Problem 1.

(a)

$$f(y-\frac{1}{L}\nabla f(y)) - f(x)$$

$$\leq f(y-\frac{1}{L}\nabla f(y)) - f(y) - \nabla f(y)^T(x-y)$$

$$\leq f(y-\frac{1}{L}\nabla f(y)) - f(y) - \nabla f(y)^T(x-y)$$

$$\leq \nabla f(y)(y-\frac{1}{L}\nabla f(y)-y) + \frac{L}{2}\|y-\frac{1}{L}\nabla f(y)-y\|^2$$
$$-\nabla f(y)^T(x-y)$$

$$\leq -\frac{1}{L}\|\nabla f(y)\|^2 + \frac{1}{2L}\|\nabla f(y)\|^2 - \nabla f(y)^T(x-y)$$

$$\leq -\frac{1}{2L}\|\nabla f(y)\|^2 - \nabla f(y)^T(x-y)$$ #

(b) using the results in (a)

$$f(x_{t+1}) - f(x_t) \leq -\frac{1}{2L}\|\nabla f(y_t)\|^2 - \nabla f(y_t)(x_t - y_t)$$

According to (1), we have $x_{t+1} = y_t - \frac{1}{L}\nabla f(y_t)$

$$\Rightarrow \nabla f(y_t) = L\cdot(y_t - x_{t+1})$$

so it will be $\leq -\frac{L}{2}\|x_{t+1}-y_t\|^2 + L\cdot(x_{t+1}-y_t)^T(x_t-y_t)$ #

(c) As we do not use any property of x_t we can simply use what we get on (b) by substitute x_t with x^*

$$\Rightarrow f(x_{t+1}) - f(x^*) \leq -\frac{L}{2}\|x_{t+1}-y_t\|^2 + L\cdot(x_{t+1}-y_t)^T(x^*-y_t)$$

\#

(d)

L.H.S : From (3) we have $\theta_{t+1}^2 - \theta_{t+1} - \theta_t^2 = 0$

$$\Rightarrow \theta_{t+1} = \theta_{t+1}^2 - \theta_t^2$$

$$\theta_t(\theta_t - 1)[f(x_{t+1}) - f(x_t)] + \theta_t[f(x_{t+1}) - f(x^*)]$$

$$= \theta_t^2[f(x_{t+1}) - f(x_t)] + \theta_t[\cancel{f(x_{t+1})} - f(x^*) \cancel{- f(x_{t+1})} + f(x_t)]$$

$$= \theta_t^2[f(x_{t+1}) - f(x_t)] + (\theta_t^2 - \theta_{t-1}^2)[f(x_t) - f(x^*)]$$

$$= \theta_t^2[f(x_{t+1}) \cancel{- f(x_t) + f(x_t)} - f(x^*)] - \theta_{t-1}^2(f(x_t) - f(x^*))$$

$$= \theta_t^2 \Delta_{t+1} - \theta_{t-1}^2 \Delta_t$$

R.H.S :

$$\theta_t(\theta_t - 1)\left[-\frac{L}{2}\|x_{t+1}-y_t\|^2 + L(x_{t+1}-y_t)^T(x_t - y_t)\right]$$

$$+ \theta_t\left[-\frac{L}{2}\|x_{t+1}-y_t\|^2 + L(x_{t+1}-y_t)^T(x^* - y_t)\right]$$

$$= -\frac{L}{2}\|\theta_t(x_{t+1}-y_t)\|^2$$

$$+ L\,\theta_t(x_{t+1}-y_t)^T[\theta_t(x_t-y_t) - x_t + y_t + x^* - y_t]$$

$$= -\frac{L}{2}\left(\|\theta_t(x_{t+1}-y_t)\|^2 + 2\theta_t(x_{t+1}-y_t)^T(\theta_t y_t - (\theta_t-1)x_t - x^*)\right)$$

Combine both Side:

$$\theta_t^2\Delta_{t+1} - \theta_{t-1}^2\Delta_t \le -\frac{L}{2}(\|\theta_t(x_{t+1}-y_t)\|^2 + 2\theta_t(x_{t+1}-y_t)^T(\theta_t y_t - (\theta_t-1)x_t - x^*))$$

#

(e)

on R.H.S.

$$\|\theta_t(x_{t+1}-y_t)\|^2 + 2\theta_t(x_{t+1}-y_t)^T(\theta_t y_t - (\theta_t-1)x_t - x^*))$$

$$= (\theta_t x_{t+1} - \theta_t y_t)^T(\theta_t x_{t+1} - \theta_t y_t + 2\theta_t y_t - 2\theta_t x_t + 2x_t - 2x^*)$$

$$= \|\theta_t x_{t+1} - (\theta_t-1)x_t - x^*\|^2 - \|\theta_t y_t - (\theta_t-1)x_t - x^*\|^2$$

From (2) we have

$$\theta_{t+1}y_{t+1} - (\theta_{t+1}-1)x_{t+1} = \theta_t x_{t+1} - (\theta_t-1)x_t$$

$$= \|\theta_{t+1}y_{t+1} - (\theta_{t+1}-1)x_{t+1} - x^*\|^2 - \|\theta_t y_t - (\theta_t-1)x_t - x^*\|^2$$

$$= \|\phi_{t+1}\|^2 - \|\phi_t\|^2$$

And we will have

$$\theta_t^2\Delta_{t+1} - \theta_{t-1}^2\Delta_t \le \frac{L}{2}(\|\phi_t\|^2 - \|\phi_{t+1}\|^2)$$ #

(f) With telescoping Sum and $\begin{cases} \theta_0 = 0 \\ x_1 = y_1 \end{cases}$

$$\Rightarrow \theta_{t-1}^2 \Delta_t - \theta_0^2 \Delta_1 \le \frac{L}{2}(\|\phi_1\|^2 - \|\phi_t\|^2) \le \frac{L}{2}\|\phi_1\|^2$$

$$\phi_1 = \theta_1 y_1 - (\theta_1 - 1)x_1 - x^* = x_1 - x^*, \quad \theta_0^2 \Delta_1 = 0$$

$$\Rightarrow \theta_{t-1}^2 \Delta_t \le \frac{L}{2}\|x_1 - x^*\|^2$$

Assume $\theta_t \ge \frac{t+1}{2}$

we have (3) $\theta_t = \frac{1+\sqrt{1+4\theta_{t-1}^2}}{2}$

By induction

$$\theta_{t+1} = \frac{1+\sqrt{1+(t+1)^2}}{2} \ge \frac{1+\sqrt{(t+1)^2}}{2}$$

$$\Rightarrow \Delta_t = f(x_t) - f(x^*) \le \frac{2L\|x_1 - x^*\|^2}{t^2}$$

as we have $\theta_{t-1} \ge \frac{t}{2}$

#

Problem 2.

We have

$$D_\phi(z\|\bar{x}) + D_\phi(\bar{x}\|x)$$

$$= \phi(z) - \cancel{\phi(\bar{x})} - \nabla\phi(\bar{x})(z-\bar{x})$$

$$+ \cancel{\phi(\bar{x})} - \phi(x) - \nabla\phi(x)(\bar{x}-x)$$

$$= \phi(z) - \phi(x) - \nabla\phi(x)(z-x) + \nabla\phi(x)(z-x)$$

$$- \nabla\phi(\bar{x})(z-\bar{x}) - \nabla\phi(x)(\bar{x}-x)$$

$$\leq D_\phi(z\|x) + \nabla\phi(x)(z-x) + \nabla\phi(\bar{x})(z-\bar{x}) + \nabla\phi(x)(\bar{x}-x)$$

$$\leq D_\phi(z\|x) + \phi(z) - \phi(x) - \phi(z) + \phi(\bar{x}) - \phi(\bar{x}) + \phi(x)$$

$$= D_\phi(z\|x)$$

#