A. Proof of Lemma 4

Proof. Combining Lemma 2 and Lemma 3, we can obtain the concrete recursion
\[
    h_{t+1,i} \leq (1 - \sigma_{t,i})h_{t,i} + \sigma_{t,i}^2 D^2 + \eta t (1 + \sigma_2(P) \sqrt{n} + 1) L \sqrt{h_{t+1,i}}.
\]
As the parameters \( \eta_t \) and \( \sigma_{t,i} \) are chosen such that \( \eta_t (1 + \sigma_2(P) \sqrt{n} + 1) L \sqrt{h_{t+1,i}} \leq \sigma_{t,i}^2 D^2 \), we can then obtain the following compact recursion
\[
    h_{t+1,i} \leq (1 - \sigma_{t,i})h_{t,i} + 2D^2 \sigma_{t,i}^2.
\]
Now based on this recursion, we can prove the bound in the lemma by induction.

First, the base case of induction is true for \( t = 1 \) since by definition we have
\[
    h_{1,i} = F_{1,i}(x_i(1)) - F_{1,i}(x_i^*(1)) = \|x_i(1) - x_i^*(1)\|^2 - \|x_i^*(1) - x_1(1)\|^2 \leq 2D^2 \leq 4D^2 \sigma_{1,i}.
\]
Second, assuming that the bound is true for \( t \), we now show that it also holds for \( t + 1 \):
\[
    h_{t+1,i} \leq (1 - \sigma_{t,i})h_{t,i} + 2D^2 \sigma_{t+1,i}^2
    \leq 4D^2 \sigma_{t,i}^2(1 - \sigma_{t,i}) + 2D^2 \sigma_{t,i}^2
    = 4D^2 \sigma_{t,i}^2(1 - \sigma_{t,i} + \frac{\sigma_{t,i}}{2})
    = 4D^2 \sigma_{t,i}^2(1 - \frac{\sigma_{t,i}}{2})
    \leq 4D^2 \sigma_{t+1,i}.
\]
The last inequality follows from the definition of \( \sigma_{t,i} \), which can be proved in the following section. \( \Box \)

B. Proof of the last inequality in Lemma 4

For the sequence \( \sigma_{t,i} = \frac{1}{\sqrt{t}}, \) \( t = 1, 2, \ldots, T \), the following inequality holds
\[
    \sigma_{t,i} (1 - \frac{\sigma_{t,i}}{2}) \leq \sigma_{t+1,i}.
\]
Proof. The inequality we need to prove is
\[
    \frac{1}{\sqrt{t}} (1 - \frac{1}{2\sqrt{t}}) \leq \frac{1}{\sqrt{t+1}}.
\]
Note that, for the right side, we have the following identity
\[
    \frac{1}{\sqrt{t+1}} = \frac{1}{\sqrt{t}} \frac{1}{\sqrt{t+1}}.
\]
Thus, dividing both sides by the common \( \frac{1}{\sqrt{t}} \), we reach the following equivalent inequality
\[
    1 - \frac{1}{2\sqrt{t}} \leq \frac{\sqrt{t}}{\sqrt{t+1}}.
\]
By rewriting, we have
\[
    1 - \frac{1}{2\sqrt{t}} \leq 1 - \frac{\sqrt{t+1} - \sqrt{t}}{\sqrt{t+1}}.
\]
It then follows that
\[
    \frac{\sqrt{t+1} - \sqrt{t}}{\sqrt{t+1}} \leq \frac{1}{2\sqrt{t}}.
\]
Multiplying \( \sqrt{t+1} \sqrt{t} \) in both sides, we obtain
\[
    (\sqrt{t+1} - \sqrt{t}) \sqrt{t} \leq \frac{\sqrt{t+1}}{2},
\]
which is equivalent to the following
\[
    \sqrt{t^2 + t} \leq \frac{\sqrt{t+1}}{2} + t.
\]
Squaring both sides, we have
\[
    t^2 + t \leq t^2 + \frac{t+1}{4} + t\sqrt{t+1}.
\]
Clearly, this inequality holds for any \( t = 1, \ldots, T \), since
\[
    t \leq \frac{t+1}{4} + t\sqrt{t+1}.
\]
\( \Box \)

C. Proof of Lemma 6

Proof. We adopt the same notations used in the proof of Lemma 3. From there, we have
\[
    z_i(t) = \sum_{r=1}^{t-1} \sum_{j=1}^{n} P_{ij}^{t-r-1} g_j(r).
\]
To proceed, we first introduce another auxiliary sequence which are composed of the averages of the subgradients over all nodes $i$ at each iteration

$$
\bar{g}(t) = \frac{1}{n} \sum_{i=1}^{n} g_i(t).
$$

Then we can show that the averaged dual variable $\bar{z}(t)$ evolves in a quite simple way

$$
\bar{z}(t+1) = \frac{1}{n} \sum_{i=1}^{n} \left( \sum_{j=1}^{n} P_{ij} z_j(t) + g_i(t) \right)
$$

$$
= \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{n} P_{ij} z_j(t) + \bar{g}(t)
$$

$$
= \bar{z}(t) + \bar{g}(t).
$$

The last equation follows from the doubly stochastic property of matrix $P$. Based on the above recursion, we can easily deduce that

$$
\bar{z}(t) = \sum_{r=1}^{t-1} \bar{g}(r) = \frac{1}{n} \sum_{r=1}^{t-1} \sum_{j=1}^{n} g_j(r).
$$

Hence,

$$
z_i(t) - \bar{z}(t) = \sum_{r=1}^{t-1} \sum_{j=1}^{n} (P_{ij}^{t-r-1} - \frac{1}{n}) g_j(r).
$$

Then using the fact that $\|g_i(t)\| \leq L$, and the properties of norm functions and matrices, we obtain

$$
\|z_i(t) - \bar{z}(t)\|
$$

$$
= \left\| \sum_{r=1}^{t-1} \sum_{j=1}^{n} (P_{ij}^{t-r-1} - \frac{1}{n}) g_j(r) \right\|
$$

$$
\leq \sum_{r=1}^{t-1} \sum_{j=1}^{n} \left| P_{ij}^{t-r-1} - \frac{1}{n} \right| \left| g_j(r) \right|
$$

$$
\leq L \sum_{i=1}^{t-1} \sum_{j=1}^{n} \left\| P_{ij}^{t-r-1} - \frac{1}{n} \right\|_1
$$

$$
= L \sum_{i=1}^{t-1} \left\| P_{ij}^{t-r-1} e_i - \frac{1}{n} \right\|_1.
$$

Since the following inequality holds for any non-negative integer $s$

$$
\left\| P^s e_i - \frac{1}{n} \right\|_1 \leq \sigma_2(P)^s \sqrt{n},
$$

we have

$$
\|z_i(t) - \bar{z}(t)\| \leq L \sum_{r=1}^{t-1} \sigma_2(P)^{t-r-1} \sqrt{n}
$$

$$
= (1 - \sigma_2(P))^{t-1} \sqrt{n} L \frac{1 - \sigma_2(P)}{1 - \sigma_2(P)}
$$

$$
\leq \sqrt{n} L \frac{1 - \sigma_2(P)}{1 - \sigma_2(P)}.
$$

The above equation and the last inequality follow respectively from the summation formula of geometric series and the fact that $\sigma_2(P) < 1$ when $P$ is a doubly stochastic matrix (Berman & Plemmons, 1979).

\[ \square \]

\[ \textbf{D. Proof of Lemma 7} \]

**Proof.** According to (Hosseini et al., 2013), the D-ODA algorithm with parameters $\alpha(t)$ applied to loss functions that are $L$-Lipschitz with respect to a general norm attains the following regret bound

$$
R_{t}^2(x_i, x) \leq \frac{L^2}{2} \sum_{t=1}^{T-1} \alpha(t) + \frac{1}{\alpha(T)} \psi(x)
$$

$$
+ L \sum_{t=1}^{T} \alpha(t) \| z_i(t) - \bar{z}(t) \|_s
$$

$$
+ \frac{2L}{n} \sum_{t=1}^{T} \alpha(t) \sum_{j=1}^{n} \| z_j(t) - \bar{z}(t) \|_s,
$$

where $\| \cdot \|_s$ denotes the corresponding dual norm.

Note that the norm we utilize is the $L_2$ norm and its dual norm is itself. Thus we can apply the bound for $\| z_i(t) - \bar{z}(t) \|$ in Lemma 6 here. Combining it with the fact that $\sum_{t=1}^{T-1} \alpha(t) \leq \sum_{t=1}^{T} \alpha(t)$, the fact that $\psi(x) = \| x - x_1 \|_2^2 \leq D^2$ and setting $\alpha(t) = \eta$ yields the stated regret bound in the lemma. \[ \square \]

\[ \textbf{E. Verification of the validity of } \eta_i \]

**Proof.** As $\eta_i = \frac{(1 + \sigma_2(P)) D}{2(\sqrt{n} + 1 + (\sqrt{n} - 1) \sigma_2(P)) L T 3^{3/4}}$, we have

$$
\eta_i \left( \frac{1 + \sigma_2(P)}{1 - \sigma_2(P)} \sqrt{n} + 1 \right) L \sqrt{h_{t+1, i}} = \frac{D \sqrt{h_{t+1, i}}}{2 T^{3/4}}.
$$

By Lemma 4 and definition of $\sigma_{t, i}$, we have

$$
h_{t+1, i} \leq 4 D^2 \sigma_{t+1, i} \leq 4 D^2 \sigma_{t, i}.
$$

It then follows that

$$
\frac{D \sqrt{h_{t+1, i}}}{2 T^{3/4}} \leq \frac{\sigma_{t, i}^{1/2}}{T^{3/4} D^2}.
$$
We thus only need to verify that the following inequality holds for any $t = 1, \cdots, T$

$$\frac{\sigma_{t,i}^{1/2}}{T^{3/4}} D^2 \leq \sigma_{t,i}^2 D^2.$$ 

This clearly holds since for any $t = 1, \cdots, T$

$$\frac{1}{T^{3/4}} \leq \sigma_{t,i}^{3/2} = \frac{1}{T^{3/4}}.$$ 

Thus, the choice of $\eta_i$ satisfies the constraint required in Lemma 4.

References
