Basic network, one learning rate | Basic network, unique learning rates | Our topology, $p = 7.56\%$

Layer pair training rate vs. Epoch

Legend:
- Input to Hidden Layer 1 ($\Delta\hat{w}^1(b)$)
- Hidden Layer 1 to Hidden Layer 2 ($\Delta\hat{w}^2(b)$)
- Hidden Layer 2 to Hidden Layer 3 ($\Delta\hat{w}^3(b)$)
- Hidden Layer 3 to Output ($\Delta\hat{w}^4(b)$)