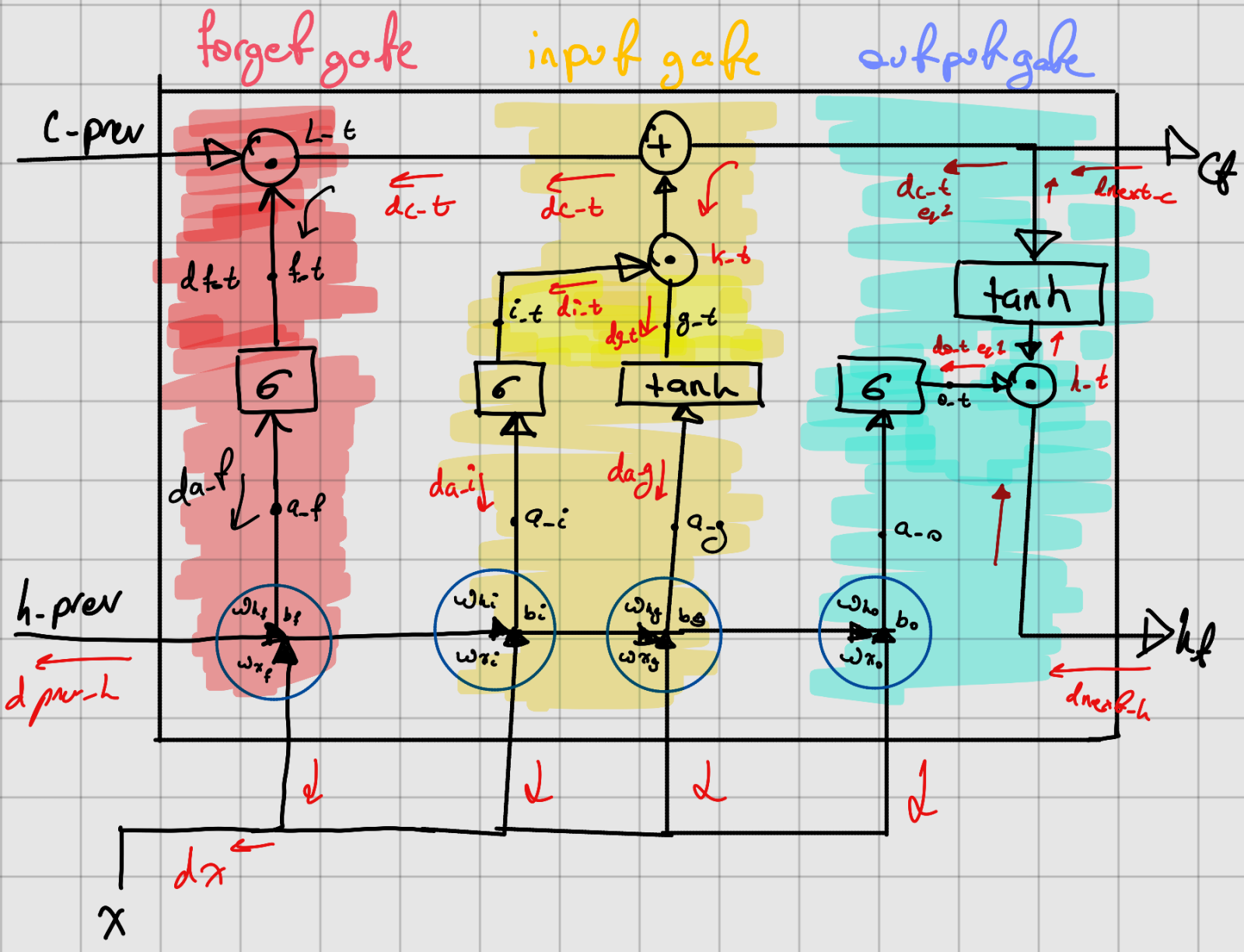

- $h_t = o_t \cdot \tanh(c_t)$

$$\therefore \frac{\partial h_t}{\partial o_t} = \tanh(c_t)$$

$$\therefore \underset{(N,H)}{do_t} = \underset{(N,H)}{d_{next\_h}} \odot \underset{(N,H)}{\tanh(c_t)} \quad \#eq\ 1$$

$\cdot \frac{\partial h_t}{\partial c_t} = o_t \cdot [1 - \tanh(c_t)]$

$\therefore \underset{(N,H)}{\partial c_t} = \underset{(N,H)}{d_{next\,h}} \odot \underset{(N,H)}{o_t} \odot (1 - \tanh^2(c_t)) + \underset{(N,H)}{d_{next\,c}}$ # eq 2

$\cdot k_t = i_t \odot g_t$

$\frac{\partial k_t}{\partial i_t} = g_t$

$\therefore \underset{(N,H)}{\partial i_t} = \underset{(N,H)}{d c_t} \odot \underset{(N,H)}{g_t}$ # eq 3

$\cdot \frac{\partial k_t}{\partial g_t} = i_t$

$\therefore \underset{(N,H)}{\partial g_t} = \partial c_t \odot \underset{(N,H)}{\partial i_t}$ # eq 4

$\cdot L_t = c_{prev} \odot f_t$

$\frac{\partial L_t}{\partial f_t} = c_{prev}$

$\therefore \underset{(N,H)}{\partial f_t} = \underset{(N,H)}{d c_t} \odot \underset{(N,H)}{c_{prev}}$ # eq 5

$\cdot \frac{\partial L_t}{\partial c_{prev}} = f_t$

$\therefore \underset{(N,H)}{\partial c_{prev}} = \underset{(N,H)}{d c_t} \odot \underset{(N,H)}{f_t}$ # eq 6

$\cdot d\sigma = \sigma(1 - \sigma)$

$\therefore da\_i = di\_t \odot i\_t \odot (1 - i\_t)$, as $i\_t = \sigma(a\_i)$ # eq 7
(N,H) (N,H) (N,H) (N,H)

$\therefore da\_f = df\_t \odot f\_t \odot (1 - f\_t)$ # eq 8
(N,H) (N,H) (N,H) (N,H)

$\therefore da\_o = do\_t \odot o\_t \odot (1 - o\_t)$ # eq 9
(N,H) (N,H) (N,H) (N,H)

- $d \tanh x = 1 - \tanh^2(x)$

$\therefore da\_g = dg\_t \odot (1 - g\_t^2)$ as $g\_t = \tanh(da\_g)$ # eq 10
(N,H) (N,H) (N,H)

- as $a = x\_t \cdot w\_x + h\_prev \cdot w\_h + b$,
(N,4H) (N,D) (D,4H) (N,H) (H,4H) (4H,-) (N,4H)

we first concatenate all da_i, da_g, da_f, da_o to form da
(N,H) (N,H) (N,H) (N,H)

then,

$$\frac{da}{dx\_t} = w\_x$$

$\therefore dx\_t = da \cdot w\_x^T$ # eq 11
(N,D) (N,4H) (4H,D)

$$\frac{da}{dw\_x} = x\_t$$

$\therefore dw\_x = x\_t^T \cdot da$ # eq 12
(D,4H) (D,N) (N,4H)

- $$\frac{da}{dw\_H} = h\_prev$$

$dw\_H = h\_prev^T \cdot da$ # eq 13
(H,4H) (H,N) (N,4H)

- $\frac{da}{dh\_prev} = w\_h$

(N, H) $(N, 4H)(4H, H)$

$dh\_prev = da \cdot w\_h^T$ # eq 14

- $\frac{\partial a}{\partial b} = 1$

$(4H, 1)$ $(4H, 1)$

$db = \sum_{n=1}^{N} da$ # eq 15

$(N, 4H)$