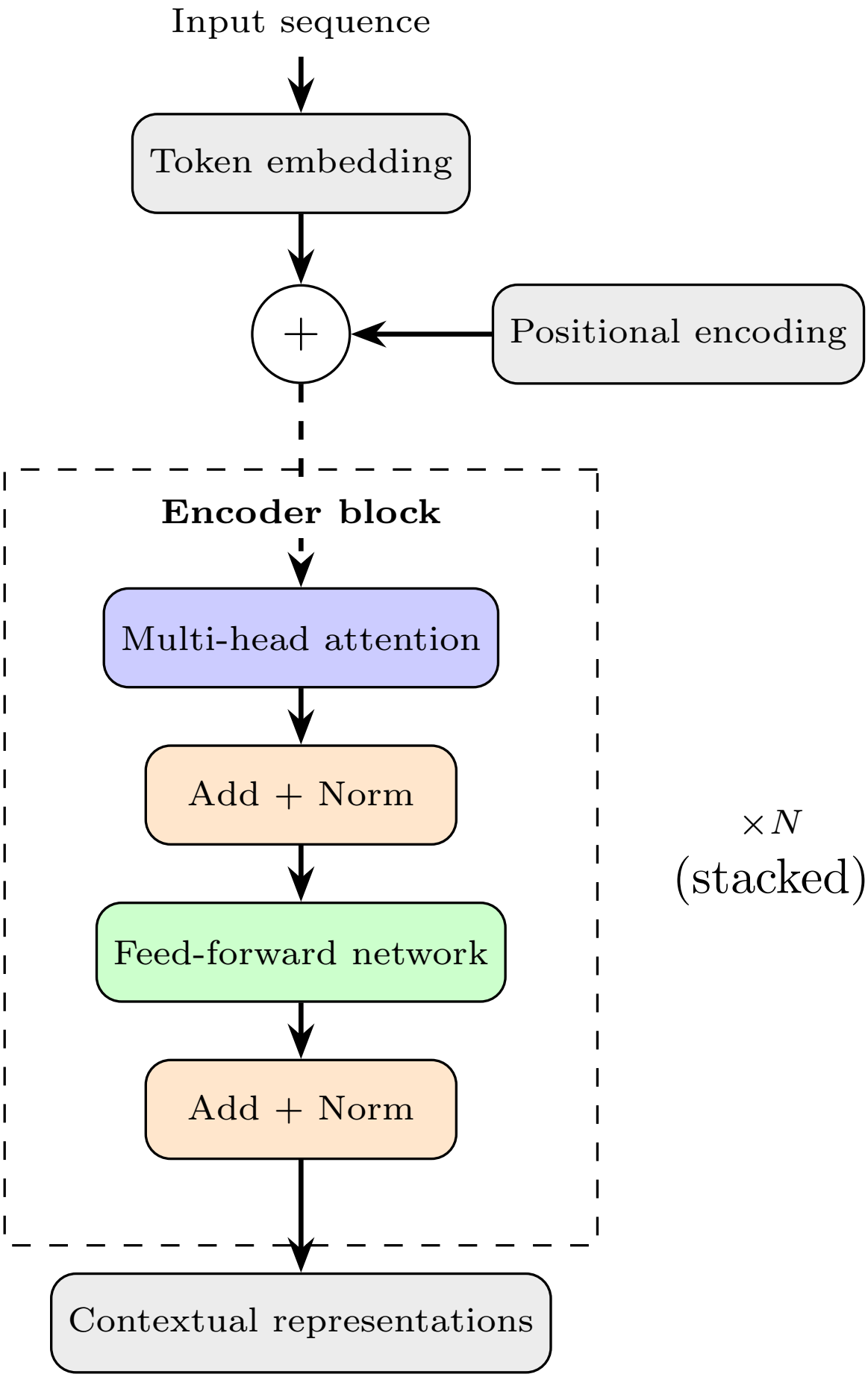
Input sequence
Token embedding
+
Positional encoding
Encoder block
Multi-head attention
Add + Norm
Feed-forward network
Add + Norm
Contextual representations
×N
(stacked)