Algorithm 1 Base Line

1: **procedure** MAIN
2:     $EPISODES \leftarrow N$
3:     $ITERATIONS \leftarrow M$
4:     Initialize simulation
5:
6:     **for** $i \leftarrow 0,\ EPISODES$ **do**
7:
8:         Start Simulation
9:
10:         **for** $j \leftarrow 1,\ ITERATIONS$ **do**
11:
12:             Sample new set of actions $\{a_{k,j}\}_{k=1}^{K}$ with $\mathcal{N}(0, \frac{constraint}{\sqrt{2}})$
13:             Interpolate trajectory between $\{a_{k,j-1}\}_{k=1}^{K}$ and $\{a_{k,j}\}_{k=1}^{K}$
14:             Add current state and action $\{s_k, a_{k,j}\}_{k=1}^{K}$ to replay buffer R
15:             Stop simulation
16:

Algorithm 2 Autoencoder

1: **procedure** MAIN
2:     $EPISODES \leftarrow N$
3:     $ITERATIONS \leftarrow M$
4:     Initialize simulation
5:     Initialize Autoencoder
6:
7:     **for** $i \leftarrow 0,\ EPISODES$ **do**
8:
9:         Start Simulation
10:
11:         **for** $j \leftarrow 1,\ ITERATIONS$ **do**
12:
13:             Calculate new set of desired actions $\{a_{k,j}\}_{k=1}^{K}$ with CMA-ES
14:             Interpolate trajectory between $\{a_{k,j-1}\}_{k=1}^{K}$ and $\{a_{k,j}\}_{k=1}^{K}$
15:             Add current state and action $\{s_k, a_{k,j}\}_{k=1}^{K}$ to replay buffer R
16:             Stop simulation
17:
18:     Initialize Autoencoder with all the data in the replay buffer R
19:     Train Autoencoder
20:     Save loss and hidden layer data