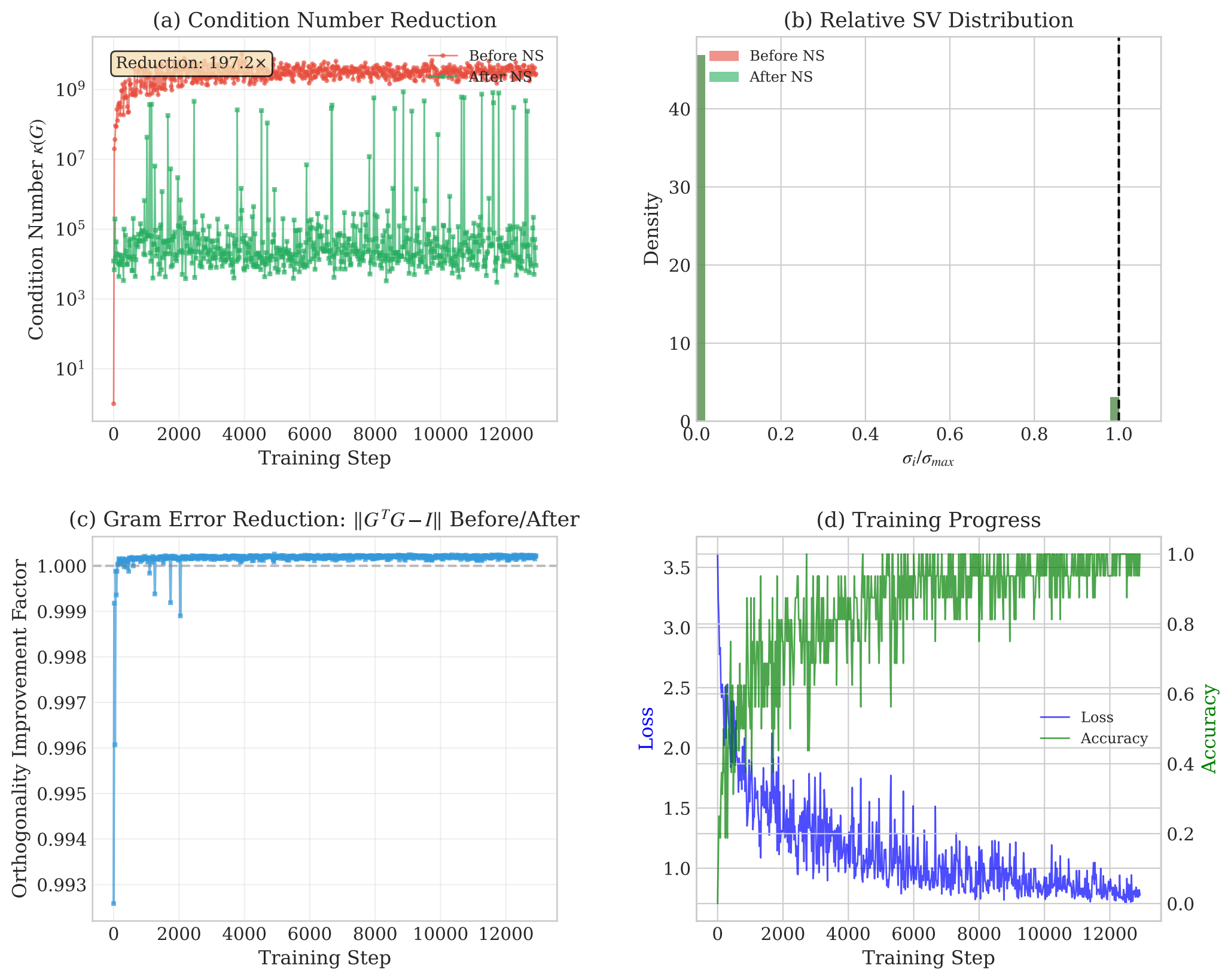
(a) Condition Number Reduction
Reduction: 197.2×
Before NS
After NS
Condition Number $\kappa(G)$
10^9
10^7
10^5
10^3
10^1
0
2000
4000
6000
8000
10000
12000
Training Step
(b) Relative SV Distribution
Before NS
After NS
Density
40
30
20
10
0
0.0
0.2
0.4
0.6
0.8
1.0
σ_i/σ_{max}
(c) Gram Error Reduction: $\|G^TG-I\|$ Before/After
Orthogonality Improvement Factor
1.000
0.999
0.998
0.997
0.996
0.995
0.994
0.993
0
2000
4000
6000
8000
10000
12000
Training Step
(d) Training Progress
Loss
Accuracy
3.5
3.0
2.5
2.0
1.5
1.0
1.0
0.8
0.6
0.4
0.2
0.0
0
2000
4000
6000
8000
10000
12000
Training Step