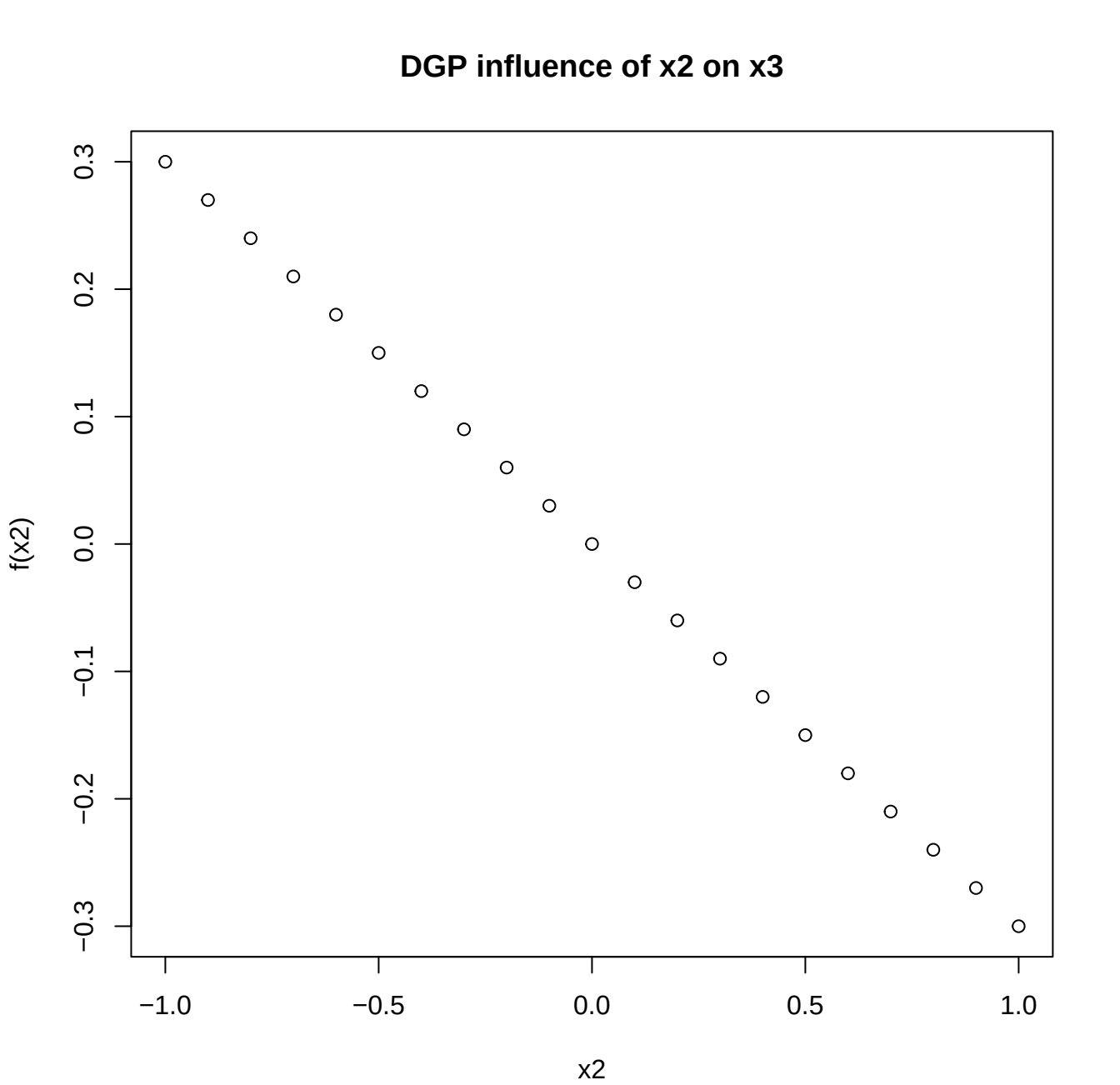
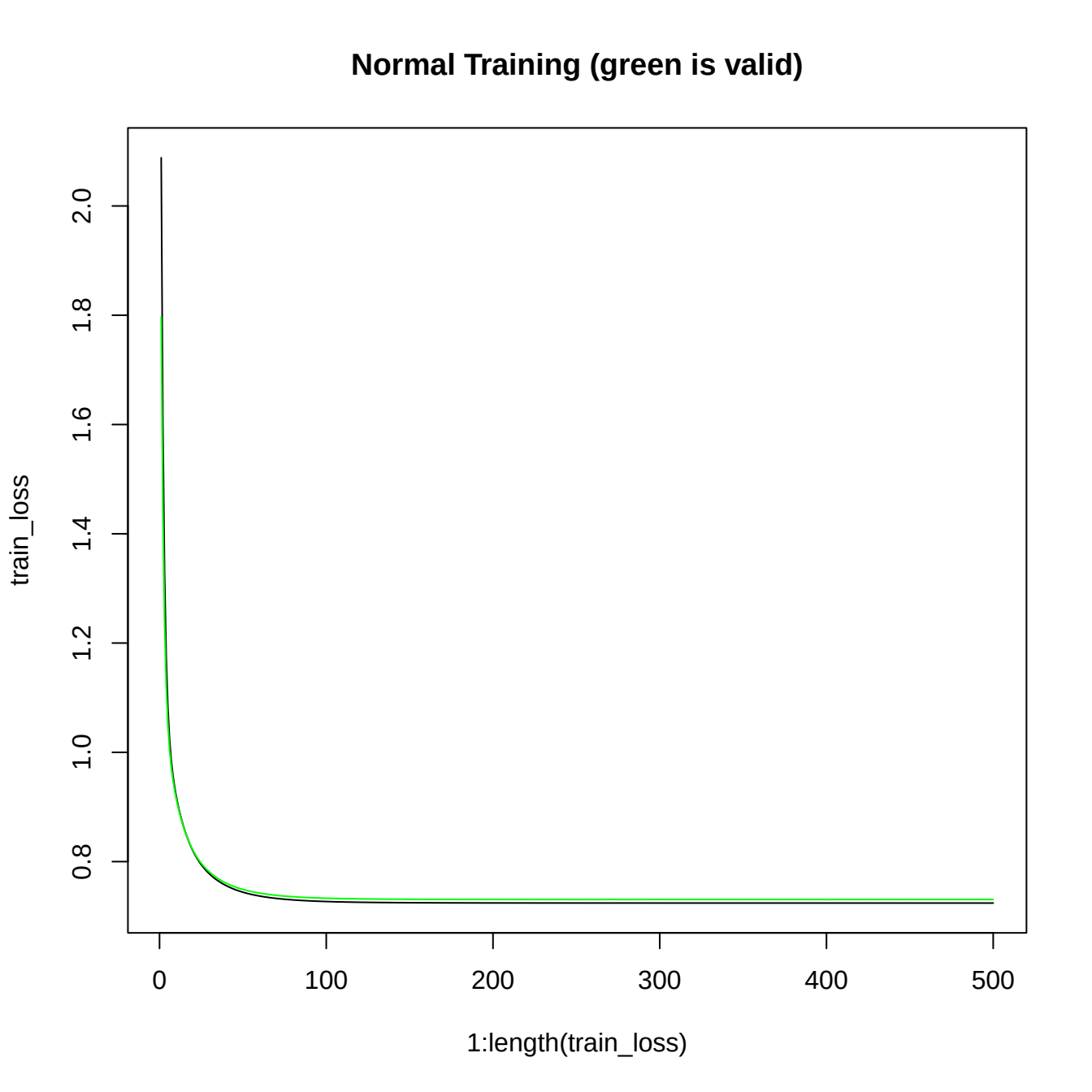
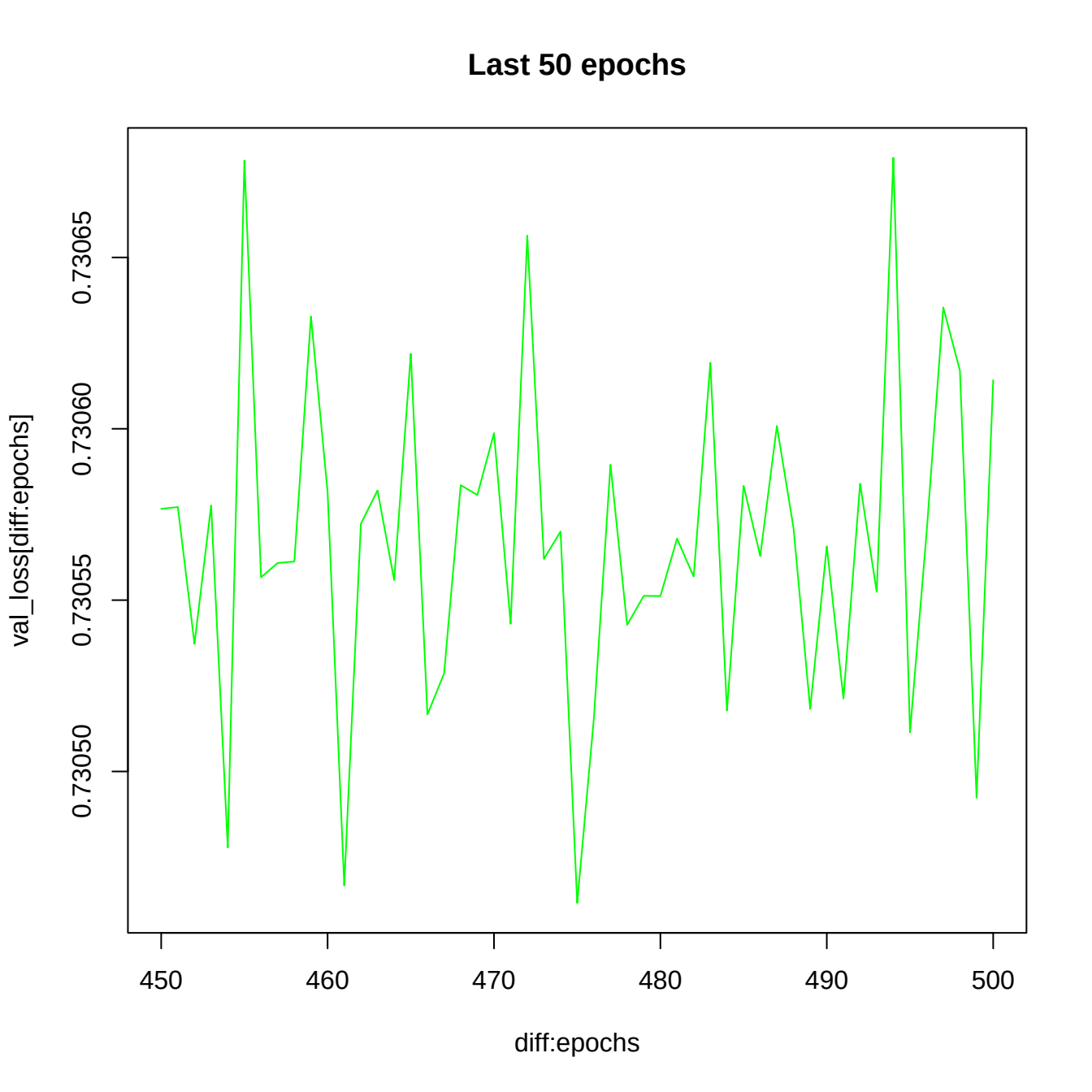
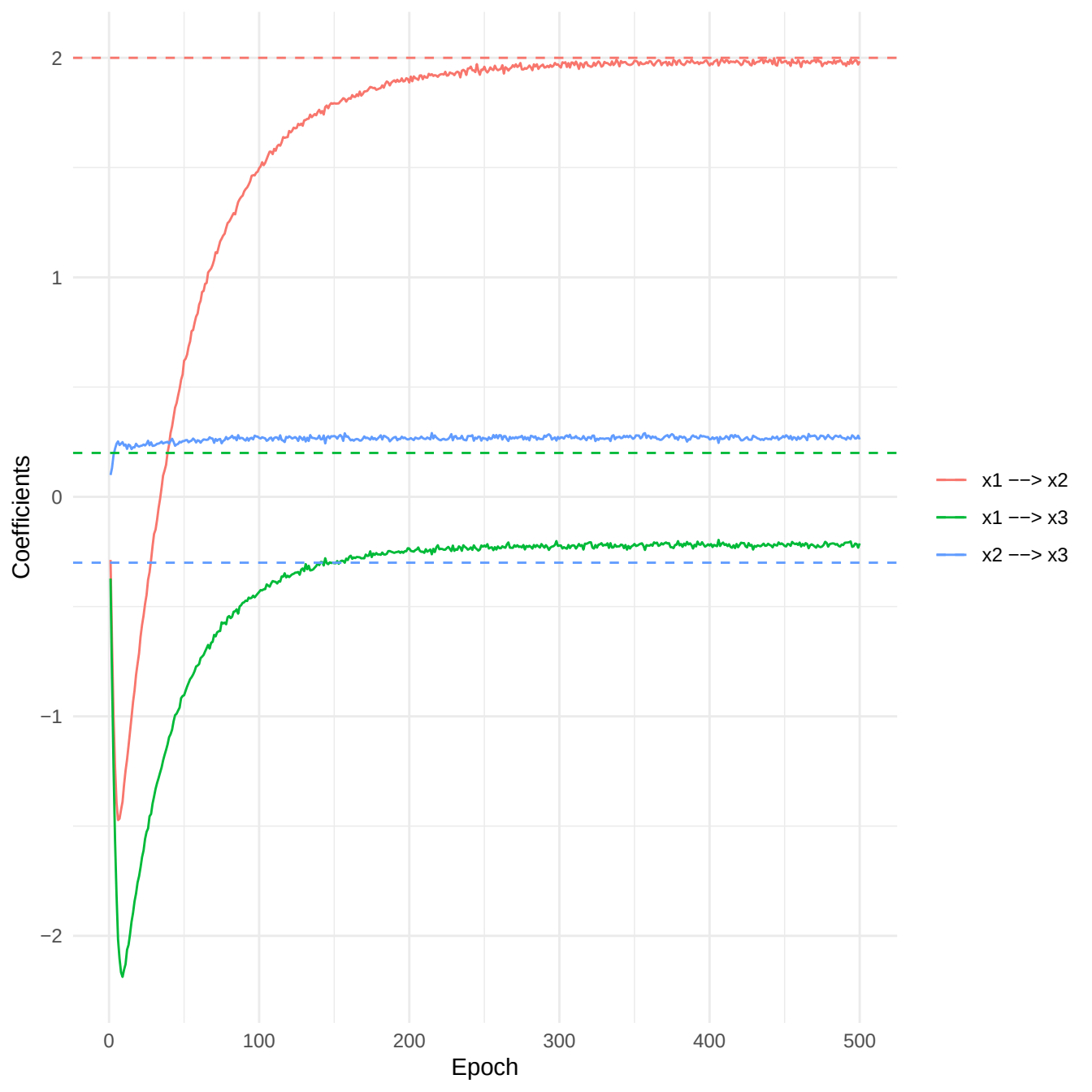
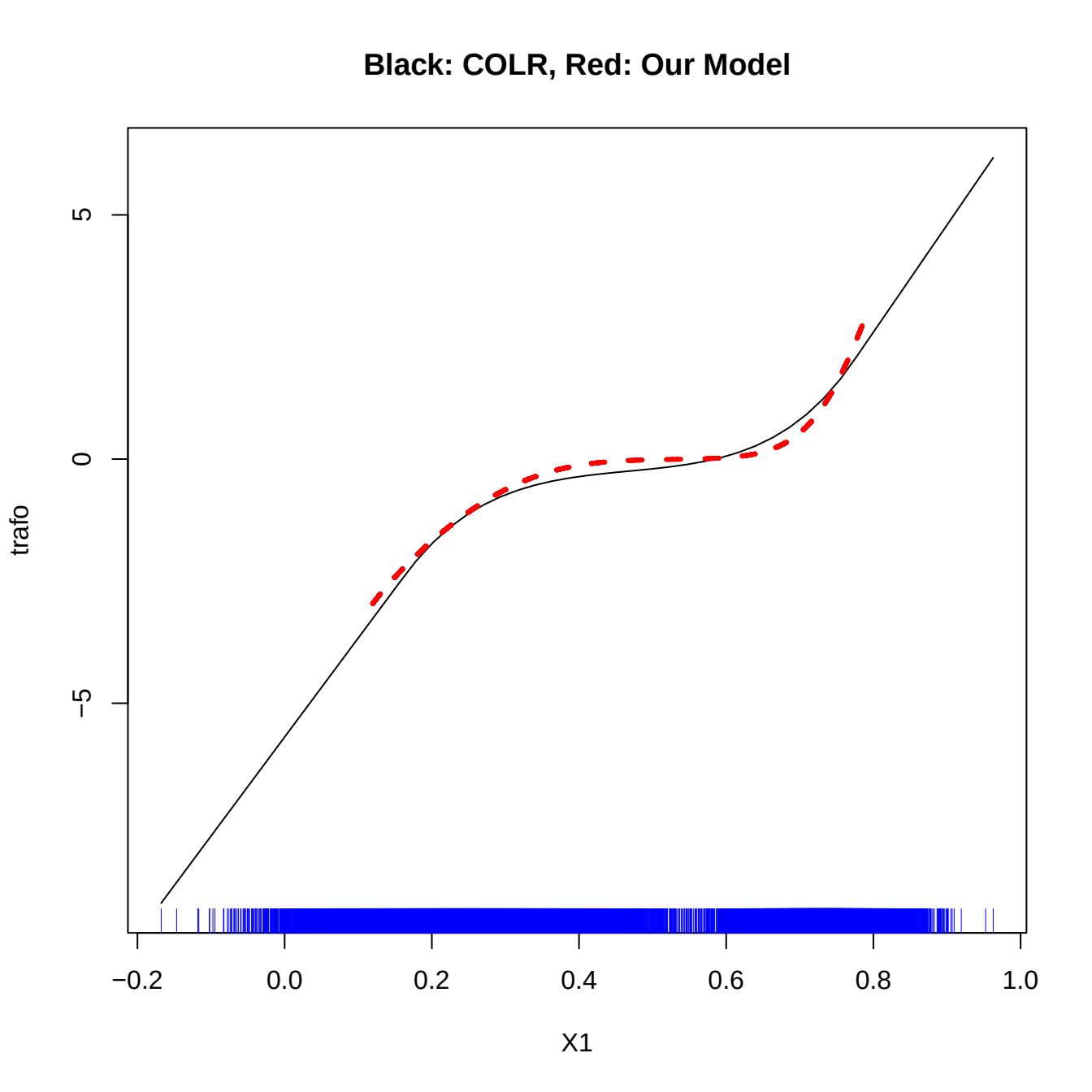


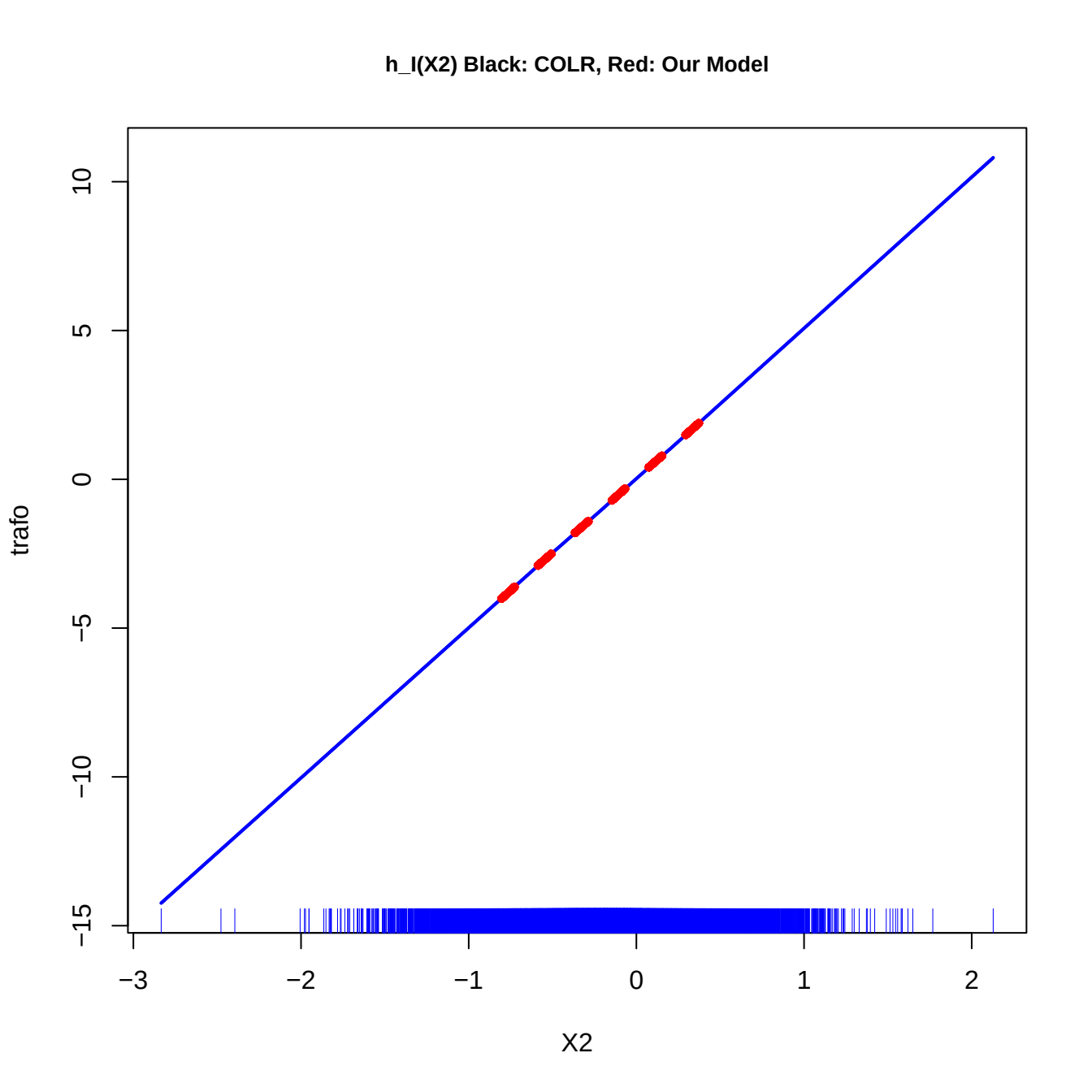
summerof24/runs/triangle_structured_continous/run_nodes25/triangle_mixed_DPGLinear_Mod

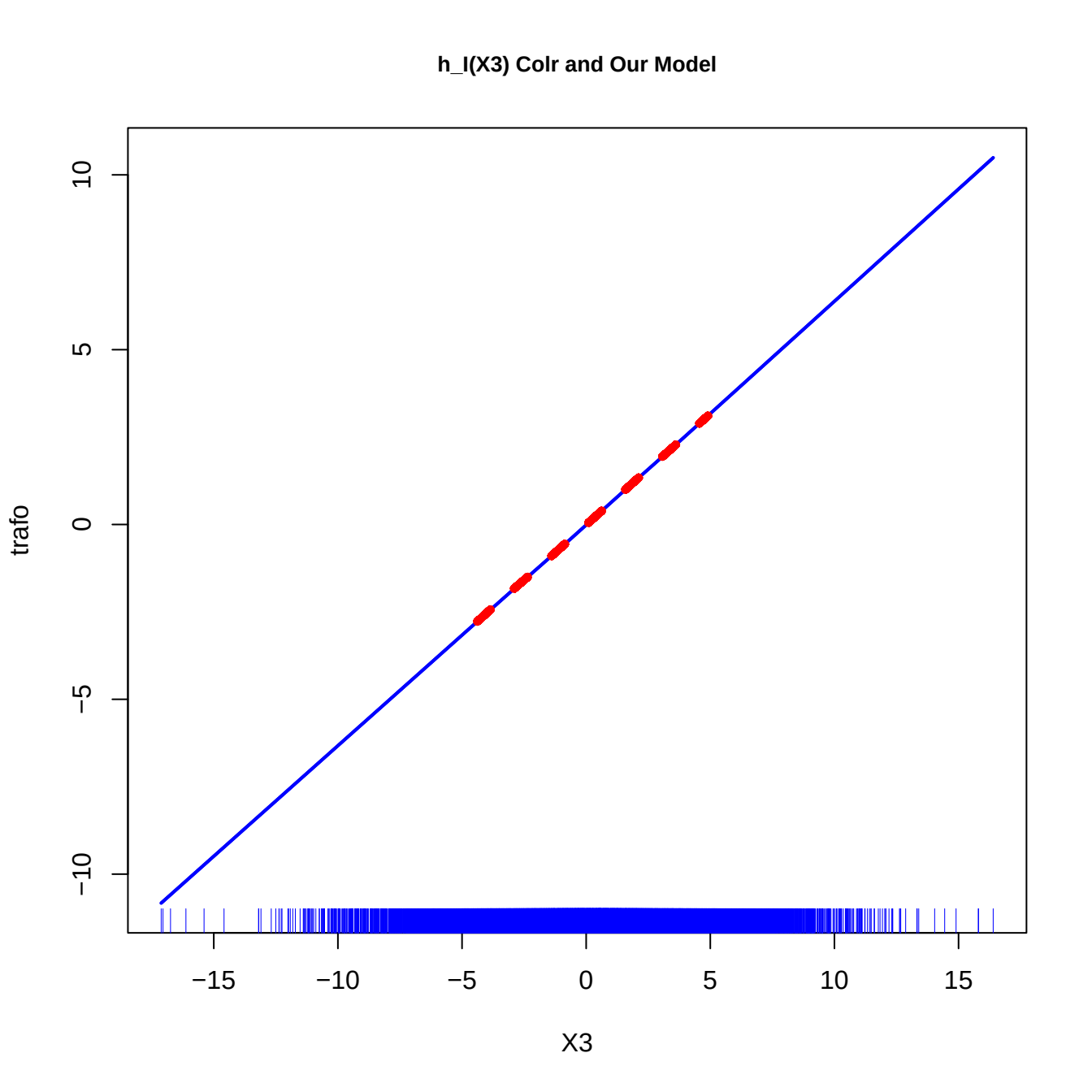

Normal Training (green is valid)
train_loss
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0
100
200
300
400
500
1:length(train_loss)

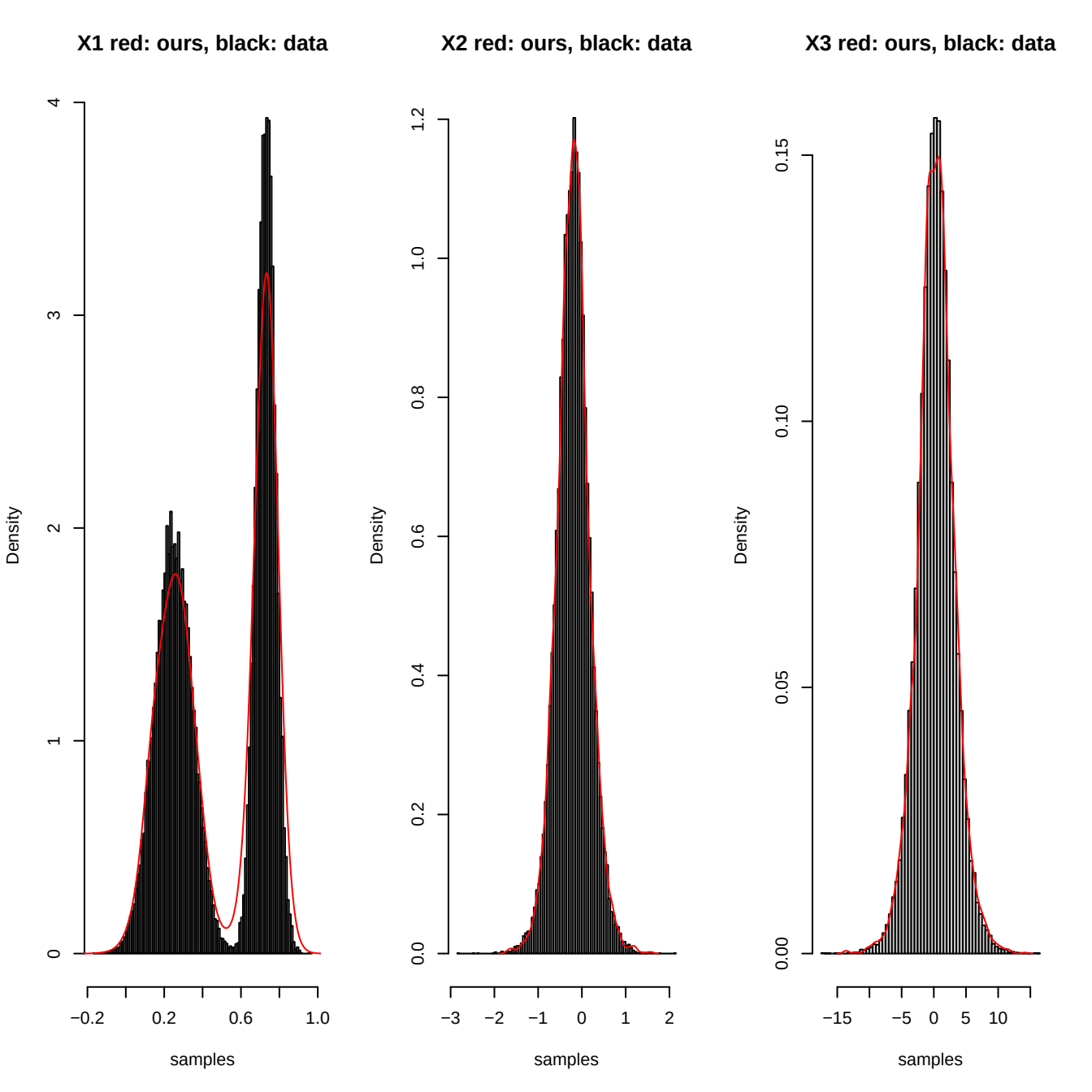

Last 50 epochs
val_loss[diff:epochs]
0.73050
0.73055
0.73060
0.73065
450
460
470
480
490
500
diff:epochs

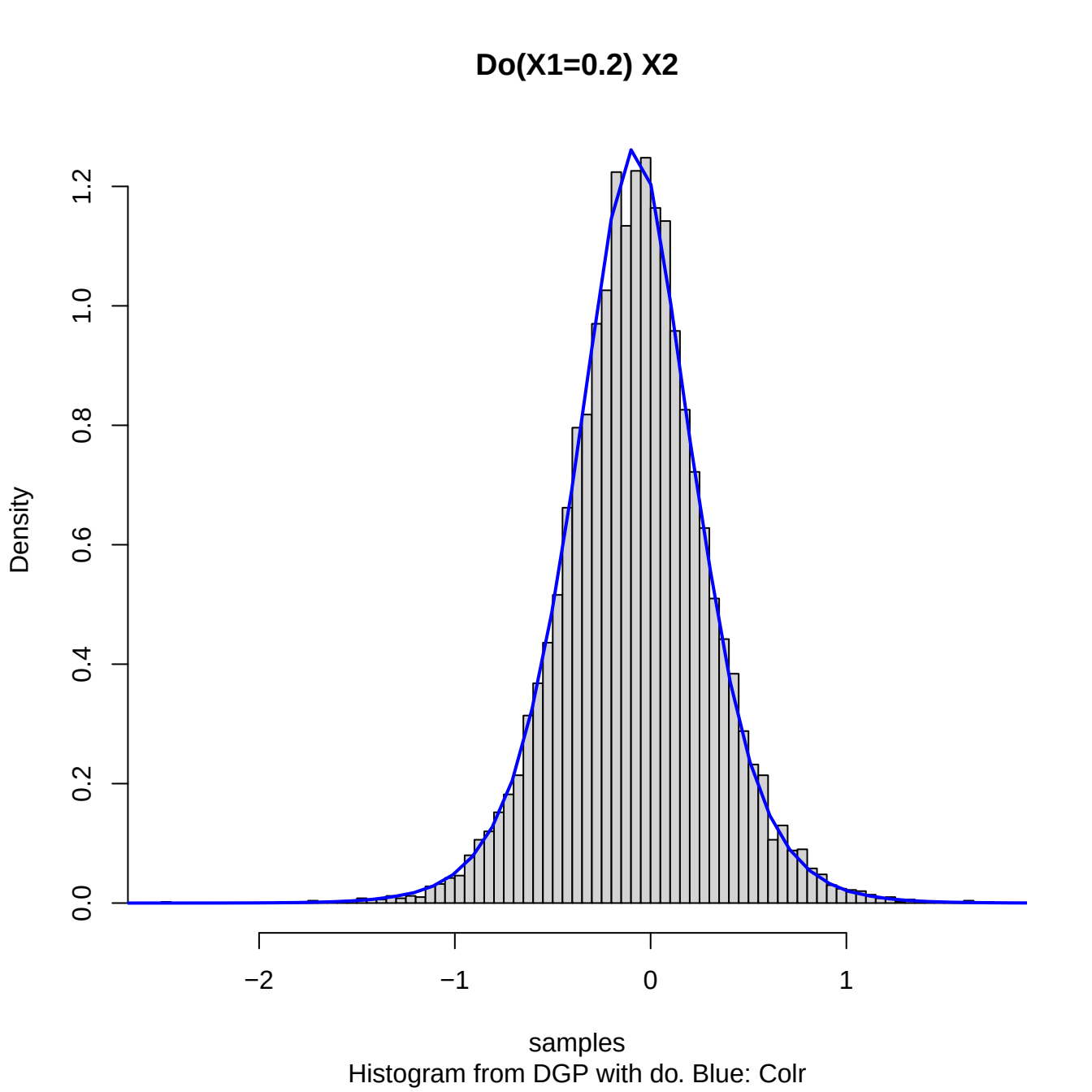

2
1
0
−1
−2
Coefficients
0
100
200
300
400
500
Epoch
x1 −−> x2
x1 −−> x3
x2 −−> x3

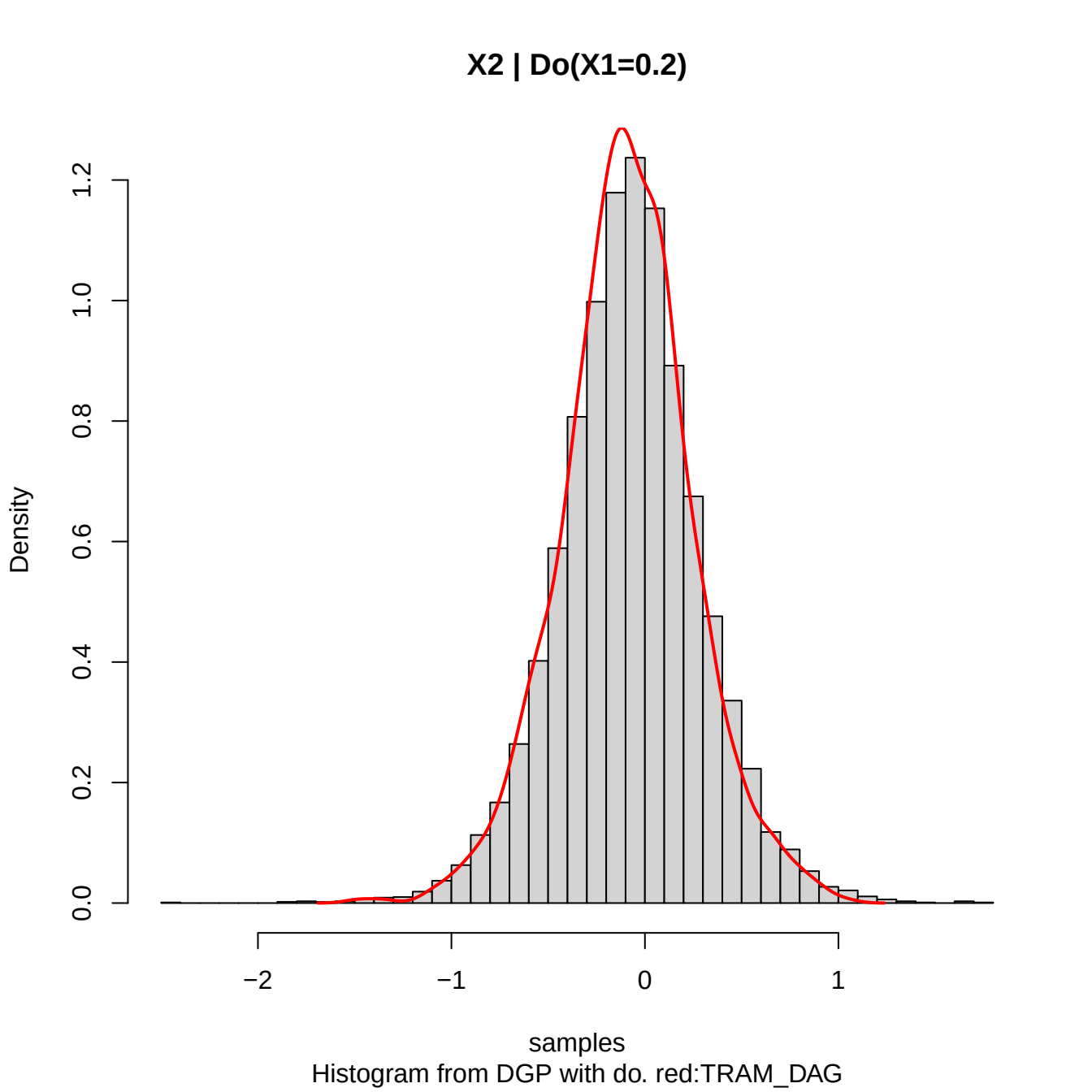

Black: COLR, Red: Our Model
trafo
5
0
−5
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
X1

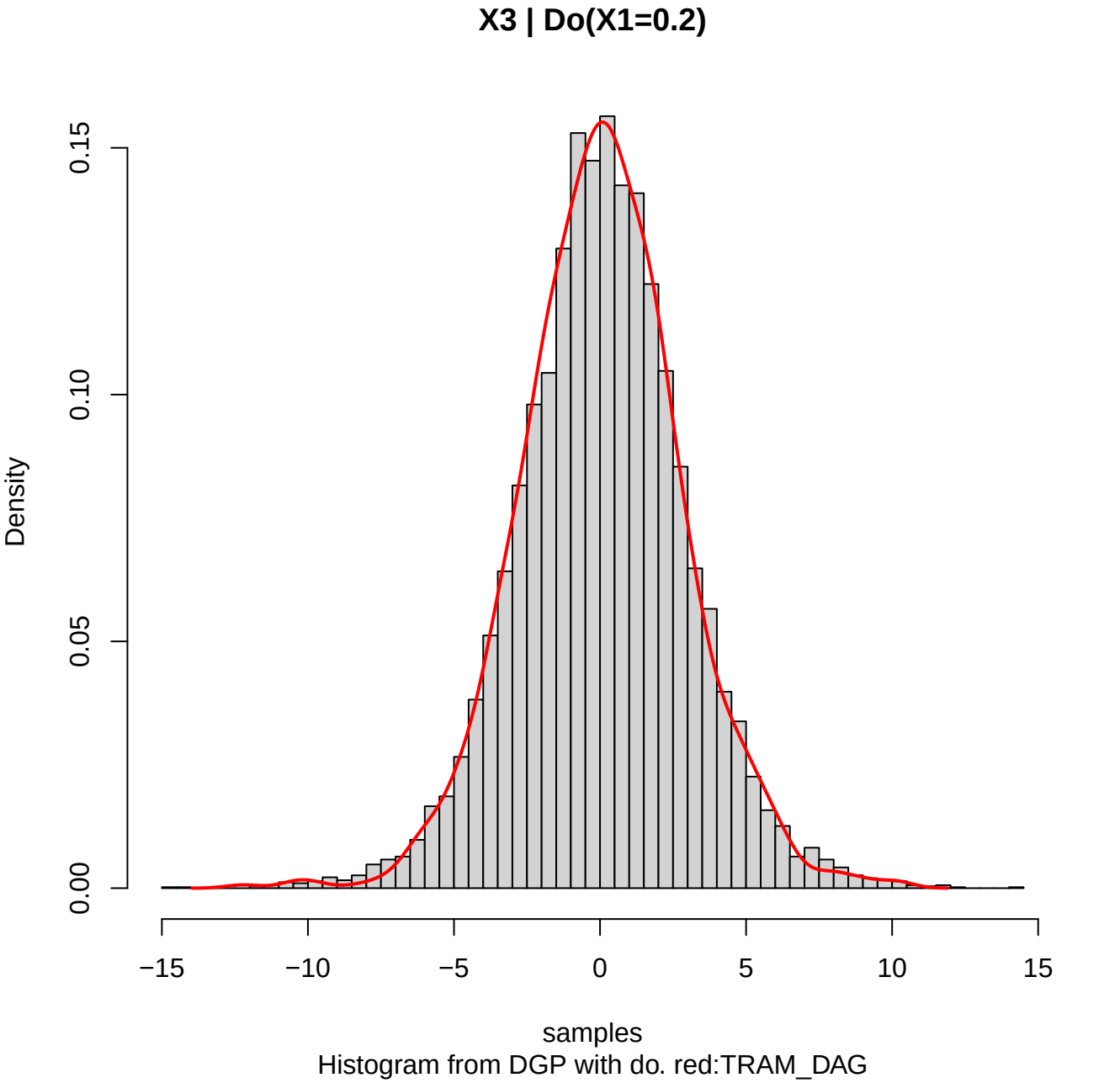

h_I(X2) Black: COLR, Red: Our Model
trafo
10
5
0
−5
−10
−15
−3
−2
−1
0
1
2
X2

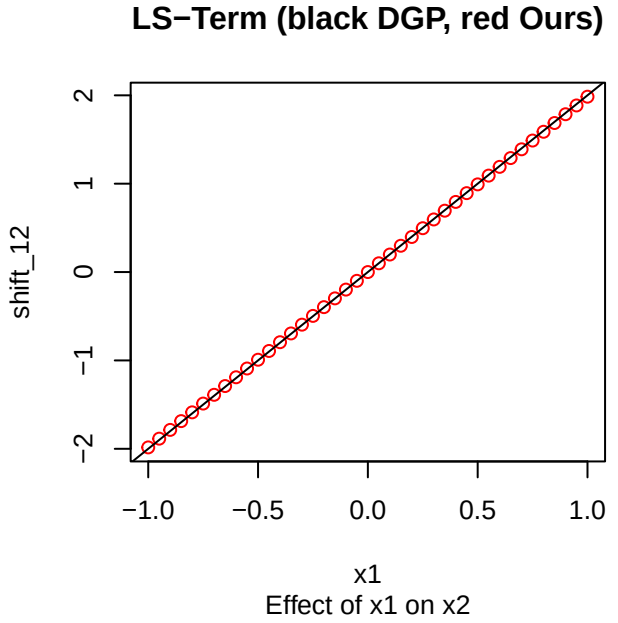

h_I(X3) Colr and Our Model
trafo
10
5
0
−5
−10
−15
−10
−5
0
5
10
15
X3

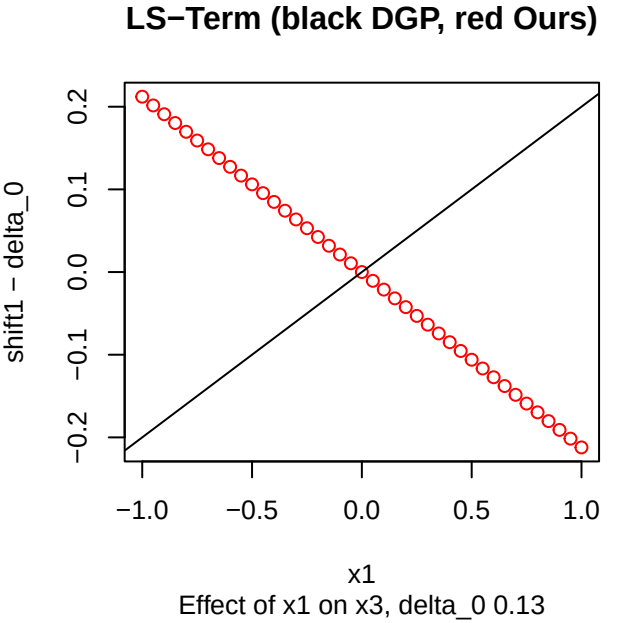

X1 red: ours, black: data
X2 red: ours, black: data
X3 red: ours, black: data
Density
samples
0
1
2
3
4
−0.2
0.2
0.6
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−3
−2
−1
0
1
2
0.00
0.05
0.10
0.15
−15
−5
0
5
10

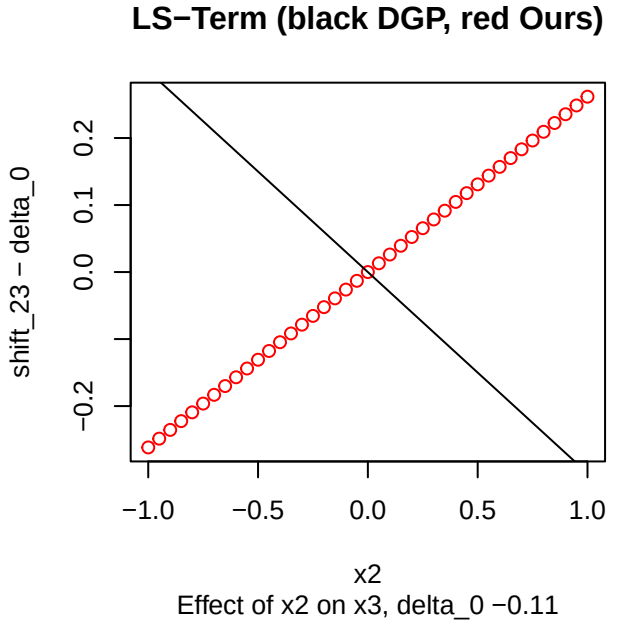

Do(X1=0.2) X2
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−2
−1
0
1
samples
Histogram from DGP with do. Blue: Colr


X2 | Do(X1=0.2)
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−2
−1
0
1
samples
Histogram from DGP with do. red:TRAM_DAG


X3 | Do(X1=0.2)
Density
0.00
0.05
0.10
0.15
−15
−10
−5
0
5
10
15
samples
Histogram from DGP with do. red:TRAM_DAG


LS−Term (black DGP, red Ours)
shift_12
2
1
0
−1
−2
−1.0
−0.5
0.0
0.5
1.0
x1
Effect of x1 on x2


LS−Term (black DGP, red Ours)
shift1 − delta_0
0.2
0.1
0.0
−0.1
−0.2
−1.0
−0.5
0.0
0.5
1.0
x1
Effect of x1 on x3, delta_0 0.13


LS−Term (black DGP, red Ours)
shift_23 − delta_0
0.2
0.1
0.0
−0.2
−1.0
−0.5
0.0
0.5
1.0
x2
Effect of x2 on x3, delta_0 −0.11