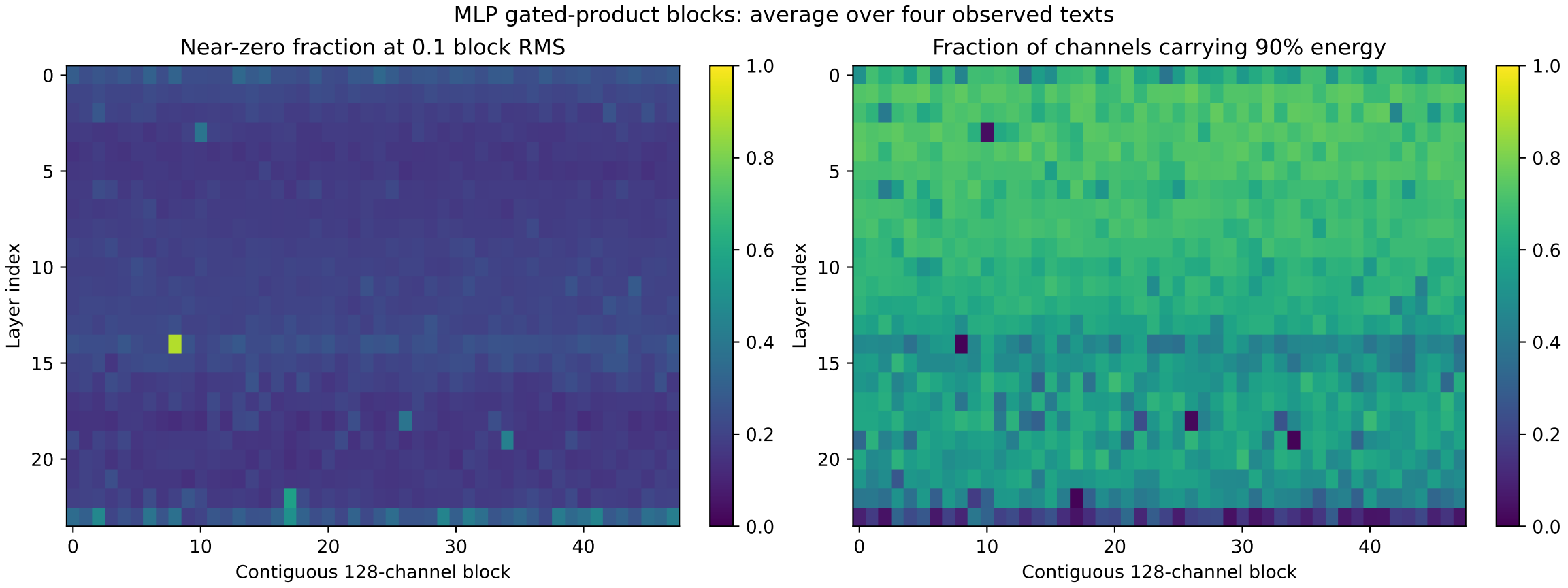

MLP gated-product blocks: average over four observed texts
Near-zero fraction at 0.1 block RMS
Fraction of channels carrying 90% energy
Layer index
Contiguous 128-channel block
0
5
10
15
20
0
10
20
30
40
0.0
0.2
0.4
0.6
0.8
1.0