Applied Estimation(EL2320) Lab 1 EKF

Jingsheng Chen
jinch@kth.se

Part 1 Preparatory Questions

Linear Kalman Filter:

1. What is the difference between a 'control' $\mathbf{u}_t$, a 'measurement' $\mathbf{z}_t$ and the state $\mathbf{x}_t$? Give examples of each?
   control u is the input of the system, measurement z is observations of the state of the system , and state x contains current information about the system.
   For example, if we want to control a car, u iscontrol signal of gas and steering wheelm, z is Odometer and GPS of the car, and x contain Speed, heading informations.

2. Can the uncertainty in the belief increase during an update? Why (or not)?
   The uncertainty will decrease. During update Σt= ¯Σt−KtCt¯Σt, so it will decrease.

3. During update what is it that decides the weighing between measurements and belief?

$$\mu_t = (I - W)\bar{\mu}_\mathbf{t} + W\mu_\mathbf{z}$$
$$W = \Sigma_t C_t^T Q_t^{-1} C_t$$
$$\text{and } C_t\mu_\mathbf{z} = (\mathbf{z}_t - \bar{\mathbf{z}}_\mathbf{t} + C_t\bar{\mu}_\mathbf{t})$$

   The relation between Σt(uncertainty) and Qt (measurement error) after the update is done decides the weighting between the measurements and the belief.

4. What would be the result of using a too large a covaraince (Q matrix) for the measurement model?
   Estimates would be pessimistic (conservative) and convergence would be slower.

5. What would give the measurements an increased effect on the updated state estimate?
   A larger Kalman gain, the small covariance matrix for the measurement error model.

6. What happens to the belief uncertainty during prediction? How can you show that?
   It will increase. According to ¯Σt=Qt+AtΣt−1AT, Each predict will introduce noise, and uncertainty is often growing.

7. How can we say that the Kalman filter is the optimal and minimum least square error estimator in the case of independent Gaussian noise and Gaus-sian priori distribution? (Just describe the reasoning not a formal proof.)

   The Kalman filter gives the true posterior distribution for a linear Gaussian system. So, suppose μbetter for the state has lower expected square error than the Gaussian mean μ.

$$\int_{-\infty}^{\infty}(x-\mu_{better})^2 G(x,\mu,\Sigma)dx$$
$$\Leftrightarrow \int_{-\infty}^{\infty}(x^2-2\mu_{better}x+\mu_{better}^2)G(x,\mu,\Sigma)dx$$
$$\Leftrightarrow \Sigma+\mu^T\mu-\mu^T\mu_{better}-\mu_{better}^T\mu+\mu_{better}^T\mu_{better}$$

Minimizing the equation by differentiating and we have:

$$-2\mu + 2\mu_{better} = 0$$
$$\mu_{better} = \mu$$

So, the Kalman filter is the optimal and minimum least square error estimator.

8.In the case of Gaussian white noise and Gaussian priori distribution, is the Kalman Filter a MLE and/or MAP estimator?
It's a MLE estimator.

Extended Kalman Filter:

9. How does the extended Kalman filter relate to the Kalman filter?
   The extended Kalman filter applies the Kalman filter to the linearized nonlinear system. Furthermore, Atxt−1 + Btut is replaced by g(ut, xt−1), At by Gt and Ct by Ht.

10. Is the EKF guaranteed to converge to a consistent solution?
    NO, the mean value of the nonlinear function is approximately linear at the estimation point, and the update depends on this linearization. If the linearization accuracy is poor, the convergence result will be affected.

11. If our filter seems to diverge often can we change any parameter to try and reduce this?
    Yes. If divergence occurs on update, we can change modeled uncertainties Q and R and increase the relative size of the measurement covariance Q. If the divergence is due to bad data association, we can change the matching threshold.

Localization:

12. If a robot is completely unsure of its location and measures the range r to a know landmark with Gaussain noise what does its posterior belief of its location $p(x, y, \theta|r)$ look like? So a formula is not needed but describe it at least.
    It will have a uniform distribution over the heading θ between−π and π. The position will be a sort of donut/ring. So a Gaussian on ρ in radial coordinate with uniform distribution on the angle φ
    So, x = ρ cos φ, y = ρ sin φ, and $e^{-\frac{(\rho-r)^2}{2\sigma_r^2}}$

13. If the above measurement also included a bearing how would the posterior look?
    The same except that the heading and angle around the ring would be Gaussian with a completely correlated covariance.

$$e^{-\left[\frac{(\rho-r)^2}{2\sigma_r^2}+\frac{(b-\phi+\theta)^2}{2\phi_b^2}\right]}$$

14. If the robot moves with relatively good motion estimation (prediction error is small) but a large initial uncertainty in heading θ how will the posterior look after traveling a long distance without seeing any features?
    It will look like a cresent/arc/C shape. The heading θ will be correlated with position along the arc.

15. If the above robot then sees a point feature and measures range and bearing to it how might the EKF update go wrong?
    The new feature will prevent the update from following the original trajectory, making it easy to diverge and cause errors.

Part 2 Matlab Excercises

2.1 Warm up problem with Standard Kalman Filter

Question 1

According to the equation $\boldsymbol{x_{k+1}}=\boldsymbol{Ax_k}+\mathrm{B}\boldsymbol{u_k}+\boldsymbol{\varepsilon_k}$, ε_k should have the same dimension as the result of Ax_k +Bu_k, which is two dimension(2*1 matrix). And δ_k should have the same dimension as the measurement z_k, which is one dimension.

To define a uniquely white Gaussian, we need to define mean value **μ** and a covariance σ^2. The covariance would be a 2*2 matrix Σ.

Question 2

Variables	Meaning
x	The actual state of system
xhat	The estimate state of system
P	Estimate error covariance matrix.
G	Identity matrix for process noise
D	Identity matrix for measurement noise
Q	Covariance matrix of noise in the measurement model
R	Covariance matrix of process noise
WStdP	The noise weight of position in the simulation
WstdV	The noise weight of velocity in the simulation
vStd	The noise weight of measurement in the simulation
u	Control signal(The acceleration in the system)
PP	The storing matrix for the position during the KF process in prediction

Question 3

The normal image, the image with the process noise increased by 100 times and the image with the measurement noise increased by 100 times are shown in the following figure respectively. I think the former will increase the Kalman gain, while the latter will decrease the Kalman gain.

By comparison, we can find that when the process noise is increased, the Kalman gain of the system becomes larger. This is because the system is more dependent on the observed data rather than the predicted data. At the same time, the prediction speed has become very unstable. I think this is because a relatively high Kalman gain will cause a large amount of change in the prediction each time it is updated. When we increase the measurement noise, the system will rely more on the data predicted by the model, so the Kalman gain will become smaller. However, due to the lack of limited observational data to update, the system has a larger position error. Compared with Figures 1 and 2, the position cannot always converge to near 0, but is maintained at about 0.4. This is in line with my prediction.

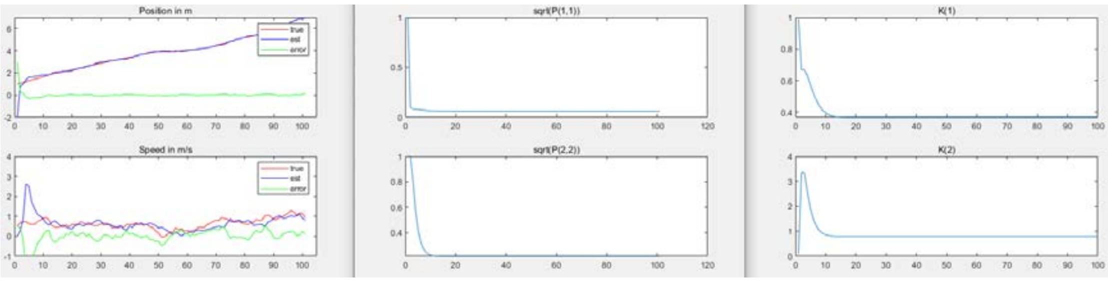

Figure 1: normal system

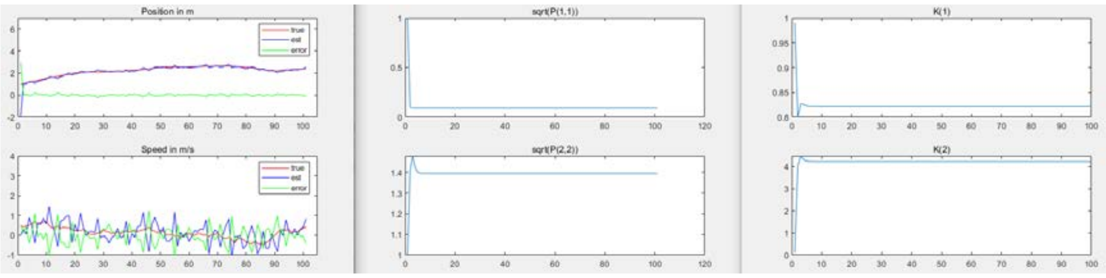

Figure 2: system with 100 times process noise

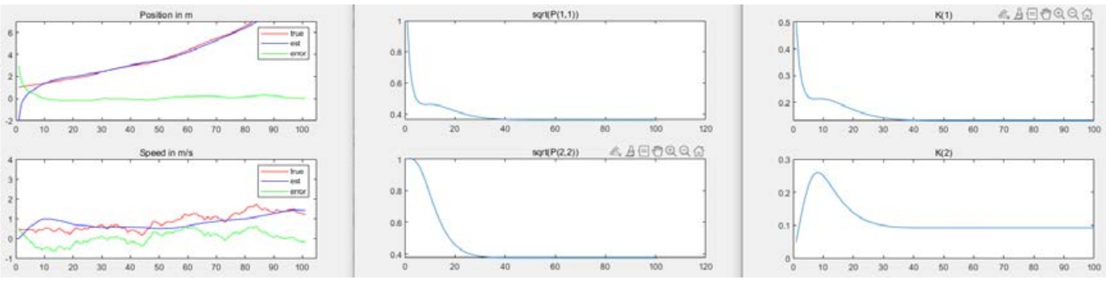

Figure 3: system with 100 times measurement noise

Question 4

From Figure 4, we can see that when P becomes very large, the system will rely heavily on observations at the beginning, so the Kalman gain will become very large at the beginning, which makes the system converge very quickly. On the contrary, when P is very small (Figure 5), the system will believe the predicted value more, the value of Kalman gain will decrease, and the system will slow down to converge to near the actual value, but the error will eventually be small.

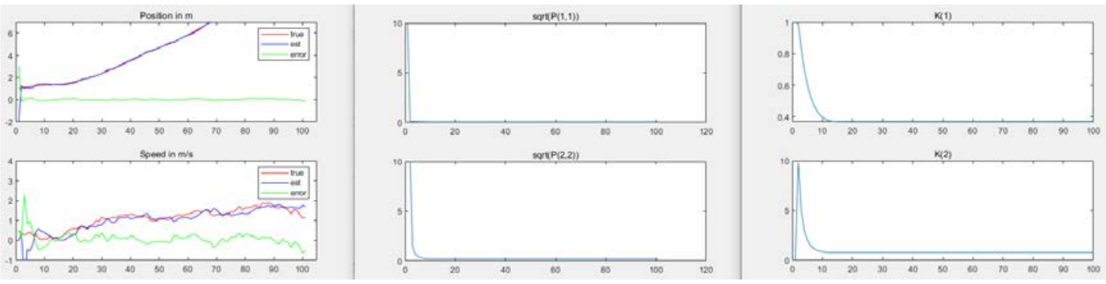

Figure 4: Increase P to 100

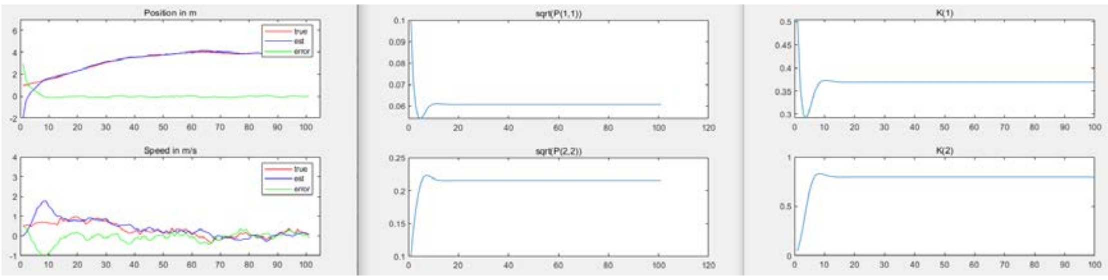

Figure 5: decrease P to 0.01

At first I increased the initial value of xhat by 100 times, but this did not bring any change, so I increased it to 1000 times, which is [1000, 500], as can be seen from Figure 6, because the initial value The deviation is large, so at first the Kalman gain is also large, and it takes a relatively long time for the system to converge to near the actual value.

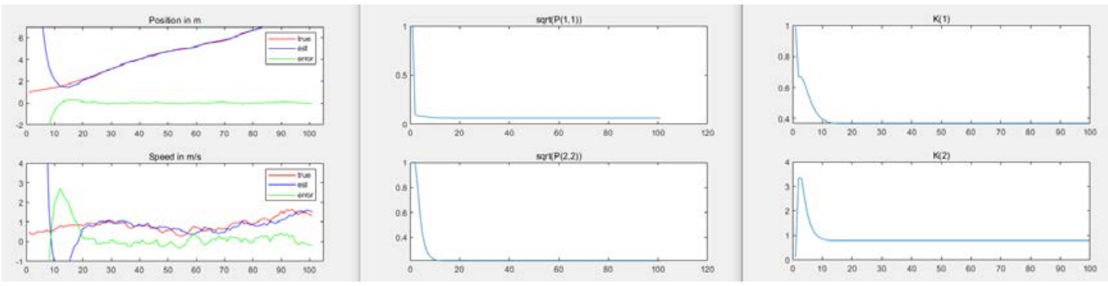

Figure 6: Increase Xhat to [1000, 500]

2.2 Main problem: EKF Localization

Question 5

The first equation of (2) is responsible for both update and prediction.
And the first equation of (3) is responsible for update, and second equation of (3) are responsible for prediction.

Question 6

Yes I think it's a valid assumption. Since each measurement is only related to the state of each time, and the noise is Gaussian, it has nothing to do with the last or next measurement. So it's independent.

Question 7

The bounde of $\delta_m\ is\ [0,1]$ as it is a probability. Increase δ_m will cause λ_m to increase. When the measured value is unreliable, λ_m should be reduced so that more outliers are rejected. And when the measurements we get from the map are reliable, there should be fewer outliers, so we can increase by λ_m.

Question 8

In sequential update, since the value of the next measurement is always estimated on the basis of the previous measurement update, the first impact always exists. If there is an error in the first measurement or there is a lot of noise, it may cause the covariance matrix to decrease, which in turn affects St, j and Mahalanobis distance, resulting in unreasonable rejection of outliers.

Question 9

There are many zero matrices in matrix multiplication. We can use the symmetry in the covariance and uncertainty matrix to reduce the computational complexity of the algorithm

Question 10

The dimension of $\bar{\boldsymbol{v}}_{\boldsymbol{t}}$ is 2nx1 and in the sequential update is 2x1.
The dimension of $\bar{\boldsymbol{H}}_{\boldsymbol{t}}$ is 2nx3 and in the sequential update is 2x3.
That means the batch update will use all features to decide outliners, which brings more calculations.

2.3 Simulating data sets

2.3.1 Dataset1

In the result of dataset 1 we can see that on all dimension the mean absolute error are less than 0.01(m, rad), which is satisfied the requirement.

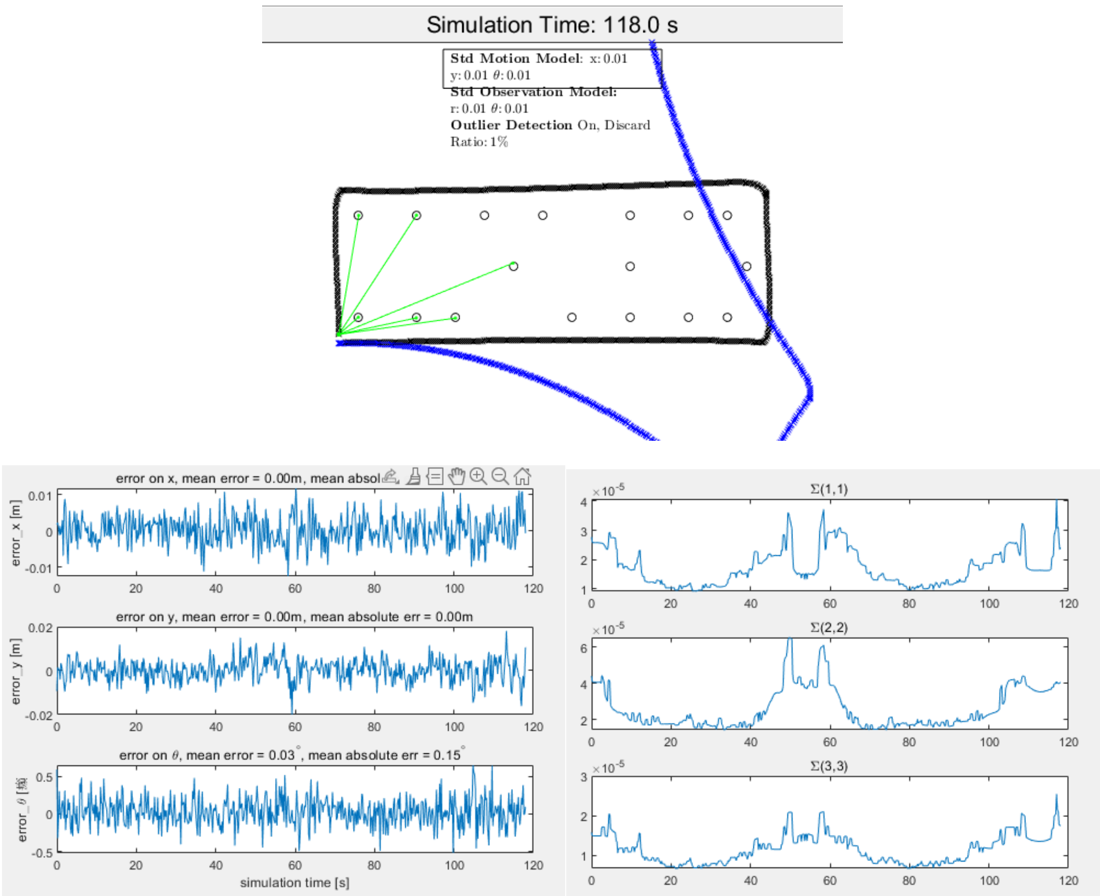

Figure 7: Result of dataset1

2.3.2 Dataset2

In the result of dataset 2 we can see that on all dimension the mean absolute error are less than 0.06(m, rad), which is satisfied the requirement.

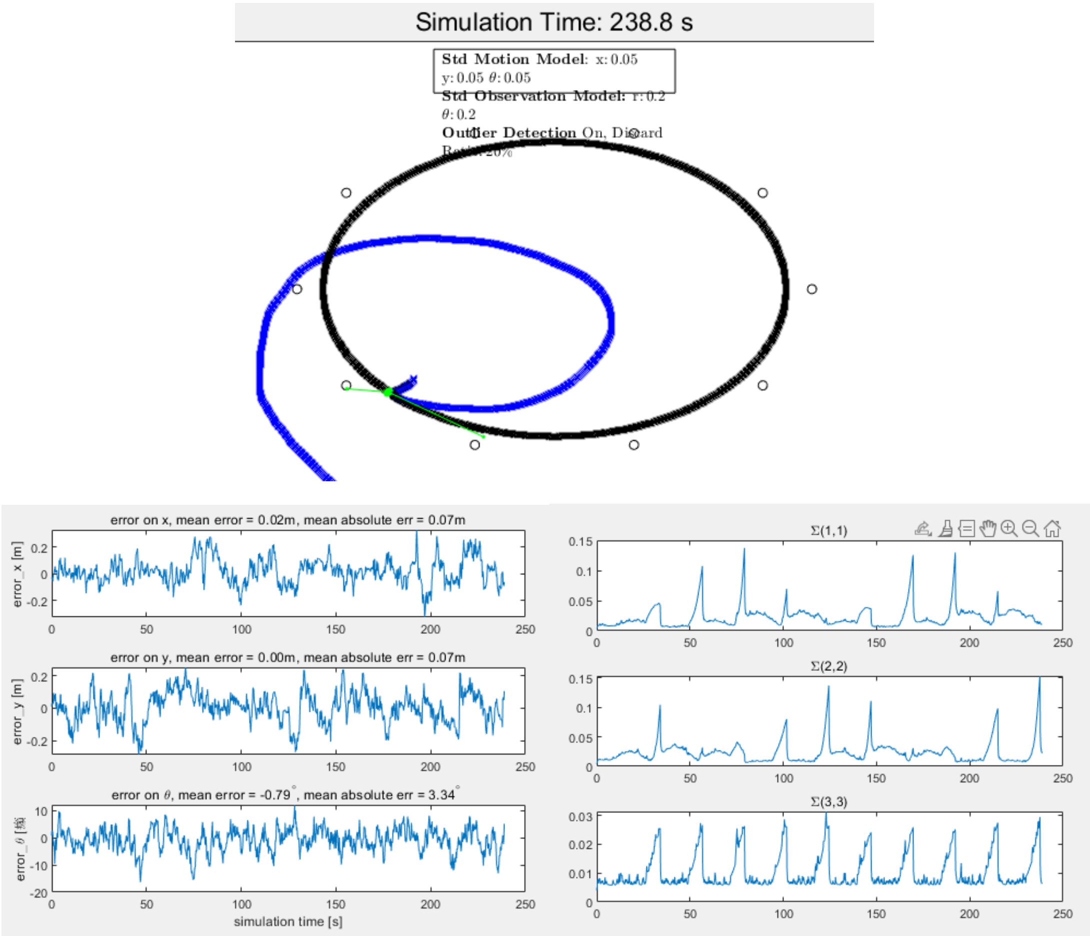

Figure 8: Result of dataset2

2.3.3 Dataset3

In the result of dataset 3 we can see that on all dimension the mean absolute error are less than 0.01(m, rad), which is satisfied the requirement.

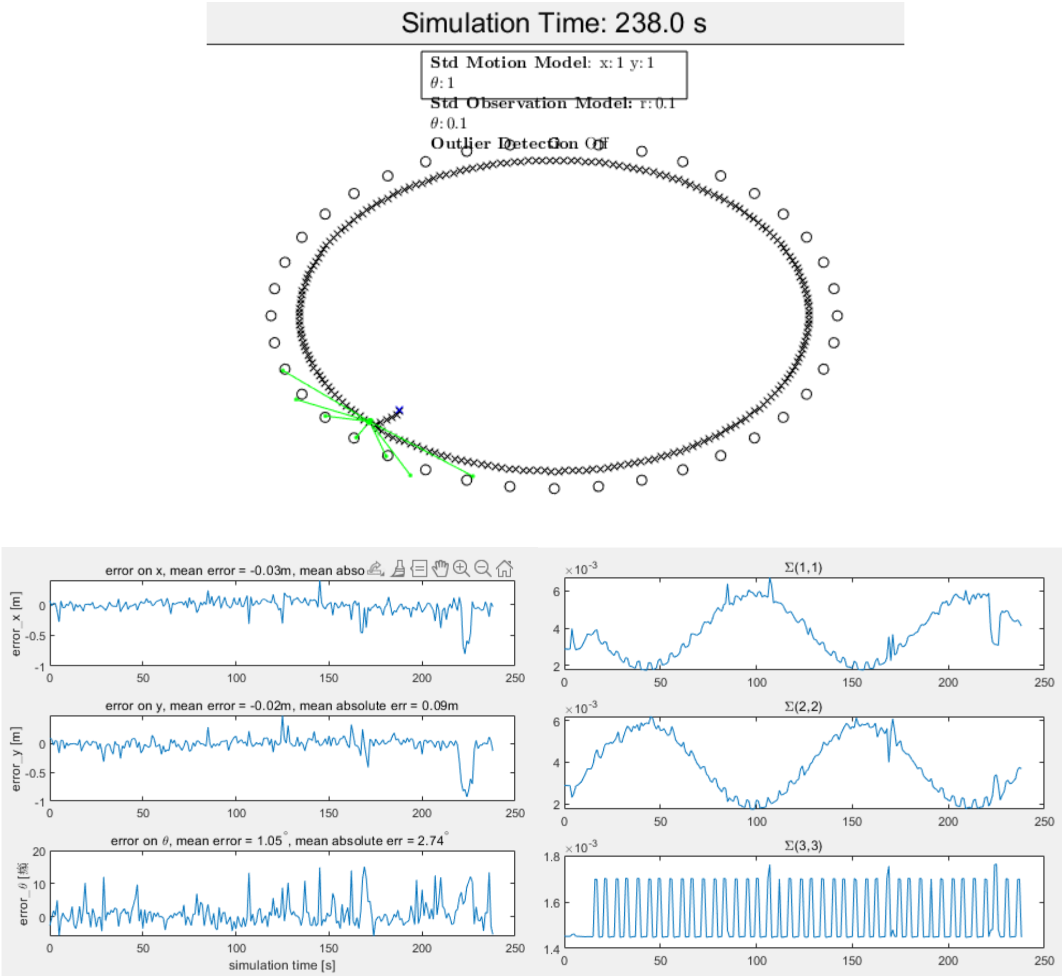

Figure 9: Result of dataset3