

prior expectations, the model of data generation encoded by the likelihood function (equation (2.45)) and the measured data. A useful conclusion beyond the gained intuition is that the regularization parameter λ should be scaled proportional to the noise variance.

2.2.4. Sampling patterns

In common MRI pulse sequences, multiple adjacent Fourier coefficients are acquired. MRI is restricted regarding the sensible choice of sampling patterns, as always a full line of k -space is acquired by frequency encoding (see section 2.1.6) and thus undersampling in this direction would not offer any acceleration. Therefore, 2D images are only accelerated in one dimension and 3D images in two dimensions reducing the undersampling possible without artifacts. The lower spatial frequencies hold most of the information contained in an image. This can be exploited by choosing a sampling pattern that is more dense close to the k -space center. Usually the very center is even sampled completely.

As outlined in section 2.2.2, incoherence of the sampling pattern is a prerequisite for the successful implementation of compressed sensing. Commonly, incoherence is achieved by random sampling¹⁴. For this purpose either polynomial sampling distributions⁴⁸ or (variable density) Poisson disk sampling⁴⁹ is used. The latter enforces a minimal distance between adjacent sampled frequencies, which is advantageous for parallel imaging*. Multiple ways of determining an optimal sampling pattern based on a given reference image have been developed. These schemes rely for example on computing the power spectrum of the reference⁴⁸ or on iteratively adding lines to the sampling pattern based on the largest error found in a previous reconstruction⁵⁰.

2.2.5. Sparsity transforms

According to a review by Jaspan et al.¹⁶ on CS-MRI in clinical applications, the most commonly used sparsifying transforms are wavelet and finite difference transforms. The latter is equivalent to total variation regularization. Wavelet transforms can yield sparse representations even of complex images as evidenced by their use in the compression format JPEG2000⁴³. For the reconstruction of much simpler ¹⁹F MR images, one can rely on the sparsity of the image itself directly and a wavelet transform is not needed^{17,18,51}. Often a total variation constraint is added regardless of the primary sparsifying transform, as it leads to improved denoising and stability of the algorithm. There exist both anisotropic and isotropic forms of the total variation. The anisotropic form was introduced in section 2.2.1 and is the ℓ^1 norm of a linear transform. Thus, it can be incorporated into the CS reconstruction directly:

$$T_{\text{aniso}}(\mathbf{r}) = \sum_{\alpha=x,y,z} \|\nabla_{\alpha}\mathbf{r}\|_1 = \left\| \begin{pmatrix} \nabla_x \\ \nabla_y \\ \nabla_z \end{pmatrix} \mathbf{r} \right\|_1, \quad (2.50)$$

where ∇_{α} is the finite differences transform matrix in direction α , such that e.g. $(\nabla_x \mathbf{r})_{i,j,k} = r_{i+1,j,k} - r_{i,j,k}$. The isotropic total variation on the other hand is nonlinear and thus leads to additional computational difficulties:

$$T_{\text{iso}}(\mathbf{r}) = \sum_{i,j,k} \sqrt{\sum_{\alpha=x,y,z} |(\nabla_{\alpha}\mathbf{r})_{i,j,k}|^2}. \quad (2.51)$$

If use of the isotropic form is desired, it is added as an additional term into the compressed sensing reconstruction:

$$\hat{\mathbf{r}} = \min_{\mathbf{r}} \left(\|F_{\text{u}}\mathbf{r} - \mathbf{y}\|_2^2 + \lambda (w_{\Phi} \|\Phi\mathbf{r}\|_1 + w_{\text{T}} T_{\text{iso}}(\mathbf{r})) \right). \quad (2.52)$$

*Parallel imaging techniques are used with multiple receive element coil arrays and utilize the spatially varying sensitivity of the individual arrays to gain additional spatial information. As they also accelerate the image acquisition by undersampling, the combination of parallel imaging and CS stands to reason, but multiple receive element arrays are usually not available for small animal MRI. Thus parallel imaging was not investigated in this thesis.

3. Methods

All data preparation, reconstructions and analysis was performed in MATLAB R2017b (The MathWorks Inc., Natick, USA) relying on the pvmatlab package for handling MR data generated by ParaVision 6.1.0 (Buker Biospin MRI GmbH, Ettlingen, Germany).

3.1. Compressed sensing reconstructions and undersampling

3.1.1. The sparsity transform

We used the ℓ^1 norm of the image itself as a regularization term to exploit the sparsity of fluorine MR images. Additionally we employed an total variation constraint to aid with denoising. The isotropic form was chosen to avoid possible artifacts due to the preference of the anisotropic form for horizontal and vertical edges [2.2.5]. The two regularization terms were given equal weight. Thus, the used reconstruction equation was equal to equation (2.52) with Φ set to I_n and $w_\Phi = w_T = 1$:

$$\hat{\mathbf{r}} = \min_{\mathbf{r}} \left(\|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \lambda (\|\mathbf{r}\|_1 + T_{\text{iso}}(\mathbf{r})) \right). \quad (3.1)$$

A further optimization of the sparsifying transforms was found to be beyond the scope of this thesis.

3.1.2. A fast reconstruction algorithm for fluorine MRI compressed sensing

As the compressed sensing reconstruction equation does not have a known analytical solution, an approximation must be found using an iterative algorithm¹⁴. The Berkeley Advanced Reconstruction Toolbox (BART) is currently the most advanced available toolbox for CS-MRI, but it does not offer regularization with the ℓ^1 norm of the image itself, which is desirable for fluorine MRI. Thus, I decided to implement an algorithm myself in MATLAB*. As the augmented Lagrangian methods are the most promising algorithms for problems involving a total variation constraint [2.2.7], I chose to develop an implementation of the accelerated version of the alternating direction method of multipliers (ADMM) proposed by Goldstein et al.⁵⁷. This accelerated version has to my knowledge not yet been adapted for CS-MRI.

One of the sub-steps of the ADMM necessitates a matrix inversion which is not feasible straightaway on normal workstations due to the problem size. I implemented a fast Fourier transform (FFT) based matrix inversion exploiting the matrix structure, which greatly accelerates the algorithm compared to the approximate solution. Diagonalization of appropriate matrices via the Fourier transform is a standard technique for convolutions in image processing⁵⁸. Nonetheless, the corresponding inversion scheme has not commonly been adopted for CS-MRI reconstruction algorithms as evidenced by the reliance of e.g. the state-of-the-art and continuously maintained BART on an approximation via conjugate gradient descent. This is partly due to the necessity of periodic boundary conditions, which can lead to wrap around artifacts in hydrogen imaging. For fluorine MRI periodic boundary conditions are not problematic, as the borders of the field of view should only contain background. Goldstein et al.⁵⁷ propose FFT-based inversion for an example similar to the context at hand. The group of Jeffrey Fessler published related algorithms for compressed sensing with parallel imaging. Ramani and Fessler⁵⁶ combined FFT-based inversion with an augmented Lagrangian method similar, but not identical to ADMM, while Le and Fessler⁵⁹ present a combination of ADMM with a tridiagonal matrix solver that does not necessitate periodic boundary conditions and for which they claim performance competitive with the groups previous FFT-based inversion scheme⁵⁶.

Section [3.3.2] describes tests evaluating the performance of the method described here relative to conventional ADMM, BART and ADMM with an approximate conjugate gradient descent substep. The corresponding results are given in section [4.1.1].

*MATLAB Release 2017a, The MathWorks, Inc., Natick, Massachusetts, United States.

To my knowledge there is currently no publication providing the update equations for the ADMM and CS-MRI covering isotropic total variation, three-dimensional data and a careful treatment of the complex nature of the MRI signal. Therefore this chapter and appendix [A.2] contain a full derivation. It was based on Goldstein et al.⁵⁷ who treat the general case of minimizing a function

$$H(\mathbf{u}) + G(\mathbf{v}) \quad \text{subject to} \quad A\mathbf{u} + B\mathbf{v} = b \quad (3.2)$$

and the CS-MRI inspired example

$$\hat{\mathbf{r}} = \arg \min_{\mathbf{r}} \left(\|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \lambda_1 \|\nabla_x \mathbf{r}\|_1 + \lambda_2 \|\nabla_y \mathbf{r}\|_2 \right). \quad (3.3)$$

Furthermore I adapted parts of the update equations regarding the isotropic total variation from Goldstein and Osher⁶⁰, where the example of isotropic total variation denoising

$$\hat{\mathbf{r}} = \arg \min_{\mathbf{r}} \left(\|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \lambda T_{\text{iso}}(\mathbf{r}) \right) \quad (3.4)$$

is discussed.

In the following, the algorithm will be developed for compressed sensing MRI with isotropic total variation regularization and an arbitrary linear sparsifying transform Φ . A solution of the update equations with exact FFT-based inversion will only be derived for the special case $\Phi = I_n$. An alternative conjugate gradient based update is described in appendix [A.2.4]. The to be solved equation is:

$$\hat{\mathbf{r}} = \arg \min_{\mathbf{r}} \left(\|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \lambda (w_\Phi \|\Phi \mathbf{r}\|_1 + w_T T_{\text{iso}}(\mathbf{r})) \right), \quad (3.5)$$

where $\hat{\mathbf{r}} \in \mathbb{C}^n$ is the reconstruction, $F_u \in \mathbb{C}^{m \times n}$ the undersampled Fourier transform, $\mathbf{y} \in \mathbb{C}^m$ the measured data and $\lambda \in \mathbb{R}$ the regularization strength. $\Phi \in \mathbb{C}^{n_\Phi \times n}$ is a matrix comprising all linear transforms used for regularization and w_Φ and w_T are weights controlling the balance between the linear constraint and the isotropic total variation T_{iso} [2.2.5].

The ADMM approaches an unconstrained optimization problem by introducing a new auxiliary variable transforming it into a simpler but constrained optimization problem. Equation (3.5) is substituted by minimizing the objective function

$$O(\mathbf{r}, \mathbf{d}) = \|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \lambda (\|\mathbf{d}_\Phi\|_1 + T_{\text{iso}}(\mathbf{d}_\nabla))$$

$$\text{subject to } \mathbf{d} = \begin{pmatrix} \mathbf{d}_\Phi \\ \mathbf{d}_\nabla \end{pmatrix} = \begin{pmatrix} \mathbf{d}_\Phi \\ \begin{pmatrix} \mathbf{d}_x \\ \mathbf{d}_y \\ \mathbf{d}_z \end{pmatrix} \end{pmatrix} = \begin{pmatrix} w_\Phi \Phi \\ \begin{pmatrix} w_T \nabla_x \\ w_T \nabla_y \\ w_T \nabla_z \end{pmatrix} \end{pmatrix} \mathbf{r}. \quad (3.6)$$

This reformulation is advantageous, because it allows separate iterative optimization regarding \mathbf{r} and \mathbf{d} . The individual steps will thus be much simpler, possibly making the use of exact instead of approximate steps possible. The downside is, that the constraint must be enforced in some way. To this end, an augmented objective functions is constructed incorporating a quadratic deviation penalty and a Lagrange multiplier:

$$\tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) = O(\mathbf{r}, \mathbf{d}) + \frac{\mu}{2} \|\mathbf{d} - \Psi \mathbf{r}\|_2^2 + \text{Re}(\mathbf{k}^\dagger (\mathbf{d} - \Psi \mathbf{r})). \quad (3.7)$$

Ψ is a shorthand for $(w_\Phi \Phi^T, w_T \nabla_x^T, w_T \nabla_y^T, w_T \nabla_z^T)^T$, μ a scalar parameter controlling the weight of the constraint enforcing terms and \mathbf{k} the complex Lagrange multiplier with dimension equal to \mathbf{d} . The ADMM solves equation (3.5) by iterating the steps

- a) $\mathbf{d}^{(j+1)} = \arg \min_{\mathbf{d}} \tilde{O}(\mathbf{r}^{(j)}, \mathbf{d}, \mathbf{k}^{(j)})$
- b) $\mathbf{r}^{(j+1)} = \arg \min_{\mathbf{r}} \tilde{O}(\mathbf{r}, \mathbf{d}^{(j+1)}, \mathbf{k}^{(j)})$
- c) $\mathbf{k}^{(j+1)} = \mathbf{k}^{(j)} + \mu (\mathbf{d}^{(j+1)} - \Psi \mathbf{r}^{(j+1)})$.

For the first iteration, the Lagrange multiplier \mathbf{k} is initialized to zero.

The improvement proposed by Goldstein et al.⁵⁷ consists simply of a predictor-corrector type acceleration step incorporating information from a previous step into the next update. This is similar to the way a conjugate gradient descent improves upon a pure gradient descent. The influence of the preceding step is controlled by the weighting parameter γ , which is initialized to one. Introducing $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{k}}$ to denote the modified iterations of \mathbf{r} and \mathbf{k} , the new update equations are

$$\begin{aligned}
\text{a)} \quad & \mathbf{d}^{(j+1)} = \arg \min_{\mathbf{d}} \tilde{O}(\tilde{\mathbf{r}}, \mathbf{d}, \tilde{\mathbf{k}}) \\
\text{b)} \quad & \mathbf{r}^{(j+1)} = \arg \min_{\mathbf{r}} \tilde{O}(\mathbf{r}, \mathbf{d}^{(j+1)}, \tilde{\mathbf{k}}) \\
\text{c)} \quad & \tilde{\mathbf{k}} = \mathbf{k}^{(j)} + \mu (\mathbf{d}^{(j+1)} - \Psi \mathbf{r}^{(j+1)}) \\
\text{d)} \quad & \gamma^{(j+1)} = \left(1 + \sqrt{1 + (2\gamma^{(j)})^2} \right) / 2 \\
\text{e)} \quad & \tilde{\mathbf{r}} = \mathbf{r}^{(j+1)} + \frac{\gamma^{(j)} - 1}{\gamma^{(j+1)}} (\mathbf{r}^{(j+1)} - \mathbf{r}^{(j)}) \\
\text{f)} \quad & \tilde{\mathbf{k}} = \mathbf{k}^{(j+1)} + \frac{\gamma^{(j)} - 1}{\gamma^{(j+1)}} (\mathbf{k}^{(j+1)} - \mathbf{k}^{(j)})
\end{aligned}$$

Goldstein et al.⁵⁷ also proposed a restart rule for the case that predictor-corrector steps e) and f) go astray. For the application at hand it was found to be not needed and, as it consumes a considerable amount of computational power, it is not used.

Update equations for steps a) and b) are derived in appendix [A.2]. As $\tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k})$ is a real-valued function with complex vector input, the stationary points of the function and thus the potential minima can not be found using the standard tools of complex analysis. Instead I rely on $\mathbb{C}\mathbb{R}$ -calculus which treats a complex number as a two-dimensional real vector and thus enables the use of results from vector analysis. Basic corollaries from $\mathbb{C}\mathbb{R}$ -calculus are given in appendix [A.2.1].

I use the partitioning $\tilde{\mathbf{k}} = (\tilde{\mathbf{k}}_{\Phi}^T, \tilde{\mathbf{k}}_x^T, \tilde{\mathbf{k}}_y^T, \tilde{\mathbf{k}}_z^T)^T$, where the dimensions are chosen analogously to the partitions of \mathbf{d} (see equation (3.6)). Furthermore, the shorthand \mathbf{s}_{β} denotes $w_{\mathbf{T}} \nabla_{\beta} \tilde{\mathbf{r}} - \frac{1}{\mu} \tilde{\mathbf{k}}_{\beta}$ for $\beta = x, y, z$. It follows for $i = 1, \dots, n$ and $\beta = x, y, z$

$$\begin{aligned}
& \mathbf{d}_{\Phi}^{(j+1)} = \text{shrink} \left(\Phi \tilde{\mathbf{r}} - \frac{1}{\mu} \tilde{\mathbf{k}}, \frac{\lambda}{\mu} \right) \\
\text{a)} \quad & (\mathbf{d}_{\beta})_i^{(j+1)} = \frac{(\mathbf{s}_{\beta})_i}{\sqrt{\sum_{\gamma=x,y,z} |(\mathbf{s}_{\gamma})_i|^2}} \max \left(\sqrt{\sum_{\gamma=x,y,z} |(\mathbf{s}_{\gamma})_i|^2} - \frac{\lambda}{\mu}, 0 \right) \\
\text{b)} \quad & \mathbf{r}^{(j+1)} = \left(2F_u^{\dagger} F_u + \mu \Psi^{\dagger} \Psi \right)^{-1} \left(\Psi^{\dagger} \left(\mu \mathbf{d}^{(j+1)} - \tilde{\mathbf{k}} \right) + 2F_u^{\dagger} \mathbf{y} \right)
\end{aligned}$$

with the definition of the shrinkage function

$$\text{shrink} : \mathbb{C}^n \times \mathbb{R} \mapsto \mathbb{C}^n, (\text{shrink}(\mathbf{z}, \delta))_i = \frac{z_i}{|z_i|} \max(|z_i| - \delta, 0). \quad (3.8)$$

Although step b) seems to break down to a simple matrix inversion, it is not feasible straight-away on a common workstation in a reasonable time-frame, as it requires the inversion of a complex $n \times n$ matrix, where n is the number of image voxels. Thus, for its explicit storage in double precision for e.g. a $128 \times 64 \times 64$ voxel 3D dataset $2 \cdot (128 \cdot 64 \cdot 64)^2 \cdot 8 \text{ B} \approx 4.4 \text{ TB}$ of memory would be needed. Furthermore, using simple Gauss elimination the inversion would have

a computational complexity of $\mathcal{O}(n^3)$. In many applications \tilde{O} is therefore instead minimized approximately using conjugate gradient descent⁵⁷.

Instead, I implemented a fast Fourier transform based inversion exploiting the matrix structure. This is always possible if the to be inverted matrix encodes a convolution (see appendix [A.2.3] for details). For the used sparsifying transforms, image space and total variation regularization, this is the case if periodic boundary conditions are chosen for the finite differences matrices $\nabla_x, \nabla_y, \nabla_z$. For $\Phi = I_n$ update step b) becomes

$$\mathbf{r}^{(j+1)} = F^\dagger (2R + w_\Phi^2 I_n + w_\Gamma^2 E_\Delta)^{-1} F \left(\Psi^\dagger (\mu \mathbf{d}^{(j+1)} - \tilde{\mathbf{k}}) + 2F_u^\dagger \mathbf{y} \right), \quad (3.9)$$

where R is a diagonal matrix such that $F_u^\dagger F_u = F^\dagger R F$ and

$$E_\Delta = F^\dagger (\Delta_x + \Delta_y + \Delta_z) F = \text{diag} \left(\sqrt{n} F (\Delta_x + \Delta_y + \Delta_z)_{*,1} \right) \quad (3.10)$$

Here $_{*,1}$ denotes the first column of a matrix and Δ_γ is the finite differences Laplace operator $\nabla_\gamma^\dagger \nabla_\gamma$ for $\gamma = x, y, z$. The diagonal entries of E_Δ are the eigenvalues of $\Delta_x + \Delta_y + \Delta_z$. Thus the inversion can be performed for the cost of two fast Fourier transforms with complexity $\mathcal{O}(n \log n)$. E_Δ and $(2R + I_n + E_\Delta)^{-1}$ can be precalculated, as they are identical on every iteration.

3.1.3. Automatic selection of the regularization strength

The regularization strength λ controls the balance between the measured data and the sparsity constraint. When dealing with noisy data without prior knowledge of the to be expected image, it is not feasible to adjust the parameter based on a visual assessment of the reconstruction. Thus it was deemed necessary to have an automatic selection of the regularization strength to enable a fair comparison between different measurement schemes and investigate a reproducible workflow independent of the operator.

Given a known noise level σ and m sampled data points, the expected deviation between the measured data \mathbf{y} and the true, but unknown image $\bar{\mathbf{r}}$ is

$$\begin{aligned} \epsilon' &= \left\langle \|\mathbf{F}_u \bar{\mathbf{r}} - \mathbf{y}\|_2^2 \right\rangle = \left\langle (\mathbf{F}_u \bar{\mathbf{r}} - \mathbf{y})^\dagger (\mathbf{F}_u \bar{\mathbf{r}} - \mathbf{y}) \right\rangle \\ &= \sum_{i=1}^m \left[\left\langle (\text{Re}(\mathbf{F}_u \bar{\mathbf{r}}) - \text{Re}(\mathbf{y}))_i^2 \right\rangle + \left\langle (\text{Im}(\mathbf{F}_u \bar{\mathbf{r}}) - \text{Im}(\mathbf{y}))_i^2 \right\rangle \right] \\ &= 2\sigma^2 m, \end{aligned} \quad (3.11)$$

as $\text{Re}(\mathbf{y})$ is distributed according to a multivariate Gaussian distribution with expectation $\text{Re}(\mathbf{F}_u \bar{\mathbf{r}})$ and covariance matrix $\sigma^2 I_m$ and $\text{Im}(\mathbf{y})$ according to a similar distribution but with expectation $\text{Im}(\mathbf{F}_u \bar{\mathbf{r}})$ [2.1.8].

To recover the true, but unknown image, the deviation between the reconstruction $\hat{\mathbf{r}}$ and the measured data $\epsilon = \|\mathbf{F}_u \hat{\mathbf{r}} - \mathbf{y}\|_2^2$ would thus have to be equal to ϵ' . In practice, it can be advantageous to set the desired deviation $\tilde{\epsilon}$ slightly below this level to avoid oversmoothing¹⁴. This is denoted by the factor $\eta > 0$ defined by $\tilde{\epsilon} = \eta \epsilon'$.

Given a dataset and a fixed sparsity transform, the compressed sensing reconstruction depends on λ alone and can be denoted as $\hat{\mathbf{r}}(\lambda)$. The function

$$f(\lambda) = \|\mathbf{F}_u \hat{\mathbf{r}}(\lambda) - \mathbf{y}\|_2^2 - \eta \epsilon' = \zeta \quad (3.12)$$

is monotonically increasing. Thus the regularization strength corresponding to $\zeta = 0$ can be easily found using a two-point bracketing method. Such methods determine the root of a function starting with a search interval enclosing a sign change. The interval is then iteratively shrunk until a desired tolerance is reached. I chose to implement a variant of the Illinois algorithm^{61,62},

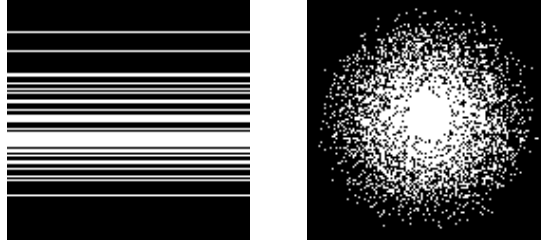


Fig. 17. **MRI sampling masks:** In MRI always a complete line of k -space is acquired. Here two sampling patterns drawn from polynomial variable density sampling distributions are shown (3.13) and (3.14)). They include a quarter of all points of a 128×128 grid. The parameter d governing the balance between the k -space center and periphery was set to 2. **(Left)** A 2D image is only undersampled in the single phase encoding direction. **(Right)** A 3D image can be undersampled in both phase encoding directions. The third dimension not depicted here is fully sampled.

which is an improved version of the false position method. Details regarding the two parts of the algorithm, finding the two initial bracketing points and the Illinois algorithm proper, can be found in appendix [A.3].

Based on simulations described in sections [3.3.1], η was set to 0.970 ± 10^{-3} for all reconstructions if not stated otherwise. The noise standard deviation σ was estimated based on a background region in a Fourier reconstruction of fully-sampled data. For prospectively undersampled acquisitions σ could be measured by a pure noise scan without excitation pulses⁶³.

3.1.4. Random variable density sampling

I implemented a polynomial variable density sampling distribution as described in Zijlstra et al.⁶⁴ and adapted it for undersampling in only one dimension, which corresponds to the acquisition of a two-dimensional MR image [2.2.4]. For two-dimensional undersampling, the sampling distribution is

$$p(k_x, k_y) \propto \max\left(1 - \sqrt{k_x^2 + k_y^2}, 0\right)^d, \quad (3.13)$$

where the coordinates k_x and k_y are scaled between -1 and 1 and $d > 0$ is a factor controlling the balance between sampling the k -space center and the periphery. For undersampling only in the k_x direction this reduces to

$$p(k_x) \propto (1 - |k_x|)^d. \quad (3.14)$$

For all patterns, 10% of the sampled coefficients were deterministically assigned to the most central k -space lines and included into the mask by default to ensure that none of the most important frequencies are missing. Typical sampling patterns are depicted in figure 17. Experiments on the influence of the pattern on the reconstruction quality are described in section [3.3.3]. If not stated otherwise, $d = 1.5$ was used for all experiments.

3.2. Image analysis

3.2.1. Noise bias correction

As described in section [2.1.8], the signal in noisy MR magnitude images follows a Rician distribution. The measured signal is biased upwards relative to the true, but unknown signal. This effect is negligible when working with high SNR data, but needs to be corrected in fluorine imaging. The expected measured signal is

$$\langle D_m \rangle = \sigma \sqrt{\frac{\pi}{2}} L_{1/2} \left(-\frac{D_t^2}{2\sigma^2} \right) \quad (3.15)$$

where D_m is the measured signal, D_t the true signal, σ the noise standard deviation of the underlying Gaussian distribution and $L_{1/2}$ a Laguerre polynomial. No inversion of (3.15) to

Enforcing sparsity in CS reconstruction leads to a downwards bias of the signal intensity^{77,78}. Thus, we also investigated the deviation of measured signal intensities from the true or reference value. The deviation was defined as just

$$\frac{\text{measured intensity}}{\text{reference intensity}} \quad (3.22)$$

for any given voxel. The deviation was then either averaged over all true positive voxels, or a signal level specific signal bias was computed similar to the TPR and FDR described in the preceding section. To separate different bias contributions, this was done in two ways:

1. For bins sorted by the reference signal and in contrast to all other analysis without prior thresholding of the reconstruction. This way there is no detection bias effect and the obtained value only depends on the reconstruction method. Similar to what was done for the TPR, results for multiple reconstructions were condensed by computation of average and standard deviation weighted by the number of positives (see equations (3.18), (3.19)).
2. For bins sorted by the measured signal. Here the metric reflects the perspective of a scientist analyzing data without prior knowledge of the signal distribution. The metric is computed for the thresholded reconstructions and only true positive voxels are taken into account. Here the average and standard deviation were computed weighted by the sum of true positives number of true positives.

Furthermore, we also explored a simple method of correcting the signal bias based on the bias observed in a comparable object. The *ex vivo* EAE data was used to correct the signal intensities in reconstructions of the *in vivo* data [3.4.3] by applying a signal level specific correction factor. For voxel i of a reconstruction according to acquisition and reconstruction method β , the corrected signal is

$$S_i^{(c)} = \frac{S_i}{d_{ex}^\beta(S_i)} \quad (3.23)$$

where S_i is the uncorrected intensity and $d_{ex}^\beta(S_i)$ is the deviation measured *ex vivo* for method β in the signal range containing S_i . Results for the remaining bias of the corrected data are computed and presented according to version 2 described above.

3.3. Tests of the compressed sensing implementation

I designed a digital phantom to test the automatic regularization strength adjustment, the performance of the developed reconstructions algorithm and to examine the effect of the sampling pattern on the reconstruction quality under low SNR conditions. It was desired that the phantom would consist of small features similar to the spots typical for fluorine MR images. They should cover a range of intensities, such that differences in sensitivity between the acquisition and reconstruction schemes manifest as different numbers of detected features. The phantom was designed such that it could be easily replicated as a real object (section [3.4.2]). We chose a 2D image with 8 concentrically arranged circles of equal diameter (see figure 21 A). The circles were homogeneous and had signal levels of $\frac{1}{8}, \frac{2}{8}, \frac{3}{8}, \dots, \frac{8}{8}$. To simulate the partial volume effects that occur when a continuous object is represented by a digital image, the phantom was generated at a 20 times higher resolution and then downsampled to 128×128 pixels* (see figure 21 B and C). The circles had a diameter of about 5 pixels.

3.3.1. Finding the optimal regularization strength

To test the automatic regularization strength adjustment and the dependence of the image quality on the chosen deviation from the measured data $\epsilon = \eta\epsilon'$, a sweep of η values was performed. For 20 values between 0.9 and 1.05, I computed reconstructions of 50 k -space datasets

*The downsampling was performed using the MATLAB function `imresize` with bilinear interpolation.

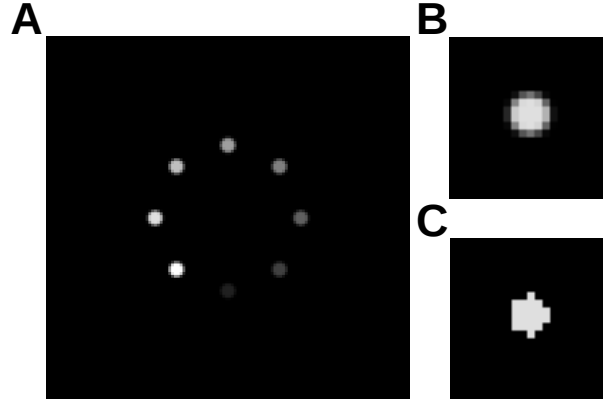


Fig. 21. **Digital phantom:** (A) Digital phantom used in experiments described in section [3.3]. The phantom has a size of 128×128 pixels. Subfigure (B) highlights the simulated partial volume effects characteristic for digitally sampled images of continuous real-world objects. The signal pixels in the image have different intensities although the underlying object is homogeneous. (C) demonstrates how the same round feature would look without taking partial volume effects into account.

sampled from the digital phantom with different undersampling masks and noise realizations. This was done on the one hand for five different noise levels ($\sigma = 0.01$ to $\sigma = 0.02$) and 4-fold undersampling and on the other hand for five different undersampling factors ($f_{\text{us}} = 1/8$ to $f_{\text{us}} = 1$) at an intermediate noise level ($\sigma = 0.04$). In contrast to all other analysis, the NRMSD was computed for the raw reconstructions without thresholding to include the denoising of the background in the metric and obtain a result independent of a threshold value. The variability of the obtained reconstruction quality was assessed by computing the standard deviation over all reconstruction with the same conditions.

3.3.2. Algorithm performance

I compared the performance of five different CS reconstruction algorithms and implementations:

1. Accelerated ADMM with analytical update steps as described in section [3.1.2]
2. ADMM with analytical update steps, but without the predictor-corrector acceleration proposed by Goldstein et al.⁵⁷
3. ADMM with an approximate conjugate-gradient update step in the implementation of the Berkeley Advanced Reconstruction Toolbox (BART)*
4. Accelerated ADMM with an approximate conjugate-gradient update step in a self-built MATLAB implementation (see appendix [A.2.4] for details)
5. ADMM with an approximate conjugate-gradient update step, but without the predictor-corrector acceleration

Algorithm number 5 was included in the comparison because BART is written in C, which is in contrast to MATLAB a compiled and not an interpreted language and thus will run the same computations faster. This is somewhat offset by the fact that in MATLAB standard routines such as the FFT are called as C or C++ programs as well.

With each algorithm, reconstructions of 40 different datasets sampled from the digital phantom with 3-fold undersampling, individual undersampling masks and noise standard deviation $\sigma = 0.02$ were computed. The regularization strength λ was set to 0.04. For each dataset the reconstructions were stopped after 2, 4, 6, 8, \dots , 60 ADMM steps to record the value of the to be minimized objective function over time, as it is not computed for the algorithm itself. As

*BART Toolbox for Computational Magnetic Resonance Imaging, DOI: 10.5281/zenodo.592960, <https://mrrecon.github.io/bart/>

BART does not offer the option to stop after a predetermined number of ADMM steps, it was stopped after 20, 40, 60, \dots , 600 conjugate gradient descent steps. The obtained values for the computation time and the objective function were averaged. To exclude influences of different initialization times, which do not reflect the performance of the algorithm itself but only details of the implementation, the first data point was discarded and the corresponding computation time subtracted from all following entries.

3.3.3. Dependence of the optimal sampling pattern on the noise level

Undersampling patterns in MRI favor the low k -space frequencies, as they contain most of the contrast information. In the used polynomial sampling distributions the balance between sampling the k -space center and periphery is governed by the parameter d . For $d = 0$ all coefficients are sampled with equal probability, whereas for $d \rightarrow \infty$ only the k -space center is recorded, which is equivalent to a measurement at a lower spatial resolution. For both undersampling in one dimension and undersampling in two dimensions, I investigated 10 different sampling distributions: $d = 0, 0.5, 1, 1.5, \dots, 4.0$ and a low resolution sampling pattern. Similarly to the experiments on the regularization strength ([3.3.1]), I considered on the one hand ten different noise levels between $\sigma = 0.01$ and $\sigma = 0.02$ with 4-fold undersampling and on the other hand five different undersampling factors between $f_{\text{us}} = 1/8$ and $f_{\text{us}} = 1/2$ at $\sigma = 0.04$. For each condition 50 datasets with different undersampling masks and noise realizations were sampled from the digital phantom. The measured NRMSDs between the reconstructions and the reference were averaged.

3.4. The application of CS to fluorine MRI

To test the main hypotheses investigated in this thesis, that CS can improve the sensitivity of fluorine MRI, we performed four different types of experiments

1. Simulations based on a digital phantom
2. MR experiments with a purpose-built capillary tube phantom
3. *In vivo* MR experiments in a mouse model of neuroinflammation
4. *Ex vivo* MR experiments with phantoms prepared from tissue samples of the aforementioned mouse model

Fourier and denoised reconstructions of conventional fully-sampled data were compared to CS reconstructions of undersampled data. Undersampling accelerates the MRI data acquisition such that in the same measurement time n times more averages of n -fold undersampled data can be acquired than of fully-sampled data. To obtain fair and meaningful results, we only compared reconstructions corresponding to equal measurement times. For the digital phantom simulations this mean that the standard deviation of the sampled noise was adjusted: $\sigma \rightarrow \sigma/\sqrt{n}$ for $f_{\text{us}} = 1/n$. For the MR experiments, the same was achieved by averaging multiple repetitions of the same fully-sampled measurement before retrospective undersampling. Thus I write of undersampling and averaging, which is denoted by the factor α . $\alpha = 2$ for example signifies that the data was 2-fold undersampled ($f_{\text{us}} = 1/2$) and 2 times more repetitions were averaged than for the Fourier reconstruction. This data preparation is illustrated in figure [22].

All MR experiments were carried out on a 9.4T horizontal bore MR system (BioSpec 94/20, Bruker BioSpin MRI, Ettlingen). This system has a bore of 120 mm, an actively-shielded B-GA12 gradient set, is equipped with a Bruker Avance III console and operated using PV6.1 software (Bruker BioSpin, Ettlingen, Germany). MR data was acquired using an in-house built, dual-tunable $^{19}\text{F}/^1\text{H}$ volume mouse head RF coil⁴.

4. Results

4.1. Tests of the compressed sensing implementation

4.1.1. Algorithm performance

The developed analytical update step sped up the CS reconstructions by more than an order of magnitude compared to an implementation of the ADMM with a conjugate gradient descent inner update. Compared to the state of the art BART, the needed reconstruction time is approximately halved, despite the advantage BART has through being implemented in a compiled language [3.3.2]. The accelerated ADMM combined with the analytical update step provides the fastest minimization of the objective function, yet the advantage gained by the predictor-corrector acceleration of Goldstein et al.⁵⁷ is small. Combined with the approximate conjugate gradient inner step the acceleration is not beneficial. The computation is slowed down and the algorithm does not find the global minimum of the objective function in 60 ADMM iterations. Here, application of the restart rule proposed by Goldstein et al.⁵⁷ would be necessary, yet this would further slow down the reconstruction. The algorithms all reach different values of the objective function in the number of iterations investigated here. If this implies significant differences in the reconstruction quality, lies beyond the scope of this thesis.

4.1.2. The optimal regularization strength

We found that a regularization strength enforcing a deviation between the measured data and the reconstruction close to the deviation between the measured data and the true, but unknown image based on the noise level reliable yielded near optimal reconstructions of the digital phantom in terms of the NRMSD from the reference. Four 4-fold undersampling and a wide range of noise levels (figure 24 A), the differences in average reconstruction quality did not exceed the noise-induced variability for values of η between 0.9 and 1. For $\eta > 1$ the reconstructions deviate from the reference increasingly. The optimal image quality is achieved for values of η between 0.94 and 0.98. The results show a clear trend towards higher values of η being optimal for higher noise levels. Figure 24 B shows results for an intermediate noise level and various undersampling factors from pure denoising ($f_{\text{us}} = 1$) to CS with $f_{\text{us}} = 1/8$. For reconstructions from completely sampled data or only 2-fold undersampled data the optimal range of η was found to be narrower with pronounced quality losses for $\eta < 0.95$ and $\eta > 1$. The minimal NRMSD is again found for all undersampling factors between $\eta = 0.94$ and $\eta = 0.98$ with the higher values corresponding to $f_{\text{us}} = 1$ and $f_{\text{us}} = 1/2$. $f_{\text{us}} = 1/6$ and $f_{\text{us}} = 1/8$ show increased variability of the results, as the sampling pattern adds a second source of variation with increased impact for severe undersampling. Based on these results, η was set to 0.97 for all following reconstructions reflecting the focus on low SNR situations and the tolerance of the more undersampled reconstructions for the higher value.

4.1.3. Dependence of the optimal sampling pattern on the noise level and undersampling factor

Our simulations showed that the optimal balance between predominantly sampling the k -space center or sampling more uniformly depends on the noise level (see figure 25). This was investigated for 4-fold undersampling ($f_{\text{us}} = 1/4$) using the digital phantom. For both undersampling in one dimension and undersampling in two dimensions, uniformly random or close to uniformly random undersampling was advantageous for a noise standard deviation $\sigma = 0.01$. In a conventional fully-sampled Fourier reconstruction this would have corresponded to an SNR of 100, which is well beyond what is achievable in typical fluorine MRI applications. On the other hand, for $\sigma = 0.2$ low resolution sampling was optimal in both cases. Between these two extremes we found a smooth transition. In general, undersampling in two dimensions showed better results

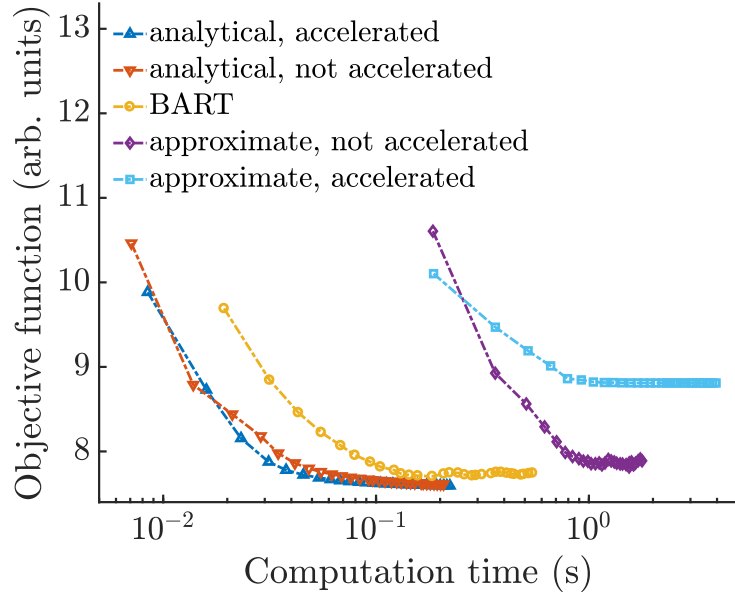


Fig. 23. **Reconstruction algorithm speed:** The to be minimized objective function is plotted versus the computation time. Every marker corresponds to two iterations of the ADMM with the exception of BART where the minimization is plotted in steps of 20 conjugate gradient descent iterations. The example used the 2D digital phantom described in section with a fixed regularization strength. [3.3]

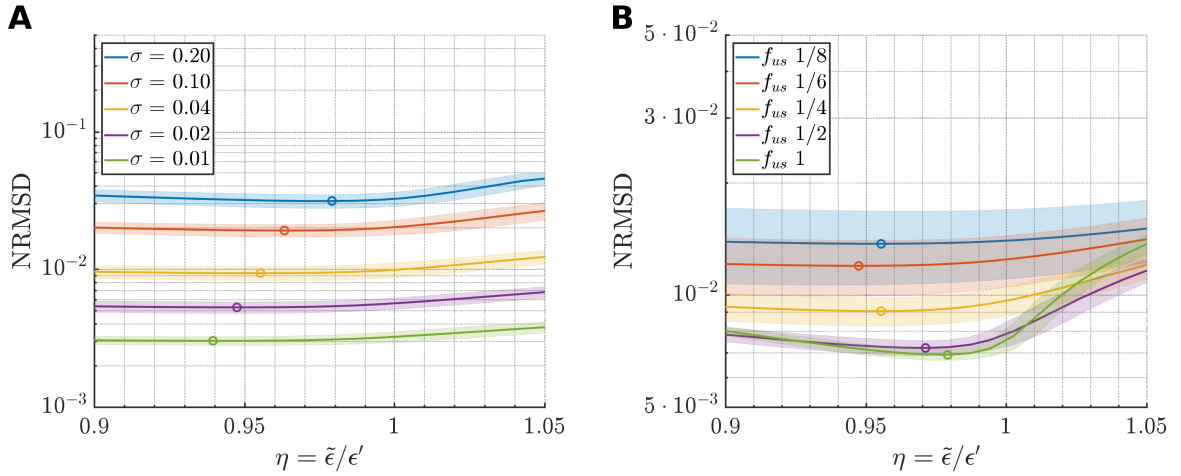


Fig. 24. **Noise level-based regularization strength adjustment:** Dependence of the image quality on η , the ratio between the demanded deviation of the reconstruction from the measured data $\tilde{\epsilon}$ and the expected deviation of the measured data from the true, but unknown image ϵ' . The NRMSD from the reference was computed for reconstructions of noisy data sampled from a digital phantom. The shaded area represents the standard deviation computed from 50 reconstructions with different noise realizations and undersampling masks. Circles mark the minima of the curves and the vertical axis is scaled logarithmically. Figure (A) shows the NRMSD for five different noise levels and $f_{us} = 1/4$. σ is the standard deviation of the complex Gaussian noise and the digital phantom was normed to one. In (B) the dependence is plotted for a fixed noise level ($\sigma = 0.04$) and five different undersampling factors between 1 and 1/8. The yellow lines in (A) and (B) are thus identical.

A. Appendix

A.1. Glossary

A.1.1. List of abbreviations

ADMM	Alternating Direction Method of Multipliers
BART	Berkeley Advanced Reconstruction Toolbox
CS	Compressed Sensing
EAE	Experimental Autoimmune Encephalomyelitis
FDR	False Discovery Rate
FOV	Field Of View
•	•
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
NRMSD	Normalized Root-Mean-Square Deviation
RMSD	Root-Mean-Square Deviation
SNR	Signal-to-Noise Ratio
TPR	True Positive Rate
•	•

A.1.2. List of symbols

The symbols are grouped by topic and given in the order of appearance. Only symbols which are repeatedly used or denote quantities central to NMR, MRI or CS are listed.

NMR:		B₀	Static magnetic field
μ	Magnetic moment	ΔE	Energy level separation
L	Angular momentum	T	Temperature
μ_p	Proton magnetic moment	T	Torque
g_p	Proton <i>g</i> -factor	B	Magnetic field
β_N	Nuclear magneton		
I	Spin angular momentum in units of ħ		
μ_N	Nuclear magnetic moment		
g_N	Nuclear <i>g</i> -factor		
I_N	Nuclear spin angular momentum in units of ħ		
γ	Gyromagnetic ratio		

A.2. Update equations for the alternating direction method of multipliers

A.2.1. \mathbb{CR} -calculus

To perform steps a) and b) we need to minimize \tilde{O}_{aniso} and \tilde{O} . These are real-valued functions with complex arguments. Formally speaking, they are functions of the type:

$$f : \mathbb{C}^n \mapsto \mathbb{R}, f(\mathbf{z}) = a, \quad (\text{A.1})$$

where $\mathbf{z} = \mathbf{x} + i\mathbf{y}$. Such functions are not analytic, i.e. they are in general not differentiable in the sense of standard complex analysis. To minimize the augmented objective functions we need to

rely on so-called \mathbb{CR} -calculus. In \mathbb{CR} -calculus complex numbers are treated as two-dimensional real vectors and functions of the type (A.1) are instead treated as functions of the form

$$f : \mathbb{R}^{2n} \mapsto \mathbb{R}, f(\mathbf{x}, \mathbf{y}) = a, \quad (\text{A.2})$$

which can be approached using the tools of real vector analysis⁹³. The possible minima of such a function are given by its stationary points, i.e. the points with vanishing gradient

$$\nabla f = \frac{\partial f}{\partial (\mathbf{x}^T, \mathbf{y}^T)^T} = 0. \quad (\text{A.3})$$

D. H. Brandwood⁹⁴ showed that this condition is equivalent to a vanishing complex (conjugate) gradient according to the following definition

$$\frac{\partial f}{\partial (\mathbf{x}^T, \mathbf{y}^T)^T} = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial \mathbf{z}} = \frac{1}{2} \left(\frac{\partial f}{\partial \mathbf{x}} - i \frac{\partial f}{\partial \mathbf{y}} \right) = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial \mathbf{z}^*} = \frac{1}{2} \left(\frac{\partial f}{\partial \mathbf{x}} + i \frac{\partial f}{\partial \mathbf{y}} \right) = 0. \quad (\text{A.4})$$

As pointed out by Brandwood, the complex conjugate gradient leads to slightly neater computations, so that it was adopted for all following optimizations. Some convenient results used in the following are⁹⁴:

$$\frac{\partial (\mathbf{a}^\dagger \mathbf{z})}{\partial \mathbf{z}^*} = 0, \quad \frac{\partial (\mathbf{z}^\dagger \mathbf{a})}{\partial \mathbf{z}^*} = \mathbf{a} \quad \text{and} \quad \frac{\partial (\mathbf{z}^\dagger R \mathbf{z})}{\partial \mathbf{z}^*} = R \mathbf{z}. \quad (\text{A.5})$$

Following Brandwood I adopted the convention to denote the gradient as a column vector. Furthermore it is:

$$\frac{\partial (\text{Re}(\mathbf{a}^\dagger \mathbf{z}))}{\partial \mathbf{z}^*} = \frac{1}{2} \frac{\partial}{\partial \mathbf{x}} (\text{Re}(\mathbf{a}^T) \mathbf{x}) + \frac{i}{2} \frac{\partial}{\partial \mathbf{y}} (\text{Im}(\mathbf{a}^T) \mathbf{y}) = \frac{1}{2} \mathbf{a}. \quad (\text{A.6})$$

In the following derivations we also assume that $d_i \neq 0$ for all elements of \mathbf{d} , which is appropriate for noisy data, so that the ℓ^1 norm is differentiable and

$$\left(\frac{\partial}{\partial \mathbf{d}^*} \|\mathbf{d}\|_1 \right)_i = \frac{\partial}{\partial d_i^*} \sqrt{d_i^* d_i} = \frac{1}{2} \frac{d_i}{|d_i|}. \quad (\text{A.7})$$

A.2.2. Update equations

We can now compute the stationary points of \tilde{O} with regard to \mathbf{r} respectively \mathbf{d} . It is

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= \frac{\partial}{\partial \mathbf{r}^*} \left(\|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \frac{\mu}{2} \|\mathbf{d} - \Psi \mathbf{r}\|_2^2 - \text{Re}(\mathbf{k}^\dagger \Psi \mathbf{r}) \right) \\ &= \frac{\partial}{\partial \mathbf{r}^*} \left((F_u \mathbf{r} - \mathbf{y})^\dagger (F_u \mathbf{r} - \mathbf{y}) + \frac{\mu}{2} (\mathbf{d} - \Psi \mathbf{r})^\dagger (\mathbf{d} - \Psi \mathbf{r}) \right) - \frac{1}{2} \Psi^\dagger \mathbf{k} \\ &= \frac{\partial}{\partial \mathbf{r}^*} \left(\mathbf{r}^\dagger F_u^\dagger F_u \mathbf{r} - 2 \text{Re}(\mathbf{y}^\dagger F_u \mathbf{r}) \right) + \frac{\mu}{2} \frac{\partial}{\partial \mathbf{r}^*} \left(\mathbf{r}^\dagger \Psi^\dagger \Psi \mathbf{r} - 2 \text{Re}(\mathbf{d}^\dagger \Psi \mathbf{r}) \right) - \frac{1}{2} \Psi^\dagger \mathbf{k} \\ &= F_u^\dagger F_u \mathbf{r} - F_u^\dagger \mathbf{y} + \frac{\mu}{2} (\Psi^\dagger \Psi \mathbf{r} - \Psi^\dagger \mathbf{d}) - \frac{1}{2} \Psi^\dagger \mathbf{k} \end{aligned} \quad (\text{A.8})$$

and thus

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= 0 \\ \Leftrightarrow 2F_u^\dagger F_u \mathbf{r} - 2F_u^\dagger \mathbf{y} + \mu \Psi^\dagger \Psi \mathbf{r} - \Psi^\dagger (\mu \mathbf{d} - \mathbf{k}) &= 0 \\ \Leftrightarrow 2F_u^\dagger F_u \mathbf{r} + \mu \Psi^\dagger \Psi \mathbf{r} &= \Psi^\dagger (\mu \mathbf{d} - \mathbf{k}) + 2F_u^\dagger \mathbf{y} \\ \Leftrightarrow \mathbf{r} &= \left(2F_u^\dagger F_u + \mu \Psi^\dagger \Psi \right)^{-1} \left(\Psi^\dagger (\mu \mathbf{d} - \mathbf{k}) + 2F_u^\dagger \mathbf{y} \right). \end{aligned} \quad (\text{A.9})$$

The gradient regarding \mathbf{d} separates: $\frac{\partial}{\partial \mathbf{d}^*} = \left(\frac{\partial}{\partial \mathbf{d}_\Phi^*} \right)$. Using $\mathbf{k} = \begin{pmatrix} \mathbf{k}_\Phi \\ \mathbf{k}_\nabla \end{pmatrix}$, it is

$$\begin{aligned} \frac{\partial}{\partial \mathbf{d}_\Phi^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= \frac{\partial}{\partial \mathbf{d}_\Phi^*} \left(\lambda \|\mathbf{d}_\Phi\|_1 + \frac{\mu}{2} \|\mathbf{d}_\Phi - w_\Phi \Phi \mathbf{r}\|_2^2 + \text{Re} \left(\mathbf{k}_\Phi^\dagger \mathbf{d}_\Phi \right) \right) \\ &= \frac{\lambda}{2} \left(\frac{(\mathbf{d}_\Phi)_1}{|(\mathbf{d}_\Phi)_1|}, \dots, \frac{(\mathbf{d}_\Phi)_{n_\Phi}}{|(\mathbf{d}_\Phi)_{n_\Phi}|} \right)^T + \frac{\mu}{2} (\mathbf{d}_\Phi - w_\Phi \Phi \mathbf{r}) + \frac{1}{2} \mathbf{k}_\Phi \end{aligned} \quad (\text{A.10})$$

and thus

$$\begin{aligned} \frac{\partial}{\partial \mathbf{d}_\Phi^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= 0 \\ \Leftrightarrow \left(\left(\frac{\lambda}{\mu} + |(\mathbf{d}_\Phi)_1| \right) \frac{(\mathbf{d}_\Phi)_1}{|(\mathbf{d}_\Phi)_1|}, \dots, \left(\frac{\lambda}{\mu} + |(\mathbf{d}_\Phi)_{n_\Phi}| \right) \frac{(\mathbf{d}_\Phi)_{n_\Phi}}{|(\mathbf{d}_\Phi)_{n_\Phi}|} \right)^T &= w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k}. \end{aligned} \quad (\text{A.11})$$

For a solution of this equation it must be for all elements of \mathbf{d}_Φ

$$(d_\Phi)_i / |(d_\Phi)_i| = \exp \left(i \arg \left(\left(w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k} \right)_i \right) \right) \quad \text{and} \quad |(d_\Phi)_i| = \left| \left(w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k} \right)_i \right| - \frac{\lambda}{\mu}. \quad (\text{A.12})$$

As $|(d_\Phi)_i| \geq 0$, $\left| \left(w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k} \right)_i \right| \geq 0$ and $\frac{\lambda}{\mu} > 0$, the second condition can not be fulfilled if $\left| \left(w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k} \right)_i \right| < \frac{\lambda}{\mu}$. In this case, we follow the sparsity enhancing approach outlined in e.g. Goldstein et al.⁵⁷, Goldstein and Osher⁶⁰, Tao et al.⁹⁵ and set $(d_\Phi)_i = 0$. Thus, the update equation for \mathbf{d}_Φ is

$$\mathbf{d}_\Phi = \text{shrink} \left(w_\Phi \Phi \mathbf{r} - \frac{1}{\mu} \mathbf{k}, \frac{\lambda}{\mu} \right), \quad (\text{A.13})$$

where the shrink function is defined as

$$\text{shrink} : \mathbb{C}^n \times \mathbb{R} \mapsto \mathbb{C}^n, (\text{shrink}(\mathbf{z}, \delta))_i = \frac{z_i}{|z_i|} \max(|z_i| - \delta, 0). \quad (\text{A.14})$$

Using $\mathbf{d}_\nabla = (\mathbf{d}_x^T, \mathbf{d}_y^T, \mathbf{d}_z^T)^T$, the second part of the derivative regarding \mathbf{d} is for $\beta = x, y, z$

$$\begin{aligned} \frac{\partial}{\partial \mathbf{d}_\beta^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= \frac{\partial}{\partial \mathbf{d}_\beta^*} \left(\lambda T_{\text{iso}}(\mathbf{d}_\nabla) + \frac{\mu}{2} \|\mathbf{d}_\beta - w_\nabla \nabla_\beta \mathbf{r}\|_2^2 + \text{Re} \left(\mathbf{k}_\beta^\dagger \mathbf{d}_\beta \right) \right) \\ &= \lambda \frac{\partial}{\partial \mathbf{d}_\beta^*} \left(\sum_{i=1}^n \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_i|^2} \right) + \frac{\mu}{2} (\mathbf{d}_\beta - w_\nabla \nabla_\beta \mathbf{r}) + \frac{1}{2} \mathbf{k}_\beta \\ &= \lambda \frac{\partial}{\partial \mathbf{d}_\beta^*} \left(\sum_{i=1}^n \sqrt{\sum_{\alpha=x,y,z} (\mathbf{d}_\alpha^*)_i (\mathbf{d}_\alpha)_i} \right) + \frac{\mu}{2} (\mathbf{d}_\beta - w_\nabla \nabla_\beta \mathbf{r}) + \frac{1}{2} \mathbf{k}_\beta \quad (\text{A.15}) \\ &= \frac{\lambda}{2} \begin{pmatrix} \frac{(\mathbf{d}_\beta)_1}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_1|^2}} \\ \vdots \\ \frac{(\mathbf{d}_\beta)_n}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_n|^2}} \end{pmatrix} + \frac{\mu}{2} (\mathbf{d}_\beta - w_\nabla \nabla_\beta \mathbf{r}) + \frac{1}{2} \mathbf{k}_\beta \end{aligned}$$

and thus

$$\begin{aligned} \frac{\partial}{\partial \mathbf{d}_\beta^*} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) &= 0 \\ \Leftrightarrow \begin{pmatrix} \left(\frac{\lambda}{\mu} + \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_1|^2} \right) \frac{(\mathbf{d}_\beta)_1}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_1|^2}} \\ \vdots \\ \left(\frac{\lambda}{\mu} + \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_n|^2} \right) \frac{(\mathbf{d}_\beta)_n}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_n|^2}} \end{pmatrix} &= w_T \nabla_\beta \mathbf{r} - \frac{1}{\mu} \mathbf{k}_\beta. \end{aligned} \quad (\text{A.16})$$

The update equations for \mathbf{d}_x , \mathbf{d}_y and \mathbf{d}_z are therefore coupled. Using the shorthand $\mathbf{s}_\beta = w_T \nabla_\beta \mathbf{r} - \frac{1}{\mu} \mathbf{k}_\beta$, we can write for the i -th entry of the three vectors

$$\begin{aligned} \left(\frac{\lambda}{\mu} + \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_i|^2} \right) \begin{pmatrix} (\mathbf{d}_x)_i \\ (\mathbf{d}_y)_i \\ (\mathbf{d}_z)_i \end{pmatrix} &= \frac{1}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_i|^2}} \begin{pmatrix} (\mathbf{d}_x)_i \\ (\mathbf{d}_y)_i \\ (\mathbf{d}_z)_i \end{pmatrix} \\ &= \frac{1}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{s}_\alpha)_i|^2}} \begin{pmatrix} (\mathbf{s}_x)_i \\ (\mathbf{s}_y)_i \\ (\mathbf{s}_z)_i \end{pmatrix}. \end{aligned} \quad (\text{A.17})$$

Having split the equation into a unit vector and a magnitude term on each side, we can equate

$$\begin{aligned} \frac{\lambda}{\mu} + \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_i|^2} &= \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{s}_\alpha)_i|^2} \\ \Leftrightarrow \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{d}_\alpha)_i|^2} &= \sqrt{\sum_{\alpha=x,y,z} |(\mathbf{s}_\alpha)_i|^2} - \frac{\lambda}{\mu} \end{aligned} \quad (\text{A.18})$$

and thus

$$\begin{pmatrix} (\mathbf{d}_x)_i \\ (\mathbf{d}_y)_i \\ (\mathbf{d}_z)_i \end{pmatrix} = \frac{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{s}_\alpha)_i|^2} - \frac{\lambda}{\mu}}{\sqrt{\sum_{\alpha=x,y,z} |(\mathbf{s}_\alpha)_i|^2}} \begin{pmatrix} (\mathbf{s}_x)_i \\ (\mathbf{s}_y)_i \\ (\mathbf{s}_z)_i \end{pmatrix}. \quad (\text{A.19})$$

A.2.3. Convolution solver

We consider the case of only images space and total variation regularization. In this scenario the combined transform matrix Ψ is

$$\Psi = \begin{pmatrix} w_\Phi I_n \\ w_T \nabla_x \\ w_T \nabla_y \\ w_T \nabla_z \end{pmatrix} \Rightarrow \Psi^\dagger \Psi = w_\Phi^2 I_n + w_T^2 (\Delta_x + \Delta_y + \Delta_z), \quad (\text{A.20})$$

where $\Delta_\alpha = \nabla_\alpha^\dagger \nabla_\alpha$ is the finite differences Laplace operator in direction α . It is e.g.

$$(\Delta_x \mathbf{r})_{i,j,k} = r(i+1, j, k) - 2r(i, j, k) + r(i-1, j, k). \quad (\text{A.21})$$

As this example already shows, the Laplace operator can be understood as a convolution if periodic boundary conditions are assumed⁹⁶.

The matrix encoding a one-dimensional convolution acting on a vector is circulant, i.e. it is completely described by one of its column (row) vectors and the next column (row) vector is just the preceding vector cyclically rotated one element down (to the right)⁹⁷. Circulant matrices can be diagonalized via the Fourier transform. For circulant matrix $C \in \mathbb{C}^{n \times n}$ it is

$$C = F^\dagger L F = F^\dagger \text{diag}(c_1, \dots, c_n) F. \quad (\text{A.22})$$

$c_i, i = 1, \dots, n$ are the eigenvalues of C , which can be computed as $\mathbf{c} = \sqrt{n}FC_{*,1}$, where $C_{*,1}$ is the first column of C and I assumed the scaling convention $F^{-1} = F^\dagger$. Similarly, the matrix encoding a two-dimensional convolution and acting on a vectorized image is block-circulant, meaning the matrix consists of blocks of circulant matrices which are themselves arranged in a circulant fashion. For a three-dimensional convolution and a vectorized three dimensional dataset the matrix becomes block-circulant with block-circulant blocks⁹⁸. In both cases the diagonalization of equation (A.22) holds if the Fourier transform operator F of the appropriate dimension is chosen⁹⁹. Using $F_u^\dagger F_u = F^\dagger R F$, where R is a diagonal matrix selecting only the rows of F sampled by F_u , and $\Delta_x + \Delta_y + \Delta_z = F^\dagger E_\Delta F$, where $E_\Delta = \text{diag}(\sqrt{n}F(\Delta_x + \Delta_y + \Delta_z)_{*,1})$, we can solve update step b) by

$$\begin{aligned} (2F_u^\dagger F_u + \mu\Psi^\dagger\Psi) \mathbf{r} &= \Psi^\dagger(\mu\mathbf{d} - \mathbf{k}) + 2F_u^\dagger \mathbf{y} \\ \Leftrightarrow F^\dagger(2R + w_\Phi^2 I_n + w_T^2 E_\Delta) F \mathbf{r} &= \Psi^\dagger(\mu\mathbf{d} - \mathbf{k}) + 2F_u^\dagger \mathbf{y} \\ \Leftrightarrow \mathbf{r} &= F^\dagger(2R + w_\Phi^2 I_n + w_T^2 E_\Delta)^{-1} F(\Psi^\dagger(\mu\mathbf{d} - \mathbf{k}) + 2F_u^\dagger \mathbf{y}). \end{aligned} \quad (\text{A.23})$$

A.2.4. Conjugate gradient approximation

Update step b) minimizes the augmented objective function $\tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k})$ regarding \mathbf{r} . This is done by performing a backtracking line search in a search direction \mathbf{m} on each step. \mathbf{m} is given by a combination of the gradient $\mathbf{g} = \frac{\partial}{\partial \mathbf{r}} \tilde{O}(\mathbf{r}, \mathbf{d}, \mathbf{k}) \Big|_{\mathbf{r}=\mathbf{r}^{(j)}}$ of the objective function at the current position and the search direction on the previous iteration. Implementations differ by the stopping criteria of the line search and the weighting of the two inputs for the new search direction. I follow the choices presented in Lustig et al.¹⁴ for a conjugate gradient solution of the complete compressed sensing reconstruction equation:

$$\text{Initialization} \quad \mathbf{r}^{(0)} = F_u^\dagger \mathbf{y}, \mathbf{m}^{(0)} = \mathbf{g}^{(0)}$$

Iterate the following steps until the maximum number of iterations is reached

$$\begin{aligned} \text{Linesearch} \quad & t = 1, \text{ while } \tilde{O}(\mathbf{r}^{(j)} + t\mathbf{m}^{(j)}) > \tilde{O}(\mathbf{r}^{(j)}) + \frac{t}{2} \text{Re}(\mathbf{g}^{(j)\dagger} \mathbf{m}^{(j)}) \text{ set } t = \frac{t}{2} \\ \text{Update step} \quad & \mathbf{r}^{(j+1)} = \mathbf{r}^{(j)} + t\mathbf{m}^{(j)} \\ \text{Compute gradient} \quad & \mathbf{g}^{(j+1)} = \frac{\partial}{\partial \mathbf{r}} \tilde{O}(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}^{(j+1)}} \\ \text{New search direction} \quad & \mathbf{m}^{(j+1)} = -\mathbf{g}^{(j+1)} + \frac{\|\mathbf{g}^{(j+1)}\|_2^2}{\|\mathbf{g}^{(j)}\|_2^2} \mathbf{m}^{(j)} \end{aligned}$$

The linesearch uses the Armijo-Goldstein¹⁰⁰ stopping condition and the new search direction is computed according the Fletcher and Reeves¹⁰¹. It is

$$\tilde{O}(\mathbf{r}) = \|F_u \mathbf{r} - \mathbf{y}\|_2^2 + \frac{\mu}{2} \|\mathbf{d} - \Psi \mathbf{r}\|_2^2 - \text{Re}(\mathbf{k}^\dagger \Psi \mathbf{r}) + \text{const.} \quad (\text{A.24})$$

and

$$\frac{\partial}{\partial \mathbf{r}^*} \tilde{O}(\mathbf{r}) = F_u^\dagger F_u \mathbf{r} - F_u^\dagger \mathbf{y} + \frac{\mu}{2} (\Psi^\dagger \Psi \mathbf{r} - \Psi^\dagger \mathbf{d}) - \frac{1}{2} \Psi^\dagger \mathbf{k}. \quad (\text{A.25})$$

A.3. Automatic selection of the regularization strength: the Illinois algorithm

The two initial bracketing points $\lambda_a^{(0)}, \lambda_b^{(0)}$ need to satisfy

$$f(\lambda_a^{(0)}) = \zeta_a^{(0)} < 0 < f(\lambda_b^{(0)}) = \zeta_b^{(0)}. \quad (\text{A.26})$$

First the reconstruction and deviation for a given starting value of the search $\lambda_s^{(0)}$ are computed. If $\zeta_s^{(0)} < 0$, the next value $\lambda_s^{(1)}$ is set to $2\lambda_s^{(0)}$. Else, $\lambda_s^{(1)}$ is set to $\frac{1}{2}\lambda_s^{(0)}$. This is repeated until a suitable interval is found.

In the simple false position method, the new candidate is computed based on a linear interpolation. For step k it is

$$\lambda_c^{(k)} = \lambda_b^{(k)} - \zeta_b^{(k)} \frac{\lambda_b^{(k)} - \lambda_a^{(k)}}{\zeta_b^{(k)} - \zeta_a^{(k)}}. \quad (\text{A.27})$$

If $\zeta_c^{(k)} < 0$, the left bracket point is updated ($\lambda_a^{(k+1)} = \lambda_c^{(k)}$) and the right bracket point is retained ($\lambda_b^{(k+1)} = \lambda_b^{(k)}$). Else, the right bracket point is updated ($\lambda_b^{(k+1)} = \lambda_c^{(k)}$) and the left bracket point is retained ($\lambda_a^{(k+1)} = \lambda_a^{(k)}$). This yields a new interval surrounding 0 and the step is repeated until a desired accuracy $|\zeta_c|/\epsilon' < \tau$ is reached.

The Illinois algorithm modifies this scheme to avoid cases where the interval is repeatedly shrunk on only one side. If e.g. $\zeta_c^{(k-2)} > 0$ and $\zeta_c^{(k-1)} > 0$, the next candidate value is set to

$$\lambda_c^{(k)} = \lambda_b^{(k)} - \zeta_b^{(k)} \frac{\lambda_b^{(k)} - \lambda_a^{(k)}}{\zeta_b^{(k)} - \frac{1}{2}\zeta_a^{(k)}} \quad (\text{A.28})$$

down-weighting $\zeta_a^{(k)}$ to shift $\lambda_c^{(k)}$ to the left side. For the opposite case, both occurrences of $\zeta_b^{(k)}$ in equation (A.27) are multiple with $1/2$.

A.4. Supplementary results on the EAE mouse model

This appendix gathers results on the EAE mouse model data that were not shown in section [4.3] due to space constraints. Figure 31 shows results regarding the detection performance of CS and conventional reconstructions in *in vivo* and *ex vivo* in 400 μm isotropic resolution. The corresponding results regarding the signal intensity bias are shown in figure 32. Figure 33 collects detection and quantification results for *ex vivo* data in the highest resolution (267 μm isotropic).

References

- ¹**G. Holland, P. A. Bottomley, and W. Hinshaw**, “19f magnetic resonance imaging”, *Journal of Magnetic Resonance* (1969) **28**, 133–136 (1977).
- ²**E. T. Ahrens, and U. Flögel**, *Fluorine magnetic resonance imaging* (Pan Stanford, 2017).
- ³**C. Faber, and F. Schmid**, “Pulse sequence considerations and schemes”, in *Fluorine magnetic resonance imaging*, edited by E. T. Ahrens, and U. Flögel, (Pan Stanford, 2017) Chap. 1, pp. 3–27.
- ⁴**H. Waiczies, S. Lepore, S. Drechsler, F. Qadri, B. Purfürst, K. Sydow, M. Dathe, A. Kühne, T. Lindel, W. Hoffmann, et al.**: “Visualizing brain inflammation with a shingled-leg radio-frequency head probe for 19f/1h mri”, *Scientific reports* **3** (2013).
- ⁵**U. Nöth, S. Morrissey, R. Deichmann, S. Jung, H. Adolf, A. Haase, and J. Lutz**, “Perfluoro-15-crown-5-ether labelled macrophages in adoptive transfer experimental allergic encephalomyelitis”, *Artificial Cells, Blood Substitutes, and Biotechnology* **25**, 243–254 (1997).
- ⁶**U. Flögel, Z. Ding, H. Hardung, S. Jander, G. Reichmann, C. Jacoby, R. Schubert, and J. Schrader**, “In vivo monitoring of inflammation after cardiac and cerebral ischemia by fluorine magnetic resonance imaging”, *Circulation* **118**, 140–148 (2008).
- ⁷**T. Niendorf, Y. Ji, and S. Waiczies**, “Fluorinated natural compounds and synthetic drugs”, in *Fluorine magnetic resonance imaging*, edited by F. U. Ahrens ET, (Pan Stanford, 2017) Chap. 11, pp. 311–342.
- ⁸**S. Waiczies, J. M. Millward, L. Starke, P. R. Delgado, T. Huelnhagen, C. Prinz, D. Marek, D. Wecker, R. Wissmann, S. P. Koch, et al.**: “Enhanced fluorine-19 mri sensitivity using a cryogenic radiofrequency probe: technical developments and ex vivo demonstration in a mouse model of neuroinflammation”, *Scientific Reports* **7**, 9808 (2017).
- ⁹**S. Waiczies, S. Lepore, K. Sydow, S. Drechsler, M.-C. Ku, C. Martin, D. Lorenz, I. Schütz, H. M. Reimann, B. Purfürst, et al.**: “Anchoring dipalmitoyl phosphoethanolamine to nanoparticles boosts cellular uptake and fluorine-19 magnetic resonance signal”, *Scientific reports* **5** (2015).
- ¹⁰**E. J. Candès, J. Romberg, and T. Tao**, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information”, *IEEE Transactions on information theory* **52**, 489–509 (2006).
- ¹¹**D. L. Donoho**, “Compressed sensing”, *IEEE Transactions on information theory* **52**, 1289–1306 (2006).
- ¹²**M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly**, “Compressed sensing mri”, *IEEE signal processing magazine* **25**, 72–82 (2008).
- ¹³**E. J. Candès**, “Lecture upon receiving the dannie-heinemann-preis: a short tour of compressive sensing”, (2013).
- ¹⁴**M. Lustig, D. Donoho, and J. M. Pauly**, “Sparse mri: the application of compressed sensing for rapid mr imaging”, *Magnetic resonance in medicine* **58**, 1182–1195 (2007).
- ¹⁵**S. H. GmbH**, *Compressed sensing: beyond speed*, (Jan. 2018) <https://www.healthcare.siemens.de/magnetic-resonance-imaging/clinical-specialities/compressed-sensing>.
- ¹⁶**O. N. Jaspán, R. Fleysher, and M. L. Lipton**, “Compressed sensing mri: a review of the clinical literature”, *The British journal of radiology* **88**, 20150487 (2015).
- ¹⁷**J. Zhong, P. H. Mills, T. K. Hitchens, and E. T. Ahrens**, “Accelerated fluorine-19 mri cell tracking using compressed sensing”, *Magnetic resonance in medicine* **69**, 1683–1690 (2013).

- ¹⁸**S. Liang, T. Dresselaers, K. Louchami, C. Zhu, Y. Liu, and U. Himmelreich**, “Comparison of different compressed sensing algorithms for low snr 19f mri applications—imaging of transplanted pancreatic islets and cells labeled with perfluorocarbons”, *NMR in Biomedicine* **30** (2017).
- ¹⁹**K. S. Krane, and D. Halliday**, *Introductory nuclear physics*, Vol. 465 (Wiley New York, 1988).
- ²⁰**E. Caurier, G. Martinez-Pinedo, F. Nowacki, A. Poves, and A. Zuker**, “The shell model as a unified view of nuclear structure”, *Reviews of Modern Physics* **77**, 427 (2005).
- ²¹**A. Carrington, and A. D. McLachlan**, “Introduction to magnetic resonance: with applications to chemistry and chemical physics”, (1967).
- ²²**J. D. Jackson**, *Classical electrodynamics* (John Wiley & Sons, 2007).
- ²³**J. J. Sakurai**, *Advanced quantum mechanics* (Pearson Education India, 1967).
- ²⁴**G. H. Fuller**, “Nuclear spins and moments”, *Journal of Physical and Chemical Reference Data* **5**, 835–1092 (1976).
- ²⁵**P. Imhof**, *Statistical physics* (lecture notes, Freie Universität Berlin, 2014/15).
- ²⁶**L. G. Hanson**, “Is quantum mechanics necessary for understanding magnetic resonance?”, *Concepts in Magnetic Resonance Part A* **32**, 329–340 (2008).
- ²⁷**R. P. Feynman, F. L. Vernon Jr, and R. W. Hellwarth**, “Geometrical representation of the schrödinger equation for solving maser problems”, *Journal of Applied Physics* **28**, 49–52 (1957).
- ²⁸**F. Bloch**, “Nuclear induction”, *Physical review* **70**, 460 (1946).
- ²⁹**M. Goldman**, “Formal theory of spin–lattice relaxation”, *Journal of Magnetic Resonance* **149**, 160–187 (2001).
- ³⁰**D. Higgins**, *Revise mri*, (Jan. 2018) <http://www.reviseMRI.com/>.
- ³¹**R. W. Brown, E. M. Haacke, Y.-C. N. Cheng, M. R. Thompson, and R. Venkatesan**, *Magnetic resonance imaging: physical principles and sequence design* (John Wiley & Sons, 2014).
- ³²**V. G. Kiselev**, “Fundamentals of diffusion mri physics”, *NMR in Biomedicine* **30** (2017).
- ³³**E. Petersen, I. Zimine, Y. L. Ho, and X. Golay**, “Non-invasive measurement of perfusion: a critical review of arterial spin labelling techniques”, *The British journal of radiology* **79**, 688–701 (2006).
- ³⁴**A. Haase, J. Frahm, D. Matthaei, W. Hanicke, and K.-D. Merboldt**, “Flash imaging. rapid nmr imaging using low flip-angle pulses”, *Journal of Magnetic Resonance* (1969) **67**, 258–266 (1986).
- ³⁵**D. W. McRobbie, E. A. Moore, and M. J. Graves**, *Mri from picture to proton* (Cambridge university press, 2017).
- ³⁶**S. Aja-Fernández, and A. Tristán-Vega**, “A review on statistical noise models for magnetic resonance imaging”, LPI, ETSI Telecomunicacion, Universidad de Valladolid, Spain, Tech. Rep (2013).
- ³⁷**H. Gudbjartsson, and S. Patz**, “The rician distribution of noisy mri data”, *Magnetic resonance in medicine* **34**, 910–914 (1995).
- ³⁸**M. D. Robson, J. C. Gore, and R. T. Constable**, “Measurement of the point spread function in mri using constant time imaging”, *Magnetic resonance in medicine* **38**, 733–740 (1997).
- ³⁹**M. N. Hood, V. B. Ho, J. G. Smirniotopoulos, and J. Szumowski**, “Chemical shift: the artifact and clinical tool revisited”, *Radiographics* **19**, 357–371 (1999).
- ⁴⁰**A. D. Elster**, *Questions and answers in mri: gibbs (truncation) artifact*, (Jan. 2018) <http://MRIquestions.com/gibbs-artifact.html>.

- ⁴¹**K. T. Block, M. Uecker, and J. Frahm**, “Undersampled radial mri with multiple coils. iterative image reconstruction using a total variation constraint”, *Magnetic resonance in medicine* **57**, 1086–1098 (2007).
- ⁴²**M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok**, “Introduction to compressed sensing”, preprint **93**, 2 (2011).
- ⁴³**L. Hochstrat**, *Signalverarbeitung und Wavelets: Bilddatenkompression mit JPEG und JPEG2000* (lecture notes, RWTH Aachen, 2009).
- ⁴⁴**G. Kutyniok**, “Theory and applications of compressed sensing”, *GAMM-Mitteilungen* **36**, 79–101 (2013).
- ⁴⁵**B. Roman, A. Bastounis, B. Adcock, and A. C. Hansen**, “On fundamentals of models and sampling in compressed sensing”, Preprint (2015).
- ⁴⁶**S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee**, “Compressive sensing: from theory to applications, a survey”, *Journal of Communications and networks* **15**, 443–456 (2013).
- ⁴⁷**S. Ji, Y. Xue, and L. Carin**, “Bayesian compressive sensing”, *IEEE Transactions on Signal Processing* **56**, 2346–2356 (2008).
- ⁴⁸**F. Knoll, C. Clason, C. Diwoky, and R. Stollberger**, “Adapted random sampling patterns for accelerated mri”, *Magnetic resonance materials in physics, biology and medicine* **24**, 43–50 (2011).
- ⁴⁹**S. Vasanawala, M. Murphy, M. T. Alley, P. Lai, K. Keutzer, J. M. Pauly, and M. Lustig**, “Practical parallel imaging compressed sensing mri: summary of two years of experience in accelerating body mri of pediatric patients”, in *Biomedical imaging: from nano to macro, 2011 IEEE International Symposium on* (IEEE, 2011), pp. 1039–1043.
- ⁵⁰**D.-d. Liu, D. Liang, X. Liu, and Y.-t. Zhang**, “Under-sampling trajectory design for compressed sensing mri”, in *Engineering in medicine and biology society (EMBC), 2012 Annual International Conference of the IEEE* (IEEE, 2012), pp. 73–76.
- ⁵¹**C. Gnahn, M. Bock, P. Bachert, W. Semmler, N. G. Behl, and A. M. Nagel**, “Iterative 3d projection reconstruction of 23na data with an 1h mri constraint”, *Magnetic resonance in medicine* **71**, 1720–1732 (2014).
- ⁵²**A. Beck, and M. Teboulle**, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”, *SIAM journal on imaging sciences* **2**, 183–202 (2009).
- ⁵³**M. Blasche, C. Forman, W. Strugnell, T. Okada, B. Wintersperger, K. Antwi, and P. Riffel**, “Compressed sensing supplement”, *Siemens Healthcare GmbH: MAGNETOM Flash - The Magazine of MRI* **66** (3/2016).
- ⁵⁴**U. Kamilov**, “Minimizing isotropic total variation without subiterations”, (2016).
- ⁵⁵**Wikipedia contributors**, *Augmented lagrangian method — Wikipedia, the free encyclopedia*, [Online; accessed 01-May-2018], 2018.
- ⁵⁶**S. Ramani, and J. A. Fessler**, “Parallel mr image reconstruction using augmented lagrangian methods”, *IEEE Transactions on Medical Imaging* **30**, 694–706 (2011).
- ⁵⁷**T. Goldstein, B. O’Donoghue, S. Setzer, and R. Baraniuk**, “Fast alternating direction optimization methods”, *SIAM Journal on Imaging Sciences* **7**, 1588–1623 (2014).
- ⁵⁸**V. Podlozhnyuk**, “Fft-based 2d convolution”, *NVIDIA white paper*, 32 (2007).
- ⁵⁹**M. Le, and J. A. Fessler**, “Efficient, convergent sense mri reconstruction for nonperiodic boundary conditions via tridiagonal solvers”, *IEEE transactions on computational imaging* **3**, 11–21 (2017).
- ⁶⁰**T. Goldstein, and S. Osher**, “The split bregman method for l1-regularized problems”, *SIAM journal on imaging sciences* **2**, 323–343 (2009).
- ⁶¹**M. Dowell, and P. Jarratt**, “A modified regula falsi method for computing the root of an equation”, *BIT Numerical Mathematics* **11**, 168–174 (1971).

- ⁶²**J. Ford**, “Improved algorithms of illinois-type for the numerical solution of nonlinear equations”, University of Essex, Department of Computer Science (1995).
- ⁶³**N. E. M. Association, et al.**: “Determination of signal-to-noise ratio (snr) in diagnostic magnetic resonance imaging”, NEMA Standards Publication MS 1-2001 (2001).
- ⁶⁴**F. Zijlstra, M. A. Viergever, and P. R. Seevinck**, “Evaluation of variable density and data-driven k-space undersampling for compressed sensing magnetic resonance imaging”, *Investigative radiology* **51**, 410–419 (2016).
- ⁶⁵**J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck**, “Maximum-likelihood estimation of rician distribution parameters”, *IEEE Transactions on Medical Imaging* **17**, 357–361 (1998).
- ⁶⁶**C. G. Koay, and P. J. Basser**, “Analytically exact correction scheme for signal extraction from noisy magnitude mr signals”, *Journal of magnetic resonance* **179**, 317–322 (2006).
- ⁶⁷**S. Aja-Fernández, C. Alberola-López, and C.-F. Westin**, “Noise and signal estimation in magnitude mri and rician distributed images: a lmmse approach”, *IEEE transactions on image processing* **17**, 1383–1398 (2008).
- ⁶⁸**R. M. Henkelman**, “Measurement of signal intensities in the presence of noise in mr images [med. phys. 12, 232 (1985)]”, *Medical physics* **13**, 544–544 (1986).
- ⁶⁹**R. Colotti, J. A. Bastiaansen, A. Wilson, U. Flögel, C. Gonzales, J. Schwitter, M. Stuber, and R. B. van Heeswijk**, “Characterization of perfluorocarbon relaxation times and their influence on the optimization of fluorine-19 mri at 3 tesla”, *Magnetic resonance in medicine* **77**, 2263–2271 (2017).
- ⁷⁰**Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli**, “Image quality assessment: from error visibility to structural similarity”, *IEEE transactions on image processing* **13**, 600–612 (2004).
- ⁷¹**M. Seeger, H. Nickisch, R. Pohmann, and B. Schölkopf**, “Optimization of k-space trajectories for compressed sensing by bayesian experimental design”, *Magnetic resonance in medicine* **63**, 116–126 (2010).
- ⁷²**A. Buades, B. Coll, and J.-M. Morel**, “Image denoising methods. a new nonlocal principle”, *SIAM review* **52**, 113–147 (2010).
- ⁷³**J. V. Manjon, P. Coupe, L. Marti-Bonmati, D. L. Collins, and M. Robles**, “Adaptive non-local means denoising of mr images with spatially varying noise levels”, *Journal of Magnetic Resonance Imaging* **31**, 192–203 (2010).
- ⁷⁴**C. Gnahn, and A. M. Nagel**, “Anatomically weighted second-order total variation reconstruction of 23 na mri using prior information from 1 h mri”, *NeuroImage* **105**, 452–461 (2015).
- ⁷⁵**R. Parikh, A. Mathai, S. Parikh, G. C. Sekhar, and R. Thomas**, “Understanding and using sensitivity, specificity and predictive values”, *Indian journal of ophthalmology* **56**, 45 (2008).
- ⁷⁶**A. Heckert, and J. Filliben**, *Dataplot reference manual; nist handbook 148*, 1995.
- ⁷⁷**S. R. Becker**, “Practical compressed sensing: modern data acquisition and signal processing”, PhD thesis (California Institute of Technology, 2011).
- ⁷⁸**S. Hu, M. Lustig, A. Balakrishnan, P. E. Larson, R. Bok, J. Kurhanewicz, S. J. Nelson, A. Goga, J. M. Pauly, and D. B. Vigneron**, “3d compressed sensing for highly accelerated hyperpolarized 13c mrsi with in vivo applications to transgenic mouse models of cancer”, *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **63**, 312–321 (2010).
- ⁷⁹**S. Azwai**, *Immunology lecture notes: antigens and antigenicity*, Feb. 2017.
- ⁸⁰**S. D. Miller, W. J. Karpus, and T. S. Davidson**, “Experimental autoimmune encephalomyelitis in the mouse”, *Current protocols in immunology* **88**, 15–1 (2010).

- ⁸¹ *Guidelines for the use of adjuvants in research*, 2016.
- ⁸² **C. S. Constantinescu, N. Farooqi, K. O'brien, and B. Gran**, "Experimental autoimmune encephalomyelitis (eae) as a model for multiple sclerosis (ms)", *British journal of pharmacology* **164**, 1079–1106 (2011).
- ⁸³ **H. H. Hofstetter, C. L. Shive, and T. G. Forsthuber**, "Pertussis toxin modulates the immune response to neuroantigens injected in incomplete freund's adjuvant: induction of th1 cells and experimental autoimmune encephalomyelitis in the presence of high frequencies of th2 cells", *The Journal of Immunology* **169**, 117–125 (2002).
- ⁸⁴ **S. Osher, F. Ruan, J. Xiong, Y. Yao, and W. Yin**, "Sparse recovery via differential inclusions", *Applied and Computational Harmonic Analysis* **41**, 436–469 (2016).
- ⁸⁵ **S. Geethanath, R. Reddy, A. S. Konar, S. Imam, R. Sundaresan, and R. Venkatesan**, "Compressed sensing mri: a review", *Critical ReviewsTM in Biomedical Engineering* **41** (2013).
- ⁸⁶ **D. Sodickson**, *Daniel Sodickson plenary talk: the rapid imaging renaissance*, SPIE. the international society for optics and photonics, conference on medical imaging, (2015) https://www.youtube.com/watch?v=oypRDAENn_E, (visited on 2018-08-20).
- ⁸⁷ **S. S. Vasanawala, M. T. Alley, B. A. Hargreaves, R. A. Barth, J. M. Pauly, and M. Lustig**, "Improved pediatric mr imaging with compressed sensing", *Radiology* **256**, 607–616 (2010).
- ⁸⁸ **A. Chambolle, V. Duval, G. Peyré, and C. Poon**, "Geometric properties of solutions to the total variation denoising problem", *Inverse Problems* **33**, 015002 (2016).
- ⁸⁹ **N. Bolo, Y. Hodé, and J.-P. Macher**, "Long-term sequestration of fluorinated compounds in tissues after fluvoxamine or fluoxetine treatment: a fluorine magnetic resonance spectroscopy study in vivo", *Magnetic Resonance Materials in Physics, Biology and Medicine* **16**, 268–276 (2004).
- ⁹⁰ **H. Pan, J. W. Myerson, L. Hu, J. N. Marsh, K. Hou, M. J. Scott, J. S. Allen, G. Hu, S. San Roman, G. M. Lanza, et al.**: "Programmable nanoparticle functionalization for in vivo targeting", *The FASEB Journal* **27**, 255–264 (2013).
- ⁹¹ **A. Keliris, I. Mamedov, G. E. Hagberg, N. K. Logothetis, K. Scheffler, and J. Engelmann**, "A smart 19f and 1h mri probe with self-immolative linker as a versatile tool for detection of enzymes", *Contrast Media & Molecular Imaging* **7**, 478–483 (2012).
- ⁹² **A. Bar-Shir, N. N. Yadav, A. A. Gilad, P. C. Van Zijl, M. T. McMahon, and J. W. Bulte**, "Single 19f probe for simultaneous detection of multiple metal ions using micest mri", *Journal of the American Chemical Society* **137**, 78–81 (2014).
- ⁹³ **K. Kreutz-Delgado**, "The complex gradient operator and the cr-calculus", arXiv preprint arXiv:0906.4835 (2009).
- ⁹⁴ **D. Brandwood**, "A complex gradient operator and its application in adaptive array theory", in *Iee proceedings f-communications, radar and signal processing*, Vol. 130, 1 (IET, 1983), pp. 11–16.
- ⁹⁵ **M. Tao, J. Yang, and B. He**, "Alternating direction algorithms for total variation deconvolution in image reconstruction", TR0918, Department of Mathematics, Nanjing University (2009).
- ⁹⁶ **Wikipedia contributors**, *Discrete laplace operator — Wikipedia, the free encyclopedia*, [Online; accessed 04-May-2018], 2018.
- ⁹⁷ **R. M. Gray, et al.**: "Toeplitz and circulant matrices: a review", *Foundations and Trends® in Communications and Information Theory* **2**, 155–239 (2006).
- ⁹⁸ **E. Sutton**, *Lecture notes in image processing & computer vision*, 2003.
- ⁹⁹ **M. Chen**, "On the solution of circulant linear systems", *SIAM Journal on Numerical Analysis* **24**, 668–683 (1987).

- ¹⁰⁰**L. Armijo**, “Minimization of functions having lipschitz continuous first partial derivatives”, Pacific Journal of mathematics **16**, 1–3 (1966).
- ¹⁰¹**R. Fletcher, and C. M. Reeves**, “Function minimization by conjugate gradients”, The computer journal **7**, 149–154 (1964).