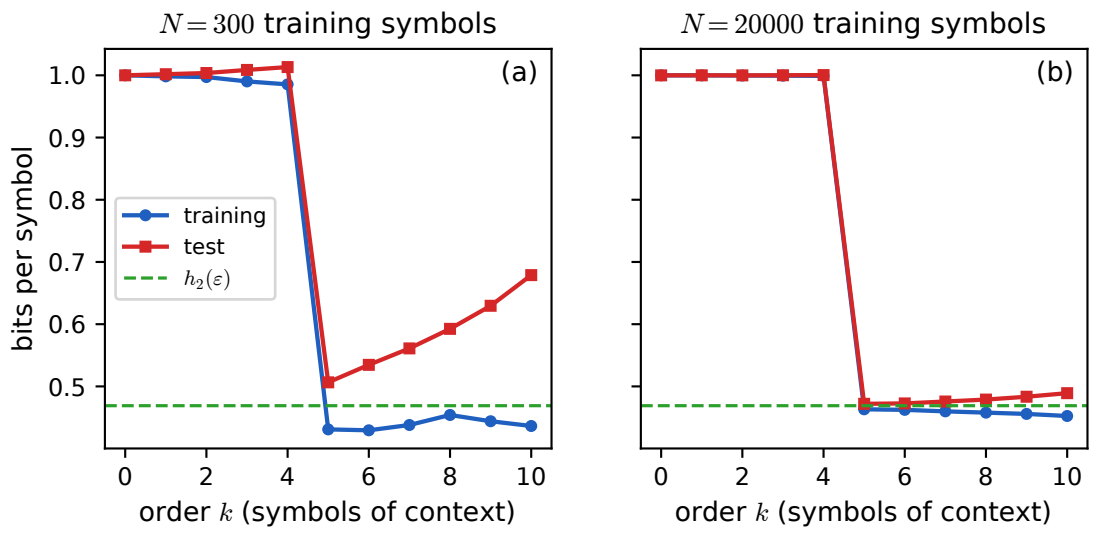

$N = 300$ training symbols
$N = 20000$ training symbols
(a)
(b)
bits per symbol
1.0
0.9
0.8
0.7
0.6
0.5
training
test
$h_2(\varepsilon)$
0
2
4
6
8
10
order k (symbols of context)
order k (symbols of context)