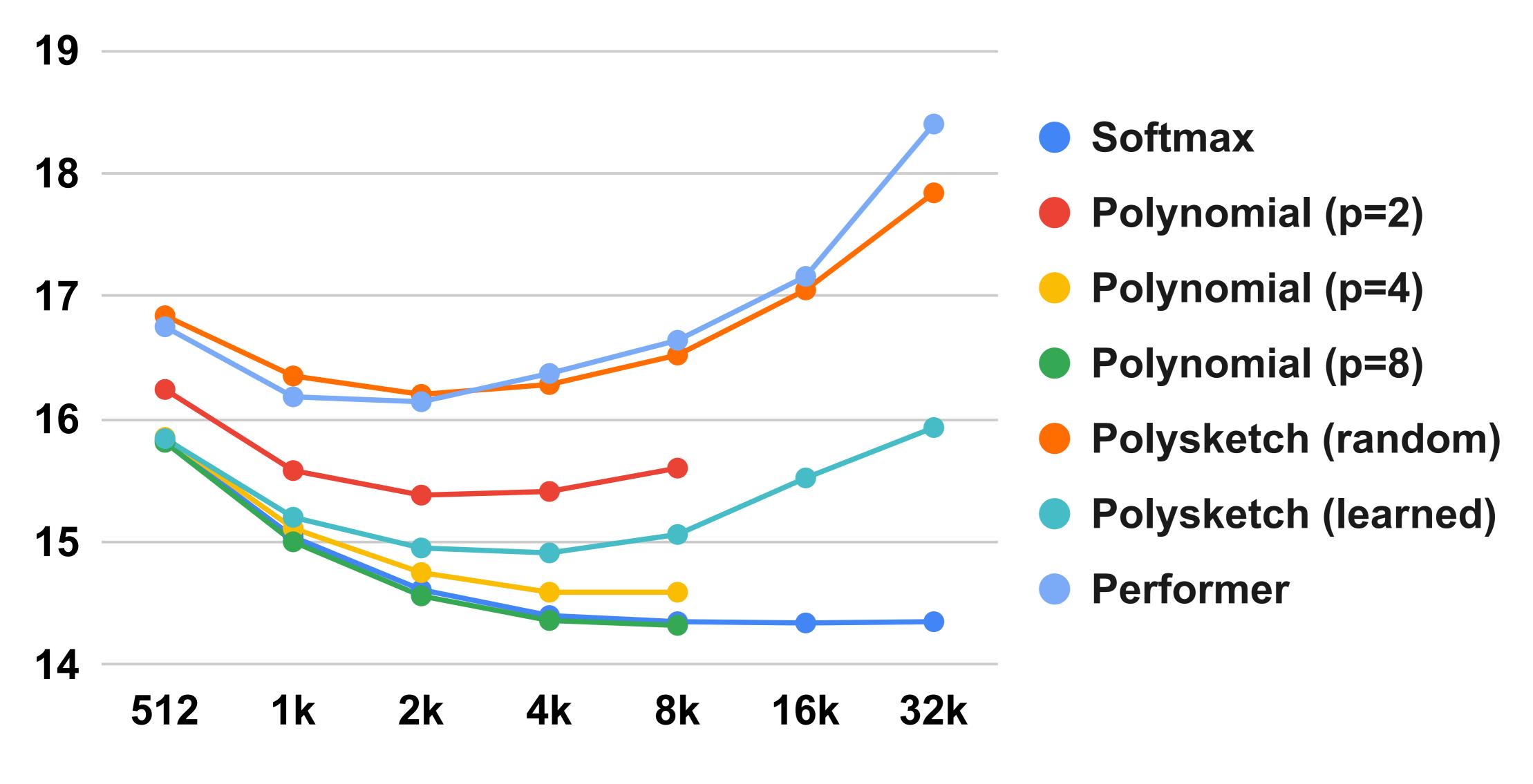
Perplexities

Perplexities on Wiki-40B

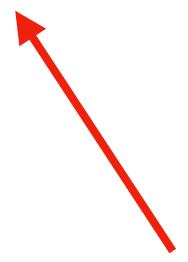


Context Length

We will later show how the the quality can be further improved

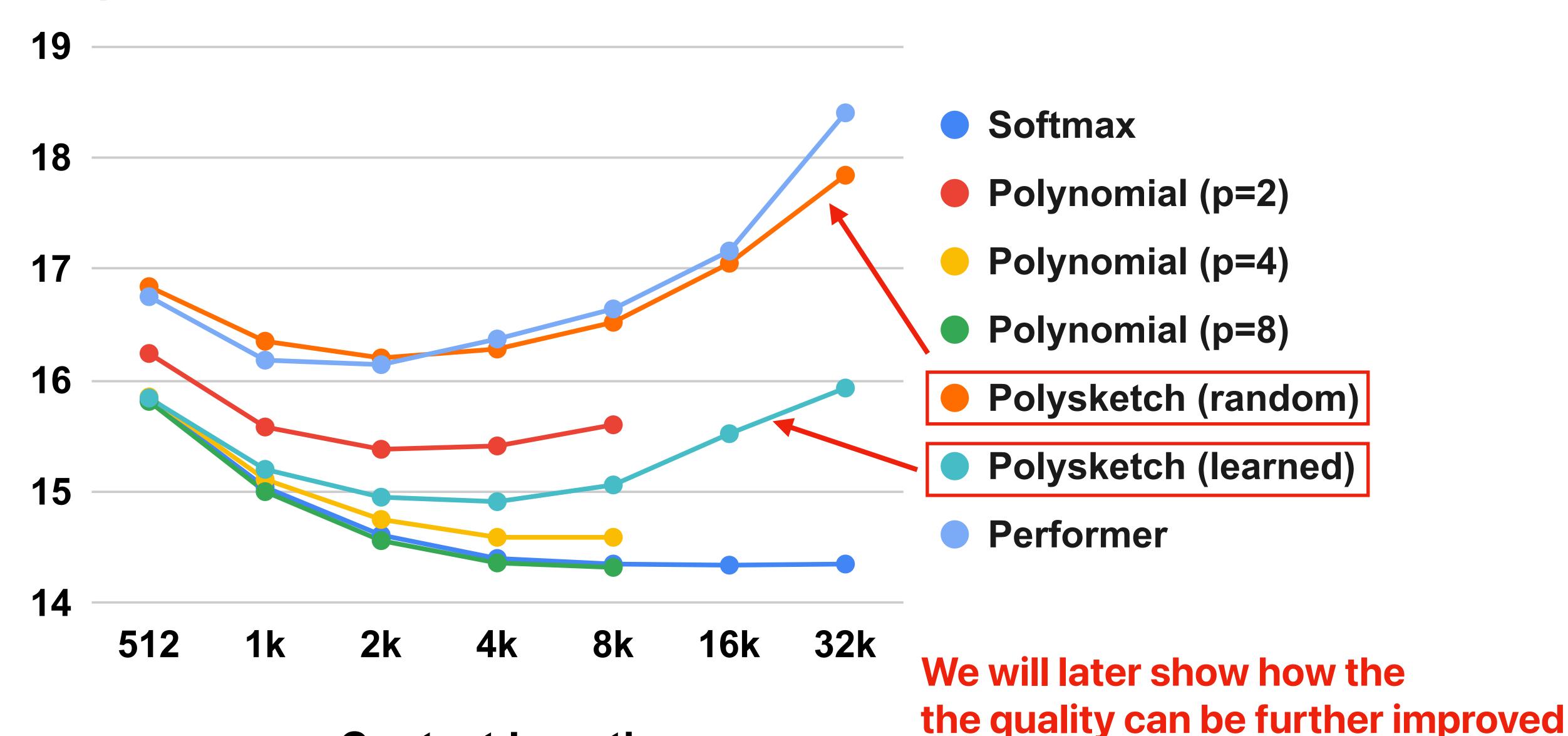








Perplexities on Wiki-40B



Context Length

32

Computing $LT(A \cdot B^T) \cdot C$