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 \( (i^{k+1}, p^{k+1}) \). Recall the definition of filtration \( \mathcal{F}_k \) which is generated by the random variable \( i(0), i(1), \ldots, i(k) \), i.e.,

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

Additionally, \( \mathcal{F} = (\mathcal{F}_k)_{k \in \mathbb{N}} \), \( \mathbb{E}_{\mathcal{F}_{k+1}} = \mathbb{E}(\cdot | \mathcal{F}_k) \) 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:

\[
\mathbb{E}_{i(k)}(\nabla_{i(k)} G(u^k), (u^k - u)_{i(k)}) = \frac{1}{N} (\nabla G(u^k), u^k - u) \geq \frac{1}{N} [G(u^k) - G(u)], \quad (A.1)
\]

\[
\mathbb{E}_{i(k)}[J_{i(k)}(u^k_{i(k)}) - J_{i(k)}(u_{i(k)})] = \frac{1}{N} [J(u^k) - J(u)], \quad (A.2)
\]

and

\[
\mathbb{E}_{i(k)}(q^k, A_{i(k)}(u^k - u)_{i(k)}) = \frac{1}{N} (q^k, A(u^k - u)). \quad (A.3)
\]

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 \( w^k \),

\[
\begin{align*}
\text{APP-AL} & \quad \left\{ \begin{array}{l}
T_u(w^k) = \arg \min_{u \in \mathbb{U}} \langle \nabla G(u^k), u \rangle + J(u) + \langle q^k, Au \rangle + \frac{1}{N} D(u, u^k); \\
T_p(w^k) = p^k + \gamma [AT_u(w^k) - b],
\end{array} \right.
\end{align*}
\]

with \( q^k = p^k + \gamma (Au^k - b) \), 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.

\[
\mathbb{E}_{i(k)} u^{k+1} = \frac{1}{N} T_u(w^k) + (1 - \frac{1}{N}) u^k, \quad (A.4)
\]

or

\[
T_u(w^k) = N \mathbb{E}_{i(k)} u^{k+1} - (N-1) u^k. \quad (A.5)
\]

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

\[
\begin{align*}
\left\{ \begin{array}{l}
\langle \nabla G(u^k), u - T_u(w^k) \rangle + J(u) - J(T_u(w^k)) + \langle q^k, A(u - T_u(w^k)) \rangle \\
+ \frac{1}{N} \langle \nabla K(T_u(w^k)) - \nabla K(u^k), u - T_u(w^k) \rangle \geq 0,
\end{array} \right.
\end{align*} \quad (A.6)
\]

\[
\gamma [AT_u(w^k) - b] = T_p(w^k) - p^k.
\]
1. Proof of Lemma 1

Proof. Take \( w' = w^* \) in (9), we have that

\[
\Lambda(w, w^*) = D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 + \frac{\epsilon(N - 1)}{N} [L(u, p) - L(u^*, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2
\]

(i) By Young’s inequality, (A.7) follows that

\[
\Lambda(w, w^*) \geq D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 + \frac{\epsilon(N - 1)}{N} [L(u, p^*) - L(u^*, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2.
\]

From Assumption 2, we have \( D(u^*, u) \geq \frac{\beta}{2} \| u - u^* \|^2 \). Together with the fact \( Au^* = b \) and \( \rho < \frac{2\gamma}{2N-1} \), above inequality follows that

\[
\Lambda(w, w^*) \geq d_1 \| w - w^* \|^2,
\]

with \( d_1 = \min \left\{ \frac{1}{2N} [N\beta - \epsilon\gamma\lambda_{\max}(A^TA)], \frac{\epsilon}{4N\gamma} \right\} \).

(ii) By Young’s inequality, (A.7) follows that

\[
\Lambda(w, w^*) \leq D(u^*, u) + \frac{\epsilon}{2N\rho} \| p - p^* \|^2 + \frac{\epsilon(N - 1)}{N} [L(u, p^*) - L(u^*, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2.
\]

From Assumption 2, we have \( D(u^*, u) \leq \frac{B}{2} \| u - u^* \|^2 \). 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 \( d_2 = \max \left\{ \frac{(4N - 3)\epsilon}{(2N - 2)N\rho}, \frac{NB + \epsilon(2N - 3)\gamma\lambda_{\max}(A^TA)}{2N} \right\} \).

(iii) By the definition of \( \Lambda(w, w') \), we have

\[
\Lambda(w, w') \geq \frac{\epsilon(N - 1)}{N} [L(u, p) - L(u^*, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2
\]

\[
= \frac{\epsilon(N - 1)}{N} [L(u, p) - L(u, p^*)] + \frac{\epsilon(N - 1)}{N} [L(u, p^*) - L(u^*, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2
\]

\[
\geq \frac{\epsilon(N - 1)}{N} [L(u, p) - L(u, p^*)] + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2
\]

\[
= \frac{\epsilon(N - 1)}{N} \langle p - p^*, Au - b \rangle + \frac{\epsilon(N - 2)\gamma}{2N} \| Au - b \|^2
\]

\[
\geq -d_3 \| p - p^* \|^2,
\]

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

\( \square \)
2. Proof of Lemma 2

Proof. Step 1: Estimate \( \frac{1}{N} \mathbb{E}_{i(k)}[L(u^{k+1}, q^k) - L(u, q^k)] \);

For all \( u \in U \), the unique solution \( u^{k+1} \) of the primal problem of RPDC is characterized by the following variational inequality:

\[
\langle \nabla_i G(u^k), (u^{k+1} - u)_{i(k)} \rangle + J_i(u_{i(k)}^{k+1}) - J_i(u_{i(k)}) + \langle q^k, A_i(u_{i(k)}^{k+1} - u)_{i(k)} \rangle
\]

\[+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u^{k+1} - u \rangle \leq 0,
\]

which follows that

\[
\langle \nabla_i G(u^k), (u^k - u - (u^k - u^{k+1}))_{i(k)} \rangle + J_i(u_{i(k)}^k) - J_i(u_{i(k)}) - \langle q^k, A_i(u - u^{k+1})_{i(k)} \rangle
\]

\[+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle \leq 0.
\] (A.9)

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

Therefore, (A.9) follows that

\[
\langle \nabla_i G(u^k), (u^k - u - (u^k - u^{k+1}))_{i(k)} \rangle + J_i(u_{i(k)}^k) - J_i(u_{i(k)}) - \langle q^k, A_i(u - u^{k+1})_{i(k)} \rangle
\]

\[+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle \leq 0.
\] (A.10)

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

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

\[+ \langle q^k, A(u^k - u^{k+1}) \rangle + \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle \right\}. \] (A.11)

or

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

\[+ \langle q^k, A(u^k - u^{k+1}) \rangle + \frac{1}{N} \left[ L(u^{k+1}, q^k) - L(u^k, q^k) \right]
\]

\[+ \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle \right\}. \] (A.12)

By the gradient Lipschitz of \( G \), term \( \alpha_1 \) in (A.12) is bounded by

\[
\alpha_1 = \langle \nabla G(u^k), u^k - u^{k+1} \rangle \leq G(u^k) - G(u^{k+1}) + \frac{B_G}{2} \| u^k - u^{k+1} \|^2.
\] (A.13)

The simple algebraic operation and Assumption 2 follows that

\[
\alpha_2 = \frac{1}{\epsilon} \langle \nabla K(u^{k+1}) - \nabla K(u^k), u - u^{k+1} \rangle = \frac{1}{\epsilon} \left[ D(u, u^k) - D(u, u^{k+1}) - D(u^{k+1}, u^k) \right]
\]

\[\leq \frac{1}{\epsilon} \left[ D(u, u^k) - D(u, u^{k+1}) \right] - \frac{\beta}{2\epsilon} \| u^k - u^{k+1} \|^2.
\] (A.14)

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

\[
\frac{1}{N} \mathbb{E}_{i(k)} \left[ L(u^{k+1}, q^k) - L(u, q^k) \right] \leq \left[ D(u, u^k) - \mathbb{E}_{i(k)} D(u, u^{k+1}) \right] + \mathbb{E}_{i(k)} \left\{ \frac{\epsilon(N - 1)}{N} \left[ L(u^k, q^k) - L(u^{k+1}, q^k) \right] \right.
\]

\[+ \frac{\beta - \epsilon B_G}{2} \| u^k - u^{k+1} \|^2 \right\}. \] (A.15)
Supplementary material for the paper: "Linear Convergence of RPDC Method for Large-scale LCCP"

Since \( p^{k+1} = p^k + \rho (A u^{k+1} - b) \) and \( q^k = p^k + \gamma (A u^k - b) \), term \( a_3 \) in (A.15) follows that

\[
a_3 = L(u^k, q^k) - L(u^{k+1}, q^k) \\
= L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \langle q^k - p^k, A u^k - b \rangle + \langle p^{k+1} - q^k, A u^{k+1} - b \rangle \\
= L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \gamma \|Au^k - b\|^2 + \rho \|Au^{k+1} - b\|^2 - \gamma (Au^k - b, Au^{k+1} - b) \\
= L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \gamma 2 \|Au^k - b\|^2 + (\rho - \gamma 2) \|Au^{k+1} - b\|^2 + \gamma 2 \|A(u^k - u^{k+1})\|^2 \\
\leq L(u^k, p^k) - L(u^{k+1}, p^{k+1}) + \gamma 2 \|Au^k - b\|^2 + (\rho - \gamma 2) \|Au^{k+1} - b\|^2 \\
+ \frac{\gamma \lambda_{\text{max}}(A^T A)}{2} \|u^k - u^{k+1}\|^2. \tag{A.16}
\]

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

\[
\frac{\epsilon}{N} \mathbb{E}_{i(k)} [L(u^{k+1}, q^k) - L(u, q^k)] \leq \left[ D(u, u^k) - \mathbb{E}_{i(k)} D(u, u^{k+1}) \right] + \mathbb{E}_{i(k)} \left\{ \frac{\epsilon(N - 1)}{N} [L(u^k, p^k) - L(u^{k+1}, p^{k+1})] \\
- \beta - \epsilon \frac{N - 1}{N} \frac{\gamma \lambda_{\text{max}}(A^T A)}{2} \|u^k - u^{k+1}\|^2 + \frac{\epsilon \gamma (N - 1)}{2N} \|Au^k - b\|^2 \\
+ \beta - \epsilon \frac{2N - \gamma (N - 1)}{2N} \|Au^{k+1} - b\|^2 \right\} \tag{A.17}
\]

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

\[
L(u^{k+1}, p) - L(u^{k+1}, q^k) = \langle p - q^k, A u^{k+1} - b \rangle \\
= \frac{1}{\rho} \langle p - p^k, p^{k+1} - p^k \rangle - \gamma (Au^k - b, Au^{k+1} - b) \\
= \frac{1}{2\rho} \left[ \|p - p^k\|^2 - \|p - p^{k+1}\|^2 + \|p^k - p^{k+1}\|^2 \right] - \gamma (Au^k - b, Au^{k+1} - b) \\
= \frac{1}{2\rho} \left[ \|p - p^k\|^2 - \|p - p^{k+1}\|^2 + \|p^k - p^{k+1}\|^2 \right] + \frac{\gamma}{2} \|A(u^k - u^{k+1})\|^2 \\
- \frac{\gamma}{2} \|Au^k - b\|^2 - \frac{\gamma}{2} \|Au^{k+1} - b\|^2 \\
= \frac{1}{2\rho} \left[ \|p - p^k\|^2 - \|p - p^{k+1}\|^2 \right] + \frac{\gamma}{2} \|A(u^k - u^{k+1})\|^2 \\
- \frac{\gamma}{2} \|Au^k - b\|^2 + \frac{\rho - \gamma}{2} \|Au^{k+1} - b\|^2 \quad \text{since } p^{k+1} = p^k + \rho (A u^{k+1} - b). \\
\leq \frac{1}{2\rho} \left[ \|p - p^k\|^2 - \|p - p^{k+1}\|^2 \right] + \frac{\gamma \lambda_{\text{max}}(A^T A)}{2} \|u^k - u^{k+1}\|^2 \\
- \frac{\gamma}{2} \|Au^k - b\|^2 + \frac{\rho - \gamma}{2} \|Au^{k+1} - b\|^2 \tag{A.18}
\]

Multiply \( \frac{\epsilon}{N} \) on both side of above inequality, we obtain that: \( \forall \rho \in \mathbb{R}^m \)

\[
\frac{\epsilon}{N} \mathbb{E}_{i(k)} [L(u^{k+1}, p) - L(u^{k+1}, q^k)] \leq \frac{\epsilon}{2N \rho} \left[ \|p - p^k\|^2 - \|p - p^{k+1}\|^2 \right] + \frac{\epsilon \gamma \lambda_{\text{max}}(A^T A)}{2} \|u^k - u^{k+1}\|^2 \\
- \frac{\epsilon \gamma}{2N} \|Au^k - b\|^2 + \frac{\epsilon (\rho - \gamma)}{2N} \|Au^{k+1} - b\|^2. \tag{A.19}
\]

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

\[
\frac{\epsilon}{N} \mathbb{E}_{i(k)} [L(u^{k+1}, p) - L(u^{k+1}, q^k)] \leq \frac{\epsilon}{2N \rho} \left[ \|p - p^k\|^2 - \mathbb{E}_{i(k)} \|p - p^{k+1}\|^2 \right] + \frac{\epsilon \gamma \lambda_{\text{max}}(A^T A)}{2} \mathbb{E}_{i(k)} \|u^k - u^{k+1}\|^2 \\
- \frac{\epsilon \gamma}{2N} \|Au^k - b\|^2 + \frac{\epsilon (\rho - \gamma)}{2N} \mathbb{E}_{i(k)} \|Au^{k+1} - b\|^2. \tag{A.20}
\]
The result of Lemma 2:

Then we have the result of Lemma 2.

**Step 3: Estimate the variance of** $\Lambda(w^k, w)$. 

Summing inequalities (A.17) and (A.20), with $d_4 = \frac{\min \left\{ \frac{\beta - \epsilon [B_0 + \gamma \lambda_{\text{max}}(A^T A)]}{\sqrt{\max \{N^2 + 2\gamma^2(N^2 + 2)\lambda_{\text{max}}(A^T A), 4\gamma^2\}}}, \frac{\epsilon (2\gamma - (2N - 1)\rho)}{2N} \right\}}{\max \{N^2 + 2\gamma^2(N^2 + 2)\lambda_{\text{max}}(A^T A), 4\gamma^2\}}$, we have that 

$$
\Lambda(w^k, w) - \mathbb{E}_{i(k)}\Lambda(w^{k+1}, w) 
\geq \mathbb{E}_{i(k)} \left\{ \frac{\epsilon}{N} [L(u^{k+1}, p) - L(u, q^k)] + \frac{\beta - \epsilon [B_0 + \gamma \lambda_{\text{max}}(A^T A)]}{2} \| u^k - u^{k+1} \|^2 + \epsilon [2\gamma - (2N - 1)\rho] \| Au^{k+1} - b \|^2 \right\} 
\geq \mathbb{E}_{i(k)} \left\{ \frac{\epsilon}{N} [L(u^{k+1}, p) - L(u, q^k)] + d_4 [(N^2 + 2\gamma^2(N^2 + 2)\lambda_{\text{max}}(A^T A))] \| u^k - u^{k+1} \|^2 + 4\gamma^2 \| Au^{k+1} - b \|^2 \right\} 
\geq \mathbb{E}_{i(k)} \left\{ \frac{\epsilon}{N} [L(u^{k+1}, p) - L(u, q^k)] + d_4 [(1 + 2\gamma^2\lambda_{\text{max}}(A^T A))] N^2 \| u^k - u^{k+1} \|^2 + 4\gamma^2 \| Au^{k+1} - b \|^2 \right\} 
= \frac{\epsilon}{N} \mathbb{E}_{i(k)} [L(u^{k+1}, p) - L(u, q^k)] + d_4 [(1 + 2\gamma^2\lambda_{\text{max}}(A^T A))] N^2 \mathbb{E}_{i(k)} \| u^k - u^{k+1} \|^2 + 2\gamma^2 \| Au^{k+1} - b \|^2]. \tag{A.21}
$$

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

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

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

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

Since $\lambda_{\text{max}}(A^T A)\| u^k - T_u(w^k) \|^2 \geq \| A[u^k - T_u(w^k)] \|^2$ and $T_p(w^k) - p^k = \gamma [AT_u(w^k) - b]$, (A.23) follows that 

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

Then we have the result of Lemma 2. \qed

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

\[
\Lambda(w^k, w^*) \geq \mathbb{E}_{i(k)}\Lambda(w^{k+1}, w^*) + \frac{\epsilon}{N}\mathbb{E}_{i(k)}[L(u^{k+1}, p^*) - L(u^*, q^k)] + d_4\|w^k - T(w^k)\|^2. \tag{A.24}
\]

Observe that \( L(u^{k+1}, p^*) - L(u^*, q^k) \geq 0 \). From statement (i) of Lemma 1, we have that \( \Lambda(w^k, w^*) \) is nonnegative. 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 \( h(w, w') = \Lambda(w, w') + \frac{\epsilon}{N} \Lambda(w, w^*) \). By statement (i) and (iii) in Lemma 1, we have \( h(w, w') \geq 0 \). From Lemma 2, we obtain that

\[
\mathbb{E}_{i(k)}\frac{\epsilon}{N}[L(u^{k+1}, p) - L(u, q^k)] \leq \Lambda(w^k, w) - \mathbb{E}_{i(k)}\Lambda(w^{k+1}, w)
\]

Taking expectation with respect to \( \mathcal{F}_t, t > k \) for above inequality, we obtain that

\[
\frac{\epsilon}{N}\mathbb{E}_{\mathcal{F}_t}[L(u^{k+1}, p) - L(u, q^k)] \leq \mathbb{E}_{\mathcal{F}_t}[\Lambda(w^k, w) - \Lambda(w^{k+1}, w)]. \tag{A.25}
\]

Take \( w = w^* \) in (A.25), we obtain

\[
0 \leq \mathbb{E}_{\mathcal{F}_t}[\Lambda(w^k, w^*) - \Lambda(w^{k+1}, w^*)]. \tag{A.26}
\]

By the combination of (A.25) and (A.26), it follows

\[
\frac{\epsilon}{N}\mathbb{E}_{\mathcal{F}_t}[L(u^{k+1}, p) - L(u, q^k)] \leq \mathbb{E}_{\mathcal{F}_t}[h(u^k, w) - h(w^k, w)] \tag{A.27}
\]

From the definition of \( \bar{u}_t \) and \( \bar{p}_t \), we have \( \bar{u}_t \in \mathcal{U} \) and \( \bar{p}_t \in \mathbb{R}^{m_t} \). From the convexity of set \( \mathcal{U}, \mathbb{R}^{m_t} \) and the function \( L(u', p') - L(u, p') \) is convex in \( u' \) and linear in \( p' \), for all \( u \in \mathcal{U} \) and \( p \in \mathbb{R}^{m_t} \), we have that

\[
\mathbb{E}_{\mathcal{F}_t}[L(\bar{u}_t, p) - L(u, \bar{p}_t)] \leq \mathbb{E}_{\mathcal{F}_t}\frac{1}{t+1} \sum_{k=0}^{t} [L(u^{k+1}, p) - L(u, q^k)] \leq \frac{N h(u^0, w)}{\epsilon(t+1)}. \tag{A.28}
\]
(ii) If $\mathbb{E}_{\mathcal{F}_t} \|A\tilde{u}_t - b\| = 0$, statement (ii) is obviously. Otherwise, $\mathbb{E}_{\mathcal{F}_t} \|A\tilde{u}_t - b\| \neq 0$ i.e., there is set $\mathcal{W}$ such that $\mathbb{P}\{\omega \in \mathcal{W} | \|A\tilde{u}_t - b\| \neq 0\} > 0$. Let $\hat{p}$ be a random vector:

$$\hat{p}(\omega) = \begin{cases} 0 & \omega \notin \mathcal{W} \\ M\bar{A}\tilde{u}_t - b & \omega \in \mathcal{W}. \end{cases} \quad (A.29)$$

Noted that for $\omega \notin \mathcal{W}$, we have $\hat{p}(\omega) = 0$ and $\|A\tilde{u}_t - b\| = 0$. Thus

$$\langle \hat{p}(\omega), A\tilde{u}_t - b \rangle = M\|A\tilde{u}_t - b\| = 0. \quad (A.30)$$

Otherwise, for $\omega \in \mathcal{W}$, we have that

$$\langle \hat{p}(\omega), A\tilde{u}_t - b \rangle = M\|A\tilde{u}_t - b\|. \quad (A.31)$$

Together (A.30) and (A.31), we have

$$\langle \hat{p}, A\tilde{u}_t - b \rangle = M\|A\tilde{u}_t - b\| \quad (A.32)$$

Moreover, since $Au^* = b$, we have

$$L(\tilde{u}_t, \hat{p}) - L(u^*, \hat{p}_t) = F(\tilde{u}_t) + \langle \hat{p}, A\tilde{u}_t - b \rangle - F(u^*) = F(\tilde{u}_t) - F(u^*) + M\|A\tilde{u}_t - b\|. \quad (A.33)$$

Moreover, by taking $u = \tilde{u}_t$ in the right hand side of saddle point inequality, we have

$$F(\tilde{u}_t) - F(u^*) \geq -(p^*, A\tilde{u}_t - b) \geq -\|p^*\|\|A\tilde{u}_t - b\|. \quad (A.34)$$

Combine (A.33) and (A.34), we have that

$$\|A\tilde{u}_t - b\| \leq \frac{L(\tilde{u}_t, \hat{p}) - L(u^*, \hat{p}_t)}{M - \|p^*\|}. \quad (A.35)$$

Take expectation on both side of above inequality, we have that

$$\mathbb{E}_{\mathcal{F}_t} \|A\tilde{u}_t - b\| \leq \frac{\mathbb{E}_{\mathcal{F}_t}[L(\tilde{u}_t, \hat{p}) - L(u^*, \hat{p}_t)]}{M - \|p^*\|} \leq \frac{\mathbb{E}_{\mathcal{F}_t}\left[N \bar{h}(u^0, (u^*, \hat{p}))\right]}{(M - \|p^*\|) \epsilon(t + 1)} \quad \text{(by (i))}

\begin{align*}
\leq \frac{\mathbb{E}_{\mathcal{F}_t} N \bar{d}_5}{(M - \|p^*\|) \epsilon(t + 1)} \\
\end{align*}$$

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.

\[\square\]

5. Proof of Lemma 3

Proof.

(i) This statement directly follows from the definition of $\phi(w, w^*)$ and statement (i) in Lemma 1.

(ii) This statement directly follows from the definition of $\phi(w, w^*)$ and statement (ii) in Lemma 1.

(iii) By the definition of $\phi(w, w^*)$, we have that.

$$\phi(w^k, w^*) - \mathbb{E}_{i(k)} \phi(w^{k+1}, w^*)$$

$$= \Lambda(w^k, w^*) - \mathbb{E}_{i(k)} \left\{ \Lambda(w^{k+1}, w^*) + \frac{\epsilon}{N} [L(w^k, p^*) - L(u^*, p^*)] - \frac{\epsilon}{N} [L(w^{k+1}, p^*) - L(u^*, p^*)] \right\}$$

$$\geq \Lambda(w^k, w^*) - \mathbb{E}_{i(k)} \left\{ \Lambda(w^{k+1}, w^*) + \frac{\epsilon}{N} [L(w^k, p^*) - L(u^*, p^*)] - \frac{\epsilon}{N} [L(w^{k+1}, p^*) - L(u^*, q^k)] \right\}$$

(by the definition of saddle point.)

$$\geq d_4 \|w^k - T(w^k)\|^2 + \frac{\epsilon}{N} [L(w^k, p^*) - L(u^*, p^*)]. \quad \text{(by Lemma 2)}$$

\[\square\]
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

\[
\begin{aligned}
0 &\in \nabla G(w^k) + \partial J(T(w^k)) + A^T q^k + \frac{1}{\epsilon} [\nabla K(T_u(w^k)) - \nabla K(u^k)] + N_{\Omega}(T_u(w^k)) \\
0 &\in b - AT_u(w^k) + \frac{1}{\epsilon} [T_p(w^k) - p^k]
\end{aligned}
\]  

(A.36)

Thus

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

From Assumption 1 and 2, there is $\delta > 0$ such that

\[
\|v(T(w^k))\|^2 \leq \delta \|w^k - T(w^k)\|^2.
\]

(A.37)

Since $H(w)$ is global strong metric subregular at $w^*$ for 0, then

\[
\|T(w^k) - w^*\| \leq \text{dist}(0, H(T(w^k))) \leq c \|v(T(w^k))\| \leq c \sqrt{\delta} \|w^k - T(w^k)\|.
\]

(A.38)

Since $\|w^k - w^*\| \leq \|T(w^k) - w^*\| + \|w^k - T(w^k)\|$, we have

\[
\|w^k - w^*\| \leq (c \sqrt{\delta} + 1) \|w^k - T(w^k)\|.
\]

(A.39)

From statement (iii) of Lemma 3, we have that

\[
\phi(w^k, w^*) - \mathbb{E}_{\xi(k)} \phi(w^{k+1}, w^*) \geq d_1 \|w^k - T(w^k)\|^2 + \frac{\epsilon}{N} [L(u^k, p^*) - L(u^*, p^*)]
\]

\[
\geq \frac{1}{(c \sqrt{\delta} + 1)^2} [\|w^k - w^*\|^2 + \frac{\epsilon}{N} [L(u^k, p^*) - L(u^*, p^*)]] \quad \text{(by (A.39))}
\]

\[
\geq \delta' \{d_1 \|w^k - w^*\|^2 + \epsilon [L(u^k, p^*) - L(u^*, p^*)]\}
\]

\[
\geq \delta' \phi(w^k, w^*). 
\]

(A.40)

where $\delta' = \min\{\frac{d_1}{\max\{d_2(c \sqrt{\delta} + 1)^2, d_4\}}, \frac{1}{N+1}\} < 1$. It follows that

\[
\mathbb{E}_{\xi(k)} \phi(w^{k+1}, w^*) \leq \alpha \phi(w^k, w^*).
\]

(A.41)

where $\alpha = 1 - \delta' \in (0, 1)$. Taking expectation with respect to $\mathcal{F}_k$ for above inequality, we obtain that

\[
\mathbb{E}_{\mathcal{F}_k} \phi(w^{k+1}, w^*) \leq \alpha^{k+1} \phi(w^0, w^*).
\]

(A.42)

\square

7. Proof of Corollary 1 (R-linear rate of the sequence $\{\mathbb{E}_{\mathcal{F}_k} w^k\}$)

Proof. By statement (i) in Lemma 3, we have that $\phi(w, w^*) \geq d_1 \|w - w^*\|^2$. By Theorem 3, we have that

\[
\mathbb{E}_{\mathcal{F}_k} \phi(w^k, w^*) \leq \alpha^k \phi(w^0, w^*).
\]

Then we have that

\[
\mathbb{E}_{\mathcal{F}_k} \|w^k - w^*\|^2 \leq \frac{\alpha^k \phi(w^0, w^*)}{d_1}.
\]

By convexity of $\| \cdot \|^2$ and Jensen’s inequality, we obtain that

\[
\|\mathbb{E}_{\mathcal{F}_k} w^k - w^*\| \leq \hat{M} (\sqrt{\alpha})^k \quad \text{with} \quad \hat{M} = \sqrt{\frac{\phi(w^0, w^*)}{d_1}}.
\]

This shows that the sequence $\{\mathbb{E}_{\mathcal{F}_k} w^k\}$ converges to the desired saddle point $w^*$ at R-linear rate; i.e.,

\[
\lim_{k \to \infty} \sup \sqrt{\|\mathbb{E}_{\mathcal{F}_k} w^k - w^*\|} = \sqrt{\alpha} < 1.
\]

\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

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

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

$$\begin{align*}
Q_iu^* - 1 + y_ip &= 0 \\
Q_iu^* - 1 + y_ip' &= 0
\end{align*} \tag{A.43}$$

We conclude that $p = p'$. Therefore $H(w)$ is global 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 pare of multipliers $(p_1, p_2)$ and $(p'_1, p'_2)$, thus we have

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

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

$$\begin{align*}
\Sigma_i u^* + \lambda \xi_i + \mu_i p_1 + p_2 &= 0 \tag{A.44} \\
\Sigma_i u^* + \lambda \xi_i + \mu_i p'_1 + p'_2 &= 0 \\
\Sigma_j u^* + \lambda \xi_j + \mu_j p_1 + p_2 &= 0 \tag{A.45} \\
\Sigma_j u^* + \lambda \xi_j + \mu_j p'_1 + p'_2 &= 0
\end{align*}$$

It follows that

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

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

References

