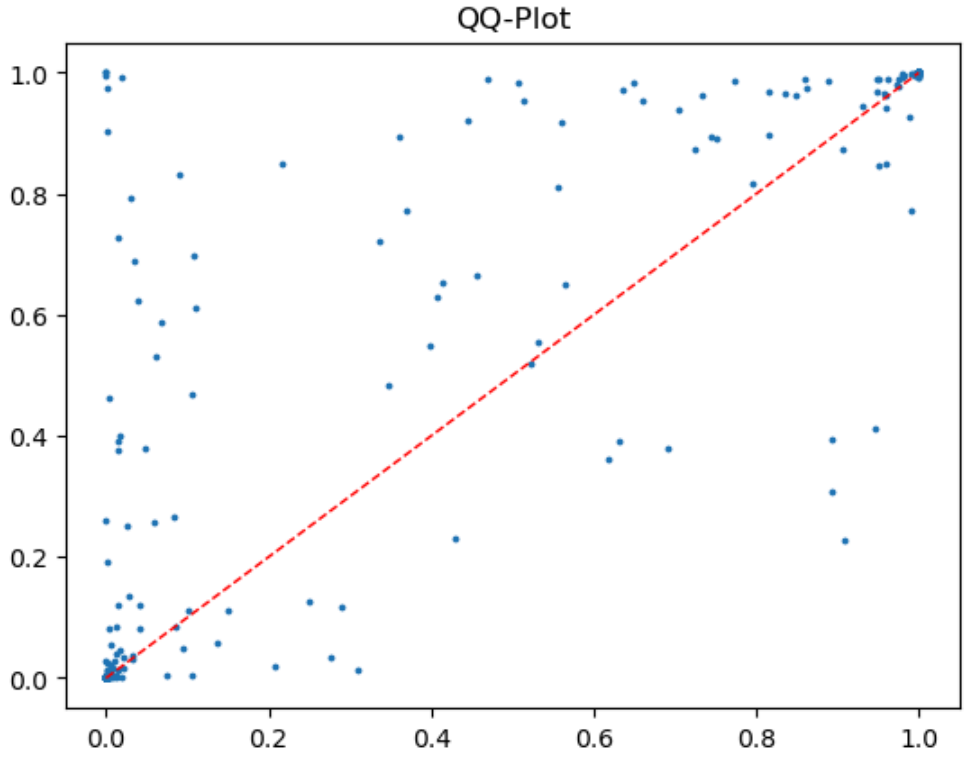


In [ ]:

lr=0.01

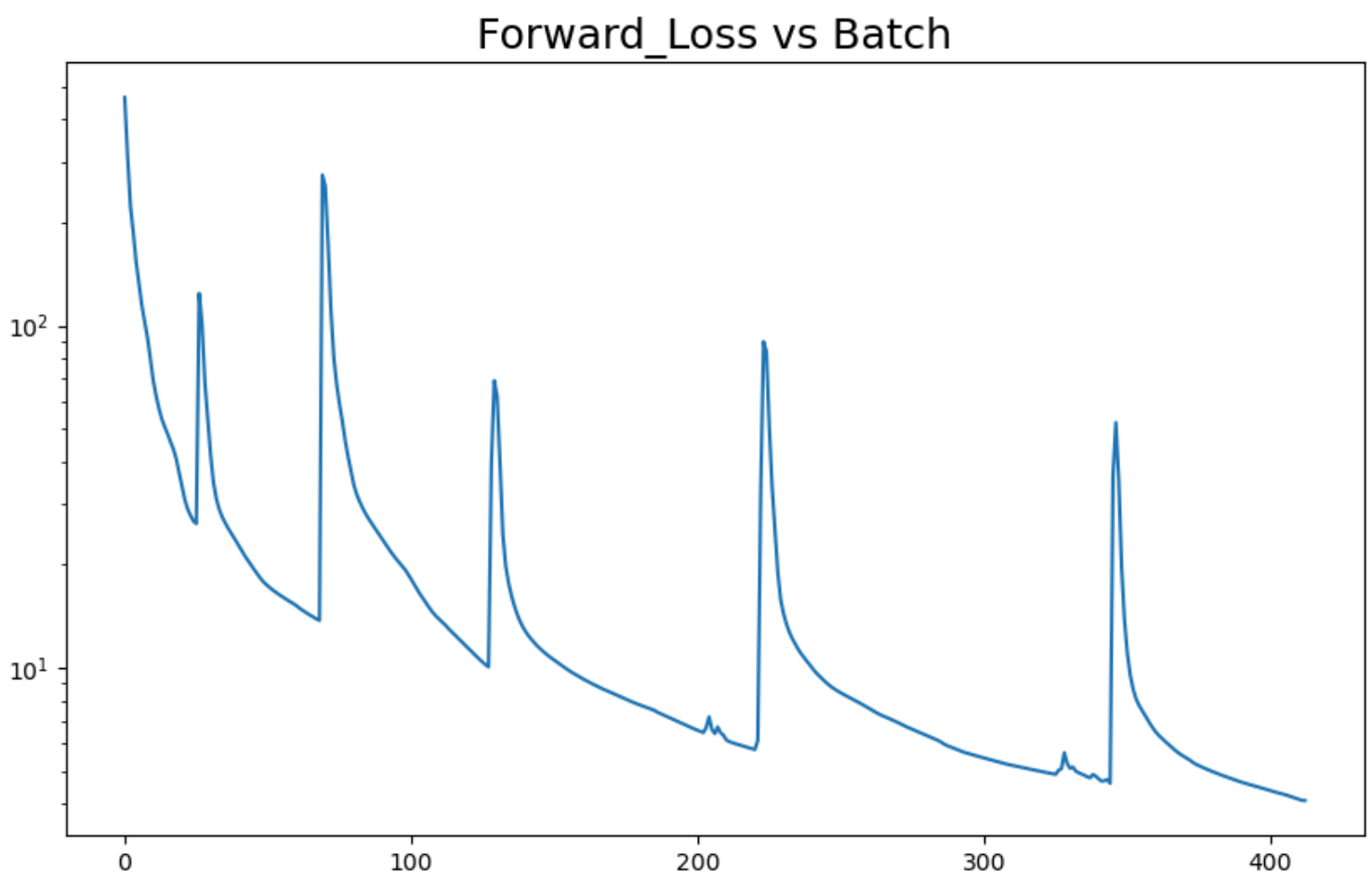
target: sigma=0.03

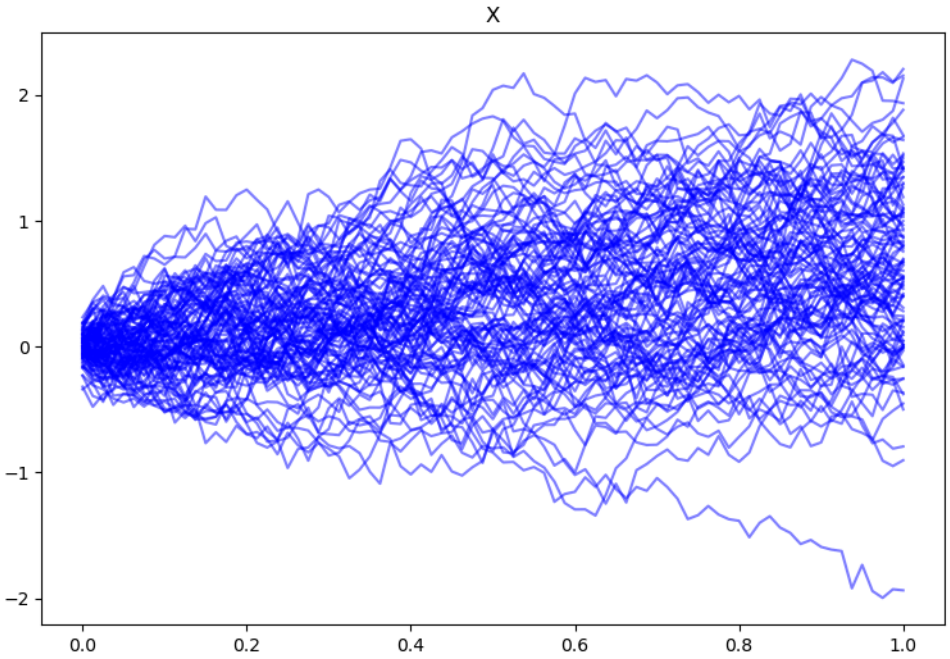
optimizer: AdamW(weight_decay=5e-4)

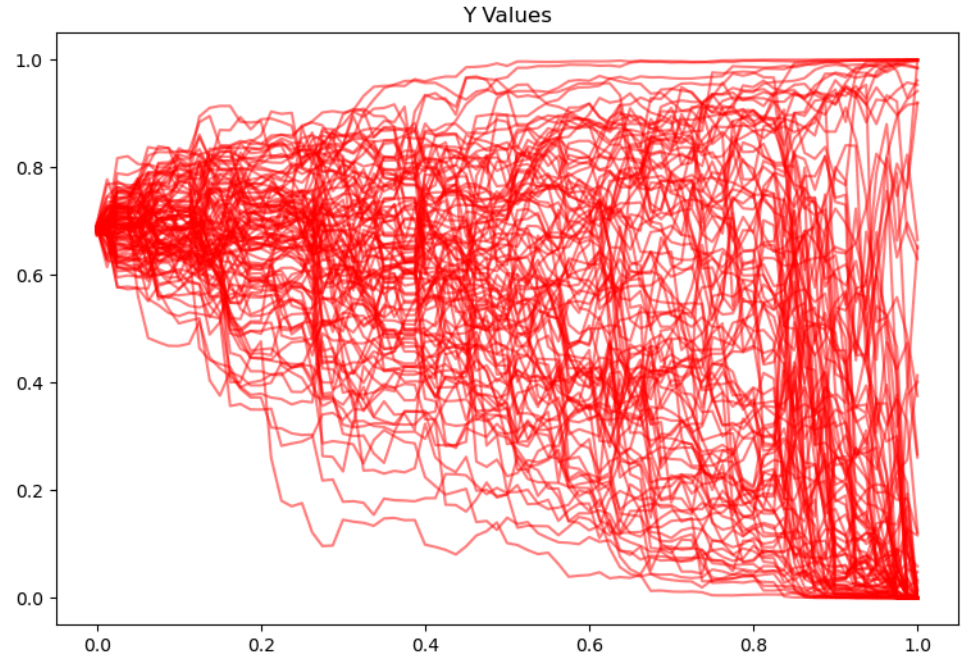
scheduler: StepLR(step=50, gamma=0.995)

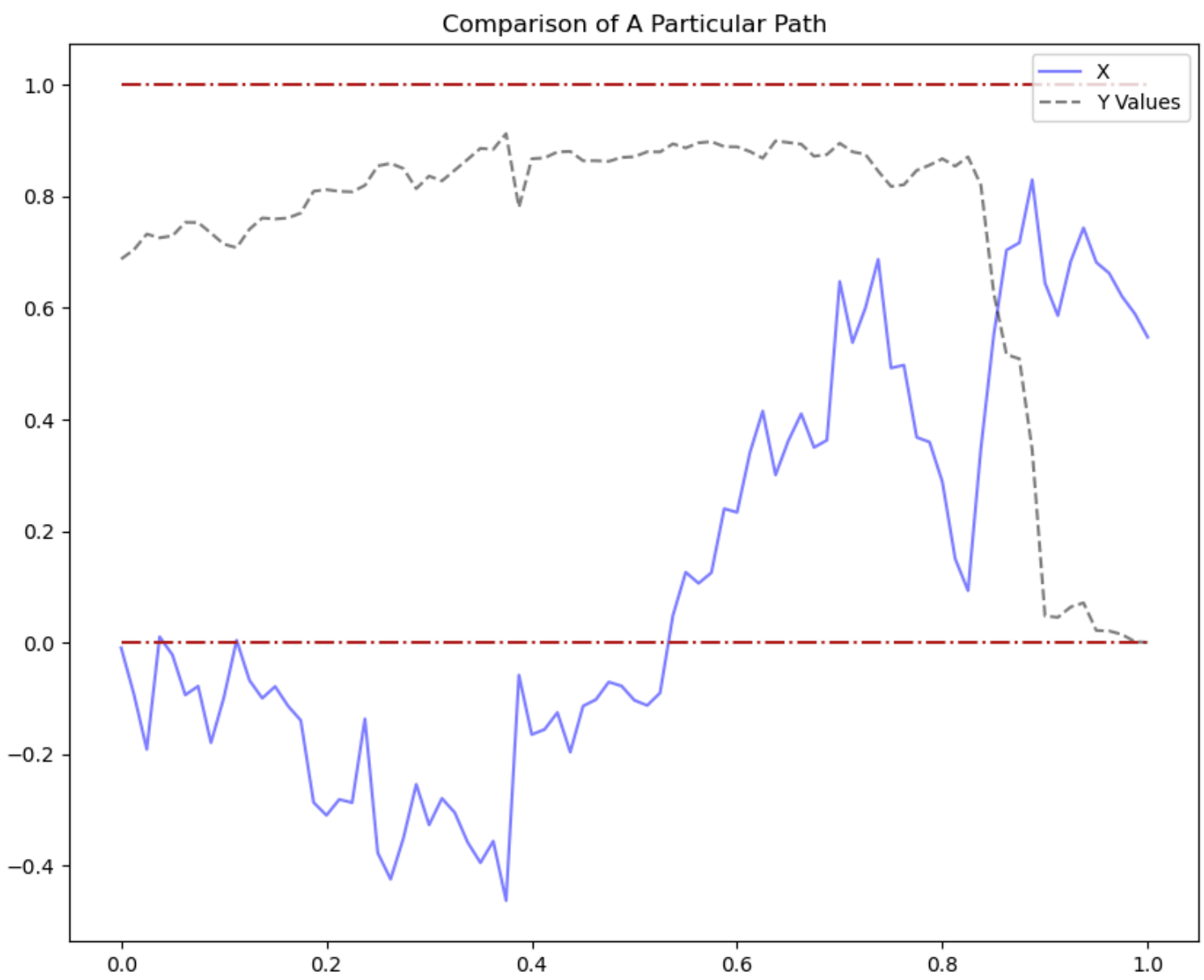
MaxBatch=500

OptimStep=20


Forward_Loss vs Batch
10^2
10^1
0
100
200
300
400


X
2
1
0
−1
−2
0.0
0.2
0.4
0.6
0.8
1.0


Y Values
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


Comparison of A Particular Path
X
Y Values
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0