

Training Loss (method=vanilla, lr=0.001 (constant), mu=1)

