

HYPERPARAMETERS:

random\_string: T1  
optimizer: Adam  
noise\_amplitude: 0.1  
clip\_gradients: True  
max\_gradient\_norm: 10  
regularization: None  
regularization\_lambda: 0  
batch\_size: 64  
note\_error\_every\_steps: 50  
train\_for\_steps: 5000  
save\_network\_every\_steps: 5000  
learning\_rate: 0.0001  
random\_seed: 4324655  
stop\_training\_at: 10

TASK PARAMETERS:

task\_name: 2DIR10  
input\_direction\_units: 100  
delay0\_from: 10  
delay0\_to: 20  
delay1\_from: 10  
delay1\_to: 90  
delay2\_from: 120  
delay2\_to: 160  
show\_direction\_for: 10  
show\_cue\_for: 100  
dim\_input: 101  
dim\_output: 2  
distractor\_probability: 1.0

MODEL PARAMETERS:

model\_name: backpropCTRNN  
dim\_input: 101  
dim\_output: 2  
dim\_recurrent: 100  
tau: 10  
nonlinearity: retanh  
input\_bias: True  
output\_bias: False  
connectivity\_cos\_exponent: 1

ADDITIONAL COMMENTS:

Training criterion: MSE loss  
Noise added at every timestep of the trial  
Inputs NOT discretized  
Output sin/cos  
Networks trained with Adam from a random initialization.

NETWORK PERFORMANCE:

O1: mse= 0.0135, error= 8.78 deg (with noise), mse= 0.0051, error= 5.16 deg (no noise)  
O2: mse= 0.0000, error= 0.00 deg (with noise), mse= 0.0000, error= 0.00 deg (no noise)