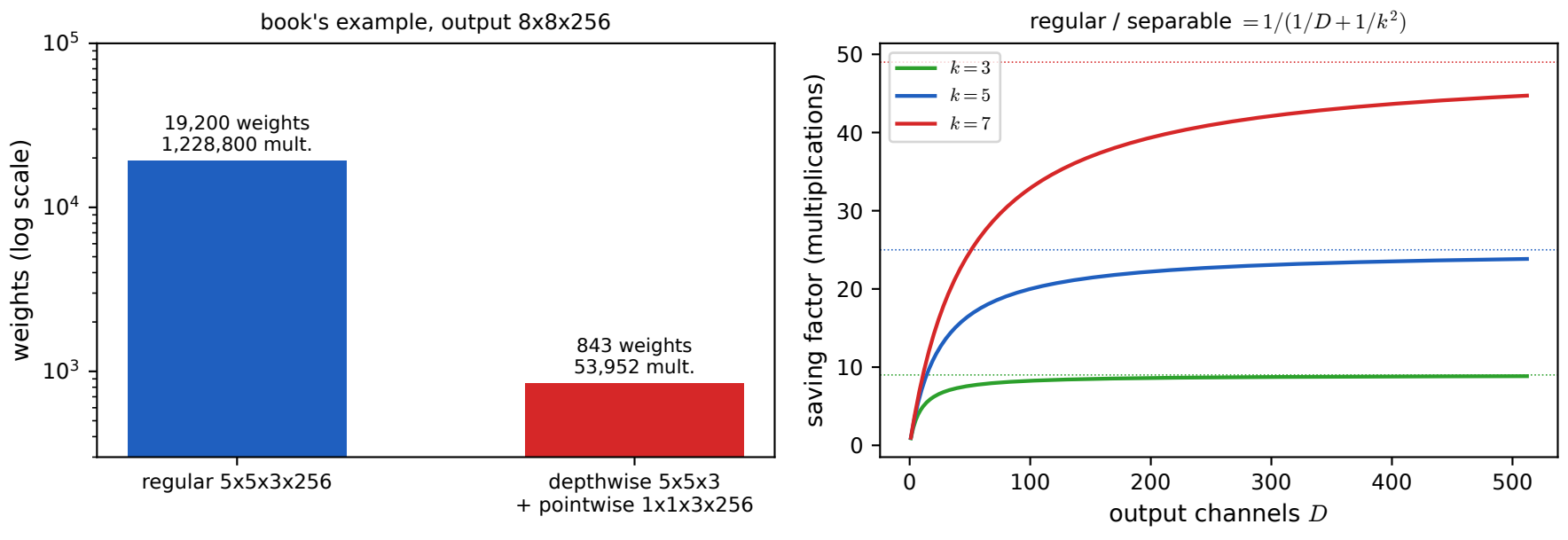

book's example, output 8x8x256
weights (log scale)
10^5
10^4
10^3
19,200 weights
1,228,800 mult.
843 weights
53,952 mult.
regular 5x5x3x256
depthwise 5x5x3
+ pointwise 1x1x3x256
regular / separable $=1/(1/D+1/k^2)$
saving factor (multiplications)
0
10
20
30
40
50
$k=3$
$k=5$
$k=7$
0
100
200
300
400
500
output channels D