PF 2014 *

5a) $\mathcal{E}(h_0) \leq \mathcal{E}(h^*) + \eta$ (because optimal h^* better than $\hat{h}$)

$\mathcal{E}(h_0) \leq \mathcal{E}(\hat{h}) + \eta$

Know: $\mathcal{E}(\hat{h}) \leq \hat{\mathcal{E}}(\hat{h}) + \gamma$ wp. $p \geq 1-\delta$

$\Rightarrow \mathcal{E}(h_0) \leq \hat{\mathcal{E}}(\hat{h}) + \eta + \gamma$

$\mathcal{E}(h_0) + \gamma \leq \hat{\mathcal{E}}(\hat{h}) + \eta + 2\gamma$

Know: $\hat{\mathcal{E}}(h_0) \leq \mathcal{E}(h_0) + \gamma$ (uniform convergence)

~~or~~ $\cancel{\mathcal{E}(h_0) \leq \hat{\mathcal{E}}(h_0) + \gamma}$

$\cancel{\Rightarrow \mathcal{E}(h_0) - \gamma \leq \hat{\mathcal{E}}(h_0) \leq}$

$\cancel{\Rightarrow}$

$\Rightarrow \hat{\mathcal{E}}(h_0) \leq \hat{\mathcal{E}}(\hat{h}) + \eta + 2\gamma$ wp $\geq 1-\delta$

$\Rightarrow \boxed{\hat{\mathcal{E}}(h_0) > \hat{\mathcal{E}}(\hat{h}) + \eta + 2\gamma \quad \text{wp} \leq \delta}$

PE 2014

b) If $\hat{\mathcal{E}}(h_0) < \hat{\mathcal{E}}(\hat{h}) + \eta - 2\gamma \Rightarrow$ yes

$$\mathcal{E}(h_0) > \mathcal{E}(h^*) + \eta$$

Know: $\hat{\mathcal{E}}(h) \leq \mathcal{E}(h) + \gamma$ (uniform conv)

$\Rightarrow \hat{\mathcal{E}}(h) - \gamma \leq \mathcal{E}(h)$ w.p. $1-\delta$

$\Rightarrow \mathcal{E}(h_0) > \hat{\mathcal{E}}(h^*) - \gamma + \eta$ wp $1-\delta$

$\Rightarrow \mathcal{E}(h_0) > \hat{\mathcal{E}}(\hat{h}) - \gamma + \eta$ (bc $\hat{h}$ has lower error on training set)

$\cancel{= \mathcal{E}(h_0) + \gamma > \hat{\mathcal{E}}(\hat{h})}$

Know: $\hat{\mathcal{E}}(h) \leq \mathcal{E}(h) + \gamma$

also, analog: $\mathcal{E}(h) \leq \hat{\mathcal{E}}(h) + \gamma$

$\Rightarrow$

$$\mathcal{E}(h_0) - \gamma > \hat{\mathcal{E}}(\hat{h}) - 2\gamma + \eta$$

$\Rightarrow \hat{\mathcal{E}}(h) + \gamma - \gamma \geq \hat{\mathcal{E}}(\hat{h}) + \eta - 2\gamma$ wp $1-\delta$

$\Rightarrow \boxed{\hat{\mathcal{E}}(h_0) < \hat{\mathcal{E}}(\hat{h}) + \eta - 2\gamma \text{ wp } \delta}$