

[Case i] The following error bounds hold true for some $M > 0$.

(i-1) $0 \leq \mathbb{E}[h(\hat{x}_\eta)] - h^* \leq \mathbf{Err}_\eta + \eta M$.

(i-2) $\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] - \|\nabla f(x^*)\| \mathbb{E}[\text{dist}(\hat{x}_\eta, X_h^*)] \leq \mathbb{E}[f(\hat{x}_\eta)] - f^* \leq \frac{1}{\eta} \mathbf{Err}_\eta$.

(i-3) If $\mu_f > 0$, then $\mathbb{E}[\|\hat{x}_\eta - x^*\|^2] \leq \frac{2}{\mu_f} \left(\|\nabla f(x^*)\| \mathbb{E}[\text{dist}(\hat{x}_\eta, X_h^*)] + \frac{1}{\eta} \mathbf{Err}_\eta \right)$.

[Case ii] Suppose X_h^* is α -weak sharp with order $\kappa \geq 1$. Then, the following holds for some $M > 0$.

(ii-1) $0 \leq \mathbb{E}[h(\hat{x}_\eta)] - h^* \leq \mathbf{Err}_\eta + \eta M$.

(ii-2) $\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] - \|\nabla f(x^*)\| \sqrt{\frac{1}{\alpha} (\mathbf{Err}_\eta + \eta M)} \leq \mathbb{E}[f(\hat{x}_\eta)] - f^* \leq \frac{1}{\eta} \mathbf{Err}_\eta$.

(i-3) If $\mu_f > 0$, then $\mathbb{E}[\|\hat{x}_\eta - x^*\|^2] \leq \frac{2}{\mu_f} \left(\|\nabla f(x^*)\| \sqrt{\frac{1}{\alpha} (\mathbf{Err}_\eta + \eta M)} + \frac{1}{\eta} \mathbf{Err}_\eta \right)$.

[Case iii] If X_h^* is α -weak sharp with order $\kappa = 1$ and $\eta \leq \frac{\alpha}{2\|\nabla f(x^*)\|}$, then the following holds.

(iii-1) $0 \leq \mathbb{E}[h(\hat{x}_\eta)] - h^* \leq 2 \mathbf{Err}_\eta$.

(iii-2) $\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] - \frac{2\|\nabla f(x^*)\|}{\alpha} \mathbf{Err}_\eta \leq \mathbb{E}[f(\hat{x}_\eta)] - f^* \leq \frac{1}{\eta} \mathbf{Err}_\eta$.

(iii-3) If $\mu_f > 0$, then $\mathbb{E}[\|\hat{x}_\eta - x^*\|^2] \leq \frac{2}{\eta\mu_f} \mathbf{Err}_\eta$.

Proof. (i-1) In view of the definition of \mathbf{Err}_η and f_η , we have $\mathbb{E}[h(\hat{x}_\eta) + \eta f(\hat{x}_\eta)] - f_\eta^* \leq \mathbf{Err}_\eta$. Also, from the definition of f_η^* , we have $f_\eta^* \leq f_\eta(x^*) = h^* + \eta f^*$. From the preceding two relations, we obtain

$$\mathbb{E}[h(\hat{x}_\eta)] - h^* + \eta(\mathbb{E}[f(\hat{x}_\eta)] - f^*) \leq \mathbf{Err}_\eta. \quad (2)$$

By invoking Assumption 1, we have $\mathbb{E}[f(\hat{x}_\eta)] > -\infty$. Therefore, there exists some $M > 0$ such that $f^* - \mathbb{E}[f(\hat{x}_\eta)] < M$. As a result, we obtain $\mathbb{E}[h(\hat{x}_\eta)] - h^* \leq \mathbf{Err}_\eta + \eta M$. From $\hat{x}_\eta \in X$, we also have $\mathbb{E}[h(\hat{x}_\eta)] - h^* \geq 0$. This completes the proof of (i-1).

(i-2) The upper bound holds in view of (2) and that $\mathbb{E}[h(\hat{x}_\eta)] - h^* \geq 0$. To show the lower bound, from the convexity of f , we may write

$$\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] + \nabla f(x^*)^\top \mathbb{E}[(\hat{x}_\eta - x^*)] \leq \mathbb{E}[f(\hat{x}_\eta)] - f^*,$$

where the expectation is taken with respect to the random variables generated in the method \mathcal{M} . Note that $\nabla f(x^*)^\top \mathbb{E}[(\hat{x}_\eta - x^*)]$ is not necessarily nonnegative. We may write

$$\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] + \nabla f(x^*)^\top \mathbb{E}[(\hat{x}_\eta - \Pi_{X_h^*}[\hat{x}_\eta] + \Pi_{X_h^*}[\hat{x}_\eta] - x^*)] \leq \mathbb{E}[f(\hat{x}_\eta)] - f^*.$$

In view of $\Pi_{X_h^*}[\hat{x}_\eta] \in X_h^*$, we have $\nabla f(x^*)^\top \mathbb{E}[(\Pi_{X_h^*}[\hat{x}_\eta] - x^*)] \geq 0$. We obtain

$$\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] + \nabla f(x^*)^\top \mathbb{E}[(\hat{x}_\eta - \Pi_{X_h^*}[\hat{x}_\eta])] \leq \mathbb{E}[f(\hat{x}_\eta)] - f^*$$

Invoking the Cauchy-Schwarz inequality, we obtain

$$\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] - \|\nabla f(x^*)\| \left\| \mathbb{E}[\hat{x}_\eta - \Pi_{X_h^*}[\hat{x}_\eta]] \right\| \leq \mathbb{E}[f(\hat{x}_\eta)] - f^*.$$

Invoking the Jensen's inequality, we obtain

$$\frac{\mu_f}{2} \mathbb{E}[\|\hat{x}_\eta - x^*\|^2] - \|\nabla f(x^*)\| \mathbb{E} \left[\left\| \hat{x}_\eta - \Pi_{X_h^*}[\hat{x}_\eta] \right\| \right] \leq \mathbb{E}[f(\hat{x}_\eta)] - f^*.$$

Noting that $\mathbb{E}[\text{dist}(\hat{x}_\eta, X_h^*)] = \mathbb{E} \left[\left\| \hat{x}_\eta - \Pi_{X_h^*}[\hat{x}_\eta] \right\| \right]$, we obtain the lower bound in (i-2).

(i-3) This result follows directly from (i-2).

(ii-1) This result is identical to (i-1).