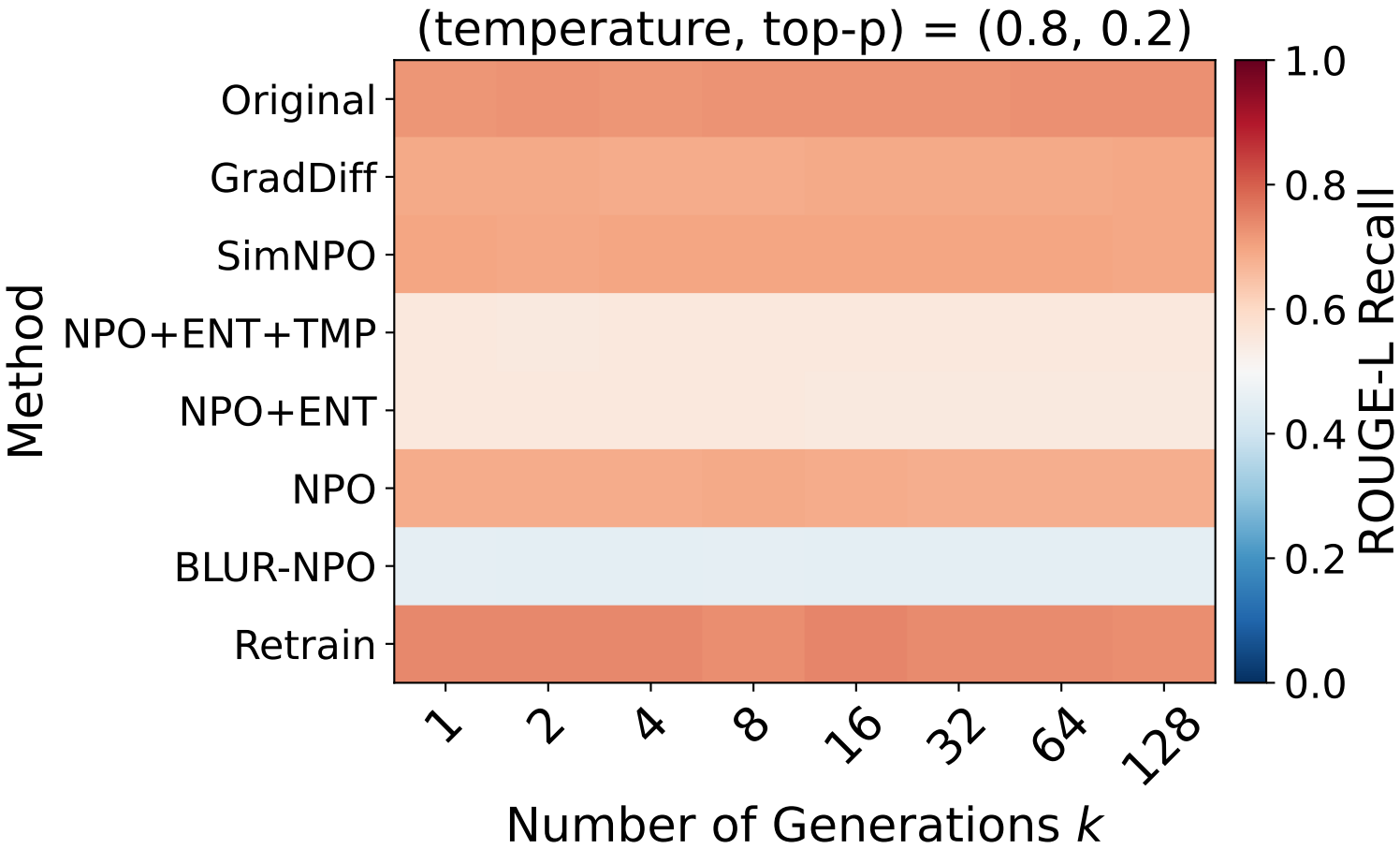
(temperature, top-p) = (0.8, 0.2)
Method
Original
GradDiff
SimNPO
NPO+ENT+TMP
NPO+ENT
NPO
BLUR-NPO
Retrain
1
2
4
8
16
32
64
128
Number of Generations k
ROUGE-L Recall
1.0
0.8
0.6
0.4
0.2
0.0