Research Article

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

The alternating direction method of multipliers (ADMM) has been widely explored due to its broad applications, and its convergence has been gotten in the real field. In this paper, an ADMM is presented for separable convex optimization of real functions in complex variables. First, the convergence of the proposed method in the complex domain is established by using the Wirtinger Calculus technique. Second, the basis pursuit (BP) algorithm is given in the form of ADMM in which the projection algorithm and the soft thresholding formula are generalized from the real case. The numerical simulations on the reconstruction of electroencephalogram (EEG) signal are provided to show that our new ADMM has better behavior than the classic ADMM for solving separable convex optimization of real functions in complex variables.

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 attached 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 f at z_0 so that the series converges to $f(z)$ in a neighborhood of z_0. A common solution is to convert the optimization problem to the real domain by referring to f as a function of the real and imaginary parts of z. Reformulating an optimization problem which is inherently complex to the real domain would miss important insights on the structure of the problem that might otherwise be exploited [28]. Even so, there are many primal-dual optimization methods for optimization problems with the complex variable in the literatures. The usual method analyzing complex-valued optimization problem is to separate it into the real part and the imaginary part and then to recast it into an equivalent real-valued optimization problem by doubling the size of the constraint conditions; see [29–31] and the references therein. For some other related references on optimization problems in complex variables, we refer to [28, 32].

To overcome the above-mentioned difficulties, the purpose of the paper is to generalize ADMM for separable convex optimization in the real number domain to the complex number domain. The concepts of convex function and subgradient are expanded from the real field to the complex field. By introducing the augmented complex variable, the definition of the augmented Lagrange function in complex variables is given. Under some mild assumptions, we establish the convergence of the proposed method. For the applications, we consider the BP algorithm which concludes projection algorithm and the soft thresholding operator in the complex field. Some numerical simulation results are reported to show that the proposed algorithm is indeed more efficient and more robust.

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.

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 kth entry of a vector z is denoted by z_k and element (i, j) of a matrix A by a_{ij}. The subscripts $\cdot_{\text{re}}$ and $\cdot_{\text{im}}$ denote the real and imaginary parts, respectively; for example, $z_{\text{re}} = \text{Re}\{z\}$ and $A_{\text{im}} = \text{Im}\{A\}$. The superscripts $\cdot^T, \bar{\cdot}, \cdot^H$ and $\cdot^{-1}$ are used for the transpose, conjugate, Hermitian conjugate, and matrix inverse. The $\text{dom} f$ 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, $z^R = (z_{\text{re}}^T, z_{\text{im}}^T)^T \in R^{2n}$, obtained by stacking z_{re} on the top of z_{im}. The notation ∂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:

$$\underline{z} = \left(z^T, z^H\right)^T \in C^{2n}, \tag{1}$$

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

$$J_n = \begin{pmatrix} I_n & jI_n \\ I_n & -jI_n \end{pmatrix} \in C^{2n \times 2n} \tag{2}$$

is unitary up to a factor of 2; that is, $J_n J_n^H = J_n^H J_n = 2I_{2n}$. 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*

$$Ax + Bz = c \iff \underline{A}\,\underline{x} + \underline{B}\,\underline{z} = \underline{c}, \tag{3}$$

where

$$\underline{A} = \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} \in C^{2p \times 2n},$$

$$\underline{B} = \begin{pmatrix} B & 0 \\ 0 & \overline{B} \end{pmatrix} \in C^{2p \times 2m}. \tag{4}$$

Proof. Since

$$\begin{aligned} \underline{A}\,\underline{x} + \underline{B}\,\underline{z} - \underline{c} &= \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} \begin{pmatrix} x \\ \overline{x} \end{pmatrix} + \begin{pmatrix} B & 0 \\ 0 & \overline{B} \end{pmatrix} \begin{pmatrix} z \\ \overline{z} \end{pmatrix} - \begin{pmatrix} c \\ \overline{c} \end{pmatrix} \\ &= \begin{pmatrix} Ax + Bz - c \\ \overline{Ax + Bz - c} \end{pmatrix}, \end{aligned} \tag{5}$$

then we have

$$Ax + Bz = c \iff \underline{A}\,\underline{x} + \underline{B}\,\underline{z} = \underline{c}. \tag{6}$$

This completes the proof. □

Consider a complex-valued function

$$f(z) = u(z_{\mathrm{re}}, z_{\mathrm{im}}) + jv(z_{\mathrm{re}}, z_{\mathrm{im}}), \tag{7}$$

where $z = z_{\mathrm{re}} + jz_{\mathrm{im}}$, $f : C^n \to C$, and $u, v : R^n \times R^n \to R$. 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 Δz approaches zero in the complex plane. This requires that the Cauchy-Riemann equations

$$\begin{aligned} \frac{\partial u}{\partial z_{\mathrm{re}}} &= \frac{\partial v}{\partial z_{\mathrm{im}}}, \\ \frac{\partial u}{\partial z_{\mathrm{im}}} &= -\frac{\partial v}{\partial z_{\mathrm{re}}} \end{aligned} \tag{9}$$

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 $v(z_{\mathrm{re}}, z_{\mathrm{im}}) \equiv 0$, and thus cost functions are not analytic. These conditions imply complex differentiability which are quite stringent and impose a strong structure on $u(z_{\mathrm{re}}, z_{\mathrm{im}})$ and $v(z_{\mathrm{re}}, z_{\mathrm{im}})$ and, consequently, on $f(z)$. Obviously, most cost functions do not satisfy the Cauchy-Riemann equations as these functions are typically $f : C^n \to R$ with $v(z_{\mathrm{re}}, z_{\mathrm{im}}) = 0$.

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 $u(z_{\mathrm{re}}, z_{\mathrm{im}})$ and $v(z_{\mathrm{re}}, z_{\mathrm{im}})$ are differentiable as the functions of real-valued variables z_{re} and z_{im}. Then, one can write the two real variables as $z_{\mathrm{re}} = (z + \overline{z})/2$ and $z_{\mathrm{im}} = -j(z - \overline{z})/2$ 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 $\overline{z}$ 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 $f : R^{2n} \to C$ by rewriting it as $f(z) = f(u, v)$ and make use of the underlying R^{2n} structure. The function $f(\cdot)$ can be regarded as either $f(z_{\mathrm{re}}, z_{\mathrm{im}})$ with variables z_{re} and z_{im} or $f(z, \overline{z})$ with variables z and $\overline{z}$, 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:

$$f(z, \overline{z}) = f(z_{\mathrm{re}}, z_{\mathrm{im}}) = f(z). \tag{10}$$

The main result in this context is stated by Brandwood in [36].

Theorem 2. *Let $f : R^{2n} \to C$ be a function of real variables z_{re} and z_{im} such that $f(z) = f(z_{\mathrm{re}}, z_{\mathrm{im}})$, where $z = z_{\mathrm{re}} + jz_{\mathrm{im}}$, and that f is analytic with respect to z and $\overline{z}$ independently. Then, consider the following:*
(1) The partial derivatives

$$\begin{aligned} \frac{\partial f}{\partial z} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_{\mathrm{re}}} - j\frac{\partial f}{\partial z_{\mathrm{im}}}\right), \\ \frac{\partial f}{\partial \overline{z}} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_{\mathrm{re}}} + j\frac{\partial f}{\partial z_{\mathrm{im}}}\right) \end{aligned} \tag{11}$$

can be computed by treating $\overline{z}$ 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

$$\frac{\partial f}{\partial z} = 0. \tag{12}$$

Similarly,

$$\frac{\partial f}{\partial \overline{z}} = 0 \tag{13}$$

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:

$$f(x, z) = 2\mathrm{Re}\left\{y^H (Ax + Bz - c)\right\}, \tag{14}$$

where $x \in C^n$, $z \in C^m$, $y \in C^p$, $c \in C^p$, $A \in C^{p \times n}$, and $B \in C^{p \times m}$.

It follows from Theorem 2 that

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{2}\left(\frac{\partial f}{\partial x_{\mathrm{re}}} - j\frac{\partial f}{\partial x_{\mathrm{im}}}\right) \\ &= (y_{\mathrm{re}}A_{\mathrm{re}} + y_{\mathrm{im}}A_{\mathrm{im}}) + j(y_{\mathrm{re}}A_{\mathrm{im}} - y_{\mathrm{im}}A_{\mathrm{re}}) \\ &= y^H A. \end{aligned} \tag{15}$$

Similarly, we have

$$\begin{aligned} \frac{\partial f}{\partial z} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_{\mathrm{re}}} - j\frac{\partial f}{\partial z_{\mathrm{im}}}\right) \\ &= (y_{\mathrm{re}}B_{\mathrm{re}} + y_{\mathrm{im}}B_{\mathrm{im}}) + j(y_{\mathrm{re}}B_{\mathrm{im}} - y_{\mathrm{im}}B_{\mathrm{re}}) = y^H B. \end{aligned} \tag{16}$$

Example 4. Consider the real function in complex variables as follows:

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

where $x \in C^n$, $z \in C^m$, $c \in C^p$, $A \in C^{p \times n}$, and $B \in C^{p \times m}$.

We have

$$
\begin{aligned}
f(x,z) &= (Ax + Bz - c)^H (Ax + Bz - c) \\
&= (r_{re} - jr_{im})(r_{re} + jr_{im}) = r_{re}^2 + r_{im}^2,
\end{aligned} \tag{18}
$$

where $r = Ax + Bz - c$. Then

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{1}{2}\left(\frac{\partial f}{\partial x_{re}} - j\frac{\partial f}{\partial x_{im}}\right) \\
&= \frac{1}{2}\left(\frac{\partial (r_{re}^2 + r_{im}^2)}{\partial x_{re}} - j\frac{\partial (r_{re}^2 + r_{im}^2)}{\partial x_{im}}\right) \\
&= (r_{re}A_{re} + r_{im}A_{im}) + j(r_{re}A_{im} - r_{im}A_{re}) = r^H A,
\end{aligned} \tag{19}
$$

$$
\begin{aligned}
\frac{\partial f}{\partial z} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_{re}} - j\frac{\partial f}{\partial z_{im}}\right) \\
&= \frac{1}{2}\left(\frac{\partial (r_{re}^2 + r_{im}^2)}{\partial z_{re}} - j\frac{\partial (r_{re}^2 + r_{im}^2)}{\partial z_{im}}\right) \\
&= (r_{re}B_{re} + r_{im}B_{im}) + j(r_{re}B_{im} - r_{im}B_{re}) = r^H B.
\end{aligned}
$$

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, if for any $z_1, z_2 \in A$ and any $\theta \in R$ with $0 \le \theta \le 1$, then

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

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

$$\frac{\partial}{\partial \underline{z}} = \left(\frac{\partial}{\partial z}, \frac{\partial}{\partial \overline{z}}\right). \tag{21}$$

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

$$\frac{\partial}{\partial z^R} = J_n^T \frac{\partial}{\partial \underline{z}}. \tag{22}$$

For real function in complex variable $f : C^n \to R$, 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 $f : C^n \to R$ is convex if $\mathrm{dom}\, f$ is a convex set and if for any $x, y \in \mathrm{dom}\, f$ and any $\theta \in R$ with $0 \le \theta \le 1$, then

$$f(\theta x + (1 - \theta) y) \le \theta f(x) + (1 - \theta) f(y). \tag{23}$$

Definition 8. A real function in complex variable $f : C^n \to R$ is proper if its effective domain is nonempty and it never attains $-\infty$.

Definition 9. A real function in complex variable $f : C^n \to R$ is closed if, for each $\alpha \in R$, the sublevel set $\{x \in \mathrm{dom}\, f \mid f(x) \le \alpha\}$ is a closed set.

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

$$f(z) \ge f(z_0) + 2\mathrm{Re}\left\{v^H (z - z_0)\right\}. \tag{24}$$

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:

$$\text{minimize} \quad \{f(x) + g(z): Ax + Bz = c, \ x \in \chi_1, \ z \in \chi_2\}, \tag{25}$$

where $\chi_1 \subset R^n$ and $\chi_2 \subset R^m$ are closed convex sets; $A \in R^{p \times n}$ and $B \in R^{p \times m}$ are given matrices; $c \in R^p$ is a given vector; and $f : R^n \to R \cup \{+\infty\}$ and $g : R^m \to R \cup \{+\infty\}$ are proper, closed, and convex functions.

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

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

$$
\begin{aligned}
L_\rho(x,y,z) &= f(x) + g(z) + y^T (Ax + Bz - c) \\
&\quad + \frac{\rho}{2}\|Ax + Bz - c\|_2^2,
\end{aligned} \tag{27}
$$

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

$$
\begin{aligned}
x^{k+1} &= \underset{x}{\mathrm{argmin}}\, L_\rho(x, z^k, y^k), \\
z^{k+1} &= \underset{z}{\mathrm{argmin}}\, L_\rho(x^{k+1}, z, y^k), \\
y^{k+1} &= y^k + \rho(Ax^{k+1} + Bz^{k+1} - c).
\end{aligned} \tag{28}
$$

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

Assumption 11. The (extended-real-valued) functions $f : R^n \to R \cup \{+\infty\}$ and $g : R^m \to R \cup \{+\infty\}$ 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

$$L_0(x^*, z^*, y) \le L_0(x^*, z^*, y^*) \le L_0(x, z, y^*) \qquad (29)$$

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. $r^k = Ax^k + Bz^k - c \rightarrow 0$ *as* $k \rightarrow \infty$; *that is, the iterates approach feasibility.*

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

(3) Dual Variable Convergence. $y^k \rightarrow y^*$ *as* $k \rightarrow \infty$, *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 $f : C^n \rightarrow R \cup \{+\infty\}$ and $g : C^m \rightarrow R \cup \{+\infty\}$. Then, the convex separable optimization of real functions in complex variables becomes

$$\text{minimize} \quad \{f(x) + g(z) : Ax + Bz = c, \ x \in \chi_1, \ z \in \chi_2\}, \quad (30)$$

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; $A \in C^{p \times n}$ and $B \in C^{p \times m}$ 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:

$$\text{minimize} \quad \{f(x, \overline{x}) + g(z, \overline{z}) : \underline{A}\,\underline{x} + \underline{B}\,\underline{z} = \underline{c}, \ x \in \chi_1, \ z \in \chi_2\}. \qquad (31)$$

The Lagrangian function of (31) is

$$\begin{aligned}
L_0(x, z, y) &= f(x, \overline{x}) + g(z, \overline{z}) + \underline{y}^H(\underline{A}\,\underline{x} + \underline{B}\,\underline{z} - \underline{c}) \\
&= f(x, \overline{x}) + g(z, \overline{z}) \\
&\quad + 2\text{Re}\{y^H(Ax + Bz - c)\},
\end{aligned} \qquad (32)$$

where $y \in C^p$. Then, the augmented Lagrangian function of (31) is

$$\begin{aligned}
L_\rho(x, z, y) &= f(x, \overline{x}) + g(z, \overline{z}) + \underline{y}^H(\underline{A}\,\underline{x} + \underline{B}\,\underline{z} - \underline{c}) \\
&\quad + \frac{\rho}{2}\|\underline{A}\,\underline{x} + \underline{B}\,\underline{z} - \underline{c}\|_2^2 \\
&= f(x, \overline{x}) + g(z, \overline{z}) \\
&\quad + 2\text{Re}\{y^H(Ax + Bz - c)\} \\
&\quad + \rho\|Ax + Bz - c\|_2^2,
\end{aligned} \qquad (33)$$

where $\rho > 0$ is called the penalty parameter. The ADMM for complex convex separable optimization is composed of the iterations

$$\begin{aligned}
x^{k+1} &= \underset{x}{\arg\min}\, L_\rho(x, z^k, y^k), \\
z^{k+1} &= \underset{z}{\arg\min}\, L_\rho(x^{k+1}, z, y^k), \\
y^{k+1} &= y^k + \rho(Ax^{k+1} + Bz^{k+1} - c).
\end{aligned} \qquad (34)$$

Let $r = Ax + Bz - c$. Then we have

$$2\text{Re}\{y^H r\} + \rho\|r\|_2^2 = \rho\|r + u\|_2^2 - \rho\|u\|_2^2, \qquad (35)$$

where $u = (1/\rho)y$ is the scaled dual variable. Using the scaled dual variable, we can express the ADMM iterations (34) as

$$x^{k+1} = \underset{x}{\arg\min}\left\{f(x, \overline{x}) + \rho\|Ax + Bz^k - c + u^k\|_2^2\right\},$$

$$z^{k+1}$$
$$= \underset{z}{\arg\min}\left\{g(z, \overline{z}) + \rho\|Ax^{k+1} + Bz - c + u^k\|_2^2\right\},$$

$$u^{k+1} = u^k + Ax^{k+1} + Bz^{k+1} - c. \qquad (36)$$

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

$$Ax^* + Bz^* - c = 0, \qquad (37)$$

and dual feasibility,

$$0 \in \partial f(x^*) + (y^*)^H A, \qquad (38)$$

$$0 \in \partial g(z^*) + (y^*)^H B. \qquad (39)$$

Because z^{k+1} minimizes $L_\rho(x^{k+1}, z, y^k)$, we have

$$\begin{aligned}
0 &\in \partial g(z^{k+1}) + (y^k)^H B + \rho(r^{k+1})^H B \\
&= \partial g(z^{k+1}) + (y^{k+1})^H B.
\end{aligned} \qquad (40)$$

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

$$0 \in \partial f\left(x^{k+1}\right) + \left(y^k\right)^H A + \rho\left(Ax^{k+1} + Bz^k - c\right)^H A$$
$$= \partial f\left(x^{k+1}\right) + \left(y^k\right)^H + \rho\left(r^{k+1}\right)^H A \qquad (41)$$
$$+ \rho\left(B\left(z^k - z^{k+1}\right)\right)^H A$$

or equivalently

$$\rho\left(B\left(z^{k+1} - z^k\right)\right)^H A \in \partial f\left(x^{k+1}\right) + \left(y^{k+1}\right)^H A. \qquad (42)$$

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), $r^{k+1} = Ax^{k+1} + Bz^{k+1} - c$ 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 $f : C^n \to R \cup \{+\infty\}$ and $g : C^m \to R \cup \{+\infty\}$ 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

$$L_0\left(x^*, z^*, y\right) \le L_0\left(x^*, z^*, y^*\right) \le L_0\left(x, z, y^*\right) \qquad (43)$$

holds for all x, z, and y.

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

(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.*

(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 $q^* = f(x^*) + g(x^*)$. Then we have

$$L_0\left(x^*, z^*, y^*\right) \le L_0\left(x^{k+1}, z^{k+1}, y^*\right). \qquad (44)$$

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

$$q^* \le q^{k+1} + 2\text{Re}\left\{\left(y^*\right)^H\left(Ax^{k+1} + Bz^{k+1} - c\right)\right\}. \qquad (45)$$

From Theorem 2 and Examples 3 and 4, we get

$$\frac{\partial L_\rho}{\partial x} = \frac{\partial f\left(x^{k+1}, \overline{x}^{k+1}\right)}{\partial x} + \left(y^k\right)^H A \qquad (46)$$
$$+ \rho\left(Ax^{k+1} + Bz^k - c\right)^H A.$$

Note that L_ρ is a real-valued function; then we get

$$\frac{\partial L_\rho}{\partial \overline{x}} = \overline{\left(\frac{\partial L_\rho}{\partial x}\right)}. \qquad (47)$$

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

$$0 \in \partial L_\rho\left(x^{k+1}, z^k, y^k\right)$$
$$= \partial f\left(x^{k+1}, \overline{x}^{k+1}\right) + \left(y^k\right)^H A \qquad (48)$$
$$+ \rho\left(Ax^{k+1} + Bz^k - c\right)^H A.$$

Since

$$y^{k+1} = y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right), \qquad (49)$$

we have

$$0 \in \partial f\left(x^{k+1}, \overline{x}^{k+1}\right) + \left(y^{k+1} - \rho B\left(z^{k+1} - z^k\right)\right)^H A, \qquad (50)$$

which implies that x^{k+1} minimizes

$$f(x, \overline{x}) + 2\text{Re}\left\{\left(y^{k+1} - \rho B\left(z^{k+1} - z^k\right)\right)^H Ax\right\}. \qquad (51)$$

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

$$g(z, \overline{z}) + 2\text{Re}\left\{\left(y^{k+1}\right)^H Bz\right\}. \qquad (52)$$

From (51) and (52), we have

$$f\left(x^{k+1}, \overline{x}^{k+1}\right)$$
$$+ 2\text{Re}\left\{\left(y^{k+1} - \rho B\left(z^{k+1} - z^k\right)\right)^H Ax^{k+1}\right\}$$
$$\le f\left(x^*, \overline{x}^*\right)$$
$$+ 2\text{Re}\left\{\left(y^{k+1} - \rho B\left(z^{k+1} - z^k\right)\right)^H Ax^*\right\}, \qquad (53)$$
$$g\left(z^{k+1}, \overline{z}^{k+1}\right) + 2\text{Re}\left\{\left(y^{k+1}\right)^H Bz^{k+1}\right\} \le g\left(z^*, \overline{z}^*\right)$$
$$+ 2\text{Re}\left\{\left(y^{k+1}\right)^H Bz^*\right\}.$$

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

$$p^{k+1} - p^* \le 2\text{Re}\left\{-\left(y^{k+1}\right)^H r^{k+1}\right.$$
$$\left. - \left(\rho B\left(z^{k+1} - z^k\right)\right)^H\left(-r^{k+1} + B\left(z^{k+1} - z^*\right)\right)\right\}, \qquad (54)$$

where $r^{k+1} = Ax^{k+1} + Bz^{k+1} - c$.

Adding (45) and (54), we get

$$4\text{Re}\left\{\left(y^{k+1} - y^*\right)^H r^{k+1} - \left(\rho B\left(z^{k+1} - z^k\right)\right)^H r^{k+1}\right.$$
$$\left. + \left(\rho B\left(z^{k+1} - z^k\right)\right)^H \left(B\left(z^{k+1} - z^*\right)\right)\right\} \le 0. \tag{55}$$

Let

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

By following manipulation and rewriting of (55), we have

$$w^{k+1} \le w^k - \rho\left\|r^{k+1}\right\|_2^2 - \rho\left\|B\left(z^{k+1} - z^k\right)\right\|_2^2. \tag{57}$$

Rewriting the first term of (55) as

$$4\text{Re}\left\{\left(y^{k+1} - y^*\right)^H r^{k+1} - \left(y^k + \rho r^{k+1} - y^*\right)^H r^{k+1}\right.$$
$$\left. + \left(y^k - y^*\right)^H r^{k+1}\right\} + \rho\left\|r^{k+1}\right\|_2^2 + \rho\left\|r^{k+1}\right\|_2^2 \tag{58}$$

and then substituting

$$r^{k+1} = \left(\frac{1}{\rho}\right)\left(y^{k+1} - y^k\right) \tag{59}$$

into the first two terms in (58) give

$$2\text{Re}\left\{\left(\frac{2}{\rho}\right)\left(y^k - y^*\right)^H\left(y^{k+1} - y^k\right)\right\}$$
$$+ \left(\frac{1}{\rho}\right)\left\|y^{k+1} - y^k\right\|_2^2 + \rho\left\|r^{k+1}\right\|_2^2. \tag{60}$$

Since

$$y^{k+1} - y^k = \left(y^{k+1} - y^*\right) - \left(y^k - y^*\right), \tag{61}$$

this can be expressed as

$$\frac{1}{\rho}\left(\left\|y^{k+1} - y^*\right\|_2^2 - \left\|y^k - y^*\right\|_2^2\right) + \rho\left\|r^{k+1}\right\|_2^2. \tag{62}$$

Now let us regroup the remaining terms, that is,

$$\rho\left\|r^{k+1}\right\|_2^2 - 2\text{Re}\left\{2\rho\left(B\left(z^{k+1} - z^k\right)\right)^H r^{k+1}\right.$$
$$\left. + 2\rho\left(B\left(z^{k+1} - z^k\right)\right)^H\left(B\left(z^{k+1} - z^*\right)\right)\right\}, \tag{63}$$

where $\rho\|r^{k+1}\|_2^2$ is taken from (62). Substituting

$$z^{k+1} - z^* = \left(z^{k+1} - z^k\right) + \left(z^k - z^*\right) \tag{64}$$

into the last term in (63) yields

$$\rho\left\|r^{k+1} - B\left(z^{k+1} - z^k\right)\right\|_2^2 + \rho\left\|B\left(z^{k+1} - z^k\right)\right\|_2^2$$
$$+ 4\rho\text{Re}\left\{\left(B\left(z^{k+1} - z^k\right)\right)^H B\left(z^k - z^*\right)\right\} \tag{65}$$

and substituting

$$z^{k+1} - z^k = \left(z^{k+1} - z^*\right) - \left(z^k - z^*\right) \tag{66}$$

into the last two terms in (63), we get

$$\rho\left\|r^{k+1} - B\left(z^{k+1} - z^k\right)\right\|_2^2$$
$$+ \rho\left(\left\|B\left(z^{k+1} - z^k\right)\right\|_2^2 - \left\|B\left(z^k - z^*\right)\right\|_2^2\right). \tag{67}$$

It implies that (55) can be expressed as

$$w^k - w^{k+1} \ge \rho\left\|r^{k+1} - B\left(z^{k+1} - z^k\right)\right\|_2^2. \tag{68}$$

To obtain (57), it suffices to show that the middle term

$$-4\text{Re}\left\{\rho\left(r^{k+1}\right)^H\left(B\left(z^{k+1} - z^k\right)\right)\right\} \tag{69}$$

of the expanded right-hand side of (68) is positive. To understand this, by reviving that z^{k+1} minimizes $g(z, \bar{z}) + 2\text{Re}\{(y^{k+1})^H Bz\}$ and z^k minimizes $g(z, \bar{z}) + 2\text{Re}\{(y^k)^H Bz\}$, we can add

$$g\left(z^{k+1}, \bar{z}^{k+1}\right) + 2\text{Re}\left\{\left(y^{k+1}\right)^H Bz^{k+1}\right\}$$
$$\le g\left(z^k, \bar{z}^k\right) + 2\text{Re}\left\{\left(y^{k+1}\right)^H Bz^k\right\},$$
$$g\left(z^k, \bar{z}^k\right) + 2\text{Re}\left\{\left(y^k\right)^H Bz^k\right\}$$
$$\le g\left(z^{k+1}, \bar{z}^{k+1}\right) + 2\text{Re}\left\{\left(y^k\right)^H Bz^{k+1}\right\} \tag{70}$$

to obtain that

$$2\text{Re}\left\{\left(y^{k+1} - y^k\right)^H B\left(z^{k+1} - z^k\right)\right\} \le 0. \tag{71}$$

Since $\rho > 0$, if we substitute

$$y^{k+1} - y^k = \rho r^{k+1}, \tag{72}$$

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 $w^k \le w^0$, it follows that y^k and Bz^k are bounded. Iterating the inequality above gives

$$\rho\sum_{k=0}^{\infty}\left(\left\|r^{k+1}\right\|_2^2 + \left\|B\left(z^{k+1} - z^k\right)\right\|_2^2\right) \le w_0, \tag{73}$$

implying that $r^k \to 0$ and $B(z^{k+1} - z^k) \to 0$ as $k \to \infty$. From (45), we have

$$f\left(x^k, \bar{x}^k\right) + g\left(z^k, \bar{z}^k\right) \longrightarrow f\left(x^*, \bar{x}^*\right) + g\left(z^*, \bar{z}^*\right) \tag{74}$$

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

$$-r^{k+1} + B\left(z^{k+1} - z^k\right) = -A\left(x^{k+1} - x^*\right). \tag{75}$$

Substituting this into (54), we get

$$
\begin{aligned}
p^{k+1} - p^* \\
\leq 2\mathrm{Re}\left\{-\left(y^{k+1}\right)^H r^{k+1} + \left(x^{k+1} - x^*\right)^H s^{k+1}\right\}.
\end{aligned} \tag{76}
$$

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, $\|r^{k+1}\|_2 \leq \varepsilon^{\mathrm{pri}}$ and $\|s^{k+1}\|_2 \leq \varepsilon^{\mathrm{dual}}$, where $\varepsilon^{\mathrm{pri}}$ and $\varepsilon^{\mathrm{dual}}$ 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

$$\text{minimize} \quad \left\{\|x\|_1 : Ax = b, \ x \in C^n\right\}, \tag{77}$$

where $A \in C^{p \times n}$ is a given matrix, $\mathrm{Rank}(A) = p$, and $b \in C^p$ is a given vector.

Recall that

$$\|x\|_1 = \sum_{k=1}^n |x_k| = \sum_{k=1}^n \sqrt{(x_k)_{\mathrm{re}}^2 + (x_k)_{\mathrm{im}}^2}. \tag{78}$$

Then

$$\|\underline{x}\|_1 = 2\|x\|_1. \tag{79}$$

In the form of the ADMM, the BP method can be expressed as

$$\text{minimize} \quad \left\{f(x) + \|z\|_1 : x = z, \ x, z \in C^n\right\}, \tag{80}$$

where f is the indicator function of $X = \{x \in C^n \mid Ax = b\}$; 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:

$$
\begin{aligned}
x^{k+1} &= \arg\min_x \left\{f(x) + \rho\left\|x - z^k + u^k\right\|_2^2\right\}, \\
z^{k+1} &= \arg\min_z \left\{\|z\|_1 + \rho\left\|x^{k+1} - z + u^k\right\|_2^2\right\}, \\
u^{k+1} &= u^k + x^{k+1} - z^{k+1}.
\end{aligned} \tag{81}
$$

The x-update, which involves solving a linearly constrained minimum Euclidean norm problem, can be written as

$$\text{minimize}_{x^{k+1}} \quad \left\{\frac{1}{2}\left\|x^{k+1} - \left(z^k - u^k\right)\right\|_2^2 : Ax = b\right\}. \tag{82}$$

Let $\tilde{x} = x^{k+1} - (z^k - u^k)$. Then, (82) is equivalent to

$$\text{minimize}_{\tilde{x}} \quad \left\{\frac{1}{2}\|\tilde{x}\|_2^2 : A\tilde{x} = \tilde{b}\right\}, \tag{83}$$

where $\tilde{b} = b - A(z^k - u^k)$.

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

Theorem 18. *The x-update of (81) is*

$$
\begin{aligned}
x^{k+1} &= \left(I - A^H\left(AA^H\right)^{-1}A\right)\left(z^k - u^k\right) \\
&\quad + A^H\left(AA^H\right)^{-1}b := \Pi\left(z^k - u^k\right).
\end{aligned} \tag{84}
$$

Proof. As A is of full row rank, its full-rank factorization is $A_{m \times n} = B_{m \times m}C_{m \times n}$. Then it yields that [41]

$$A^\dagger = C^H\left(CC^H\right)^{-1}\left(B^H B\right)^{-1}B^H. \tag{85}$$

From Lemma 17, we have

$$\tilde{x} = C^H\left(CC^H\right)^{-1}\left(B^H B\right)^{-1}B^H \tilde{b}; \tag{86}$$

that is,

$$
\begin{aligned}
x^{k+1} - \left(z^k - u^k\right) \\
= C^H\left(CC^H\right)^{-1}\left(B^H B\right)^{-1}B^H\left(b - A\left(z^k - u^k\right)\right).
\end{aligned} \tag{87}
$$

Since B is of full rank, we can rearrange (87) to obtain

$$
\begin{aligned}
x^{k+1} &= \left(I - A^H\left(AA^H\right)^{-1}A\right)\left(z^k - u^k\right) \\
&\quad + A^H\left(AA^H\right)^{-1}b.
\end{aligned} \tag{88}
$$

This completes the proof. $\square$

If problem (82) is in the real number domain, we have

$$
\begin{aligned}
x^{k+1} &= \left(I - A^T\left(AA^T\right)^{-1}A\right)\left(z^k - u^k\right) \\
&\quad + A^T\left(AA^T\right)^{-1}b,
\end{aligned} \tag{89}
$$

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 $t = 1/(2\rho)$. 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

$$
\begin{aligned}
S_t(a) &= \max(0, a - t) - \max(0, -a - t) \\
&= \begin{cases} a - t, & a > t; \\ 0, & |a| \leq t; \\ a + t, & a < -t. \end{cases}
\end{aligned} \tag{90}
$$

(2) If $x^{k+1} + u^k$ is purely imaginary, that is, $x^{k+1} + u^k = bj$, the soft thresholding operator is

$$S_t(bj) = j(\max(0, b - t) - \max(0, -b - t))$$

$$= \begin{cases} (b - t)j, & b > t; \\ 0, & |b| \le t; \\ (b + t)j, & b < -t. \end{cases} \tag{91}$$

(3) If $x^{k+1} + u^k = a + bj$, the soft thresholding operator is

$$S_t(a + bj)$$

$$= \max(0, a - t_{re}) - \max(0, -a - t_{re})$$

$$+ j(\max(0, b - t_{im}) - \max(0, -b - t_{im}))$$

$$= \begin{cases} 0, & \sqrt{a^2 + b^2} \le t; \\ (a - t_{re}) + j(b - t_{im}), & \sqrt{a^2 + b^2} > t, \ a > 0, \ b > 0; \\ (a + t_{re}) + j(b - t_{im}), & \sqrt{a^2 + b^2} > t, \ a < 0, \ b > 0; \\ (a - t_{re}) + j(b + t_{im}), & \sqrt{a^2 + b^2} > t, \ a > 0, \ b < 0; \\ (a + t_{re}) + j(b + t_{im}), & \sqrt{a^2 + b^2} > t, \ a < 0, \ b < 0, \end{cases} \tag{92}$$

where $t_{re} = t\sqrt{a^2/(a^2 + b^2)}$ 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:

$$F(z) = \|z\|_1 + \rho\|z - a\|_2^2$$

$$= \sum_{k=1}^{n} \sqrt{(z_{re})_k^2 + (z_{im})_k^2}$$

$$+ \rho\sum_{k=1}^{n} \left(((z_{re})_k - a_k)^2 + (z_{im})_k^2\right). \tag{93}$$

From Theorem 2, we have

$$\frac{\partial F}{\partial(z_{im})_k} = \frac{(z_{im})_k}{\sqrt{(z_{re})_k^2 + (z_{im})_k^2}} + \rho(z_{im})_k = 0. \tag{94}$$

This implies that $(z_{im})_k = 0$. Then (93) can be rewritten as

$$F(z) = \|z_{re}\|_1 + \rho\|z_{re} - a\|_2^2$$

$$= \sum_{k=1}^{n} |(z_{re})_k| + \rho\sum_{k=1}^{n} \left((z_{re})_k - a_k\right)^2. \tag{95}$$

When $a_k \ge 0$, $(z_{re})_k$ should be positive. Then, we have

$$F(z) = \sum_{k=1}^{n} (z_{re})_k + \rho\sum_{k=1}^{n} \left((z_{re})_k - a_k\right)^2. \tag{96}$$

It is clear that this is a simple parabola, and its results are

$$\text{If } a_k > t = \frac{1}{(2\rho)}, \quad (z_{re})_k = (z_{re})_k - t_k;$$

$$\text{if } 0 \le a_k \le t, \quad (z_{re})_k = 0. \tag{97}$$

When $a_k < 0$, we can get the similar results as follows:

$$\text{If } a_k < -t, \quad (z_{re})_k = (z_{re})_k + t_k;$$

$$\text{if } -t \le a_k < 0, \quad (z_{re})_k = 0. \tag{98}$$

From the above discussion, we can complete the proof.

(2) Assume that $x^{k+1} + u^k = bj$; we have

$$F(z) = \|z\|_1 + \rho\|z - bj\|_2^2$$

$$= \sum_{k=1}^{n} \sqrt{(z_{re})_k^2 + (z_{im})_k^2}$$

$$+ \rho\sum_{k=1}^{n} \left((z_{re})_k^2 + (z_{im})_k - b_k^2\right), \tag{99}$$

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

$$F(z) = \|z\|_1 + \rho\|z - a + bj\|_2^2$$

$$= \sum_{k=1}^{n} \sqrt{(z_{re})_k^2 + (z_{im})_k^2}$$

$$+ \rho\sum_{k=1}^{n} \left(((z_{re})_k - a)^2 + ((z_{im})_k - b_k)^2\right). \tag{100}$$

It follows from Theorem 2 that

$$\frac{\partial f}{\partial(z_{re})_k} = \frac{(z_{re})_k}{\sqrt{(z_{re})_k^2 + (z_{im})_k^2}} + 2\rho((z_{re})_k - a_k) = 0,$$

$$\frac{\partial f}{\partial(z_{im})_k} = \frac{(z_{im})_k}{\sqrt{(z_{re})_k^2 + (z_{im})_k^2}} + 2\rho((z_{im})_k - b_k) = 0. \tag{101}$$

By resolving it, we may get $(z_{re})_k = a_k - (t_{re})_k$, $(z_{im})_k = b_k - (t_{im})_k$, where $t_{re} = t\sqrt{a^2/(a^2 + b^2)}$ 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. $\square$

From what has been discussed above on x-update and z-update, the iteration of the BP algorithm is

$$x^{k+1} = \Pi(z^k - u^k),$$

$$z^{k+1} = S_{1/(2\rho)}(x^{k+1} + u^k), \tag{102}$$

$$u^{k+1} = u^k + x^{k+1} - z^{k+1},$$

where Π is projection onto $\{x \in C^n \mid Ax = b\}$ 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.
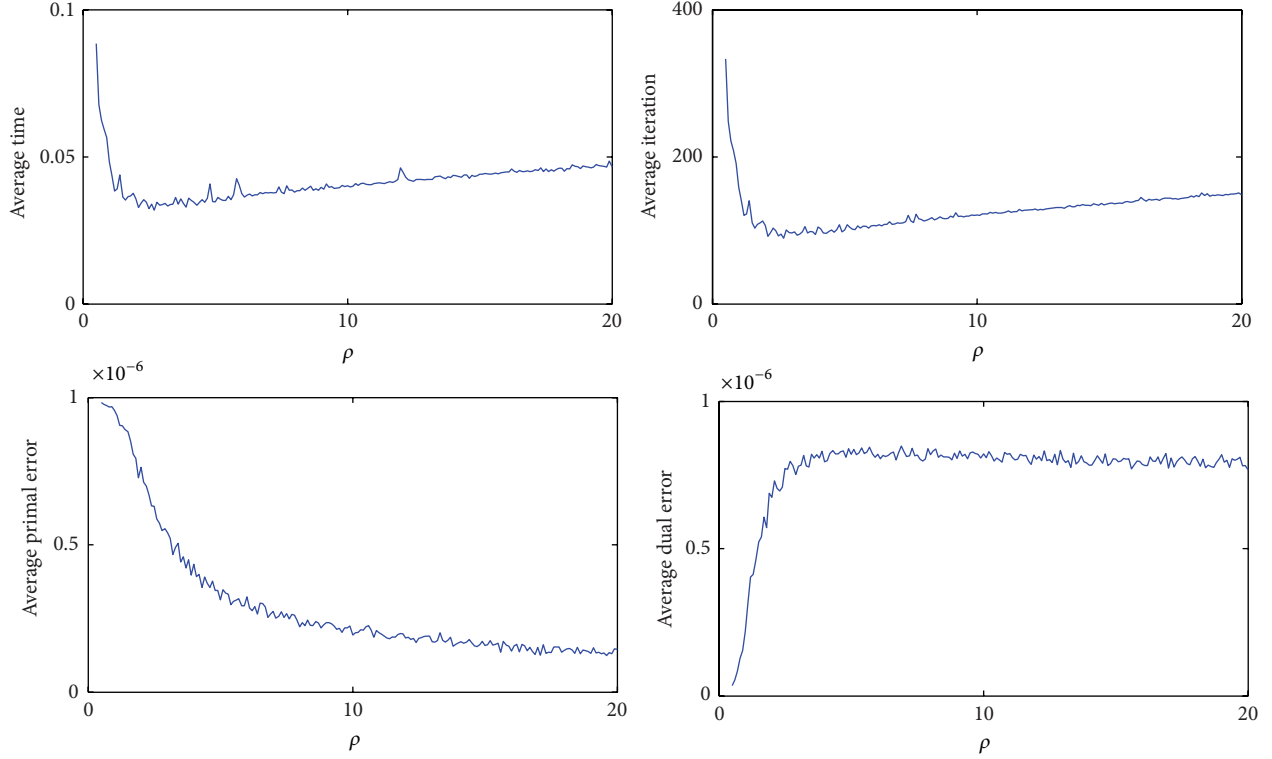
FIGURE 1: Comparison with different parameter ρ.

5.1. Numerical Simulation of the BP Algorithm with Random Data. Assume that $x \in C^n$ 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 $A_{p \times n} x = b$. Hence reconstructing signal x from measurement 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:

$$\text{minimize} \quad \{\|x\|_0 : Ax = b, \ x \in C^n\}, \tag{103}$$

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 ρ. 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 ρ on the convergence, we set the penalty

parameter ρ from 0.1 to 20 with the step 0.1. We have repeated the same experiment 100 times with the same parameter ρ. 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 ρ 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 ρ 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 ϵ^{pri} and the dual residuals ϵ^{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 ϵ, 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

TABLE 1: Numerical results on the different tolerances for residuals.

n	ϵ	Iteration	Time	n	ϵ	Iteration	Time
400	10^{-1}	21.56	0.0459	600	10^{-1}	23.34	0.1182
400	10^{-2}	37.80	0.0482	600	10^{-2}	40.52	0.1474
400	10^{-3}	55.68	0.0571	600	10^{-2}	59.20	0.1720
400	10^{-4}	78.04	0.0603	600	10^{-4}	92.46	0.2238
400	10^{-5}	119.80	0.0951	600	10^{-5}	114.98	0.2590
400	10^{-6}	139.00	0.0956	600	10^{-6}	149.26	0.3063
400	10^{-7}	161.64	0.1056	600	10^{-7}	230.24	0.4346
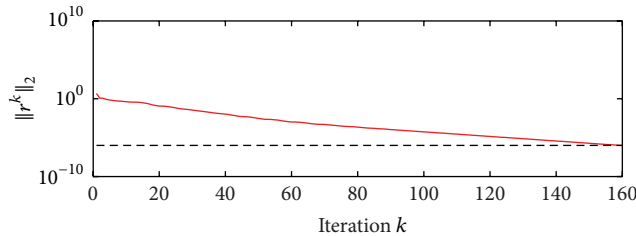
(a)
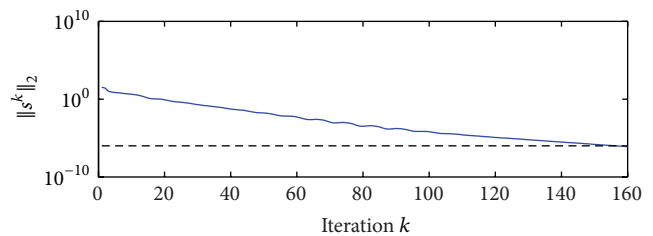
(b)

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

changes of the dual residuals s^k. The dotted lines in Figures 2(a) and 2(b) represent the original residual tolerance $\epsilon^{\mathrm{pri}} = 10^{-6}$ and dual residuals tolerance $\epsilon^{\mathrm{dual}} = 10^{-6}$, respectively. From Figure 2, we can see that the two residuals descend monotonously.

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,

$$\text{minimize} \quad \{\|\tilde{x}\|_1 : A\tilde{x} = b, \ \tilde{x} \in C^n\}, \quad (104)$$

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,

$$\tilde{s} = \mathcal{F}^{-1}\tilde{x}. \quad (105)$$

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 $\delta = \mathrm{norm}(s - \tilde{s})/\mathrm{norm}(s) = 0.1858$.

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 $b_1 = Ax_{\mathrm{re}}$ and b_2 from $b_2 = Ax_{\mathrm{im}}$, 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,

$$\text{minimize} \quad \{\|\tilde{x}_{\mathrm{re}}\|_1 : A\tilde{x}_{\mathrm{re}} = b_1, \ \tilde{x}_{\mathrm{re}} \in R^n\},$$
$$\text{minimize} \quad \{\|\tilde{x}_{\mathrm{im}}\|_1 : A\tilde{x}_{\mathrm{im}} = b_2, \ \tilde{x}_{\mathrm{im}} \in R^n\}. \quad (106)$$

(a) Original EEG   (b) Real part of DFT   (c) Imaginary part of DFT

FIGURE 3: Original EEG signal and its DFT signal.

(a) Original EEG signal

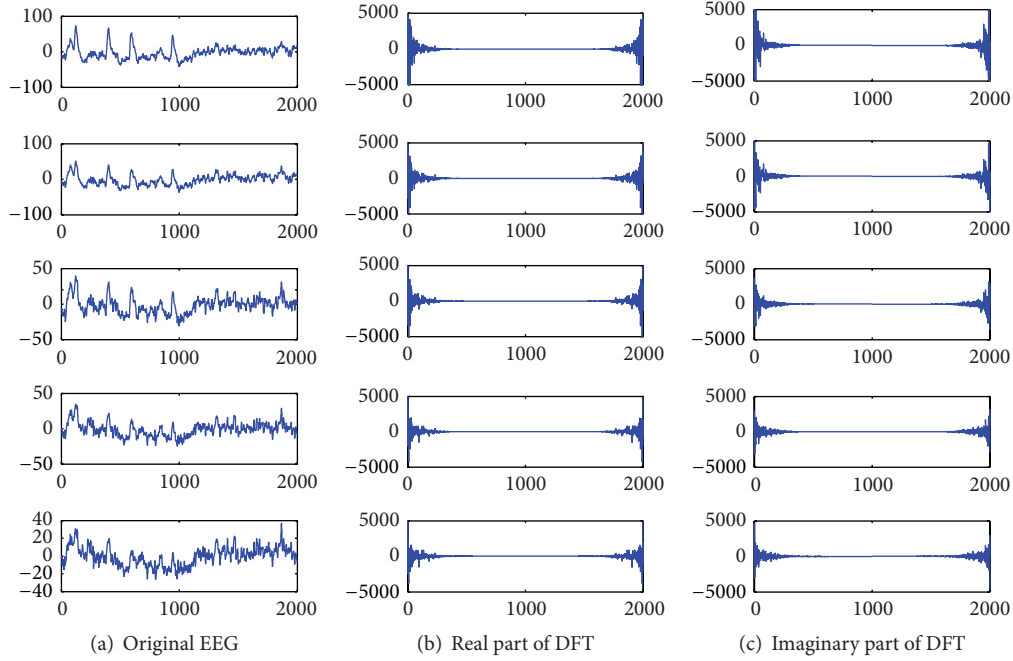(b) Reconstructed EEG signal in complex number domain

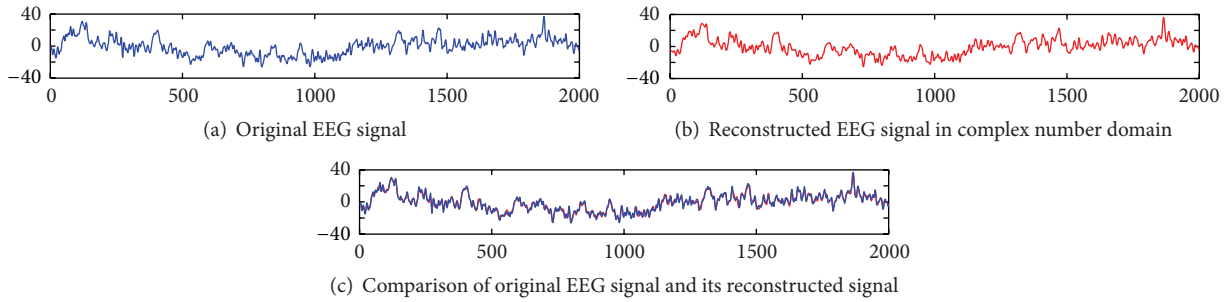(c) Comparison of original EEG signal and its reconstructed signal

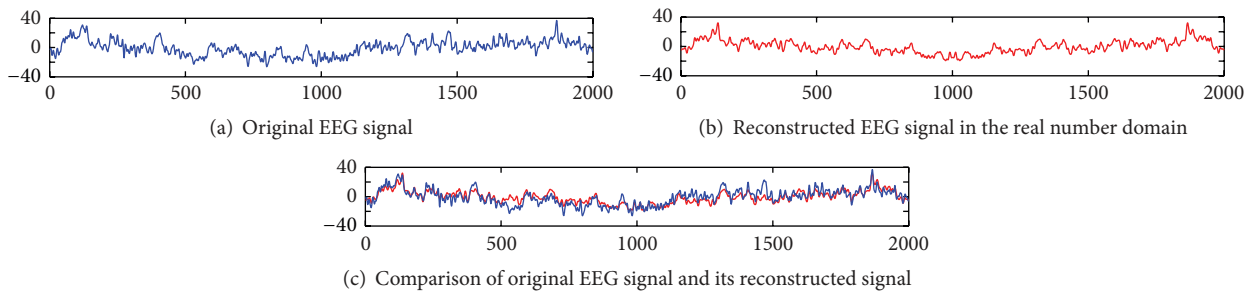FIGURE 4: Original EEG signal and its reconstructed signal in the complex number domain.

(a) Original EEG signal

(b) Reconstructed EEG signal in the real number domain

(c) Comparison of original EEG signal and its reconstructed signal

FIGURE 5: Original EEG signal and its reconstructed signal in the real number domain.

Although $\widetilde{x}_{re}$ and $\widetilde{x}_{im}$ can approach x_{re} and x_{im}, respectively, the reconstructed signal $\widetilde{s} = \mathcal{F}^{-1}(\widetilde{x}_{re} + i\widetilde{x}_{im})$ is not consistent with the original signal s; see Figure 5. The relative error $\delta = \text{norm}(s - \widetilde{s})/\text{norm}(s) = 0.5879$. 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

References

[1] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, no. 5, pp. 303–320, 1969.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[3] M. K. Ng, P. Weiss, and X. M. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.

[4] K. Scheinberg, S. Q. Ma, and D. Goldfarb, "Sparse inverse covariance selection via alternating linearization methods," in *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS '10)*, Vancouver, Canada, December 2010.

[5] R. D. C. Monteiro and B. F. Svaiter, "Implementation of a block-decomposition algorithm for solving large-scale conic semidefinite programming problems," *Computational Optimization and Applications*, vol. 57, no. 1, pp. 45–69, 2014.

[6] B. S. He and X. M. Yuan, "On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal on Numerical Analysis*, vol. 50, no. 2, pp. 700–709, 2012.

[7] M. Tao and X. M. Yuan, "On the $O(1/t)$ convergence rate of alternating direction method with logarithmic quadratic proximal regularization," *SIAM Journal on Optimization*, vol. 22, no. 4, pp. 1431–1448, 2012.

[8] C. H. Chen, B. S. He, Y. Ye, and X. M. Yuan, "The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent," *Mathematical Programming*, 2015.

[9] M. Li, X. Li, and X. M. Yuan, "Convergence analysis of the generalized alternating direction method of multipliers with logarithmic-quadratic proximal regularization," *Journal of Optimization Theory and Applications*, vol. 164, no. 1, pp. 218–233, 2015.

[10] C. H. Chen, M. Li, and X. M. Yuan, "Further study on the convergence rate of alternating direction method of multipliers with logarithmic-quadratic proximal regularization," *Journal of Optimization Theory and Applications*, vol. 166, no. 3, pp. 906–929, 2015.

[11] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[12] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1–3, pp. 293–318, 1992.

[13] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[14] Y. Bai and K. Shen, "Alternating direction method of multipliers for ℓ_1-ℓ_2-regularized logistic regression model," *Journal of the Operations Research Society of China*, pp. 1–11, 2015.

[15] T. Y. Lin, S. Q. Ma, and S. Z. Zhang, "On the sublinear convergence rate of multi-block ADMM," *Journal of the Operations Research Society of China*, vol. 3, no. 3, pp. 251–274, 2015.

[16] S. Q. Ma, "Alternating direction method of multipliers for sparse principal component analysis," *Journal of the Operations Research Society of China*, vol. 1, no. 2, pp. 253–274, 2013.

[17] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[18] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[19] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "*K*-*t* FOCUSS: a general compressed sensing framework for high resolution dynamic MRI," *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 103–116, 2009.

[20] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[21] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[22] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[23] K. C. Kwak and W. Pedrycz, "Face recognition using an enhanced independent component analysis approach," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 530–541, 2007.

[24] V. D. Calhoun and T. Adali, "Unmixing fMRI with independent component analysis," *IEEE Engineering in Medicine and Biology Magazine*, vol. 25, no. 2, pp. 79–90, 2006.

[25] J. Annemüller, T. J. Sejnowski, and S. Makeig, "Complex spectral domain independent component analysis of electroencephalographic data," in *Proceedings of the 4th International Symposium*

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

[26] C. N. K. Mooers, "A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants," *Deep-Sea Research and Oceanographic Abstracts*, vol. 20, no. 12, pp. 1129–1141, 1973.

[27] G. Tauböck, "Complex noise analysis of DMT," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5739–5754, 2007.

[28] L. Sorber, M. V. Barel, and L. D. Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM Journal on Optimization*, vol. 22, no. 3, pp. 879–898, 2012.

[29] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11/12, no. 1–4, pp. 625–653, 1999.

[30] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—a MATLAB software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11/12, no. 1–4, pp. 545–581, 1999.

[31] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proceedings of the IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289, IEEE, Taipei, Taiwan, September 2004.

[32] B. Jiang, S. Q. Ma, and S. Z. Zhang, "Alternating direction method of multipliers for real and complex polynomial optimization models," *Optimization*, vol. 63, no. 6, pp. 883–898, 2014.

[33] W. Wirtinger, "Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.

[34] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, Cambridge, UK, 2010.

[35] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: the proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, 2011.

[36] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings H: Microwaves, Optics and Antennas*, vol. 130, no. 1, pp. 11–16, 1983.

[37] H. Groemer, "On an inequality of Minkowski for mixed volumes," *Geometriae Dedicata*, vol. 33, no. 1, pp. 117–122, 1990.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[39] D. Zhang, *The p-perfectly convex set and absolutely p-persectly convex set [M.S. thesis]*, Shanghai University, 2010, (Chinese).

[40] P. L. Combettes and J. C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49, pp. 185–212, Springer, Berlin, Germany, 2011.

[41] G. R. Wang, Y. Wei, and S. Z. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, China, 2003.

[42] E. D. Übeyli and İ. Güler, "Features extracted by eigenvector methods for detecting variability of EEG signals," *Pattern Recognition Letters*, vol. 28, no. 5, pp. 592–603, 2007.

[43] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

[44] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2009–2015, 2013.

[45] J. Jin, B. Z. Allison, Y. Zhang, X. Y. Wang, and A. Cichocki, "An Erp-based BCI using an oddball paradigm with different faces and reduced errors in critical functions," *International Journal of Neural Systems*, vol. 24, no. 8, Article ID 1450027, 14 pages, 2014.

[46] J. Jin, I. Daly, Y. Zhang, X. Y. Wang, and A. Cichocki, "An optimized ERP brain-computer interface based on facial expression changes," *Journal of Neural Engineering*, vol. 11, no. 3, Article ID 036004, 11 pages, 2014.