```
    allow_broadcast=True)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/constant_op.py", line 290, in _constant_impl
    name=name).outputs[0]
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/util/deprecation.py", line 507, in new_func
    return func(*args, **kwargs)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 3617, in create_op
    self._create_op_helper(ret, compute_device=compute_device)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 3702, in _create_op_helper
    all_colocation_groups.extend(colocation_op.colocation_groups())
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 2095, in colocation_groups
    class_attr = self.get_attr("_class")
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 2640, in get_attr
    data = c_api.TF_GetBuffer(buf)
  File "/home/ubuntu/anaconda3/lib/python3.7/contextlib.py", line 158, in __exit__
    if sys.exc_info()[1] is value:
KeyboardInterrupt
Traceback (most recent call last):
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 2639, in get_attr
    c_api.TF_OperationGetAttrValueProto(self._c_op, name, buf)
tensorflow.python.framework.errors_impl.InvalidArgumentError: Operation 'w_3/batchnorm/add_1' has no attr named '_XlaCompile'.

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 398, in _MaybeCompile
    xla_compile = op.get_attr("_XlaCompile")
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 2643, in get_attr
    raise ValueError(str(e))
ValueError: Operation 'w_3/batchnorm/add_1' has no attr named '_XlaCompile'.

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 70, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/simu_funs.py", line 84, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 252, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 158, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 158, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 87, in getOneRegionValue
    epsilon = epsilon)
  File "/home/ubuntu/main.py", line 247, in getWeight
    w_hidden = w_hidden, Learning_rate = lr, reg_weight = reg_weight)
  File "/home/ubuntu/weight.py", line 93, in __init__
    self.train_op = tf.train.AdamOptimizer(Learning_rate).minimize(self.loss) # mute the later part?+reg_weight * self.reg_loss # not me
ntioned in the paper
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/training/optimizer.py", line 403, in minimize
    grad_loss=grad_loss)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/training/optimizer.py", line 512, in compute_gradients
    colocate_gradients_with_ops=colocate_gradients_with_ops)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_impl.py", line 158, in gradients
    unconnected_gradients)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 731, in _GradientsHelper
    lambda: grad_fn(op, *out_grads))
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 403, in _MaybeCompile
    return grad_fn()  # Exit early
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 731, in <lambda>
    lambda: grad_fn(op, *out_grads))
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/math_grad.py", line 1008, in _AddGrad
    gx = array_ops.reshape(math_ops.reduce_sum(grad, rx), sx)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/util/dispatch.py", line 180, in wrapper
    return target(*args, **kwargs)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/math_ops.py", line 1458, in reduce_sum
    name=name))
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gen_math_ops.py", line 10935, in _sum
    name=name)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/op_def_library.py", line 664, in _apply_op_helper
    for attr in op_def.attr:
KeyboardInterrupt
ubuntu@ip-172-31-6-10:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:20, 04/13; num of cores:96
sd_u_0_30

Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, t_func] = [16, None, None, 40, 0.5, 1.5, 20, None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[85, 95, 100, 110, 120, 125], [0, 20, 40], [7], [[0.0001, 5e-05], [0.0001, 5e-05]]
]

-------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 0]
```

```
max(u_0) =  191.7
O_threshold = 85
means of Order:

83.3 97.7 14.6 165.6 28.3

66.3 120.1 50.2 57.7 63.6

122.1 191.7 101.7 55.3 121.6

76.2 99.2 147.0 70.1 100.4

64.9 93.7 110.3 60.4 86.4

target policy:

0 1 0 1 0

0 1 0 0 0

1 1 1 0 1

0 1 1 0 1

0 1 1 0 1

number of reward locations:  13
O_threshold = 95
number of reward locations:  11
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  7
O_threshold = 120
number of reward locations:  6
O_threshold = 125
number of reward locations:  3
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

----------------------------------------
Value of Behaviour policy:43.474
O_threshold = 85
MC for this TARGET:[55.52, 0.062]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.19, -0.51, -1.7]][[-0.54, -55.52, -12.05]]
std:[[0.38, 0.38, 0.22]][[0.34, 0.0, 0.22]]
MSE:[[0.42, 0.64, 1.71]][[0.64, 55.52, 12.05]]
MSE(-DR):[[0.0, 0.22, 1.29]][[0.22, 55.1, 11.63]]
***
==============
O_threshold = 95
MC for this TARGET:[54.029, 0.06]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.15, -1.37, -1.91]][[-2.11, -54.03, -10.56]]
std:[[0.41, 0.42, 0.23]][[0.33, 0.0, 0.22]]
MSE:[[1.22, 1.43, 1.92]][[2.14, 54.03, 10.56]]
MSE(-DR):[[0.0, 0.21, 0.7]][[0.92, 52.81, 9.34]]
***
==============
O_threshold = 100
MC for this TARGET:[55.41, 0.054]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.06, -5.24, -5.57]][[-6.88, -55.41, -11.94]]
std:[[0.52, 0.52, 0.31]][[0.35, 0.0, 0.22]]
MSE:[[5.09, 5.27, 5.58]][[6.89, 55.41, 11.94]]
MSE(-DR):[[0.0, 0.18, 0.49]][[1.8, 50.32, 6.85]]
***
==============
O_threshold = 110
MC for this TARGET:[55.095, 0.056]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.68, -7.79, -7.99]][[-10.42, -55.1, -11.62]]
std:[[0.57, 0.58, 0.27]][[0.3, 0.0, 0.22]]
MSE:[[7.7, 7.81, 7.99]][[10.42, 55.1, 11.62]]
MSE(-DR):[[0.0, 0.11, 0.29]][[2.72, 47.4, 3.92]]
***
==============
O_threshold = 120
MC for this TARGET:[53.394, 0.061]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.23, -7.32, -7.57]][[-10.67, -53.39, -9.92]]
std:[[0.63, 0.64, 0.34]][[0.29, 0.0, 0.22]]
MSE:[[7.26, 7.35, 7.58]][[10.67, 53.39, 9.92]]
MSE(-DR):[[0.0, 0.09, 0.32]][[3.41, 46.13, 2.66]]
```

***
==========
O_threshold = 125
MC for this TARGET:[51.795, 0.061]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-9.21, -9.2, -9.31]][[-15.23, -51.8, -8.32]]
std:[[0.97, 1.01, 0.28]][[0.3, 0.0, 0.22]]
MSE:[[9.26, 9.26, 9.31]][[15.23, 51.8, 8.32]]
MSE(-DR):[[0.0, 0.0, 0.05]][[5.97, 42.54, -0.94]]
***
==========
******************** THIS SETTING IS GOOD ********************
[[0.42 0.64 1.71 0.64 55.52 12.05]
 [1.22 1.43 1.92 2.14 54.03 10.56]
 [5.09 5.27 5.58 6.89 55.41 11.94]
 [7.7   7.81 7.99 10.42 55.1 11.62]
 [7.26 7.35 7.58 10.67 53.39 9.92]
 [9.26 9.26 9.31 15.23 51.8   8.32]]

time spent until now: 17.2 mins

14:37, 04/13

[pattern_seed, day, sd_R] = [2, 7, 20]

max(u_O) = 191.7
O_threshold = 85
means of Order:

83.3 97.7 14.6 165.6 28.3

66.3 120.1 50.2 57.7 63.6

122.1 191.7 101.7 55.3 121.6

76.2 99.2 147.0 70.1 100.4

64.9 93.7 110.3 60.4 86.4

target policy:

0 1 0 1 0

0 1 0 0 0

1 1 1 0 1

0 1 1 0 1

0 1 1 0 1

number of reward locations: 13
O_threshold = 95
number of reward locations: 11
O_threshold = 100
number of reward locations: 9
O_threshold = 110
number of reward locations: 7
O_threshold = 120
number of reward locations: 6
O_threshold = 125
number of reward locations: 3
^CProcess Process-188:
Process Process-169:
Process Process-159:
Process Process-161:
Process Process-118:
Process Process-163:
Traceback (most recent call last):
  File "EC2.py", line 103, in <module>
Process Process-184:
Process Process-126:
Process Process-177:
Process Process-129:
Process Process-136:
Process Process-149:
Process Process-145:
Process Process-150:
Process Process-113:
Process Process-142:
Process Process-157:
Process Process-122:
Process Process-146:
Process Process-178:
Process Process-190:
Process Process-116:
Process Process-114:

```
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 70, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/simu_funs.py", line 84, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 252, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 158, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 158, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 114, in getOneRegionValue
    spatial = False)
  File "/home/ubuntu/main.py", line 257, in getWeight
    w_hidden = w_hidden, Learning_rate = lr, reg_weight = reg_weight)
  File "/home/ubuntu/weight.py", line 93, in __init__
    self.train_op = tf.train.AdamOptimizer(Learning_rate).minimize(self.loss) # mute the later part?+reg_weight * self.reg_loss # not me
ntioned in the paper
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/training/optimizer.py", line 403, in minimize
    grad_loss=grad_loss)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/training/optimizer.py", line 512, in compute_gradients
    colocate_gradients_with_ops=colocate_gradients_with_ops)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_impl.py", line 158, in gradients
    unconnected_gradients)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 731, in _GradientsHelper
    lambda: grad_fn(op, *out_grads))
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 403, in _MaybeCompile
    return grad_fn()  # Exit early
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/gradients_util.py", line 731, in <lambda>
    lambda: grad_fn(op, *out_grads))
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/math_grad.py", line 1002, in _AddGrad
    sx = array_ops.shape(x)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/array_ops.py", line 330, in shape
    return shape_internal(input, name, optimize=True, out_type=out_type)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/ops/array_ops.py", line 357, in shape_internal
    return constant(input_shape.as_list(), out_type, name=name)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/constant_op.py", line 246, in constant
    allow_broadcast=True)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/constant_op.py", line 290, in _constant_impl
    name=name).outputs[0]
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/util/deprecation.py", line 507, in new_func
    return func(*args, **kwargs)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 3616, in create_op
    op_def=op_def)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/ops.py", line 2005, in __init__
    self._traceback = tf_stack.extract_stack()
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/util/tf_stack.py", line 58, in extract_stack
    lineno = f.f_lineno
KeyboardInterrupt
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 70, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/simu_funs.py", line 84, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 252, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 158, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 158, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 114, in getOneRegionValue
    spatial = False)
  File "/home/ubuntu/main.py", line 262, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 286, in train
    subsamples = np.random.choice(N, batch_size)
KeyboardInterrupt
Exception ignored in: <function ScopedTFGraph.__del__ at 0x7fc881555d90>
Traceback (most recent call last):
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/framework/c_api_util.py", line 48, in __del__
    def __del__(self):
KeyboardInterrupt
ubuntu@ip-172-31-6-10:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:40, 04/13; num of cores:96
sd_u_0_30

Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, t_func] = [16, None, None, 40, 0.5, 1.5, 20, None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[85, 95, 100, 110, 120, 125], [0, 20, 40], [14], [[0.0001, 5e-05], [0.0001, 5e-05]
]]

--------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 0]

max(u_0) =  191.7
```

```
O_threshold = 85
means of Order:

83.3 97.7 14.6 165.6 28.3

66.3 120.1 50.2 57.7 63.6

122.1 191.7 101.7 55.3 121.6

76.2 99.2 147.0 70.1 100.4

64.9 93.7 110.3 60.4 86.4

target policy:

0 1 0 1 0

0 1 0 0 0

1 1 1 0 1

0 1 1 0 1

0 1 1 0 1

number of reward locations:  13
O_threshold = 95
number of reward locations:  11
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  7
O_threshold = 120
number of reward locations:  6
O_threshold = 125
number of reward locations:  3
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

----------------------------------------
Value of Behaviour policy:43.495
O_threshold = 85
MC for this TARGET:[55.519, 0.034]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.1, -0.23, -1.61]][[-0.61, -55.52, -12.02]]
std:[[0.36, 0.37, 0.21]][[0.22, 0.0, 0.15]]
MSE:[[0.37, 0.44, 1.62]][[0.65, 55.52, 12.02]]
MSE(-DR):[[0.0, 0.07, 1.25]][[0.28, 55.15, 11.65]]
***
==============
O_threshold = 95
MC for this TARGET:[54.03, 0.034]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.08, -1.33, -1.84]][[-2.14, -54.03, -10.53]]
std:[[0.4, 0.42, 0.24]][[0.19, 0.0, 0.15]]
MSE:[[1.15, 1.39, 1.86]][[2.15, 54.03, 10.53]]
MSE(-DR):[[0.0, 0.24, 0.71]][[1.0, 52.88, 9.38]]
***
==============
O_threshold = 100
MC for this TARGET:[55.416, 0.038]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.3, -5.49, -5.51]][[-6.92, -55.42, -11.92]]
std:[[0.37, 0.36, 0.28]][[0.2, 0.0, 0.15]]
MSE:[[5.31, 5.5, 5.52]][[6.92, 55.42, 11.92]]
MSE(-DR):[[0.0, 0.19, 0.21]][[1.61, 50.11, 6.61]]
***
==============
O_threshold = 110
MC for this TARGET:[55.097, 0.042]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-8.11, -8.26, -7.91]][[-10.38, -55.1, -11.6]]
std:[[0.28, 0.27, 0.23]][[0.17, 0.0, 0.15]]
MSE:[[8.11, 8.26, 7.91]][[10.38, 55.1, 11.6]]
MSE(-DR):[[0.0, 0.15, -0.2]][[2.27, 46.99, 3.49]]
**
==============
O_threshold = 120
MC for this TARGET:[53.391, 0.046]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.73, -7.84, -7.57]][[-10.65, -53.39, -9.9]]
std:[[0.39, 0.39, 0.22]][[0.15, 0.0, 0.15]]
MSE:[[7.74, 7.85, 7.57]][[10.65, 53.39, 9.9]]
MSE(-DR):[[0.0, 0.11, -0.17]][[2.91, 45.65, 2.16]]
**
```