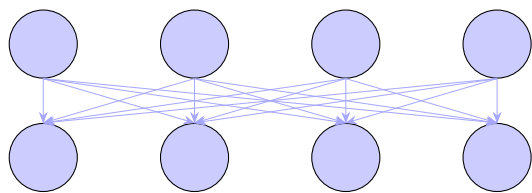
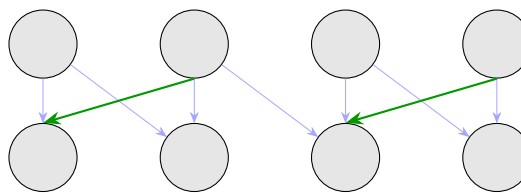


Standard Transformer

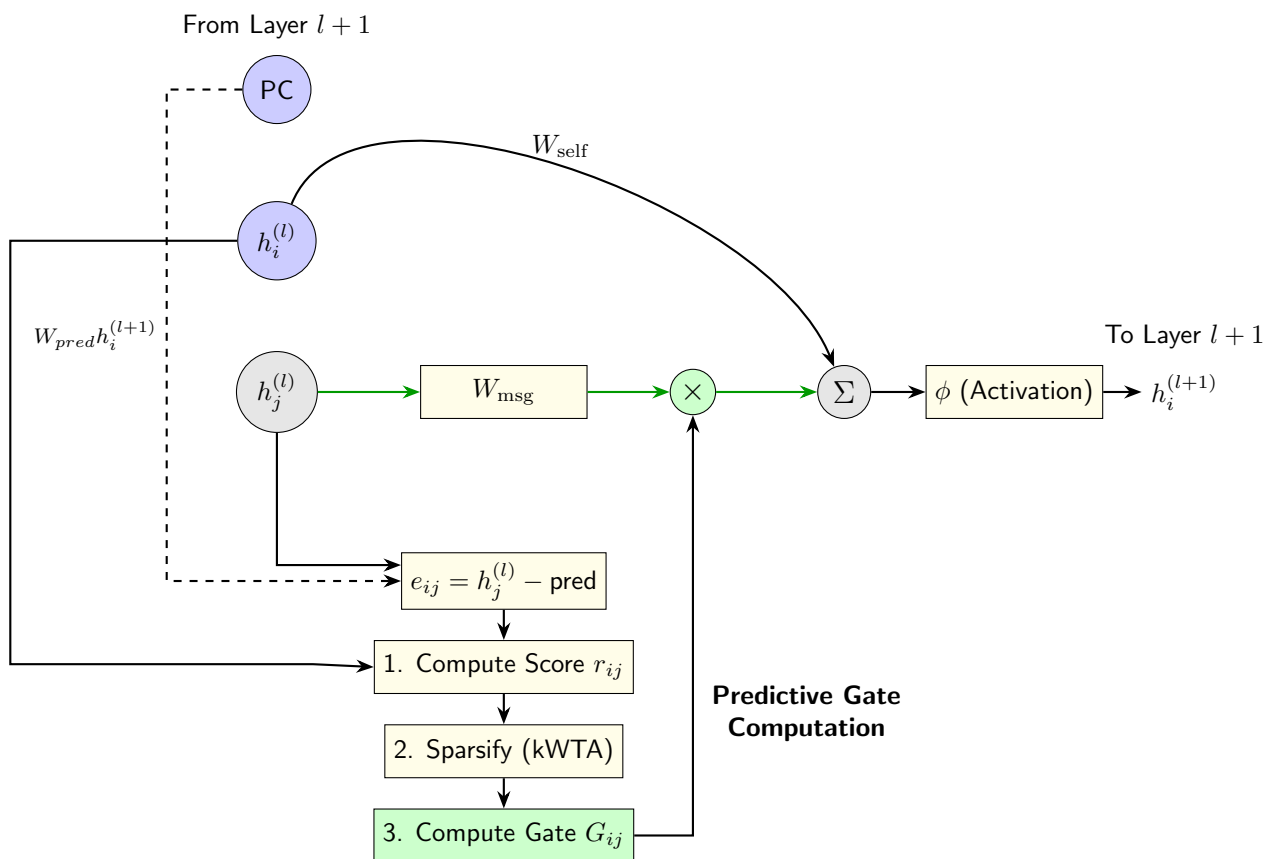


Dense, All-to-All Broadcast
Quadratic $O(N^2)$ Complexity

AFMP



Sparse, Predictive Routing
Guaranteed $O(N)$ Complexity



**Attention-Free Perceptual
Capsule Unit (PCU) Update**