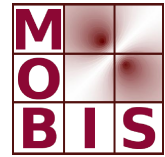




SpezialForschungsBereich F 32



Karl-Franzens Universität Graz  
Technische Universität Graz  
Medizinische Universität Graz



# Parallel Imaging with Nonlinear Reconstruction using Variational Penalties

Florian Knoll      Christian Clason  
Kristian Bredies      Martin Uecker  
Rudolf Stollberger

SFB-Report No. 2010-035

October 2010

A-8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the  
Austrian Science Fund (FWF)



SFB sponsors:

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



# PARALLEL IMAGING WITH NONLINEAR RECONSTRUCTION USING VARIATIONAL PENALTIES\*

Florian Knoll<sup>†</sup>    Christian Clason<sup>‡</sup>    Kristian Bredies<sup>‡</sup>  
Martin Uecker<sup>§</sup>    Rudolf Stollberger<sup>†</sup>

October 1, 2010

Nonlinear inversion was recently proposed for autocalibrated parallel imaging and shown to yield improved reconstruction quality. In addition, it has been shown that the aliasing arising from certain undersampled trajectories can be removed when using additional prior knowledge about the structure of the solution. Nonlinear inversion can be applied to arbitrary sampling trajectories, but the latter option was not yet exploited for this algorithm. In this work, it is demonstrated that nonlinear inversion can be extended with regularization terms that make use of such prior knowledge. The presented algorithms make use of the iteratively regularized Gauss-Newton method with additional variational constraints of total variation and total generalized variation type. Experimental results are presented for phantom and in-vivo measurements of undersampled radial and pseudorandom trajectories. The proposed approach yields results with reduced noise and undersampling artifacts in all cases when compared to conventional reconstruction with nonlinear inversion employing standard quadratic constraints.

Key words: Parallel Imaging, Radial Sampling, Pseudorandom Sampling, Nonlinear Inversion, Total Variation, Total Generalized Variation.

---

\*This work was supported by the Austrian Science Fund (FWF) under grant SFB F3209-18. (SFB “Mathematical Optimization and Applications in Biomedical Sciences”)

<sup>†</sup>Institute of Medical Engineering, Graz University of Technology, Kronesgasse 5, A-8010 Graz, Austria (florian.knoll@tugraz.at)

<sup>‡</sup>Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria

<sup>§</sup>Biomedizinische NMR Forschungs GmbH, Max-Planck-Institut für biophysikalische Chemie, Göttingen, Germany

## 1 INTRODUCTION

It was shown recently that nonlinear inversion can be applied successfully to image reconstruction of undersampled data from multiple coils, i.e., parallel imaging [1, 2, 3, 4]. The joint estimation of images and coil sensitivities, which can be achieved with an iteratively regularized Gauss-Newton (IRGN) method, leads to more accurate estimations of the coil sensitivities, and therefore yields results with improved image quality. It was also demonstrated that this method can be extended to non-Cartesian imaging [5, 6]. In particular, radial sampling has the advantage that the sampling pattern automatically leads to an oversampling of the central frequencies of  $k$ -space, which eliminates the need to acquire additional reference lines when performing auto-calibrated parallel imaging.

Another important characteristic of radial sampling is the fact that aliasing artifacts, which are introduced in the case of undersampling, have a very different appearance than the image content. For this reason, it is possible to remove these so-called streaking artifacts during reconstruction with the integration of specific a-priori information about the structure of the image. For radial MRI, this has been demonstrated in conventional (linear) parallel imaging with the penalization of the total variation (TV) of the image [7]. The general strategy, which relies on the sparsity of the solution in some transform domain, has been formulated for pseudorandom sampling patterns under the name of compressed sensing [8, 9, 10]; see also [11] for a related idea.

As demonstrated in this work, the formulation of autocalibrated parallel imaging as a nonlinear inversion problem yields a general framework that allows the integration of such additional a-priori information. This is illustrated with the integration of a TV penalty in the IRGN method. The method then combines both concepts, the integration of a-priori knowledge with appropriate regularization terms and the joint estimation of coil sensitivities and image content. In general, the integration of a TV term helps to suppress the noise amplification [12] that seriously limits parallel imaging with conventional methods. When applied to data sets obtained with radial or pseudorandom sampling patterns, this also leads to enhanced removal of artifacts. Additionally, it is shown that the flexibility provided by this framework easily allows replacing TV with another regularization functional in cases when the TV assumption of piecewise constant images is not reasonable; for instance, total generalized variation (TGV) of second order [13].

## 2 THEORY

Mathematically, parallel MR imaging can be formulated as a nonlinear inverse problem where the sampling operator  $\mathcal{F}_S$  (defined by the  $k$ -space trajectory, e.g., Fourier transform followed by multiplication with a binary mask for standard Cartesian subsampling) and the correspondingly acquired  $k$ -space data  $g = (g_1, \dots, g_N)^T$  from  $N$  receiver coils are given, and the spin density  $u$  and the unknown (or not perfectly known) set of coil sensitivities

$c = (c_1, \dots, c_N)^T$  have to be found such that

$$(2.1) \quad F(u, c) := (\mathcal{F}_S(u \cdot c_1), \dots, \mathcal{F}_S(u \cdot c_N))^T = g$$

holds. As was shown in [3, 5], this problem can be solved using the iteratively regularized Gauss-Newton (IRGN) method [14, 15, 16, 17], i.e., computing in each step  $k$  for given  $x^k := (u^k, c^k)$  the minimum  $\delta x := (\delta u, \delta c)$  of

$$(2.2) \quad \min_{\delta x} \frac{1}{2} \|F'(x^k) \delta x + F(x^k) - g\|^2 + \frac{\alpha_k}{2} \mathcal{W}(c^k + \delta c) + \beta_k \mathcal{R}(u^k + \delta u)$$

for given  $\alpha_k, \beta_k > 0$ , and then setting  $x^{k+1} := x^k + \delta x$ ,  $\alpha_{k+1} := q_\alpha \alpha_k$  and  $\beta_{k+1} := q_\beta \beta_k$  with  $0 < q_\alpha, q_\beta < 1$ . Here,  $F'(x^k)$  is the Fréchet derivative of  $F$  evaluated at  $x^k$ . The term  $\mathcal{W}(c) = \|Wc\|^2 = \|w \cdot \mathcal{F}c\|^2$  is a penalty on the high Fourier coefficients of the sensitivities and  $\mathcal{R}$  is a regularization term on the image. So far, the application of the IRGN method to parallel imaging has been formulated with a conventional  $L^2$  penalty [3, 5] (i.e.,  $\mathcal{R}(u) = \frac{1}{2} \|u\|^2$ ). As demonstrated in this work, the IRGN method can also be used with other regularization terms, which can be chosen dependent on the application. For example, the stability of the method with respect to noise can be improved: Since  $\alpha_k$  and  $\beta_k$  are decreasing during the iteration, the problem in (2.2) will become increasingly ill-conditioned. This leads to noise amplification, which can be counteracted by using a regularization term with stronger noise removal properties than the  $L^2$  penalty. A possible choice is the total variation (TV) of the image, i.e.,

$$(2.3) \quad \mathcal{R}(u) = \int |\nabla u|_2 dx,$$

where  $|\cdot|_2$  denotes the Euclidean norm in  $\mathbb{R}^2$ . To calculate the solution of (2.2), we make use of the dual characterization of the TV seminorm:

$$(2.4) \quad \beta \int |\nabla u|_2 dx = \sup_{p \in C_\beta} \langle u, -\operatorname{div} p \rangle,$$

where  $p = (p_1, p_2)^T$ ,  $\operatorname{div} p = \partial_x p_1 + \partial_y p_2$  with appropriate boundary conditions and

$$(2.5) \quad C_\beta = \{p \in L^2(\Omega; \mathbb{C}^2) : \operatorname{div} p \in L^2(\Omega; \mathbb{C}), |p(x)|_2 \leq \beta \text{ for almost all } x \in \Omega\}.$$

The problem in (2.2) then becomes a non-smooth convex-concave saddle-point problem

$$(2.6) \quad \min_{\delta u, \delta c} \max_{p \in C_{\beta_k}} \frac{1}{2} \|F'(x^k) \delta x + F(x^k) - g\|^2 + \frac{\alpha_k}{2} \mathcal{W}(c^k + \delta c) + \langle u^k + \delta u, -\operatorname{div} p \rangle,$$

which can be solved efficiently using a projected primal-dual extra-gradient method [18, 19], given as Algorithm 1. Since this requires only application of  $F'(x^k)$  and its adjoint  $F'(x^k)^*$ , the algorithm can be implemented efficiently on modern multi-core hardware such as graphics processing units (GPUs). Due to the bilinear structure of  $F$ , the action of  $F'(x^k)$  and  $F'(x^k)^*$

---

**Algorithm 1** Solution of TV sub-problem (2.6)

---

```

1: function TVSOLVE( $u, c, g, \alpha, \beta$ )
2:    $\delta u, \overline{\delta u}, \delta c, \overline{\delta c}, p \leftarrow 0$ , choose  $\sigma, \tau > 0$ 
3:   repeat
4:      $p \leftarrow \text{proj}_\beta(p + \tau \nabla(u + \overline{\delta u}))$ 
5:      $\delta u_{old} \leftarrow \delta u, \delta c_{old} \leftarrow \delta c$ 
6:      $\delta u \leftarrow \delta u - \sigma(\sum_{i=1}^N c_i^* \cdot \mathcal{F}_s^*(\mathcal{F}_s(u \cdot \overline{\delta c_i} + c_i \cdot \overline{\delta u}) + F(u, c) - g) - \text{div } p)$ 
7:      $\delta c \leftarrow \delta c - \sigma(u^* \cdot \mathcal{F}_s^*(\mathcal{F}_s(u \cdot \overline{\delta c_i} + c_i \cdot \overline{\delta u}) + F(u, c) - g) + \alpha W^* W(c_i + \overline{\delta c_i}))$ 
8:      $\overline{\delta u} \leftarrow 2\delta u - \delta u_{old}$ 
9:      $\overline{\delta c} \leftarrow 2\delta c - \delta c_{old}$ 
10:  until convergence
11:  return  $\delta u, \delta c$ 
12: end function

```

---

can be calculated explicitly in terms of the subsampling operator  $\mathcal{F}_s$  and its adjoint  $\mathcal{F}_s^*$ . The projection onto the convex set  $C_\beta$  can be calculated pointwise by setting for all  $x \in \Omega$

$$(2.7) \quad \text{proj}_\beta(q)(x) = \frac{q(x)}{\max(1, \beta^{-1}(|q(x)|_2)}.$$

Since TV regularization is known to introduce staircasing artifacts if the penalty parameter is large, we also investigated second order total generalized variation (TGV), which is a generalization of TV that avoids the staircasing in regions of smooth signal change [13, 20]. This amounts to setting

$$(2.8) \quad \beta \mathcal{R}(u) = \inf_v \beta \|\nabla u - v\| + 2\beta \|\mathcal{E}v\|,$$

where  $\mathcal{E}v = \frac{1}{2}(\nabla v + \nabla v^T)$  denotes the symmetrized gradient of the complex-valued vector field  $v \in \mathcal{C}^1(\Omega; \mathbb{C}^2)$ . We refer to [13, 20] for a detailed description of this functional and an explanation of its properties. Using again the dual representation of the norms, the Gauss-Newton step (2.2) is equivalent to the saddle point problem

$$(2.9) \quad \min_{\delta u, \delta c, v} \max_{p \in C_{\beta_k}, q \in C_{\beta_k}^2} \frac{1}{2} \|F'(x^k) \delta x + F(x^k) - g\|^2 + \frac{\alpha_k}{2} \mathcal{W}(c^k + \delta c) + \langle \nabla u^k + \delta u - v, p \rangle + \langle \mathcal{E}v, q \rangle,$$

where

$$(2.10) \quad C_\beta^2 = \left\{ q \in \mathcal{C}_c(\Omega, \mathcal{S}^{2 \times 2}) : (|q_{11}(x)|^2 + |q_{22}(x)|^2 + 2|q_{12}(x)|^2)^{1/2} \leq 2\beta \ \forall x \in \Omega \right\}$$

and  $\mathcal{S}^{2 \times 2}$  denotes the set of symmetric complex matrices. The corresponding extra-gradient method is given as Algorithm 2, where the projection  $\text{proj}_\beta^2$  onto  $C_\beta^2$  can again be computed pointwise.

---

**Algorithm 2** Solution of TGV sub-problem (2.9)

---

```
1: function TGV SOLVE( $u, c, g, \alpha, \beta$ )
2:    $\delta u, \overline{\delta u}, \delta c, \overline{\delta c}, v, \overline{v}, p, q \leftarrow 0$ , choose  $\sigma, \tau > 0$ 
3:   repeat
4:      $p \leftarrow \text{proj}_\beta(p + \tau(\nabla(u + \overline{\delta u}) - v))$ 
5:      $q \leftarrow \text{proj}_\beta^2(q + \tau(\mathcal{E}v))$ 
6:      $\delta u_{old} \leftarrow \delta u, \delta c_{old} \leftarrow \delta c, v_{old} \leftarrow v$ 
7:      $\delta u \leftarrow \delta u - \sigma(\sum_{i=1}^N c_i^* \cdot \mathcal{F}_s^*(\mathcal{F}_s(u \cdot \overline{\delta c_i} + c_i \cdot \overline{\delta u}) + F(u, c) - g) - \text{div } p)$ 
8:      $\delta c \leftarrow \delta c - \sigma(u^* \cdot \mathcal{F}_s^*(\mathcal{F}_s(u \cdot \overline{\delta c_i} + c_i \cdot \overline{\delta u}) + F(u, c) - g) + \alpha W^* W(c_i + \overline{\delta c_i}))$ 
9:      $\overline{v} \leftarrow v - \sigma(-p + \mathcal{E}^* q)$ 
10:     $\overline{\delta u} \leftarrow 2\delta u - \delta u_{old}$ 
11:     $\overline{\delta c} \leftarrow 2\delta c - \delta c_{old}$ 
12:     $\overline{v} \leftarrow 2v - v_{old}$ 
13:  until convergence
14:  return  $\delta u, \delta c$ 
15: end function
```

---

### 3 MATERIALS AND METHODS

#### 3.1 DATA ACQUISITION

Experiments were performed for 2D radial as well as 3D pseudorandom sampling patterns. All measurements were performed on a clinical 3T system (Siemens Magnetom TIM Trio, Erlangen, Germany). Written informed consent was obtained from all volunteers prior to the examination.

Pseudorandom sampling was tested with phantom experiments and in-vivo measurements of the brain of a healthy volunteer. A receive-only 12 channel head coil was used, and an SVD based coil compression [21] was applied to reduce the data to 8 virtual channels for the phantom measurements and 9 in the case of the in-vivo experiments. Measurements were performed with a 3D FLASH sequence with the following sequence parameters: TR=20ms, TE=5ms, flip angle 18°, matrix size 256 × 256 × 256. The pulse sequence was modified to include a binary 2D mask defining the subsampling of both phase-encoding directions. Pseudorandom patterns were generated with an adaptive procedure to account for the energy distribution in  $k$ -space [22]. For the phantom experiments, a resolution of 1mm × 1mm × 5mm was used. The in-vivo measurements were obtained with an isotropic spatial resolution of 1mm. Raw data was exported from the scanner, a 1D Fourier transform was performed along the readout direction, and partitions orthogonal to this axis were reconstructed.

Radial sampling experiments were performed with an rf-spoiled radial FLASH sequence with sequence parameters TR=2.0ms, TE=1.3ms, and a flip angle of 8°. Images of a water phantom and of the heart of a healthy volunteer were made. The acquisition of the in-vivo data was performed without cardiac gating and during free breathing [23, 6, 24] using a 32-channel

body array coil. After the acquisition the data was compressed to 12 virtual channels for the in-vivo experiments and 8 virtual channels for the phantom experiments. An in-plane resolution of  $2\text{mm} \times 2\text{mm}$  and a slice thickness of  $8\text{mm}$  was used in combination with  $128 \times 128$  image matrices. Due to the two-fold oversampling, this resulted in 256 sample points for each radial spoke.

### 3.2 NONLINEAR RECONSTRUCTION

All reconstructions were performed offline using a Matlab (R2010a, The MathWorks, Natick, MA, USA) implementation of the described nonlinear inversion method. For the reconstruction of radial data sets, Fessler and Sutton's NUFFT [25] code was used. To facilitate comparison, the solution of (2.2) with  $\mathcal{R}(v) = \frac{1}{2}\|v\|^2$  was computed using the same extra-gradient scheme, which can be obtained from Algorithm 1 by removing step 4 and replacing the term  $-\text{div } p$  with  $+\beta u + \overline{\delta u}$  in step 6. In the following, we will refer to the Gauss-Newton reconstruction using an  $L^2$ -penalty simply as IRGN, while the reconstruction using TV and TGV penalties will be denoted by IRGN-TV and IRGN-TGV, respectively.

### 3.3 PARAMETER CHOICE

The parameters in Algorithm 1 were chosen according to the convergence theory for the projected extra-gradient scheme. The step lengths  $\sigma$  and  $\tau$  were chosen such that  $\sigma\tau L^2 < 1$  holds, where  $L$  is the Lipschitz constant of the gradient of the functional to be minimized. This constant depends on the subsampling strategy and the iterates  $u^k, c^k$ , but can be estimated using a few iterations of the power method to approximately compute the norms of the partial Fréchet derivatives of the linearized operator  $F'(x)$ . As the norm of the finite difference approximation of the divergence and gradient operators with mesh size 1 is  $\sqrt{8}$ , we set  $\tau = \sigma = 1/\sqrt{8 + 2 \max(|\tilde{L}_u|, |\tilde{L}_c|)}$  in Algorithm 1, where  $\tilde{L}_u, \tilde{L}_c$  are the estimates from the power method. The step lengths in Algorithm 2 were set to  $\tau = \sigma = 1/\sqrt{12 + 2 \max(|\tilde{L}_u|, |\tilde{L}_c|)}$  based on the norm of the linear operator involving the symmetrized derivative. The iteration was terminated after a fixed number of iterations, since the efficiency estimate for the extra-gradient method gives an upper bound on the required number of iterations to achieve a given accuracy. Since a high accuracy is not necessary during the initial Gauss-Newton iterations with large penalties, we started with  $N_0 = 20$  iterations and set  $N_k = 2N_{k+1}$ . These choices were stable and yielded good results for all data sets.

The parameters in the Gauss-Newton iteration were chosen according to a quasi-optimality criterion. The initial penalties  $\alpha_0, \beta_0$  were chosen such that the norm of the residual  $\|F(u^1, c^1) - g\|$  after the first iteration was roughly  $3/4$  of the initial residual, and the reduction factors  $q_\alpha, q_\beta$  were set such that each further iteration roughly reduced the residual by a factor of  $1/2$ . The iteration was terminated once the achieved reduction factor fell below  $3/4$ . This lead to the choice  $\alpha_0 = 1, \beta_0 = 2, q_\alpha = q_\beta = 1/10$ , and 5 Gauss-Newton iterations for the radial data set. For the pseudorandom data set,  $\beta_0 = 1, q_\beta = 1/5$  and 6 Gauss-Newton iterations were used.

Since the TV regularization parameter is continually decreased during the Gauss-Newton



iteration, the final reconstruction will typically not show strong signs of TV-filtering such as a cartoon-like appearance. A more pronounced TV-effect can be achieved if the decrease of the regularization parameter is stopped at the desired level. To illustrate this, we will also show reconstructions where we have set  $\beta_{k+1} = \max(5 \cdot 10^{-3}, q_\beta \beta_k)$  for  $L^2$ , TV and TGV regularization (with otherwise unchanged parameters).

## 4 RESULTS

Figure 1 shows a partition of a water phantom reconstructed with IRGN and IRGN-TV from pseudorandomly subsampled 3D data using acceleration factors  $R = 4$  and  $R = 10$ . The reduced noise amplification and artifact removal characteristics of IRGN-TV are clearly visible for both acceleration factors. In the case of moderate acceleration with  $R = 4$ , the final TV regularization parameter  $\beta_{min}$  was set to zero. For  $R = 10$ ,  $\beta_{min} = 5 \cdot 10^{-3}$  was used to achieve a stronger TV regularization and a better removal of artifacts. It must be noted that the phantom only consists of regions that are piecewise constant, and therefore the underlying assumption of the TV penalty is fulfilled. However, in the case of  $R = 10$  with increased TV regularization, staircasing artifacts can be observed in regions where modulations caused by the inhomogeneity of the coil sensitivities affect the reconstructed image.

The above findings are confirmed for in-vivo conditions. A partition of a human brain reconstructed from pseudorandomly subsampled measurements is displayed in Fig. 2. In this example, using a moderate acceleration of  $R = 4$ ,  $\beta_{min}$  was again set to zero. Similar to the experiments with the phantom, a pronounced reduction of the noise can be observed in the image reconstructed with IRGN-TV as compared to the conventional IRGN reconstruction. As can be seen best in the enlarged image regions, image details are preserved well in the IRGN-TV reconstruction.

Figure 3 shows two slices of a water phantom acquired with subsampled radial measurements and reconstructed with IRGN and IRGN-TV. Here, 25 spokes were acquired to reconstruct a  $128 \times 128$  matrix, corresponding to an undersampling factor of approximately 8 in comparison to a fully sampled radial data set ( $128 \cdot \frac{\pi}{2} \approx 201$  spokes). Because of the high undersampling, an increased value of  $\beta_{min} = 5 \cdot 10^{-3}$  was used in this example. Both slices show reduced noise and streaking artifacts when IRGN-TV is used. However, one of the images (top row in Fig. 3) again exhibits staircasing artifacts for the IRGN-TV solution.

Reconstructions of real-time images of the beating heart are displayed in Fig. 4. Results for 25 ( $R \approx 8.0$ ), 21 ( $R \approx 9.6$ ) and 19 ( $R \approx 10.6$ ) acquired spokes are shown, corresponding to image acquisition times of 50ms, 41ms and 38ms. For all reconstructions,  $\beta_{min} = 5 \cdot 10^{-3}$  was used. The image reconstructed from the 25 spokes data set does not show streaking artifacts for both IRGN and IRGN-TV. However, noise amplification is much stronger in case of IRGN. In contrast, reconstructions from 21 and 19 spokes show residual streaking artifacts due to increased subsampling, which are again reduced in the images reconstructed with IRGN-TV.

Finally, the effect of TGV regularization is demonstrated in Fig. 5. It shows highlighted

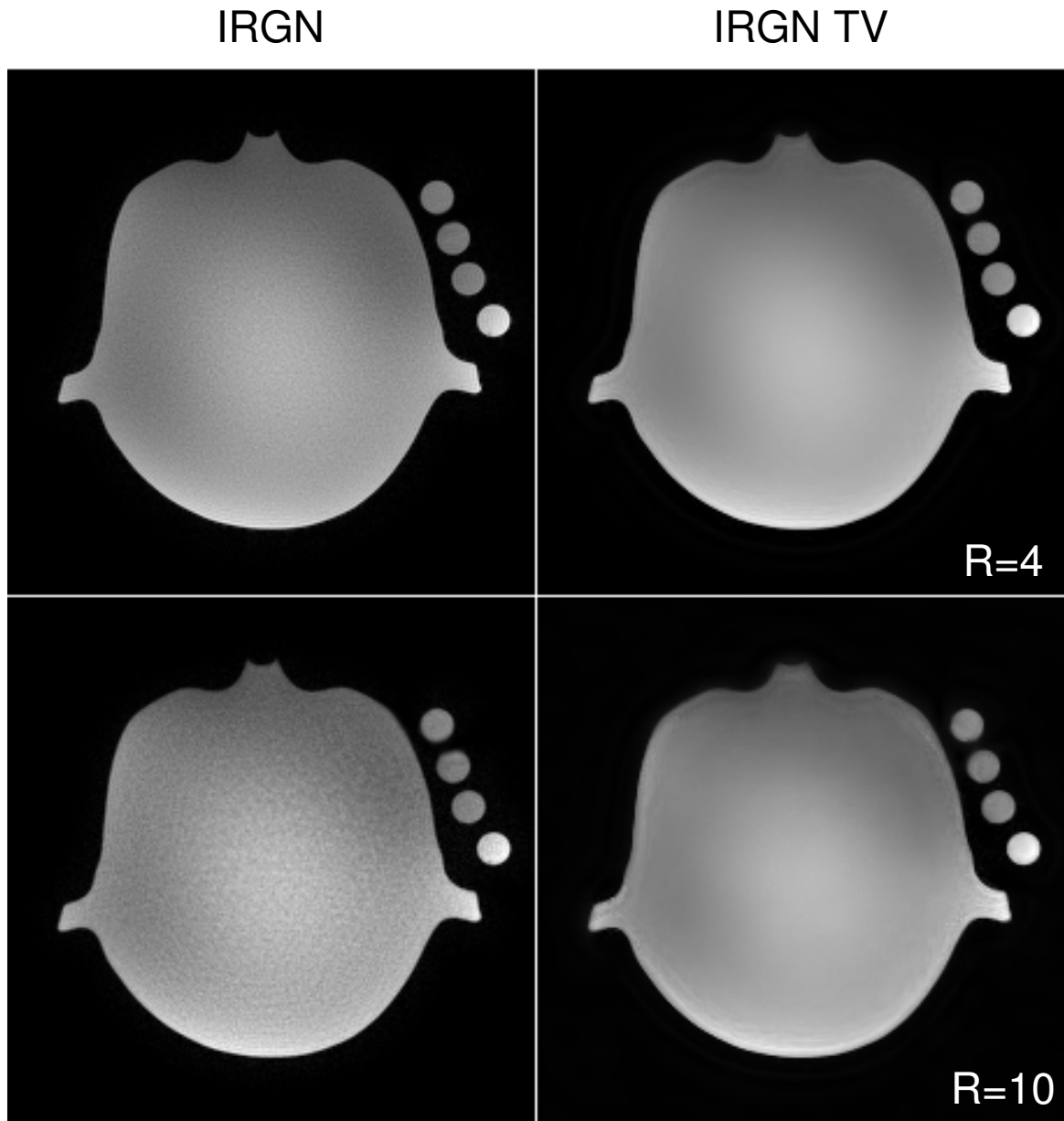


Figure 1: Reconstructions of pseudorandom sampling, phantom data set. IRGN (left), IRGN-TV (right). Acceleration factors  $R = 4$  (top,  $\beta_{min} = 0$ ) and  $R = 10$  (bottom,  $\beta_{min} = 5 \cdot 10^{-3}$ ).

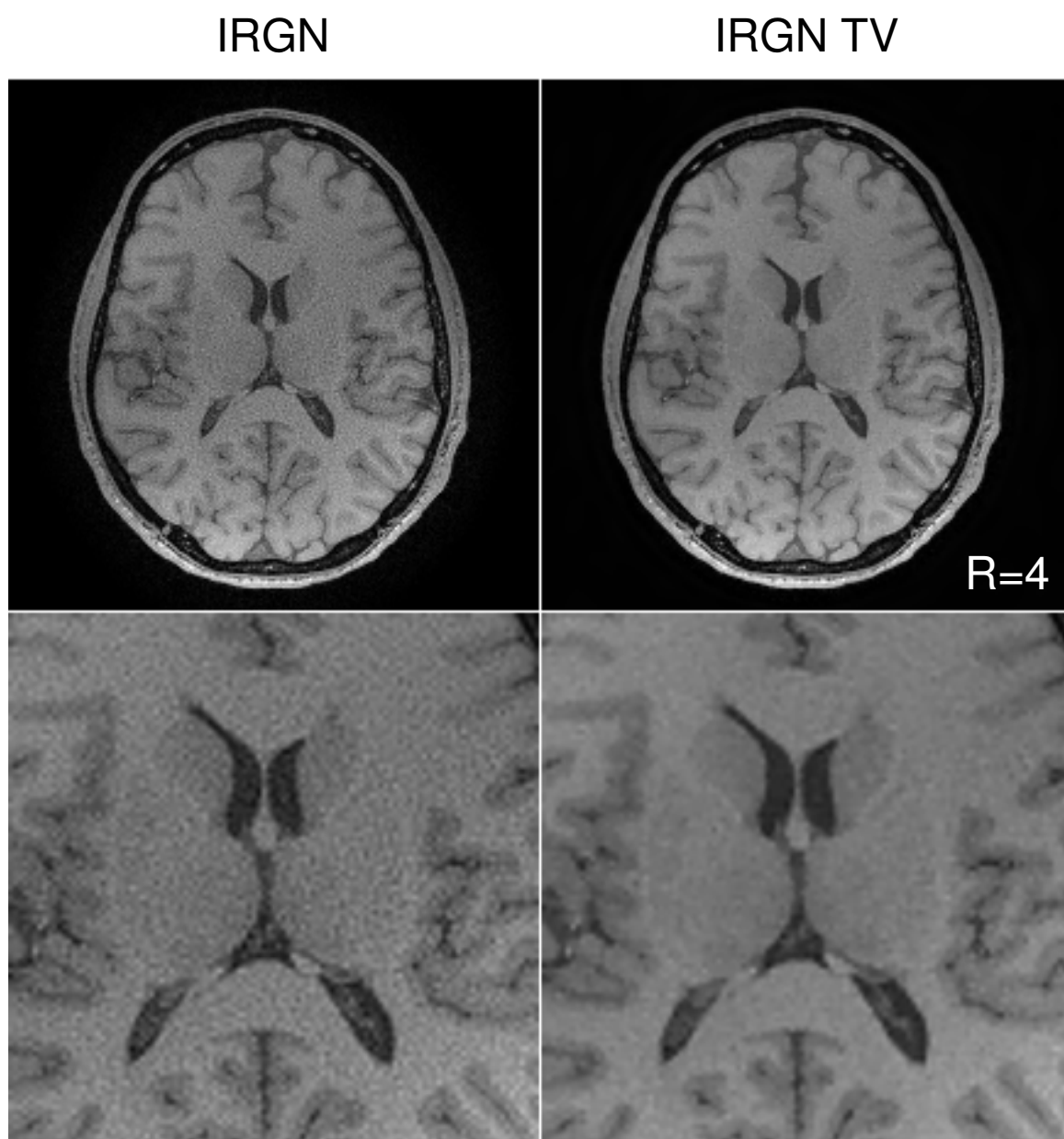


Figure 2: Reconstructions of pseudorandom sampling, brain data set. IRGN (left) and IRGN-TV (right). Acceleration factor  $R = 4$ ,  $\beta_{min} = 0$ .

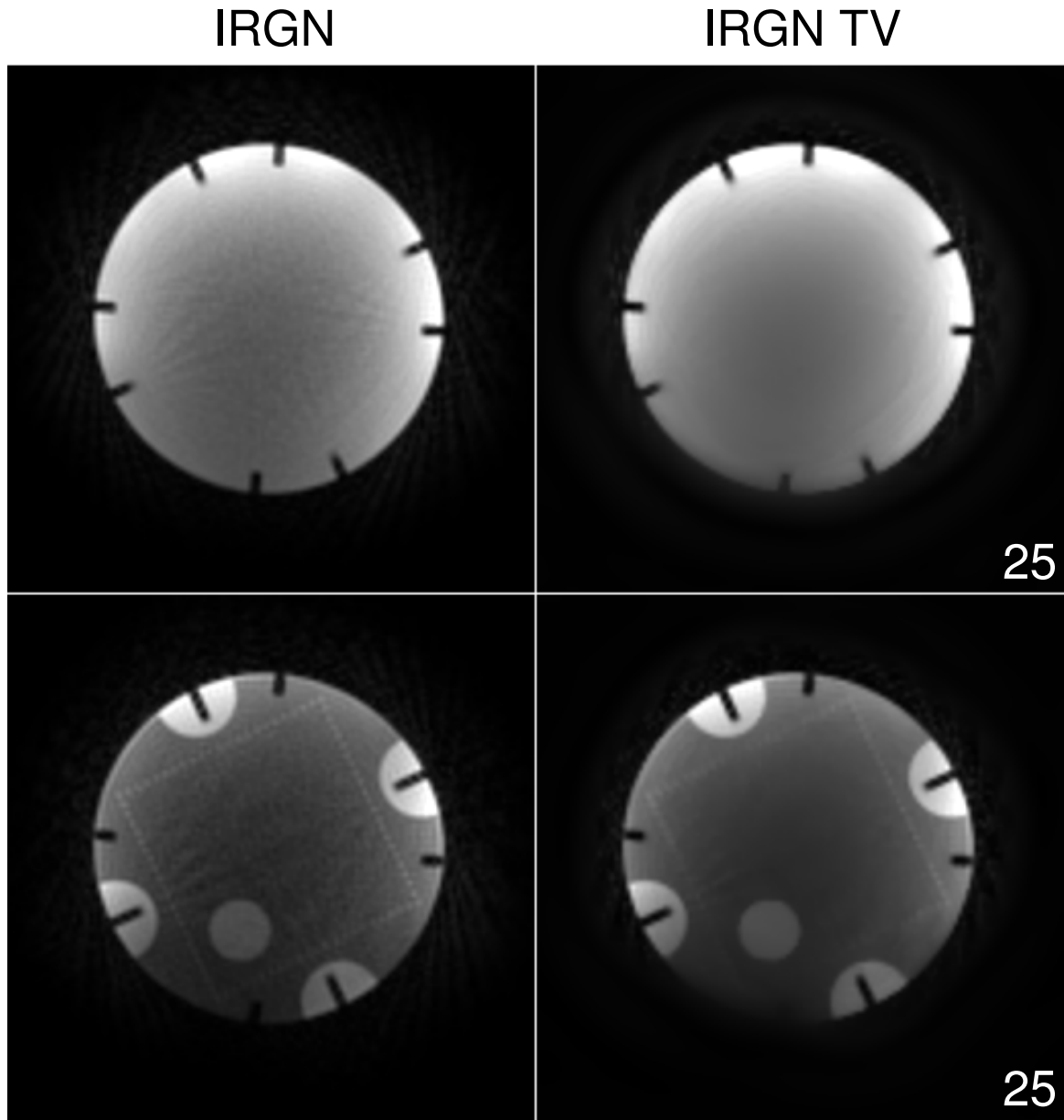


Figure 3: Reconstructions of radial sampling, phantom data, two different slices. IRGN (left) and IRGN-TV (right). 25 spokes were used for the reconstruction of a  $128 \times 128$  image,  $\beta_{min} = 5 \cdot 10^{-3}$ .

IRGN

IRGN TV

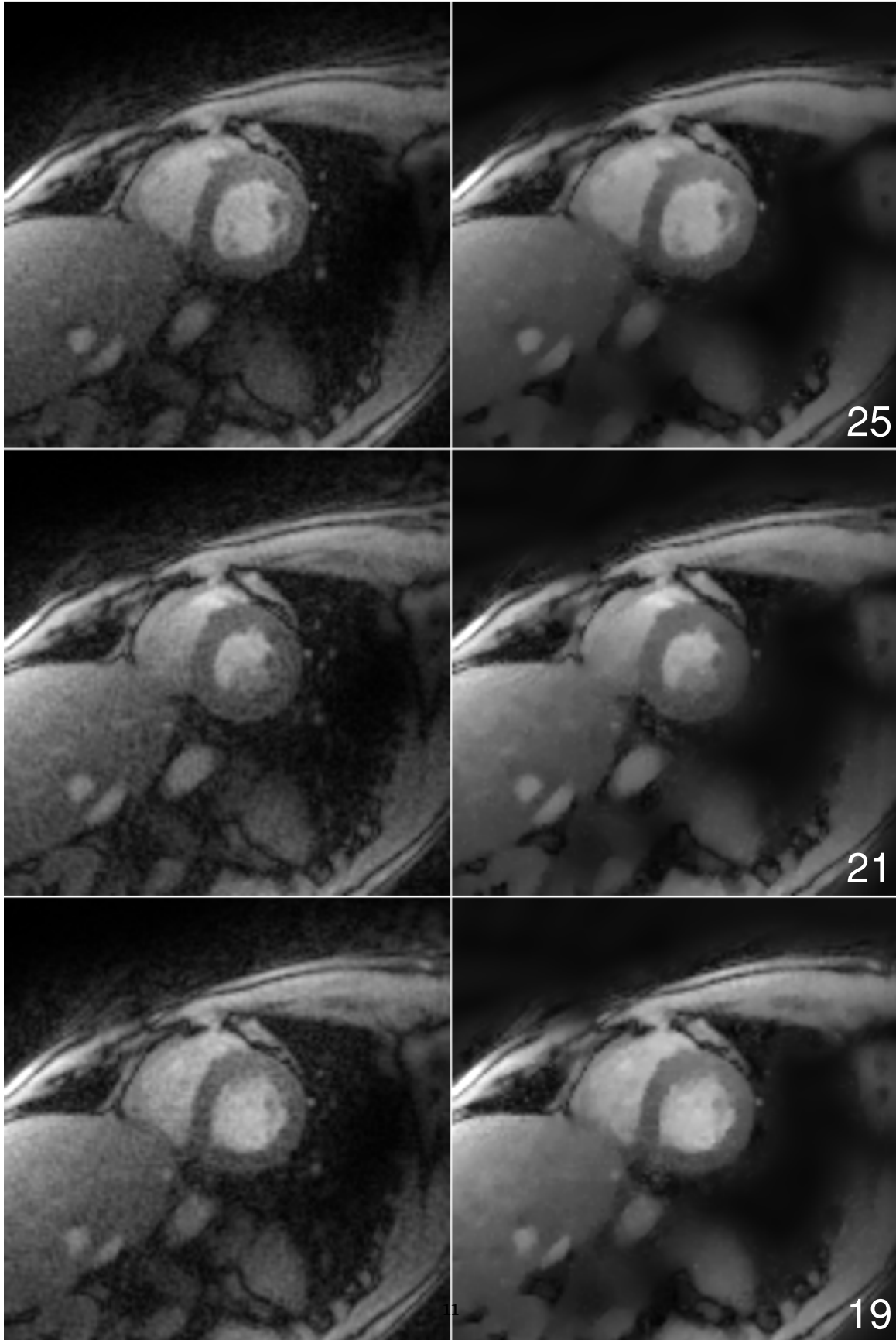


Figure 4: IRGN (left) and IRGN-TV (right) reconstructions of images of a human heart. 25 (top), 21 (middle), and 19 (bottom) spokes were used for the reconstruction of a  $128 \times 128$  image,  $\beta_{min} = 5 \cdot 10^{-3}$ .

regions of both phantom images affected by staircasing artifacts (pseudorandom and radial sampling, see Figs. 1 and 3) using TV as well as TGV regularization, both with  $\beta_{min} = 5 \cdot 10^{-3}$ . The staircasing artifact is completely removed in the reconstruction with the IRGN-TGV method.

## 5 DISCUSSION

The results from this work demonstrate that pronounced improvements in reconstruction quality of parallel imaging with nonlinear inversion can be achieved with TV and TGV based regularization instead of a simple  $L^2$  penalty. If moderate acceleration is used (Figs. 1, case of  $R = 4$ , and 2), TV serves as a stabilization term against noise amplification, which, otherwise, limits the practical use of parallel imaging to low acceleration factors. In cases where acceleration is pushed to its limits (Figs. 1, case of  $R = 10$ , 3 and 4), TV also leads to an additional removal of undersampling artifacts when combined with trajectories that produce incoherent aliasing. However, it must be noted that in this case, small image features with low contrast may also be removed during the reconstruction. This effect can be observed for some smaller vessels in Fig. 4.

Our Matlab implementations reconstructed a single slice in a few minutes, where IRGN-TV took roughly 10% more time than IRGN (and similarly, IRGN-TGV was about 10% slower than IRGN-TV). It is possible to exploit the differentiability of the data consistency term to apply more efficient minimization algorithms such as the method of conjugate gradients in the case of IRGN or order optimal convex minimization methods such as those investigated in [19] for IRGN-TV and IRGN-TGV. Because this work focused on the effect of the different regularization techniques on image quality, the same primal-dual extra-gradient method was used as inner algorithm in all cases to allow a direct comparison for identical parameter choices. In this context, it should be noted that the parameters for the iteratively regularized Gauss-Newton method were independent of the chosen regularization term, and only depended on the trajectory type. Similarly, since the norm of the forward operator is estimated in the algorithm, the fixed parameters for the primal-dual extra-gradient methods were independent of the data set.

In this work, the flexibility to include different regularization terms was demonstrated on 2D examples, where each image (slice, i.e., 2D partition of a 3D data set, or frame of a time series) had been reconstructed individually. While computationally more demanding, the extension of the penalties into a third space or time dimension should further improve the image quality. For example, earlier work has shown that residual streaking artifacts in radial imaging can be removed with the use of a median filter in the time dimension when using an interleaved k-space sampling scheme [6]. Since the median filter can be interpreted as solving an  $L^1$  minimization problem, it is expected that the inclusion of a corresponding penalty – either in the form of an  $L^1$  penalty on the difference between the current and previous slice or frame, or of a higher-dimensional T(G)V penalty on the full data set – will yield even better results. The resulting convex minimization problems can be solved using the same

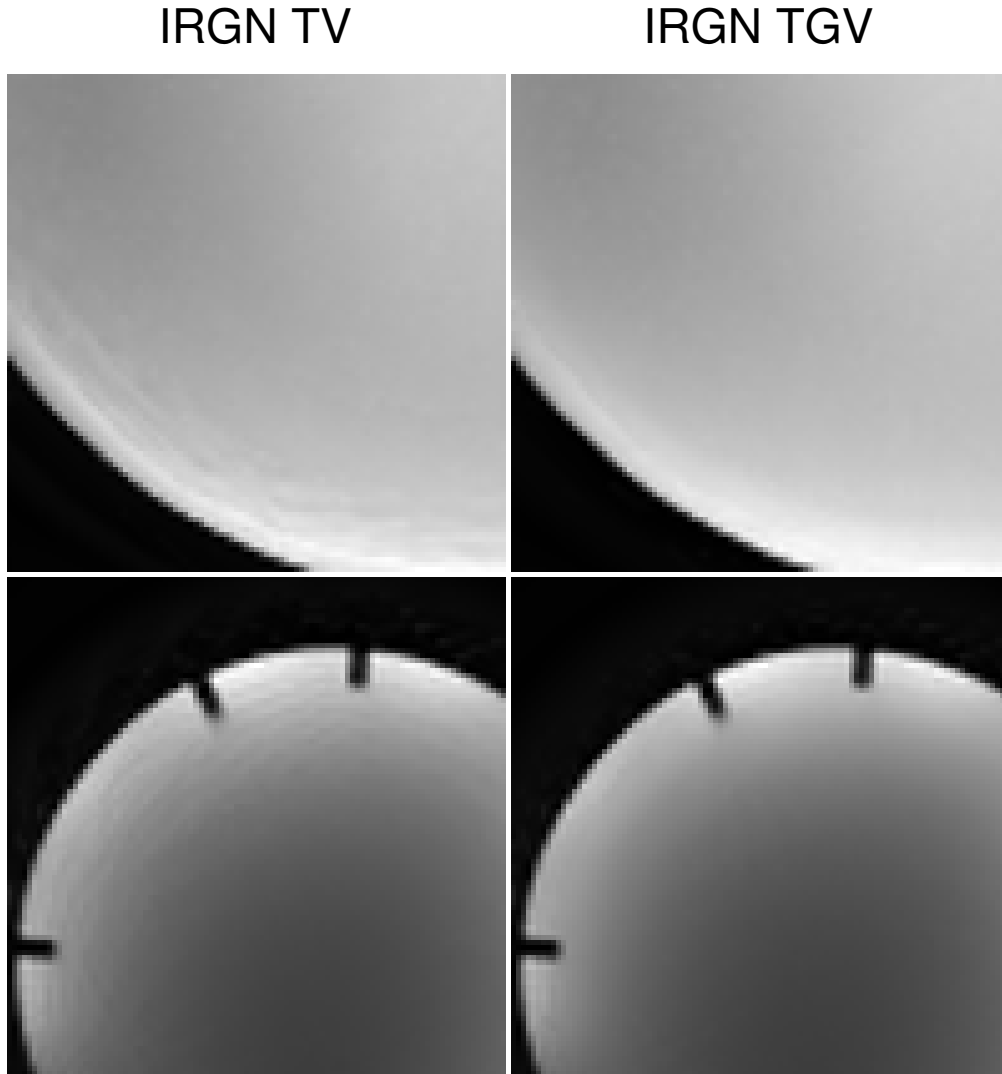


Figure 5: Comparison of conventional TV (left) and TGV (right) regularization for phantom experiments where modulations from the coil sensitivities lead to pronounced staircasing artifacts when TV is applied. Left: magnified regions from Figs. 1 (R=10, top) and 3 (bottom); right: IRGN-TGV reconstruction of the same data set (with  $\beta_{min} = 5 \cdot 10^{-3}$ ).

primal-dual extra-gradient method as employed in this work.

Another possible extension of this work is the integration of additional information about the physical signal model into the functional in (2.2). An important application is mapping of relaxation parameters in multi echo sequences [26, 27]. Here, the framework of nonlinear inverse problems makes it straightforward to perform parameter identification during the reconstruction. Since the parameters are thus estimated directly from the raw data, instead of from reconstructed images, better quantification is possible.

## 6 CONCLUSIONS

This work describes an approach to include additional variational penalties in parallel imaging with nonlinear inversion. The presented algorithms combine the advantages of nonlinear inversion, i.e., improved image quality through a better estimation of the coil sensitivities, with the advantageous properties of TV-based regularization terms. In addition to reducing the noise, the regularization is able to remove undersampling artifacts when combined with sampling strategies that produce incoherent aliasing, such as radial and pseudorandom sampling.

## REFERENCES

- [1] Bauer F, Kannengiesser S. An alternative approach to the image reconstruction for parallel data acquisition in MRI. *Math Meth Appl Sci* 2007;30(12):1437–1451.
- [2] Ying L, Sheng J. Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). *Magn Reson Med* 2007;57(6):1196–1202.
- [3] Uecker M, Hohage T, Block KT, Frahm J. Image reconstruction by regularized nonlinear inversion–joint estimation of coil sensitivities and image content. *Magn Reson Med* 2008;60(3):674–682.
- [4] Uecker M, Karaus A, Frahm J. Inverse reconstruction method for segmented multishot diffusion-weighted MRI with multiple coils. *Magn Reson Med* 2009;62(5):1342–1348.
- [5] Knoll F, Clason C, Uecker M, Stollberger R. Improved reconstruction in non-cartesian parallel imaging by regularized nonlinear inversion. In *Proceedings of the 17th Scientific Meeting and Exhibition of ISMRM, Honolulu, HI. 2009*; .
- [6] Uecker M, Zhang S, Frahm J. Nonlinear inverse reconstruction for real-time MRI of the human heart using undersampled radial FLASH. *Magn Reson Med* 2010;63(6):1456–1462.



- [7] Block KT, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med* 2007; 57(6):1086–1098.
- [8] Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 2006;52(2):489–509.
- [9] Donoho DL. Compressed sensing. *IEEE Transactions on Information Theory* 2006; 52(4):1289–1306.
- [10] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* 2007;58(6):1182–1195.
- [11] Xiang QS. Accelerating MRI by skipped phase encoding and edge deghosting (SPEED). *Magn Reson Med* 2005;53(5):1112–1117.
- [12] Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Phys D* 1992;60(1-4):259–268.
- [13] Bredies K, Kunisch K, Pock T. Total generalized variation. *SIAM Journal on Imaging Sciences* 2010;3(3):492–526.
- [14] Bakushinsky AB, Kokurin MY. Iterative methods for approximate solution of inverse problems, volume 577 of *Mathematics and Its Applications (New York)*. Dordrecht: Springer, 2004.
- [15] Engl HW, Hanke M, Neubauer A. Regularization of inverse problems, volume 375 of *Mathematics and its Applications*. Dordrecht: Kluwer Academic Publishers Group, 1996.
- [16] Blaschke B, Neubauer A, Scherzer O. On convergence rates for the iteratively regularized Gauss-Newton method. *IMA J Numer Anal* 1997;17(3):421–436.
- [17] Hohage T. Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems* 1997;13(5):1279–1299.
- [18] Pock T, Cremers D, Bischof H, Chambolle A. An algorithm for minimizing the Mumford-Shah functional. In *International Conference on Computer Vision (ICCV)*. 2009; 1133 – 1140.
- [19] Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. Technical Report 685, Ecole Polytechnique, Centre de Mathématiques Appliquées, 2010.
- [20] Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. *Magn Reson Med* 2010; accepted for publication, in production.

- [21] Buehrer M, Pruessmann KP, Boesiger P, Kozerke S. Array compression for MRI with large coil arrays. *Magn Reson Med* 2007;57(6):1131–1139.
- [22] Knoll F, Clason C, Stollberger R. Tailored 3D random sampling patterns for nonlinear parallel imaging. In *Proc. Intl. Soc. Mag. Reson. Med.*, 18. 2010; 2876.
- [23] Zhang S, Block KT, Frahm J. Magnetic resonance imaging in real time: advances using radial FLASH. *J Magn Reson Imaging* 2010;31(1):101–109.
- [24] Uecker M, Zhang S, Voit D, Karaus A, Merboldt KD, Frahm J. Real-time MRI at a resolution of 20 ms. *NMR Biomed* 2010;23:986–994.
- [25] Fessler JA, Sutton BP. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Transactions on Signal Processing* 2003;51(2):560–574.
- [26] Block KT, Uecker M, Frahm J. Model-based iterative reconstruction for radial fast spin-echo MRI. *IEEE Trans Med Imaging* 2009;28(11):1759–1769.
- [27] Doneva M, Boernert P, Eggers H, Stehning C, Senegas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magn Reson Med* 2010;.