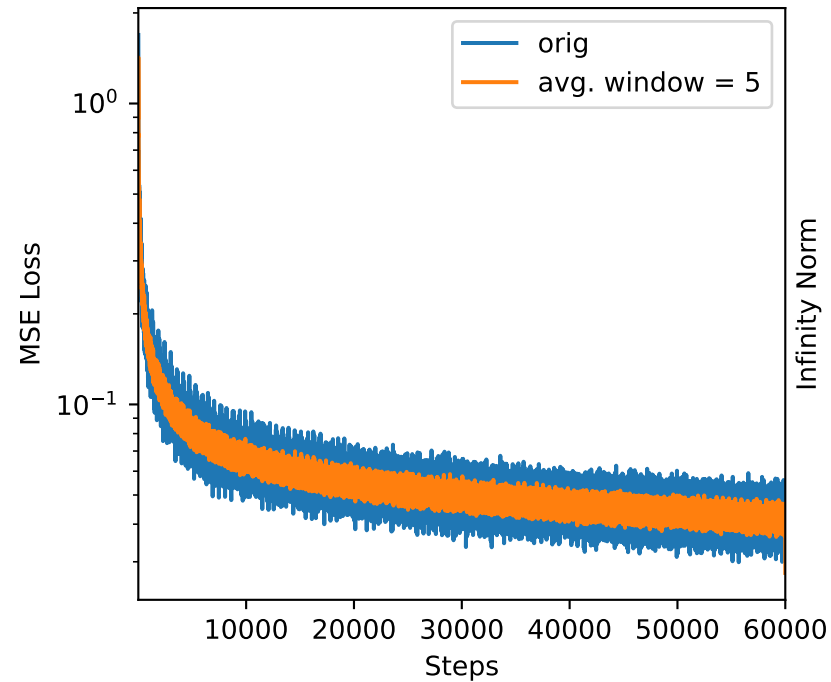
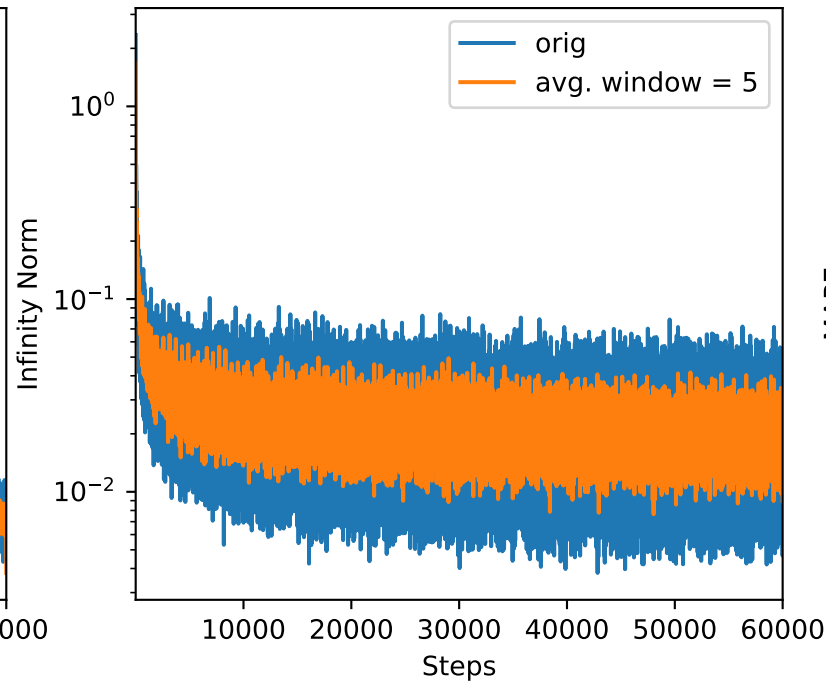


Transformer: 42 NY Clients, SGD, lr:0.001, steps:60000, features:8 BIG_BATCH 2048
Gradient Norm
Test set.

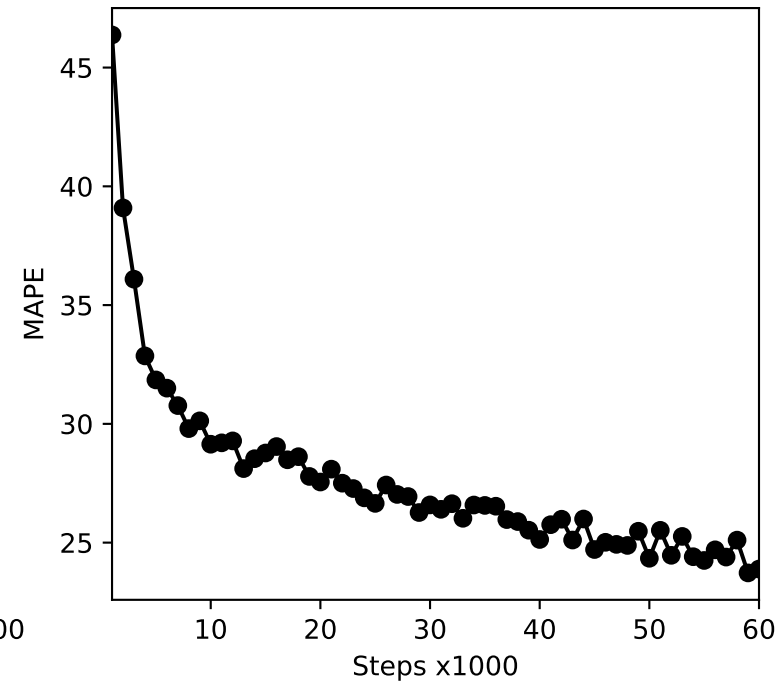
Train Loss



Gradient Norm



Test set.



Reconstruction

