Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 19, No. 1, pp. 141–159, 2023

PRIMAL AND DUAL ALTERNATING DIRECTION METHODS OF MULTIPLIERS FOR COMPRESSIVE SENSING IMAGE RECONSTRUCTION CORRUPTED BY IMPLUSIVE NOISE*

Chenchen Lian, Yanyun Ding and Yunhai Xiao†

Abstract: It is known that, the compressive sensing theories offered the possibilities of accurately reconstructing images from highly undersampled data and simultaneously correcting the possible noise. It is also known that, if the undersampled data is corrupted by white Gaussian noise, the state-of-the-art solver RecPF can be employed successfully, but for impulsive noise, it is not quite suitable. As a remedy, in this paper, we concentrate on the impulsive noise case, and particularly focus on the cost function being the sum of a total variation regularized term, an ℓ_1-norm regularized term, and an ℓ_1-norm measured data fidelity term. However, this cost function cause a little more challenges for minimizing because of these non-differentiable terms. To tackle this difficulty, this paper presents a pair of efficient algorithms from two different aspects: (1) employing an alternating direction method of multipliers (ADMM) to solve the primal problem in a straightforward way; (2) proposing an ADMM to the dual problem whose objective function contains four blocks of variables and three blocks of non-differentiable terms. In dual cases, a symmetric Gauss-Seidel technique is employed to decompose the involved bigger subproblem into some smaller ones. It should be emphasized that the most remarkable feature of our proposed algorithms is that each subproblem is easily implementable by making full use of the favorable structures, such as the fast Fourier transforms, the proximal mapping and the Moreau decomposition of ℓ_1-norm function. We do extensive numerical simulations using some magnetic resonance images which demonstrate that the algorithm based on dual model is evidently efficient.

Key words: *total variation, alternating direction method of multipliers, symmetric Gauss-Seidel, Lagrangian dual problem, optimality conditions*

Mathematics Subject Classification: *90C26, 90C90*

1 Introduction

Image reconstruction plays a crucial part in diverse applications, especially in magnetic resonance imaging (MRI). Let $\bar{u} = (u_1, u_2, \ldots, u_n)^\top \in \mathbb{R}^n$ be an unknown two-dimensional grayscale digital image (with n pixels) expected to possess piecewise continuous behavior. As a priori information, we assume that $\bar{u}$ is compressible or (approximately) sparse under a suitable basis (e.g. Fourier or wavelet basis). Let $\Psi = (\psi_1, \psi_2, \ldots, \psi_n) \in \mathbb{C}^{n\times n}$ be a certain orthogonal basis, i.e., $\bar{u} = \Psi\bar{x}$ or $\bar{x} = \Psi^*\bar{u}$, where Ψ^* represents the conjugate transpose of Ψ, then $\bar{u}$ is said to be K-sparse under Ψ if the number of nonzero entries in $\bar{x}$ is K, and $\bar{u}$

*This work is supported by the National Natural Science Foundation of China (No. 11971149).

†Corresponding author. The work of Y. Xiao is supported the National Natural Science Foundation of China (Grant No. 11971149).

is said to be compressible if $\bar{x}$ has only a few large (in magnitude) components. The idea of compressive sensing (CS) [4, 5, 8] in data acquisition and reconstruction is firstly projecting $\bar{u}$ onto a certain subspace via a linear operator Φ, i.e, $b = \Phi\bar{u} = \Phi\Psi\bar{x}$ with $b \in \mathbb{C}^m$ and $m < n$, and then reconstructing $\bar{x}$ (or equivalently $\bar{u}$) from the undersampled measurement b. For Φ, we focus on the partial (undersampled) Fourier transform in the form of $\Phi = \mathcal{P}\mathcal{F}$, where $\mathcal{F} \in \mathbb{C}^{n\times n}$ is a specific two-dimensional Fourier transform matrix, and $\mathcal{P} \in \mathbb{R}^{m\times n}$ be a selection matrix containing m rows of the identity matrix of order n. Besides, in the process of acquiring, storing, transmitting, or displaying, the undersampled data may be inevitably degenerated by noise, i.e.,

$$b = \Phi\bar{u} + \omega, \tag{1.1}$$

where $\omega \in \mathbb{C}^m$ represents noise, e.g., the impulsive noise and white Gaussian noise.

It is noted that the MR images have piecewise continuous structures, which indicates there exists a total finite difference operator $D = [D_1; \ldots; D_n] \in \mathbb{R}^{2n\times n}$ with a local finite difference operator $D_i \in \mathbb{R}^{2\times n}$ such that $D_i u \in \mathbb{R}^2$ is a discrete gradient of u at pixel i, and $Du \in \mathbb{R}^{2n}$ contains an overwhelming majority of zeros. Traditionally, the sum $\sum_i^n \|D_i u\|_2$ is named the discrete isotropic total variation (TV) [29] of u under certain boundary conditions if the ℓ_2-norm $\|\cdot\|_2$ (Euclidean norm) is used. In the early work, Candés et al. [5] minimized the TV regularization to reconstruct $\bar{u}$ from the undersampled measurement b at the noiseless case. However, the TV regularization itself will not be a good sparse transformation because the fact that the MR image itself is sparse under a certain basis is neglected. The pioneered works of Lustig et al. [24] and He et al. [16] characterized the image reconstruction as a linear combination of wavelet sparsity and TV regularization:

$$\min_u \sum_i^n \|D_i u\|_2 + \tau\|\Psi^* u\|_1 + \frac{\mu}{2}\|\Phi u - b\|_2^2, \tag{1.2}$$

where $\|\cdot\|_1$ is the ℓ_1-norm defined as the sum of absolute values of all entries of a vector, and τ, $\mu > 0$ are weighting parameters to balance each term for minimization. The problem (1.2) is generally more difficult to solve than any of those with a single TV regularization problem, e.g., [42, 34, 39, 1, 36, 37]. For solving (1.2), the method of He et al. [16] is a nonlinear inverse scale space method, but its speeds are not impressive. The method of Ma et al. [25] is actually an operator splitting approach to the inclusion problem resulting from the first-order optimality condition of (1.2). The method of Yang et al. [41] is an alternating direction multiplier method (ADMM), which is used to solve the equivalent form of problem (1.2) containing some auxiliary variables. The method of Goldstein and Osher [15] is a splitting Bregman method to solve a constrained optimization variant of (1.2). The method of Li et al. [23] is a two-step fixed-point proximity algorithm to a noiseless variant of (1.2) from a dual approach. The method of Ding et al. [9] is also taking a dual approach but using an ADMM instead. All the reviewed methods are different in convergence speed, ease of implementations, and practical applicabilities, but there is no evidence can verify that which approach outperforms the others under all scenarios.

There are many works on the reconstruction of image which is only corrupted by white Gaussian noise, while little is about the impulsive noises, such as the salt-and-pepper (*resp.* random-valued) noise where the noisy pixels take only (*resp.* have random values uniformly distributed) on the maximum and the minimum values in a range. The type of this noise is often due to bit errors in transmission, faulty memory locations, errors in analog-to-digital conversion, malfunctioning pixel elements in camera sensors. It has been highlighted in the literature that using the ℓ_1-norm as the measure of the data fidelity may exhibit enormous

advantages, and this fact has been made clear by some impressive numerical results in [26]. Indeed, the TV regularization with ℓ_1-norm data fidelity was studied frequently in the context of image denoising and deblurring, e.g., [3, 26, 6, 2, 40, 38, 30]. Inspired by the favorable influence of the ℓ_1-norm data fidelity under impulsive noise, in this paper, we replace the ℓ_2-norm in (1.2) with ℓ_1-norm and restrict our attention to the following reconstruction model:

$$\min_u \sum_{i=1}^{n} \|D_i u\|_2 + \tau\|\Psi^* u\|_1 + \mu\|\Phi u - b\|_1. \tag{1.3}$$

This model has latent capacity of eliminating the underlying sparse noise, and simultaneously preserving sparsity with respect to discrete gradient transform and wavelet transform. More than such attractive properties, though, solving (1.3) is really more challenging due to these three non-differentiable terms. To the best of our knowledge, there is no algorithm available to solve (1.3) in the field of CS. One might think that the ℓ_1-regularized terms can be combined together, and then solved by some existing solvers, such as [22, 27]. However, the inherent separable-structures may not be fully utilized. Hence, efficient algorithms that make full use of the favorable structure of each non-differentiable term are deserve investigating.

It should be noted that the main contribution of this paper is to propose a model (1.3) to solve the image reconstruction problem with impulse noise from highly undersampled data. We show that the proposed model can be solved by using existing algorithms from the perspective of primal and dual. The first algorithm focuses the primal model by making the best use of its separable structures. More accurately, problem (1.3) is firstly converted into an equivalent convex constrained minimization problem consisting of four blocks of variables and three blocks of non-differentiable functions, and then solved by ADMM directly. In contrast, the second algorithm focuses on the dual problem whose objective function contains three blocks of non-differentiable terms. We use ADMM to the dual problem and employ an ingenious symmetric Gauss-Seidel (sGS) technique [19, 20] to split its augmented Lagrangian function into some smaller parts. We show that each subproblem involved in both algorithms is easily performed via the favourable structures of Fourier transforms, wavelet transforms, and the Frenchel conjugate ℓ_1-norm. Experiments using simulated phantom and clinical images demonstrate that dual algorithm has superior numerical efficiency.

The remaining parts of this paper are organized as follows. In Section 2, we quickly review the key ingredients needed to design for our subsequent developments. In Section 3, we apply the ADMM to solve the model (1.3) and list its convergence result. In Section 4, we turn our attention to the applications of sGS-ADMM to the dual formulation of (1.3). In Section 5, we report numerical experiments to show the efficiency of all algorithms. Finally, we conclude our paper in Section 6.

2 Preliminaries

2.1 Basic concepts

For a nonempty closed convex set $\mathcal{C}$, we use the symbol $\delta_{\mathcal{C}}(x)$ to represent the indicator function over $\mathcal{C}$ such that $\delta_{\mathcal{C}}(x) = 0$ if $x \in \mathcal{C}$ and $+\infty$ otherwise. A subset $\mathcal{K}$ of $\mathcal{X}$ is called a cone if it is closed under positive scalar multiplication, i.e., $\lambda x \in \mathcal{K}$ when $x \in \mathcal{K}$ and $\lambda > 0$ [28]. The normal cone of $\mathcal{K}$ at point $x \in \mathcal{K}$ is defined by $\mathcal{N}_{\mathcal{K}}(x) = \{y \in \mathcal{X} \mid \langle y, z - x\rangle \leq 0,\ \forall z \in \mathcal{K}\}$. Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function. The subdifferential of f at $x \in \mathrm{dom}(f)$ is defined as $\partial f(x) = \{x^* \mid f(z) \geq f(x) + \langle x^*, z - x\rangle,\ \forall z \in \mathcal{X}\}$. Obviously,

$\partial f(x)$ is a closed convex set when it is not empty [28]. Let $\|\cdot\|$ be a norm function defined on $\mathcal{X}$. Then its dual norm $\|\cdot\|_*$ is defined as:

$$\|x\|_* = \sup_y\{x^\top y \mid \|y\| \le 1\}.$$

It is easy to see that the dual norm of ℓ_1-norm is ℓ_∞-norm, and the dual norm of ℓ_2-norm is ℓ_2-norm itself. The Fenchel conjugate of a convex f at $x \in \mathcal{X}$ is defined as

$$f^\star(y) := \sup_x\{\langle x, y\rangle - f(x)\} = -\inf_x\{f(x) - \langle x, y\rangle\}, \quad \forall y \in \mathcal{X}.$$

It is well known that the conjugate function $f^\star(y)$ is always convex and closed, proper if and only if f is proper [28]. For any $x \in \mathcal{X}$, there exists a $y \in \mathcal{X}$ such that $y \in \partial f(x)$ or equivalently $x \in \partial f^\star(y)$ due to a fact of f being closed and convex [28, Theorem 23.5].

Using the definition of the dual norm, it is easy to deduce that the Fenchel conjugate of $\|x\|_1$ is $\|x\|_1^\star = \delta_{\mathcal{B}_\infty^{(1)}}(x)$ where $\mathcal{B}_\infty^{(1)} = \{x \mid \|x\|_\infty \le 1\}$ is a closed and convex set. Given $x \in \mathcal{X}$, it is known that the orthogonal projection onto the ℓ_∞-norm ball $\mathcal{B}_\infty^{(r)}(x)$ and ℓ_2-norm ball $\mathcal{B}_2^{(r)}(x)$ with radius $r > 0$ can be expressed explicitly. The Moreau-Yosida regularization of a closed proper convex function f at $x \in \mathcal{X}$ with positive scalar $\beta > 0$ is defined by

$$\varphi_f^\beta(x) := \min_{y\in\mathcal{X}}\Big\{f(y) + \frac{1}{2\beta}\|y - x\|_2^2\Big\}. \tag{2.1}$$

Moreover, the problem (2.1) has an unique optimal solution, which is known as the proximal mapping of x associated with f, i.e.,

$$\mathcal{P}_f^\beta(x) := \arg\min_{y\in\mathcal{X}}\Big\{f(y) + \frac{1}{2\beta}\|y - x\|_2^2\Big\}.$$

It is shown that the proximal mapping of the ℓ_1-norm function at point x obeys the following form

$$\mathcal{P}_{\|\cdot\|_1}^\beta(x) = \mathrm{sgn}(x) \odot \max\big\{|x| - \beta, 0\big\},$$

where $\odot$ is Hadamard product, and the operations of sign function "sgn$(\cdot)$" and absolute value function "$|\cdot|$" are component-wises. It is also well known that $\mathcal{P}_f^\beta(\cdot)$ is firmly nonexpansive and globally Lipschitz continuous with modulus 1. For any $x \in \mathcal{X}$, the Moreau decomposition is expressed as $x = \mathcal{P}_f^\beta(x) + \mathcal{P}_{f^\star}^\beta(x)$. Taking the ℓ_1-norm function as an example, its proximal mapping at x can be expressed as $\mathcal{P}_{\|\cdot\|_1}^\beta(x) = x - \mathcal{P}_{\delta_{\mathcal{B}_\infty^{(1)}}}^\beta(x) = x - \Pi_{\mathcal{B}_\infty^{(1)}}(x)$, which also indicates $\mathcal{P}_{\|\cdot\|_\infty}^\beta(x) = x - \Pi_{\mathcal{B}_1^{(1)}}(x)$.

2.2 Semi-proximal ADMM

Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be finite dimensional real Euclidian spaces. Consider the convex optimization problem with the following two-block separable structure

$$\begin{array}{ll}\min\limits_{y,z} & f(y) + g(z)\\ \text{s.t.} & \mathcal{A}^*y + \mathcal{B}^*z = c,\end{array} \tag{2.2}$$

where $f : \mathcal{Y} \to (-\infty, +\infty]$ and $g : \mathcal{Z} \to (-\infty, +\infty]$ are closed proper convex functions, $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{B} : \mathcal{X} \to \mathcal{Z}$ are given linear maps, and $c \in \mathcal{X}$ is given data. The Karush-Kuhn-Tucker (KKT) system of problem (2.2) is given by

$$0 \in \mathcal{A}x + \partial f(y), \quad 0 \in \mathcal{B}x + \partial g(z), \quad \text{and} \quad \mathcal{A}^*y + \mathcal{B}^*z = c.$$

The augmented Lagrangian function associated with (2.2) is given by

$$\mathcal{L}_\sigma(y,z;x) = f(y) + g(z) + \langle x, \mathcal{A}^*y + \mathcal{B}^*z - c\rangle + \frac{\sigma}{2}\|\mathcal{A}^*y + \mathcal{B}^*z - c\|^2,$$

where $x \in \mathcal{X}$ is a multiplier, and $\sigma > 0$ be a given penalty parameter. Staring from an initial point $(x^0, y^0, z^0) \in \mathcal{X} \times (\mathrm{dom}\ f) \times (\mathrm{dom}\ g)$, the iterations of the semi-proximal ADMM of Fazel, Pong, Sun & Tseng [11] for solving (2.2) is summarized as

$$\left\{\begin{array}{ll} y^{k+1} & = \arg\min_y \left\{\mathcal{L}_\sigma(y, z^k; x^k) + \frac{\sigma}{2}\|y - y^k\|^2_{\mathcal{T}_f}\right\}, \\ z^{k+1} & = \arg\min_z \left\{\mathcal{L}_\sigma(y^{k+1}, z; x^k) + \frac{\sigma}{2}\|z - z^k\|^2_{\mathcal{T}_g}\right\}, \\ x^{k+1} & = x^k + \xi\sigma\big(\mathcal{A}^*y^{k+1} + \mathcal{B}^*z^{k+1} - c\big), \end{array}\right. \tag{2.3}$$

where $\mathcal{T}_f$ and $\mathcal{T}_g$ are positive semi-definite and the step-length ξ is chosen in the interval $(0, (1+\sqrt{5})/2)$. When $\mathcal{T}_f = 0$ and $\mathcal{T}_g = 0$, the semi-proximal ADMM (2.3) reduces to the classical ADMM in the mid-1970s. The following theorem is selected from the convergence Theorem B.1 in [11]. For more details, one can refer to [11] and the references therein.

3 ADMM Based on Primal Formulation

3.1 Iterative framework

This section is devoted to the first assignment of this paper, which aims to apply the ADMM to solve the problem (1.3) from a primal perspective. In order to facilitate the analysis and design, we introduce a triple of auxiliary variables $w_i \in \mathbb{R}^2$, $z \in \mathbb{R}^n$, and $v \in \mathbb{R}^m$ such that $w_i = D_i u$, $z = \Psi^* u$, and $v = \Phi u - b$. Therefore, the problem (1.3) is equivalently transformed into

$$\begin{array}{ll} \min\limits_{w,z,u,v} & \sum_{i=1}^n \|w_i\|_2 + \tau\|z\|_1 + \mu\|v\|_1 \\ \text{s.t.} & w_i = D_i u, \quad i = 1,\ldots,n, \\ & z = \Psi^* u, \\ & v = \Phi u - b, \end{array} \tag{3.1}$$

which contains four separable blocks of variables with respect to w, z, u, and v, and three non-smooth blocks of functions. The augmented Lagrangian function of (3.1) is given by

$$\begin{array}{ll} & \mathcal{L}_\beta(w,z,u,v;\lambda_1,\lambda_2,\lambda_3) \\ = & \sum_{i=1}^n \|w_i\|_2 + \tau\|z\|_1 + \mu\|v\|_1 - \sum_i \langle (\lambda_1)_i, w_i - D_i u\rangle - \langle \lambda_2, z - \Psi^* u\rangle \\ & -\langle \lambda_3, v - (\Phi u - b)\rangle + \frac{\beta}{2}\sum_i \|w_i - D_i u\|_2^2 + \frac{\beta}{2}\|z - \Psi^* u\|_2^2 + \frac{\beta}{2}\|v - (\Phi u - b)\|_2^2, \end{array} \tag{3.2}$$

where $(\lambda_1)_i \in \mathbb{R}^2$, $\lambda_2 \in \mathbb{R}^n$, and $\lambda_3 \in \mathbb{R}^m$ are multipliers and $\beta > 0$ is a penalty parameter.

If we treat (w, z, v) as one group and u itself as another, then starting from $(\omega^0, z^0, u^0, v^0)$, it is well known that the classical ADMM of Glowinski and Marroco [12] takes the following form:

$$\left\{\begin{array}{l} (w^{k+1}, z^{k+1}, v^{k+1}) = \arg\min\limits_{w,z,v} \mathcal{L}_\beta(w, z, u^k, v; \lambda_1^k, \lambda_2^k, \lambda_3^k), \\ u^{k+1} = \arg\min\limits_{u} \mathcal{L}_\beta(w^{k+1}, z^{k+1}, u, v^{k+1}; \lambda_1^k, \lambda_2^k, \lambda_3^k), \\ (\lambda_1^{k+1})_i = (\lambda_1^k)_i - \xi\beta(w_i^{k+1} - D_i u^{k+1}), \qquad \forall i = 1,\ldots,n, \\ \lambda_2^{k+1} = \lambda_2^k - \xi\beta(z^{k+1} - \Psi^* u^{k+1}), \\ \lambda_3^{k+1} = \lambda_3^k - \xi\beta(v^{k+1} - \Phi u^{k+1} + b), \end{array}\right. \tag{3.3}$$

where $\xi \in (0, (1+\sqrt{5})/2)$. For more detail on the convergence of ADMM with the iterative scheme (3.3), one may refer to Gabay [14] and Gabay & Mercier [13].

3.2 Subproblems' solving

This subsection aims to solve the subproblems involved in iterative framework (3.3) by making the best use of their favourable structures. We now show that simple closed-form solutions are permitted for each subproblem which indicates that all the subproblems are easily performed.

We first focus on the solving of the (w, z, v)-subproblems by utilizing its separable structures. With fixed u^k, and λ_i^k for $i = 1, 2, 3$, the next iterates $(w^{k+1}, z^{k+1}, v^{k+1})$ at the first implementation of (3.3) with respect to (w, z, v) can be expressed as

$$\begin{aligned}
&(w^{k+1}, z^{k+1}, v^{k+1}) \\
= \ & \arg\min_{w,z,v} \Big\{ \textstyle\sum_{i=1}^n \|w_i\|_2 + \tau\|z\|_1 - \sum_{i=1}^n \langle (\lambda_1)_i^k, w_i - D_i u^k\rangle \\
& \qquad -\langle \lambda_2^k, z - \Psi^* u^k\rangle + \tfrac{\beta}{2}\textstyle\sum_{i=1}^n \big\|w_i - D_i u^k\big\|_2^2 + \tfrac{\beta}{2}\big\|z - \Psi^* u^k\big\|_2^2 \\
& \qquad\qquad +\mu\|v\|_1 - \langle \lambda_3^k, v - (\Phi u^{k+1/2} - b)\rangle + \tfrac{\beta}{2}\big\|v - (\Phi u^{k+1/2} - b)\big\|_2^2 \Big\} \\
= \ & \Big(\arg\min_{w} \big\{ \textstyle\sum_{i=1}^n (\|w_i\|_2 + \tfrac{\beta}{2}\big\|w_i - \big(D_i u^k + \tfrac{(\lambda_1)_i^k}{\beta}\big)\big\|_2^2)\big\}, \\
& \qquad \arg\min_{z} \big\{\tau\|z\|_1 + \tfrac{\beta}{2}\big\|z - \big(\Psi^* u^k + \tfrac{\lambda_2^k}{\beta}\big)\big\|_2^2\big\}, \\
& \qquad\qquad \arg\min_{v} \Big\{\mu\|v\|_1 + \tfrac{\beta}{2}\big\|v - (\Phi u^{k+1/2} - b + \tfrac{\lambda_3^k}{\beta})\big\|_2^2\Big\}\Big).
\end{aligned} \tag{3.4}$$

From some simple calculations (see [39, 40, 41]), one may show that each subproblem obeys the forms of proximal mapping, and hence, they can be expressed explicitly as follows

$$\begin{cases}
w_i^{k+1} = \max\Big\{\big\|D_i u^k + (\lambda_1)_i^k/\beta\big\|_2 - \frac{1}{\beta}, 0\Big\} \odot \operatorname{sgn}(D_i u^k + (\lambda_1)_i^k/\beta), \qquad i = 1, \ldots, n, \\
z_{k+1} = \max\Big\{\big|\Psi^* u^k + \lambda_2^k/\beta\big| - \frac{\tau}{\beta}, 0\Big\} \odot \operatorname{sgn}(\Psi^* u^k + \lambda_2^k/\beta), \\
v_{k+1} = \max\Big\{|\Phi u^{k+1/2} - b + \lambda_3^k/\beta| - \frac{\mu}{\beta}, 0\Big\} \odot \operatorname{sgn}(\Phi u^{k+1/2} - b + \lambda_3^k/\beta),
\end{cases} \tag{3.5}$$

where the operations "max" and $\cdot/\cdot$ are component-wises, and $0/0 = 0$ is followed. This calculation process indicates that the (w, z, v)-subproblems can be partitioned into a triple of lower-dimensional subproblems concerning each variable in a parallel way to make it is more easier to implement.

With the latest values of w^{k+1}, z^{k+1} and v^{k+1}, and λ_i^k for $i = 1, 2, 3$, the u^{k+1} in (3.3) coincides with the following form

$$\begin{aligned}
u^{k+1} = \arg\min_{u} \Big\{ & -\sum_{i=1}^n \langle ((\lambda_1)_i^k, w_i^{k+1} - D_i u\rangle - \langle \lambda_2^k, z^{k+1} - \Psi^* u\rangle - \langle \lambda_3^k, v^{k+1} - (\Phi u - b)\rangle \\
& + \frac{\beta}{2}\sum_{i=1}^n \|w_i^{k+1} - D_i u\|_2^2 + \frac{\beta}{2}\|z^{k+1} - \Psi^* u\|_2^2 + \frac{\beta}{2}\|v^{k+1} - (\Phi u - b)\|_2^2\Big\}.
\end{aligned} \tag{3.6}$$

To simplify presentation, we denote $D := (D_{(1)}; D_{(2)}) \in \mathbb{R}^{2n\times n}$, where $D_{(1)} \in \mathbb{R}^{n\times n}$ and $D_{(2)} \in \mathbb{R}^{n\times n}$ contains, respectively, the first and second rows of D_i for all i. Correspondingly,

we also denote $w := (w_{(1)}; w_{(2)}) \in \mathbb{R}^{2n}$, where $w_{(1)} := D_{(1)}u \in \mathbb{R}^n$ and $w_{(2)} := D_{(2)}u \in \mathbb{R}^n$, that is, $w_{(1)}$ (*resp.* $w_{(2)}$) is a vector constructed by the first (*resp.* second) component of $w_i \in \mathbb{R}^2$ for $i = 1, \ldots, n$. Meanwhile, $\lambda_1 \in \mathbb{R}^{2n}$ is also reformulated as $(\lambda_{1(1)}; \lambda_{1(2)}) \in \mathbb{R}^{2n}$ with $\lambda_{1(1)} \in \mathbb{R}^n$ and $\lambda_{1(2)} \in \mathbb{R}^n$ corresponding to the structure of w. With these notations, the u-subproblem can be rewritten as

$$\arg\min_u \Big\{ -(\lambda_1^k)^\top (w^{k+1} - Du) + \frac{\beta}{2}\|w^{k+1} - Du\|_2^2 - (\lambda_2^k)^\top (z^{k+1} - \Psi^* u) + \frac{\beta}{2}\|z^{k+1} - \Psi^* u\|_2^2 \\ - \langle \lambda_3^k, v^{k+1} - (\Phi u - b)\rangle + \frac{\beta}{2}\|v^{k+1} - (\Phi u - b)\|_2^2 \Big\},$$

which is an unconstrained quadratic programming. Therefore, from the first-order optimality condition and the orthonormality of Ψ, we know that solving this problem is equivalent to finding a solution to the following normal equation:

$$Au = y^k, \tag{3.7}$$

where

$$A = D_{(1)}^\top D_{(1)} + D_{(2)}^\top D_{(2)} + I + \Phi^*\Phi,$$

and

$$y^k = D_{(1)}^\top \Big(w_{(1)}^{k+1} - \frac{\lambda_{1(1)}^k}{\beta}\Big) + D_{(2)}^\top \Big(w_{(2)}^{k+1} - \frac{\lambda_{1(2)}^k}{\beta}\Big) + \Psi\Big(z^{k+1} - \frac{\lambda_2^k}{\beta}\Big) + \Phi^*\Big((v^{k+1} + b) - \frac{\lambda_3^k}{\beta}\Big).$$

Since $D_{(1)}$ and $D_{(2)}$ are finite difference operators, under the periodic boundary conditions for u, they are circulant matrices and can be diagonalized by the Fourier transform $\mathcal{F}$. It is worth pointing out that if $\mathcal{T}$ is a discrete cosine transform, the same result holds under the symmetric boundary conditions. The attractive feature indicates that the linear system (3.7) is easily solved by taking a fast Fourier transform to both sides, and then taking an inverse version. It should be noted that these techniques are from [39, 40, 41], and they appeared here is to make this part more easier to follow.

In light of all above derivations, we summarize the steps of ADMM for solving problem (3.1) as follows. The implemented details concerning on stopping criterion will be stated in the numerical parts later.

Algorithm: ADMM

Step 0. Input problem data b; choose model parameters τ, $\mu > 0$, and constants $\beta > 0$, $\xi \in (0, (1+\sqrt{5})/2)$. Initialize w_i^0, z^0, u^0, v^0, λ_1^0, λ_2^0, and λ_3^0. For $k = 0, 1, \ldots$, do the following operations iteratively.

Step 1. Compute w^{k+1}, z^{k+1} and v^{k+1} via (3.5), respectively.

Step 2. Compute u^{k+1} via (3.7) with fixed w^{k+1}, z^{k+1}, v^{k+1} and λ_i^k for $i = 1, 2, 3$.

Step 3. Update multipliers λ_i^{k+1} for $i = 1, 2, 3$ separately via (3.7) with fixed w^{k+1}, z^{k+1}, v^{k+1} and u^{k+1}.

Step 4. Set $k := k + 1$, go to Step 1.

---

4 sGS-ADMM Based on Dual Formulation

4.1 Lagrangian dual problem and optimality condition

This section aims to construct a sGS-ADMM frame algorithm to solve the problem (3.1) from a dual perspective. The Lagrangian function associated with (3.1) is defined by

$$\begin{aligned}\mathcal{L}(w,z,u,v;\lambda_1,\lambda_2,\lambda_3) =& \sum_{i=1}^{n}\|w_i\|_2+\tau\|z\|_1+\mu\|v\|_1-\sum_{i=1}^{n}\langle(\lambda_1)_i,w_i-D_iu\rangle\\ &-\langle\lambda_2,z-\Psi^*u\rangle-\langle\lambda_3,v-(\Phi u-b)\rangle,\end{aligned}$$

where the vectors $(\lambda_1)_i\in\mathbb{R}^2$, $\lambda_2\in\mathbb{R}^n$, $\lambda_3\in\mathbb{R}^m$ are multipliers or dual variables associated with the constraints in (3.1). It is not a hard work to deduce that the Lagrangian dual problem of the original problem (3.1) takes the following equivalent form:

$$\begin{aligned}\min_{\lambda_1,\lambda_2,\lambda_3}\quad & \langle\lambda_3,b\rangle\\ \text{s.t.}\quad & D^\top\lambda_1+\Psi\lambda_2+\Phi^\top\lambda_3=0,\\ & (\lambda_1)_i\in\mathcal{B}_2^{(1)},\lambda_2\in\mathcal{B}_\infty^{(\tau)},\lambda_3\in\mathcal{B}_\infty^{(\mu)},i=1,\ldots,n,\end{aligned}\tag{4.1}$$

where $\mathcal{B}_2^{(1)}:=\{(\lambda_1)_i\in\mathbb{R}^2\mid\|(\lambda_1)_i\|_2\le 1\}$, $\mathcal{B}_\infty^{(\tau)}:=\{\lambda_2\in\mathbb{R}^n\mid\|\lambda_2\|_\infty\le\tau\}$, and $\mathcal{B}_\infty^{(\mu)}:=\{\lambda_3\in\mathbb{R}^m\mid\|\lambda_3\|_\infty\le\mu\}$. Assuming that both the optimal solutions to the primal problem (1.3) and its dual problem (4.1) exist, then $(\hat\lambda_1,\hat\lambda_2,\hat\lambda_3)$ is said to be an optimal solution of problem (4.1) if there exists a $(\hat w_i,\hat z,\hat u,\hat v)$ be an optimal solution of (1.3) such that the following system is satisfied

$$\begin{cases} w_i-D_iu=0,\ z-\Psi^*u=0,\ v-(\Phi u-b)=0,\\ D^\top\lambda_1+\Psi\lambda_2+\Phi^\top\lambda_3=0,\\ (\lambda_1)_i\in\partial\|w_i\|_2,\ \lambda_2\in\tau\partial\|z\|_1,\ \lambda_3\in\mu\partial\|v\|_1.\end{cases}\tag{4.2}$$

The above system is the well-known Karush-Kuhn-Tucker (KKT) condition in optimization community.

4.2 Algorithm's construction and convergence result

We now focus on the applications of sGS-ADMM for solving of the dual problem (4.1). For the convenience of calculation, we introduce an auxiliary variable $x\in\mathbb{R}^n$ to make an equivalent formulation:

$$\begin{aligned}\min_{\lambda_1,\lambda_2,\lambda_3,x}\quad & \langle\lambda_3,b\rangle+\delta_{B_2^{(1)}}(\lambda_1)+\delta_{B_\infty^{(\tau)}}(x)+\delta_{B_\infty^{(\mu)}}(\lambda_3)\\ \text{s.t.}\quad & D^\top\lambda_1+\Psi\lambda_2+\Phi^\top\lambda_3=0,\\ & \lambda_2-x=0,\end{aligned}\tag{4.3}$$

where $\delta_{\mathcal{C}}(\cdot)$ represents an indicator function on a closed and convex set $\mathcal{C}$. Given $\beta > 0$, the augmented Lagrangian function associated with (4.3) is defined by

$$\begin{aligned} & \mathcal{L}_\beta(\lambda_1, \lambda_2, \lambda_3, x; u, z) \\ = \; & \langle \lambda_3, b\rangle + \delta_{B_2^{(1)}}(\lambda_1) + \delta_{B_\infty^{(\tau)}}(x) + \delta_{B_\infty^{(\mu)}}(\lambda_3) - \langle u, D^\top\lambda_1 + \Psi\lambda_2 + \Phi^\top\lambda_3\rangle - \langle z, \lambda_2 - x\rangle \\ & + \tfrac{\beta}{2}\|D^\top\lambda_1 + \Psi\lambda_2 + \Phi^\top\lambda_3\|_2^2 + \tfrac{\beta}{2}\|\lambda_2 - x\|_2^2, \end{aligned} \tag{4.4}$$

where u and z are multipliers, or the primal variables in problem (3.1). To apply sGS-ADMM, we view (λ_1, λ_2) as one group and (λ_3, x) as another. Noting that λ_1 and λ_2 are coupled together while λ_3 and x are separated from each other, we adapt the sGS technique in the (λ_1, λ_2) group with order $\lambda_2 \to \lambda_1 \to \lambda_2$, and take a parallel manner in the group (λ_3, x). More precisely, the $(\lambda_1, \lambda_2, \lambda_3, x)$-subproblems can be solved individually via the following scheme:

$$\left\{\begin{array}{l} \lambda_2^{k+1/2} = \arg\min_{\lambda_2}\big\{\mathcal{L}_\beta(\lambda_1^k, \lambda_2, \lambda_3^k, x^k; u^k, z^k) + \frac{1}{2}\|\lambda_2 - \lambda_2^k\|_{\mathcal{T}_2}\big\}, \\ \lambda_1^{k+1} = \arg\min_{\lambda_1}\big\{\mathcal{L}_\beta(\lambda_1, \lambda_2^{k+1/2}, \lambda_3^k, x^k; u^k, z^k) + \frac{1}{2}\|\lambda_1 - \lambda_1^k\|_{\mathcal{T}_1}\big\}, \\ \lambda_2^{k+1} = \arg\min_{\lambda_2}\big\{\mathcal{L}_\beta(\lambda_1^{k+1}, \lambda_2, \lambda_3^k, x^k; u^k, z^k) + \frac{1}{2}\|\lambda_2 - \lambda_2^{k+1/2}\|_{\mathcal{T}_2}\big\}, \\ \lambda_3^{k+1} = \arg\min_{\lambda_3}\big\{\mathcal{L}_\beta(\lambda_1^{k+1}, \lambda_2^{k+1}, \lambda_3, x^k; u^k, z^k) + \frac{1}{2}\|\lambda_3 - \lambda_3^k\|_{\mathcal{T}_3}\big\}, \\ x^{k+1} = \arg\min_{x}\big\{\mathcal{L}_\beta(\lambda_1^{k+1}, \lambda_2^{k+1}, \lambda_3^{k+1}, x; u^k, z^k) + \frac{1}{2}\|x - x^k\|_{\mathcal{T}_x}\big\}, \end{array}\right. \tag{4.5}$$

where $\mathcal{T}_x$ and $\mathcal{T}_i$ for $i = 1, 2, 3$ are self-adjoint positive semi-definite linear operators.

To enjoy closed-form solution for each subproblem, we set $\mathcal{T}_2 \equiv \mathcal{T}_x \equiv 0$, $\mathcal{T}_1 := \beta(\alpha I - DD^\top)$ and $\mathcal{T}_3 := \beta(\eta I - \Phi\Phi^\top)$ where the positive scalars α and η are chosen particularly to ensure that $\mathcal{T}_1$ and $\mathcal{T}_3$ are positive semi-definite. Generally, we can choose α and η such that $\alpha \geq \rho(DD^\top)$ and $\eta \geq \rho(\Phi\Phi^\top)$, where $\rho(\cdot)$ denotes a spectral radius of a given matrix. The following lemma plays a key role in convergence analysis, which is just an application of [20, Theorem 1] to a special practical problem.

Lemma 4.1. *([20, Theorem 1]) Suppose that $\mathcal{T}_2 \equiv \mathcal{T}_x \equiv 0$, and $\mathcal{T}_1 := \beta(\alpha I - DD^\top)$ and $\mathcal{T}_3 := \beta(\eta I - \Phi\Phi^\top)$ are self-adjoint positive semi-definite linear operators with positive scalars α and η. Then the iterative scheme (4.5) is equivalent to*

$$\left\{\begin{array}{l} \begin{pmatrix} \lambda_1^{k+1} \\ \lambda_2^{k+1} \end{pmatrix} = \arg\min\limits_{\lambda_1,\lambda_2} \mathcal{L}_\beta\big((\lambda_1, \lambda_2), (\lambda_3^k, x^k); u^k, z^k\big) + \frac{1}{2}\left\| \begin{array}{c} \lambda_1 - \lambda_1^k \\ \lambda_2 - \lambda_2^k \end{array} \right\|_{\mathcal{Q}_1}^2, \\ \begin{pmatrix} \lambda_3^{k+1} \\ x^{k+1} \end{pmatrix} = \arg\min\limits_{\lambda_3,x} \mathcal{L}_\beta\big((\lambda_1^{k+1}, \lambda_2^{k+1}), (\lambda_3, x); u^k, x^k\big) + \frac{1}{2}\left\| \begin{array}{c} \lambda_3 - \lambda_3^k \\ x - x^k \end{array} \right\|_{\mathcal{Q}_2}^2, \end{array}\right.$$

where $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are defined as, respectively,

$$\mathcal{Q}_1 = \frac{\beta}{2}\begin{pmatrix} DD^\top & 0 \\ 0 & 0 \end{pmatrix}, \quad \textit{and} \quad \mathcal{Q}_2 = \beta\begin{pmatrix} \eta I - \Phi\Phi^\top & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof. The augmented Lagrangian function given in (4.4) with regard to variables λ_1 and λ_2 is a quadratic function with quadratic term

$$\mathcal{H}_1 = \beta\begin{pmatrix} DD^\top & D\Psi \\ \Psi^* D^\top & 2I \end{pmatrix}.$$

Define

$$\mathcal{H}_2 = \begin{pmatrix} \mathcal{T}_1 & 0 \\ 0 & \mathcal{T}_2 \end{pmatrix} = \beta\begin{pmatrix} \alpha I - DD^\top & 0 \\ 0 & 0 \end{pmatrix},$$

and denote

$$\mathcal{H} := \mathcal{H}_1 + \mathcal{H}_2 = \beta\begin{pmatrix} \alpha I & D\Psi \\ \Psi^* D^\top & 2I \end{pmatrix}.$$

Moreover, we decompose $\mathcal{H}$ into three pars with form

$$\mathcal{J} = \beta\begin{pmatrix} \alpha I & 0 \\ 0 & 2I \end{pmatrix}, \quad \mathcal{U} = \beta\begin{pmatrix} 0 & D\Psi \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad \mathcal{U}^* = \beta\begin{pmatrix} 0 & 0 \\ \Psi^* D^\top & 0 \end{pmatrix}.$$

It follows from [20, Theorem 1], that, the desired $\mathcal{Q}_1$ obeys the following form

$$\mathcal{Q}_1 = \mathcal{U}\mathcal{J}^{-1}\mathcal{U}^* = \frac{\beta}{2}\begin{pmatrix} DD^\top & 0 \\ 0 & 0 \end{pmatrix}.$$

The explicit form of $\mathcal{Q}_2$ is easily from the fact that the variables λ_3 and x are separated from each other. □

This equivalence property is enough to ensure the convergence theory from [11, Theorem B.1] which can be reported as follows.

Theorem 4.2. *Suppose that $\mathcal{T}_2 \equiv \mathcal{T}_x \equiv 0$, and $\mathcal{T}_1 \triangleq \beta(\alpha I - DD^\top)$ and $\mathcal{T}_3 \triangleq \beta(\eta I - \Phi\Phi^\top)$ are self-adjoint positive semi-definite linear operators with positive scalars α and η. Let $\{\lambda_1^k, \lambda_2^k, \lambda_3^k, x^k\}$ be the sequence generated by the algorithm associated with the iterative scheme (4.5), then it converges to the optimal solution of problem (3.1).*

4.3 Subproblems' solving and algorithm's steps

We now pay our attention on the solving of the subproblems involved in scheme (4.5). Firstly, fix the values of other variables and notice that $\mathcal{T}_2 \equiv 0$, the λ_2-subproblem is expressed as

$$\lambda_2^{k+1/2} = \arg\min_{\lambda_2} \Big\{ -\langle u^k, D^\top\lambda_1^k + \Psi\lambda_2 + \Phi^\top\lambda_3^k\rangle - \langle z^k, \lambda_2 - x^k\rangle\} \\ + \frac{\beta}{2}\|D^\top\lambda_1^k + \Psi\lambda_2 + \Phi^\top\lambda_3^k\|_2^2 + \frac{\beta}{2}\|\lambda_2 - x^k\|^2\Big\}.$$

Finding the minimizer $\lambda_2^{k+1/2}$ is equivalent to finding the solution of the following system

$$\Psi^* u^k + z^k = \beta\Psi^*(D^\top\lambda_1^k + \Psi\lambda_2 + \Phi^\top\lambda_3^k) + \beta(\lambda_2 - x^k).$$

Hence, the solution $\lambda_2^{k+1/2}$ is given explicitly by

$$\lambda_2^{k+1/2} = \frac{1}{2\beta}\Big((\Psi^* u^k + z^k) + \beta x^k - \beta\Psi^*(D^\top\lambda_1^k + \Phi^\top\lambda_3^k)\Big). \tag{4.6}$$

Secondly, choose $\alpha > 0$ such that $\mathcal{T}_1 = \beta(\alpha I - DD^\top)$ is positive semi-definite, it is easy to deduce that the λ_1-subproblem reduces to

$$\begin{aligned}\lambda_1^{k+1} &= \arg\min_{\lambda_1} \Big\{\delta_{\mathcal{B}_2^{(1)}}(\lambda_1) - \langle u^k, D^\top\lambda_1\rangle + \frac{\beta}{2}\|D^\top\lambda_1 \\ &\qquad + \Psi\lambda_2^{k+1/2} + \Phi^\top\lambda_3^k\|_2^2 + \frac{1}{2}\|\lambda_1 - \lambda_1^k\|^2_{\beta(\alpha I - DD^\top)}\Big\} \\ &= \Pi_{\mathcal{B}_2^{(1)}}\Big(\lambda_1^k - \frac{G^k}{\alpha} + \frac{Du^k}{\alpha\beta}\Big),\end{aligned} \tag{4.7}$$

where $G^k = D(D^\top\lambda_1^k + \Psi\lambda_2^{k+1/2} + \Phi^\top\lambda_3^k)$, and $\Pi_{\mathcal{B}_2^{(1)}}(\cdot)$ is a metric projection onto the closed convex set $B_2^{(1)}$. Thirdly, choose $\eta > 0$ such that $\mathcal{T}_3 = \beta(\eta I - \Phi\Phi^\top)$ is positive semi-definite, the λ_3-subproblem can be solved as follows:

$$\begin{aligned}\lambda_3^{k+1} &= \arg\min_{\lambda_3}\Big\{\langle\lambda_3, b\rangle + \delta_{B_\infty^{(\mu)}}(\lambda_3) - \langle u^k, \Phi^\top\lambda_3\rangle + \frac{\beta}{2}\big\|D^\top\lambda_1^{k+1} + \Psi\lambda_2^{k+1} + \Phi^\top\lambda_3\big\|^2 \\ &\qquad + \frac{1}{2}\|\lambda_3 - \lambda_3^k\|^2_{\beta(\eta I - \Phi\Phi^\top)}\Big\} \\ &= \Pi_{\mathcal{B}_\infty^{(\mu)}}\Big(\lambda_3^k - \frac{M^k}{\eta} + \frac{\Phi u^k - b}{\eta\beta}\Big), \end{aligned} \tag{4.8}$$

where $M^k = \Phi(D^\top\lambda_1^{k+1} + \Psi\lambda_2^{k+1} + \Phi^\top\lambda_3^k)$. Finally, choosing $\mathcal{T}_x \equiv 0$, the x-subproblem can be reorganized as

$$x^{k+1} = \arg\min_x\Big\{\delta_{\mathcal{B}_\infty^{(\tau)}}(x) - \langle z^k, \lambda_2^{k+1} - x\rangle + \frac{\beta}{2}\|\lambda_2^{k+1} - x\|^2\Big\} = \Pi_{\mathcal{B}_\infty^{(\tau)}}\Big(\lambda_2^{k+1} - \frac{z^k}{\beta}\Big). \tag{4.9}$$

In light of the above analysis, we summarize the steps of the sGS-ADMM for solving the dual problem (4.1) as follows:

Algorithm: sGS-ADMM_d

Step 0. Input problem data b; input positive parameters τ, $\mu > 0$, and constants $\beta > 0$, $\xi \in (0, (1+\sqrt{5})/2)$; input $\alpha \geq \rho(DD^\top)$ and $\eta \geq \rho(\Phi\Phi^\top)$. Initialize z^0, u^0, λ_1^0, λ_2^0, and λ_3^0. For $k = 0, 1, \ldots$, do the following operations iteratively.

Step 1. Compute λ_2^{k+1} and λ_1^{k+1}:

- Compute $\lambda_2^{k+1/2}$ via (4.6);
- Compute λ_1^{k+1} via (4.7);
- Compute λ_2^{k+1} via solving (4.6) with the newest λ_1^{k+1}.

Step 2. Compute λ_3^{k+1} and x^{k+1}:

- Fixed λ_3^{k+1} via solving (4.8);
- Fixed x^{k+1} via solving (4.9).

Step 3. Updating multipliers u^{k+1} and z^{k+1}:

$$u^{k+1} = u^k - \xi\beta(D^\top\lambda_1^{k+1} + \Psi\lambda_2^{k+1} + \Phi^\top\lambda_3^{k+1});$$

$$z^{k+1} = z^k - \xi\beta(\lambda_2^{k+1} - x^{k+1}).$$

Step 4. Set $k := k + 1$, go to Step 1.

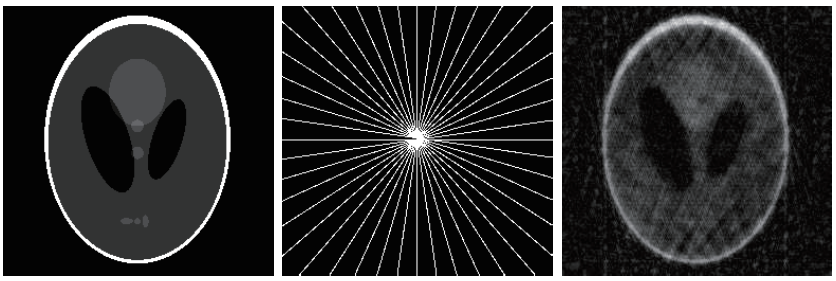
5 Numerical Experiments

In this section, we present MR image reconstruction results to evaluate the practical performances of all the presented algorithms for solving (1.3) from the primal and dual respects.

All the experiments are performed with Microsoft Windows 10 and MATLAB R2018a, and run on a PC with an Intel Core i7 CPU at 1.80 GHz and 8 GB of memory.

5.1 General descriptions

We conduct experiments on two types of MRI data: a simulated "Shepp-Logan" phantom, and some clinical MR images. The pixels values of all tested images are all re-scaled into $[0, 1]$ for convenience. We generate the data b by the undersampling Fourier coefficients of these images followed by an adding impulsive noise. The k-space undersampling is simulated by using the masks of pseudo radial sampling [32], which sampled the Fourier domain along a number of radial lines spread out from the center. The orthogonal transform Ψ and its inverse Ψ^* is generated by using the the Rice Wavelet Toolbox which is available at `https://github.com/ricedsp/rwt`. The Toolbox is a collection of functions for 1D and 2D wavelet and filter bank design, analysis, and processing. The undersampled data $\Phi\bar{u}$ in (1.1) is assumed to be corrupted, with different noisy densities, by two types of impulse noise, i.e., the random valued noise and salt-and-pepper noise. More precisely, let "min" and "max" be the minimum and the maximum element values of vector $\Phi\bar{u}$, respectively. In salt-and-pepper noise, the corrupted pixels take the "min" vale or the "max" value randomly, while in random valued noise the corrupted pixels take the random values at the range of ["min", "max"] uniformly. We conduct the following experiments on data corroded by salt-and-pepper noise and its level of impulse noise in each test is set as 0.1.



(a) Shepp-Logan phantom (b) 22 sampled lines (c) Corrupted phantom

Figure 5.1: *(a) The test* 256×256 *Shepp-Logan phantom; (b) Fourier domain sampling positions with* 22 *radial lines; (c) corrupted by impulsive noise phantom image.*

For comparison in a relatively fair way, we measure the reconstruction solutions using the relative error (RLNE) defined as

$$\text{RLNE} := \frac{\|\hat{u} - \bar{u}\|_2}{\|\bar{u}\|_2},$$

where $\hat{u}$ and $\bar{u}$ are the reconstructed and ground truth images, respectively. Besides, we also use the peak signal-to-noise ratio (PSNR) in the unit of dB defined as follows to measure the quality of the re-solutions

$$\text{PSNR} := 10 \log_{10} \frac{255\sqrt{n}}{\|\hat{u} - \bar{u}\|_2} (\text{dB}).$$

The iterative process of each algorithm is terminated if RLNE is sufficiently small or the maximum iteration number is achieved. The parameter is tuned as $\xi = 1.618$ in each tested algorithm. For the sGS-ADMM_d, the values of the parameters $\alpha = 8$ and $\eta = 10/9$ are

fixed. The values of β and weighting parameters μ and τ will be determined adaptively with respect to each experiment. We also try different starting points for each algorithm and find that all of them are insensitive towards starting points. Therefore, we initialize the starting points by zero in all experiments.

5.2 Test on "Shepp-Logan" phantom image

In this part, we test against the state-of-the-art solver RecPF [41] on phantom data. The Matlab package of RecPF is provided by the authors of [41] where the parameters including μ and τ are left to be its default settings. Before conducting numerical experiments, we must emphasize once again that RecPF [41], the most advanced solver, is only designed for under-sampled data corrupted by white Gaussian noise, but not the impulsive noise considered in this paper. We used RecPF for comparison because no solver available to solve the proposed model (1.3) at present. We choose the regularization parameters τ = 7E-6 and μ = 7E+2 which are suitable for three algorithms and $\beta = 20$ in this part. A "Shepp-Logan" phantom image and corrupted by impulsive noise phantom image with size 256×256 is shown in Figure 5.1 (a) and (c), and the Fourier domain sampling positions with 22 radial lines (sampling rate 9.36%) is listed in Figure 5.1 (b). Figure 5.2 displays the results within 3000 iterations derived by algorithms ADMM, sGS-ADMM_d and RecPF, respectively. At first glance, we see that tested algorithms work successfully except for RecPF which fails to produce acceptable reconstructions. More accurately, the RLNE value abstained by RecPF is 29.07% which is more than 1E+6 times larger than the one obtained by sGS-ADMM_d.

To further visibly illustrate the superiority of all the proposed algorithms, we draw the 5

(a) RLNE= 23.99% (b) RLNE=1.02E-6% (c) RLNE= 29.07%

Figure 5.2: *Reconstructing results of the* 256×256 *Shepp-Logan phantom image using 22 lines: (a) the reconstructed image by ADMM; (b) the reconstructed image by sGS-ADMM_d; (c) the reconstructed image by RecPF.*

times scaled difference images of Figure 5.2 to the ground true image. The compared heat images are listed in Figure 5.3. We observe from these figures that sGS-ADMM_d is the best one, followed by ADMM, and then sGS-ADMM_p, but RecPF is a loser.

To more visually examine the algorithms' performance, we also draw the curves of PSNR values and RLNE values with the increases of the running time and the iteration numbers in Figure 5.4. The stopping criteria in this test is chosen as the condition RLNE $\leq$ 1E-8 is achieved or the maximum iteration number 6000 is reached. As can be seen from the two plots at the lefthand side of this figure that, the curves derived by RecPF are always at the bottom after a certain number of iterations, which indicates that RecPF is the inferior among these algorithms. While we shift our attention to the both plots of the right hand sides, we see that the RLNE derived by sGS-ADMM_d decreases faster than that derived by

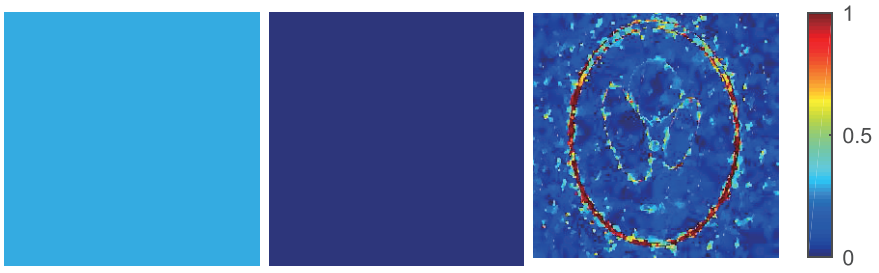


Figure 5.3: *The 5× scaled difference images of reconstructed images of ADMM, sGS-ADMM_d and RecPF to the ground truth image.*

other algorithms. Moreover, we also see that the curve derived by RecPF always at the top of both plots, which once again indicates that RecPF is the loser. From this simple test, we conclude that our proposed algorithms are effective in rebuilding phantom images corrupted by impulsive norse and perform better than RecPF significantly. It is worth mentioning that the dual algorithm has obvious advantages in both time and iteration number compared with the other two algorithms.

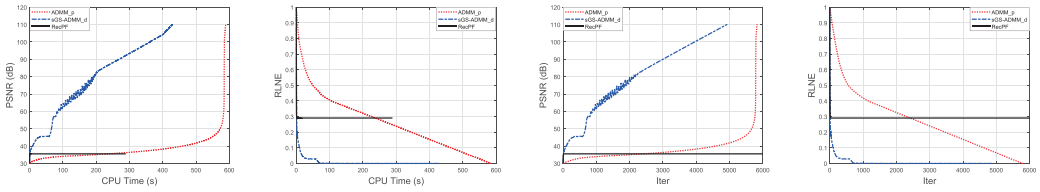
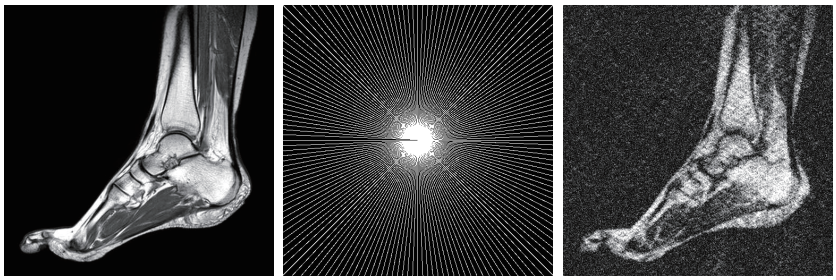


Figure 5.4: *Reconstruction results of the phantom image using 22 sampled lines: (a) PSNR versus the computational time; (b) RLNE versus the computational time; (c) PSNR versus the number of iterations; (e) RLNE versus the number of iterations.*

5.3 Test on a real foot MR image

In this subsection, we further investigate the validness of proposed algorithms by using a real foot MR image. The true image located at the leftmost side of Figure 5.5 is one slice of sagittal T1-weighted foot MR image with size 512 × 512 which is available at `http://www.mr-tip.com`. The impulsive noise corrupted image is listed in Figure 5.5 (c). In this test, we use the measurements from 84 radial sampling lines (sampling rate 17.74%) as shown at the middle plot in Figure 5.5. We choose the regularization parameters as μ=1E+3, τ=1E-3 and $\beta = 5$ which are better choices for these algorithms.

In this test, we choose the stopping criteria RLNE $\leq$ 1E-3, and set the maximum iteration number as 300. We run ADMM, sGS-ADMM_d and RecPF again and list the recovering images derived by each algorithm in Figure 5.6. From these figures, we clearly see that only the sGS-ADMM_d achieved acceptable reconstruction while ADMM is a complete failure. Moreover, the RLNE value derived by sGS-ADMM_d is RLNE = 6.23%, which is much smaller than that derived by other two algorithms. To further investigate the algorithms'



(a) The foot image (b) 84 sampled lines (c) Corrupted foot image

Figure 5.5: *(a) Real foot MR image with size* 512×512 *; (b) Fourier domain sampling positions with* 84 *radial lines; (c) corrupted by impulsive noise foot image.*

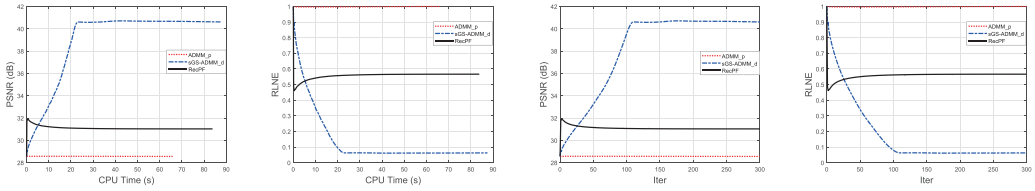
convergence behavior, we also draw some curves regrading PSNR and RLNE as the computing time and iteration numbers increase in Figure 5.7. From the plots (a) and (b) of Figure 5.7, we clearly see that sGS-ADMM_d is the fastest to increase the PSNR and decrease RLNE, while ADMM and RecPF are the slowest. Similarly, from the remaining two plots, we see that sGS-ADMM_d is also the winner in sense of requiring fewer number of iterations. From this test, we conclude that all the implemented algorithms are really efficient in recovering the foot MR image and sGS-ADMM_d is the most promising.

(a) RLNE= 99.9% (b) RLNE= 6.23% (c) RLNE= 56.64%

Figure 5.6: *Reconstructing results of the* 512×512 *real foot MR image using 84 lines: (a) reconstructed image by ADMM; (b) reconstructed image by sGS-ADMM_d; (c) reconstructed image by RecPF.*

6 Conclusions

The compressive sensing theories offered the possibility to reconstruct the MR images from highly undersampled data accurately. If the undersampled data is corrupted by the Gaussian noise, the reconverting task can be accomplished preferably and efficiently by the famous solver RecPF [41]. However, for the impulsive noise case, as far as we know, no solver can be employed in this literature. To remedy this slight neglect, this paper proposed a reconstruction model of being the sum of three non-differentiable terms. However, these non-differentiable terms caused more challenges for using some traditional algorithms. Based on the primal and dual problem's formulations, this paper employed a pair of ADMMs, gave



(a) (b) (c) (d)

Figure 5.7: *Reconstruction results of a* 512×512 *real foot MR image using 84 lines: (a) PSNR versus computational time; (b) RLNE versus computational time; (c) PSNR versus number of iterations; (e) RLNE versus number of iterations.*

their convergence, and tested their numerical performance. The numerical experiments on MR images demonstrated that the sGS-ADMM_d based on dual formulation performed better than ADMM and RecPF. Last but not least, the enhancement of sGS-ADMM_d make us to believe that the dual approach is valid for MR image reconstructing and possibility it may have its own extraordinary potency in other related problems.

Acknowledgements

We would like to thank the editor and two anonymous referees for their useful comments and suggestions which improved this paper greatly.

References

[1] M.V. Afonso, J.M. Bioucas-Dias and M.A. T. Figueiredo, An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems, *IEEE Trans. Image Process.* 20 (2011) 681–695.

[2] M.R. Bai, X.J. Zhang and Q.Q. Shao, Adaptive correction procedure for TVL1 image debluring under impulse noise, *Inverse Problems* 32 (2016): 085004.

[3] J.F. Cai, R.H. Chan and M. Nikolova, Fast two-phase image deblurring under impulse noise, *J. Math. Imaging Vision* 36 (2010) 46–53.

[4] E. Candés, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate information, *Commun. Pure Appl. Math.* 59 (2005) 1207–1233.

[5] E. Candés, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory* 52 (2006) 489–509.

[6] T. F. Chan and S. Esedoglu, Aspects of total variation regularized L1 function approximation, *SIAM J. Appl. Math.* 65 (2005) 1817–1837.

[7] L. Chen, D.F. Sun and K.-C. Toh, An efficient inexact symmetric Gauss-Seidel based majorized ADMM for high-dimensional convex composite conic programming, *Math. Program.* 161 (2017) 237–270.

[8] D. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory* 52 (2006) 1289–1306.

[9] Y. Ding, P. Li, Y. Xiao and H. Zhang, Symmetric Gauss-Seidel alternating direction methods of multipliers for sparse compressive sensing magnetic resonance imaging reconstruction based on dual formulation, submitted.

[10] Y. Ding and Y. Xiao, Symmetric Gauss-Seidel technique-based alternating direction methods of multipliers for transform invariant low-rank textures problem, *J. Math. Imaging Vision* 60 (2018) 1220–1230.

[11] M. Fazel, T.K. Pong, D.F. Sun and P. Tseng, Hankel matrix rank minimization with applications in system identification and realization, *SIAM J. Matrix Anal. Appl.* 34 (2013) 946–977.

[12] R. Glowinski and A. Marroco, Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires, *Revue française d'atomatique,Informatique Recherche Opérationelle. Analyse Numérique* 9 (1975) 41–76.

[13] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Comput. Math. Appl.* 2 (1976) 17–40.

[14] D. Gabay, Applications of the method of multipliers to variational inequalities in augmented Lagrangian methods: applications to the numerical solution of boundary-value problems, in: *Stud. Math. Appl.,* M. Fortin, R. Glowinski (eds.), vol.15, 1983, pp. 299–331.

[15] T. Goldstein and S. Osher, The Split Bregman method for L1 regularized problems, *SIAM J. Sci. Comput.* 2 (2009) 323–343.

[16] L. He, T.C. Chang, S. Osher, T. Fang and P. Speier, MR image reconstruction from undersampled data by using the iterative refinement procedure, *Proc. Appl. Math. Mech.* 7 (2007) 1011207–1011208.

[17] P. Li and Y. Xiao, An efficient algorithm for sparse inverse covariance matrix estimation based on dual formulation, *Comput. Statist. Data Anal.* 128 (2018) 292–307.

[18] P. Li, Y. Xiao and S. Lu, Efficient algorithms for sparse inverse covariance matrices estimation with explicit eigenvalue constraints, *J. Oper. Res. Soc. China,* DOI: 10.1007/s40305-021-00351-y.

[19] X.D. Li, D.F. Sun and K.-C. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions, *Math. Program.* 155 (2016) 333–373.

[20] X.D. Li, D.F. Sun and K.-C. Toh, A block symmtric Gauss-Seidel decomposition theorem for convex composite quadratic programming and its applications, *Math. Program.* 175 (2018) 395–418.

[21] X.D. Li, D.F. Sun and K.-C. Toh, QSDPNAL: A two-phase augmented Lagrangian method for convex quadratic semidefinite programming, *Math. Prog. Comp.* 10 (2018) 703–743.

[22] Q. Li, L. Shen, Y. Xu and N. Zhang, Multi-step fixed-point proximity algorithms for solving a class of optimization problems arising from image processing, *Adv. Comput. Math.* 41 (2015) 387-422.

[23] Q. Li, Y. Xu and N. Zhang, Two-step fixed-point proximity algorithms for multi-block separable convex problems, *J. Sci. Comput.* 70 (2016) 1204–1228.

[24] M. Lustig, D. Donoho and J.M. Pauly, Sparse MRI: the application of compressed sensing for rapid MR imaging, *Magn. Reson. Med.* 58 (2007) 1182–1195.

[25] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, An efficient algorithm for compressed MR imaging using total variation and wavelets, in: *IEEE Conference on Computer Vision and Pattern Recognition,* 2008, pp. 1–8.

[26] M. Nikolova, A variational approach to remove outliers and impulsive noise, *J. Math. Imaging Vision* 20 (2004) 99–120.

[27] T. Pock and A. Chambolle, Diagonal preconditioning for first order primal-dual algorithms in convex optimization, in: *Proc. IEEE International Conference on Computer Vision (ICCV),* 2011, pp. 1762–1769.

[28] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[29] L.I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Phys. D* 60 (1992) 259–268.

[30] N. Shen, Z.-F. Jin and Q. Wang, Nested alternating direction method of multipliers to low-rank and sparse-column matrices recovery, *Chinese Quart. J. Math.* 36 (2021): 90.

[31] D.F. Sun, K.-C. Toh and L.Q. Yang, A convergent proximal alternating direction method of multipliers for conic programming with 4-block constraints, *SIAM J. Optim.* 25 (2015), 882–915.

[32] J. Trzasko and A. Manduca, Highly undersampled magnetic resonance image reconstruction via homotopic l_0-minimization, *IEEE Trans. Med. Imaging,* 28 (2008) 106–121.

[33] S. Wang, Y. Xiao and Z. Jin, An efficient algorithm for batch images alignment with adaptive rank-correction term, *J. Comput. Appl. Math.* 346 (2019) 171–183.

[34] Y. Wang, J. Yang, W. Yin and Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM J. Imag. Science* 1 (2008) 248–272.

[35] Y. Xiao, L. Chen and D. Li, A generalized alternating direction method of multipliers with semi-proximal terms for convex composite conic programming, *Math. Prog. Comp.* 10 (2018) 533–555.

[36] Y. Xiao and H. Song, An inexact alternating directions algorithm for constrained total variation regularized compressive sensing problems, *J. Math. Imaging Vision* 44 (2012) 114–127.

[37] Y. Xiao, J. Yang and X. Yuan, A fast algorithm for total variation image reconstruction from random projections, *Inverse Probl. Imaging* 6 (2012) 547–563.

[38] J. Yang, G. Gu and S. Jiang, A TVSCAD approach for image deblurring with impulsive noise, *Inverse Problems* 33 (2017): 125008

[39] J. Yang, W. Yin, Y. Zhang and Y. Wang, A fast algorithm for edge preserving variational multichannel image restoration, *SIAM J. Imag. Science* 2 (2009) 569–592.

[40] J. Yang, Y. Zhang and W. Yin, An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise, *SIAM J. Sci. Comput.* 31 (2009) 2842–2865.

[41] J. Yang, Y. Zhang and W. Yin, A fast alternating direction method for TVL1-L2 signal reconstruction from partial fourier data, *IEEE Journal of Selected Topics in Signal Processing* 4 (2010) 288–297.

[42] W. Yin, D. Goldfarb, and S. Osher, The total variation regularized L1 model for multiscale decomposition, *Multiscale Model. Simul.* 6 (2006), 190–211.

Manuscript received 9 July 2021
revised 23 September 2021
accepted for publication 16 October 2021

Chenchen Lian
College of Geography and Environmental Science
Henan University, Kaifeng 475000, China
E-mail address: E-mail: macclian@163.com

Yanyun Ding
Department of Operations Research and Information Engineering
Beijing University of Technology, Beijing 100124, China
E-mail address: dingyanyunhenu@163.com

Yunhai Xiao
Center for Applied Mathematics of Henan Province
Henan University, Kaifeng 475000, China
E-mail address: E-mail: yhxiao@henu.edu.cn