**Proof** With stepsize $\eta_t = \frac{1}{\lambda t}$ plugged into (4.6), we have

$$\mathbb{E}[f(\boldsymbol{w}^t)] - f(\boldsymbol{w}^*) \leq \frac{1}{2\lambda t}\mathbb{E}[\|\boldsymbol{g}_{i_t}^t\|^2] + \frac{\lambda(t-1)}{2}\mathbb{E}[\|\boldsymbol{w}^t - \boldsymbol{w}^*\|^2] - \frac{\lambda t}{2}\mathbb{E}[\|\boldsymbol{w}^{t+1} - \boldsymbol{w}^*\|^2] \tag{4.7}$$

We use convexity of the function $f$, as given by Jensen's inequality:

$$\mathbb{E}\left[f\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{w}^t\right)\right] - f(\boldsymbol{w}^*) \overset{Jensen}{\leq} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} f(\boldsymbol{w}^t)\right] - f(\boldsymbol{w}^*)$$

Summing up (4.7) over all steps $t = 1 \dots T$, we can bound the right hand side of the above inequality

$$= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[f(\boldsymbol{w}^t)] - f(\boldsymbol{w}^*) \leq \frac{1}{T}\sum_{t=1}^{T}\frac{\mathbb{E}[\|\boldsymbol{g}_{i_t}^t\|^2]}{2\lambda}\frac{1}{t} - \frac{\lambda}{2}\mathbb{E}[\|\boldsymbol{w}^{T+1} - \boldsymbol{w}^*\|^2]$$

(where we have used the telescoping sum of the norm terms.)

Re-arranging terms, and trivially bounding the left hand side of Jensen's inequality by $0 \leq \mathbb{E}\left[f\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{w}^t\right)\right] - f(\boldsymbol{w}^*)$, we obtain the claimed bound

$$\mathbb{E}[\|\boldsymbol{w}^{T+1} - \boldsymbol{w}^*\|^2] \leq \frac{1}{\lambda^2 T}\sum_{t=1}^{T}\frac{\mathbb{E}[\|\boldsymbol{g}_{i_t}^t\|]^2}{t}.$$

$\square$