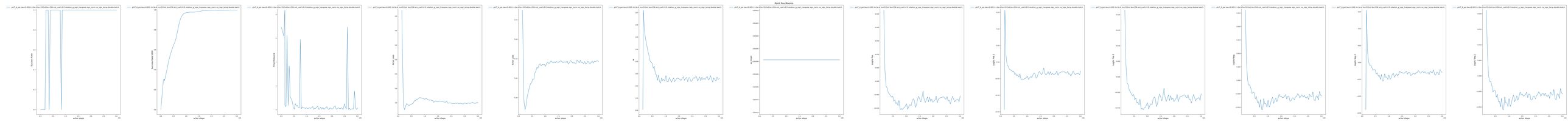

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps 1e5

0.0

0.2

0.4

0.6

0.8

1.0

Success Rate

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.0

0.2

0.4

0.6

0.8

1.0

Success Rate 1000

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0

2

4

6

8

Final Distance

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

4.9

5.0

5.1

5.2

5.3

5.4

5.5

Actor Loss

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

5.50

5.51

5.52

5.53

5.54

Critic Loss

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.99

1.00

1.01

1.02

1.03

1.04

1.05

1.06

1.07

w

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.00370

0.00375

0.00380

0.00385

0.00390

0.00395

0.00400

0.00405

0.00410

w_mean

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.010

0.005

0.000

0.005

0.010

0.015

0.020

0.025

Logits Pos

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.03

0.02

0.01

0.00

0.01

0.02

0.03

Logits Pos 1

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.010

0.005

0.000

0.005

0.010

0.015

0.020

Logits Pos 2

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.010

0.005

0.000

0.005

0.010

0.015

0.020

0.025

Logits Neg

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.0 0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.03

0.02

0.01

0.00

0.01

0.02

0.03

Logits Neg 1

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

0.5 1.0 1.5 2.0 2.5 3.0

actor steps

1e5

0.010

0.005

0.000

0.005

0.010

0.015

0.020

Logits Neg 2

phiT_A_psi tau=0.005 lr=3e-4 hs=512x4 bs=256 ent_coef=0.0 rotation_g_repr_tranpose repr_norm no_repr_temp double batch

Point FourRooms