31] and (iteratively) by the MAPE methods employed two steps. First, a low-resolution image is reconstructed from the symmetrically sampled k-space portion. This image is upsampled by a 1D-Hann k-space window in the phase encoding direction to obtain the low-resolution phase map on the full grid. Second, the unsampled k-space portion is recovered by enforcing conjugate symmetry after correcting for the phase map. For POCS-ICE, the low-resolution phase map is obtained in a post-processing step using 1D-Hann k-space windowing, followed by the homodyne reconstruction.

The data for the deep learning task were generated from the 4- and 5-shot DWI scans of the 7 subjects with 26 slices and 3 diffusion directions. Thus, in total 1,092 raw images were available (7 subjects × 2 segmentation factors × 26 slices × 3 DWI directions). Following the data synthesis described above, 715 input and target images were created for each of the 14 subject-segmentation combinations obtaining approximately 10,000 images in total. A subject-wise training/test split of 6/1 was used to separate training and test data. The full simulation dataset was jointly normalized so that the absolute values lie in the interval $[0,1]$. Finally, the images were padded to (240, 240) image size.

RESULTS

U-NET PERFORMANCE

The U-Net's ability to reduce shot phase-related artifacts in the joint images is crucial to provide valuable image priors for the shot phase estimation. The performance of the NN on the test dataset is analyzed in Fig. 3.3 in comparison to the target images.

Figure 3.3a contains the target image along with the U-Net input and output for two example slices with different levels of corruption. The input image of Example 1 appears similar to the target image and does not contain visible shot phase-related ghosting artifacts, only a slightly decreased SNR. The noise level appears slightly reduced in the U-Net output yielding a nRMSE reduction from 11% to 8%. Example 2 is subject to severe ghosting artifacts from the shot phase variations (orange arrows). The U-Net largely suppresses the signal cancellations and recovers the overall object support well while bringing down the nRMSE from 46% to 19%. Nevertheless, some image features are missing and the NN introduces blurring in the areas with strong artifacts (blue arrow). Figure 3.3b shows a nRMSE boxplot statistic over the complete test set with 1400 images. The median nRMSE is reduced from 17.5% to 12.2% and the lower/upper quartiles are reduced from 13.3% / 29,0% to 9.8% / 16.6%.

The U-Net consistently improves the joint image estimates measured by nRMSE over the testset. Although the image quality of the one-time U-Net output is not sufficient in case of phase-corrupted datasets, it is still capable to recover the overall shape already after one application. It thus provides a more consistent object support for subsequent shot phase estimations. For uncorrupted or converging joint image estimates, the impact of the network is small and a denoising characteristic is observable on the test data, which are important features to maintain convergence in an iterative process.

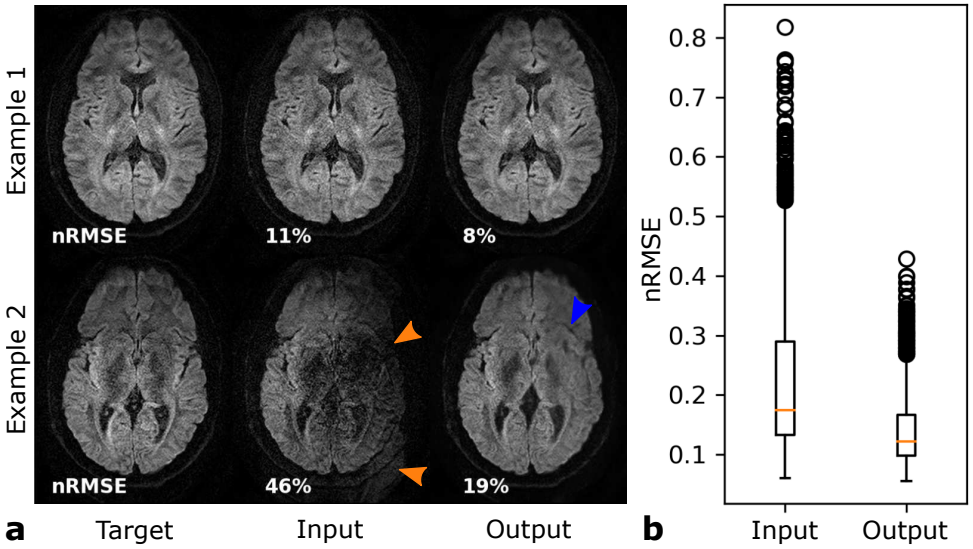


Figure 3.3: U-Net performance on the test dataset. a: Network input and output images of two example slices from the test dataset compared to the target image. The nRMSEs with respect to the target are indicated below the images. Example 1 appears free from input phase corruptions and the U-net obtains a similar but slightly denoised output. Example 2 contains strong ghosting artifacts (orange arrows) in the input image that are effectively mitigated in the output at the cost of blurring (blue arrow). b: nRMSE statistic over the full test set. The U-Net Output shows significantly reduced errors.

Multi-shot DWI Reconstructions - Moderate Segmentations

In-vivo multi-shot data from the test set was reconstructed by the four methods to evaluate the performance of MAPE and MAPE+U-Net, where the latter algorithm involves the iterative U-Net applications.

The multi-shot reconstruction performance over the test set is compiled in Tab. 3.2 considering the nRMSE to the IRIS reference and the reconstruction time per phase iteration. The overall nRMSE is reduced for the MAPE methods, whereby MAPE+U-NET outperforms MAPE. The reconstruction times per phase iteration are comparable for POCS-ICE and MAPE+U-Net and reduced to approximately two thirds for MAPE without the U-Net. Each U-Net application takes on average approximately 1.2 s, leading to about 0.24 s time difference per phase iteration between MAPE and MAPE+U-Net (as the NN is applied every five phase updates).

Table 3.2: nRMSE and duration mean and standard deviations of 78 self-navigated multi-shot DWI reconstructions per segmentation factor. The nRMSE was calculated with respect to the navigated IRIS algorithm.

Algorithm	**#shots**	**POCS-ICE**	**MAPE**	**MAPE+U-Net**
nRMSE / %	4 shots	19.24±8.40	15.29±5.51	15.10±5.14
	5 shots	23.86±6.97	19.37±5.78	18.82±4.98
Duration per phase update / s	4 shots	0.41±0.12	0.26±0.08	0.49±0.13
	5 shots	0.53±0.12	0.29±0.06	0.53±0.09

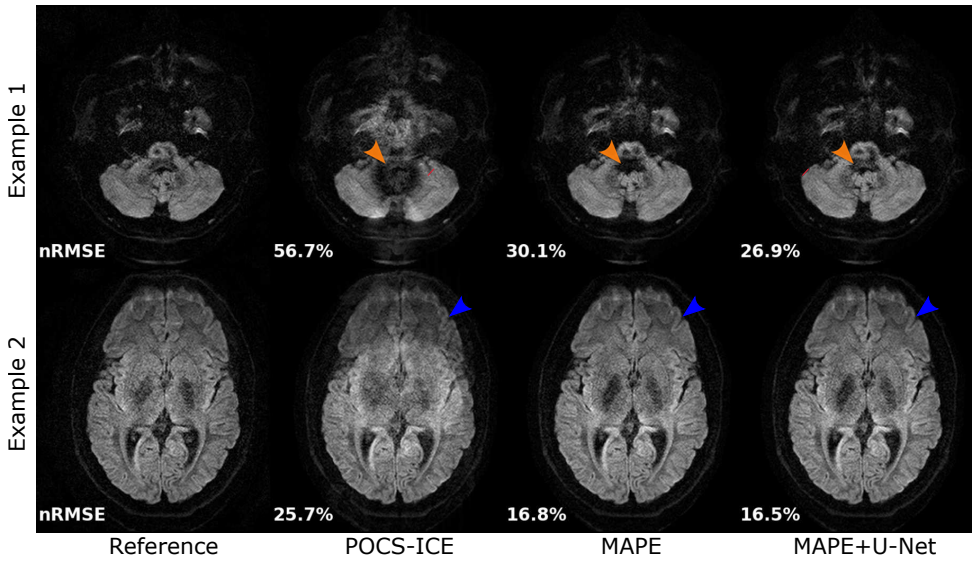


Figure 3.4: Comparison of the self-navigated iterative reconstruction algorithms to the navigated IRIS method for challenging 5-shot DWI examples at $b = 1000$ s/mm^2. The final joint images are shown along with the nRMSE with respect to the navigated IRIS (to be interpreted with care). Example 1 in the brain base is subject to strong phase disturbances (orange arrows), which are improved for MAPE and substantially mitigated for MAPE+U-Net. The residual ghosting visible for POCS-ICE in Example 2 is suppressed by both MAPE methods.

Figure 3.4 shows two example slices from the 5-shot DWI reconstructions. Example 1 is a slice from the brain base that comprises considerable shot phase variations. Compared to the navigated IRIS reference, POCS-ICE suffers from residual ghosting and signal cancellations, while MAPE reduces the dropouts and MAPE+U-Net appears closest to the reference (orange arrows). The respective nRMSE dropped accordingly from 56.7% over 30.1% to 26.9%. Example 2 shows residual ghosting at the frontal lobe and an increased nRMSE of 25.7% for POCS-ICE, while the MAPE methods appear ghost-free at similar nRMSE of 16.8% and 16.5% (blue arrows).

The MAPE framework outperforms POCS-ICE providing more reliable shot phase estimates. For challenging datasets, MAPE+U-Net benefits from the improved image priors of the U-Net leading to more consistent phase estimates and thereby reduces the ghosting artifacts considerably. On average, MAPE+U-Net slightly improves on MAPE, which shows that the U-Net is especially helpful in ill-conditioned situations.

MULTI-SHOT DWI RECONSTRUCTIONS - HIGH SEGMENTATION

To evaluate the characteristics under strongly ill-conditioned in-vivo settings, the self-navigated algorithms were applied to 8-shot datasets with 500 phase updates. The final joint images are visualized along with three example shot phase maps in Fig. 3.5.

POCS-ICE contains residual aliasing artifacts after 500 iterations (orange arrow) and achieves a nRMSE of 52% with respect to the navigated IRIS reference. The estimated shot phase maps strongly deviate from the reference navigator phases in the central brain areas and contain residual aliasing artifacts (red arrows). MAPE without the U-Net

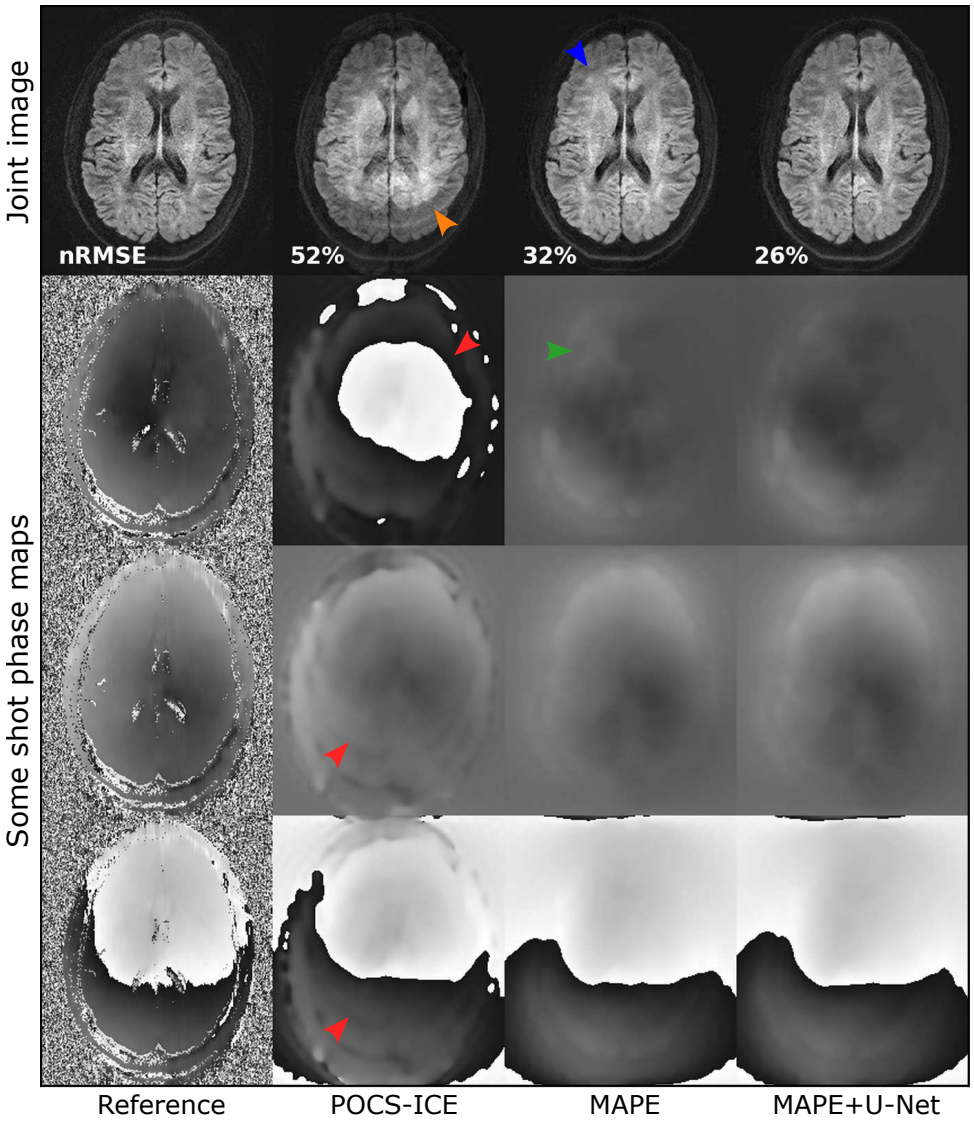


Figure 3.5: Self-navigated in-vivo results for highly segmented 8-shot echo-planar DWI with $b = 1000$ s/mm^2. POCS-ICE is subject to severe residual aliasing artifacts, caused by erroneous shot phase estimates (red arrows) from imperfect pixel unfolding. MAPE reduces the aliasing artifacts, but still comprises pronounced shading artifacts in the frontal lobe. Herein, the U-Net support greatly improves the capability of MAPE to capture important phase characteristics for the highly segmented data.

considerably mitigates the artifacts compared to POCS-ICE and achieves 32% nRMSE, but remains with some artifacts at the frontal lobe (blue arrow) caused by residual phase inconsistencies for ill-conditioned shots (green arrow). MAPE+U-Net appears aliasing-free and achieves the lowest error.

The U-Net, which was trained for 4- and 5-shot data, successfully supports the image reconstruction for higher segmentation. The deep learning module proofs to be effective in ill-conditioned situations and obtains artifact-free self-navigated reconstructions even for highly segmented 8-shot DWI. For such high segmentations, the individual shot problems are extremely undersampled (R = 8) and SENSE-based shot phase navigation therefore gets very ill-conditioned, especially for shots whose sampling trajectory omit the high-energy k-space center. Even iterative methods like POCS-ICE comprise shot phase artifacts from incomplete unfolding, especially in the central brain, where the coil sensitivity distinction is poor resulting in a high g-factor. The fixed joint image magnitude and the low-resolution reconstruction of MAPE provide valuable regularization for the individual shot phase estimations. The U-Net furthermore improves the quality of the joint image priors for the phase estimation. Reducing ghosting artifacts and providing object support in areas with strong signal cancellations, the U-Net support results in more effective phase updates, in particular for ill-conditioned situations.

MULTI-SHOT DTI RECONSTRUCTIONS

The 5-shot 15-direction in-vivo DTI data was reconstructed for each diffusion direction using the navigated IRIS reference algorithm and the three self-navigated approaches. The diffusion tensors were estimated by a linear least squares fit over all diffusion directions using the *Dipy* package for python [68]. Figure 3.6 compares the colored fractional anisotropy maps of POCS-ICE and MAPE+U-NET to the reference algorithm.

Example 1 in Fig. 3.6 shows a slice through the lateral ventricles comprising the strongly anisotropic diffusion characteristic of the corpus callosum, which connects the two brain hemispheres. POCS-ICE and MAPE+U-Net result in FA maps that are very similar to the navigated reference algorithm. Both self-navigated algorithms deliver comparable high-quality results for this well-conditioned dataset. Example 2 is a transversal slice through the third ventricle involving more challenging shot phase variations. POCS-ICE suffers from residual phase-induced signal cancellations in the underlying multi-shot DWI images that spoil the tensor estimates in the central brain area, whereas MAPE+U-Net mitigates the DWI artifacts leading to diffusion tensor results comparable to the reference (white arrows). By this, MAPE+U-Net also shows its generalization capabilities over the diffusion directions, as the network was trained on three orthogonal diffusion directions only. Around the third ventricle, the self-navigated algorithms present structural FA and SNR differences compared to the IRIS reference (orange arrows) that are unrelated to shot phase inconsistencies in the underlying DWI images.

DISCUSSION

The main contributions of the MAPE framework are threefold. First, it adapts the alternating phase cycling algorithm proposed by Ong et al. [83] for multi-shot DWI and thereby provides self-navigated high-quality image reconstructions. Second, the framework offers a comprehensible integration of deep learning to support the ill-conditioned shot phase estimation. Third, the proposed neural network training is based on navigated data to overcome the problem of missing reference data for DWI. Overall, the U-Net effectively regularizes the shot phase updates by improved joint image priors. Working on the joint image level, it is not bound to a fixed segmentation factor. MAPE+U-Net

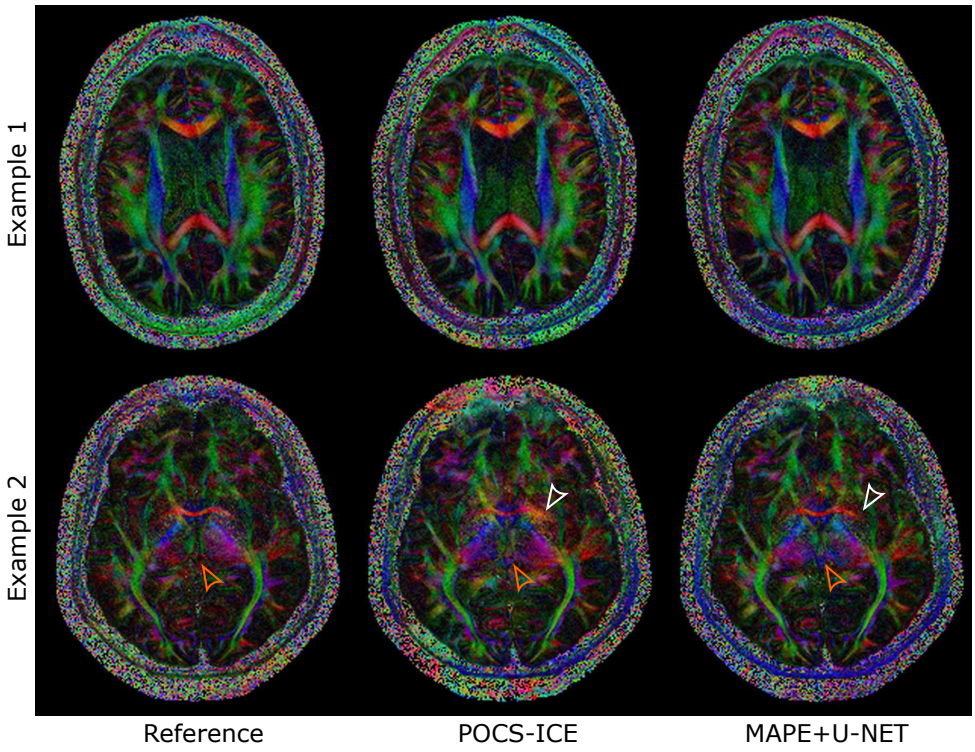


Figure 3.6: Colored fractional anisotropy maps for 5-shot in-vivo DTI datasets with 15 diffusion directions and $b = 1000$ s/mm^2. For Example 1, the self-navigated algorithms obtain similar high-quality results compared to the IRIS reference. For Example 2, POCS-ICE structurally differs from the reference in central brain areas caused by residual shot phase inconsistencies, while the MAPE+U-Net results nicely coincide with the navigated reference in this area (white arrows). Nevertheless, the self-navigated algorithms show structural differences to the navigated reference (orange arrows) requiring clinical evaluation of the methods apart from the shot phase inconsistencies.

enables self-navigation for high segmentation factors and challenging datasets. At the same time, the U-Net is kept isolated from the final image reconstruction, which facilitates the interpretation of potential artifacts by means of the underlying shot phase estimates. For low segmentations or small phase variations, MAPE+U-Net performs similar to MAPE, because the individual shot problems are well-conditioned and the regularization imposed by the U-Net's joint image priors is not critical.

The shot phase estimation of MAPE depends on three main components: the encoding model, the phase smoothness filter and the joint image prior. The encoding model consists of the SENSE and the Fourier encoding, whose properties determine the condition of the shot forward model. The SENSE encoding could be improved by coil designs with higher channel numbers and optimized sensitivity profiles [12]. Regarding the Fourier encoding, non-Cartesian sampling trajectories like the multi-shot spiral provide a more randomized sampling that improves the denoising potential [45]. Moreover, the spiral shots contain valuable low-resolution information from the k-space center for

each shot. Conversely, non-Cartesian trajectories come with more expensive image reconstructions and are more prone to blurring artifacts [57].

Phase smoothness is imposed by low-resolution windowing functions in k-space for both POCS-ICE and MAPE. The central k-space has a higher SNR compared to outer regions for normal MRI images. Appropriate filtering can therefore improve the SNR for the shot problem at the cost of missing important high-resolution phase effects [34]. MAPE can convert the low-resolution conditions to immediate computational gains, because the shot phase updates are directly performed on reduced resolution. In contrast, POCS-ICE iteratively projects the shot image estimates onto the sampled shot data and averages the shot images to a joint image after suppressing the low-resolution shot phases [39]. By this, the shot phase resolution can not be reduced without changing the resolution of the joint image.

Finally, MAPE holds the image magnitude fixed for the proximal gradient updates that are just performed on the reduced set of phase parameters [83]. The quality of the joint image priors is crucial to pose the gradient phase updates well, because a magnitude misfit affects the phase gradients through a non-linear relation. Hence, alternating updates of both the shot phases and the joint image are required to obtain high-quality image reconstructions.

The current major limitations are the small homogeneous dataset used in this study and the accessibility of reliable reference data. The current results represent a proof-of-concept study on a relatively small dataset from a single scanner. To achieve clinical acceptance, implementation and testing on larger multi-scanner and -site datasets are required. The generalization capabilities to untrained diffusion directions and segmentation factors were shown for the DTI data and 8-shot DWI data, respectively. Nevertheless, the effects of varying scan parameters, such as the b-value, the echo times, repetition times, resolution or other anatomies, on the network performance are subject to future work and represent an important concern for clinical adoption.

The self-navigated algorithms were evaluated with respect to the navigated IRIS algorithm [31] as a gold standard. IRIS-based methods have been authorized for clinical use and have been commercially available for a few years. Nevertheless, the DTI results of the navigated and self-navigated approaches show different SNR and structural DTI properties as outlined for Fig. 3.6. The results suggest further critical evaluations of the differences between navigated and self-navigated methods and, in general, limit the validity of the navigated reference method as a gold standard for structural DTI evaluations, because it relies on navigator information that can be compromised by the high navigator echo time. There might be a transitions zone where self-navigated approaches are not only more scan time-efficient but also more reliable than their navigated counterparts. However, for the present evaluation of shot phase estimates and the amount of artifacts in the resulting joint images, the navigated IRIS method represents a reliable reference method.

The comparison of MAPE+U-Net to other deep learning-based multi-shot DWI methods is another major issue. The NEATR algorithm [77] performs image enhancement on MUSSELS reconstructions [40], followed by a final phase estimation and image reconstruction. A balanced reference dataset for a performance comparison of the methods is missing, apart from the accessibility of the algorithms for comparison.

Future research projects could integrate the deep learning support with convex reformulations of the shot phase problem like MUSSELS [40]. The proximal gradient descent algorithm for the shot phase estimation is based on a non-convex problem formulation. In the recent years, low-rank approaches have achieved convex reformulations of the shot phase-corrected multi-shot DWI reconstruction leveraging either local k-space [40] or image-space [41] correlations of the shots or exploiting the q-space similarity [42] over multiple diffusion directions in DTI. Extensions of the proposed method to incorporate low-rank constraints could provide convergence guarantees and benefit from the learning-based regularization in ill-conditioned cases.

Conclusions

To summarize, the proposed magnitude-regularized phase estimation (MAPE) offers an effective iterative framework to improve the shot phase estimation in self-navigated multi-shot DWI. Moreover, it supports segmentation-independent deep learning integration for the SNR-critical shot phase estimation, while keeping a conventional, interpretable joint image production. Future work is required to validate the deep learning integration on larger datasets.

3.2

Self-navigated Rigid In-plane Motion Correction

Multi-shot techniques offer improved resolution and signal-to-noise ratio for diffusion-weighted imaging, but make the acquisition vulnerable to shot-specific phase variations and inter-shot macroscopic motion. Several model-based reconstruction approaches with iterative phase correction have been proposed, but robust macroscopic motion estimation is still challenging. Segmented diffusion imaging with iterative motion corrected reconstruction (SEDIMENT) uses iteratively refined data-driven shot navigators based on sensitivity encoding (SENSE) to cure phase and rigid in-plane motion artifacts. The iterative scheme is compared in simulations and in-vivo to a non-iterative reference algorithm for echo-planar imaging with up to six-fold segmentation. The SEDIMENT framework supports partial Fourier acquisitions and furthermore includes options for data rejection and learning-based modules to improve robustness and convergence.

This section is based on a publication in NMR in Biomedicine [49].
(Volume: 33, Issue: 12, 2020)

INTRODUCTION

Besides the inevitable shot-specific phase variations in multi-shot DWI, macroscopic inter-shot motion during the lengthy diffusion scans produces mismatches of individual segments and severely compromises image quality. Multi-shot DWI with shot phase and rigid in-plane motion correction has already been addressed by two methods using either self-navigation [43] or additional MR navigators [44]. Navigated techniques provide easily accessible navigation signals, but require a prolonged sequence leading to less SNR-efficient sampling. Therefore, this work focuses on data-driven approaches for shot phase and macroscopic motion correction in multi-shot DWI. A detailed review of the multi-shot DWI algorithms is provided in Section 2.2.

The self-navigated augmented MUSE (AMUSE) algorithm extends the two-step MUSE framework [35], which reconstructs individual SENSE-based shot navigators and uses the filtered shot phases for a multi-shot reconstruction. AMUSE furthermore uses the SENSE navigators to obtain shot-wise macroscopic motion states by rigid registration and a tensor estimate over all diffusion directions. Thus, the algorithm corrects also for diffusion contrast variations, which arise from the change in the effective diffusion direction according to the shot-specific rotations. Analogous to the one-time phase estimation, the registration and tensor estimates are sensitive to noise propagation effects within the SENSE-based shot navigators, which could be enhanced by iteratively reinforcing data consistency. For non-DWI multi-shot applications, iterative rigid motion estimations have achieved improved image quality by enforcing SENSE-based data consistency [84, 86, 87, 97–99], but the iterative integration of both physiological phase and macroscopic motion correction for DWI has still been challenging.

The proposed SEDIMENT method, short for segmented diffusion imaging with iterative motion corrected reconstruction, is a self-navigated alternating algorithm that subsequently improves the SENSE-based shot navigators, the shot parameters and the joint multi-shot image. The proposed contrast corrections have been shown to improve the DTI results, but the robust estimation of the initial diffusion tensor poses another challenging problem for higher segmentations apart from the motion estimates. Therefore, the diffusion contrast corrections were excluded from this work. The SEDIMENT method is presented for brain EPI. Moreover, the framework enables partial Fourier (PF) reconstructions [2] to harvest SNR by reducing the echo times for EPI.

MODEL-BASED IMAGE RECONSTRUCTION

The SEDIMENT framework addresses the multi-shot DWI problem with shot phase variations and macroscopic in-plane motion between the shots as visualized in Fig. 2.6. The joint multi-shot DWI problem optimizes for the joint image $\boldsymbol{\rho}$, the shot phase operators Φ_s and the shot-wise macroscopic motion operators Ω_s similar to Eq. 2.30:

$$\begin{aligned} &\text{minimize} && \sum_{s \in S'} \| M_s F C \, \boldsymbol{\rho}_s - \mathbf{d}_s \|_2^2 \\ &\text{subject to} && \boldsymbol{\rho}_s = \Omega_s \Phi_s \, \boldsymbol{\rho}, \quad s \in S'. \end{aligned} \tag{3.6}$$

s is the shot index from the set of included shots S'. Note that the SEDIMENT framework was set up with reversed shot operators Φ_s and Ω_s compared to the multi-shot problem

formulation in Eq. 2.30. This means that the phase maps are defined in rigid alignment with the joint image. Considering in-plane motion and assuming that the object stays within the FOV, the phase content is well represented in both ways. The shot images, be it navigated or self-navigated, are thus aligned to the joint image by Ω_s^H before the phase map extraction.

Plugging the equality constraints into the objective function yields an unconstrained formulation of the multi-shot problem:

$$\underset{\boldsymbol{\rho},\Phi_s,\Omega_s}{\text{minimize}} \quad \sum_{s\in S'} \| F\, C\, \Omega_s\, \Phi_s\, \boldsymbol{\rho} - \boldsymbol{d}_s \|_2^2. \tag{3.7}$$

SEDIMENT FRAMEWORK

SEDIMENT adopts SENSE navigation to estimate both motion-induced phase variations and macroscopic motion for multi-shot DWI reconstruction and embeds it into an iterative scheme. Inspired by POCS-ICE [39], the algorithm alternates between the shot-wise data consistency in k-space (objective function of Eq. 3.6) and the joint image constraints in image space (constraint functions in Eq. 3.6), but extends this scheme by continuous macroscopic motion corrections and optionally includes partial Fourier acquisitions and data rejection. SEDIMENT and the reference algorithm MC-SENSE+CG (Macroscopic motion corrected SENSE+CG) are shown in Fig. 3.7. Iteration superscripts were dropped to provide an uncluttered notation. The operations are always performed on the up-to-date estimates.

Initialization

SEDIMENT estimates the first shot images $\boldsymbol{\rho}_s^{init} \in \mathbb{C}^{N_p}$ using *initial CG-SENSE* [13]. The conjugate gradient (CG) method includes an intrinsic regularization attenuating low singular values [24]. The iteration number controls the regularization and thus balances aliasing and noise propagation [100].

Symmetry projection

Partial Fourier techniques assume that the spatial frequency content of the object's phase is limited in phase-encoding direction [2]. The reconstruction therefore imposes another low-resolution constraint on the phase using a conventional phase projection operator [101]. For this, the signal phase is substituted by the constrained low-resolution phase vector $\boldsymbol{\theta}_s \in \mathbb{R}^{N_p}$ with N_p image voxels:

$$\boldsymbol{\theta}_s = \measuredangle\left({F_{pe}}^H\, V_1\, F_{pe}\, \boldsymbol{\rho}_s^{init} \right), \qquad s \in S', \tag{3.8}$$

$$\boldsymbol{\rho}_s = \text{abs}(\boldsymbol{\rho}_s^{init}) \circ e^{j\boldsymbol{\theta}_s}, \qquad s \in S'. \tag{3.9}$$

The operator F_{pe} ($N_p \times N_p$) denotes the 1D-FFT in phase-encoding direction and V_1 ($N \times N$) is a 2D window with 1D Hann shape in phase-encoding direction. $\measuredangle$ extracts the element-wise phase, j is the imaginary number, the abs function takes the element-wise absolute value, $\circ$ is the (element-wise) Hadamard product and $\boldsymbol{\rho}_s \in \mathbb{C}^{N_p}$ is the s-th shot image estimate. The superscript H denotes the Hermitian of an operator. For full Fourier acquisitions, this step is skipped by setting $\boldsymbol{\rho}_s = \boldsymbol{\rho}_s^{init}$. Note that the partial Fourier phase constraint also limits the maximum frequency content of the motion-induced shot phase estimates.

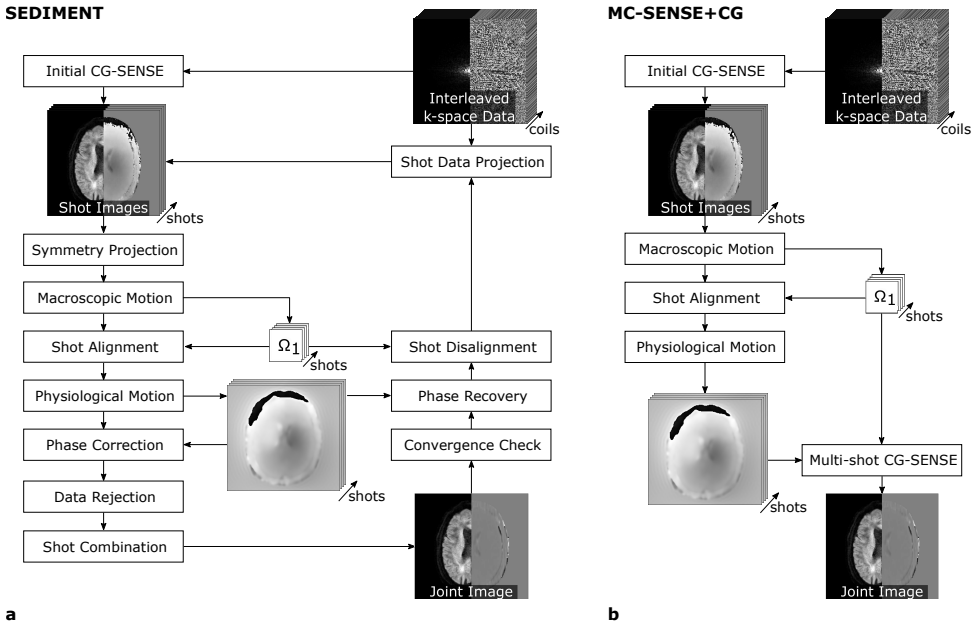


Figure 3.7: Schematic diagrams of the presented motion corrected multi-shot DWI algorithms. **a:** SEDIMENT. **b:** MC-SENSE+CG (non-iterative reference algorithm). Both schemes estimate complex-valued shot images from the data by initial CG-SENSE. The shots navigate macroscopic and physiological motion estimations yielding macroscopic parameters and phase maps. SEDIMENT optionally includes symmetry projections for partial Fourier acquisition and data rejection before shot combination and feedback to the data projection. MC-SENSE+CG employs a final multi-shot CG to solve for the joint image.

Motion estimation

The shot-image guesses navigate macroscopic and physiological motion estimations. The shot-wise operator for *macroscopic motion* Ω_s allows to include various motion models, such as rigid, affine or elastic transformations. In the context of brain imaging, this work performs rigid registration of the shot magnitude images as presented in AMUSE [43]. The shot with the highest total correlation to all other shots is chosen as registration reference with index s_{ref} once in advance. The shots are then aligned by applying the estimated motion transformation. Next, the *physiological motion* correction requires the estimation of the motion-induced shot phase variations $\boldsymbol{\phi}_\mathsf{s} \in \mathbb{R}^{N_p}$. Assuming spatial smoothness, the phase maps are extracted from low-resolution data using a 2D triangular k-space window [39]:

$$\boldsymbol{\phi}_\mathsf{s} = \measuredangle\left({F_{2d}}^H \, V_2 \, F_{2d} \, \Omega_\mathsf{s}^H \boldsymbol{\rho}_\mathsf{s}\right), \qquad \mathsf{s} \in \mathsf{S}'. \tag{3.10}$$

Here, $\Omega_\mathsf{s}^H \boldsymbol{\rho}_\mathsf{s}$ is the aligned shot image, F_{2d} $(N_p \times N_p)$ the 2D-FFT operator and V_2 $(N_p \times N_p)$ the window function. The phase operator is constructed by $\Phi_\mathsf{s} = \mathrm{diag}(e^{j\boldsymbol{\phi}_\mathsf{s}})$. After correcting for macroscopic and physiologic motion, a joint image estimate $\hat{\boldsymbol{\rho}}_\mathsf{s} = \Phi_\mathsf{s}^H \Omega_\mathsf{s}^H \boldsymbol{\rho}_\mathsf{s}$ can be obtained for each shot.

Optional data rejection

The initial set of shots S can be reduced to make the reconstruction robust against unmodeled artifacts by data rejection. Besides shot mismatches due to through-plane motion or failed registration, patient motion can affect the MR signal evolution and the diffusion encoding in various ways [22] leading to severely corrupted shot data. By exclusion from the current set S', corrupted shots can be rejected before shot combination by analyzing k-space peak broadenings [30, 37], correlation measures [102] or other means.

In this work, the normalized root-mean-square error (nRMSE) of shot s with respect to the registration reference s_{ref} is evaluated in every iteration:

$$\mathrm{nRMSE}_{\mathsf{s}} = \frac{\|\hat{\boldsymbol{\rho}}_{\mathsf{s}} - \hat{\boldsymbol{\rho}}_{\mathsf{s}_{ref}}\|_2^2}{\|\hat{\boldsymbol{\rho}}_{\mathsf{s}_{ref}}\|_2^2}. \tag{3.11}$$

A shot index is rejected from S' when $\mathrm{nRMSE}_{\mathsf{s}}$ exceeds a given threshold η. This measure includes both residual magnitude and phase variations.

Shot combination

The joint image $\boldsymbol{\rho}$ is obtained by complex averaging of the motion corrected shot images:

$$\boldsymbol{\rho} = \frac{1}{|\mathsf{S}'|} \sum_{\mathsf{s}\in\mathsf{S}'} \hat{\boldsymbol{\rho}}_{\mathsf{s}}. \tag{3.12}$$

The operator $|\mathsf{S}'|$ is the cardinality of the set S' representing the number of averaged shots. For known motion operators, the joint image constraints in Eq. 3.6 can be interpreted as a consensus constraint for the unconstrained joint problem in Eq. 3.7. This is related to an average projection operator [53].

Apart from the projection, the joint image could also be retrieved by solving the extended SENSE problem in Eq. 3.7 using CG as in AMUSE [43], but the complex averaging involves relatively low computational loads for the use in an iterative scheme and reuses the previous shot-wise joint image guesses $\hat{\boldsymbol{\rho}}_{\mathsf{s}}$.

Shot data projection

The *shot data projection* recovers shot images $\boldsymbol{\rho}_{\mathsf{s}}^{proj}$ from the joint image $\boldsymbol{\rho}$ and reinforces SENSE-based data consistency for each shot [39]. For Cartesian trajectories, the data projection just substitutes the estimated data points by actually measured ones, whereas unsampled k-space positions remain unchanged [101]:

$$\boldsymbol{\rho}_{\mathsf{s}}^{proj} = \Omega_{\mathsf{s}}\,\Phi_{\mathsf{s}}\boldsymbol{\rho}, \qquad \mathsf{s}\in\mathsf{S}', \tag{3.13}$$

$$\boldsymbol{\rho}_{\mathsf{s}}^{init} = \boldsymbol{\rho}_{\mathsf{s}}^{proj} + C^H F^H M_{\mathsf{s}}^H (\mathbf{d}_{\mathsf{s}} - M_{\mathsf{s}} F C\, \boldsymbol{\rho}_{\mathsf{s}}^{proj}), \qquad \mathsf{s}\in\mathsf{S}'. \tag{3.14}$$

Hence, the shot estimates are updated by the joint image and fed back into the iterative scheme which is repeated until convergence.

ACCELERATION STRATEGIES

The SEDIMENT scheme is furthermore accelerated using coil compression [95, 103, 104] and the point spread function (PSF) of regularly undersampled Cartesian EPI [12, 99]. As the SENSE matrix size scales linearly with the number of coils N_c, the computational expense of the SENSE updates can be reduced by coil compression.

Moreover, the PSF of Cartesian trajectories can be used to substitute the k-space undersampling [39, 101] by pure image space operations. In general, Cartesian undersampling corresponds to a convolution with a shah-shaped PSF in image space [12].

Regarding multiple uniformly undersampled EPI interleaves, the trajectories are equal, except for a shot-specific shift from the k-space origin. The k-space undersampling thus involves three concatenated image space operations. Firstly, the trajectory is shifted back to the origin by multiplying a phase ramp Φ_s^{shift} $(N_c N_p \times N_c N_p)$ in image space. Secondly, the k-space undersampling of the centered trajectory is applied by PSF convolution represented by PSF $(N_c N_p \times N_c N_p)$. Finally, the trajectory offset is reversed by the conjugate phase ramp to obtain the original shifted trajectory for each shot:

$$F^H M_s^H M_s F = \Phi_s^{shift\,H} \,\mathrm{PSF}\, \Phi_s^{shift}. \tag{3.15}$$

This PSF formulation is an identity, not an approximation, for substituting the FFTs. Note that the image space operations require that the number of pixels in phase-encoding direction is divisible by the reduction factor. Furthermore, partial Fourier acquisitions exclude this acceleration as the truncated sampling affects the Cartesian PSF.

METHODS

REFERENCE ALGORITHMS

The proposed algorithm was compared to the MC-SENSE+CG reference scheme and another variant of the SEDIMENT framework. All algorithms start by the CG-SENSE initialization, followed by macroscopic motion estimation, shot alignment and physiological motion estimation as shown in Fig. 3.7.

MC-SENSE+CG is a non-iterative algorithm that solves the joint multi-shot diffusion problem in Eq. 3.7 using CG. The method can be interpreted as SENSE+CG [36] extended by macroscopic motion estimation. Phase unwrapping and 2D median filtering were used for physiological motion estimation. MC-SENSE+CG is also basically similar to AMUSE [43] as it obtains shot estimates using SENSE, extracts the motion parameters and calculates the final image once. AMUSE further corrects for diffusion contrast variations caused by rotational motion which is neglected in this work. Regarding the scope of this work, a dedicated partial Fourier scheme for MC-SENSE+CG was excluded.

Prior-MC SEDIMENT (Prior Macroscopic Motion Corrected SEDIMENT) adapts the iterative procedure of SEDIMENT, but skips the macroscopic motion estimation after the initial estimate. This variant was implemented to evaluate the necessity of combined iterative physiological and macroscopic motion correction. The repetitive registration should yield performance gains to justify the increased computational load.

NUMERICAL SIMULATIONS

The algorithms were evaluated both in simulations and in-vivo. The simulations were performed using the BrainWeb phantom [105] from the T_1-weighted normal brain database and $1 \times 1 \times 1\,\text{mm}^3$ resolution. The phantom was padded to a matrix size of $256 \times \lfloor 256/N_{shots} \rfloor N_{shots}$ in readout and phase-encoding direction, respectively, ensuring equal but shifted trajectories for all EPI shots.

The simulation data was prepared by five steps according to the forward model. First, the motion-induced phase variations were created and applied for each shot as random functions of second spatial order as presented by Hu et al.[106]. Second, shot-wise rigid in-plane motion was uniformly sampled from a range of ± 5 pix (equals ± 5 mm) and ± 10 deg and the shot data was transformed accordingly. Third, 12 2D Gaussian sensitivity maps were arranged circularly around the image center. The disturbed shot data was multiplied by the CSMs to obtain multi-shot multi-coil data. Fourth, complex Gaussian noise with zero mean and equal variance for the real and imaginary parts was added in image-space according to the predefined SNR. Finally, the data was undersampled in k-space according to the shot trajectories (optionally also including partial Fourier acquisition).

The BrainWeb phantom was prepared for $\{2,3,4,5,6\}$ shots and SNRs of $\{5,10,15,20\}$ without partial Fourier trajectories. The simulation data was reconstructed by the three algorithms without data rejection. The nRMSE and reconstruction time were used to measure performance. Total performance was measured as the average over 10 random simulation cases for each shot-SNR pair.

IN-VIVO EXPERIMENTS

The in-vivo experiments were executed on a 3 Tesla Philips Ingenia Scanner (Philips Healthcare, Best, The Netherlands) using a head coil with 13 channels. The data was obtained from six healthy volunteers. Informed consent was attained according to the rules of the institution.

The multi-shot echo-planar brain DWI experiments were performed using conventional Stejskal-Tanner diffusion encoding within a spin echo sequence [5] and magnetization prepared fat suppression [93]. The DWI data was obtained in both full and partial Fourier acquisitions with 4 and 6 shots for a b-value of 1000 s/mm^2 in three orthogonal directions. The subjects were asked to perform random in-plane motion from shot to shot within the head coil. DTI experiments were executed with 4 and 5 shots for a b-value of 1000 s/mm^2 in 15 diffusion directions. Here, both static and gross motion-corrupted data were acquired. Coil sensitivity maps were acquired by precalibration. Relevant parameter settings are listed in Table 3.3.

The multi-shot datasets were reconstructed using the three presented algorithms. The data rejection threshold was set to $\eta = \{0.4, 0.55, 0.7\}$ for segmentations of $\{4,5,6\}$, respectively, if not stated otherwise. For DTI, the joint images from the individual multi-shot reconstructions for each diffusion direction were aligned with the non-DWI ($b = 0$ s/mm^2) image using an affine motion model. The registration, tensor estimation and fractional anisotropy (FA) calculations [4] were performed using the *Dipy* library [68].

Table 3.3: In-vivo measurement settings for DWI and DTI experiments with varying segmentation and partial Fourier strategies.

Experiment	**DWI**				**DTI**	
Acquisition	Full Fourier		Partial Fourier		Partial Fourier	
Number of shots	4	6	4	6	4	5
TE/ms	98	89	65	61	68	66
TR/s	5	5	5	5	5	5
Matrix size R	232	232	232	232	232	232
Matrix size P	228	222	228	222	228	225
In-plane resolution / mm	0.957	0.957	1.000	1.000	1.000	1.000
Slice thickness / mm	4	4	4	4	4	4
Partial factor	1.000	1.000	0.632	0.622	0.632	0.622
b-value / s/mm^2	1,000	1,000	1,000	1,000	1,000	1,000
#diffusion directions	3	3	3	3	15	15

IMPLEMENTATION DETAILS

The reconstructions were conducted using Python 3.6.5 on a system with a 2.7 GHz Intel Core i7 4-core CPU and 16 GB RAM. The preparations of the coil sensitivity maps included masking and coil compression [95]. The sensitivities were masked [12] by a magnitude signal thresholding with a body coil reference image. The threshold was set to 10 % of the body coil magnitude maximum value. In addition, binary closing (10 iterations) and binary dilations (5 iterations) were performed using the *scipy* library. In simulations, the phantom was used as the reference image for thresholding after smoothing by a 2D Gaussian filter ($\sigma = 15$). Coil compression was performed by principal component analysis with 99 %-threshold using the singular value decomposition in *numpy*. This resulted in a reduction from 13 to 7 coils.

Non-uniform sampling on EPI ramps was adjusted by gridding in advance and an odd/even correction based on an EPI reference scan was used to prevent Nyquist ghosting [94]. Fourier transforms were performed using the FFT of the *numpy* library. The PSF-based undersampling was used whenever possible (not for partial Fourier acquisition) to avoid FFTs including data projections and CG gradient computations.

The CG iterations were stopped if the residual norm criterion [13] dropped below the tolerance $\tau_r = 10^{-4}$. In addition, a maximum iteration count was empirically set to 12 and 10 for single- and multi-shot reconstructions, respectively. Coil sensitivity normalization [13] was applied for all CG methods.

The macroscopic motion was estimated using the rigid registration described in the fast elastic image registration [107] framework. After an exhaustive presearch, the registration performs a Gauss-Newton scheme with Armijo's step size rule. A normalized gradient field [85] metric was used to stabilize the registration against intensity variations in the g-factor areas [12].

To stabilize convergence, the shots were aligned at their joint average location after each registration by subtracting the mean rigid parameters of all included shots (in S') [86, 108]. Furthermore, registration parameters below 0.01 pix (about 10 μm) and 0.01 deg were ignored and set to zero. Registration accuracies below this threshold were assumed immoderate hampering convergence.

The rigid shot alignment with Ω_s was performed using the k-space formulation proposed by Cordero-Grande et al. [86]. Translations were applied in k-space by multiplying phase ramps according to the Fourier shift theorem. Rotations were implemented as a concatenation of three shears applied in k-space [109]. The action of Ω_s was thus implemented by subsequent rotational and translational operators.

The SEDIMENT phase filters were k-space window functions avoiding phase unwrapping. For physiological motion estimation, the shot phases $\boldsymbol{\phi}_s$ were smoothed by a 2D triangular window in k-space using 2D-FFTs [39]. For full Fourier acquisition, the window size was scaled to half the image size. For partial Fourier acquisitions, the range in phase-encoding direction was limited to the symmetric area. MC-SENSE+CG used the phase unwrapping and 2D median filter from the *scipy* library. The median filter kernel was set to 9 x 9 pixels applied on unwrapped full resolution phases. Phase estimation was disabled for non-DWI datasets ($b_0 = 0\,\mathrm{s/mm^2}$). Partial Fourier projection phases $\boldsymbol{\theta}_s$ were obtained by a 2D window with 1D Hann shape in phase-encoding direction scaled to the size of the symmetric sampling area in the k-space center.

The iterative algorithms were stopped either by a convergence criterion or by a maximum iteration count. The mean-square difference criterion of subsequent iterations in Eq. 2.37 was used with a tolerance $\tau_s = 10^{-6}$. The maximum iteration count was set to 200 iterations.

Results

Simulation Results

The outcomes of the multi-shot DWI BrainWeb simulations are compiled in Fig. 3.8. The bar plots in Figs. 3.8a and b show the nRMSE and the durations over varying segmentations for all three methods, namely MC-SENSE+CG, Prior-MC SEDIMENT and SEDIMENT. The algorithms perform similarly for low segmentations, whereas, for more than three shots, the iterative methods achieve significantly lower errors than MC-SENSE+CG. Moreover, SEDIMENT outperforms Prior-MC SEDIMENT in terms of quality, but the reconstructions are prolonged by the registration step in each iteration compared to Prior-MC SEDIMENT.

A BrainWeb reconstruction example and the corresponding convergence are visualized in Figs. 3.8c and d. The SEDIMENT image (Fig. 3.8c) appears artifact-free compared to the ground truth, while the others comprise residual aliasing in the center from registration mismatch. The MC-SENSE+CG reconstruction additionally contains signal dropouts indicating phase-induced errors. The nRMSE of 0.1165, 0.092 and 0.0255 for MC-SENSE+CG, Prior-MC SEDIMENT and SEDIMENT matches the visual impression. The algorithms took 7.25 s, 24.95 s and 52.77 s, respectively. The nRMSE convergence (Fig. 3.8d) of the iterative methods drops under the MC-SENSE+CG error level after few iterations. Prior-MC SEDIMENT converges steadily to 0.092 after 45 iterations, whereas the convergence of SEDIMENT appears bumpier reaching 0.0255 in iteration 64.

The BrainWeb simulations emphasize the importance of iterative reconstruction, especially for high segmentations. Exceeding three shots, the SENSE-induced g-factor penalty [12] increasingly deteriorates the shot navigators resulting in corrupted final images. The iterative phase estimation realized by Prior-MC SEDIMENT effectively cures

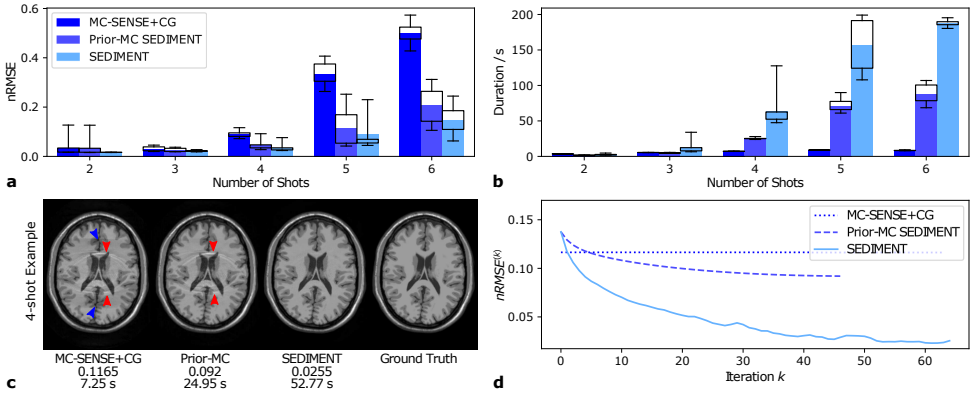


Figure 3.8: Evaluation of three multi-shot DWI algorithms in BrainWeb [105] simulations. **a:** nRMSE. **b:** Duration. The nRMSE and the duration are compiled with SNR = 10 for varying segmentations. The bars indicate 10-fold averaged mean values, while the box plot whiskers contain the full value range. **c:** A 4-shot example with SNR = 10 is shown with the final images next to the static ground truth, the nRMSE and the durations. The SEDIMENT image contains no visible artifacts, whereas Prior-MC (SEDIMENT) and MC-SENSE+CG comprise residual aliasing (red arrows) from insufficient macroscopic motion estimates. The MC-SENSE+CG has shading artifacts (blue arrows) related to deteriorated phase guesses. **d:** The nRMSE convergence is evaluated and the error level of the non-iterative MC-SENSE+CG is visualized for comparison.

ghosting and shading artifacts compared to MC-SENSE+CG, but in the presence of rigid motion, Prior-MC SEDIMENT relies on the initial macroscopic motion guess, which has proven to be occasionally insufficient.

SEDIMENT's repeated gross motion estimation refines the associated parameters at the cost of increased computational load. The convergence with in-plane rigid motion estimation is more unsteady due to its non-convex nature, but, nevertheless, SEDIMENT decently reconstructs even severely corrupted datasets. Besides, there are cases in which the present motion is not completely corrected by SEDIMENT. Convergence was stabilized by averaging the shot motion parameters [108] and limiting the registration to ± 0.01 pix and ± 0.01 deg.

IN-VIVO RESULTS

The in-vivo performance of the three algorithms for full and partial Fourier multi-shot DWI is compared for 6-shot datasets in Fig. 3.9. The full Fourier reconstructions in Fig. 3.9a show the final shot and joint magnitude images of MC-SENSE+CG, Prior-MC SEDIMENT, and SEDIMENT. The dataset contains only minor inter-shot gross motion. MC-SENSE+CG uses the initial CG-SENSE shot images and performs no further shot updates. These shot images thus also represent the shot initializations of the other two methods. The individual shots comprise strong noise propagation artifacts in areas where the signal is suppressed due to the CG regularization by early stopping [24]. The MC-SENSE+CG joint image is strongly corrupted by ghosting and signal dropout, whereby the anatomical structures, such as the interhemispheric fissure, appear unblurred. This suggests sufficiently accurate rigid motion estimation in the first step, but deficient phase estimation. The iterative recoveries appear artifact-free and are hardly distinguishable. The convergence according to Eq. 2.37 in Fig. 3.9c is almost congruent.

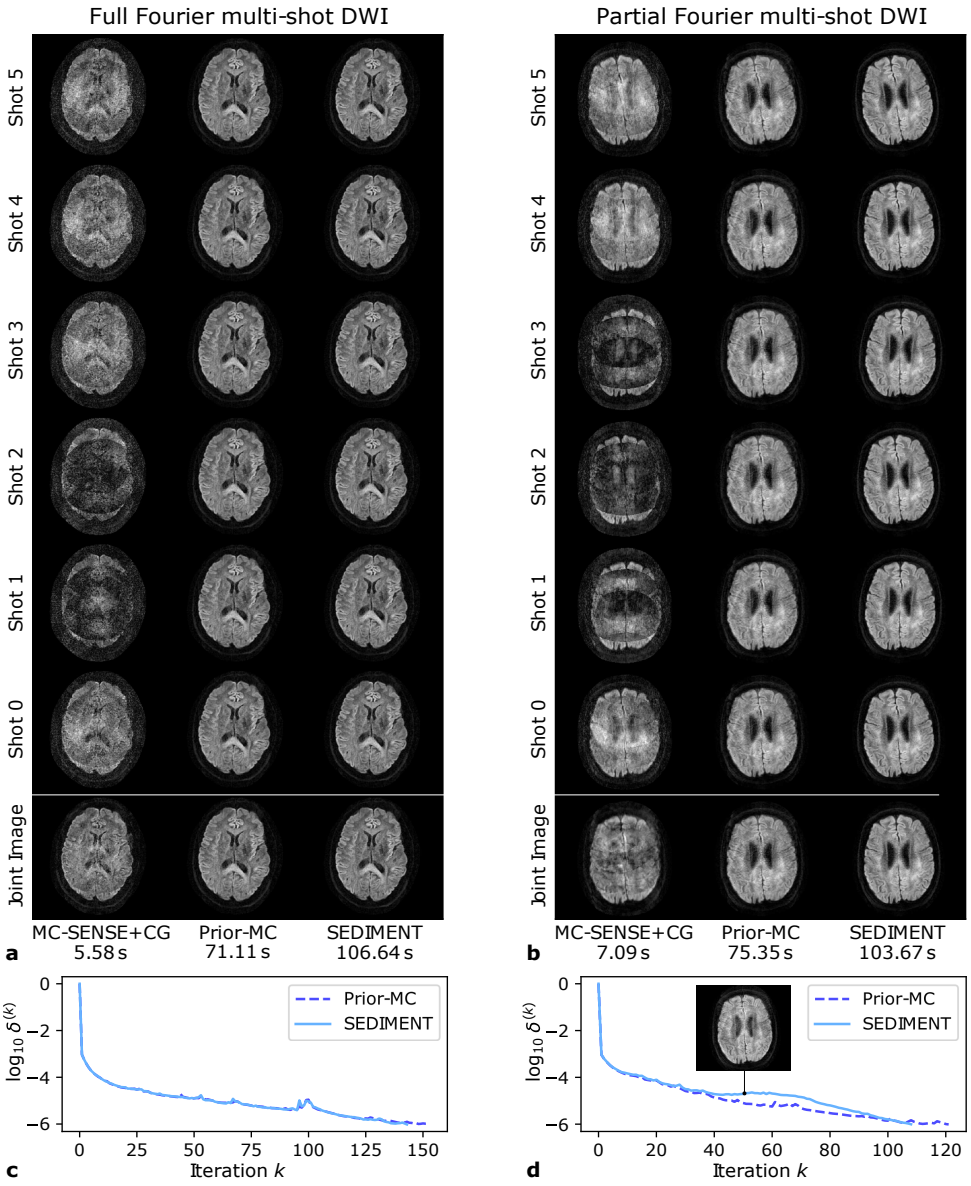


Figure 3.9: Full and partial Fourier reconstructions of 6-shot motion corrupted in-vivo DWI datasets for the three algorithms. **a:** Full Fourier images. **b:** Partial Fourier images. The full and partial reconstructions contain the final shot and joint magnitude images for the diffusion-weighted ($b = 1000\,\mathrm{s/mm^2}$) datasets with 13 coils. Reconstruction times are indicated below the algorithm names. The partial Fourier factor was 0.622. Although full and partial Fourier results of two different volunteers and slices are compared, one can appreciate the varying T_2-weighting of the partial Fourier approach. **c:** Convergence of full Fourier acquisition. **d:** Convergence of partial Fourier acquisition. The convergence progression is almost congruent in **c**, compared to the varying shape in **d**. Moreover, the joint magnitude image of SEDIMENT in iteration 50 is inserted in **d**.

The partial Fourier dataset in Fig. 3.9b appears severely affected by inter-shot gross motion. The MC-SENSE+CG shot images contain similar artifacts as in Fig. 3.9a as well as blurring in phase-encoding direction from uncorrected partial Fourier acquisition. The MC-SENSE+CG joint image is severely deteriorated. Prior-MC SEDIMENT cured the shot phase-related artifacts by iterating over the phase, but the images still comprise strong blurring from inadequate registration. SEDIMENT yields decent final images.

The partial Fourier convergence criterion in Fig. 3.9d drops unsteadily for both iterative methods. SEDIMENT needed less iterations (but more time) due to the enhanced consistency by repeated image registration. The joint image of SEDIMENT at iteration 50 was inserted to analyze the convergence unsteadiness. Most phase-related artifacts have already been corrected at this point, but the image is still gross motion corrupted, which is visibly reduced in the final image. The figure shows two slices from different subjects to capture the diversity and the differing levels of motion corruption in the data. Thus, a direct comparison is not possible, but, in general, partial Fourier data comprise less T_2-weighting and can potentially increase SNR compared to full Fourier scans.

The in-vivo results demonstrate successful image reconstructions for up to six shots with the given receive array. If no macroscopic motion is present or if the first registration provides sufficiently accurate parameters, no quality issue can be observed for Prior-MC SEDIMENT. Conversely, SEDIMENT has been shown to recover decent images even for inaccurate initial gross motion guesses as demonstrated in Fig. 3.9b. The SENSE-enabled registration apparently casts the algorithm into a sufficiently accurate (but still possibly local) minimum.

IN-VIVO DATA REJECTION

For some datasets, the motion-corrected SEDIMENT reconstruction failed due to the non-convex nature of the problem. In these cases, the SNR of the shot images normally impedes the estimation of sufficiently accurate rigid motion parameters. Figure 3.10 shows a final joint image example and the convergence criteria, respectively, of a 6-shot SEDIMENT reconstruction with and without data rejection.

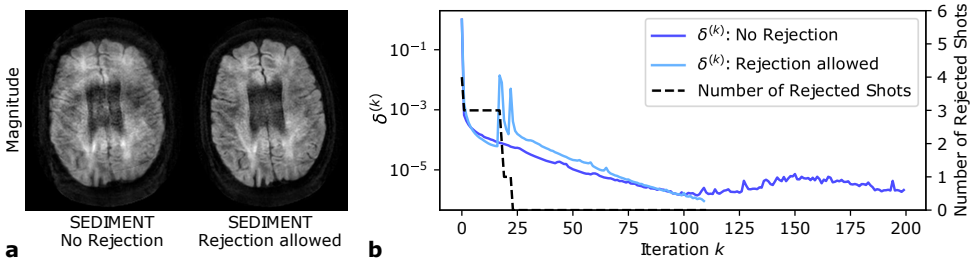


Figure 3.10: SEDIMENT reconstruction of 6-shot motion corrupted in-vivo DWI data with and without shot rejection. **a:** The final joint magnitude images with and without rejection are displayed. **b:** The respective convergence criteria are compared and the black dashed line indicates the current number of rejected shots. SEDIMENT with rejection starts with just 2 shots (4 rejected) and sequentially incorporates further shots continuously checking the rejection criteria. In this example, all shots are finally included and contribute to the joint image reconstruction with improved SNR.

The SEDIMENT reconstruction without rejection is blurred by rigid shot motion mismatches. The iterative shot data rejection with a tolerance of $\eta = 0.46$ results in an uncorrupted image. The convergence of SEDIMENT with shot rejection comprises two strong strikes compared to the relatively continuous evolution without rejection. The black dotted line indicates the number of rejected shots in each iteration.

The iterative data rejection of SEDIMENT initially excludes four shots by the nRMSE consistency measure and combines just the two remaining shots. These two shots contribute to the joint reconstruction and together enhance SNR. This, in turn, improves the conditions for image registration in subsequent iterations so that more and more shots can be included by consistency until, after about 25 iterations, all shots contribute to the final image. The convergence spikes in Fig. 3.10b coincide with the inclusion of further shots. Hence, iterative rejection improves the overall convergence by selectively including shots based on a consistency measure.

In-vivo DTI Results

For the in-vivo DTI data, the multi-shot datasets for each diffusion direction were reconstructed by SEDIMENT. Using *Dipy* [68], the resulting joint images per direction were aligned by affine registration to subsequently estimate the tensors.

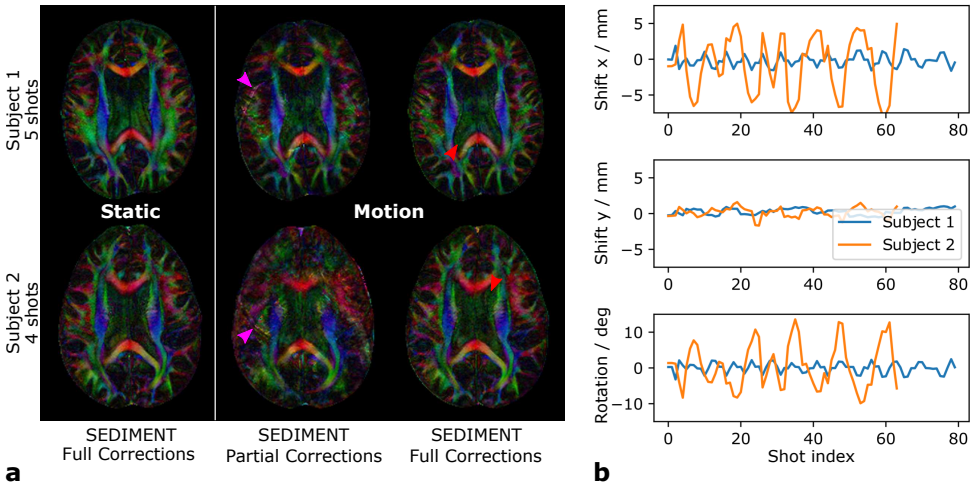


Figure 3.11: FA maps and estimated rigid motion parameters of SEDIMENT reconstructions for static and macroscopic motion-corrupted in-vivo multi-shot DTI data. **a:** FA maps of two subjects with 5-fold (upper row) and 4-fold (bottom row) segmentation. The left column shows the SEDIMENT-based FA maps acquired under static conditions (no voluntary motion). The remaining data was acquired during in-plane motion. The center column shows SEDIMENT results without gross motion correction and data rejection. Subject 1 only performed small gross motion during the scan, which produced blurring and kinks in the FA maps (pink arrow). Subject 2 performed strong motion during the scan (about ±10 deg and ±5 mm). The uncorrected mismatch severely deteriorated the SEDIMENT reconstructions resulting in a poor FA map. The iterative rigid correction is able to mitigate the motion-induced artifacts, except for minor blurring and noise-like artifacts (red arrows). Note that the static and motion case of Subject 1 show slightly different slices. **b:** Estimated rigid motion parameters for the motion case over the full data acquisition (shot index is a surrogate of time).

The resulting FA maps for a 4-shot and a 5-shot case from two different subjects are presented in Fig. 3.11a. The left column contains FA maps for static reference acquisitions without voluntary gross motion. The center and the right columns show FA maps under voluntary gross motion. In the center column, macroscopic motion correction and data rejection were disabled and only the phase correction was active (Partial Corrections). The right column shows SEDIMENT-based FA maps, in which all corrections were applied (Full Corrections). Note, that the aforementioned SEDIMENT options just affect the reconstructions of the individual diffusion directions. The calculation of the presented FA maps always involved prior affine registration of the different direction-wise images, which may or may not contain gross motion artifacts. Figure 3.11b shows the rigid motion parameters estimated by SEDIMENT over the full data acquisition. The shot index indicates the chronological excitation number in the experiment and therefore represents a surrogate of time.

Subject 1 moved just slightly during the scan (about ±2 deg and ±2 mm), which produced blurring and kinks from the rigid mismatch, which propagate through the direction-wise reconstructions into the tensor estimates (pink arrow). The motion correction greatly reduces these artifacts, except for minor blurring and noise-like structures (red arrows), and recovers important white matter structures. In contrast, Subject 2 performed large movements during the scan (about ±10 deg and ±5 mm). The strong misalignment led to heavy gross motion artifacts in the direction-wise reconstructions, which render the FA maps unusable. The iterative motion correction is able to align the shot data sufficiently to restore the overall structures and the SNR of the motion-corrupted data compared to the static reference.

DISCUSSION

SEDIMENT provides decent multi-shot DWI reconstructions in the presence of physiological and inter-shot in-plane macroscopic motion. The model-based algorithm successfully combines multiple segments and augments the coverage of motion corruption scenarios shown in Fig. 2.6. The iterative scheme achieves superior image quality compared to non-iterative methods like AMUSE [43] making higher EPI segmentations feasible, albeit the achievable segmentation is still limited for two reasons. Firstly, the g-factor penalty affects the SENSE-based shot motion navigators, which impedes data consistency for high segmentations and impairs robustness. Secondly, the computational load increases with the number of shots.

The initial CG-SENSE shot-image guesses crucially affect the robustness to deal with the non-convex macroscopic motion estimation. SEDIMENT generally uses convex optimization strategies so that it can only produce acceptable results for sufficiently precise initial guesses. Herein, SENSE provides a reliable technique to resolve the undersampling and thereby empowers the algorithm to leverage non-convex motion effects. SEDIMENT has even shown convergent behavior with acceptable results for visibly inaccurate initial motion estimates. In general, the condition of the individual SENSE-based shot problems determines both convergence and the feasibility itself.

SENSE-based shot navigators generally benefit from enhanced coil orthogonality, more channels or proper regularization [110]. Moreover, the trajectory and its undersampling pattern affect the g-factor characteristics. Spirals [39], variable-density spi-

rals [33] or Cartesian trajectories with extra lines for self-navigation [32, 34] provide crucial low-resolution signal for motion estimation. In contrast, the EPI interleaves acquire varying amounts of central k-space energy resulting in different degradations of the shot reconstructions. As an example, the single shots 0 to 5 in Fig. 3.9b contain about 25, 9, 8, 5, 26, 27 % of the total signal energy, which qualitatively matches the visual impression. Moreover, randomized sampling schemes enable supportive sparsity-enforcing regularization [45] of the joint problem in Eq. 3.6, but this is difficult to be realized for EPI.

In addition, partial Fourier acquisition allows for significant echo time reductions, which, at the same time, improves the feasibility of T_2-critical applications for abdominal diffusion. The rapid T_2 signal decay experienced, for example, in prostate DWI complicates image recovery for conventional EPI trajectories. Partial Fourier techniques shorten the k-space trajectory so that the echo top is reached earlier. The echo time reduction can help to increase SNR and reduces T_2 imprinting issues in DWI. Nevertheless, partial Fourier reconstructions require the signal phase to be slowly varying in phase-encoding direction. This assumption might be unjustified due to, for example, susceptibility variations, strong field inhomogeneities or flow [2] and must be reconsidered for each application scenario.

Iterative data rejection is a crucial reconstruction element to improve robustness by including only consistent shot data. As an example, the shot combination might become unfavorable for failed registrations, through-plane motion or contrast variations [22] when intra-shot motion occurs during the diffusion encoding process. For this purpose, the iterative SEDIMENT approach allows one to selectively include datasets for consistent reconstruction improving convergence and reliability.

Compared to AMUSE [43], which is mimicked by MC-SENSE+CG in the present work, the current SEDIMENT implementation for multi-shot DWI neglects the implications of rotational motion onto the diffusion contrast and is thus only valid for small rotations under the assumption of a smooth q-space. Considering the integration of SEDIMENT into a DTI framework, the refined motion estimates could be leveraged to achieve more accurate tensor estimates. However, the initial tensor estimation becomes more challenging for higher segmentations as well. Using the initial SENSE-based shot estimates, the g-factor penalty complicates appropriate initial tensor guesses. In contrast, the consistency-based integration into the iterative framework further increases the computational load and raises questions about the numerical stability of the tensor model. Hence, the integration of SEDIMENT into a DTI framework is subject to future research.

Currently, SEDIMENT neglects macroscopic motion-induced variations of the coil sensitivity maps by making several assumptions. The use of equal CSMs firstly implies that the electromagnetic properties at each point in space remain unchanged after small macroscopic object motion. The coil setup is thereby assumed to be at rest and independent of the body, in contrast to setups that are attached. Secondly, macroscopic motion disturbance of the low-resolution SENSE reference scan is neglected. Thirdly, the coarse SENSE reference scans cover the whole motion range of the subject within the coil. As an alternative approach, sensitivity estimation could be incorporated into the model-based images reconstruction [75].

The presented retrospective algorithms could further be fruitfully fused with prospective motion correction approaches [21]. Strong subject motion can corrupt the

DWI data in various ways stressing the implemented motion models. Prospectively navigated acquisitions reduce the prevalent artifacts in the data providing enhanced conditions for retrospective corrections.

Possible future extensions to SEDIMENT involve learning-based modules for rigid motion estimation, phase denoising and data rejection. Hitherto, the rigid registration and the phase estimation ignore the SENSE-related g-factor distribution over the image, which could be used to suppress areas of high noise propagation. Moreover, their implementation normally includes filters, thresholds and weighting factors that are manually determined and which are sensitive to SNR and, thus, the segmentation.

Feature-based deep learning modules could provide enhanced motion parameter estimates and rejection tools, which are separable from the joint image recovery. As an example, Bilgic et al. [77] proposed to jointly denoise the magnitudes of segmented multi-echo reconstructions using a neural network and to use the denoised magnitudes to regularize phase estimation using phase cycling [83]. The enhanced phase maps are then used for conventional reconstruction. In this way, the influence of neural networks is restricted to the estimation of motion parameters and erroneous estimates appear with well-defined artifact shapes in the conventional physics-based reconstruction. This modular inclusion opposes fully end-to-end approaches like Variational Networks [47] or AUTOMAP [46].

CONCLUSIONS

SEDIMENT is a SENSE-navigated iterative scheme that improves state-of-the-art multi-shot DWI reconstruction, correcting for motion-induced phase variations and rigid inter-shot in-plane motion. The continuously refined shot motion estimates enable the consistent combination of multiple shot datasets, thereby making high EPI segmentations feasible in the presence of macroscopic motion. The algorithm supports partial Fourier reconstructions and strategies for shot data rejection to boost SNR and robustness. The presented scheme provides an adjustable modular framework enhancing reconstruction quality and speed to ease clinical adoption of the method for DWI.

4

Navigated Multi-shot DWI

Navigated approaches derive the shot parameters of the optimization problem from additional MR signals that are integrated into the scan sequence. The navigator signals commonly involve benign reconstruction problems to access the shot parameters circumventing the joint estimations of image and shot parameters, which are generally non-convex and computationally demanding. As a major drawback, navigation reduces the scan efficiency through the additional sampling and it relies on the accuracy of the navigation signal for the shot parameter estimation. Nevertheless, navigated approaches for multi-shot DWI provide robust imaging solutions that are widely used for shot phase correction in clinical practice. This chapter extends the multi-shot DWI models to macroscopic motion correction and evaluates the potential of the shot navigator signals to address this challenging but important clinical problem of subject motion. The first section of this chapter investigates a navigated multi-shot DWI reconstruction including rigid in-plane corrections and compares the performance to the self-navigated approach of Section 3.2. The method presented in the second section integrates simultaneous multi-slice acquisitions into the multi-shot problem and leverages the data support in slice direction to estimate and correct for three-dimensional rigid motion.

4.1

Navigated Rigid In-plane Motion Correction

Macroscopic motion is a costly problem for clinical MRI and represents a serious image quality factor for multi-shot DWI, which is especially prone to shot-wise signal variations. This section investigates the potential of 2D navigators in multi-shot DWI acquisitions for joint shot phase and rigid in-plane motion corrections. A navigated SENSE-based reconstruction algorithm for multi-shot DWI is proposed that includes navigator-based corrections of the shot phase, rigid in-plane motion and optionally motion-induced shot contrast variations. The reconstruction furthermore utilizes a data rejection scheme to encounter through-plane motion and irreversible signal loss. The navigated method is compared to a self-navigated algorithm for 4- and 5-shot in-vivo DTI. The motion estimation from low-resolution navigator data enables high-quality image reconstructions comparable to self-navigated results but with faster reconstruction times. In the evaluations over multiple subjects and segmentations, the navigated reconstructions are found to be more robust against strong motion corruption and high segmentations.

This section is based on a conference submission to the 2020 ISMRM & SMRT Virtual Conference & Exhibition. (Abstract number: 4339, 2020, Ref. [50])

INTRODUCTION

The shot-to-shot phase variations in multi-shot DWI have been successfully studied and addressed in the last decades by a multitude of navigated [17, 18, 25–31], self-navigated [32–41] and navigator-free [40, 41] methods allowing to exploit the resolution and SNR gains of segmented acquisitions. Macroscopic motion during the lengthy DWI scans entails a challenging problem extension that has been approached just recently by self-navigated [43, 49] and navigated [44, 50] reconstruction models.

AMUSE [43] is a self-navigated method using SENSE-based [12] shot images from the interleaved data to estimate the shot phases, rigid in-plane motion and even diffusion tensors for shot contrast correction maps. Depending on the SENSE-based shot images, AMUSE suffers from the g-factor penalty [12] for increasing segmentations, which is reduced by the iterative SEDIMENT algorithm [49] at the cost of an increased computational burden. Dong et al. [44], in contrast, used a navigated acquisition to robustly estimate the shot phase and rigid in-plane motion parameters and proposed a motion-corrected GRAPPA-based [14] reconstruction. Nevertheless, comparisons of navigated and self-navigated methods are missing.

This work builds on the navigated IRIS method [31], short for image reconstruction using image-space sampling function, that uses 2D navigators to estimate the shot phase maps. The proposed algorithm, termed IRIS with rigid in-plane correction (IRIS-RiCo), extends the use of the navigators by rigid in-plane registration of the magnitude images and embeds the obtained rigid parameters into an efficient SENSE-based image reconstruction. The proposed navigated IRIS-RiCo algorithm is compared to the self-navigated SEDIMENT method [49] for in-vivo multi-shot DTI data. Furthermore, a navigator-based correction of the diffusion contrast variations from head rotations is described, but the resulting image quality improvements require further evaluation.

MODEL-BASED IMAGE RECONSTRUCTION

The general multi-shot DWI problem with macroscopic motion correction is given by Eq. 2.31. The method proposed here uses MR navigation signals to estimate the shot variation parameters of each shot s, i.e. the phase map for the shot phase operator Φ_s, the 2D rigid motion parameters for the macroscopic motion operator Ω_s and the weighting map for the contrast-reweighting operator Υ_s. The navigation reduces the demanding non-convex problem in Eq. 2.31 to a convex problem for the joint image vector $\boldsymbol{\rho}$:

$$\underset{\boldsymbol{\rho}}{\text{minimize}} \qquad \frac{1}{2} \sum_{s \in S'} \| M_s F C \Phi_s \Omega_s \Upsilon_s \boldsymbol{\rho} - \mathbf{d}_s \|_2^2 + \lambda \| \boldsymbol{\rho} \|_2^2. \tag{4.1}$$

A Tikhonov regularization with weight λ was added to stabilize the reconstruction. The set of included shots S' is determined by a correlation-based shot rejection [32]. In analogy to Eq. 2.33, the formulation with multi-shot operators is:

$$\underset{\boldsymbol{\rho}}{\text{minimize}} \qquad \frac{1}{2} \| \grave{M} \grave{F} \grave{C} \grave{\Phi} \grave{\Omega} \grave{\Upsilon} P \boldsymbol{\rho} - \grave{\mathbf{d}} \|_2^2 + \lambda \| \boldsymbol{\rho} \|_2^2, \tag{4.2}$$

which is summarized by the multi-shot forward operator $A = \grave{M} \grave{F} \grave{C} \grave{\Phi} \grave{\Omega} \grave{\Upsilon} P$ for given shot parameters in the following. P provides a joint image copy for each shot to match

the processing of the block-diagonal multi-shot operators indicated by grave accents. This model is solved separately for every diffusion direction q of N_q total directions.

This work compares two versions of the navigated algorithm with rigid motion correction, which are visualized in Fig. 4.1. The first version is termed *IRIS with rigid in-plane correction* (IRIS-RiCo) and neglects the shot contrast variations by settings $\Upsilon_\mathsf{s} = I_{N_p}$ with the identity matrix I_{N_p}. N_p is the number of pixels. The phase variations and the rigid shot alignment, on the other hand, remain active. The second version is termed *IRIS with rigid in-plane and contrast correction* (IRIS-CoCo) and involves all operators including Υ_s. Compared to IRIS-RiCo, IRIS-CoCo involves the additional navigator-based tensor estimation to estimate the shot contrast variations on low resolution, which is indicated by a blue color code and dashed arrows in Fig. 4.1.

IRIS WITH RIGID IN-PLANE CORRECTION (IRIS-RICO)

The description of the proposed IRIS-RiCo algorithm is separated into three parts: the *navigator reconstruction*, the *shot parameter estimation* and the *joint image reconstruction* from the interleaved high-resolution data.

Navigator reconstruction

The complex navigator data is upsampled to full resolution using a triangular k-space window to reduce Gibb's ringing [3]. The upsampled navigator data is then reconstructed using SENSE [12] with a Tikhonov regularization [24] yielding the shot navigator images $\boldsymbol{\rho}_\mathsf{s}$. The regularization weight was empirically set to $\lambda_{nav} = 10^{-2}$. As the coil sensitivity maps are assumed invariant to head motion, the inversion matrices of SENSE can be calculated once for all navigators of a slice.

Shot parameter estimation

The next step estimates the shot parameters from the navigator images $\boldsymbol{\rho}_\mathsf{s}$ to tune the multi-shot reconstruction involving *macroscopic motion* estimation, *phase extraction* and *data rejection*. The *phase extraction* obtains the shot phase maps $\boldsymbol{\phi}_\mathsf{s}$ by taking the pixel-wise argument of the complex-valued images. The shot phase operators are then constructed by $\Phi_\mathsf{s} = \mathrm{diag}(e^{j\boldsymbol{\phi}_\mathsf{s}})$ for each shot s.

The shot-wise *macroscopic motion* is parameterized by a rigid in-plane model. The three rigid motion parameters (two translations and one rotation) per shot are estimated from the shot magnitude images by rigid registration. The reference shot index s_{ref} for the registration is chosen once in advance by analyzing the mutual shot correlations in the unaligned state. The shot with the highest overall correlation to all other shots is chosen as the reference. The macroscopic motion operator Ω_s is implemented as proposed by Cordero Grande et al [86]. The rigid resampling employs a k-space formulation using the Fourier shift theorem for translations and a factorization into three shears [109] for rotations. The resampling is based on 1D-FFTs and requires a sufficiently large FOV to avoid wrap-around effects at the image edges for large in-plane motion.

The *data rejection* excludes significantly deteriorated shot signals from the joint image reconstruction based on a correlation measure [32]. For this, the corrected shot navigators $\hat{\boldsymbol{\rho}}_\mathsf{s}$ are calculated by $\hat{\boldsymbol{\rho}}_\mathsf{s} = \Upsilon_\mathsf{s}^H \Omega_\mathsf{s}^H \Phi_\mathsf{s}^H \boldsymbol{\rho}_\mathsf{s}$. Note again that IRIS-RiCo effectively drops the shot contrast-reweighting operator Υ_s by setting it to the identity $\Upsilon_\mathsf{s} = I_{N_p}$. The reference shot is selected once more as the one with the highest correlation to all other shots. The correlation threshold for shot rejection was set to 95%.

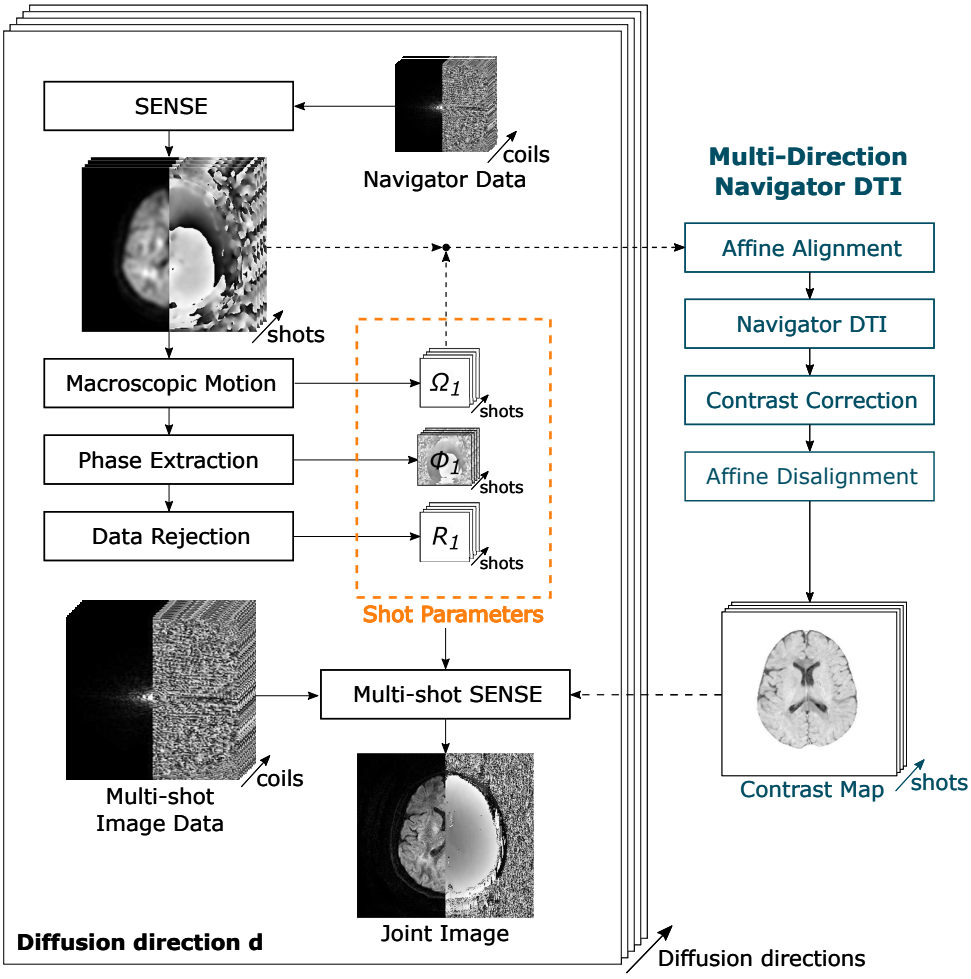


Figure 4.1: SENSE-based reconstruction scheme for multi-shot DWI with navigated motion and contrast correction. First, the moderately subsampled navigator images are recovered by SENSE [12]. Second, the navigators are used to calibrate shot-wise macroscopic motion parameters, phase maps and data rejection criteria. For optional shot contrast corrections (shown on the right-hand side with dashed lines and blue color code), navigator diffusion tensors are estimated from all diffusion directions after affine alignment to produce shot contrast-reweighting maps according to the shot rotation. The shot parameters including the optional contrast maps are used in the multi-shot SENSE model to reconstruct the joint image.

Joint image reconstruction

The joint multi-shot image reconstruction is performed individually for each diffusion direction using the conjugate gradients (CG) method [24]. The intensity correction operator $\tilde{I}$ proposed by Pruessmann et al. [13] is included for preconditioning. Compiling the multi-shot forward model A as in Eq. 4.2, the normal equations to solve by CG are:

$$(\tilde{I}A^H A\tilde{I} + \lambda\tilde{I}\tilde{I})\,(\tilde{I}^{-1}\boldsymbol{\rho}) = \tilde{I}A^H\mathbf{d}. \tag{4.3}$$

Note that the intensity correction is approximative in the presence of head movements, because motion leads to a shot-specific coil intensity imprint. However, the sensitivities are smooth and the static coil weighting was found to approximate the motion case well for preconditioning. With intensity correction, the CG effectively solves for $(\tilde{I}^{-1}\boldsymbol{\rho})$ so that the final image is recovered by: $\boldsymbol{\rho} = \tilde{I}\,(\tilde{I}^{-1}\boldsymbol{\rho})$. The regularization weight was empirically set to $\lambda = 10^{-3}$. The CG is stopped if either the residual norm [13] drops below $\tau_r = 10^{-6}$ or 30 iterations are exceeded.

Add-on: IRIS-RiCo with Contrast Correction (IRIS-CoCo)

Besides the macroscopic misalignment, subject rotations change the effective diffusion-weighting direction in the patient frame, leading to shot contrast variations. IRIS-CoCo estimates a tensor map from the low-resolution navigators and determines shot contrast-reweighting maps for the joint reconstruction. The extended processing is indicated in Fig. 4.1 by dotted arrows and the blue color. Rejected shots are excluded.

Shot contrast-reweighting

The estimation of shot contrast-reweighting maps consists of four steps. First, the navigator magnitudes are aligned using an affine in-plane registration. The affine motion operator of shot s is denoted by $\Omega_{A,\mathsf{s}}$. The additional zoom and shear parameters of the affine mapping are considered to encounter potential eddy current-induced distortions between the diffusion directions [9].

Second, the navigator DTI analysis is performed using all single-shot navigators ($N_\mathsf{q} N_i$ shots without rejection) to obtain a tensor estimate $\hat{\mathbf{D}}_\mathsf{p}$ for each voxel p. The diffusion direction $\mathbf{g}_\mathsf{s}$ of each shot s is corrected beforehand using the shot rotation parameters $\mathbf{g}_{R,\mathsf{s}} = T_{R,\mathsf{s}}^{-1}(\mathbf{g}_\mathsf{s})$, where the rotational transformation $T_{R,\mathsf{s}}$ is determined from the concatenated rigid and affine transformations. By this, the tensor estimation takes into account the corrected diffusion direction for each single-shot navigator avoiding biases in the tensor data from contrast mixing of the multi-shot datasets as in Ref. [44].

Third, the shot contrast correction maps $\boldsymbol{\upsilon}_\mathsf{s}$ are calculated per voxel p as in Eq. 2.28 using the tensor estimate $\hat{\mathbf{D}}_\mathsf{p}$ and the shot-wise directional difference $\tilde{\mathbf{g}}_\mathsf{s} = T_{R,\mathsf{s}}^{-1}(\mathbf{g}_\mathsf{s}) - \mathbf{g}_\mathsf{s}$:

$$(\boldsymbol{\upsilon}_\mathsf{s})_\mathsf{p} = e^{-b\;\tilde{\mathbf{g}}_\mathsf{s}^T\,\hat{\mathbf{D}}_\mathsf{p}\,\tilde{\mathbf{g}}_\mathsf{s}} \tag{4.4}$$

Finally, the adjoint affine mapping $\Omega_{A,\mathsf{s}}^H$ transforms the contrast correction maps $\boldsymbol{\upsilon}_\mathsf{s}$ back to be in alignment with the joint image. The contrast-reweighting operators are constructed by $\Upsilon_\mathsf{s} = \mathrm{diag}(\Omega_{A,\mathsf{s}}^H \boldsymbol{\upsilon}_\mathsf{s})$. The contrast-reweighting operators are then employed in the joint image reconstruction as part of the forward model A in Eq. 4.3.

Methods

Data acquisition and preprocessing

Multi-shot echo-planar DTI data with 4 and 5 shots were acquired using a 2D-navigated Stejskal-Tanner diffusion sequence [31] as presented in Fig. 2.5. The data was obtained from 8 healthy volunteers on a 3T Philips Ingenia Scanner using a 13-channel head coil. Informed consent was obtained according to the rules of the institution. Spectral pre-saturation with inversion recovery (SPIR) was used for fat suppression [93]. Relevant scan parameters are listed in Table 4.1.

Table 4.1: Sampling parameters of the diffusion tensor imaging scans.

Parameters	**DTI scans**	
#interleaves N_i	4	5
TR	5000 ms	5000 ms
TE image \navigator echo	68\139 ms	66\126 ms
FOV (R × P × S)	$232 \times 228 \times 16\,\text{mm}^3$	$232 \times 225 \times 16\,\text{mm}^3$
Resolution image	$1.0 \times 1.0 \times 4.0\,\text{mm}^3$	$1.0 \times 1.0 \times 4.0\,\text{mm}^3$
Resolution navigator	$5.0 \times 5.0 \times 4.0\,\text{mm}^3$	$5.0 \times 5.0 \times 4.0\,\text{mm}^3$
#slices N_z	4	4
Partial Fourier factor	0.614	0.622
b-value	$1{,}000\,\text{s/mm}^2$	$1{,}000\,\text{s/mm}^2$
#diffusion directions N_d	15	15
#T_2-weighted acquisitions	2	2
Scan time [m:ss]	5:50	7:15

The DTI experiments described in Table 4.1 were done twice with different motion-related instructions. In the first scan, the volunteers were asked to remain still. For the second scan, they were asked to perform continuous in-plane motion with a "no" head shake trajectory. A gradient-echo prescan was used to acquire coil sensitivity maps once in advance [12]. EPI ramp samples were gridded to the Cartesian sampling position prior to the proposed reconstruction and, to reduce Nyquist ghosting, EPI reference data was acquired for odd/even phase correction [94].

EXPERIMENTAL DESIGN

The image reconstructions were performed on a system with a 2.7 GHz Intel Core i7 4-core CPU and 16 GB RAM. The algorithms were implemented in Python 3.6.9. IRIS-RiCo and IRIS-CoCo were compared to two reference algorithms to evaluate the macroscopic motion correction performance for the in-vivo multi-shot DTI data. The first reference method was the basic IRIS method [31], which performs the shot phase corrections while neglecting macroscopic motion. The phase correction and data rejection are performed as described for the proposed methods in this work. The second reference method is the self-navigated SEDIMENT algorithm [49] presented in Section 3.2, which also performs rigid in-plane and shot phase corrections.

In this work, the evaluation of the navigated and self-navigated methods is based on a comparison of the image reconstruction results for static and the associated motion-disturbed datasets. Therefore, the final images of all methods are registered to the high-resolution IRIS reconstruction of the static T_2-weighted case ($b = 0\,\text{s/mm}^2$). The final alignment uses an affine multi-scale registration with a mutual information metric and is resampled by linear interpolation. The diffusion gradient directions for the DTI processing are corrected for rotations of the concatenated joint affine alignment and the mean rigid parameters of the shots involved in the multi-shot image reconstruction.

The performance of the rigid motion correction was compared qualitatively by analyzing the DWI image results of the different methods. Tensor estimates and FA maps were calculated using *Dipy* [68]. A quantitative nRMSE assessment is provided by comparing the reconstructed images under motion conditions with the associated static images from the same algorithm to avoid biases towards a specific algorithm. It should be noted that the evaluation under motion conditions is challenging. The proposed comparison with a static reference suffers from inevitable through-plane motion that occurs during the macroscopic motion-disturbed scans. Therefore, the validity of the static images as reference data is limited. This has also impeded evaluations of the subtle contrast corrections in IRIS-CoCo, which are superimposed by the anatomical variations due to through-plane motion in this study. However, the static data still serves as a sufficiently accurate reference for the nRMSE evaluation of the DWI images.

Implementation Details

The sensitivity and Fourier operator implementations follow the descriptions of SEDIMENT [49]. The sensitivity operator C involves masking and coil compression [95]. The mask is obtained from a body coil reference image by thresholding the magnitude image at 10% of its maximum. Then, the mask is processed by 10 binary closings and 5 dilations. The coil compression uses a principal component analysis with a 99% threshold.

The aliasing for EPI manifests in phase encoding direction so that the PSF is one-dimensional without support in the readout direction. Therefore, the EPI data is brought to hybrid k_y-x-space and the Fourier operator F involves only 1D-FFTs in phase encoding direction. Without partial Fourier sampling, the undersampling can be completely represented in image-space, which can further reduce the computational demands [49].

For partial Fourier undersampling, the proposed *joint image reconstruction* of the interleaved data is embedded into a homodyne two-step strategy [2]. In the first step, the image reconstruction is performed from the low-resolution k-space center, which is symmetrically covered by the sampling. For this, the data is downsampled by a k-space window function with 1D-Hann shape in phase encoding direction. The low-resolution multi-shot problem is solved by CG and the phase map is extracted yielding the low-resolution object phase. Second, the full interleaved data is pre-weighted by an asymmetric s-shape window in k-space [2] followed by the CG reconstruction. The s-shape window obtains a uniform weighting of the signal partition with conjugate even symmetry (Eq. 2.7) in k-space, which is related to the real-valued signal in image space by Fourier symmetry properties. The final image is then recovered from the joint image of the second step by eliminating the low-resolution phase from the first step and taking the real part. This processing enforces the conjugate even symmetry in k-space after the image-space correction of the estimated low-resolution phase.

The rigid and affine navigator registrations employ the preregistration of the fast elastic image registration (FEIR) framework [107] with a normalized gradient field metric [85]. After the rigid registration, the motion states are centered within each multi-shot experiment by subtracting the mean from the rigid shot motion parameters.For the affine navigator alignment in IRIS-CoCo, the registration reference is the first shot of the non-DWI ($b = 0$ s/mm^2) images. The affine alignment uses linear interpolation, which is the default processing of *Dipy* [68] for tensor estimation.

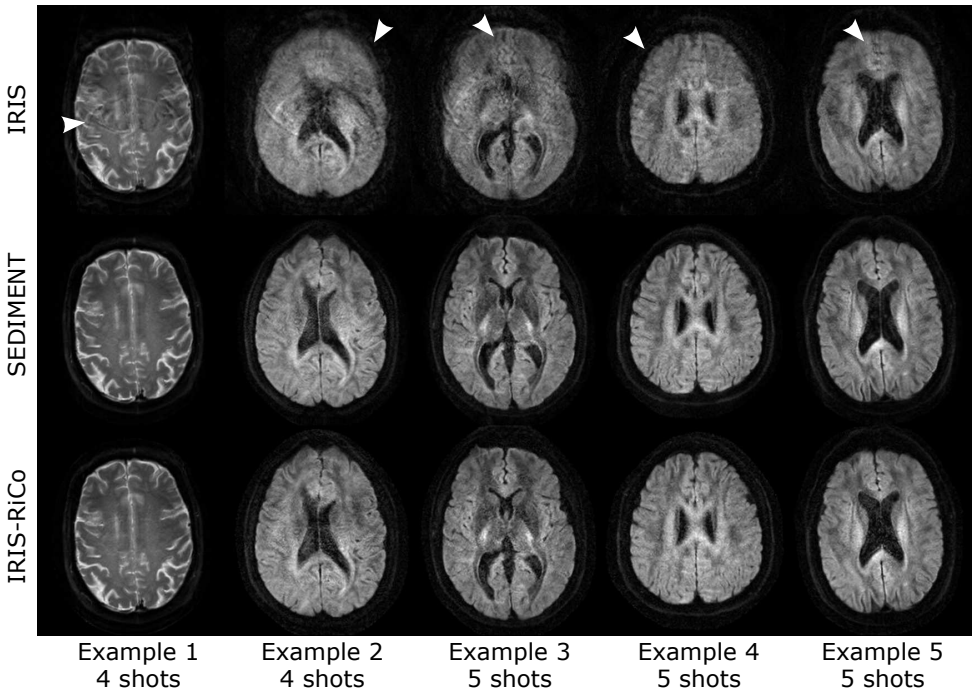


Figure 4.2: Five in-vivo multi-shot DWI examples reconstructed with the different motion correction algorithms. (Top row): IRIS corrects for shot-to-shot phase variations only. (Middle row): SEDIMENT applies self-navigated shot phase and rigid in-plane motion correction. (Bottom row): IRIS-RiCo performs navigated shot phase and rigid in-plane motion correction. All methods suppress ghosting artifacts from the shot phase variations. The IRIS results contain strong image artifacts from macroscopic subject motion (white arrows), whereas SEDIMENT and IRIS-RiCo reduce the artifacts by rigid in-plane modeling in the reconstruction. The method is applicable to both non-DWI (left column) and DWI (other columns) multi-shot data.

RESULTS

MOTION CORRECTION FOR IN-VIVO DWI

Figure 4.2 shows five DWI examples from different subjects to qualitatively assess the image reconstructions under gross motion conditions. IRIS [31] (upper row) performs no rigid motion correction, while SEDIMENT [49] (middle row) uses a self-navigated and IRIS-RiCo (bottom row) the proposed navigated rigid motion correction approach.

All reconstruction methods employ shot phase corrections leading to DWI results that are free from the characteristic phase-induced ghosting artifacts and signal dropouts. The IRIS results comprise strong blurring artifacts (white arrows) for the macroscopic motion-corrupted datasets, which are well resolved by the self-navigated SEDIMENT and the navigated IRIS-RiCo algorithm, which both perform 2D rigid motion corrections. The results include successful image reconstructions of both T_2-weighted (non-DWI) and DWI data with different segmentation factors.

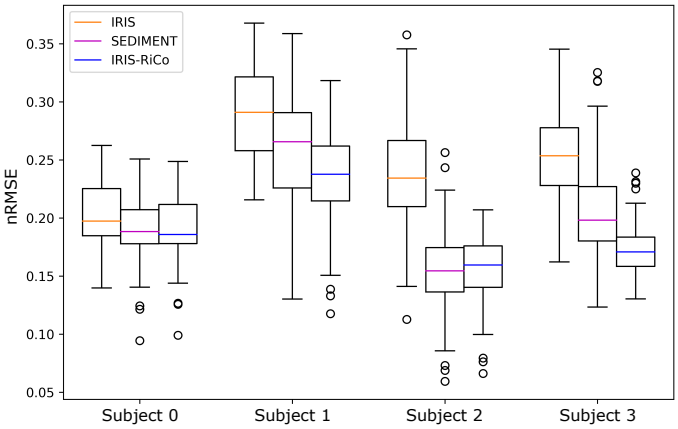


Figure 4.3: Normalized root-mean-square error (nRMSE) statistics per subject comparing the image reconstruction methods under macroscopic motion conditions. The nRMSE was calculated with respect to the reference images under static conditions after rigid alignment of the datasets. One subject was excluded from the statistics, because the table top moved between the scans under static and motion conditions causing a through-plane offset. Both rigid in-plane corrected methods improve on the IRIS results, whereby IRIS-RiCo produces lower median values in three of four subjects and generally shows smaller inter-quartile ranges.

Image Quality Evaluation for in-vivo DWI

The nRMSE performance of the self-navigated SEDIMENT [49] and the navigated IRIS-RiCo methods are quantitatively compared over all DWI reconstructions in Fig. 4.3. During the DTI processing, all reconstructed images were aligned to the first T_2-weighted IRIS image under static conditions by rigid image registration. The nRMSE is then calculated for the aligned image magnitudes and compares each motion-disturbed image to the respective image under static conditions, which is reconstructed by the same algorithm (IRIS, SEDIMENT, IRIS-RiCo) to avoid biases towards specific algorithms. The evaluation is visualized per subject to analyze the performance separately for different volunteers with their respective motion profiles. Please note that the availability of reference data and thus the nRMSE quantification are problematic and should only indicate overall trends.

The rigid in-plane corrected methods SEDIMENT (purple) and IRIS-RiCo (blue) consistently reduce the nRMSEs over all subjects compared to the IRIS method [31] (orange), which neglects macroscopic motion. The median value is lower for IRIS-RiCo than for SEDIMENT in three of four subjects. Moreover, the navigated IRIS-RiCo shows less variation in the image reconstruction results, as measured by the inter-quartile ranges of the box-plots, compared to the self-navigated SEDIMENT.

The reconstruction times per slice of the navigated IRIS-RiCo method were on average approximately 15 s, while IRIS-CoCo with its additional navigator DTI evaluation took 17 s. The self-navigated SEDIMENT algorithm, which optimizes the shot parameters and the joint image in an alternating fashion, required on average 52 s and took 55 iterations. Neglecting macroscopic motion in the reconstruction model, IRIS allows for an efficient image-space solution that took only 3.5 s on average, but it shows compromised image quality for motion-disturbed datasets.

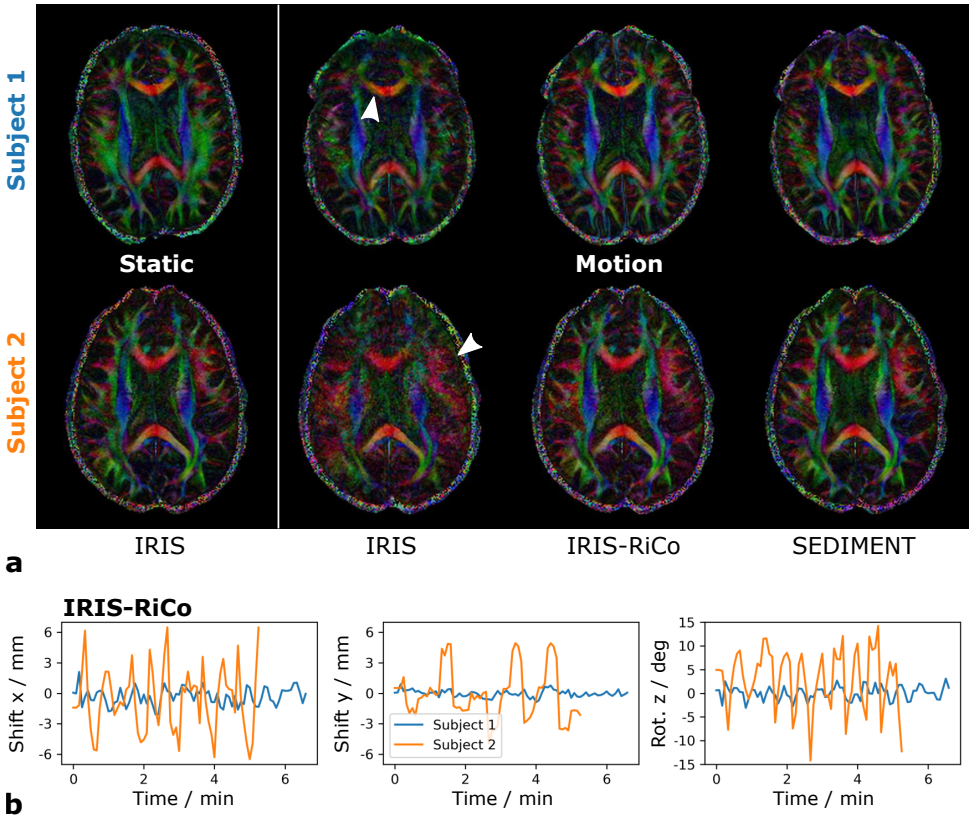


Figure 4.4: In-vivo DTI reconstruction results for two gross motion-corrupted datasets compared to static reference data. (a): Colored FA maps for a static (left column) and a motion-corrupted case (other columns), reconstructed once with IRIS, the proposed navigated IRIS-RiCo and self-navigated SEDIMENT. b: Rigid motion estimates of IRIS-RiCo for both subjects. Subject 1 performed minor gross motion (±1 mm, ±2 deg), while Subject 2 moved heavily throughout the scan (±6 mm, ±14 deg) causing strong gross motion artifacts for IRIS (white arrows). Both the navigated and the self-navigated method with rigid in-plane motion correction recover great parts of the structural information. Please note that the static and motion cases might show slightly different structures due to through-plane motion between the scans.

MOTION CORRECTION FOR IN-VIVO DTI

Figure 4.4 shows the different DTI reconstruction results for two motion-disturbed datasets in comparison to a static reference. Colored fractional anisotropy (FA) maps for two subjects are presented in Fig. 4.4a, along with the estimated rigid motion parameters of IRIS-RiCo in Fig. 4.4b.

IRIS provides decent image quality for the static datasets with high SNR and sharp anatomical delineation. Under macroscopic motion conditions, the IRIS reconstructions suffer from strong artifacts that propagate into the tensor estimates and severely degrade the FA maps. Both rigid in-plane motion-corrected methods successfully reduce the motion artifacts in the underlying DWI images recovering the structural information and a reasonable SNR level under continuous subject motion throughout the whole scan.

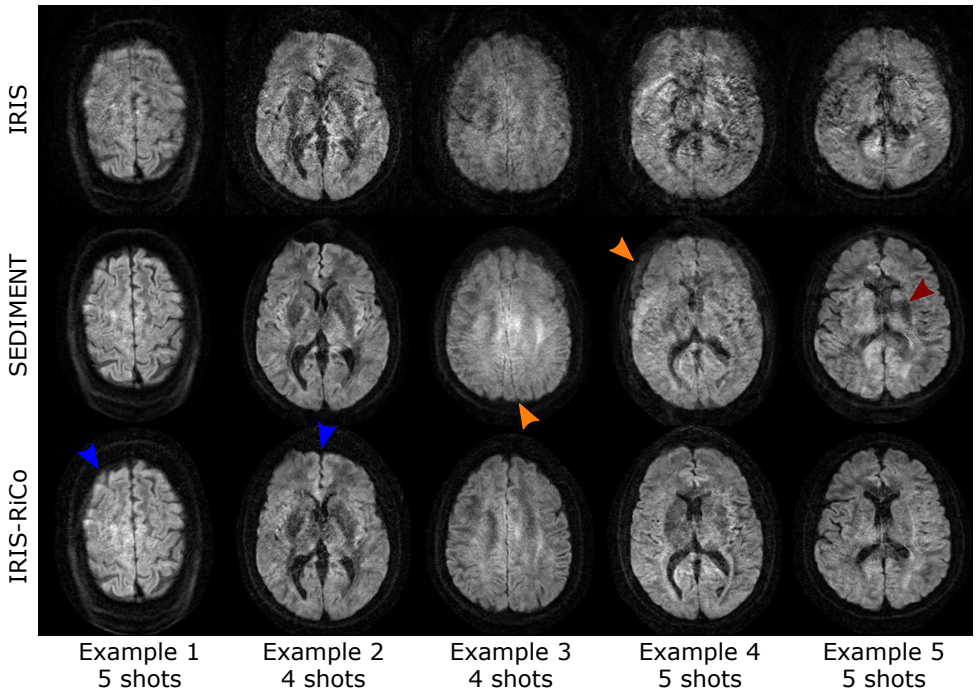


Figure 4.5: Residual artifacts for the motion-corrected multi-shot DWI reconstructions in challenging in-vivo examples. IRIS-RiCo and SEDIMENT generally improve the visual appearance of the images, but residual artifacts might persist in the reconstructed images. Some IRIS-RiCo reconstructions show minor residual blurring artifacts (blue arrows). SEDIMENT reconstructions might not be able to resolve the strong shot misalignments (orange arrows) in challenging motion cases and signal cancellation artifacts caused by residual inconsistencies of the self-navigated shot phase estimation might persist.

Residual Artifact Evaluation

Macroscopic motion estimation is a challenging non-convex task, especially in the low-SNR diffusion realm. Moreover, the signal perturbations might not be sufficiently captured by the employed rigid in-plane model. Figure 4.5 gives an overview of residual artifacts that persist in the reconstructions for SEDIMENT and IRIS-RiCo.

The IRIS reconstructions in Fig. 4.5 show 5 slice examples with different types of macroscopic motion corruption. Example 1 and 2 contain relatively small initial motion corruptions and are well recovered by SEDIMENT, whereas IRIS-RiCo comprises minor residual blurring of the fine structures (blue arrows). Nevertheless, the reconstructed images of IRIS-RiCo are substantially improved compared to IRIS. Example 3 and 4 are heavily affected by gross motion in the IRIS reconstructions. Here, IRIS-RiCo greatly reduces the image artifacts so that the inter-hemispheric fissures appear well-resolved. For these challenging cases, SEDIMENT is not able to estimate appropriate rigid shot parameters, so that several misaligned shot ghosts remain in the resulting images (orange arrows). In addition, insufficient shot phase estimates of SEDIMENT lead to residual signal dropouts in some cases (red arrows).

Within the studied data, 3.75% of the IRIS-RiCo images appeared with residual blurring artifacts as indicated by blue arrows in Fig 4.5. SEDIMENT was prone to stronger misalignment artifacts as indicated by the orange arrows that occured in 6.25% of the cases, especially for the higher 5-shot segmentation.

CONTRAST CORRECTION EXAMPLE

IRIS with rigid in-plane and contrast correction (IRIS-CoCo) employs the low-resolution navigators for the challenging diffusion tensor pre-estimation from motion-corrupted data. However, the differences between IRIS-RiCo and IRIS-CoCo are subtle and, at the same time, the requested in-plane assumption of the motion study design is not perfectly possible leading to through-plane motion components in practice. Reliable ground truth reference data for a precise contrast evaluation is thus missing and the adverse effects are biasing the small contrast corrections. Therefore, Figure 4.6 shows a reconstruction example of IRIS-CoCo with navigated contrast correction. A thorough evaluation is subject to future work.

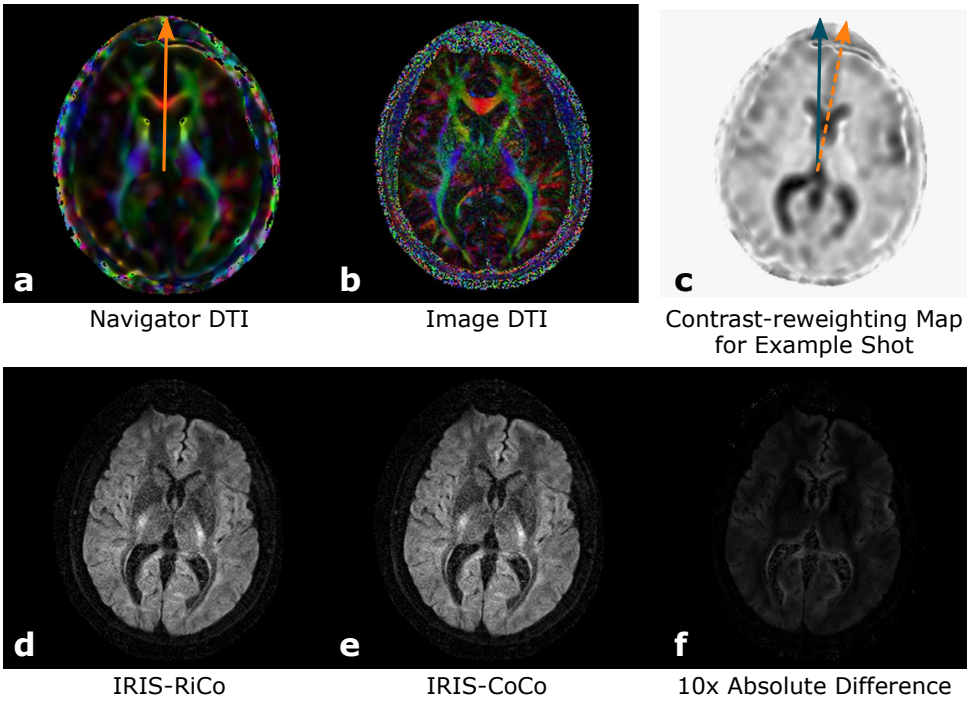


Figure 4.6: Example of a navigator-based shot contrast correction for 4-shot DTI data. (a) Low-resolution navigator FA map. (b) High-resolution image-echo FA map from IRIS-CoCo. (c) Contrast correction map derived from the navigator tensor estimates for a shot with approximately 10 degrees clockwise in-plane rotation. By this, an intended diffusion gradient direction along the inter-hemispheric fissure in the static patient frame (orange arrows) effectively weights the anatomy in the rotated shot frame (blue arrow). (d, e) The associated multi-shot reconstructions without and with contrast correction, respectively. (f) The 10-fold magnified difference image reveals variations below 5% signal magnitude, especially in highly anisotropic brain regions. The accuracy evaluation of this approach in simulations is subject to future research.

The navigator-based fractional anisotropy map (Fig. 4.6a) captures the structures of the image-echo FA map (Fig. 4.6b) reconstructed by IRIS-CoCo at a coarse level. The navigator tensor estimates are used to calculate contrast-reweighting maps for each shot according to the rotation parameters (Fig. 4.6c), which are then used in the high-resolution image-echo reconstruction of IRIS-CoCo (Fig. 4.6e). The resulting corrections are below 5% of the image magnitudes even for this 4-shot dataset with considerable in-plane rotations of approximately 10 degrees. In addition to the relatively small contrast variations resulting from the rotations that are feasible within a head coil, the contrast errors are averaged over all shots in a segmented acquisition.

Discussion

IRIS-RiCo demonstrates the feasibility of navigated multi-shot brain DWI in the presence of shot-to-shot phase variations and macroscopic in-plane motion and allows for fast image reconstructions. The low-resolution navigators enable robust estimations of shot phases, rigid in-plane motion and shot rejection criteria. The proposed navigated method achieves comparable image quality with state-of-the-art self-navigated image reconstructions, while showing improved robustness in challenging imaging situations, such as heavy motion corruption or high segmentation. However, a direct comparison of navigated and self-navigated reconstructions is challenging due to the lack of reliable reference data and depends on the respective sequence implementations.

Navigation versus Self-navigation - A Difficult Comparison

IRIS-RiCo uses single-shot low-resolution navigators to estimates the shot phase and rigid motion parameters that are required for the multi-shot DWI reconstruction model. Navigation prolongs the scan time, but provides valuable low-resolution information for each shot of the interleaved scan and, by this, circumvents demanding alternating optimizations. In contrast, self-navigated methods like SEDIMENT [49] estimate the shot parameters from the undersampled image-echo interleaves and, thus, depend on the conditioning of the undersampled problem, which is further complicated by the non-convexity of motion estimation. With increasing segmentation, the noise propagation dramatically spoils the self-navigated estimation of phase maps and rigid parameters requiring many iterations to achieve acceptable image quality, whereas the navigator unfolding is independent of the image-echo segmentation.

A comparison of navigated and self-navigated methods depends on the effective sequence properties and is thus specific to the use case. Besides several sampling-specific factors, the SNR of the shot navigators depends on the achievable echo time. This, in turn, is bounded by the duration of the diffusion-weighting process, which is related to its strength and the employed gradient system, and the image-echo sampling, which might employ partial Fourier or other acceleration strategies. Therefore, the navigator quality varies with the sequence settings for a DWI scan and influences the relative performance of navigation and self-navigation. Especially for non-brain applications like abdominal DWI, the fast signal decay precludes late navigator samplings with high echo times. In the same way, self-navigation is affected by the image-echo time and the effective shot undersampling of the interleaved scan. Hence, the presented results do not directly generalize to other sequence setups and anatomies.

Another challenging problem is the reliability of the reference data for the method comparison in the presence of motion. In this work, the image reconstructions of macroscopic motion-disturbed data were compared to results from static data. However, the instructed in-plane motion profile is commonly mixed with through-plane components in practice, which corrupt the image quality assessment and spoil the alignment of the static and motion-corrupted data. Thus, the image quality measures must be interpreted with care in these settings. As an alternative, simulations could create corrupted data from clean static image reconstructions giving access to the ground truth for the image quality assessment. Nevertheless, the simulations require a realistic matching of the image-echo and navigator-echo SNR properties for the specific setup.

NAVIGATED VERSUS SELF-NAVIGATED IMAGE RECONSTRUCTION

The properties of the navigation and self-navigation signals have various effects on the shot parameter estimation and, by this, influence the characteristics of the image reconstruction results. As outlined in Table 2.1, navigated methods commonly involve one single shot parameter estimation from the reconstructed navigators. Some self-navigated methods also use one shot parameter estimation step, but for segmentations of four and above alternating optimizations as for SEDIMENT [49] are normally beneficial to achieve high image quality.

The navigated IRIS-RiCo algorithm achieved relatively robust shot parameter estimates obtaining reasonable rigid shot motion parameters even for challenging datasets. The shot navigators are reconstructed from single-shot acquisitions with moderate undersampling and equal trajectories providing reliable shot navigators even for high image-echo segmentation. The shot phase extraction and the rigid registration therefore remain relatively benign problems leading to robust image quality improvements with less variability as discussed for Fig. 4.3. As a major disadvantage, the navigated processing assumes that the shot parameter estimation is sufficiently captured on a coarse resolution. In some occasions, minor residual blurring therefore remained in the reconstructed images as presented in Fig. 4.5. Nevertheless, the low-resolution navigators have overall proven valuable capturing the shot phases and rigid motion profiles with sufficient accuracy in the presented study.

In contrast, the self-navigated SEDIMENT algorithm [49] alternately optimizes the shot parameters and the joint image on the full resolution. By this, the algorithm shows good convergence properties for minor residual misalignments persisting between the shots, so that the repeated image registrations resolve small blurring artifacts well. As the shot models comprise relatively high undersampling, the self-navigated shot parameter estimation suffers from noise propagation effects and therefore crucially depends on a reliable parameter initialization for the non-convex problem. Moreover, the optimization tries to resolve the superimposed phase-related and misalignment-related artifacts at the same time. Hence, SEDIMENT requires many iterations for challenging datasets and phase artifacts might remain in the image reconstructions (see Fig. 4.5).

The navigated IRIS-RiCo method provides faster reconstructions with comparable image quality to SEDIMENT at the cost of a less efficient sampling. Future modeling frameworks could combine the advantages of both approaches using navigator-based reconstructions to initialize the self-navigated algorithm if residual artifacts persist.

Evaluation of the Contrast Correction

The contrast correction by IRIS-CoCo was presented in a qualitative manner and requires further evaluation to confirm the benefits of the navigator-based tensor estimation. Comparisons to IRIS-RiCo contained signal magnitude differences below 5% even for strong in-plane rotations above 10 degrees, as the variations are subtle and the multi-shot combination further averages the contrast variations. The single-shot navigators provide unbiased contrast properties and require only a correction of the diffusion direction for motion-corrected tensor estimates. Although the low-resolution navigators might not capture fine structural variations, it at least includes the anisotropic structures on a coarse level for this challenging problem.

In the current study, the evaluation with respect to the static data was impeded by the presence of through-plane motion, which overlaid the subtle corrections by anatomical through-plane variations. To overcome the reference data problem, future projects could employ simulations or extended motion models to cope with through-plane motion and achieve meaningful contrast correction evaluations.

Model Limitations

The current model includes only in-plane macroscopic motion and makes several assumptions about its effects on the signal encoding and the data sampling. The rigid in-plane model provides an effective though rather simple model for the complex implications of motion onto the MR sequence [22]. The proposed framework deals with through-plane motion only by data rejection excluding shots from the multi-shot reconstruction that drop below the correlation threshold. The same similarity-based rejection can identify compromised shots that suffer from significant signal loss caused by intra-shot motion during the sensitive DWI sequence. As an important alternative or extension, prospective methods [21] allow for online rigid motion corrections of the sampling [80] and the diffusion gradients [88], for example, based on cameras [111].

Besides the direct effects, the imaging signals are affected by several secondary motion artifacts that spoil the image reconstruction and the shot image registrations. Geometric distortions caused by off-resonance or susceptibility variations mainly manifest in the phase encoding direction, which changes under motion conditions. Shot combinations and similarity metrics for image registration suffer from these shot-wise signal variations, which could be addressed by additional B_0 map modeling [112]. Moreover, the coil sensitivities are assumed to be invariant to head motion and sufficiently covered by a coarse gradient echo prescan, which is less affected by geometric distortions than EPI. To overcome these limiting factors, the sensitivity information could be incorporated into the model [75] and by this derived from the EPI data itself.

Conclusion

The proposed navigated multi-shot DWI approach efficiently recovers high-quality DWI images and represents a robust alternative to self-navigated algorithms, especially for high segmentations and strong in-plane motion. The low-resolution navigators have proven useful to capture and monitor motion-induced shot variations and provide beneficial data support to address the demanding non-convex motion problems of DWI.

4.2

Simultaneous Multi-slice and 3D Rigid Motion Correction

Besides in-plane rigid motion, through-plane motion components pose an important and challenging issue in clinical practice. This section aims at improving the robustness of diffusion-weighted imaging data acquired with segmented simultaneous multi-slice echo-planar imaging against in-plane and through-plane rigid motion. The proposed algorithm incorporates a 3D rigid motion correction and wavelet denoising into the image reconstruction of segmented SMS-EPI diffusion data. Low-resolution navigators are used to estimate 3D rigid motion parameters by simultaneous-multi-slice-to-volume registration and the shot diffusion phase corruptions. The shot-wise rigid motion and phase parameters are then integrated into a SENSE-based full-volume reconstruction for each diffusion direction. The algorithm is compared to a navigated SMS reconstruction without gross motion correction in simulations and in-vivo studies with 4-fold interleaved 3-SMS diffusion tensor acquisitions. Simulations demonstrate high fidelity achieved in the simultaneous-multi-slice-to-volume registration, with sub-millimeter registration errors and improved image reconstruction quality. In-vivo experiments validate successful artifact reduction in 3D motion-compromised in-vivo scans with a temporal motion resolution of approximately 0.3 s. This work demonstrates the feasibility of retrospective 3D rigid motion correction from shot navigators for segmented simultaneous multi-slice DWI.

This section is based on a manuscript under revision submitted to MRM [51].
(revision submitted on Feb 24, 2021)

INTRODUCTION

Multi-shot imaging [15, 16] is a promising technique for DWI to overcome the shortcomings of single-shot EPI, but it increases the sensitivity to shot-to-shot variations. Besides the phase variations [17–19], the lengthy DWI scans are prone to macroscopic inter-shot motion that has so far been addressed only by in-plane rigid modeling in retrospective motion-corrected image reconstructions using either navigators [44, 50] or self-navigation [43, 49]. However, through-plane motion poses an important practical problem for image quality and also for intelligent patient tracking to enable adaptive sampling approaches. Prospective methods [21] have already been proposed for 3D rigid motion-corrected multi-shot DWI using camera-based extra-navigation [80, 88], but they require extra hardware and depend on the sensor accuracy. Therefore, retrospective corrections offer an important alternative, which represents a challenging estimation task for multi-shot DWI requiring data support in the slice dimension.

Besides the in-plane accelerations through multi-shot imaging, simultaneous multi-slice [54] (SMS) approaches offer signal support in the slice direction and scan time reductions without an immediate loss of SNR. Controlled aliasing in parallel imaging (CAIPI) techniques [61, 62] improved the coil encoding efficiency and were successfully combined with EPI by blipped-CAIPI [63] and SENSE [66, 67]. Combining SMS and multi-shot DWI, several algorithms have been proposed for shot-to-shot phase corrections with navigated [64], self-navigated [79, 113] and navigator-free [82] approaches. Further macroscopic motion corrections have been developed using external tracking devices and intermediate non-diffusion-weighted navigators [81]. Recent retrospective approaches exploit SMS acceleration for single-shot EPI to obtain signal support in the slice direction and propose SMS-to-volume registration for 3D rigid motion-corrected fMRI [114, 115] and diffusion tensor fitting [116].

This work extends the navigated IRIS approach for segmented DWI [31] to SMS [64] and investigates the potential of 3D rigid motion-corrected DWI reconstructions. IRIS [31] uses 2D low-resolution navigators to sense the shot-to-shot phase variations and integrates them into an efficient SENSE-based [12] algorithm. The proposed algorithm, termed motion-aware SMS-accelerated and interleaved image creation (MoSaIC) for DWI, estimates shot phase maps and 3D rigid motion from low-resolution SMS navigators and reconstructs motion-corrected DWI volumes per diffusion direction. Furthermore, an SMS extension for the IRIS [31] algorithm, termed SMS-IRIS, is implemented serving as a phase-navigated reference algorithm without rigid motion correction.

MODEL-BASED IMAGE RECONSTRUCTION

SEGMENTED SIMULTANEOUS MULTI-SLICE DWI SAMPLING

The proposed method is based on the navigated Stejskal-Tanner spin-echo sequence for DWI [31] shown in Fig. 2.5. Here, this sequence is extended to SMS and blipped-CAIPI encoding as shown in Fig. 4.7, following the descriptions by Dai et al. [64] (see also Fig. 2.4 on blipped-CAIPI SMS encoding). The first spin echo samples the interleaved high-resolution *image echo*, while a second spin echo acquires a low-resolution *navigator echo* at a lower undersampling rate. The echo-spacings are adjusted to match the off-resonance induced distortions in phase encoding direction for both samplings.

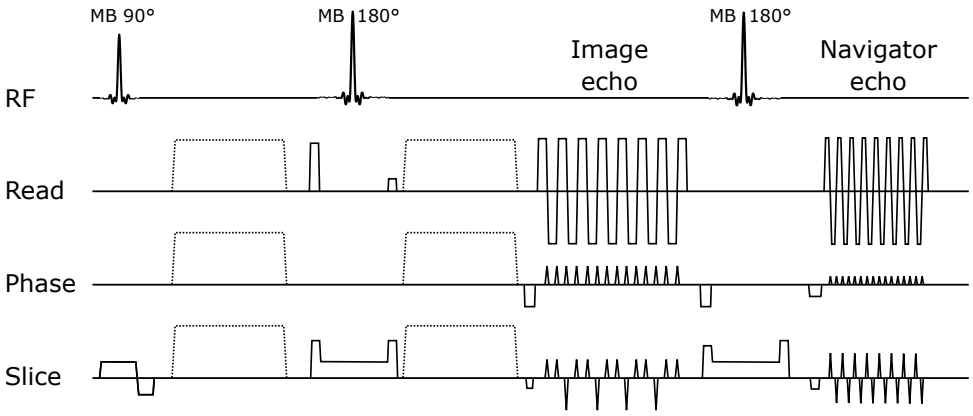


Figure 4.7: Stejskal-Tanner spin-echo sequence diagram for navigated multi-shot DWI with SMS extension. The sequence uses multi-band radio frequency (RF) excitation (MB 90°) and refocusing (MB 180°) pulses to form the image and navigator echo for SMS acquisition. The dotted lines indicate the diffusion-sensitizing gradients. In-plane accelerations are achieved using interleaved EPI for the image echo and single-shot EPI with moderate regular undersampling for the navigator echo. Slice gradient blips are added for blipped-CAIPI encoding with slice-specific shifts, which significantly improve the coil encoding capabilities. The image echo is sampled with a FOV/3 shift between adjacent slices, while the navigator slices are shifted by FOV/2.

The reconstruction problem for segmented SMS DWI under shot-to-shot phase variations und macroscopic motion is shown in Figure 4.8. The high-resolution SMS images per interleave in Fig. 4.8b were produced from an in-vivo case using the motion-informed forward model. The shot phase and macroscopic motion variations cause severe ghosting and blurring for the motion-unaware 2D-SENSE [67] in Fig. 4.8c, while MoSaIC suppresses the artifacts through shot phase and 3D rigid motion estimation.

Model formulation for 3D motion corrected DWI

The model extension to address through-plane motion requires a full-volume reconstruction framework as used in Ref. [87], rather than the standard slice-by-slice reconstruction. The slice positions of an SMS stack in the scanner frame are determined by the slice gradients and the RF excitation, being thus independent of subject motion in the first place. For pure in-plane motion, the anatomies therefore stay in the same SMS stack as long as the object remains within the FOV. For through-plane motion, new anatomies enter the slice positions decoupling the strict assignment of anatomies to SMS slice stacks so that a full-volume reconstruction is required to resolve the motion. The model furthermore assumes that the signal encoding remains unchanged in the scanner frame and is explicitly unaffected by macroscopic subject motion. Thus, the spatial profiles of the coil sensitivities stay valid during the whole scan.

To facilitate the model description, some notations are given here in advance. N_x, N_y and N_z are the number of voxels in read-out, phase, and slice encoding direction giving in total $N_\rho = N_x N_y N_z$ voxels. The volume is sampled with N_c coils, N_{MB} simultaneously acquired slices and N_i interleaves. A full volume is covered by $N_g = N_z / N_{MB}$ SMS slice group excitations. Each of the N_d diffusion directions is thus sampled by $N_{shots} = N_i N_g$ shots. For full interleaved sampling, the number of samples is $N_b = N_c N_{shots} (N_y / R_P) N_x$ with a shot reduction $R_P = N_i$ in phase encoding direction.

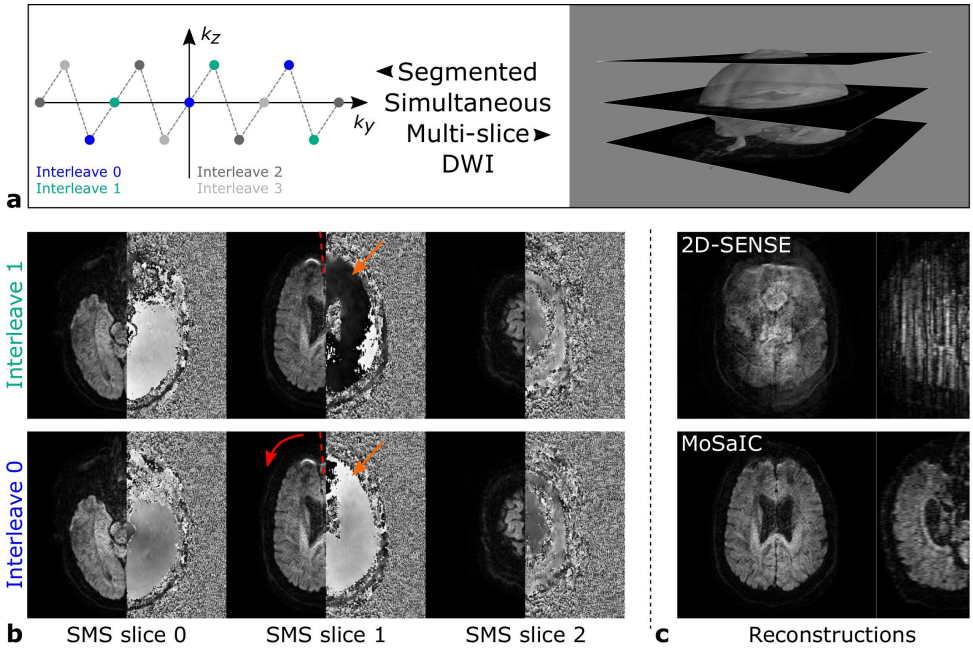


Figure 4.8: Segmented SMS DWI sampling and reconstruction problem for a 4-fold interleaved 3-SMS DWI acquisition. (a) This example visualizes the segmented SMS sampling. (b) The variations of two interleaves from the same SMS slice group are shown (please note that here magnitude and phase are illustrated for each dataset) along with two SENSE-based reconstructions given in (c). The two interleaves contain diffusion-related shot-to-shot phase variations (orange arrows) causing ghosting and signal dropouts in a CAIPI-adapted 2D-SENSE reconstruction. Secondly, macroscopic motion, here in-plane, (red arrow) is a common problem, which blurs anatomical structures. Moreover, through-plane motion mixes the anatomies of different slice groups. MoSaIC is a navigated technique that produces full volume reconstructions per diffusion direction with 3D rigid motion and shot phase corrections. This can be seen in (c) for the transversal example slice as well at the coronal reformat next to it.

The individual shot operators of the forward model are visualized in Fig. 4.9. The individual (and parallelizable) shot operators are stacked in block-diagonal structure into the following multi-shot operators. Based on the formulated assumptions, the forward model relates the multi-shot and multi-coil data $\grave{\mathbf{d}} \in \mathbb{C}^{N_b}$ to the image volume $\boldsymbol{\rho} \in \mathbb{C}^{N_\rho}$ by several linear operators and a complex Gaussian white noise vector $\boldsymbol{\eta} \in \mathbb{C}^{N_b}$:

$$\grave{\mathbf{d}} = \grave{M}\grave{F}\grave{\Theta}\grave{C}\grave{\Phi}\grave{M}^{sms}\Omega\,\boldsymbol{\rho} + \boldsymbol{\eta}. \tag{4.5}$$

First, the macroscopic motion operator Ω ($N_{shots}N_zN_yN_x \times N_zN_yN_x$) resamples the image volume according to the macroscopic shot motion parameters. Second, the slice-sampling operator $\grave{M}^{sms}$ ($N_{shots}N_{MB}N_yN_x \times N_{shots}N_zN_yN_x$) selects the SMS slices excited for each shot. Third, the physiological motion operator $\grave{\Phi}$ ($N_{shots}N_{MB}N_yN_x \times N_{shots}N_{MB}N_yN_x$) applies the shot- and slice-specific phase variations, followed by the coil weighting of the sensitivity operator $\grave{C}$ ($N_cN_{shots}N_{MB}N_yN_x \times N_{shots}N_{MB}N_yN_x$). Fifth, the CAIPI operator $\grave{\Theta}$ ($N_cN_{shots}N_yN_x \times N_cN_{shots}N_{MB}N_yN_x$) combines the SMS slices including CAIPI shifts. Finally, the coil images are Fourier transformed by $\grave{F}$ to sample the shot-specific trajectories using $\grave{M}$ ($N_cN_{shots}(N_y/R_P)N_x \times N_cN_{shots}N_yN_x$). Here, Ω directly includes the shot copy operator P by $\Omega = \grave{\Omega}P$ and Υ is neglected.

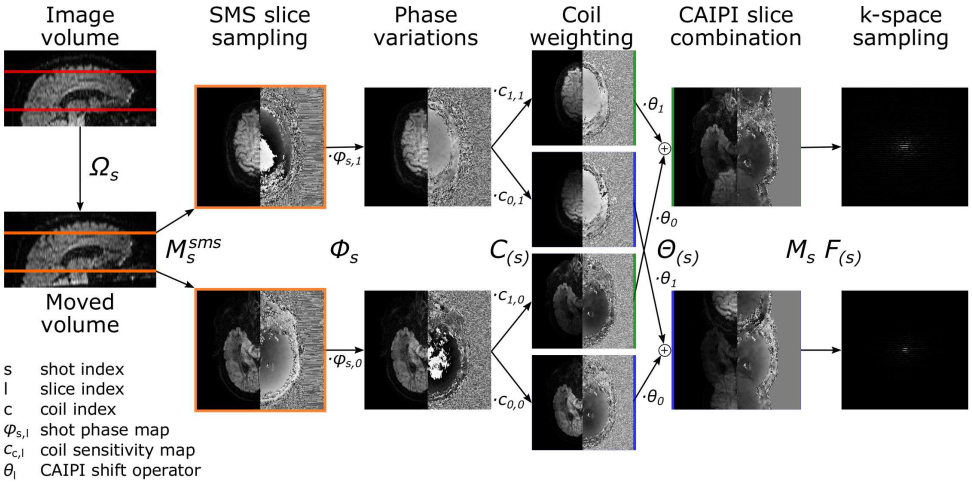


Figure 4.9: Illustration of the shot-specific forward model employed for motion-corrected segmented SMS DWI (using a SMS factor of 2). The image volume is resampled by the macroscopic motion operator Ω_s according to the motion transformation of shot s. The slice-sampling operator M_s^{sms} selects the shot-specific slices. The signal is weighted by the diffusion shot phases and coil sensitivities in Φ_s and $C_{(s)}$. $\Theta_{(s)}$ applies the blipped-CAIPI encoding and slice combination followed by the Fourier operator $F_{(s)}$ and shot-specific sampling M_s. The subscript (s) with shot index s in parentheses indicates shot operators that are identical for all shots. The multi-shot operators (Eq. 4.5) are obtained by embedding the shot operators in block-diagonal structure.

SMS-adapted Image Reconstruction Problem

The optimization problem for motion-corrected segmented SMS DWI is stated as a regularized data discrepancy minimization:

$$\underset{\boldsymbol{\rho},\grave{\Phi},\grave{\Omega}}{\text{minimize}} \quad \|\grave{M}\grave{F}\grave{\Theta}\grave{C}\grave{\Phi}\grave{M}^{sms}\grave{\Omega}\,\boldsymbol{\rho} - \grave{\mathbf{d}}\|^2_{\Psi^{-1}} + \lambda R(\boldsymbol{\rho}), \tag{4.6}$$

where the weighted norm $\|\cdot\|^2_{\Psi^{-1}} = (\cdot)^H \Psi^{-1} (\cdot)$ integrates the noise covariance matrix Ψ for SNR-optimal reconstruction as used for SENSE [12]. R is a regularization function and λ a weighting factor. The proposed MoSaIC algorithm employs an $\ell 1$-norm regularization $R(\boldsymbol{\rho}) = \|V\boldsymbol{\rho}\|_1$ in a wavelet domain [45] with transform operator V ($N_p \times N_p$). The image volume $\boldsymbol{\rho}$, the shot phase operator $\grave{\Phi}$ and the macroscopic motion operator $\grave{\Omega}$ are considered unknown. The problem is difficult to optimize due to the non-convexity associated with both $\grave{\Omega}$ and $\grave{\Phi}$.

Navigation approaches are used to linearize and simplify the non-convex optimization by estimating reconstruction parameters from an additional signal [31, 44, 64]. The proposed method leverages the low-resolution navigator echo to estimate the shot phase operator $\grave{\Phi}$ and the macroscopic motion operator $\grave{\Omega}$. For known $\grave{\Phi}$ and $\grave{\Omega}$, Eq. with $\ell 1$-norm regularization represents a convex optimization problem for $\boldsymbol{\rho}$, which can be readily solved using fast gradient projections (FGP) [96].

METHODS

DATA ACQUISITION AND PREPROCESSING

DTI data were acquired using a navigated Stejskal-Tanner spin-echo sequence with blipped-CAIPI [63] simultaneous multi-slice excitation similar to the work by Dai et al. [64]. Unlike this publication, the prewinder phase offset φ_{offset} was implemented as originally proposed [63], as for this SENSE implementation the samples are not required to lie on an integer-valued k-space grid. Thus, the k_z-sampling is centered around zero by:

$$\varphi_{\text{offset}} = -\frac{(D-1)}{2}\frac{2\pi}{D}, \quad D \in \mathbb{N}, \tag{4.7}$$

for a FOV_y/D CAIPI shift between two adjacent SMS slices. The first image echo was sampled at high-resolution in an interleaved fashion with a CAIPI shift of FOV_y/N_{MB} ($D_1 = N_{MB}$). The second echo used a moderately accelerated low-resolution sampling with a fixed CAIPI shift of $\text{FOV}_y/2$ ($D_2 = 2$). The navigator CAIPI shift was fixed for simplicity reasons. Given the z-encoding capabilities of the used 32-channel coil, the CAIPI encoding showed sufficient slice disentangling capabilities. The SMS slice groups were sampled in an interleaved ordering to reduce slice crosstalk [54]. Fat suppression was performed by spectral pre-saturation with inversion recovery (SPIR) [93]. The SMS slices were excited with slice-specific RF phases to reduce the peak B_1 [54]. Further sampling parameters are listed in Table 4.2.

Table 4.2: Sampling parameters of the diffusion tensor imaging scans. *The EPI factor is given for the k-space sampling without partial Fourier reduction.

Parameters	**DTI scan**
TR	3000 ms
TE image \navigator echo	70 \145 ms
FOV (R × P × S)	$232 \times 228 \times 120\,\text{mm}^3$
Resolution image \navigator	$1.0 \times 1.0 \times 4.0$ \$5.0 \times 5.0 \times 4.0\,\text{mm}^3$
#slices N_z	30
Multi-band factor N_{MB}	3
CAIPI shift image \navigator	(FOV / 3) \(FOV / 2)
#interleaves N_i	4
In-plane reduction R_P image \navigator	4 \1.62
Partial Fourier factor image \navigator	0.632 \1.000
Echo spacing image \navigator	1.3361 \0.5829 ms
EPI factor* image \navigator	57 \27
b-value	$1{,}000\,\text{s/mm}^2$
#diffusion directions N_d	15
#T_2-weighted acquisitions	2
Scan time [m:ss]	3:30

The data were acquired on a 3T Philips Ingenia Scanner (Best, The Netherlands) using a 32-channel head coil. The scan session was carried out on five healthy volunteers. Informed consent was attained according to the rules of the institution. The DTI scan with 15 diffusion directions was performed twice under different motion conditions:

1. Static: no voluntary motion
2. Motion: separated into four parts:
   a. T_2-weighted images: no voluntary motion (for $b = 0$ s/mm^2 data)
   b. DW directions #0-4: no voluntary motion
   c. DW directions #5-9: continuous in-plane motion ('head shake ')
   d. DW directions #10-14: continuous through-plane motion ('nodding ')

The subjects were asked to move at moderate rates of change to avoid the dominance of macroscopic intra-shot motion effects on the sensitive DWI sequence and provide sufficient non-compromised data. Coil sensitivity maps were acquired once in advance using a gradient-echo prescan [12]. The EPI data preprocessing involved gridding the ramp samples in the readout direction and applying odd/even echo Nyquist ghost correction, based on EPI reference data jointly acquired for all SMS slices [94]. The SMS phase from isocenter offsets in the slice direction was corrected for both the image and the navigator echo in advance [63, 66].

Proposed algorithm: MoSaIC

The proposed navigated algorithm, termed motion-aware SMS-accelerated and interleaved image creation (MoSaIC) for DWI, is described in the following four sections: navigator reconstruction, navigator analysis, linear operator construction and full-volume image reconstruction. The first three steps are performed for the full DTI dataset, while the latter can be parallelized over the diffusion directions. Figure 4.10 provides an overview of the reconstruction pipeline.

Navigator reconstruction

The navigator data are upsampled to the high-resolution voxel size of the image echo using zero padding in k-space and a triangular apodization window to reduce Gibbs ringing [3]. The navigators are reconstructed by 2D-SENSE [67] with Tikhonov regularization (regularization parameter $\lambda_{nav} = 0.05$) recovering the unfolded slice groups of each shot, which are shown in the green box on the bottom left of Fig. 4.10. Finally, residual slice-specific phases from RF excitation and blipped-CAIPI slice encoding are corrected.

Navigator analysis

The navigator analysis contains a shot *phase extraction*, a *shot rejection* and a *macroscopic motion* estimation. The *phase extraction* module obtains the phase maps $\boldsymbol{\phi}_{s,l}$ of shot s and slice l for all diffusion-weighted shots from the low-resolution navigators.

Compromised shots are rejected using robust statistics on the $\ell 2$-norm of the diffusion-weighted navigator shot data. The shot rejection module uses the median absolute deviation (MAD) over all diffusion-weighted shot datasets and excludes shots whose energy, measured by the $\ell 2$-norm, drops below the threshold of $5 \cdot$MAD from the

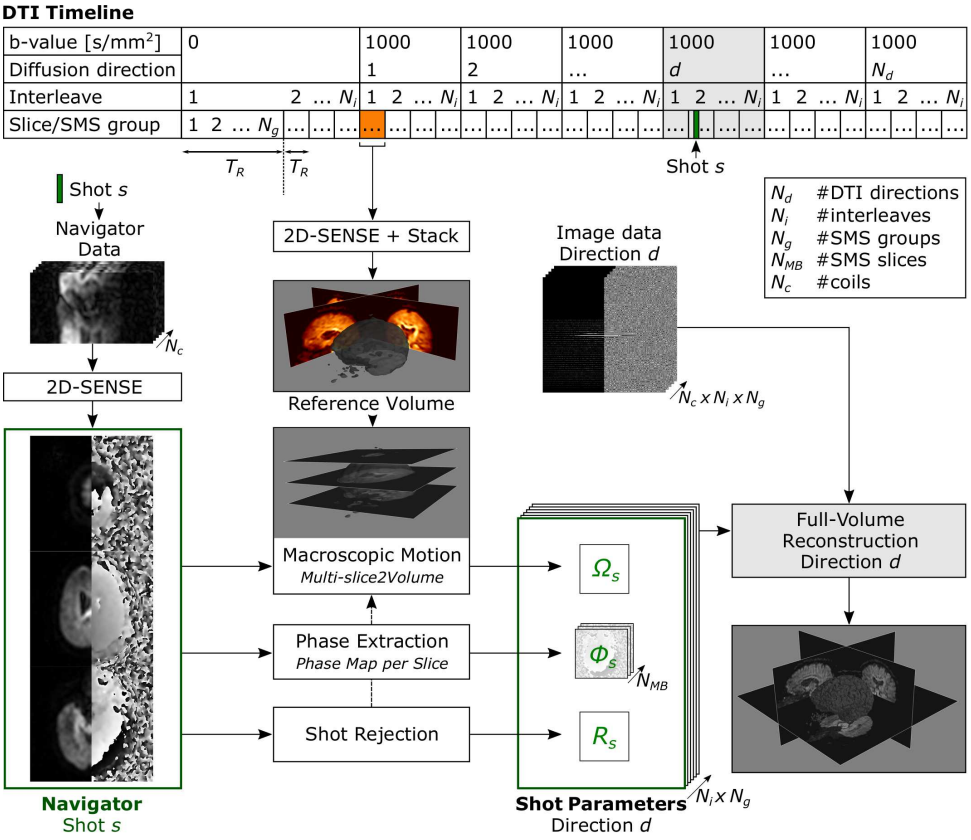


Figure 4.10: Overview of the proposed MoSaIC algorithm for a DTI acquisition. The DTI acquisition timeline loops over the diffusion-/T_2-weightings, the interleaves and the SMS groups (listed from outer to inner loops). All shot navigators are reconstructed by 2D-SENSE [67] (left green box) and the shot navigators from the reference TR (orange) are stacked to a full reference volume for registration. Next, the navigators are used to calculate shot rejection criteria, phase maps per shot and slice, and macroscopic motion by multi-slice-to-volume registration. The full-volume reconstruction uses the shot-specific parameters (for Ω_s, Φ_s, R_s) to reconstruct motion-corrected high-resolution image volumes separately for each diffusion direction.

median. The threshold was empirically set from analyzing the motion-corrupted navigators over multiple subjects and diffusion directions. The MAD criterion filters shots, which suffer from a significant SNR loss, while remaining robust against such outliers. Rejected shots are excluded from the registration strategy and from the full-volume image reconstruction.

The *macroscopic motion* module estimates 3D rigid motion parameters using a shot-wise multi-slice-to-volume registration together with a motion detection strategy. The navigators are downsampled to roughly isotropic resolution of $(4\text{mm})^3$. The first full diffusion-weighted navigator volume (orange square in Fig. 4.10) is used as the reference volume and is set as the moving image for the registration. This choice avoids resampling issues of the discrete simultaneous multi-slice data [117].

The multi-slice-to-volume registration was implemented in SimpleITK [118] with the following subtasks:

1. Volume-to-volume (Vol2Vol) registration of full navigator volumes per TR
   a. Method of moments pre-alignment
   b. Gradient descent-based registration
2. Motion detection per diffusion direction (from Vol2Vol parameters)
3. SMS-to-volume (SMS2Vol) registration per shot (in acquisition order) if motion is detected
   a. Linear interpolation of Vol2Vol parameters to shot time points
   b. Warm start: choose either Vol2Vol or previous SMS2Vol shot parameters by similarity evaluation through the registration metric
   c. Gradient descent-based registration
2. Median filter

The method of moments estimates rigid translations and rotations from the first and second statistical moments of the volumes. The registration uses the mutual information metric with 25 histogram bins, a 3D rigid versor transform and B-Spline interpolation. The maximum number of iterations for gradient optimizations was 1000. For the Vol2Vol registration, the gradient descent uses a regular step size strategy with a minimum step of 10^{-6}.

The motion detection checks for each diffusion direction whether any rigid Vol2Vol parameter deviates more than 0.2 mm or 0.2 deg from the median among the N_i subvolumes (TRs). This threshold was empirically set from accuracy analyses in registration simulations. If no motion is detected for a diffusion direction, the macroscopic motion operator $\grave{\Omega}$ is replaced by the identity operator I. Otherwise, the Vol2Vol preregistration parameters of the full TR volume are linearly interpolated to the shot time points. Then, a warm start strategy is employed that compares the registration metric for the interpolated Vol2Vol parameters and the SMS2Vol parameters of the previous shot. The subsequent gradient-based optimization is started from the preferred initial parameters and uses a line search strategy with at most 20 (line search) iterations. For SMS2Vol registration, the metric is evaluated only at SMS slice positions using a metric mask. Finally, a median filter (kernel size 5) is applied to each rigid motion parameter over time to ensure smoothness and to filter outliers.

Linear operator construction

The macroscopic motion operator $\grave{\Omega}$ for the high-resolution reconstruction is implemented as described by Cordero-Grande et al. [86]. Translations use the Fourier shift theorem, whereas rotations are factored into three shears [109], which are implemented via three 1D-FFTs. Moreover, the image $\boldsymbol{\rho}$ is reconstructed on an extended FOV to handle the FFT-related interpolation boundary conditions.

The operator $\grave{\Phi}$ multiplies the sampled slices by the shot- and slice-specific navigator phase maps $\boldsymbol{\phi}_{s,l}$. Moreover, encoding phases from the blipped-CAIPI image-space representation, off-isocenter encoding and slice-specific RF excitation phases φ_l^{RF} are

integrated into $\grave{\Phi}$, ensuring a continuous volume phase for the complex-valued interpolation in $\grave{\Omega}$. The coil sensitivities in the operator C are masked using a reference image and compressed [95] using a PCA with a threshold of 97% (resulting in 13 from initially 32 channels). The CAIPI operator $\grave{\Theta}$ applies the integer-valued shifts in image space and adds up the SMS group. The EPI data are then masked in k_y-x-space using a 1D-FFT $\grave{F}$, where it is masked according to the shot-specific EPI sampling trajectories by $\grave{M}$.

Full-volume image reconstruction

The full-volume reconstruction is based on FGP [96]. The algorithm is initialized by the motion-unaware SMS-IRIS solution described below. Then, gradient updates and soft thresholding in the Daubechies 4 wavelet domain with transform V are iteratively performed. λ was empirically set to 30. The FGP is stopped if either 100 iterations are exceeded or the normalized $\ell 2$-norm difference of two subsequent iterations δ_s (Eq. 2.37) drops below the tolerance $\tau_s = 10^{-5}$.

As the estimation of the Lipschitz constant for the FGP algorithm is demanding, a backtracking strategy was implemented [96]. The Lipschitz constant is initialized by the squared maximum absolute value of the coil sensitivity profiles. The Lipschitz constant is then increased by a factor of 1.5 if the objective function for the image estimate exceeds a quadratic approximation bound [96].

REFERENCE ALGORITHM: SMS-IRIS

The reference algorithm omits macroscopic motion during the image reconstruction. The static conditions, represented by an identity matrix I for the macroscopic motion operator $\grave{\Omega} = I$, decouple the SMS slice group signals from each other and reduce the SENSE problem to small groups of aliasing pixels.

IRIS [31], short for image reconstruction using image-space sampling function, is a SENSE formalism incorporating shot-to-shot phase variations from a navigator for interleaved single-slice EPI. In this work, a SMS extension, termed SMS-IRIS, is introduced, which requires the adaption of IRIS to the interleaved phase encoding in k_z-k_y-space [67]. SMS-IRIS can also be interpreted as a SENSE-based formulation of the navigated method by Dai et al. [64] or as a navigated version of SMS-based MUSE [79].

For SMS-IRIS, the shot- and slice-specific phase maps $\boldsymbol{\phi}_{\mathbf{s},l}$ are estimated as described for MoSaIC. The non-iterative algorithm incorporates a weighted Tikhonov regularization $R_{\boldsymbol{\rho}} = \|W\boldsymbol{\rho}\|_2^2$ with weighting matrix W ($N_p \times N_p$). W is constructed from the inverse absolute values of a motion-free T_2-weighted image filtered by a triangular window of about 4 mm isotropic k-space extent. $\lambda = 10^{-2}$ was set empirically. Moreover, the shot rejection is adopted to exclude severely compromised shots. This navigated algorithm yields time-efficient reconstructions without macroscopic motion correction providing the MoSaIC initialization and a reference method in this work.

EXPERIMENTAL DESIGN

The proposed algorithms were evaluated in simulation and in-vivo studies. The static and motion-compromised in-vivo data were reconstructed using SMS-IRIS and MoSaIC. Unsampled k-space portions from partial Fourier acquisitions are recovered by projections onto convex sets (POCS) [2]. The algorithms were implemented in Python 3.6.9. Computations were performed on a cluster node with 48 GB RAM.

A tensor model was fit to the reconstructed images using *Dipy* [68] after affine registration, rotation correction of the diffusion directions and PCA-based DTI denoising [119]. The registration was performed in two subsequent steps using a rigid and an affine preregistration of the fast elastic image registration (FEIR) framework [107] with a normalized gradient field metric [85]. The first rigid alignment uses the FFT-based resampling [86] described above. The second affine alignment is resampled with the SimpleITK [118] B-Spline interpolation. The first T_2-weighted image of the static dataset was set as the registration reference. The diffusion direction per image volume (after the multi-shot reconstruction) was corrected for the estimated rotations from the average MoSaIC and the DTI alignment parameters [44]. The PCA thresholding τ-factor was set the default value 2.3 and the SNR was estimated using *Dipy*'s PCA noise estimation. Fractional anisotropy (FA) maps [4] and isotropic diffusion images [7] were calculated.

For the simulations, the motion-free SMS-IRIS reconstructions were used as ground truth data. Ten rigid motion trajectories were simulated by Gaussian processes for four motion scenarios, namely no, rigid in-plane, rigid through-plane and fully 3D rigid motion. The variances were set to 0.01 rad^{2}and 0.5 mm^2. Diffusion phase maps were created by 3D second-order polynomials with random polynomial weights sampled according to Hu et al. [106]. The sampling was adapted to the in-vivo data with 4-shot 3-SMS acquisition as well as FOV/3 and FOV/2 CAIPI shifts for image and navigator echo, respectively. The simulation data were created by selecting random diffusion directions from the ground truth data, applying the forward model and adding uncorrelated complex Gaussian noise in k-space, whereby the SNR was matched to the in-vivo data.

The simulation data were recovered by SMS-IRIS and four MoSaIC variants to assess the registration components. All variants exclude rejected shots beforehand:

a. *MoSaIC Static*: No macroscopic motion correction ($\grave{\Omega} = I$),
b. *MoSaIC Vol2Vol*: Rigid parameters of Vol2Vol registration,
c. *MoSaIC SMS2Vol*: Rigid parameters of Vol2Vol and SMS2Vol registration,
d. *MoSaIC* (default): Rigid parameters of Vol2Vol and SMS2Vol registration if motion was detected and *MoSaIC Static* otherwise.

The simulation results of all methods were registered to the ground truth by FEIR [107] with a rigid model, as the final volume registration of SMS-IRIS was considered standard DWI processing for a fair nRMSE comparison.

The simulation results were compared by the target registration error (TRE), normalized root-mean-square error (nRMSE) and reconstruction time. The TRE [120] evaluates the mean Euclidian distance over the registered coordinates u_r of target points r:

$$\mathrm{TRE}(T'_\mathsf{s}, T^*_\mathsf{s}) = \frac{\sum_{(\mathsf{s},r)\in\mathrm{Target}} \| T'_\mathsf{s}(u_r) - T^*_\mathsf{s}(u_r)\|_2}{|\mathrm{Target}|}, \tag{4.8}$$

where T'_s is the estimated and T^*_s the true coordinate transformation of shot s. The target was defined as a mask derived from the ground truth by thresholding the absolute values at 5% of its maximum and selecting the SMS slices of each shot s. The cardinality |Target| is the number of target points over all shots and slices (per diffusion direction).

RESULTS

SIMULATED DWI RESULTS

The DWI motion simulations provide a quantitative assessment of the registration and reconstruction performance. Figure 4.11 gives an overview of the simulation results for *SMS-IRIS* and the four *MoSaIC* variants.

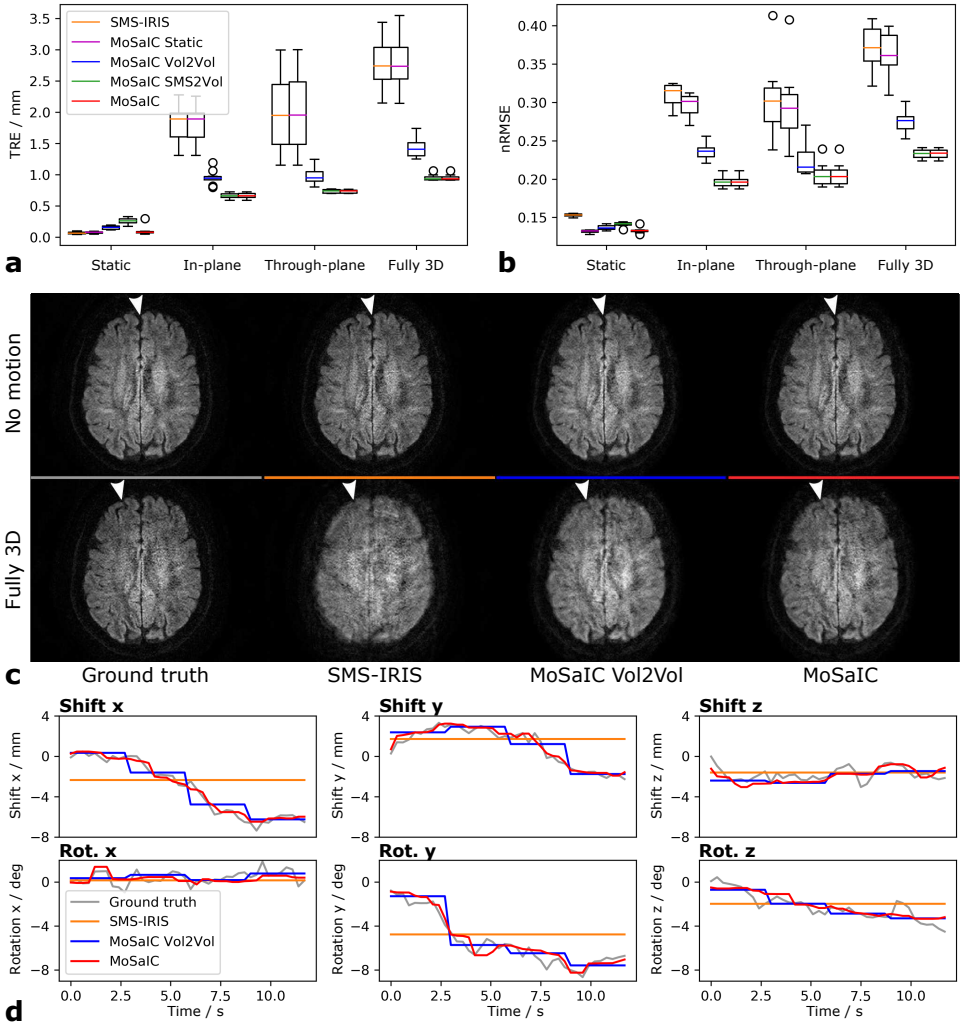


Figure 4.11: Full-volume reconstruction results of simulated 4-shot 3-SMS data with random rigid trajectories and diffusion phases. Ten cases were simulated for different motion states, namely static, in-plane, through-plane and fully 3D rigid motion. (a, b) show the target registration error (TRE) and nRMSE, respectively, for SMS-IRIS and four MoSaIC variants as standard boxplots with whiskers of 1.5 times the interquartile range. (c) Reconstruction examples without motion and with heavy fully 3D rigid motion are compared. (d) Rigid motion estimates including the final full-volume registration to the reference (leading also to non-zero motion parameters for SMS-IRIS and MoSaIC static). Without motion, all methods provide results with similar visual appearance (arrows). If motion is present, MoSaIC with its SMS2Vol registration improves anatomical delineation (arrows) by taking into account sub-volume (sub-TR) shot-to-shot motion.

For all algorithms, TRE and nRMSE increase in Figs. 4.11a and 4.11b with the degrees of freedom of the simulated motion profile. *MoSaIC Static* with wavelet denoising shows similar TRE and reduced nRMSE compared to *SMS-IRIS*. If motion is present, SMS2Vol outperforms Vol2Vol registration in terms of TRE and nRMSE, which itself improves on the static methods. *MoSaIC* provides similar results to *MoSaIC SMS2Vol*. For the static case, *SMS-IRIS*, *MoSaIC Static* and *MoSaIC* (with its motion detection) have a low TRE, which increases slightly from *MoSaIC Vol2Vol* to *SMS2Vol*. Figure 4.11c shows two transversal slice examples. For the first motion-free case, the results appear similar to the ground truth. The interhemispheric fissure is well resolved for all methods. For the second 3D motion case, *SMS-IRIS* shows heavy blurring artifacts from inter-shot motion. *MoSaIC* is able to mitigate the motion artifacts and shows less blurring than *MoSaIC Vol2Vol*. The underlying motion estimates for the fully 3D rigid motion case are provided in Fig. 4.11d.

The simulations show that the navigator with $5 \times 5 \times 4\,\mathrm{mm}^3$ resolution enables submillimeter TREs and improved image quality of the high-resolution full-volume reconstructions. The achieved registration accuracy does not visibly differ between in- and through-plane motion (Fig. 4.11a), although the nRMSE shows increased interquartile ranges for through-plane motion implying higher variation (Fig. 4.11b). This could be related to the increased susceptibility to interpolation errors in the coarse slice direction. If motion is present, the high temporal resolution of the SMS2Vol outperforms the Vol2Vol registration, which, in turn, improves on *SMS-IRIS*. Thus, the shot-to-shot motion estimation per SMS slice group captures continuous motion trajectories better, while having the same reconstruction complexity as *MoSaIC Vol2Vol*. The rigid motion estimates in Fig. 4.11d support this observation overall. Nevertheless, the motion parameters can contain temporary discrepancies above the sampling time resolution at some points, e.g. for the y-rotation at 5.0 s. If no motion is present, the SMS2Vol registration (used for *MoSaIC SMS2Vol* and *MoSaIC*) is more instable and introduces higher errors, whereby the visual appearance is not visibly degraded (Fig. 4.11c). *MoSaIC* uses the motion detection that switches off the SMS2Vol registration for small Vol2Vol estimates and thereby mitigates this downside by incorporating awareness of the achievable registration accuracy. The nRMSE improvements for through-plane motion demonstrate the benefits of the 3D motion over the 2D in-plane motion correction.

In-vivo DWI Results

An overview of the in-vivo full-volume DWI reconstructions with different motion types is given in Fig. 4.12. Reconstruction examples of *SMS-IRIS* and *MoSaIC* in Fig. 4.12a are related to the associated rigid motion estimates of *MoSaIC* in Fig. 4.12b by a color code. The images appear similar for the static case, whereas *MoSaIC* mitigates motion artifacts for the remaining cases (arrows). The motion detection was triggered for all datasets, whereby the SMS2Vol registration parameters remain almost constant for the static case (blue). Strong in-plane (z) rotations are detected for the orange and green datasets, leading to reduced blurring for *MoSaIC*. Light (red) and heavy (purple) nodding motion with rotations about the right-left axis (x) smear the structures in the coarse through-plane direction for *SMS-IRIS*, which are reduced by *MoSaIC*. Despite the 3D rigid corrections, residual artifacts remain for cases with heavy motion and the image quality of the static

datasets might not be fully recoverable. For the static/motion datasets of the five subjects, the data rejection excluded 1/23, 1/9, 2/24, 0/0, 1/3 of 680 total shots.

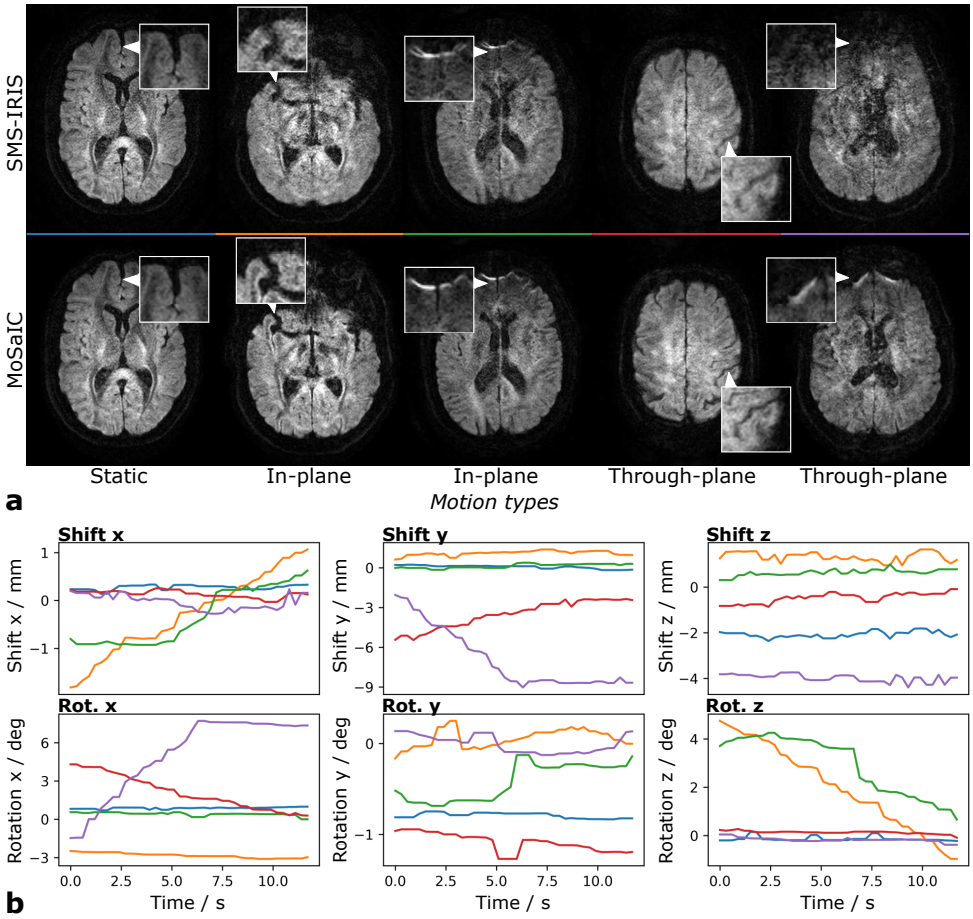


Figure 4.12: Segmented SMS image reconstructions without and with motion correction for in-vivo data. (a) Example slices of the full-volume reconstructions are shown for SMS-IRIS and MoSaIC, whereby different motion types have been present during the associated acquisitions. (b) The estimated rigid trajectories of MoSaIC are plotted and associated to the images by a color code. MoSaIC provides similar image quality for the static dataset (blue) and mitigates macroscopic motion artifacts in the presence of both in-plane (orange and green) and through-plane (red and purple) motion (compare close-ups). The proposed motion detection was triggered for all cases.

IN-VIVO DTI RESULTS

Figure 4.13 compares static and motion-disturbed DTI results for *SMS-IRIS* and *MoSaIC*. The estimated shift and rotation parameters of *MoSaIC* are displayed in Figs. 4.13a and b, respectively, and rejected shots are indicated. The time resolution is $\mathrm{TR}/N_g = 3/10\,\mathrm{s}$ with the number of SMS groups N_g yielding >3 Hz motion sampling frequency. For the static DTI dataset, the affine alignment resulted in 1.0042 ± 0.0047 and $-3 \cdot 10^{-5} \pm 0.0035$ for the non-rigid zoom and shear parameters, respectively.

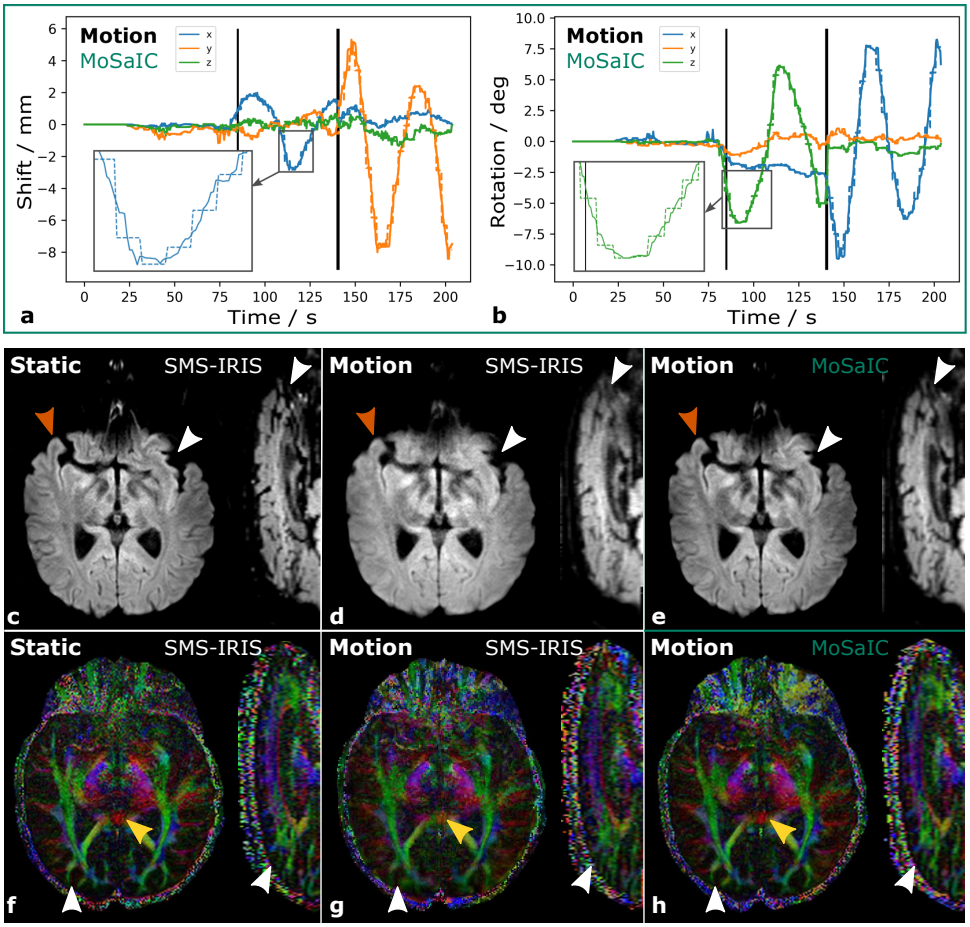
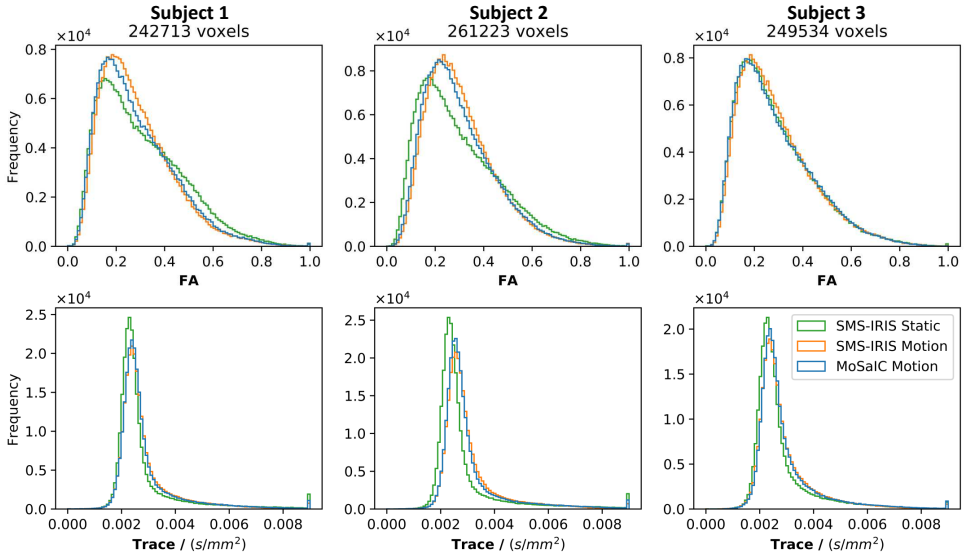


Figure 4.13: SMS-IRIS and MoSaIC reconstructions of in-vivo DTI datasets measured under static and motion-disturbed conditions. (a, b) The rigid shifts and rotations for the motion-disturbed dataset determined by MoSaIC are shown (mint green color code). The dashed lines with the same color in the background (see close-ups) provide the underlying Vol2Vol registration, whereas the black vertical lines indicate shots that were automatically rejected. (c) Isotropic DWI images of SMS-IRIS are shown under static and (d) motion conditions, along with the MoSaIC reconstruction in (e). (f-h) Colored FA maps are displayed in the same order. MoSaIC recovers blurred structures (white arrows), improves the grey to white matter differentiation in the isotropic DWIs (orange arrows) and improves the directional fidelity in the colored FA maps (yellow arrows).

The DTI analysis underlines the benefits of shot-to-shot motion correction for in-vivo data. The static reconstructions comprise high SNR and clear delineation of the brain gyri and fractional anisotropies. The estimated motion trajectories reflect the motion study design, comprising roughly no motion in the first, in-plane in the second and through-plane motion in the last third. The shot rejection is active at 85 s and 140 s. Compared to *SMS-IRIS*, *MoSaIC* reduces blurring of motion-corrupted brain structures and improves the gray-to-white matter differentiation (orange arrows) as well as the directional accuracy of the tensor results (yellow arrows).

A quantitative histogram evaluation of the FAs and tensor traces between static and motion-corrupted reconstructions is presented for three subjects in Fig. 4.14. The overlaps of the histograms from motion-corrupted datasets with the histograms from the static datasets are evaluated using the Kullback-Leibler divergence (KLD) [24]. *MoSaIC* reduces the dissimilarity measured by the KLD over all subjects.



a

Parameter	Method	Kullback–Leibler divergence		
		Subject 1	Subject 2	Subject 3
FA	SMS-IRIS	0.0280	0.0534	0.0030
	MoSaIC	0.0134	0.0304	0.0011
Trace	SMS-IRIS	0.0494	0.1836	0.0346
	MoSaIC	0.0389	0.1446	0.0344

b

Figure 4.14: Quantitative histogram comparisons of SMS-IRIS and MoSaIC results between static and motion-disturbed datasets for three subjects. (a): Fractional anisotropy (upper row) and trace (bottom row) histograms are shown for three subjects (columns). The static SMS-IRIS (green) results are compared to SMS-IRIS (orange) and MoSaIC (blue) under motion conditions. The MoSaIC histograms for motion datasets (blue) are overall closer to SMS-IRIS for static datasets (green) than the SMS-IRIS histograms for motion datasets (orange). However, MoSaIC still shows significant differences to the static cases. Please note that the comparison involves a resampling of the motion datasets from their specific reconstruction grids and a direct comparison is not trivial. (b): Histogram overlap evaluations of the motion cases to the associated static cases by the Kullback-Leibler divergence (KLD). The histogram distances of MoSaIC, measured by the KLD, decrease for all subjects compared to SMS-IRIS.

The image quality improvements of MoSaIC come at the cost of higher computational complexity. The durations of the main processing steps are given in Table 4.3 for the DTI dataset in Fig. 4.13. The highest computational cost of MoSaIC is placed by the FFT-based interpolation requiring multiple 1D-FFTs over the 3D data per shot and FGP iteration, which can be drastically reduced if no motion is detected.

Table 4.3: Reconstruction times of main processing steps. The numbers are given for a dataset containing 17 volumes (15 diffusion- + two T_2-weighted contrasts) with 256 × 240 × 30 size (R × P × S), multi-band 3, 10 SMS slice groups, 4 interleaves and 32 coils. Besides the total CPU times on an Intel Xeon Silver 4214 CPU @ 2.20GHz, the key performance indicators (KPIs) provide insights into the processing times per relevant cost driver indicating also potentials for parallelization. The MoSaIC durations are considered with 3D rigid correction and without ($\grave{\Omega} = I$), depending on the motion detection switch.

Processing step	**Reference quantity**	**CPU Time** [h:mm:ss]	**KPI** s / cost driver	
Navigator 2D-SENSE	All navigators	0:04:28	26.8	s/SMS group
Navigator Vol2Vol	All navigators	0:01:53	1.66	s/TR volume
Navigator SMS2Vol	All navigators	0:55:26	4.89	s/shot
MoSaIC - 3D rigid correction	DWI direction	6:11:11	317.02	s/FGP iteration
MoSaIC - no rigid correction	DWI direction	0:04:22	17.96	s/FGP iteration
SMS-IRIS per volume	DWI direction	0:02:03	12.32	s/SMS group

Residual Artifact Evaluation

The *MoSaIC* reconstructions are affected by several residual artifact types. Figure 4.15 compiles four types of artifacts (white arrows) that were encountered for MoSaIC. In all four cases, MoSaIC improves the blurring from head motion, but some artifacts remain, such as residual blurring of the interhemispheric fissure (Fig. 4.15a), overlays of different susceptibility-induced deformations (Fig. 4.15b), residual signal shadings (Fig. 4.15c) and low SNR as well as speckled noise structures for strong motion cases (Fig. 4.15d).

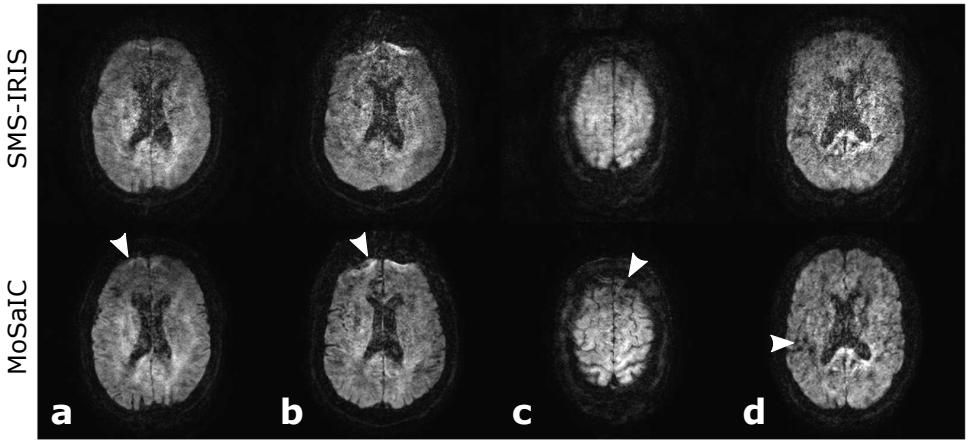


Figure 4.15: Examples of residual artifacts in MoSaIC reconstructions. MoSaIC mitigates macroscopic motion artifacts compared to SMS-IRIS, but some residual problems can remain. (a) Blurring might remain if the registration accuracy is insufficient. (b) Rotations cause shot-specific effects from off-resonances like susceptibility artifacts that are erroneously combined. (c) Ghosting and slice leakage can result from residual shot-to-shot phase inconsistencies. (d) Spoiled navigator diffusion phase estimates degrade the image quality.

SAMPLING DENSITY EVALUATION

Another problem arises from the fact that some anatomies might remain unsampled if they have left the FOV. Especially for through-plane motion, lower brain areas might leave or enter the FOV in feet-head direction. Nevertheless, the motion estimates give valuable information about the sampling locations, which can be translated to sampling densities in image space. The motion-aware sampling density estimation presented in Fig. 4.16 counts how many samples fall into the 3D cuboid Voronoi area of each voxel, given the rigid parameter estimates. The result can be smoothened by a Gaussian in-plane kernel ($\sigma = 2$). The sampling density provides beneficial information, which could be used for dedicated reacquisition strategies or to tag low SNR regions.

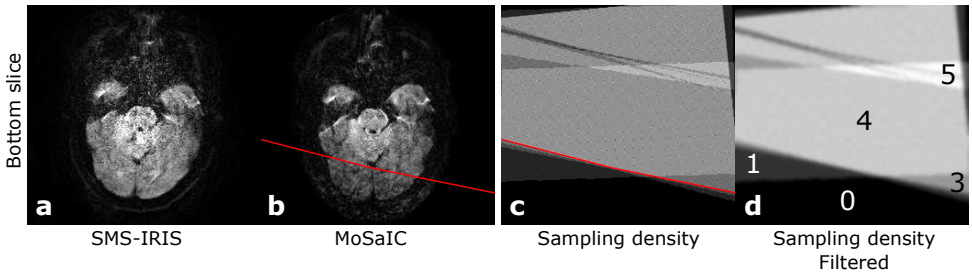


Figure 4.16: Image-space sampling density estimation by Voronoi partitions for regular SMS EPI under shot-specific motion. (a) SMS-IRIS shows strong motion-induced blurring from both in- and through-plane rotations for the presented dataset, which is resolved by MoSaIC (b). Nevertheless, the SNR falls off towards the right occipital lobe. (c) Sampling density in the image domain. (d) Sampling density after smoothing with a 2D Gaussian window. The numbers on the filtered sampling density plot indicate, how often a particular area in the image domain has been sampled in presence of motion. The analysis indicates that the affected area below the red line remained effectively unsampled, as it left the FOV due to nodding-type head motion.

DISCUSSION

MoSaIC provides 3D rigid motion-corrected full-volume reconstructions for navigated DWI with SMS and interleaved EPI sampling. The navigator shot images are employed to estimate the shot-specific phase variations, rigid motion states, and data rejection criteria. The multi-band excitation not only increases the SNR [54] per unit time by $\sqrt{N_{MB}}$, but also provides valuable data support in the slice direction enabling high temporal resolution through-plane motion evaluation. The performance of MoSaIC was evaluated in simulations and in-vivo for DTI scans with 4 shots and $N_{MB} = 3$ yielding shot motion estimates at roughly 3 Hz temporal resolution.

UNMODELED SHOT VARIATIONS

Several unmodeled shot variations [8] spoil the shot similarity for image registration and multi-shot image reconstructions. First, local off-resonance effects manifest as geometric distortions in phase encoding direction for EPI. As head rotations vary the effective phase encoding direction, the local distortions differ per shot (Fig. 4.15b). Geometric distortions can be reduced by readout-segmented acquisitions [29, 30] or corrected by model extensions accounting for off-resonance effects using B_0 map estimates [112].

This improves the shot navigator similarity for registration and offers higher local shot consistency for multi-shot reconstructions. In the current setting, the distortions also introduce mismatches of the EPI data and the gradient-echo sensitivity pre-scan, which could be avoided estimating the sensitivities from the non-DWI EPI data by ESPIRiT [59]. Second, field inhomogeneities can introduce slice-specific gradient timing offsets that are not captured by the current pre-scan and introduce Nyquist ghosting. The slice-adaptive odd-even corrections could be addressed by model-based estimations [121]. Third, eddy currents especially from the strong diffusion-sensitizing gradients induce generally affine transformations and exceed rigid modeling. Although the zoom and shear parameters were rather small in this study, they depend on the sequence and the system tuning. The field deviations can be reduced by twice-refocused sequences [27] or corrected by model-based reconstructions [23]. Alternatively, the motion model could be extended to affine transformations. Fourth, the effective diffusion direction is affected by subject rotations leading to shot-wise diffusion contrast variations [43]. The presented framework assumes for each diffusion direction that the induced contrast variations are negligible under sufficiently small rotations. Contrast corrections can be introduced by imposing q-space relations [8, 42, 43] between the shots.

Navigator Analysis

Navigation is prone to potential signal differences between the image and the navigator data. Intra-shot motion might spoil the acquisition windows differently, affect the navigator refocusing and lead to a geometric mismatch [22]. A central assumption of the current multi-shot model is that the object comprises one consistent phase map, whose shot variations are completely described by the navigator phase and the encoding model. However, residual phase inconsistencies and noise from intra-shot motion might impede proper pixel unfolding in the full-volume image reconstruction, which can lead to ghosting (Fig. 4.15c) and speckled noise (Fig. 4.15d) artifacts. Self-navigation is currently intractable for the studied motion-disturbed data due to the large g-factor penalty. The lack of viable shot reference data from the image echo impedes the evaluation of intra-shot motion leaving this as an open issue for further studies. The resolution of the single-shot navigator introduces another discrepancy for phase and macroscopic motion estimation.

The SMS-to-volume registration is itself a challenging problem [117] that requires careful design of the parameter initialization and gradient optimization. Regarding the warm start strategy, the exploitation of time correlations in combination with the challenging optimization landscape can cause temporal discrepancies of the registration parameters as described for Fig. 4.11. The registration accuracy is further impeded by the low SNR of DWI, the coarse navigator resolution, intra-shot motion and the unmodeled shot variations. Thus, MoSaIC might suffer from residual blurring artifacts (Fig. 4.15a).

Full-volume Image Reconstruction

The encoding model mainly determines the quality of the shot navigators and the full-volume image reconstructions comprising the k-space sampling trajectory and the sensitivity encoding. The deployment of other potentially non-Cartesian trajectories and improved coil setups might be beneficial [13]. As part of the sampling trajectory, this also applies to optimizations of the navigator CAIPI shift, which was FOV/2 in this work.

MoSaIC solves the multi-shot problem with $\ell 1$-norm regularization by fast gradient projections, and the sparsity-enforcing regularization was shown to reduce the nRMSE in simulations. With an $\ell 2$-norm regularization, the multi-shot problem is tractable by the conjugate gradients method [24], whereby both methods require one application of the forward model and its adjoint per iteration.

The choice of interpolation is a crucial trade-off balancing image quality and computational cost. The macroscopic motion operator $\grave{\Omega}$ resamples the image volume for all shots, twice per FGP iteration. GPU-based implementations [86] can significantly accelerate this processing. Furthermore, the motion detection of MoSaIC analyses the volumetric preregistration parameters for significant variations and allows circumventing the expensive image resampling.

MoSaIC is currently restricted to anisotropic resolution with coarse slice thickness, which might be overcome by improved slice encodings. The reduction of the slice thickness quickly becomes SNR-critical for the proposed acquisition and requires enhanced slice encoding, such as simultaneous 3D multi-slab approaches [113] or radio-frequency slice encoding [122]. Thinner slices should ease the relative interpolation burden, but, at the same time, they are affected more severely by motion.

INFORMED SAMPLING AND RECONSTRUCTION

MoSaIC uses a rather simple shot rejection analyzing the navigators energy content similar to Ref. [26]. More elaborate criteria to detect degenerate signals involve correlation measures [32] or k-space signal moments [30]. The reacquisition of deteriorated signals [30] presents a valuable extension to exclude malicious data without sacrificing SNR. Through the 3D rigid motion estimation, MoSaIC is furthermore aware of the spatial sampling density making reacquisitions of undersampled areas in image-space feasible as discussed for Fig. 4.16.

Prospective methods [21] represent an important alternative for macroscopic motion corrections. Prospective control allows for affine online corrections of the gradient coordinate system mitigating the motion artifacts directly in the acquisition and thereby facilitates the computational burden. The online adaption of the diffusion gradient [88] and phase encoding direction can thus reduce shot contrast and distortion variations. Conversely, prospective methods depend on the functionality of the external sensor system and cannot account for local motion transformations and phase variations [21]. In synergistic use, prospective methods could improve the database and ease the interpolation by reducing the residual motion, while the retrospective corrections cover remaining artifacts to enable highly motion-robust DWI.

CONCLUSIONS

The proposed MoSaIC framework improves diffusion-weighted imaging quality in the presence of head motion. The use of navigation and SMS acquisitions enable 3D rigid registration for motion-corrected full-volume reconstructions, which was presented for 4-shot 3-SMS DTI. The presented algorithms make combined use of SMS and interleaving allowing flexible balancing of SNR and scan time. This model-based strategy offers potentials for smart motion-aware image sampling and reconstruction to improve the robustness of DWI in clinical practice.

5

General Discussion and Conclusions

Robustness against subject motion in MRI is a multi-faceted problem that involves approaches from patient coaching over fast acquisitions to adaptive reconstructions [22] and needs to trade off the sequence and processing requirements for realistic use cases. DWI is a particularly motion-sensitive sequence [9] putting high demands on both hardware and software to enable diagnostic imaging. Multi-shot DWI entails shot-to-shot phase inconsistencies and macroscopic misalignments, but also comes with multiple secondary motion-related effects including DWI contrast variations, EPI artifacts like geometric distortions and Nyquist ghosting, and signal deteriorations from intra-shot motion [8]. Furthermore, the complex interaction of motion with the sequence [22] requires a careful balancing of the SNR properties of image and potential navigator signals, acquisition and reconstruction times, and the model complexity.

Thesis Contributions

This dissertation contributes to the field of multi-shot DWI reconstructions by investigating the potentials of different navigation strategies and motion models with increasing complexity as briefly outlined in Chapter 1. The four algorithms were shown to significantly improve image quality in the presence of shot-to-shot phase variations and macroscopic subject motion addressing important practical problems. An overview of the algorithms developed in this thesis is presented in Table 5.1.

Table 5.1: Model-based image reconstructions for multi-shot DWI developed as part of this dissertation.

Algorithm	Section	Acquisition		Motion Correction		Deep
		Navigated	SMS	Phase Maps	Rigid	Learning
MAPE+U-Net [48]	3.1	×	×	✓	×	✓
SEDIMENT [49]	3.2	×	×	✓	2D	×
IRIS-RiCo [50]	4.1	✓	×	✓	2D	×
MoSaIC [51]	4.2	✓	✓	✓	3D	×

Shot-to-shot phase variations [17–19] pose a challenging high-dimensional problem requiring a full phase map per shot. MAPE+U-Net [48] uses a neural network that effectively reduces phase-related artifacts in common SENSE reconstructions providing enhanced magnitude image priors to stabilize the ill-conditioned shot phase estimation. The magnitude-regularized phase estimation is embedded into an iterative algorithm which alternately improves the shot phase maps and the joint image, where the joint image reconstruction is kept isolated without deep learning. The U-Net for self-navigated usage is trained on simulation data from navigated image reconstructions [31].

In contrast to the NEATR approach [77] which uses the neural network as a post-processing step to improve a preceding MUSSELS reconstruction [40, 82], MAPE integrates the U-Net as an iterative regularization module into the alternating reconstruction. Nevertheless, the presented work represents a proof-of-concept study on a small training set and requires further validation and a comparison to NEATR [77] on a larger database. A related method, called NAMER [72], similarly performs image enhancement on multi-shot images with a U-Net to provide improved image priors for the rigid pa-

rameter estimation. For a combination of MAPE and NAMER, the current U-Net could be retrained on DWI data with simulated shot phase and rigid motion corruption.

Multi-shot DWI reconstructions have been extended to rigid in-plane motion corrections by the self-navigated AMUSE algorithm [43], which derives the shot parameters once from SENSE-based [12] shot image reconstructions. However, the g-factor penalty of SENSE [12] impedes high segmentation factors. The proposed SEDIMENT [49] algorithm therefore includes the shot phase and rigid in-plane corrections into an alternating reconstruction. The repeated feedback over the joint multi-shot image stabilizes the SENSE-based shot problems [39, 101] and improves the shot phase and rigid motion estimates making higher segmentations feasible.

Besides the self-navigation, a navigated GRAPPA-based [14] image reconstruction with shot phase and rigid in-plane motion correction has been presented by Dong et al. [44], but a comparison between the navigated and self-navigated methods is missing so far. This work proposes a navigated SENSE-based algorithm termed IRIS-RiCo [50], which is compared to the self-navigated SEDIMENT [49] algorithm for 4- and 5-shot DTI evaluating the implications of the navigation signals onto the image reconstruction. Both navigation strategies enable effective motion artifact reductions. However, the navigated approach is more robust for challenging datasets with strong motion or high segmentation, while the iterative self-navigation is more prone to fail disentangling the shot motion. Conversely, the coarse navigator resolution can affect the image registration accuracy leading to minor residual blurring in some datasets.

Through-plane motion is an important limitation of previous retrospective multi-shot DWI methods, which is addressed only through data rejection. To solve this important practical problem, the MoSaIC algorithm [51] adapts the navigated multi-shot DWI acquisition with simultaneous multi-slice sampling [64]. The resulting SMS navigators are used for shot phase estimation and multi-slice-to-volume [117] registration enabling navigated full-volume reconstructions with 3D rigid shot motion correction. By this, MoSaIC is an MR image-based alternative to the existing 3D rigid motion corrections based on optical cameras [80, 81], which require the extra sensor equipment. Furthermore, the shot phase variations are intractable by camera-based navigation and are therefore obtained by SENSE-based self-navigation, which is limited by the g-factor. A joint comparison of the navigated and extra-navigated methods is therefore required. MoSaIC could further be implemented as a prospective correction method, following the work for fMRI by Hoinkiss et al. [115]. This, however, poses online requirements on the navigator registration, which is infeasible in the current implementation.

The four presented algorithms exhibit varying abilities to be integrated into current clinical routines. IRIS-RiCo [50] is based on the clinically used IRIS method [31] and the extension proposed in this work includes common rigid in-plane registrations. The reconstruction generally follows a similar work flow as IRIS with a single successive optimization of the shot parameters and the joint image facilitating the practical implementation. In contrast, the alternating optimizations of MAPE+U-Net [48] and SEDIMENT [49] offer image quality improvements at the cost of increased computational and stability requirements, representing higher hurdles for practical uses. Finally, the reconstruction times of the MoSaIC method with 3D rigid motion corrections are currently infeasible for practical use, requiring further research to enable clinical adoption.

Modeling Limitations

Model-based image reconstructions provide effective means for motion artifact reduction in multi-shot DWI and build an important cornerstone to achieve robust MRI. However, the challenging non-convex optimization problems constitute several limitations of the presented methods that are discussed with respect to the navigation strategies, the motion models and the computational demands.

Navigation strategies represent a practical and efficient way to deal with the non-convexity of the joint multi-shot image and parameter reconstruction. Tracking devices like cameras for extra-navigation have been shown to capture rigid motion parameters well [21], while MR navigators are capable of sensing phase variations [26] and macroscopic motion [44]. Although these navigation strategies provide robust indicators for the signal variations, they remain problem substitutes of the image data consistency functional. As a major source of discrepancy, the presented navigated methods assume that the phase variations and the macroscopic motion are accurately captured on the coarse navigator resolution. This might lead to insufficient registration accuracy (Figs. 4.5 and 4.15a) or residual phase inconsistencies (Fig. 4.15c), but the navigator sampling has to trade off image resolution and SNR for proper navigation. Intra-shot motion or different off-resonance induced distortions constitute further signal discrepancy sources. In contrast, self-navigated methods directly optimize the image data consistency avoiding the signal discrepancy, but they generally suffer from the noise propagation effects [12] of the individual undersampled shot problems. At the same time, the comparison of navigated and self-navigated method is hindered by the lack of reliable reference data, as discussed fir IRIS-RiCo on page 81. This issue also impedes the evaluation of the structural tensor differences presented in Fig. 3.6 for MAPE+U-Net.

Being developed for the brain, the application of the methods for body anatomies like in abdominal and prostate imaging changes the SNR relations and motion conditions. Body imaging is characterized through a fast T_2 signal decay [4], which might render navigator signals from the second spin-echo unusable due to the high echo time. This issue favors the use of self-navigation in body applications, especially for sequences with strong diffusion encoding and long sampling windows. Regularized reconstructions like the U-Net supported MAPE [48], presented in Section 3.1, can thus help reducing the SNR issue. In addition, the macroscopic motion profiles in the body do usually require affine or even elastic components, which entails a larger parameter space and requires adapted optimization and interpolation strategies. For this, SEDIMENT [49] could be extended to non-rigid motion models. Moreover, the role of through-plane motion along with the increased respiratory activity [4] potentially supports applications of self-navigated SMS-to-volume registrations [114, 115] similar to MoSaIC [51].

Following the discussion for MoSaIC [51] in Section 4.2, the macroscopic motion models investigated in this thesis neglect several secondary sampling-related variations, which degrade both the image and shot parameter estimations. First, EPI is prone to off-resonance induced geometric distortions in the low-bandwidth phase encoding direction, which obstructs the overall geometric fidelity of the images. The head motion moreover varies the phase encoding direction inducing shot-specific distortions, which spoil the shot consistency for both the joint image reconstructions and image registrations. The distortions can be reduced using readout segmented acquisitions [29, 30],

prospective motion correction [21] or by including an estimated B_0 off-resonance map into the image reconstruction [112]. Second, eddy currents introduce mainly affine image transformations, which are observable between different diffusion directions and could be further reduced using twice-refocused DWI sequences [27] or by adapted reconstruction models [23]. Third, Nyquist ghosting was corrected in this work using EPI prescans, but head motion and off-resonance effects might lead to odd/even variations during the scan, which are addressable by adaptive modeling [121].

The shot-specific diffusion contrast variations from subject rotations [43] were mostly neglected as a secondary motion artifact in the present work, considering that rotations in common head coils are limited and the contrast variations are averaged over the shots. Nevertheless, these variations present a confounding factor for high-fidelity tensor applications. Prospective corrections of the gradient direction [88] offer a convenient solution to this problem. Despite this, retrospective contrast corrections [43] are possible but require a tensor estimate, which suffers from noise amplification in case of high segmentations. The potential of navigator-based low-resolution tensor estimates for contrast correction, as described in Section 4.1, requires further validation.

The employed motion models focus on inter-shot variations, establishing motion awareness on the time scale of the sequence duration with approximately 100 ms. Signal deteriorations and spin history effects resulting from the complex, non-linear interactions of motion [22] with the sequence are a major problem for motion-corrupted DWI, which is normally handled by similarity-based data rejections. Recent methods like NAMER [72] address intra-shot motion by separating the EPI train into multiple intra-shot segments for the motion estimation of inconsistent shots, which are optimized in a deep learning-supported image reconstruction. Such methods pave the way towards intra-shot motion and potentially avoid losing data by rejection. Nevertheless, the benefits of such methods must be critically balanced with the computational demands considering adaptive resampling strategies as potentially beneficial alternatives.

Finally, another limitation is the computational load that is involved especially with macroscopic motion-corrections and volumetric image reconstructions. The application of the large forward models represents the main computational expense of the presented algorithms, which was shown to be reducible by dedicated GPU implementations as shown by Cordero-Grande et al. [86]. At the same time, improvements in the field of algorithmic solvers, such as the FGP algorithm [96], boost the expensive iterative optimizations involved for motion correction. In addition, effective heuristics like the motion detection of MoSaIC [51] have proven valuable avoiding the costly resampling operations if the dataset appears motion-free.

Fostering Acquisition and Reconstruction Synergies

The main limitations of the model-based image reconstructions can be related to the achievable signal-to-noise ratios, the non-convexity of the optimization problem and the high computational demands. Although the limitations are challenging for the reconstruction on its own, joint developments with acquisition technologies have the potential to innovate the future MR image production using improvements of the signal encoding, the optimization methods, and regularizations enabling the design of adaptive image acquisition and reconstruction strategies for robust MRI.

The SNR of the image and navigator reconstructions is mainly determined by the encoding-related conditioning of the forward model. The Fourier operator can be combined with non-Cartesian trajectories like variable density spirals. The sampling order from low to high frequencies offers motion-robustness and the oversampling of the k-space center provides data redundancies supporting self-navigated methods [57]. Moreover, the PSF of spirals spreads the aliasing more inconsistently over the FOV improving the g-factor properties of the image encoding [57]. The extension to non-Cartesian trajectories requires interpolation procedures like gridding [58], which can be included into the k-space sampling operator $\grave{M}$ as in CG-SENSE [13]. The major drawbacks of non-Cartesian acquisitions are the higher susceptibility to gradient- and off-resonance-related blurring and the increased computational load from the gridding operator [57].

Besides the sampling trajectory, the employed forward models further rely on the encoding capabilities of the receive coil setup expressed through the orthogonality of the sensitivity operator [11]. Hence, improved coil design and more channels help to increase the capabilities to disentangle aliased pixels. In the presence of motion, the common usage of coil sensitivities from a prescan must be carefully reconsidered, as the sensitivities have been effectively measured only in the initial position and must be extrapolated to non-object regions. Moreover, non-Cartesian sampling techniques like WAVE-CAIPI [123] are able spread the aliasing more evenly in the three spatial dimensions allowing to make more efficient use of the 3D coil sensitivity profiles.

To achieve full brain coverage, the methods in this work compile independent single-slice or simultaneous multi-slice acquisitions to subsequently fill the full volume stack. However, the employed encoding with slice-selective excitations is incapable to achieve isotropic resolution, because the reduction of the slice thickness quickly becomes SNR-critical for the presented methods. Some recent approaches combine SMS with a phase encoding in the slice direction [113, 124], so that multiple thick slabs are excited and the additional slice encoding allows to resolve thin slices within each slab. Alternatively, RF encoding strategies [122, 125] can also provide an improved slice encoding to achieve isotropic DWI. Nevertheless, reducing the slice thickness further intensifies the motion sensitivity, so that the proposed methods become even more relevant.

Navigation is a key feature for both prospective and retrospective corrections to achieve motion-robust multi-shot DWI reconstructions. Navigators, be it camera-based tracking of rigid motion [111] or image-based phase navigation [26], provide useful guidance for the non-convex motion-corrected multi-shot DWI problem. A continuous motion awareness enables prospective scan adaption directly providing a database with reduced errors [21] and an improved problem conditioning for a subsequent retrospective correction of the residual artifacts. Severely deteriorated shots can be selectively reacquired [30] to stabilize the image reconstruction without SNR losses as experienced for data rejection. In addition, navigated reconstructions like IRIS-RiCo [50] can be used as a starting point for self-navigated reconstructions like SEDIMENT [49], providing valuable initializations for the ill-conditioned non-convex search. This kick-start of the joint optimization helps balancing the acquisition and reconstruction times and improves the convergence properties. The navigators provide continuous awareness of the scan situation and thus enable intelligent acquisition and reconstruction strategies.

REFERENCES

[1] D. G. Nishimura, *Principles of magnetic resonance imaging*. Stanford University, 1996.

[2] Z.-P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. SPIE Optical Engineering Press, 2000.

[3] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: physical principles and sequence design*, 2nd ed. John Wiley & Sons, 2014.

[4] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: concepts and applications," *JMRI*, vol. 13, no. 4, pp. 534–546, 2001.

[5] R. Bammer, "Basic principles of diffusion-weighted imaging," *European journal of radiology*, vol. 45, no. 3, pp. 169–184, 2003.

[6] L. Minati and W. P. Wglarz, "Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain: A review," *Concepts in Magnetic Resonance Part A*, vol. 30A, no. 5, pp. 278–307, Sep. 2007.

[7] P. Mukherjee, J. Berman, S. Chung, C. Hess, and R. Henry, "Diffusion Tensor MR Imaging and Fiber Tractography: Theoretic Underpinnings," *American Journal of Neuroradiology*, vol. 29, no. 4, pp. 632–641, Apr. 2008.

[8] W. Wu and K. L. Miller, "Image formation in diffusion MRI: A review of recent technical developments: Review of Image Formation in dMRI," *JMRI*, vol. 46, no. 3, pp. 646–662, Sep. 2017.

[9] D. Le Bihan, C. Poupon, A. Amadon, and F. Lethimonnier, "Artifacts and pitfalls in diffusion MRI," *JMRI*, vol. 24, no. 3, pp. 478–488, Sep. 2006.

[10] P. Mansfield, "Multi-planar image formation using NMR spin echoes," *Journal of Physics C: Solid State Physics*, vol. 10, no. 3, pp. L55–L58, Feb. 1977.

[11] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR phased array," *Magnetic Resonance in Medicine*, vol. 16, no. 2, pp. 192–225, 1990.

[12] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.

[13] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magnetic Resonance in Medicine*, vol. 46, no. 4, pp. 638–651, 2001.

[14] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, Jun. 2002.

[15] G. C. McKinnon, "Ultrafast interleaved gradient-echo-planar imaging on a standard scanner," *Magnetic Resonance in Medicine*, vol. 30, no. 5, pp. 609–616, Nov. 1993.

[16] F. Hennel, "Multiple-shot echo-planar imaging," *Concepts in Magnetic Resonance*,

vol. 9, no. 1, pp. 43–58, 1997.

[17] A. W. Anderson and J. C. Gore, "Analysis and correction of motion artifacts in diffusion weighted imaging," *Magnetic Resonance in Medicine*, vol. 32, no. 3, pp. 379–387, 1994.

[18] K. L. Miller and J. M. Pauly, "Nonlinear phase correction for navigated diffusion imaging," *Magnetic Resonance in Medicine*, vol. 50, no. 2, pp. 343–353, Aug. 2003.

[19] R. L. O'Halloran, S. Holdsworth, M. Aksoy, and R. Bammer, "Model for the correction of motion-induced phase errors in multishot diffusion-weighted-MRI of the head: Are cardiac-motion-induced phase errors reproducible from beat-to-beat?" *Magnetic Resonance in Medicine*, vol. 68, no. 2, pp. 430–440, Aug. 2012.

[20] J. B. Andre, B. W. Bresnahan, M. Mossa-Basha, M. N. Hoff, C. P. Smith, Y. Anzai, and W. A. Cohen, "Toward Quantifying the Prevalence, Severity, and Cost Associated With Patient Motion During Clinical MR Examinations," *Journal of the American College of Radiology*, vol. 12, no. 7, pp. 689–695, Jul. 2015.

[21] J. Maclaren, M. Herbst, O. Speck, and M. Zaitsev, "Prospective motion correction in brain imaging: A review," *Magnetic Resonance in Medicine*, vol. 69, no. 3, pp. 621–636, Mar. 2013.

[22] M. Zaitsev, J. Maclaren, and M. Herbst, "Motion artifacts in MRI: A complex problem with many partial solutions: Motion Artifacts and Correction," *Journal of Magnetic Resonance Imaging*, vol. 42, no. 4, pp. 887–901, Oct. 2015.

[23] J. Fessler, "Model-Based Image Reconstruction for MRI," *IEEE Signal Processing Magazine*, vol. 27, no. 4, pp. 81–89, Jul. 2010.

[24] C. R. Vogel, *Computational methods for inverse problems*. Society for Industrial and Applied Mathematics, 2002.

[25] K. Butts, A. de Crespigny, J. M. Pauly, and M. Moseley, "Diffusion-weighted interleaved echo-planar imaging with a pair of orthogonal navigator echoes," *Magnetic Resonance in Medicine*, vol. 35, no. 5, pp. 763–770, May 1996.

[26] K. Butts, J. Pauly, A. De Crespigny, and M. Moseley, "Isotropic diffusion-weighted and spiral-navigated interleaved EPI for routine imaging of acute stroke," *Magnetic Resonance in Medicine*, vol. 38, no. 5, pp. 741–749, Nov. 1997.

[27] D. Atkinson, D. A. Porter, D. L. G. Hill, F. Calamante, and A. Connelly, "Sampling and reconstruction effects due to motion in diffusion-weighted interleaved echo planar imaging," *Magnetic Resonance in Medicine*, vol. 44, pp. 101–109, 2000.

[28] D. Atkinson, S. Counsell, J. V. Hajnal, P. G. Batchelor, D. L. Hill, and D. J. Larkman, "Nonlinear phase correction of navigated multi-coil diffusion images," *Magnetic Resonance in Medicine*, vol. 56, no. 5, pp. 1135–1139, 2006.

[29] S. J. Holdsworth, S. Skare, R. D. Newbould, R. Guzmann, N. H. Blevins, and R. Bammer, "Readout-segmented EPI for rapid high resolution diffusion imaging at 3T," *European Journal of Radiology*, vol. 65, no. 1, pp. 36–46, Jan. 2008.

[30] D. A. Porter and R. M. Heidemann, "High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition," *Magnetic Resonance in Medicine*, vol. 62, no. 2, pp. 468–475, Aug. 2009.

[31] H.-K. Jeong, J. C. Gore, and A. W. Anderson, "High-resolution human diffusion tensor imaging using 2-D navigated multishot SENSE EPI at 7 T," *Magnetic Resonance*

in Medicine, vol. 69, no. 3, pp. 793–802, Mar. 2013.

[32] J. G. Pipe, V. G. Farthing, and K. P. Forbes, "Multishot diffusion-weighted FSE using PROPELLER MRI," *Magnetic Resonance in Medicine*, vol. 47, no. 1, pp. 42–52, Jan. 2002.

[33] C. Liu, M. E. Moseley, and R. Bammer, "Simultaneous phase correction and SENSE reconstruction for navigated multi-shot DWI with non-cartesian k-space sampling," *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1412–1422, Dec. 2005.

[34] M. Uecker, A. Karaus, and J. Frahm, "Inverse reconstruction method for segmented multishot diffusion-weighted MRI with multiple coils," *Magnetic Resonance in Medicine*, vol. 62, no. 5, pp. 1342–1348, Nov. 2009.

[35] N.-k. Chen, A. Guidon, H.-C. Chang, and A. W. Song, "A robust multi-shot scan strategy for high-resolution diffusion weighted MRI enabled by multiplexed sensitivity-encoding (MUSE)," *NeuroImage*, vol. 72, pp. 41–47, May 2013.

[36] T.-K. Truong and A. Guidon, "High-resolution multishot spiral diffusion tensor imaging with inherent correction of motion-induced phase errors: Inherent Motion Correction for Multishot Spiral DTI," *Magnetic Resonance in Medicine*, vol. 71, no. 2, pp. 790–796, Feb. 2014.

[37] Z. Zhang, F. Huang, X. Ma, S. Xie, and H. Guo, "Self-feeding MUSE: A robust method for high resolution diffusion imaging using interleaved EPI," *NeuroImage*, vol. 105, pp. 552–560, Jan. 2015.

[38] M.-L. Chu, H.-C. Chang, H.-W. Chung, T.-K. Truong, M. R. Bashir, and N.-k. Chen, "POCS-based reconstruction of multiplexed sensitivity encoded MRI (POCSMUSE): A general algorithm for reducing motion-related artifacts: POCSMUSE Reconstruction for Motion Artifact Removal," *Magnetic Resonance in Medicine*, vol. 74, no. 5, pp. 1336–1348, Nov. 2015.

[39] H. Guo, X. Ma, Z. Zhang, B. Zhang, C. Yuan, and F. Huang, "POCS-enhanced inherent correction of motion-induced phase errors (POCS-ICE) for high-resolution multishot diffusion MRI: High Resolution Multishot DWI without Navigator," *Magnetic Resonance in Medicine*, vol. 75, no. 1, pp. 169–180, Jan. 2016.

[40] M. Mani, M. Jacob, D. Kelley, and V. Magnotta, "Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS): Annihilating Filter K-Space Formulation for Multi-Shot DWI Recovery," *Magnetic Resonance in Medicine*, vol. 78, no. 2, pp. 494–507, Aug. 2017.

[41] Y. Hu, E. G. Levine, Q. Tian, C. J. Moran, X. Wang, V. Taviani, S. S. Vasanawala, J. A. McNab, B. A. Daniel, and B. L. Hargreaves, "Motion-robust reconstruction of multishot diffusion-weighted images without phase estimation through locally low-rank regularization," *Magnetic Resonance in Medicine*, vol. 81, no. 2, pp. 1181–1190, Oct. 2018.

[42] Y. Hu, X. Wang, Q. Tian, G. Yang, B. Daniel, J. McNab, and B. Hargreaves, "Multi-shot diffusion-weighted MRI reconstruction with magnitude-based spatial-angular locally low-rank regularization (SPA-LLR)," *Magnetic Resonance in Medicine*, vol. 83, no. 5, pp. 1596–1607, May 2020.

[43] S. Guhaniyogi, M.-L. Chu, H.-C. Chang, A. W. Song, and N.-k. Chen, "Motion immune diffusion imaging using augmented MUSE for high-resolution multi-shot EPI: Motion Immune Diffusion Imaging Using AMUSE," *Magnetic Resonance in*

Medicine, vol. 75, no. 2, pp. 639–652, Feb. 2016.

[44] Z. Dong, F. Wang, X. Ma, E. Dai, Z. Zhang, and H. Guo, "Motion-corrected k-space reconstruction for interleaved EPI diffusion imaging: Motion Correction for iEPI DWI," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 1992–2002, Apr. 2018.

[45] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[46] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.

[47] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, pp. 3055–3071, Jun. 2018.

[48] M. Steinhoff, A. Mertins, and P. Börnert, "Magnitude-regularized Phase Estimation (MAPE) with U-Net Support for Self-navigated Multi-shot Echo-planar DWI in the Brain," in *Proceedings of the 2020 ISMRM & SMRT Virtual Conference & Exhibition*, 2020, Program no. 4367.

[49] M. Steinhoff, K. Nehrke, A. Mertins, and P. Börnert, "Segmented diffusion imaging with iterative motion-corrected reconstruction (SEDIMENT) for brain echo-planar imaging," *NMR in Biomedicine*, vol. 33, no. 12, p. e4185, Dec. 2020.

[50] M. Steinhoff, A. Mertins, and P. Börnert, "SENSE-based Multi-shot DWI Reconstruction with Extra-navigated Rigid Motion and Contrast Correction for Brain EPI," in *Proceedings of the 2020 ISMRM & SMRT Virtual Conference & Exhibition*, 2020, Program no. 4339.

[51] M. Riedel (né Steinhoff), K. Setsompop, A. Mertins, and P. Börnert, "Segmented simultaneous multi-slice diffusion-weighted imaging with navigated 3D rigid motion correction," *Under review: submitted to Magnetic Resonance in Medicine on Feb 24, 2021*.

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

[53] N. Parikh, "Proximal Algorithms," *Foundations and Trendső in Optimization*, vol. 1, pp. 127–239, 2014.

[54] M. Barth, F. Breuer, P. J. Koopmans, D. G. Norris, and B. A. Poser, "Simultaneous multislice (SMS) imaging techniques: SMS Imaging," *Magnetic Resonance in Medicine*, vol. 75, no. 1, pp. 63–81, Jan. 2016.

[55] W. A. Edelstein, J. M. S. Hutchison, G. Johnson, and T. Redpath, "Spin warp NMR imaging and applications to human whole-body imaging," *Physics in Medicine and Biology*, vol. 25, no. 4, pp. 751–756, Jul. 1980.

[56] M. H. Buonocore and L. Gao, "Ghost artifact reduction for echo planar imaging using image phase correction," *Magnetic Resonance in Medicine*, vol. 38, no. 1, pp. 89–100, Jul. 1997.

[57] B. M. A. Delattre, R. M. Heidemann, L. A. Crowe, J.-P. Vallée, and J.-N. Hyacinthe, "Spiral demystified," *Magnetic Resonance Imaging*, p. 20, 2010.

[58] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski, "Selection of a Convo-

lution Function for Fourier Inversion Using Gridding," *IEEE Transactions on Medical Imaging*, vol. 10, no. 3, pp. 473–478, 1991.

[59] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT-an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magnetic Resonance in Medicine*, vol. 71, no. 3, pp. 990–1001, Mar. 2014.

[60] E. Wong, "Optimized phase schedules for minimizing peak RF power in simultaneous multi-slice RF excitation pulses. Abstract 2209." in *Proceedings of the 20th Annual Meeting of the ISMRM*, Melbourne, Australia, 2012.

[61] F. A. Breuer, M. Blaimer, R. M. Heidemann, M. F. Mueller, M. A. Griswold, and P. M. Jakob, "Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging," *Magnetic Resonance in Medicine*, vol. 53, no. 3, pp. 684–691, Mar. 2005.

[62] F. A. Breuer, M. Blaimer, M. F. Mueller, N. Seiberlich, R. M. Heidemann, M. A. Griswold, and P. M. Jakob, "Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA)," *Magnetic Resonance in Medicine*, vol. 55, no. 3, pp. 549–556, Mar. 2006.

[63] K. Setsompop, B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen, and L. L. Wald, "Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty," *Magnetic Resonance in Medicine*, vol. 67, no. 5, pp. 1210–1224, May 2012.

[64] E. Dai, X. Ma, Z. Zhang, C. Yuan, and H. Guo, "Simultaneous multislice accelerated interleaved EPI DWI using generalized blipped-CAIPI acquisition and 3D K-space reconstruction: SMS Accelerated iEPI DWI," *Magnetic Resonance in Medicine*, vol. 77, no. 4, pp. 1593–1605, Apr. 2017.

[65] M. Weiger, K. P. Pruessmann, and P. Boesiger, "2D sense for faster 3D MRI," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 14, no. 1, pp. 10–19, 2002.

[66] B. Zahneisen, B. A. Poser, T. Ernst, and V. A. Stenger, "Three-dimensional Fourier encoding of simultaneously excited slices: Generalized acquisition and reconstruction framework: 3D Fourier Encoding for SMS Acquisitions," *Magnetic Resonance in Medicine*, vol. 71, no. 6, pp. 2071–2081, Jun. 2014.

[67] B. Zahneisen, T. Ernst, and B. A. Poser, "SENSE and simultaneous multislice imaging: SENSE and Simultaneous Multislice Imaging," *Magnetic Resonance in Medicine*, vol. 74, no. 5, pp. 1356–1362, Nov. 2015.

[68] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. Van Der Walt, M. Descoteaux, and I. Nimmo-Smith, "Dipy, a library for the analysis of diffusion MRI data," *Frontiers in neuroinformatics*, vol. 8, p. 8, 2014.

[69] P. G. Batchelor, D. Atkinson, P. Irarrazaval, D. L. G. Hill, J. Hajnal, and D. Larkman, "Matrix description of general motion correction applied to multishot images," *Magnetic Resonance in Medicine*, vol. 54, no. 5, pp. 1273–1280, Nov. 2005.

[70] C. Liao, X. Cao, J. Cho, Z. Zhang, K. Setsompop, and B. Bilgic, "Highly efficient MRI through multi-shot echo planar imaging," in *Wavelets and Sparsity XVIII*, Y. M. Lu, M. Papadakis, and D. Van De Ville, Eds. San Diego, United States: SPIE, Sep. 2019, p. 43.

[71] Z. Hu, Y. Wang, Z. Dong, and H. Guo, "Water/fat separation for distortion-free EPI with point spread function encoding," *Magnetic Resonance in Medicine*, vol. 82, no. 1, pp. 251–262, Jul. 2019.

[72] M. W. Haskell, S. F. Cauley, B. Bilgic, J. Hossbach, D. N. Splitthoff, J. Pfeuffer, K. Setsompop, and L. L. Wald, "Network Accelerated Motion Estimation and Reduction (NAMER): Convolutional neural network guided retrospective motion correction using a separable motion model," *Magnetic Resonance in Medicine*, vol. 82, no. 4, pp. 1452–1461, 2019.

[73] F. Wang, Z. Dong, T. G. Reese, B. Bilgic, M. Katherine Manhard, J. Chen, J. R. Polimeni, L. L. Wald, and K. Setsompop, "Echo planar time-resolved imaging (EPTI)," *Magnetic Resonance in Medicine*, vol. 81, no. 6, pp. 3599–3615, Jun. 2019.

[74] S. J. Vannesjo, N. N. Graedel, L. Kasper, S. Gross, J. Busch, M. Haeberlin, C. Barmet, and K. P. Pruessmann, "Image reconstruction using a gradient impulse response model for trajectory prediction: GIRF-Based Image Reconstruction," *Magnetic Resonance in Medicine*, vol. 76, no. 1, pp. 45–58, Jul. 2016.

[75] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)," *Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1196–1202, Jun. 2007.

[76] V. B. Xie, M. Lyu, Y. Liu, Y. Feng, and E. X. Wu, "Robust EPI Nyquist ghost removal by incorporating phase error correction with sensitivity encoding (PEC-SENSE): Nyquist Ghost Removal by PEC-SENSE," *Magnetic Resonance in Medicine*, vol. 79, no. 2, pp. 943–951, Feb. 2018.

[77] B. Bilgic, I. Chatnuntawech, M. K. Manhard, Q. Tian, C. Liao, S. S. Iyer, S. F. Cauley, S. Y. Huang, J. R. Polimeni, L. L. Wald, and K. Setsompop, "Highly accelerated multishot echo planar imaging through synergistic machine learning and joint reconstruction," *Magnetic Resonance in Medicine*, vol. 82, no. 4, pp. 1343–1358, May 2019.

[78] J. P. Haldar, "Low-Rank Modeling of Local k-Space Neighborhoods (LORAKS) for Constrained MRI," *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 668–681, Mar. 2014.

[79] H.-C. Chang, S. Guhaniyogi, and N.-k. Chen, "Interleaved diffusion-weighted improved by adaptive partial-Fourier and multiband multiplexed sensitivity-encoding reconstruction: Reconstruction Framework for Artifact-Free DWI," *Magnetic Resonance in Medicine*, vol. 73, no. 5, pp. 1872–1884, May 2015.

[80] M. Herbst, B. Zahneisen, B. Knowles, M. Zaitsev, and T. Ernst, "Prospective motion correction of segmented diffusion weighted EPI: Prospective Motion Correction of Segmented Diffusion Weighted EPI," *Magnetic Resonance in Medicine*, vol. 74, no. 6, pp. 1675–1681, Dec. 2015.

[81] M. Herbst, B. Poser, A. Singh, W. Deng, B. Knowles, M. Zaitsev, V. Stenger, and T. Ernst, "Motion correction for diffusion weighted SMS imaging," *Magnetic Resonance Imaging*, vol. 38, pp. 33–38, May 2017.

[82] M. Mani, M. Jacob, G. McKinnon, B. Yang, B. Rutt, A. Kerr, and V. Magnotta, "SMS MUSSELS: A navigator-free reconstruction for simultaneous multi-slice-accelerated multi-shot diffusion weighted imaging," *Magnetic Resonance in Medicine*, vol. 83, no. 1, pp. 154–169, Jan. 2020.

[83] F. Ong, J. Y. Cheng, and M. Lustig, "General phase regularized reconstruction using phase cycling: General Phase Regularized Reconstruction Using Phase Cycling," *Magnetic Resonance in Medicine*, vol. 80, no. 1, pp. 112–125, Jul. 2018.

[84] R. Bammer, M. Aksoy, and C. Liu, "Augmented generalized SENSE reconstruction to correct for rigid body motion," *Magnetic Resonance in Medicine*, vol. 57, no. 1, pp. 90–102, Jan. 2007.

[85] E. Haber and J. Modersitzki, "Beyond Mutual Information: A Simple and Robust Alternative," in *Bildverarbeitung für die Medizin 2005*, H.-P. Meinzer, H. Handels, A. Horsch, and T. Tolxdorff, Eds. Berlin/Heidelberg: Springer-Verlag, 2005, pp. 350–354.

[86] L. Cordero-Grande, R. P. A. G. Teixeira, E. J. Hughes, J. Hutter, A. N. Price, and J. V. Hajnal, "Sensitivity Encoding for Aligned Multishot Magnetic Resonance Reconstruction," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 266–280, Sep. 2016.

[87] L. Cordero-Grande, E. J. Hughes, J. Hutter, A. N. Price, and J. V. Hajnal, "Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to neonatal brain imaging: Aligned Multi-Shot Multi-Slice MRI," *Magnetic Resonance in Medicine*, vol. 79, no. 3, pp. 1365–1376, Jun. 2017.

[88] M. Herbst, J. Maclaren, M. Weigel, J. Korvink, J. Hennig, and M. Zaitsev, "Prospective motion correction with continuous gradient updates in diffusion weighted imaging: Prospective Motion Correction with Continuous Gradient Updates," *Magnetic Resonance in Medicine*, vol. 67, no. 2, pp. 326–338, Feb. 2012.

[89] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[90] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778.

[91] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167 [cs]*, Feb. 2015, arXiv: 1502.03167.

[92] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980 [cs]*, Jan. 2017, arXiv: 1412.6980.

[93] C. H. Oh, S. K. Hilal, and Z. H. Cho, "Selective partial inversion recovery (SPIR) in steady state for selective saturation magnetic resonance imaging (MRI)," in *Book of Abstracts: Seventh Annual Meeting of the Society of Magnetic Resonance in Medicine*, vol. 2, San Francisco, 1988, p. 1042.

[94] H. Bruder, H. Fischer, H.-E. Reinfelder, and F. Schmitt, "Image reconstruction for echo planar imaging with nonequidistant k-space sampling," *Magnetic Resonance in Medicine*, vol. 23, no. 2, pp. 311–323, 1992.

[95] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke, "Array compression for MRI with large coil arrays," *Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1131–1139, Jun. 2007.

[96] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for

Linear Inverse Problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[97] M. Bydder, D. Larkman, and J. Hajnal, "Detection and elimination of motion artifacts by regeneration ofk-space," *Magnetic Resonance in Medicine*, vol. 47, no. 4, pp. 677–686, Apr. 2002.

[98] F. Odille, P.-A. Vuissoz, P.-Y. Marie, and J. Felblinger, "Generalized Reconstruction by Inversion of Coupled Systems (GRICS) applied to free-breathing MRI," *Magnetic Resonance in Medicine*, vol. 60, no. 1, pp. 146–157, Jul. 2008.

[99] T. Nielsen and P. Börnert, "Iterative motion compensated reconstruction for parallel imaging using an orbital navigator," *Magnetic Resonance in Medicine*, vol. 66, no. 5, pp. 1339–1345, Nov. 2011.

[100] P. Qu, K. Zhong, B. Zhang, J. Wang, and G. X. Shen, "Convergence behavior of iterative SENSE reconstruction with non-Cartesian trajectories," *Magnetic Resonance in Medicine*, vol. 54, no. 4, pp. 1040–1045, 2005.

[101] A. A. Samsonov, E. G. Kholmovski, D. L. Parker, and C. R. Johnson, "POCSENSE: POCS-based reconstruction for sensitivity encoded magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 52, no. 6, pp. 1397–1406, Dec. 2004.

[102] K. Kim, P. Habas, F. Rousseau, O. Glenn, A. Barkovich, and C. Studholme, "Intersection Based Motion Correction of Multislice MRI for 3-D *in Utero* Fetal Brain Image Formation," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 146–158, Jan. 2010.

[103] F. Huang, S. Vijayakumar, Y. Li, S. Hertel, and G. R. Duensing, "A software channel compression technique for faster reconstruction with many channels," *Magnetic Resonance in Medicine*, vol. 26, no. 1, pp. 133–141, Jan. 2008.

[104] T. Zhang, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "Coil compression for accelerated imaging with Cartesian sampling," *Magnetic Resonance in Medicine*, vol. 69, no. 2, pp. 571–582, Feb. 2013.

[105] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," *NeuroImage*, vol. 5, p. 425, 1997.

[106] Z. Hu, X. Ma, T.-K. Truong, A. W. Song, and H. Guo, "Phase-updated regularized SENSE for navigator-free multishot diffusion imaging: PR-SENSE for DWI," *Magnetic Resonance in Medicine*, vol. 78, no. 1, pp. 172–181, Jul. 2017.

[107] S. Kabus and C. Lorenz, "Fast elastic image registration," *Medical Image Analysis for the Clinic: A Grand Challenge*, vol. 89, pp. 81–89, 2010.

[108] K. Bhatia, J. Hajnal, B. Puri, A. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No. 04EX821)*, vol. 2. Arlington, VA, USA: IEEE, 2004, pp. 908–911.

[109] M. Unser, P. Thevenaz, and L. Yaroslavsky, "Convolution-based interpolation for fast, high-quality rotation of images," *IEEE Transactions on image processing*, vol. 4, no. 10, pp. 1371–1381, 1995.

[110] D. J. Larkman and R. G. Nunes, "Parallel magnetic resonance imaging," *Physics in Medicine and Biology*, vol. 52, no. 7, pp. R15–R55, Apr. 2007.

[111] M. Zaitsev, C. Dold, G. Sakas, J. Hennig, and O. Speck, "Magnetic resonance imag-

ing of freely moving objects: prospective real-time motion correction using an external optical motion tracking system," *NeuroImage*, vol. 31, no. 3, pp. 1038–1050, Jul. 2006.

[112] J. L. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging," *NeuroImage*, vol. 20, no. 2, pp. 870–888, Oct. 2003.

[113] S. Moeller, S. Ramanna, C. Lenglet, P. K. Pisharady, E. J. Auerbach, L. Delabarre, X. Wu, M. Akcakaya, and K. Ugurbil, "Self-navigation for 3D multishot EPI with data-reference," *Magnetic Resonance in Medicine*, vol. 84, no. 4, pp. 1747–1762, Oct. 2020.

[114] J. R. Teruel, J. M. Kuperman, A. M. Dale, and N. S. White, "High temporal resolution motion estimation using a self-navigated simultaneous multi-slice echo planar imaging acquisition: Motion Estimation Using Self-Navigated SMS-EPI Acquisition," *Journal of Magnetic Resonance Imaging*, vol. 48, no. 3, pp. 780–787, Sep. 2018.

[115] D. C. Hoinkiss, P. Erhard, N.-J. Breutigam, F. von Samson-Himmelstjerna, M. Günther, and D. A. Porter, "Prospective motion correction in functional MRI using simultaneous multislice imaging and multislice-to-volume image registration," *NeuroImage*, vol. 200, pp. 159–173, Oct. 2019.

[116] B. Marami, B. Scherrer, S. Khan, O. Afacan, S. P. Prabhu, M. Sahin, S. K. Warfield, and A. Gholipour, "Motion-robust diffusion compartment imaging using simultaneous multi-slice acquisition," *Magnetic Resonance in Medicine*, vol. 81, no. 5, pp. 3314–3329, May 2019.

[117] E. Ferrante and N. Paragios, "Slice-to-volume medical image registration: a survey," *Medical Image Analysis*, vol. 39, pp. 101–123, Jul. 2017, arXiv: 1702.01636.

[118] R. Beare, B. Lowekamp, and Z. Yaniv, "Image Segmentation, Registration and Characterization in *R* with **SimpleITK**," *Journal of Statistical Software*, vol. 86, no. 8, 2018.

[119] J. V. Manjón, P. Coupé, L. Concha, A. Buades, D. L. Collins, and M. Robles, "Diffusion Weighted Image Denoising Using Overcomplete Local PCA," *PLoS ONE*, vol. 8, no. 9, p. e73021, Sep. 2013.

[120] M. Kuklisova-Murgasova, G. Quaghebeur, M. A. Rutherford, J. V. Hajnal, and J. A. Schnabel, "Reconstruction of fetal brain MRI with intensity matching and complete outlier removal," *Medical Image Analysis*, vol. 16, no. 8, pp. 1550–1564, Dec. 2012.

[121] S. Clare, "Iterative Nyquist ghost correction for single and multi-shot EPI using an entropy measure," in *Proc 11th Annual Meeting ISMRM*, 2003, p. 1041, event-place: Toronto.

[122] K. Setsompop, Q. Fan, J. Stockmann, B. Bilgic, S. Huang, S. F. Cauley, A. Nummenmaa, F. Wang, Y. Rathi, T. Witzel, and L. L. Wald, "High-resolution in vivo diffusion imaging of the human brain with generalized slice dithered enhanced resolution: Simultaneous multislice (gSlider-SMS): High-Resolution Diffusion Imaging With gSlider-SMS," *Magnetic Resonance in Medicine*, vol. 79, no. 1, pp. 141–151, Jan. 2018.

[123] B. Bilgic, B. A. Gagoski, S. F. Cauley, A. P. Fan, J. R. Polimeni, P. E. Grant, L. L. Wald,

and K. Setsompop, "Wave-CAIPI for highly accelerated 3D imaging: Wave-CAIPI for Highly Accelerated 3D Imaging," *Magnetic Resonance in Medicine*, vol. 73, no. 6, pp. 2152–2162, Jun. 2015.

[124] I. P. Bruce, H.-C. Chang, C. Petty, N.-K. Chen, and A. W. Song, "3D-MB-MUSE: A robust 3D multi-slab, multi-band and multi-shot reconstruction approach for ultrahigh resolution diffusion MRI," *NeuroImage*, vol. 159, pp. 46–56, Oct. 2017.

[125] F. Wang, B. Bilgic, Z. Dong, M. K. Manhard, N. Ohringer, B. Zhao, M. Haskell, S. F. Cauley, Q. Fan, T. Witzel, E. Adalsteinsson, L. L. Wald, and K. Setsompop, "Motion-robust sub-millimeter isotropic diffusion imaging through motion corrected generalized slice dithered enhanced resolution (MC-gSlider) acquisition," *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 1891–1906, Nov. 2018.

SUMMARY

Multi-shot acquisitions offer improvements in SNR and resolution for diffusion-weighted imaging (DWI), but the contrast-related sensitivity to motion complicates the combination of the individual shot datasets. The main signal variations for DWI include the inevitable shot-to-shot phase variations and macroscopic subject motion. Model-based image reconstructions incorporate the motion-induced signal variations into the optimization problem and jointly estimate the image along with the motion parameters. However, the joint optimizations are typically non-convex, ill-posed and computationally demanding, depending particularly on the motion model. Therefore, several navigation strategies exist to sense the motion parameters from different types of signals. This dissertation proposes four reconstruction models for motion-corrected multi-shot DWI that are either navigated, using additional MR signals in the sequence, or self-navigated, using the imaging data itself. The four algorithms investigate the use of deep learning to improve the shot phase estimation and evaluate different navigation strategies for retrospective corrections of 2D in-plane and 3D rigid patient motion.

The self-navigated estimation of the shot phase maps from the segmented image data is a high-dimensional problem, which suffers from noise amplification for increasing segmentations. This work evaluates the use of a U-Net to suppress ghosting artifacts for phase-corrupted multi-shot combinations and stabilize the shot phase estimation in an alternating optimization. The U-Net is shown to provide enhanced image magnitude priors for the shot phase estimation, yielding improved image quality for highly segmented multi-shot DWI. By design, the U-Net interacts only with the shot phase estimation, restricting the influence on the joint images to phase-related ghosting.

A recently published multi-shot DWI reconstruction further includes macroscopic motion by rigid in-plane modeling between the shot datasets and derives the rigid motion parameters from the undersampled shot data. To encounter the SNR penalty for increasing segmentations, the algorithm proposed in this work performs alternating optimizations of the shot phase, rigid in-plane and joint image parameters instead of a single estimation. The repeated multi-shot updates reduce the noise amplification for the individual undersampled shot problems achieving refined shot parameters and improved image quality by self-navigation in the presence of subject motion.

An alternative approach uses additional low-resolution MR navigator signals to sense the shot phase and rigid in-plane parameters, but a side-by-side comparison of the navigation strategies is missing so far. This work proposes a motion-corrected image reconstruction algorithm that derives the shot phase and rigid in-plane motion parameters once from shot navigator images. The navigated and self-navigated strategies are compared for in-vivo multi-shot data, proving both effective in reducing in-plane motion artifacts. The navigated strategy shows improved robustness to high segmentations and strong motion, but occasionally suffers from small residual blurring, which demonstrates the difficulties of one-time navigator image registrations at low resolution.

Through-plane motion is another major practical problem that is not captured by the rigid in-plane motion models. This work employs a navigated multi-shot DWI sequence with simultaneous multi-slice (SMS) acquisition to increase the sampling efficiency and obtain data support in the slice direction for 3D motion estimation with high temporal resolution. The low-resolution SMS navigators are used to estimate the shot-specific phase maps and 3D rigid motion states using a multi-slice-to-volume registration. The shot parameters then guide a model-based full-volume reconstruction per diffusion direction including wavelet denoising. The proposed method is validated in simulations and in-vivo, achieving sub-millimeter target registration errors and improved image quality with reduced motion artifacts.

This work is dedicated to model-based image reconstructions which provide effective means to address the shot-to-shot phase variations and macroscopic motion in multi-shot DWI. The proposed methods offer powerful data-driven tools to exploit the joint information in large imaging datasets. Although further technical developments are still necessary, this dissertation presents new motion correction strategies that are directly applicable to state-of-the-art imaging methods in clinical practice, and also contributes to the research field evaluating demanding iterative algorithms to improve the robustness against noise and through-plane motion. The model parameterizations, the navigation strategies and regularizations by deep learning or wavelet denoising were shown to play a major role to effectively constrain the ill-posed motion correction problems. The described models allow extensions for further signal variations and novel encoding schemes. Overall, the model-based motion navigation builds the basis for adaptive image acquisition and reconstruction strategies paving the way for robust, efficient, and high-quality diffusion-weighted imaging.

ZUSAMMENFASSUNG

Segmentierte Akquisitionstechniken ermöglichen Verbesserungen des Signal-Rausch-Verhältnisses und der Auflösung in der diffusionsgewichteten Magnetresonanztomographie (dMRT). Die Kombination der einzelnen Datensegmente wird aber durch die Bewegungsempfindlichkeit des Diffusionskontrastes signifikant beeinträchtigt. Zu den wichtigsten Signalvariationen der dMRT zählen die unvermeidbaren segmentspezifischen Phasenvariationen und die makroskopischen Patientenbewegungen. Modellbasierte Bildrekonstruktionen berücksichtigen diese Signalvariationen und schätzen die Bild- und Bewegungsparameter gemeinsam. Allerdings sind solche gemeinsamen Optimierungsprobleme häufig nichtkonvex, schlecht gestellt und rechenaufwendig. Aus diesem Grund finden diverse Navigationsstrategien Anwendung, die die Bewegungsparameter aus unterschiedlichen Signaltypen ableiten. Diese Dissertation erarbeitet vier neue bewegungskorrigierte Rekonstruktionsmethoden für die segmentierte dMRT. Dabei kommen sowohl navigierte Methoden zum Einsatz, die zusätzliche MRT-Signale in der Messsequenz für die Bewegungsschätzung nutzen, als auch selbstnavigierte Methoden, die die segmentierten Bilddaten dafür verwenden. Auf Grundlage der vier Algorithmen wird zum einen die Unterstützung der Phasenschätzungen durch maschinelles Lernen untersucht, zum anderen werden die unterschiedlichen Navigationsstrategien zur Korrektur rigider 2D- und 3D-Bewegungsmuster evaluiert.

Die selbstnavigierte Schätzung der Phasenvariationen aus den Datensegmenten ist ein hochdimensionales Problem, das für zunehmende Segmentierung unter Rauschpropagationseffekten leidet. Diese Arbeit untersucht die Fähigkeit eines neuronalen Netzwerks, des sog. U-Nets, zur Unterdrückung phaseninduzierter Replika-Artefakte in segmentierten Datensätzen mit dem Ziel, die segmentspezifischen Phasenschätzungen in einem alternierenden Algorithmus zu stabilisieren. Es wird gezeigt, dass das U-Net wertvolle Startschätzungen der gemeinsamen Bildmagnitude bereitstellt, wodurch die Phasenschätzung stabilisiert und die finale Bildqualität für hochsegmentiertes dMRT verbessert wird. Das U-Net interagiert dabei nur mit der Phasenschätzung, wodurch das neuronale Netzwerk von der Bildproduktion entkoppelt bleibt, was potenzielle Störungen auf phaseninduzierte Artefakte beschränkt.

Ein kürzlich publizierter Algorithmus integriert makroskopische Patientenbewegungen in eine segmentierte dMRT-Rekonstruktion, indem rigide 2D-Bewegungen in der Bildebene zwischen den Datensegmenten modelliert und einmalig aus den unterabgetasteten Bilddaten geschätzt werden. Um den Rauschpropagationseffekten für hohe Segmentierungen zu begegnen, wird in dieser Arbeit ein alternierender Algorithmus vorgestellt, der die Phasenvariationen, die rigiden Bewegungsparameter und das gemeinsame Bild in einem iterativen Optimierungsprozess bestimmt. Die wiederholte Einbindung aller Datensegmente reduziert die Rauschverstärkung der einzelnen unterabgetasteten Segmentrekonstruktionen und verbessert dadurch sowohl die Bewegungsparameter als auch die Bildqualität im Falle makroskopischer Patientenbewegung.

Ein alternativer Ansatz nutzt zusätzliche MRT-Navigatorsignale mit niedriger Auflösung, um die Phasenvariationen und die rigiden Bewegungsparameter zu schätzen. Ein direkter Vergleich der unterschiedlichen Navigationsstrategien fehlt allerdings. Diese Dissertation beschreibt dazu einen navigierten Algorithmus, der die Phasenvariationen und die rigiden 2D-Bewegungsparameter für jedes Segment aus den Navigatorbildern bestimmt. Der navigierte und der selbstnavigierte Algorithmus werden anhand von segmentierten In-Vivo-Daten verglichen. Dabei erweisen sich beide als effektiv, um Artefakte aus 2D-Bewegungen zu reduzieren. Die navigierte Methode ist robuster gegenüber hohen Segmentierungen und starken Bewegungen, leidet jedoch vereinzelt unter Unschärfeartefakten, die auf die Schwierigkeit einmaliger Bewegungsschätzungen aus niedrigaufgelösten Navigatorbildern zurückzuführen sind.

Bewegungen, die orthogonal zur Schichtrichtung der Bildkodierung stattfinden, stellen ein weiteres praktisches Problem dar, das von den rigiden 2D-Bewegungsmodellen in der Bildebene nicht erfasst wird. Ein in dieser Arbeit entwickelter Algorithmus verbindet eine segmentierte und navigierte dMRT-Sequenz mit einer Technologie zur simultanen Akquisition mehrerer Schichten, um damit nicht nur die Messeffizienz zu steigern, sondern auch Datenunterstützung in der Schichtrichtung für die 3D-Bewegungsschätzungen mit hoher Zeitauflösung zu erhalten. Die mehrschichtigen Navigatoren mit niedriger Auflösung werden zur Schätzung der Phasenvariationen und der rigiden 3D-Bewegungskomponenten mittels einer Mehrschicht-zu-Volumen-Registrierung verwendet. Die ermittelten Parameter steuern dann eine modellbasierte Volumenrekonstruktion pro Diffusionsrichtung mit einer Regularisierung für spärlich-verteilte Signale. Simulationen und In-Vivo-Rekonstruktionen ergeben für die vorgeschlagene Methode Registrierungsfehler im Submillimeterbereich und eine Verbesserung der Bildqualität durch die Reduktion der Bewegungsartefakte.

Diese Dissertation leistet Beiträge zur modellbasierten Bildrekonstruktion von segmentierten dMRT-Daten und beschreibt effektive Algorithmen zur Korrektur von segmentspezifischen Phasenvariationen und makroskopischen Patientenbewegungen. Die Methoden erweisen sich dabei als leistungsfähige, datengetriebene Werkzeuge zur gemeinsamen Auswertung großer Bildgebungsdatensätze. Diese Arbeit stellt Strategien zur Bewegungskorrektur vor, die unmittelbar in der klinischen Praxis anwendbar sind, beschreibt aber auch aufwendigere, iterative Algorithmen, die robuster gegen Rauschen und 3D-Bewegungsmuster sind, jedoch weiterer technischer Entwicklung für den klinischen Einsatz bedürfen. Die Leistungsfähigkeit der Bewegungskorrekturen hängt maßgeblich von den Modellparametrisierungen, den Navigationsstrategien und effektiven Regularisierungen auf Basis des maschinellen Lernens und der Spärlichkeit (engl. sparsity) der Signale ab. Die untersuchten Modelle sind außerdem auf weitere Signalvariationen und Kodierungstechnologien erweiterbar. Zusammenfassend zeigt sich die modellbasierte Bewegungsnavigation als Kernelement für zukünftige, adaptive Strategien zur Bildakquisition und -rekonstruktion, die eine robuste, effiziente und hochwertige diffusionsgewichtete Bildgebung ermöglichen.

Abbreviations

ADC	Apparent diffusion coefficient
BN	Batch normalization
CAIPI	Controlled aliasing in parallel imaging
CG	Conjugate gradients
DL	Deep learning
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
EPI	Echo-planar imaging
FA	Fractional anisotropy
FFT	Fast Fourier transform
FGP	Fast gradient projections
FISTA	Fast iterative shrinkage-thresholding algorithm
GRAPPA	Generalized autocalibrating partially parallel acquisitions
IRIS	Image reconstruction using image-space sampling function
MAD	Median absolute deviation
MB	Multi-band
MRI	Magnetic resonance imaging
NN	Neural network
nRMSE	Normalized root-mean-square error
PF	Partial Fourier
POCS	Projections onto convex sets
PSF	Point spread function
ReLU	Rectified Linear Unit
RF	Radio frequency
SB	Single-band
SENSE	Sensitivity encoding
SMS	Simultaneous multi-slice
SNR	Signal-to-noise ratio
SPIR	Spectral pre-saturation with inversion recovery
SSFP	Steady-state free precession
STEAM	Stimulated echo acquisition mode
TRE	Target registration error
TV	Total variation
VOI	Volume of interest

Acknowledgments

My doctoral project has been a wonderful and intense period thanks to a lot of inspiring people, critical discussions, stunning talks, joint failures and successes. I thank Prof. Dr. Peter Börnert for the close collaboration and the continuous support during the whole project. It was a pleasure to work in the Philips Research laboratory in Hamburg and I appreciate the supportive and inspiring atmosphere cultivated by all the employees and students. I would like to thank Philips Research for funding the project with the University of Lübeck and all people in both institutions who were involved setting up this collaboration.

Although my everyday life took place in Hamburg, I felt just as welcome at the Institute of Signal Processing in Lübeck. I would like to thank Prof. Dr. Alfred Mertins for enriching discussions, the ongoing support and for making the collaboration project possible. The cross-institutional work was furthermore greatly facilitated by Iris Kruck, who helped me through the organizational formalities and kept me close to the daily life in Lübeck.

On the personal side, I thank my wife who constantly supports me with her joyful nature and who found an excellent balance keeping me on and off the track during this time. I also appreciate the unlimited backing of my parents and the joint activities as well as the honest feedback of my siblings. And, last but not least, I am grateful to have such marvelous friends around me who make both work and life easy and fun.

Malte
Hamburg, March, 2021