Asia-Pacific Journal of Operational Research
Vol. 32, No. 3 (2015) 1550024 (19 pages)

DOI: 10.1142/S0217595915500244



# A Convergent 3-Block Semi-Proximal ADMM for Convex Minimization Problems with One Strongly Convex Block

Min Li

*School of Economics and Management*
*Southeast University*
*Nanjing 210096, P. R. China*
*limin@seu.edu.cn*

Defeng Sun

*Department of Mathematics and Risk Management Institute*
*National University of Singapore*
*10 Lower Kent Ridge Road, Singapore*
*matsundf@nus.edu.sg*

Kim-Chuan Toh

*Department of Mathematics*
*National University of Singapore*
*10 Lower Kent Ridge Road, Singapore*
*mattohkc@nus.edu.sg*

Received 29 October 2014
Revised 23 November 2014
Accepted 25 November 2014
Published 19 January 2015

In this paper, we present a semi-proximal alternating direction method of multipliers (sPADMM) for solving 3-block separable convex minimization problems with the second block in the objective being a strongly convex function and one coupled linear equation constraint. By choosing the semi-proximal terms properly, we establish the global convergence of the proposed sPADMM for the step-length $\tau \in (0, (1+\sqrt{5})/2)$ and the penalty parameter $\sigma \in (0, +\infty)$. In particular, if $\sigma > 0$ is smaller than a certain threshold and the first and third linear operators in the linear equation constraint are injective, then all the three added semi-proximal terms can be dropped and consequently, the convergent 3-block sPADMM reduces to the directly extended 3-block ADMM with $\tau \in (0, (1+\sqrt{5})/2)$.

*Keywords*: Convex minimization problems; alternating direction method of multipliers; semi-proximal; strongly convex.

AMS Subject Classifications: 90C25, 90C33, 65K05

## 1. Introduction

We consider the following separable convex minimization problem whose objective function is the sum of three functions without coupled variables:

$$\min_{x_1,x_2,x_3} \{\theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3) \mid A_1^* x_1 + A_2^* x_2 + A_3^* x_3 = c, \\ x_i \in \mathcal{X}_i, i = 1, 2, 3\}, \tag{1}$$

where $\mathcal{X}_i$ $(i = 1, 2, 3)$ and $\mathcal{Z}$ are real finite-dimensional Euclidean spaces each equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$, $\theta_i : \mathcal{X}_i \to (-\infty, +\infty]$ $(i = 1, 2, 3)$ are closed proper convex functions, $A_i^* : \mathcal{X}_i \to \mathcal{Z}$ is the adjoint of the linear operator $A_i : \mathcal{Z} \to \mathcal{X}_i, i = 1, 2, 3$, and $c \in \mathcal{Z}$. Since $\theta_i, i = 1, 2, 3$, are closed proper convex functions, there exist self-adjoint and positive semi-definite operators $\Sigma_i, i = 1, 2, 3$, such that

$$\langle \hat{x}_i - x_i, \hat{w}_i - w_i \rangle \geq \langle \hat{x}_i - x_i, \Sigma_i(\hat{x}_i - x_i) \rangle, \quad \forall\, \hat{x}_i, x_i \in \mathrm{dom}(\theta_i), \\ \hat{w}_i \in \partial\theta_i(\hat{x}_i), \;\; w_i \in \partial\theta_i(x_i), \tag{2}$$

where $\partial\theta_i$ is the sub-differential mapping of $\theta_i, i = 1, 2, 3$. The solution set of problem (1) is assumed to be nonempty throughout our discussions in this paper.

Let $\sigma > 0$ be a given penalty parameter and $z \in \mathcal{Z}$ be the Lagrange multiplier associated with the linear equality constraint in problem (1). For any $(x_1, x_2, x_3) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$, write $x \equiv (x_1, x_2, x_3), \theta(x) \equiv \theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3)$ and $A^*x \equiv A_1^* x_1 + A_2^* x_2 + A_3^* x_3$. Then the augmented Lagrangian function for problem (1) is defined by

$$\mathcal{L}_\sigma(x_1, x_2, x_3; z) := \theta(x) + \langle z, A^*x - c \rangle + \frac{\sigma}{2}\|A^*x - c\|^2 \tag{3}$$

for any $(x_1, x_2, x_3, z) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{Z}$. The direct extension of the classical alternating direction method of multipliers (ADMM) for solving problem (1) consists of the following iterations for $k = 0, 1, \ldots$

$$\begin{cases} x_1^{k+1} := \underset{x_1 \in \mathcal{X}_1}{\operatorname{argmin}}\{\mathcal{L}_\sigma(x_1, x_2^k, x_3^k; z^k)\}, \\ x_2^{k+1} := \underset{x_2 \in \mathcal{X}_2}{\operatorname{argmin}}\{\mathcal{L}_\sigma(x_1^{k+1}, x_2, x_3^k; z^k)\}, \\ x_3^{k+1} := \underset{x_3 \in \mathcal{X}_3}{\operatorname{argmin}}\{\mathcal{L}_\sigma(x_1^{k+1}, x_2^{k+1}, x_3; z^k)\}, \\ z^{k+1} := z^k + \tau\sigma(A^*x^{k+1} - c), \end{cases} \tag{4}$$

where $\tau > 0$ is the step-length. Different from the 2-block ADMM whose convergence has been established for a long time (Glowinski and Marrocco, 1975; Gabay and Mercier, 1976; Glowinski, 1980; Fortin and Glowinski, 1983; Gabay, 1983; Eckstein and Bertsekas, 1992), the 3-block ADMM may not converge in general, which was demonstrated by Chen *et al.* (2014) using counterexamples. Nevertheless, if all the functions $\theta_i, i = 1, 2, 3$, are strongly convex, Han and Yuan (2012) proved the global convergence of the 3-block ADMM scheme (4) with $\tau = 1$ (Han and Yuan actually considered the general $m$-block case for any $m \geq 3$. Here and below we focus on the 3-block case only) under the condition that

$$\Sigma_i = \mu_i I \succ 0, \quad i = 1, 2, 3, \quad 0 < \sigma < \min_{i=1,2,3} \left\{ \frac{\mu_i}{3\lambda_{\max}(A_i A_i^*)} \right\},$$

where $\lambda_{\max}(S)$ is the largest eigenvalue of a given self-adjoint linear operator $S$. Hong and Luo (2013) proposed to adopt a small step-length $\tau$ when updating the Lagrange multiplier $z^{k+1}$ in (4). Chen *et al.* (2013) proposed the following sufficient condition

$$A_1^* \text{ is injective}, \quad \Sigma_i = \mu_i I \succ 0, \quad i = 2, 3 \quad \text{and}$$

$$0 < \sigma < \frac{\mu_2}{\lambda_{\max}(A_2 A_2^*)}, \quad \sigma \leq \frac{\mu_3}{\lambda_{\max}(A_3 A_3^*)}$$

for the global convergence of the directly extended 3-block ADMM with $\tau = 1$ for solving problem (1). Closely related to the work of Chen *et al.* (2013), Lin *et al.* (2014a) provided an analysis on the iteration complexity for the same method under the condition

$$\Sigma_i = \mu_i I \succ 0, \quad i = 2, 3 \quad \text{and} \quad 0 < \sigma \leq \min \left\{ \frac{\mu_2}{2\lambda_{\max}(A_2 A_2^*)}, \frac{\mu_3}{2\lambda_{\max}(A_3 A_3^*)} \right\}.$$

In Lin *et al.* (2014b), under additional assumptions including some smoothness conditions, the same group of authors further proved the global linear convergence of the mentioned method.

The purpose of this work is to extend the 2-block semi-proximal ADMM (sPADMM) studied in Fazel *et al.* (2013) to deal with problem (1) by only assuming $\theta_2$ to be strongly convex, i.e., $\Sigma_2 \succ 0$. Note that the sPADMM with $\tau > 1$ often works better in practice than its counterpart with $\tau \leq 1$. So it is desirable to establish the convergence of the proposed sPADMM that allows $\tau$ to stay in the larger region $(0, (1 + \sqrt{5})/2)$.

One of our motivating examples is the following convex quadratic conic programming

$$\begin{aligned} \min \quad & \frac{1}{2}\langle X, \mathcal{Q}X\rangle + \langle C, X\rangle \\ \text{s.t.} \quad & \mathcal{A}X \geq b, \quad X \in \mathcal{K}, \end{aligned} \tag{5}$$

where $\mathcal{K}$ is a nonempty closed convex cone in a finite-dimensional real Euclidean space $\mathcal{X}$ endowed with an inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$, $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ is a self-adjoint and positive semi-definite linear operator, $\mathcal{A} : \mathcal{X} \to \Re^m$ is a linear map, $C \in \mathcal{X}$ and $b \in \Re^m$ are given data. The dual of problem (5) takes the form of

$$\begin{aligned} \max \quad & -\frac{1}{2}\langle X', \mathcal{Q}X'\rangle + \langle b, y\rangle \\ \text{s.t.} \quad & \mathcal{A}^*y - \mathcal{Q}X' + S = C, \quad y \geq 0, \quad S \in \mathcal{K}^*, \end{aligned} \tag{6}$$

where $\mathcal{K}^* := \{v \in \mathcal{X} : \langle v, w\rangle \geq 0, \forall\, w \in \mathcal{K}\}$ is the dual cone of $\mathcal{K}$. Since $\mathcal{Q}$ is self-adjoint and positive semi-definite, $\mathcal{Q}$ can be decomposed as $\mathcal{Q} = \mathcal{L}^*\mathcal{L}$ for some linear map $\mathcal{L}$. By introducing a new variable $\Xi = -\mathcal{L}X'$, we can re-write problem (6) equivalently as

$$\begin{aligned} \min \quad & \delta_{\Re^m_+}(y) - \langle b, y\rangle + \frac{1}{2}\|\Xi\|^2 + \delta_{\mathcal{K}^*}(S) \\ \text{s.t.} \quad & \mathcal{A}^* y + \mathcal{L}^*\Xi + S = C, \end{aligned} \tag{7}$$

where $\delta_{\Re^m_+}(\cdot)$ and $\delta_{\mathcal{K}^*}(\cdot)$ are the indicator functions of $\Re^m_+$ and $\mathcal{K}^*$, respectively. As one can see, problem (7) has only one strongly convex block, i.e., the block with respect to $\Xi$. Consequently, the results in the aforementioned papers for the convergence analysis of the directly extended 3-block ADMM applied to solving problem (7) are no longer valid. We shall show in the next section that our proposed 3-block sPADMM can exactly solve this kind of problems. When $\mathcal{K} = \mathcal{S}^n_+$, the cone of symmetric and positive semi-definite matrices in the space $\mathcal{S}^n$ of $n \times n$ symmetric matrices, problem (7) is a convex quadratic semi-definite programming problem that has been extensively studied both theoretically and numerically in Li *et al.* (2014); Nie and Yuan (2001); Qi (2009); Qi and Sun (2011); Sun (2006); Sun *et al.* (2008); Sun and Zhang (2010); Toh (2008); Toh *et al.* (2007); Zhao (2009), to name only a few.

The remaining parts of this paper are organized as follows. In Sec. 2, we first present our 3-block sPADMM and then provide the main convergence results. We give some concluding remarks in Sec. 3.

**Notation.**

- The effective domain of a function $f : \mathcal{X} \to (-\infty, +\infty]$ is defined as $\mathrm{dom}(f) := \{x \in \mathcal{X} \mid f(x) < +\infty\}$. The set of all relative interior points of a given nonempty convex set $\mathcal{C}$ is denoted by $\mathrm{ri}(\mathcal{C})$.
- For convenience, for any given $x$, we use $\|x\|^2_G$ to denote $\langle x, Gx\rangle$ if $G$ is a self-adjoint linear operator in a given finite-dimensional Euclidean space $\mathcal{X}$. If $\Sigma : \mathcal{X} \to \mathcal{X}$ is a self-adjoint and positive semi-definite linear operator, we use $\Sigma^{\frac{1}{2}}$ to denote the unique self-adjoint and positive semi-definite square root of $\Sigma$.
- Denote

$$x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad u := \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, \quad A := \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}, \quad B := \begin{pmatrix} A_2 \\ A_3 \end{pmatrix}.$$

- Let $\alpha \in (0, 1]$ be given. Denote

$$M := \begin{pmatrix} (1-\alpha)\Sigma_2 + T_2 & 0 \\ 0 & \Sigma_3 + T_3 \end{pmatrix} + \sigma B B^*, \tag{8}$$

$$H := \begin{pmatrix} \frac{5(1-\alpha)}{2}\Sigma_2 + T_2 & 0 \\ 0 & \frac{5}{2}\Sigma_3 + T_3 - \frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*) \end{pmatrix} + \min(\tau, 1 + \tau - \tau^2)\sigma B B^*. \tag{9}$$

## 2. A 3-Block SPADMM

Based on our previous introduction and motivation, we propose our 3-block sPADMM for solving problem (1) in the following:

**Algorithm sPADMM: A 3-block sPADMM for solving problem (1).** Let $\sigma \in (0, +\infty)$ and $\tau \in (0, +\infty)$ be given parameters. Let $T_i$, $i = 1, 2, 3$, be given self-adjoint and positive semi-definite linear operators defined on $\mathcal{X}_i$, $i = 1, 2, 3$, respectively. Choose $(x_1^0, x_2^0, x_3^0, z^0) \in \mathrm{dom}(\theta_1) \times \mathrm{dom}(\theta_2) \times \mathrm{dom}(\theta_3) \times \mathcal{Z}$ and set $k = 0$.

**Step 1.** Compute

$$\begin{cases} x_1^{k+1} := \operatorname*{argmin}_{x_1 \in \mathcal{X}_1} \left\{ \mathcal{L}_\sigma(x_1, x_2^k, x_3^k; z^k) + \frac{1}{2}\|x_1 - x_1^k\|_{T_1}^2 \right\}, \\ x_2^{k+1} := \operatorname*{argmin}_{x_2 \in \mathcal{X}_2} \left\{ \mathcal{L}_\sigma(x_1^{k+1}, x_2, x_3^k; z^k) + \frac{1}{2}\|x_2 - x_2^k\|_{T_2}^2 \right\}, \\ x_3^{k+1} := \operatorname*{argmin}_{x_3 \in \mathcal{X}_3} \left\{ \mathcal{L}_\sigma(x_1^{k+1}, x_2^{k+1}, x_3; z^k) + \frac{1}{2}\|x_3 - x_3^k\|_{T_3}^2 \right\}, \\ z^{k+1} := z^k + \tau\sigma(A^* x^{k+1} - c). \end{cases} \tag{10}$$

**Step 2.** If a termination criterion is not met, set $k := k+1$ and then go to Step 1.

In order to analyze the convergence properties of Algorithm sPADMM, we make the following assumptions.

**Assumption 2.1.** The convex function $\theta_2$ satisfies (2) with $\Sigma_2 \succ 0$.

**Assumption 2.2.** The self-adjoint and positive semi-definite operators $T_i, i = 1, 2, 3$, are chosen such that the sequence $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ generated by Algorithm sPADMM is well defined.

**Assumption 2.3.** There exists $x' = (x_1', x_2', x_3') \in ri(\mathrm{dom}(\theta_1) \times \mathrm{dom}(\theta_2) \times \mathrm{dom}(\theta_3)) \cap P$, where

$$P := \{x := (x_1, x_2, x_3) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \mid A^* x = c\}.$$

Under Assumption 2.3, it follows from Corollaries 28.2.2 and 28.3.1 of Rockafellar (1970) that $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$ is an optimal solution to problem (1) if and only if there exists a Lagrange multiplier $\bar{z} \in \mathcal{Z}$ such that

$$-A_i \bar{z} \in \partial\theta_i(\bar{x}_i), \quad i = 1, 2, 3 \quad \text{and} \quad A^*\bar{x} - c = 0. \tag{11}$$

Moreover, any $\bar{z} \in \mathcal{Z}$ satisfying (11) is an optimal solution to the dual of problem (1).

Let $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$ and $\bar{z} \in \mathcal{Z}$ satisfy (11). For the sake of convenience, define for $(x_1, u, z) := (x_1, (x_2, x_3), z) \in \mathcal{X}_1 \times (\mathcal{X}_2 \times \mathcal{X}_3) \times \mathcal{Z}, \alpha \in (0, 1]$ and $k = 0, 1, \ldots,$ the following quantities

$$\phi_k(x_1, u, z) := (\sigma\tau)^{-1}\|z^k - z\|^2 + \|x_1^k - x_1\|_{\Sigma_1 + T_1}^2 + \|u^k - u\|_M^2$$

and

$$\begin{cases} x_{ie}^k := x_i^k - \bar{x}_i, \quad i = 1,2,3, \quad u_e^k := u^k - \bar{u}, \quad z_e^k := z^k - \bar{z}, \\ \Delta x_i^k := x_i^{k+1} - x_i^k, \quad i = 1,2,3, \quad \Delta u^k := u^{k+1} - u^k, \quad \Delta z^k := z^{k+1} - z^k, \\ \overline{\phi}_k := \phi_k(\bar{x}_1, \bar{u}, \bar{z}) = (\sigma\tau)^{-1}\|z_e^k\|^2 + \|x_{1e}^k\|^2_{\Sigma_1 + T_1} + \|u_e^k\|^2_M, \\ \xi_{k+1} := \|\Delta x_2^k\|^2_{T_2} + \|\Delta x_3^k\|^2_{T_3 + \frac{\sigma^2}{\alpha}(A_2 A_3^*)^* \Sigma_2^{-1}(A_2 A_3^*)}, \\ s_{k+1} := \|\Delta x_1^k\|^2_{\frac{1}{2}\Sigma_1 + T_1} + \|\Delta x_2^k\|^2_{\frac{1-\alpha}{2}\Sigma_2 + T_2} \\ \qquad\qquad + \|\Delta x_3^k\|^2_{\frac{1}{2}\Sigma_3 + T_3 - \frac{\sigma^2}{2\alpha}(A_2 A_3^*)^* \Sigma_2^{-1}(A_2 A_3^*)} \\ \qquad\qquad + \sigma\|A_1^* x_1^{k+1} + B^* u^k - c\|^2, \\ t_{k+1} := \|\Delta x_1^k\|^2_{\frac{1}{2}\Sigma_1 + T_1} + \|\Delta u^k\|^2_H, \\ r^k := A^* x^k - c. \end{cases} \tag{12}$$

To prove the convergence of Algorithm sPADMM for solving problem (1), we first present some useful lemmas.

**Lemma 2.1.** *Assume that Assumptions* 2.1–2.3 *hold. Let* $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ *be generated by Algorithm sPADMM. Then, for any* $\tau \in (0, +\infty)$ *and integer* $k \geq 0$, *we have*

$$\overline{\phi}_k - \overline{\phi}_{k+1} \geq (1-\tau)\sigma\|r^{k+1}\|^2 + s_{k+1}, \tag{13}$$

*where* $\overline{\phi}_k, s_{k+1}$ *and* $r^{k+1}$ *are defined as in* (12).

**Proof.** The sequence $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ is well defined under Assumption 2.2. Notice that the iteration scheme (10) of Algorithm sPADMM can be re-written as for $k = 0, 1, \ldots$ that

$$\begin{cases} -A_1\left[z^k + \sigma\left(A_1^* x_1^{k+1} + \sum_{j=2}^{3} A_j^* x_j^k - c\right)\right] - T_1(x_1^{k+1} - x_1^k) \in \partial\theta_1(x_1^{k+1}), \\ -A_2\left[z^k + \sigma\left(\sum_{j=1}^{2} A_j^* x_j^{k+1} + A_3^* x_3^k - c\right)\right] - T_2(x_2^{k+1} - x_2^k) \in \partial\theta_2(x_2^{k+1}), \\ -A_3[z^k + \sigma(A^* x^{k+1} - c)] - T_3(x_3^{k+1} - x_3^k) \in \partial\theta_3(x_3^{k+1}), \\ z^{k+1} := z^k + \tau\sigma(A^* x^{k+1} - c). \end{cases} \tag{14}$$

Combining (2) with (11) and (14), and using the definitions of $x_{ie}^{k+1}$ and $\Delta x_i^k$, for $i = 1, 2, 3$, we have

$$\left\langle x_{ie}^{k+1}, A_i\bar{z} - A_i z^k - \sigma A_i \left( \sum_{j=1}^{i} A_j^* x_j^{k+1} + \sum_{j=i+1}^{3} A_j^* x_j^k - c \right) - T_i \Delta x_i^k \right\rangle$$
$$\geq \|x_{ie}^{k+1}\|_{\Sigma_i}^2. \tag{15}$$

For any vectors $a, b, d$ in the same Euclidean vector space and any self-adjoint linear operator $G$, we have the identity

$$\langle a - b, G(d - a)\rangle = \frac{1}{2}(\|d - b\|_G^2 - \|a - b\|_G^2 - \|a - d\|_G^2).$$

Taking $a = x_i^{k+1}, b = \bar{x}_i, d = x_i^k$ and $G = T_i$ in the above identity, and using the definitions of $x_{ie}^{k+1}$ and $\Delta x_i^k$, we get

$$\langle x_{ie}^{k+1}, -T_i \Delta x_i^k\rangle = \frac{1}{2}(\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2 - \|\Delta x_i^k\|_{T_i}^2), \quad i = 1, 2, 3. \tag{16}$$

Let

$$\tilde{z}^{k+1} = z^k + \sigma(A^* x^{k+1} - c) = z^k + \sigma(A_1^* x_1^{k+1} + B^* u^{k+1} - c). \tag{17}$$

Substituting (16) and (17) into (15) and using the definition of $\Delta x_j^k$, for $i = 1, 2$, we have

$$\left\langle x_{ie}^{k+1}, A_i\bar{z} - A_i\tilde{z}^{k+1} + \sigma A_i \sum_{j=i+1}^{3} A_j^* \Delta x_j^k \right\rangle + \frac{1}{2}(\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2)$$
$$\geq \frac{1}{2}\|\Delta x_i^k\|_{T_i}^2 + \|x_{ie}^{k+1}\|_{\Sigma_i}^2 \tag{18}$$

and

$$\langle x_{3e}^{k+1}, A_3\bar{z} - A_3\tilde{z}^{k+1}\rangle + \frac{1}{2}(\|x_{3e}^k\|_{T_3}^2 - \|x_{3e}^{k+1}\|_{T_3}^2) \geq \frac{1}{2}\|\Delta x_3^k\|_{T_3}^2 + \|x_{3e}^{k+1}\|_{\Sigma_3}^2. \tag{19}$$

Adding (18) for $i = 1, 2$ to (19), we get

$$\sum_{i=1}^{3} \langle x_{ie}^{k+1}, A_i\bar{z} - A_i\tilde{z}^{k+1}\rangle + \sigma \left\langle x_{1e}^{k+1}, A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \right\rangle + \sigma\langle x_{2e}^{k+1}, A_2 A_3^* \Delta x_3^k\rangle$$
$$+ \frac{1}{2}\sum_{i=1}^{3}(\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2)$$
$$\geq \frac{1}{2}\sum_{i=1}^{3} \|\Delta x_i^k\|_{T_i}^2 + \sum_{i=1}^{3} \|x_{ie}^{k+1}\|_{\Sigma_i}^2. \tag{20}$$

By simple manipulations and using $A_1^* x_{1e}^{k+1} = A_1^* x_1^{k+1} - A_1^* \bar{x}_1 = B^* \bar{u} + (A_1^* x_1^{k+1} - c)$, we get

$$\begin{aligned} &\sigma \left\langle x_{1e}^{k+1}, A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \right\rangle \\ &\quad = \sigma \langle -x_{1e}^{k+1}, -A_1 B^* \Delta u^k \rangle \\ &\quad = \sigma \langle -A_1^* x_{1e}^{k+1}, B^* u^k - B^* u^{k+1} \rangle \\ &\quad = \sigma \langle (-B^* \bar{u}) - (A_1^* x_1^{k+1} - c), (-B^* u^{k+1}) - (-B^* u^k) \rangle. \end{aligned} \tag{21}$$

For any vectors $a, b, d, e$ in the same Euclidean vector space, we have the identity

$$\langle a - b, d - e \rangle = \frac{1}{2}(\|a - e\|^2 - \|a - d\|^2) + \frac{1}{2}(\|b - d\|^2 - \|b - e\|^2). \tag{22}$$

In the above identity, by taking $a = -B^* \bar{u}, b = A_1^* x_1^{k+1} - c, d = -B^* u^{k+1}$ and $e = -B^* u^k$, and applying it to the right-hand side of (21), we obtain from the definitions of $u_e^k$ and $\tilde{z}^{k+1}$ that

$$\begin{aligned} \sigma \left\langle x_{1e}^{k+1}, A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \right\rangle &= \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2) \\ &\quad + \frac{\sigma}{2}(\|A_1^* x_1^{k+1} + B^* u^{k+1} - c\|^2 \\ &\quad - \|A_1^* x_1^{k+1} + B^* u^k - c\|^2) \\ &= \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2) \\ &\quad + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2 - \frac{\sigma}{2}\|A_1^* x_1^{k+1} + B^* u^k - c\|^2. \end{aligned} \tag{23}$$

Using the Cauchy–Schwarz inequality, for the parameter $\alpha \in (0, 1]$, we get

$$\begin{aligned} \sigma \langle x_{2e}^{k+1}, A_2 A_3^* \Delta x_3^k \rangle &= 2 \left\langle (\alpha \Sigma_2)^{\frac{1}{2}} x_{2e}^{k+1}, \frac{\sigma}{2} (\alpha \Sigma_2)^{-\frac{1}{2}} A_2 A_3^* \Delta x_3^k \right\rangle \\ &\leq \alpha \|x_{2e}^{k+1}\|_{\Sigma_2}^2 + \frac{\sigma^2}{4\alpha} \|\Delta x_3^k\|_{(A_2 A_3^*)^* \Sigma_2^{-1} (A_2 A_3^*)}^2. \end{aligned} \tag{24}$$

It follows from (17) that

$$\begin{aligned} \sum_{i=1}^{3} \langle x_{ie}^{k+1}, A_i \bar{z} - A_i \tilde{z}^{k+1} \rangle &= \left\langle \bar{z} - \tilde{z}^{k+1}, \sum_{i=1}^{3} A_i^* x_{ie}^{k+1} \right\rangle \\ &= \frac{1}{\sigma} \langle \bar{z} - \tilde{z}^{k+1}, \tilde{z}^{k+1} - z^k \rangle. \end{aligned} \tag{25}$$

Substituting (23)–(25) into (20), we obtain

$$\frac{1}{\sigma}\langle \bar{z} - \tilde{z}^{k+1}, \tilde{z}^{k+1} - z^k\rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2 + \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2) + \frac{1}{2}\sum_{i=1}^{3}(\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2)$$
$$\geq \frac{\sigma}{2}\|A_1^* x_1^{k+1} + B^* u^k - c\|^2 + \frac{1}{2}\sum_{i=1}^{3}\|\Delta x_i^k\|_{T_i}^2 + \sum_{i=1,i\neq 2}^{3}\|x_{ie}^{k+1}\|_{\Sigma_i}^2 + (1-\alpha)\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \frac{\sigma^2}{4\alpha}\|\Delta x_3^k\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)}. \tag{26}$$

From the elementary inequality $\|a\|^2 + \|b\|^2 \geq \|a-b\|^2/2$ and $x_{ie}^{k+1} - x_{ie}^k = \Delta x_i^k$, it follows that

$$\sum_{i=1,i\neq 2}^{3}\|x_{ie}^{k+1}\|_{\Sigma_i}^2 + (1-\alpha)\|x_{2e}^{k+1}\|_{\Sigma_2}^2$$
$$= \frac{1}{2}\sum_{i=1,i\neq 2}^{3}(\|x_{ie}^{k+1}\|_{\Sigma_i}^2 + \|x_{ie}^k\|_{\Sigma_i}^2) + \frac{1}{2}\sum_{i=1,i\neq 2}^{3}(\|x_{ie}^{k+1}\|_{\Sigma_i}^2 - \|x_{ie}^k\|_{\Sigma_i}^2) + \frac{1-\alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 + \|x_{2e}^k\|_{\Sigma_2}^2) + \frac{1-\alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \|x_{2e}^k\|_{\Sigma_2}^2)$$
$$\geq \frac{1}{4}\sum_{i=1,i\neq 2}^{3}\|\Delta x_i^k\|_{\Sigma_i}^2 + \frac{1}{2}\sum_{i=1,i\neq 2}^{3}(\|x_{ie}^{k+1}\|_{\Sigma_i}^2 - \|x_{ie}^k\|_{\Sigma_i}^2) + \frac{1-\alpha}{4}\|\Delta x_2^k\|_{\Sigma_2}^2 + \frac{1-\alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \|x_{2e}^k\|_{\Sigma_2}^2). \tag{27}$$

By simple manipulations and using the definition of $z_e^k$, we get

$$\frac{1}{\sigma}\langle \bar{z} - \tilde{z}^{k+1}, \tilde{z}^{k+1} - z^k\rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2$$
$$= \frac{1}{\sigma}\langle \bar{z} - z^k, \tilde{z}^{k+1} - z^k\rangle + \frac{1}{\sigma}\langle z^k - \tilde{z}^{k+1}, \tilde{z}^{k+1} - z^k\rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2$$
$$= \frac{1}{\sigma}\langle -z_e^k, \tilde{z}^{k+1} - z^k\rangle - \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2$$
$$= \frac{1}{2\sigma\tau}(\|z_e^k\|^2 - \|z_e^k + \tau(\tilde{z}^{k+1} - z^k)\|^2) + \frac{\tau - 1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2. \tag{28}$$

By using (14), (17) and the definitions of $z_e^k$ and $r^{k+1}$, we have

$$z_e^{k+1} = z_e^k + \tau(\tilde{z}^{k+1} - z^k) \quad \text{and} \quad z^k - \tilde{z}^{k+1} = -\sigma r^{k+1},$$

which, together with (28), imply

$$\frac{1}{\sigma}\langle \bar{z} - \tilde{z}^{k+1}, \tilde{z}^{k+1} - z^k \rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2$$

$$= \frac{1}{2\sigma\tau}(\|z_e^k\|^2 - \|z_e^{k+1}\|^2) + \frac{(\tau-1)\sigma}{2}\|r^{k+1}\|^2. \tag{29}$$

Substituting (27) and (29) into (26), and using the definitions of $\overline{\phi}_k$, $s_{k+1}$ and $r^{k+1}$, we get the assertion (13). The proof is complete. □

**Lemma 2.2.** *Assume that Assumptions* 2.1 *and* 2.2 *hold. Let* $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ *be generated by Algorithm sPADMM. Then, for any* $\tau \in (0, +\infty)$ *and integer* $k \geq 1$, *we have*

$$-\sigma\langle B^*\Delta u^k, r^{k+1}\rangle \geq -(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle$$

$$+ \frac{1}{2}\sum_{i=2}^{3}(\|\Delta x_i^k\|_{\mathcal{T}_i+2\Sigma_i}^2 - \|\Delta x_i^{k-1}\|_{\mathcal{T}_i}^2)$$

$$+ \sigma\langle A_2^*\Delta x_2^k, A_3^*(\Delta x_3^{k-1} - \Delta x_3^k)\rangle, \tag{30}$$

*where* $\Delta u^k, \Delta x_i^k$ $(i = 2, 3)$ *and* $r^{k+1}$ *are defined as in* (12).

**Proof.** Let

$$v^{k+1} := z^k + \sigma\left(\sum_{j=1}^{2} A_j^* x_j^{k+1} + A_3^* x_3^k - c\right).$$

By using (14) and the definition of $\Delta x_2^k$, we have

$$-A_2 v^{k+1} - \mathcal{T}_2\Delta x_2^k \in \partial\theta_2(x_2^{k+1}) \quad \text{and} \quad -A_2 v^k - \mathcal{T}_2\Delta x_2^{k-1} \in \partial\theta_2(x_2^k).$$

Thus, we obtain from (2) that

$$\langle \Delta x_2^k, (A_2 v^k + \mathcal{T}_2\Delta x_2^{k-1}) - (A_2 v^{k+1} + \mathcal{T}_2\Delta x_2^k)\rangle \geq \|\Delta x_2^k\|_{\Sigma_2}^2.$$

By using the Cauchy–Schwarz inequality, we obtain

$$\langle \Delta x_2^k, \mathcal{T}_2(\Delta x_2^k - \Delta x_2^{k-1})\rangle = \|\Delta x_2^k\|_{\mathcal{T}_2}^2 - \langle \Delta x_2^k, \mathcal{T}_2\Delta x_2^{k-1}\rangle$$

$$\geq \frac{1}{2}\|\Delta x_2^k\|_{\mathcal{T}_2}^2 - \frac{1}{2}\|\Delta x_2^{k-1}\|_{\mathcal{T}_2}^2.$$

Adding up the above two inequalities, we get

$$\langle A_2^*\Delta x_2^k, v^k - v^{k+1}\rangle \geq \frac{1}{2}\|\Delta x_2^k\|_{\mathcal{T}_2+2\Sigma_2}^2 - \frac{1}{2}\|\Delta x_2^{k-1}\|_{\mathcal{T}_2}^2. \tag{31}$$

Using $z^{k-1} - z^k = -\tau\sigma r^k$ and the definitions of $v^k$ and $r^k$, we have

$$v^k - v^{k+1} = (1-\tau)\sigma r^k - \sigma r^{k+1} - \sigma A_3^*(\Delta x_3^{k-1} - \Delta x_3^k).$$

Substituting the above equation into (31), we get

$$\sigma\langle -A_2^*\Delta x_2^k, r^{k+1}\rangle \geq -(1-\tau)\sigma\langle A_2^*\Delta x_2^k, r^k\rangle + \sigma\langle A_2^*\Delta x_2^k, A_3^*(\Delta x_3^{k-1} - \Delta x_3^k)\rangle + \frac{1}{2}(\|\Delta x_2^k\|^2_{T_2+2\Sigma_2} - \|\Delta x_2^{k-1}\|^2_{T_2}). \quad (32)$$

Similarly as for deriving (32), we can obtain that

$$\sigma\langle -A_3^*\Delta x_3^k, r^{k+1}\rangle \geq -(1-\tau)\sigma\langle A_3^*\Delta x_3^k, r^k\rangle + \frac{1}{2}(\|\Delta x_3^k\|^2_{T_3+2\Sigma_3} - \|\Delta x_3^{k-1}\|^2_{T_3}).$$

Adding up the above inequality and (32), and using the definitions of $B^*$ and $u$, we get the assertion (30). The proof is complete. □

**Lemma 2.3.** *Assume that Assumptions* 2.1 *and* 2.2 *hold. Let* $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ *be generated by Algorithm sPADMM. For any* $\tau \in (0, +\infty)$ *and integer* $k \geq 1$, *we have*

$$(1-\tau)\sigma\|r^{k+1}\|^2 + s_{k+1} \geq t_{k+1} + \max(1-\tau, 1-\tau^{-1})\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) + \min(\tau, 1+\tau-\tau^2)\sigma\tau^{-1}\|r^{k+1}\|^2 + (\xi_{k+1} - \xi_k), \quad (33)$$

*where* $s_{k+1}, t_{k+1}, \xi_{k+1}$ *and* $r^{k+1}$ *are defined as in* (12).

**Proof.** By simple manipulations and using the definition of $r^{k+1}$, we obtain

$$\|A_1^*x_1^{k+1} + B^*u^k - c\|^2 = \|r^{k+1} - B^*\Delta u^k\|^2 = \|r^{k+1}\|^2 - 2\langle B^*\Delta u^k, r^{k+1}\rangle + \|B^*\Delta u^k\|^2. \quad (34)$$

It follows from (30) and (34) that

$$(1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|A_1^*x_1^{k+1} + B^*u^k - c\|^2 \geq \sigma\|B^*\Delta u^k\|^2 + (2-\tau)\sigma\|r^{k+1}\|^2 - 2(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle + 2\sigma\langle A_2^*\Delta x_2^k, A_3^*(\Delta x_3^{k-1} - \Delta x_3^k)\rangle + \sum_{i=2}^{3}(\|\Delta x_i^k\|^2_{T_i+2\Sigma_i} - \|\Delta x_i^{k-1}\|^2_{T_i}). \quad (35)$$

By the Cauchy–Schwarz inequality, for the parameter $\alpha \in (0,1]$, we have

$$
\begin{aligned}
&2\sigma\langle A_2^*\Delta x_2^k, A_3^*(\Delta x_3^{k-1} - \Delta x_3^k)\rangle \\
&\quad = 2\langle(\alpha\Sigma_2)^{\frac{1}{2}}\Delta x_2^k, \sigma(\alpha\Sigma_2)^{-\frac{1}{2}}(A_2A_3^*)\Delta x_3^{k-1}\rangle \\
&\qquad - 2\langle(\alpha\Sigma_2)^{\frac{1}{2}}\Delta x_2^k, \sigma(\alpha\Sigma_2)^{-\frac{1}{2}}(A_2A_3^*)\Delta x_3^{k}\rangle \\
&\quad \geq -\alpha\|\Delta x_2^k\|_{\Sigma_2}^2 - \frac{\sigma^2}{\alpha}\|\Delta x_3^{k-1}\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\qquad - \alpha\|\Delta x_2^k\|_{\Sigma_2}^2 - \frac{\sigma^2}{\alpha}\|\Delta x_3^{k}\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\quad = -2\alpha\|\Delta x_2^k\|_{\Sigma_2}^2 - \frac{\sigma^2}{\alpha}(\|\Delta x_3^{k-1}\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\qquad + \|\Delta x_3^{k}\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)}).
\end{aligned}
$$

Substituting the above inequality into (35), we get

$$
\begin{aligned}
&(1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|A_1^*x_1^{k+1} + B^*u^k - c\|^2 \\
&\quad \geq \sigma\|B^*\Delta u^k\|^2 + (2-\tau)\sigma\|r^{k+1}\|^2 - 2(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle \\
&\qquad + (\|\Delta x_2^k\|_{T_2}^2 - \|\Delta x_2^{k-1}\|_{T_2}^2) + (\|\Delta x_3^k\|^2_{T_3 + \frac{\sigma^2}{\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\qquad - \|\Delta x_3^{k-1}\|^2_{T_3 + \frac{\sigma^2}{\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)}) + 2(1-\alpha)\|\Delta x_2^k\|_{\Sigma_2}^2 \\
&\qquad + 2\|\Delta x_3^k\|_{\Sigma_3}^2 - \frac{2\sigma^2}{\alpha}\|\Delta x_3^k\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)}. 
\end{aligned} \tag{36}
$$

By using the definitions of $s_{k+1}$ and $t_{k+1}$, and the fact that

$$
\begin{aligned}
\|\Delta u^k\|_H^2 &= \|\Delta x_2^k\|^2_{\frac{5(1-\alpha)}{2}\Sigma_2 + T_2} + \|\Delta x_3^k\|^2_{\frac{5}{2}\Sigma_3 + T_3 - \frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\quad + \min(\tau, 1+\tau-\tau^2)\sigma\|B^*\Delta u^k\|^2,
\end{aligned}
$$

we have

$$
\begin{aligned}
&2(1-\alpha)\|\Delta x_2^k\|_{\Sigma_2}^2 + 2\|\Delta x_3^k\|_{\Sigma_3}^2 - \frac{2\sigma^2}{\alpha}\|\Delta x_3^k\|^2_{(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)} \\
&\quad = -s_{k+1} + t_{k+1} - \min(\tau, 1+\tau-\tau^2)\sigma\|B^*\Delta u^k\|^2 \\
&\qquad + \sigma\|A_1^*x_1^{k+1} + B^*u^k - c\|^2.
\end{aligned}
$$

Substituting the above equation into (36) and using the definition of $\xi_{k+1}$, we get

$$
\begin{aligned}
&(1-\tau)\sigma\|r^{k+1}\|^2 + s_{k+1} - t_{k+1} + \min(\tau, 1+\tau-\tau^2)\sigma\|B^*\Delta u^k\|^2 \\
&\quad \geq \sigma\|B^*\Delta u^k\|^2 + (2-\tau)\sigma\|r^{k+1}\|^2 - 2(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle \\
&\qquad + (\xi_{k+1} - \xi_k).
\end{aligned} \tag{37}
$$

By using the Cauchy–Schwarz inequality, we get

$$\begin{cases} -2(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle \geq -(1-\tau)\sigma\|B^*\Delta u^k\|^2 - (1-\tau)\sigma\|r^k\|^2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } \tau \in (0,1], \\ -2(1-\tau)\sigma\langle B^*\Delta u^k, r^k\rangle \geq (1-\tau)\tau\sigma\|B^*\Delta u^k\|^2 + \dfrac{(1-\tau)\sigma}{\tau}\|r^k\|^2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } \tau \in (1,+\infty). \end{cases} \tag{38}$$

Substituting (38) into (37), we obtain from simple manipulations that

$$\begin{aligned} &(1-\tau)\sigma\|r^{k+1}\|^2 + s_{k+1} - t_{k+1} + \min(\tau, 1+\tau-\tau^2)\sigma\|B^*\Delta u^k\|^2 \\ &\quad \geq \max(1-\tau, 1-\tau^{-1})\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) \\ &\qquad + \min(\tau, 1+\tau-\tau^2)\sigma(\tau^{-1}\|r^{k+1}\|^2 + \|B^*\Delta u^k\|^2) + (\xi_{k+1} - \xi_k). \end{aligned}$$

The assertion (33) is proved immediately. □

Now, we are ready to prove the convergence of the sequence $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ generated by Algorithm sPADMM.

**Theorem 2.1.** *Assume that Assumptions* 2.1–2.3 *hold. Let* $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ *be generated by Algorithm sPADMM. Then, for any* $\tau \in (0, +\infty)$ *and integer* $k \geq 1$, *we have*

$$\begin{aligned} &(\overline{\phi}_k + \max(1-\tau, 1-\tau^{-1})\sigma\|r^k\|^2 + \xi_k) \\ &\qquad - (\overline{\phi}_{k+1} + \max(1-\tau, 1-\tau^{-1})\sigma\|r^{k+1}\|^2 + \xi_{k+1}) \\ &\quad \geq t_{k+1} + \min(\tau, 1+\tau-\tau^2)\sigma\tau^{-1}\|r^{k+1}\|^2, \end{aligned} \tag{39}$$

*where* $\overline{\phi}_k, \xi_{k+1}, t_{k+1}$ *and* $r^k$ *are defined as in* (12). *Assume that* $\tau \in (0, (1+\sqrt{5})/2)$. *If for some* $\alpha \in (0,1]$ *it holds that*

$$\frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0, \quad H \succ 0 \quad \text{and} \quad M \succ 0, \tag{40}$$

*then the whole sequence* $\{(x_1^k, x_2^k, x_3^k)\}$ *converges to an optimal solution to problem* (1) *and* $\{z^k\}$ *converges to an optimal solution to the dual of problem* (1).

**Proof.** By substituting (33) into (13), we can easily get (39).

Assume that $\tau \in (0, (1+\sqrt{5})/2)$. Since (40) holds for some $\alpha \in (0,1]$, we have $\min(\tau, 1+\tau-\tau^2) > 0, H \succ 0$ and $M \succ 0$. From (39), we see immediately that the sequence $\{\overline{\phi}_{k+1}\}$ is bounded, $\lim_{k\to\infty} t_{k+1} = 0$ and $\lim_{k\to\infty} \|r^{k+1}\| = 0$, i.e.,

$$\begin{aligned} &\lim_{k\to\infty} \|\Delta x_1^k\|^2_{\frac{1}{2}\Sigma_1 + T_1} = 0, \quad \lim_{k\to\infty} \|\Delta u^k\|^2_H = 0, \\ &\lim_{k\to\infty} \|r^{k+1}\| = \lim_{k\to\infty} (\tau\sigma)^{-1}\|\Delta z^k\| = 0. \end{aligned} \tag{41}$$

Since $H \succ 0$, we also have that

$$\lim_{k\to\infty} \|\Delta x_2^k\| = 0, \quad \lim_{k\to\infty} \|\Delta x_3^k\| = 0 \tag{42}$$

and thus

$$\begin{aligned} \|A_1^* \Delta x_1^k\| &= \left\| r^{k+1} - r^k - \left( \sum_{j=2}^{3} A_j^* \Delta x_j^k \right) \right\| \\ &\le \|r^{k+1}\| + \|r^k\| + \sum_{j=2}^{3} \|A_j^* \Delta x_j^k\| \to 0 \end{aligned} \tag{43}$$

as $k \to \infty$. Now from (41) and (43), we obtain

$$\lim_{k\to\infty} \|\Delta x_1^k\|^2_{(\frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^*)} = \lim_{k\to\infty} (\|\Delta x_1^k\|^2_{\frac{1}{2}\Sigma_1 + T_1} + \sigma \|A_1^* \Delta x_1^k\|^2) = 0. \tag{44}$$

Recall that $\frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0$. Thus it follows from (44) that

$$\lim_{k\to\infty} \|\Delta x_1^k\| = 0. \tag{45}$$

By the definition of $\overline{\phi}_{k+1}$, we see that the three sequences $\{\|z_e^{k+1}\|\}, \{\|x_{1e}^{k+1}\|_{\Sigma_1+T_1}\}$, and $\{\|u_e^{k+1}\|_M\}$ are all bounded. Since $M \succ 0$, the sequences $\{\|x_2^{k+1}\|\}$ and $\{\|x_3^{k+1}\|\}$ are also bounded. Furthermore, by using

$$\|A_1^* x_{1e}^{k+1}\| = \|A^* x^{k+1} - A^* \bar{x} - B^* u_e^{k+1}\| \le \|r^{k+1}\| + \|B^* u_e^{k+1}\|, \tag{46}$$

we also know that the sequence $\{\|A_1^* x_{1e}^{k+1}\|\}$ is bounded, and so is the sequence $\{\|x_{1e}^{k+1}\|_{(\Sigma_1+T_1+\sigma A_1 A_1^*)}\}$. This shows that the sequence $\{\|x_1^{k+1}\|\}$ is also bounded as the operator $\Sigma_1 + T_1 + \sigma A_1 A_1^* \succeq \frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0$. Thus, the sequence $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ is bounded.

Since the sequence $\{(x_1^k, x_2^k, x_3^k, z^k)\}$ is bounded, there is a subsequence $\{(x_1^{k_i}, x_2^{k_i}, x_3^{k_i}, z^{k_i})\}$ which converges to a cluster point, say $\{(x_1^\infty, x_2^\infty, x_3^\infty, z^\infty)\}$. Taking limits on both sides of (14) along the subsequence $\{(x_1^{k_i}, x_2^{k_i}, x_3^{k_i}, z^{k_i})\}$, using (41), (42) and (45), we obtain that

$$-A_j z^\infty \in \partial\theta_j(x_j^\infty), \quad j = 1,2,3 \quad \text{and} \quad A^* x^\infty - c = 0,$$

i.e., $(x_1^\infty, x_2^\infty, x_3^\infty, z^\infty)$ satisfies (11). Thus $\{(x_1^\infty, x_2^\infty, x_3^\infty)\}$ is an optimal solution to (1) and $z^\infty$ is an optimal solution to the dual of problem (1).

To complete the proof, we show next that $(x_1^\infty, x_2^\infty, x_3^\infty, z^\infty)$ is actually the unique limit of $\{(x_1^k, x_2^k, x_3^k, z^k)\}$. Replacing $(\bar{x}_1, \bar{u}, \bar{z}) := (\bar{x}_1, (\bar{x}_2, \bar{x}_3), \bar{z})$ by $(x_1^\infty, u^\infty, z^\infty) := (x_1^\infty, (x_2^\infty, x_3^\infty), z^\infty)$ in (39), for any integer $k \ge k_i$, we have

$$\begin{aligned} &\phi_{k+1}(x_1^\infty, u^\infty, z^\infty) + \max(1-\tau, 1-\tau^{-1})\sigma\|r^{k+1}\|^2 + \xi_{k+1} \\ &\quad \le \phi_{k_i}(x_1^\infty, u^\infty, z^\infty) + \max(1-\tau, 1-\tau^{-1})\sigma\|r^{k_i}\|^2 + \xi_{k_i}. \end{aligned} \tag{47}$$

Note that

$$\lim_{i\to\infty} (\phi_{k_i}(x_1^\infty, u^\infty, z^\infty) + \max(1-\tau, 1-\tau^{-1})\sigma\|r^{k_i}\|^2 + \xi_{k_i}) = 0.$$

Therefore, from (47) we get

$$\lim_{k\to\infty} \phi_{k+1}(x_1^\infty, u^\infty, z^\infty) = 0,$$

i.e.,

$$\lim_{k\to\infty} ((\sigma\tau)^{-1}\|z^{k+1} - z^\infty\|^2 + \|x_1^{k+1} - x_1^\infty\|^2_{\Sigma_1+T_1} + \|u^{k+1} - u^\infty\|^2_M) = 0.$$

Since $M \succ 0$, we also have that $\lim_{k\to\infty} u^k = u^\infty$, that is $\lim_{k\to\infty} x_2^k = x_2^\infty$ and $\lim_{k\to\infty} x_3^k = x_3^\infty$. Using the fact that $\lim_{k\to\infty} \|r^{k+1}\| = 0$ and $\lim_{k\to\infty} \|u^{k+1} - u^\infty\| = 0$, we get from (46) that $\lim_{k\to\infty} \|A_1^*(x_1^{k+1} - x_1^\infty)\| = 0$. Thus

$$\lim_{k\to\infty} \|x_1^{k+1} - x_1^\infty\|^2_{\Sigma_1+T_1+\sigma A_1A_1^*} = 0.$$

Since $\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0$, we also obtain that $\lim_{k\to\infty} x_1^k = x_1^\infty$. Therefore, we have shown that the sequence $\{(x_1^k, x_2^k, x_3^k)\}$ converges to an optimal solution to (1) and $\{z^k\}$ converges to an optimal solution to the dual of problem (1) for any $\tau \in (0, (1+\sqrt{5})/2)$. The proof is complete. □

**Remark 2.1.** Assume that $(1-\alpha)\Sigma_2 + \sigma A_2 A_2^*$ is invertible for some $\alpha \in (0,1]$. Set $\tau = 1$ (the case that $1 \neq \tau \in (0,(1+\sqrt{5})/2)$ can be discussed in a similar but slightly more complicated manner) and $T_2 = 0$ in (8) and (9). Then the assumptions $H \succ 0$ and $M \succ 0$ in (40) reduce to

$$\begin{pmatrix} \frac{5(1-\alpha)}{2}\Sigma_2 + \sigma A_2 A_2^* & \sigma A_2 A_3^* \\ \sigma A_3 A_2^* & \frac{5}{2}\Sigma_3 + T_3 + \sigma A_3 A_3^* - \frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*) \end{pmatrix} \succ 0$$

and

$$\begin{pmatrix} (1-\alpha)\Sigma_2 + \sigma A_2 A_2^* & \sigma A_2 A_3^* \\ \sigma A_3 A_2^* & \Sigma_3 + T_3 + \sigma A_3 A_3^* \end{pmatrix} \succ 0,$$

which are, respectively, equivalent to

$$\frac{5}{2}\Sigma_3 + T_3 + \sigma A_3 A_3^* - \frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*) - \sigma^2(A_3A_2^*)\left(\frac{5(1-\alpha)}{2}\Sigma_2 + \sigma A_2A_2^*\right)^{-1}(A_2A_3^*) \succ 0 \tag{48}$$

and

$$\Sigma_3 + T_3 + \sigma A_3 A_3^* - \sigma^2(A_3A_2^*)((1-\alpha)\Sigma_2 + \sigma A_2A_2^*)^{-1}(A_2A_3^*) \succ 0 \tag{49}$$

in terms of the Schur-complement format. The conditions (48) and (49) can be satisfied easily by choosing a proper $T_3$ for given $\alpha \in (0,1]$ and $\sigma \in (0,+\infty)$. Evidently, with a fixed $\alpha, T_3$ can take a smaller value with a smaller $\sigma$ and $T_3$ can even take the zero operator for any $\sigma > 0$ smaller than a certain threshold if $\Sigma_3 + (1-\alpha)\sigma A_3 A_3^* \succ 0$. To see this, let us consider the following example constructed in Chen *et al.* (2014):

$$\begin{aligned} \min \quad & \frac{1}{20}x_1^2 + \frac{1}{20}x_2^2 + \frac{1}{20}x_3^2 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0, \end{aligned} \tag{50}$$

which is a convex minimization problem with three strongly convex functions. In Chen *et al.* (2014) showed that the directly extended 3-block ADMM scheme (4) with $\tau = \sigma = 1$ applied to problem (50) is divergent. For problem (50), $\Sigma_1 = \Sigma_2 = \Sigma_3 = \frac{1}{10}, A_1 = (1,1,1), A_2 = (1,1,2)$ and $A_3 = (1,2,2)$. From (48) and (49), by taking $\alpha = 1$, we have that $T_3$ and $\sigma$ should satisfy the following conditions

$$\frac{1}{4} + T_3 - 1225\sigma^2 + \frac{5}{6}\sigma > 0 \quad \text{and} \quad \frac{1}{10} + T_3 + \frac{5}{6}\sigma > 0,$$

which hold true, in particular, if $T_3 = 0$ and $\sigma < \frac{1+\sqrt{1765}}{2940} \approx 0.015$ or if $\sigma = 1$ and $T_3 > \frac{14687}{12} \approx 1223.92$.

**Remark 2.2.** If $A_2^*$ is vacuous, then for any integer $k \geq 0$, we have that $x_2^{k+1} = x_2^0 = \bar{x}_2$, the 3-block sPADMM is just a 2-block sPADMM, and condition (40) reduces to

$$\frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0, \quad \Sigma_3 + T_3 + \sigma A_3 A_3^* \succ 0 \quad \text{and}$$

$$\frac{5}{2}\Sigma_3 + T_3 + \min(\tau, 1 + \tau - \tau^2)\sigma A_3 A_3^* \succ 0,$$

which is equivalent to

$$\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0 \quad \text{and} \quad \Sigma_3 + T_3 + \sigma A_3 A_3^* \succ 0 \tag{51}$$

since $\Sigma_1 \succeq 0, T_1 \succeq 0, \Sigma_3 \succeq 0$ and $T_3 \succeq 0$. Condition (51) is exactly the same as the one used in Theorem B.1 in Fazel *et al.* (2013).

## 3. Conclusions

In this paper, we provided a convergence analysis about a 3-block sPADMM for solving separable convex minimization problems with the condition that the second block in the objective is strongly convex.[a] The step-length $\tau$ in our proposed sPADMM is allowed to stay in the desirable region $(0, (1+\sqrt{5})/2)$. From Remark 2.1, we know that with a fixed parameter $\alpha \in (0, 1]$, the added semi-proximal terms can be chosen to be small if the penalty parameter $\sigma$ is small. If $A_1^*$ and $A_3^*$ are both injective and $\sigma > 0$ is taken to be smaller than a certain threshold, then the convergent 3-block sPADMM includes the directly extended 3-block ADMM with

[a]One can prove similar results if the third instead of the second block is strongly convex.

$\tau \in (0, (1+\sqrt{5})/2)$ by taking $T_i, i = 1, 2, 3$, to be zero operators. With no much difficulty, one could extend our 3-block sPADMM to deal with the $m$-block ($m \geq 4$) separable convex minimization problems possessing $m - 2$ strongly convex blocks and provide the iteration complexity analysis for the corresponding algorithm in the sense of He and Yuan (2012). In this work, we choose not to do the extension because we are not aware of interesting applications of the $m$-block ($m \geq 4$) separable convex minimization problems with $m - 2$ strongly convex blocks. While our sufficient condition bounding the range of values for $\sigma$ and $T_3$ is quite flexible, it may have one potential limitation: $T_3$ can be very large if $\sigma$ is not small as shown in Remark 2.1. Since a larger $T_3$ can potentially make the algorithm converge slower, in our future research we shall first study how this limitation can be circumvented before we study other important issues such as the iteration complexity.[b]

## Acknowledgments

The first author was supported by the National Natural Science Foundation of China (Grant No. 11001053), Program for New Century Excellent Talents in University (Grant No. NCET-12-0111) and Qing Lan Project. This research was supported in part by the Ministry of Education, Singapore, Academic Research Fund (Grant No. R-146-000-194-112).

## References

Cai, XJ, DR Han and XM Yuan (2014). The direct extension of ADMM for three-block separable convex minimization models is convergent when one function is strongly convex. Optimization Online, http://www.optimization-online.org/DB_FILE/2014/11/4644.pdf.

Chen, CH, BS He, YY Ye and XM Yuan (2014). The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Mathematical Programming*, *Series A*, doi: 10.1007/s10107-014-0826-5.

Chen, CH, Y Shen and YF You (2013). On the convergence analysis of the alternating direction method of multipliers with three blocks. *Abstract and Applied Analysis*, 2013, Article ID 183961, 7 pp.

Eckstein, J and DP Bertsekas (1992). On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55, 293–318.

Fazel, M, TK Pong, DF Sun and P Tseng (2013). Hankel matrix rank minimization with applications to system identification and realization. *SIAM Journal on Matrix Analysis and Applications*, 34(3), 946–977.

Fortin, M and R Glowinski (1983). *Augmented Lagrangian Methods*: *Applications to the Numerical Solution of Boundary-value Problems*, Studies in Mathematics and its Applications, Vol. 15, Amsterdam: North-Holland, Translated from the French by B. Hunt and D. C. Spicer.

[b]In a recent report Cai *et al.* (2014) independently proved a result similar to Theorem 2.1 for the directly extended ADMM (i.e., all the three semi-proximal terms $T_1, T_2$ and $T_3$ disappear) with $\tau = 1$ and provided an analysis on the iteration complexity.

Gabay, D (1983). Applications of the method of multipliers to variational inequalities. In *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, M Fortin and R Glowinski (eds.), Studies in Mathematics and Its Applications, Vol. 15, pp. 299–331. Amesterdam: North-Holland.

Gabay, D and B Mercier (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers and Mathematics with Applications*, 2, 17–40.

Glowinski, R (1980). *Lectures on Numerical Methods for Nonlinear Variational Problems*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, Vol. 65. Bombay: Tata Institute of Fundamental Research, Notes by M.G. Vijayasundaram and M Adimurthi.

Glowinski, R and A Marrocco (1975). Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de dirichlet non linéares. *Revue Francaise d'Automatique, Informatique et Recherche Opérationelle*, 9, 41–76.

Han, DR and XM Yuan (2012). A note on the alternating direction method of multipliers. *Journal of Optimization Theory and Applications*, 155(1), 227–238.

He, BS and XM Yuan (2012). On the $O(1/n)$ convergence rate of the Douglas–Rachford alternating direction method. *SIAM Journal on Numerical Analysis*, 50, 700–709.

Hong, MY and ZQ Luo (2013). On the linear convergence of the alternating direction method of multipliers. arXiv:1208.3922v3.

Li, XD, DF Sun and K-C Toh (2014). A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions. *Mathematical Programming*, Series A, doi: 10.1007/s10107-014-0850-5.

Lin, TY, SQ Ma and SZ Zhang (2014a). On the convergence rate of multi-block ADMM. arXiv:1408.4265v1.

Lin, TY, SQ Ma and SZ Zhang (2014b). On the global linear convergence of the ADMM with multi-block variables. arXiv:1408.4266v1.

Nie, JW and YX Yuan (2001). A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps. *Annals of Operations Research*, 103, 115–133.

Qi, HD (2009). Local duality of nonlinear semidefinite programming. *Mathematics of Operations Research*, 34(1), 124–141.

Qi, HD and DF Sun (2011). An augmented Lagrangian dual approach for the H-weighted nearest correlation matrix problem. *IMA Journal of Numerical Analysis*, 31, 491–511.

Rockafellar, RT (1970). *Convex Analysis*. Princeton: Princeton University Press.

Sun, DF (2006). The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. *Mathematics of Operations Research*, 31, 761–776.

Sun, DF, J Sun and LW Zhang (2008). The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Mathematical Programming*, 114, 349–391.

Sun, J and S Zhang (2010). A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs. *European Journal of Operational Research*, 207, 1210–1220.

Toh, K-C (2008). An inexact primal-dual path-following algorithm for convex quadratic SDP. *Mathematical Programming*, 112, 221–254.

Toh, K-C, RH Tütüncü and MJ Todd (2007). Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems. *Pacific Journal of Optimization*, 3, 135–164.

Zhao, XY (2009). A semismooth Newton-CG augmented Lagrangian method for large scale linear and convex quadratic SDPs. Ph.D. thesis, Department of Mathematics, National University of Singapore.

## Biography

**Min Li** received her BS and PhD degrees in 2002 and 2007, respectively, both from Department of Mathematics, Nanjing University, China. She joined School of Economics and Management, Southeast University in 2007, where she is currently a full professor of Management Science and Engineering. Her research interests include numerical optimization and its applications.

**Defeng Sun** is Professor at Department of Mathematics, National University of Singapore. He received his PhD in Operations Research and Control Theory from the Institute of Applied Mathematics, Chinese Academy of Sciences, China in 1995. He completed his post-doctoral training at the University of New South Wales, Australia. His research interests are mainly on Optimization, a subject of studying best decision-making with limited resources, with side interest in financial risk management. He currently serves as an Associate Editor to Mathematical Programming Series A and Series B, SIAM Journal on Optimization and Journal of China Operations Research Society.

**Kim-Chuan Toh** is a Professor at Department of Mathematics, National University of Singapore. He received his BS and MS degrees in Mathematics from the National University of Singapore, and PhD degree in Applied Mathematics from Cornell University, Ithaca, NY. His research focuses on algorithms for large scale conic programming, in particular, semidefinite programming. He is currently serving as an Area Editor of Mathematical Programming Computation, an Associate Editor of SIAM Journal on Optimization, and a Subject Editor of Optimization and Engineering.