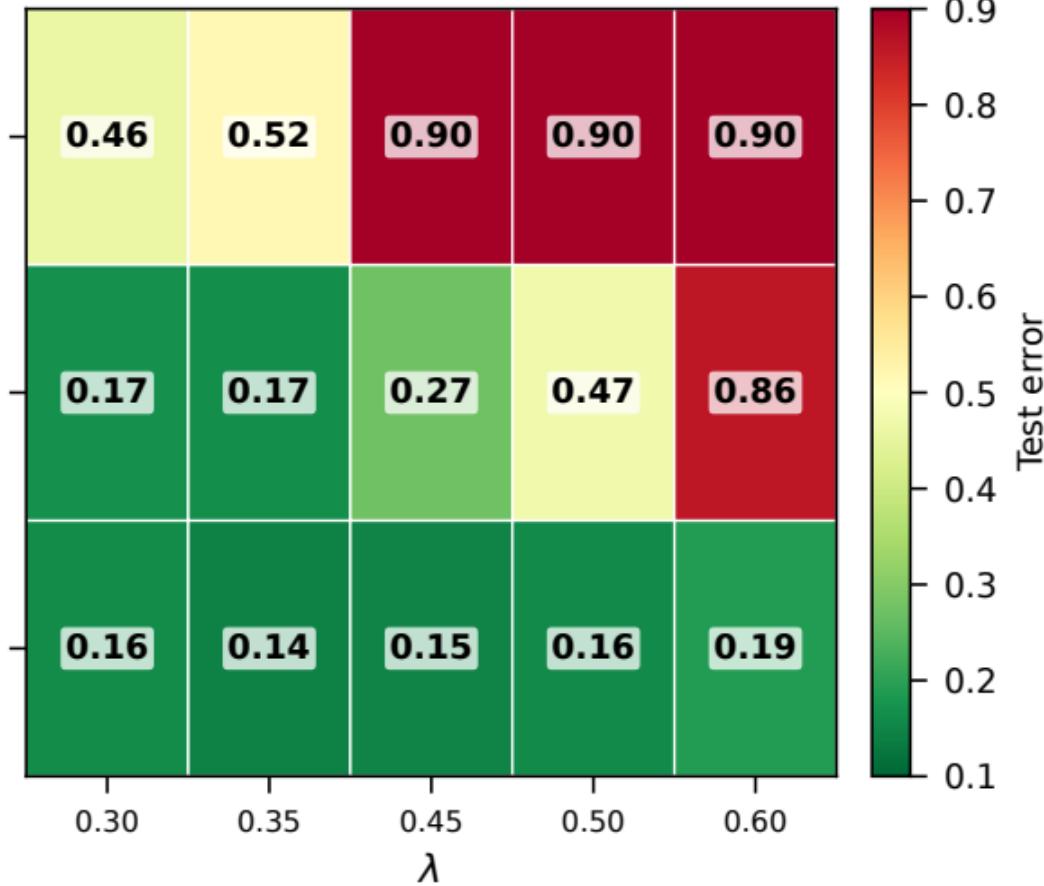


aLR- λ interaction (Q3, $\eta=0.4$)

Learning rate

**b**

Gradient alignment over training (Q4)

 C_{sv} (gradient cosine)