



**Diagram of one GATv2 layer with  $K$  attention heads**  
 (shapes of matrices are indicated in grey)