Figure showing training curves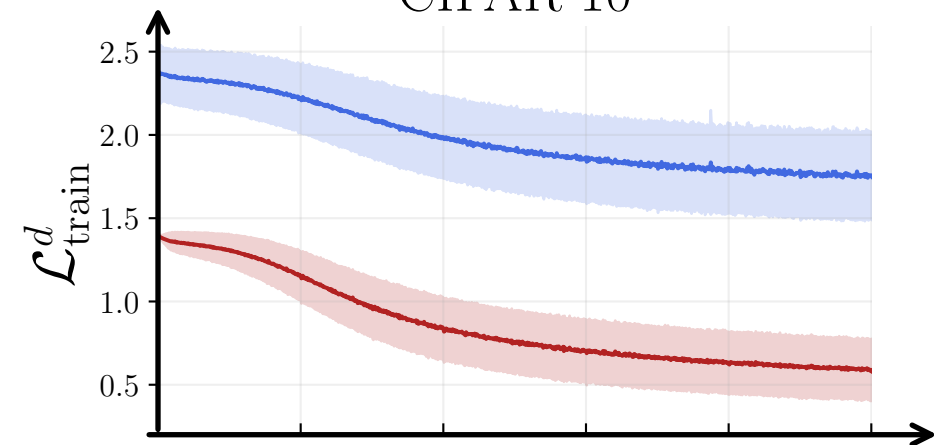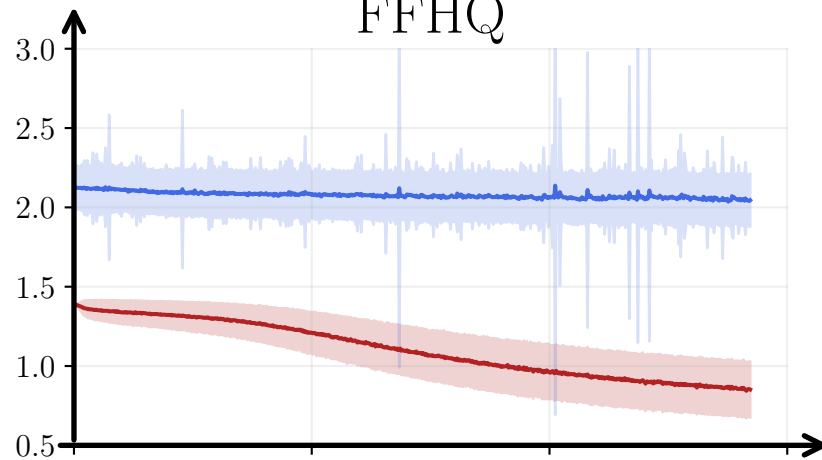 for CIFAR-10, FFHQ, and AFHQv2 comparing DG: $\mathcal{L}_{\text{train}}^d = \mathcal{L}_{\text{CE}}^d$ and Ours: $\mathcal{L}_{\text{train}}^d = \mathcal{L}_{\text{MSE}}^d + \gamma \mathcal{L}_{\text{CE}}^d$