

Operator norms for Attention & MLP

