Supplementary Material of Fast Proximal Linearized Alternating Direction Method of Multiplier with Parallel Splitting

Canyi Lu\textsuperscript{1}, Huan Li\textsuperscript{2}, Zhouchen Lin\textsuperscript{2,3,∗}, Shuicheng Yan\textsuperscript{1}

\textsuperscript{1} Department of Electrical and Computer Engineering, National University of Singapore
\textsuperscript{2} Key Laboratory of Machine Perception (MOE), School of EECS, Peking University
\textsuperscript{3} Cooperative Medianet Innovation Center, Shanghai Jiaotong University
canyilu@gmail.com, lihuan_ss@126.com, zlin@pku.edu.cn, eleyans@nus.edu.sg

This documents provides the proof details of the convergence results of our proposed fast methods. First, in Section 1, we give some useful results which are useful for the convergence analysis of Fast PALM in Section 2 and Fast PL-ADMM-PS in Section 3.

1. Some Lemmas

Lemma 1 \cite{2} Let \( g : \mathbb{R}^m \to \mathbb{R} \) be a continuously differentiable function with Lipschitz continuous gradient and Lipschitz constant \( L \). Then, for any \( x, y \in \mathbb{R}^m \),

\[
g(x) \leq g(y) + \langle x - y, \nabla g(y) \rangle + \frac{L}{2} ||x - y||^2. \tag{1}
\]

Lemma 2 Given any \( a, b, c, d \in \mathbb{R}^m \), we have

\[
\langle a - b, a - c \rangle = \frac{1}{2} (||a - b||^2 + ||a - c||^2 - ||b - c||^2). \tag{2}
\]

\[
\langle a - b, c - d \rangle = \frac{1}{2} (||a - d||^2 - ||a - c||^2 - ||b - d||^2 + ||b - c||^2). \tag{3}
\]

Lemma 3 Assume the sequences \( \{a^{(k)}\} \) and \( \{b^{(k)}\} \) satisfy \( a^{(0)} = 1, 0 < a^{(k+1)} - a^{(k)} \leq 1 \) and \( b^{(k)} > 0 \). Then we have

\[
\sum_{k=0}^{K} a^{(k)} (b^{(k)} - b^{(k+1)}) \leq \sum_{k=0}^{K} b^{(k)}. \tag{4}
\]

Proof. We deduce

\[
\sum_{k=0}^{K} a^{(k)} (b^{(k)} - b^{(k+1)}) = a^{(0)} b^{(0)} + \sum_{k=0}^{K-1} (a^{(k+1)} - a^{(k)}) b^{(k+1)} - a^{(K)} b^{(K+1)} \\
\leq b^{(0)} + \sum_{k=0}^{K-1} b^{(k+1)} = \sum_{k=0}^{K} b^{(k)}. \tag{5}
\]

Lemma 4 Define the sequence \( \{\theta^{(k)}\} \) as \( \theta^{(0)} = 1, \frac{1 - \theta^{(k+1)}}{(\theta^{(k+1)})^2} = \frac{1}{(\theta^{(k)})^2} \) and \( \theta^{(k)} > 0 \). Then we have the following properties

\[
\theta^{(k+1)} = \frac{-(\theta^{(k)})^2 + \sqrt{(\theta^{(k)})^4 + 4(\theta^{(k)})^2}}{2}, \tag{5}
\]

*Corresponding author.
\[
\sum_{k=0}^{K} \frac{1}{\theta^{(k)}} = \frac{1}{(\theta^{(K)})^2},
\]
(6)

\[
0 < \frac{1}{\theta^{(k+1)}} - \frac{1}{\theta^{(k)}} < 1,
\]
(7)

\[
\theta^{(k)} \leq \frac{2}{k + 2},
\]
(8)

and
\[
\theta^{(k)} \leq 1.
\]
(9)

Proof. From the definition of \(\theta^{(k+1)}\), it is easy to get that
\[
\theta^{(k+1)} = -\frac{(\theta^{(k)})^2 + \sqrt{(\theta^{(k)})^4 + 4(\theta^{(k)})^2}}{2}.
\]
This implies that \(\theta^{(k)}\) is well defined for any \(k \geq 0\). Furthermore, since
\[
\frac{1}{\theta^{(k+1)}} = \frac{1}{(\theta^{(k)})^2} - \frac{1}{(\theta^{(k)})^2} \quad \text{and} \quad \theta^{(0)} = 1,
\]
we have
\[
\sum_{k=0}^{K} \frac{1}{\theta^{(k)}} = \frac{1}{\theta^{(0)}/(\theta^{(k)})^2} - \frac{1}{\theta^{(0)}/(\theta^{(k)})^2} = \frac{1}{\theta^{(k)}},
\]
(10)

From \(\frac{1}{\theta^{(k+1)}} = \frac{1}{(\theta^{(k)})^2} - \frac{1}{(\theta^{(k)})^2}\), \(\theta^{(k)} > 0\) and \(\theta^{(k-1)} > 0\), we can easily get
\[
\frac{1}{\theta^{(k+1)}} - \frac{1}{\theta^{(k)}} > 0,
\]
(11)

and
\[
\frac{1}{\theta^{(k+1)}} - \frac{1}{\theta^{(k)}} = \frac{1}{\theta^{(k+1)}} - \frac{\sqrt{1 - \theta^{(k+1)}}}{\theta^{(k+1)}} = \frac{1 - \sqrt{1 - \theta^{(k+1)}}}{\theta^{(k+1)}} = \frac{1}{1 + \sqrt{1 - \theta^{(k+1)}}} < 1.
\]
(12)

Next we proof \(\theta^{(k)} \leq \frac{2}{k + 2}\) by induction. First \(\theta^{(0)} = 1 \leq \frac{2}{0 + 2}\). Now assume that \(\theta^{(k)} \leq \frac{2}{k + 2}\) and we prove \(\theta^{(k+1)} \leq \frac{2}{k + 3}\). We deduce
\[
\theta^{(k+1)} = -\frac{(\theta^{(k)})^2 + \sqrt{(\theta^{(k)})^4 + 4(\theta^{(k)})^2}}{2} = \frac{2(\theta^{(k)})^2}{(\theta^{(k)})^2 + \sqrt{(\theta^{(k)})^4 + 4(\theta^{(k)})^2}}
\]
\[
\leq \frac{2}{1 + \sqrt{1 + \frac{4}{(\theta^{(k)})^2}}} \leq \frac{2}{1 + \sqrt{1 + (k + 2)^2}} \leq \frac{2}{k + 3}.
\]

So (8) holds. Note that \(\theta^{(k)}\) is decreasing by (7) and \(\theta^{(0)} = 1\), we have (9). The proof is completed. \(\blacksquare\)

2. Convergence Analysis of Fast PALM

In this section, we give the convergence analysis of Fast PALM for solving the following problem
\[
\min_{x} \ f(x), \quad s.t. \quad A(x) = b,
\]
(13)

where \(f(x) = g(x) + h(x)\), both \(g\) and \(h\) are convex, and \(g \in C^{1,1}\):
\[
\|\nabla g(x) - \nabla g(y)\| \leq L \|x - y\|, \quad \forall x, y.
\]
(14)

For the completeness, we give the Fast PALM in Algorithm 1.

It is worth mentioning that the definition of \(\theta^{(k+1)}\) in (20) is equivalent to \(\theta^{(0)} = 1\), \(\frac{1 - \theta^{(k+1)}}{(\theta^{(k)})^2} = \frac{1}{(\theta^{(k)})^2}\) and \(\theta^{(k)} > 0\) in Lemma 4. Such a property will be used in the following analysis several times.

The analysis of our algorithms is based on the following property:

Lemma 5 [1] \(\tilde{x}\) is an optimal solution to (13) if and only if there exists \(\alpha > 0\), such that
\[
f(\tilde{x}) - f(x^*) + \langle \lambda^*, A(\tilde{x}) - b \rangle + \frac{\alpha}{2} \|A(\tilde{x}) - b\|^2 = 0.
\]
(15)
Initialize: \( x^0, z^0, \lambda^0, \beta(0) = \theta(0) = 1 \).

for \( k = 0, 1, 2, \cdots \) do

\[
\begin{align*}
y^{k+1} &= (1 - \theta(k))x^k + \theta(k)z^k; \\
z^{k+1} &= \arg\min_x g(y^{k+1}) + \langle \nabla g(y^{k+1}), x - y^{k+1} \rangle + h(x) \\
&\quad + \langle \lambda^k, A(x) - b \rangle + \frac{\beta(k)}{2} \|A(x) - b\|^2 + \frac{L\theta(k)}{2} \|x - z^k\|^2; \\
x^{k+1} &= (1 - \theta(k))x^k + \theta(k)z^{k+1}; \\
\lambda^{k+1} &= \lambda^k + \beta(k)(A(z^{k+1}) - b); \\
\theta(k+1) &= \frac{-(\theta(k))^2 + \sqrt{(\theta(k))^4 + 4(\theta(k))^2}}{2}; \\
\beta(k+1) &= \frac{1}{\theta(k+1)}.
\end{align*}
\]

end

Algorithm 1: Fast PALM Algorithm

Proposition 1 In Algorithm 1, for any \( x \), we have

\[
\frac{1 - \theta(k+1)}{(\theta(k+1))^2} \left( f(x^{k+1}) - f(x) \right) - \frac{1}{\theta(k)} \langle A^T(\lambda^{k+1}), x - z^{k+1} \rangle \\
\leq \frac{1 - \theta(k)}{(\theta(k))^2} \left( f(x^{k}) - f(x) \right) + \frac{L}{2} \left( \|z^k - x\|^2 - \|z^{k+1} - x\|^2 \right).
\]

Proof. From the optimality of \( z^{k+1} \) to (17), we have

\[
0 \in \partial h(z^{k+1}) + \nabla g(y^{k+1}) + A^T(\lambda^k) + \beta(k)A^T(A(z^{k+1}) - b) + L\theta(k)(z^{k+1} - z^k) \\
= \partial h(z^{k+1}) + \nabla g(y^{k+1}) + A^T(\lambda^{k+1}) + L\theta(k)(z^{k+1} - z^k),
\]

where (23) uses (19). From the convexity of \( h \), we have

\[
b(x) - b(z^{k+1}) \geq \langle -\nabla g(y^{k+1}) - A^T(\lambda^{k+1}) - L\theta(k)(z^{k+1} - z^k), x - z^{k+1} \rangle.
\]
On the other hand,

\[ f(x^{k+1}) \leq g(y^{k+1}) + \langle \nabla g(y^{k+1}), x^{k+1} - y^{k+1} \rangle + \frac{L}{2} \| x^{k+1} - y^{k+1} \|^2 + h(x^{k+1}) \]  
(25)

\[ = g(y^{k+1}) + \langle \nabla g(y^{k+1}), (1 - \theta(k))x^k + \theta(k)z^{k+1} - y^{k+1} \rangle + \frac{L}{2} \| (1 - \theta(k))x^k + \theta(k)z^{k+1} - y^{k+1} \|^2 + h \left( (1 - \theta(k))x^k + \theta(k)z^{k+1} \right) \]  
(26)

\[ \leq (1 - \theta(k)) \left( g(y^{k+1}) + \langle \nabla g(y^{k+1}), x^k - y^{k+1} \rangle + h(x^k) \right) + \frac{L(\theta(k))^2}{2} \| z^{k+1} - z^k \|^2 \]  
(27)

\[ = (1 - \theta(k)) f(x^k) + \theta(k) \left( g(x) + \langle \nabla g(y^{k+1}), z^{k+1} - x \rangle + h(z^{k+1}) \right) + \frac{L(\theta(k))^2}{2} \| z^{k+1} - z^k \|^2 \]  
(28)

\[ \leq (1 - \theta(k)) f(x^k) + \theta(k) \left( g(x) + h(x) + \langle A^T(\lambda^{k+1}) + L\theta(k)(z^{k+1} - z^k), x - z^{k+1} \rangle \right) + \frac{L(\theta(k))^2}{2} \| z^{k+1} - z^k \|^2 \]  
(29)

\[ = (1 - \theta(k)) f(x^k) + \theta(k) f(x) + \theta(k) \langle A^T(\lambda^{k+1}), x - z^{k+1} \rangle - \frac{L(\theta(k))^2}{2} \left( \| z^{k+1} - x \|^2 - \| z^k - x \|^2 \right) \]  
(30)

where (25) uses (1), (26) uses (18), (27) is from the convexity of $h$, (28) is from the convexity of $g$, (29) uses (24) and (30) uses (2). Reranging the above inequality leads to

\[ (f(x^{k+1}) - f(x)) - \theta(k) \langle A^T(\lambda^{k+1}), x - z^{k+1} \rangle \leq (1 - \theta(k)) (f(x^k) - f(x)) + \frac{L(\theta(k))^2}{2} \left( \| z^k - x \|^2 - \| z^{k+1} - x \|^2 \right). \]  
(31)

Diving both sides of the above inequality by $(\theta(k))^2$ leads to

\[ \frac{1}{\theta(k)^2} (f(x^{k+1}) - f(x)) - \frac{1}{\theta(k)} \langle A^T(\lambda^{k+1}), x - z^{k+1} \rangle \leq \frac{1 - \theta(k)}{(\theta(k))^2} (f(x^k) - f(x)) + \frac{L}{2} \left( \| z^k - x \|^2 - \| z^{k+1} - x \|^2 \right). \]  

The proof is completed by using the property of $\theta(k)$ in Lemma 4. \qed

**Proposition 2** In Algorithm 1, the following result holds for any $\lambda$

\[ \langle A(z^{k+1}) - b, \lambda - \lambda^{k+1} \rangle + \frac{\beta(k)}{2} \| A(z^{k+1}) - b \|^2 \]

\[ = \frac{1}{2\beta(k)} \left( \| \lambda^k - \lambda \|^2 - \| \lambda^{k+1} - \lambda \|^2 \right) \]  
(32)

**Proof.** By using (19) and (2), we have

\[ \langle A(z^{k+1}) - b, \lambda - \lambda^{k+1} \rangle \]

\[ = \frac{1}{\beta(k)} \langle \lambda^{k+1} - \lambda, \lambda - \lambda^{k+1} \rangle \]

\[ = \frac{1}{2\beta(k)} \left( \| \lambda^k - \lambda \|^2 - \| \lambda^{k+1} - \lambda \|^2 - \| \lambda^{k+1} - \lambda^k \|^2 \right) \]

\[ = \frac{1}{2\beta(k)} \left( \| \lambda^k - \lambda \|^2 - \| \lambda^{k+1} - \lambda \|^2 \right) - \frac{\beta(k)}{2} \| A(z^{k+1}) - b \|^2. \]
The proof is completed. ■

Theorem 1 In Algorithm 1, for any $K > 0$, we have
\[
f(x^{K+1}) - f(x^*) + \langle \lambda^*, A(x^{K+1}) - b \rangle + \frac{1}{2} \|A(x^{K+1}) - b\|^2 \leq \frac{2}{(K+2)^2} (LD^2_{\lambda,*} + D^2_{\lambda,*}).
\] (34)

\section{Proof}
Let $x = x^*$ and $\lambda = \lambda^*$ in (22) and (32). We have
\[
1 - \frac{\theta^{(k+1)}}{(\theta^{(k+1)})^2} (f(x^{k+1}) - f(x^*)) - \frac{1 - \theta^{(k)}}{(\theta^{(k)})^2} (f(x^k) - f(x^*)) + \frac{1}{\theta^{(k)}} \langle \lambda^*, A(z^{k+1}) - b \rangle
\] (35)
\[
\leq \frac{L}{2} \left( \|z^k - x^*\|^2 - \|z^{k+1} - x^*\|^2 \right) + \frac{1}{\theta^{(k)}} \langle \lambda^* - \lambda^{k+1}, A(z^{k+1}) - b \rangle
\] (36)
\[
\leq \frac{L}{2} \left( \|z^k - x^*\|^2 - \|z^{k+1} - x^*\|^2 \right) + \frac{1}{2\theta^{(k)}\beta^{(k)}} \left( \|\lambda^k - \lambda^{k+1}\|^2 - \|\lambda^{k+1} - \lambda^*\|^2 \right) - \frac{\beta^{(k)}}{2\theta^{(k)}\beta^{(k)}} \|A(z^{k+1}) - b\|^2
\] (37)
\[
= \frac{L}{2} \left( \|z^k - x^*\|^2 - \|z^{k+1} - x^*\|^2 \right) + \frac{1}{2} \left( \|\lambda^k - \lambda^{k+1}\|^2 - \|\lambda^{k+1} - \lambda^*\|^2 \right) - \frac{1}{2\theta^{(k)}\beta^{(k)}} \|A(z^{k+1}) - b\|^2
\] (38)

where (36) uses the fact $A(x^*) = b$, (37) uses (32) and (38) uses $\beta^{(k)} = \frac{1}{\theta^{(k)}}$.

Summing (35)-(38) from $k = 0$ to $K$, we have
\[
\frac{1}{\theta^{(K+1)}} \sum_{k=0}^{K} \frac{1 - \theta^{(k)}}{\theta^{(k)}(\theta^{(k)})^2} \left( f(x^{k+1}) - f(x^*) \right) - \frac{1 - \theta^{(0)}}{\theta^{(0)}(\theta^{(0)})^2} \left( f(x^{0}) - f(x^*) \right) + \sum_{k=0}^{K} \frac{1}{\theta^{(k)}} \langle \lambda^*, A(z^{k+1}) - b \rangle
\] (39)
\[
\leq \frac{L}{2} \sum_{k=0}^{K} \|z^0 - x^*\|^2 + \frac{1}{2} \sum_{k=0}^{K} \|\lambda^k - \lambda^{k+1}\|^2 \|A(z^{k+1}) - b\|^2
\] (40)

where (40) uses (9). Also note that $\theta^{(0)} = 1$. So the second term of (39) disappears.

On the other hand, by the property of $\theta^{(k)}$ in Lemma 4 and (18), we have
\[
\sum_{k=0}^{K} \frac{z^{k+1}}{\theta^{(k)}} = \sum_{k=0}^{K} \left( \frac{1}{\theta^{(k)}(\theta^{(k)})^2} x^{k+1} - \frac{1 - \theta^{(k)}}{\theta^{(k)}(\theta^{(k)})^2} x^k \right)
\]
\[
= \sum_{k=0}^{K} \left( \frac{1 - \theta^{(k+1)}}{\theta^{(k+1)}(\theta^{(k+1)})^2} x^{k+1} - \frac{1 - \theta^{(k)}}{\theta^{(k)}(\theta^{(k)})^2} x^k \right)
\]
\[
= \frac{1 - \theta^{(K+1)}}{(\theta^{(K+1)})^2} x^{K+1} - \frac{1 - \theta^{(0)}}{(\theta^{(0)})^2} x^0
\]
\[
= \frac{1 - \theta^{(K+1)}}{(\theta^{(K+1)})^2} x^{K+1} - \frac{1}{\theta^{(K)^2}} x^{K+1}
\]
\[
= \frac{1}{\theta^{(K)^2}} x^{K+1}.
\] (41)

So
\[
\sum_{k=0}^{K} \frac{1}{\theta^{(k)}} \langle \lambda^*, A(z^{k+1}) - b \rangle = \frac{1}{\theta^{(K)^2}} \langle \lambda^*, A(x^{K+1}) - b \rangle.
\] (42)
By the convexity of \( \| \cdot \|^2 \), we have

\[
\sum_{k=0}^{K} \frac{1}{2\theta(k)} \|A(z^{k+1}) - b\|^2 = \frac{1}{2(\theta(K))^2} \sum_{k=0}^{K} \theta(k)^2 \|A(z^{k+1}) - b\|^2 \\
\geq \frac{1}{2(\theta(K))^2} \|A(x^{K+1}) - b\|^2,
\]

where (43) uses (6) and (44) uses (41).

Substituting (42) into (39) and (44) into (40) respectively, we obtain

\[
\frac{1 - \theta(k+1)}{(\theta(K+1))^2} (f(x^{k+1}) - f(x^*)) + \frac{1}{(\theta(K))^2} \langle A^*, A(x^{K+1}) - b \rangle + \frac{1}{2(\theta(K))^2} \|A(x^{K+1}) - b\|^2 \leq \frac{L}{2}\|z^0 - x^*\|^2 + \frac{1}{2}\|\lambda^0 - \lambda^*\|^2.
\]

(46)

Multiplying (45) and (46) by \((\theta(K))^2\) and using (8) leads to

\[
f(x^{K+1}) - f(x^*) + \langle \lambda^*, A(x^{K+1}) - b \rangle + \frac{1}{2}\|A(x^{K+1}) - b\|^2 \leq \frac{2}{(K + 2)^2} (L\|z^0 - x^*\|^2 + \|\lambda^0 - \lambda^*\|^2) \\
= \frac{2}{(K + 2)^2} (L\lambda^2 + D_{\lambda^2}).
\]

The proof is completed.

\[\blacksquare\]

3. Convergence Analysis of Fast PL-ADMM-PS

In this section, we give the convergence analysis of Fast PL-ADMM-PS for solving the following problem

\[
\min_{x_1, \ldots, x_n} \sum_{i=1}^{n} f_i(x_i), \quad s.t. \quad \sum_{i=1}^{n} A_i(x_i) = b, \quad (47)
\]

where \(f_i(x_i) = g_i(x_i) + h_i(x_i)\), both \(g_i\) and \(h_i\) are convex, and \(g_i \in C^{1,1}\). The whole procedure of Fast PL-ADMM-PS is shown in Algorithm 2.

**Proposition 3** In Algorithm 2, for any \(x_i\), we have

\[
\frac{1 - \theta(k+1)}{(\theta(K+1))^2} (f_i(x_i^{k+1}) - f_i(x_i)) - \frac{1}{\theta(k)} \langle A_i^T(\lambda^{k+1}), x_i - z_i^{k+1} \rangle \\
\leq \frac{1 - \theta(k)}{(\theta(K))^2} (f_i(x_i^k) - f_i(x_i)) + \frac{L_i}{2} (\|z_i^k - x_i\|^2 - \|z_i^{k+1} - x_i\|^2) \\
+ \frac{\beta(k) \eta_i}{2(\theta(K))} (\|z_i^k - x_i\|^2 - \|z_i^{k+1} - x_i\|^2 - \|z_i^{k+1} - z_i^k\|^2),
\]

where

\[
\lambda^{k+1} = \lambda^k + \beta(k) (A(z^k) - b).
\]

(53)
On the other hand, from the optimality of $z_i^{k+1}$ to (49), we have

$$0 \in \partial h_i(z_i^{k+1}) + \nabla g_i(y_i^{k+1}) + A_i^T (\lambda^k) + \beta(k) A_i^T (A(z^k) - b) + (L_t \theta(k) + \beta(k) \eta_t) (z_i^{k+1} - z_i^k)$$

$$= \partial h_i(z_i^{k+1}) + \nabla g_i(y_i^{k+1}) + A_i^T (\lambda^{k+1}) + (L_t \theta(k) + \beta(k) \eta_t) (z_i^{k+1} - z_i^k),$$

where (55) uses (54). From the convexity of $h_i$, we have

$$h_i(x_i) - h_i(z_i^{k+1}) \geq -\langle \nabla g_i(y_i^{k+1}) - A_i^T (\lambda^{k+1}) - (L_t \theta(k) + \beta(k) \eta_t) (z_i^{k+1} - z_i^k), x_i - z_i^{k+1} \rangle.$$  

On the other hand,

$$f_i(x_i^{k+1}) \leq g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), x_i^{k+1} - y_i^{k+1} \rangle + \frac{L_t}{2} \| x_i^{k+1} - y_i^{k+1} \|^2 + h_i(x_i^{k+1})$$

$$= g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), (1 - \theta(k)) x_i^k + \theta(k) z_i^{k+1} - y_i^{k+1} \rangle$$

$$+ \frac{L_t}{2} \| (1 - \theta(k)) x_i^k + \theta(k) z_i^{k+1} - y_i^{k+1} \|^2 + h_i \left( (1 - \theta(k)) x_i^k + \theta(k) z_i^{k+1} \right)$$

$$\leq \left( 1 - \theta(k) \right) \left( g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), x_i - y_i^{k+1} \rangle + h_i(x_i^k) \right)$$

$$+ \theta(k) \left( g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), z_i^{k+1} - y_i^{k+1} \rangle + h_i(z_i^{k+1}) \right) + \frac{L_t (\theta(k))^2}{2} \| z_i^{k+1} - z_i^k \|^2$$

$$= \left( 1 - \theta(k) \right) \left( g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), x_i - y_i^{k+1} \rangle + h_i(x_i^k) \right)$$

$$+ \theta(k) \left( g_i(y_i^{k+1}) + \langle \nabla g_i(y_i^{k+1}), x_i - y_i^{k+1} \rangle + \langle \nabla g_i(y_i^{k+1}), z_i^{k+1} - x_i \rangle + h_i(z_i^{k+1}) \right)$$

$$+ \frac{L_t (\theta(k))^2}{2} \| z_i^{k+1} - z_i^k \|^2$$

$$\leq \left( 1 - \theta(k) \right) f_i(x_i^k) + \theta(k) \left( g_i(x_i) + \langle \nabla g_i(y_i^{k+1}), z_i^{k+1} - x_i \rangle + h_i(z_i^{k+1}) \right) + \frac{L_t (\theta(k))^2}{2} \| z_i^{k+1} - z_i^k \|^2$$

$$\leq \left( 1 - \theta(k) \right) f_i(x_i^k) + \theta(k) \left( g_i(x_i) + h_i(x_i) + \langle A_i^T (\lambda^{k+1}) + (L_t \theta(k) + \beta(k) \eta_t) (z_i^{k+1} - z_i^k), x_i - z_i^{k+1} \rangle \right)$$

$$+ \frac{L_t (\theta(k))^2}{2} \| z_i^{k+1} - z_i^k \|^2$$

$$= \left( 1 - \theta(k) \right) f_i(x_i^k) + \theta(k) f_i(x_i) + \theta(k) \langle A_i^T (\lambda^{k+1}), x_i - z_i^{k+1} \rangle$$

$$- \frac{L_t (\theta(k))^2 + \theta(k) \beta(k) \eta_t}{2} \left( \| z_i^{k+1} - x_i \|^2 - \| z_i^k - x_i \|^2 + \| z_i^{k+1} - z_i^k \|^2 \right) + \frac{L_t (\theta(k))^2}{2} \| z_i^{k+1} - z_i^k \|^2,$$
where (57) uses (1), (58) uses (50), (59) is from the convexity of \( h_i \), (61) is from the convexity of \( g_i \), (62) uses (56) and (63) uses (3). Rearranging the above inequality leads to

\[
(f_i(x_i^{k+1}) - f_i(x_i)) - \theta^{(k)} \left< A_i^T(\lambda^{k+1}), x_i - z_i^{k+1} \right> \\
\leq (1 - \theta^{(k)}) (f_i(x_i^k) - f_i(x_i)) + \frac{L_i (\theta^{(k)})^2}{2} (\| z_i^k - x_i \|^2 - \| z_i^{k+1} - x_i \|^2) \\
+ \frac{\theta^{(k)} \beta^{(k)} \eta_i}{2} (\| z_i^k - x_i \|^2 - \| z_i^{k+1} - x_i \|^2 - \| z_i^{k+1} - z_i^k \|^2)
\]

(64)

Dividing both sides of the above inequality by \((\theta^{(k)})^2\) leads to

\[
\frac{1}{(\theta^{(k)})^2} (f_i(x_i^{k+1}) - f_i(x_i)) - \frac{1}{\theta^{(k)}} \left< A_i^T(\lambda^{k+1}), x_i - z_i^{k+1} \right> \\
\leq \frac{1 - \theta^{(k)}}{(\theta^{(k)})^2} (f_i(x_i^k) - f_i(x_i)) + \frac{L_i}{2} (\| z_i^k - x_i \|^2 - \| z_i^{k+1} - x_i \|^2) \\
+ \frac{\beta^{(k)} \eta_i}{2\theta^{(k)}} (\| z_i^k - x_i \|^2 - \| z_i^{k+1} - x_i \|^2 - \| z_i^{k+1} - z_i^k \|^2)
\]

The proof is completed by using \( \frac{1 - \theta^{(k+1)}}{(\theta^{(k+1)})^2} = \frac{1}{(\theta^{(k)})^2} \).

\[\Box\]

**Proposition 4** In Algorithm 2, the following result holds for any \( \lambda \)

\[
\left< A(z^{k+1}) - b, \lambda - \hat{\lambda}^{k+1} \right> + \frac{\beta^{(k)} \alpha}{2} \| A(z^{k+1}) - b \|^2 \\
\leq \frac{1}{2\beta^{(k)}} (\| \lambda^k - \lambda \|^2 - \| \lambda^{k+1} - \lambda \|^2) + \frac{\beta^{(k)}}{2} \sum_{i=1}^n \eta_i \| z_i^{k+1} - z_i^k \|^2,
\]

(65)

where \( \alpha = \min \left\{ \frac{1}{n+1}, \left\{ \frac{n_i - n_i \| A_i \|^2}{(n+1)\| A_i \|^2}, i = 1, \ldots, n \right\} \right\} \).

**Proof.** Using (51) and (3), we have

\[
\left< A(z^{k+1}) - b, \lambda - \hat{\lambda}^{k+1} \right> \\
= \frac{1}{\beta^{(k)}} (\lambda^{k+1} - \lambda^k, \lambda - \hat{\lambda}^{k+1}) \\
= \frac{1}{2\beta^{(k)}} (\| \lambda^k - \hat{\lambda}^{k+1} \|^2 - \| \lambda - \lambda^{k+1} \|^2) - \frac{1}{2\beta^{(k)}} (\| \lambda^{k+1} - \lambda \|^2 - \| \lambda^{k+1} - \hat{\lambda}^{k+1} \|^2),
\]

(66)
Now, consider the last two terms in the above inequality. We deduce
\[
\begin{align*}
\frac{1}{2\beta(k)} \left( ||\hat{\lambda}^{k+1} - \lambda^k||^2 - ||\lambda^{k+1} - \hat{\lambda}^{k+1}||^2 \right) \\
= \frac{\beta(k)}{2} \left( \left\| \sum_{i=1}^{n} A_i(z_i^k) - b \right\|^2 - \left\| \sum_{i=1}^{n} A_i(z_i^{k+1} - z_i^k) \right\|^2 \right) \\
\geq \frac{\beta(k)}{2} \left( \left\| \sum_{i=1}^{n} A_i(z_i^k) - b \right\|^2 - \sum_{i=1}^{n} \|A_i\|^2 ||z_i^{k+1} - z_i^k||^2 \right) \\
= \frac{\beta(k)}{2} \left( \left\| \sum_{i=1}^{n} A_i(z_i^k) - b \right\|^2 + \sum_{i=1}^{n} \eta_i \|A_i\|^2 \|A_i z_i^{k+1} - z_i^k\|^2 - \sum_{i=1}^{n} \eta_i ||z_i^{k+1} - z_i^k||^2 \right) \\
\geq \frac{\beta(k)}{2} \left( \alpha(n+1) \left( \left\| \sum_{i=1}^{n} A_i(z_i^k) - b \right\|^2 + \sum_{i=1}^{n} \|A_i\|^2 \|A_i z_i^{k+1} - z_i^k\|^2 \right) - \sum_{i=1}^{n} \eta_i ||z_i^{k+1} - z_i^k||^2 \right) \\
= \frac{\beta(k)}{2} \alpha \|A(z^{k+1}) - b\|^2 - \frac{\beta(k)}{2} \sum_{i=1}^{n} \eta_i ||z_i^{k+1} - z_i^k||^2 
\end{align*}
\]
(67)

(68)
where (71) uses the fact \( A(x^*) = b \) and (72) uses (65).

Summing (70)-(73) from \( k = 0 \) to \( K \) and fixing \( \beta^{(k)} = \beta > 0 \), we have

\[
\begin{align*}
\frac{1 - \theta^{(K+1)}}{(\theta^{(K+1)})^2} & \left( \sum_{i=1}^n (f_i(x_i^{K+1}) - f_i(x_i^*)) \right) - \frac{1 - \theta^{(0)}}{(\theta^{(0)})^2} \left( \sum_{i=1}^n (f_i(x_i^0) - f_i(x_i^*)) \right) + \sum_{k=0}^K \frac{1}{\theta^{(k)}} \left( \lambda^*, A(z^{k+1}) - b \right) \\
\leq & \frac{1}{2} \sum_{i=1}^n L_i \|z_i^0 - x_i^*\|^2 + \frac{1}{2} \sum_{k=0}^K \sum_{i=1}^n \beta \eta_k \left( z_i^k - x_i^* \right)^2 \quad \text{for all } 1 \leq \eta_k \leq \infty \\
& + \sum_{i=1}^n K \beta \eta_{\max} D_X^2 + \frac{1}{\beta} K D_A - \beta \alpha \sum_{k=0}^K \frac{1}{\theta^{(k)}} \|A(z^{k+1}) - b\|^2 \\
\leq & \frac{1}{2} \left( L_{\max} D_X^2 + K \beta \eta_{\max} D_X^2 + \frac{1}{\beta} K D_A - \beta \alpha \sum_{k=0}^K \frac{1}{\theta^{(k)}} \|A(z^{k+1}) - b\|^2 \right) \\
\end{align*}
\]

(74)

(75)

(76)

where (75) uses (4). Also note that \( \theta^{(0)} = 1 \). So the second term of (74) disappears.

Note that (42) and (44) also hold here. Substituting (42) into (74) and (44) into (76) respectively and using \( \frac{1 - \theta^{(K+1)}}{(\theta^{(K+1)})^2} = \frac{1 - \theta^{(0)}}{(\theta^{(0)})^2} \), we obtain

\[
\begin{align*}
\frac{1}{(\theta^{(K)})^2} & \left( \sum_{i=1}^n (f_i(x_i^{K+1}) - f_i(x_i^*)) \right) + \frac{1}{(\theta^{(0)})^2} \left( \lambda^*, A(x^{K+1}) - b \right) + \frac{\beta \alpha}{2(\theta^{(k)})^2} \|A(x^{K+1}) - b\|^2 \\
\leq & \frac{1}{2} \left( L_{\max} D_X^2 + K \beta \eta_{\max} D_X^2 + \frac{1}{\beta} K D_A \right).
\end{align*}
\]

The proof is completed by multiplying both sides of the above inequality with \( \theta^{(K)} \) and using (8). \( \square \)

**Theorem 3** Assume the mapping \( A(x_1, \ldots, x_n) = \sum_{i=1}^n A_i(x_i) \) is onto\(^1 \), the sequence \( \{z^k\} \) is bounded, \( \partial h(x) \) and \( \nabla g(x) \) are bounded if \( x \) is bounded, then \( \{x^k\}, \{y^k\} \) and \( \{\lambda^k\} \) are bounded.

**Proof.** Assume \( \|z^k\| \leq C_1 \) for all \( k \) and \( \|x^0\| \leq C_1 \). Then from (50) we can easily get \( \|x^k\| \leq C_1 \) for all \( k \). Then from (48) we have \( \|y^k\| \leq C_1 \) for all \( k \). Assume \( \|\partial h(x)\| \leq C_2 \) and \( \|\nabla g(x)\| \leq C_2 \) if \( \|x\| \leq C_1 \). Then from (55), we have

\[
0 \in \partial h(z^{k+1}) + \nabla g(y^{k+1}) + A^T(\lambda^k) + \beta^{(k)} A^T(A(z^k) - b) + \left[ (L_1 \theta^{(k)} + \beta^{(k)} \eta_1)(z_1^{k+1} - z_1^k) \right] \quad \text{for all } 1 \leq \eta_1 \leq \infty \\
\vdots \\
\left[ (L_n \theta^{(k)} + \beta^{(k)} \eta_n)(z_n^{k+1} - z_n^k) \right],
\]

and

\[
-A A^T(\lambda^k) \in A \left( \partial h(z^{k+1}) + \nabla g(y^{k+1}) + \beta^{(k)} A^T(A(z^k) - b) + \left[ (L_1 \theta^{(k)} + \beta^{(k)} \eta_1)(z_1^{k+1} - z_1^k) \right] \quad \text{for all } 1 \leq \eta_1 \leq \infty \\
\vdots \\
\left[ (L_n \theta^{(k)} + \beta^{(k)} \eta_n)(z_n^{k+1} - z_n^k) \right].
\]

\(^1\)This assumption is equivalent to that the matrix \( A \equiv (A_1, \ldots, A_n) \) is of full row rank, where \( A_i \) is the matrix representation of \( A_i \).
So we have

\[
\|\lambda^k\| \leq \left\| (AA^T)^{-1}A \left( \partial h(z^{k+1}) + \nabla g(y^{k+1}) + \beta^k A^T (Az^k - b) + \begin{bmatrix}
(L_1 \theta(k) + \beta^k \eta_1)(z_1^{k+1} - z_1^k)
\vdots
(L_i \theta(k) + \beta^k \eta_i)(z_i^{k+1} - z_i^k)
\vdots
(L_n \theta(k) + \beta^k \eta_n)(z_n^{k+1} - z_n^k)
\end{bmatrix}
\right) \right\| \\
\leq \| (AA^T)^{-1}A \| \left( \| \partial h(z^{k+1}) \| + \| \nabla g(y^{k+1}) \| + \| \beta^k A^T Az^k \| + \| \beta^k A^T b \| + (L_{\max} \theta(k) + \beta^k \eta_{\max}) (\| z^{k+1} \| + \| z^k \|) \right) \\
\leq \| (AA^T)^{-1}A \| \left( 2C_2 + \beta^k \| A^T A \| C_1 + \beta^k \| A^T b \| + 2(L_{\max} \theta(k) + \beta^k \eta_{\max})C_1 \right). 
\]

for all \( k \).

\[ \Box \]

References
