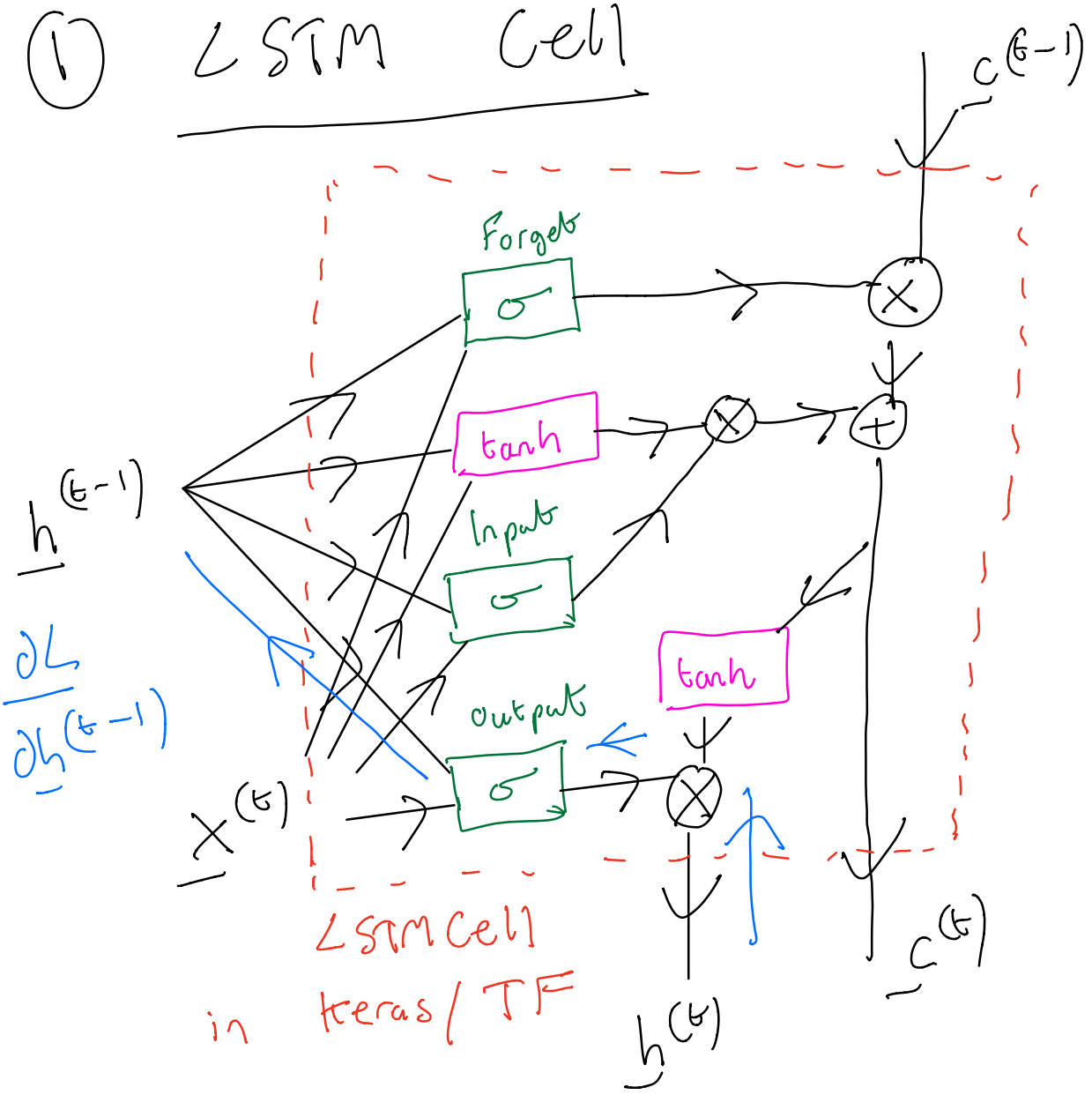
# (1) LSTM Cell




Assume $\frac{\partial L}{\partial h^{(t)}} \neq 0$ $\frac{\partial L}{\partial c^{(t)}} = 0$

Output gate $\underline{o}^{(t)} = \sigma(\underline{z}^{(t)})$

where $\underline{z}^{(t)} = \underline{W}_o^T \underline{h}^{(t-1)} + \underline{U}_o^T \underline{x}^{(t)} + \underline{b}_o$

Note: do not confuse $\underline{z}$ with reset gate in GRU, this is just the pre-activation value

$$\underline{h}^{(t)} = \underline{o}^{(t)} \odot \tanh(\underline{c}^{(t)})$$

$$\frac{\partial L}{\partial \underline{o}^{(t)}} = \frac{\partial L}{\partial \underline{h}^{(t)}} \frac{\partial \underline{h}^{(t)}}{\partial \underline{o}^{(t)}}$$

$$= \frac{\partial L}{\partial \underline{h}^{(t)}} \odot \tanh(\underline{c}^{(t)})$$

$$\frac{\partial L}{\partial \underline{z}^{(t)}} = \frac{\partial L}{\partial \underline{o}^{(t)}} \frac{\partial \underline{o}^{(t)}}{\partial \underline{z}^{(t)}}$$

$$= \frac{\partial L}{\partial h^{(t)}} \odot \tanh(c^{(t)}) \odot \sigma(z^{(t)}) \odot (1 - \sigma(z^{(t)}))$$

$$\frac{\partial L}{\partial h^{(t-1)}} = \frac{\partial L}{\partial z^{(t)}} \frac{\partial z^{(t)}}{\partial h^{(t-1)}}$$

$$= W_o \left[ \frac{\partial L}{\partial h^{(t)}} \odot \tanh(c^{(t)}) \odot \sigma(z^{(t)}) \odot (1 - \sigma(z^{(t)})) \right]$$