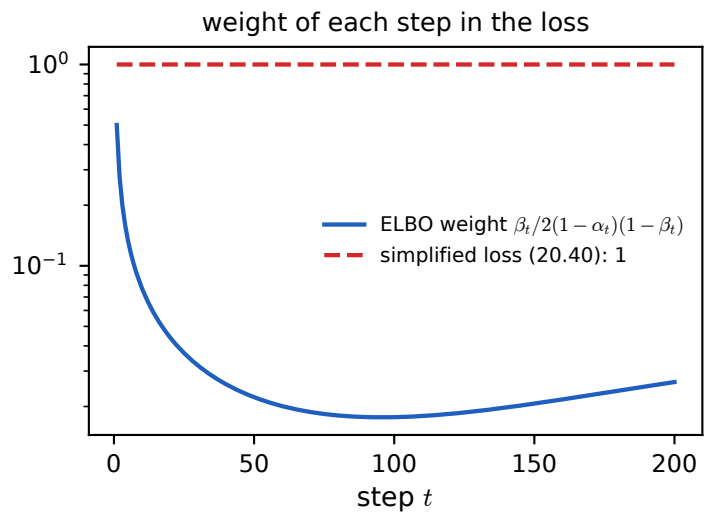
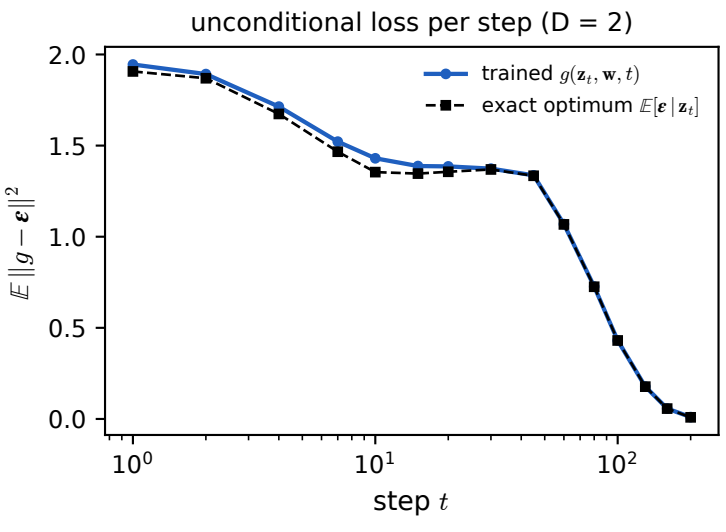
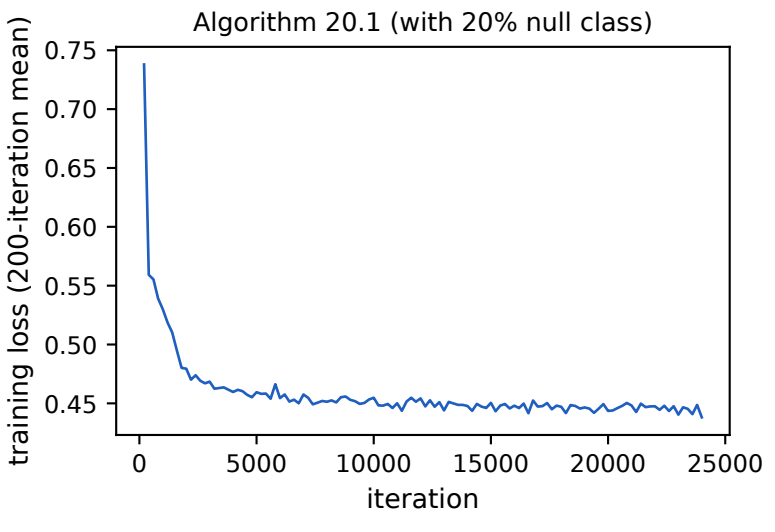

weight of each step in the loss
10^0
10^{-1}
ELBO weight $\beta_t/2(1-\alpha_t)(1-\beta_t)$
simplified loss (20.40): 1
0
50
100
150
200
step t


unconditional loss per step (D = 2)
$\mathbb{E}\,\|g-\boldsymbol{\epsilon}\|^2$
2.0
1.5
1.0
0.5
0.0
trained $g(\mathbf{z}_t, \mathbf{w}, t)$
exact optimum $\mathbb{E}[\boldsymbol{\epsilon}\,|\,\mathbf{z}_t]$
10^0
10^1
10^2
step t


Algorithm 20.1 (with 20% null class)
training loss (200-iteration mean)
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0
5000
10000
15000
20000
25000
iteration