```
bias: [1.097 1.123 1.082 1.138 1.121]
                                                                    std: [0.333 0.329 0.332 0.33 0.328]
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
                                                                                                                                      0 1 0 1 0
w_spatial, l, T] = [0.01, 30, 0, 3, 0, 10, 1, 5,
                                                                    MSE: [1.146 1.17 1.132 1.185 1.168]
                                                                   time spent until now: 19.5 mins
                                                                                                                                      10100
23:00, 03/12; num of cores:16
1 1 1 1 0
                                                                                                                                      1 1 0 1 1
                                                                   [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0, w_spatial, l, T] = [0.01, 30, 3, 3, 0, 10, 1, 5,
10110
                                                                                                                                      MC-based mean [average reward] and its std: [10.846
                                                                                                                                      0.398]
1 1 1 1 0
                                                                                                                                      DR, DR2, IS, Susan, DR_NS
                                                                   23:20, 03/12; num of cores:16
                                                                                                                                       bias: [0.454 0.502 0.442 0.513 0.479]
0 0 1 1 1
                                                                   1 1 0 1 1
                                                                                                                                       std: [0.353 0.335 0.351 0.337 0.351]
                                                                                                                                       MSE: [0.575 0.604 0.564 0.614 0.594]
1 1 0 1 0
                                                                   10000
                                                                                                                                      time spent until now: 38.9 mins
MC-based mean [average reward] and its std: [10.964
                                                                   00101
0.335]
DR, DR2, IS, Susan, DR_NS
                                                                   10100
                                                                                                                                       [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
bias: [0.345 0.379 0.331 0.392 0.368]
                                                                                                                                      w_{spatial}, l, T] = [0.01, 30, 2, 3, 2, 10, 1, 5,
std: [0.325 0.33 0.326 0.33 0.324]
MSE: [0.474 0.503 0.465 0.512 0.49 ]
                                                                   0 1 0 0 1
                                                                                                                                      6721
                                                                                                                                      23:39, 03/12; num of cores:16
time spent until now: 6.7 mins
                                                                   MC-based mean [average reward] and its std: [11.76
                                                                                                                                      00100
                                                                   0.461]
                                                                   DR, DR2, IS, Susan, DR NS
                                                                                                                                      00100
                                                                    bias: [0.463 0.416 0.475 0.404 0.434]
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0, w_spatial, l, T] = [0.01, 30, 1, 3, 0, 10, 1, 5,
                                                                    std: [0.332 0.325 0.327 0.33 0.338]
                                                                                                                                      1 1 0 1 0
                                                                    MSE: [0.57 0.528 0.577 0.522 0.55 ]
                                                                   time spent until now: 26.0 mins
                                                                                                                                      1 1 0 1 0
23:07, 03/12; num of cores:16
                                                                                                                                      10000
0 1 0 0 0
00001
                                                                                                                                      MC-based mean [average reward] and its std: [10.219
0 1 0 1 0
                                                                                                                                      DR, DR2, IS, Susan, DR_NS
                                                                                                                                      bias: [1.1 1.125 1.085 1.141 1.12 ]
std: [0.342 0.336 0.342 0.337 0.343]
10100
                                                                    [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
                                                                                                                                       MSE: [1.152 1.174 1.138 1.19 1.171]
                                                                   w_{spatial}, l, T] = [0.01, 30, 0, 3, 2, 10, 1, 5,
                                                                                                                                      time spent until now: 45.3 mins
1 1 0 1 1
                                                                   6721
                                                                   23:26, 03/12; num of cores:16
MC-based mean [average reward] and its std: [10.845]
                                                                   1 1 1 1 0
DR, DR2, IS, Susan, DR_NS
                                                                                                                                       [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
                                                                   10110
bias: [0.451 0.498 0.437 0.511 0.476]
                                                                                                                                      w_{spatial}, l, T] = [0.01, 30, 3, 3, 2, 10, 1, 5,
 std: [0.338 0.327 0.336 0.33 0.342]
                                                                                                                                      672]
                                                                   1 1 1 1 0
MSE: [0.564 0.596 0.551 0.608 0.586]
                                                                                                                                      23:45, 03/12; num of cores:16
time spent until now: 13.1 mins
                                                                                                                                      11011
                                                                   0 0 1 1 1
                                                                                                                                      10000
                                                                   1 1 0 1 0
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0, w_spatial, l, T] = [0.01, 30, 2, 3, 0, 10, 1, 5,
                                                                                                                                      00101
                                                                   MC-based mean [average reward] and its std: [10.965
                                                                   0.333]
                                                                                                                                      10100
                                                                   DR, DR2, IS, Susan, DR_NS
23:13, 03/12; num of cores:16
                                                                    bias: [0.348 0.38 0.334 0.394 0.374]
00100
                                                                                                                                      0 1 0 0 1
                                                                    std: [0.334 0.338 0.335 0.337 0.337]
                                                                    MSE: [0.482 0.509 0.473 0.518 0.503]
00100
                                                                                                                                      MC-based mean [average reward] and its std: [11.761
                                                                   time spent until now: 32.4 mins
                                                                                                                                      0.456]
1 1 0 1 0
                                                                                                                                      DR, DR2, IS, Susan, DR_NS
bias: [0.46 0.414 0.472 0.401 0.428]
                                                                                                                                       std: [0.336 0.331 0.33 0.337 0.344]
1 1 0 1 0
                                                                   [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0, w_spatial, l, T] = [0.01, 30, 1, 3, 2, 10, 1, 5,
                                                                                                                                       MSE: [0.57 0.53 0.576 0.524 0.549]
                                                                                                                                      time spent until now: 51.7 mins
10000
                                                                   23:33, 03/12; num of cores:16
MC-based mean [average reward] and its std: [10.218]
                                                                   0 1 0 0 0
DR, DR2, IS, Susan, DR_NS
                                                                   00001
```

```
_____
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
w_spatial, l, T] = [0.01, 30, 0, 3, 0, 15, 1, 5,
23:52. 03/12: num of cores:16
1 1 1 1 0
10110
1 1 1 1 0
00111
1 1 0 1 0
MC-based mean [average reward] and its std: [16.74
0.6391
DR, DR2, IS, Susan, DR_NS
bias: [0.226 0.285 0.2 0.311 0.262]
std: [0.96 0.997 0.955 1.002 0.96]
MSE: [0.986 1.037 0.976 1.049 0.995]
time spent until now: 58.2 mins
 [lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
w_{spatial}, l, T] = [0.01, 30, 1, 3, 0, 15, 1, 5,
23:58, 03/12; num of cores:16
0 1 0 0 0
00001
0 1 0 1 0
10100
1 1 0 1 1
MC-based mean [average reward] and its std: [16.948
0.521]
DR, DR2, IS, Susan, DR_NS
bias: [0.037 0.078 0.014 0.102 0.08 ]
 std: [0.983 1.007 0.986 1.003 0.98 ]
MSE: [0.984 1.01 0.986 1.008 0.983]
time spent until now: 64.6 mins
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
w_{spatial}, l, T] = [0.01, 30, 2, 3, 0, 15, 1, 5,
00:05, 03/13; num of cores:16
00100
00100
1 1 0 1 0
1 1 0 1 0
10000
```

```
MC-based mean [average reward] and its std: [16.024
0.279]
DR, DR2, IS, Susan, DR_NS
bias: [0.946 1.001 0.92 1.027 1.001]
 std: [0.989 1.001 0.986 1.004 0.987]
 MSE: [1.369 1.416 1.349 1.436 1.406]
time spent until now: 71.1 mins
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0,
w_{spatial}, l, T] = [0.01, 30, 3, 3, 0, 15, 1, 5,
6721
00:11, 03/13; num of cores:16
1 1 0 1 1
10000
00101
10100
0 1 0 0 1
MC-based mean [average reward] and its std: [16.81
1.756]
DR, DR2, IS, Susan, DR_NS
bias: [0.172 0.22 0.151 0.24 0.224]
 std: [0.994 0.993 0.984 1.003 0.99 ]
 MSE: [1.009 1.017 0.996 1.031 1.015]
time spent until now: 77.5 mins
_____
______
[lam, w_hidden, pattern_seed, sd_D,sd_R, u_0, w_spatial, l, T] = [0.01, 30, 0, 3, 2, 15, 1, 5,
00:18, 03/13; num of cores:16
11110
10110
11110
0 0 1 1 1
1 1 0 1 0
```