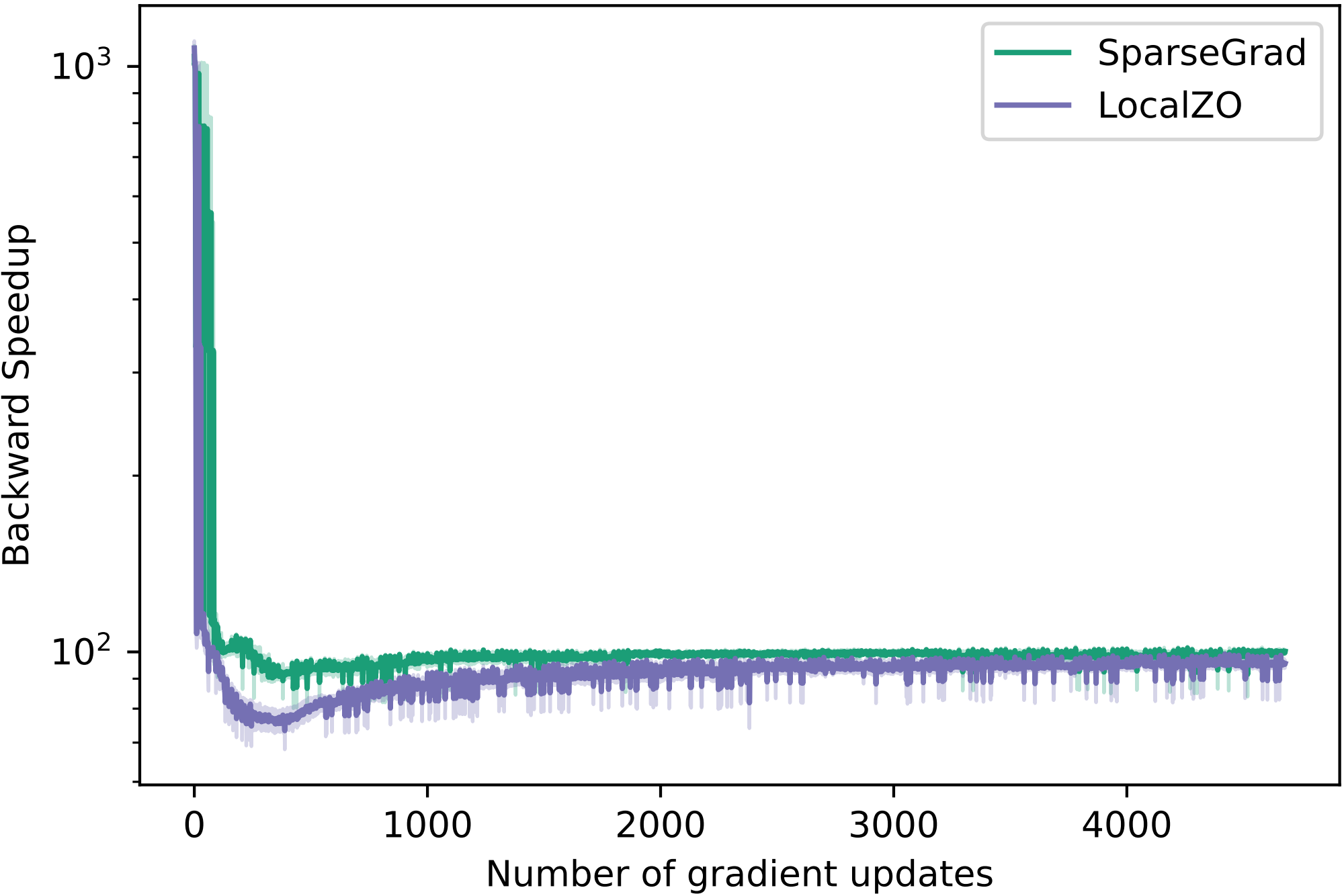
SparseGrad
LocalZO
10^3
10^2
Backward Speedup
0
1000
2000
3000
4000
Number of gradient updates