Lab 2: PF
EL2320 Applied Estimation

Ning Wang
ningwa@kth.se
980902-2453

December 3, 2021

Part I

1.1 What are the particles of the particle filter?

Particles are the samples of the posterior distribution.
The particles are a representation of a probability density function.
The particles are samples from a probability density function over the state.

1.2 What are importance weights, target distribution, and proposal distribution and what is the relation between them?

The importance weights are the probability of measuring a particle when applying the measurement model.

$$w(t) = p(\{z(t)\}|x(t))$$

The target distribution is the true posterior $p(x(t)|\{z(t)\}, \{u(t)\})$, that is to be approximated with particles.
The proposal distribution is the sampled distribution $p(x(t)|\{u(t)\}, \{z(t-1)\})$.
The relationship can be denoted as:

$$importance\ weight = target\ distribution/proposal\ distribution$$

1.3 What is the cause of particle deprivation and what is the danger?

Particle deprivation means having no particles at the place where the state has non-zero probability of being. It is caused by random sampling and is more likely to occur when a smaller number of samples are applied to the high likelihood region. When particle deprivation happens, we can run into the danger of getting a entirely wrong estimation.

1.4 Why do we resample instead of simply maintaining a weight for each particle always.

Maintaining a weight for each particle will approximate the posterior, but most of the particles will end up in low probability posterior regions. There are highly unlikely particles that need to be removed, and the new state might not be included in the present particle set.
The diffusion stage will not be able to populate the likely regions with dense samples unless we make copies of the good particles and eliminate the bad ones.

1.5 Give some examples of the situations which the average of the particle set is not a good representation of the particle set.

If the distribution has multiple peaks, the average can lie between different peaks, in a low probability region. A local peak far from the ground truth can impact heavily on the average. In a ring-like distribution, the average can lie in the low probability central area. Also for highly skewed distributions.

1.6 How can we make inferences about states that lie between particles.

By using density extraction, we can convert a particle set into continuous density. Another method is to use kernel density estimation.

1.7 How can sample variance cause problems and what are two remedies?

The two sampling steps in the particle filter result in the particle distribution being close but not identical to the posterior. The difference is called sample variance. It can cause particle deprivation. Remedies:
Low variance sampling;
Stratified sampling, to do a separate resample of particles of each peak.

1.8 For robot localization for a given quality of posterior approximation, how are the pose uncertainty (spread of the true posteriori) and number of particles we chose to use related.

With a higher pose uncertainty, we need a larger number of particles to cover the region and represent the true posterior.

Part II

2.1 What are the advantages/drawbacks of using (6) compared to (8)? Motivate.

Advantage: Using the same value for θ, dx and dy at each time step, and it is in fact only 2-dimensional, which will make computation faster.
Drawback: It can only represent motions with constant linear and angular velocity well, not able to predict more complex motions. When dealing with noise, it will result in a wave-like pattern.

2.2 What types of circular motions can we model using (9)? What are the limitations(what do we need to know/fix in advance)?

We can use (9) to model uniform circular motion with constant magnitude of velocity and constant rotational velocity. The magnitude of velocity v_0 and the rotational velocity ω_0 need to be fixed.

2.3 What is the purpose of keeping the constant part in the denominator of (10)?

The purpose is to scale the likelihood and make the integral over the whole distribution equal to 1. Only in this way can it be a probability function.

2.4 How many random numbers do you need to generate for the Multinomial re-sampling method? How many do you need for the Systematic re-sampling method?

Assume the number of particles to be M, we need M random numbers to generate for the Multinomial re-sampling method, while only 1 random number is needed for the Systematic re-sampling method.

2.5 With what probability does a particle with weight $\omega = \frac{1}{M} + \epsilon$ survive the re-sampling step in each type of re-sampling (vanilla and systematic)? What is this probability for a particle with $0 \leq \omega < \frac{1}{M}$? What does this tell you? (Hint: it is easier to reason about the probability of not surviving, that is M failed binary selections for vanilla, and then subtract that amount from 1.0 to find the probability of surviving.

In vanilla re-sampling, for particle p with weight ω, the probability of being chosen at each iteration is ω. After M iterations the probability of not being chosen is $(1-\omega)^M$. So the probability of survival is $1-(1-\omega)^M$, which holds for both cases.
In systematic re-sampling, the step size is $\frac{1}{M}$, if $\omega = \frac{1}{M} + \epsilon$, it must be chosen so the surviving probability will be 1. If $0 < \omega < \frac{1}{M}$, the probability of surviving will be $\frac{\omega}{\frac{1}{M}} = M\omega$.

2.6 Which variables model the measurement noise/process noise models?

Variable Sigma_R models the process noise while Variable Sigma_Q models the measurement noise.

2.7 What happens when you do not perform the diffusion step? (You can set the process noise to 0)

If the diffusion step is not performed, the predicted state will be equal to applying motions. After several re-samplings, the particles will becomes M copies of the same particle that is the closest to the position where we apply the motion without noise at the beginning. Gradually, this particle will become far from the ground truth.

2.8 What happens when you do not re-sample? (set RESAMPLE_MODE=0)

If re-sampling is not performed, the particles will not converge to the target position and stay at the initial place. The error will become very large.

2.9 What happens when you increase/decrease the standard deviations(diagonal elements of the covariance matrix) of the observation noise model? (try values between 0.0001 and 10000)

when we increase the standard deviations of the observation noise model, the tolerance of difference between prediction and measurement will increase and we will trust more on our observations. The particles will converge to the correct state more quickly and tightly. However, if the standard deviations of the observation noise model are set to large, the error will also increase as the particles will spread with a larger variance.
when we decrease the standard deviations of the observation noise model, the prediction might be trusted too much and the particles might converge to a region far from the ground truth. In this way the error will also increase.

2.10 What happens when you increase/decrease the standard deviations(diagonal elements of the covariance matrix) of the process noise model? (try values between 0.0001 and 10000)

The diversity of particles will increase when we increase the standard deviations of the process noise model. The particles will spread to a wider region and will be more likely to cover the true position. However, the cluster will also become larger and there might be a larger error around the ground truth. Otherwise, the cluster will be small and condense, but not guaranteed to converge to the ground truth.

2.11 How does the choice of the motion model affect a reasonable choice of process noise model?

If the motion model is more accurate, a smaller process noise should be chosen. Otherwise, a larger process noise should be chosen so that the particles can estimate the true position.

2.12 How does the choice of the motion model affect the precision/accuracy of the results? How does it change the number of particles you need?

A more accurate motion model will make the results more precise and accurate. And fewer particles will be needed because of the high accuracy of the particles. If a less accurate motion model is chosen, we will need more particles to cover the target region and track the true position.

2.13 What do you think you can do to detect the outliers in third type of measurements? Hint: what happens to the likelihoods of the observation when it is far away from what the filter has predicted?

The likelihoods of the observation will be small when it is far away from what the filter has predicted. We can set a threshold for the calculated average likelihoods of the observation. If the likelihood is smaller than the threshold, it will be classified as an outlier.

2.14 Using 1000 particles, what is the best precision you get for the second type of measurements of the object moving on the circle when modeling a fixed, a linear or a circular motion(using the best parameter setting)? How sensitive is the filter to the correct choice of the parameters for each type of motion?

Motion Type	Sigma_Q	Sigma_R	Measurements Error	Estimates Error
Fixed	diag([80, 80])	diag([150, 150, 1.5])	1.2 ± 0.7	1.3 ± 0.7
Linear	diag([70, 70])	diag([150, 150, 1.5])	1.2 ± 0.7	1.3 ± 0.6
Circular	diag([50, 50])	diag([150, 150, 1.5])	1.2 ± 0.7	1.2 ± 0.6

Fixed motion is the most sensitive to the correct choice of process noise. Linear motion is more sensitive to the correct choice of measurement noise, and less sensitive to process noise. Circular notion is the most accurate because the object is moving on a circle, and it is not very sensitive to the choice of parameters.

2.15 What parameters affect the mentioned outlier detection approach? What will be the result of the mentioned method if you model a very weak measurement noise $|\mathbf{Q}| \to 0$?

The threshold λ_{Psi} will affect the outlier detection approach. The larger λ_{Psi} is, the more measurements will be classified as outliers. The measurement noise Q will also affect the outlier detection. If we model a very weak measurement noise $|\mathrm{Q}| \to 0$, the probability function will concentrate on the prediction, and the tolerance of difference between prediction and measurement will decrease, resulting in more measurements classified as outliers.

2.16 What happens to the weight of the particles if you do not detect outliers?

If we do not detect outliers, all measurements will be considered valid and will be trusted equally. The particles near an outlier will be given a larger weight than it should have. After the re-sampling we might get a wrong particle set and it will become more difficult to converge to the true position.

Simulation results

Dataset 4

At the beginning, there are four valid hypotheses for these 4 landmarks. Gradually it drops to two. When using 1 000 particles, the filter cannot keep track of all hypotheses and will converge to one of them quickly, sometimes correct while sometimes wrong. Increasing M to 10000 will make the filter able to keep track of more hypotheses. Multinomial re-sampling doesn't have abetter performance in keeping track of multiple hypotheses since it introduces random values for each particle. With a stronger measurement noise, more hypotheses will be preserved with lower accuracy. The hypothesis closest to the ground truth will be less likely to be discarded.

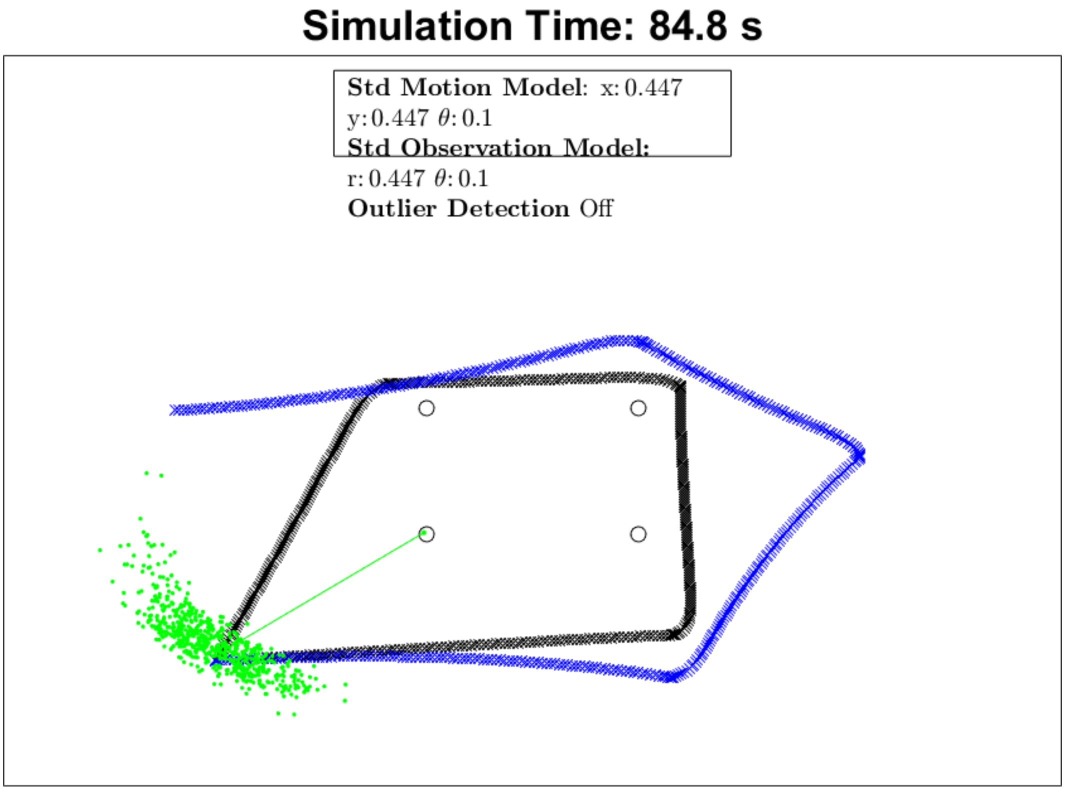

Figure 1: The path in the simulation for dataset 4 with 1000 particles

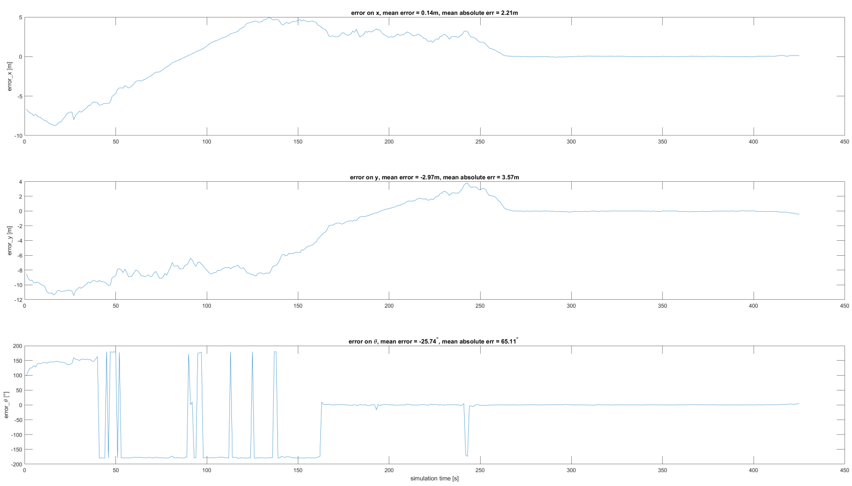

Figure 2: The error in the simulation for dataset 4 with 1000 particles

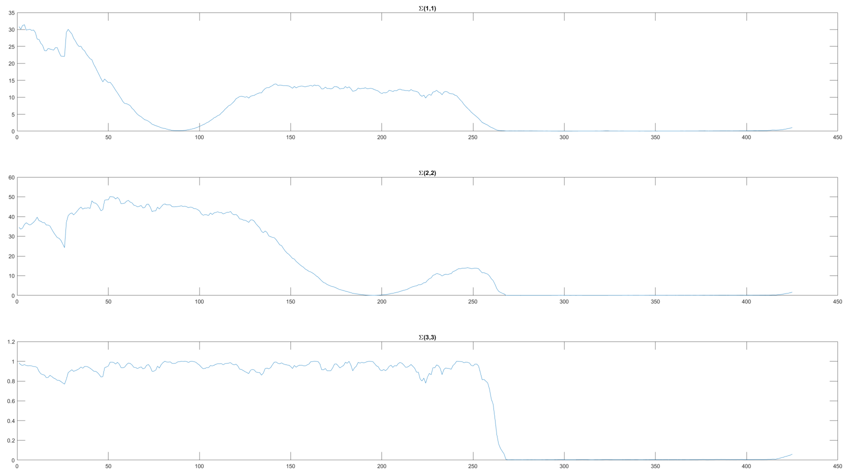

Figure 3: The covariance in the simulation for dataset 4 with 1000 particles

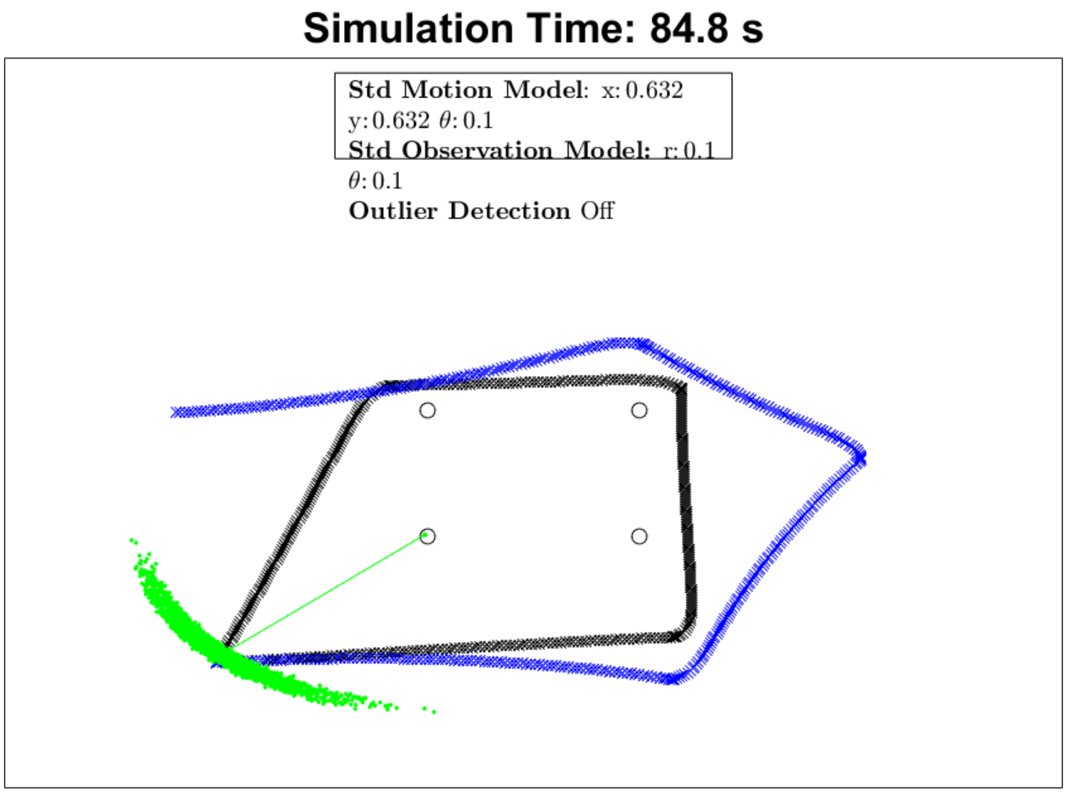

Figure 4: The path in the simulation for dataset 4 with 10000 particles

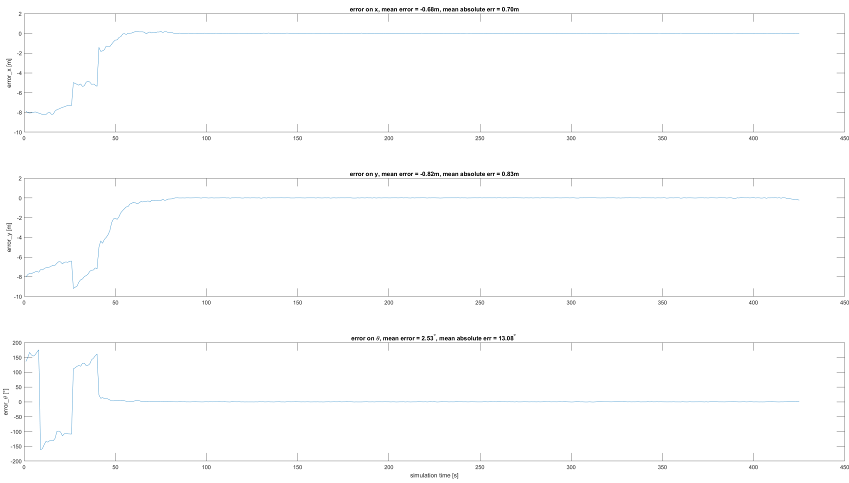

Figure 5: The error in the simulation for dataset 4 with 10000 particles

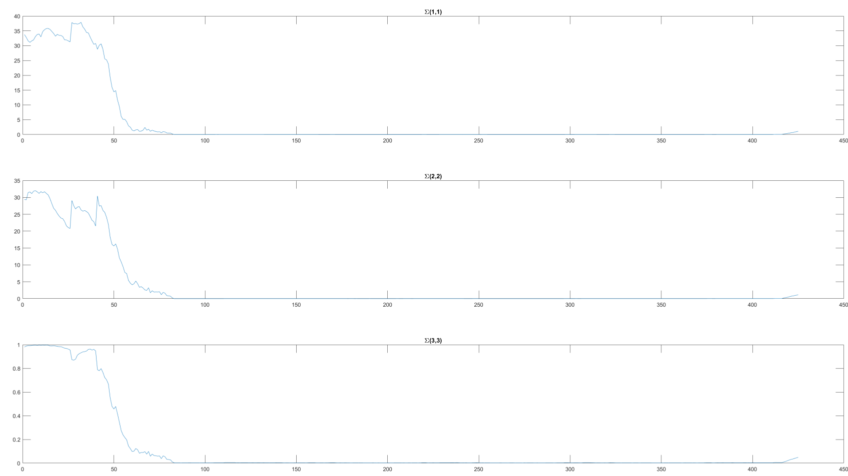

Figure 6: The covariance in the simulation for dataset 4 with 10000 particles

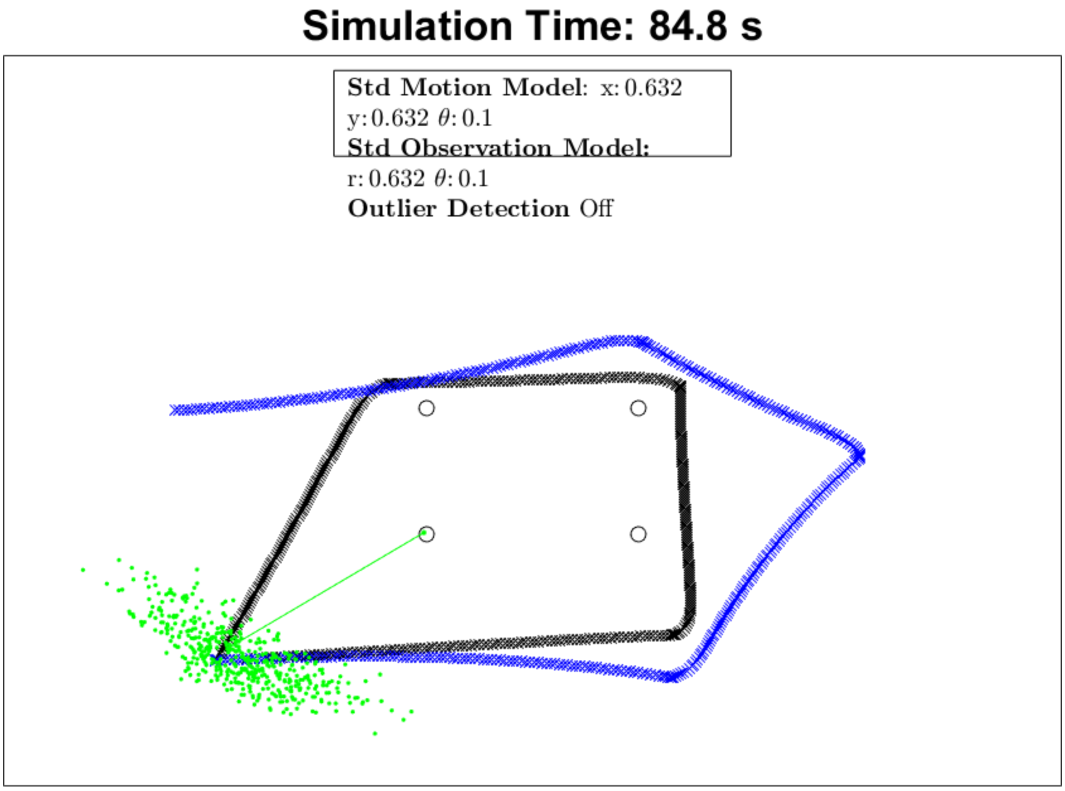

Figure 7: The path in the simulation for dataset 4 with 1000 particles and Multinomial re-sampling

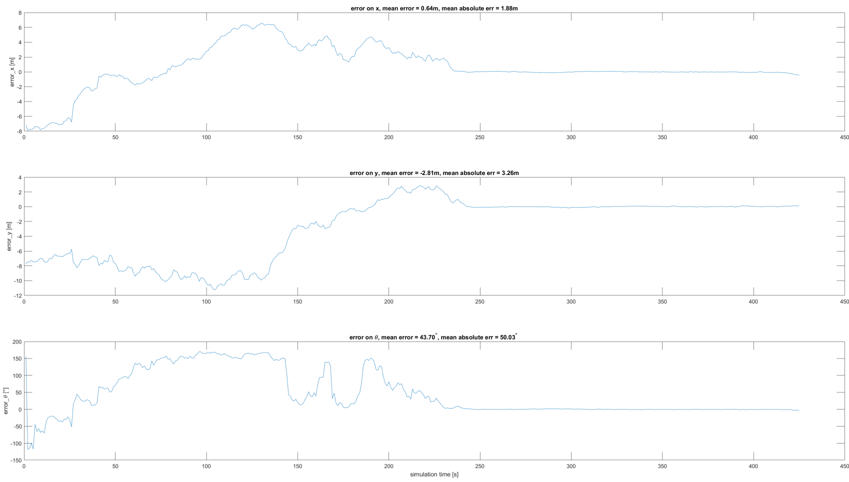

Figure 8: The error in the simulation for dataset 4 with 1000 particles and Multinomial re-sampling

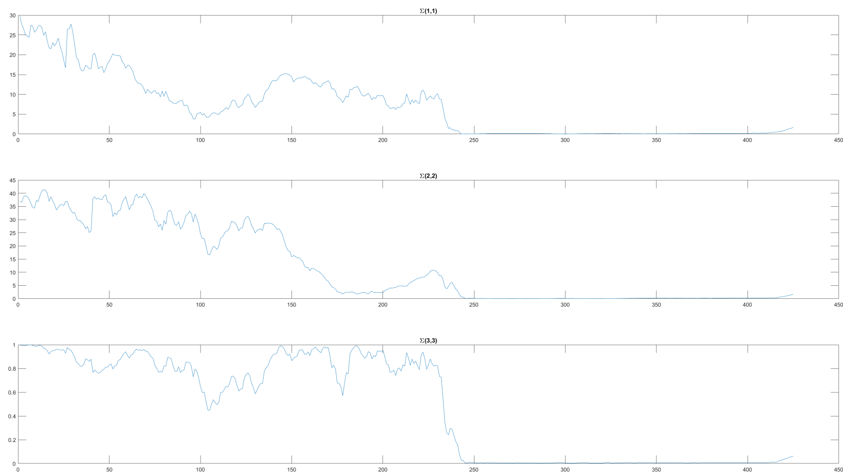

Figure 9: The covariance in the simulation for dataset 4 with 1000 particles and Multinomial re-sampling

Dataset 5

It is shown that the filter can converge to the ground truth after the new observation is obtained.

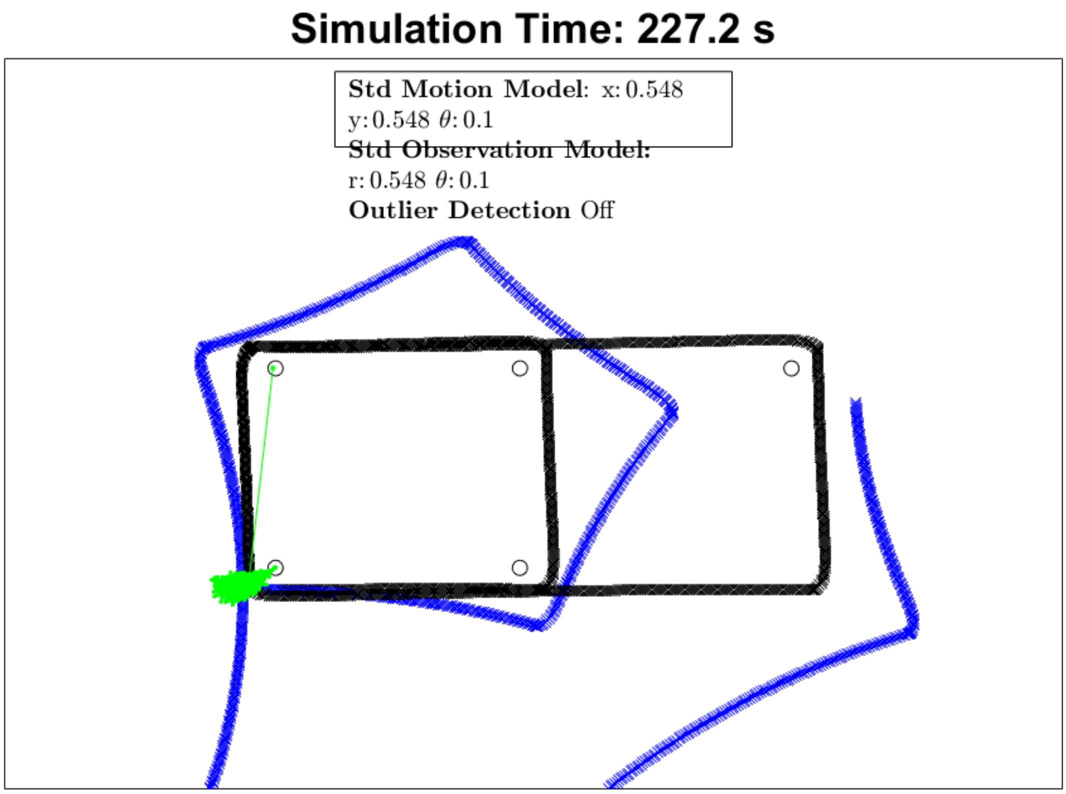

Figure 10: The path in the simulation for dataset 5 with 10000 particles

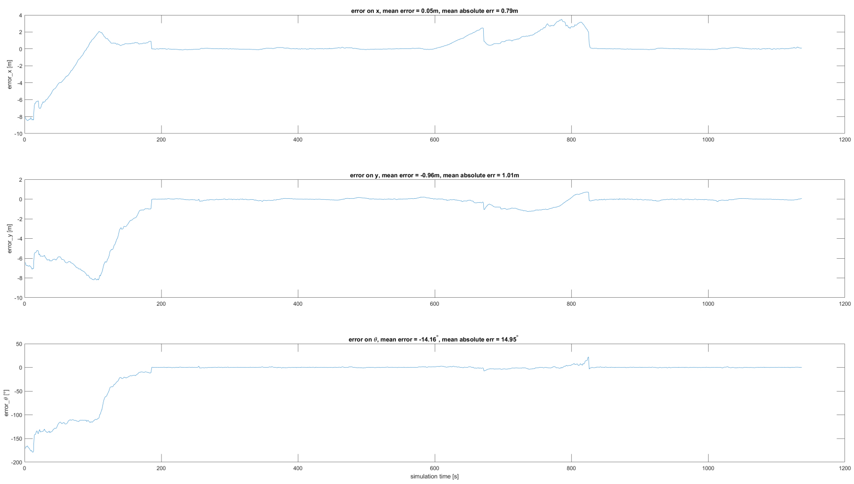

Figure 11: The error in the simulation for dataset 5 with 10000 particles

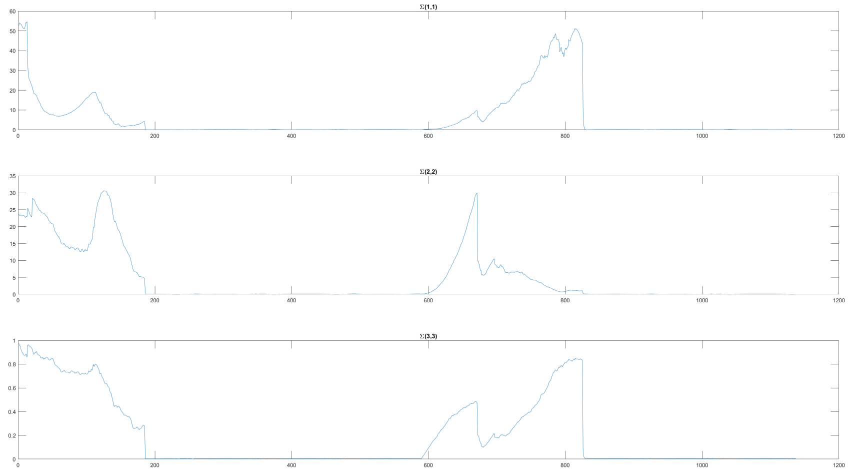

Figure 12: The covariance in the simulation for dataset 5 with 10000 particles

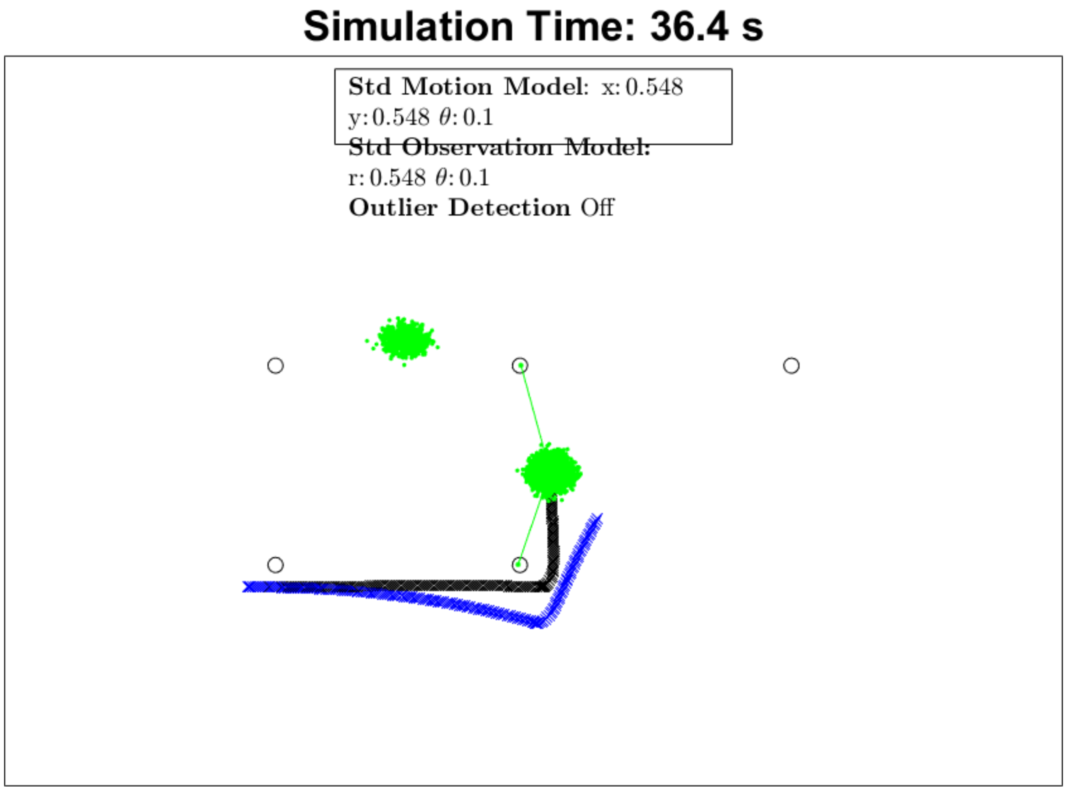

Figure 13: Simulation for dataset 5 before convergence

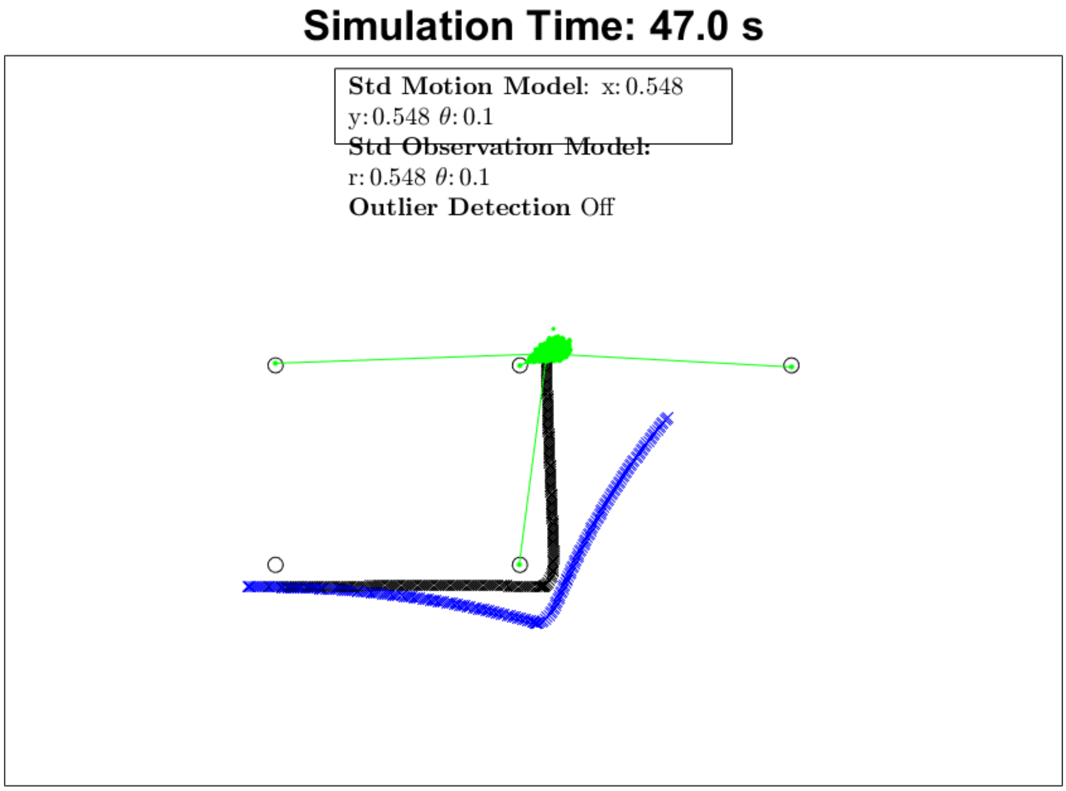

Figure 14: Simulation for dataset 5 after convergence