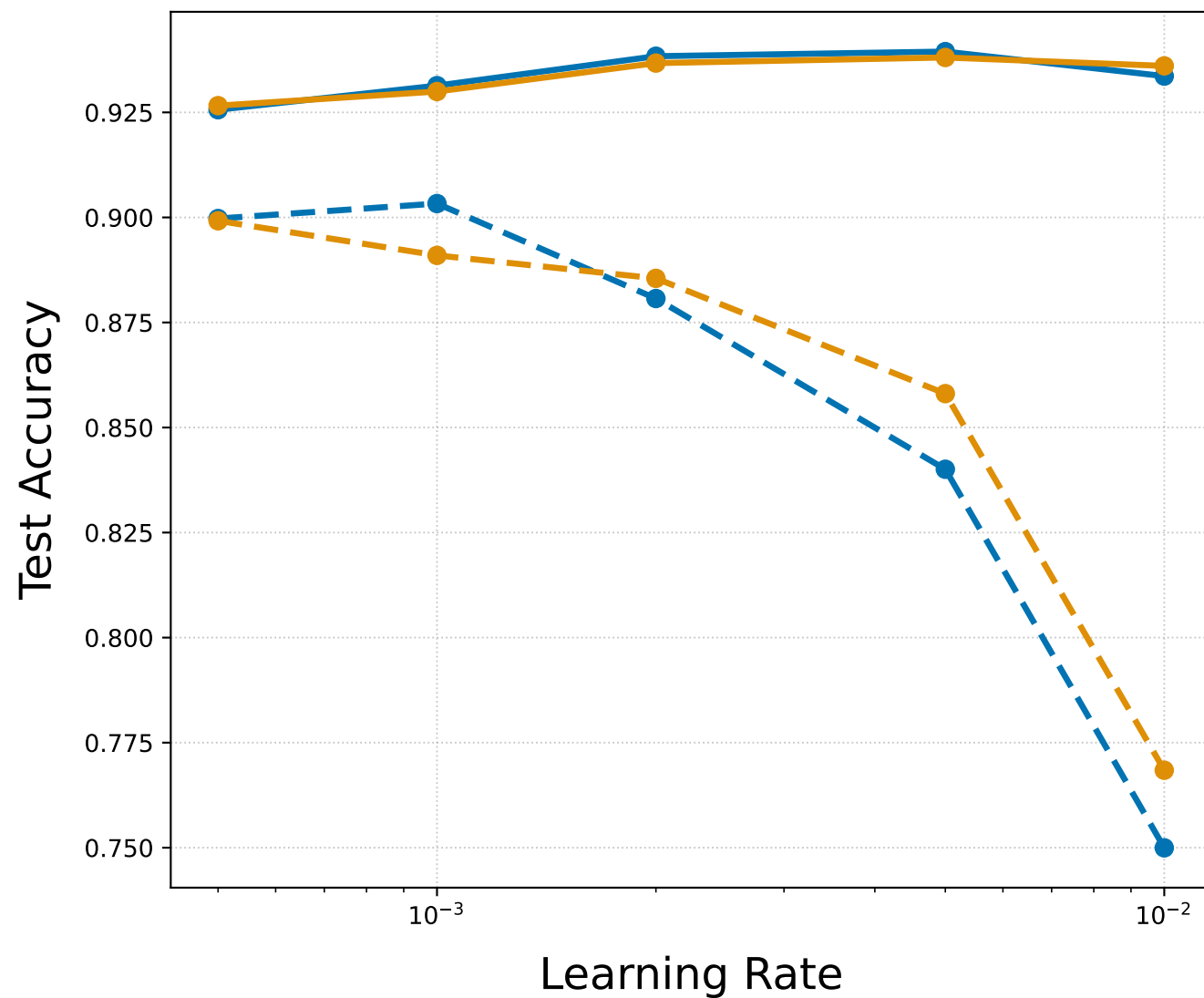
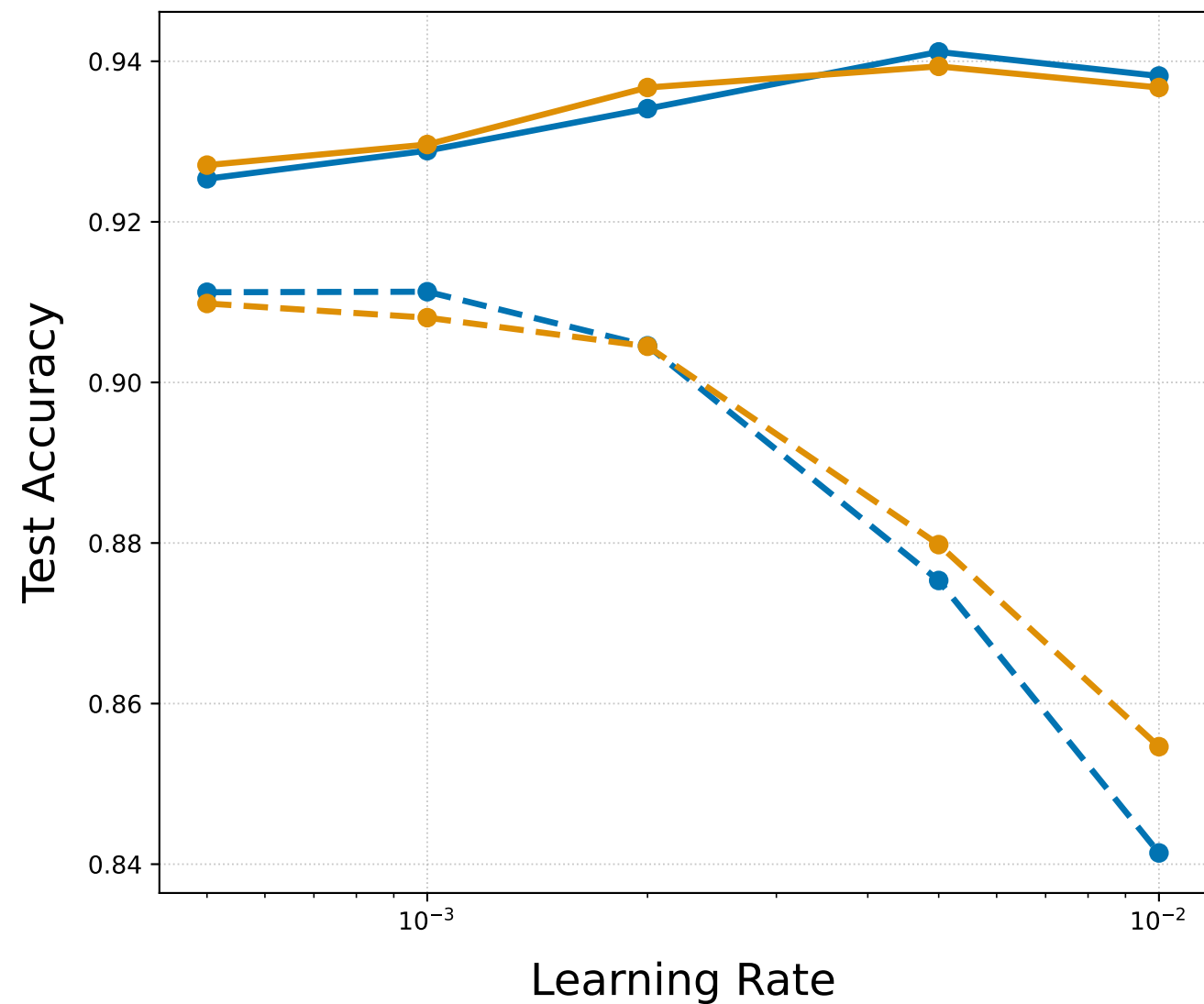


$$\beta_1 = 0.9, \beta_2 = 0.999$$



$$\beta_1 = 0.95, \beta_2 = 0.95$$



— Bias Correction, Warmup Cosine

— No Bias Correction, Warmup Cosine

- - Bias Correction, Constant LR

- - No Bias Correction, Constant LR