attention layer 11  $10^{5}$ Addition Shuffled Addition Multiplication  $10^{3}$ Addition and Multiplication Singular value Random Tokens  $10^{1}$ Training Subset Random Token Addition  $10^{-3}$ 100 200 300 400 500 700 600 800 Rank