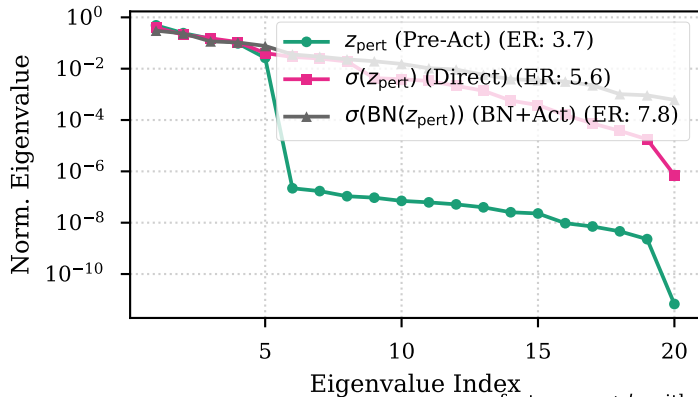


Synergy of Normalization and Activation ($k_{in}=5$, $d_{full}=20$)

ReLU (σ)



z_{pert} : per-feature $a_j x_j + b_j$ with $a_j \sim U[0, 2]$, $b_j \sim U[0, 1]$

Tanh (σ)

