

Plug-and-Play ADMM for MRI Reconstruction With Convex Nonconvex Sparse Regularization

JINCHENG LI, JINLAN LI, ZHAOYANG XIE, AND JIAN ZOU^{1b}

School of Information and Mathematics, Yangtze University, Jingzhou, Hubei 434020, China

Corresponding author: Jian Zou (zoujian@yangtzeu.edu.cn)

ABSTRACT Traditional ℓ_1 -regularized compressed sensing magnetic resonance imaging (CS-MRI) model tends to underestimate the fine textures and edges of the MR image, which play important roles in clinical diagnosis. In contrast, the convex nonconvex (CNC) strategy allows the use of nonconvex regularization while maintaining the convexity of the total objective function. Plug-and-play (PnP) algorithm is a powerful framework for sparse regularization problems, which plug any advanced denoiser into traditional proximal algorithms. In this paper, we propose a PnP-ADMM algorithm for CS-MRI reconstruction with CNC sparse regularization. We first obtain the proximal operator for CNC sparse regularization. Then we present PnP-ADMM algorithm by replacing the proximal operator of ADMM with properly pre-trained denoisers. Furthermore, we conduct comparison experiments using various denoisers under different sampling templates for different images. The experimental results verify the effectiveness of the proposed algorithm with both numerical criteria and visual effects.

INDEX TERMS Plug-and-play method, ADMM, convex nonconvex sparse regularization, compressed sensing, MRI reconstruction.

I. INTRODUCTION

MRI is one of the most powerful clinical diagnostic tools due to its non-invasive and non-ionizing characteristics. However, one of the primary limiting factors of MRI application is its relatively long image acquisition time. In order to reduce the scanning time, MRI data are usually undersampled. How to reconstruct MR images from undersampled measurements without degrading quality has been a concern [1]–[4].

CS-MRI is one of the most important breakthroughs in fast MR imaging [5]. Mathematically, for a single-coil CS-MRI model, the acquired K-space data $\mathbf{y} \in \mathbb{C}^m$ is given by

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^n$ is the ground truth MR image, $\boldsymbol{\eta} \in \mathbb{C}^m$ denotes measurement noise or disturbance, $\mathbf{A} \in \mathbb{C}^{m \times n}$ ($m < n$) represents the CS sampling matrix. Usually, we have $\mathbf{A} = \mathbf{P} \times \mathbf{F}$ where $\mathbf{P} \in \mathbb{R}^{m \times n}$ is an undersampling matrix, and $\mathbf{F} \in \mathbb{C}^{n \times n}$ is the Fourier transform.

The ground truth MR image \mathbf{x} can be obtained by solving the following regularized inverse problem

$$\min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{P}\mathbf{F}\mathbf{x}\|_2^2 + \lambda\phi(\mathbf{x}) \right\}, \quad (2)$$

The associate editor coordinating the review of this manuscript and approving it for publication was Ehab Elsayed Elattar^{1b}.

where the nonsmooth regularization term $\phi(\mathbf{x})$ is used to present some prior knowledge of \mathbf{x} , λ is the regularization parameter. The most used prior for MR image is sparsity in the wavelet domain, which can be included as the regularization term $\phi(\mathbf{x}) = \|\Psi\mathbf{x}\|_1$. Here, Ψ is a discrete wavelet transform, and the ℓ_1 norm is used to promote sparsity. The convexity of the ℓ_1 norm can ensure the efficient solvability of the objective function (2). However, ℓ_1 regularization is a biased estimation, which tends to underestimate the high-frequency part of \mathbf{x} in the wavelet domain. The high-frequency part of the wavelet domain corresponds to the details of the image, such as fine textures and edges, which is very important in clinical diagnosis. In contrast, existing nonconvex regularizers (such as ℓ_q norm $\phi(\mathbf{x}) = \|\Psi\mathbf{x}\|_q$ ($0 < q < 1$), Capped ℓ_1 [6], Smoothly Clipped Absolute Deviation (SCAD) [7] and Minimax Concave Penalty (MCP) [8], etc.) can yield a more accurate estimate of high-amplitude components. However, the nonconvex regularizer will make the objective function of the problem (2) nonconvex and will produce suboptimal local solutions [7]–[11].

Recently, Selesnick proposed a new CNC sparse regularizer formed by subtracting the Moreau envelope of ℓ_1 norm from itself [12]–[16]. The CNC sparse regularizer is nonconvex that can reduce the estimation bias. Meanwhile, the convexity of the whole objective function (2) can maintain by

adjusting the parameter settings. The CNC strategy is widely used in image reconstruction [17], [18], image denoising [12], [13], [19], and fault detection [20]–[22].

How to solve sparse regularization problem (2) is a challenging topic. PnP algorithm is a powerful framework for sparse regularization problems, which plug any advanced denoiser into traditional proximal algorithms. Ahmad *et al.* proposed PnP-ADMM for MRI reconstruction with ℓ_1 -regularization [3]. In this paper, we propose a novel PnP-ADMM algorithm to solve problem (2) with the CNC sparse regularization. The contributions of this paper are summarized as follows: Firstly, we propose the proximal operator for CNC sparse regularizer in iterative form. This iterative proximal operator can be used in other generalized CNC regularizers, which closed-form solution is challenging to obtain. Secondly, by replacing the proximal operator with properly pre-trained denoisers, we obtain the PnP-ADMM algorithm for the problem (2). Finally, we apply the proposed algorithm to MRI reconstruction. Experimental results demonstrate the superiority of the proposed algorithm.

The rest of this paper is organized as follows. In the next section, we present some basic concepts of the proposed method. In Section III, we introduce the CNC sparse regularizer and give its proximal operator in iterative form. ADMM and PnP-ADMM algorithms are proposed to solve the CNC sparse regularization model in Section IV. The experimental results are demonstrated in Section V. Finally, the conclusion is presented in Section VI.

A. RELATED WORK

MRI reconstruction algorithms can be roughly divided into three categories: Optimization-based, neural network-based and combination algorithms with optimization and neural network.

Traditional first-order iterative optimization algorithms, e.g. iterative shrinkage/thresholding algorithm (ISTA) [23], alternating direction method of multipliers (ADMM) [24], primal-dual hybrid gradient (PDHG) [25] and proximal gradient descent (PGD) [26], [27], have achieved great success in the past. Proximal operators are often used to deal with the nonsmooth regularization terms. Recent works have shown the effectiveness of the above algorithms when a simple closed-form solution of the proximal operator can be computed easily [28], [29]. However, the iterative algorithms may be hard to build when the proximal operator is not easy to evaluate. Furthermore, iterative algorithms often need hundreds of iterations to converge to an optimal solution.

In recent years, with the rapid rise of deep learning, neural networks are also used in MRI reconstruction. Neural networks, such as U-Net [30], [31] and generative adversarial network (GAN) [32], are used to learn the mapping between the k-space data and the ground truth MR image. The Network parameters are learned end-to-end. The main drawbacks of these methods are the huge network parameters and lack of interpretability.

Incorporating deep learning techniques into traditional optimization algorithms attracts increasing attention. Roughly, the so-called learning-based or data-driven algorithms implement in two ways. The first approach is algorithm unrolling (or unfolding), which unrolls an iterative algorithm to a deep neural network [33], [34]. The deep neural network architecture is constructed by mapping one iteration to a network layer and then stacking all layers together. The network parameters are learned from real-world training data sets through end-to-end training. Existing algorithm unrolling includes Learning ISTA (LISTA) [35], ADMM-Net [36], PGD-Net [37], Primal-Dual-Net (PD-Net) [38], etc.

The second approach is PnP, which plugs any advanced denoiser into traditional proximal algorithms. Existing PnP algorithms include PnP-ADMM, PnP-FBS, and PnP-ISTA, have shown their superiority in image reconstruction, image restoration, and other inverse imaging fields [3], [39]–[45]. Compare with algorithm unrolling, PnP has several advantages. First of all, when there is not enough data for end-to-end training, any traditional or pre-trained denoiser can be used as a modular component under the PnP framework. Commonly used denoisers include traditional hand-designed denoisers, such as soft threshold, Total Variation (TV) [46], Block-Matching 3D (BM3D) [47], and deep neural network denoisers [48]–[51]. Secondly, the convergence of PnP algorithms can guarantee theoretically under some general assumptions. For example, [39] used the bounded noise reducer hypothesis to analyze the convergence. Reference [40] studied the convergence of some PnP frameworks under a specific Lipschitz condition on the denoisers. Reference [41] proved the convergence of PnP algorithms under the assumption that the denoiser is nonexpansive.

II. PRELIMINARIES

In this section, we recall some basic concepts of proximal operator and monotone operator theory, which will be helpful in the rest of the paper. These concepts can be found in classical convex optimization literature [52], [53].

For a proper, closed, convex function $g(\mathbf{x})$, its proximal operator \mathbf{prox} is defined as

$$\mathbf{prox}_{\lambda g}(\mathbf{y}) = \arg \min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda g(\mathbf{x}) \right\}. \quad (3)$$

As a special case, if $g(\mathbf{x}) = \|\mathbf{x}\|_1$, then its proximal operator $\mathbf{prox}_{\lambda \|\cdot\|_1}$ is the element-wise soft threshold

$$\begin{aligned} \mathbf{prox}_{\lambda \|\cdot\|_1}(\mathbf{y})_i &= [|\mathbf{y}_i| - \lambda]_+ \mathbf{sgn}(\mathbf{y}_i) \\ &= \begin{cases} \mathbf{y}_i - \lambda, & \mathbf{y}_i \geq \lambda, \\ 0, & |\mathbf{y}_i| < \lambda, \\ \mathbf{y}_i + \lambda, & \mathbf{y}_i \leq -\lambda. \end{cases} \end{aligned} \quad (4)$$

where \mathbf{y}_i is the i -th element of \mathbf{y} . The soft threshold can be viewed as a denoiser since it can reduce the elements of \mathbf{y} towards zero.

For a general regularized model

$$\min_{\mathbf{x}} \{f(\mathbf{x}) + \lambda g(\mathbf{x})\}, \quad (5)$$

where $f(\mathbf{x})$ is a convex and differentiable data-fidelity term, $g(\mathbf{x})$ is a convex but nonsmooth regularization term.

By using the proximal operator $\mathbf{prox}_{\alpha g}$, PGD iteration for (5) is given by

$$\mathbf{x}^{k+1} = \mathbf{prox}_{\alpha g}(\mathbf{x}^k - \alpha \nabla f(\mathbf{x}^k)), \quad (6)$$

where $\nabla f(\mathbf{x})$ is the gradient of $f(\mathbf{x})$ and α is the step size.

The main iteration steps of ADMM for the problem (5) can be summarized as

$$\mathbf{x}^{k+1} = \mathbf{prox}_{\frac{1}{\beta} f} \left(\mathbf{z}^k - \frac{1}{\beta} \mathbf{u}^k \right), \quad (7)$$

$$\mathbf{z}^{k+1} = \mathbf{prox}_{\frac{1}{\beta} g} \left(\mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right), \quad (8)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \beta \left(\mathbf{x}^{k+1} - \mathbf{z}^{k+1} \right). \quad (9)$$

Since $f(\mathbf{x})$ is a convex and differentiable, subproblem (7) have analytic solutions. If the proximal operator $\mathbf{prox}_{\frac{1}{\beta} g}$ can be expressed in a simple closed-form, ADMM can solve (5) efficiently.

The key idea behind PnP-ADMM is to plug in a powerful pre-trained denoiser in place of the proximal operator $\mathbf{prox}_{\frac{1}{\beta} g}$. In this case, subproblem (8) becomes

$$\mathbf{z}^{k+1} = H_\sigma \left(\mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right), \quad (10)$$

where H_σ denotes a pre-trained denoiser with denoising parameter σ .

Figure 1 provides an intuitionistic illustration of ADMM and PnP-ADMM framework.

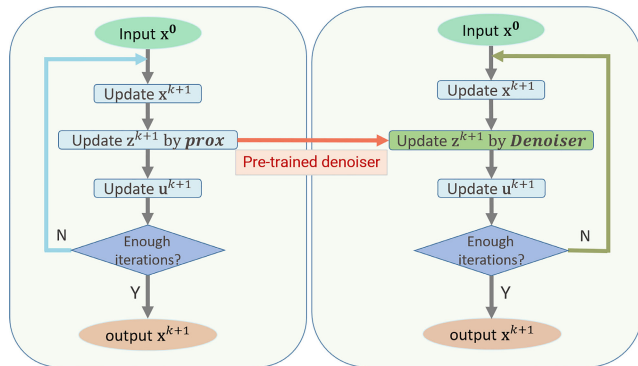


FIGURE 1. ADMM (left) and PnP-ADMM (right).

III. CNC SPARSE REGULARIZATION AND ITS PROXIMAL OPERATOR

Recently, Selesnick proposed a new CNC sparse regularizer formed by subtracting the Moreau envelope of ℓ_1 norm from itself [12]–[16]. To be specific, the CNC sparse regularizer is represented by

$$\phi_b(\mathbf{x}) = \|\mathbf{x}\|_1 - S_b(\mathbf{x}), \quad (11)$$

where $S_b(\mathbf{x}) = \min_{\mathbf{v}} \{ \|\mathbf{v}\|_1 + \frac{b^2}{2} \|\mathbf{x} - \mathbf{v}\|_2^2 \}$ is the Moreau envelope of $\|\mathbf{x}\|_1$, b is the scaling parameter that can control

the nonconvexity of $\phi_b(\mathbf{x})$. $\phi_b(\mathbf{x})$ is a nonconvex regularizer that can reduce the estimation bias. Meanwhile, the convexity of the whole objective function (2) can maintain by adjusting the parameter settings. Particularly, the objective function in (2) is convex if $b^2 \leq \frac{\mathbf{A}^T \mathbf{A}}{\lambda}$ and is strictly convex if $b^2 < \frac{\mathbf{A}^T \mathbf{A}}{\lambda}$ [14].

By the definition of the proximal operator, the proximal operator of the CNC sparse regularizer can be defined as

$$\mathbf{prox}_{\lambda \phi_b}(\mathbf{y}) = \arg \min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \phi_b(\mathbf{x}) \right\}. \quad (12)$$

The proximal operator of $\phi_b(\mathbf{x})$ may not have a simple explicit formula since $\phi_b(\mathbf{x})$ is nonconvex. However, we can still obtain the PGD iteration by rearranging the formula of the objective function.

The objective function of equation (12) can be expressed as

$$F(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 - \lambda S_b(\mathbf{x}). \quad (13)$$

As noted by Theorem 6.60 in [52], $S_b(\mathbf{x})$ is differentiable and its gradient is given by

$$\nabla S_b(\mathbf{x}) = b^2 \left[\mathbf{x} - \mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1}(\mathbf{x}) \right]. \quad (14)$$

Let $h(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 - \lambda S_b(\mathbf{x})$ and $g(\mathbf{x}) = \lambda \|\mathbf{x}\|_1$. $h(\mathbf{x})$ is also differentiable, and its gradient is

$$\nabla h(\mathbf{x}) = \mathbf{x} - \mathbf{y} - \lambda b^2 \left[\mathbf{x} - \mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1}(\mathbf{x}) \right]. \quad (15)$$

Furthermore, the proximal operator of $g(\mathbf{x}) = \lambda \|\mathbf{x}\|_1$ is the soft threshold. So the PGD iteration for (13) can express as

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{prox}_{\alpha g} \left(\mathbf{x}^k - \alpha \nabla h(\mathbf{x}^k) \right) \\ &= \mathbf{prox}_{\alpha \lambda \|\cdot\|_1} \left\{ (1 - \alpha) \mathbf{x}^k + \alpha \mathbf{y} \right. \\ &\quad \left. + \alpha \lambda b^2 \left[\mathbf{x}^k - \mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1}(\mathbf{x}^k) \right] \right\}. \end{aligned} \quad (16)$$

Finally, we can get the iterative expression of $\mathbf{prox}_{\lambda \phi_b}$ as follows

$$\mathbf{prox}_{\lambda \phi_b}(\mathbf{y}) = \mathbf{prox}_{\alpha \lambda \|\cdot\|_1} \left\{ (1 - \alpha) \mathbf{x}^k + \alpha \mathbf{y} \right. \\ \left. + \alpha \lambda b^2 \left[\mathbf{x}^k - \mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1}(\mathbf{x}^k) \right] \right\}. \quad (17)$$

We also give a numerical example to verify the validity of the iterative expression of $\mathbf{prox}_{\lambda \phi_b}$. When the variables in (12) reduce to scalar (one-dimensional) and $b^2 \leq 1/\lambda$, $\mathbf{prox}_{\lambda \phi_b}$ can be expressed as firm threshold function [10], [14], i.e.,

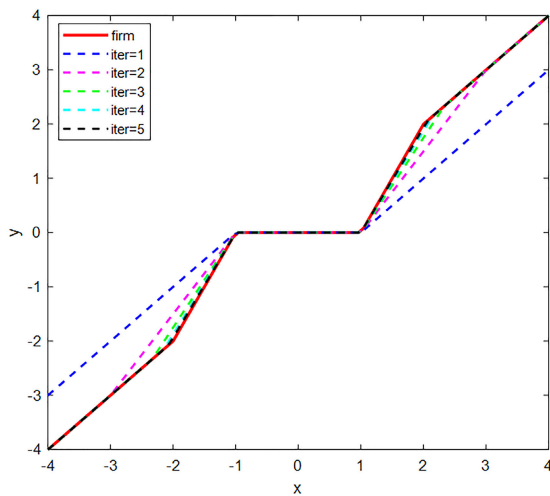
$$\mathbf{prox}_{\lambda \phi_b}(\mathbf{y}) = \mathbf{firm}(\mathbf{y}; \lambda, 1/b^2), \quad (18)$$

where

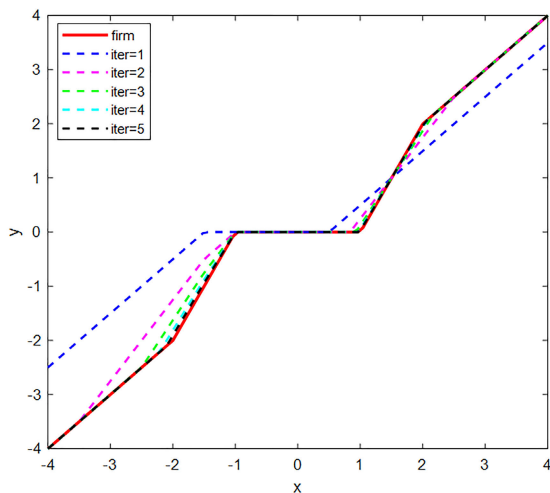
$$\mathbf{firm}(\mathbf{y}; \lambda, \mu) = \begin{cases} 0, & |\mathbf{y}| \leq \lambda, \\ \frac{\mu(|\mathbf{y}| - \lambda)}{\mu - \lambda} \text{sign}(\mathbf{y}), & \lambda \leq |\mathbf{y}| \leq \mu, \\ \mathbf{y}, & |\mathbf{y}| \geq \mu. \end{cases} \quad (19)$$

Firm threshold is a generalization of the hard threshold and soft threshold. When $\mu \rightarrow \lambda$ and $\mu \rightarrow \infty$, the firm threshold function is close to the hard threshold function and the soft threshold function, respectively.

We set $\alpha = 1$, $\lambda = 1$ and $b = 1/\sqrt{2}$, plot firm threshold function (18) and the iterative solutions (17) with different initial values $x^{(0)}$ in Figure 2. As shown in Figure 2, although the initial values are different, the iterative solutions converge to the firm threshold after a few iterations. This means that the iterative expression of the proximal operator is equally valid when the closed-form expression is difficult to obtain.



(a) $x^{(0)} = 0$



(b) $x^{(0)} = 1$

FIGURE 2. Firm threshold and iterative solutions with different initial values.

IV. PnP-ADMM ALGORITHM

A. ADMM ITERATION FOR CNC SPARSE REGULARIZATION

This subsection gives the ADMM algorithm for sparse regularization model (2) with CNC regularizer. In this case,

problem (2) can be expressed as

$$\min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{P}\mathbf{F}\mathbf{x}\|_2^2 + \lambda \phi_b(\Psi\mathbf{x}) \right\}, \quad (20)$$

while $\phi_b(\Psi\mathbf{x})$ is defined as (11).

Set $\mathbf{z} = \Psi\mathbf{x}$, then the augmented Lagrangian form of (20) is given by

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{u}) = & \frac{1}{2} \|\mathbf{y} - \mathbf{P}\mathbf{F}\mathbf{x}\|_2^2 + \lambda \phi_b(\mathbf{z}) + \mathbf{u}^T (\Psi\mathbf{x} - \mathbf{z}) \\ & + \frac{\beta}{2} \|\mathbf{z} - \Psi\mathbf{x}\|_2^2. \end{aligned} \quad (21)$$

According to ADMM iteration, the minimizer of function (21) can be found by solving the following sub-problems.

Step 1: Update \mathbf{x}^{k+1} with \mathbf{z}^k and \mathbf{u}^k fixed:

$$\begin{aligned} \mathbf{x}^{k+1} &= \arg \min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{P}\mathbf{F}\mathbf{x}\|_2^2 + (\mathbf{u}^k)^T \Psi\mathbf{x} + \frac{\beta}{2} \|\mathbf{z} - \Psi\mathbf{x}\|_2^2 \right\} \\ &= \arg \min_{\mathbf{x}} \left\{ \frac{\beta}{2} \left\| \Psi\mathbf{x} - \left(\mathbf{z}^k - \frac{1}{\beta} \mathbf{u}^k \right) \right\|_2^2 + \frac{1}{2} \|\mathbf{y} - \mathbf{P}\mathbf{F}\mathbf{x}\|_2^2 \right\} \\ &= \left(\mathbf{F}^T \mathbf{P}^T \mathbf{P}\mathbf{F} + \beta \Psi^T \Psi \right)^{-1} \left(\mathbf{F}^T \mathbf{P}^T \mathbf{y} + \beta \Psi^T \mathbf{z}^k - \Psi^T \mathbf{u}^k \right). \end{aligned} \quad (22)$$

Step 2: Update \mathbf{z}^{k+1} with \mathbf{x}^{k+1} and \mathbf{u}^k fixed:

$$\begin{aligned} \mathbf{z}^{k+1} &= \arg \min_{\mathbf{z}} \left\{ \lambda \phi_b(\mathbf{z}) - (\mathbf{u}^k)^T \mathbf{z} + \frac{\beta}{2} \|\mathbf{z} - \Psi\mathbf{x}^{k+1}\|_2^2 \right\} \\ &= \arg \min_{\mathbf{z}} \left\{ \frac{\beta}{2} \left\| \mathbf{z} - \left(\Psi\mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right) \right\|_2^2 + \lambda \phi_b(\mathbf{z}) \right\} \\ &= \mathbf{prox}_{\frac{\lambda}{\beta} \phi_b} \left(\Psi\mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right) \\ &= \mathbf{prox}_{\frac{\alpha\lambda}{\beta} \|\cdot\|_1} \left\{ (1 - \alpha)\mathbf{z}^k + \alpha \left(\Psi\mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right) \right. \\ & \quad \left. + \frac{\alpha\lambda b^2}{\beta} \left[\mathbf{z}^k - \mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1} (\mathbf{z}^k) \right] \right\}. \end{aligned} \quad (23)$$

The last equation is by definition of $\mathbf{prox}_{\lambda\phi_b}(\cdot)$ in 17.

Step 3: Update \mathbf{u}^{k+1} with \mathbf{x}^{k+1} and \mathbf{z}^{k+1} fixed:

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \beta \left(\Psi\mathbf{x}^{k+1} - \mathbf{z}^{k+1} \right). \quad (24)$$

Finally, by rearranging equations (22)-(24), we can get the ADMM algorithm for CNC sparse regularization (ADMM-CNC) as Algorithm 1.

B. PnP-ADMM ITERATION FOR CNC SPARSE REGULARIZATION

As mentioned in the section IV-A, the proximal operators $\mathbf{prox}_{\frac{1}{b^2} \|\cdot\|_1}$ and $\mathbf{prox}_{\alpha\rho\lambda \|\cdot\|_1}$ in (23) are soft thresholds with different parameters $1/b^2$ and $\alpha\rho\lambda$. Furthermore, they can be viewed as denoisers with varying abilities of denoising.

The basic idea behind the original PnP framework is to plug in powerful denoisers in place of the proximal operators. If we

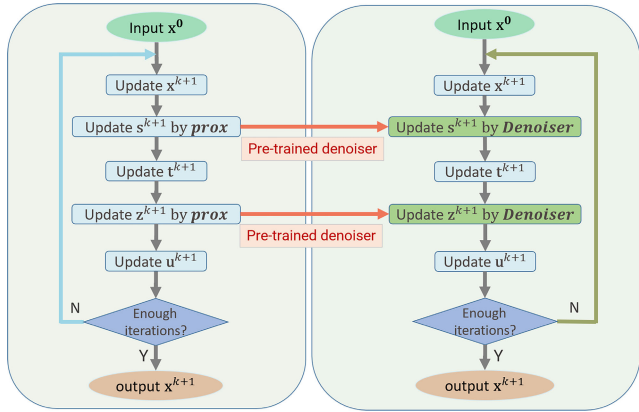


FIGURE 3. ADMM-CNC algorithm (left) and PnP-ADMM-CNC algorithm (right).

replace the proximal operators $\text{prox}_{\frac{1}{b^2} \|\cdot\|_1}$ and $\text{prox}_{\alpha\rho\lambda\|\cdot\|_1}$ with two appropriate denoisers H_{σ_1} and H_{σ_2} , where σ_1 and σ_2 are denoising parameters. The larger the value of σ_1 and σ_2 is, the greater the denoising intensity will be. Then (23) can be rewritten as

$$\mathbf{z}^{k+1} = H_{\sigma_2} \left\{ (1 - \alpha)\mathbf{z}^k + \alpha \left(\Psi \mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right) + \frac{\alpha\lambda b^2}{\beta} \left[\mathbf{z}^k - H_{\sigma_1}(\mathbf{z}^k) \right] \right\} \quad (25)$$

Algorithm 1 ADMM-CNC

Require: $f, \mathbf{x}^0, \alpha > 0, \lambda > 0, \beta > 0, b > 0$.

Ensure: \mathbf{x} .

while “stopping criterion is not met” **do**
 $\mathbf{x}^{k+1} = (\mathbf{F}^T \mathbf{P}^T \mathbf{P} \mathbf{F} + \beta \Psi^T \Psi)^{-1} \cdot (\mathbf{F}^T \mathbf{P}^T \mathbf{y} + \beta \Psi^T \mathbf{z}^k - \Psi^T \mathbf{u}^k);$
 $\mathbf{s}^{k+1} = \text{prox}_{\frac{1}{b^2} \|\cdot\|_1}(\mathbf{z}^k);$
 $\mathbf{t}^{k+1} = (1 - \alpha)\mathbf{z}^k + \alpha \left(\Psi \mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right) + \frac{\alpha\lambda b^2}{\beta} (\mathbf{z}^k - \mathbf{s}^{k+1});$
 $\mathbf{z}^{k+1} = \text{prox}_{\frac{\alpha\lambda}{\beta} \|\cdot\|_1}(\mathbf{t}^{k+1});$
 $\mathbf{u}^{k+1} = \mathbf{u}^k + \beta (\Psi \mathbf{x}^{k+1} - \mathbf{z}^{k+1}).$
end while

Finally, we can get plug-and-play ADMM algorithm for CNC sparse regularization (PnP-ADMM-CNC) as Algorithm 2. Figure 3 provides a intuitionistic illustration of ADMM-CNC and PnP-ADMM-CNC framework.

V. NUMERICAL EXPERIMENT

A. IMPLEMENTATION DETAILS

In this section, we empirically validate the proposed PnP-ADMM-CNC algorithm in the context of CS-MRI reconstruction. We consider 15 widely used MR images and three sampling templates, as shown in Figure 4 and Figure 5.

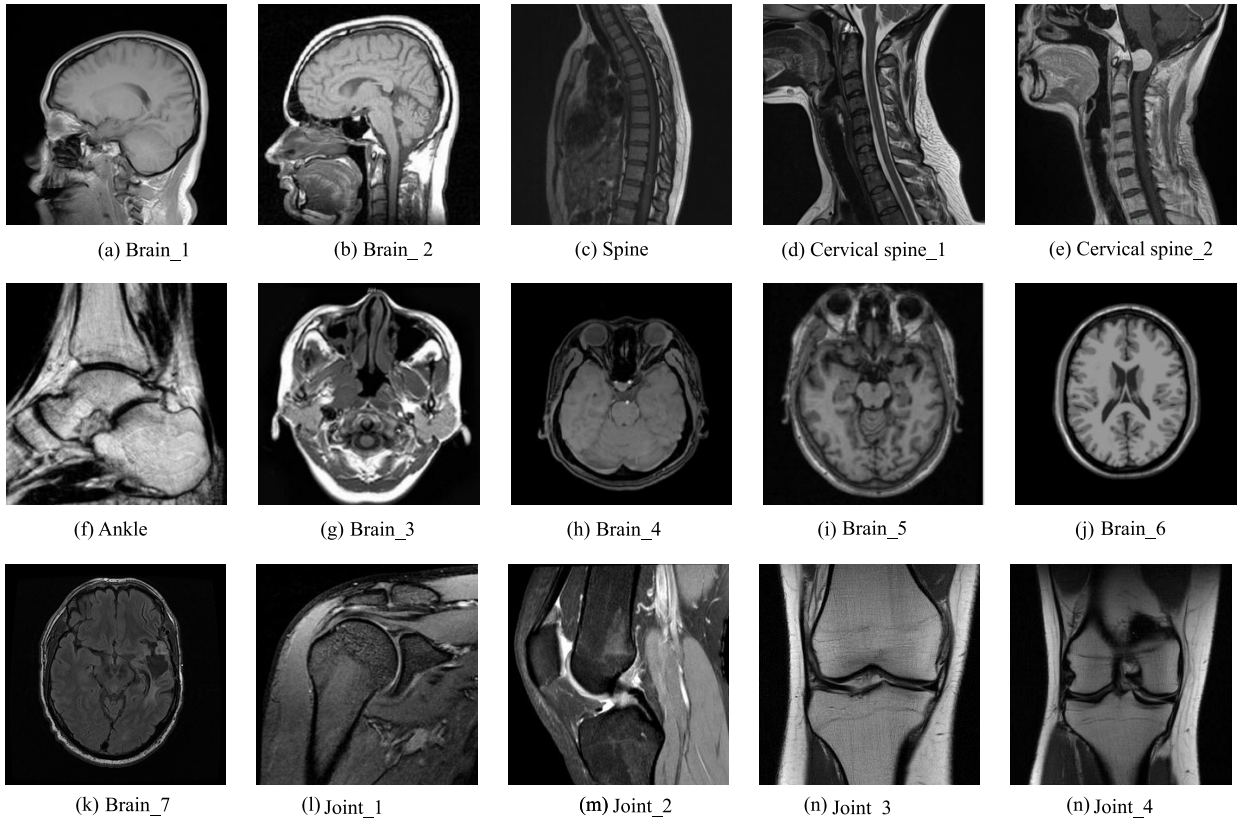


FIGURE 4. MR images.

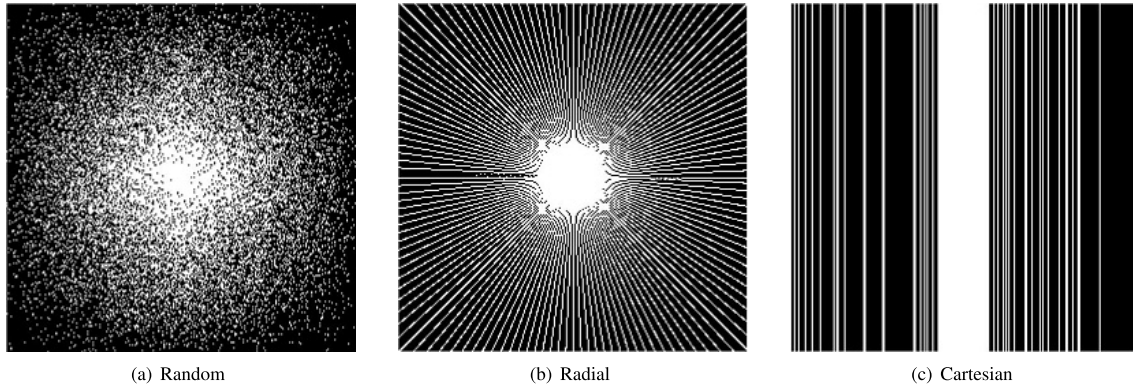


FIGURE 5. Sampling templates.

TABLE 1. Denoisers under the PnP framework.

| Denoisers | Brief Description |
|-----------|---|
| BM3D | Sparse 3D transform-domain collaborative filtering, filtering by finding similar blocks in the image. |
| DnCNN | Learned the residual mapping of the 17-layer convolutional neural network and has Batch Normalization (BN) modules. |
| FDnCNN | Compared with DnCNN, it lacks residual learning and BN modules. |
| IRCNN | With HQS, the 7-layer denoising network with dilated convolution is introduced. |
| FFDNet | Adjustable noise level map as input, learning 15 layers of convolutional neural network. |
| DRUNet | Combination of U-Net and ResNet, the network structure is symmetrical. |

TABLE 2. Quantitative results of different algorithms. The best results are highlighted in bold.

| | Random | | | Radial | | | Cartesian | | | |
|--------------------|--------|--------------|---------------|--------------|--------------|---------------|--------------|--------------|---------------|--------------|
| | RE | PSNR | SSIM | RE | PSNR | SSIM | RE | PSNR | SSIM | |
| Zero-Filling | 0.218 | 20.641 | 0.438 | 0.214 | 20.878 | 0.435 | 0.240 | 20.253 | 0.392 | |
| BM3D-MRI | 0.186 | 24.441 | 0.610 | 0.140 | 24.239 | 0.572 | 0.210 | 23.556 | 0.578 | |
| ADMM- ℓ_1 | 0.193 | 24.259 | 0.567 | 0.197 | 24.178 | 0.561 | 0.223 | 23.112 | 0.518 | |
| ADMM-CNC | 0.182 | 24.768 | 0.501 | 0.185 | 24.636 | 0.492 | 0.210 | 23.535 | 0.452 | |
| U-Net | 0.106 | 29.290 | 0.758 | 0.111 | 29.062 | 0.743 | 0.120 | 28.354 | 0.712 | |
| PD-Net | 0.100 | 30.042 | 0.751 | 0.105 | 29.497 | 0.733 | 0.115 | 28.764 | 0.722 | |
| PnP-ADMM- ℓ_1 | BM3D | 0.191 | 24.697 | 0.616 | 0.199 | 24.387 | 0.605 | 0.210 | 23.887 | 0.580 |
| | FDnCNN | 0.125 | 28.132 | 0.720 | 0.129 | 27.802 | 0.711 | 0.150 | 26.505 | 0.671 |
| | DnCNN | 0.122 | 28.299 | 0.722 | 0.126 | 28.044 | 0.713 | 0.144 | 26.876 | 0.681 |
| | FFDNet | 0.114 | 28.964 | 0.736 | 0.117 | 28.554 | 0.727 | 0.140 | 27.123 | 0.692 |
| | IRCNN | 0.112 | 29.127 | 0.788 | 0.115 | 28.787 | 0.782 | 0.135 | 27.284 | 0.708 |
| | DRUNet | 0.110 | 29.275 | 0.751 | 0.112 | 28.892 | 0.740 | 0.130 | 27.749 | 0.712 |
| PnP-ADMM-CNC | BM3D | 0.151 | 26.549 | 0.669 | 0.157 | 26.308 | 0.657 | 0.169 | 25.699 | 0.633 |
| | FDnCNN | 0.111 | 29.117 | 0.728 | 0.116 | 28.629 | 0.716 | 0.134 | 27.281 | 0.679 |
| | DnCNN | 0.110 | 29.252 | 0.731 | 0.112 | 28.668 | 0.717 | 0.136 | 27.548 | 0.685 |
| | FFDNet | 0.099 | 30.097 | 0.755 | 0.102 | 29.511 | 0.739 | 0.120 | 28.120 | 0.706 |
| | IRCNN | 0.097 | 30.393 | 0.772 | 0.101 | 29.839 | 0.758 | 0.118 | 28.332 | 0.722 |
| | DRUNet | 0.094 | 30.612 | 0.758 | 0.097 | 30.156 | 0.746 | 0.110 | 29.114 | 0.716 |

The grey value range of all MR images is [0, 1], and the size is set to 256×256 , the undersampling rate is 30%, the measurement noise $\eta_e \sim N(0, \sigma_e \mathbf{I}_k)$ with noise level $\sigma_e = 15/255$. All experiments are executed on a PC with Intel(R) Core(TM) i7-8750HM, 2.20GHz CPU, 16GB memory, and NVIDIA GeForce GTX 1060.¹

We compared the proposed PnP-ADMM-CNC algorithm with different kinds of MRI reconstruction methods: (1) Zero-Filling, which simply applies the inverse Fourier

¹The source code is available at https://github.com/zj15001/PnP_ADMM_CNC_MRI

transform with the unsampled K-space data filled with zeros. (2) BM3D-MRI, which uses decoupled iterations alternating over a BM3D denoising step [54]. (3) U-Net, the state-of-the-art network architecture for end-to-end MRI reconstruction [31]. (4) PD-Net, the state-of-the-art unrolling algorithm for MRI reconstruction [38]. (5) ADMM- ℓ_1 and ADMM-CNC, traditional iterative algorithms for ℓ_1 and CNC sparse regularization, respectively. Additionally, U-Net and PD-Net are trained on fastMRI dataset with Cartesian under-sampling schemes, which training details can be found in [4].

It has been shown that the choice of denoiser affects the performance of the PnP algorithms [43], [44]. For

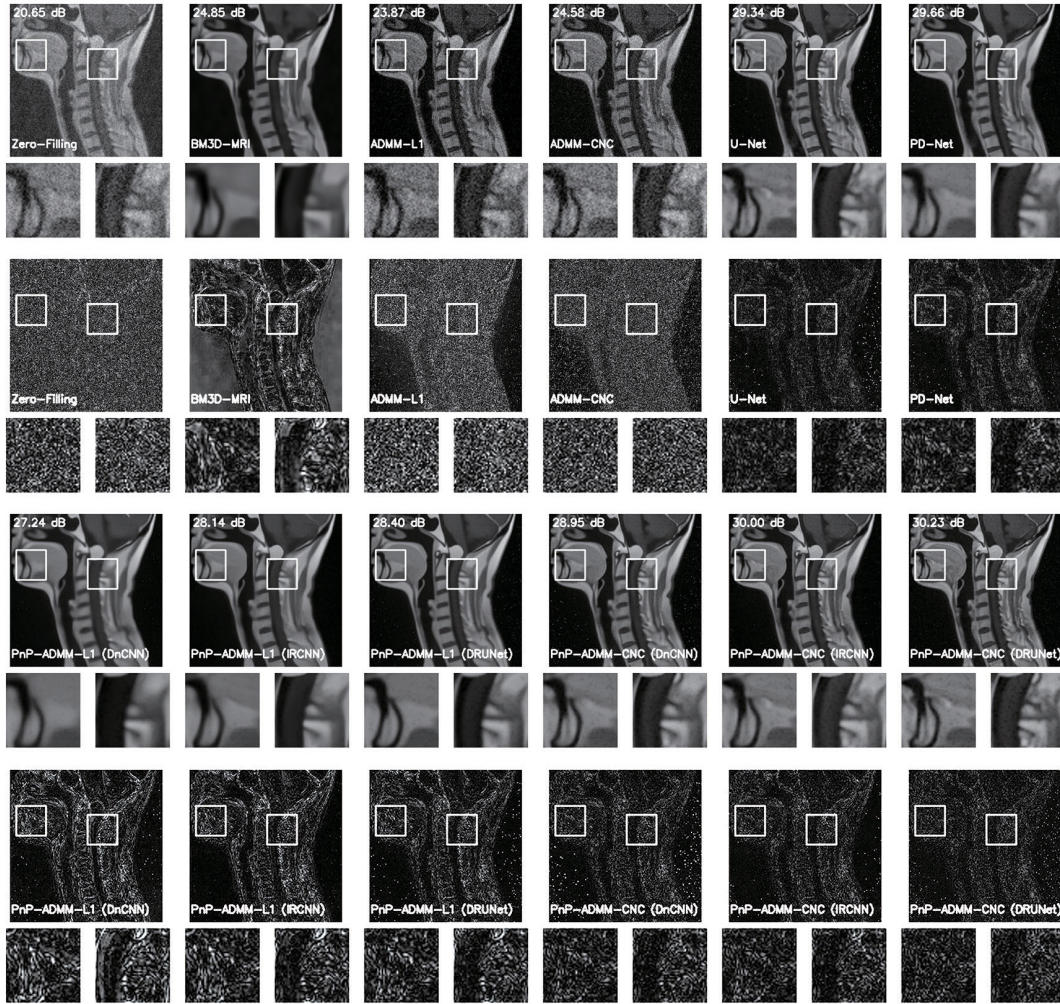


FIGURE 6. Visual comparison results of different algorithms on Cervical spine_2 MR image with random sampling template. The first and third rows show the reconstructed images. The second and fourth rows show the difference between the ground truth image and the reconstructed images.

Algorithm 2 PnP-ADMM-CNC

Require: $f, \mathbf{x}^0, \alpha > 0, \lambda > 0, \beta > 0, b > 0$.

Ensure: \mathbf{x} .

```

while “stopping criterion is not met” do
 $\mathbf{x}^{k+1} = (\mathbf{F}^T \mathbf{P}^T \mathbf{P} \mathbf{F} + \beta \Psi^T \Psi)^{-1}$ 
 $\cdot (\mathbf{F}^T \mathbf{P}^T \mathbf{y} + \beta \Psi^T \mathbf{z}^k - \Psi^T \mathbf{u}^k);$ 
 $\mathbf{s}^{k+1} = H_{\sigma_1}(\mathbf{z}^k);$ 
 $\mathbf{t}^{k+1} = (1 - \alpha)\mathbf{z}^k + \alpha \left( \Psi \mathbf{x}^{k+1} + \frac{1}{\beta} \mathbf{u}^k \right)$ 
 $+ \frac{\alpha \lambda b^2}{\beta} (\mathbf{z}^k - \mathbf{s}^{k+1});$ 
 $\mathbf{z}^{k+1} = H_{\sigma_2}(\mathbf{t}^{k+1});$ 
 $\mathbf{u}^{k+1} = \mathbf{u}^k + \beta (\Psi \mathbf{x}^{k+1} - \mathbf{z}^{k+1}).$ 
end while
    
```

PnP-ADMM- ℓ_1 and PnP-ADMM-CNC, we consider several state-of-the-art denoiser, including model-based denoiser BM3D [47] and five neural network-based denoisers (DnCNN [48], FDnCNN [50], IRCNN [49], FFDNet [50] and DRUNet [51]) are used as denoisers under the

PnP framework. Table 1 lists the characteristics of the above denoisers briefly. As shown in [44], DRUNet achieves superior performance against other deep neural network denoisers. In particular, ADMM-CNC and ADMM- ℓ_1 can also be considered as special PnP algorithms if we view the soft threshold function as a special denoiser.

As mentioned in [40] and [55], PnP method has high generalizability due to its plug-and-play nature. It does not even require problem-specific training. Neural network denoisers that learn from natural images could also obtain satisfactory results in MRI reconstruction and seismic interpolation. In our experiment, we also do not train the denoisers on the MRI image dataset. The five neural network-based denoisers used in the experiment are downloaded from <https://github.com/cszn/KAIR>. These denoisers are trained on natural images datasets, e.g., BSD dataset [56], DIV2K dataset [57] and Flick2K dataset [58].

Furthermore, we only consider two proximal operators to replace with the same kind of denoiser in the experiment of this paper. As mentioned in [44], DnCNN needs

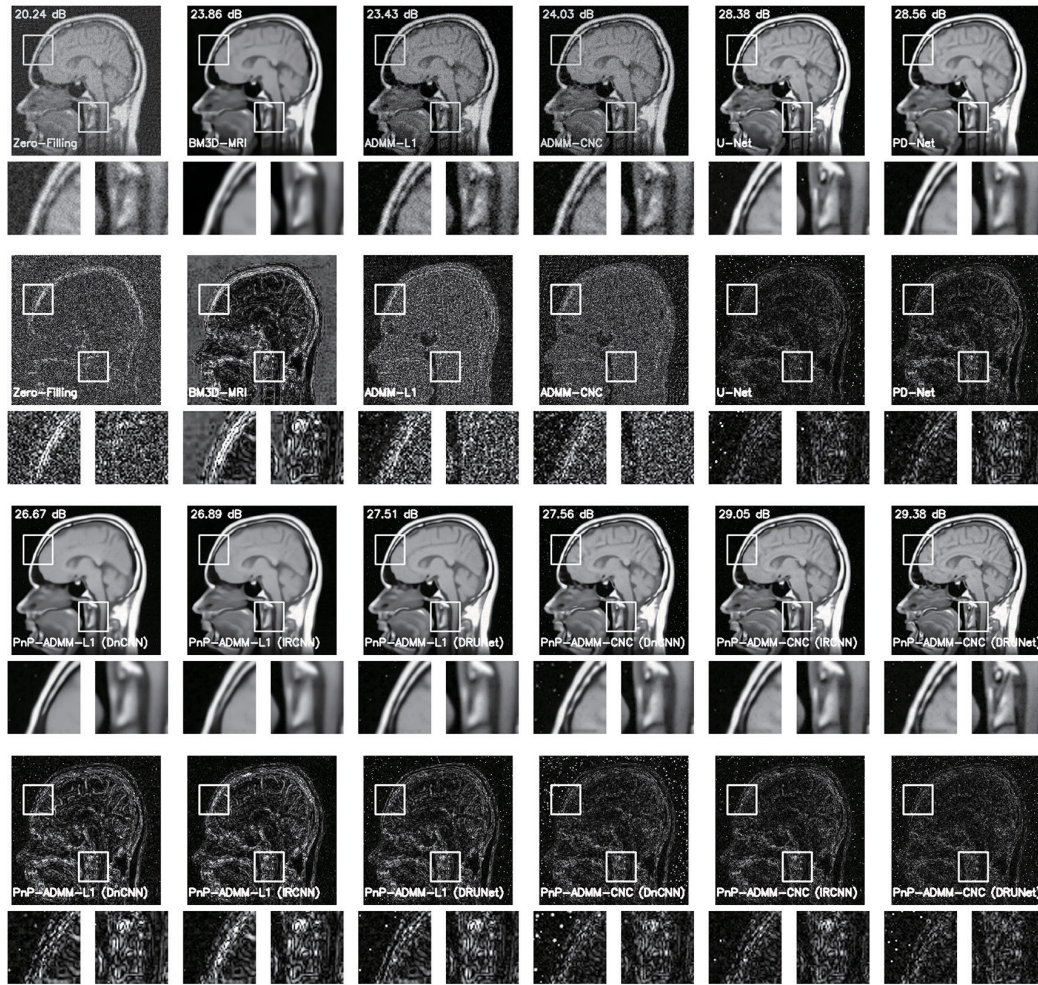


FIGURE 7. Visual comparison results of different algorithms on Brain_2 MR image with radial sampling template. The first and third rows show the reconstructed images. The second and fourth rows show the difference between the ground truth image and the reconstructed images.

to learn a single model for each noise level separately, and other denoisers can deal with various noise levels via a single model. Therefore, we set denoising parameters $\sigma_1 = 5\sigma_\epsilon/3 = 25/255$, $\sigma_2 = \sigma_\epsilon = 15/255$ for DnCNN and $\sigma_1 = \sigma_2 = \sigma_\epsilon = 15/255$ for the other four neural network-based denoisers. Other algorithmic hyperparameters, including the step size α , regularization parameter λ and nonconvexity control parameter b , are properly tuned for the best reconstruction performance with respect to the ground truth MR image.

B. EXPERIMENTAL EVALUATION CRITERIA

To evaluate the quality of the reconstruction, relative error (RE), peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used.

RE is defined as

$$RE = \frac{\|\mathbf{x} - \mathbf{x}_0\|_2}{\|\mathbf{x}_0\|_2}, \quad (26)$$

where \mathbf{x} and \mathbf{x}_0 are the reconstructed and the ground truth image, respectively.

PSNR is defined as

$$PSNR = 10 \cdot \log_{10} \left(\frac{\text{MAX}_{\mathbf{x}_0}^2}{\text{MSE}} \right), \quad (27)$$

where $\text{MSE} = \frac{1}{mn} \|\mathbf{x} - \mathbf{x}_0\|_2^2$, and $\text{MAX}_{\mathbf{x}_0}$ is the maximum value of the ground truth MR image data.

RE and PSNR are sometimes mismatched to perceive visual quality. In contrast, SSIM measures the structural similarity of two images, which is defined as

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (28)$$

where μ_x and μ_y are the means of \mathbf{x} and \mathbf{y} , σ_x and σ_y are the variances of \mathbf{x} and \mathbf{y} , σ_{xy} is the covariance of \mathbf{x} and \mathbf{y} , c_1 and c_2 are constants that maintain stability. Lower RE, higher PSNR and SSIM mean better quantitative performance.

C. IMPLEMENTATION RESULTS

We first compare different algorithms on 15 MR images with three sampling templates, as shown in Figure 4 and 5. The

average values of RE, PSNR, and SSIM are summarized in Table 2. As one can see, the algorithms for CNC regularization outperform the algorithms for regularization since CNC regularization can reduce the reconstruction bias. On the other hand, PnP-PGD significantly outperforms traditional PGD since more powerful denoisers are embedded into the PnP framework. In addition, DRUNet has achieved the best performance in most of the evaluation criteria among all the advanced denoisers. Specifically, compared with Zero-Filling, the average PSNR and SSIM of PnP-PGD-CNC with DRUNet denoiser are increased by 9.37 dB and 0.32, and the average RE is reduced by 0.12. U-Net and PD-Net perform well under Cartesian sampling, but not so well under other sampling since they are trained with Cartesian sampling schemes.

To get a more intuitive understanding of the reconstruction performance, we also design two visual comparisons of different algorithms on Cervical spine_2 (Figure 4(e)) and Brain_2 (Figure 4(b)) MR image with the random sampling template (Figure 5(a)) and the radial sampling template (Figure 5(b)), respectively. The reconstruction results are shown in Figure 6 and 7. To show the difference more clearly, we enlarge two local regions, display the PSNR values on the reconstructed images and set the grey value range of the difference images to [0,0.2]. From Figure 6 and 7, it can be seen that PnP-PGD-CNC with the DRUNet denoiser can reconstruct more fine details and contrasts, which leads to the lower error level around edges and high-frequency areas.

VI. CONCLUSION

In this work, we propose a novel PnP algorithm for CNC sparse regularization. CNC sparse regularization can reduce the reconstruction bias and ensure the global convexity of the objective function. PnP algorithm is obviously superior to the traditional optimization-based algorithm, which is mainly benefited from the powerful pre-trained denoisers. Experiments results have demonstrated that the proposed PnP-ADMM-CNC algorithm with a powerful deep denoiser can improve the reconstruction result.

In this work we only consider a single-coil CS-MRI model with simulated MR data. As parallel MRI can lead higher image quality, it is important to apply the proposed algorithm on multi-coil CS-MRI model with acquired MR data. we will employ the proposed algorithm to the practical MR application in our future work. In addition, we would like to extend the proposed algorithm to other tasks, such as seismic interpretation [55] and Electrical impedance tomography [59].

REFERENCES

- [1] F. Knoll, K. Hammernik, C. Zhang, S. Moeller, and M. Akcakaya, "Deep learning methods for parallel magnetic resonance image reconstruction," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 128–140, Jan. 2020.
- [2] D. Liang, J. Cheng, Z. Ke, and L. Ying, "Deep magnetic resonance image reconstruction: Inverse problems meet neural networks," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 141–151, Jan. 2020.
- [3] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, S. Liu, E. T. Reehorst, and P. Schniter, "Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 105–116, Jan. 2020.
- [4] Z. Ramzi, P. Ciuciu, and J.-L. Starck, "Benchmarking MRI reconstruction neural networks on large public datasets," *Appl. Sci.*, vol. 10, no. 5, p. 1816, 2020.
- [5] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
- [6] T. Zhang, "Analysis of multi-stage convex relaxation for sparse regularization," *J. Mach. Learn. Res.*, vol. 11, no. 3, pp. 1081–1107, 2010.
- [7] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its Oracle properties," *J. Amer. Statist. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.
- [8] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 2010.
- [9] Z. Xu, X. Chang, F. Xu, and H. Zhang, " $L_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.
- [10] J. Woodworth and R. Chartrand, "Compressed sensing recovery via non-convex shrinkage penalties," *Inverse Problems*, vol. 32, no. 7, Jul. 2016, Art. no. 075004.
- [11] F. Wen, L. Chu, P. Liu, and R. C. Qiu, "A survey on nonconvex regularization-based sparse and low-rank recovery in signal processing, statistics, and machine learning," *IEEE Access*, vol. 6, pp. 69883–69906, 2018.
- [12] I. Selesnick, "Total variation denoising via the Moreau envelope," *IEEE Signal Process. Lett.*, vol. 24, no. 2, pp. 216–220, Feb. 2017.
- [13] A. Lanza, S. Morigi, and F. Sgallari, "Convex image denoising via non-convex regularization with parameter selection," *J. Math. Imag. Vis.*, vol. 56, no. 2, pp. 195–220, Oct. 2016.
- [14] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4481–4494, Sep. 2017.
- [15] I. Selesnick, A. Lanza, S. Morigi, and F. Sgallari, "Non-convex total variation regularization for convex denoising of signals," *J. Math. Imag. Vis.*, vol. 62, pp. 825–841, Jan. 2020.
- [16] A. Lanza, S. Morigi, I. W. Selesnick, and F. Sgallari, "Sparsity-inducing nonconvex nonseparable regularization for convex image processing," *SIAM J. Imag. Sci.*, vol. 12, no. 2, pp. 1099–1134, Jan. 2019.
- [17] Y. Liu, H. Du, Z. Wang, and W. Mei, "Convex MR brain image reconstruction via non-convex total variation minimization," *Int. J. Imag. Syst. Technol.*, vol. 28, no. 4, pp. 246–253, Dec. 2018.
- [18] M. Shen, J. Li, T. Zhang, and J. Zou, "Magnetic resonance imaging reconstruction via non-convex total variation regularization," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 1, pp. 412–424, 2020.
- [19] J. Zou, M. Shen, Y. Zhang, H. Li, G. Liu, and S. Ding, "Total variation denoising with non-convex regularizers," *IEEE Access*, vol. 7, pp. 4422–4431, 2018.
- [20] S. Wang, I. Selesnick, G. Cai, Y. Feng, X. Sui, and X. Chen, "Non-convex sparse regularization and convex optimization for bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7332–7342, Sep. 2018.
- [21] Z. Diwu, H. Cao, L. Wang, and X. Chen, "Collaborative double sparse period-group lasso for bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
- [22] W. Huang, N. Li, I. Selesnick, J. Shi, J. Wang, L. Mao, X. Jiang, and Z. Zhu, "Nonconvex group sparsity signal decomposition via convex optimization for bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4863–4872, Jul. 2020.
- [23] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [25] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.
- [26] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.

- [27] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Exact worst-case convergence rates of the proximal gradient method for composite convex minimization," *J. Optim. Theory Appl.*, vol. 178, no. 2, pp. 455–476, Aug. 2018.
- [28] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.
- [29] N. G. Polson, J. G. Scott, and B. T. Willard, "Proximal algorithms in statistics and machine learning," *Stat. Sci.*, vol. 30, no. 4, pp. 559–581, 2015.
- [30] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2015, pp. 234–241.
- [31] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.*, vol. 63, no. 13, Jun. 2018, Art. no. 135007.
- [32] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1488–1497, Jun. 2018.
- [33] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.
- [34] H. Li, Y. Yang, D. Chen, and Z. Lin, "Optimization algorithm inspired deep neural network structure design," in *Proc. Asian Conf. Mach. Learn.*, 2018, pp. 614–629.
- [35] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, 2010, pp. 399–406.
- [36] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 10–18.
- [37] M. Mardani, Q. Sun, D. Donoho, V. Pappas, H. Monajemi, S. Vasanawala, and J. Pauly, "Neural proximal gradient descent for compressive imaging," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, pp. 9573–9583.
- [38] J. Cheng, H. Wang, L. Ying, and D. Liang, "Model learning: Primal dual networks for fast mr imaging," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2019, pp. 21–29.
- [39] S. H. Chan, X. Wang, and O. A. Elgandy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 84–98, Jan. 2017.
- [40] E. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin, "Plug-and-play methods provably converge with properly trained denoisers," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 5546–5557.
- [41] Y. Sun, B. Wohlberg, and U. S. Kamilov, "An online plug-and-play algorithm for regularized image reconstruction," *IEEE Trans. Comput. Imag.*, vol. 5, no. 3, pp. 395–408, Sep. 2019.
- [42] Y. Sun, Z. Wu, X. Xu, B. Wohlberg, and U. Kamilov, "Scalable plug-and-play ADMM with convergence guarantees," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 849–863, 2021.
- [43] K. Wei, A. Aviles-Rivero, J. Liang, Y. Fu, C.-B. Schönlieb, and H. Huang, "Tuning-free plug-and-play proximal algorithm for inverse imaging problems," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 10158–10169.
- [44] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jun. 14, 2021, doi: 10.1109/TPAMI.2021.3088914.
- [45] T. Chen, X. Chen, W. Chen, H. Heaton, J. Liu, Z. Wang, and W. Yin, "Learning to optimize: A primer and a benchmark," 2021, *arXiv:2103.12828*.
- [46] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [47] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [48] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [49] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3929–3938.
- [50] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
- [51] M. Jafari, D. Auer, S. Francis, J. Garibaldi, and X. Chen, "DRU-Net: An efficient deep convolutional neural network for medical image segmentation," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1144–1148.
- [52] A. Beck, *First-Order Methods in Optimization*. Philadelphia, PA, USA: SIAM, 2017.
- [53] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York, NY, USA: Springer, 2011.
- [54] E. M. Eksioğlu, "Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI," *J. Math. Imag. Vis.*, vol. 56, no. 3, pp. 430–440, Nov. 2016.
- [55] H. Zhang, X. Yang, and J. Ma, "Can learning from natural image denoising be used for seismic data interpolation?" *Geophysics*, vol. 85, no. 4, pp. 1–142, 2020.
- [56] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2, Jul. 2001, pp. 416–423.
- [57] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1122–1131.
- [58] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1132–1140.
- [59] S. Liu, J. Jia, Y. D. Zhang, and Y. Yang, "Image reconstruction in electrical impedance tomography based on structure-aware sparse Bayesian learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 9, pp. 2090–2102, Sep. 2018.



JINCHENG LI received the B.S. degree in mathematics and applied mathematics from China West Normal University, in 2019. She is currently pursuing the master's degree with the School of Information and Mathematics, Yangtze University. Her research interests include optimization algorithms and digital image processing.



JINLAN LI received the B.S. degree in information and computing science from Yangtze University, in 2020, where she is currently pursuing the master's degree with the School of Information and Mathematics. Her research interests include optimization algorithms and compressed sensing.



ZHAOYANG XIE received the B.S. degree in information and computer science from Wuhan Textile University, in 2020. He is currently pursuing the master's degree with the School of Information and Mathematics, Yangtze University. His research interests include image processing and machine learning.



JIAN ZOU was born in Hubei, China, in 1983. He received the M.S. degree in applied mathematics from Yangtze University, in 2007, and the Ph.D. degree in signal and information processing from the South China University of Technology, in 2013. He is currently an Associate Professor with Yangtze University. His research interests include adaptive signal processing and numerical optimization.

...