Rebuttal Plots

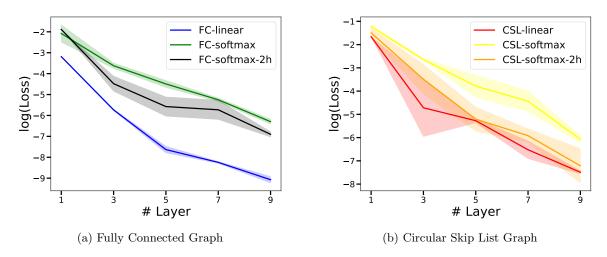


Figure 1: Comparison of loss for the Electric Flow Problem (Figure 1(a) of paper) against number of layers, taken at convergence. "linear" denotes 1-head linear Transformer. "softmax" denotes 1-head softmax-activated Transformer. "softmax-2h" denotes 2-head softmax-activated Transformer.

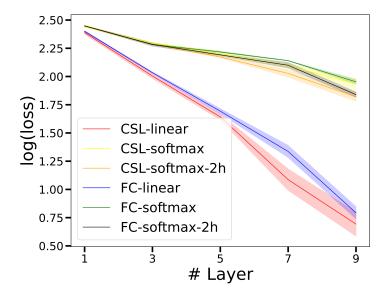


Figure 2: Comparison of loss for the Eigenvector Problem (Figure 2(d) of paper) against number of layers, taken at convergence. "linear" denotes 1-head linear Transformer. "softmax" denotes 1-head softmax-activated Transformer. "softmax-2h" denotes 2-head softmax-activated Transformer.