

*timestep in decoder*

$\mathbf{b}^{t_d}$

$h_\ell$

*unit index in layer  $\ell$*