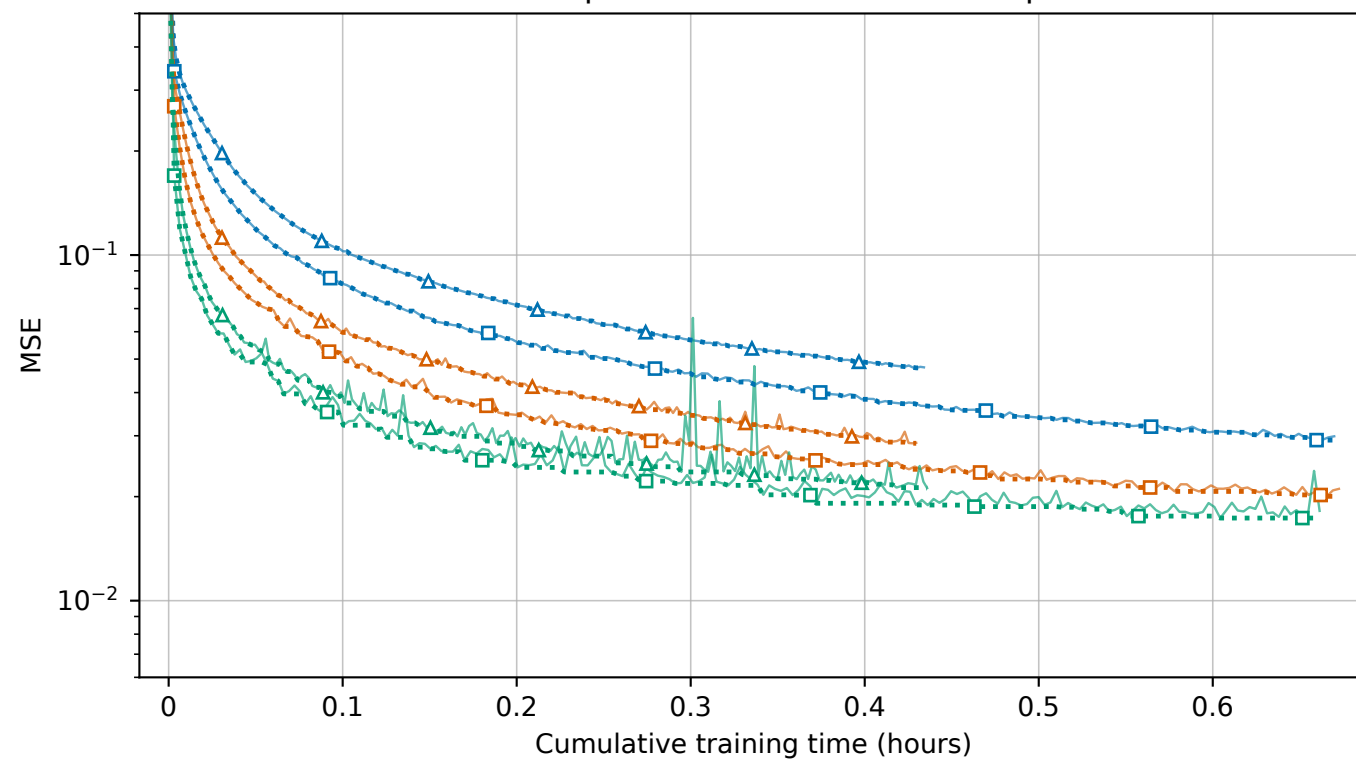


Training and validation: pooled MSE and R²

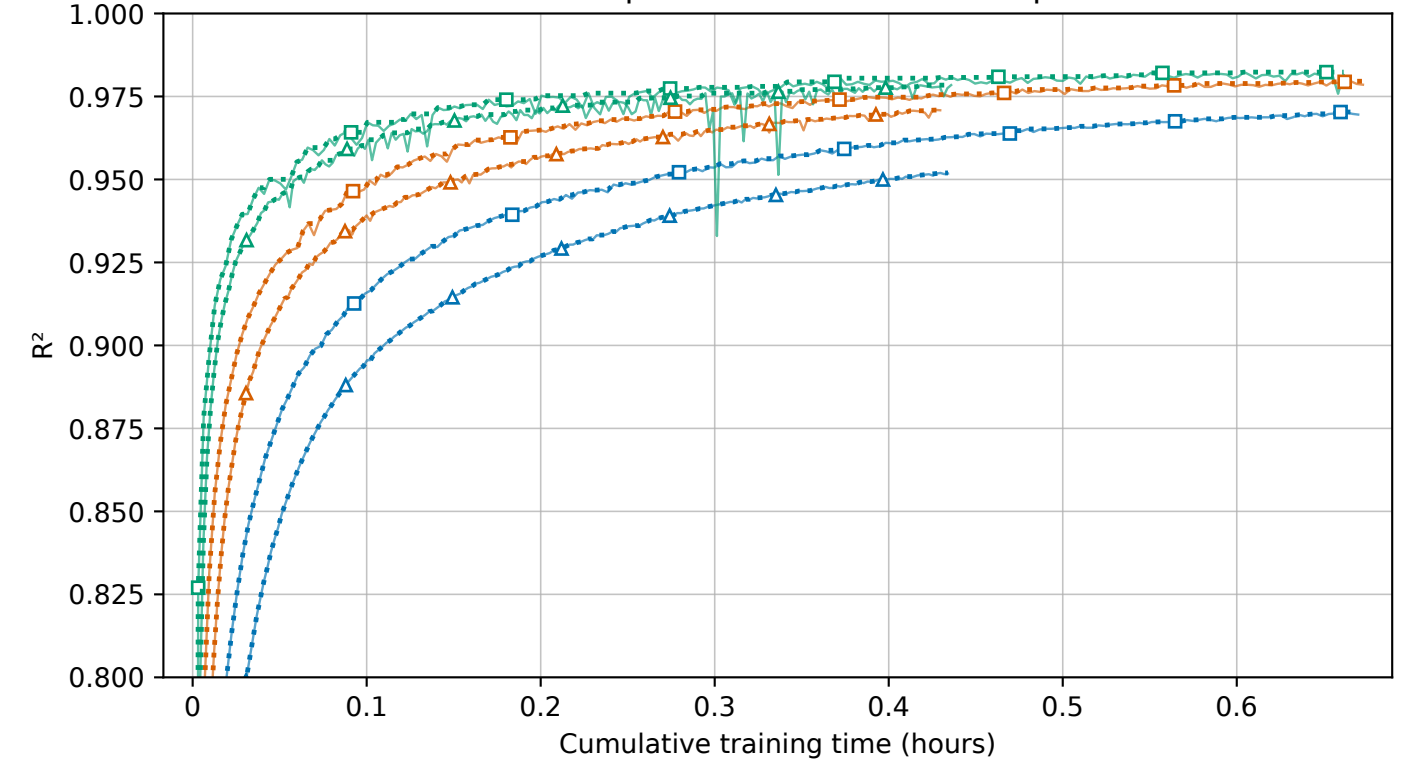
Initialization seed 125672 | 6 runs | through epoch 2,000

— valid ···· best_valid — train ···· best_train

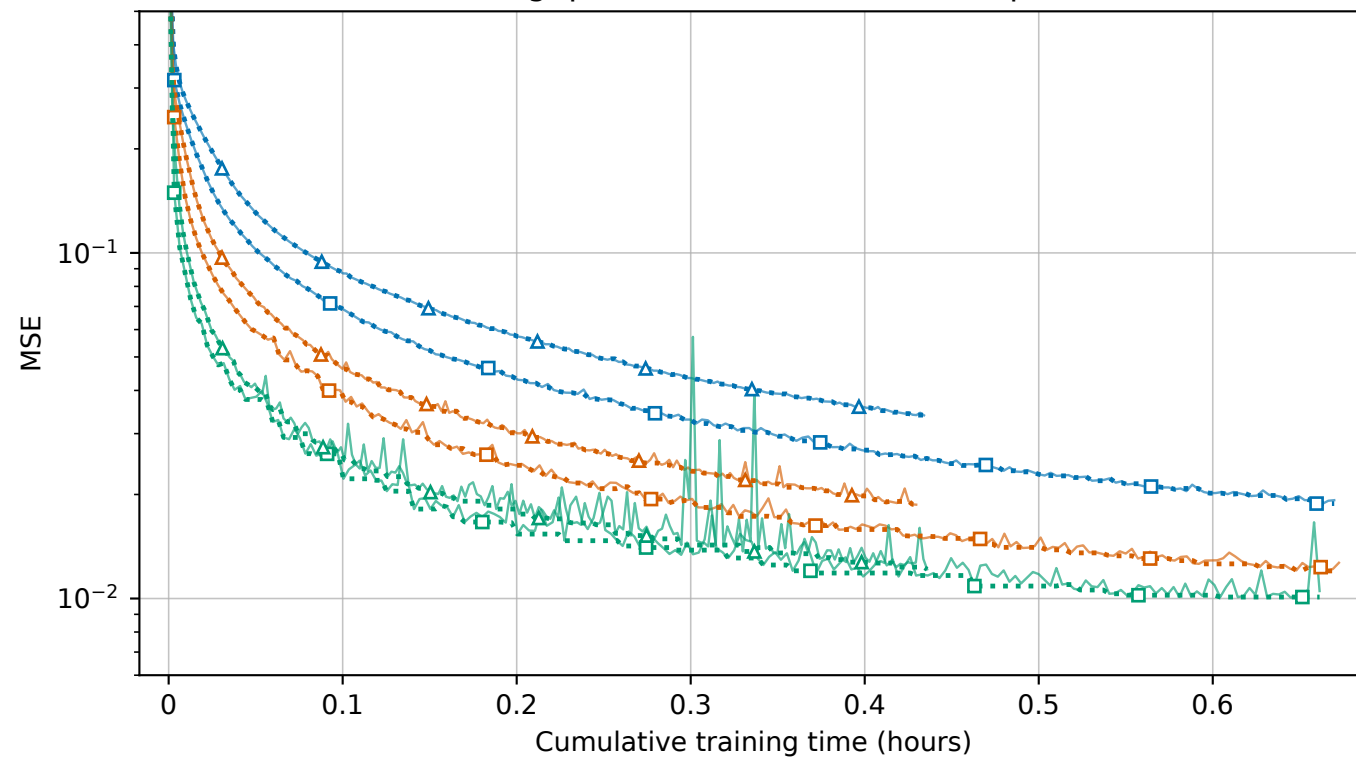
Validation: pooled MSE in normalized space



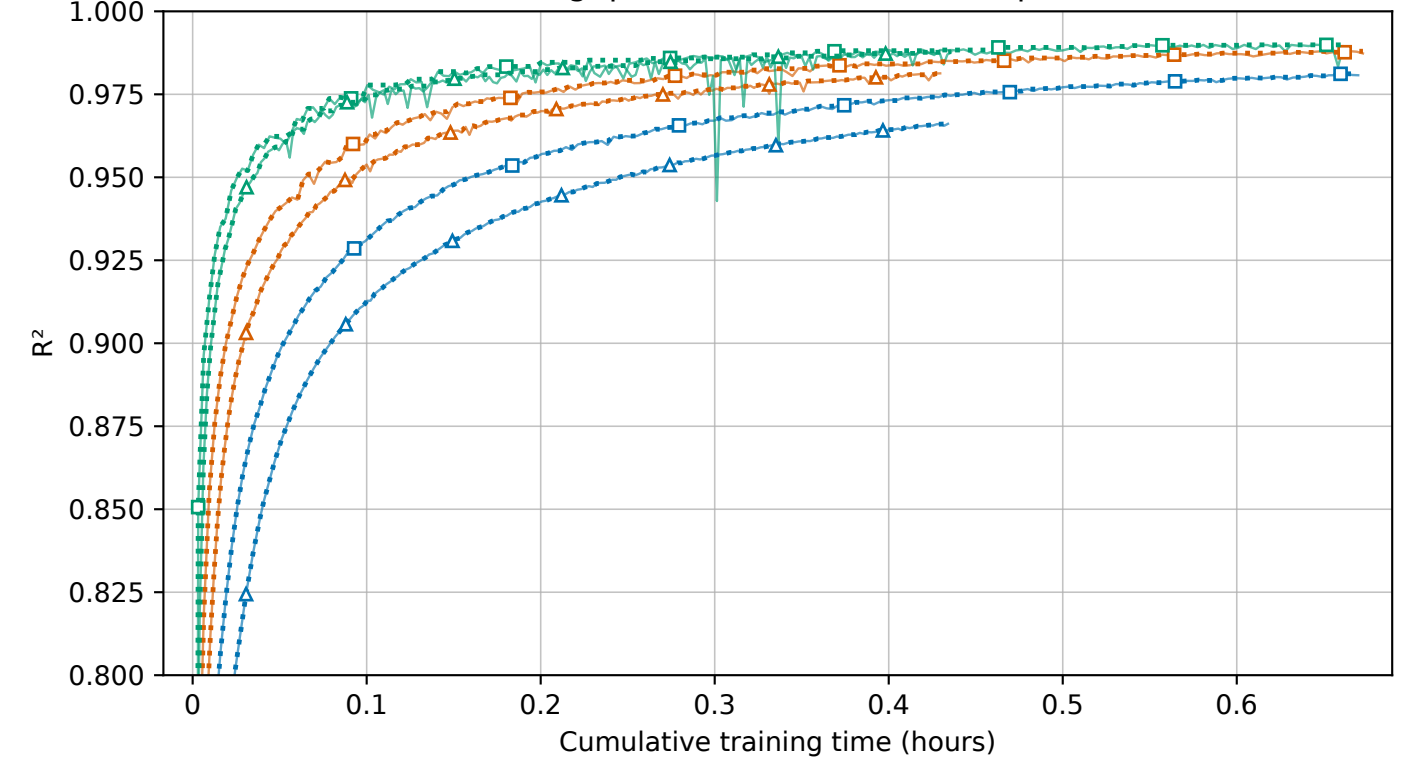
Validation: pooled R² in normalized space



Training: pooled MSE in normalized space



Training: pooled R² in normalized space



····□···· B=16,384, LR=0.0001 ····□···· B=16,384, LR=0.0003 ····□···· B=16,384, LR=0.001
 ····△···· B=65,536, LR=0.0001 ····△···· B=65,536, LR=0.0003 ····△···· B=65,536, LR=0.001

best_valid and best_train evaluate the SAME checkpoint, selected by lowest pooled validation MSE.