Transformer test loss on random tasks with different regular alphabet sizes 2.00 – n8 – n 32 1.75 n 128 n 256 1.50 n 512 n 1024 1.25 n 2048 n 8192 Test loss 1.00 0.75 0.50 0.25 0.00 0.50 0.00 0.25 0.75 1.00 1.25 1.50 1.75 2.00 regular alphabet size