Lab 2 Part I

1. **What are the particles of the particle filter?**

   Particles are the different possible states of a system. Particles are the samples of a posterior distribution.

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

   Proposal distribution: The distribution from which the particles are sample. The distribution of particles obtained after the prediction step, that is, after applying the motion model to a given state space.

   $$p(x_t|u_t, x_{t-1})bel(x_{t-1})$$

   Weights: Weights denote the probability of measuring a particle when the measurement model is applied. Higher the probability for a particle being measured, higher the weight allocated for the particle.

   $$w_t = p(z_t|x_t)$$

   Target distribution: The true posterior that is to be approximated with particles. The distribution of state space after incorporating the measurement model.

   $$p(x_t|u_t, z_t)$$

   $$weights = \frac{target\ distribution}{proposal\ distribution}$$

3. **What is the cause of particle deprivation and what is the danger?**

   Random resampling causes particle deprivation. Due to random resampling – a series of unlucky random numbers, there is a high chance that all the particles near the true state can be wiped out.

4. **Why do we resample instead of simply maintaining a weight for each particle?**

   - Maintaining a weight will still approximate the posterior, but most of the particles will end up in low probability posterior regions. So we might need more particles to represent the posterior.
   - To remove highly unlikely particles. A particle is deemed unlikely based on the weight assigned to it. As time progresses, these unlikely particles transition to unlikely states, thereby not contributing towards the representation of the pdf. Hence, we resample the particles to represent the pdf as accurately as possible.

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

   Multi-modal distribution, ring like distribution, skewed distributions.

6. **How can we make inferences about states that lie between particles?**

   Using interpolation or density extraction.

7. **How can sample variance cause problems and what are two remedies?**

   Sampling variance: The variance of the particle set as an estimator of the true belief is known as sampling variance. The resampling process reduces the diversity in the particle set. This causes the variance within the particle set to reduce but the variance of the particle set as an estimator to the true belief increases.

   Remedies:
   1) Reducing the rate at which resampling takes place. Eg. When there is no state change, resampling shouldn't take place.
   2) Low variance sampling: Instead of selecting M random particles, one random number is chosen in the interval 0 to M^{-1}, the particles are chosen based on this number with a probability proportional to the weight.
   3) Stratified sampling: Particles are grouped into clusters, over the time number of particles within a cluster should remain the same irrespective of the total weight of the particles within a cluster.

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

   Higher the uncertainty of the true posterior, more number of particles are needed to represent the true posterior.

Part II

1. **What are the advantages/drawbacks of using (5) compared to (7)? Motivate.**

   - Equation 5 represents the state space with two state variables and equation 7 represents the state space with 3 variables. Using three state variables makes the state space representation more generic and easy to adapt to different types of target, i.e., target on a line (where the third state variable (θ) doesn't change over the course of time) and target on a circle (where third state variable (θ) changes).
   - Since the motion model has noise, using the 3D space representation provides better results as the control model will use the θ which is an estimate (considering the motion model noise) from the previous time step.

2. **What types of circular motions can we model using (8)? What are the limitations (what do we need to know/fix in advance)?**

   Circular motion with constant linear and angular velocity can be modelled using Equation 8. v_o *and* ω_o are constant, we need to set them in advance.

3. **What is the purpose of keeping the constant part in the denominator of (10)?**

   The denominator serves as a normalization factor, $p(z|x, \Sigma_Q, \bar{C})$ is a probability distribution whose integral should be 1.

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?**

   For Multinomial re-sampling M (number of particles required) random numbers are required whereas for Systematic re-sampling, one random number is required.

5. **With what probability does a particle with weight $\boldsymbol{w = \frac{1}{M} + \epsilon}$ survive the re-sampling step in each type of re-sampling (vanilla and systematic)? What is this probability for a particle with $0 \leq w < \frac{1}{M}$? 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 resampling:
   Probability of not being drawn in any M chances = $(1 - w)^M$
   Probability of being drawn in M chances = 1 - $(1 - w)^M$
   where $\boldsymbol{w = \frac{1}{M} + \epsilon}$ or $\mathbf{0 \leq w < \frac{1}{M}}$

   Systematic resampling:
   For $\boldsymbol{w = \frac{1}{M} + \epsilon}$ the probability of survival is 1, as the random number generated is $rand(0 \leq ro \leq \frac{1}{M})$ which is less than $\boldsymbol{w}$**.**

   For $\mathbf{0 \leq w < \frac{1}{M}}$, the probability of survival depends on both size of M and weight. As M increases, the CDF increases and chances of CDF being greater than the random number generated. Hence, the probability of survival is M*w.

6. **Which variables model the measurement noise/process noise models?**

   Variable Q – measurement noise covariance matrix
   Variable R – process noise covariance matrix

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

   When there is no process noise, M copies of a single particle represents the true posterior after few resampling steps. This single particle is the closest to the ground truth, and assigned a higher weight.

8. **What happens when you do not re-sample? (set RESAMPLE MODE=0)**

   The initial (t=1) uniform particle distribution remains the same for all time steps. Each particle in the uniform distribution moves based on the motion model but do not converge as the weights assigned due to the measurement model are not used to resample the particles.

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 the measurement standard deviation is very low ~0.0001, it implies that the sensors are very accurate. The motion model on the other hand is not so accurate, it has considerable noise. For this situation, the proposal and target distributions deviate as $p(z_t|x_t)$ will be very low. The probability of an accurate z_t matching an inaccurate x_t is less.
- When the measurement standard deviation is comparatively close to the motion standard deviation, Q ~ 10, the convergence of the particles to the ground truth is faster.
- When Q ~ 10000, the uncertainty about the measurement it too high, hence the uncertainty of the particles around the ground truth is high.

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)

For smaller process standard deviation, the true state estimate is slow and the particle cloud is concentrated. For larger process standard deviation, the particle cloud is spread out and the true state estimation is fast.

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

If the motion model is accurate, then the process noise covariance matrix can have low values. In case the motion model is not accurate, the inaccuracy can be represented with high values in the noise covariance matrix.

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?

An inaccurate model will have large uncertainty, therefore requires a large number of particles and the spread of the particles will be large.

13. What do you think you can do to detect the outliers in the 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 likelihood function value will be low when the observation is far away from the predicted value, a threshold can be set on the likelihood function value. If the likelihood for a particular observed and predicted state is less than the threshold, then the observation can be deemed as an outlier.

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 modelling 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:

Q – measurement noise	R- process noise	Estimation error
$\begin{matrix} 100 & 0 \\ 0 & 100 \end{matrix}$	$\begin{matrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0.01 \end{matrix}$	26.3 +- 9.0

200 0 0 200	10 0 0 0 10 0 0 0 0.01	12.2 +- 5.4

Linear motion:

Q – measurement noise	R- process noise	Estimation error
100 0 0 100	2 0 0 0 2 0 0 0 0.01	12.5 +- 4.9
300 0 0 300	5 0 0 0 5 0 0 0 0.01	9.0 +- 4.1

Circular motion:

Q – measurement noise	R- process noise	Estimation error
100 0 0 100	2 0 0 0 2 0 0 0 0.01	8.7 +- 4.6
200 0 0 200	2 0 0 0 2 0 0 0 0.001	6.5 +- 3.3

The fixed motion model has a large process noise compared to other motion models in order to reduce the estimation error. Of all the motion models, circular motion model has the least estimation error. The linear model required more measurement noise compared to the circular motion model, even though they are comparable.

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| → 0?

For the mentioned outlier detection approach, parameters – measurement noise, the likelihood threshold affect the most. A weak measurement noise will lead to more measurements being classified as outliers. Similarly, higher the value of likelihood threshold, higher the number of measurements being considered as outliers.

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

Since weight distribution is proportional to $p(z_t|x_t)$, an outlier will result is erroneous weight distribution. The true state estimation might end up being inaccurate.

Simulation results

Dataset 4:

- **Tracking using 1000 particles:**
  The filter performs well for the tracking problem as the initial position of the robot is known. Two types of resampling methods were tried for tracking and the results can be seen below.

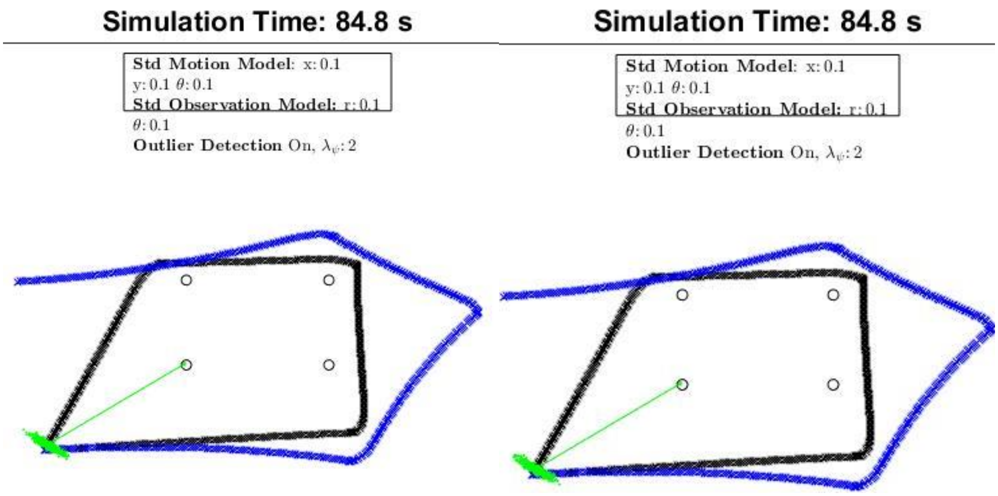


Fig 1. Tracking output – Systematic resampling (left), Multinomial resampling (right)

From the outputs obtained, it can be understood that for tracking problem systematic resampling and multinomial resampling produces the same results. There are slight differences in the error evolution, but the error mean are the same.

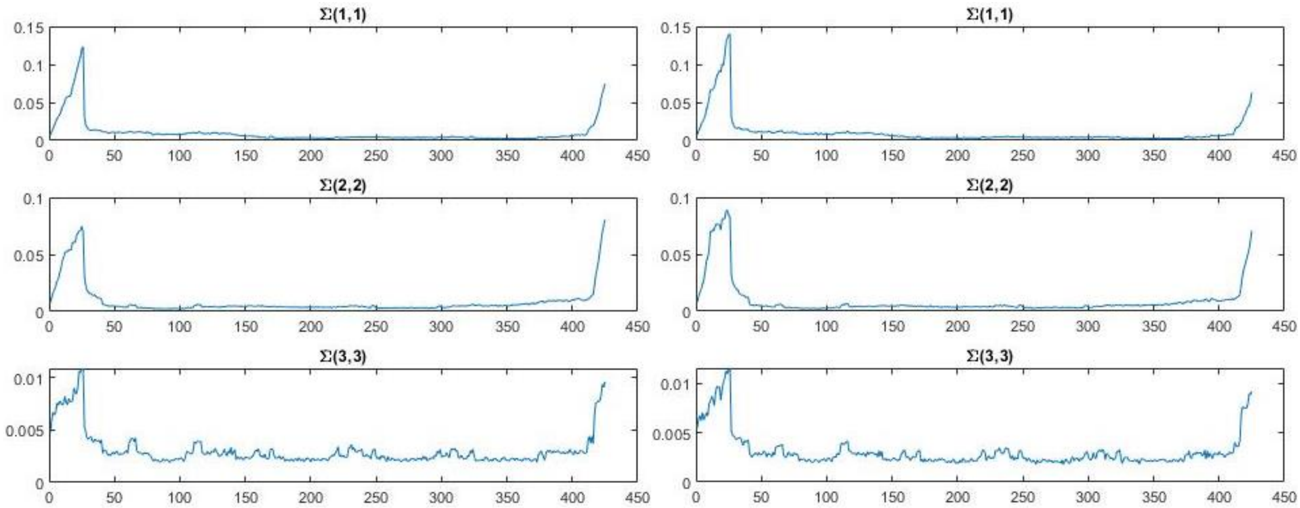


Fig 2. Tracking covariance evolution – Systematic resampling (left), Multinomial resampling (right)

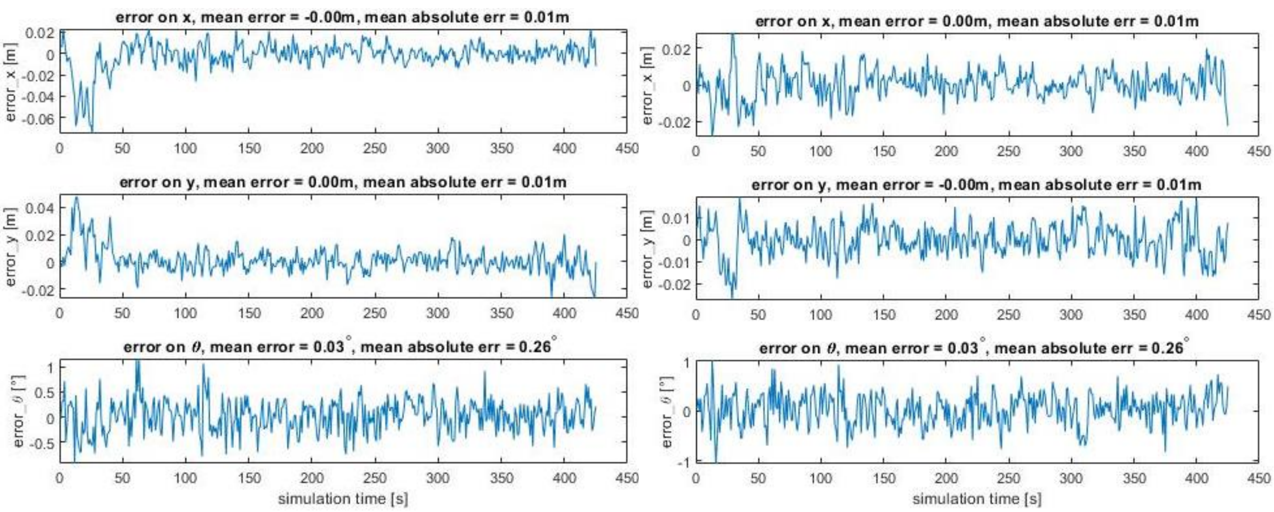


Fig 3. Tracking mean errors – Systematic resampling (left), Multinomial resampling (right)

- **Localization using 1000 particles using Systematic resampling**

In localization, at the start of the filter the particles are spread throughout the map signifying the robot's uncertainty in its position. As the time progresses, resampling removes particles and at the end of the simulation only the possible hypothesis are represented by the particles. In this symmetric map, there are four possible hypothesis.

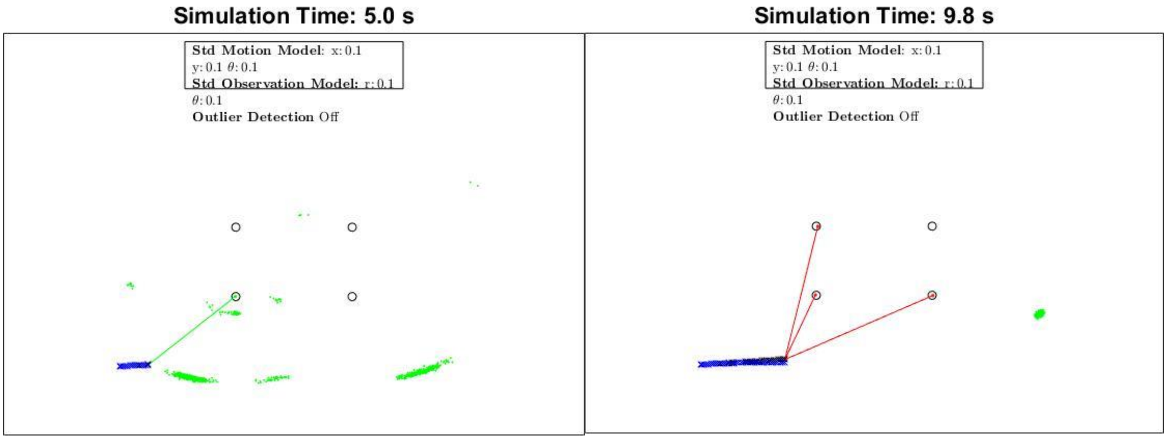
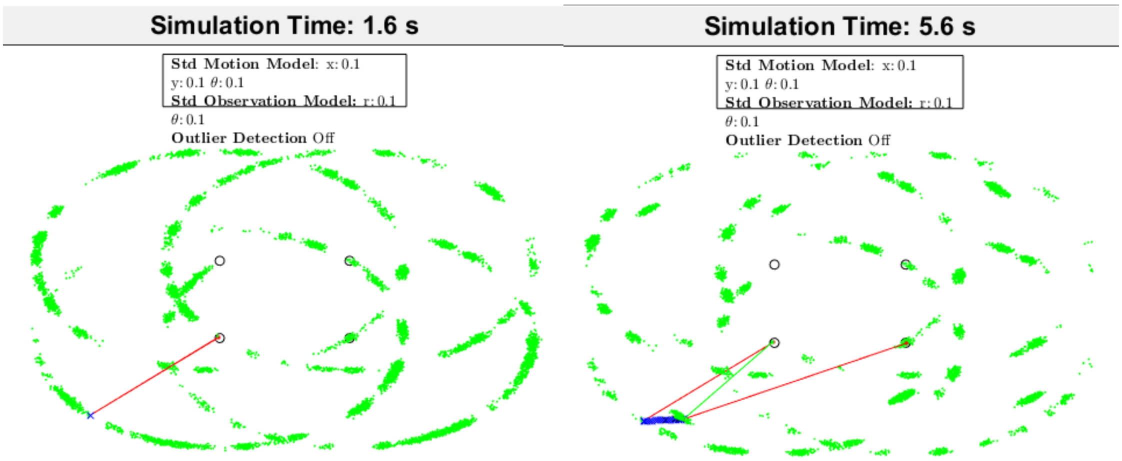


Fig 4. Localization particle (1000) spread at different time instances

There are 4 hypothesis available for this symmetric map. But not all of them are preserved using 1000 particles, particle deprivation occurs. In order to preserve the hypothesis the number of particles has to be increased.

- **Localization using 10000 particles using Systematic resampling**



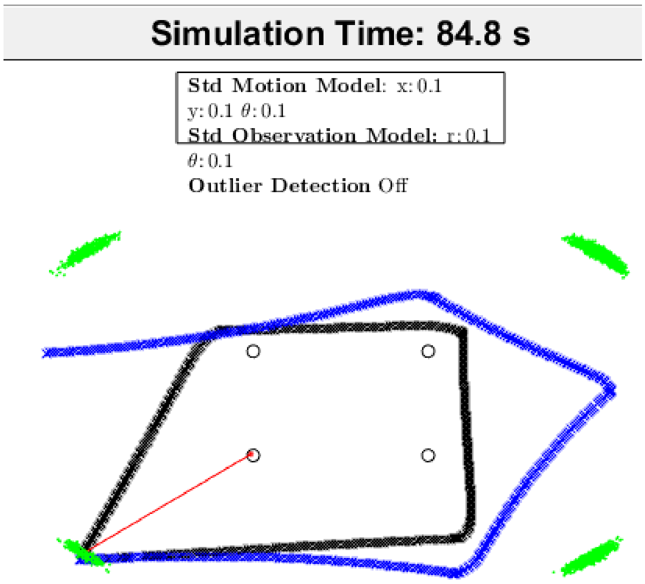
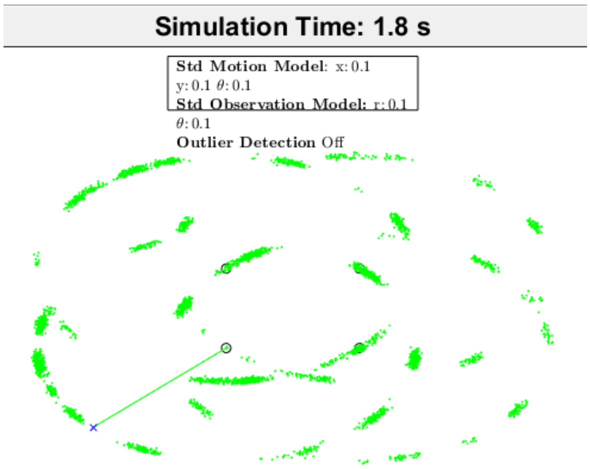


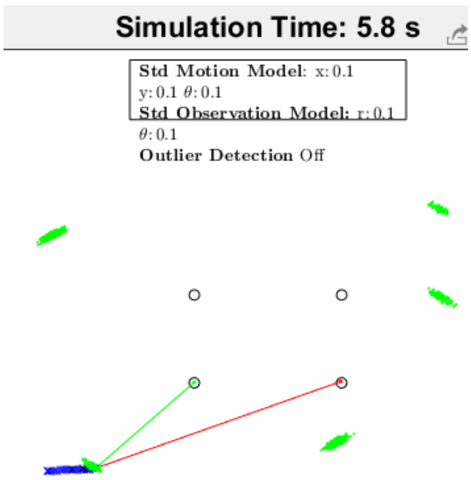
Fig 5. Localization particle (10000) spread using Systematic resampling - at different time instances

Using more number of particles for the localization task prevents particle deprivation and helps in representing the 4 hypothesis over longer period of time.

- **Localization using 10000 particles using Multinomial resampling**

Hypothesis disappear faster in multinomial resampling even though there are 10000 partices. Multinomial resampling produces more sampling variance than systematic resampling. The results can be seen below.





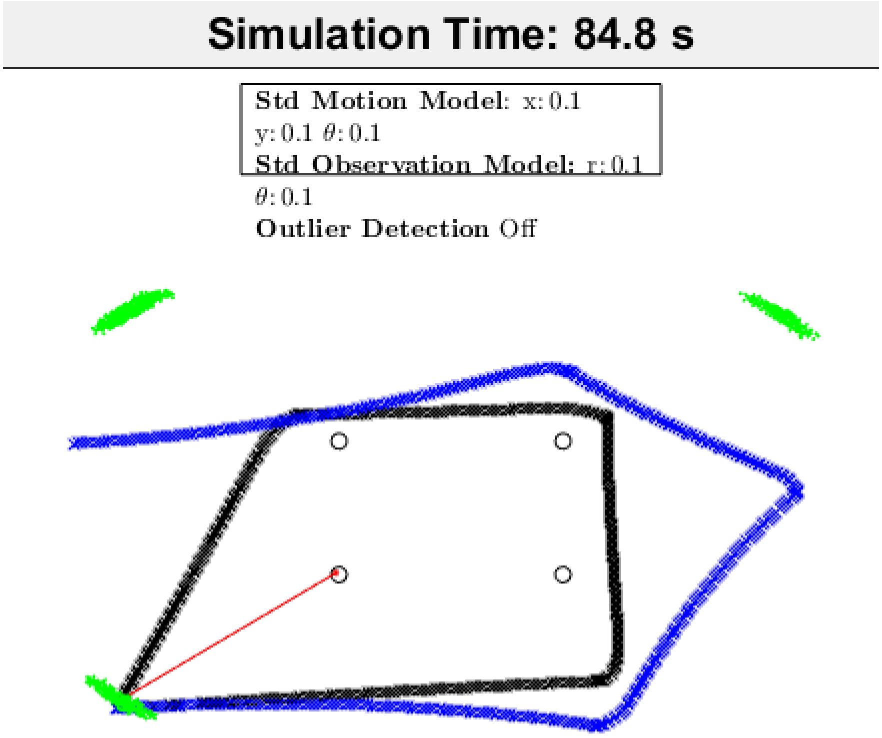
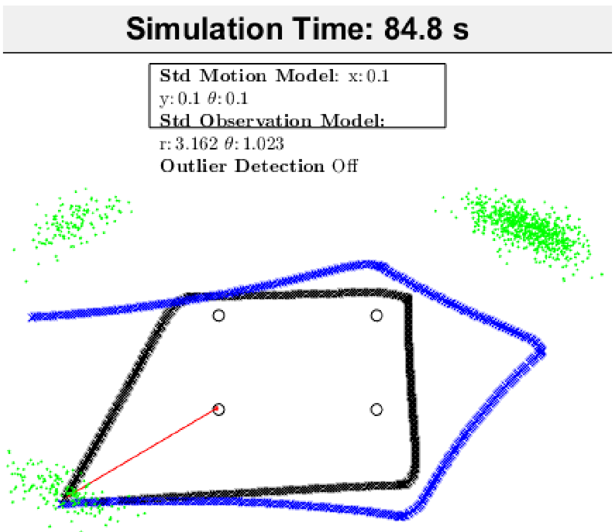
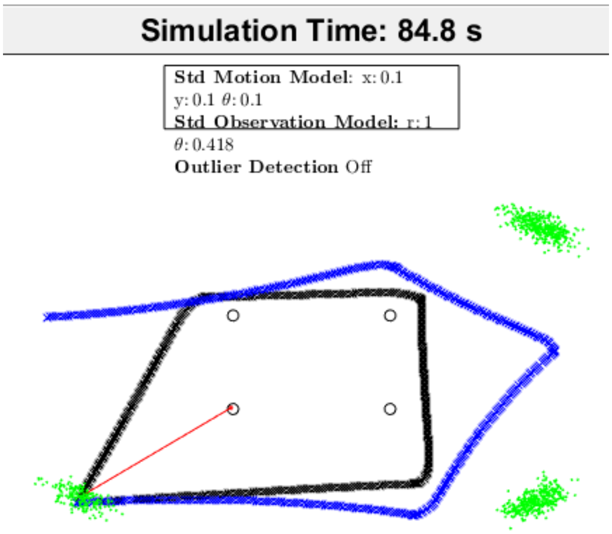


Fig 6. Localization particle (10000) spread using Multinomial resampling - at different time instances

- **Localization using 10000 particles using Systematic resampling – Measurement noise change**

  Output of particle filter for different measurement noises can be seen below. When the measurement noise is increased, the hypothesis are preserved better. When the measurement noise is low, it implies that the sensors are accurate and the probability of an accurate z_t matching an inaccurate x_t is less. This will reduce the weight assigned to each particle causing the particles to vanish. Thus, poor hypothesis preservation in case of low sensor noise.





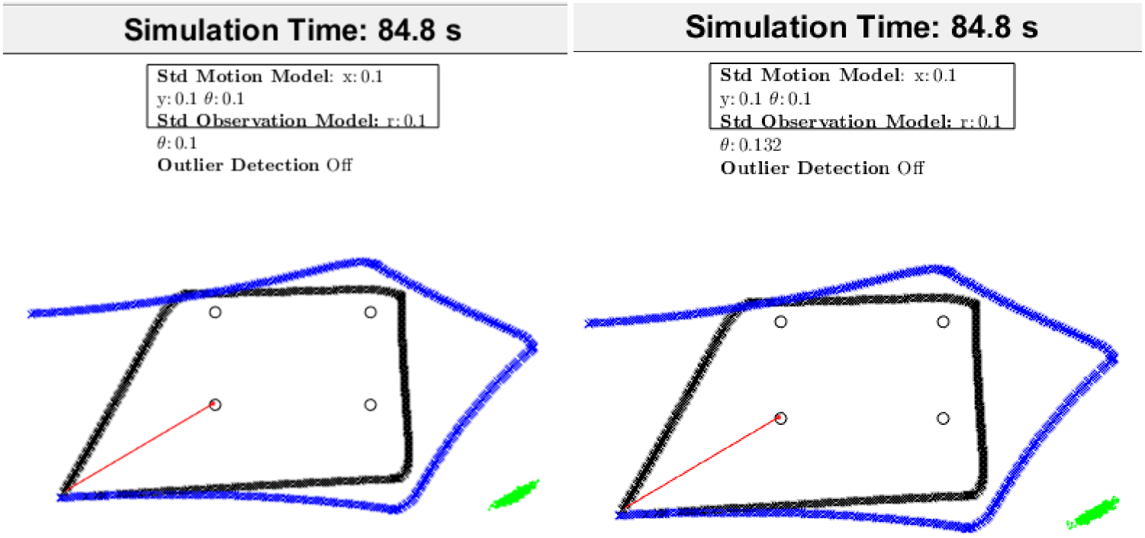
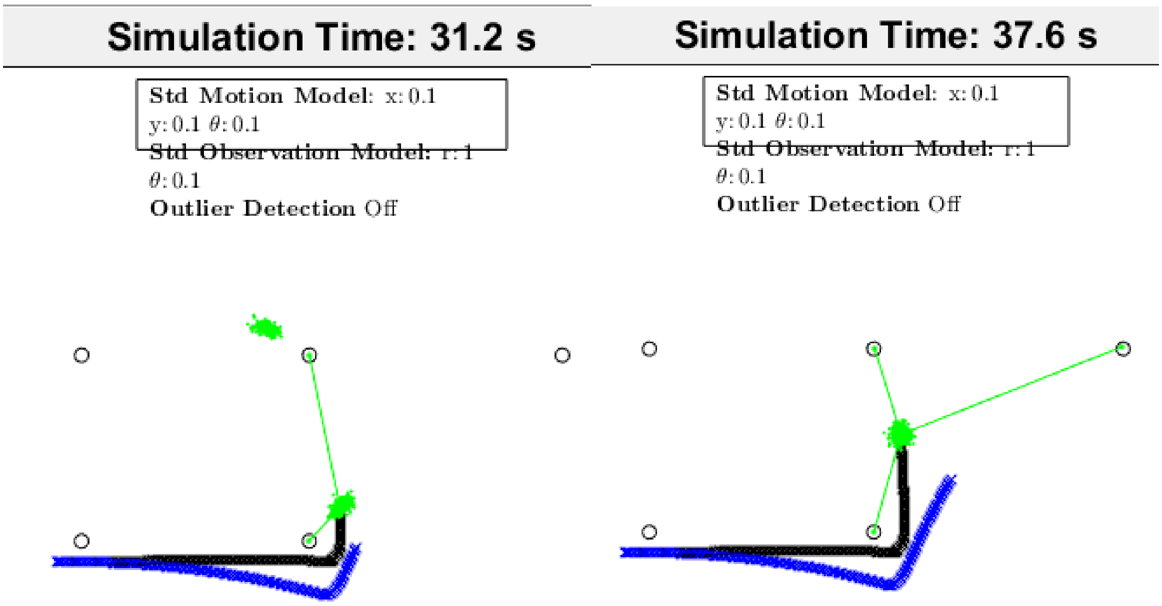


Fig 7. Localization particle (10000) spread using Systematic resampling – different measurement noises

Dataset 5:

For this dataset, as it is mentioned in the manual, after 37s when the robot observes the far right point the symmetry of the map is broken thus causing multiple hypothesis to converge into single hypothesis. The localisation problem behaves like a tracking problem. The below mentioned images describe this behaviour. To obtain these results the measurement noise was increased from the default settings.



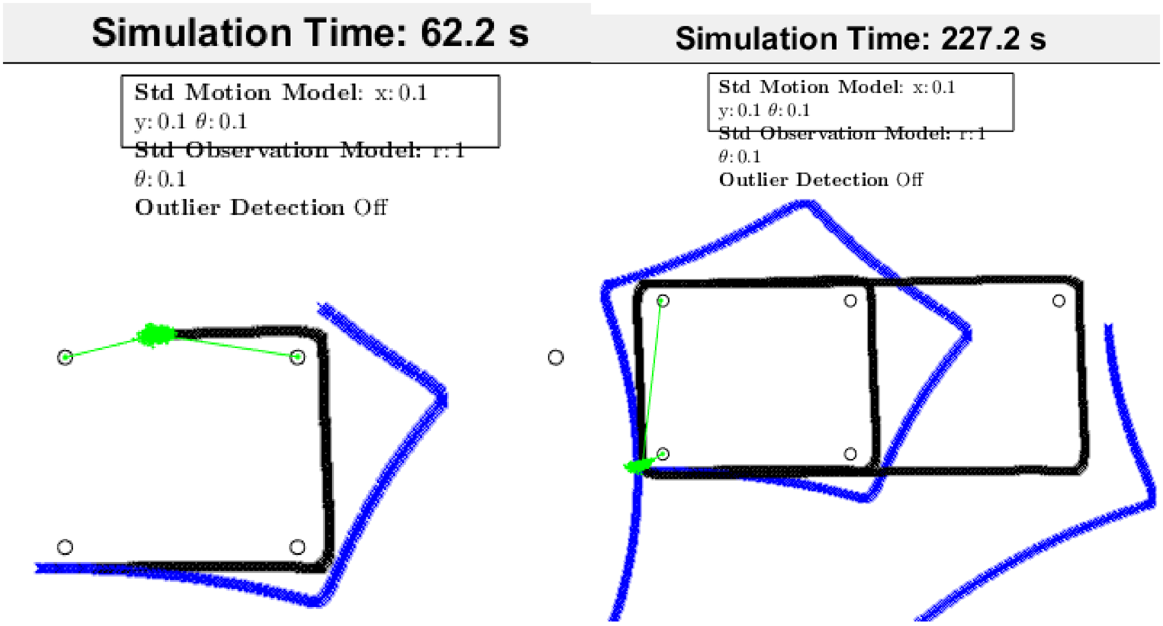


Fig 8. Localisation output for dataset 5 at different time instances