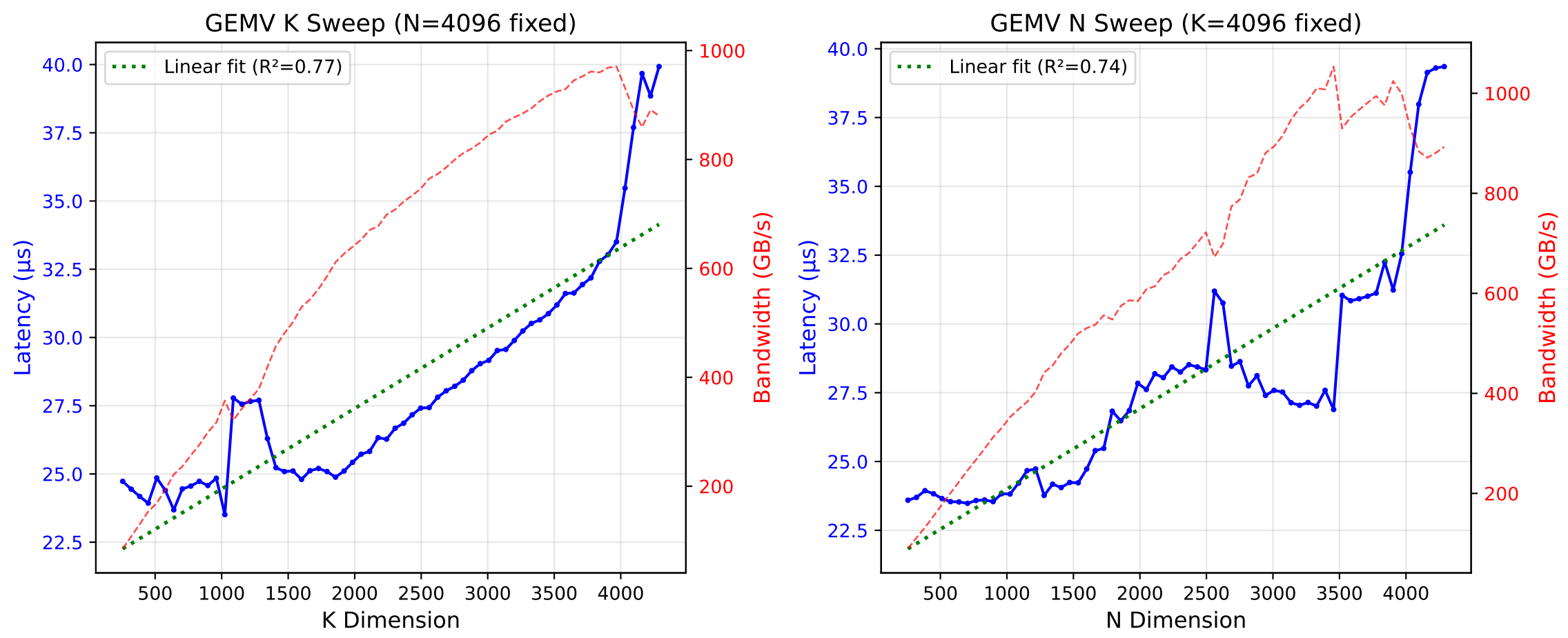
GEMV K Sweep (N=4096 fixed)
Linear fit (R²=0.77)
Latency (µs)
Bandwidth (GB/s)
K Dimension
40.0
37.5
35.0
32.5
30.0
27.5
25.0
22.5
1000
800
600
400
200
500
1000
1500
2000
2500
3000
3500
4000
GEMV N Sweep (K=4096 fixed)
Linear fit (R²=0.74)
Latency (µs)
Bandwidth (GB/s)
N Dimension