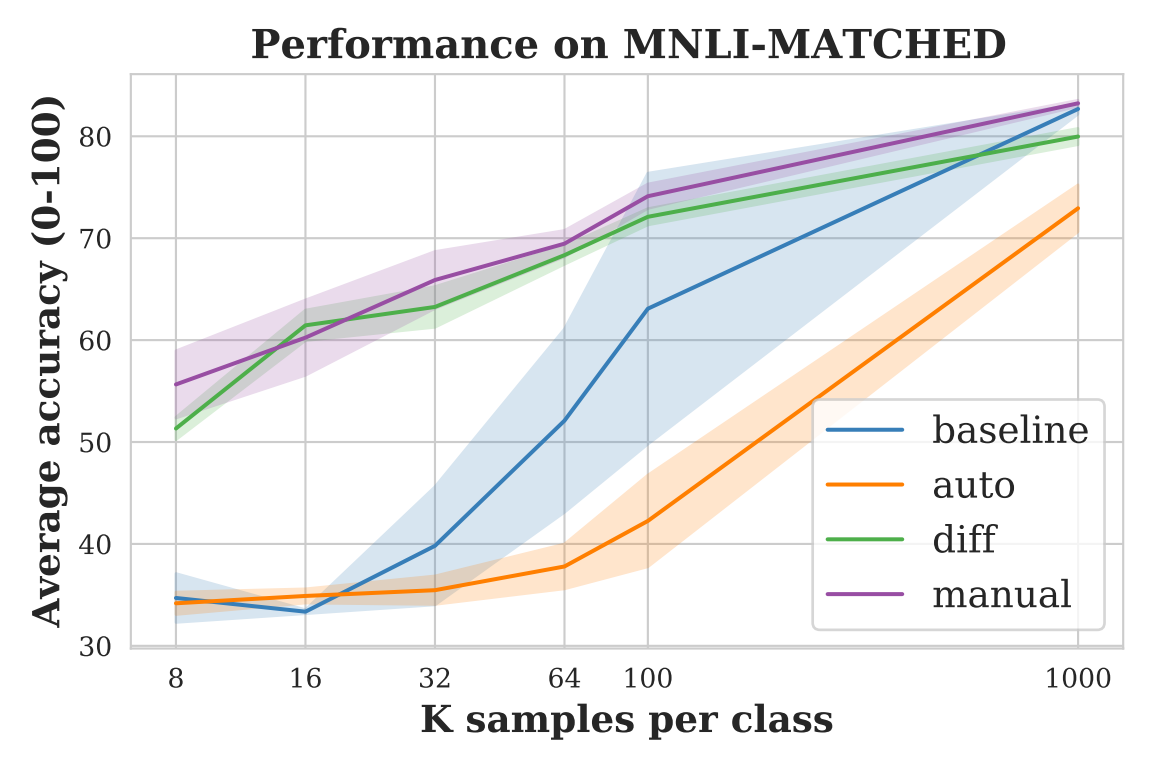
Performance on MNLI-MATCHED
Average accuracy (0-100)
30
40
50
60
70
80
8
16
32
64
100
1000
K samples per class
baseline
auto
diff
manual