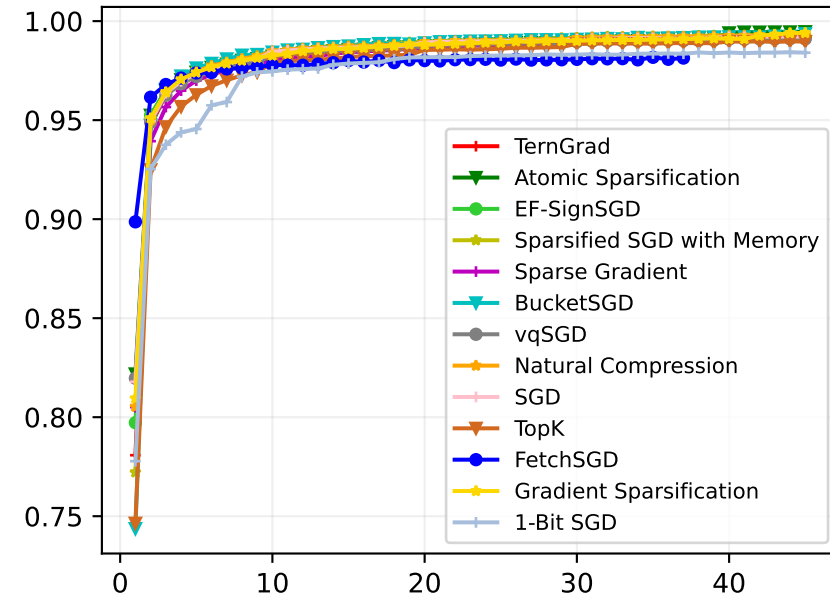
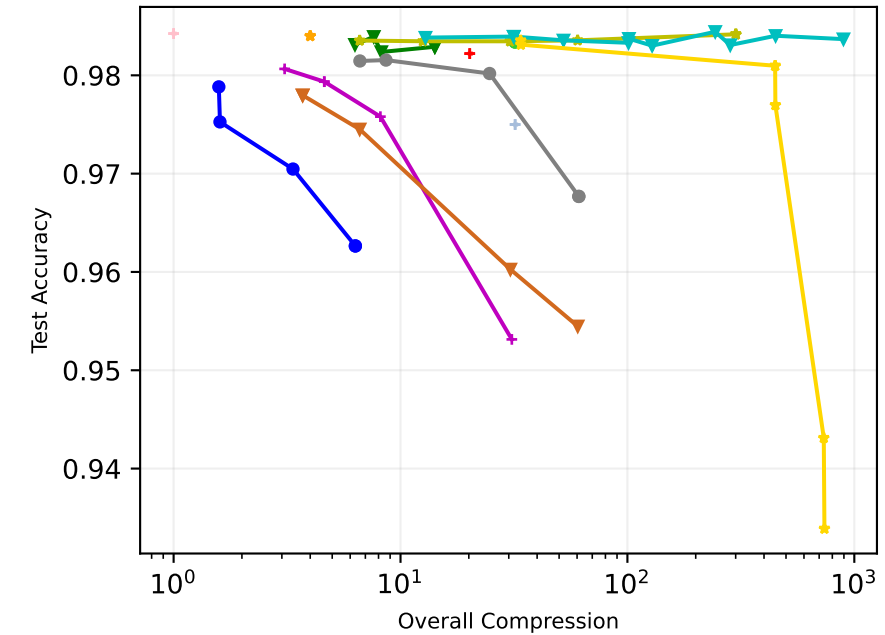
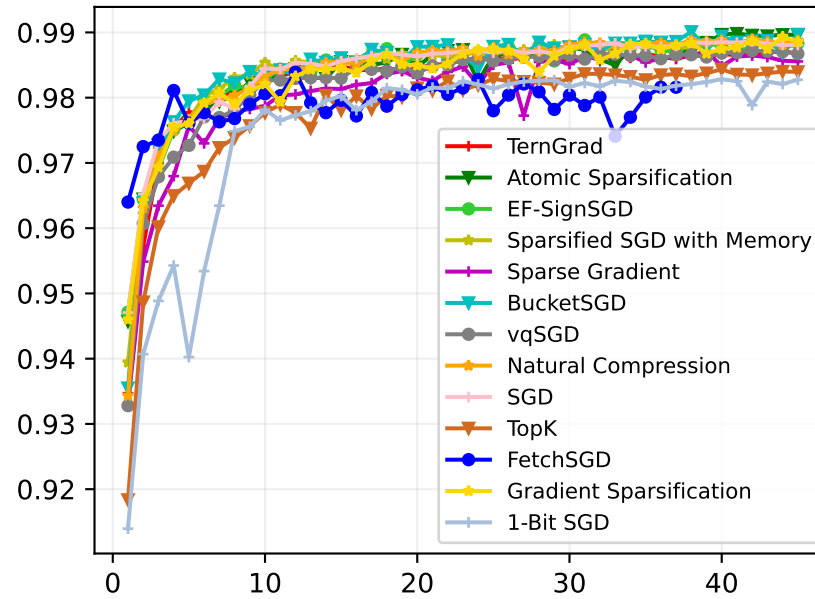


BASELINE

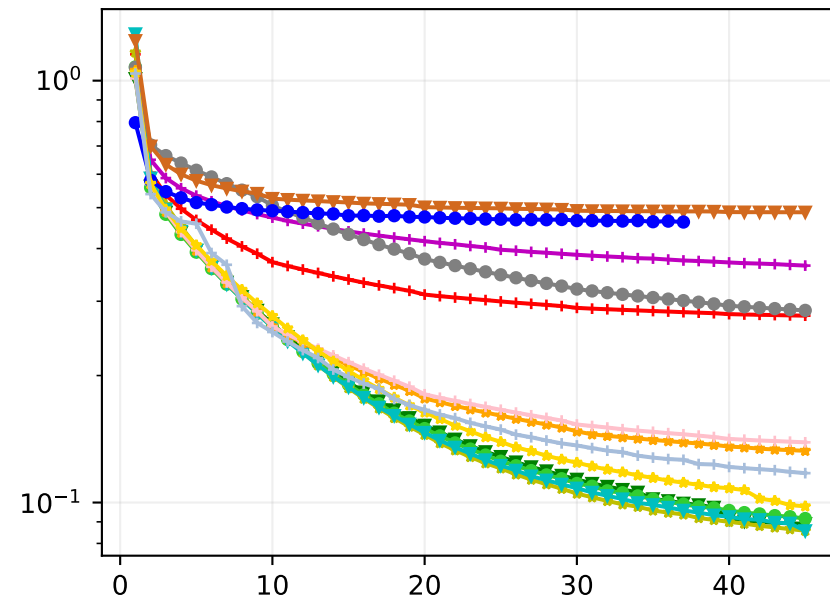
Training Accuracy



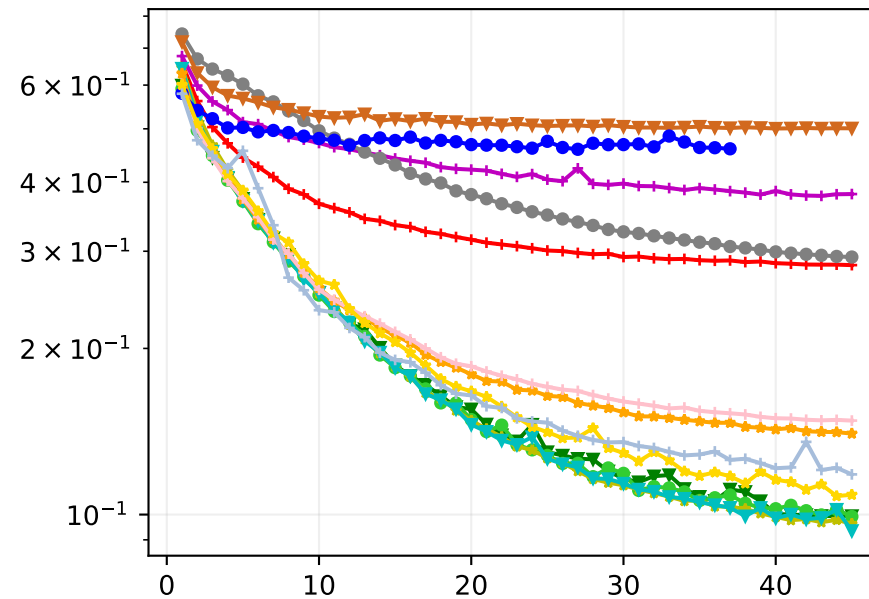
Test Accuracy



Training Loss



Test Loss



| Method | Mean Test Acc | Compression Ratio |
|----------------------------|---------------|-------------------|
| BucketSGD | 98.443 | 243.76 |
| SGD | 98.425 | 1.0 |
| Sparsified SGD with Memory | 98.417 | 300.47 |
| Natural Compression | 98.4 | 4.0 |
| Atomic Sparsification | 98.392 | 7.63 |
| Gradient Sparsification | 98.355 | 33.86 |
| EF-SignSGD | 98.338 | 32.0 |
| TernGrad | 98.221 | 20.19 |
| vqSGD | 98.155 | 8.63 |
| Sparse Gradient | 98.065 | 3.08 |
| FetchSGD | 97.883 | 1.58 |
| TopK | 97.797 | 3.7 |
| 1-Bit SGD | 97.5 | 32.0 |