CIFAR10 Tuning Results

Jerry Chee

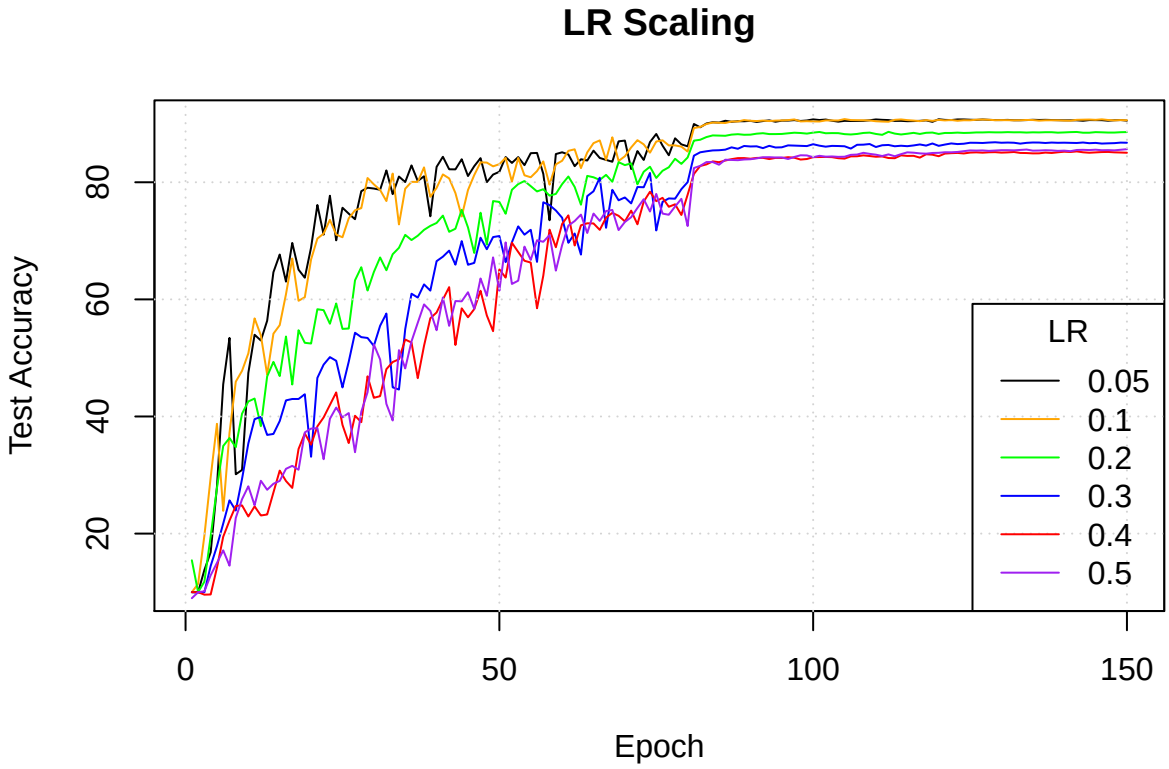
7/17/2019

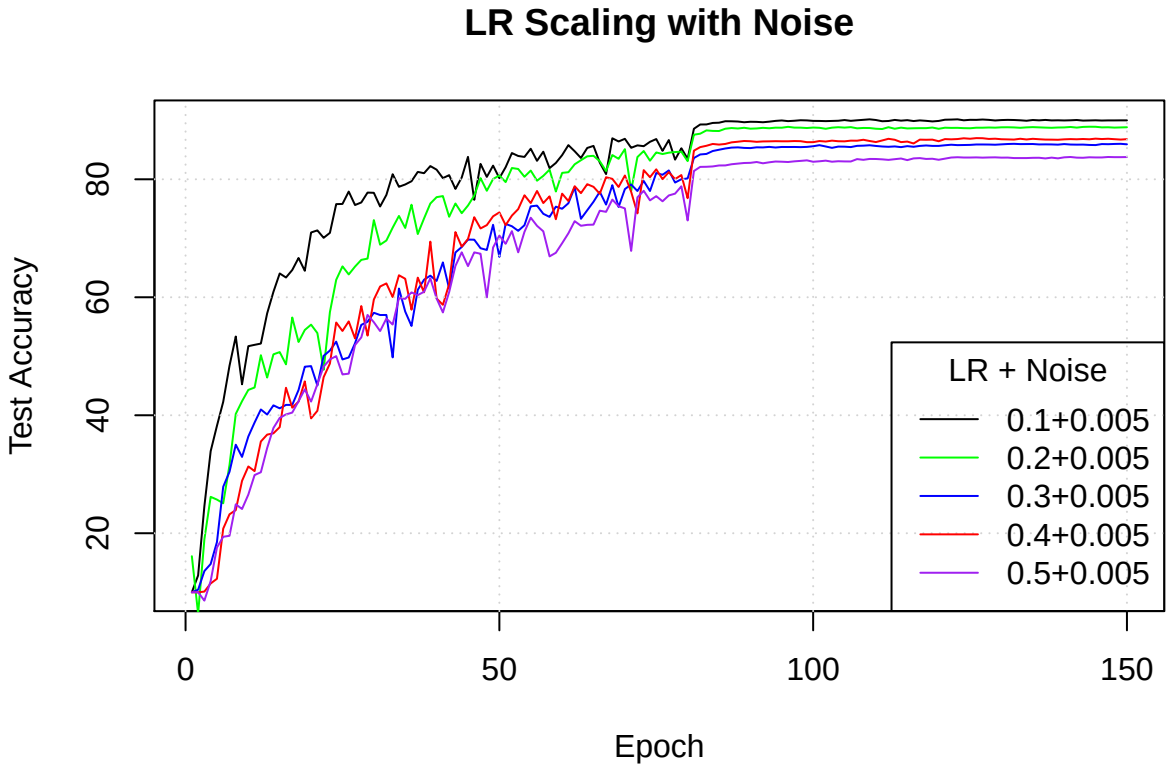
Tuning experiments for ResNet18 on CIFAR10. Adding mean zero gaussian noise to the gradient. Extensive experiments conducted on batch size 8192, 4096, 2048. Plots only shared for 8192, to see the rest of the experimental data see `checkpoint/` folder. Extensive tuning proved unsuccessful in improving performance on test with added gradient noise.

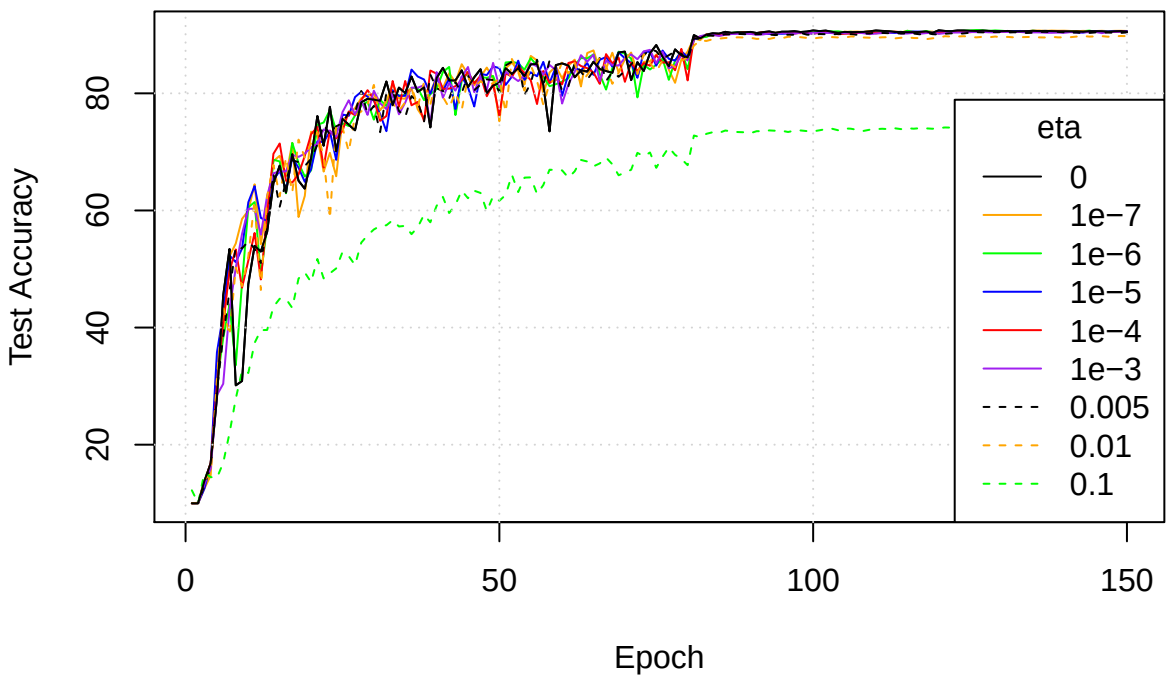
Batch Size 8192

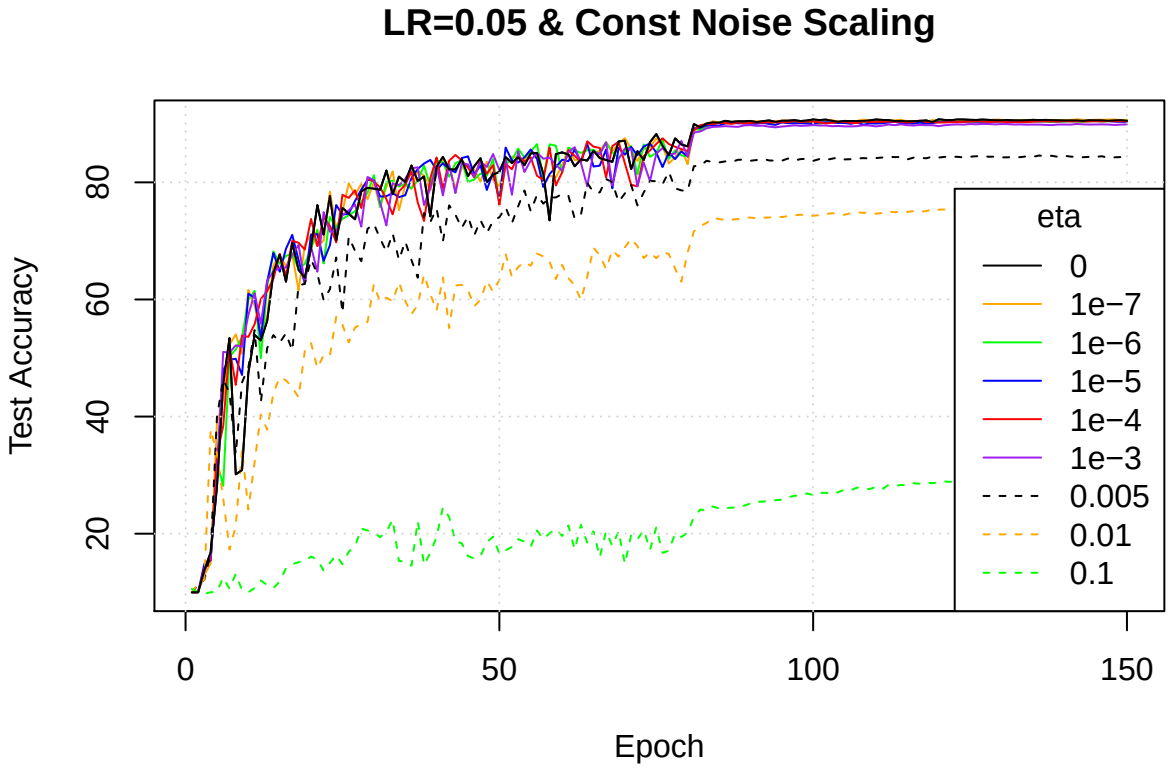
Some observations: so far the most promising this is that adding noise prevents the decay in performance as the learning rate increases. At learning rate 0.2 and above, adding noise improves the generalization performance. But, learning rate 0.1 is still best. Again, I am not sure why the performance decreases with increasing the learning rate. A starting LR=0.1 is very common for small batch sizes, as I understand.

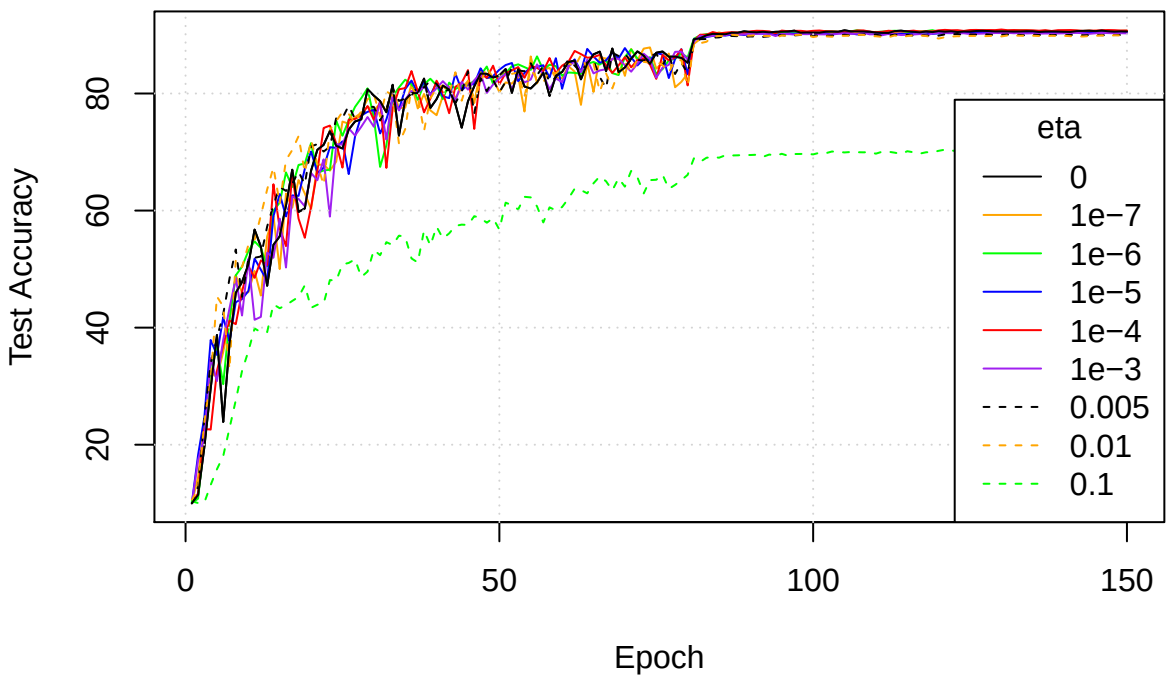
Smaller noise does better, larger noise does worse. A decreasing noise does better than a constant noise

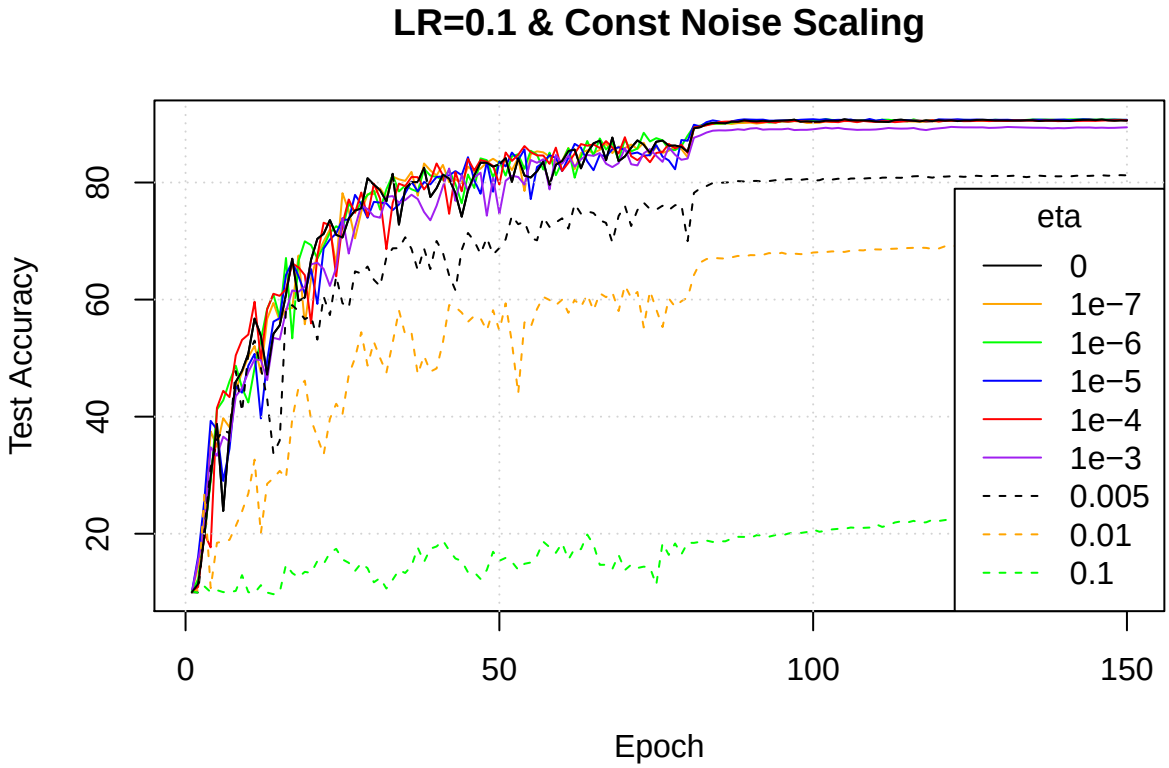


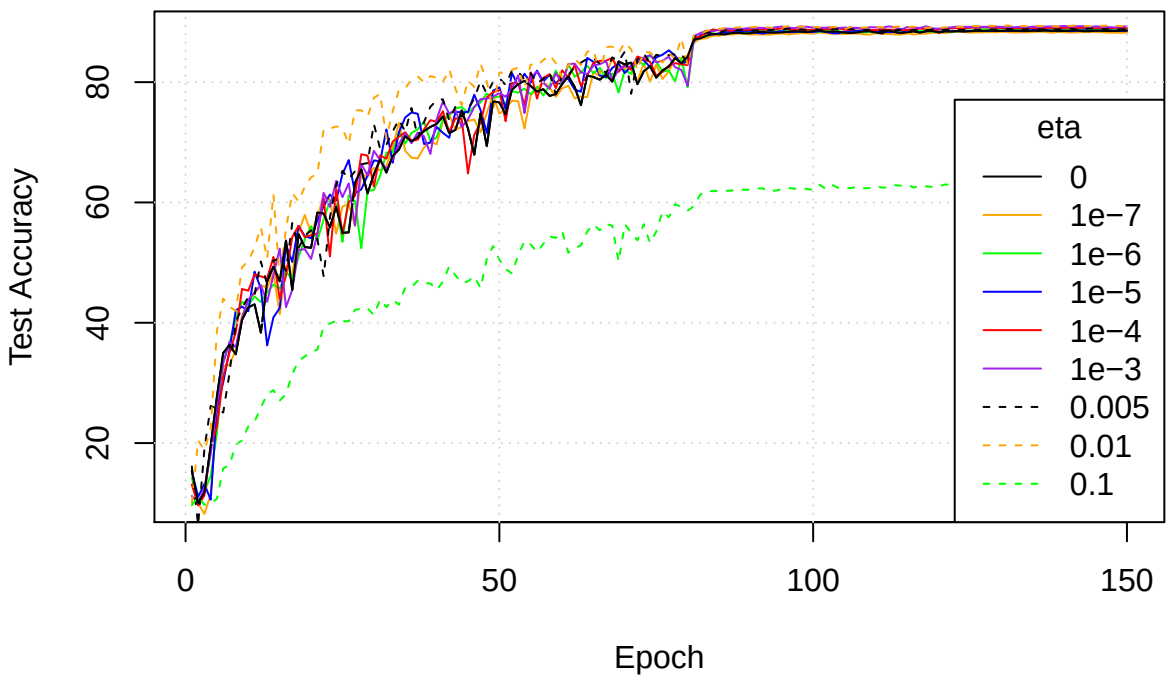

LR Scaling with Noise
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
LR + Noise
0.1+0.005
0.2+0.005
0.3+0.005
0.4+0.005
0.5+0.005

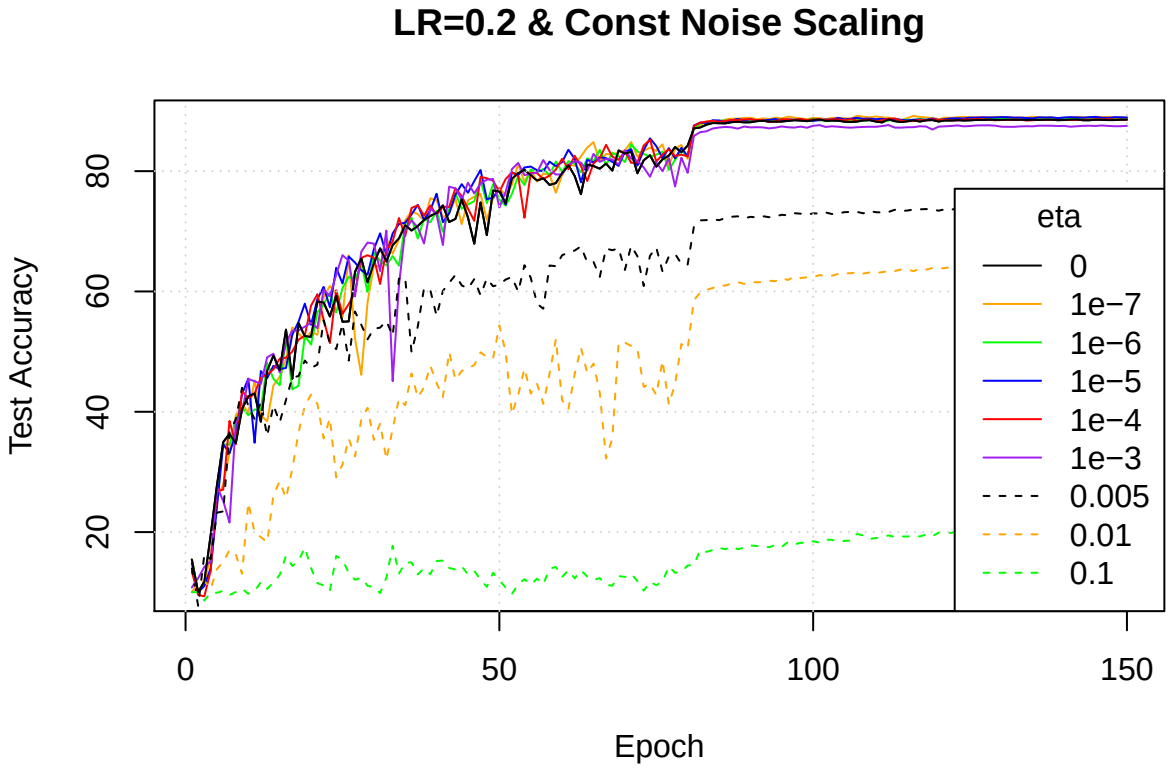

LR=0.05 & Decr Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1

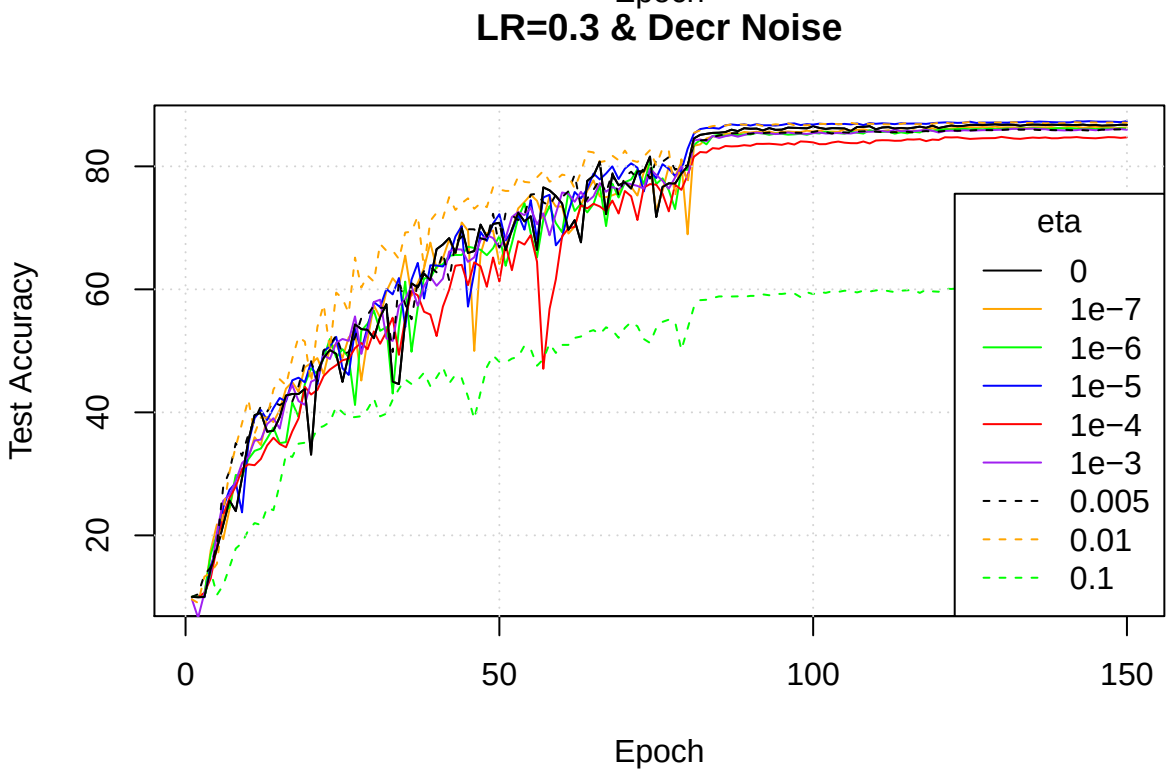

LR=0.05 & Const Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1

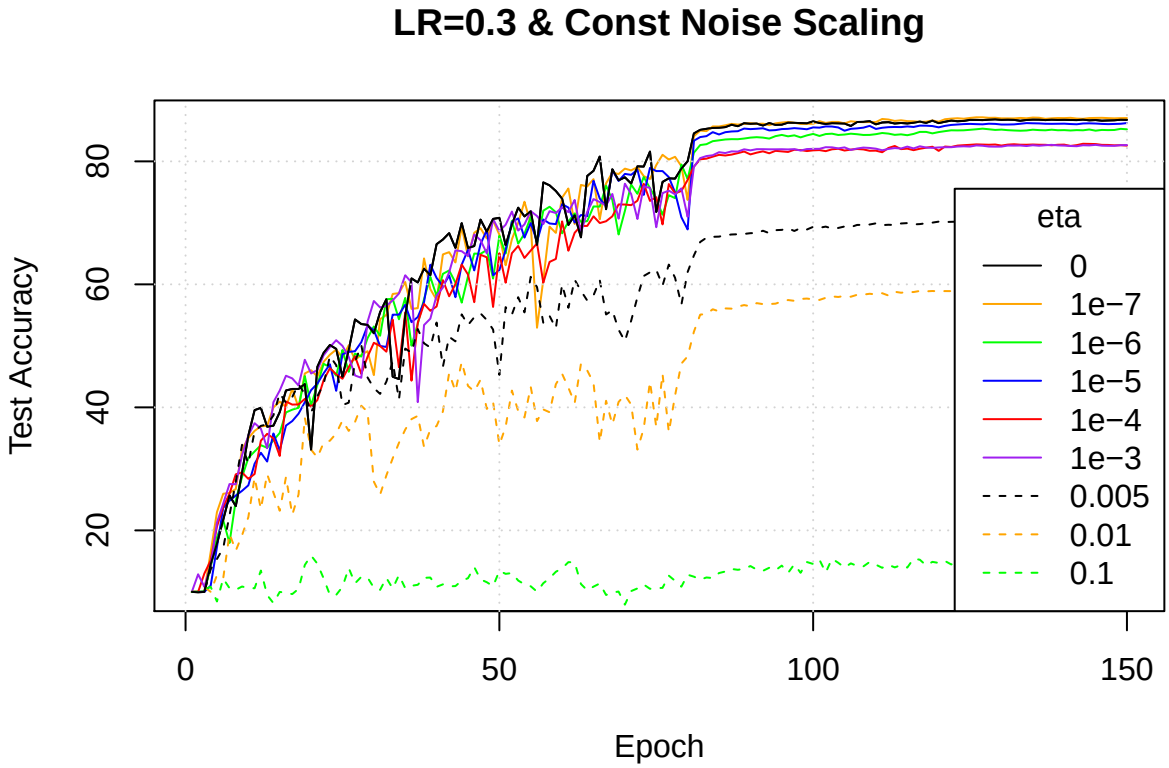

LR=0.1 & Decr Noise
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1

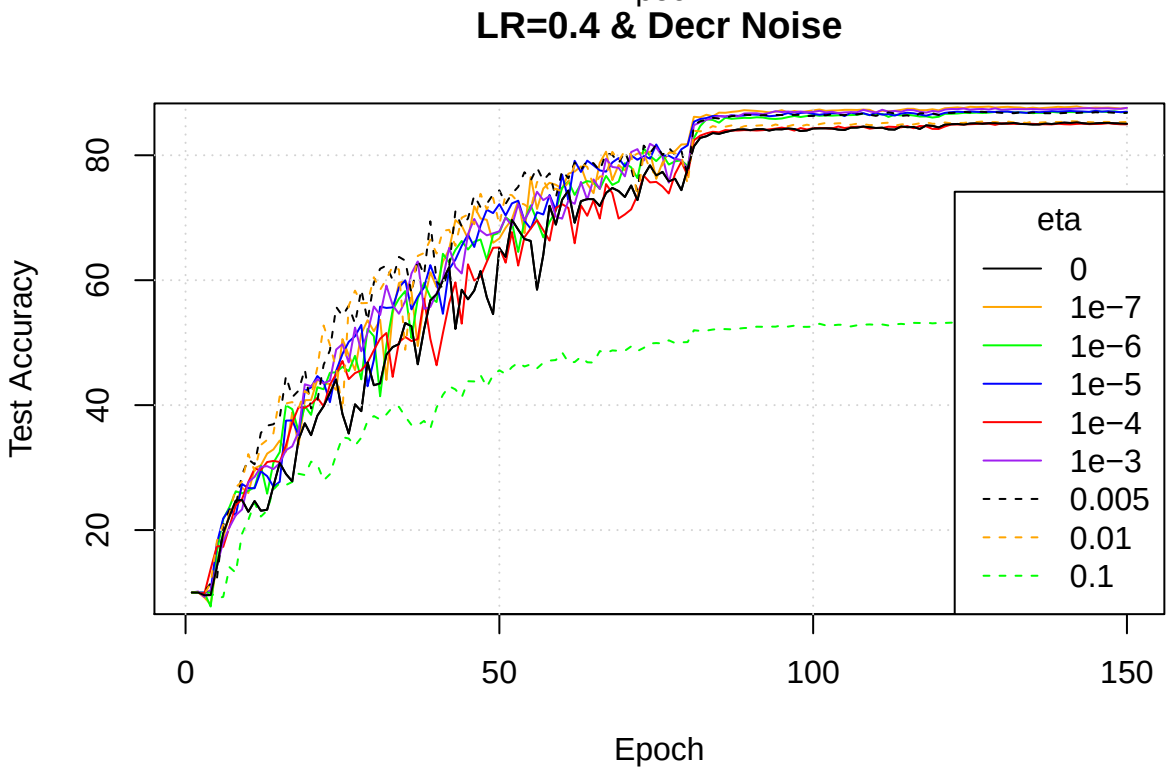

LR=0.1 & Const Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1

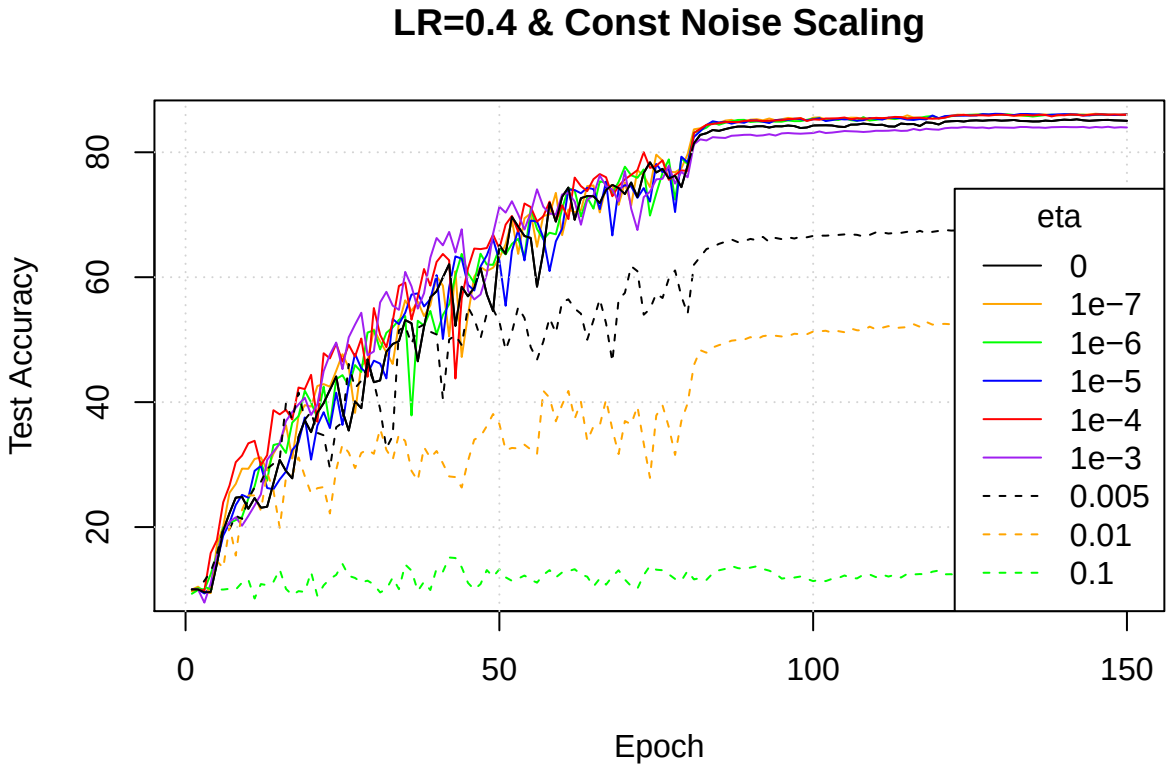

LR=0.2 & Decr Noise
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1

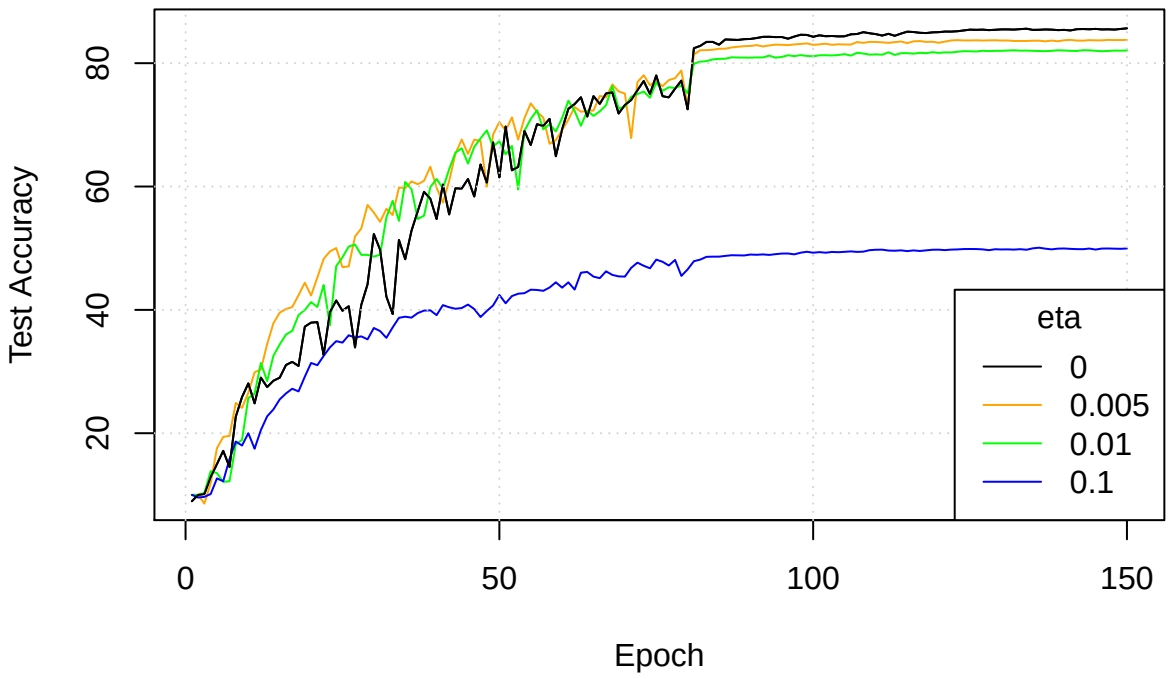

LR=0.2 & Const Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1


LR=0.3 & Decr Noise
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1


LR=0.3 & Const Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1


LR=0.4 & Decr Noise
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1


LR=0.4 & Const Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
1e−7
1e−6
1e−5
1e−4
1e−3
0.005
0.01
0.1


LR=0.5 & Decr Noise Scaling
Test Accuracy
20
40
60
80
0
50
100
150
Epoch
eta
0
0.005
0.01
0.1