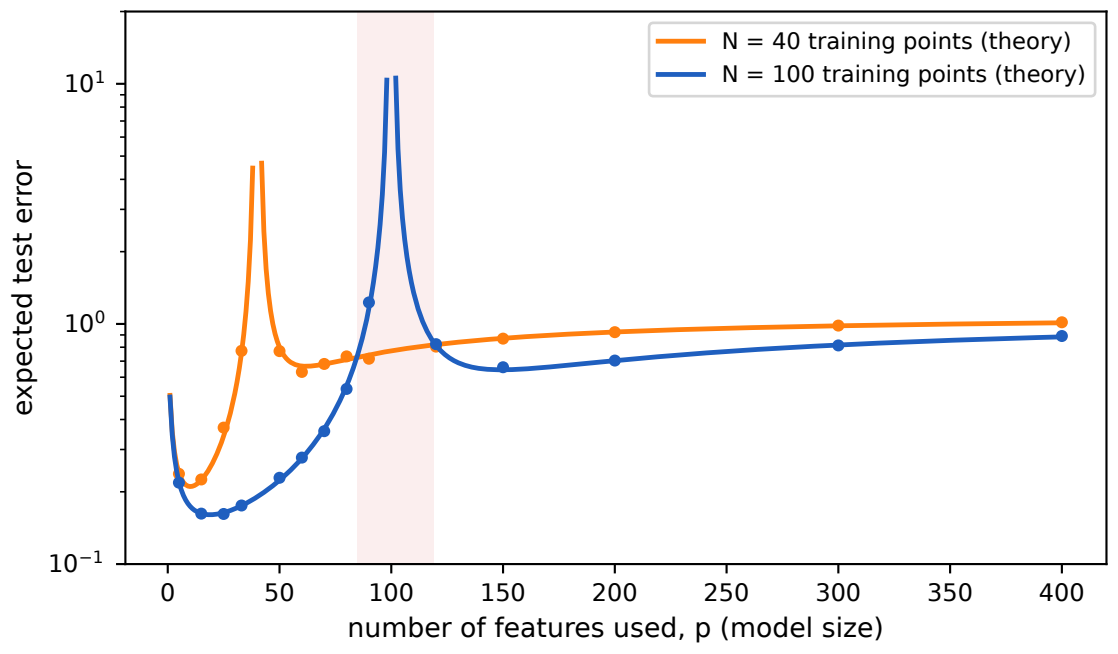

N = 40 training points (theory)
N = 100 training points (theory)
expected test error
10^1
10^0
10^{-1}
0
50
100
150
200
250
300
350
400
number of features used, p (model size)