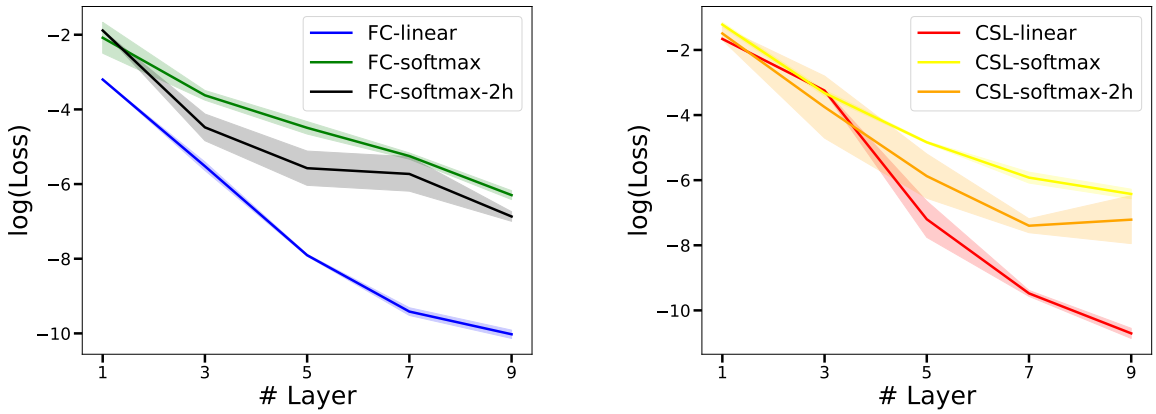
Rebuttal Plots

1 Comparison with Softmax



(a) Fully Connected Graph

(b) Circular Skip List Graph

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.

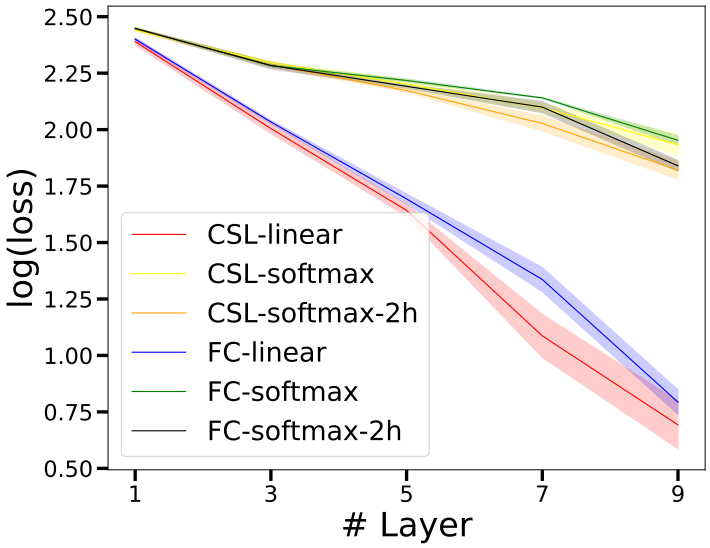
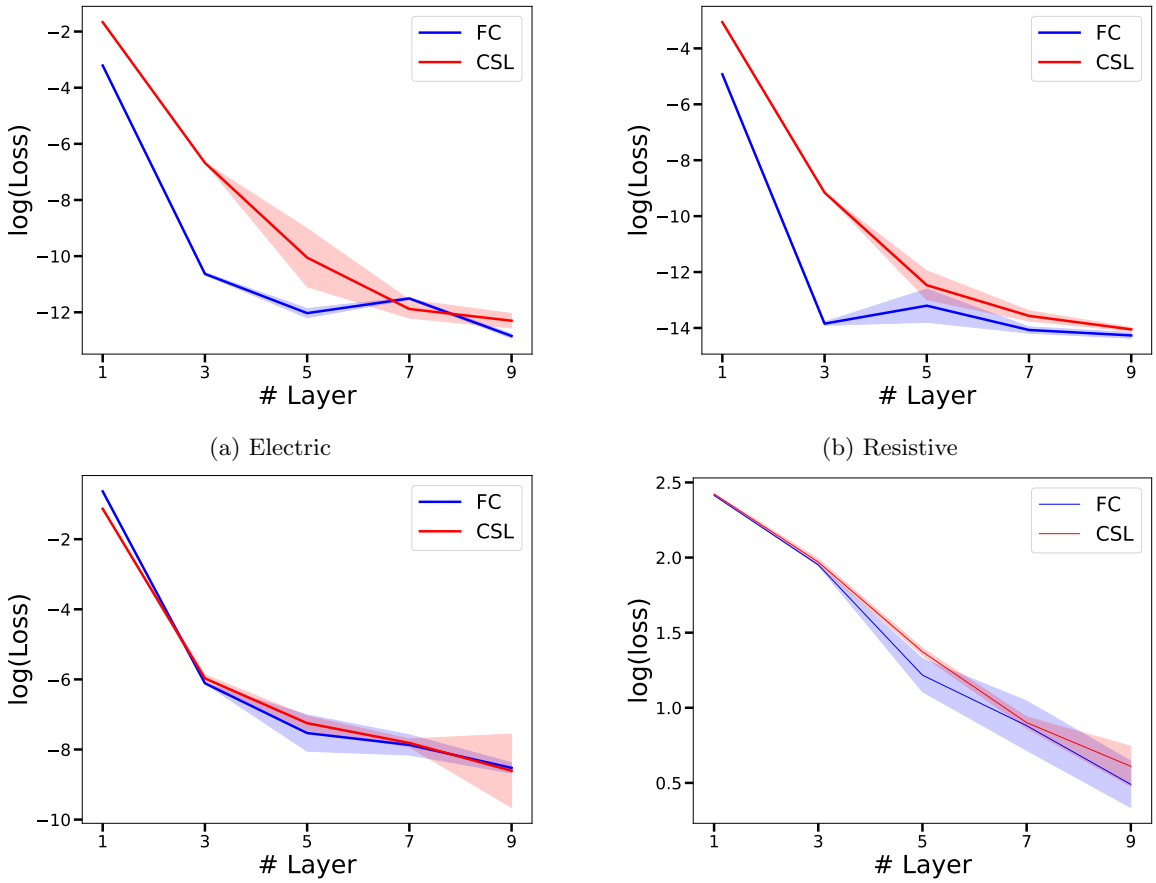


Figure 2: Comparison of loss for the Eigenvector Problem (Figure 2(d) of paper) against number of layers, taken at convergence. "linear" denotes 1-head linear Transformer. "softmax" denotes 1-head softmax-activated Transformer. "softmax-2h" denotes 2-head softmax-activated Transformer.

2 Evaluating Efficient Implementation on Synthetic Data



(a) Electric

(b) Resistive

(c) Heat

(d) Eigenvector

Figure 3: log(loss) against number of layers for the Electric, Resistive, Heat, Eivengector problems. The Transformer implementation is the Efficient Implementation described in Section 5. All values are taken after convergence.

Figure 3(a) maps to Figure 1(a) of the paper.
Figure 3(b) maps to Figure 1(b) of the paper.
Figure 3(c) maps to Figure 1(c) of the paper.
Figure 3(d) maps to Figure 2(d) of the paper.