

Question 25.d)

Convergence of the gradient norm for random initialization

