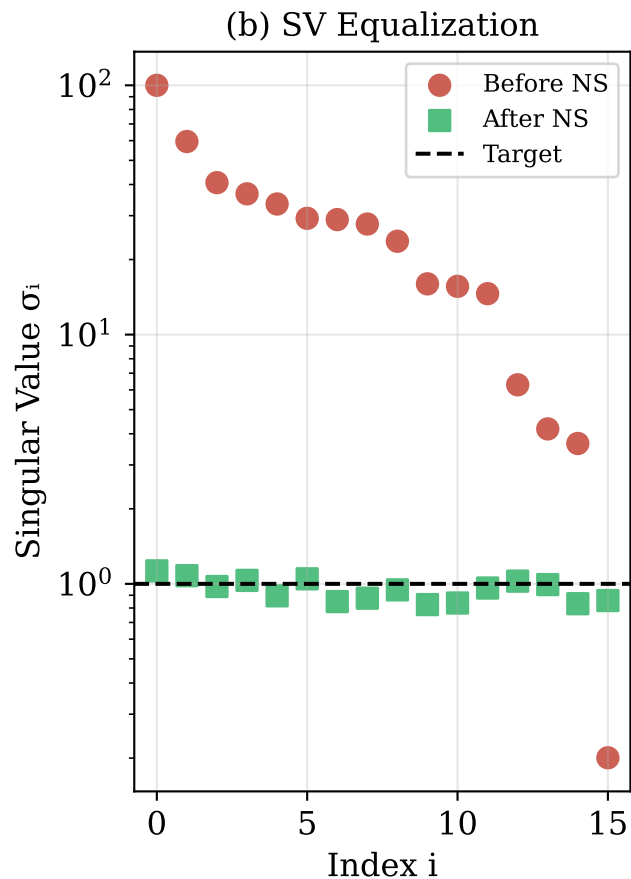
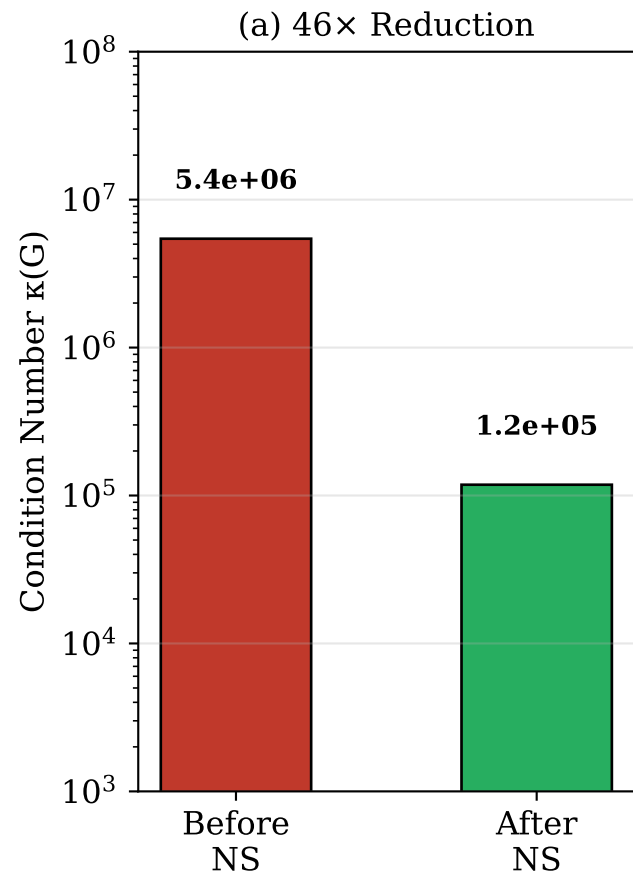


Newton-Schulz Orthogonalization in MuonLonghorn



(c) Benefits

Newton-Schulz Benefits:

1. Numerical Stability
 κ : 5443515 \rightarrow 118303
(46× improvement)
2. Gradient Flow
 - Prevents exploding gradients
 - Stable training dynamics
3. Memory Efficiency
 - Applied in constant memory
 - 0(1) iterations (typically 3)