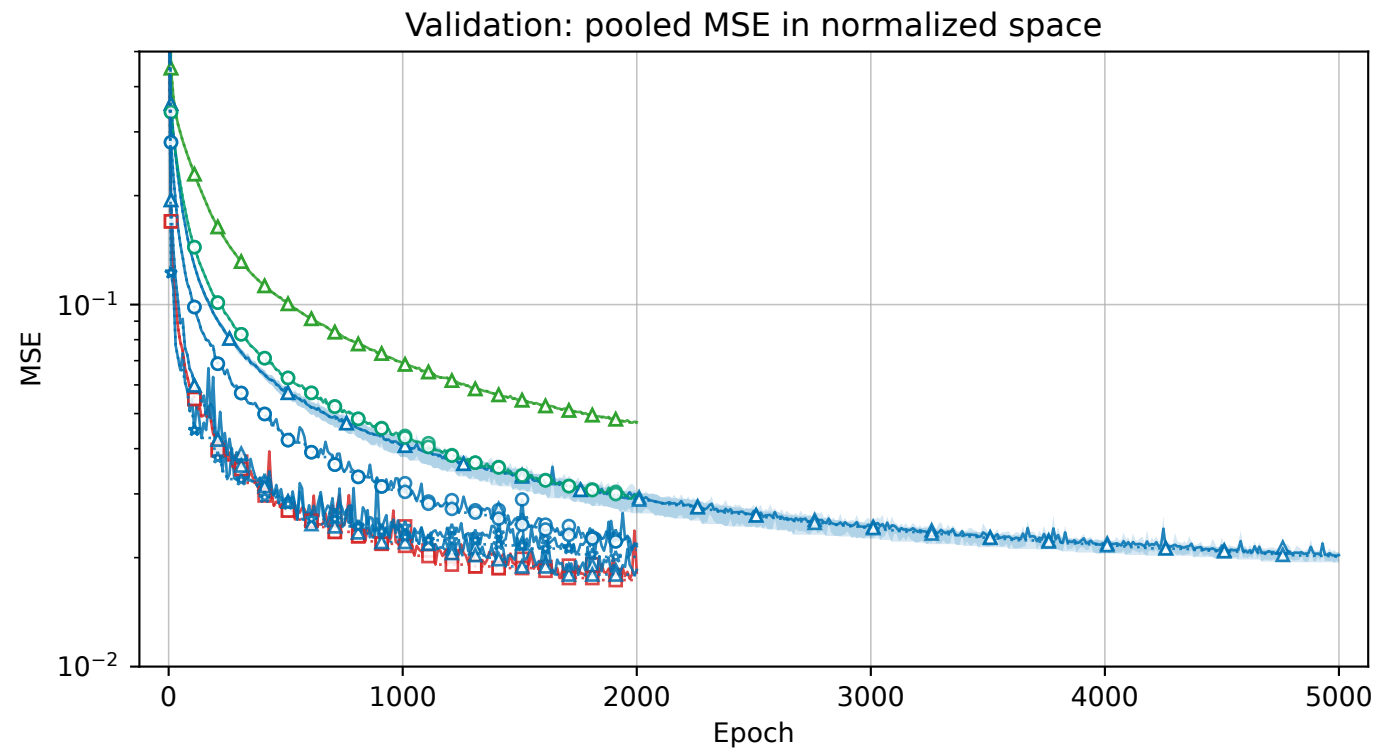
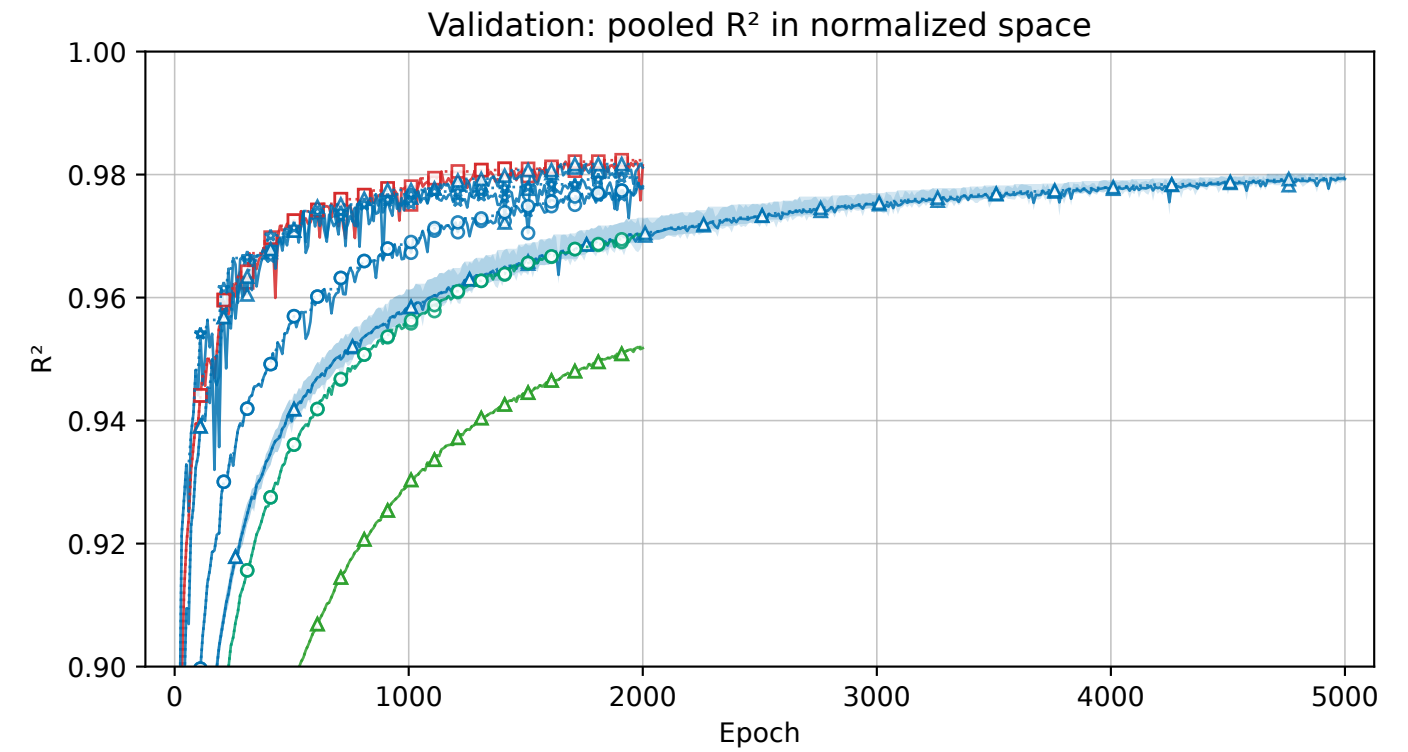
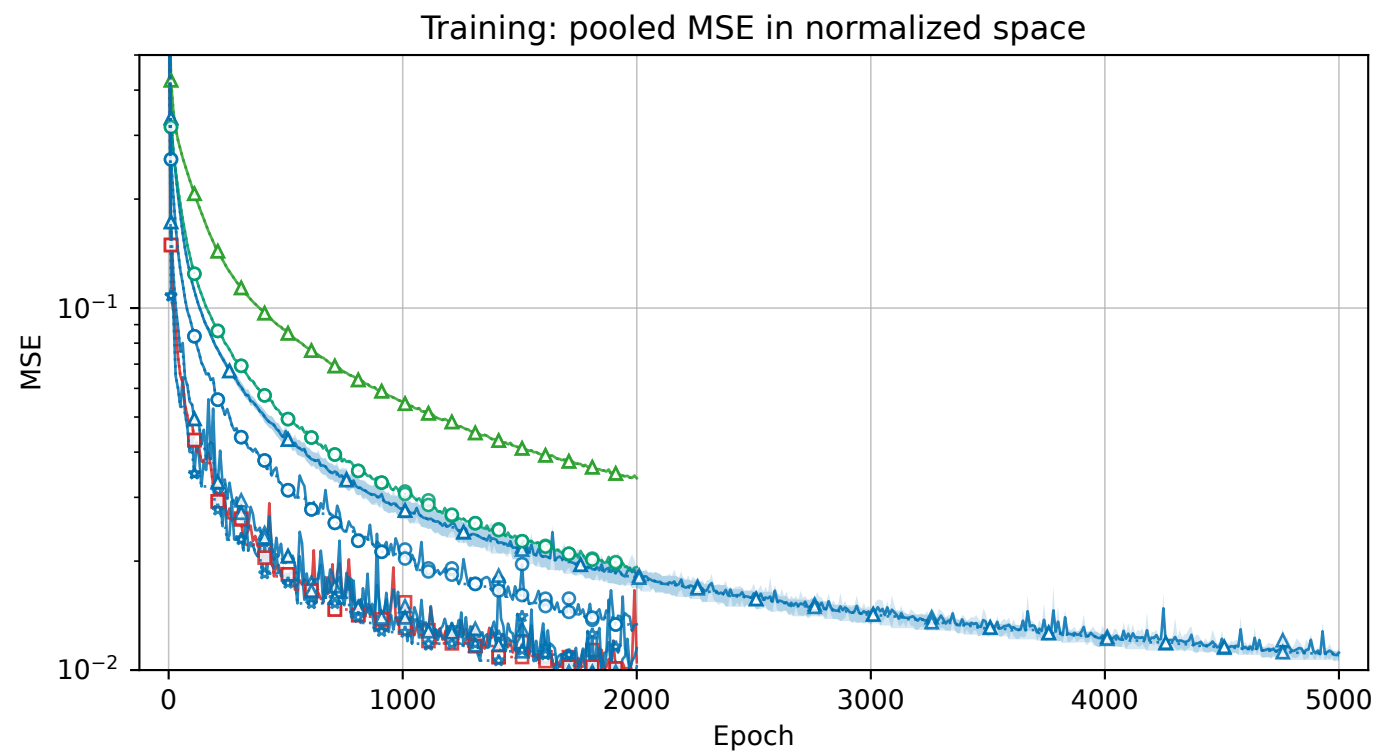
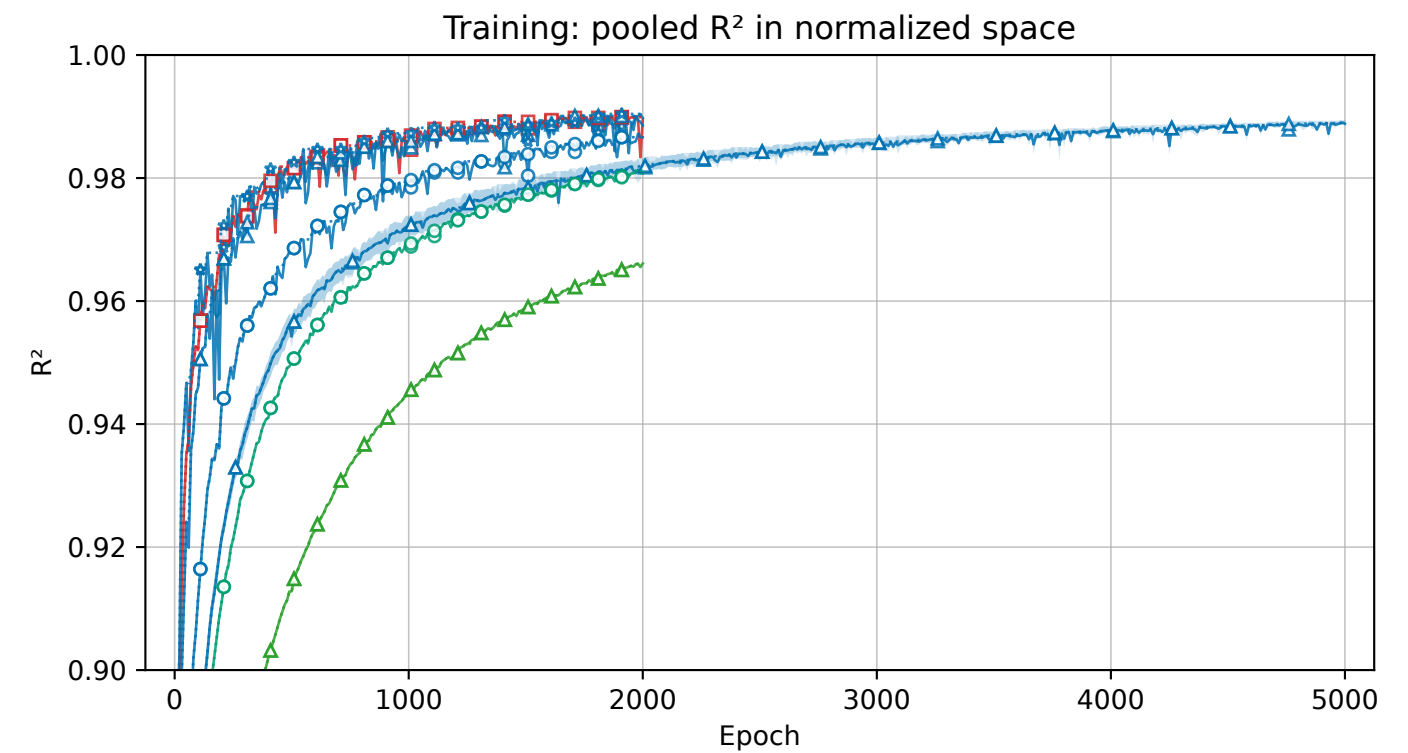


Initialization ensembles: pooled training and validation metrics

7 experiments | ensemble means and min/max ranges

— valid best_valid — train best_train



.....△..... LR:1e-4, B:65536
□..... LR:1e-3, B:16384
△..... LR:3e-4, B:4096
.....△..... LR:3e-4, B:65536
.....○..... LR:1e-4, B:4096
.....★..... LR:1e-3, B:4096
.....○..... LR:1e-4, B:16384

Shading: member min/max (not a confidence interval); no shading for a single member.

best_train uses each member's validation-selected best snapshot.

Epochs: common recorded observations within each ensemble.