

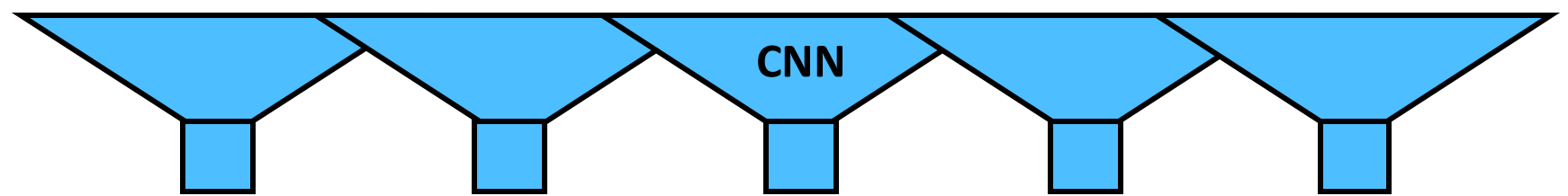
Recording
(raw audio data)

\mathcal{X}



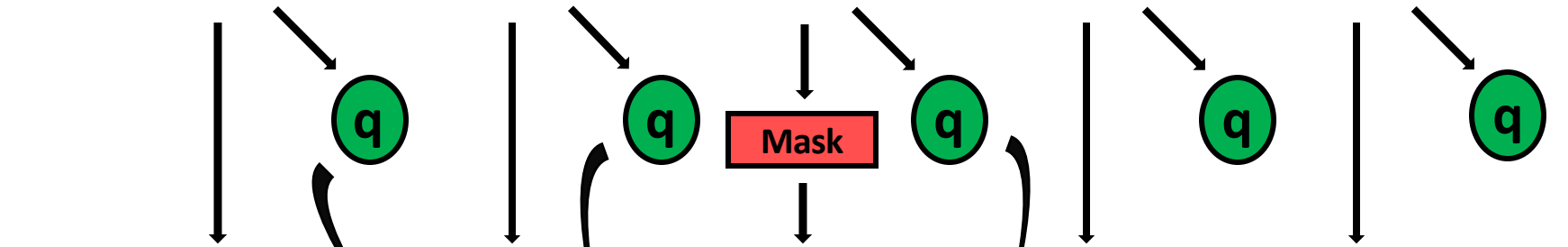
Feature
Encoder

\mathcal{Z}



Quantization
Module

\mathcal{Q}



Context
Network

\mathcal{C}

