EL2320 - Applied Estimation - Lab 2

Particle Filter (PF)

PART II

Alexandra Hotti

November 2019

1 Warm-up Problem: Particle Filter

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

One advantage of using (5) is that its dimensionality is smaller compared to the representation in (7) and thus using (5) requires less computations in the PF.

A disadvantage of (5) is that the heading is set to a constant value throughout the entire motion. Therefore, since x and y are affected by noise the motion around the line becomes "jerky" in the heading direction around the line. In (7) θ is not fixed instead its value at time t depends on its value from the previous time step $x_{t-1,\theta}$. Therefore, the bearing is allowed to vary and the motion becomes continuous and smooth.

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

In equation (8) there are two fixed parameters that we need to know in advance: the angular velocity ω_0 and the translational velocity v_0. Thus, we can only model circular motion with constant angular and translational velocity.

Question 3: What is the purpose of keeping the constant part in the denominator of (10)?

The purpose of the constant denominator in the likelihood expression is normalization. For every observation z we vary the parameters that we condition on and therefore we get likelihoods instead of a valid probability distribution. To normalize the likelihoods into a valid distribution, i.e. to make them sum to 1, we divide the likelihood values by the sum over all the likelihoods.

Question 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?

In the Multinomial Re-Sampling method one random number is generated from a uniform distribution for every single particle. Thus, M random numbers are generated.

In the Systematic Re-Sampling method one single random number is generated and used in the re-sampling of all the M particles.

Question 5: With what probability does a particle with weight $w = \frac{1}{M} + \epsilon$ 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.

For multinomial sampling the probability of sampling a certain particle is w. Thus, since we can only sample or not sample a certain probability, the probability of not sampling a particle becomes: $1 - w$. Then the probability of not sampling a specific particle after M re-sampling steps becomes: $(1 - w)^M$. Then the probability of surviving M sampling rounds becomes: $1 - (1 - w)^M$.

For systematic sampling first a random number r is drawn from a uniform distribution in the range: $[0, \frac{1}{M}]$. From this number r we walk along the weights with a step size of $\frac{1}{M}$. Thus if a particle has a weight larger than the r we sample: $w = \frac{1}{M} + \epsilon$ it will survive with a probability of 1 since we will always end up in its bin. However, if the weight of a particle is in the range: $[0, \frac{1}{M}]$, its probability of surviving the sampling becomes at least: $M \cdot w$. In the smallest probability case the probability of being selected simply depends on the number of bins or equivalently the number of particles. Meaning that if a particle has a small weight the probability of it being resampled decreases with the number of bins/particles.

Question 6: Which variables model the measurement noise/process noise models?

- The process noise: $params.Sigma\_R$
- The measurement noise $params.Sigma\_Q$

Try to adjust the number of particles for a 2D state space to have a precise estimate for the fixed meas 1 measurement set (precise: mean absolute error should be less than 1 pixels). Compare now to the 3D state space without changing the parameters.

Number of Particles [M]	2D - Mean Absolute Error - Estimates	3D - Mean Absolute Error - Estimates
100	1.2 $\pm$0.9	1.2 $\pm$0.7
500	0.8 $\pm$1.3	0.8 $\pm$1.1
1000	0.6 $\pm$0.7	0.6 $\pm$0.4

Table 1: Mean absolute error evolution as the number of particles is varied. In the middle column a 2 dimensional motion model is used and in the last column a 3 dimensional state space is used. It appears as if the 3D model gives lower variance in the error and as if both the mean and the standard deviation of the error decreases as the number of particles increase.

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

When the process noise is set to 0 the number of particles rapidly decreases such that we are left with only one particle or one ancestor particle and its decedents. These particles will all be located in the exact same spot since we have removed the diffusion in the model transition. See Figure 1. Therefore, where this single point ends up depends on which points get resampled in the early time steps such that they get more descendants and thus particles in that spot are more likely to be resampled.

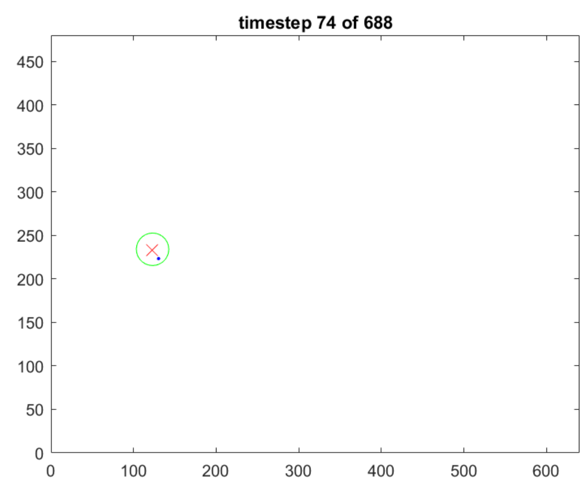


Figure 1: When the process noise is set to 0 the number of particles rapidly decreases such that we are left with only one particle or one ancestor particle and its decedents.

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

When we do not resample we keep all the initial particles that are spread out over all the possible states. The particles move around slightly from their initial state due to the transition noise but do not get far from their initial position and therefore many of these particles are located in low probability regions and therefore we cannot find the target location see Figure 2. Since the particles at the end are still spread out quite equally the mean of the posterior ends up close to the center of the image see Figure 3.

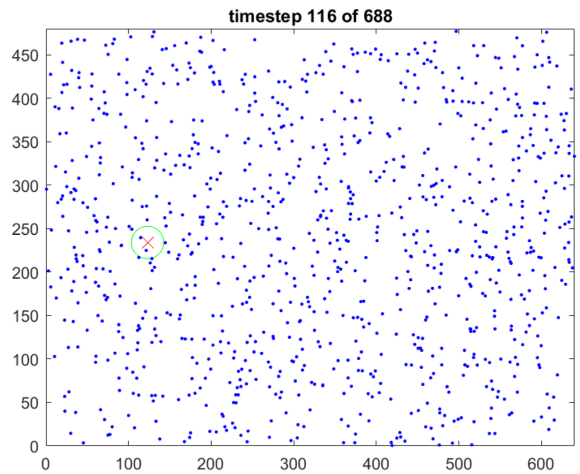


Figure 2: When we do not resample we keep all the initial particles that are spread out over all the possible states.

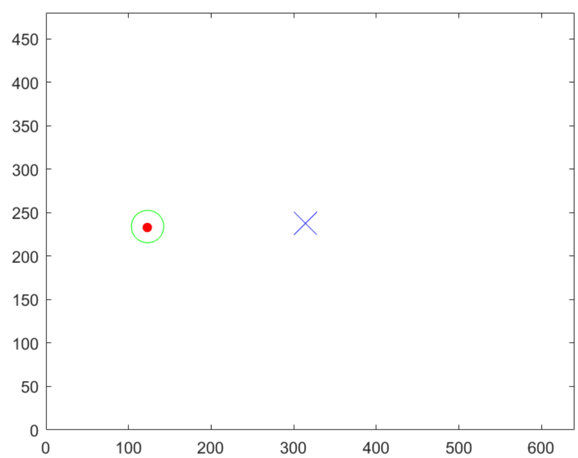


Figure 3: The center of the estimated posterior distribution.

Question 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 observations noise is set to a really low value the particle filter will not converge. See Figure 4. The low covariance affects the likelihood that we compute for the measurements, such that we become more strict about which points are regarded as inliers. Here we only accept very small deviations between the true and the predicted measurements for a point to not be regared as an outlier.

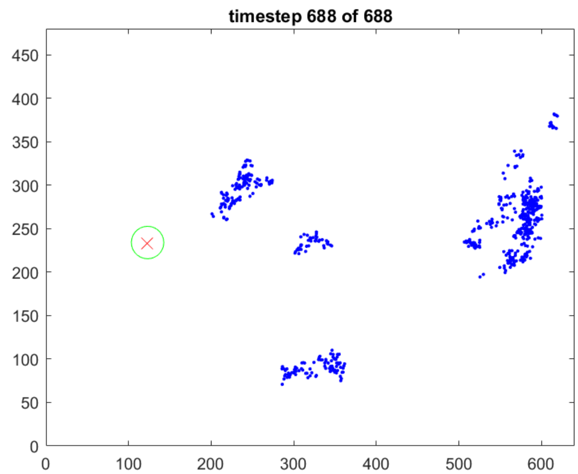


Figure 4: Low modeled observation noise means that almost all observations will be regarded as outliers and therefore the PF cannot converge.

On the other hand, when the elements in the covariance matrix are very large the particles converge quite rapidly to the true state. However, since the observation noise is so large we get a large spread around the true target state. See Figure 5.

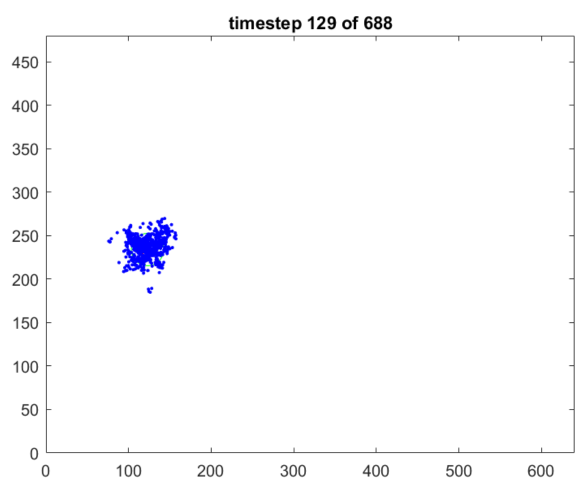


Figure 5: High modeled observation covariance noise results in rapid convergence. However, it comes at the cost of a large spread around the true state.

Question 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)

When a very small value is used for the modeled process noise the particles do not move around as much. They quickly close in on a few particles and whether or not these particles are close to the true state is affected by the particle initialisation. Convergence becomes slow.

When the process noise is very large the particle filter quickly converges to the true state. However, since the transition noise is large the particles jump around with a large spread around the true state. See Figure 6.

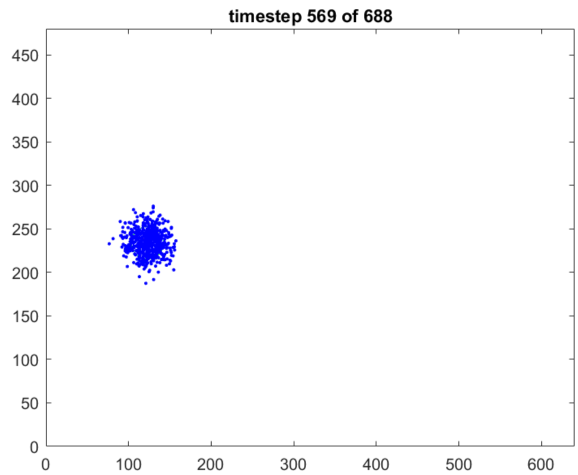


Figure 6: Large process noise.

Now, try the moving targets. For the next three questions try to think about the answer and write what you think will happen.

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

How much process noise we need in the motion model depends on how closely the motion model used actually models the true process. So if we have a nearly perfect model which depicts the true transitions in the motion, then we need very little noise. However, if we use a model which is not as good then we need to account for deviations from the model by incorporating more process noise.

Question 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 motion model that closely models the true transition process will yield more accurate results and need fewer particles since we are less uncertain about where the true state lies.

If the model instead is a simplified version of the true motion then we get less accurate results. Also, more noise is needed to account for deviations from the model and therefore we also need more particles since the distribution gets a larger variation around the modeled motion.

Question 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?

To detect outliers in general one can compute the likelihood for the observations conditioned on the particles at the current time step. In fact, this is already incorporated into the warm up code. In the code a measurement is regarded as an outlier if its mean likelihood is below a certain likelihood threshold. This threshold can be adjusted and should be set rather high since the third dataset contains many outliers.

Question 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?

Based on my previous answers, the circular motion model should be best at modeling these observations. This was in line with the error in the results, as can be seen in table 2

The linear and the circular motion models were both quite insensitive to the choice of parameters. However, the circular motion model was slightly less sensitive compared to the linear model. The most sensitive model was the fixed target motion model. Small changes in the parameter setting of the fixed motion model had a large impact on the model's error.

	Fixed Target Model	Linear Motion Model	Circular Motion Model
Error	13.1 ±5.2	8.6 ±4.4	7.9 ±4.2
Process Noise	diag(8, 8, 0.064)	diag(6.5, 6.5, 0.052)	diag(4.5, 4.5, 0.036)
Observation Noise	diag(250, 250)	diag(300, 300)	diag(300, 300)

Table 2: Best parameter setting for a fixed, linear and circular model for the dataset: *circ_meas_2*.

2 Main problem: Monte Carlo Localization

Question 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 $|Q| \to 0$?

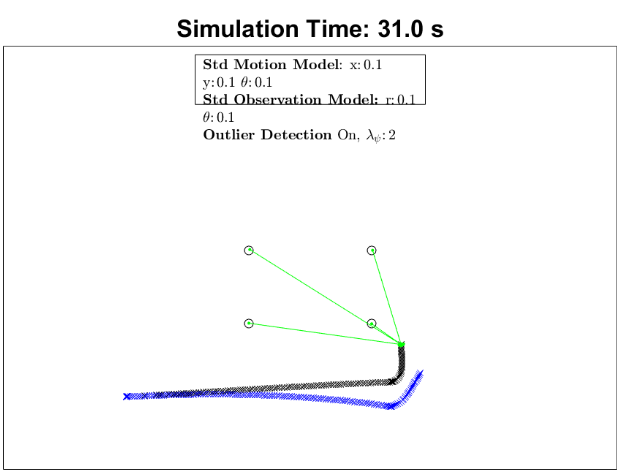
Two parameters affect the outlier detection: the likelihood threshold λ and the model observation noise Q.
When $|Q| \to 0$ the covariance of the likelihood decreases such that only measurement with very small deviations from the true observations are regarded as inliers. Also, we will assign really high likelihoods to measurements close to the true observation. If we were to set $Q = 0$, then we would only accept measurements that exactly correspond to the true observation.

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

If we do not remove the outliers before we calculate the weights then these unlikely measurements will be used to update the weights which ultimately will disrupt the posterior estimation.

Data Set 4

In the tracking problem the initial state is known and therefore the number of particles used is not as important compared to the localization problem. When the number of particles is increased from $M = 1000$ to $M = 10000$ the mean absolute error only slightly changes for the tracking problem.



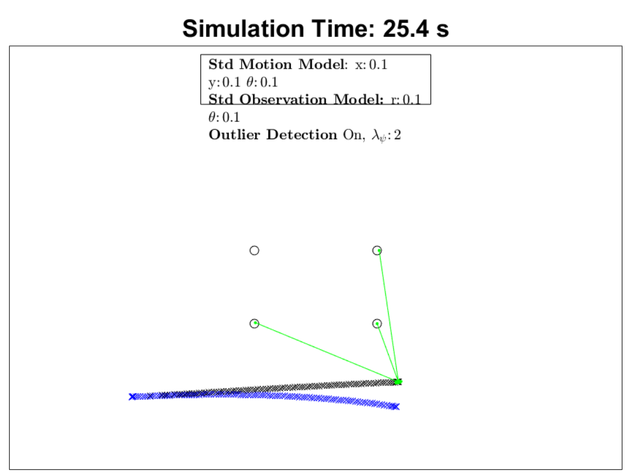


Figure 7: This is the results from a tracking problem with systematic resampling. To the left 10 000 particles are used and to the right 1000 particles are used in the filter.

In the global localization problem the number of particles used is of greater importance. In this environment their are 4 symmetric areas and therefore 4 hypothesis which we want the filter to reliably track. When 10 000 particles are used the filter can easily track the 4 hypotheses. See the left hand side in Figure 8. However, when only 1000 particles are used the number of particles are to few to cover all likely states, the filter suffers from particle deprivation and not all of the 4 hypotheses can be tracked. See the right hand side in Figure 8. Here only 2 hypothesis are left.

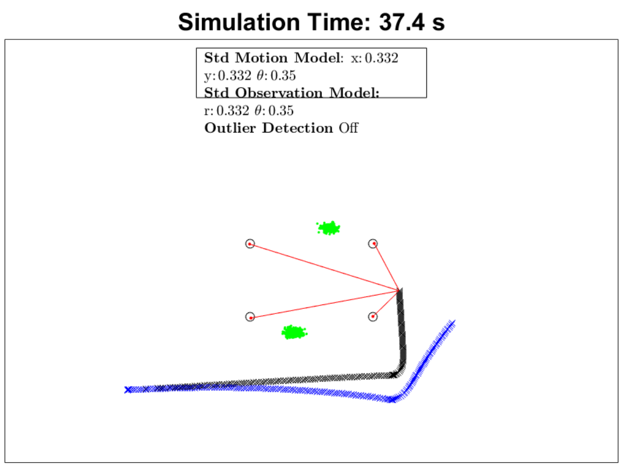


Figure 8: This is the results from a Global Localization problem with systematic resampling. To the left 1000 particles are used and to the right 10 000 particles are used in the filter.

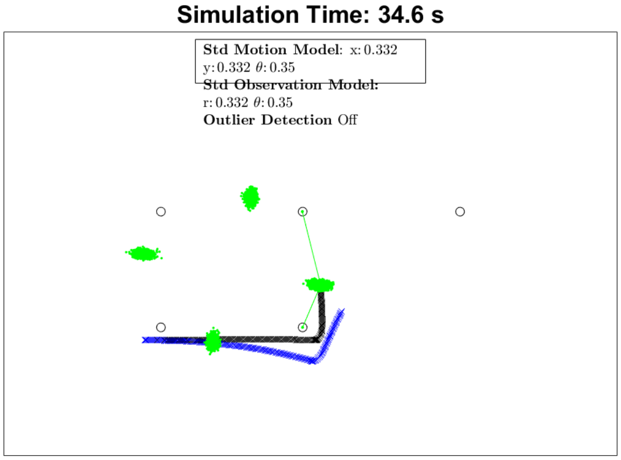
When multinomial sampling is used instead of systematic resampling, the filter becomes less reliable at tracking the 4 hypotheses.

Systematic resampling is better at preserving the clusters since it preserves the number of particles in each cluster by sampling the same number of times from each cluster as the number of particles in the cluster. In multinomial sampling on the other hand each particle will not be sampled from. Instead a larger weight will increase the chance of sampling such that some particles might be sampled from multiple times. Therefore, the number of particles in each cluster will likely not be preserved with multinomial sampling and therefore it is not as good at preserving the 4 hypotheses as to the systematic sampler.

When the measurement noise is higher the the particle filter is better at preserving the four likely state locations. However, the uncertainty around the possible states increases and thus the spread around the estimated states increases. When the measurement noise is weak it is difficult for the filter to preserve all 4 hypothesis.

Data Set 5

This is the results from a Global Localization problem with 10000 particles with systematic resampling on Dataset 5. Right before the robot has seen the landmark to the right, the four hypotheses all look probable. See the left hand side of Figure 9. In the right figure the robot has just seen the landmark to the right and thus only the true pose is probable out of the 4 hypotheses. Thus, the filter converges to the true hypothesis. See the right hand side of Figure 9.



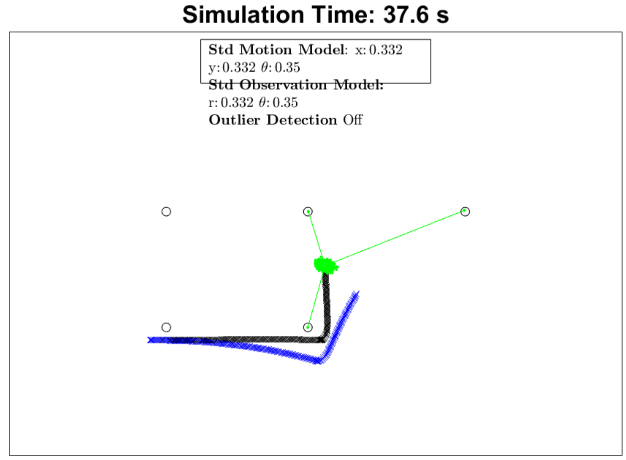


Figure 9: This is the results from a Global Localization problem with 10000 particles with systematic resampling on Dataset 5. Right before the robot has seen the landmark to the right, the four hypotheses are all probable. This is seen in the left figure. In the right figure the robot has just seen the landmark to the right and thus only the true pose is probable out of the 4 hypotheses.