Student Name: Mohammad Mohammad Beigi
Student ID: 99102189
Subject: Motivation and classical conditioning

Advanced Topics in Neuroscience - Dr. Ali Ghazizadeh
Assignment 5

---

1. Aim: Modelling learning in classical conditioning paradigms

   Description:

   We often need to rapidly learn about the value of new stimuli that we encounter or be ready for changes to familiar stimulus values that we were familiar without prior notice. Classical conditioning includes a large group of paradigms where different combination of reward histories with stimuli can form our judgement about their current values.

   A well know model in classical conditioning and reinforcement learning is the Rescola-Wagner. This model predicts that violations of our expected reward for each stimuli or combination of stimuli causes incremental changes in our belief about their values.

   Instructions:

   Use the model

   $$v = wu$$

   Where, W is the weight, v is the expected reward and u is a binary variable that represents the presence or absence of the stimulus.

   And update w with the Rescola-Wagner (RW) rule:

   $$w \rightarrow w + \epsilon \delta u \textbf{ with } \delta = r - v$$

   Where ϵ is the learning rate.

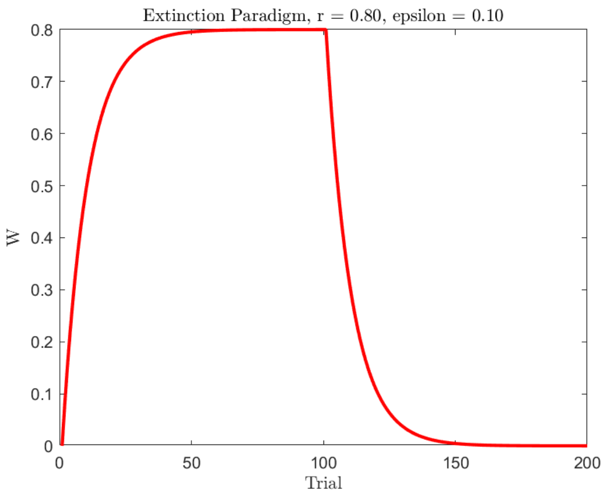
   Below are some of the well-known classical conditioning paradigms and the behavioral results of each paradigm as shown by experiments:

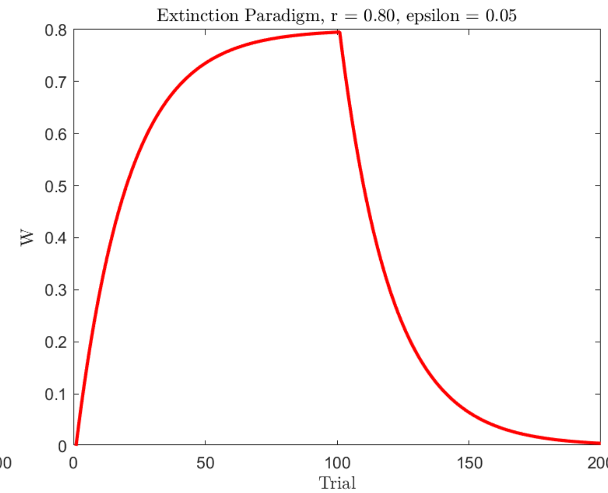
Paradigm	Pre-Train	Train	Result
Pavlovian		$s \rightarrow r$	$s \rightarrow$ 'r'
Extinction	$s \rightarrow r$	$s \rightarrow \cdot$	$s \rightarrow$ '$\cdot$'
Partial		$s \rightarrow r \quad s \rightarrow \cdot$	$s \rightarrow \alpha$ 'r'
Blocking	$s_1 \rightarrow r$	$s_1 + s_2 \rightarrow r$	$s_1 \rightarrow$ 'r' $\quad s_2 \rightarrow$ '$\cdot$'
Inhibitory		$s_1 + s_2 \rightarrow \cdot \quad s_1 \rightarrow r$	$s_1 \rightarrow$ 'r' $\quad s_2 \rightarrow -$'r'
Overshadow		$s_1 + s_2 \rightarrow r$	$s_1 \rightarrow \alpha_1$'r' $\quad s_2 \rightarrow \alpha_2$'$r$'
Secondary	$s_1 \rightarrow r$	$s_2 \rightarrow s_1$	$s_2 \rightarrow$ 'r'

- Using RW rule, simulate and plot the outcome of following paradigms: extinction, partial, blocking, inhibitory and overshadow. Assume a fixed learning rate and number trials in each phase to be such that learning almost saturates at the each of each phase. Which one the predictions of RW rule match the above table?

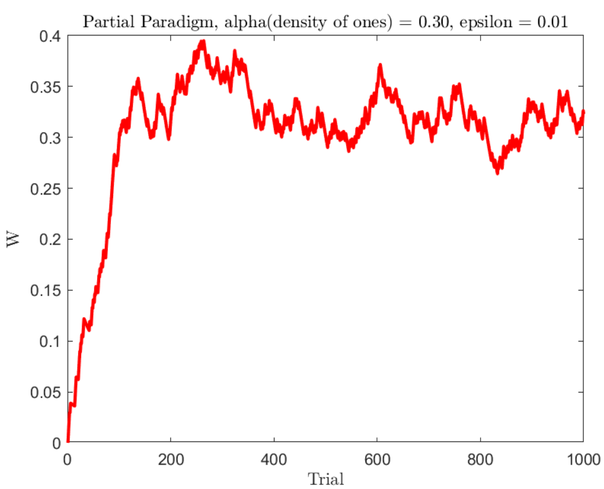
  Answer:

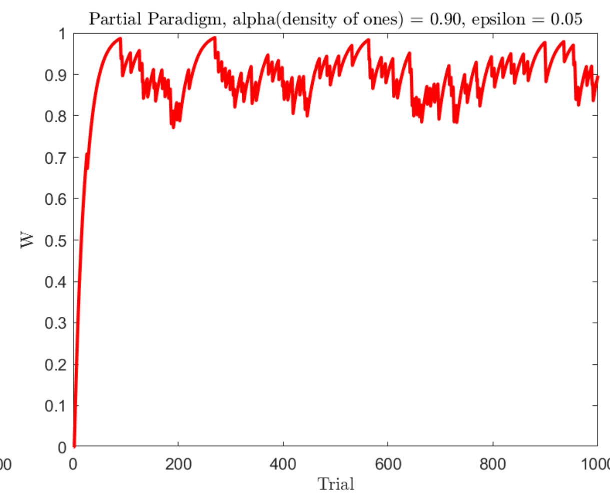
  Extinction Paradigm:

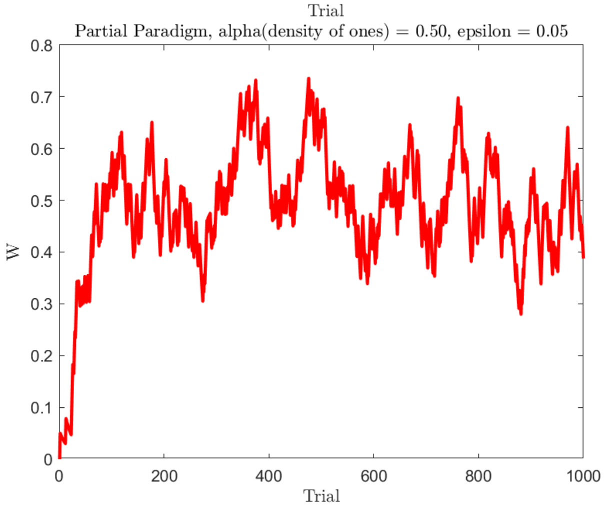


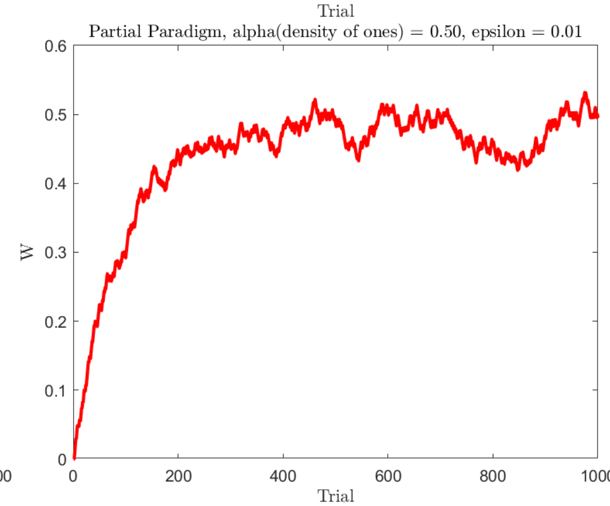


  Partial Paradigm:

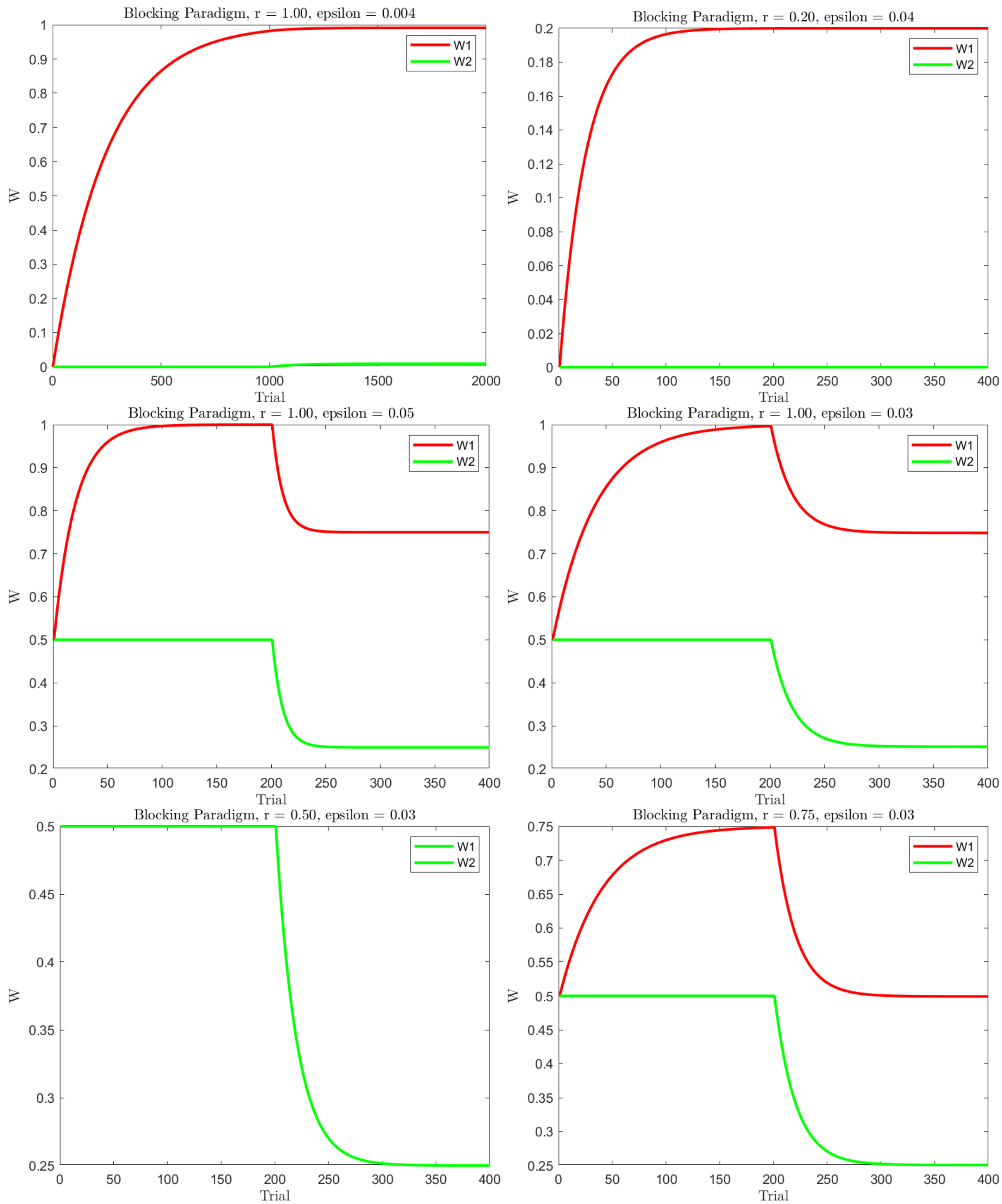


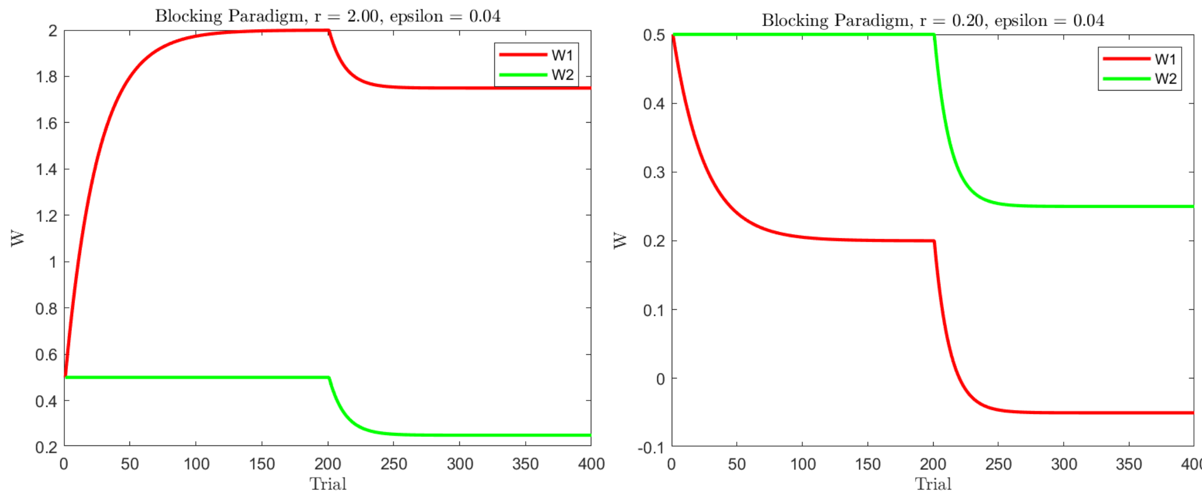




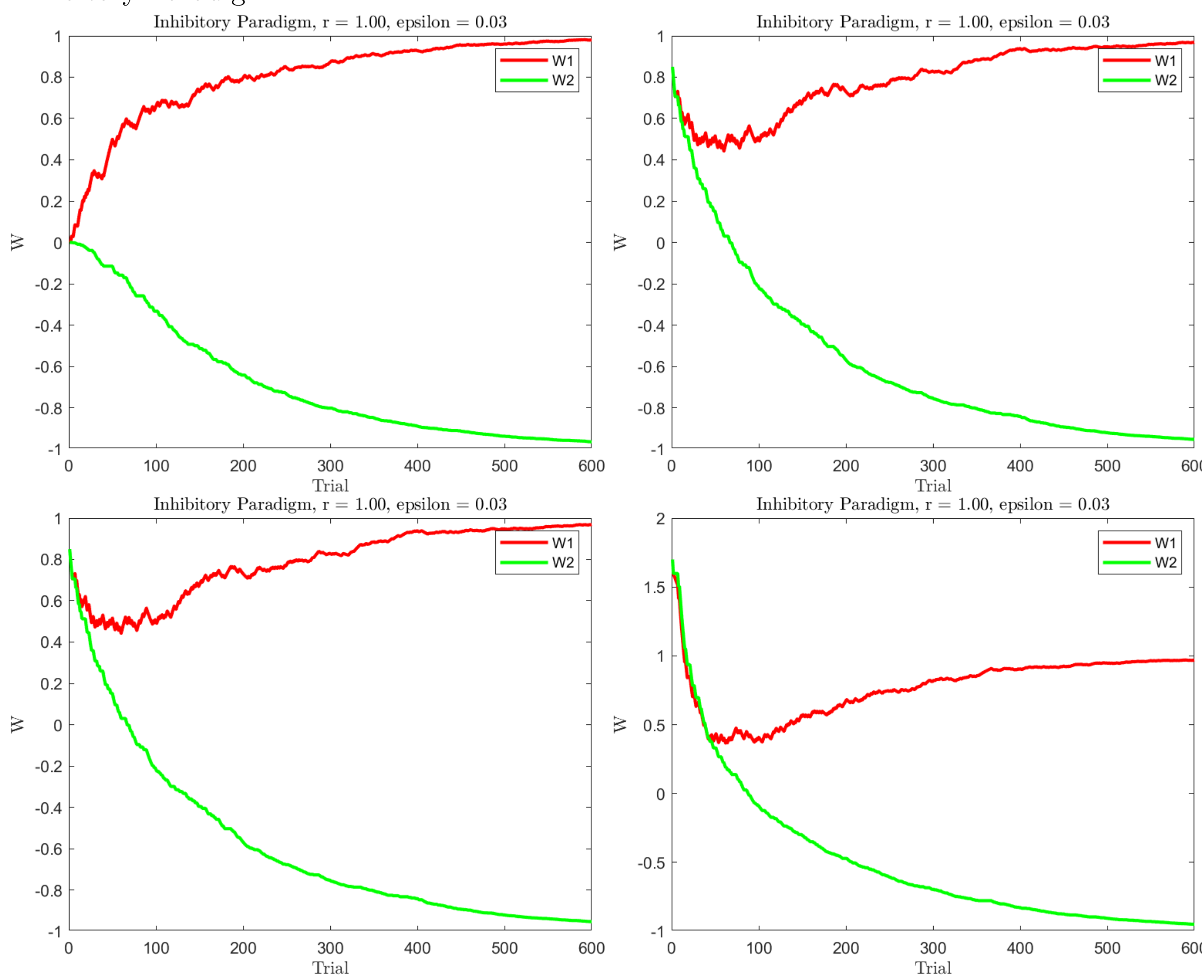


Blocking Paradigm:



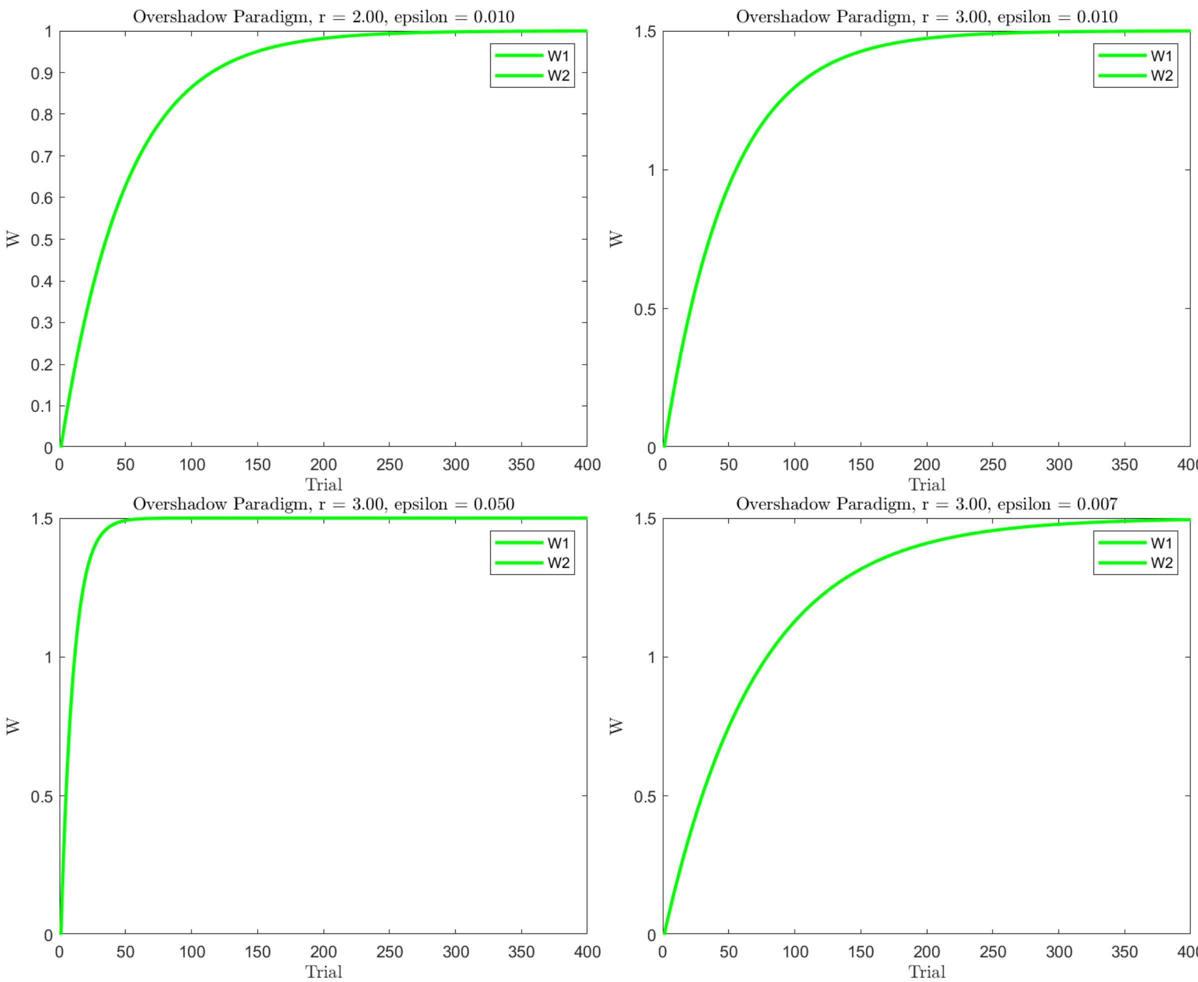


Inhibitory Paradigm:



Overshadow Paradigm:

In this paradigm first we plot when S_1 and S_2 are presented in all trials. To do this, we can consider α_1 and α_2 equal to 1. The result will be as below:

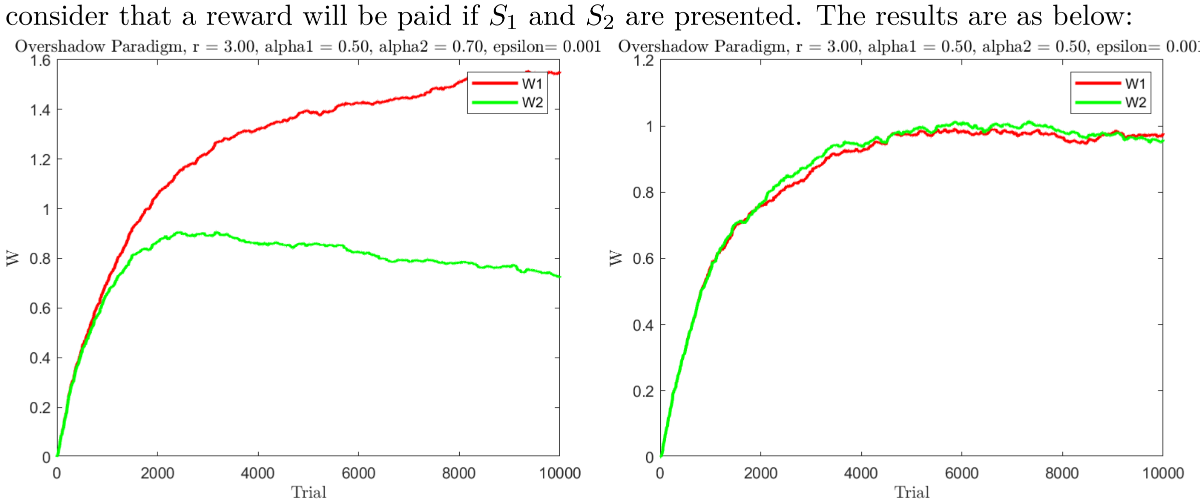


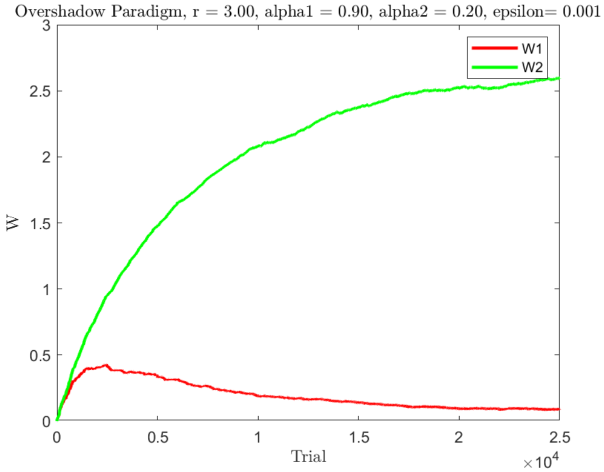
- For overshadow condition, how can one have different amount of learned value for each stimuli? The ambiguity is a form of a concept known as 'credit assignment' in reinforcement learning literature.

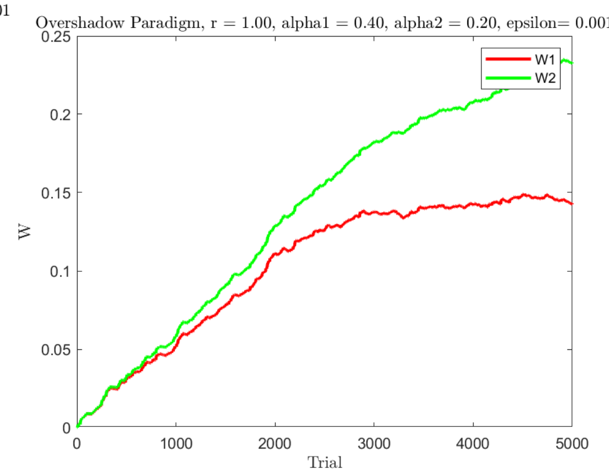
  To implement Overshadow condition by considering credit assignment, let:

$$0 < \alpha_1, \alpha_2 < 1$$

  Which implies that in $100\alpha_1$ percent of trials S_1 is presented and in $100\alpha_2$ percent of trials S_2 is presented. Note that it is not necessary to have: $\alpha_1 + \alpha_2 = 1$. In this case we should consider that a reward will be paid if S_1 and S_2 are presented. The results are as below:

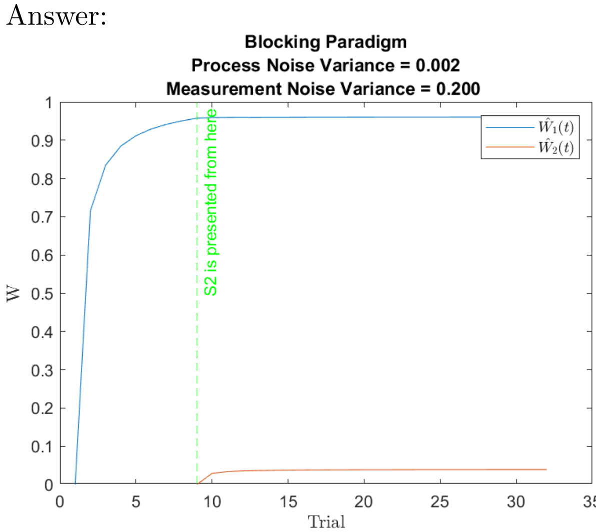


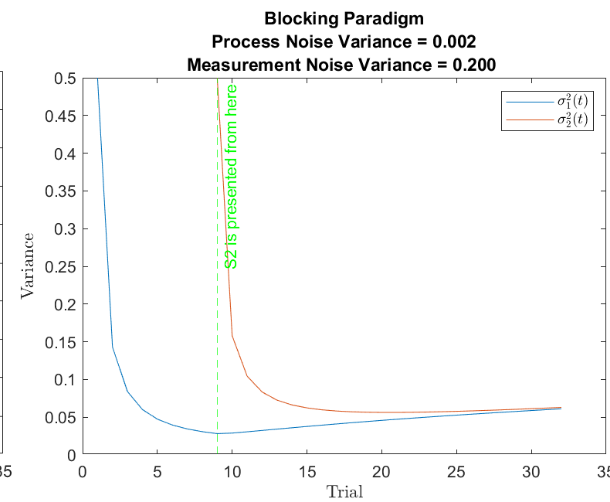


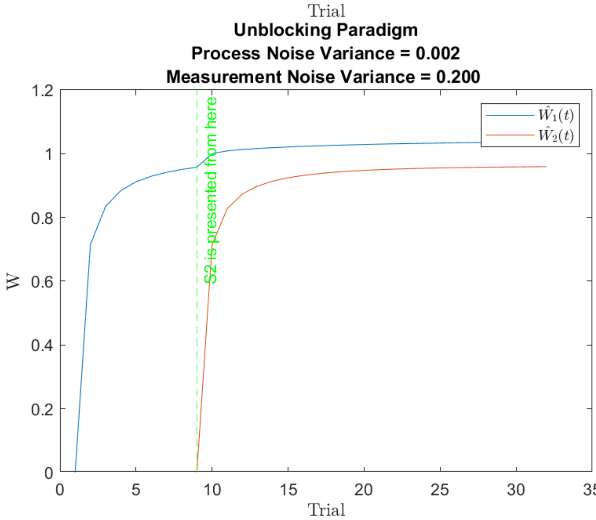


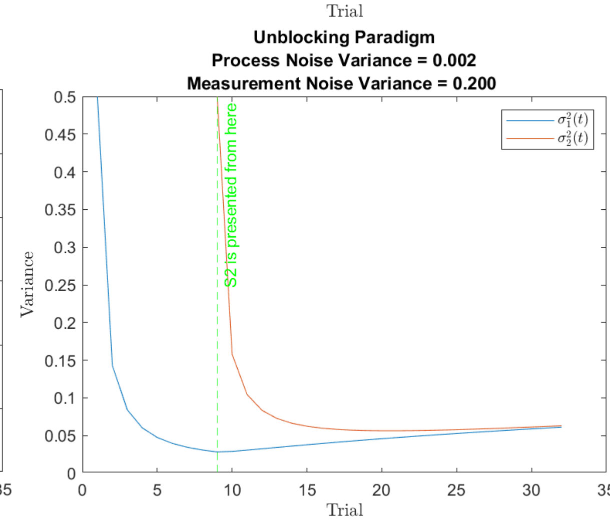
2. According to the paper, 'Uncertainty and Learning', implement Kalman filter method to explain blocking and unblocking in conditioning.

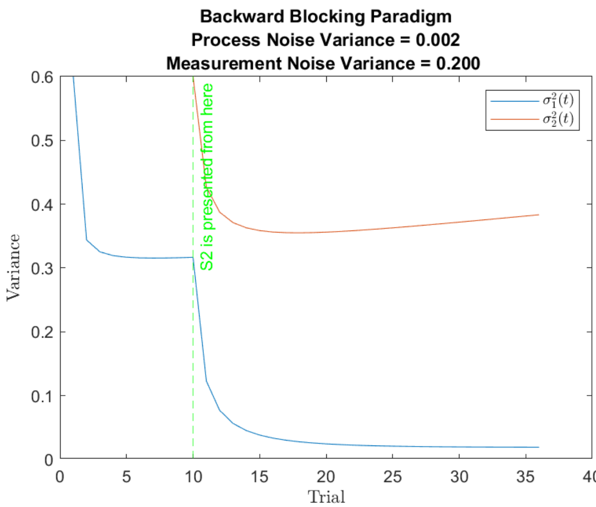
   (a) Simulate the results shown in figures 1-2 in the Dayan and Yu paper:
   Answer:

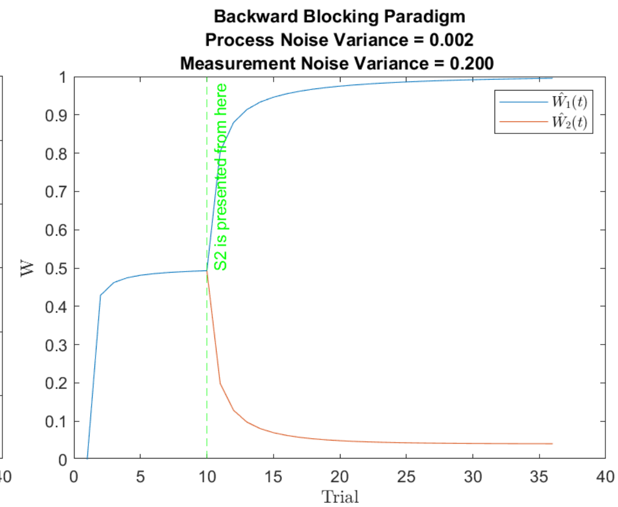


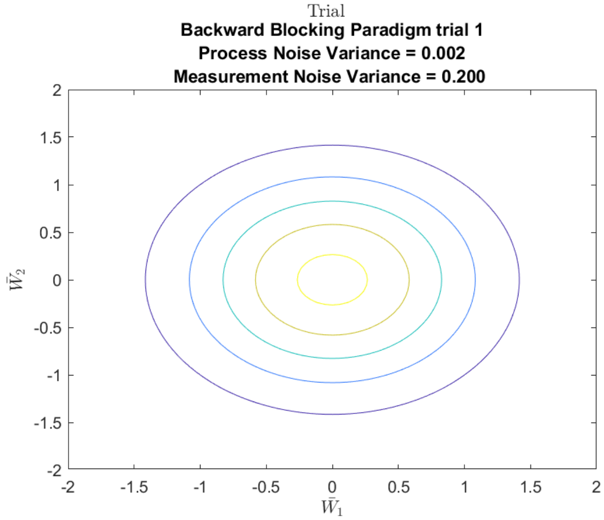


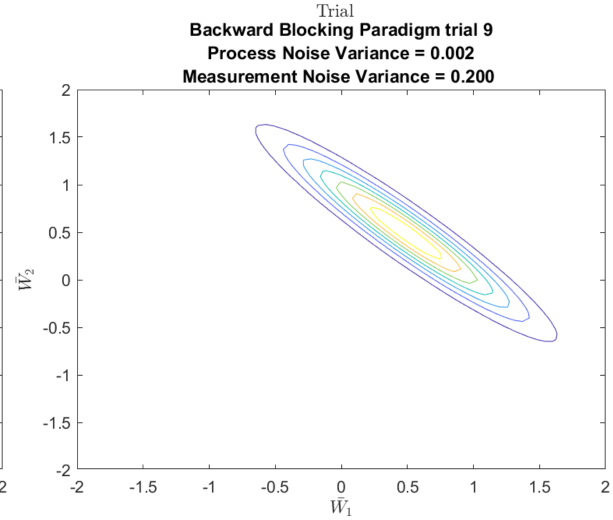


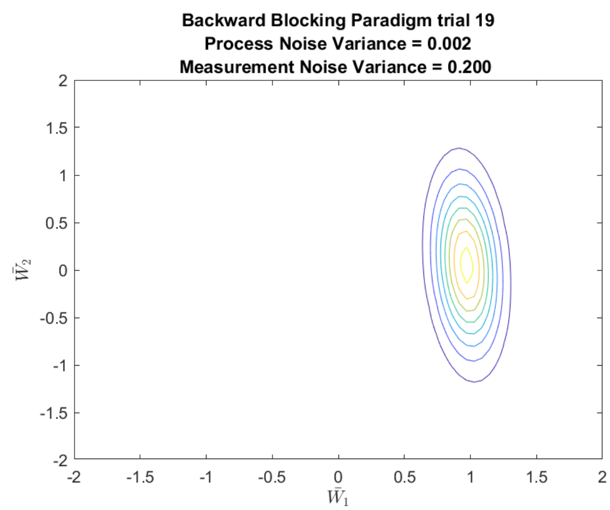






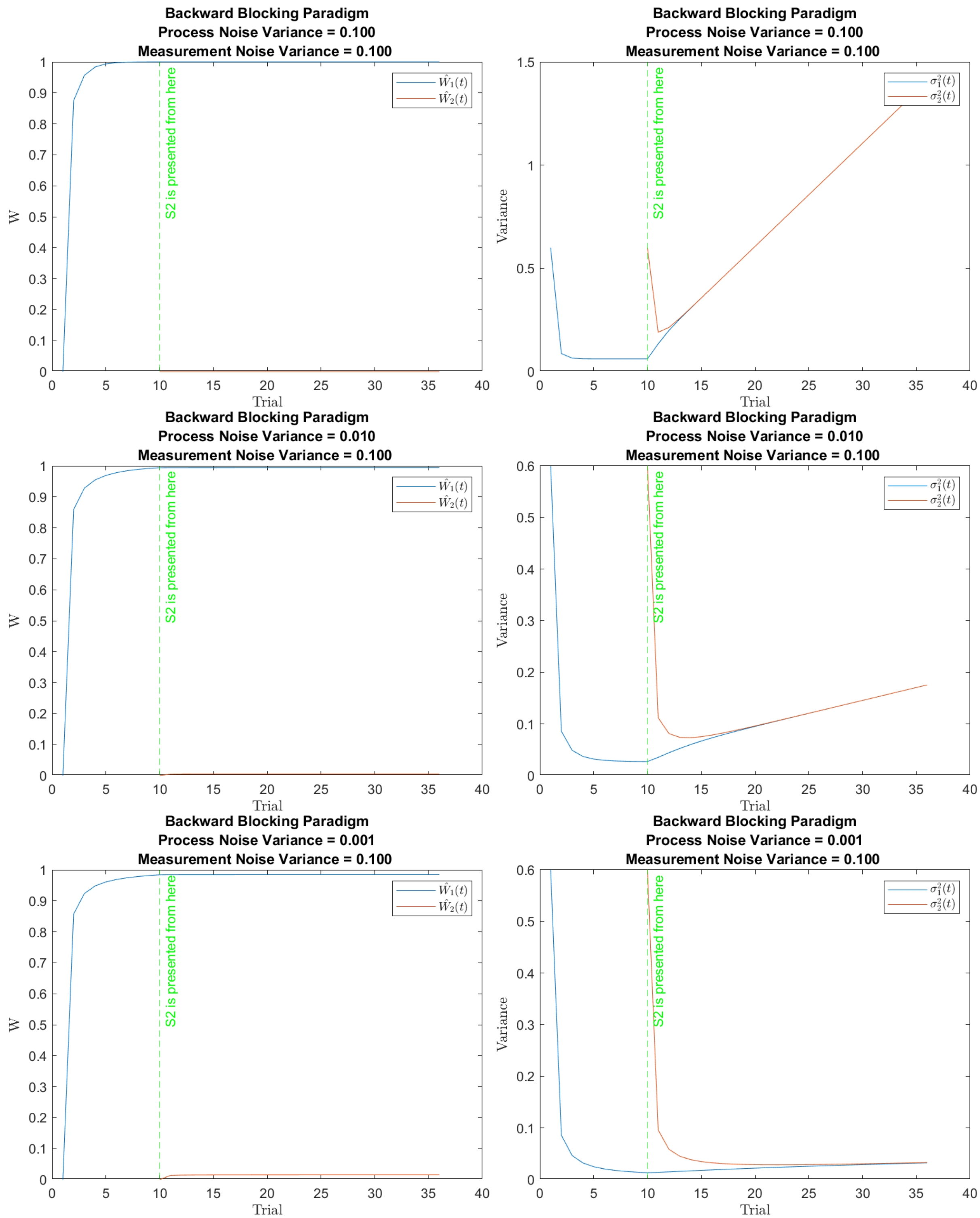


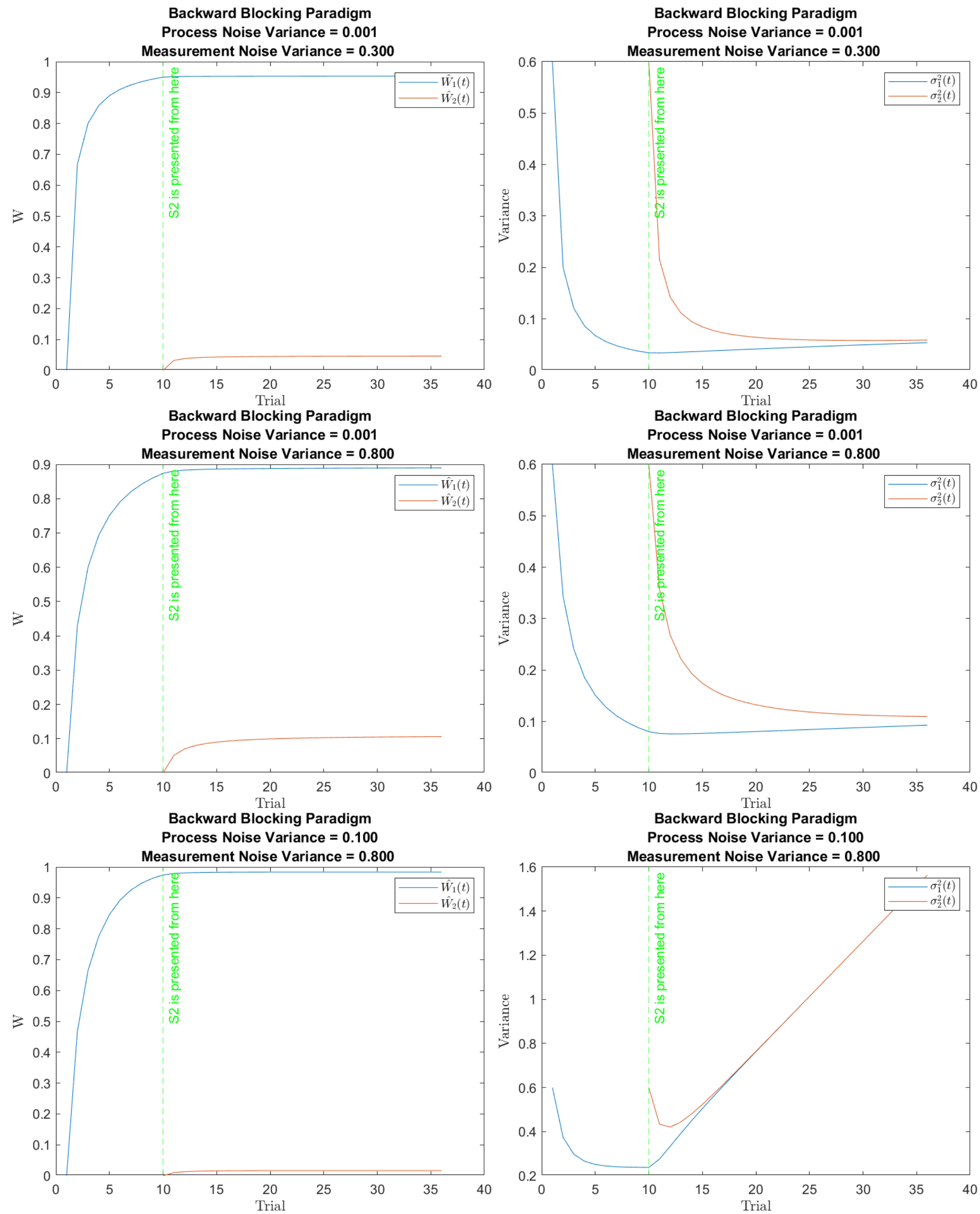




(b) How does result change by changing process noise vs measurement noise?

As you see in the following figures, by increasing process noise, uncertainty grows when both stimuli are presented. Also by increasing measurement noise, steady state values of weights won't be the results that we expected and we have some error in the results. The error increases by increasing measurement noise. If both process noise and measurement noise increase we have error in the steady state values and also we have higher uncertainty.


Backward Blocking Paradigm
Process Noise Variance = 0.100
Measurement Noise Variance = 0.100
Backward Blocking Paradigm
Process Noise Variance = 0.010
Measurement Noise Variance = 0.100
Backward Blocking Paradigm
Process Noise Variance = 0.001
Measurement Noise Variance = 0.100
W
Variance
Trial
S2 is presented from here
$\hat{W}_1(t)$
$\hat{W}_2(t)$
$\sigma_1^2(t)$
$\sigma_2^2(t)$


Backward Blocking Paradigm
Process Noise Variance = 0.001
Measurement Noise Variance = 0.300
W
Trial
$\hat{W}_1(t)$
$\hat{W}_2(t)$
S2 is presented from here
Backward Blocking Paradigm
Process Noise Variance = 0.001
Measurement Noise Variance = 0.300
Variance
Trial
$\sigma_1^2(t)$
$\sigma_2^2(t)$
S2 is presented from here
Backward Blocking Paradigm
Process Noise Variance = 0.001
Measurement Noise Variance = 0.800
W
Trial
$\hat{W}_1(t)$
$\hat{W}_2(t)$
S2 is presented from here
Backward Blocking Paradigm
Process Noise Variance = 0.001
Measurement Noise Variance = 0.800
Variance
Trial
$\sigma_1^2(t)$
$\sigma_2^2(t)$
S2 is presented from here
Backward Blocking Paradigm
Process Noise Variance = 0.100
Measurement Noise Variance = 0.800
W
Trial
$\hat{W}_1(t)$
$\hat{W}_2(t)$
S2 is presented from here
Backward Blocking Paradigm
Process Noise Variance = 0.100
Measurement Noise Variance = 0.800
Variance
Trial
$\sigma_1^2(t)$
$\sigma_2^2(t)$
S2 is presented from here

(c) What factors determine the value of Kalman gain at steady state? Can you derive an approximate relationship between steady state Kalman gain and the model parameters?

If A is identical, then the Kalman filter is a stationary system and hence, the steady-state error covariance matrix P_∞ can be obtained by solving the following algebraic Riccati equation:

$AP_\infty + P_\infty A^T - P_\infty H^T(HP_\infty H^T + R)^{-1}HP_\infty + Q = 0$

If A is identical, then it simplifies the Riccati equation to:

$2AP_\infty - P_\infty H^T(HP_\infty H^T + R)^{-1}HP_\infty + Q = 0$

The steady-state error covariance matrix is obtained by solving this algebraic Riccati equation for P_∞.

Once P_∞ is obtained, the steady-state gain K_∞ can be computed using the expression:

$K_\infty = P_\infty H^T(HP_\infty H^T + R)^{-1}$

Note that the steady-state gain provides the optimal estimate of the system state given the measurements, and it minimizes the mean square error between the true state of the system and its estimate.

$$k \to +\infty \quad \Rightarrow \quad P_k = P_{k-1} = P_\infty$$

P_∞ depends on both process noise and measurement noise. So Gain of Kalman depends on these two parameters.

Another approach is:

Interpretation of Innovation/Gain

- Model
- Observation

$$\underline{\hat{X}}_k = \underline{\hat{X}}_k^- + G_k(\underline{Z}_k - H_k\underline{\hat{X}}_k^-) = F_k\underline{\hat{X}}_{k-1} + \left(P_k^- H_k^T(H_kP_k^-H_k^T + R_k)^{-1}\right)I_k \cong F_k\underline{\hat{X}}_{k-1} + \left(P_kH_kR_k^{-1}\right)I_k$$

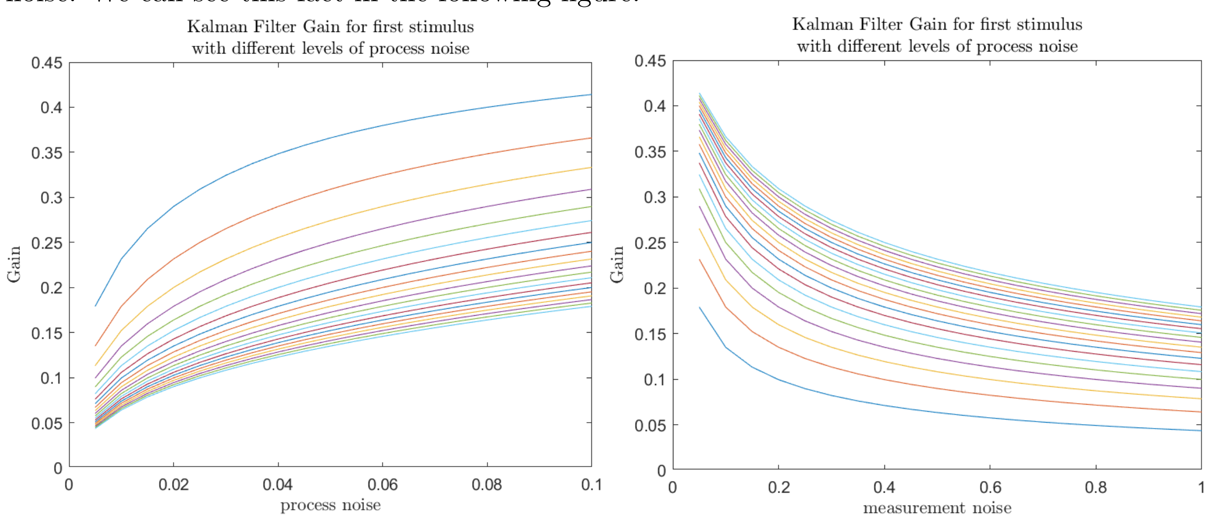
- Particular case1: scaler case

$$\hat{X}_k = \hat{X}_k^- + G_k(Z_k - H_k\hat{X}_k^-) = F_k\hat{X}_{k-1} + \frac{P_k^-H_k}{H_kP_k^-H_k + R_k}I_k \cong F_k\hat{X}_{k-1} + \frac{P_kH_k}{R_k}I_k \cong F_k\hat{X}_{k-1} + \frac{\sigma_w^2}{\sigma_v^2}H_kI_k$$

- Particular case2

$$\begin{cases} R_k = \sigma_v^2 I \\ P_k = \sigma_e^2 I \cong \sigma_w^2 I \end{cases} \Rightarrow \quad \underline{\hat{X}}_k \cong F_k\underline{\hat{X}}_{k-1} + \frac{\sigma_w^2}{\sigma_v^2}H_k\underline{I}_k$$

Which implies that kalman filter Gain in infinity depends on process noise and measurement noise. We can see this fact in the following figure:



So:

What really matters about the gain in steady state is to build an intuition. We can easily understand the following rules of thumb: 1. The more noisy our measurement is, the less precise it is. Hence, the innovation that represents the bias in our measurement may not be real innovation but rather an artifact of the measurement noise. Noise or uncertainty is directly captured by variance. Thus the larger the measurement variance noise, the lower the Kalman gain should be.

2. The more noisy our process state is, the more important the innovation should be taken into account. Hence, the larger the process state variance, the larger the Kalman gain should be.

Summarizing these two intuitions, we would expect the Kalman gain to be:

$$G_\infty \propto \frac{\text{process noise}}{\text{measurement noise}}$$

(d) Does the change in uncertainty depend on the errors made in each trial or changes in the learning context?

By the following equations:

$$\Sigma(t+1)^- = A\Sigma(t)A^T + W = \Sigma(t) + W$$

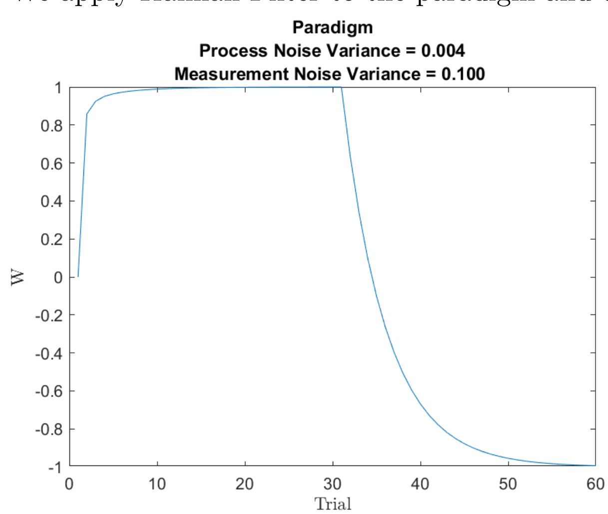
$$G = \Sigma(t+1)^- C^T (C\Sigma(t+1)^- C^T + \tau^2)^{-1}$$

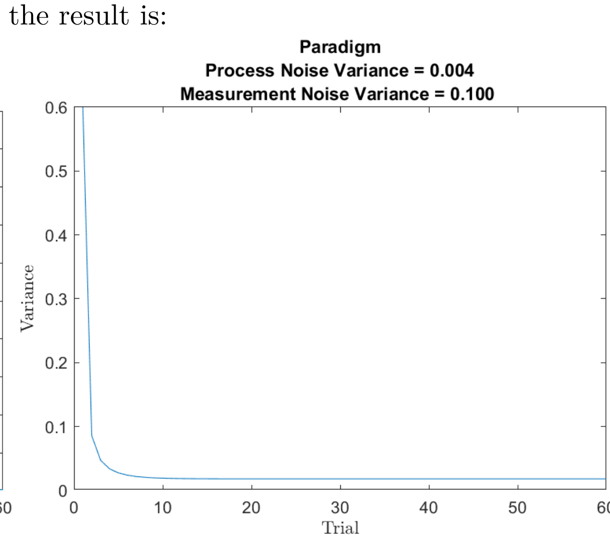
$$\Sigma(t+1) = \Sigma(t+1)^- - GC\Sigma(t+1)^-$$

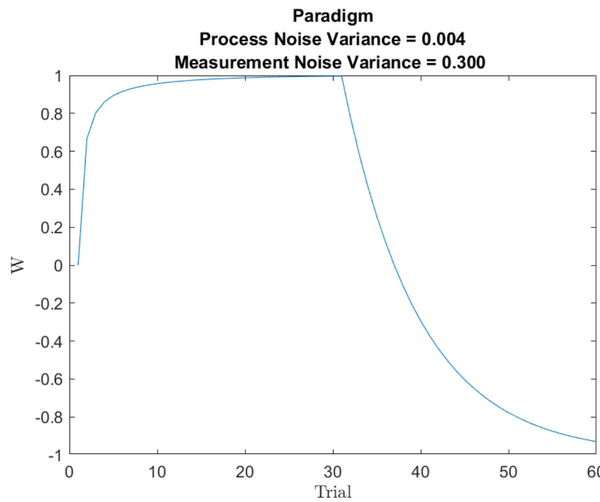
we can see that the reward is not needed to calculate $\sum$ in each step. So the uncertainty depends on changes in the learning context and does not depend on errors made in each trial.

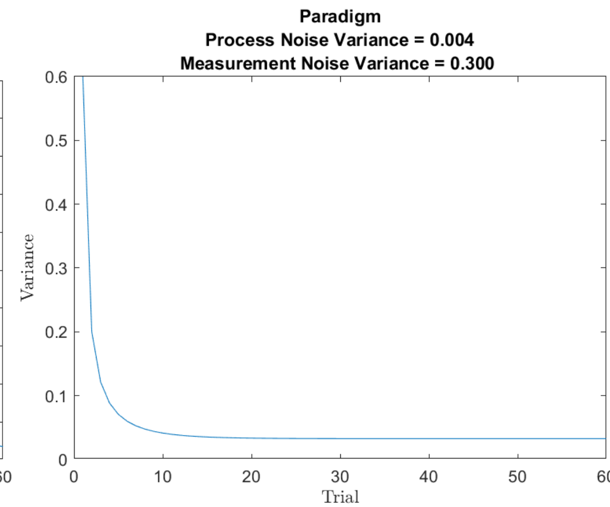
(e) Using this paradigm if we learn S1-> 'r' first and then S1->'-r'. how does learning the second context compares to the first context? What happens to the learning rate?

We apply Kalman Filter to the paradigm and the result is:









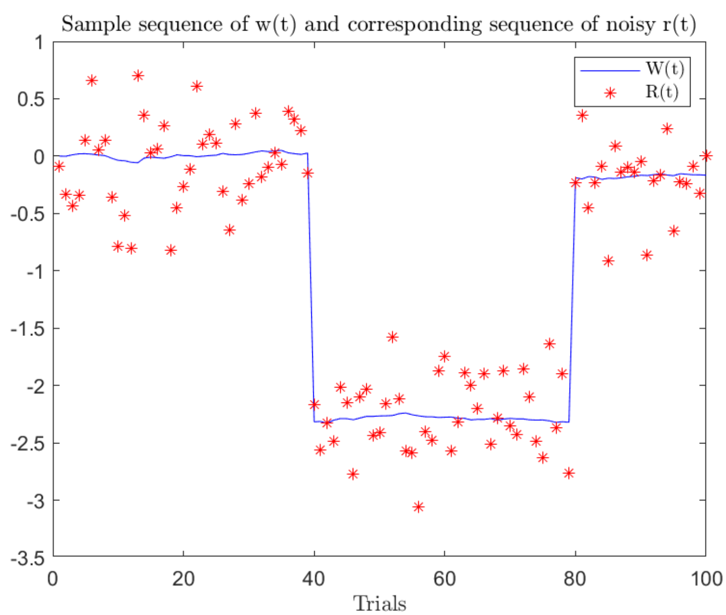
As you see in the figures the paradigm after pre-train by S1-> 'r needs more trials to adapt with S1->'-r' and thus learning rate has been decreased.

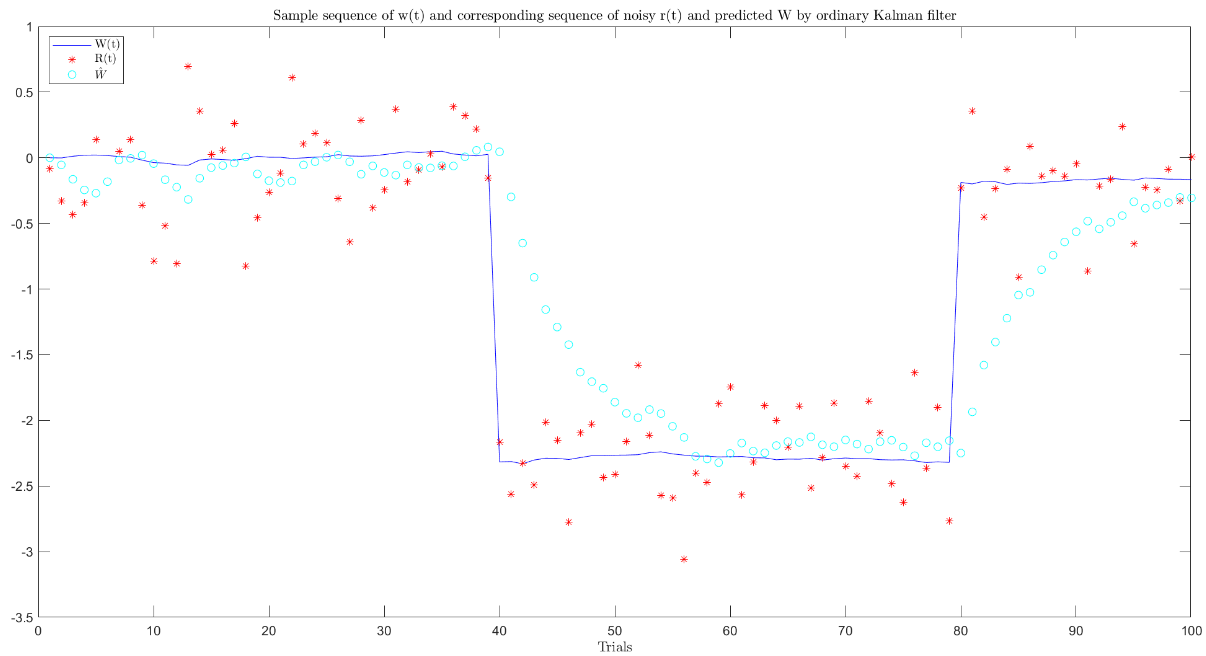
3. The uncertainty modelled by Kalman filter is referred to as 'known uncertainty'. This is the uncertainty about the value of the stimulus for which the agent has some estimate. However there are times when we don't even know how much we do not know about value of a stimulus. This is referred to as 'ambiguity' or 'unknown uncertainty'.

   If we do not account for the ambiguity, then our estimate of uncertainty gets smaller over time and we cannot learn about new changes in the environment. To account for this Dayan and Yu made their model sensitive to the error magnitude. Large errors served to reset the uncertainty about the values to promote learning according to thresholding this value:

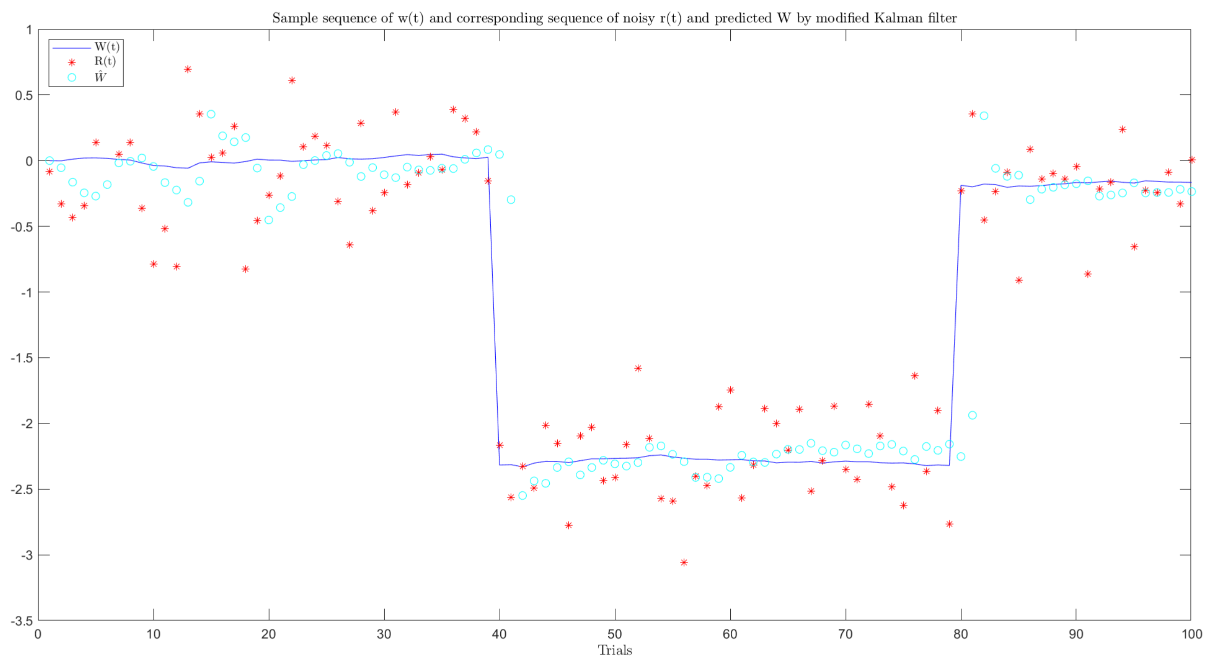
$$\beta(t) = (r(t) - x(t) \cdot \hat{\mathbf{w}}(t))^2 / (\mathbf{x}(t)^T \Sigma(t) \mathbf{x}(t) + \tau^2)$$

   - Simulate the results shown in Figure 3 of the paper.
     We show the results for our specified parameters:

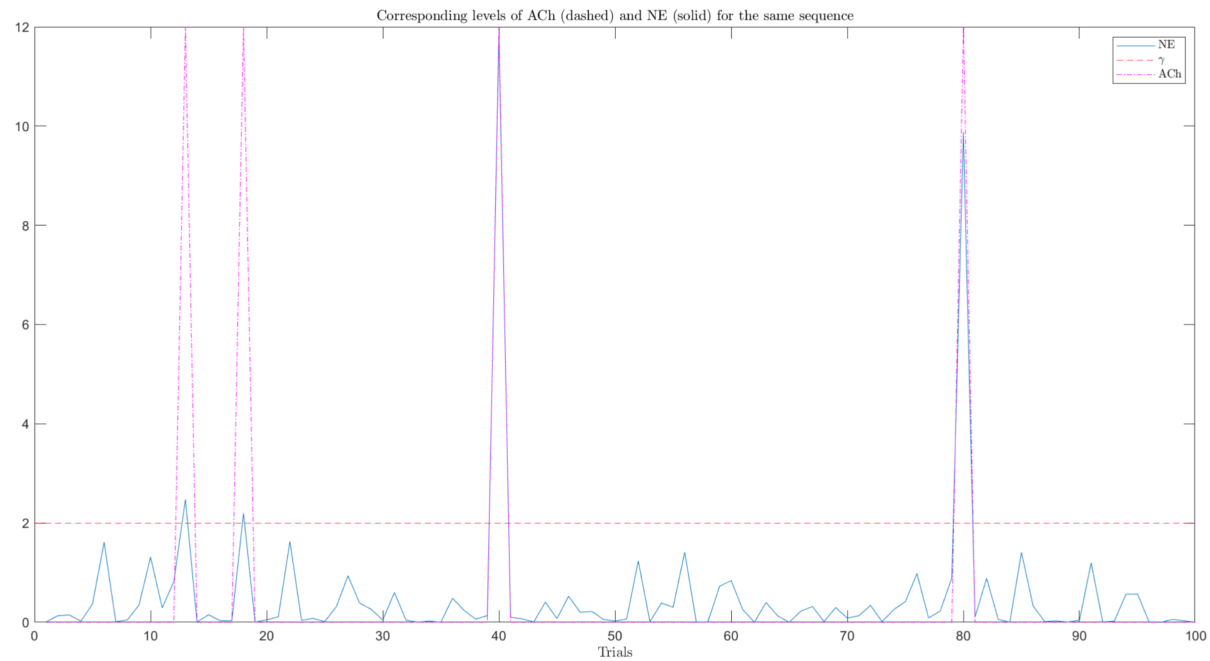




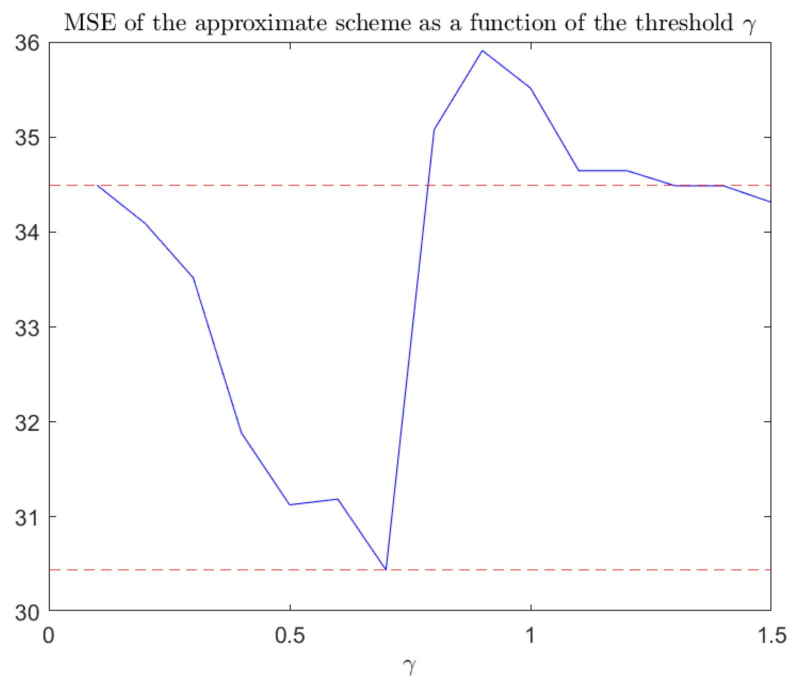
Sample sequence of w(t) and corresponding sequence of noisy r(t) and predicted W by ordinary Kalman filter
W(t)
R(t)
$\hat{W}$
Trials



Sample sequence of w(t) and corresponding sequence of noisy r(t) and predicted W by modified Kalman filter
W(t)
R(t)
$\hat{W}$
Trials



Now we plot MSE per different amounts of Threshold:



Note that by running the codes in my Zip file you will get different plots from the above plots because of unknown uncertainty. Also you can change the values of process and measurement noise or $\hat{\sigma}$(here this parameter is 50) which is added to σ when β is higher than threshold to change the results.

By changing the parameters you can see high MSEs which have been quartered or even lower when you use modified Kalman Filter.