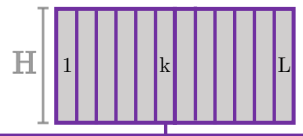


Layer input, $\mathbf{u}_{1:L}$



Legend:

Learnable parameter

Operation

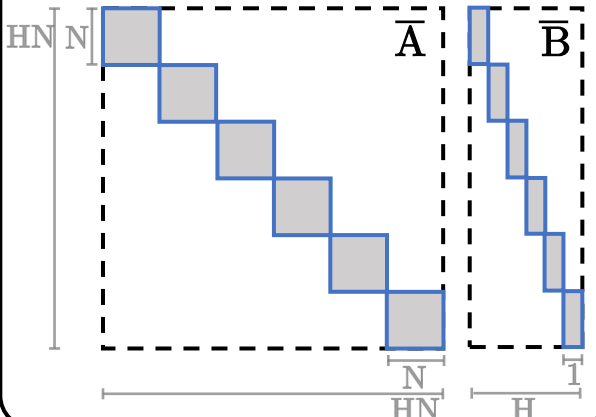
Instantiated variable

Message between layers

\mathbf{u}_k

H independent S4 SSMs:

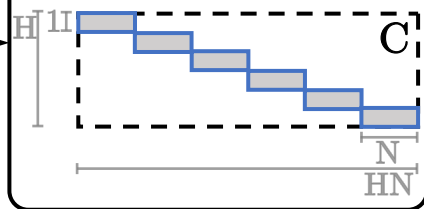
$$\mathbf{x}_k^{(1:H)} = \bar{\mathbf{A}}\mathbf{x}_{k-1}^{(1:H)} + \bar{\mathbf{B}}\mathbf{u}_k$$



$\mathbf{x}_k^{(1:H)}$

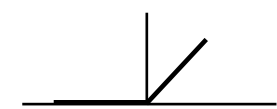
Emission:

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k^{(1:H)}$$



\mathbf{y}_k

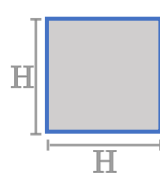
Nonlinearity



\mathbf{y}'_k

Mixing layer:

$$\mathbf{u}'_k = \mathbf{f}(\mathbf{y}'_k)$$



\mathbf{u}'_k

Single S4 layer

Layer output, $\mathbf{u}'_{1:L}$

