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

Min Li,[*] Defeng Sun[†] and Kim-Chuan Toh[‡]

Abstract

In this paper, we present a semi-proximal alternating direction method of multipliers (AD-MM) 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 semi-proximal ADMM 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 semi-proximal ADMM 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:

$$\min_{x_1,x_2,x_3} \left\{ \theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3) \;\middle|\; A_1^* x_1 + A_2^* x_2 + A_3^* x_3 = c, \; x_i \in \mathcal{X}_i, \; i = 1, 2, 3 \right\}, \qquad (1)$$

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 θ_i, $i = 1, 2, 3$, are closed proper convex functions, there exist self-adjoint and positive semi-definite operators Σ_i, $i = 1, 2, 3$, such that

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

where $\partial\theta_i$ is the sub-differential mapping of θ_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 $k = 0, 1, \ldots$

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

where $\tau > 0$ is the step-length. Different from the 2-block ADMM whose convergence has been established for a long time [10, 8, 9, 6, 7, 4], the 3-block ADMM may not converge in general, which was demonstrated by Chen, He, Ye and Yuan [2] using counterexamples. Nevertheless, if all the functions θ_i, $i = 1, 2, 3$, are strongly convex, Han and Yuan [11] 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

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

where $\lambda_{\max}(S)$ is the largest eigenvalue of a given self-adjoint linear operator S. Hong and Luo [13] proposed to adopt a small step-length τ when updating the Lagrange multiplier z^{k+1} in (4). Chen, Shen and You [3] proposed the following sufficient condition

$$A_1^* \text{ is injective, } \ \Sigma_i = \mu_i I \succ 0, \ i = 2, 3 \ \text{ and } \ 0 < \sigma < \frac{\mu_2}{\lambda_{\max}(A_2 A_2^*)}, \ \ \sigma \leq \frac{\mu_3}{\lambda_{\max}(A_3 A_3^*)}$$

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, Shen and You [3], in [15], Lin, Ma and Zhang provided an analysis on the iteration complexity for the same method under the condition

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

2

In [16], 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 studied in [5] to deal with problem (1) by only assuming θ_2 to be strongly convex, i.e., $\Sigma_2 \succ 0$. Note that the semi-proximal ADMM 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 semi-proximal ADMM that allows τ 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 Ξ. 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 semi-proximal ADMM 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 semidefinite programming problem that has been extensively studied both theoretically and numerically in the literature [14, 17, 18, 19, 21, 22, 23, 24, 25, 26], to name only a few.

The remaining parts of this paper are organized as follows. In the next section, we first present our 3-block semi-proximal ADMM and then provide the main convergence results. We give some concluding remarks in the final section.

Notation.

- The effective domain of a function $f\colon \mathcal{X} \to (-\infty, +\infty]$ 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 $\|x\|_G^2$ 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 Σ.

- Denote

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

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

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

$$
H := \begin{pmatrix} \frac{5(1-\alpha)}{2}\Sigma_2 + T_2 & 0 \\ 0 & \frac{5}{2}\Sigma_3 + T_3 - \frac{5\sigma^2}{2\alpha}(A_2 A_3^*)^* \Sigma_2^{-1}(A_2 A_3^*) \end{pmatrix} + \min(\tau, 1+\tau-\tau^2)\sigma BB^*. \tag{9}
$$

2   A 3-Block Semi-Proximal ADMM

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

Algorithm sPADMM: A 3-block semi-proximal ADMM 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

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

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

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

Assumption 2.1 *The convex function θ_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.*

4

Assumption 2.3 *There exists* $x' = (x'_1, x'_2, x'_3) \in ri(dom(\theta_1) \times dom(\theta_2) \times dom(\theta_3)) \bigcap P$, *where*

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

Under Assumption 2.3, it follows from [20, Corollary 28.2.2] and [20, Corollary 28.3.1] 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

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

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

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

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, 2.2 and 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

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

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

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

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

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

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 Δx_i^k, we get

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

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 Δx_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

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

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

$$\sum_{i=1}^{3} \left\langle x_{ie}^{k+1}, \ A_i \bar{z} - A_i \tilde{z}^{k+1} \right\rangle + \sigma \left\langle x_{1e}^{k+1}, \ A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \right\rangle + \sigma \left\langle x_{2e}^{k+1}, \ A_2 A_3^* \Delta x_3^k \right\rangle$$

$$+ \frac{1}{2} \sum_{i=1}^{3} \left(\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2 \right) \geq \frac{1}{2} \sum_{i=1}^{3} \|\Delta x_i^k\|_{T_i}^2 + \sum_{i=1}^{3} \|x_{ie}^{k+1}\|_{\Sigma_i}^2. \tag{20}$$

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

$$\sigma \left\langle x_{1e}^{k+1}, \ A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \right\rangle = \sigma \left\langle -x_{1e}^{k+1}, \ -A_1 B^* \Delta u^k \right\rangle = \sigma \left\langle -A_1^* x_{1e}^{k+1}, \ B^* u^k - B^* u^{k+1} \right\rangle$$

$$= \sigma \left\langle (-B^* \bar{u}) - (A_1^* x_1^{k+1} - c), \ (-B^* u^{k+1}) - (-B^* u^k) \right\rangle. \tag{21}$$

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

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

6

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

$$
\sigma \langle x_{1e}^{k+1}, \ A_1 \sum_{j=2}^{3} A_j^* \Delta x_j^k \rangle
$$

$$
= \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2) + \frac{\sigma}{2}(\|A_1^* x_1^{k+1} + B^* u^{k+1} - c\|^2 - \|A_1^* x_1^k + B^* u^k - c\|^2)
$$

$$
= \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2) + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2 - \frac{\sigma}{2}\|A_1^* x_1^{k+1} + B^* u^k - c\|^2. \tag{23}
$$

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

$$
\sigma \langle x_{2e}^{k+1}, \ A_2 A_3^* \Delta x_3^k \rangle = 2 \langle (\alpha \Sigma_2)^{\frac{1}{2}} x_{2e}^{k+1}, \ \frac{\sigma}{2}(\alpha \Sigma_2)^{-\frac{1}{2}} A_2 A_3^* \Delta x_3^k \rangle
$$

$$
\leq \alpha \|x_{2e}^{k+1}\|_{\Sigma_2}^2 + \frac{\sigma^2}{4\alpha}\|\Delta x_3^k\|_{(A_2 A_3^*)^* \Sigma_2^{-1}(A_2 A_3^*)}^2. \tag{24}
$$

It follows from (17) that

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

Substituting (23), (24) and (25) into (20), we obtain

$$
\frac{1}{\sigma} \langle \bar{z} - \tilde{z}^{k+1}, \ \tilde{z}^{k+1} - z^k \rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2 + \frac{\sigma}{2}(\|B^* u_e^k\|^2 - \|B^* u_e^{k+1}\|^2)
$$

$$
+ \frac{1}{2} \sum_{i=1}^{3} (\|x_{ie}^k\|_{T_i}^2 - \|x_{ie}^{k+1}\|_{T_i}^2)
$$

$$
\geq \frac{\sigma}{2}\|A_1^* x_1^{k+1} + B^* u^k - c\|^2 + \frac{1}{2} \sum_{i=1}^{3} \|\Delta x_i^k\|_{T_i}^2 + \sum_{i=1, i \neq 2}^{3} \|x_{ie}^{k+1}\|_{\Sigma_i}^2
$$

$$
+ (1 - \alpha)\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \frac{\sigma^2}{4\alpha}\|\Delta x_3^k\|_{(A_2 A_3^*)^* \Sigma_2^{-1}(A_2 A_3^*)}^2. \tag{26}
$$

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

$$
\sum_{i=1, i \neq 2}^{3} \|x_{ie}^{k+1}\|_{\Sigma_i}^2 + (1 - \alpha)\|x_{2e}^{k+1}\|_{\Sigma_2}^2
$$

$$
= \frac{1}{2} \sum_{i=1, i \neq 2}^{3} (\|x_{ie}^{k+1}\|_{\Sigma_i}^2 + \|x_{ie}^k\|_{\Sigma_i}^2) + \frac{1}{2} \sum_{i=1, i \neq 2}^{3} (\|x_{ie}^{k+1}\|_{\Sigma_i}^2 - \|x_{ie}^k\|_{\Sigma_i}^2)
$$

$$
+ \frac{1 - \alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 + \|x_{2e}^k\|_{\Sigma_2}^2) + \frac{1 - \alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \|x_{2e}^k\|_{\Sigma_2}^2)
$$

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

$$
+ \frac{1 - \alpha}{2}(\|x_{2e}^{k+1}\|_{\Sigma_2}^2 - \|x_{2e}^k\|_{\Sigma_2}^2). \tag{27}
$$

7

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

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

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

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

which, together with (28), imply

$$\frac{1}{\sigma}\langle \bar{z} - \tilde{z}^{k+1},\, \tilde{z}^{k+1} - z^k \rangle + \frac{1}{2\sigma}\|z^k - \tilde{z}^{k+1}\|^2 = \frac{1}{2\sigma\tau}(\|z_e^k\|^2 - \|z_e^{k+1}\|^2) + \frac{(\tau - 1)\sigma}{2}\|r^{k+1}\|^2. \quad (29)$$

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. $\qquad\square$

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*

$$-\sigma\langle B^*\Delta u^k,\, r^{k+1}\rangle \geq -(1-\tau)\sigma\langle B^*\Delta u^k,\, r^k\rangle + \frac{1}{2}\sum_{i=2}^{3}(\|\Delta x_i^k\|_{T_i + 2\Sigma_i}^2 - \|\Delta x_i^{k-1}\|_{T_i}^2)$$
$$+ \sigma\langle A_2^*\Delta x_2^k,\, A_3^*(\Delta x_3^{k-1} - \Delta x_3^k)\rangle, \qquad (30)$$

where Δu^k, Δx_i^k $(i = 2, 3)$ and r^{k+1} are defined as in (12).

Proof. Let

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

By using (14) and the definition of Δx_2^k, we have

$$-A_2 v^{k+1} - T_2\Delta x_2^k \in \partial\theta_2(x_2^{k+1}) \quad \text{and} \quad -A_2 v^k - T_2\Delta x_2^{k-1} \in \partial\theta_2(x_2^k).$$

Thus, we obtain from (2) that

$$\langle \Delta x_2^k,\, (A_2 v^k + T_2\Delta x_2^{k-1}) - (A_2 v^{k+1} + T_2\Delta x_2^k)\rangle \geq \|\Delta x_2^k\|_{\Sigma_2}^2.$$

By using the Cauchy-Schwarz inequality, we obtain

$$\langle \Delta x_2^k,\, T_2(\Delta x_2^k - \Delta x_2^{k-1})\rangle = \|\Delta x_2^k\|_{T_2}^2 - \langle \Delta x_2^k,\, T_2\Delta x_2^{k-1}\rangle \geq \frac{1}{2}\|\Delta x_2^k\|_{T_2}^2 - \frac{1}{2}\|\Delta x_2^{k-1}\|_{T_2}^2.$$

8

Adding up the above two inequalities, we get

$$\langle A_2^* \Delta x_2^k, \ v^k - v^{k+1} \rangle \geq \frac{1}{2} \|\Delta x_2^k\|_{T_2 + 2\Sigma_2}^2 - \frac{1}{2} \|\Delta x_2^{k-1}\|_{T_2}^2. \tag{31}$$

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

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

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

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

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

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*

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

where s_{k+1}, t_{k+1}, ξ_{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

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

It follows from (30) and (34) that

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

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

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

9

Substituting the above inequality into (35), we get

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

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

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

we have

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

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

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

By using the Cauchy-Schwarz inequality, we get

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

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

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

The assertion (33) is proved immediately. $\qquad\square$

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, 2.2 and 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*

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

where $\overline{\phi}_k$, ξ_{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.,

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

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

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

and thus

$$\|A_1^* \Delta x_1^k\| = \left\| r^{k+1} - r^k - (\sum_{j=2}^{3} A_j^* \Delta x_j^k) \right\| \leq \|r^{k+1}\| + \|r^k\| + \sum_{j=2}^{3} \|A_j^* \Delta x_j^k\| \to 0 \quad (43)$$

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

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

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

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

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

$$\|A_1^* x_{1e}^{k+1}\| = \left\| A^* x^{k+1} - A^* \bar{x} - B^* u_e^{k+1} \right\| \leq \|r^{k+1}\| + \|B^* u_e^{k+1}\|, \quad (46)$$

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

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

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^∞ 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 $k \geq k_i$, we have

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

11

Note that

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

Therefore, from (47) we get

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

i.e.,

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

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

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

Since $\Sigma_1+T_1+\sigma A_1A_1^*\succ 0$, 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. $\square$

Remark 2.1 *Assume that $(1-\alpha)\Sigma_2+\sigma A_2A_2^*$ 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*

$$\begin{pmatrix}\frac{5(1-\alpha)}{2}\Sigma_2+\sigma A_2A_2^* & \sigma A_2A_3^* \\ \sigma A_3A_2^* & \frac{5}{2}\Sigma_3+T_3+\sigma A_3A_3^*-\frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)\end{pmatrix}\succ 0$$

and

$$\begin{pmatrix}(1-\alpha)\Sigma_2+\sigma A_2A_2^* & \sigma A_2A_3^* \\ \sigma A_3A_2^* & \Sigma_3+T_3+\sigma A_3A_3^*\end{pmatrix}\succ 0,$$

which are, respectively, equivalent to

$$\frac{5}{2}\Sigma_3+T_3+\sigma A_3A_3^*-\frac{5\sigma^2}{2\alpha}(A_2A_3^*)^*\Sigma_2^{-1}(A_2A_3^*)-\sigma^2(A_3A_2^*)\Big(\frac{5(1-\alpha)}{2}\Sigma_2+\sigma A_2A_2^*\Big)^{-1}(A_2A_3^*)\succ 0 \quad (48)$$

and

$$\Sigma_3+T_3+\sigma A_3A_3^*-\sigma^2(A_3A_2^*)\big((1-\alpha)\Sigma_2+\sigma A_2A_2^*\big)^{-1}(A_2A_3^*)\succ 0 \quad (49)$$

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 α, T_3 can take a smaller value with a smaller σ and T_3 can even take the zero operator for any $\sigma>0$ smaller than a certain threshold if $\Sigma_3+(1-\alpha)\sigma A_3A_3^*\succ 0$. To see this, let us consider the following example constructed in [2]:

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

which is a convex minimization problem with three strongly convex functions. In [2], Chen, He, Ye and Yuan 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 σ should satisfy the following conditions

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

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*

$$\frac{1}{2}\Sigma_1 + T_1 + \sigma A_1 A_1^* \succ 0, \quad \Sigma_3 + T_3 + \sigma A_3 A_3^* \succ 0 \quad and \quad \frac{5}{2}\Sigma_3 + T_3 + \min(\tau, 1 + \tau - \tau^2)\sigma A_3 A_3^* \succ 0,$$

which is equivalent to

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

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 [5].

3   Conclusions

In this paper, we provided a convergence analysis about a 3-block semi-proximal ADMM for solving separable convex minimization problems with the condition that the second block in the objective is strongly convex[1]. The step-length τ in our proposed semi-proximal ADMM 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 σ 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 semi-proximal ADMM includes the directly extended 3-block ADMM with $\tau \in (0, (1 + \sqrt{5})/2)$ by taking T_i, $i = 1, 2, 3$, to be zero operators. With no much difficulty, one could extend our 3-block semi-proximal ADMM 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 [12]. 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 σ and T_3 is quite flexible, it may have one potential limitation: T_3 can be very large if σ 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[2].

---

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

[2]In a recent report [1], Cai, Han and Yuan 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.

References

[1] X. J. Cai, D. R. Han and X. M. Yuan, The direct extension of ADMM for three-block separable convex minimization models is convergent when one function is strongly convex, Optimization Online, (2014).

[2] 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, DOI 10.1007/s10107-014-0826-5, (2014).

[3] 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 (2013), Article ID 183961, 7 pages.

[4] J. Eckstein and D. P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Mathematical Programming, 55 (1992), pp. 293–318.

[5] M. Fazel, T. K. Pong, D. F. Sun and P. Tseng, Hankel matrix rank minimization with applications to system identification and realization, SIAM Journal on Matrix Analysis and Applications, 34(3) (2013), pp. 946–977.

[6] M. Fortin and R. Glowinski, Augmented Lagrangian methods, vol. 15 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.

[7] D. Gabay, 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., vol. 15 of Studies in Mathematics and Its Applications, Elsevier, (1983), pp. 299–331.

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

[9] R. Glowinski, Lectures on numerical methods for nonlinear variational problems, vol. 65 of Tata Institute of Fundamental Research Lectures on Mathematics and Physics, Tata Institute of Fundamental Research, Bombay, 1980. Notes by M. G. Vijayasundaram and M. Adimurthi.

[10] R. Glowinski and A. Marrocco, 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 (1975), pp. 41–76.

[11] D. R. Han and X. M. Yuan, A note on the alternating direction method of multipliers, Journal of Optimization Theory and Applications, 155(1) (2012), pp. 227–238.

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

[13] M. Y. Hong and Z. Q. Luo. On the linear convergence of the alternating direction method of multipliers, arXiv:1208.3922v3, 2013.

[14] X. D. Li, D. F. Sun, K.-C. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions, accepted by Mathematical Programming, (2014).

[15] T. Y. Lin, S. Q. Ma and S. Z. Zhang, On the convergence rate of multi-block ADMM, arXiv:1408.4265v1, (2014).

[16] T. Y. Lin, S. Q. Ma and S. Z. Zhang, On the global linear convergence of the ADMM with multi-block variables, arXiv:1408.4266v1, (2014).

[17] J. W. Nie and Y. X. Yuan, A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps, Annals of Operations Research, 103 (2001), pp. 115–133.

[18] H. D. Qi, Local duality of nonlinear semidefinite programming, Mathematics of Operations Research, 34(1) (2009), pp. 124–141.

[19] H. D. Qi and D. F. Sun, An augmented Lagrangian dual approach for the H-weighted nearest correlation matrix problem, IMA Journal of Numerical Analysis, 31 (2011), pp. 491–511.

[20] R. T. Rockafellar, Convex Analysis. Princeton University Press, Princeton (1970).

[21] D. F. Sun, The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications, Mathematics of Operations Research, 31 (2006), pp. 761–776.

[22] D. F. Sun, J. Sun and L. W. Zhang, The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming, Mathematical Programming, 114 (2008), pp. 349–391.

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

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

[25] K.-C. Toh, R. H. Tütüncü and M. J. Todd, Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems, Pacific Journal of Optimization, 3 (2007), pp. 135–164.

[26] X. Y. Zhao, A semismooth Newton-CG augmented Lagrangian method for large scale linear and convex quadratic SDPs, PhD thesis, Department of Mathematics, National University of Singapore, 2009.