Estimating $\mathrm{tr}(\mathrm{NTK})$: Speedup v. Accuracy

A Example 1
MLP Model (n=3.2k, n_θ=64.7k)

B Example 2
GRU Model (n=48k, n_θ=14.5k)