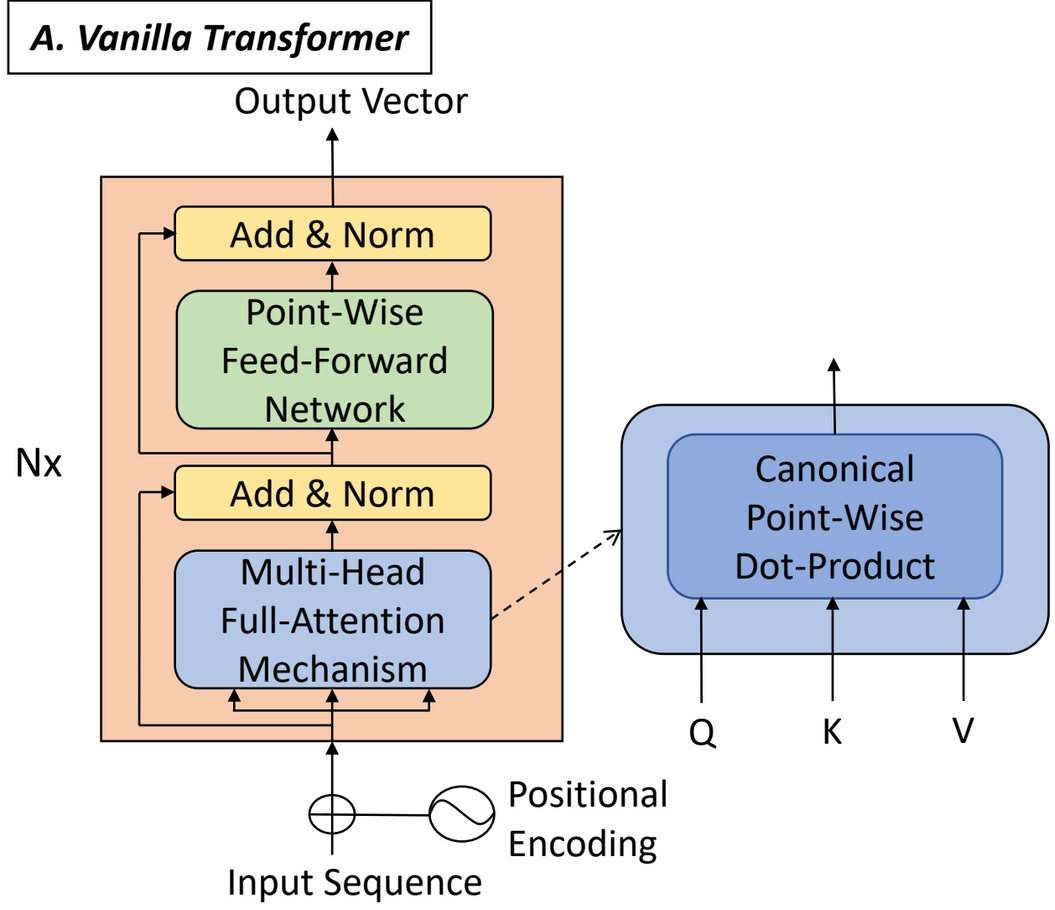
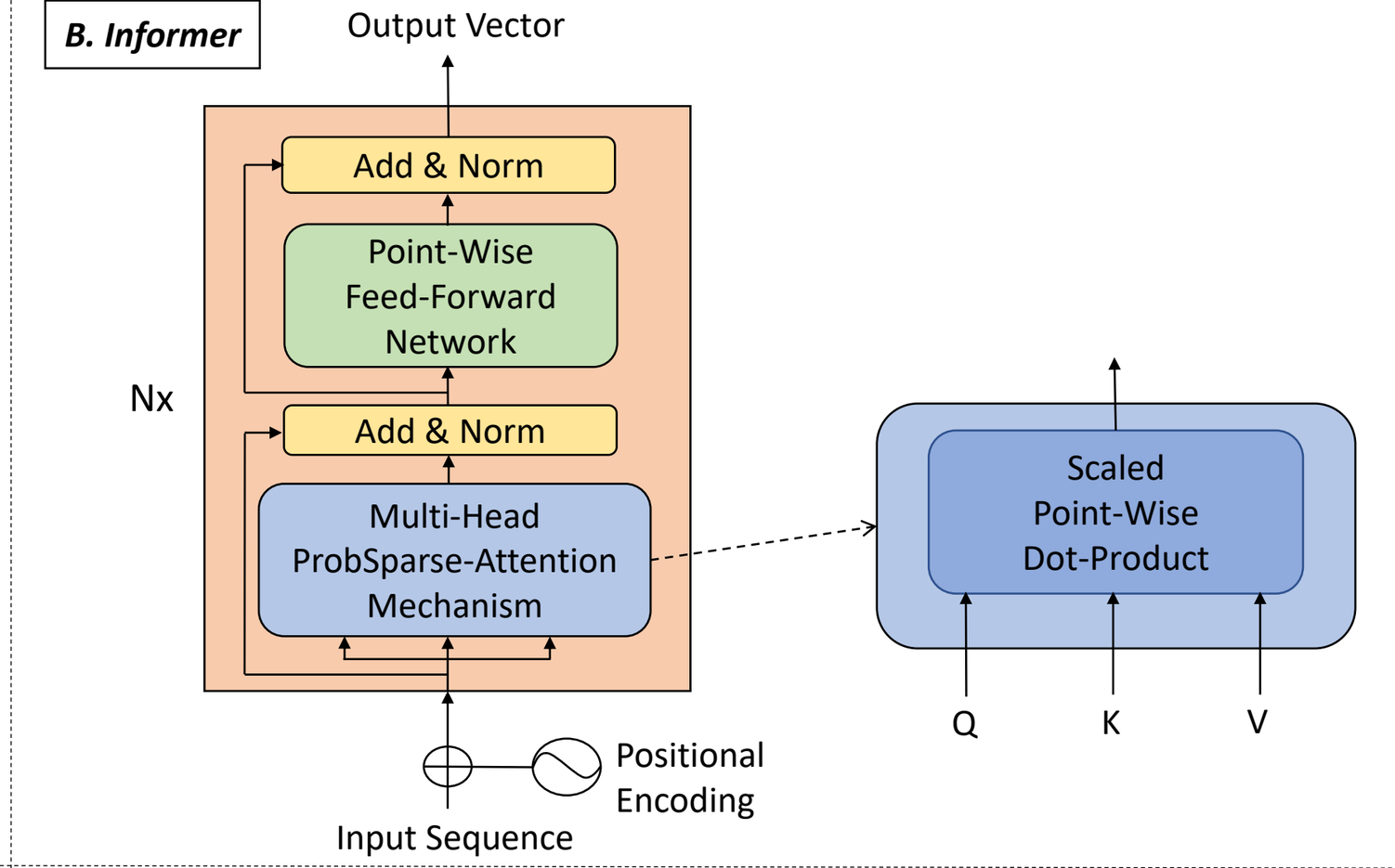
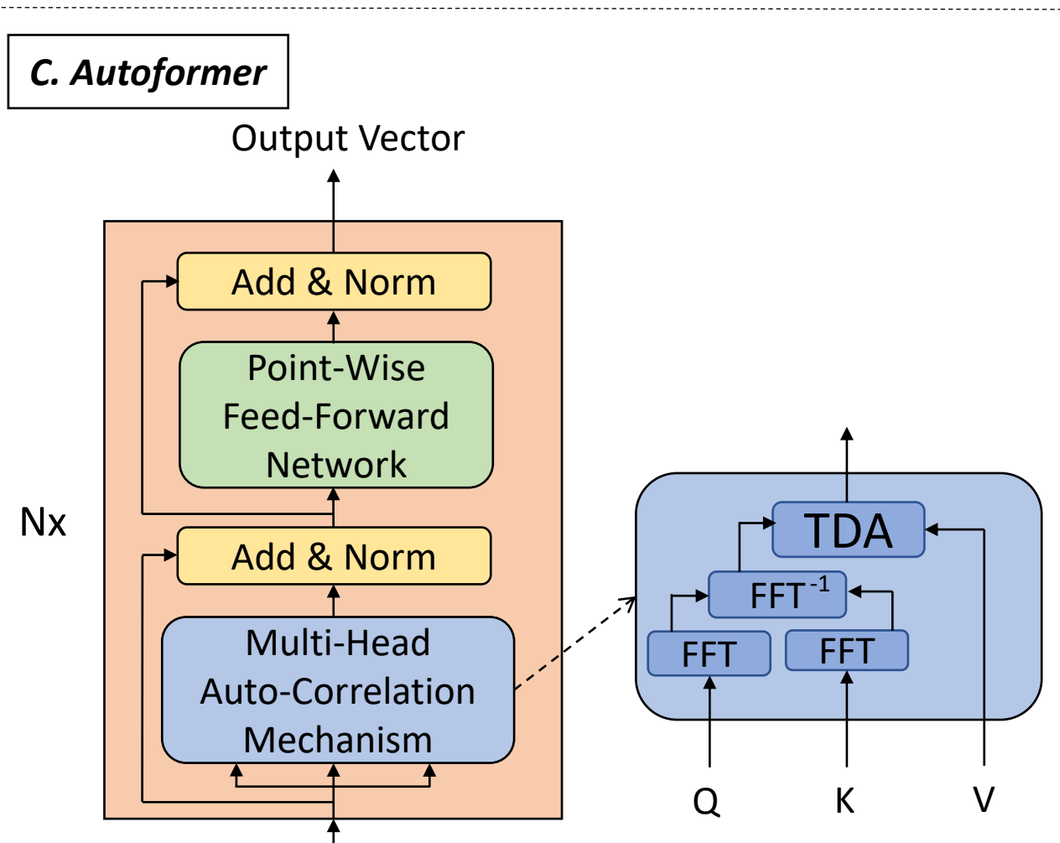
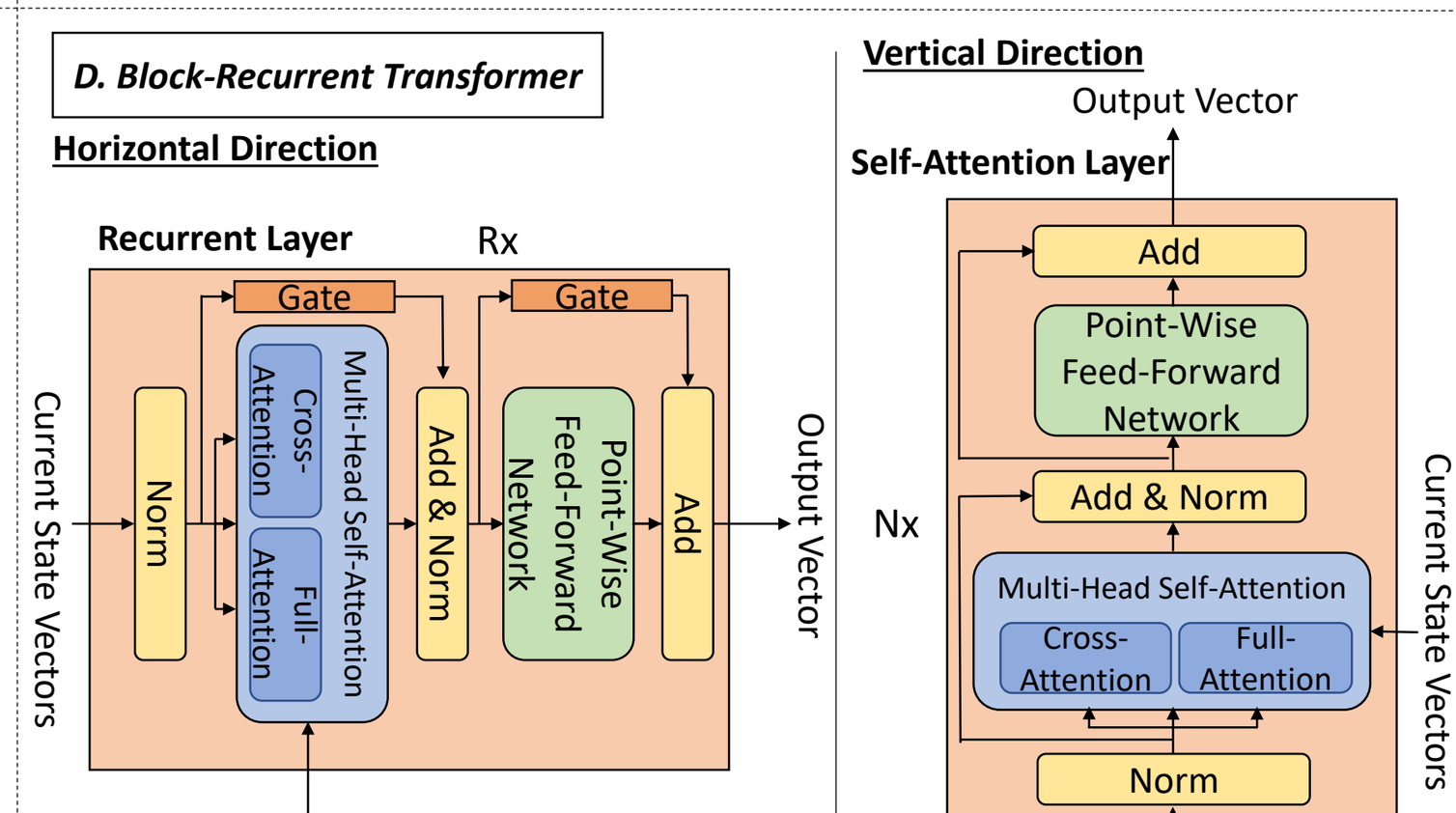
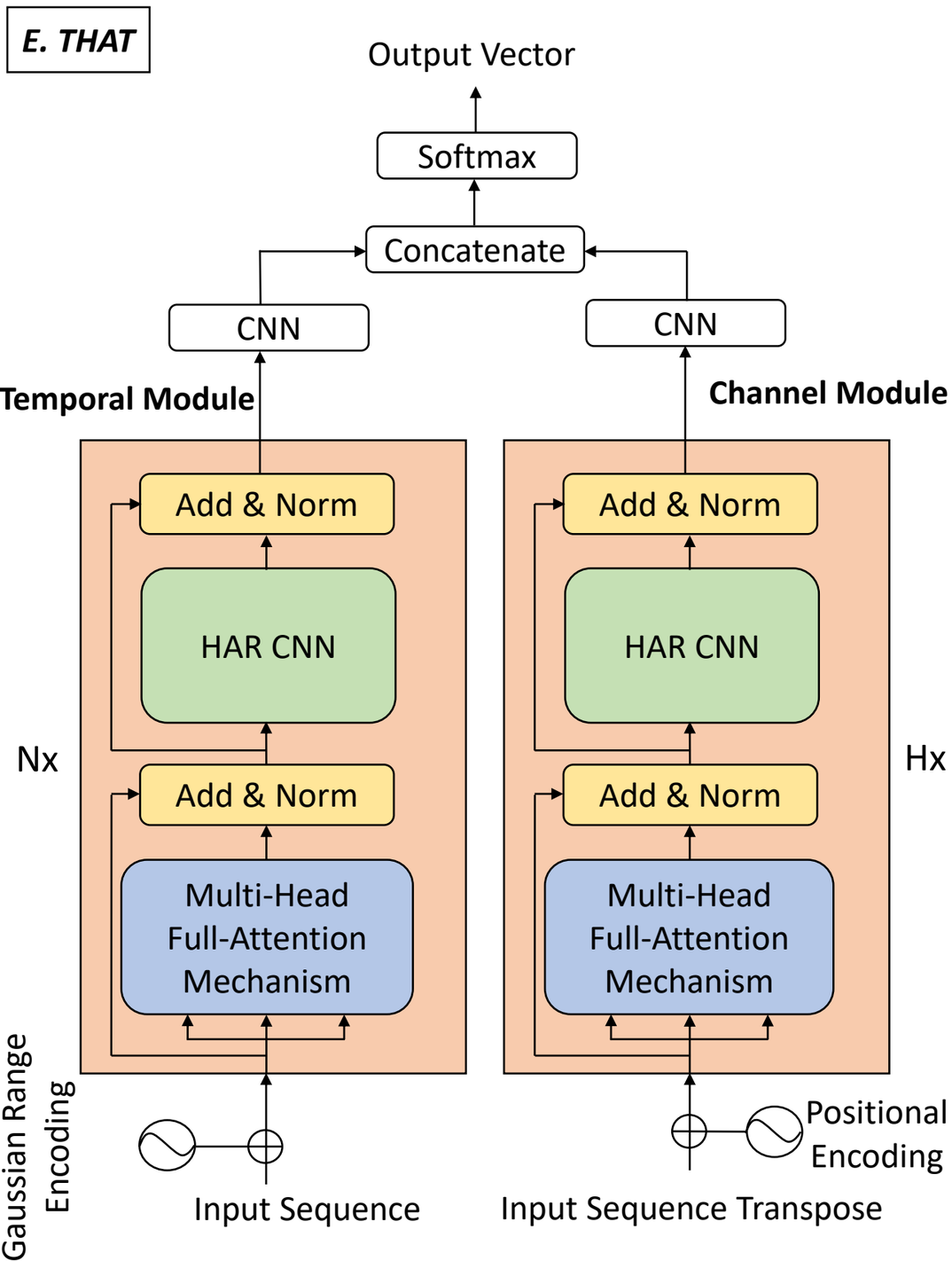
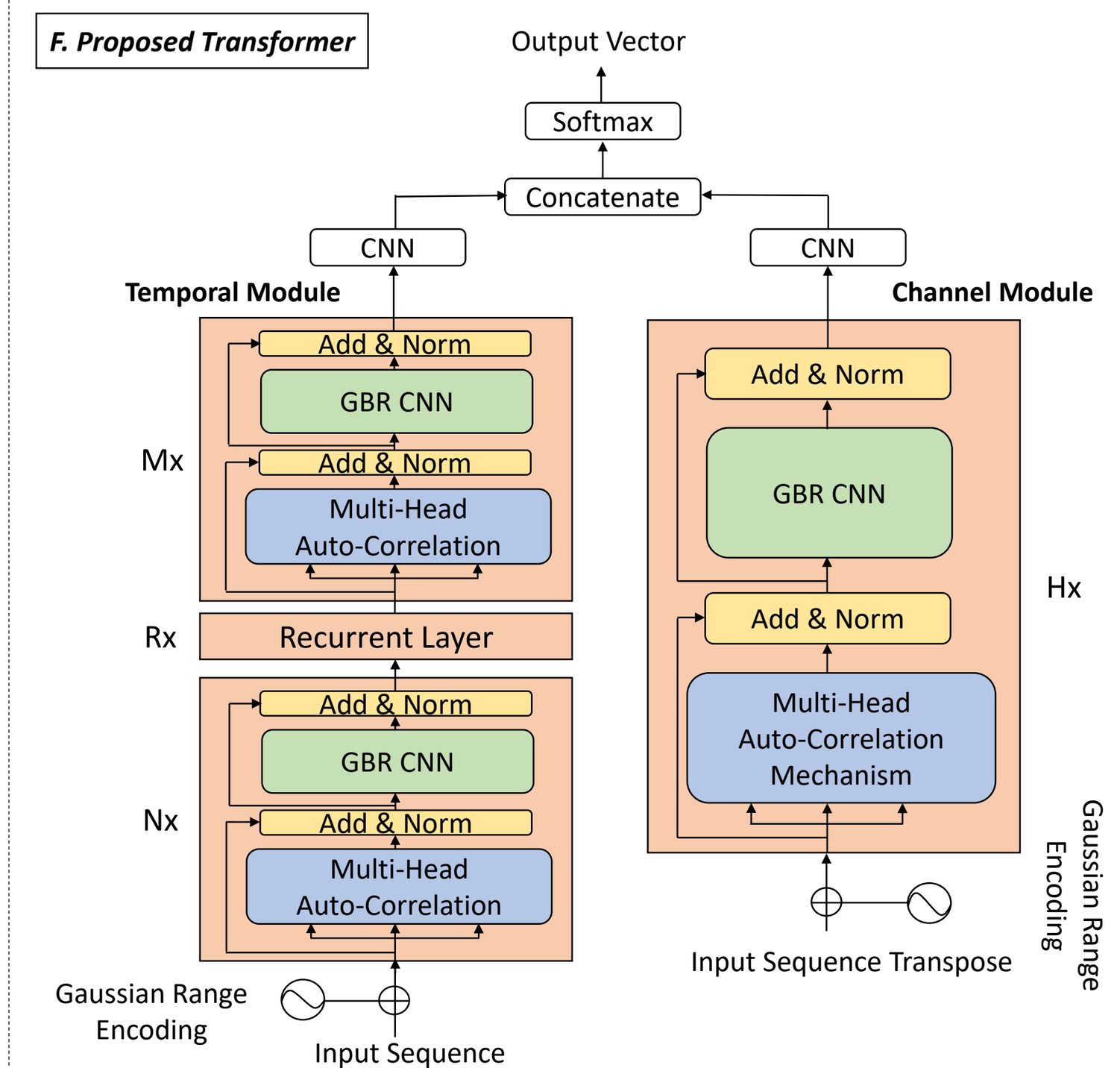

A. Vanilla Transformer
Output Vector
Add & Norm
Point-Wise Feed-Forward Network
Add & Norm
Multi-Head Full-Attention Mechanism
Nx
Canonical Point-Wise Dot-Product
Q
K
V
Positional Encoding
Input Sequence


B. Informer
Output Vector
Add & Norm
Point-Wise Feed-Forward Network
Add & Norm
Multi-Head ProbSparse-Attention Mechanism
Nx
Scaled Point-Wise Dot-Product
Q
K
V
Positional Encoding
Input Sequence


C. Autoformer
Output Vector
Add & Norm
Point-Wise Feed-Forward Network
Add & Norm
Multi-Head Auto-Correlation Mechanism
Nx
TDA
FFT-1
FFT
FFT
Q
K
V
Positional Encoding
Input Sequence


D. Block-Recurrent Transformer
Horizontal Direction
Recurrent Layer
Rx
Gate
Gate
Norm
Cross-Attention
Full-Attention
Multi-Head Self-Attention
Add & Norm
Point-Wise Feed-Forward Network
Add
Current State Vectors
Output Vector
Positional Encoding
Input Sequence
Vertical Direction
Output Vector
Self-Attention Layer
Add
Point-Wise Feed-Forward Network
Add & Norm
Multi-Head Self-Attention
Cross-Attention
Full-Attention
Norm
Nx
Current State Vectors
Positional Encoding
Input Sequence


E. THAT
Output Vector
Softmax
Concatenate
CNN
CNN
Temporal Module
Channel Module
Add & Norm
HAR CNN
Add & Norm
Multi-Head Full-Attention Mechanism
Nx
Hx
Gaussian Range Encoding
Positional Encoding
Input Sequence
Input Sequence Transpose


F. Proposed Transformer
Output Vector
Softmax
Concatenate
CNN
CNN
Temporal Module
Channel Module
Add & Norm
GBR CNN
Add & Norm
Multi-Head Auto-Correlation
Mx
Rx
Recurrent Layer
Add & Norm
GBR CNN
Add & Norm
Multi-Head Auto-Correlation
Nx
Add & Norm
GBR CNN
Add & Norm
Multi-Head Auto-Correlation Mechanism
Hx
Gaussian Range Encoding
Input Sequence
Input Sequence Transpose
Gaussian Range Encoding