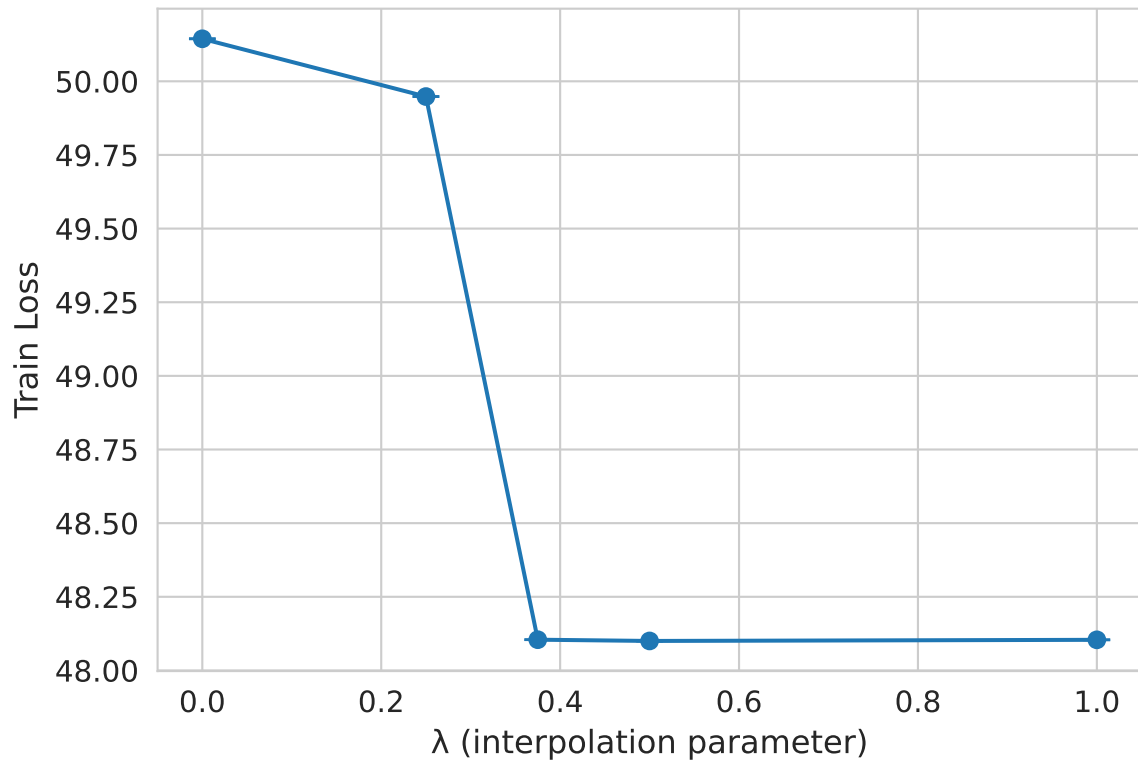


Initialization Effect on Distillation
($\theta_s = \lambda \cdot \theta_t + (1-\lambda) \cdot \theta_r$, Self-distillation, PreResNet-20)

Final Train Loss



Final Train Agreement

