

EL2320 - Applied Estimation

Lab II

Alexandros Filotheou

1 Part I

1.1 Question 1

Particles are called the samples from a posterior distribution.

1.2 Question 2

An *importance weight* $w_t^{[m]}$ is the probability of a measurement z_t regarding a specific particle $x_t^{[m]}$: $w_t^{[m]} = p(z_t|x_t^{[m]})$.

In the context of particle filters, *target distribution* is called the belief $bel(x_t)$, while *proposal distribution* is called the $\overline{bel}(x_t)$.

The target distribution cannot be sampled directly, however, it can be sampled indirectly by sampling from the proposal distribution. Each sample has a probability of being drawn equal to its importance weight.

1.3 Question 3

Even though a small number of particles results in a higher probability of particle deprivation, the main cause of *particle deprivation* is the variance introduced by random sampling. The danger behind particle deprivation is that there may be no particles near the true state of the system.

1.4 Question 4

The resampling step is necessary in order to estimate the target distribution: it forces the particles to be distributed according to the posterior $bel(x_t)$. If, on the other hand, no resampling took place, more particles would reside in areas with a low posterior probability, and would thus be wasted to regions of no real interest.

1.5 Question 5

In the case of a multimodal posterior, the average of the particle set would fail to account for the multiple regions of true interest, i.e. the regions where the true state of the system has the highest probability of being.

1.6 Question 6

Inferences about states that lie between particles can be made using histograms, where a bin is comprised by a collection of adjacent particles, or Gaussian kernels, where each particle is used as the center of a Gaussian kernel.

1.7 Question 7

High sample variance can cause an incorrect estimation about the true state of a system. One solution is to reduce the resampling frequency. The second one uses a sequential stochastic process instead of drawing samples independently from each other.

1.8 Question 8

Pose uncertainty and number of particles used are proportionally related. The higher the uncertainty about the true state of the system, the larger the spread of the posterior distribution, hence the more particles are needed to estimate the potentially different modes of the distribution.

2 Part II

2.1 Question 1

In 2D space, equation 6 uses the exact same value for the angle θ_0 for all timesteps t . This equation would accurately estimate the line of travel if no noise was introduced to the predicted state. However, because noise is present, the estimate of the motion will result in a zig-zag pattern, unlike the one in 3D space. In 3D space, instead of a fixed angle, the control model uses the angle of travel from the previous timestep, which was infected by noise, hence its estimate of the motion can be smoother than in the previous case.

2.2 Question 2

Since the initial translational and angular velocities u_o, ω_0 need to be set in advance, the only circular motions that equation 9 can model are instances of a specific circle, with a specific direction of travel.

2.3 Question 3

The constant part in equation 10 serves for normalization purposes. The integral over the entire distribution needs to equal to 1.0, since the distribution represents a probability distribution.

2.4 Question 4

Algorithm 2 illustrates the need for M random numbers to be generated for the Multinomial re-sampling, where M is the number of particles used. In contrast, algorithm 3 shows that only one random number needs to be generated in the context of Systematic re-sampling.

2.5 Question 5

In Multinomial re-sampling, a particle $p^{[k]}$ with an importance weight $w^{[k]}$ attached to it has a probability of being selected equal to $w^{[k]}$, and a probability of not being selected equal to $1 - w^{[k]}$. The probability of it not being selected N consecutive times, since re-sampling takes place with replacement, is $(1 - w^{[k]})^N$. Hence, the probability of it being selected N consecutive times is $1 - (1 - w^{[k]})^N$.

In Systematic re-sampling, a particle $p^{[k]}$ with a weight $w^{[k]} = M^{-1} + \epsilon, 0 < \epsilon \ll 1$ attached to it will always be selected since the selection step size $M^{-1} < M^{-1} + \epsilon$. However, if $0 \leq w^{[k]} < M^{-1}$, the probability of its survival is $w^{[k]} \cdot M$. Since the random number r is selected in the interval $[0, M^{-1}]$, assuming a uniform distribution for the its selection, and given the fact that the probability of selection of a particle is proportional to its weight, the probability of its selection is $\frac{w^{[k]} - 0}{M^{-1} - 0} = w^{[k]} \cdot M$.

2.6 Question 6

Measurement noise is modelled by $\text{Sigma_}Q$, while process noise is modelled by $\text{Sigma_}R$.

2.7 Question 7

If the process noise is set to zero, then the population of particles ends up being formed by M copies of the same particle. This is reasonable, since the process noise is essentially a factor which contributes to particle diversity.

2.8 Question 8

If no re-sampling is performed, then the population of particles is not transformed in any way. The consistency of the populations stays the same throughout the experiment, and is comprised by the random particle set initialized at $t = 1$. The motion of each particle simulates the estimation of the true state of the system from a different initial state and no convergence is made.

2.9 Question 9

Low values for the standard deviation of the measurement noise ($Q(1,1) = Q(2,2) \sim 0.0001$) mean that we expect the estimate of the measurements to be very accurate. If these estimates do not follow the true measurements closely, then they are regarded as outliers. The end result is that the estimate does not converge to the true state. Increasing the standard deviation to values $Q(1,1) = Q(2,2) \sim 10000$ means that our uncertainty about the measurements increases, and so does the variance of the particles around the true state.

2.10 Question 10

Low values for the standard deviation of the process noise ($Q(1,1) = Q(2,2) \sim 0.0001$) decrease the diversity of the particle population. This is exhibited in the formation of a dense particle cloud which oscillates around the true state, with convergence not being guaranteed. As the standard deviation increases, so does the probability of particles gaining more weight because of the larger variance in their position, which makes them estimate the true state quicker. The end result is a particle population spread evenly around the true state. The radius of the particle cloud is proportionally related to the variance of the process noise.

2.11 Question 11

If the motion model models the true motion accurately, then the process noise can be kept low, since there is enough confidence in the estimate of the true motion. On the other hand, if the motion model is known to be inaccurate, then the process noise should be larger, in order for particles to be able to estimate the true motion.

2.12 Question 12

A motion model that describes the actual motion in an accurate manner will have higher precision and accuracy. Furthermore, fewer particles will be needed, since the overall accuracy of the particles will be high. In contrast, an inaccurate motion model needs a larger particle population, so that the system can converge to the actual state.

2.13 Question 13

When there is a large discrepancy between a measurement and its prediction, the likelihood function takes a small value. Hence, we could install a threshold to monitor this discrepancy and discard what we assume to be outliers.

2.14 Question 14

The following three sections illustrate the best settings for the covariance matrices R, Q . As expected, the fixed motion model required the injection of higher levels of process noise in order for the particles to be able to follow the true state, and was the most sensitive to it. The linear and circular models were comparable, although the former required more measurement noise in order to be close to the latter. Furthermore, the latter's best approximation of the true motion allowed for overall lower levels of noise.

2.14.1 Fixed Motion

Adjusting the covariance matrices R, Q as follows, it is possible to minimize the error to a value of 11.7 ± 5.4 .

$$R = \begin{bmatrix} \mathbf{10} & 0 & 0 \\ 0 & \mathbf{10} & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$$
$$Q = \begin{bmatrix} 200 & 0 \\ 0 & 200 \end{bmatrix}$$

2.14.2 Linear Motion

Adjusting the covariance matrices R, Q as follows, it is possible to minimize the error to a value of 7.1 ± 3.8 .

$$R = \begin{bmatrix} \mathbf{3} & 0 & 0 \\ 0 & \mathbf{3} & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$$
$$Q = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix}$$

2.14.3 Circular Motion

Adjusting the covariance matrices R, Q as follows, it is possible to minimize the error to a value of 6.8 ± 3.7 .

$$R = \begin{bmatrix} \mathbf{2} & 0 & 0 \\ 0 & \mathbf{2} & 0 \\ 0 & 0 & \mathbf{0.005} \end{bmatrix}$$
$$Q = \begin{bmatrix} 200 & 0 \\ 0 & 200 \end{bmatrix}$$

3 Part III

3.1 Question 15

The two main parameters that have a direct impact on the outlier detection approach are the set threshold λ_Ψ and the measurement noise Q . As $|Q| \rightarrow 0$, the confidence in the correctness of the measurements increases, and the level of tolerance for discrepancies between the prediction of measurements and their true values decreases. This means that more and more measurements are classified as outliers. Furthermore, the same result is observed for higher values of λ_Ψ . The higher its value is, the higher we demand the likelihood a particular measurement to be, and the more probable it is that measurements are to be discarded as outliers.

3.2 Question 16

If no outliers are discarded, then all measurements are considered to be valid. However, in the general case, not all measurements are to be equally trusted. Since even erroneous measurements will be taken into account, the weights of the particles are going to be erroneously distributed and the estimation of the true state will be more incorrect than it would have been otherwise.

3.3 `map_sym2 + so_sym2_nk`

Looking at the code provided, `part_bound` seems to influence the x, y coordinates of the initial position of each particle in the global localization case. The initial position for a particle is given by

$$init(a) = r \cdot l_{max} + (1 - r) \cdot l_{min} + \text{part_bound} \cdot (2r - 1) \quad (1)$$

where a, l represent either x, y coordinates of a particle and a landmark respectively, and $r \in [0, 1]$ is a random value. From equation 1 we infer that the higher the value of `part_bound`, the larger the variance of the initial position of a particle around the outermost landmarks of map W . Hence, the higher the value of `part_bound`, the higher the probability that particles will occupy more spread-out areas near these symmetrically placed landmarks, which will result in a higher probability of the particles being able to distinguish them and keep them as reliable hypotheses.

Figures 1, 2 and 3 illustrate the resulting trajectories for tracking with $M = 1000$ and $M = 10000$ particles, using the default noise settings.

For the rest of the analysis we shall model stronger process noise with process noise and measurement covariance matrices R, Q :

$$R = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$Q = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

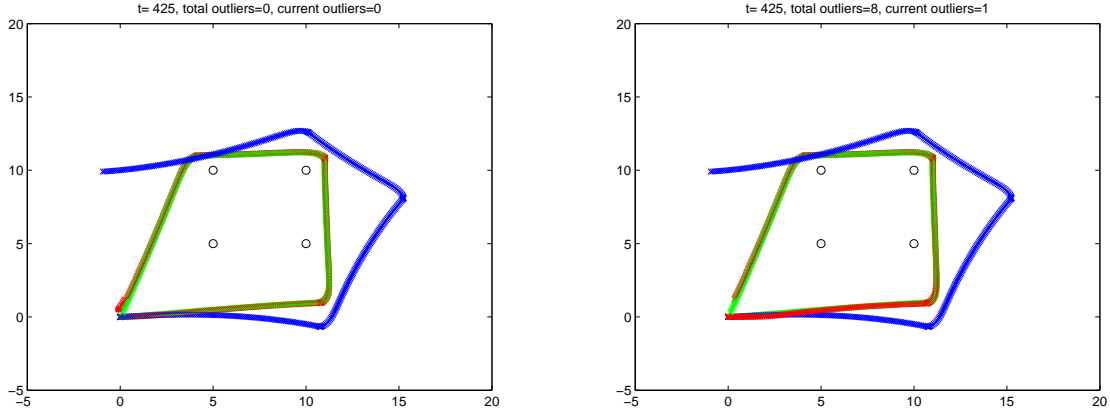


Figure 1: Resulting trajectories for $M = 1000$ (left) and $M = 10000$ (right).

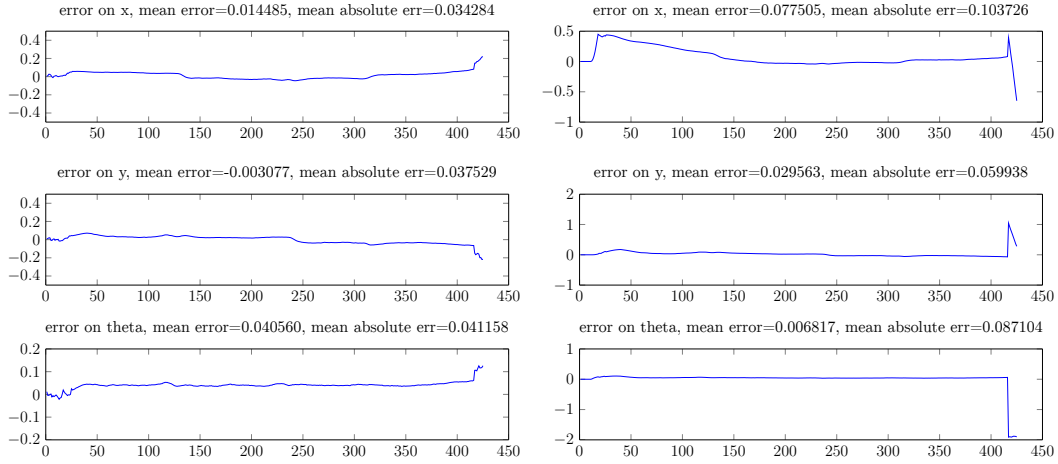


Figure 2: Mean pose errors for $M = 1000$ (left) and $M = 10000$ (right).

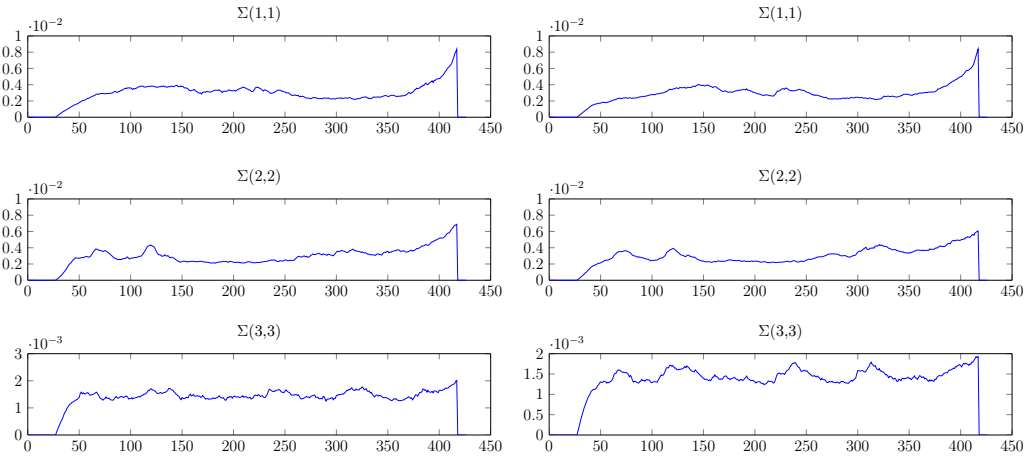


Figure 3: Uncertainties in pose for $M = 1000$ (left) and $M = 10000$ (right).

In this map setting there are 4 hypotheses, one for every landmark. Increasing the number of particles M to $M = 10000$ from $M = 1000$ results in these hypotheses being kept for longer and with more reliability. The main effect observed with a smaller number of particles is particle deprivation. This effect is proportionally related to the level of process noise: the stronger the process noise, the less particle diversity is observed, hence the more acute is the particle deprivation effect.

Regarding multinomial sampling, it is more or less evident that it is not as robust or reliable in preserving multiple hypotheses, since it introduces sample variance and sequential sampling tackles this problem effectively.

Finally, the higher the measurement noise, the more spread-out are the particles around the hypotheses, which are still preserved, but with lower accuracy. On the other hand, the lower the measurement noise, the less the degree in which the hypotheses are kept. When it reaches a certain threshold, we observe no relation between the particle distribution and the given hypotheses.

3.4 map_sym3 + so_sym3_nk

Figure 4 illustrates the convergence of the particle filter. Until timestep $t = 180$ the upper right landmark has not been observed, hence the robot keeps all 4 hypotheses as to where it might be due to the symmetry of the 4 landmarks it has seen. However, when it observes the 5th landmark, this symmetry is resolved and it can reliably locate itself.

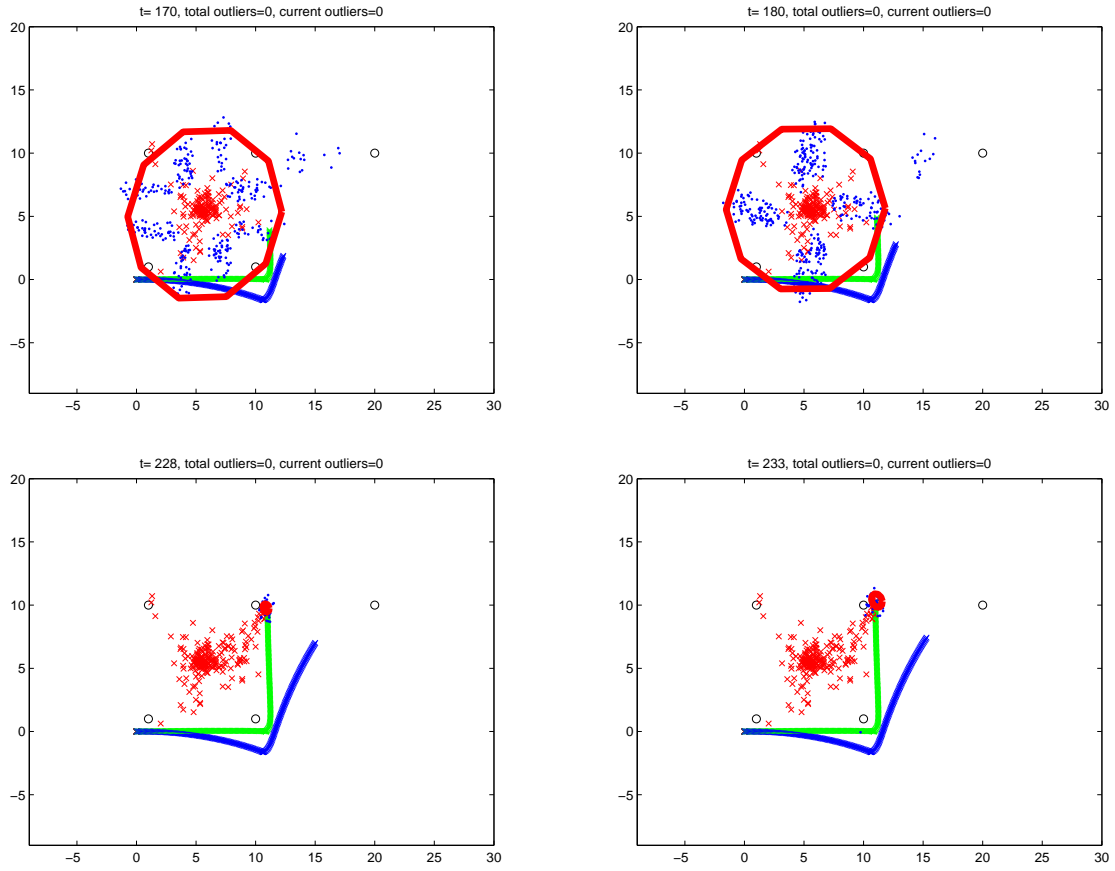


Figure 4: Convergence of the PF to one single hypothesis due to symmetry resolution.