

$$(F_1, \phi_1) := \operatorname{argmin}_{F, \phi} \mathbb{E}_{x \in \mathcal{D}} \mathcal{L}_1(F \circ \phi(x), F^* \circ \phi^*(x))$$

$$(F_2, \phi_2, K_2) := \operatorname{argmin}_{F, \phi} \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F \circ \phi(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K \circ \phi(x), K^* \circ \phi^*(x))]$$

$$K_3 := \operatorname{argmin}_K \mathbb{E}_{x \in \mathcal{D}} \mathcal{L}_2(K \circ \phi_1(x), K^* \circ \phi^*(x))$$

$$\begin{aligned} & \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F_2 \circ \phi_2(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K_2 \circ \phi_2(x), K^* \circ \phi^*(x))] \\ = & \min_{F, \phi, K} \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F \circ \phi(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K \circ \phi(x), K^* \circ \phi^*(x))] \\ \leq & \min_{K, \phi = \phi_1, F = F_1} \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F \circ \phi(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K \circ \phi(x), K^* \circ \phi^*(x))] \\ = & \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F_1 \circ \phi_1(x), F^* \circ \phi^*(x))] + \min_K \mathbb{E}_{x \in \mathcal{D}} \alpha \cdot \mathcal{L}_2(K \circ \phi_1(x), K^* \circ \phi^*(x)) \\ = & \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F_1 \circ \phi_1(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K_3 \circ \phi_1(x), K^* \circ \phi^*(x))] \\ \leq & \mathbb{E}_{x \in \mathcal{D}} [\mathcal{L}_1(F_2 \circ \phi_2(x), F^* \circ \phi^*(x)) + \alpha \cdot \mathcal{L}_2(K_3 \circ \phi_1(x), K^* \circ \phi^*(x))] \end{aligned}$$

The second inequality holds because of the first definition above.