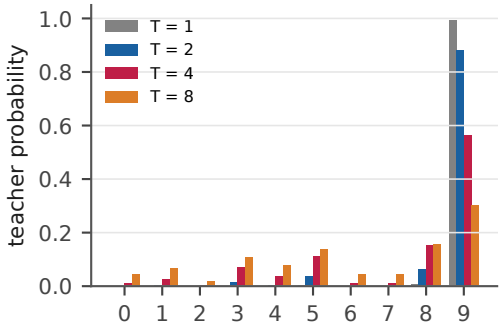
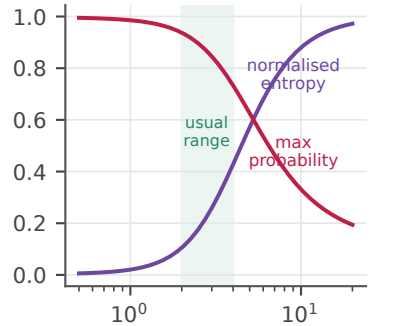
input: a "9"




same logits, four temperatures
T = 1
T = 2
T = 4
T = 8
teacher probability
1.0
0.8
0.6
0.4
0.2
0.0
0 1 2 3 4 5 6 7 8 9


averaged over 450 test digits
normalised entropy
usual range
max probability
1.0
0.8
0.6
0.4
0.2
0.0
10^0
10^1
temperature T