Generated distributions when increasing pretraining epochs Distribution after 5 pretraining epochs Distribution after 10 pretraining epochs 150 frequency Distribution after 10 pretraining epochs + Ir decay Distribution after 5 pretraining epochs + Ir decay frequency epoch 0 epoch 8 epoch 10 epoch 2 Target dist. epoch 4 epoch 6 yaw yaw