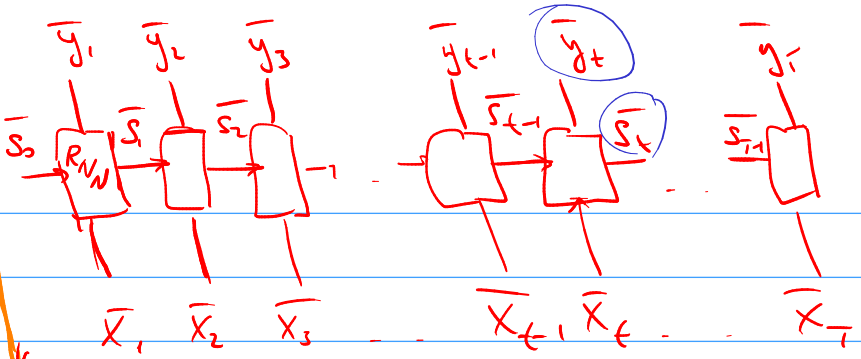
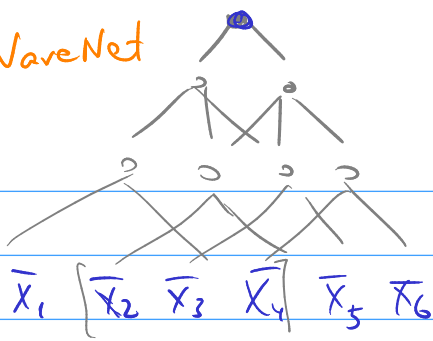
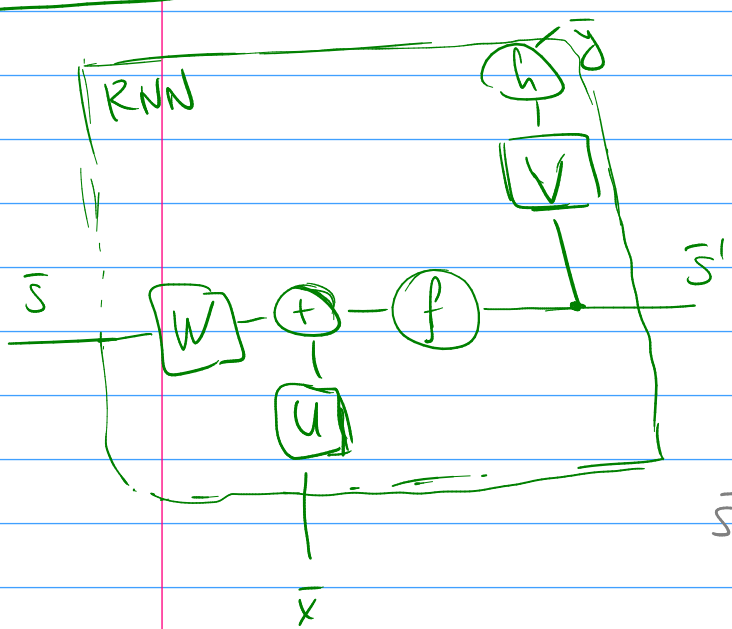
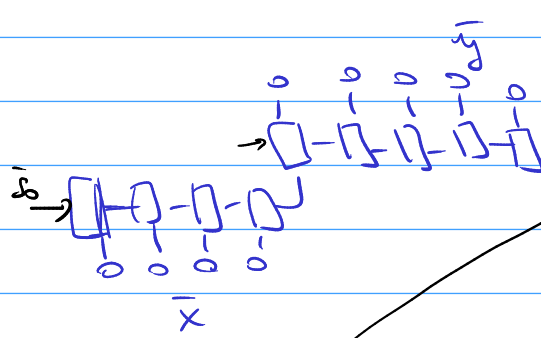
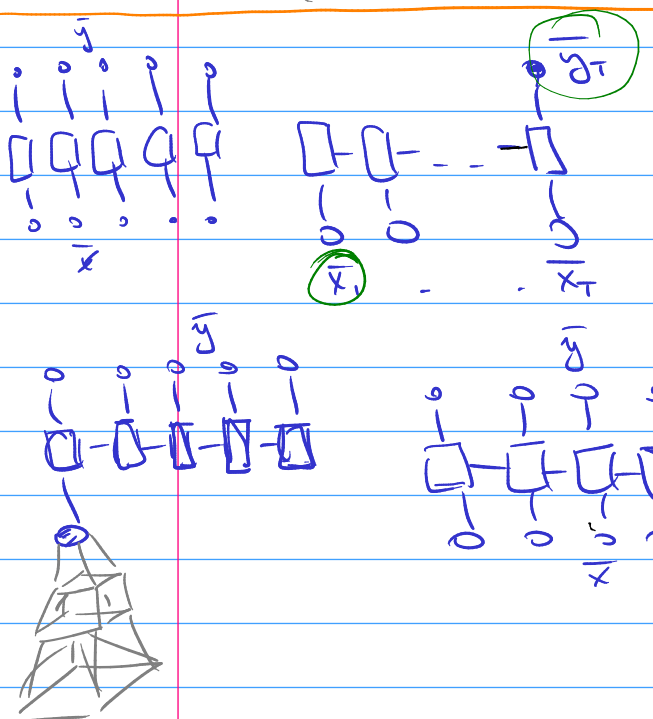


WaveNet



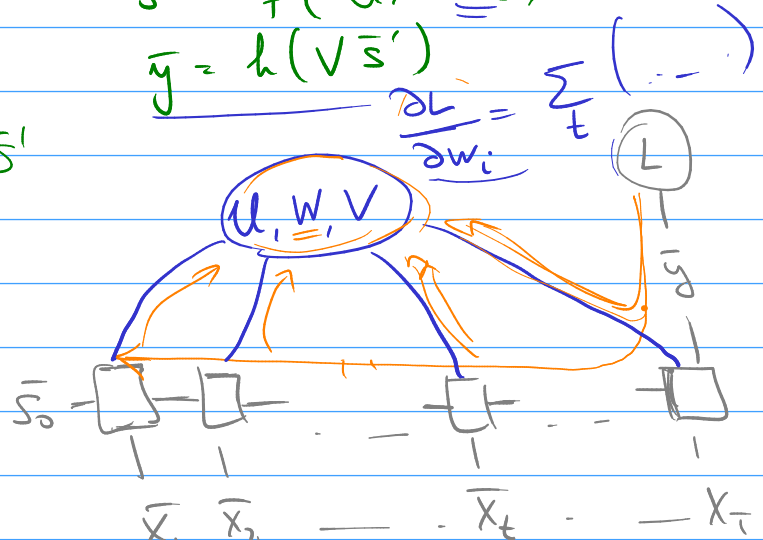
$$p(\bar{y}_t | \bar{x}_1, \dots, \bar{x}_t, \bar{s}_0, \dots, \bar{s}_{t-1}) = p(\bar{y}_t | \bar{x}_t, \bar{s}_{t-1})$$

$$p(\bar{s}_t | \dots) = p(\bar{s}_t | \bar{x}_t, \bar{s}_{t-1})$$

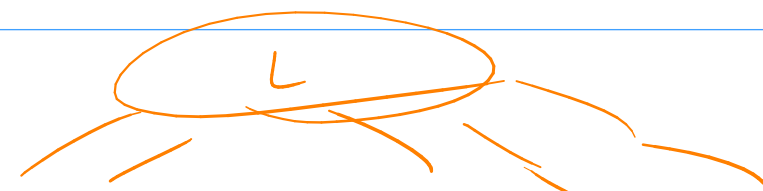


$$\bar{s}' = f(U\bar{x} + W\bar{s})$$

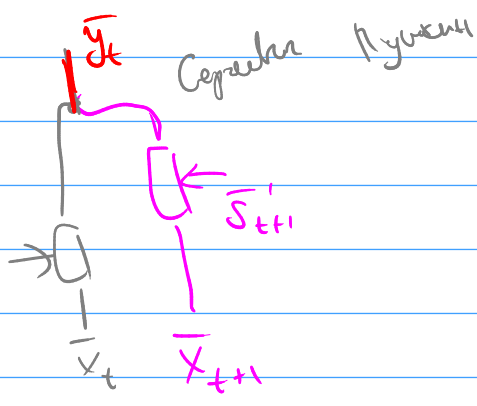
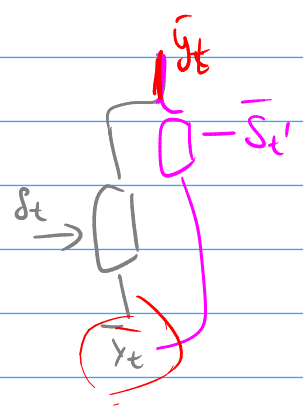
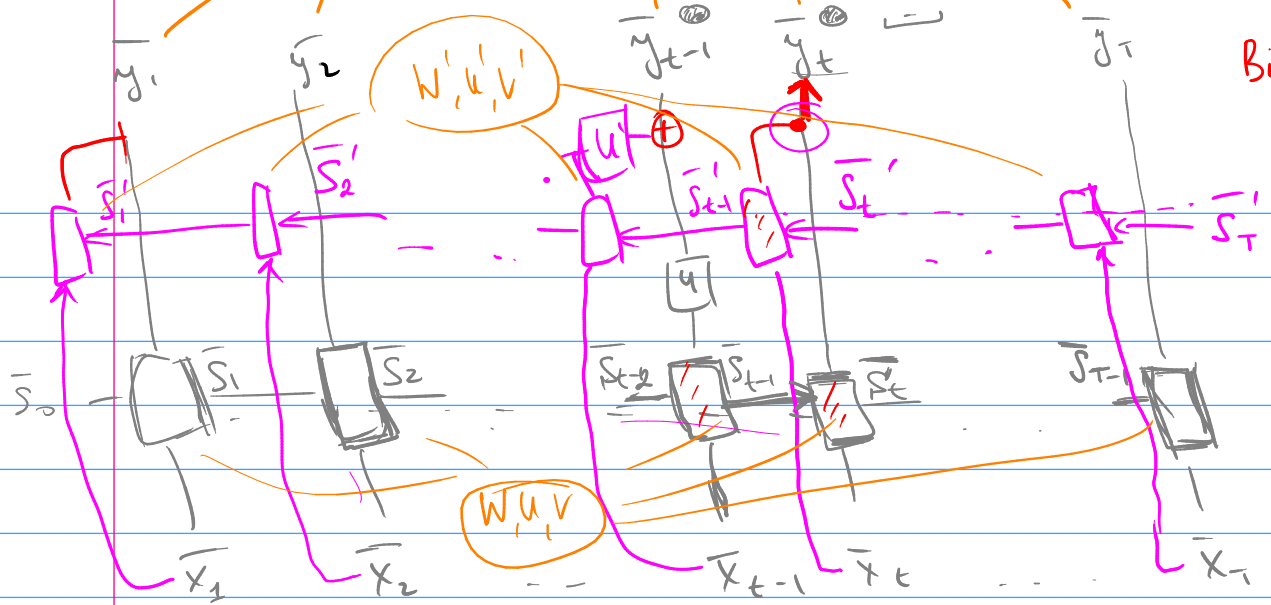
$$\bar{y} = h(V\bar{s}')$$



$\bar{s}_t \sim W^t \bar{s}_0$ *vanishing grad.* $\|W\| < 1 \Rightarrow |\bar{s}_t| \rightarrow 0$
exploding gradients $\|W\| > 1 \Rightarrow |\bar{s}_t| \rightarrow \infty$

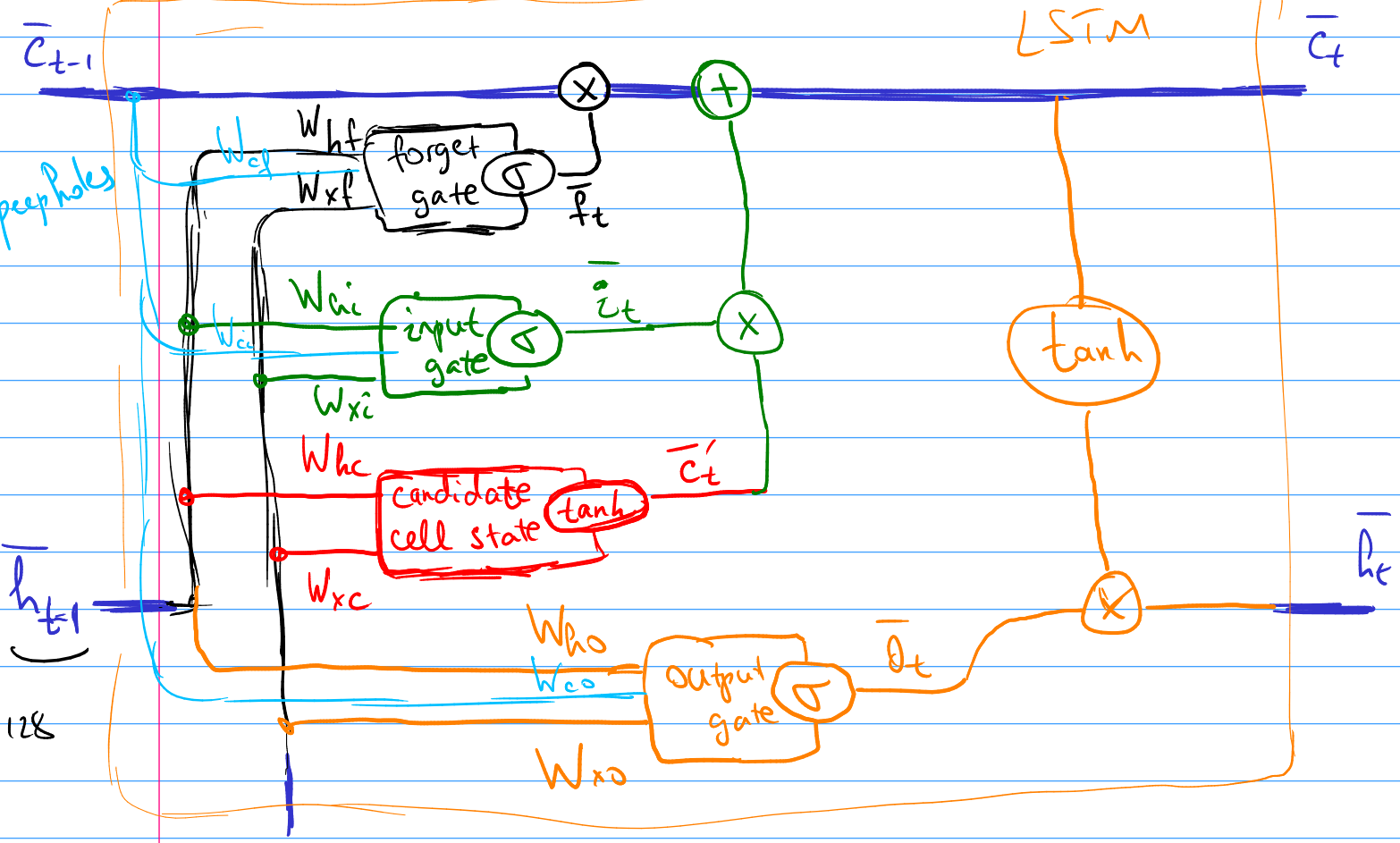


Bidirectional RNN



$$\bar{z} = f(A\bar{x} + B\bar{y})$$

LSTM - long short-term memory



$$\bar{c}_t = \bar{i}_t \odot \bar{c}'_t + \bar{f}_t \odot \bar{c}_{t-1}$$

$$\bar{z}_t = \sigma \left(\underbrace{W_{hi}}_{d_x \times d_h} \bar{h}_{t-1} + \underbrace{W_{xi}}_{d_x \times d_h} \bar{x}_t + \underbrace{W_{ci}}_{128 \times 128} \bar{c}_{t-1} \right)$$