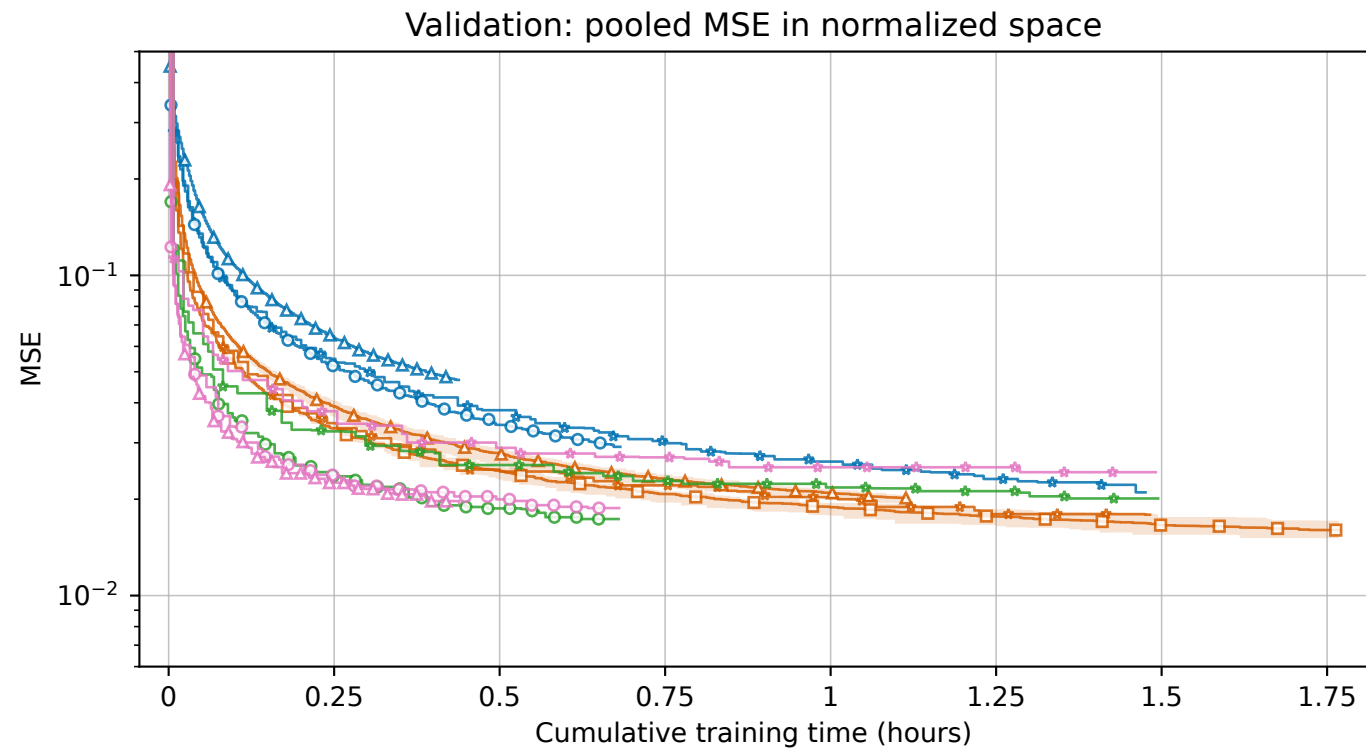
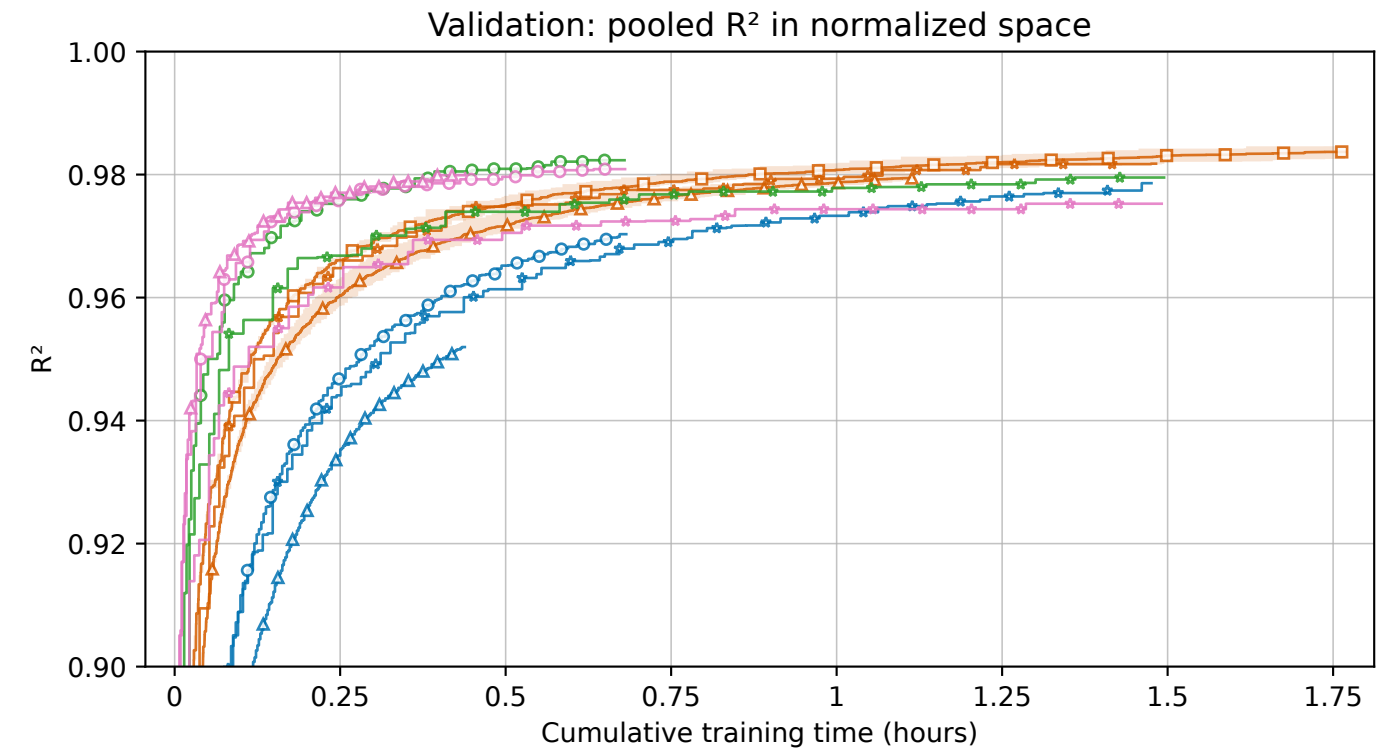
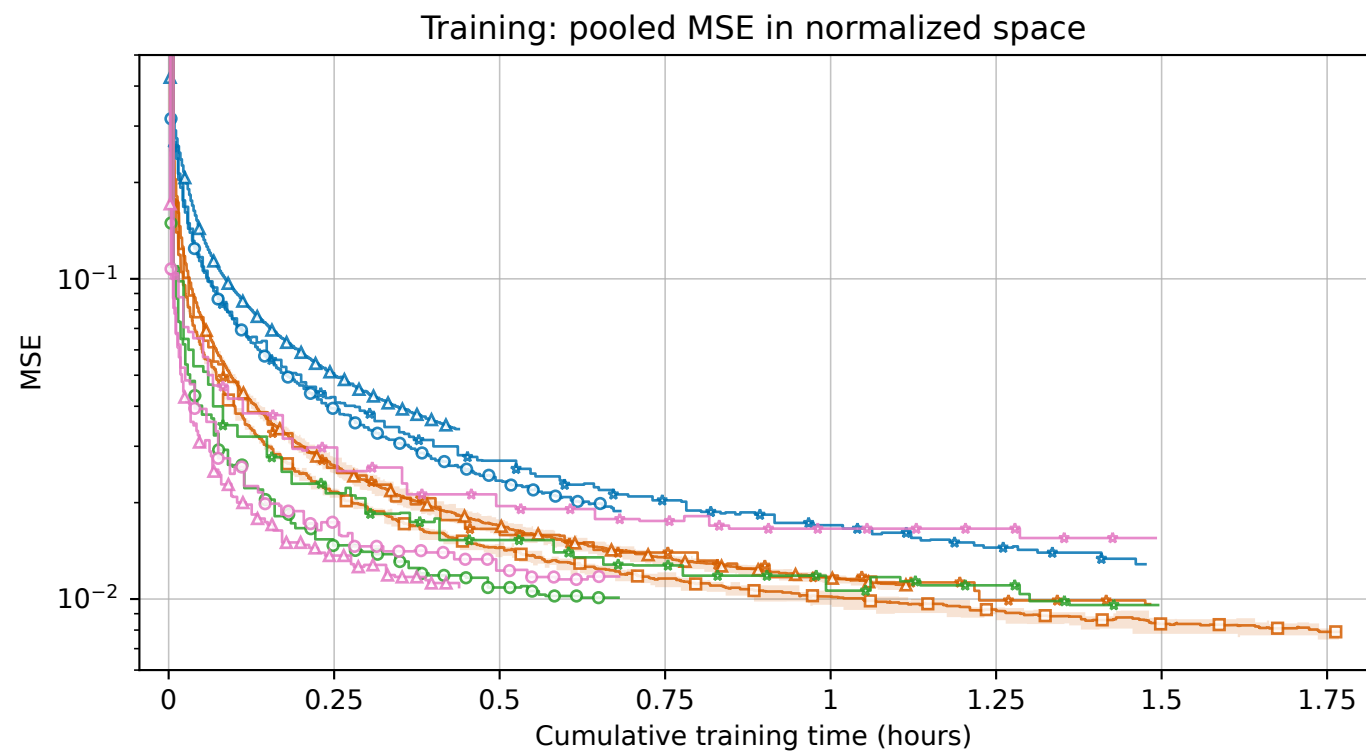
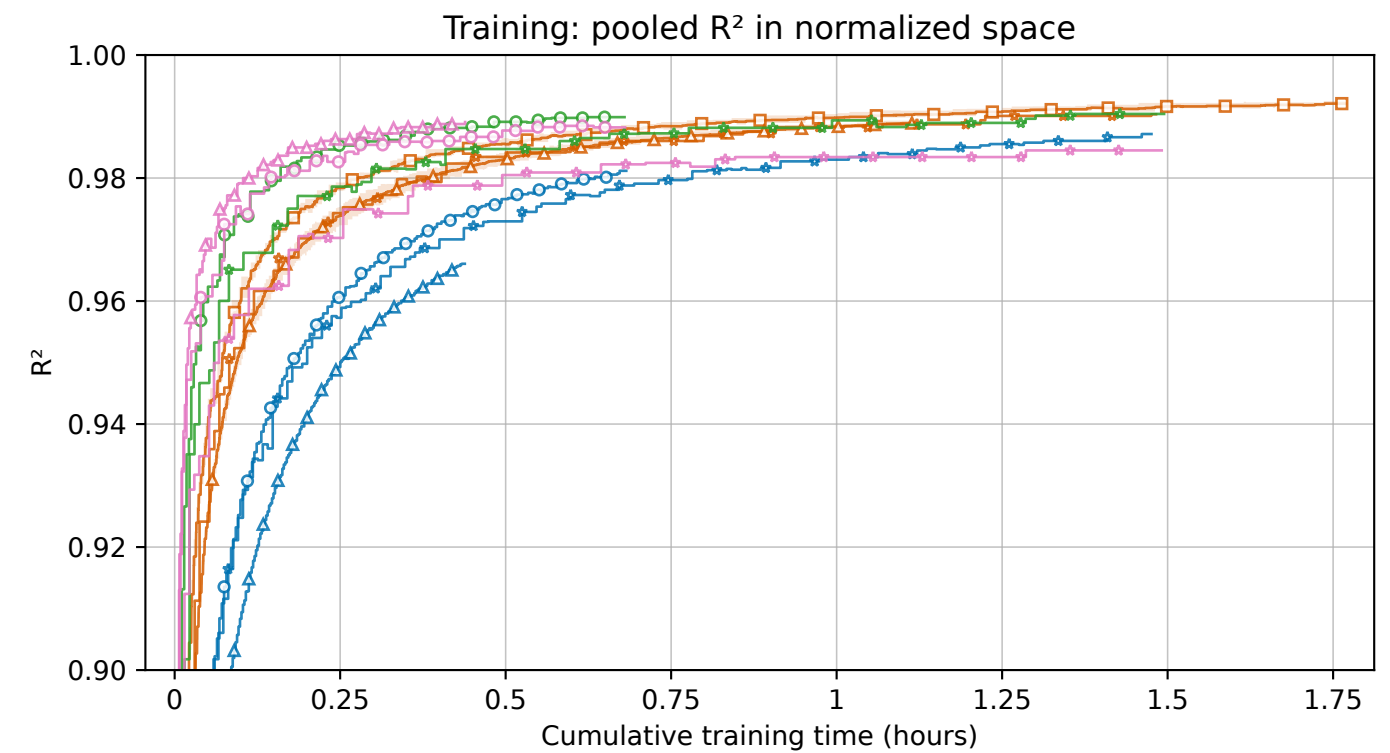


Initialization ensembles: pooled training and validation metrics

11 experiments | ensemble means and min/max ranges

— best_valid — best_train



- | | | |
|----------------------|----------------------|----------------------|
| —*— LR:1e-4, B:04096 | —□— LR:3e-4, B:16384 | —*— LR:3e-3, B:04096 |
| —○— LR:1e-4, B:16384 | —△— LR:3e-4, B:65536 | —○— LR:3e-3, B:16384 |
| —△— LR:1e-4, B:65536 | —*— LR:1e-3, B:04096 | —△— LR:3e-3, B:65536 |
| —*— LR:3e-4, B:04096 | —○— LR:1e-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.