Alternating Direction Method of Multipliers for Separable Convex Optimization of Real Functions in Complex Variables

Lu Li, 1,2 Xingyu Wang, 1 and Guoqiang Wang 2

1 School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
2 College of Fundamental Studies, Shanghai University of Engineering Science, Shanghai 201620, China

Correspondence should be addressed to Xingyu Wang; xywang@ecust.edu.cn

Received 31 July 2015; Revised 4 November 2015; Accepted 26 November 2015

1. Introduction

The augmented Lagrangian methods (ALMs) are a certain class of algorithms for solving constrained optimization problems, which were originally known as the method of multipliers in 1969 [1], and were studied much in the 1970s and 1980s as a good alternative to penalty methods. They have similarities to penalty methods in that they replace a constrained optimization problem by a series of unconstrained problems and add a penalty term to the objective. In particular, a variant of the standard ALMs that uses partial updates (similar to the Gauss-Seidel method for solving linear equations) known as the alternating direction method of multipliers (ADMM) gained some attention [2]. The ADMM has been extensively explored in recent years due to broad applications and empirical performance in a wide variety of problems such as image processing [3], applied machine learning and statistics [4], sparse optimizations, and other relevant fields [2]. Specifically, an advantage of the ADMM is that it can handle linear equality constraint of the form \( \{(x, z) \mid Ax + Bz = c\} \), which makes distributed optimization by variable splitting in a batch setting straightforward.

Recently, the convergence rates of order \( O(1/k) \) for the real case are considered under some additional assumptions; see, for example, [5–10]. For a survey on the ALMs and the ADMM, we refer to the references [1, 2, 11–16].

Compressed sensing (CS) is a signal processing technique for efficiently acquiring and reconstructing a signal by finding solutions to underdetermined linear systems. In the CS processing, the sparsity of a signal can be exploited to recover it from samples far fewer than required by the Shannon-Nyquist sampling theorem. The idea of CS got a new life in 2006 when Candès et al. [17] and Donoho [18] gave important results on the mathematical foundation of CS. This methodology has attracted much attention from applied mathematicians, computer scientists, and engineers for a variety of applications in biology [19], medicine [20], and radar [21], and so forth. Algorithms for signal reconstruction in a CS framework are expressed as sparse signal reconstruction algorithms. One of the most successful algorithms, known as basis pursuit (BP), is on the basis of constrained \( l_1 \)-norm minimization [22]. Most of the work is focused on the optimization in the real number field.
Signals in complex variables emerge in many areas of science and engineering and have become the objects of signal processing. There have been many works on the signal processing in complex variable. For example, independent component analysis (ICA) for separating complex-valued signals has found utility in many applications such as face recognition [23], analysis of functional magnetic resonance imaging [24], and electroencephalograph [25]. Taking impropriety and noncircularity of complex-valued signals into consideration, the right type of processing can give significant performance gains [26]. Methods of digital modulation schemes that produce improper complex signals have been studied in [27], such as binary phase shift keying and pulse amplitude modulation. In these researches, most nonlinear optimization methods use the first-order or second-order approximation of the objective function to create a new step or a descent direction, where the approximation is either updated or recomputed in every iteration. Unfortunately, all these functions do not satisfy the Cauchy-Riemann conditions. There exists no Taylor series of these functions do not satisfy the Cauchy-Riemann conditions. Since $A\alpha + B\beta - c = (A_0\alpha + A\beta - c_0 \alpha + A\beta - c) = (A\alpha + B\beta - c)$, the identity matrix of order $n$ is denoted by $I_n$. The one-norm and two-norm are denoted by $\| \cdot \|_1$ and $\| \cdot \|_2$, respectively. $z^R$ denotes the real composite $2n$-dimensional vector; for example, $z^R = (z^r_z, z^i_z)^T \in R^{2n}$, obtained by stacking $z_{r_z}$ on the top of $z_{i_z}$. The notation $\partial f$ denotes the set of all subgradients of $f$. 

2. Preliminaries

In this section, we first give some notations used. Vectors are denoted by lower case, for example, $z$, and matrices are denoted by capital letters, for example, $A$. The $k$th entry of a vector $z$ is denoted by $z_k$ and element $(i, j)$ of a matrix $A$ by $a_{ij}$. The subscripts ‘r’ and ‘i’ denote the real and imaginary parts, respectively; for example, $z_{r_z} = \text{Re}[z]$ and $A_{i_m} = \text{Im}[A]$. The superscripts $^T, \cdot, ^H$ and $^{-1}$ are used for the transpose, conjugate, Hermitian conjugate, and matrix inverse. The dom$f$ denotes the domain of function $f$. The identity matrix of order $n$ is denoted by $I_n$. The one-norm and two-norm are denoted by $\| \cdot \|_1$ and $\| \cdot \|_2$, respectively. $z^R$ denotes the real composite $2n$-dimensional vector; for example, $z^R = (z^r_z, z^i_z)^T \in R^{2n}$, obtained by stacking $z_{r_z}$ on the top of $z_{i_z}$. The notation $\partial f$ denotes the set of all subgradients of $f$.

2.1. Wirtinger Calculus. We next recall some well-known concepts and results on the complex analysis and Wirtinger calculus which will be used in our future analysis. A comprehensive treatment of Wirtinger calculus can be found in [33, 34].

Define the complex augmented vector as follows:

$$z = (z^T, z^H)^T \in C^{2n}, \quad (1)$$

which is obtained by stacking $z$ on the top of its complex conjugate $\overline{z}$. The complex augmented vector $\overline{z} \in C^{2n}$ is related to the real composite vector $z^R \in R^{2n}$ as $\overline{z} = J_n z^R$ and $z^R = (1/2)J_n^H \overline{z}$, where the real-to-complex transformation

$$J_n = \begin{pmatrix} I_n & j I_n \\ I_n & -j I_n \end{pmatrix} \in C^{2nx2n} \quad (2)$$

is unitary up to a factor of $2$; that is, $J_n J_n^H = I_n^2 I_n = 2I_{2n}$. The linear map $J_n$ is an isomorphism map from $R^{2n}$ to $C^{2n}$ and its inverse is given by $(1/2)J_n^H$.

**Lemma 1.** Let $A \in C^{p \times n}$, $B \in C^{p \times m}$, and $c \in C^{p}$. Then

$$Ax + Bz = c \iff A\overline{x} + B\overline{z} = \overline{c}, \quad (3)$$

where

$$A = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \in C^{2px2n}, \quad (4)$$

$$B = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} \in C^{2px2m}. \quad (5)$$

**Proof.** Since

$$Ax + Bz - c = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} x \\ \overline{x} \end{pmatrix} + \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} z \\ \overline{z} \end{pmatrix} - \begin{pmatrix} c \\ \overline{c} \end{pmatrix}$$

$$= \begin{pmatrix} Ax + Bz - c \\ A\overline{x} + B\overline{z} - \overline{c} \end{pmatrix} \quad (6)$$

The outline of the paper is as follows. In Section 2, we recall some elementary theories and methods of the complex analysis and Wirtinger calculus. The ADMM for complex separable convex optimization and its convergence are presented in Section 3. In Section 4, we study the BP algorithm for the equality-constrained $l_1$ minimization problem in the form of ADMM. In Section 5, some numerical simulations are provided. Finally, some conclusions are drawn in Section 6.
then we have
\[ \begin{align*}
Ax + Bz &= c \iff \begin{cases} \Delta x + B \bar{z} &= c, \end{cases} \\
\text{This completes the proof.} \end{align*} \tag{6}
\]

Consider a complex-valued function
\[ f(z) = u(\re z, \im z) + jv(\re z, \im z), \tag{7} \]
where \( z = \re z + j\im z, f : \mathbb{C}^n \to \mathbb{C}, \) and \( u, v : \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}. \) The definition of complex differentiability requires that the derivatives be defined as the limit
\[ f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \tag{8} \]
is independent of the direction in which \( \Delta z \) approaches zero in the complex plane. This requires that the Cauchy-Riemann equations
\[ \begin{align*}
\frac{\partial u}{\partial \re z} &= \frac{\partial v}{\partial \im z}, \\
\frac{\partial u}{\partial \im z} &= -\frac{\partial v}{\partial \re z} \tag{9}
\end{align*} \]
should be satisfied \[35\]. These conditions are necessary for \( f(z) \) to be complex-differentiable. A function which is complex-differentiable on its entire domain is called analytic or holomorphic. Clearly, the Cauchy-Riemann conditions do not hold for real-valued functions which are \( v(\re z, \im z) \equiv 0, \) and thus cost functions are not analytic. These conditions imply complex differentiability which are quite stringent and impose a strong structure on \( u(\re z, \im z) \) and \( v(\re z, \im z) \) and, consequently, on \( f(z). \) Obviously, most cost functions do not satisfy the Cauchy-Riemann equations as these functions are typically \( f : \mathbb{C}^n \to \mathbb{R} \) with \( v(\re z, \im z) = 0. \)

To overcome such a difficulty, a sound approach in \[33\] relaxes this strong requirement for differentiability and defines a less stringent form for the complex domain. More importantly, it describes how this new definition can be used for defining complex differential operators that allow computation of derivatives in a very straightforward manner in the complex number domain, by simply using real differentiation results and procedures. A function is called real differentiable when \( u(\re z, \im z) \) and \( v(\re z, \im z) \) are differentiable as the functions of real-valued variables \( \re z \) and \( \im z. \) Then, one can write the two real variables as \( \re z = (z + \bar{z})/2 \) and \( \im z = -j(z - \bar{z})/2 \) and use the chain rule to derive the operators for differentiation given in the theorem below. The key point in the derivation is regarding the two variables \( z \) and \( \bar{z} \) as independent variables, which is also the main approach allowing us to make use of the elegance of Wirtinger calculus.

In view of this, we consider the function \(7\) as \( f : \mathbb{R}^{2n} \to \mathbb{C} \) by rewriting it as \( f(z) = f(u, v) \) and make use of the underlying \( \mathbb{R}^{2n} \) structure. The function \( f(z) \) can be regarded as either \( f(\re z, \im z) \) with variables \( \re z \) and \( \im z \) or \( f(z, \bar{z}) \) with variables \( z \) and \( \bar{z}, \) and it can be simply written as \( f(z). \) The functions may take different forms; however, they are equally valued. For convenience, we use the same function \( f \) to denote them as follows:
\[ f(z, \bar{z}) = f(\re z, \im z) = f(z). \tag{10} \]
The main result in this context is stated by Brandwood in \[36\].

**Theorem 2.** Let \( f : \mathbb{R}^{2n} \to \mathbb{C} \) be a function of real variables \( \re z \) and \( \im z \) such that \( f(z) = f(\re z, \im z), \) where \( z = \re z + j\im z, \) and that \( f \) is analytic with respect to \( z \) and \( \bar{z} \) independently. Then, consider the following:

1. The partial derivatives
\[
\begin{align*}
\frac{\partial f}{\partial z} &= \frac{1}{2} \left( \frac{\partial f}{\partial \re z} - j \frac{\partial f}{\partial \im z} \right), \\
\frac{\partial f}{\partial \bar{z}} &= \frac{1}{2} \left( \frac{\partial f}{\partial \re z} + j \frac{\partial f}{\partial \im z} \right)
\end{align*} \tag{11}
\]
can be computed by treating \( \bar{z} \) as a constant in \( f \) and \( z \) as a constant, respectively.

2. A necessary and sufficient condition for \( f \) to have a stationary point is that
\[ \frac{\partial f}{\partial \bar{z}} = 0. \tag{12} \]
Similarly,
\[ \frac{\partial f}{\partial z} = 0 \tag{13} \]
is also a necessary and sufficient condition.

As for the applications of Wirtinger derivatives, we consider the following two examples, which will be used in the subsequent analysis.

**Example 3.** Consider the real function in complex variables as follows:
\[ f(x, z) = 2 \Re \left\{ y^H (Ax + Bz - c) \right\}, \tag{14} \]
where \( x \in \mathbb{C}^n, z \in \mathbb{C}^m, y \in \mathbb{C}^p, c \in \mathbb{C}^p, A \in \mathbb{C}^{pn}, \) and \( B \in \mathbb{C}^{pm}. \)

It follows from Theorem 2 that
\[ \begin{align*}
\frac{\partial f}{\partial x} &= \frac{1}{2} \left( \frac{\partial f}{\partial \re x} - j \frac{\partial f}{\partial \im x} \right) \\
&= (y^H A_{\re} + y^H A_{\im}) + j (y^H A_{\im} - y^H A_{\re}) \\
&= y^H A. \tag{15}
\end{align*} \]

Similarly, we have
\[ \begin{align*}
\frac{\partial f}{\partial z} &= \frac{1}{2} \left( \frac{\partial f}{\partial \re z} - j \frac{\partial f}{\partial \im z} \right) \\
&= (y^H B_{\re} + y^H B_{\im}) + j (y^H B_{\im} - y^H B_{\re}) = y^H B \tag{16}
\end{align*} \]
Example 4. Consider the real function in complex variables as follows:

$$f(x, z) = \|Ax + Bz - c\|_2^2,$$  \hspace{1cm} (17)

where \(x \in \mathbb{C}^n, z \in \mathbb{C}^m, A \in \mathbb{C}^{n \times n}, \text{ and } B \in \mathbb{C}^{m \times m}\).

We have

$$f(x, z) = (Ax + Bz - c)^H(Ax + Bz - c)$$

$$= (r_{re} - jr_{im})(r_{re} + jr_{im}) = r_{re}^2 + r_{im}^2$$

where \(r = Ax + Bz - c\). Then

$$\frac{df}{dx} = \frac{1}{2} \left( \frac{\partial f}{\partial x_{re}} - j \frac{\partial f}{\partial x_{im}} \right)$$

$$= \frac{1}{2} \left( \frac{\partial (r_{re}^2 + r_{im}^2)}{\partial x_{re}} - j \frac{\partial (r_{re}^2 + r_{im}^2)}{\partial x_{im}} \right)$$

$$= (r_{re}A_{re} + r_{im}A_{im}) + j(r_{re}A_{im} - r_{im}A_{re}) = r^H A,$$

$$\frac{df}{dz} = \frac{1}{2} \left( \frac{\partial f}{\partial z_{re}} - j \frac{\partial f}{\partial z_{im}} \right)$$

$$= \frac{1}{2} \left( \frac{\partial (r_{re}^2 + r_{im}^2)}{\partial z_{re}} - j \frac{\partial (r_{re}^2 + r_{im}^2)}{\partial z_{im}} \right)$$

$$= (r_{re}B_{re} + r_{im}B_{im}) + j(r_{re}B_{im} - r_{im}B_{re}) = r^H B.$$  \hspace{1cm} (19)

2.2. Convex Analysis in the Complex Number Domain. In order to meet the demands of next work, we give some definitions in the complex number domain.

Definition 5 (see [37]). A set \(A\) is convex if the line segment between any two points in \(A\) lies in \(A\); that is, for any \(z_1, z_2 \in A\) and any \(\theta \in \mathbb{R} \) with \(0 \leq \theta \leq 1\), then

$$\theta z_1 + (1 - \theta) z_2 \in A.$$  \hspace{1cm} (20)

Definition 6 (see [34]). Let \(z = z_{re} + jz_{im} \in \mathbb{C}^m\). The complex gradient operator \(\partial / \partial z\) is defined by

$$\frac{\partial}{\partial z} = \left( \frac{\partial}{\partial z_{re}}, \frac{\partial}{\partial z_{im}} \right).$$  \hspace{1cm} (21)

The linear map \(J_n\) also defines a one-to-one correspondence between the real gradient \(\partial / \partial z^R\) and the complex gradient \(\partial / \partial z\); namely,

$$\frac{\partial}{\partial z^R} = J_n^T \frac{\partial}{\partial z}.$$  \hspace{1cm} (22)

For real function in complex variable \(f : \mathbb{C}^n \rightarrow \mathbb{R}\), it has an equivalent form as \(f(z) = u(z_{re}, z_{im})\) according to (10). So we can similarly extend some concepts of the functions in the real number domain [38, 39] to the complex number domain.

Definition 7. A real function in complex variable \(f : \mathbb{C}^n \rightarrow \mathbb{R}\) is convex if \(\text{dom} f\) is a convex set and if for any \(x, y \in \text{dom} f\) and any \(\theta \in \mathbb{R}\) with \(0 \leq \theta \leq 1\), then

$$f(\theta x + (1 - \theta) y) \leq \theta f(x) + (1 - \theta) f(y).$$  \hspace{1cm} (23)

Definition 8. A real function in complex variable \(f : \mathbb{C}^n \rightarrow \mathbb{R}\) is proper if its effective domain is nonempty and it never attains \(-\infty\).

Definition 9. A real function in complex variable \(f : \mathbb{C}^n \rightarrow \mathbb{R}\) is closed if, for each \(\alpha \in \mathbb{R}\), the sublevel set \([x \in \text{dom} f | f(x) \leq \alpha]\) is a closed set.

Definition 10. Given any real function in complex variable \(f : \mathbb{C}^n \rightarrow \mathbb{R} \cup (+\infty)\), a vector \(v \in \mathbb{C}^n\) is said to be a subgradient of \(f\) at \(z_0\) if

$$f(z) \geq f(z_0) + 2 \text{Re} \{v^H(z - z_0)\}.$$  \hspace{1cm} (24)

3. ADMM for Convex Separable Optimization

In this section, we will first recall the ADMM for real convex separable optimization. Then we will study the ADMM for convex separable optimization of real functions in complex variables.

3.1. ADMM for Real Convex Separable Optimization. The ADMM has been well studied for the following linearly constrained separable convex programming whose objective function is separated into two individual convex functions with nonoverlapping variables as follows:

$$\text{minimize} \quad \{f(x) + g(z) : Ax + Bz = c, \ x \in \mathbb{X}_1, \ z \in \mathbb{X}_2\},$$  \hspace{1cm} (25)

where \(\mathbb{X}_1 \subset \mathbb{R}^n\) and \(\mathbb{X}_2 \subset \mathbb{R}^m\) are closed convex sets; \(A \in \mathbb{R}^{n \times m}\) and \(B \in \mathbb{R}^{m \times m}\) are given matrices; \(c \in \mathbb{R}^n\) is a given vector; and

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \cup (+\infty)$$

and

$$g : \mathbb{R}^m \rightarrow \mathbb{R} \cup (+\infty)$$

are proper, closed, and convex functions.

More specifically, the Lagrangian function and the augmented Lagrangian function of (25) are given by

$$L_0(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c),$$  \hspace{1cm} (26)

$$L_\rho(x, y, z) = f(x) + g(z) + y^T(Ax + Bz - c)$$

$$+ \frac{\rho}{2} \|Ax + Bz - c\|_2^2,$$  \hspace{1cm} (27)

respectively. Then the iterative scheme of the ADMM for solving (25) is given by

$$x^{k+1} = \arg\min_x L_\rho(x, z^k, y^k),$$

$$z^{k+1} = \arg\min_z L_\rho(x^{k+1}, z, y^k),$$  \hspace{1cm} (28)

$$y^{k+1} = y^k + \rho(Ax^{k+1} + Bz^{k+1} - c).$$

Without loss of generality, we give the following two assumptions.

Assumption II. The (extended-real-valued) functions \(f : \mathbb{R}^n \rightarrow \mathbb{R} \cup (+\infty)\) and \(g : \mathbb{R}^m \rightarrow \mathbb{R} \cup (+\infty)\) are proper, closed, and convex.
Assumption 12. The Lagrangian function \( L_0 \) has a saddle point; that is, there exists \((x^*, z^*, y^*)\), not necessarily unique, for which
\[
L_0(x^*, z^*, y) \leq L_0(x^*, z^*, y^*) \leq L_0(x, z, y^*)
\]
holds for all \( x, z, \) and \( y \).

The convergence of the ADMM for real convex separable optimization is established in the following theorem.

**Theorem 13** (Section 3.2.1 in [2]). Under Assumptions 11 and 12, the ADMM iterates (28) satisfy the following.

1. **Residual Convergence.** \( r^k = Ax^k + Bz^k - c \to 0 \) as \( k \to \infty \); that is, the iterates approach feasibility.

2. **Objective Convergence.** \( f(x^k) + g(z^k) \to f(x^*) + g(z^*) \) as \( k \to \infty \); that is, the objective function of the iterates approaches the optimal value.

\[
\min \quad \{ f(x, x) + g(z, z) : Ax + Bz = c, \ x \in \chi_1, \ z \in \chi_2 \}.
\]  

(3) Dual Variable Convergence. \( y^k \to y^* \) as \( k \to \infty \), where \( y^* \) is a dual optimal point.

### 3.2. ADMM for Complex Convex Separable Optimization.

According to (10), we can consider the real functions in complex variables \( f : C^n \to R \cup \{+\infty\} \) and \( g : C^m \to R \cup \{+\infty\} \). Then, the convex separable optimization of real functions in complex variables becomes

\[
\min \quad \{ f(x) + g(z) : Ax + Bz = c, \ x \in \chi_1, \ z \in \chi_2 \},
\]

where \( f \) and \( g \) are proper, closed, and convex functions; \( \chi_1 \subset C^n \) and \( \chi_2 \subset C^m \) are closed convex sets; \( A \in C^{pn} \) and \( B \in C^{pm} \) are given matrices; and \( c \in C^p \) is a given vector.

From (10) and Lemma 1, we can conclude that the complex convex separable optimization (30) is equivalent to the following convex separable optimization problem:

\[
\min \quad \{ f(x, x) + g(z, z) : A^* x + B^* z = c, \ x \in \chi_1, \ z \in \chi_2 \}.
\]  

(31)

The Lagrangian function of (31) is
\[
L_p(x, z, y) = f(x, x) + g(z, z) + y^H (A^* x + B^* z - c) + \frac{\rho}{2} \| A^* x + B^* z - c \|^2_z,
\]

where \( y \in C^p \). Then, the augmented Lagrangian function of (31) is
\[
L_p(x, z, y) = f(x, x) + g(z, z) + y^H (A^* x + B^* z - c) + \frac{\rho}{2} \| A^* x + B^* z - c \|^2_z,
\]

(32)

where \( y \in C^p \). Then, the augmented Lagrangian function of (31) is
\[
L_p(x, z, y) = f(x, x) + g(z, z) + y^H (A^* x + B^* z - c) + \frac{\rho}{2} \| A^* x + B^* z - c \|^2_z,
\]

(33)

where \( \rho > 0 \) is called the penalty parameter. The ADMM for complex convex separable optimization is composed of the iterations
\[
x^{k+1} = \arg \min_x L_p(x, z^k, y^k),
\]

\[
z^{k+1} = \arg \min_z L_p(x^{k+1}, z, y^k),
\]

\[
y^{k+1} = y^k + \rho (A^{k+1} x^k + B^{k+1} z - c).
\]  

(34)

Let \( r = Ax + Bz - c \). Then we have
\[
2 \text{Re} \{ y^H r \} + \rho \| r \|^2_z = \rho \| r + u \|^2_z - \rho \| u \|^2_z,
\]

(35)

where \( u = (1/\rho) y \) is the scaled dual variable. Using the scaled dual variable, we can express the ADMM iterations (34) as
\[
x^{k+1} = \arg \min_x \{ f(x, x) + \rho \| A^* x + B^* z - c + u^k \|^2_z \},
\]

\[
z^{k+1} = \arg \min_z \{ g(z, z) + \rho \| A^{k+1} x - B^* z - c + u^k \|^2_z \},
\]

\[
u^{k+1} = u^k + A^{k+1} x^{k+1} + B^* z^{k+1} - c.
\]  

### 3.3. Optimality Conditions.

The necessary and sufficient optimality conditions for the ADMM problem (31) are primal feasibility,

\[
Ax^* + Bz^* - c = 0,
\]

(37)

and dual feasibility,

\[
0 \in \partial f(x^*) + (y^*)^H A,
\]

\[
0 \in \partial g(z^*) + (y^*)^H B.
\]

(38)

(39)

Because \( z^{k+1} \) minimizes \( L_p(x^{k+1}, z, y^k) \), we have
\[
0 \in \partial g(z^{k+1}) + (y^k)^H B + \rho (r^{k+1})^H B
\]

\[
= \partial g(z^{k+1}) + (y^k)^H B.
\]

(40)

This means that \( z^{k+1} \) and \( y^{k+1} \) always satisfy (39); thus attaining optimality leads to satisfying (37) and (38).
Because $x^{k+1}$ minimizes $L_\rho(x, z^k, y^k)$, we have

$$0 \in \partial f(x^{k+1}) + (y^k)^H A + \rho \left( Ax^{k+1} + B z^k - c \right)^H A$$

$$= \partial f(x^{k+1}) + (y^k)^H A + \rho (r^{k+1})^H A$$

(41)

or equivalently

$$\rho \left( B (z^k - z^{k+1}) \right)^H A \in \partial f(x^{k+1}) + (y^{k+1})^H A.$$  

(42)

From (38), $s^{k+1} = \rho (B (z^{k+1} - z^k))^H A$ can be viewed as a residual for the dual feasibility condition. By (37), $r^{k+1} = Ax^{k+1} + B z^k - c$ can be viewed as a residual for the primal feasibility condition. These two residuals converge to zero as the ADMM proceeds.

3.4. Convergence. Similar to the ADMM for separable convex optimization in the real number domain, we can establish the convergence of the ADMM for complex separable convex optimization.

In this paper, we make the following two assumptions on the separable convex optimization of the real functions in complex variables.

Assumption 14. The (extended-real-valued) functions $f : C^n \to \mathbb{R} \cup \{+\infty\}$ and $g : C^m \to \mathbb{R} \cup \{+\infty\}$ are proper, closed, and convex.

Assumption 15. The Lagrangian function $L_0$ (32) has a saddle point; that is, there exists $(x^*, z^*, y^*)$, not necessarily unique, for which

$$L_0(x^*, z^*, y) \leq L_0(x, z, y^*) \leq L_0(x, z, y^*)$$  

(43)

holds for all $x, z, y$.

Theorem 16. Under Assumptions 14 and 15, the ADMM iterations (36) have the following conclusions.

(1) Residual Convergence. $r^k = Ax^k + B z^k - c \to 0$ as $k \to \infty$; that is, the iterates approach feasibility.

(2) Objective Convergence. $f(x^k) + g(z^k) \to f(x^*) + g(z^*)$ as $k \to \infty$; that is, the objective function of the iterates approaches the optimal value.

(3) Dual Variable Convergence. $y^k \to y^*$ as $k \to \infty$, where $y^*$ is a dual optimal point.

Proof. Let $(x^*, z^*, y^*)$ be the saddle point for $L_0$ and $q^* = f(x^*) + g(z^*)$. Then we have

$$L_0(x^*, z^*, y^*) \leq L_0(x^{k+1}, z^{k+1}, y^*).$$  

(44)

Since $Ax^* + B z^* - c = 0$ and $L_0(x^*, z^*, y^*)$ is equivalent to $q^*$, then we have

$$q^* \leq q^{k+1} + 2 \text{Re} \left\{ (y^*)^H (Ax^{k+1} + B z^{k+1} - c) \right\}.$$  

(45)

From Theorem 2 and Examples 3 and 4, we get

$$\frac{\partial L_\rho}{\partial x} = \frac{\partial f(x^{k+1}, \bar{x}^{k+1})}{\partial x} + (y^k)^H A + \rho \left( Ax^{k+1} + B z^k - c \right)^H A.$$  

(46)

Note that $L_\rho$ is a real-valued function; then we get

$$\frac{\partial L_\rho}{\partial x} = \left( \frac{\partial L_\rho}{\partial x} \right).$$  

(47)

By (36), $x^{k+1}$ minimizes $L_\rho(x, z^k, y^k)$ for $f$ is convex, which is subdifferentiable, and so is $L_\rho(x, z, y)$. Based on Theorem 2, the optimality condition is

$$0 \in \partial f(x^{k+1}, z^k, y^k)$$

$$= \partial f(x^{k+1}, z^{k+1}) + (y^k)^H A + \rho \left( Ax^{k+1} + B z^k - c \right)^H A.$$  

(48)

Since

$$y^{k+1} = y^k + \rho \left( Ax^{k+1} + B z^{k+1} - c \right),$$  

(49)

we have

$$0 \in \partial f(x^{k+1}, z^{k+1}) + (y^{k+1} - \rho B(z^{k+1} - z^k))^H A,$$  

(50)

which implies that $x^{k+1}$ minimizes

$$f(x, \bar{x}) + 2 \text{Re} \left\{ (y^{k+1} - \rho B(z^{k+1} - z^k))^H Ax \right\}.$$  

(51)

Similarly, we may have that $z^{k+1}$ minimizes

$$g(z, \bar{z}) + 2 \text{Re} \left\{ (y^{k+1})^H Bz \right\}.$$  

(52)

From (51) and (52), we have

$$f(x^{k+1}, \bar{x}^{k+1}) + 2 \text{Re} \left\{ (y^{k+1} - \rho B(z^{k+1} - z^k))^H Ax^{k+1} \right\} \leq f(x^*, \bar{x}^*) + 2 \text{Re} \left\{ (y^*)^H Bz^* \right\},$$  

(53)

$$g(z^{k+1}, \bar{z}^{k+1}) + 2 \text{Re} \left\{ (y^{k+1})^H Bz^{k+1} \right\} \leq g(z^*, \bar{z}^*),$$  

(54)

From (53) and $Ax^* + B z^* - c = 0$, we can make the conclusion that

$p^{k+1} - p^* \leq 2 \text{Re} \left\{ - (y^{k+1})^H r^{k+1} \right\} \leq \left( \rho B(z^{k+1} - z^k))^H (-r^{k+1} + B(z^{k+1} - z^*)) \right\},$  

(54)

where $r^{k+1} = Ax^{k+1} + B z^{k+1} - c.$
Adding (45) and (54), we get
\[
4 \Re \left\{ \left( y^{k+1} - y^* \right)^H r^{k+1} - \left( \rho B \left( z^{k+1} - z^k \right) \right)^H r^{k+1} \right\}
+ \left( \rho B \left( z^{k+1} - z^k \right) \right)^H B \left( z^{k+1} - z^k \right) \leq 0.
\] (55)

Let
\[
w^k = \left( \frac{1}{\rho} \right) \left\| y^k - y^* \right\|_2^2 + \rho \left\| B (z^k - z^*) \right\|_2^2.
\] (56)

By following manipulation and rewriting of (55), we have
\[
w^{k+1} \leq w^k - \rho \left\| r^{k+1} \right\|_2^2 - \rho \left\| B (z^{k+1} - z^k) \right\|_2^2.
\] (57)

Rewriting the first term of (55) as
\[
4 \Re \left\{ \left( y^{k+1} - y^* \right)^H r^{k+1} - \left( y^k + \rho r^{k+1} - y^* \right)^H r^{k+1} \right\}
+ \left( y^k - y^* \right)^H r^{k+1} \right\} + \rho \left\| r^{k+1} \right\|_2^2 + \rho \left\| r^{k+1} \right\|_2^2
\] (58)

and then substituting
\[
y^{k+1} = \left( \frac{1}{\rho} \right) \left( y^{k+1} - y^k \right)
\] (59)

into the first two terms in (58) give
\[
2 \Re \left\{ \left( \frac{2}{\rho} \right) \left( y^k - y^* \right)^H \left( y^{k+1} - y^k \right) \right\}
+ \left( \frac{1}{\rho} \right) \left\| y^{k+1} - y^k \right\|_2^2 + \rho \left\| r^{k+1} \right\|_2^2.
\] (60)

Since
\[
y^{k+1} - y^k = (y^{k+1} - y^*) - (y^k - y^*),
\] (61)

can be expressed as
\[
\frac{1}{\rho} \left( \left\| y^{k+1} - y^* \right\|_2^2 - \left\| y^k - y^* \right\|_2^2 \right) + \rho \left\| r^{k+1} \right\|_2^2.
\] (62)

Now let us regroup the remaining terms, that is,
\[
\rho \left\| r^{k+1} \right\|_2^2 - 2 \Re \left\{ 2 \rho \left( B \left( z^{k+1} - z^k \right) \right)^H r^{k+1} \right\}
+ 2 \rho \left( B \left( z^{k+1} - z^k \right) \right)^H B \left( z^{k+1} - z^k \right) \right\}.
\] (63)

where \( \rho \left\| r^{k+1} \right\|_2^2 \) is taken from (62). Substituting
\[
z^{k+1} - z^* = (z^{k+1} - z^k) + (z^k - z^*)
\] (64)

into the last term in (63) yields
\[
\rho \left\| r^{k+1} - B \left( z^{k+1} - z^k \right) \right\|_2^2 + \rho \left\| B \left( z^{k+1} - z^k \right) \right\|_2^2
+ 4 \rho \Re \left\{ \left( B \left( z^{k+1} - z^k \right) \right)^H B \left( z^k - z^* \right) \right\}
\] (65)

and substituting
\[
z^{k+1} - z^k = (z^{k+1} - z^*) - (z^k - z^*)
\] (66)

into the last two terms in (63), we get
\[
\rho \left\| r^{k+1} - B \left( z^{k+1} - z^k \right) \right\|_2^2
+ \rho \left\| B \left( z^{k+1} - z^k \right) \right\|_2^2.
\] (67)

It implies that (55) can be expressed as
\[
w^k - w^{k+1} \geq \rho \left\| r^{k+1} - B \left( z^{k+1} - z^k \right) \right\|_2^2.
\] (68)

To obtain (57), it suffices to show that the middle term
\[
-4 \Re \left\{ \left( r^{k+1} \right)^H B \left( z^{k+1} - z^k \right) \right\}
\] (69)

of the expanded right-hand side of (68) is positive. To understand this, by reviving that \( z^{k+1} \) minimizes \( g(z, \overline{z}) + 2 \Re \{ (y^{k+1})^H B z \} \) and \( z^k \) minimizes \( g(z, \overline{z}) + 2 \Re \{ (y^k)^H B z \} \), we can add
\[
g \left( z^{k+1}, z^{k+1} \right) + 2 \Re \left\{ \left( y^{k+1} \right)^H B z^{k+1} \right\}
\leq g \left( z^k, z^k \right) + 2 \Re \left\{ \left( y^k \right)^H B z^k \right\},
\] (70)

\[
g \left( z^k, \overline{z} \right) + 2 \Re \left\{ \left( y^k \right)^H B \overline{z} \right\}
\leq g \left( z^{k+1}, z^{k+1} \right) + 2 \Re \left\{ \left( y^{k+1} \right)^H B z^{k+1} \right\}
\] to obtain that
\[
2 \Re \left\{ \left( y^{k+1} - y^k \right)^H B \left( z^{k+1} - z^k \right) \right\} \leq 0.
\] (71)

Since \( \rho > 0 \), if we substitute
\[
y^{k+1} - y^k = \rho r^{k+1},
\] (72)

we can get (57).

This means that \( w^k \) decreases in each iteration by an amount depending on the norm of the residual and on the change in \( z \) over one iteration. Since \( w^k \leq w^0 \), it follows that \( y^k \) and \( B \) are bounded. Iterating the inequality above gives
\[
\rho \sum_{k=0}^{\infty} \left( \left\| r^{k+1} \right\|_2^2 + \left\| B \left( z^{k+1} - z^k \right) \right\|_2^2 \right) \leq w_0,
\] (73)

implying that \( r^k \to 0 \) and \( B(z^{k+1} - z^k) \to 0 \) as \( k \to \infty \). From (45), we have
\[
f \left( x^k, \overline{x}^k \right) + g \left( z^k, \overline{z} \right) \to f \left( x^*, \overline{x}^* \right) + g \left( z^*, \overline{z}^* \right)
\] (74)
as \( k \to \infty \). Furthermore, since \( w^k \to 0 \) as \( k \to \infty \), we have \( y^k \to y^* \) as \( k \to \infty \). This completes the proof. □
3.5. Stopping Criterion. We can find that
\[ -r^{k+1} + B(z^{k+1} - z^k) = -A(x^{k+1} - x^*). \] (75)
Substituting this into (54), we get
\[ p^{k+1} - p^* \leq 2 \text{Re} \left\{ -(y^{k+1})^H r^{k+1} + (x^{k+1} - x^*)^H s^{k+1} \right\}. \] (76)
This means that when the two residuals are small, the error must be small. Thus an appropriate termination criterion is that the primal residuals \( r^{k+1} \) and dual residuals \( s^{k+1} \) are small simultaneously; that is, \( \|r^{k+1}\|_2 \leq \varepsilon_{\text{primal}} \) and \( \|s^{k+1}\|_2 \leq \varepsilon_{\text{dual}} \), where \( \varepsilon_{\text{primal}} \) and \( \varepsilon_{\text{dual}} \) are tolerances for the primal and dual feasibility, respectively.

4. Basis Pursuit with Complex ADMM
Consider the equality-constrained \( l_1 \) minimization problem in the complex number domain
\[
\begin{aligned}
\text{minimize} & \quad \|x\|_1 : Ax = b, \ x \in \mathbb{C}^n, \\
\end{aligned}
\] (77)
where \( A \in \mathbb{C}^{m \times n} \) is a given matrix, and \( b \in \mathbb{C}^m \) is a given vector.
Recall that
\[
\|x\|_1 = \sum_{k=1}^{n} |x_k| = \sum_{k=1}^{n} \sqrt{(x_k)_{\text{re}}^2 + (x_k)_{\text{im}}^2},
\] (78)
Then
\[
\|(z)\|_1 = 2 \|x\|_1.
\] (79)
In the form of the ADMM, the BP method can be expressed as
\[
\begin{aligned}
\text{minimize} & \quad \{ f(x) + \|z\|_1 : x = z, \ x, z \in \mathbb{C}^n, \},
\end{aligned}
\] (80)
where \( f \) is the indicator function of \( X = \{x \in \mathbb{C}^n \mid Ax = b\} \); that is, \( f(x) = 0 \) for \( x \in X \) and \( f(x) = +\infty \) otherwise. Then, with the idea in [40], the ADMM iterations are provided as follows:
\[
\begin{aligned}
x^{k+1} &= \arg \min_x \left\{ f(x) + \rho \|x - z^k + u^k\|_2^2 \right\}, \\
z^{k+1} &= \arg \min_z \left\{ \|z\|_1 + \rho \|x^{k+1} - z + u^k\|_2^2 \right\}, \\
u^{k+1} &= u^k + x^{k+1} - z^{k+1}.
\end{aligned}
\] (81)
The \( x \)-update, which involves solving a linearly constrained minimum Euclidean norm problem, can be written as
\[
\begin{aligned}
\text{minimize} & \quad \left\{ \frac{1}{2} \|x^{k+1} - (z^k - u^k)\|_2^2 : Ax = b \right\}.
\end{aligned}
\] (82)
Let \( \bar{x} = x^{k+1} - (z^k - u^k) \). Then, (82) is equivalent to
\[
\begin{aligned}
\text{minimize} & \quad \left\{ \frac{1}{2} \|\bar{x}\|_2^2 : A\bar{x} = \bar{b} \right\},
\end{aligned}
\] (83)
where \( \bar{b} = b - A(z^k - u^k) \).

**Lemma 17** (see [41]). The minimum-norm least-squares solution of \( A\bar{x} = \bar{b} \) is \( \bar{x} = A^\dagger \bar{b} \), where \( A^\dagger \) is the Moore-Penrose inverse of matrix \( A \).

**Theorem 18.** The \( x \)-update of (81) is
\[
\begin{aligned}
x^{k+1} &= \left( I - A^H (AA^H)^{-1} A \right) (z^k - u^k) \\ &\quad + A^H (AA^H)^{-1} b = \Pi (z^k - u^k).
\end{aligned}
\] (84)
**Proof.** As \( A \) is of full row rank, its full-rank factorization is \( A_{\text{mox}} = B_{\text{mox}} C_{\text{mox}} \). Then it yields that [41]
\[
A^+ = C^H (CC^H)^{-1} (B^HB)^{-1} B^H.
\] (85)
From Lemma 17, we have
\[
\bar{x} = C^H (CC^H)^{-1} (B^HB)^{-1} B^H \bar{b},
\] (86)
that is,
\[
\begin{aligned}
x^{k+1} &= (z^k - u^k) \\ &= C^H (CC^H)^{-1} (B^HB)^{-1} B^H (b - A(z^k - u^k)).
\end{aligned}
\] (87)
Since \( B \) is of full rank, we can rearrange (87) to obtain
\[
\begin{aligned}
x^{k+1} &= \left( I - A^H (AA^H)^{-1} A \right) (z^k - u^k) \\ &\quad + A^H (AA^H)^{-1} b.
\end{aligned}
\] (88)
This completes the proof.

If problem (82) is in the real number domain, we have
\[
\begin{aligned}
x^{k+1} &= \left( I - A^T (AA^T)^{-1} A \right) (z^k - u^k) \\ &\quad + A^T (AA^T)^{-1} b,
\end{aligned}
\] (89)
which is the same one obtained in Section 6.2 [2].

The \( z \)-update can be solved by the soft thresholding operator \( S \) in the following theorem, which is a generalization of the soft thresholding in [2].

**Theorem 19.** Let \( t = 1/(2p) \). Then one has the following.
(1) If \( x^{k+1} + u^k \) is real-valued, that is, \( x^{k+1} + u^k = a \), the soft thresholding operator is
\[
S_t(a) = \max(0, a-t) - \max(0, -a-t)
\] (90)
\[
= \begin{cases} 
  a-t, & a > t; \\
  0, & |a| \leq t; \\
  a+t, & a < -t.
\end{cases}
\]
(2) If \( x^{k+1} + u^k \) is purely imaginary, that is, \( x^{k+1} + u^k = bj \), the soft thresholding operator is
\[
S_j(bj) = \begin{cases} 
(j \text{ max } (0, b - t) - \text{ max } (0, -b - t)) & \text{if } b > t; \\
0 & \text{if } |b| \leq t; \\
(j \text{ max } (0, b - t) - \text{ max } (0, -b - t)) & \text{if } b < -t.
\end{cases}
\]

(3) If \( x^{k+1} + u^k = a + bj \), the soft thresholding operator is
\[
S_k(a + bj) = \begin{cases} 
\text{max } (0, a - t_re) - \text{max } (0, -a - t_re) & \text{if } t > 0, b > 0; \\
\text{max } (0, a - t_re) + t_re & \text{if } t > 0, b < 0; \\
\text{max } (0, a - t_re) + t_re & \text{if } t < 0, b > 0; \\
\text{max } (0, a - t_re) - t_re & \text{if } t < 0, b < 0,
\end{cases}
\]
where \( t_re = t \sqrt{a^2 + b^2} \) and \( t_im = t \sqrt{b^2/(a^2 + b^2)} \).

Proof. (1) Assume that \( x^{k+1} + u^k = a \), the updating of \( z \) becomes minimizing the following function:
\[
F(z) = \|z\|_1 + \rho \|z - d\|_2^2 + \sum_{k=1}^{n} \sqrt{(z_{re})^2_k + (z_{im})^2_k} + \rho \sum_{k=1}^{n} \left( ((z_{re})_k - a_k)^2 + ((z_{im})_k - b_k)^2 \right).
\]

From Theorem 2, we have
\[
\frac{\partial F}{\partial (z_{im})_k} = \frac{(z_{im})_k}{\sqrt{(z_{re})^2_k + (z_{im})^2_k}} + \rho (z_{im})_k = 0.
\]

This implies that \( (z_{im})_k = 0 \). Then (93) can be rewritten as
\[
F(z) = \|z_{re}\|_1 + \rho \|z_{re} - d\|_2^2 + \sum_{k=1}^{n} \left( ((z_{re})_k - a_k)^2 \right).
\]

When \( a_k \geq 0 \), \( (z_{re})_k \) should be positive. Then, we have
\[
F(z) = \sum_{k=1}^{n} ((z_{re})_k - a_k)^2.
\]
It is clear that this is a simple parabola, and its results are
\[
\text{If } a_k > t = \frac{1}{2\rho}, \quad (z_{re})_k = (z_{re})_k - t_k;
\]
\[
\text{if } 0 \leq a_k \leq t, \quad (z_{re})_k = 0.
\]

When \( a_k < 0 \), we can get the similar results as follows:
\[
\text{If } a_k < -t, \quad (z_{re})_k = (z_{re})_k + t_k;
\]
\[
\text{if } -t \leq a_k < 0, \quad (z_{re})_k = 0.
\]

From the above discussion, we can complete the proof.

(2) Assume that \( x^{k+1} + u^k = bj \); we have
\[
F(z) = \|z\|_1 + \rho \|z - bj\|_2^2 + \sum_{k=1}^{n} \sqrt{(z_{re})^2_k + (z_{im})^2_k} + \rho \sum_{k=1}^{n} \left( ((z_{re})_k - a_k)^2 + ((z_{im})_k - b_k)^2 \right).
\]

and by adopting the same approach in the above (1), we can get the results.

(3) Assume that \( x^{k+1} + u^k = a + bj \) and satisfy \( \sqrt{a^2 + b^2} > t \), \( a > 0, b > 0; \) then
\[
F(z) = \|z\|_1 + \rho \|z - a + bj\|_2^2 + \sum_{k=1}^{n} \sqrt{(z_{re})^2_k + (z_{im})^2_k} + \rho \sum_{k=1}^{n} \left( ((z_{re})_k - a)^2 + ((z_{im})_k - b_k)^2 \right).
\]

It follows from Theorem 2 that
\[
\frac{\partial f}{\partial (z_{re})_k} = \frac{(z_{re})_k}{\sqrt{(z_{re})^2_k + (z_{im})^2_k}} + 2\rho ((z_{re})_k - a_k) = 0,
\]
\[
\frac{\partial f}{\partial (z_{im})_k} = \frac{(z_{im})_k}{\sqrt{(z_{re})^2_k + (z_{im})^2_k}} + 2\rho ((z_{im})_k - b_k) = 0.
\]

By resolving it, we may get \( (z_{re})_k = a_k - (t_{re})_k, (z_{im})_k = b_k - (t_{im})_k \), where \( t_{re} = t \sqrt{a^2 + b^2} \) and \( t_{im} = t \sqrt{b^2/(a^2 + b^2)} \). Other cases can be discussed similarly. Thus we omit the proof here. This completes the proof.

From what has been discussed above on \( x \)-update and \( z \)-update, the iteration of the BP algorithm is
\[
x^{k+1} = \Pi \left( z^k - u^k \right),
\]
\[
z^{k+1} = S_{1/2\rho} \left( x^{k+1} + u^k \right),
\]
\[
u^{k+1} = u^k + x^{k+1} - z^{k+1},
\]
where \( \Pi \) is projection onto \( \{ x \in \mathbb{R}^n \mid Ax = b \} \) and \( S \) is the soft thresholding operator in the complex number domain.

5. Numerical Simulation

We give two numerical simulations with random data and EEG data. All our numerical experiments are carried out on a PC with Intel Core i7-4710MQ CPU at 2.50 GHz and 8 GB of physical memory. The PC runs MATLAB Version: R2013a on Window 7 Enterprise 64-bit operating system.
5.1. Numerical Simulation of the BP Algorithm with Random Data. Assume that $x \in \mathbb{C}^n$ is a discrete complex signal interested. $x$ itself is $r$-sparse, which contains (at most) $r$ nonzero entries with $r \ll n$. Select $p \ (p < n)$ measurements uniformly at random matrix $A_{p \times n}$ via $A_{p \times n}x = \mathbf{b}$. Hence reconstructing signal $x$ from measurement $\mathbf{b}$ is generally an ill-posed problem which is an undetermined system of linear equations. However, the sparsest solution can be obtained by solving the constrained optimization problem:

$$\min \{ \|x\|_0: Ax = \mathbf{b}, \ x \in \mathbb{C}^n \} ,$$

(103)

where $\|x\|_0$ is the $l_0$-norm of $x$. Unfortunately, (103) is a combinatorial optimization problem of which the computational complexity grows exponentially with the signal size $n$. A key result in [17, 18] is that if $x$ is sparse, the sparsest solution of (103) can be obtained with overwhelming probability by solving the convex optimization problem (77). Next, we consider CS which is actually a kind of application of the BP method in complex variables.

5.1.1. The Effects of the Parameter $\rho$. We demonstrate the complex signal sampling and recovery techniques with a discrete-time complex signal $x$ of length $n = 300$ with sparsity $r = 30$ which is generated randomly. $A_{p \times n}$ is a random sensing matrix with $p = 120$. The variables $u_0$ and $z_0$ are initialized to be zero. We set the two tolerances of primal and dual residuals equal to $10^{-6}$. In order to understand the effects of the parameter $\rho$ on the convergence, we set the penalty parameter $\rho$ from 0.1 to 20 with the step 0.1. We have repeated the same experiment 100 times with the same parameter $\rho$. The average runtime, the average numbers of iterations, and the average primal and dual errors of the ADMM for the different choices of the parameter $\rho$ are presented in Figure 1.

It is clear from Figure 1 (top) that when $2 \leq \rho \leq 4$, the average runtime and the average iterations are reasonable. From Figure 1 (bottom), we can observe that, with the parameter $\rho$ increasing, the primal error decreases while the dual error becomes bigger. Numerical simulations suggest that choosing $\rho \in [2, 4]$ could accelerate the convergence of the ADMM.

5.1.2. The Effects of Tolerances for the Primal and Dual Residuals. Now, we take the different tolerances for the primal residuals $\epsilon_{\text{pri}}$ and the dual residuals $\epsilon_{\text{dual}}$ to analyze the performance of the ADMM, where sparse $r = 0.1n$, measurements $p = 4r$, the penalty parameter $\rho = 2$, and $\epsilon_{\text{pri}} = \epsilon_{\text{dual}} = \epsilon$. We take two different signal lengths $n = 400$ and $n = 600$. We have repeated the same experiment 100 times by a set of randomly generated data. For different choices of $\epsilon$, the average numbers of iterations and the executing time of the above ADMM algorithm (102) are presented in Table 1. It is shown that, with the increasing of precision, the number of iterations increases accordingly while increasing of executing time is not obvious.

In Figure 2(a), the full line describes the changes of the primal residuals $r^k$. In Figure 2(b), the full line describes the

Figure 1: Comparison with different parameter $\rho$. 
Table 1: Numerical results on the different tolerances for residuals.

<table>
<thead>
<tr>
<th>n</th>
<th>ε</th>
<th>Iteration</th>
<th>Time</th>
<th>n</th>
<th>ε</th>
<th>Iteration</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>10^{-1}</td>
<td>21.56</td>
<td>0.0459</td>
<td>600</td>
<td>10^{-1}</td>
<td>23.34</td>
<td>0.1182</td>
</tr>
<tr>
<td>400</td>
<td>10^{-2}</td>
<td>37.80</td>
<td>0.0482</td>
<td>600</td>
<td>10^{-2}</td>
<td>40.52</td>
<td>0.1474</td>
</tr>
<tr>
<td>400</td>
<td>10^{-3}</td>
<td>55.68</td>
<td>0.0571</td>
<td>600</td>
<td>10^{-3}</td>
<td>59.20</td>
<td>0.1720</td>
</tr>
<tr>
<td>400</td>
<td>10^{-4}</td>
<td>78.04</td>
<td>0.0603</td>
<td>600</td>
<td>10^{-4}</td>
<td>92.46</td>
<td>0.2238</td>
</tr>
<tr>
<td>400</td>
<td>10^{-5}</td>
<td>119.80</td>
<td>0.0951</td>
<td>600</td>
<td>10^{-5}</td>
<td>114.98</td>
<td>0.2590</td>
</tr>
<tr>
<td>400</td>
<td>10^{-6}</td>
<td>139.00</td>
<td>0.0956</td>
<td>600</td>
<td>10^{-6}</td>
<td>149.26</td>
<td>0.3063</td>
</tr>
<tr>
<td>400</td>
<td>10^{-7}</td>
<td>161.64</td>
<td>0.1056</td>
<td>600</td>
<td>10^{-7}</td>
<td>230.24</td>
<td>0.4346</td>
</tr>
</tbody>
</table>

Figure 2: Norms of primal residual (a) and dual residual (b) versus iteration.

5.2. Reconstruction of Electroencephalogram Signal by Using Complex ADMM. Electroencephalogram (EEG) signal is a weak bioelectricity of brain cells group, which can be recorded by placing the available electrodes on the scalp or intracranial detect. The EEG signal could reflect the brain bioelectricity rhythmic activity regularity of random nonstationary signal. In this area, much is known for clinical diagnosis and brain function [42]. Because EEG data is large, a very meaningful work is to compress EEG data. It is to effectively reduce the amount of data at the same time and to guarantee that the main features basically remain unchanged [43, 44].

In this paper, EEG signals are recorded with a g.USBamp and a g.EEGcap (Guger Technologies, Graz, Austria) with a sensitivity of 100 V, band pass filtered between 0.1 and 30 Hz, and sampled at 256 Hz. Data are recorded and analyzed using the ECUST BCI platform software package developed through East China University of Science and Technology [45, 46].

We get the complex signal \( x_o \) by performing the discrete Fourier transform (DFT) on EEG signal \( s \); that is, \( x_o = \mathcal{F}(s) \), with signal \( s \) with length \( n = 2000 \). Original EEG signal \( s \) is not sparse, but its DFT signal \( x_o \) becomes the approximate sparse signal; see Figure 3. The hard threshold of \( x_o \) is properly set, which leads to \( x \) with 90 percent zero valued entries, and \( x \) is the approximation of \( x_o \).

To get the compression of sparse signal \( x \), we can first calculate \( b \) from \( b = Ax \). \( A \) is a random complex matrix of size \( p \times n \), where \( p = 800 \), \( n = 2000 \), and the sampling rate is 40%. Substituting \( A \) and \( b \) into the optimization model as presented in (77), that is,

\[
\text{minimize} \quad \| \tilde{x} \|_1 : A\tilde{x} = b, \tilde{x} \in \mathbb{C}^n, \tag{104}
\]

we can obtain the sparse optimal solution \( \tilde{x} \) by employing the ADMM algorithm (102) in Section 4, in which \( \tilde{x} \) is a good approximation of \( x \). By applying the inverse discrete Fourier transform (IDFT) on \( \tilde{x} \), we can get the approximation \( \tilde{s} \) of original signal \( s \); that is,

\[
\tilde{s} = \mathcal{F}^{-1}\tilde{x}. \tag{105}
\]

The original signal \( s \) and its reconstruction signal \( \tilde{s} \) can be seen in Figure 4, in which (a) is the original signal \( s \), (b) is the reconstruction signal \( \tilde{s} \), and (c) is the comparison of them. With the comparison of (a), (b), and (c) in Figure 4, we can observe that the reconstructed signal is in good agreement with the original signal and retains the leading characteristic. The relative error \( \delta = \text{norm}(s - \tilde{s})/\text{norm}(s) = 0.1858 \).

Now we separate the complex signal \( x \) into the real part \( x_{re} \) and the imaginary part \( x_{im} \) and then recast it into an equivalent real-valued optimization problem. We can calculate \( b_1 \) from \( b_1 = Ax_{re} \) and \( b_2 \) from \( b_2 = Ax_{im} \), respectively. Here \( A \) is a random real matrix of size \( p \times n \), where \( p = 800 \), \( n = 2000 \), and the sampling rate is 40% which is similar to the one used in the complex number domain. Substitute \( A \) and \( b_1,b_2 \) into the optimization model as presented in (77); that is,

\[
\text{minimize} \quad \| \bar{x}_{re} \|_1 : A\bar{x}_{re} = b_1, \bar{x}_{re} \in \mathbb{R}^n, \tag{106a}
\]

\[
\text{minimize} \quad \| \bar{x}_{im} \|_1 : A\bar{x}_{im} = b_2, \bar{x}_{im} \in \mathbb{R}^n. \tag{106b}
\]
Although $\tilde{x}_{\text{re}}$ and $\tilde{x}_{\text{im}}$ can approach $x_{\text{re}}$ and $x_{\text{im}}$, respectively, the reconstructed signal $\tilde{s} = \mathcal{F}^{-1}(\tilde{x}_{\text{re}} + i\tilde{x}_{\text{im}})$ is not consistent with the original signal $s$; see Figure 5. The relative error $\delta = \text{norm}(s - \tilde{s})/\text{norm}(s) = 0.5879$. It can be seen that our new ADMM proposed in this paper performs better than the classic ADMM.

6. Conclusions

In this paper, the ADMM for separable convex optimization of real functions in complex variables has been studied. By using Wirtinger calculus, we have established the convergence of the algorithm, which is the generalization of
the one obtained in real variables. Furthermore, the BP algorithm is given in the form of the ADMM, in which projection algorithm and the soft thresholding formula are generalized from the real number domain to the complex case. The simulation results demonstrate that the ADMM can quickly solve convex optimization problems in complex variables within the scopes of the signal compression and reconstruction, which is better than the results in the real number domain.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank Dr. Shouwei Zhao and Dr. Zhongtuan Zheng for pointing out some typos and inconsistencies in an earlier version of the paper. This research was supported by Shanghai Natural Science Fund Project (no. 14ZR148900), the National Natural Science Foundation of China (no. 11471211), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References

on Independent Component Analysis and Blind Signal Separation (ICA’03), Nara, Japan, April 2003.


