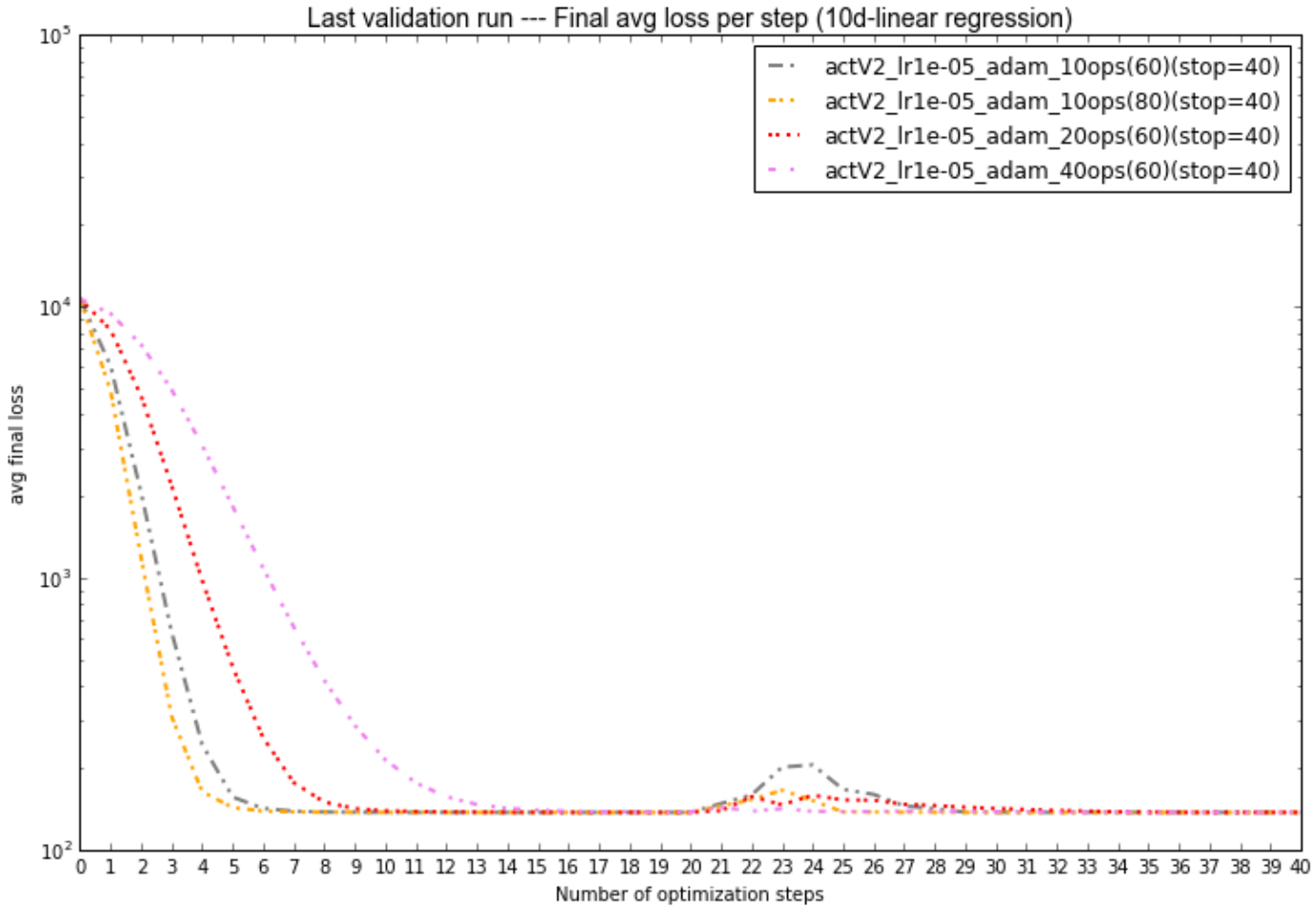
Main results of experiments with 10d-regression functions

ACT optimizer trained with average horizon T equal to {10, 20, 40} optimization steps for 60 epochs

Figure showing average loss for 3 different training horizons **T: 10, 20 and 40**

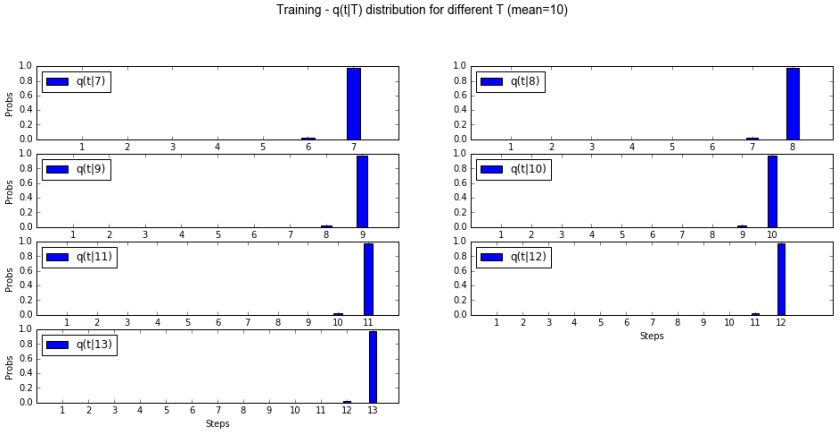
Please note: the yellow curve is from an ACT model with E[T]=10 but trained for **80 epochs** instead of 60. I therefore think that I still need to train the ACT models for more epochs

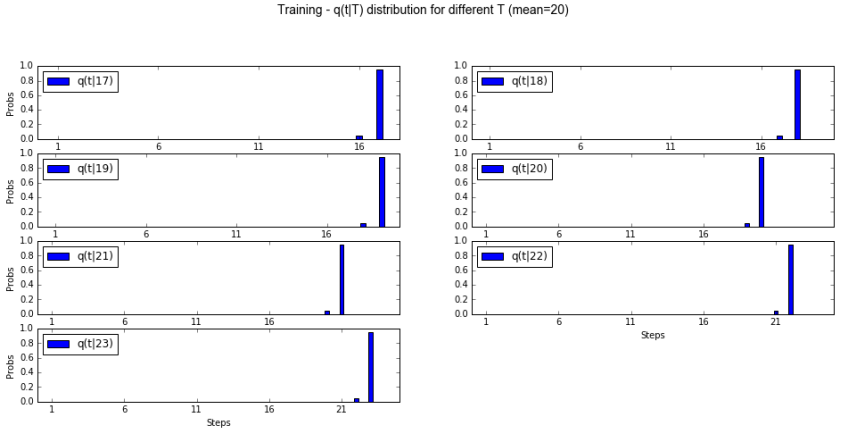


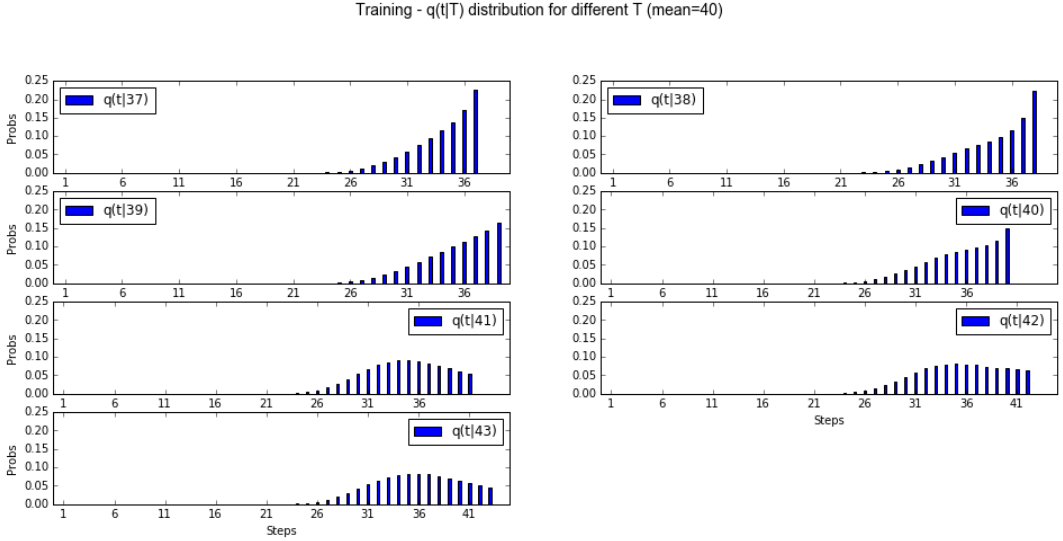
Approximated q(t|T) for models trained with different horizons T - computed during training

Note each figure shows the results for an ACT optimizer trained with a different horizon T.

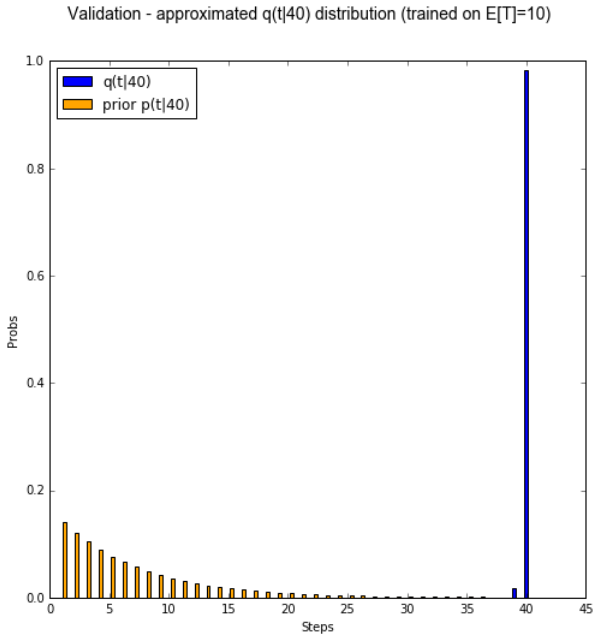
Each figure shows 7 different q(t | T) approximations, in which the horizon T varies around the E[T] for that model

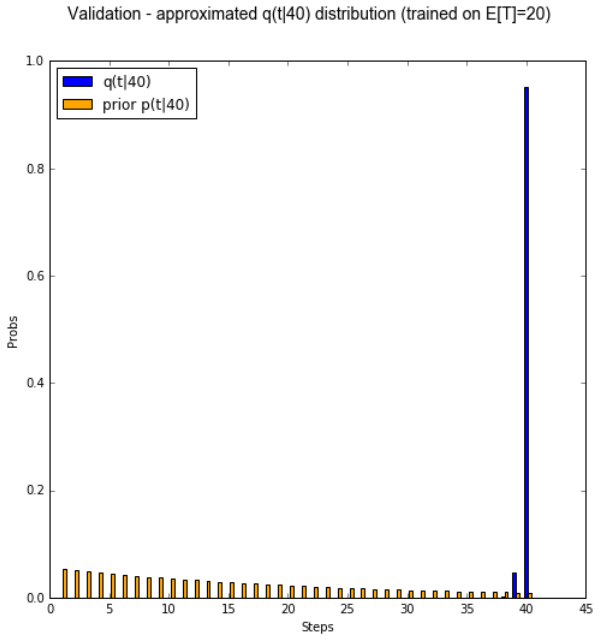


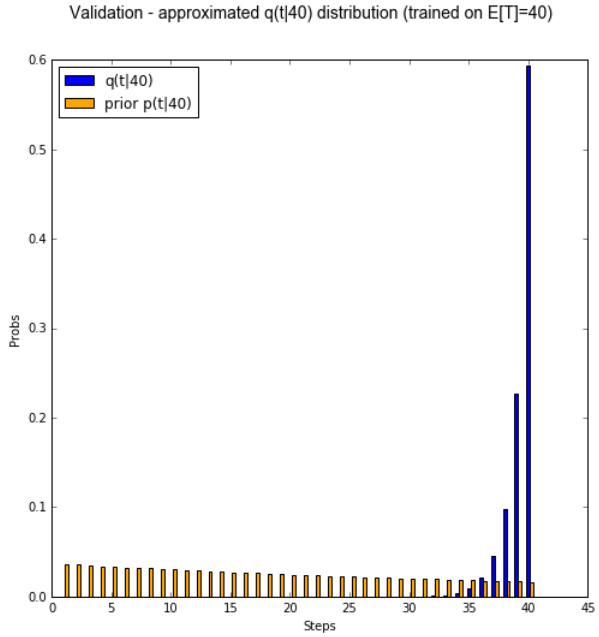




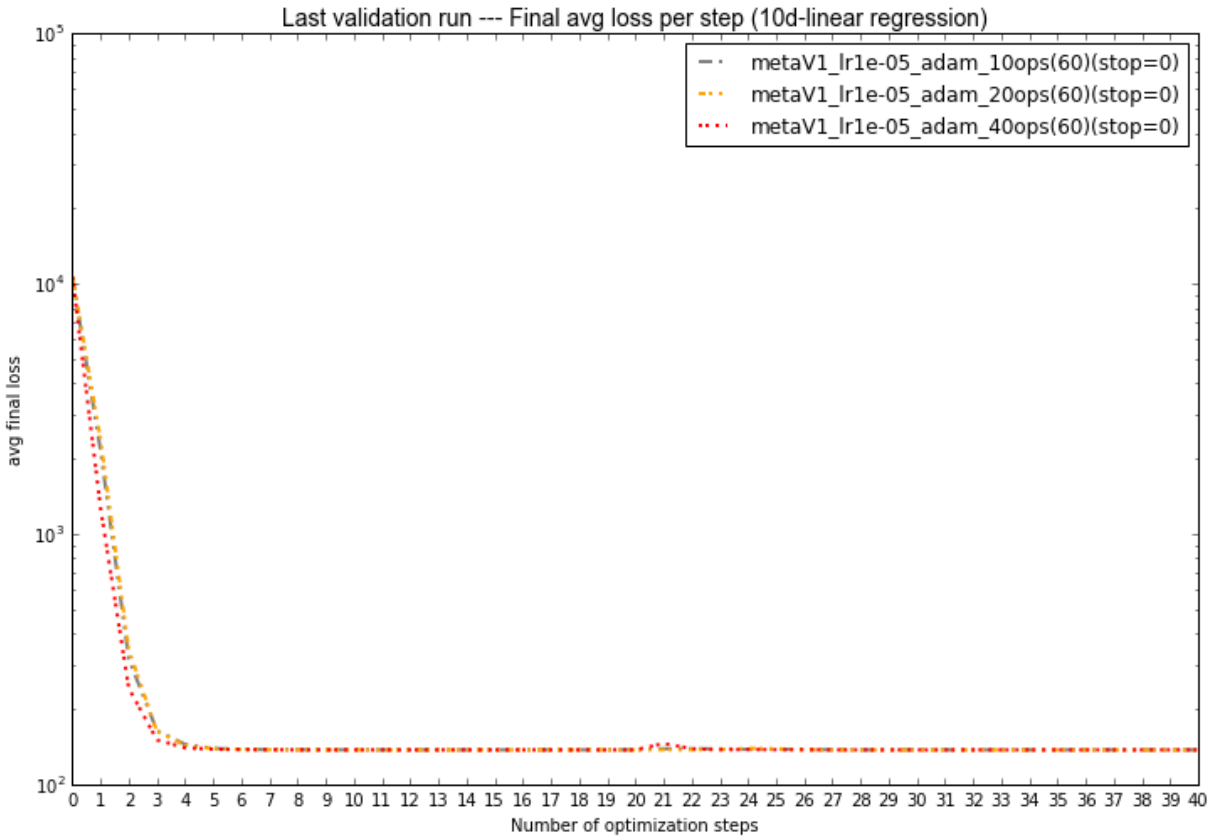
Approximated q(t|T) for models trained with different horizons T - computed during validation (unrolled for 40 steps)







LSTM optimizer trained with constant horizon T equal to {10, 20, 40} optimiziation steps for 60 epochs



Performance comparison betweeen LSTM and ACT optimizer