Train Perplexity ChristBERT $ChristBERT_{\mathsf{scratch}}$ ChristBERT_{BPE} 10^{4} perplexity (log) 10^{2} 10^{1} 60000 100000 20000 40000 80000 steps