A MODIFIED STRICTLY CONTRACTIVE PEACEMAN-RACHFORD SPLITTING METHOD FOR MULTI-BLOCK SEPARABLE CONVEX PROGRAMMING

SU-HONG JIANG AND MIN LI∗

School of Management and Engineering
Nanjing University, Nanjing 210093, China
(Communicated by Aviv Gibali)

Abstract. We propose a modified splitting method for a linearly constrained minimization model whose objective function is the sum of three convex functions without coupled variables. Our work is mainly inspired by the recently proposed strictly contractive Peaceman-Rachford splitting method (SC-PRSM) for a two-block separable convex minimization model. For the new method, we prove its convergence and estimate its convergence rates measured by iteration complexity in the nonergodic sense. We also test the SC-PRSM on the continuous resource allocation problem, and the numerical results show that our method has a competitive performance with the direct extension of ADMM which usually works well in practice but may fail to converge in theory.

1. Introduction. We consider the following convex minimization model with linear constraints and a separable objective function that is the sum of three convex functions without coupled variables:

$$\begin{align*}
\min & \quad \theta_1(x) + \theta_2(y) + \theta_3(z) \\
\text{s.t.} & \quad Ax + By + Cz = b,
\end{align*}$$

where $A \in \mathbb{R}^{m \times n_1}$, $B \in \mathbb{R}^{m \times n_2}$, $C \in \mathbb{R}^{m \times n_3}$, $b \in \mathbb{R}^m$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $\Omega^*$ is assumed to be nonempty.

This kind of separable convex programming models arises frequently in a broad spectrum of application domains such as inverse problems, statistical learning, image processing, computer vision, wireless communication network and so on. For such applications, usually each function in the objective function represents some specific explanation such as a data-fidelity or regularization term. It is thus a common sense that we should take full advantage of these functions’ properties/features.

2010 Mathematics Subject Classification. Primary: 90C25; Secondary: 90B50.

Keywords and phrases. Convex programming, multi-block, Peaceman-Rachford splitting method, convergence rate, resource allocation.

This research was supported by National Science Foundation of China (Grant No. 11401300, 71602083, 71671085 and 11001053) and Program for New Century Excellent Talents in University (Grant No. NCET-12-0111).

∗ Corresponding author: MIN LI.
individually when designing efficient algorithms, instead of treating all the functions/variables as a whole. For a similar separable model but with only a two-block objective function:

$$\begin{align*}
\min & \quad \theta_1(x) + \theta_2(y) \\
\text{s.t.} & \quad Ax + By = b, \\
& \quad x \in \mathcal{X}, y \in \mathcal{Y},
\end{align*}$$

i.e., $\theta_3(z) = 0$ and $C = 0$ in (1), the alternating direction method of multipliers (ADMM), which was proposed originally in [11], becomes extremely popularly recently, because of its simplicity of implementation, wide applicability in various areas and numerical efficiency. The iterative scheme of ADMM for (2) reads as

$$\begin{align*}
x^{k+1} &= \arg \min \{ L(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\
y^{k+1} &= \arg \min \{ L(x^{k+1}, y, \lambda^k) \mid y \in \mathcal{Y} \}, \\
\lambda^{k+1} &= \lambda^k - \beta (Ax^{k+1} + By^{k+1} - b),
\end{align*}$$

where

$$L(x, y, \lambda) := \theta_1(x) + \theta_2(y) - \lambda^T (Ax + By - b) + \frac{\beta}{2} \| Ax + By - b \|^2$$

is the augmented Lagrangian function of (2) with the Lagrange multiplier $\lambda$ and the penalty parameter $\beta > 0$. Note that ADMM originates from the Douglas-Rachford splitting method (DRSM) in [8, 23], see e.g., [9]. Another equally important splitting method in PDE literature is the Peaceman-Rachford splitting method (PRSM) in [23, 27]. Applying the PRSM to the model (2), we obtain the scheme

$$\begin{align*}
x^{k+1} &= \arg \min \{ L(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\
\lambda^{k+1} &= \lambda^k - \alpha \beta (Ax^{k+1} + By^k - b), \\
y^{k+1} &= \arg \min \{ L(x^{k+1}, y, \lambda^{k+1}) \mid y \in \mathcal{Y} \}, \\
\lambda^{k+1} &= \lambda^{k+1} - \alpha \beta (Ax^{k+1} + By^{k+1} - b).
\end{align*}$$

Although the scheme (4) differs from (3) for (2) only in that the Lagrangian multiplier is updated twice in a symmetric way at each iteration and the scheme indeed works well for some applications (see [10]), it is elaborated in [5] that the scheme (4) is not necessarily strictly contractive because the sequence generated by (4) is not necessarily strictly contractive. This difficulty inspired the so-called strictly contractive PRSM (SC-PRSM) method in [14]:

$$\begin{align*}
x^{k+1} &= \arg \min \{ L(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\
\lambda^{k+1} &= \lambda^k - \alpha \beta (Ax^{k+1} + By^k - b), \\
y^{k+1} &= \arg \min \{ L(x^{k+1}, y, \lambda^{k+1}) \mid y \in \mathcal{Y} \}, \\
\lambda^{k+1} &= \lambda^{k+1} - \alpha \beta (Ax^{k+1} + By^{k+1} - b).
\end{align*}$$

Note that the parameter $\alpha \in (0, 1)$ in (5) serves as a factor enforcing the strict contraction of the iterative sequence and thus ensuring the convergence. It has been shown in [14] that the SC-RPSM usually outperforms ADMM for a wide range of applications.

The success of ADMM and SC-PRSM encourages us to extend them to the three-block model (1) in the straightforward way and thus obtain the schemes...
Algorithm 1 A Modified SC-PRSM Algorithm

where the scheme (5) while its convergence is guaranteed under mild assumptions. The Rachford splitting method (SC-PRSM) – Algorithm 1 that inherits the advantages in infancy. Research on extending the scheme (5) to multi-block separable convex models is still numerical performance is inherited, see, e.g., [12, 15, 16, 20]. As comparison, the so that the convergence is ensured while the simplicity of implementation and good are a set of papers focusing on how to modify the direct extension of ADMM (6) when extending (3) and (5) to the multi-block model (1). In the literature, there is not straightforward and we should discuss the algorithmic design sophisticatedly the well-known and simple schemes (3) and (5) to the more complicated model (1). For (7), it is also shown in [13] that the convergence is not necessarily guaranteed, either.

Despite of some empirical efficiency in e.g., [28, 29], the scheme (6) has been proved in [3] to be not necessarily convergent. Some study on how to ensure the convergence of the scheme (6) under more assumptions can be found in the literature, e.g., [2, 4, 19, 21, 22]. For (7), it is also shown in [13] that the convergence is not necessarily guaranteed, either.

The failure of convergence of the schemes (6) and (7) indicates that how to extend the well-known and simple schemes (3) and (5) to the more complicated model (1) is not straightforward and we should discuss the algorithmic design sophisticatedly when extending (3) and (5) to the multi-block model (1). In the literature, there are a set of papers focusing on how to modify the direct extension of ADMM (6) so that the convergence is ensured while the simplicity of implementation and good numerical performance is inherited, see, e.g., [12, 15, 16, 20]. As comparison, the research on extending the scheme (5) to multi-block separable convex models is still in infancy.

In the following, we will propose a modified strictly contractive Peaceman-Rachford splitting method (SC-PRSM) – Algorithm 1 that inherits the advantages of the scheme (5) while its convergence is guaranteed under mild assumptions. Denote

\[ \mathbf{v} := \begin{pmatrix} y \\ z \\ \lambda \end{pmatrix}. \quad (8) \]

**Algorithm 1 A Modified SC-PRSM Algorithm**

Let \( \alpha \in (0, 1/2) \), \( \gamma \in (0, 1/2) \) and \( \beta \in (0, +\infty) \) be given parameters. Choose \( \mathbf{v}^0 := (y^0, z^0, \lambda^0) \in \mathcal{Y} \times \mathcal{Z} \times \mathbb{R}^m \) and set \( k = 0 \).

**Step 1.** Find \( x^{k+1} \in \mathcal{X} \) such that

\[ x^{k+1} := \arg \min \{ L(x, y^k, z^k, \lambda^k) \mid x \in \mathcal{X} \} \]. \quad (9)

**Step 2.** Update the Lagrange multiplier

\[ \lambda^{k+\frac{1}{2}} := \lambda^k - \alpha \beta (Ax^{k+1} + By^k + Cz^k - b). \] \quad (10)
Step 3. Compute $\tilde{v}^k := (\tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in Y \times Z \times \mathbb{R}^m$ parallelly
\[
\begin{aligned}
\tilde{y}^k &:= \arg\min \{ L(x^{k+1}, y, z, \lambda^{k+\frac{1}{2}}) \mid y \in Y \}, \\
\tilde{z}^k &:= \arg\min \{ L(x^{k+1}, y^k, z, \lambda^{k+\frac{1}{2}}) \mid z \in Z \}, \\
\tilde{\lambda}^k &:= \lambda^k - \beta (Ax^{k+1} + By^k + Cz^k - b).
\end{aligned}
\tag{11}
\]

Step 4. Set $\tilde{w}^k := (x^{k+1}, \tilde{v}^k)$. If a termination criterion is met, stop, otherwise go to Step 5.

Step 5. Compute $v^{k+1} := (y^{k+1}, z^{k+1}, \lambda^{k+1}) \in Y \times Z \times \mathbb{R}^m$ by
\[
v^{k+1} := v^k - \gamma M(v^k - \tilde{v}^k),
\tag{12}
\]
where
\[
M := \begin{pmatrix}
I_{n_2} & 0 & 0 \\
0 & I_{n_3} & 0 \\
-\alpha \beta B & -\alpha \beta C & 2\alpha I_m
\end{pmatrix},
\tag{13}
\]
and 0 is a matrix with all zero entries in appropriate dimensionality. Set $k := k + 1$ and go to Step 1.

Our modification in the new algorithm is to twist the scheme (7) slightly and then employ a simple correction step with a constant step size to correct the output. The iterative sequence is thus proved to be strictly contractive with respect to the solution set of (1); furthermore, the strict contraction property ensures the convergence of the new algorithm and enables us to estimate its worst-case convergence rate in terms of iteration complexity.

The rest of this paper is organized as follows. In Section 2, we summarize some preliminaries and develop some simple results that are useful for further analysis. Then, we prove the contractive properties and global convergence for the sequence generated by Algorithm 1 in Section 3. Section 4 is devoted to establishing the convergence rate in the nonergodic sense for Algorithm 1. In Section 5, we investigate the numerical performance of our modified SC-PRSM applied to the resource allocation problem. Finally, we draw some conclusions in Section 6.

2. Preliminaries. In this section, we summarize some useful preliminaries known in the literature and prove some simple conclusions for further analysis.

2.1. Variational reformulation of (1). First, as the work [17, 18] for analyzing the convergence rate, we need a variational inequality (VI) reformulation of the model (1) and a characterization of its solution set. More specifically, solving (1) is equivalent to finding $w^* = (x^*, y^*, z^*, \lambda^*) \in \Omega := X \times Y \times Z \times \mathbb{R}^m$ such that
\[
\text{VI}(\Omega, F, \theta) : \quad \theta(u) - \theta(w^*) + (w - w^*)^T F(w^*) \geq 0 \quad \forall w \in \Omega, \tag{14a}
\]
where
\[
\begin{aligned}
\theta(u) &:= \theta_1(x) + \theta_2(y) + \theta_3(z), \\
F(w) &:= \begin{pmatrix}
-A^T \lambda \\
-B^T \lambda \\
-C^T \lambda \\
Ax + By + Cz - b
\end{pmatrix},
\end{aligned}
\tag{14b}
\]
and
\[
\begin{aligned}
F(w) &:= \begin{pmatrix}
-A^T \lambda \\
-B^T \lambda \\
-C^T \lambda \\
Ax + By + Cz - b
\end{pmatrix},
\end{aligned}
\tag{14b}
\]
and
\[
\theta(u) := \theta_1(x) + \theta_2(y) + \theta_3(z). \tag{14c}
\]
Since the mapping $F(w)$ defined in (14b) is affine with a skew-symmetric matrix, it is monotone.
Lemma 1.2 in [7] is useful for establishing a worst-case $o(1/t)$ convergence rate on the consecutive iterates distance. We list it as follows.

**Lemma 2.1.** If a sequence $\{a_t\} \subseteq \mathbb{R}$ obeys: (1) $a_t \geq 0$; (2) $\sum_{t=0}^{\infty} a_t < +\infty$; (3) $a_t \leq a_{t-1}$ for any integer $t \geq 1$, then we have $a_t = o(1/t)$.

### 2.2. Some notations

As mentioned in [1] for ADMM, the variable $x$ is an intermediate variable during the iteration since it essentially requires only $v_k = (y^k, z^k, \lambda^k)$ in (9)-(12) to generate the $(k+1)$-th iterate. For this reason, we define the notation

$$V^* := \{v^* = (y^*, z^*, \lambda^*) \mid w^* = (x^*, y^*, z^*, \lambda^*) \in \Omega^*\};$$

and it suffices to analyze the convergence rate of the sequence $\{v_k\}$ to the set $V^*$ in order to study the convergence rate of the sequence $\{w_k\}$ generated by (9)-(12).

Note $V^*$ is also closed and convex.

Then, we define some matrices in order to present our analysis in a compact way. Let

$$Q := \begin{pmatrix}
\beta B^T B & 0 & -\alpha B^T \\
0 & \beta C^T C & -\alpha C^T \\
-B & -C & \frac{1}{2} I_m
\end{pmatrix} \tag{15}$$

and

$$H := \frac{1}{2} \begin{pmatrix}
(2 - \alpha)\beta B^T B & -\alpha \beta B^T C & -B^T \\
-\alpha \beta C^T B & (2 - \alpha)\beta C^T C & -C^T \\
-B & -C & \frac{1}{\alpha \beta} I_m
\end{pmatrix}. \tag{16}$$

Below we prove some assertions regarding the matrices just defined. These assertions will be used in our theoretical analysis about the convergence rate of (9)-(12); their role is to make our proof presentable in compact notation.

**Lemma 2.2.** The matrix $H$ defined in (16) is positive definite for $\alpha \in (0, 1/2)$.

**Proof.** It’s easy to prove that the matrix $H$ has the following decomposition form

$$H = \frac{1}{2\beta} \begin{pmatrix}
\beta B & 0 & 0 \\
0 & \beta C & 0 \\
0 & 0 & I_m
\end{pmatrix}^T \begin{pmatrix}
(2 - \alpha)I_m & -\alpha I_m & -I_m \\
-\alpha I_m & (2 - \alpha)I_m & -I_m \\
-I_m & -I_m & \frac{1}{\alpha \beta} I_m
\end{pmatrix} \begin{pmatrix}
\beta B & 0 & 0 \\
0 & \beta C & 0 \\
0 & 0 & I_m
\end{pmatrix}.$$

It is not difficult to verify that the matrix

$$\begin{pmatrix}
2 - \alpha & -\alpha & -1 \\
-\alpha & 2 - \alpha & -1 \\
-1 & -1 & \frac{1}{\alpha \beta}
\end{pmatrix}$$

is positive definite if and only if $\alpha \in (0, 1/2)$. Note that $B$ and $C$ are full column rank. The assertion is thus proved. \qed

**Lemma 2.3.** The matrices $M$, $Q$ and $H$ defined respectively in (13), (15) and (16) have the following relationships:

$$HM = Q \tag{17}$$

and for given $\alpha \in (0, 1/2)$ and $\gamma \in (0, 1/2)$, it holds that

$$G := Q^T + Q - \gamma M^T HM > 0. \tag{18}$$
Proof. Using the definitions of the matrices $M$, $Q$ and $H$, by a simple manipulation, we obtain

$$HM = \frac{1}{2} \begin{pmatrix} (2 - \alpha)\beta B^T B & -\alpha \beta B^T C & -B^T \\ -\alpha \beta C^T B & (2 - \alpha)\beta C^T C & -C^T \\ -B & -C & \frac{1}{\alpha \beta} I_m \end{pmatrix} \begin{pmatrix} I_{n_2} & 0 & 0 \\ 0 & I_{n_3} & 0 \\ -\alpha \beta B & -\alpha \beta C & 2\alpha I_m \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 2\beta B^T B & 0 & -2\alpha B^T \\ 0 & 2\beta C^T C & -2\alpha C^T \\ -2B & -2C & \frac{2}{\alpha \beta} I_m \end{pmatrix} = Q.$$

The first assertion (17) is proved. Consequently, we get

$$M^T HM = M^T Q = \begin{pmatrix} I_{n_2} & 0 & -\alpha \beta B^T \\ 0 & I_{n_3} & -\alpha \beta C^T \\ 0 & 0 & 2\alpha I_m \end{pmatrix} \begin{pmatrix} \beta B^T B & 0 & -\alpha B^T \\ 0 & \beta C^T C & -\alpha C^T \\ -B & -C & \frac{1}{\alpha \beta} I_m \end{pmatrix}$$

Using (15) and the above equation, we have

$$Q^T + Q - \gamma M^T HM = \begin{pmatrix} (2 - \gamma (1 + \alpha))\beta B^T B & -\gamma \alpha \beta B^T C & (2\gamma \alpha - (\alpha + 1))B^T \\ -\gamma \alpha \beta C^T B & (2 - \gamma (1 + \alpha))\beta C^T C & (2\gamma \alpha - (\alpha + 1))C^T \\ (2\gamma \alpha - (\alpha + 1))B & (2\gamma \alpha - (\alpha + 1))C & \frac{2 - 2\gamma \alpha}{\alpha \beta} I_m \end{pmatrix}$$

$$= \frac{1}{\beta} \begin{pmatrix} \beta B & 0 & 0 \\ 0 & \beta C & 0 \\ 0 & 0 & I_m \end{pmatrix}^T \begin{pmatrix} (2 - \gamma (1 + \alpha))I_m & -\gamma \alpha I_m & (2\gamma \alpha - (\alpha + 1))I_m \\ -\gamma \alpha I_m & (2 - \gamma (1 + \alpha))I_m & (2\gamma \alpha - (\alpha + 1))I_m \\ (2\gamma \alpha - (\alpha + 1))I_m & (2\gamma \alpha - (\alpha + 1))I_m & (2 - 2\gamma \alpha)I_m \end{pmatrix}$$

$$\times \begin{pmatrix} \beta B & 0 & 0 \\ 0 & \beta C & 0 \\ 0 & 0 & I_m \end{pmatrix}.$$  (19)

Denote $H_{\alpha,\gamma}$ by

$$H_{\alpha,\gamma} = \begin{pmatrix} (2 - \gamma (1 + \alpha)) & -\gamma \alpha & (2\gamma \alpha - (\alpha + 1)) \\ -\gamma \alpha & (2 - \gamma (1 + \alpha)) & (2\gamma \alpha - (\alpha + 1)) \\ (2\gamma \alpha - (\alpha + 1)) & (2\gamma \alpha - (\alpha + 1)) & (2 - 2\gamma \alpha) \end{pmatrix}.$$  

By direct computation, we are able to show all the leading principle minors are positive definite when $\alpha \in (0, 1/2)$ and $\gamma \in (0, 1/2)$, and consequently $H_{\alpha,\gamma} \succ 0$. Note that $B$ and $C$ are full column rank. We get the assertion (18). The lemma is proved.  

3. Convergence analysis. In this section, we analyze the contraction property and global convergence for the sequence $\{u^k\}$ generated by the proposed algorithm. From now on, we assume the parameters $\alpha$ and $\gamma$ are chosen in the interval $(0, 1/2)$. Then according to Lemmas 2.2 and 2.3, the matrices $H$ and $G$ are positive definite.
To further alleviate the notation in our analysis, first, we need to define an auxiliary sequence \( \{\tilde{w}^k\} \) as

\[
\tilde{w}^k := \begin{pmatrix}
\tilde{x}^k \\
\tilde{y}^k \\
\tilde{z}^k \\
\tilde{\lambda}^k
\end{pmatrix} := \begin{pmatrix}
x^{k+1} \\
y^{k+1} \\
z^{k+1} \\
\lambda^{k+1}
\end{pmatrix},
\]

(20)

where \( (x^{k+1}, y^{k+1}, z^{k+1}, \lambda^{k+1}) \) is generated by (9) and (11). With the notation of \( \tilde{w}^k \), the proposed algorithm can be written as

\[
\begin{aligned}
\tilde{x}^k &:= \text{arg min} \left\{ \theta_1(x) - (\lambda^k)^T Ax + \frac{\beta}{2}\|Ax + By^k + Cz^k - b\|^2 \mid x \in \mathcal{X} \right\}, \\
\tilde{y}^k &:= \text{arg min} \left\{ \theta_2(y) - [\lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)]^T By + \frac{\beta}{2}\|A\tilde{x}^k + By + Cz^k - b\|^2 \mid y \in \mathcal{Y} \right\}, \\
\tilde{z}^k &:= \text{arg min} \left\{ \theta_3(z) - [\lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)]^T Cz + \frac{\beta}{2}\|A\tilde{x}^k + By^k + Cz - b\|^2 \mid z \in \mathcal{Z} \right\}, \\
\tilde{\lambda}^k &:= \lambda^k - \beta(A\tilde{x}^k + By^k + Cz^k - b).
\end{aligned}
\]

(21)

Now, we start to prove some properties for the sequence \( \{\tilde{w}^k\} \) defined in (20).

**Lemma 3.1.** For given \( v^k \in \mathcal{Y} \times \mathcal{Z} \times \mathbb{R}^m \), let \( w^{k+1} \) and \( \tilde{w}^k \) be generated by Algorithm 1. Then, we have \( \tilde{w}^k \in \Omega \) and

\[
\theta(\tilde{w}^k) - \theta(w^k) + (\tilde{w}^k - w^k)^T F(\tilde{w}^k) \leq -(v^k - \tilde{v}^k)^T Q(v^k - \tilde{v}^k) \quad \forall w \in \Omega,
\]

(22)

where the matrix \( Q \) is defined in (15).

**Proof.** Since \( x^{k+1} = \tilde{x}^k \), by deriving the first-order optimality condition of the \( x \)-minimization problem in (21), we have

\[
\theta_1(x) - \theta_1(\tilde{x}^k) + (x - \tilde{x}^k)^T \{A^T [\beta(A\tilde{x}^k + By^k + Cz^k - b) - \lambda^k]\} \geq 0 \quad \forall x \in \mathcal{X}.
\]

(23)

Using the notation of \( \tilde{\lambda}^k \) in (21), the inequality (23) can be written as

\[
\theta_1(x) - \theta_1(\tilde{x}^k) + (x - \tilde{x}^k)^T (-A^T \tilde{\lambda}^k) \geq 0 \quad \forall x \in \mathcal{X}.
\]

(24)

Similarly, by deriving the first-order optimality condition of the \( y \)-minimization problem in (21), we get

\[
\begin{aligned}
\theta_2(y) - \theta_2(\tilde{y}^k) + (y - \tilde{y}^k)^T B^T \{\beta(A\tilde{x}^k + By^k + Cz^k - b) - [\lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)]\} &\geq 0 \quad \forall y \in \mathcal{Y}.
\end{aligned}
\]

(25)

Again, using the notation of \( \tilde{\lambda}^k \) in (21), we have

\[
\begin{aligned}
\beta(A\tilde{x}^k + By^k + Cz^k - b) - [\lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)]
&= -[\tilde{\lambda}^k - \lambda^k] + \beta B(\tilde{y}^k - y^k) - [\lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)]
&= -[\tilde{\lambda}^k + \beta B(\tilde{y}^k - y^k) - \alpha(\lambda^k - \tilde{\lambda}^k)].
\end{aligned}
\]

Consequently, it follows from (25) that

\[
\begin{aligned}
\theta_2(y) - \theta_2(\tilde{y}^k) + (y - \tilde{y}^k)^T \{-B^T \tilde{\lambda}^k + \beta B^T B(\tilde{y}^k - y^k) - \alpha B^T (\tilde{\lambda}^k - \lambda^k)\} &\geq 0 \quad \forall y \in \mathcal{Y}.
\end{aligned}
\]

(26)
Similarly, we have
\[
\theta_3(z) - \theta_3(\tilde{z}^k) + (z - \tilde{z}^k)^T (-C^T\tilde{\lambda}^k + \beta C^T C(\tilde{z}^k - z^k) - \alpha C^T (\tilde{\lambda}^k - \lambda^k)) \geq 0 \quad \forall z \in \mathbb{Z}.
\] (27)

In addition, based on (21) we have
\[
(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) - B(\tilde{y}^k - y^k) - C(\tilde{z}^k - z^k) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) = 0.
\] (28)

Combining (24), (26), (27) and (28) together, we get \(\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega; \) and for any \(w = (x, y, z, \lambda) \in \Omega,\) it holds
\[
\theta(u) - \theta(\tilde{w}^k) + \left( \begin{array}{c}
x - \tilde{x}^k \\
y - \tilde{y}^k \\
z - \tilde{z}^k \\
\lambda - \tilde{\lambda}^k
\end{array} \right)^T \left( \begin{array}{c}
-A^T \tilde{\lambda}^k \\
-B^T \tilde{\lambda}^k \\
-C^T \tilde{\lambda}^k \\
A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b
\end{array} \right) + \left( \begin{array}{c}
-\alpha B^T B(\tilde{y}^k - y^k) - \alpha C^T C(\tilde{z}^k - z^k) \\
-\beta B^T B(\tilde{y}^k - y^k) - \alpha C^T (\tilde{\lambda}^k - \lambda^k)
\end{array} \right) \geq 0.
\] (29)
By using the notations of \(Q\) in (15), \(w\) and \(F\) in (14b) and \(v\) in (8), from the above inequality we have the assertion (22). The proof is complete.

**Lemma 3.2.** Let the sequences \(\{w^k\}\) and \(\{\tilde{w}^k\}\) be generated by Algorithm 1, and \(H\) be defined in (16). Then, \(\tilde{w}^k\) is a solution of VI(\(\Omega, F, \theta\)) if \(\|v^k - v^{k+1}\|_H = 0.\)

**Proof.** It follows from (22), (17) and (12) that
\[
\theta(\tilde{w}^k) - \theta(u) + (\tilde{w}^k - v^k)^T F(\tilde{w}^k) \leq -(v - \tilde{v}^k)^T H(v^k - \tilde{v}^k) = -\frac{1}{\gamma}(v - \tilde{v}^k)^T H(v^k - v^{k+1}) \quad \forall w \in \Omega.
\]
If \(\|v^k - v^{k+1}\|_H = 0,\) we have
\[
\theta(u) - \theta(\tilde{w}^k) + (w - \tilde{w}^k)^T F(\tilde{w}^k) \geq 0 \quad \forall w \in \Omega,
\]
which implies that \(\tilde{w}^k\) is a solution of VI(\(\Omega, F, \theta\)). The proof is complete.

The following lemma aims at further bounding the term \((v - \tilde{v}^k)^T Q(v^k - \tilde{v}^k)\) found in Lemma 3.1 by the difference of two quadratic terms involving two consecutive iterates of the sequence \(\{v^k\}\) and a quadratic term involving \(v^k\) and the auxiliary iterate \(\tilde{v}^k\). This refined bound is convenient for the manipulation over the whole sequence \(\{v^k\}\) recursively and thus for establishing the convergence rate of \(\{v^k\}\) in the nonergodic sense.

**Lemma 3.3.** Let the sequences \(\{w^k\}\) and \(\{\tilde{w}^k\}\) be generated by Algorithm 1, and \(M, Q\) and \(H\) be defined in (13), (15) and (16), respectively. Then we have
\[
(v - \tilde{v}^k)^T Q(v^k - \tilde{v}^k) \geq \frac{1}{2\gamma}(\|v - v^{k+1}\|_H^2 - \|v - v^k\|_H^2) + \frac{1}{2}\|v^k - \tilde{v}^k\|_G^2 \quad \forall v \in \mathcal{V}.
\] (30)

**Proof.** By using \(Q = HM\) and \(M(v^k - \tilde{v}^k) = \frac{1}{\gamma}(v^k - v^{k+1})\) (see (12)), it follows that
\[
(v - \tilde{v}^k)^T Q(v^k - \tilde{v}^k) = (v - \tilde{v}^k)^T HM(v^k - \tilde{v}^k) = \frac{1}{\gamma}(v - \tilde{v}^k)^T H(v^k - v^{k+1}).
\] (31)
For the vectors $a, b, c, d$ in the same space and a matrix $H$ with appropriate dimensionality, we have the identity
\[(a - b)^T H (c - d) = \frac{1}{2} (\|a - d\|_H^2 - \|a - c\|_H^2) + \frac{1}{2} (\|c - b\|_H^2 - \|d - b\|_H^2).
\]
In this identity, we take
\[a = v, \quad b = \tilde{v}^k, \quad c = v^k \quad \text{and} \quad d = v^{k+1},\]
and submit it to the right-hand side of (31). The resulting equation is
\[\begin{align*}
(v - \tilde{v}^k)^T Q (v^k - \tilde{v}^k) &= \frac{1}{2\gamma} (\|v - v^{k+1}\|_H^2 - \|v - v^k\|_H^2) + \frac{1}{2\gamma} (\|v^k - \tilde{v}^k\|_H^2 - \|v^{k+1} - \tilde{v}^k\|_H^2).
\end{align*}
\]
Now, we deal with the last term of the right-hand side of (32). By using (12) and (17), we get
\[\begin{align*}
\|v^k - \tilde{v}^k\|_H^2 - \|v^{k+1} - \tilde{v}^k\|_H^2 &= \|v^k - \tilde{v}^k\|_H^2 - \|(v^k - \tilde{v}^k) - (v^k - v^{k+1})\|_H^2 \\
\quad \text{(12)} &= \|v^k - \tilde{v}^k\|_H^2 - \|(v^k - \tilde{v}^k) - \gamma M (v^k - \tilde{v}^k)\|_H^2 \\
\quad \text{(12)} &= 2\gamma (v^k - \tilde{v}^k)^T H M (v^k - \tilde{v}^k) - \gamma^2 (v^k - \tilde{v}^k)^T H M (v^k - \tilde{v}^k) \\
\quad \text{(17)} &= \gamma (v^k - \tilde{v}^k)^T (Q^T + Q - \gamma M^T H M) (v^k - \tilde{v}^k) \\
\quad \text{(18)} &= \gamma \|v^k - \tilde{v}^k\|_G^2.
\end{align*}\]
Substituting it in (32), we obtain the assertion (30). The proof is complete. \(\square\)

Now we are ready to present an inequality which shows the contraction property for the iterative sequence.

**Theorem 3.4.** For given $v^k \in \mathcal{V} \times \mathcal{Z} \times \mathbb{R}^m$, let $w^{k+1}$ be generated by Algorithm 1, and $H$ be defined in (16). Then, there is a constant $c_0 > 0$ such that
\[\|v^{k+1} - v^*\|^2_H \leq \|v^k - v^*\|^2_H - \frac{c_0}{\gamma} \|v^k - v^{k+1}\|^2_H \quad \forall v^* \in \mathcal{V}^*.
\]

**Proof.** Replacing the right-hand side term in (22) with the inequality (30), and using
\[(\hat{w}^k - w)^T F(w) \leq (\hat{w}^k - w)^T F(\hat{w}^k),\]
we obtain
\[\theta(\hat{w}^k) - \theta(u) + (\hat{w}^k - w)^T F(w) \leq \frac{1}{2\gamma} (\|v - v^k\|_H^2 - \|v - v^{k+1}\|_H^2) - \frac{1}{2} \|v^k - \tilde{v}^k\|_G^2 \quad \forall w \in \Omega.
\]
Setting $w = w^*$ in the above inequality, where $w^*$ being an arbitrary solution point in $\Omega^*$, we get
\[\|v^k - v^*\|^2_H - \|v^{k+1} - v^*\|^2_H \geq \gamma \|v^k - \tilde{v}^k\|^2_G + 2\gamma \{\theta(\hat{w}^k) - \theta(u^*) + (\hat{w}^k - w^*)^T F(w^*)\},
\]
which implies that
\[\|v^{k+1} - v^*\|^2_H \leq \|v^k - v^*\|^2_H - \gamma \|v^k - \tilde{v}^k\|^2_G.
\]


Furthermore, by using (11) and \(Ax\), we have
\[
\|v^{k+1} - v^*\|_H^2 \leq \|v^k - v^*\|_H^2 - \gamma c_0 \|v^k - \tilde{v}^k\|_{M^T H M}^2
\]
\[
= \|v^k - v^*\|_H^2 - \frac{c_0}{\gamma} \|\gamma M(v^k - \tilde{v}^k)\|_H^2
\]
\[
= \|v^k - v^*\|_H^2 - \frac{c_0}{\gamma} \|v^k - v^{k+1}\|_H^2 \quad \forall v^* \in V^*.
\]

Thus, the assertion (33) is proved.

**Theorem 3.5.** Let the sequences \(\{w^k\}\) and \(\{\tilde{w}^k\}\) be generated by Algorithm 1. Then, we have that the sequences \(\{w^k\}\) and \(\{\tilde{w}^k\}\) both converge to a point in \(\Omega^*\).

**Proof.** It follows from (35) that the sequence \(\{v^k\}\) is Fejér-monotone with respect to \(V^*\), and
\[
\sum_{k=0}^{\infty} \gamma \|v^k - \tilde{v}^k\|_G^2 < +\infty.
\]

Since \(H > 0\) and \(G > 0\), the sequences \(\{v^k\}\) and \(\{\tilde{v}^k\}\) are bounded, and
\[
\lim_{k \to \infty} (v^k - \tilde{v}^k) = 0. \tag{36}
\]

Furthermore, by using (11) and \(Ax^* + By^* + Cz^* = b\), for any \(w^* \in \Omega^*\), we know
\[
\|x^{k+1} - x^*\|_{A^T A}^2 = \|Ax^{k+1} - Ax^*\|^2
\]
\[
= \|\frac{1}{\beta}(\lambda^k - \tilde{\lambda}^k) - (By^k - By^*) - (Cz^k - Cz^*)\|^2, \tag{37}
\]
and thus \(\{\|x^{k+1} - x^*\|_{A^T A}\}\) is bounded. This shows that the sequence \(\{x^k\}\) is also bounded as \(A\) is full column rank. Therefore, the sequence \(\{\tilde{w}^k = (x^{k+1}, \tilde{\omega}^k)\}\) is bounded.

Since the sequence \(\{\tilde{w}^k\}\) is bounded, there is a subsequence \(\{\tilde{w}^{k_j}\}\) which converges to a cluster point, say \(w^\infty := (x^\infty, y^\infty, z^\infty, \lambda^\infty) \in \Omega\). Taking limits on both sides of (22) along the subsequences \(\{\tilde{w}^{k_j}\}\) and \(\{w^{k_j}\}\), and using (36), we obtain that
\[
\theta(u) - \theta(w^\infty) + (w - w^\infty)^T F(w^\infty) \geq 0 \quad \forall w \in \Omega,
\]
i.e., \(w^\infty \in \Omega^*\).

To complete the proof, we show next that \(w^\infty\) is actually the unique limit of \(\{w^k\}\) and \(\{\tilde{w}^k\}\). Replacing \(v^*\) by \(v^\infty\) in (33), for any integer \(k \geq k_j\), we have
\[
\|v^{k+1} - v^\infty\|_H \leq \|v^{k_j} - v^\infty\|_H. \tag{38}
\]

Note that
\[
\lim_{j \to \infty} \|v^{k_j} - v^\infty\|_H = \lim_{j \to \infty} \|(v^{k_j} - \tilde{v}^{k_j}) + (\tilde{v}^{k_j} - v^\infty)\|_H = 0.
\]

Therefore, from (36) and (38), we get
\[
\lim_{k \to \infty} \|v^k - v^\infty\|_H = 0,
\]
i.e.,
\[
\lim_{k \to \infty} v^k = v^\infty.
\]
Combining this with (36), we also have
\[
\lim_{k \to \infty} \tilde{v}^k = v^\infty.
\]
It follows from (36), (37) and \( w^\infty \in \Omega^* \) that
\[
\lim_{k \to \infty} \|x^{k+1} - x^\infty\|_A r_A = \lim_{k \to \infty} \frac{1}{\beta} (\lambda_k - \tilde{\lambda}_k) = B(y^k - y^\infty) - C(z^k - z^\infty) = 0. \tag{39}
\]
Since \( A \) is full column rank, we also obtain that \( \lim_{k \to \infty} x^k = x^\infty \). Therefore, we have shown that the sequences \( \{w^k := (x^k, v^k)\} \) and \( \{\tilde{w}^k := (x^{k+1}, \tilde{v}^k)\} \) both converge to \( w^\infty \in \Omega^* \). The proof is complete. \( \square \)

4. Convergence rate in the nonergodic sense. In this section, we show that the worst-case convergence rate of Algorithm 1 is \( o(1/t) \) in the nonergodic sense. Our starting point for the analysis is the inequality (33), and a crucial property is the monotonicity of the sequence \( \{\|v^k - v^{k+1}\|_H^2\} \). That is, we will prove that
\[
\|v^{k+1} - v^{k+2}\|_H^2 \leq \|v^k - v^{k+1}\|_H^2, \quad \forall k \geq 0.
\]

We first take a closer look at the assertion (22) in Lemma 3.1.

**Lemma 4.1.** Let the sequences \( \{w^k\} \) and \( \{\tilde{w}^k\} \) be generated by Algorithm 1, and the matrix \( Q \) be defined in (15). Then, we have
\[
\theta(\tilde{w}^{k+1}) - \theta(\tilde{w}^k) + (\tilde{w}^{k+1} - \tilde{w}^k)^T F(\tilde{w}^k) \geq \tilde{v}^{k+1} - \tilde{v}^k)^T Q (v^k - \tilde{v}^k). \tag{41}
\]

**Proof.** Setting \( w = \tilde{w}^{k+1} \) in (22), we have
\[
\theta(\tilde{w}^{k+1}) - \theta(\tilde{w}^k) + (\tilde{w}^{k+1} - \tilde{w}^k)^T F(\tilde{w}^k) \geq \tilde{v}^{k+1} - \tilde{v}^k)^T Q (v^k - \tilde{v}^k). \tag{41}
\]
Note that (22) is also true for \( k := k + 1 \) and thus
\[
\theta(v^k) - \theta(\tilde{w}^{k+1}) + (w - \tilde{w}^{k+1})^T F(\tilde{w}^{k+1}) \geq (v - \tilde{w}^{k+1})^T Q (v^{k+1} - \tilde{w}^{k+1}) \quad \forall w \in \Omega.
\]
Setting \( w = \tilde{w}^k \) in the above inequality, we obtain
\[
\theta(\tilde{w}^k) - \theta(\tilde{w}^{k+1}) + (\tilde{w}^k - \tilde{w}^{k+1})^T F(\tilde{w}^{k+1}) \geq (\tilde{v}^k - \tilde{w}^{k+1})^T Q (v^{k+1} - \tilde{w}^{k+1}), \tag{42}
\]
Adding (41) and (42) and using the monotonicity of \( F \), we get (40) immediately. \( \square \)

**Lemma 4.2.** Let the sequences \( \{w^k\} \) and \( \{\tilde{w}^k\} \) be generated by Algorithm 1, and the matrices \( M, Q \) and \( H \) be defined in (13), (15) and (16), respectively. Then, we have
\[
(v^k - \tilde{v}^k)^T M^T H M \{(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\} \geq \frac{1}{2\gamma} \|(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\|_{(Q^2 + Q)}^2. \tag{43}
\]

**Proof.** Adding the equation
\[
[(v^k - v^{k+1}) - (\tilde{v}^k - \tilde{v}^{k+1})]^T Q \{(v^k - v^{k+1}) - (\tilde{v}^k - \tilde{v}^{k+1})\} = \frac{1}{2} \|(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\|_{(Q^2 + Q)}^2
\]
to both sides of (40), we get
\[
(v^k - v^{k+1})^T Q \{(v^k - v^{k+1}) - (\tilde{v}^k - \tilde{v}^{k+1})\} \geq \frac{1}{2} \|(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\|_{(Q^2 + Q)}^2. \tag{44}
\]
By using (see (12) and (17))
\[
v^k - v^{k+1} = \gamma M(v^k - \tilde{v}^k) \quad \text{and} \quad Q = HM,
\]
to the term \( v^k - v^{k+1} \) in the left-hand side of (44), we obtain
\[
\gamma(v^k - \tilde{v}^k)^T M^T H M \{(v^k - v^{k+1}) - (\tilde{v}^k - \tilde{v}^{k+1})\} \geq \frac{1}{2} \|(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\|_{(Q^2 + Q)}^2,
\]
and the lemma is proved. □

Let the sequences \( \{ w^k \} \) and \( \{ \tilde{w}^k \} \) be generated by Algorithm 1, and the matrix \( H \) be defined in (16). Then, we have

\[
\| v^{k+1} - v^k \|^2_H \leq \| v^k - v^{k+1} \|^2_H. \tag{45}
\]

Proof. Setting \( a = M(v^k - \tilde{v}^k) \) and \( b = M(v^{k+1} - \tilde{v}^{k+1}) \) in the identity

\[
\|a\|^2_H - \|b\|^2_H = 2a^T H (a - b) = \|a - b\|^2_H,
\]

we obtain

\[
\| M(v^k - \tilde{v}^k) \|^2_H - \| M(v^{k+1} - \tilde{v}^{k+1}) \|^2_H = 2(v^k - \tilde{v}^k)^T M^T H M \{ (v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1}) \} - \| M(v^k - \tilde{v}^k) - M(v^{k+1} - \tilde{v}^{k+1}) \|^2_H.
\]

Inserting (43) into the first term of the right-hand side of the last equality, we obtain

\[
\| M(v^k - \tilde{v}^k) \|^2_H - \| M(v^{k+1} - \tilde{v}^{k+1}) \|^2_H \\
\geq \frac{1}{\gamma} \| (v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1}) \|^2_{Q^T + Q} - \| (v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1}) \|^2_{\gamma M^T H M} \\
= \frac{1}{\gamma} \| (v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1}) \|^2_{(Q^T + Q - \gamma M^T H M)}.
\]

The assertion (45) follows from the above inequality, the relationship (12) and (18) immediately. The proof is complete. □

From (33), it is easy to see that

\[
\sum_{k=0}^{\infty} \| v^k - v^{k+1} \|^2_H \leq \frac{\gamma}{c_0} \| v^0 - v^* \|^2_H \quad \forall v^* \in \mathcal{V}^*.
\tag{46}
\]

With Lemma 4.3, we can immediately derive a worst-case \( O(1/t) \) convergence rate in the nonergodic sense for the proposed algorithm.

Theorem 4.4. Let \( \{ w^t \} \) be the sequence generated by Algorithm 1. For any \( v^* \in \mathcal{V}^* \), we have

\[
\| v^t - v^{t+1} \|^2_H \leq \frac{\gamma}{(t+1)c_0} \| v^0 - v^* \|^2_H. \tag{47}
\]

Notice that \( \mathcal{V}^* \) is convex and closed. Let \( v^0 = (y^0, z^0, \lambda^0) \) be the initial iterate and \( d := \inf \{ \| v^0 - v^* \|_H | v^* \in \mathcal{V}^* \} \). Then, for any given \( \epsilon > 0 \), Theorem 4.4 shows that the proposed algorithm needs at most \( [\gamma d^2/(\epsilon c_0) - 1] \) iterations to ensure that \( \| v^k - v^{k+1} \|^2_H \leq \epsilon \). It follows from Lemma 3.2 that \( \tilde{v}^k \) is a solution of VI(\( \Omega, F, \theta \)) if \( \| v^k - \tilde{v}^k \|^2_H \leq 0 \). A worst-case \( O(1/t) \) convergence rate in the nonergodic sense for the proposed algorithm is thus established in Theorem 4.4.

Furthermore, we will prove a worst-case \( o(1/t) \) convergence rate in the nonergodic sense for the proposed algorithm.

Corollary 4.5. Let \( \{ w^t \} \) be the sequence generated by Algorithm 1. For any \( v^* \in \mathcal{V}^* \), we have

\[
\| v^t - v^{t+1} \|^2_H = o(1/t). \tag{48}
\]
allocation problem could be modeled as the total cost subject to the capacity constraints of the resources. Clearly, the manager wants to obtain a distributing strategy of the allocation which minimizes as \( \theta \) allocated by 3 users. The cost function of the allocation to the \( n \) SC-PRSM algorithm. Specifically, we consider a market in which Numerical results.

5. Corollary 4.6. Let \( \{\hat{w}^t := (\hat{x}^t, \hat{y}^t, \hat{z}^t, \hat{\lambda}^t)\} \) be defined by (20). Then we have
\[
\text{dist}^2(0, \partial \theta_1(\hat{x}^t) - A^T \hat{\lambda}^t) + \text{dist}^2(0, \partial \theta_2(\hat{y}^t) - B^T \hat{\lambda}^t) + \text{dist}^2(0, \partial \theta_3(\hat{z}^t) - C^T \hat{\lambda}^t) + \|A\hat{x}^t + B\hat{y}^t + C\hat{z}^t - b\|^2 = o(1/t),
\]
where “dist” denotes the distance between a point to a convex set.

Proof. From (29), we see clearly that
\[
\begin{pmatrix}
0 \\
Q(v^t - \hat{v}^t)
\end{pmatrix} \in \begin{pmatrix}
\partial \theta_1(\hat{x}^t) - A^T \hat{\lambda}^t \\
\partial \theta_2(\hat{y}^t) - B^T \hat{\lambda}^t \\
\partial \theta_3(\hat{z}^t) - C^T \hat{\lambda}^t \\
A\hat{x}^t + B\hat{y}^t + C\hat{z}^t - b
\end{pmatrix},
\]
and thus
\[
\text{dist}^2(0, \partial \theta_1(\hat{x}^t) - A^T \hat{\lambda}^t) + \text{dist}^2(0, \partial \theta_2(\hat{y}^t) - B^T \hat{\lambda}^t) + \text{dist}^2(0, \partial \theta_3(\hat{z}^t) - C^T \hat{\lambda}^t) + \|A\hat{x}^t + B\hat{y}^t + C\hat{z}^t - b\|^2 \leq \|Q(v^t - \hat{v}^t)\|^2 = \|HM(v^t - \hat{v}^t)\|^2 \leq \frac{1}{r^2}\|H(v^t - v^{t+1})\|^2.
\]
This together with Lemma 4.5 proves this corollary.

Note that the left side in (49) is the square residual of the KKT system at \( \hat{w}^t \). Hence Corollary 4.6 actually establishes a worst case \( o(1/t) \) convergence rate measured by the KKT square residual for Algorithm 1.

5. Numerical results. In this section, we test an allocation problem arising from market mechanisms (see, e.g., [6, 26, 30]) to verify the efficiency of the modified SC-PRSM algorithm. Specifically, we consider a market in which \( n \) resources are allocated by 3 users. The cost function of the allocation to the \( i \)-th user is defined as \( \theta_i \) and the amount of each resource is defined as \( b_j (j = 1, 2, \ldots, n) \). The market manager wants to obtain a distributing strategy of the allocation which minimizes the total cost subject to the capacity constraints of the resources. Clearly, the allocation problem could be modeled as
\[
\min \{\theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3) \mid x_1 + x_2 + x_3 = b, x_1, x_2, x_3 \geq 0\}.
\]
As [6], the cost functions are constructed based on the one-dimensional function \( \phi \) listed in Table 1. Specifically, for any given function \( \phi(s_i) \), we can easily extend it to \( n \) dimensions by defining
\[
\Phi(s) := \sum_{i=1}^{n} \phi(s_i),
\]
where \( s = (s_1, s_2, \ldots, s_n) \), \( w = (w_1, w_2, \ldots, w_n) \) and \( k = (k_1, k_2, \ldots, k_n) \). Note that for each cost function in Table 1, its proximity operator either has a closed form solution or can be computed by solving a one-dimensional polynomial system. As a result, the proximal operator of \( \theta_i \) can also be computed by a fast procedure and hence all the subproblems in the modified PRSM (or ADMM) can be solved with a fast speed.
Table 1. The function $\phi(s)$ for generating the cost function

| Name                        | $\phi(s_i) : \mathbb{R}_+ \to [0, \infty)$ | Parameters          |
|-----------------------------|-------------------------------------------|---------------------|
| Linear cost                 | $w_i s_i$                                 | $w_i \in U(1, 5)$  |
| Power cost                  | $k_i s_i^2 / s_i$ for $s_i \leq w_i / \sqrt{2 k_i}$; $w_i \sqrt{2 k_i} - w_i^2 / 2$ otherwise. | $k_i \in U(1, 5)$  |
| Piecewise quadratic cost    |                                           |                     |

For the allocation problem (51), we test the proposed modified SC-PRSM (abbreviated as MC-ADM), the alternating proximal gradient method [24] (abbreviated as APGM) and the direct extension of ADMM (6) (abbreviated as EADMM). Although (6) could fail to converge, its efficiency has been verified numerically in many applications and it could be served as a benchmark for measuring the performance of the ADMM type methods. In the implementation, we set $b = nI$ in the model; $\beta = 1$ and $\alpha = 0.49$ for the MC-ADM; $\beta = 1$, $r_1 = 1/6$ and $r_2 = 1/6$ for the APGM; and $\beta = 1$ for the EADMM. Figure 1 depicts the numerical performance of the three methods for the allocation problems with $n = 100, 400$ and $800$. Clearly, the EADMM (without theoretical convergence) performs the best for all cases. Our proposed MC-ADM with provable convergence outperforms the APGM and has a competitive performance with the EADMM. Specifically, the MC-ADM runs a bit slower than the EADMM at the first few iterations but they almost attain the optimal value at the same time.

![Figure 1. Evolutions of objective function values w.r.t. CPU time](image)

6. **Conclusions.** This paper proposes a modified SC-PRSM for solving multi-block separable convex programming problems. The proposed algorithm perfectly inherits the advantage of the direct extension of ADMM (6) and SC-PRSM (7), and it can fully exploit the properties of $\theta_i$’s. We also discuss the convergence and convergence rate in the nonergodic sense of the proposed algorithm under mild conditions and verify the efficiency of our algorithm by the resource allocation problem.

**Acknowledgments.** The authors thank Dr. Wenxing Zhang from University of Electronic Science and Technology of China for the discussion on the numerical
implementation of the algorithms. The authors also would like to thank the anonymous referees and the associate editor for their thoughtful and insightful comments, which improved the paper greatly.

REFERENCES

[1] 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*, 3 (2010), 1–122.

[2] X. J. Cai, D. R. Han and X. M. Yuan, On the convergence of the direct extension of ADMM for three-block separable convex minimization models with one strongly convex function, *Computational Optimization and Applications*, 66 (2017), 39–73.

[3] C. H. Chen, B. S. He, Y. Y. Ye and X. M. Yuan, The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent, *Mathematical Programming Ser. A*, 155 (2016), 57–79.

[4] C. H. Chen, Y. Shen and Y. F. You, On the convergence analysis of the alternating direction method of multipliers with three blocks. *Abstract and Applied Analysis*, (2013), Article ID 183961, 7 pages.

[5] E. Corman and X. M. Yuan, A generalized proximal point algorithm and its convergence rate, *SIAM Journal on Optimization*, 24 (2014), 1614–1638.

[6] Y. H. Dai, D. R. Han, X. M. Yuan and W. X. Zhang, A sequential updating scheme of the Lagrange multiplier for separable convex programming, *Mathematics of Computation*, 86 (2017), 315–343.

[7] W. Deng, M.-J. Lai, Z. M. Peng and W. T. Yin, Parallel multi-block ADMM with $o(1/k)$ convergence, *Journal of Scientific Computing*, 71 (2017), 712–736.

[8] J. Douglas and H. H. Rachford, On the numerical solution of the heat conduction problem in 2 and 3 space variables, *Transactions of the American Mathematical Society*, 82 (1956), 421–439.

[9] D. Gabay, Applications of the method of multipliers to variational inequalities, in *Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems* (eds. M. Fortin and R. Glowinski), North Holland, Amsterdam, The Netherlands, (1983), 299–331.

[10] R. Glowinski, T. Kärkkäinen and K. Majava, On the convergence of operator-splitting methods, in *Numerical Methods for Scientific Computing, Variational Problems and Applications* (eds. Y. Kuznetsov, P. Neittanmaki and O. Pironneau), Barcelona, (2003).

[11] R. Glowinski and A. Marrocco, Approximation par éléments finis d’ordre un et résolution par pénalisation-dualité d’une classe de problèmes non linéaires, *R.A.I.R.O.*, 9 (1975), 41–76.

[12] D. R. Han, X. M Yuan, W. X. Zhang and X. J. Cai, An ADM-based splitting method for separable convex programming, *Computational Optimization and Applications*, 54 (2013), 343–369.

[13] B. S. He, H. Liu, J. W. Lu and X. M. Yuan, Application of the strictly contractive Peaceman-Rachford splitting method to multi-block separable convex programming, manuscript, in *Splitting Methods in Communication and Imaging, Science, and Engineering* (eds. R. Glowinski, S. Osher and W. Yin), Springer, Switzerland, (2016), 195–235.

[14] B. S. He, H. Liu, Z. R. Wang and X. M. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, *SIAM Journal on Optimization*, 24 (2014), 1101–1140.

[15] B. S. He, M. Tao and X. M. Yuan, A splitting method for separable convex programming, *IMA Journal of Numerical Analysis*, 35 (2015), 394–426.

[16] B. S. He, M. Tao and X. M. Yuan, Alternating direction method with Gaussian back substitution for separable convex programming, *SIAM Journal on Optimization*, 22 (2012), 313–340.

[17] B. S. He and X. M. Yuan, On the $O(1/n)$ convergence rate of Douglas-Rachford alternating direction method, *SIAM Journal on Numerical Analysis*, 50 (2012), 700–709.

[18] B. S. He and X. M. Yuan, On nonergodic convergence rate of Douglas-Rachford alternating direction method of multipliers, *Numerische Mathematik*, 30 (2015), 567–577.

[19] M. Li, D. F. Sun and K.-C. Toh, A convergent 3-block semi-proximal ADMM for convex minimization problems with one strongly convex block, *Asia Pacific Journal of Operational Research*, 32 (2015), 1550024, 19 pp.
[20] X. D. Li, D. F. Sun and K.-C. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions, *Mathematical Programming Ser. A*, 155 (2016), 333–373.

[21] T. Y. Lin, S. Q. Ma and S. Z. Zhang, On the global linear convergence of the ADMM with multi-block variables, *SIAM Journal on Optimization*, 25 (2015), 1478–1497.

[22] 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*, 3 (2015), 251–274.

[23] P. L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM Journal on Numerical Analysis*, 16 (1979), 964–979.

[24] S. Q. Ma, Alternating proximal gradient method for convex minimization, *Journal of Scientific Computing*, 68 (2016), 546–572.

[25] Y. E. Nesterov, Gradient methods for minimizing composite objective function, *Mathematical Programming Ser. B*, 140 (2013), 125–161.

[26] M. Patriksson, A survey on the continuous nonlinear resource allocation Problem, *European Journal of Operations Research*, 185 (2008), 1–46.

[27] D. H. Peaceman and H. H. Rachford, The numerical solution of parabolic elliptic differential equations, *SIAM Journal on Applied Mathematics*, 3 (1955), 28–41.

[28] Y. G. Peng, A. Ganesh, J. Wright, W. L. Xu and Y. Ma, Robust alignment by sparse and low-rank decomposition for linearly correlated images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (2012), 2233–2246.

[29] M. Tao and X. M. Yuan, Recovering low-rank and sparse components of matrices from incomplete and noisy observations, *SIAM Journal on Optimization*, 21 (2011), 57–81.

[30] H. Uzawa, Market mechanisms and mathematical programming, *Econometrica*, 28 (1960), 872–881.

Received June 2016; revised December 2016.

*E-mail address:* dg1415005@smail.nju.edu.cn

*E-mail address:* liminnju@yahoo.com