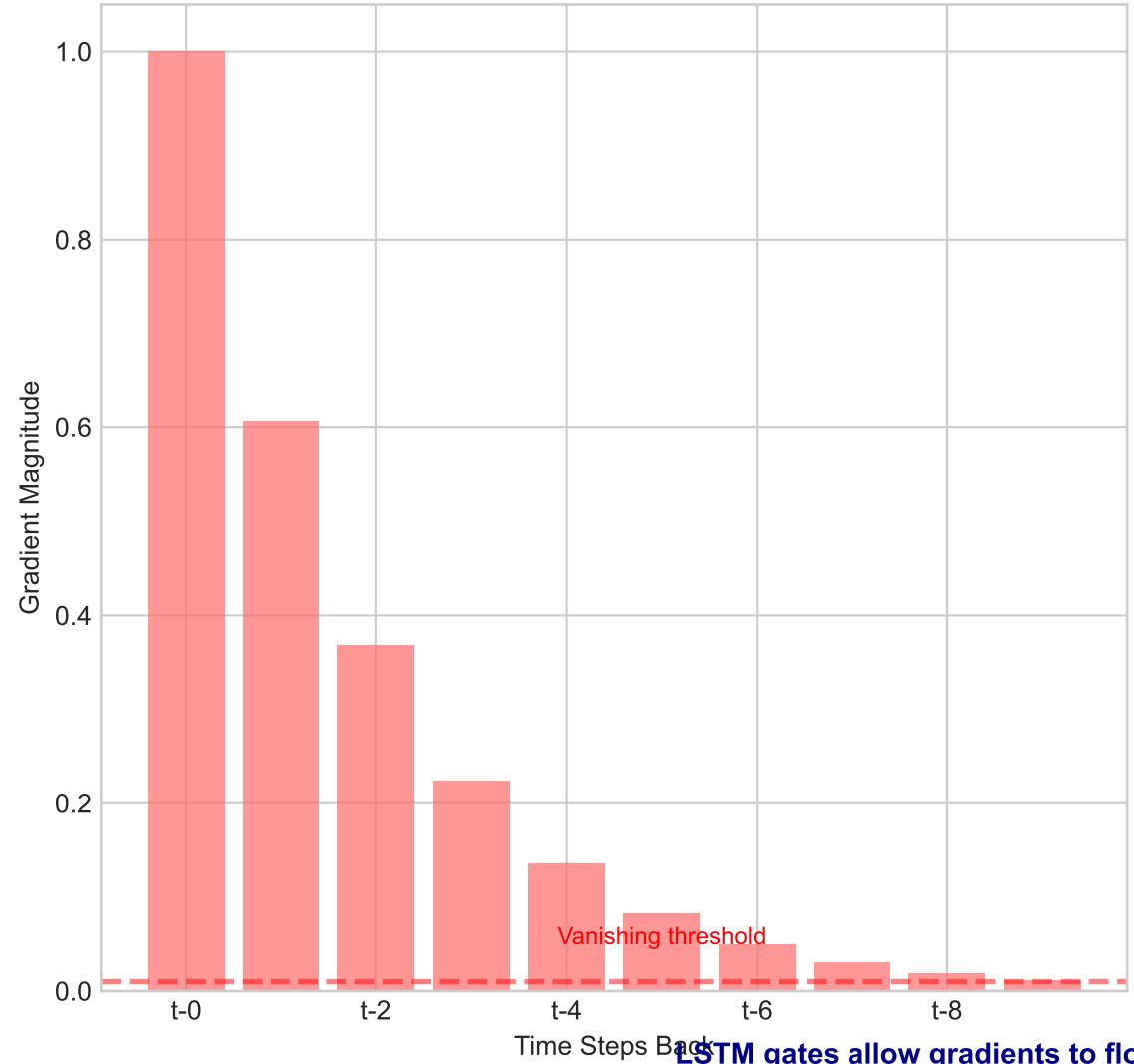
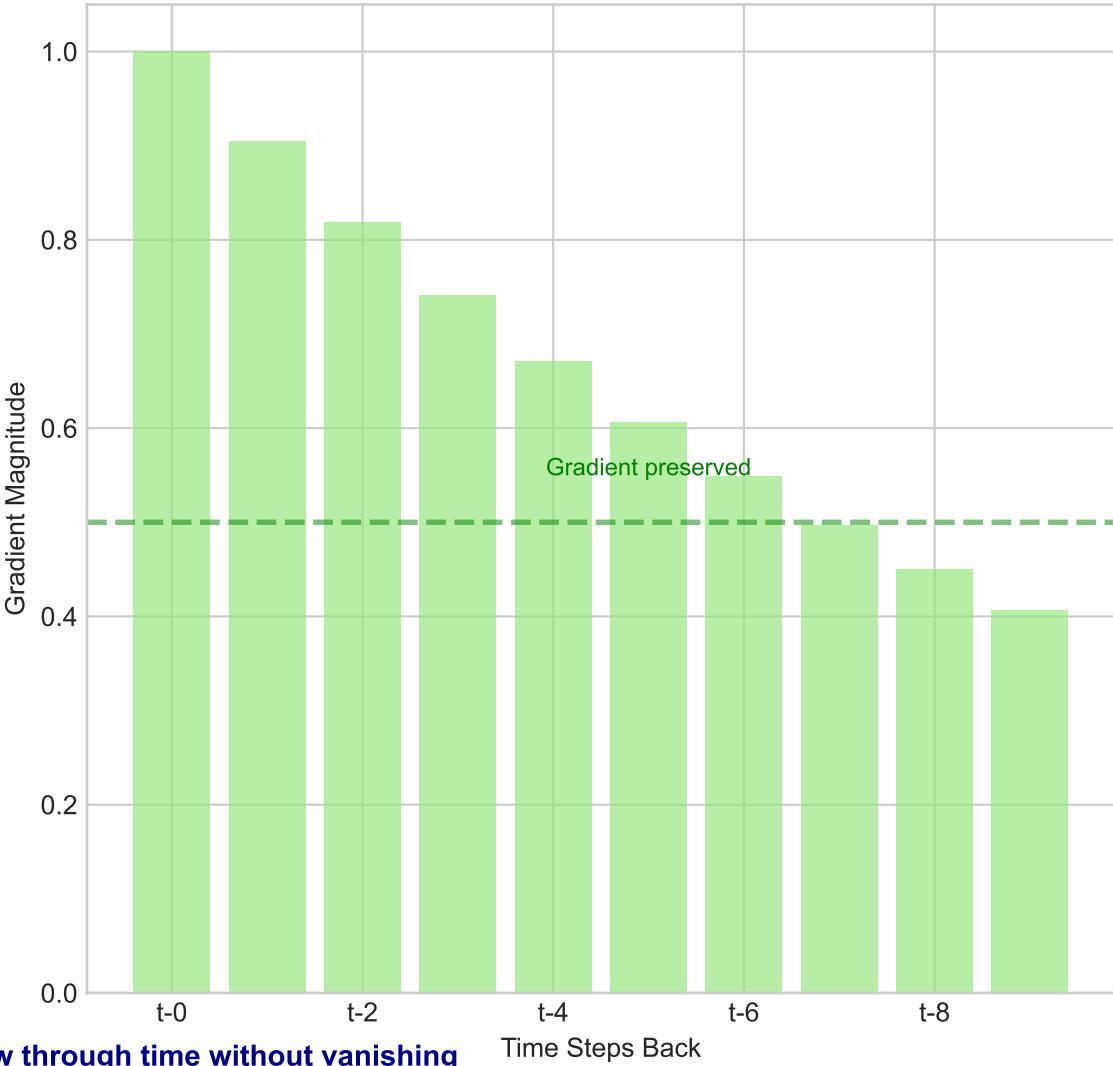


Vanilla RNN: Vanishing Gradients



LSTM: Preserved Gradients



LSTM gates allow gradients to flow through time without vanishing