Training Loss over Epochs

- With Newton-Schulz
- Without Newton-Schulz
- Adaptive Newton-Schulz