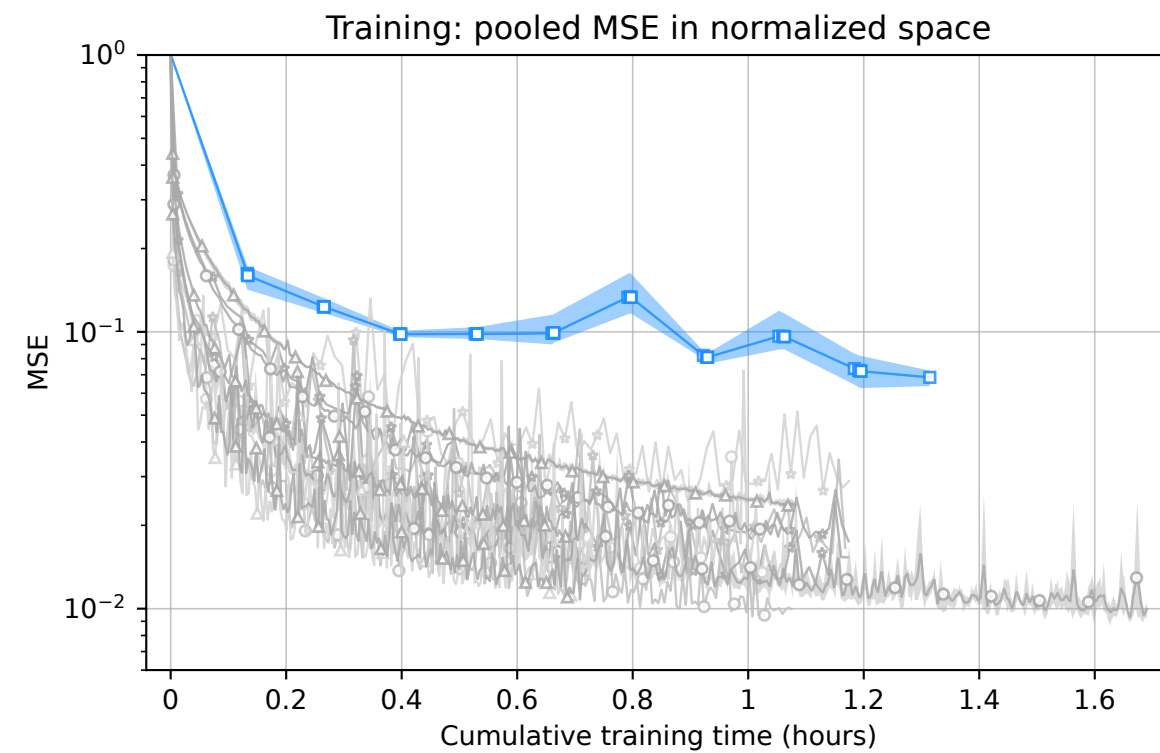
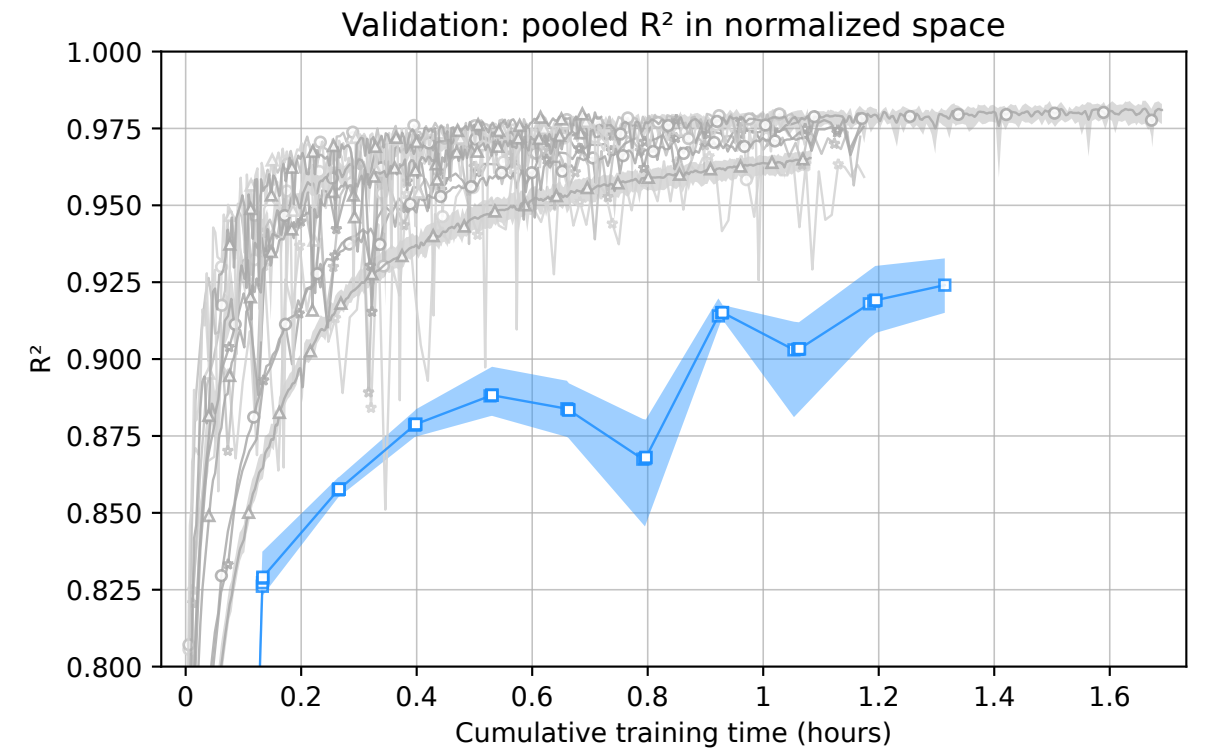
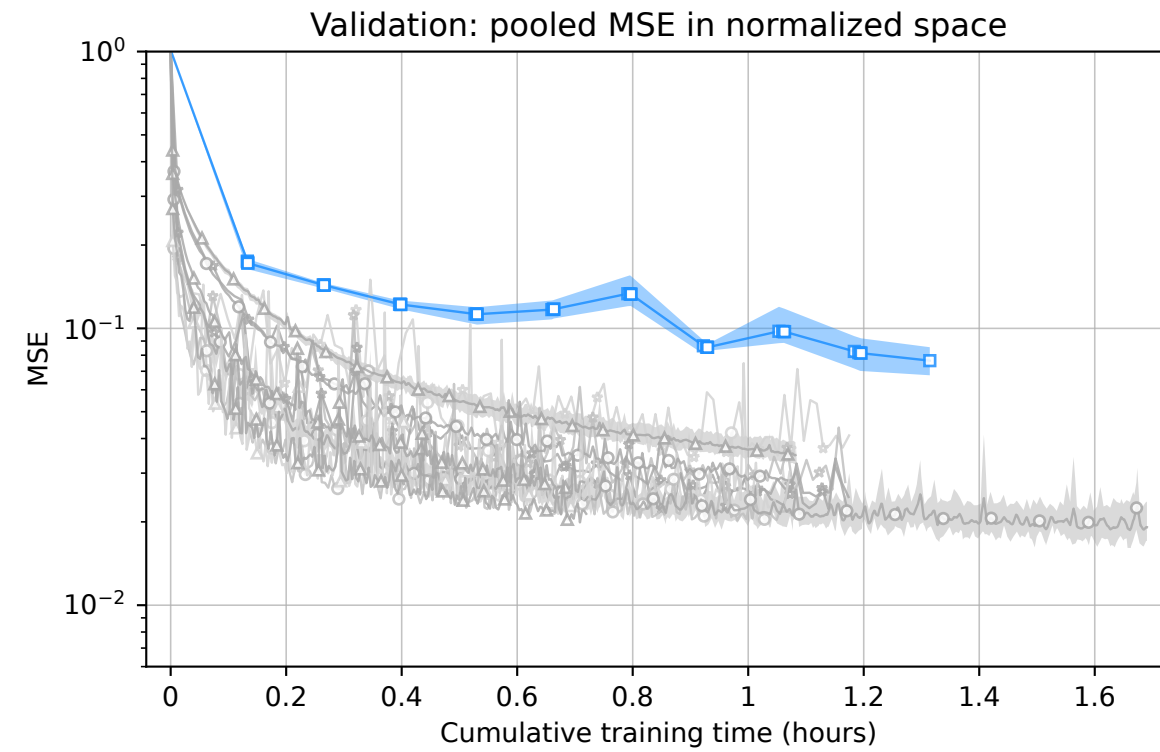


# Initialization ensembles: pooled training and validation metrics

13 experiments | ensemble means and min/max ranges

— valid — train



- |                      |                      |                      |
|----------------------|----------------------|----------------------|
| —*— LR 3e-3, B 4096  | —○— LR 1e-3, B 16384 | —△— LR 3e-4, B 65536 |
| —*— LR 1e-3, B 4096  | —○— LR 3e-4, B 16384 | —△— LR 1e-4, B 65536 |
| —*— LR 3e-4, B 4096  | —○— LR 1e-4, B 16384 | —□— LR 1e-3, B 256   |
| —*— LR 1e-4, B 4096  | —△— LR 3e-3, B 65536 | —△— LR 1e-3, B 65536 |
| —○— LR 3e-3, 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.

Time: common range within each ensemble; current metrics interpolated linearly, best metrics held between observations.