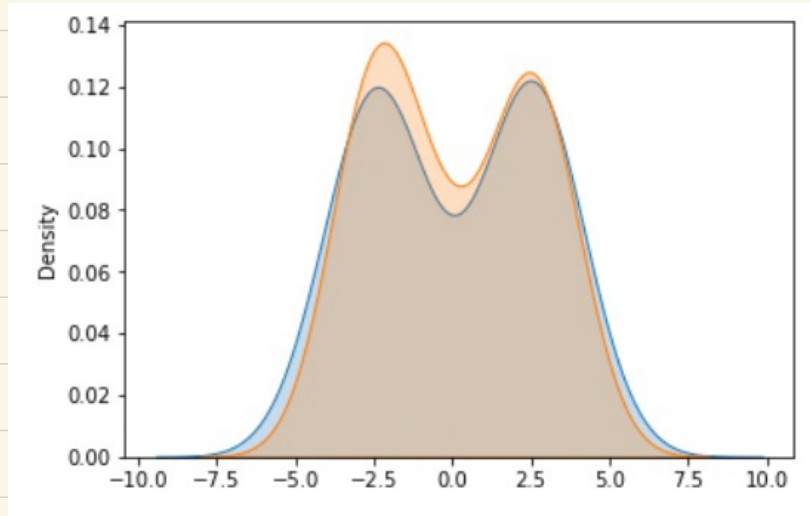
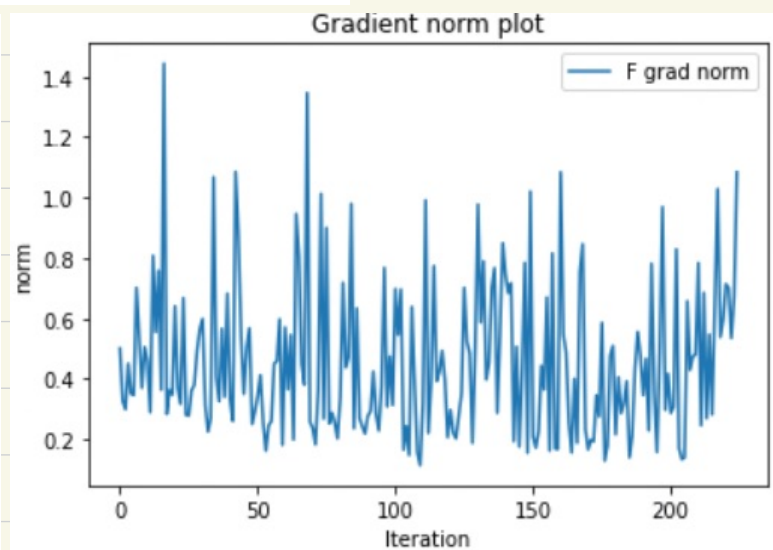
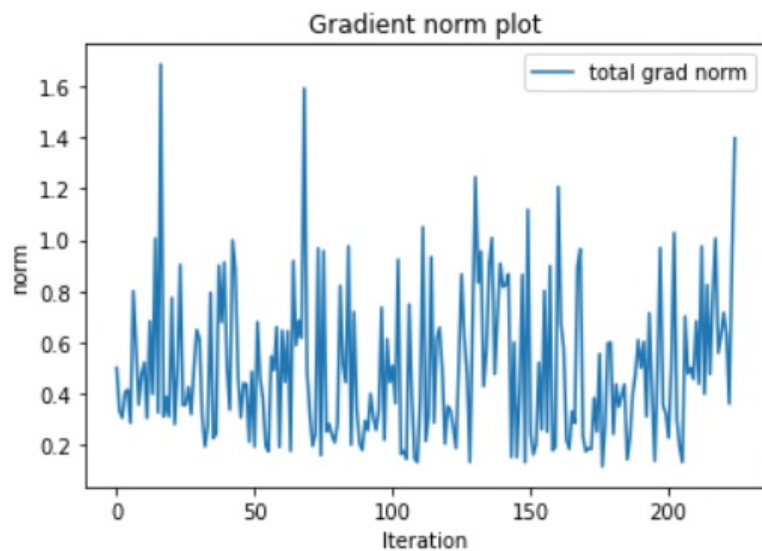
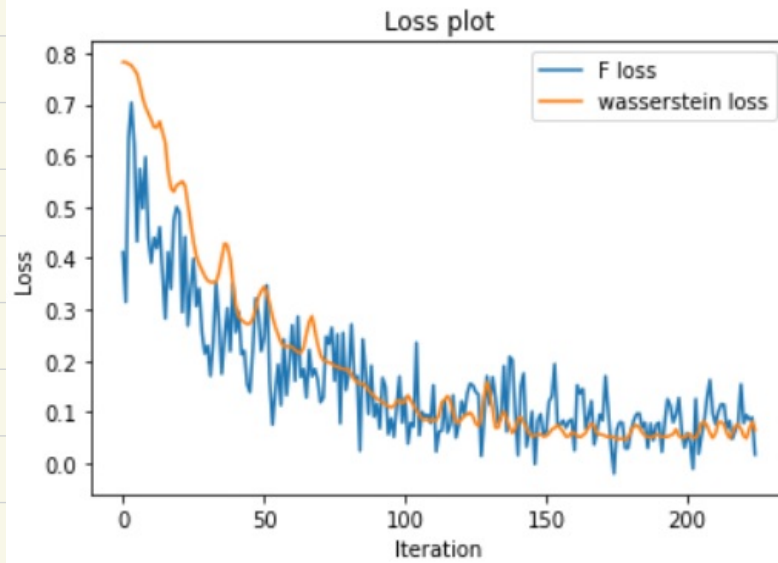


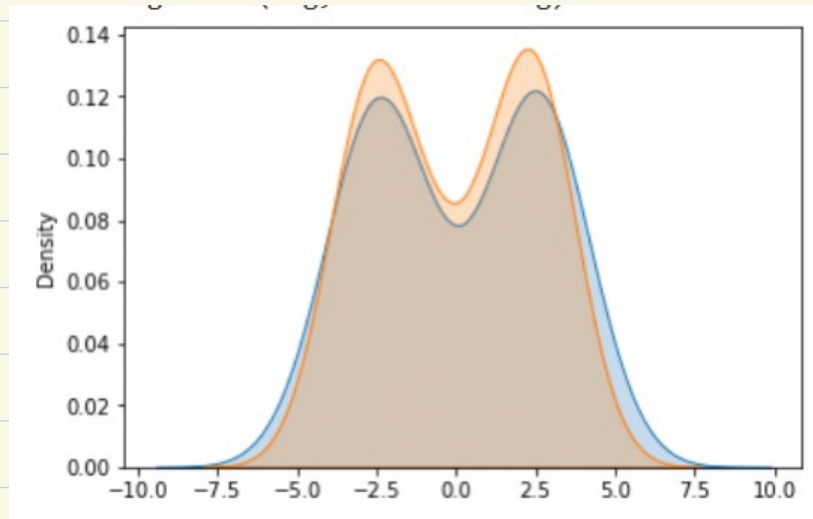
Nesterov = False, SGD, momentum = 0.9, exponential decay after a number of iterations.



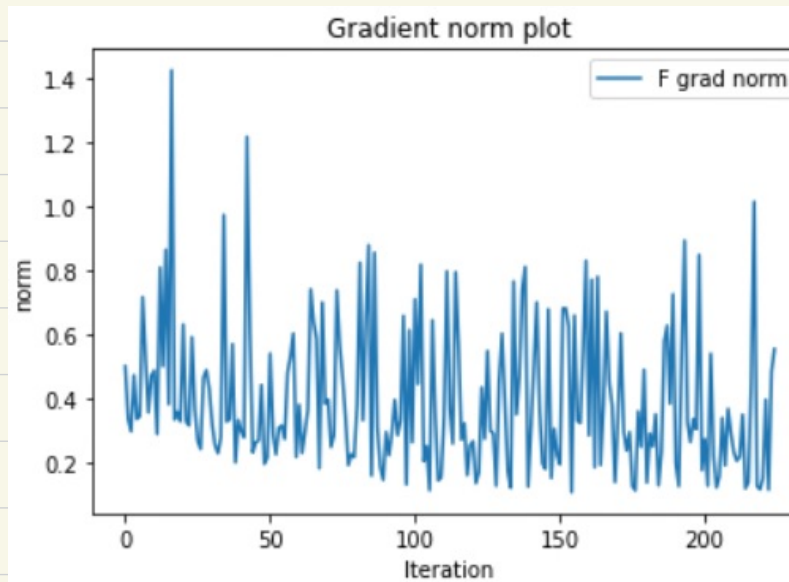
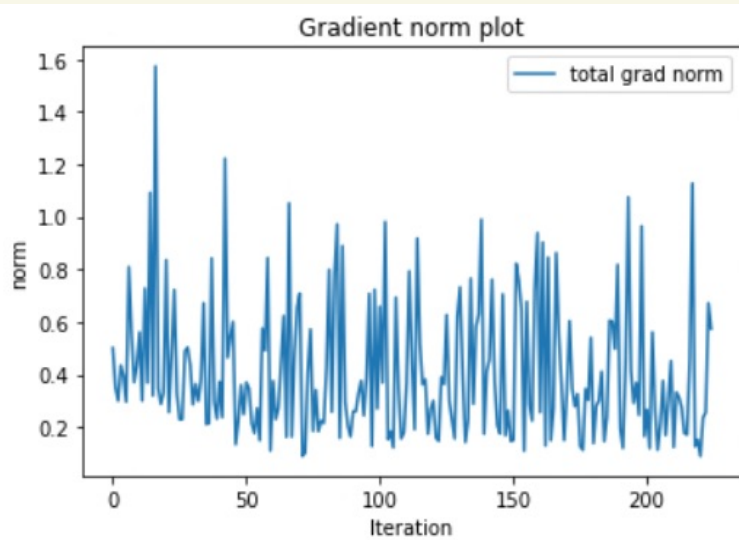
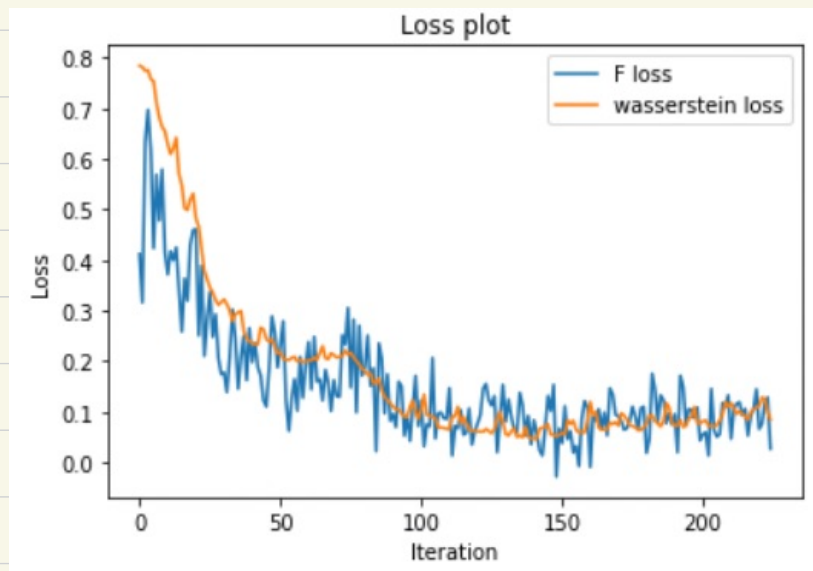
$$w_2 = 100.78977$$



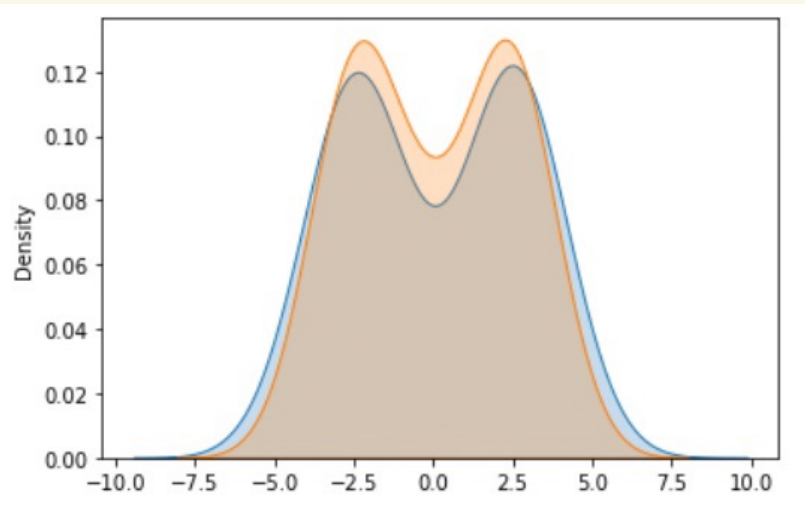
Nesterov = True , SGD, momentum = 0.9, exponential decay after a number of iterations.



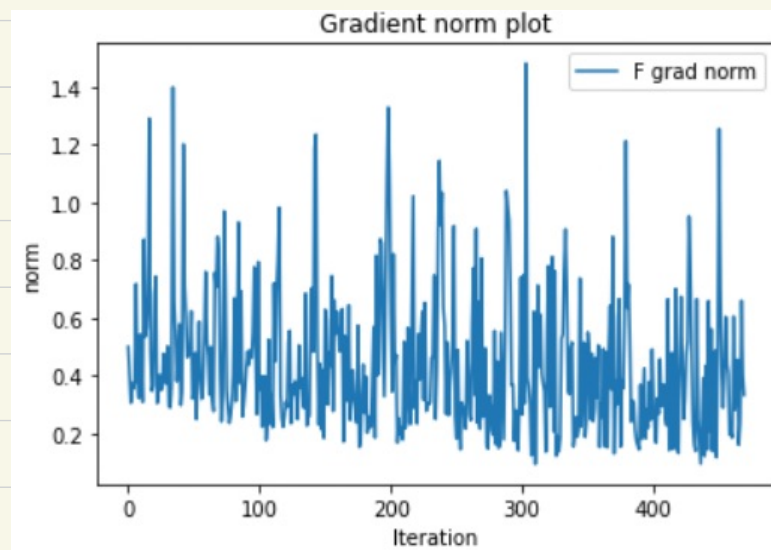
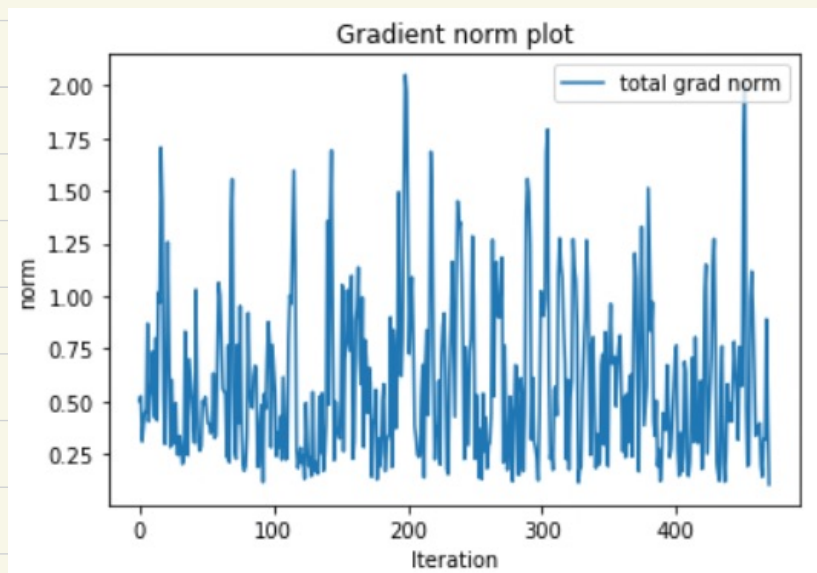
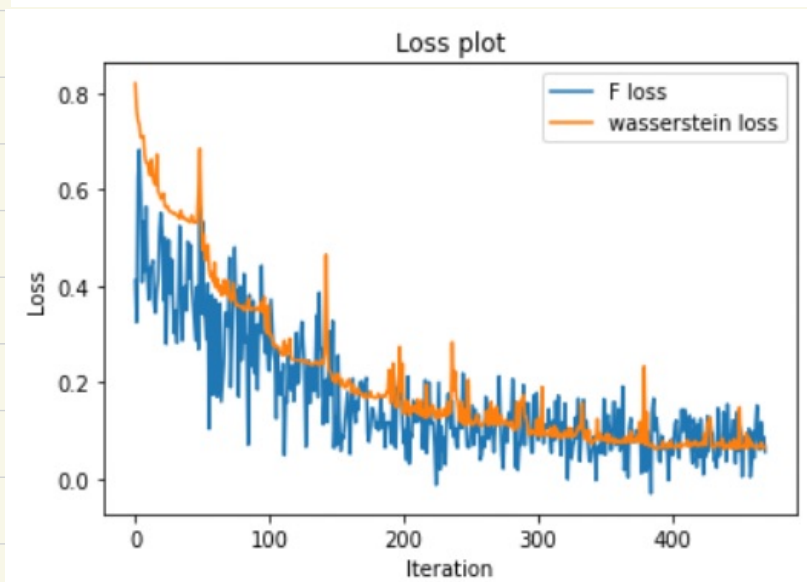
$W_2 = 12237124$



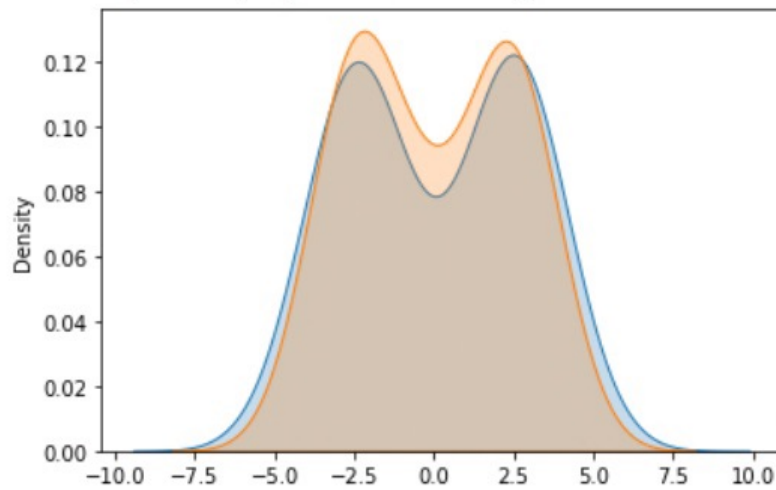
$\frac{A_1}{B_1 + C_1 \cdot \text{iteration}}$ for lr and $\frac{A_2}{B_2 + C_2 \cdot \text{iteration}}$ for jko. Nesterov = False, SGD, momentum = 0.9



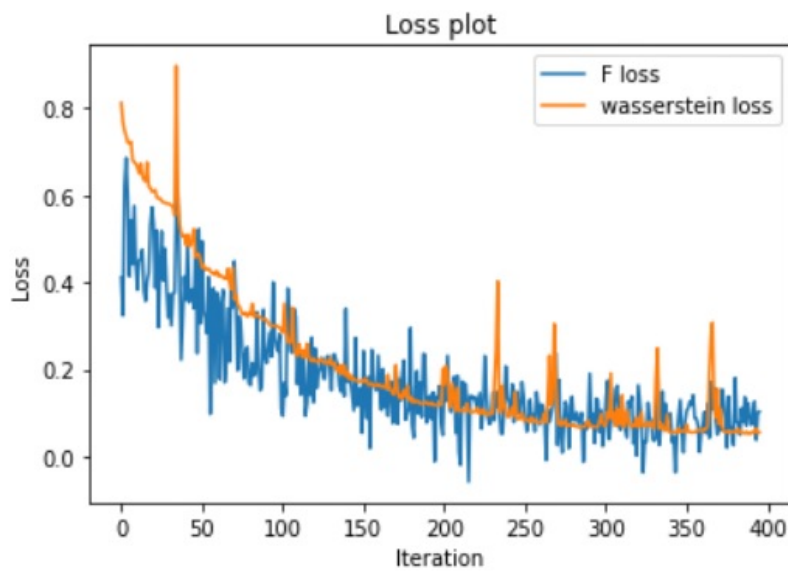
$$w_2 = 0.11113917$$



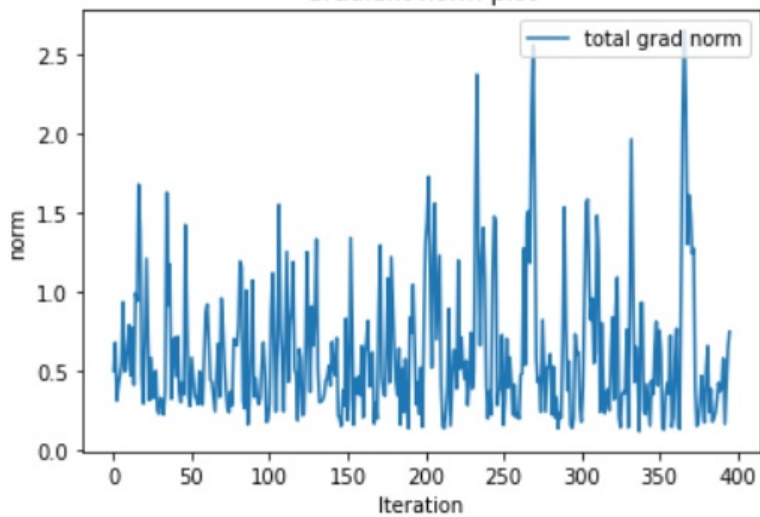
Nesterov = True , SGD, momentum = 0.9



$w_2 = 0.10824274$



Gradient norm plot



Gradient norm plot

