Applied estimation lab2

Zhengming Zhu

November 23, 2023

1 Part1

1.What are the particles of the particle filter?
The every particle represent individual hypotheses or samples of the possible state of the system at time t. Particle filter use a set of particles to estimate the posterior.

$$x_t^{[m]} - p(x_t|z_{1:t}, u_{1:t})$$

2.What are importance weights, target distribution, and proposal distribution and what is the relation between them?
The particle filters is based on Monte Carlo methods. Importance weights measures how well a sample from the proposal distribution represents the target distribution. Target distribution is often the posterior distribution of the system's stare given all the observations, which usually non-linear, complex. Proposal distribution is a distribution from which we can easily draw samples, the goal is make the proposal distribution similar to the target distribution. The relationship indicates that importance weight is how probable the particles are under the target distribution.

3.What is the cause of particle deprivation and what is the danger?
The cause can be large weight disparities which leads to the particles with high weights copied many times whereas the particles with low weights are discarded. Particle deprivation has a lot of dangers like loosing diversity, overconfident estimation, performance degradation.

4. Why do we resample instead of simply maintaining a weight for each particle always.
- avoid particle degeneracy: most particles might end up having negligible weights due to the multiplication of small likelihoods.
- maintaining diversity: resampling helps exploration and maintaining diversity of particles.
- avoid weight collapse: resampling helps in redistributing the weights more evenly among particles.

5.Give some examples of the situations which the average of the particle set is not a good representation of the particle set.
When the distribution of particles are Non-Gaussian or Multi-peaks distribution, the average of the particles will fall between these peaks, which is not the accurate estimation of the true state of system.

6.How can we make inferences about states that lie between particles.
Because every particle represent the hypothesis about the state of system, also, every particle has corresponding weights. Therefore, there are some techniques to inferences the state of system. The most simple way is weighted average. Every particle contributes by its importance weight. There are some other way to inferences as follows:
- create bin to count how many particles in the bin to form histogram.
- Using Gaussian kernel:

$$P(x) \propto \sum_{m=1}^{M} G(x - x^{(m)}, \sigma^2)$$

7.How can sample variance cause problems and what are two remedies?
The high sample variance can cause inaccurate estimation like weight degeneracy, the filter's estimate becomes overly reliant on a small subset of particles, reducing its accuracy. Also, high variance can

mean that the particle set does not represent the entire state space well.
Remedies:
- simpling using more particles which can reduce this type of error.
- using resampling techniques to reduce sample variance like stratified or systematic resampling.
- when: $\sum(\omega^(m) - \overline{\omega}) >= threshold$, then resampling.

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.
In general, the more particles we use, the lower pose uncertainty we will get. Because the filter can capture finer details of the distribution. However, there is a trade-off between the efficiency and the quality, more particles mean greater computation cost.

2 Part2

2.1 Prediction

1.What are the advantages/drawbacks of using (6) compared to (8)? Motivate.
Advantage:(6) uses the same initial value of the angle, it will reduce computation cost and error if there is no noise.
Disadvantage:Because there is much noise in 3D state space, using the previous angle as reference can smooth the trace. Therefore, (6) may leads to a wave path.

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 model uniform speed or uniform accelerated circular motions using (9). We need to know v_o and ω_0.

2.2 Sensor Model

3.What is the purpose of keeping the constant part in the denominator of (10)?
To normalize the likelihood function, so that it can represent the measurement probability distribution.

2.3 Re-Sampling

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?
The Multinomial re-sampling method need M random numbers whereas the Systematic re-sampling method only need one random number.

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 $0 <= \omega < \frac{1}{M}$? 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.
- vanilla re-sampling: Because the probability of surviving is equal to the weight value, no matter what the *omega* is, the particle can't be selected in N iteration is: $1 - (1 - \omega)^N$.
- Systematic re-sampling: If the $\omega = \frac{1}{M} + \epsilon$, and $\epsilon > 0$, the probability of being selected is 1. If the $0 <= \omega < \frac{1}{M}$, the probability of being selected is $p = \frac{\omega}{1/M} = M\omega$.

2.4 Experiments

6.Which variables model the measurement noise/process noise models?
R matrix models the modeled process noise and Q matrix models the observation noise. And we assume they are white Gaussian.

7.What happens when you do not perform the diffusion step?
Because the $\overline{u} = 0$ for fixed target, also, $R = 0$, there is $\overline{x}_t^m = x_{t-1}^m$, the result will the M copies of the same particle.

8.What happens when you do not re-sample? (set RESAMPLE MODE=0)
If the filter does not re-sampling, the particles can not converge according to the weight. Every particle will be remain and spread out over the whole region and move by the motion propagation.

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).
Obviously, when the standard deviation of the observation noise model become too big, the particles are hard to converge, which will cause bad estimation. On the other hand, decreasing the standard deviations can decrease the mean error, but if it is too small, the filter will depends on a small number of particles who has large weights, which may cause less usable particles, and the time of convergence will longer. Besides, more particles will be discarded as outliers.

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)
Increasing the standard deviations of the process noise model really makes particles spread out and helps exploration. But if it is too big, the convergence time will be longer and the mean error also larger. Decreasing the standard deviations of the process noise model concentrates the particles, which is not benefit for exploration.

11.How does the choice of the motion model affect a reasonable choice of process noise model?
If the motion model is accurate, we can choose a smaller process noise model, also, if the motion model can not describe accurately, we should choose a larger process noise model.

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? Obviously, if the motion model describe the real motion of the system, the precision and accuracy of the estimation will increase. If the motion model is complex and cover a large region, we may need more particles to explore the actual state of the system. However, too large number of particles leads to slow convergence.

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 further distance from the state the particles have, the smaller value the likelihoods function has, therefore, we can set a threshold to detect the outliers.

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?
- Fixed motion: Because the motion model does not describe the actual motion of the system, we should set large process noise to explore, but not too large. For many pairs of parameters, we get the best precision is 11.2 +-5.8:

$$R = \begin{bmatrix} 50 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 0.05 \end{bmatrix} Q = \begin{bmatrix} 500 & 0 \\ 0 & 500 \end{bmatrix}$$

Re-sampling mode: systematic re-sampling
Likelihood threshold: 0.0001

- Linear motion: We got the best precision is 7.8 +- 3.7:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{bmatrix} Q = \begin{bmatrix} 200 & 0 \\ 0 & 200 \end{bmatrix}$$

Re-sampling mode: systematic re-sampling
Likelihood threshold: 0.0001

- Circular motion: Because this motion model the actual system motion, we can set small covariance of the process noise and measurement model. We got the best precision is 7.5 +- 3.5:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{bmatrix} Q = \begin{bmatrix} 300 & 0 \\ 0 & 300 \end{bmatrix}$$

Re-sampling mode: vanilla re-sampling
Likelihood threshold: 0.0001

3 Main problem: Monte Carlo Localization

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 $|Q| \to 0$?
The threshold λ_Ψ affect the outlier detection. For example, λ_Ψ is high, more particles will be recognized as outliers. If $|Q| \to 0$, the value of likelihood will be large, so that less useless particles can be discarded as outlier, which is not good for convergence.

16.What happens to the weight of the particles if you do not detect outliers?
The Ψ of outlier is very small, and if do not detect outlier, the weight of particles will change a lot and not valid for estimation, which cause bad estimation.

3.1 Dataset4

3.1.1 Tracking

1.Different particles: For a tracking problem, the initial position is given, so it is easy to track and estimation state of robot.
Add particles from 1000 to 10000 will not change much about the error and convergence time because most of useless particles will be discarded early.
However, reducing the process noise and measurement noise covariance could reduce the mean error.

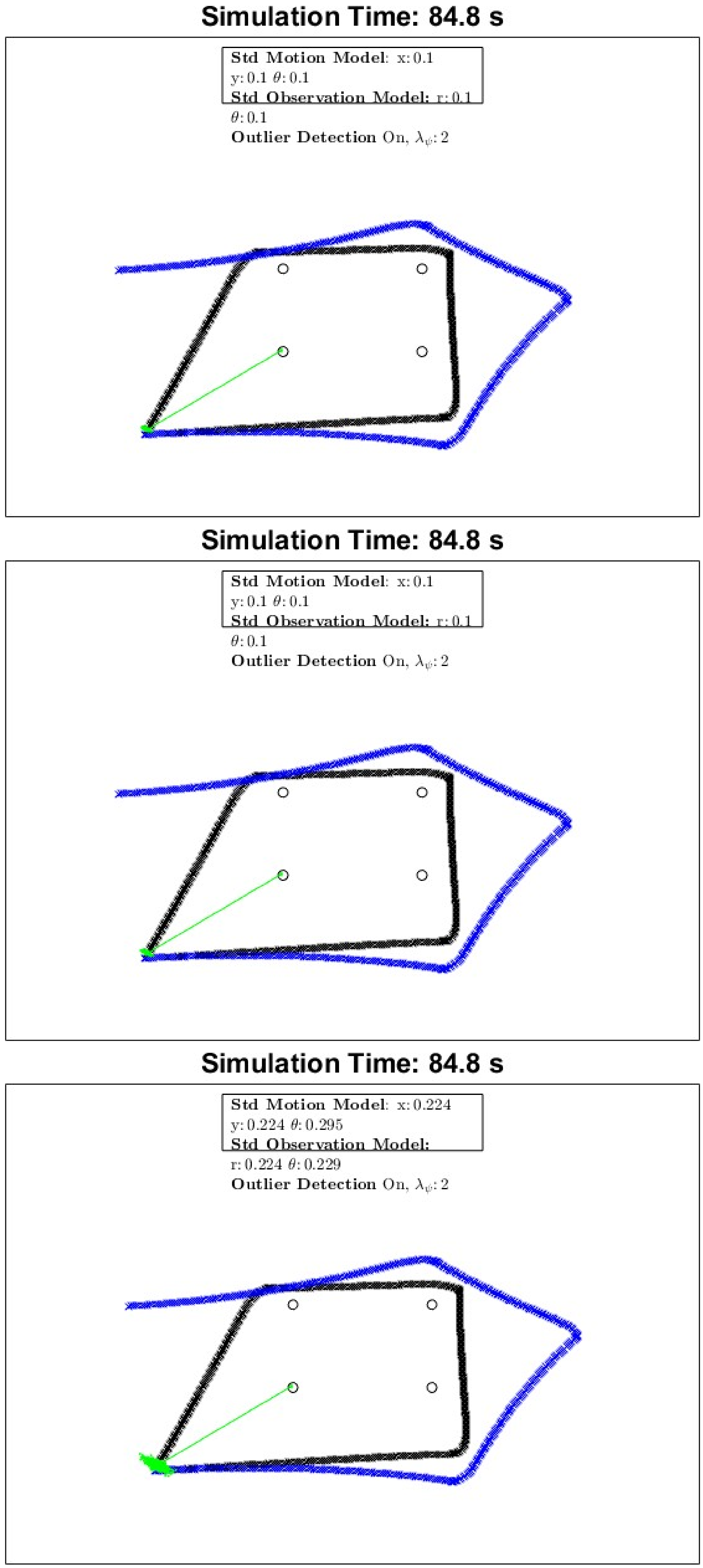


Figure 1: Default(left), M = 10000(middle), Small noise(right)

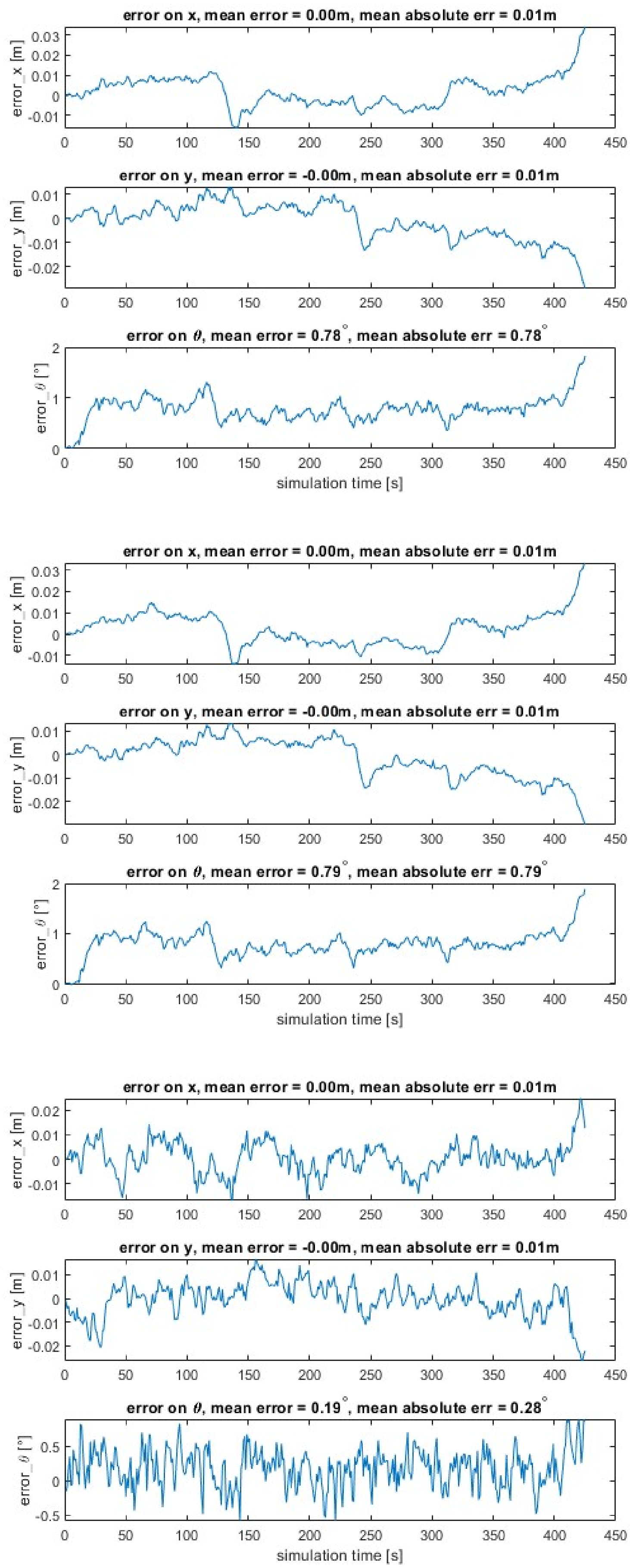


Figure 2: Error: Default(left), M = 10000(middle), Small noise(right)

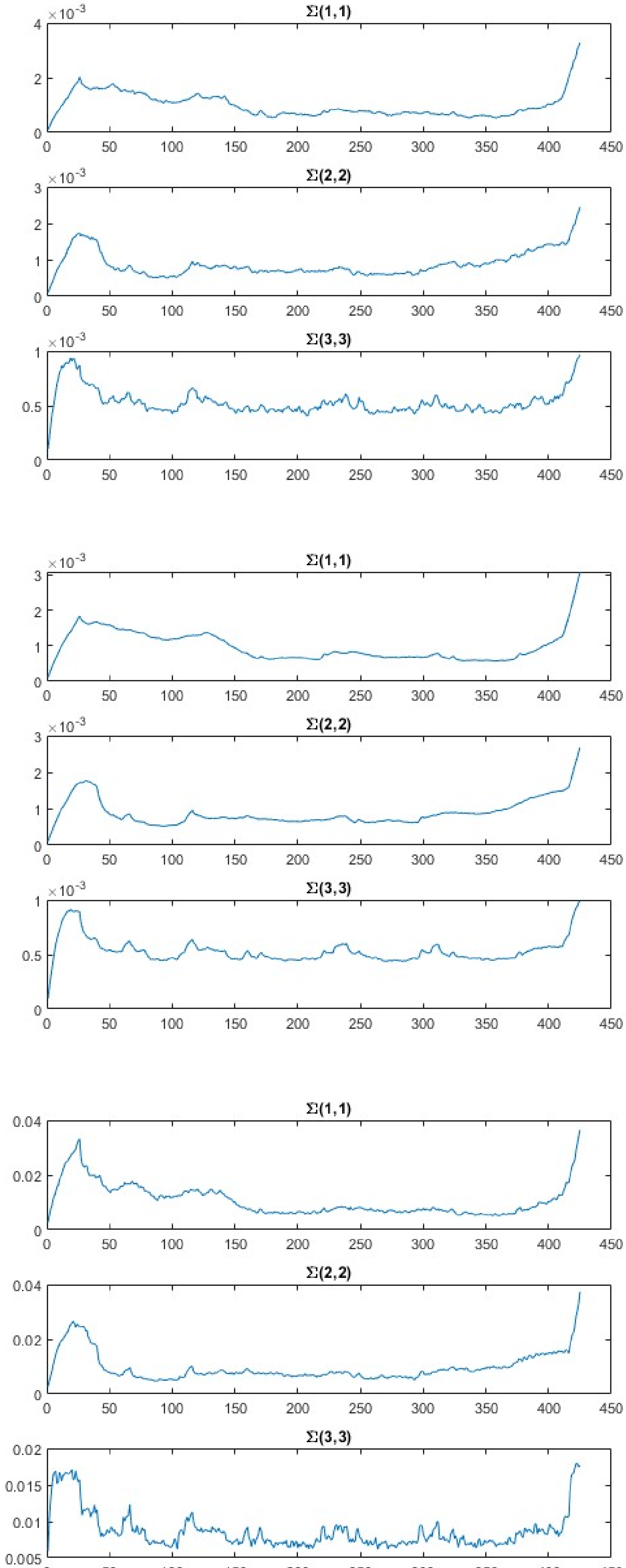
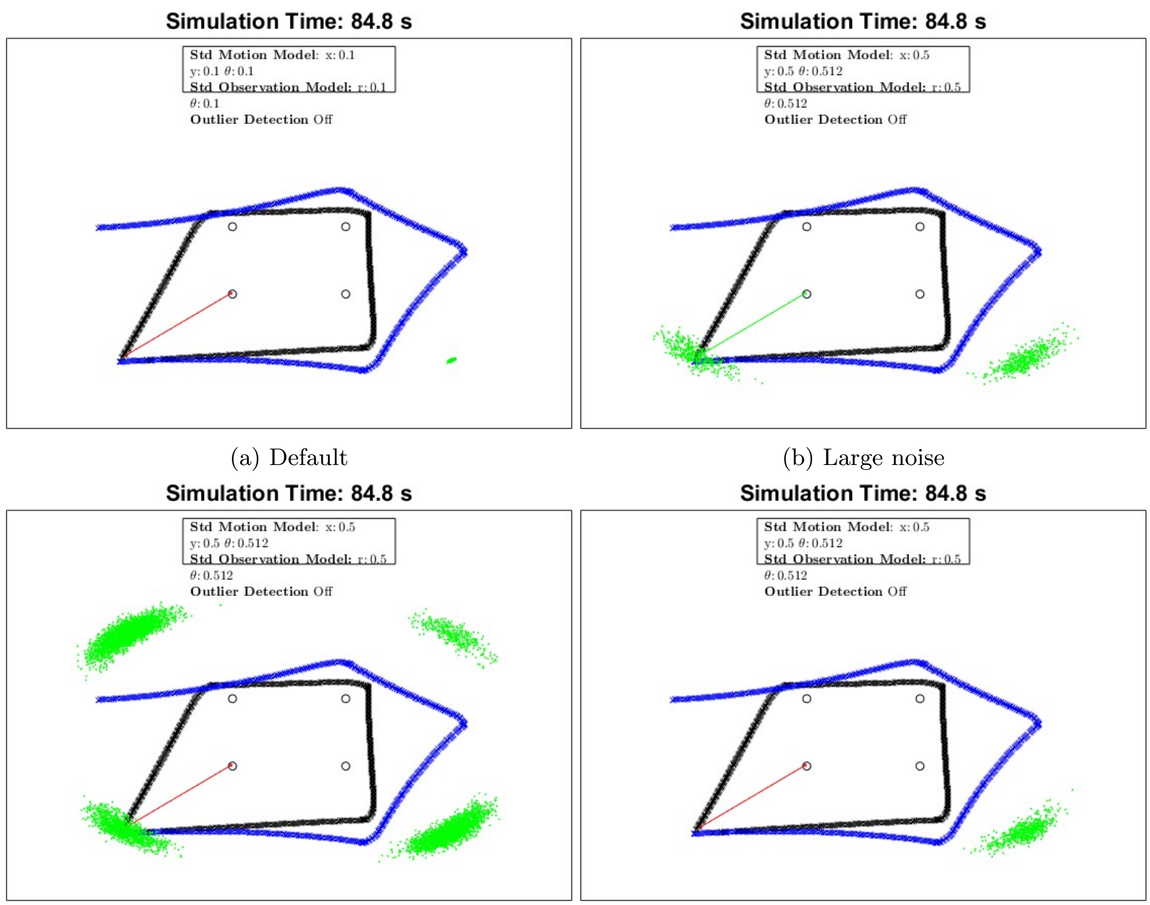


Figure 3: Covariance: Default(left), M = 10000(middle), Small noise(right)

3.1.2 Global localization

For a global localization problem, the initial position is not given, and the map is perfectly symmetric environment with 4 landmarks. So, there should be 4 hypothesis.
Also, we need large process noise and measurement noise covariance to explore the possible state.
Obviously, increasing particle number also help preserve multiple hypotheses, because the filter can cover wider region.
In practice, The multinomial resampling does not help preserve multiple hypotheses since the multinomial resampling introduces random value for each particles.



(a) Default

(b) Large noise

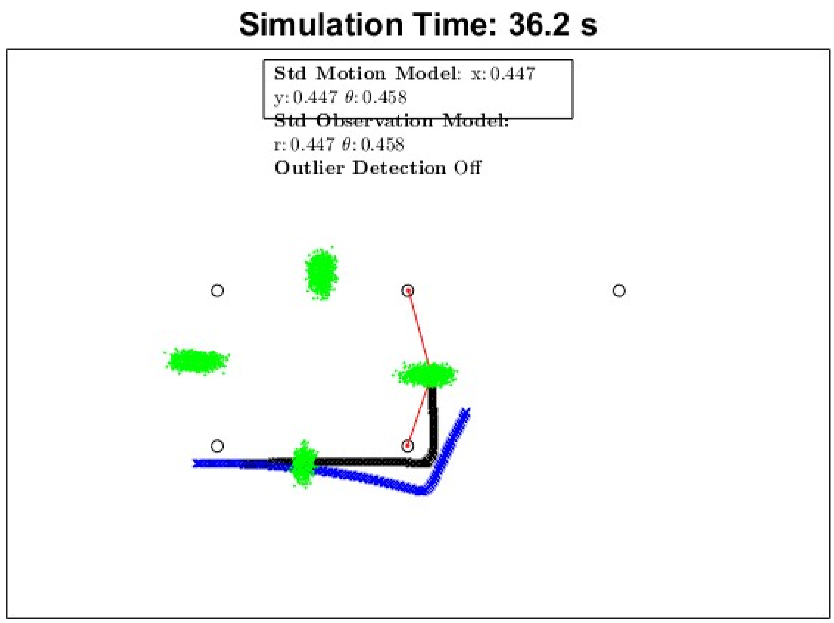
(c) M=10000

(d) Multinomial resampling

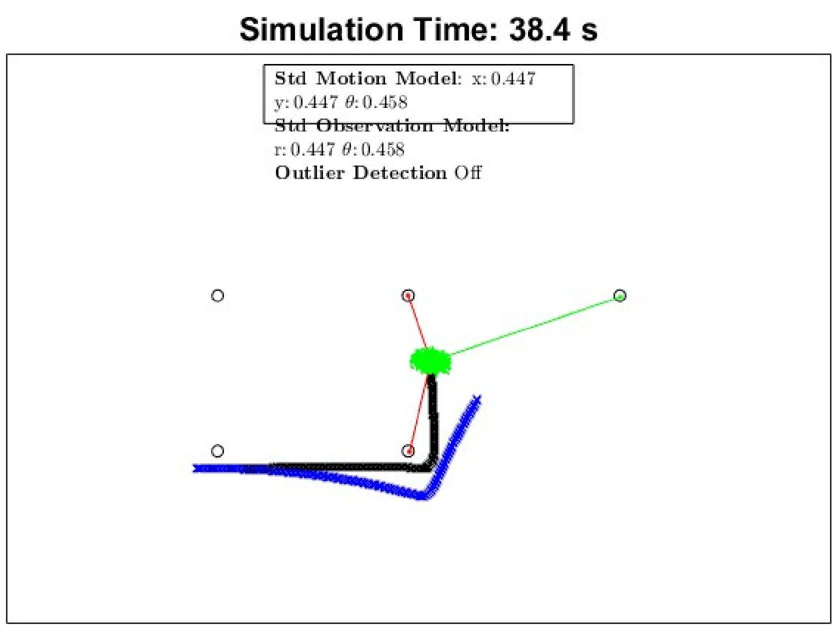
Figure 4: Comparison of different settings

3.2 Dataset5

At first, there are four valid hypotheses at time 36.2s, because the robot only measured the four symmetric landmarks. But, at time 38.4s, the robot observed the top right landmarks which breaks the symmetry, so there is only one hypotheses.



(a) Before the convergence



(b) After the convergence