

Article

# MRI Reconstruction Using Markov Random Field and Total Variation as Composite Prior

Marko Panić <sup>1,\*</sup> , Dušan Jakovetić <sup>2</sup>, Dejan Vukobratović <sup>3,\*</sup>, Vladimir Crnojević <sup>1</sup> and Aleksandra Pižurica <sup>4,\*</sup> 

<sup>1</sup> BioSense Institute, University of Novi Sad, 21000 Novi Sad, Serbia; crnojevic@biosense.rs

<sup>2</sup> Faculty of Sciences, University of Novi Sad, 21000 Novi Sad, Serbia; dusan.jakovetic@dmi.uns.ac.rs

<sup>3</sup> Faculty of Technical Sciences, University of Novi Sad, 21000 Novi Sad, Serbia

<sup>4</sup> Department of Telecommunications and Information Processing, Ghent University, 9000 Ghent, Belgium

\* Correspondence: panic@biosense.rs (M.P.); dejanv@uns.ac.rs (D.V.); Aleksandra.Pizurica@ugent.be (A.P.)

Received: 30 March 2020; Accepted: 27 May 2020; Published: 3 June 2020



**Abstract:** Reconstruction of magnetic resonance images (MRI) benefits from incorporating a priori knowledge about statistical dependencies among the representation coefficients. Recent results demonstrate that modeling intraband dependencies with Markov Random Field (MRF) models enable superior reconstructions compared to inter-scale models. In this paper, we develop a novel reconstruction method, which includes a composite prior based on an MRF model and Total Variation (TV). We use an anisotropic MRF model and propose an original data-driven method for the adaptive estimation of its parameters. From a Bayesian perspective, we define a new position-dependent type of regularization and derive a compact reconstruction algorithm with a novel soft-thresholding rule. Experimental results show the effectiveness of this method compared to the state of the art in the field.

**Keywords:** magnetic resonance imaging; Markov random field; image reconstruction

## 1. Introduction

Recent years have witnessed huge interest in accelerating Magnetic Resonance Imaging (MRI) acquisition and improving the achievable quality of MRI images based on compressed sensing [1–3]. Reconstruction of a magnetic resonance image  $\mathbf{x} \in \mathbb{C}^N$  is formulated as the following optimization problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \tau\phi(\mathbf{B}\mathbf{x}) \quad (1)$$

where  $\mathbf{y} \in \mathbb{C}^M$  are measurements obtained through partially observed Fourier transform  $\mathbf{A} \in \mathbb{C}^{M \times N}$ ,  $M \ll N$  [4].  $\mathbf{B} \in \mathbb{C}^{D \times N}$  denotes analysis operator and  $\phi$  is a regularization function, with the parameter  $\tau > 0$ .  $\phi$  is typically the  $\ell_1$ -norm:  $\phi(\boldsymbol{\theta}) = \|\boldsymbol{\theta}\|_1$  when  $\mathbf{B} = \mathbf{W}$  denotes a wavelet-like transformation of the signal  $\mathbf{x}$ , e.g.  $\boldsymbol{\theta} = \mathbf{W}\mathbf{x}$ , or the TV pseudo-norm  $\|\mathbf{x}\|_{\text{TV}}$  when  $\mathbf{B} = \mathbf{I}$  is the identity operator. With compound regularization, the problem in (1) becomes

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \sum_{i=1}^m \tau_i \phi_i(\mathbf{B}_i \mathbf{x}). \quad (2)$$

Many reported methods, including [4–7] combine  $\ell_1$  and TV regularization:  $\mathbf{B}_1 = \mathbf{I}$ ,  $\mathbf{B}_2 = \mathbf{W}$ ,  $\phi_1(\mathbf{x}) = \|\mathbf{x}\|_{\text{TV}}$ ,  $\phi_2(\mathbf{W}\mathbf{x}) = \|\boldsymbol{\theta}\|_1$ , which proved to be beneficial over either of these regularization functions alone.

Apart from CS-MRI iterative algorithms for solving the problem in (1), deep learning, which has been successful in solving many computer vision problems, has also been proved beneficial in tackling

image reconstruction problems [8,9]. In particular, deep convolutional neural networks (CNNs) are extremely powerful and efficient in learning signal representation [10]. Wang et al. in [11] train a deep CNN from downsampled reconstruction MRI images to learn a fully sampled reconstruction and then use deep learning results either as an initialization or as a regularization term (1). Following the same principle, the authors of [12] trained a reconstruction function from undersampled/full sampled image pairs, utilizing a more complex network structure based on U-net architecture [13] compared to the one used in [11]. In [14], motivated by a linkage between CNNs and Hankel matrix decomposition, the authors propose a fully data-driven deep learning algorithm for  $k$ -space interpolation of undersampled measurements. A trend of designing complex network architectures for constructing de-aliasing MRI reconstruction function was continued in [15,16] with the involvement of Generative Adversarial Network (GAN) [17] and U-net architectures and in [18,19] where a deep cascaded CNNs and its stochastic variation are used. A GAN based network architecture demonstrated also promising performances in the task of super-resolution (SR) for image enhancement which has great importance in medical image applications [20–23]. However, besides very good performances, the training procedure of deep CNNs, especially GANs, demands good hyperparameters tuning to achieve training convergence, which without experience is a quite difficult and time-consuming task. Another observation is that most of the results of trained deep CNNs are reported for the brain MRI scans, which means that their usage for the MRI images of another anatomy would possibly require a new training cycle.

Recent research showed that MRI reconstruction benefits from modelling statistically the spatial clustering properties of the important transform coefficients, i.e., modelling the coefficient support. Representatives are wavelet-tree sparsity methods [24,25] and methods that employ MRF priors [26–29]. A work of Panić et al. [29] demonstrated a superior reconstruction performance of MRF-based approaches over wavelet-tree sparsity methods. The reconstruction algorithm in [29] is an augmented Lagrangian method inspired by [30] and extended with an MRF prior. A variant of this algorithm, also reported in [29] with compound regularization consisting of a MRF-based and TV regularization terms demonstrated significant improvements over several state-of-the-art methods, including [25,31,32]. Although the algorithms in [29] perform well over a wide range of the parameters of the MRF model, on the analysed human brain images, they require parameter tuning when imaging different anatomical structures. We also found that hard-thresholding rule in [29] based on the estimated sparse signal support leads to some small oscillations around the convergence state, which do not affect noticeably the quality of the reconstruction but may impair practical determination of the stopping criterion.

In this paper, we aim to develop a more robust model, with automatic estimation of the parameters and with improved stability. We also extend the isotropic MRF model used in [29] to a more flexible non-isotropic model that can better capture the properties of various anatomic structures and adapt to these automatically through efficient parameter estimation. To solve the resulting optimisation problem, we develop a novel numerical method that involves a Nesterov acceleration step [33,34]. The main technical novelties are in: (i) developing a compressed sensing approach for MRI images with an anisotropic MRF model and deriving automatic estimation of its parameters; (ii) deriving a more stable solver, which replaces hard selection of the coefficients by a soft-thresholding (note that the resulting estimator is derived analytically by introducing the corresponding proximal operators for MRF-based and TV regularization functions); (iii) we extend this approach such that it is applicable not only for the reconstruction of the magnitude images as in [29], but also to complex images and to multi-coil data.

The paper is organized as follows. The formulation of the problem with the proposed regularization function is given in Section 2.1. In Section 2.2 we derive the efficient solver, and in Section 2.3 we develop an efficient method for the estimation of the MRF parameters. Section 2.4 introduces algorithm extensions to reconstruction of complex images from one or more coil measurements. Section 3 presents experimental results and comparison with state-of-the-art methods. In Section 4 we provide our conclusions.

## 2. Methodology

### 2.1. MRF Priors in MRI Reconstruction

Consider a Bayesian approach to recovering jointly representation coefficients  $\theta = \mathbf{W}\mathbf{x}$  and their bi-level significance map  $\mathbf{s}$  with the maximum a posteriori probability (MAP) criterion:  $P(\theta, \mathbf{s}|\mathbf{y}) \propto p(\mathbf{y}|\theta)p(\theta|\mathbf{s})P(\mathbf{s})$ . Here,  $\mathbf{s} \in \{-1, 1\}^D$  are the latent variables formed in a lattice that encode the significance of the representation coefficients:  $s_i = +1$  on those positions  $i$  in lattice where  $\theta_i$  is significant, and  $s_i = -1$  conversely.  $P(\mathbf{s})$  is the prior probability of  $\mathbf{s}$  while  $p(\mathbf{y}|\theta)$  and  $p(\theta|\mathbf{s})$  are conditional probability density functions (p.d.f.). Equivalent formulations of the graphical model were presented, e.g., in [27–29] and a similar one in [26,35] for the signal sparse in image domain. Using the negative log transformation of  $P(\theta, \mathbf{s}|\mathbf{y})$ , the MAP criterion translates into the joint minimization with respect to (w.r.t.)  $\theta$  and  $\mathbf{s}$ :

$$\min_{\theta, \mathbf{s}} -(\log p(\mathbf{y}|\theta) + \log p(\theta|\mathbf{s}) + \log P(\mathbf{s})). \quad (3)$$

We make a connection between (3) and formulations of type (1) by reordering the terms as follows:

$$\min_{\theta} \left\{ \underbrace{-\log p(\mathbf{y}|\theta)}_{f(\mathbf{W}^H\theta)} + \underbrace{\min_{\mathbf{s}} -(\log p(\theta|\mathbf{s}) + \log P(\mathbf{s}))}_{\psi_{\text{MRF}}(\mathbf{W}\mathbf{x})} \right\}. \quad (4)$$

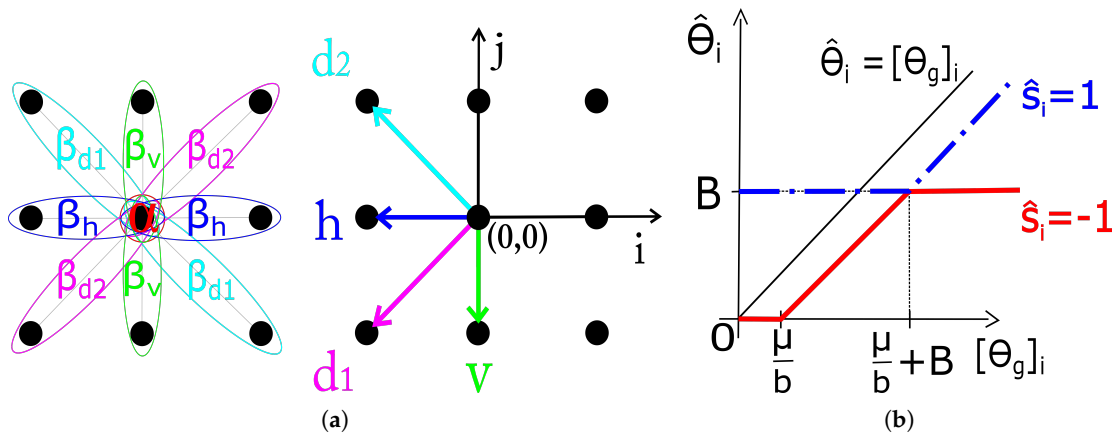
Assuming that measurements  $\mathbf{y}$  are corrupted by independent identically distributed (i.i.d.) Gaussian noise, we have that  $p(\mathbf{y}|\theta) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{P}^H\theta, \mathbf{I})$ . It then follows that the first term  $f(\mathbf{W}^H\theta)$  in (4) using a relation  $\mathbf{x} = \mathbf{W}^H\theta$  becomes equal to data fidelity part  $\frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$  in (1).  $\psi_{\text{MRF}}(\mathbf{W}\mathbf{x})$  is the new type of regularization function, which is taking the role of the second term  $\tau\phi(\mathbf{B}\mathbf{x})$  term from (1). Assuming conditional independence  $p(\theta|\mathbf{s}) = \prod_i p(\theta_i|s_i)$  and a Laplacian prior for image coefficients  $p(\theta_i)$  we define accordingly conditionals  $p(\theta_i|s_i)$  as follows:  $p(\theta_i|s_i = 1) \propto \frac{1}{2b}e^{-\frac{|\theta_i|}{b}}$  for  $|\theta_i| \geq B$  and zero otherwise while  $p(\theta_i|s_i = -1)$  is defined oppositely with respect to  $B$ . The threshold  $B$  is related to the noise level and estimated from the highest-frequency subband coefficients; the scale parameter  $b$  is adaptively estimated per subband [29]. We choose the Gibbs distribution (MRF prior) for the  $\theta$ 's support model  $P(\mathbf{s}) = \frac{1}{Z}e^{-H(\mathbf{s})/T}$ , where the normalizing constant  $Z = \sum_{\mathbf{s}} e^{-H(\mathbf{s})/T}$  represents the partition function, and the temperature  $T$  controls the peaking in the probability density [36]. We employ a common model with pair-site cliques. However, in contrast to our previous work [27–29], we allow for different interaction coefficients for cliques of different orientations, which models better the actual subband statistics. The energy function of this anisotropic MRF is

$$H(\mathbf{s}) = \sum_i \alpha s_i + \sum_{\langle i,j \rangle \in \mathcal{C}} \beta_o s_i s_j \quad (5)$$

where  $\alpha$  expresses the a priori preference for labels of one type over the other and  $\beta_o$  the interaction strength for cliques with orientation  $o \in \{h, v, d1, d2\}$  shown in Figure 1. The estimation of the MRF parameters  $(\beta_o, \alpha)$  is treated in Section 2.3.

Involving the composition of the MRF-based regularization function  $\psi_{\text{MRF}}(\mathbf{W}\mathbf{x})$  and TV norm  $\phi_{\text{TV}}(\mathbf{x}) = \|\mathbf{x}\|_{\text{TV}}$ , defined as  $\|\mathbf{x}\|_{\text{TV}} = \sum_i \sum_j \sqrt{|\mathbf{x}_{i+1,j} - \mathbf{x}_{i,j}|^2 + |\mathbf{x}_{i,j+1} - \mathbf{x}_{i,j}|^2}$ , we arrive at the proposed optimization problem formulation:

$$\min_{\mathbf{x}} \left\{ f(\mathbf{x}) + \psi_{\text{MRF}}(\mathbf{W}\mathbf{x}) + \phi_{\text{TV}}(\mathbf{x}) \right\}. \quad (6)$$



**Figure 1.** (a) Single and pair-site cliques for the second-order neighborhood with parameters and directions used in calculation of  $R_i$ . (b) Soft-thresholding rules for image coefficients  $[\theta_g]_i$  based on the estimated  $\hat{s}_i = \{1, -1\}$ .

The optimization problem (6) is non-convex in general due to the presence of the non-convex  $\psi_{\text{MRF}}$  regularization function and is very hard to globally solve exactly. In the following section we propose a computationally efficient method to suboptimally solve (6).

### 2.2. Proposed Algorithm

To solve the problem (6) we adopt the idea of fast composite splitting from [37] and extend this algorithm to deal with the MRF-based regularization. The key difference is that our model involves  $\psi_{\text{MRF}}$  instead of  $\ell_1$ -norm in [37]. Hence, we have to solve the proximal map  $\hat{\theta} = \text{prox}_{\mu}(\psi_{\text{MRF}})(\mathbf{W}\mathbf{x}_g)$ :

$$\underset{\theta}{\text{argmin}} \left( \psi_{\text{MRF}}(\theta) + \frac{1}{2\mu} \|\theta - \mathbf{W}\mathbf{x}_g\|_2^2 \right). \tag{7}$$

Note that, in notation  $\text{prox}_{\mu}(\psi_{\text{MRF}})(\mathbf{W}\mathbf{x}_g)$ ,  $\psi_{\text{MRF}}$  is the function for which the proximal map is calculated, and  $\theta_g = \mathbf{W}\mathbf{x}_g$  is the argument at which it is evaluated. Since the evaluation of this proximal map is very hard, we adopt a suboptimal, block-coordinate approach explained in the following.

The key novel ingredient of the proposed algorithm is a computationally efficient approximation of the proximal map in (7) for a fixed  $\mathbf{x}_g$ . From (4), it is clear that the minimization in (7) needs to be carried out jointly w.r.t.  $\theta$  and  $\mathbf{s}$  in order to evaluate (7) exactly. We adopt a suboptimal yet computationally efficient approach where we first minimize the second term in (4) w.r.t.  $\mathbf{s} \in \{-1, 1\}^D$  for a fixed  $\theta$ ; then we fix the obtained  $\hat{\mathbf{s}}$  and minimize the objective in (7) w.r.t.  $\theta$ . For the former minimization step, we adopt notation  $\hat{\mathbf{s}} = \text{MAP-support}\{\theta\}$ . The step is done via the Metropolis sampler procedure using a “warm-start” initial  $\mathbf{s}$  as in [28,29]. The latter minimization (w.r.t.  $\theta$  for a fixed  $\hat{\mathbf{s}}$ ), as shown here, can be done in closed-form and leads to a novel soft-thresholding operation. This is accomplished by solving the following minimization problem

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \underbrace{\sum_i -\log p(\theta_i | \hat{s}_i)}_{\psi_{\hat{\mathbf{s}}}(\theta)} + \frac{1}{2\mu} (\theta_i - [\theta_g]_i)^2. \tag{8}$$

Equation (8) is derived from (7) using a simple algebraic manipulation and omitting the terms that do not depend on  $\theta$ . With the analytic form of  $p_{\theta_i | \hat{s}_i}(\theta_i | \hat{s}_i)$  for  $\hat{s}_i = \{-1, 1\}$  (see Section 2.1), the closed form solution for each single component of  $\hat{\mathbf{s}}$  (it turns out that the solution decouples component-wise)

is derived in (9) (index  $i$  in the equation is omitted for notation simplicity). The solution (9) is illustrated in Figure 1 on range  $[0, \infty)$  due to its (odd) symmetry about the origin.

$$\hat{\theta}_{\hat{s}=1} = \begin{cases} B \cdot \text{sgn}(\theta_g), & |\theta_g| \leq \frac{\mu}{b} + B \\ \theta_g - \frac{\mu}{b} \cdot \text{sgn}(\theta_g), & |\theta_g| > \frac{\mu}{b} + B \end{cases}$$

$$\hat{\theta}_{\hat{s}=-1} = \begin{cases} 0, & |\theta_g| \leq \frac{\mu}{b} \\ \theta_g - \frac{\mu}{b} \cdot \text{sgn}(\theta_g), & \frac{\mu}{b} < |\theta_g| \leq \frac{\mu}{b} + B \\ B \cdot \text{sgn}(\theta_g), & |\theta_g| > \frac{\mu}{b} + B \end{cases} \quad (9)$$

The described block-coordinate cyclic procedure should in principle proceed in several iterative rounds. As we demonstrate here numerically, it is sufficient to perform a single cycle. Once the approximation of the proximal map in (7) is available, we incorporate it in a fast-splitting framework akin to the one in [37]. That is, we first carry out a gradient-like step on  $\mathbf{x}$  w.r.t. data fidelity term. Then, we perform in parallel the proximal maps that correspond to the two regularization functions in (6). The two outputs of the proximal mappings are then simply averaged and fed into a Nesterov-acceleration-like step. The motivation for the parallel-proximal-then-average approach is computational efficiency, as carrying out a joint proximal map (even approximate) w.r.t.  $\psi_{\text{MRF}}$  and  $\phi_{\text{TV}}$  would be very hard. The motivation for the Nesterov-acceleration step is to further speed-up the method.

The overall method is presented in Algorithm 1. Recall the definition of  $\psi_{\hat{s}}$  from (8). Note that steps 3-4 are the single block-coordinate cycle to approximate the operation in (7). The parameters  $\tau_1, \tau_2, \mu$  for simplicity are all set to 1, although these are not the optimal values, unless it is otherwise stated. We refer to the proposed method as Fast Composite Lattice and TV regularization (FCLaTV). A version without acceleration CLaTV can also be used, and it is obtained after omitting steps 7 and 8 from FCLaTV and replacing  $\mathbf{r}^{\{k\}}$  in the step 1 with  $\mathbf{x}^{\{k\}}$ . Extensive numerical studies demonstrate that Algorithm 1 always converges. The rigorous convergence analysis is left for future work.

---

#### Algorithm 1 FCLaTV

---

**Require:**  $k = 1, \mu, \tau_1, \tau_2, t^{\{1\}} = 1, \mathbf{y}, \mathbf{r}^{\{1\}} = \mathbf{x}^{\{0\}}$

1: **repeat**

2:  $\mathbf{x}_g = \mathbf{r}^{\{k\}} - \mu \mathbf{A}^H (\mathbf{A} \mathbf{r}^{\{k\}} - \mathbf{y})$

3:  $[\psi_{\hat{s}}, \hat{\mathbf{s}}] = \text{MAP-support}\{\mathbf{W} \mathbf{x}_g\}$

4:  $\mathbf{x}_1 = \mathbf{W}^H (\text{prox}_{\mu}(\tau_1 \psi_{\hat{s}})(\mathbf{W} \mathbf{x}_g))$

5:  $\mathbf{x}_2 = \text{prox}_{\mu}(\tau_2 \phi_{\text{TV}})(\mathbf{x}_g)$

6:  $\mathbf{x}^{\{k\}} = (\mathbf{x}_1 + \mathbf{x}_2) / 2$

7:  $t^{\{k+1\}} = (1 + \sqrt{1 + 4(t^{\{k\}})^2}) / 2$

8:  $\mathbf{r}^{\{k+1\}} = \mathbf{x}^{\{k\}} + \frac{t^{\{k\}} - 1}{t^{\{k+1\}}} (\mathbf{x}^{\{k\}} - \mathbf{x}^{\{k-1\}})$

9:  $k = k + 1$

10: **until** some stopping criterion is satisfied

11: **return**  $\mathbf{x} = \mathbf{x}^{\{k\}}$

---

### 2.3. Parameter Estimation for the Anisotropic MRF Prior

Here we propose a data-driven approach for specifying the parameters  $\{\alpha, \beta_h, \beta_v, \beta_{d1}, \beta_{d2}\}$  of the MRF model. The core idea is to relate the parameters of the prior  $P(\mathbf{s})$  for the support of  $\theta$  to some measurable characteristics of the observed  $\theta$ . The representation coefficients  $\theta$  are re-estimated in each iteration of Algorithm 1 and so are the MRF parameters.

The four interaction coefficients  $\{\beta_h, \beta_v, \beta_{d1}, \beta_{d2}\}$  represent the clustering strength of the coefficient labels  $\mathbf{s}$  in the corresponding four directions for the observed subband. We reason that the clustering strength in a particular direction should be proportional to the correlation among the corresponding representation coefficients  $\theta$  in that direction. Therefore, we base the estimation of interaction coefficients on magnitude of representation coefficients  $\theta$ . In order to reduce the effect the noise present in  $\theta$  on estimation of the interaction coefficients, we used squared magnitude of  $\theta$  and account only for those that were marked as significant by the initial  $\mathbf{s}$  prior to its MAP estimation  $\hat{\mathbf{s}}$ . In what follows,  $\theta^S$  denotes a subband in which all the coefficients that were selected as significant by  $\mathbf{s}$  are left unchanged ( $\theta_i^S = \theta_i$  if  $s_i = 1$ ) while others are set to zero ( $\theta_i^S = 0$  if  $s_i = -1$ ). To express two-dimensional (2D) correlation, we need to revert to 2D spatial indices  $(k, l)$ . Let  $r_{(i,j)}$  denote the correlation coefficient for squared  $\theta^S$  corresponding to the spatial shift  $(i, j)$ :

$$r_{(i,j)} = \sum_k \sum_l \left( \theta^S(k, l) \theta^S(k+i, l+j) \right)^2 \quad (10)$$

We take four correlation coefficients corresponding to the smallest possible spatial shifts in the corresponding directions, indicated by arrows in Figure 1. These are:  $r_{(-1,-1)}$ ,  $r_{(-1,0)}$ ,  $r_{(-1,1)}$  and  $r_{(0,-1)}$ . Note that by symmetry  $r_{(-1,-1)} = r_{(1,1)}$ ,  $r_{(-1,0)} = r_{(1,0)}$ ,  $r_{(-1,1)} = r_{(1,-1)}$  and  $r_{(0,-1)} = r_{(0,1)}$ . The mapping  $t = 3i + j + 4$  translates these index pairs into a single index  $t \in \{0, 1, 2, 3\}$ . With  $R_t = r_{(i,j)}$  we have that  $R_0 = r_{(-1,-1)}$  is the correlation coefficient in the  $45^\circ$  ( $d_1$ -direction),  $R_1 = r_{(-1,0)}$  the horizontal,  $R_2 = r_{(-1,1)}$  the  $135^\circ$  ( $d_2$ -direction) and  $R_3 = r_{(0,-1)}$  in the vertical direction. Now we can specify the four interaction coefficients as normalized correlation coefficients of  $\theta^S$ :

$$\begin{aligned} \beta_{d1} &= \frac{R_0}{\|\mathbf{R}\|_2}, & \beta_h &= \frac{R_1}{\|\mathbf{R}\|_2} \\ \beta_{d2} &= \frac{R_2}{\|\mathbf{R}\|_2}, & \beta_v &= \frac{R_3}{\|\mathbf{R}\|_2} \end{aligned} \quad (11)$$

The normalization by  $\ell_2$ -norm here is optional (as only the relative values of the MRF parameters with respect to each other actually matter) but we find it convenient in practice to have these parameters in the range  $[-1, 1]$  as it is now guaranteed. We also tested  $\ell_1$  and  $\ell_\infty$  for the purpose of this normalization, but  $\ell_2$  led to best performances in our experiments.

It still remains to specify the parameter  $\alpha$ , which represents a priori preference for one type of labels ( $-1$  or  $+1$ ) over the other. With  $\alpha = 0$  both labels are a priori equally likely and as  $\alpha$  increases in magnitude the more preference goes to one of these labels. For  $\alpha > 0$ , the labels  $-1$  will be favoured, which means that significant coefficients (labelled by  $+1$ ) will be sparse. We specify  $\alpha$  as the mean energy of the coefficients in  $\theta^S$  relative to the energy of the largest coefficient in that subband:

$$\alpha = \frac{1}{N} \frac{\|\theta^S\|_2^2}{\left( \|\theta^S\|_{\max} \right)^2} \quad (12)$$

Note that we have omitted the iteration indices for compactness. In fact, we have sequences  $\mathbf{s}^{(k)}$ ,  $\theta^{S(k)}$  and  $\alpha^{(k)}$ ,  $\beta_h^{(k)}$ ,  $\beta_v^{(k)}$ ,  $\beta_{d1}^{(k)}$ ,  $\beta_{d2}^{(k)}$  that get improved through iterations  $k$ .

#### 2.4. Complex Image Reconstruction and Multi-Coil Reconstruction

The proposed approach is developed for the reconstruction of MR image magnitude. Here we extend it for the reconstruction of complex images (magnitude and phase) and the reconstruction of magnitude images from multi-coil measurements. In cases where it is of interest to recover the image phase together with the magnitude, steps 4 and 5 in Algorithm 1 are repeated twice: first for the regularization of the real part and then, equivalently, for the regularization of the imaginary part of a

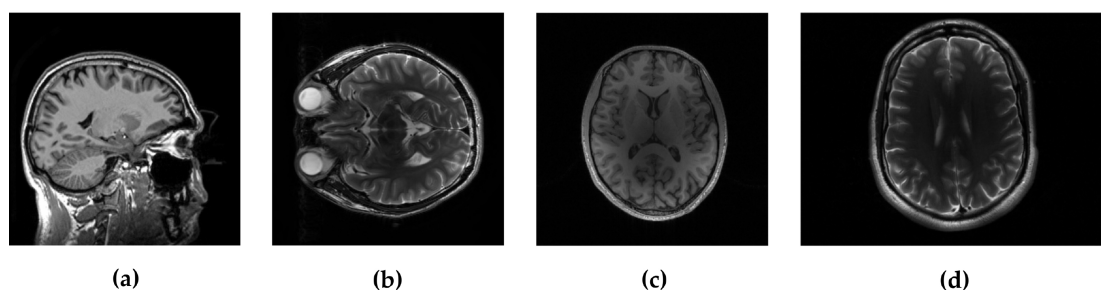
complex image  $\mathbf{x}$ . This approach showed promising performances as it will be seen in the following section. A multi-coil image reconstruction demands knowledge of sensitivity profiles for each coil and therefore different construction of the operator  $\mathbf{A}$  during the reconstruction procedure. If we denote with  $\mathbf{C}_i$  the sensitivity profile for coil  $i$  then undersampled measurements from the same coil are  $\mathbf{y}_i = \mathbf{A}\mathbf{C}_i\mathbf{x}$ . Let us collect all available measurements from  $N_c$  coils in one vector  $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_{N_c}^T]^T$  and create an augmented vectorized image  $\mathbf{x}_a = \mathbf{T}\mathbf{x}$  with the usage of matrix  $\mathbf{T}$  which is formed by stacking the identity matrix  $\mathbf{I}_{N \times N}$   $N_c$  times row-wise. Then the acquisition process is defined through the following equation

$$\mathbf{y} = \mathbf{A}_B \mathbf{C}_B \mathbf{x}_a \quad (13)$$

where  $\mathbf{A}_B$  consists of repeated  $\mathbf{A}$  along the diagonal while  $\mathbf{C}_B$  is formed by stacking coil sensitivity maps  $\mathbf{C}_i$  along diagonal. This way a multi-coil reconstruction problem can be solved using the single-coil reconstruction method developed in this paper. This generalization in the definition of the measurement operator  $\mathbf{M} = \mathbf{A}_B \mathbf{C}_B$  is easily introduced in step 2 in Algorithm 1 instead of  $\mathbf{A}$  corresponding to the single-coil reconstruction scenario. The computational complexity is marginally increased and doesn't require additional code optimization.

### 3. Results

In the experimental evaluation we used different MRI images, starting from high-resolution MRI images, shown in Figure 2, and using real data acquired in  $k$ -space. For the following experiments, we used simulated sampling trajectories on a Cartesian grid with different sampling rates, except for the last experiment, when the measurements are undersampled with the non-Cartesian radial sampling trajectory used in a real scan. First, we consider reconstruction of MR image magnitude for different sampling rate (SR) obtained using various sampling trajectories. For this we utilize dataset of 248 T1 MRI brain slices acquired on a Cartesian grid at Ghent University hospital (one *sagittal* slice is presented in Figure 2). Then we focus on the reconstruction of complex MR images from single and multi-coil undersampled measurements. Complex T2-weighted brain images *axial-1*, *axial-3* from [31] and [38] respectively (their magnitudes are presented in Figure 2) are used in experiments for the reconstruction of single-coil complex images while the T1-weighted brain image *axial-2* [38] is used in the multi-coil reconstruction experiment. We also present the reconstruction results on real radially acquired measurements in  $k$ -space on non-Cartesian grid. These data, consisting of the acquisitions of a pomelo fruit were supplied by the Bioimaging Lab in Antwerp. In our experiments for sparse signal representation we used the non-decimated wavelet transform with 3 scales. For comparison, we report the results of LaSAL and LaSAL2 from [29], FCSA [5], FCSANL [39] and WaTMRI [25] with the original implementations. All these methods, except LaSAL, employ a compound regularization. The reconstruction results for complex images from single-coil and multi-coil measurements, are compared with the corresponding results of pFISTA [31] and P-LORAKS [38,40] methods.



**Figure 2.** Test images: (a) *sagittal*, (b) *axial-1* and (c) *axial-2*, all three  $256 \times 256$ , and (d) *axial-3* with resolution  $256 \times 340$ .

For quantitative comparison between the reconstructed and the reference MR image we adopt Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR is defined as:

$$PSNR(\hat{\mathbf{x}}, \mathbf{x}_r) = 10 \log_{10} \left( \frac{MAX_{MR}^2}{MSE(\hat{\mathbf{x}}, \mathbf{x}_r)} \right) \quad (14)$$

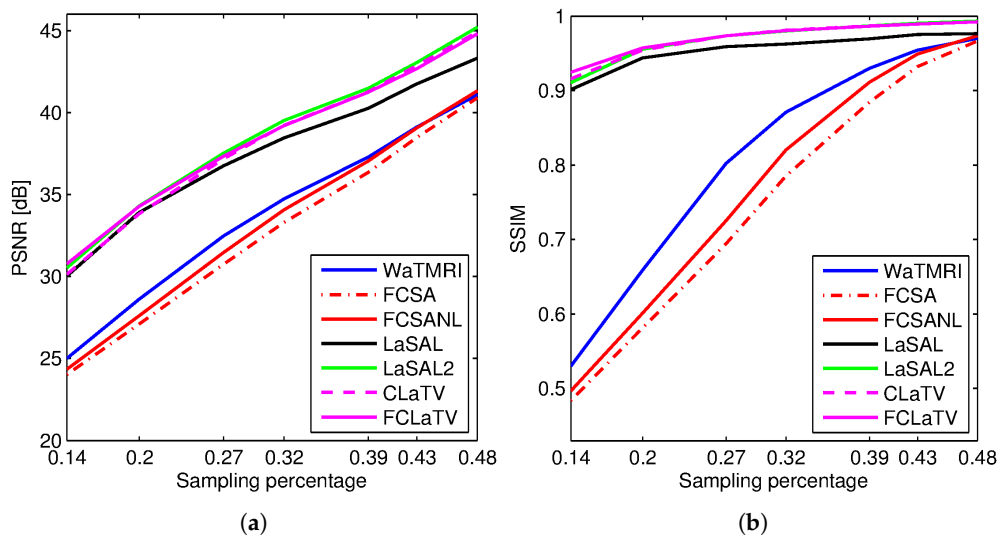
where  $MAX_{MR}^2$  denotes squared maximum possible pixel value in the magnitude of MR image and  $MSE$  is the mean squared error between magnitudes of reconstructed  $\hat{\mathbf{x}}$  and reference image  $\mathbf{x}_r$ . For the calculation of SSIM index we used the simplified equation form:

$$SSIM(\hat{\mathbf{x}}, \mathbf{x}_r) = \frac{(2\mu_{\hat{\mathbf{x}}}\mu_{\mathbf{x}_r} + C_1)(2\sigma_{\hat{\mathbf{x}}\mathbf{x}_r} + C_2)}{(\mu_{\hat{\mathbf{x}}}^2 + \mu_{\mathbf{x}_r}^2 + C_1)(\sigma_{\hat{\mathbf{x}}}^2 + \sigma_{\mathbf{x}_r}^2 + C_2)} \quad (15)$$

where  $\mu_{\hat{\mathbf{x}}}$ ,  $\mu_{\mathbf{x}_r}$ ,  $\sigma_{\hat{\mathbf{x}}}$ ,  $\sigma_{\mathbf{x}_r}$ ,  $\sigma_{\hat{\mathbf{x}}\mathbf{x}_r}$  are the local means, standard deviations, and cross-covariance for images  $\hat{\mathbf{x}}$ ,  $\mathbf{x}_r$  and  $C_1 = (0.01 * MAX_{MR})^2$  and  $C_2 = (0.03 * MAX_{MR})^2$  are regularization constants to avoid instability in index calculation. Local statistics is calculated using isotropic Gaussian function with radius 1.5. Through experiments pixel values are stored in double-precision floating-point format with dynamic range of  $[0, 255]$  unless otherwise stated.

### 3.1. Data Sets Acquired on Cartesian Grid

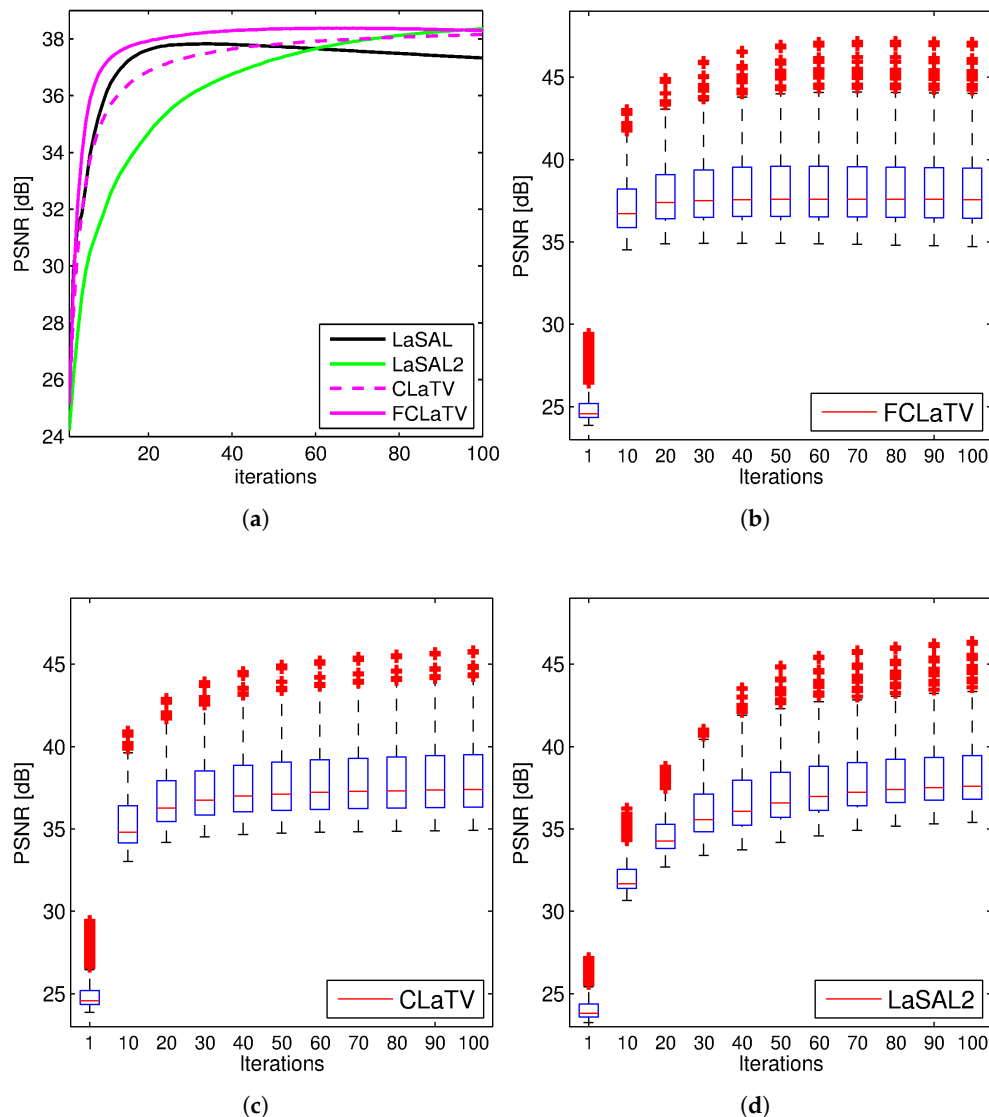
Figure 3 shows the PSNR and SSIM for the reconstructed *sagittal* MR image from radially undersampled measurements with sampling rate (SR) ranging from 14% to 48%. The MRF-based methods LaSAL, LaSAL2, CLaTV and FCLaTV achieve a consistent and significant improvement in PSNR (at some sampling rates more than 4 dB) compared to WaTMRI, FCSA and FCSANL. The proposed methods CLaTV and FCLaTV outperform LaSAL and yield only slightly lower reconstruction PSNR and equally good SSIM as the best reference method LaSAL2. These results are achieved with automatic estimation of MRF parameters and without tuning the regularization parameters (the  $\mu, \tau_1$  and  $\tau_2$  parameters are all set to 1).



**Figure 3.** (a) PSNR and (b) SSIM for the reconstructions of the test image *sagittal* at different sampling rates.

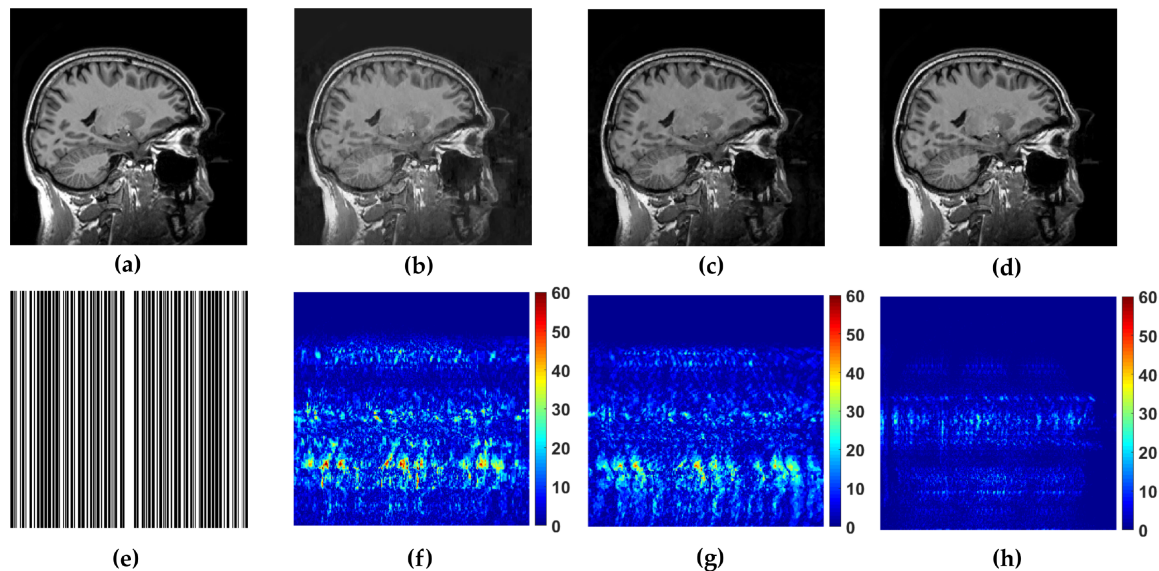


Performances of the proposed methods are further tested through reconstruction of all 248 T1 MRI slices from data set. Results are shown in Figure 4 where besides mean PSNR values through iterations we also provide distribution of PSNR values for CLaTV, FCLaTV and LaSAL2 methods. On average, FCLaTV reached the peak performance much before CLaTV and LaSAL2. The LaSAL2 achieved its highest PSNR in average after 80 iterations which is 2 times more after FCLaTV. CLaTV, FCLaTV and LaSAL2 reached the same maximal median PSNR value of 37.5 dB through all iterations. The proposed methods need less iterations than LaSAL2 to reach this value which is presented in Figure 4 with the shift of PSNR distribution towards higher values in the first 20 iterations.



**Figure 4.** PSNR values obtained from 248 MRI brain slices from the first data set, with radial sampling (SR 25%). Mean PSNR (a) and the PSNR distribution for FCLaTV (b), CLaTV (c), LaSAL2 (d). The result is presented as a box plot: the edges of the each box represents 25th and 75th percentile while the central mark (red line) in the box is median. The whiskers extend to the most extreme PSNR values which are not considered outliers while outliers are plotted separately with red crosses.

WaTMRI, LaSAL2 and FCLaTV are further tested in reconstruction from measurements undersampled with straight vertical lines in  $k$ -space which is usual sampling trajectory in real scanners. From the results of this experiment, shown in Figure 5, the FCLaTV method outperforms WaTMRI and LaSAL2 visually and in terms of PSNR and SSIM measure.



**Figure 5.** Reference image (a) *sagittal* and its reconstruction with the following methods (b) WaTMRI (PSNR = 28.70 dB, SSIM = 0.81), (c) LaSAL2 (PSNR = 29.77 dB, SSIM = 0.84) and (d) FCLaTV (PSNR = 36.49 dB, SSIM = 0.97) using 45% of measurements sampled with trajectory shown in (e). The corresponding reconstruction error of the used methods are shown in (f–h).

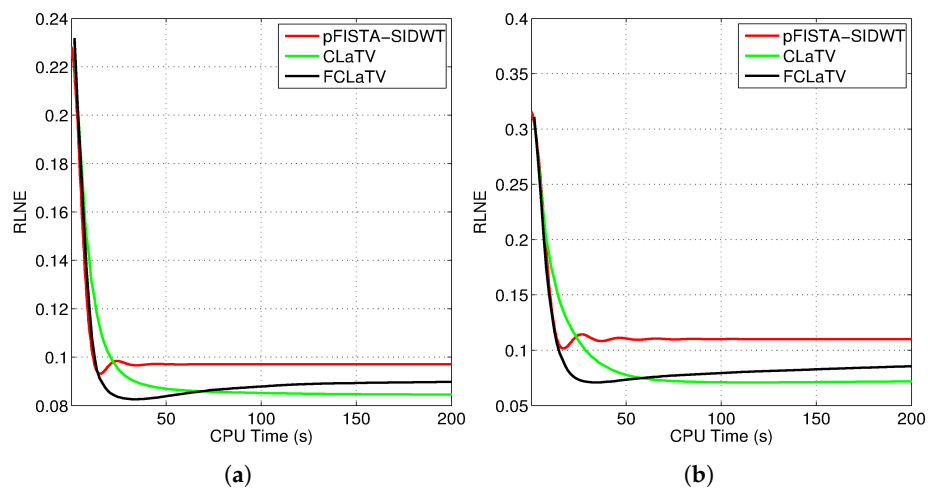
We tested the proposed methods in the reconstruction of a complex T2-weighted MR image *axial-1* slice, the magnitude of which is shown in Figure 2. For the performance measure we use relative  $\ell_2$  norm error (RLNE), defined as

$$e(\hat{\mathbf{x}}) = \frac{\|\hat{\mathbf{x}} - \mathbf{x}_r\|_2}{\|\mathbf{x}_r\|_2} \quad (16)$$

where  $\mathbf{x}_r$  is the reference image recovered from all measurements while  $\hat{\mathbf{x}}$  denotes the estimated (reconstructed) image from undersampled measurements. Algorithm steps 4 and 5, which refer to proximal operators are simultaneously applied on real and imaginary parts of the temporary reconstructed image  $\mathbf{x}_g$  obtained from step 2. We found that this adaptation of algorithm for handling reconstruction of complex images achieves the best performances. For the reference methods we use the pFISTA method from [31] with the Shift Invariant Discrete Wavelet Transform (SIDWT) as a tight frame for sparse signal representation. Results are shown in Figure 6 with respect to amount of CPU (Intel(R) Core(TM) i7-4700MQ CPU @ 2.4GHz, 8GB of RAM memory) time needed for reconstruction. For the both trajectories, CLaTV and FCLaTV outperform pFISTA-SIDWT in terms of RLNE measure. FCLaTV in less than 50s reaches the lowest RLNE for both trajectories, 0.0826 for radial and 0.0709 for random, while CLaTV needs around 70s to reach RLNE of 0.0846 for radial and 0.0711 for random trajectory.

A comparison of FCLaTV with the P-LORAKS [40] method is conducted by reconstructing the single channel T2-weighted complex brain image *axial-3* from 50% of measurements sampled by random and uniform trajectory. The results are presented in Table 1 where for both sampling strategies FCLaTV achieved lower RLNE value than the reference P-LORAKS method.

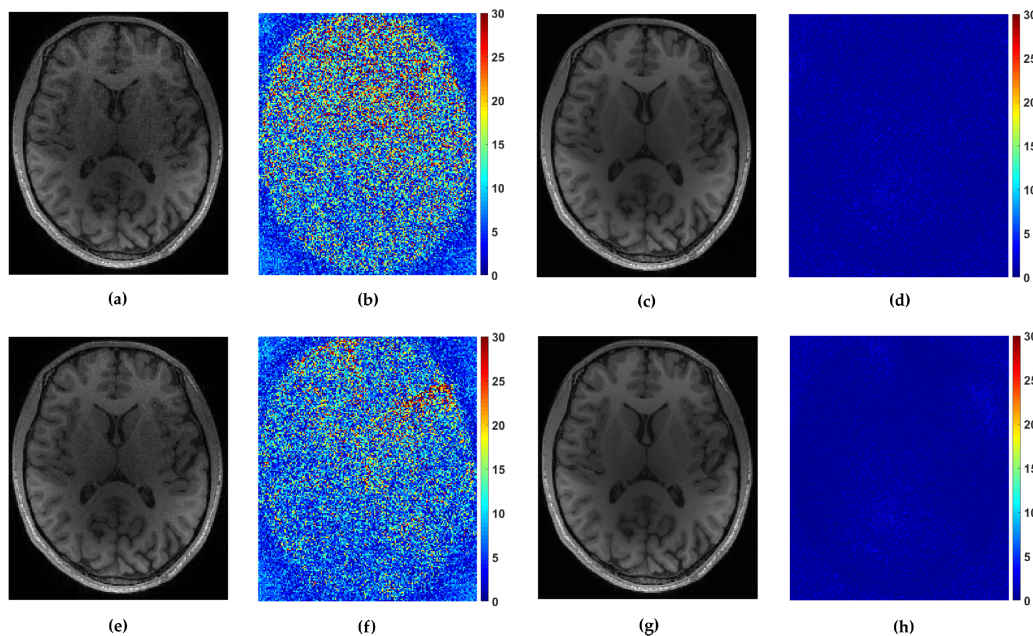
For the multi-coil reconstruction scenario, we used a T1 weighted brain image *axial-2* from [38], shown in Figure 2, and measurements gathered from 4 coils with known given sensitivity profiles. We incorporate random and uniform sampling trajectories with sampling rate of 14% per each coil. For both trajectories, the proposed FCLaTV achieved lower RLNE values reported in Table 1. Figure 7 shows visual comparison of reconstructions between FCLaTV and P-LORAKS methods.



**Figure 6.** Obtained RLNE for the reconstructed T2-weighted brain image *axial-1* from (a) radially and (b) randomly undersampled measurements with the same SR of 30%.

**Table 1.** P-LORAKS [40] and FCLaTV comparison on a T1-weighted *axial-2* and T2-weighted *axial-3* image using random and uniform sampling trajectories.

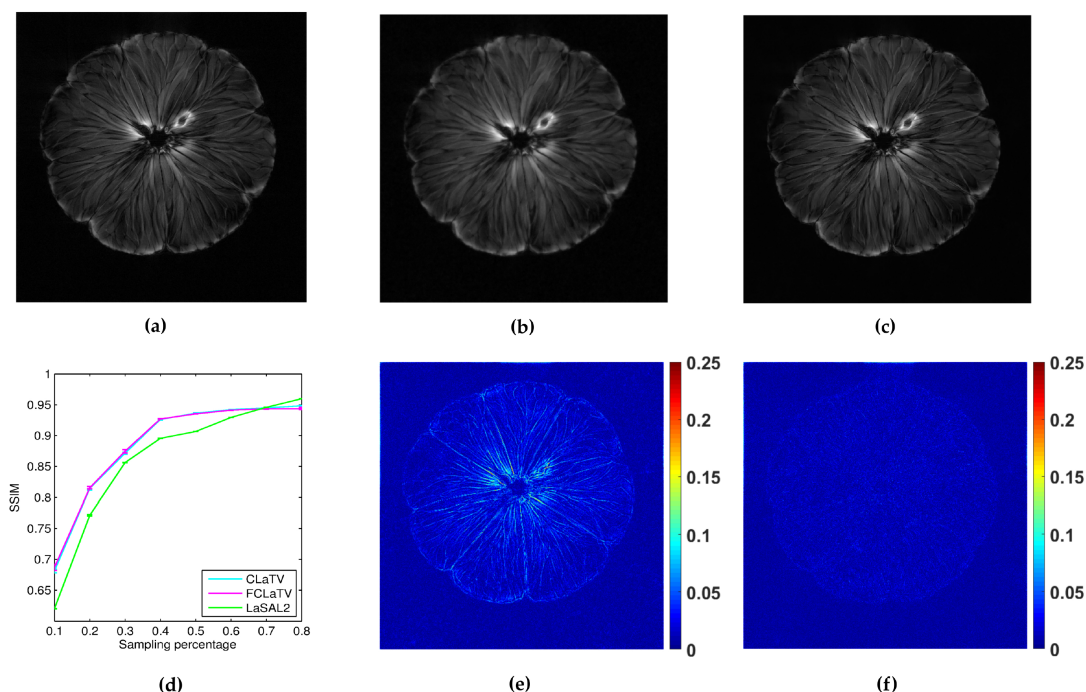
	<i>axial-2</i>		<i>axial-3</i>	
	Random	Uniform	Random	Uniform
P-LORAKS	0.06	0.07	0.083	0.097
FCLaTV	0.015	0.013	0.078	0.080



**Figure 7.** Reconstruction of *axial-2* from randomly and uniformly sampled measurements from 4 coils. (a,c) Reconstructed images with P-LORAKS and FCLaTV method respectively from randomly sampled measurements with their (b,d) corresponding errors. (e,g) Reconstructed images with P-LORAKS and FCLaTV method respectively from uniformly sampled measurements with their (f,h) corresponding errors.

### 3.2. Data Sets Acquired on Non-Cartesian Grid

In the following we present the reconstruction of a *pomelo*, acquired with radial sampling in the  $k$ -space. The data consist of 1608 radial lines, each with 1024 samples. We form undersampled versions by leaving out some of the radial lines. In particular, we aim to implement undersampling based on the golden ratio profile spacing [41], which guarantees a nearly uniform coverage of the space for an arbitrary number of the remaining radial lines. Starting from an arbitrary selected radial line, each next line is chosen by skipping an azimuthal gap of  $111.246^\circ$  until we reach desired sampling rate. In practice we cannot always achieve this gap precisely (since we have a finite, although large, number of lines to start with). Therefore we choose the nearest available radial line relative to the position obtained after moving. Since we deal here with non-uniformly sampled  $k$ -space data, we need to employ the non-uniform FFT procedures [41], which are commonly used in MRI reconstruction and readily available. For the reference method we used LaSAL2 [29]. During the reconstruction we add different amount of Gaussian noise on undersampled measurements with zero mean and the following standard deviations: 0.01, 0.02, 0.03, 0.04 and 0.05. For each amount we perform 10 experiments of reconstruction with different noise realizations. From the obtained average performances of reconstruction in terms of SSIM for different standard deviation of noise, we calculate average across all suggested standard deviations and present results in the form of mean line with standard deviation bars in Figure 8. We empirically set  $\tau_2$  to 140 and 130 for CLaTV and FCLaTV respectively due to the different dynamic range of reference image [0,1]. Regularization parameters for LaSAL2 method are taken from [29]. As expected, standard deviation of performances reduce as the sampling rate increase and the mean SSIM becomes greater for all methods. At relatively small sampling rates, up to 60% CLaTV and FCLaTV outperforms LaSAL2 which is presented in visual comparison shown in the same Figure 8. CLaTV and FCLaTV much better recover structure in the image than LaSAL2 which is demonstrated visually in the images of reconstruction error.



**Figure 8.** Reconstruction of pomelo fruit from radially sampled measurements: (a) reference image obtained by conjugate gradient method using all measurements, (b,c) pomelo reconstructions using LaSAL2 and FCLaTV methods respectively from 40% of measurements, (d) average SSIM for different sampling rate, (e,f) corresponding reconstructions errors using LaSAL2 and FCLaTV methods respectively.

#### 4. Conclusions

We considered the problem of MRI image reconstruction and formulated it as an optimization problem with a quadratic least-squares type fidelity term and a composite regularization—the sum of a total variation and a MRF regularizer. We propose a novel efficient iterative method to solve the problem that involves proximal maps and Nesterov-like acceleration. The major technical novelty is in a computationally efficient method for approximating the MRF-regularizer proximal map; the method exhibits a novel soft-thresholding rule that is different from standard  $\ell_1$ -norm-induced soft-thresholding. Another major novelty is an efficient method for estimating the MRF model parameters on the fly, during the algorithm iterations. Significant improvements in the reconstruction performance are achieved compared to the related methods. Further research direction will include an extension of evaluation metrics with perceptual quality based and semantic-based metrics [42] and involving more undersampling trajectories from the real clinical scans.

**Author Contributions:** Conceptualization, M.P., D.J. and A.P.; Methodology, M.P., D.J., D.V. and A.P.; Software, M.P.; Investigation, M.P. and V.C.; Visualization, M.P. and D.V.; Formal analysis, D.J. and A.P.; Writing—original draft preparation, M.P. and A.P.; Writing—review and editing, D.J., D.V., V.C. and A.P.; Funding acquisition, V.C. All authors have read and agreed to the published version of the manuscript

**Funding:** This work was supported by: ANTARES project funded from European Union’s Horizon 2020 research and innovation programme under grant agreement No. 739570, and national project III44006 “Development in ICT using mathematical methods with application in medicine, telecommunications, power engineering, protection of national heritage and education” from the Ministry of Education, Science and Technological Development of Republic of Serbia and in part by the the Artificial Intelligence Research Flanders program funded by the Flemish Government.

**Acknowledgments:** The authors are grateful to Karel Deblaere from the Radiology Department of UZ Gent hospital, Els Fieremans from the NYU school of Medicine and to Marleen Vehoye from the University of Antwerp for providing the data.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Deka, B.; Datta, S. *Compressed Sensing Magnetic Resonance Image Reconstruction Algorithms*; Springer Series on Bio-and Neurosystems; Springer: Singapore, 2019.
2. Fessler, J.A. Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms. *IEEE Signal Process. Mag.* **2020**, *37*, 33–40. [[CrossRef](#)] [[PubMed](#)]
3. Ahmad, R.; Bouman, C.A.; Buzzard, G.T.; Chan, S.; Liu, S.; Reehorst, E.T.; Schniter, P. Plug-and-Play Methods for Magnetic Resonance Imaging: Using Denoisers for Image Recovery. *IEEE Signal Process. Mag.* **2020**, *37*, 105–116. [[CrossRef](#)]
4. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **2007**, *58*, 1182–1195. [[CrossRef](#)] [[PubMed](#)]
5. Huang, J.; Zhang, S.; Metaxas, D. Efficient MR image reconstruction for compressed MR imaging. *Med. Image Anal.* **2011**, *15*, 670–679. [[CrossRef](#)]
6. Ma, S.; Yin, W.; Zhang, Y.; Chakraborty, A. An efficient algorithm for compressed MR imaging using total variation and wavelets. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
7. Yang, J.; Zhang, Y.; Yin, W. A fast alternating direction method for  $TV\ell_1\text{-}\ell_2$  signal reconstruction from partial Fourier data. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 288–297. [[CrossRef](#)]
8. Greenspan, H.; Van Ginneken, B.; Summers, R.M. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* **2016**, *35*, 1153–1159. [[CrossRef](#)]
9. Wang, G.; Ye, J.C.; Mueller, K.; Fessler, J.A. Image reconstruction is a new frontier of machine learning. *IEEE Trans. Med. Imaging* **2018**, *37*, 1289–1296. [[CrossRef](#)] [[PubMed](#)]
10. Ye, J.C.; Han, Y.; Cha, E. Deep convolutional framelets: A general deep learning framework for inverse problems. *SIAM J. Imaging Sci.* **2018**, *11*, 991–1048. [[CrossRef](#)]

11. Wang, S.; Su, Z.; Ying, L.; Peng, X.; Zhu, S.; Liang, F.; Feng, D.; Liang, D. Accelerating magnetic resonance imaging via deep learning. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; pp. 514–517.
12. Hyun, C.M.; Kim, H.P.; Lee, S.M.; Lee, S.; Seo, J.K. Deep learning for undersampled MRI reconstruction. *Phys. Med. Biol.* **2018**, *63*, 135007. [[CrossRef](#)]
13. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
14. Han, Y.; Sunwoo, L.; Ye, J.C. k-space deep learning for accelerated MRI. *IEEE Trans. Med. Imaging* **2019**, *39*, 377–386. [[CrossRef](#)]
15. Yang, G.; Yu, S.; Dong, H.; Slabaugh, G.; Dragotti, P.L.; Ye, X.; Liu, F.; Arridge, S.; Keegan, J.; Guo, Y.; et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans. Med. Imaging* **2017**, *37*, 1310–1321. [[CrossRef](#)]
16. Yu, S.; Dong, H.; Yang, G.; Slabaugh, G.; Dragotti, P.L.; Ye, X.; Liu, F.; Arridge, S.; Keegan, J.; Firmin, D.; et al. Deep de-aliasing for fast compressive sensing mri. *arXiv* **2017**, arXiv:1705.07137.
17. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Twenty-Eighth Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
18. Schlemper, J.; Caballero, J.; Hajnal, J.V.; Price, A.N.; Rueckert, D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imaging* **2017**, *37*, 491–503. [[CrossRef](#)] [[PubMed](#)]
19. Schlemper, J.; Yang, G.; Ferreira, P.; Scott, A.; McGill, L.A.; Khalique, Z.; Gorodezky, M.; Roehl, M.; Keegan, J.; Pennell, D.; et al. Stochastic deep compressive sensing for the reconstruction of diffusion tensor cardiac MRI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 295–303.
20. Wang, C.; Papanastasiou, G.; Tsaftaris, S.; Yang, G.; Gray, C.; Newby, D.; Macnaught, G.; MacGillivray, T. TPSDicyc: Improved Deformation Invariant Cross-domain Medical Image Synthesis. In *International Workshop on Machine Learning for Medical Image Reconstruction*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 245–254.
21. Zhu, J.; Yang, G.; Ferreira, P.; Scott, A.; Nielles-Vallespin, S.; Keegan, J.; Pennell, D.; Lio, P.; Firmin, D. A ROI Focused Multi-Scale Super-Resolution Method for the Diffusion Tensor Cardiac Magnetic Resonance. In Proceedings of the 27th Annual Meeting (ISMRM), Montréal, QC, Canada, 11–16 May 2019; International Society for Magnetic Resonance in Medicine: Concord, CA, USA, 2019; p. 1.
22. Zhu, J.; Yang, G.; Lio, P. How can we make gan perform better in single medical image super-resolution? A lesion focused multi-scale approach. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 1669–1673.
23. Zhu, J.; Yang, G.; Lio, P. Lesion focused super-resolution. In *Medical Imaging 2019: Image Processing*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 10949, p. 109491L.
24. Chen, C.; Huang, J. Compressive sensing MRI with wavelet tree sparsity. In Proceedings of the Twenty-Sixth Annual Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–8 December 2012; pp. 1115–1123.
25. Chen, C.; Huang, J. Exploiting the wavelet structure in compressed sensing MRI. *Magn. Reson. Imaging* **2014**, *32*, 1377–1389. [[CrossRef](#)] [[PubMed](#)]
26. Cevher, V.; Duarte, M.F.; Hegde, C.; Baraniuk, R. Sparse signal recovery using Markov random fields. In Proceedings of the 23rd Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; pp. 257–264.
27. Pižurica, A.; Aelterman, J.; Bai, F.; Vanlooche, S.; Luong, Q.; Goossens, B.; Philips, W. On structured sparsity and selected applications in tomographic imaging. In *SPIE Conference on Wavelets and Sparsity XIV*; International Society for Optics and Photonics: Bellingham, WA, USA, 2011; Volume 8138, pp. 333–344.
28. Panić, M.; Aelterman, J.; Crnojević, V.; Pižurica, A. Compressed Sensing in MRI with a Markov Random Field prior for spatial clustering of subband coefficients. In Proceedings of the 24th European Signal Processing Conference, EUSIPCO 2016, Budapest, Hungary, 29 August– 2 September 2016.

29. Panić, M.; Aelterman, J.; Crnojević, V.; Pižurica, A. Sparse Recovery in Magnetic Resonance Imaging With a Markov Random Field Prior. *IEEE Trans. Med. Imaging* **2017**, *36*, 2104–2115. [[CrossRef](#)] [[PubMed](#)]
30. Afonso, M.V.; Bioucas-Dias, J.M.; Figueiredo, M.A. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.* **2011**, *20*, 681–695. [[CrossRef](#)]
31. Liu, Y.; Zhan, Z.; Cai, J.; Guo, D.; Chen, Z.; Qu, X. Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging. *IEEE Trans. Med. Imaging* **2016**, *35*, 2130–2140. [[CrossRef](#)]
32. Ravishankar, S.; Bresler, Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imaging* **2011**, *30*, 1028. [[CrossRef](#)]
33. Nesterov, Y.E. A method of solving a convex programming problem with convergence rate  $O(k^2)$ . In *Doklady Akademii Nauk*; Russian Academy of Sciences: Saint Petersburg, Russia, 1983; Volume 269, pp. 543–547.
34. Nesterov, Y. *Introductory Lectures on Convex Optimization: A Basic Course*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 87.
35. Cevher, V.; Indyk, P.; Carin, L.; Baraniuk, R.G. Sparse signal recovery and acquisition with graphical models. *IEEE Signal Process. Mag.* **2010**, *27*, 92–103. [[CrossRef](#)]
36. Li, S.Z. *Markov Random Field Modeling in Image Analysis*; Springer Science & Business Media: Berlin, Germany, 2009.
37. Huang, J.; Zhang, S.; Li, H.; Metaxas, D. Composite splitting algorithms for convex optimization. *Comput. Vis. Image Underst.* **2011**, *115*, 1610–1622. [[CrossRef](#)]
38. Kim, T.H.; Setsompop, K.; Haldar, J.P. LORAKS makes better SENSE: Phase-constrained partial fourier SENSE reconstruction without phase calibration. *Magn. Reson. Med.* **2017**, *77*, 1021–1035. [[CrossRef](#)] [[PubMed](#)]
39. Huang, J.; Yang, F. Compressed magnetic resonance imaging based on wavelet sparsity and nonlocal total variation. In Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 968–971.
40. Haldar, J.P.; Zhuo, J. P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data. *Magn. Reson. Med.* **2016**, *75*, 1499–1514. [[CrossRef](#)] [[PubMed](#)]
41. Winkelman, S.; Schaeffter, T.; Koehler, T.; Eggers, H.; Doessel, O. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans. Med. Imaging* **2007**, *26*, 68–76. [[CrossRef](#)] [[PubMed](#)]
42. Seitzer, M.; Yang, G.; Schlemper, J.; Oktay, O.; Würfl, T.; Christlein, V.; Wong, T.; Mohiaddin, R.; Firmin, D.; Keegan, J.; et al. Adversarial and perceptual refinement for compressed sensing MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 232–240.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).