First of all, we have the following observations: In algorithm RPDC, the indices \( i(k) \), \( k = 0, 1, 2, \ldots \) are random variables. After \( k \) iterations, RPDC method generates a random output \((u^{k+1}, p^{k+1})\). Recall the definition of filtration \( \mathcal{F}_k \) which is generated by the random variable \( i(0), i(1), \ldots, i(k) \), i.e.,

\[
\mathcal{F}_k = \{i(0), i(1), \ldots, i(k)\}, \mathcal{F}_k \subset \mathcal{F}_{k+1}.
\]

Additionally, \( \mathcal{F} = (\mathcal{F}_k)_{k\in \mathbb{N}}, \mathbb{E} = \mathbb{E}(|\mathcal{F}_k|) \) is the conditional expectation w.r.t. \( \mathcal{F}_k \) and the conditional expectation in term of \( i(k) \) given \( i(0), i(1), \ldots, i(k-1) \) as \( \mathbb{E}_{i(k)} \).

Knowing \( \mathcal{F}_{k-1} = \{i(0), i(1), \ldots, i(k-1)\} \), we have:

\[
\mathbb{E}_{i(k)}[\langle \nabla_i G(u^k), (u^k - u)_{i(k)} \rangle] = \frac{1}{N} \langle \nabla G(u^k), u^k - u \rangle \\
\geq \frac{1}{N} \langle G(u^k) - G(u) \rangle,
\]

\[
A.1
\]

\[
\mathbb{E}_{i(k)}[J_i(k)(u_{i(k)}^k) - J_i(k)(u_{i(k)})] = \frac{1}{N} [J(u^k) - J(u)],
\]

\[
A.2
\]

and

\[
\mathbb{E}_{i(k)}[\langle q^k, A_i(k)(u^k - u)_{i(k)} \rangle] = \frac{1}{N} \langle q^k, A(u^k - u) \rangle.
\]

\[
A.3
\]

Secondly, reconsidering the point \( T(w^k) = (T_u(w^k), T_p(w^k)) \) generated by one deterministic iteration of APP-AL (Cohen & Zhu, 1984) for given \( w^k \),

\[
\begin{align*}
T_u(w^k) &= \arg\min_{u \in \mathcal{U}} \langle \nabla G(u^k), u \rangle + J(u) + \langle q^k, Au \rangle \\
T_p(w^k) &= p^k + \gamma [AT_u(w^k) - b],
\end{align*}
\]

with \( q^k = p^k + \gamma (A u^k - b) \), we have the following observations. The convex combination of \( u^k \) and \( T_u(w^k) \) provides the expected value of \( u^{k+1} \) as following.

\[
\mathbb{E}_{i(k)} u^{k+1} = \frac{1}{N} T_u(w^k) + (1 - \frac{1}{N}) u^k,
\]

\[
A.4
\]

Moreover, the point \( T(w^k) \) satisfies that: for any \( (u, p) \in \mathcal{U} \times \mathbb{R}^m \),

\[
\begin{align*}
\langle \nabla G(w^k), u - T_u(w^k) \rangle + J(u) - J(T_u(w^k)) \\
+ \langle q^k, A(u - T_u(w^k)) \rangle \\
+ \frac{1}{2} \langle \nabla K(T_u(w^k)) - \nabla K(u^k), u - T_u(w^k) \rangle \geq 0,
\end{align*}
\]

\[
A.6
\]
1. Proof of Lemma 1

Proof. Take \( w' = w^* \) in (9), we have that
\[
\Lambda(w, w^*) = D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 \\
+ \frac{\epsilon(N-1)}{N} [L(u, p) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2
\]
\[
= D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 \\
+ \frac{\epsilon(N-1)}{N} [L(u, p^*) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-1)\beta}{2N\gamma} \| p - p^* \|^2 + \frac{\gamma}{2} \| Au - b \|^2 \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2.
\] (A.7)

(i) Since \( L(u, p^*) - L(u^*, p^*) \geq 0 \) and \( \frac{1}{2\gamma} \| p - p^* \|^2 + \frac{\gamma}{2} \| Au - b \|^2 + \langle p - p^*, Au - b \rangle \geq 0 \), (A.7) follows that
\[
\Lambda(w, w^*) \geq D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 \\
- \frac{\epsilon(N-1)}{2N\gamma} \| p - p^* \|^2 - \frac{\epsilon\gamma}{2N} \| Au - b \|^2.
\]

From Assumption 2, we have \( D(u^*, u) \geq \frac{\beta}{2} \| u - u^* \|^2 \). Together with the fact \( Au^* = b \) and \( \rho < \frac{2N\gamma}{2N-1} \), above inequality follows that
\[
\Lambda(w, w^*) \geq d_1 \| w - w^* \|^2,
\]
with \( d_1 = \min \left\{ \frac{1}{2N\gamma} [N\beta - \epsilon\gamma \lambda_{\max}(A^TA)], \frac{\epsilon}{4N\gamma} \right\} \).

(ii) By Young's inequality, (A.7) follows that
\[
\Lambda(w, w^*) \\
\leq D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 \\
+ \frac{\epsilon(N-1)}{N} [L(u, p^*) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-1)}{N} \left[ \frac{1}{2\gamma} \| p - p^* \|^2 + \frac{\gamma}{2} \| Au - b \|^2 \right] \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2.
\]

From Assumption 2, we have \( D(u^*, u) \leq \frac{\beta}{2} \| u - u^* \|^2 \). Together with the fact \( Au^* = b \) and \( 2\gamma > (2N-1)\rho \), above inequality follows that
\[
\Lambda(w, w^*) \leq d_2 \| w - w^* \|^2 \\
+ \frac{\epsilon(N-1)}{N} [L(u, p^*) - L(u^*, p^*)],
\]
with \( d_2 = \max \left\{ \frac{(4N-3)\epsilon}{(4N-2)N\rho}, \frac{NB + \epsilon(2N-3)\gamma \lambda_{\max}(A^TA)}{2N} \right\} \).

(iii) By the definition of \( \Lambda(w, w') \), we have
\[
\Lambda(w, w') \geq \frac{\epsilon(N-1)}{N} [L(u, p) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2 \\
= \frac{\epsilon(N-1)}{N} [L(u, p) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2 \\
\geq \frac{\epsilon(N-1)}{N} [L(u, p) - L(u^*, p^*)] \\
+ \frac{\epsilon(N-2)\gamma}{2N} \| Au - b \|^2 \\
\geq -d_3 \| p - p^* \|^2,
\] (A.8)

with \( d_3 = \frac{\epsilon(N-1)^2}{2\gamma N(N-2)} \).

\qed
2. Proof of Lemma 2

Proof. Step 1: Estimate $\frac{1}{N} \mathbb{E}_{\tilde{i}(k)} \left[ L(u^{k+1}, q^k) - L(u, q^k) \right]$;
For all $u \in U$, the unique solution $u^{k+1}$ of the primal problem of RPDC is characterized by the following variational inequality:

$$
\langle \nabla_i(u^k) G(u^k), (u^{k+1} - u, u_{i(k)}) \rangle + J_i(u^k) - J_i(u_{i(k)})
+ (q^k, A_i(u^{k+1} - u_{i(k)}))
+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u^{k+1} - u \rangle \leq 0,
$$

which follows that

$$
\langle \nabla_i(u^k) G(u^k), (u^k - u - (u^k - u^{k+1})), u_{i(k)} \rangle
+ J_i(u^k) - J_i(u_{i(k)})
+ (q^k, A_i(u^k - u^{k+1})), u_{i(k)}
+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u^{k+1} - u \rangle \leq 0. \tag{A.9}
$$

Observing that for any separable mapping $\psi(u) = \sum_{i=1}^{N} \psi_i(u_i)$, we have $\psi_i(u^k) - \psi_i(u^{k+1}) = \psi(u^k) - \psi(u^{k+1})$. Therefore, (A.9) follows that

$$
\langle \nabla_i(u^k) G(u^k), (u^k - u - (u^k - u^{k+1})), u_{i(k)} \rangle
+ J_i(u^k) - J_i(u_{i(k)})
+ (q^k, A_i(u^k - u^{k+1})), u_{i(k)}
+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u^{k+1} - u \rangle \leq 0. \tag{A.10}
$$

Taking expectation with respect to $i(k)$ on both sides of (A.10), together the condition expectation (A.1)-(A.3), we get

$$
\frac{1}{N} \mathbb{E}_{\tilde{i}(k)} \left[ L(u^k, q^k) - L(u, q^k) \right]
\leq \mathbb{E}_{\tilde{i}(k)} \left\{ \langle \nabla G(u^k), u^k - u^{k+1} \rangle + J(u^k) - J(u^{k+1})
+ (q^k, A(u^k - u^{k+1})),
+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u^{k+1} - u \rangle \right\}. \tag{A.11}
$$

or

$$
\frac{1}{N} \mathbb{E}_{\tilde{i}(k)} \left[ L(u^{k+1}, q^k) - L(u, q^k) \right]
\leq \mathbb{E}_{\tilde{i}(k)} \left\{ \langle \nabla G(u^k), u^k - u^{k+1} \rangle + J(u^k) - J(u^{k+1})
+ (q^k, A(u^k - u^{k+1})),
+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle \right\}. \tag{A.12}
$$

By the gradient Lipschitz of $G$, term $a_1$ in (A.12) is bounded by

$$
a_1 = \langle \nabla G(u^k), u^k - u^{k+1} \rangle
\leq G(u^k) - G(u^{k+1}) + \frac{B_G}{2} \| u^k - u^{k+1} \|^2. \tag{A.13}
$$

The simple algebraic operation and Assumption 2 follows that

$$
a_2 = \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle
= \frac{1}{\epsilon} \left[ D(u, u^k) - D(u, u^{k+1}) - D(u^{k+1}, u^k) \right]
\leq \frac{1}{\epsilon} \left[ D(u, u^k) - D(u, u^{k+1}) - \frac{\beta}{2\epsilon} \| u^k - u^{k+1} \|^2 \right]. \tag{A.14}
$$

Combining (A.12)-(A.14), we obtain that

$$
\frac{\epsilon}{N} \mathbb{E}_{\tilde{i}(k)} \left[ L(u^{k+1}, q^k) - L(u, q^k) \right]
\leq \left[ D(u, u^k) - \mathbb{E}_{\tilde{i}(k)} D(u, u^{k+1}) \right]
+ \frac{\epsilon(N-1)}{N} \left[ L(u^k, q^k) - L(u^{k+1}, q^k) \right]
\leq \frac{\beta - \epsilon B_G}{2} \| u^k - u^{k+1} \|^2 \right\}. \tag{A.15}
$$
Since $p^{k+1} = p^k + \rho (Au^{k+1} - b)$ and $q^k = p^k + \gamma (Au^k - b)$, term $\alpha_3$ in (A.15) follows that

$$
\alpha_3 = L(u^k, q^k) - L(u^{k+1}, q^{k+1})
= L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \langle q^k - p^k, Au^k - b \rangle
+ \rho\|Au^{k+1} - b\|^2 - \gamma\|Au^k - b, Au^{k+1} - b\|
= L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \frac{\gamma}{2}\|Au^k - b\|^2
+ (\rho - \frac{\gamma}{2})\|Au^{k+1} - b\|^2
+ \gamma\lambda_{\text{max}}(A^TA)\|u^k - u^{k+1}\|^2. \tag{A.16}
$$

Combining (A.15)-(A.16), we have that

$$
\leq \frac{\epsilon}{N}E_{i(k)}[L(u^{k+1}, q^{k}) - L(u, q^k)]
+ \frac{\epsilon(N-1)}{2N}\|[L(u^k, p^k) - L(u^{k+1}, p^{k+1})]
- \beta - \epsilon B\epsilon + \frac{\epsilon}{2}\|u^k - u^{k+1}\|^2
+ \frac{\epsilon(\rho - \gamma)(N-1)}{2N}\|Au^{k+1} - b\|^2
+ \frac{\epsilon(\rho - \gamma)(N-1)}{2N}\|Au^{k+1} - b\|^2 \tag{A.17}
$$

Step 2: Estimate $\frac{\epsilon}{N}E_{i(k)}[L(u^{k+1}, p) - L(u^{k+1}, q^k)]$

$$
L(u^{k+1}, p) - L(u^{k+1}, q^k)
= \langle p - q^k, Au^{k+1} - b \rangle
= \frac{1}{\rho^2}[\|p - p^k\|^2 - \|p - p^{k+1}\|^2 + \|p^k - p^{k+1}\|^2]
- \gamma\|Au^k - b, Au^{k+1} - b\|
= \frac{1}{2\rho^2}[\|p - p^k\|^2 - \|p - p^{k+1}\|^2 + \|p^k - p^{k+1}\|^2]
+ \frac{\gamma}{2}\|A(u^k - u^{k+1})\|^2
- \frac{\gamma}{2}\|Au^k - b\|^2
- \frac{\gamma}{2}\|Au^{k+1} - b\|^2
= \frac{1}{2\rho^2}[\|p - p^k\|^2 - \|p - p^{k+1}\|^2 + \|p^k - p^{k+1}\|^2]
+ \frac{\gamma}{2}\|A(u^k - u^{k+1})\|^2
- \frac{\gamma}{2}\|Au^k - b\|^2
+ \frac{\gamma}{2}\|Au^{k+1} - b\|^2. \tag{A.18}
$$

Multiply $\frac{\epsilon}{N}$ on both side of above inequality, we obtain that:

$$
\leq \frac{\epsilon}{2N\rho^2}[\|p - p^k\|^2 - \|p - p^{k+1}\|^2]
+ \frac{\epsilon\gamma}{2N}\lambda_{\text{max}}(A^TA)\|u^k - u^{k+1}\|^2
- \frac{\epsilon\gamma}{2N}\|Au^k - b\|^2
+ \frac{\epsilon(\rho - \gamma)}{2N}\|Au^{k+1} - b\|^2. \tag{A.19}
$$

Taking expectation with respect to $i(k)$ on both side of inequality (A.19), we have

$$
\leq \frac{\epsilon}{2N\rho^2}[\|p - p^k\|^2 - E_{i(k)}\|p - p^{k+1}\|^2]
+ \frac{\epsilon\gamma}{2N}\lambda_{\text{max}}(A^TA)\|u^k - u^{k+1}\|^2
- \frac{\epsilon\gamma}{2N}\|Au^k - b\|^2
+ \frac{\epsilon(\rho - \gamma)}{2N}E_{i(k)}\|Au^{k+1} - b\|^2. \tag{A.20}
$$
Step 3: Estimate the variance of $\Lambda(u^k, w)$.
Summing inequalities (A.17) and (A.20), with $d_4 = \max\left\{\frac{\beta - \epsilon[B_G + \gamma \lambda_{max}(A^T A)]}{\epsilon(2\gamma - \epsilon)N}, \frac{1}{\epsilon[2\gamma - (2N - 1)\mu]}\right\}$, we have that

$$
\Lambda(u^k, w) - \mathbb{E}_{i(k)} \Lambda(u^{k+1}, w) 
\geq \mathbb{E}_{i(k)} \left\{ \frac{\epsilon}{N} \left[ L(u^{k+1}, p) - L(u, q^k) \right] + \frac{\beta - \epsilon[B_G + \gamma \lambda_{max}(A^T A)]}{2} \|u^k - u^{k+1}\|^2 
+ \frac{\epsilon(2\gamma - (2N - 1)\mu)}{2N} \|Au^{k+1} - b\|^2 \right\}
$$

Then we have the result of Lemma 2. \(\square\)

By Jensen’s inequality, (A.21) follows that

$$
\Lambda(u^k, w) - \mathbb{E}_{i(k)} \Lambda(u^{k+1}, w) 
\geq \frac{\epsilon}{N} \mathbb{E}_{i(k)} \left[ L(u^{k+1}, p) - L(u, q^k) \right] + d_4 \left\{ \frac{1 + 2\gamma^2 \lambda_{max}(A^T A)}{N^2} \|u^k - u^{k+1}\|^2 
+ \frac{2\gamma^2 \|Au^k - b\|^2}{2N} \right\}. \tag{A.22}
$$

Since $\lambda_{max}(A^T A) \|u^k - T_u(w^k)\|^2 \geq \|A[u^k - T_u(w^k)]\|^2$ and $\mathbb{E}_{i(k)} \Lambda(u^{k+1}, w)$, \(\epsilon \mathbb{E}_{i(k)} [L(u^{k+1}, p) - L(u, q^k)] \]

$$
\geq \frac{\epsilon}{N} \mathbb{E}_{i(k)} \left[ L(u^{k+1}, p) - L(u, q^k) \right] 
+ d_4 \left\{ \frac{1 + 2\gamma^2 \lambda_{max}(A^T A)}{N^2} \|u^k - u^{k+1}\|^2 
+ \frac{2\gamma^2 \|Au^k - b\|^2}{2N} \right\} \geq \frac{\epsilon}{N} \mathbb{E}_{i(k)} \left[ L(u^{k+1}, p) - L(u, q^k) \right] 
+ d_4 \|u^k - T_w(w^k)\|^2. \tag{A.23}
$$

\[\text{Since } \mathbb{E}_{i(k)} u^{k+1} - u^k = \frac{1}{N} [T_u(w^k) - w] \text{ in (A.4)}, \text{ (A.22) yields that} \]

$$
\Lambda(u^k, w) - \mathbb{E}_{i(k)} \Lambda(u^{k+1}, w) 
\geq \frac{\epsilon}{N} \mathbb{E}_{i(k)} \left[ L(u^{k+1}, p) - L(u, q^k) \right] 
+ d_4 \left\{ \frac{1 + 2\gamma^2 \lambda_{max}(A^T A)}{N^2} \|u^k - T_u(w^k)\|^2 
+ \frac{2\gamma^2 \|Au^k - b\|^2}{2N} \right\}. \tag{A.23}
$$

Supplementary material for the paper: “Linear Convergence of RPDC Method for Large-scale LCCP”
3. Proof of Theorem 1 (Almost surely convergence)

Proof.

(i) Take \( w = w^* \) in Lemma 2, we have
\[
\Lambda(w^k, w^*) \geq \frac{\epsilon}{N} \mathbb{E}_i(k) L(u^{k+1}, u^*) + d_4 \| w^k - T(w^k) \|^2.
\]
(A.24)

Observe that \( L(u^{k+1}, u^*) - L(u^*, q^k) \geq 0 \). From statement (i) of Lemma 1, we have that \( \Lambda(w^k, w^*) \) is non-negative. By the Robbins-Siegmund Lemma (Robbins & Siegmund, 1971), we obtain that \( \lim_{k \to +\infty} \Lambda(w^k, w^*) \) almost surely exists, \( \sum_{k=0}^{+\infty} \| w^k - T(w^k) \|^2 < +\infty \) a.s.

(ii) Since \( \lim_{k \to +\infty} \Lambda(w^k, w^*) \) almost surely exists, thus \( \Lambda(w^k, w^*) \) is almost surely bounded. Thanks statement (i) of Lemma 1, it implies the sequences \( \{ w^k \} \) is almost surely bounded.

(iii) From statement (i) we have that
\[
\lim_{k \to +\infty} \| w^k - T(w^k) \| = 0 \quad \text{a.s.}
\]
By variational inequality system (A.6), we have that any cluster point of a realization sequence generated by RPDC almost surely is a saddle point of Lagrangian for (P).

\[ \square \]

4. Proof of Theorem 2 (Expected primal suboptimality and expected feasibility)

Proof.

(i) Let \( h(w, w') = \Lambda(w, w') + \frac{d_4}{N} \Lambda(w, w^*) \). By statement (i) and (iii) in Lemma 1, we have \( h(w, w') \geq 0 \). From Lemma 2, we obtain that
\[
\mathbb{E}_i(k) \frac{\epsilon}{N} [L(u^{k+1}, p) - L(u, q^k)] \leq \Lambda(w^k, w) - \mathbb{E}_i(k) \Lambda(u^{k+1}, w)
\]
Taking expectation with respect to \( \mathcal{F}_t, t > k \) for above inequality, we obtain that
\[
\mathbb{E}_{\mathcal{F}_t} [L(u^{k+1}, p) - L(u, q^k)] \leq \mathbb{E}_{\mathcal{F}_t} [\Lambda(w^k, w) - \Lambda(u^{k+1}, w)]. \quad (A.25)
\]
Take \( w = w^* \) in (A.25), we obtain
\[
0 \leq \mathbb{E}_{\mathcal{F}_t} [\Lambda(w^k, w^*) - \Lambda(u^{k+1}, w^*)]. \quad (A.26)
\]
By the combination of (A.25) and (A.26), it follows
\[
\mathbb{E}_{\mathcal{F}_t} [L(u^{k+1}, p) - L(u, q^k)] \leq \mathbb{E}_{\mathcal{F}_t} [h(w^k, w) - h(u^{k+1}, w)] \quad (A.27)
\]
From the definition of \( \bar{u}_t \) and \( \bar{p}_t \), we have \( \bar{u}_t \in \mathbb{U} \) and \( \bar{p}_t \in \mathbb{R}^m \). From the convexity of set \( \mathbb{U}, \mathbb{R}^m \) and the function \( L(u', p) - L(u, p') \) is convex in \( u' \) and linear in \( p' \), for all \( u \in \mathbb{U} \) and \( p \in \mathbb{R}^m \), we have that
\[
\mathbb{E}_{\mathcal{F}_t} [L(\bar{u}_t, p) - L(u, \bar{p}_t)] \leq \mathbb{E}_{\mathcal{F}_t} \frac{1}{t+1} \sum_{k=0}^{t} [L(u^{k+1}, p) - L(u, q^k)] \leq \frac{N h(w^0, w)}{\epsilon(t+1)}. \quad (A.28)
\]

(ii) If \( \mathbb{E}_{\mathcal{F}_t} \| A\bar{u}_t - b \| = 0 \), statement (ii) is obviously. Otherwise, \( \mathbb{E}_{\mathcal{F}_t} \| A\bar{u}_t - b \| \neq 0 \) i.e., there is set \( \mathbb{W} \) such that \( P\{ \omega \in \mathbb{W} | \| A\bar{u}_t - b \| \neq 0 \} > 0 \). Let \( \hat{p} \) be a random vector:
\[
\hat{p} (\omega) = \begin{cases} 
0 & \omega \notin \mathbb{W} \\
\frac{M(A\bar{u}_t - b)}{\| A\bar{u}_t - b \|} & \omega \in \mathbb{W}
\end{cases} \quad (A.29)
\]
Noted that for \( \omega \notin \mathbb{W} \), we have \( \hat{p}(\omega) = 0 \) and \( \| A\bar{u}_t - b \| = 0 \). Thus
\[
\langle \hat{p}(\omega), A\bar{u}_t - b \rangle = M \| A\bar{u}_t - b \| = 0. \quad (A.30)
\]
Otherwise, for \( \omega \in \mathbb{W} \), we have that
\[
\langle \hat{p}(\omega), A\bar{u}_t - b \rangle = M \| A\bar{u}_t - b \|. \quad (A.31)
\]
Together (A.30) and (A.31), we have
\[ \langle \hat{p}, A \bar{u}_t - b \rangle = M \| A \bar{u}_t - b \| \] (A.32)

Moreover, since \( Au^* = b \), we have
\[
\begin{align*}
L(\bar{u}_t, \hat{p}) - L(u^*, \bar{p}_t) &= F(\bar{u}_t) + \langle \hat{p}, A \bar{u}_t - b \rangle - F(u^*) \\
&= F(\bar{u}_t) - F(u^*) + M \| A \bar{u}_t - b \|. \tag{A.33}
\end{align*}
\]

Moreover, by taking \( u = \bar{u}_t \) in the right hand side of saddle point inequality, we have
\[
F(\bar{u}_t) - F(u^*) \geq - \| p^* \| \| A \bar{u}_t - b \|. \tag{A.34}
\]

Combine (A.33) and (A.34), we have that
\[
\| A \bar{u}_t - b \| \leq \frac{L(\bar{u}_t, \hat{p}) - L(u^*, \bar{p}_t)}{(M - \| p^* \|)}. \tag{A.35}
\]

Take expectation on both side of above inequality, we have that
\[
\mathbb{E}_{\mathcal{F}_t} \| A \bar{u}_t - b \| \leq \mathbb{E}_{\mathcal{F}_t} \frac{L(\bar{u}_t, \hat{p}) - L(u^*, \bar{p}_t)}{(M - \| p^* \|)} \leq \mathbb{E}_{\mathcal{F}_t} \frac{N h(w^0, (u^*, \hat{p}))}{(M - \| p^* \|)} \epsilon(t + 1) \leq \mathbb{E}_{\mathcal{F}_t} \frac{N d_5}{(M - \| p^* \|)} \epsilon(t + 1) \tag{A.35}
\]

where \( d_5 = \sup_{\| p \| < M} h(u^0, (u^*, p)) \).

(iii) Again from (A.33), (A.34) and statement (ii), statement (iii) is coming.
6. Proof of Theorem 3 (Global strong metric subregularity of $H(w)$ implies linear convergence of RPDC)

**Proof.** Considering the reference point $T(w^k)$ associated with given point $w^k$, we have that
\[
\begin{align*}
0 & \in \nabla G(w^k) + \partial J(T_u(w^k)) + A^T q^k + \frac{1}{2} \left[ \nabla K(T_u(w^k)) - \nabla K(w^k) \right] + \mathcal{N}_U(T_u(w^k)) \\
0 & = b - AT_u(w^k) + \frac{1}{2} \begin{bmatrix} T_p(w^k) - p^k \end{bmatrix} \end{align*}
\]
(A.36)

Thus
\[
v(T(w^k)) = \left( \nabla G(T_u(w^k)) - \nabla G(w^k) + A^T(T_p(w^k) - q^k) \right) + \frac{1}{2} \left[ p^k - T_p(w^k) \right] \in H(T(w^k)).
\]

From Assumption 1 and 2, there is $\delta > 0$ such that
\[
\|v(T(w^k))\|^2 \leq \delta\|w^k - T(w^k)\|^2. \tag{A.37}
\]

Since $H(w)$ is global strong metric subregular at $w^*$ for 0, then
\[
\|T(w^k) - w^*\| \leq \epsilon\text{dist}(0, H(T(w^k)))
\leq \epsilon\|v(T(w^k))\|
\leq \epsilon\sqrt{\delta\|w^k - T(w^k)\|}. \tag{A.38}
\]

Since $\|w^k - w^*\| \leq \|T(w^k) - w^*\| + \|w^k - T(w^k)\|$, we have
\[
\|w^k - w^*\| \leq (\epsilon\sqrt{\delta} + 1)\|w^k - T(w^k)\|. \tag{A.39}
\]

From statement (iii) of Lemma 3, we have that
\[
\phi(w^k, w^*) = E_{i(k)}(w^{k+1}, w^*) \geq d_4\|w^k - T(w^k)\|^2 + \frac{\epsilon}{N} [L(u^k, p^*) - L(u^*, p^*)]
\geq d_4 \left( \epsilon\sqrt{\delta} + 1 \right)^2 \|w^k - w^*\|^2
+ \frac{\epsilon}{N} [L(u^k, p^*) - L(u^*, p^*)] \quad \text{(by (A.39))}
\geq \delta' \{d_2\|w^k - w^*\|^2 + \epsilon[L(u^k, p^*) - L(u^*, p^*)]\}
\geq \delta' \phi(w^k, w^*). \quad \text{(by (i) of Lemma 3)} \tag{A.40}
\]

where
\[
\delta' = \min\left\{ \frac{d_4}{\max\{d_2(\epsilon\sqrt{\delta} + 1)^2, d_4 + 1\}}, \frac{1}{N+1} \right\} < 1.
\]

It follows that
\[
E_{i(k)}(w^{k+1}, w^*) \leq \alpha \phi(w^k, w^*). \tag{A.41}
\]

where $\alpha = 1 - \delta' \in (0, 1)$. Taking expectation with respect to $\mathcal{F}_{k+1}$ for above inequality, we obtain that
\[
E_{\mathcal{F}_{k+1}}(w^{k+1}, w^*) \leq \alpha^{k+1} \phi(w^0, w^*). \tag{A.42}
\]

\[ \square \]
8. Proof of Proposition 1

Proof. By the piecewise linear of \( H(w) \) and Zheng and Ng (Zheng & Ng, 2014), we have that \( H(w) \) is global metric subregular at \( w^* \) for 0. Since \( Q \) is positive-definite, then problem (SVM) has unique solution \( u^* \). Hence, to show \( H(w) \) is global strongly metric subregular, we need to prove uniqueness of the Lagrangian multiplier for (SVM).

Suppose there are two multipliers \( p \) and \( p' \), thus we have

\[
\begin{align*}
0 & \in Qu^* - 1_n + py + N_{[0,c]^n}(u^*) \\
0 & \in Qu^* - 1_n + p'y + N_{[0,c]^n}(u^*)
\end{align*}
\]

Since there exists at least one component \( u_i^* \) of optimal solution \( u^* \) satisfies \( 0 < u_i^* < c \), then \( \xi_i = N_{[0,c]}(u_i^*) = 0 \).

Thus, we have that

\[
\begin{align*}
Q_j u^* - 1 + y_j p &= 0 \\
Q_j u^* - 1 + y_j p' &= 0 \quad (A.43)
\end{align*}
\]

We conclude that \( p = p' \). Therefore \( H(w) \) is globally strongly metric subregular. \( \square \)

9. Proof of Proposition 2

Proof. By the piecewise linear of \( H(w) \) and Zheng and Ng (Zheng & Ng, 2014), we have that \( H(w) \) is global metric subregular at \( w^* \) for 0. Since \( \Sigma \) is positive-definite, then problem (MLP) has unique solution \( u^* \). Hence, to show \( H(w) \) is global strongly metric subregular, we need to prove uniqueness of the Lagrangian multiplier for (MLP). Suppose there are two pair of multipliers \((p_1, p_2)\) and \((p'_1, p'_2)\), thus we have

\[
\begin{align*}
0 & \in \Sigma u^* + \lambda \partial ||u^*|| + p_1 \mu + p_2 1_n \\
0 & \in \Sigma u^* + \lambda \partial ||u^*|| + p'_1 \mu + p'_2 1_n
\end{align*}
\]

Since \( u_i^* \neq 0, u_j^* \neq 0 \), thus \( \xi_i = \partial ||u_i^*|| \) and \( \xi_j = \partial ||u_j^*|| \) are single valued and we have

\[
\begin{align*}
\Sigma_i u^* + \lambda \xi_i + \mu_i p_1 + p_2 &= 0 \\
\Sigma_i u^* + \lambda \xi_i + \mu_i p'_1 + p'_2 &= 0 \quad (A.44)
\end{align*}
\]

\[
\begin{align*}
\Sigma_j u^* + \lambda \xi_j + \mu_j p_1 + p_2 &= 0 \\
\Sigma_j u^* + \lambda \xi_j + \mu_j p'_1 + p'_2 &= 0 \quad (A.45)
\end{align*}
\]

It follows that

\[
\begin{align*}
\mu_i (p_1 - p'_1) + p_2 - p'_2 &= 0 \\
\mu_j (p_1 - p'_1) + p_2 - p'_2 &= 0 \quad (A.46)
\end{align*}
\]

Since \( \mu_i \neq \mu_j \), we conclude that \( p_1 = p'_1 \) and \( p_2 = p'_2 \). Therefore \( H(w) \) is globally strongly metric subregular. \( \square \)

References

