# Notes for Lottery Tickets and Neural Networks

## Timings for Ghost Clipping

I timed the training time spent on a single batch by different implementations of the differentially private sgd algorithm.

Below are the results for a simple conv net that consists of 2 fully connected layers and 2 convolutional layers on the MNIST dataset. The batch size used is 120. The timing is the average over 500 batches.

- functorch_dp, 0.008851111078984104
- opacus_dp, 0.007145165540161542
- mixed ghost clipping, 0.010015537165221758
- ghost clipping, 0.010277516969712451
- not mixed not ghost, 0.010565361546818168
- public training, 0.0040898429183289405

4 convolutional layers on Cifar10, batch size = 100, seconds:

- functorch_dp, 0.014373262761160731
- opacus_dp, 0.008994201084831729
- mixed ghost clipping, 0.014257113210158422
- ghost clipping, 0.01808279032376595
- not mixed not ghost, use papers authors' privacy engine: 0.014138890799833461
- public, 0.00498583750706166