The effect of unfolding a network for BPTT ($\tau = 3$).