Transformer test loss vs. learning rate and embedding dimension, at n = 1024