A. Discussions on one point convexity

If f is δ-one point strongly convex around x^* in a convex domain $\mathcal{D}$, then x^* is the only local minimum point in $\mathcal{D}$ (i.e., global minimum).

To see this, for any fixed $x \in \mathcal{D}$, look at the function $g(t) = f(tx^* + (1-t)x)$ for $t \in [0,1]$, then $g'(t) = \langle \nabla f(tx^* + (1-t)x), x^* - x \rangle$. The definition of δ-one point strongly convex implies that the right side is negative for $t \in (0,1]$. Therefore, $g(t) > g(1)$ for $t > 0$. This implies that for every point y on the line segment joining x to x^*, we have $f(y) > f(x^*)$, so x^* is the only local minimum point.

B. Proof for Lemma 5

Proof. Recall that we want to show

$$\frac{\sqrt{2}(3.5\eta^2 r^2 + 7\eta r\delta)}{\sqrt{\eta c}} \log^{\frac{1}{2}}(\zeta) + \frac{21b}{\lambda} \le \delta^2 = \frac{\mu^2 b}{\lambda} = \frac{\mu^2\eta^2 r^2(1+\eta L)^2}{\lambda}$$

On the left hand side there are three summands. Below we show that each of them is bounded by $\frac{\mu^2 b}{3\lambda}$[6].

Since $\mu \ge \max\{8, 42\log^{\frac{1}{2}}(\zeta)\}$, we know $\frac{21b}{\lambda} \le \frac{63b}{3\lambda} < \frac{8^2 b}{3\lambda} \le \frac{\mu^2 b}{3\lambda}$. Next, we have

$$\begin{aligned}
&42\log^{\frac{1}{2}}(\zeta) \le \mu \\
\Rightarrow&\sqrt{30\log^{\frac{1}{2}}(\zeta)\eta^{0.5}c^{0.5}} \le \mu \\
\Rightarrow&15\log^{\frac{1}{2}}(\zeta) \le \frac{\mu^2}{2\eta^{0.5}c^{0.5}} \\
\Rightarrow&\frac{15}{\sqrt{c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\mu^2\eta^{0.5}}{\lambda} \\
\Rightarrow&\frac{3.5\sqrt{2}\eta^{1.5}r^2}{\sqrt{c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\mu^2\eta^2 r^2}{3\lambda} \\
\Rightarrow&\frac{3.5\sqrt{2}\eta^2 r^2}{\sqrt{\eta c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\mu^2\eta^2 r^2(1+\eta L)^2}{3\lambda}
\end{aligned}$$

Finally,

$$\begin{aligned}
&42\log^{\frac{1}{2}}(\zeta) \le \mu \\
\Rightarrow&\frac{42}{\sqrt{c}}\log^{\frac{1}{2}}(\zeta) \le \mu\sqrt{\frac{1}{c}} \\
\Rightarrow&\frac{7\sqrt{2}\eta r}{\sqrt{\eta c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\mu\sqrt{\frac{\eta^2 r^2(1+\eta L)^2}{2\eta c}}}{3} \\
\Rightarrow&\frac{7\sqrt{2}\eta r}{\sqrt{\eta c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\delta}{3} \\
\Rightarrow&\frac{7\sqrt{2}\eta r\delta}{\sqrt{\eta c}}\log^{\frac{1}{2}}(\zeta) \le \frac{\delta^2}{3}
\end{aligned}$$

Adding the three summands together, we get the claim. □

[6] We made no effort to optimize the constants here.

C. Proof for Theorem 3

Proof. Recall that we have $x_{t+1} = x_t - \eta \nabla f(x_t)$. Since we have $\langle -\nabla f(x_t), x^* - x_t \rangle \leq c' \|x^* - x_t\|_2^2$, then

$$\begin{aligned}
\|x_{t+1} - x^*\|_2^2 &= \|x_t - \eta \nabla f(x_t) - x^*\|_2^2 \\
&= \|x_t - x^*\|_2^2 + \eta^2 \|\nabla f(x_t)\|_2^2 - 2\eta \langle \nabla f(x_t), x_t - x^* \rangle \\
&\geq (1 - 2\eta c') \|x_t - x^*\|_2^2 + \eta^2 \|\nabla f(x_t)\|_2^2 > \|x_t - x^*\|_2^2
\end{aligned}$$

Where the last inequality holds since we know $\eta > \frac{2c' \|x_t - x^*\|_2^2}{\|\nabla f(x_t)\|_2^2}$. □