A PROOF OF LEMMA 2

Proof. The proof technique is standard, and can be found in Zinkevich (2003); Hazan et al. (2016).

First, we prove the regret bound of (21). Note that by Definition 2, \( s_t^2(x) \) is \( 2\eta^2G^2 \)-strongly convex. For convince, we denote \( \alpha_{t+1} = 1/(2\eta^2G^2t) \), \( \lambda^s = 2\eta^2G^2 \), and define the upper bound of the gradients of \( s_t^2(x) \) as

\[
\max_{x \in D} \| \nabla s_t^2(x) \| = \max_{x \in D} \| \eta g_t + 2\eta^2G^2(x - x_t) \| \leq G\eta + 2\eta^2G^2D =: G^s.
\]

By the update rule of \( x_t^{n,s} \), we have

\[
\| x_t^{n,s} - u \| = \| \Pi_D^G (x_t^{n,s} - \alpha_{t+1} \nabla s_t^2(x_t^{n,s})) - u \|
\leq \| x_t^{n,s} - \alpha_{t+1} \nabla s_t^2(x_t^{n,s}) - u \|
\leq \| x_t^{n,s} - u \|^2 + \alpha_{t+1} \| \nabla s_t^2(x_t^{n,s}) \|^2 - 2\alpha_{t+1} (x_t^{n,s} - u)^T \nabla s_t^2(x_t^{n,s}).
\]

Hence,

\[
2(x_t^{n,s} - u)^T \nabla s_t^2(x_t^{n,s}) \leq \frac{\| x_t^{n,s} - u \| - \| x_t^{n,s} - u \|^2}{\alpha_{t+1}} + \alpha_{t+1}(G^s)^2.
\]

Summing over 1 to \( T \) and applying definition 2, we get

\[
2 \sum_{t=1}^{T} s_t^2(x_t^{n,s}) - 2 \sum_{t=1}^{T} s_t^2(u) \leq \sum_{t=1}^{T} \| x_t^{n,s} - u \|^2 \left( \frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t} - \lambda^s \right) + (G^s)^2 \sum_{t=1}^{T} \alpha_{t+1}
\leq \frac{(G^s)^2}{\lambda^s}(1 + \log T).
\]

Note that \( \eta \leq \frac{1}{5DG} \). We have

\[
(G^s)^2 = G^2\eta^2 + 4\eta^3G^3D + 4\eta^4G^4D^2 \leq G^2\eta^2 + \frac{4\eta^2G^2}{5} + \frac{4\eta^2G^2}{25} \leq 2\eta^2G^2 = \lambda^s.
\]

Next, we prove the regret bound of (22). We start with the following inequality

\[
\nabla \ell_t^0(x)(\nabla \ell_t^0(x))^T = \eta^2\xi_t g_t^T + 4\eta^3\xi_t (x - x_t)^T g_t^T + 4\eta^4\xi_t (x - x_t)(x - x_t)^T g_t^T
\leq \eta^2\xi_t g_t^T + \xi_t \left( 4\eta^3(x - x_t)^T g_t + 4\eta^4((x - x_t)^T g_t)^2 \right) g_t^T
\leq 2\eta^2\xi_t g_t^T = \nabla^2 \ell_t^0(x)
\]

where \( \nabla^2 \ell_t^0(x) \) denotes the Hessian matrix. The inequality implies that \( \nabla^2 \ell_t^0(x) \geq \nabla \ell_t^0(x)(\nabla \ell_t^0(x))^T \). According to Lemma 4.1 in Hazan et al. (2016), \( \ell_t^0(x) \) is 1-exp-concave. Next, we prove that the gradient of \( \ell_t^0(x) \) can be upper bounded as follows

\[
\max_{x \in D} \| \nabla \ell_t^0(x) \| \leq \eta G + 2\eta^2G^2D \leq \frac{7}{25D} = G^d.
\]

By Theorem 4.3 in Hazan et al. (2016), we have

\[
\sum_{t=1}^{T} \ell_t^0(x_t^{n,t}) - \sum_{t=1}^{T} \ell_t^0(u) \leq 5(1 + G^d D) d \log T \leq 10d \log T.
\]

Finally, we prove the regret bound of (23). Note that the gradient of \( c_t(x) \) is upper bounded by \( \max_{x \in D} \| \nabla c_t(x) \| \leq \eta^s G \). Define \( m_t = \frac{\ell_t^0}{\eta^s G^2} \). By the convexity of \( c_t(x) \), we have \( \forall u \in D \),

\[
c_t(x_t^{n,t}) - c_t(u) \leq (x_t^{n,t} - u)^T \nabla c_t(x_t^{n,t}).
\]
On the other hand, according to the update rule of $x^{c}_{t+1}$, we have

$$
\|x^{c}_{t+1} - u\|^2 = \|\Pi_{D}^{c}(x^{c}_{t} - m_{t}\nabla c_{t}(x^{c}_{t})) - u\|^2 \\
\leq \|x^{c}_{t} - m_{t}\nabla c_{t}(x^{c}_{t}) - u\|^2 \\
= \|x^{c}_{t} - u\|^2 + m_{t}^{2}\|\nabla c_{t}(x^{c}_{t})\|^2 - 2m_{t}(x^{c}_{t} - u)^{T}\nabla c_{t}(x^{c}_{t})
$$

(36)

where the inequality follows from Theorem 2.1 in Hazan et al. (2016). Hence,

$$
2(x^{c}_{t} - u)^{T}\nabla c_{t}(x^{c}_{t}) \leq \|x^{c}_{t} - u\|^2 - \|x^{c}_{t+1} - u\|^2 + m_{t}\|\nabla c_{t}(x^{c}_{t})\|^2 \\
\leq \|x^{c}_{t} - u\|^2 - \|x^{c}_{t+1} - u\|^2 + m_{t}(\eta^{c}G)^{2}
$$

(37)

Substituting the above inequality into (35) and summing over $T$, we have

$$
\sum_{t=1}^{T} c_{t}(x^{c}_{t}) - c_{t}(u) \leq \sum_{t=1}^{T} (x^{c}_{t} - u)^{T}\nabla c_{t}(x^{c}_{t}) \\
\leq \frac{1}{2} \sum_{t=1}^{T} \|x^{c}_{t} - u\|^2 \left( \frac{1}{m_{t}} - \frac{1}{m_{t-1}} \right) + \frac{(\eta^{c}G)^{2}}{2} \sum_{t=1}^{T} m_{t} \\
\leq D^{2} \frac{1}{2m_{T}} + \frac{(\eta^{c}G)^{2}}{2} \sum_{t=1}^{T} m_{t} \\
\leq \frac{3}{2}\eta^{c}GD\sqrt{T} \leq \frac{3}{4}
$$

(38)

where the last inequality is due to $\eta^{c} = \frac{1}{2GD\sqrt{T}}$. \qed