Preparatory Questions

Linear Kalman Filter:

1. What is the difference between a 'control' u_t, a 'measurement' z_t and the state x_t? Give examples of each?

A state is the collection of all aspects of an actor and it's environment that can impact the actor and its environment in the future, such as the pose of a robot. A control gives information about the change in state of an environment and is used to get the current state estimate (this can be a measurement, or an actual control, such as an input to a motor) and a measurement (such as a sensor reading) gives the momentary state of the environment.

2. Can the uncertainty in the belief increase during an update? Why (or not)?

No, during the update step sensor data is incorporated into the belief, which provides more information and inherently decreases the uncertainty. If the measurement is noisy, the Kalman gain will take this into account as it specifies the degree to which the measurement is incorporated into the belief.

3. During update what is it that decides the weighing between measurements and belief?

The Kalman gain

4. What would be the result of using a too large a covariance (Q matrix) for the measurement model?

If there is too large a covariance in the measurement model, then the Kalman gain will be smaller than necessary and will place a lower weight on the first measurements than needed.

5. What would give the measurements an increased effect on the updated state estimate?

A large Kalman gain, which would occur with a small Q matrix, low variability in measurements, and a high belief covariance. These indicate that the measurements can be trusted more, and that the belief is still estimated to change a lot, so the gain must be higher to accommodate this.

6. What happens to the belief uncertainty during prediction? How can you show that?

The uncertainty is represented by the following step:

$$\bar{\Sigma} = A\Sigma A^T + Q$$

A is arbitrary, and the uncertainty usually increases because the dynamic model generally adds noise to the state. The uncertainty will always increase if (A-I) > 0.

7. How can we say that the Kalman filter is the optimal and minimum least square error estimator in the case of independent Gaussian noise and Gaussian priori distribution? (Just describe the reasoning, not a formal proof.)
The update step of a Kalman filter is effectively an implementation of a recursive least squares estimator, since they both create an estimate dependent on the previous estimation, and add a factor of a gain times an innovation. Since the Kalman filter also includes a prediction step where the Kalman gain is tuned, the Kalman filter is a more accurate/optimal estimator.

8. In the case of Gaussian white noise and Gaussian priori distribution, is the Kalman Filter a MLE and/or MAP estimator?
In the case that a prior distribution is known, the Kalman filter is a MAP estimator, since it can incorporate this prior distribution into its prediction of the current state estimate.

Extended Kalman Filter:
9. How does the extended Kalman filter relate to the Kalman filter?
The EKF only differs from a standard KF in that it models the state transition and measurement model as non-linear as opposed to the linear models used in the standard KF.

10. Is the EKF guaranteed to converge to a consistent solution?
No, it is not. Since it provides an approximation of the state using linearization, if a poor model is used for the state transition or if a bad initial state estimate is used, the EFK can diverge.

11. If our filter seems to diverge often can we change any parameter to try and reduce this?
If the filter diverges often, a better state transition model or different initial state estimate should be used.

Localization:

12. If a robot is completely unsure of its location and measures the range r to a know landmark with Gaussian noise what does its posterior belief of its location $p(x,y,\theta|r)$ look like? A formula is not needed but describe it at least.

The robot can be located anywhere in a circle with radius = r around the landmark. With Gaussian noise, the distribution would look like a donut. Since there is no information on θ. it will be a uniform distribution.

13. If the above measurement also included a bearing how would the posterior look?

The posterior would look the same, except now the robot's bearing and the angle of the robot relative to the landmark will be correlated.

14. If the robot moves with relatively good motion estimation (prediction error is small) but a large initial uncertainty in heading θ how will the posterior look after traveling a long distance without seeing any features?

Since the initial heading angle is unknown, the endpoint can be anywhere along a circle path starting from the robot's initial location. It will have a good idea of the distance travelled, but not the final position.

15. If the above robot then sees a point feature and measures range and bearing to it how might the EKF update go wrong?

The EKF could linearize around the wrong point, giving a poor estimate of the state.

Lab 1

Question 1: What are the dimensions of εk and δk? What parameters do you need to define in order to uniquely characterize a white Gaussian?

εk has dimensions of 2x1, δk is 1x1. To characterize a white Gaussian, since the mean is already defined (zero mean), you just need to define the covariance.

Question 2: Make a table showing the roles/usages of the variables(x, xhat, P, G, D, Q, R, wStdP, wStdV, vStd, u, PP). To do this one must go beyond simply reading the comments in the code to seeing how the variable is used. (hint some of these are our estimation model and some are for simulating the car motion).

x	Actual state
xhat	Predicted state
P	Covariance of predicted state
G	Identity matrix used for scaling R
D	Identity matrix used for scaling Q
Q	Covariance of measurement noise
R	Covariance of modelled motion noise (process noise)
wStdP	Standard deviation of predicted state noise
wStdV	Standard deviation of motion noise
vStd	Standard deviation of measurement noise
u	Control vector
PP	Stored predicted state covariance values

Question 3: Please answer this question with one paragraph of text that summarizes broadly what you learn/deduce from changing the parameters in the code as described below. Choose two illustrative sets of plots to include as demonstration.

What do you expect if you increase/decrease the covariance matrix of the modeled (not the actual simulated) process noise/measurement noise 100 times(one change in the default parameters each time) for the same underlying system? Characterize your expectations. Confirm your expectations using the code (save the corresponding figures so you can analyze them in your report).

Do the same analysis for the case of increasing/decreasing both parameters by the same factor at the same time. (Hint: It is the mean and covariance behavior over time that we are asking about.)

If the measurement noise, Q, were to increase, then I would expect the Kalman gain to decrease as the measurements are less reliable and should be given less weight in the prediction. It is also clear from the following step in the Kalman filter algorithm that K is proportional to 1/Q:

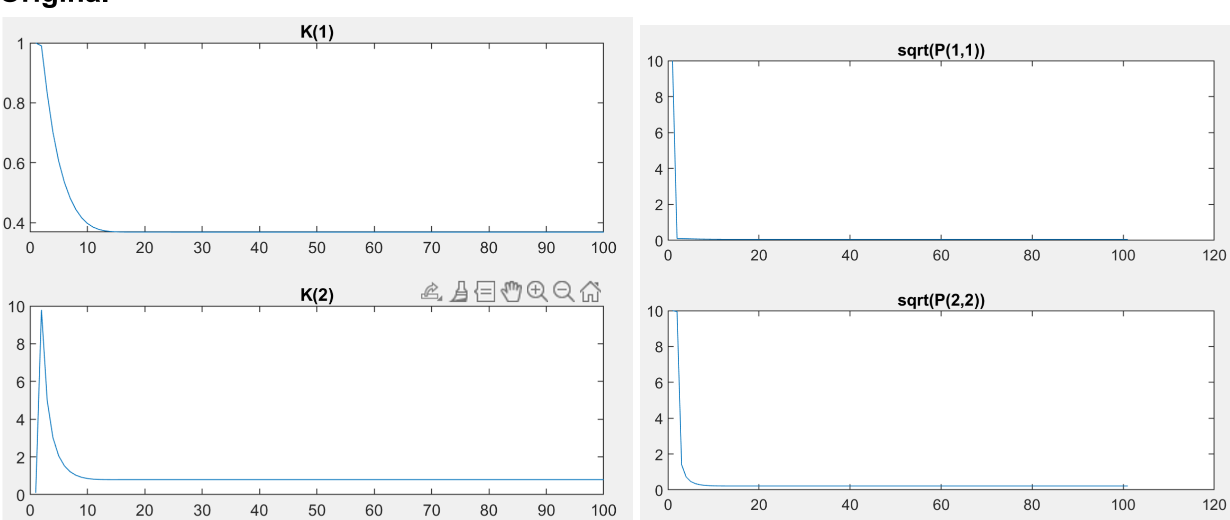
$$K_t = \bar{\Sigma}_t \, C_t^T (C_t \, \bar{\Sigma}_t \, C_t^T + Q_t)^{-1}$$

If the process noise, R, were to increase, I would expect an increase in the Kalman gain, as the measurements would be seen as more reliable than the model and should be weighted more. In other words, the prediction would be seen as more unreliable. I would also expect the covariance to increase, from the following prediction step in the Kalman filter algorithm:

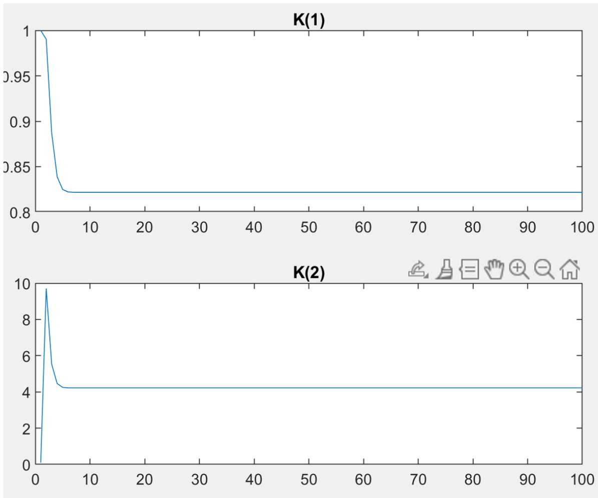
$$\bar{\Sigma}_t = A_t \, \Sigma_{t-1} \, A_t^T + R_t$$

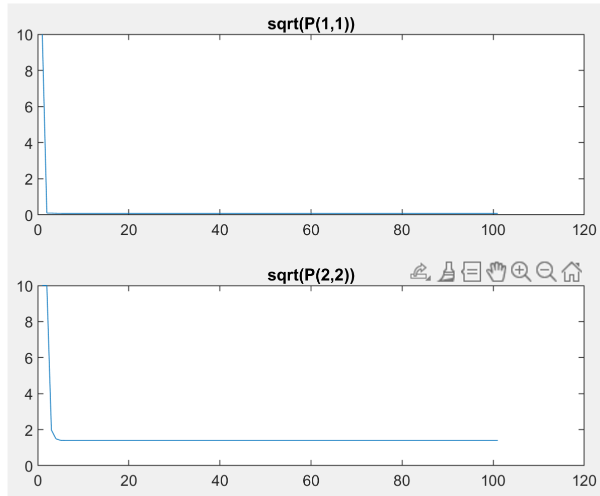
If both R and Q were to increase, I would expect the filter to take longer to converge, since both the measurement and model are unreliable, and I would expect the Kalman gain to increase, due to the Kalman gain equation above. I would expect the opposite to be true when both R and Q decrease.

Original

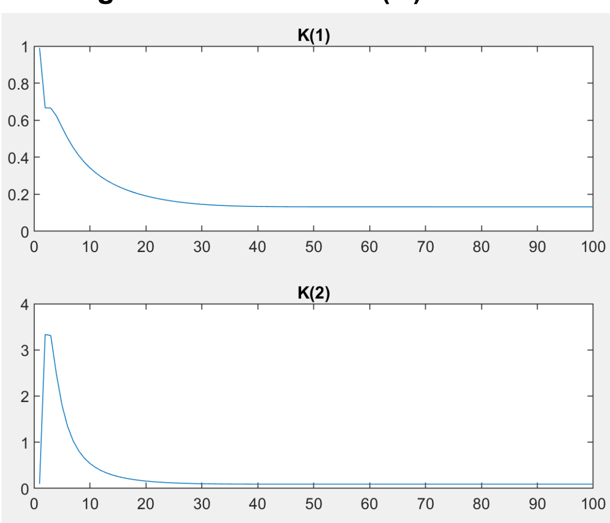


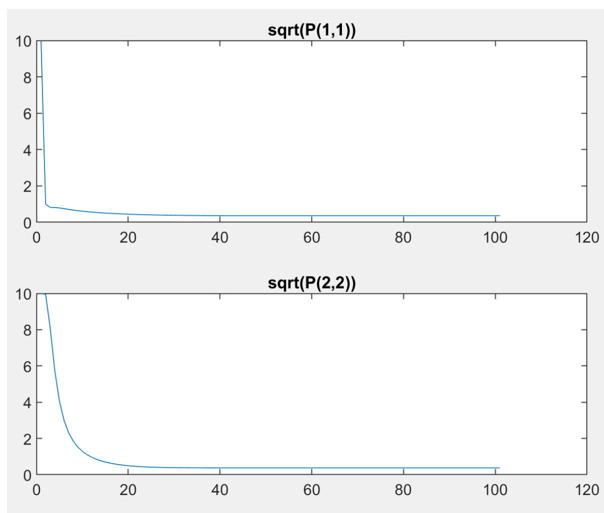
100x higher process noise (R)



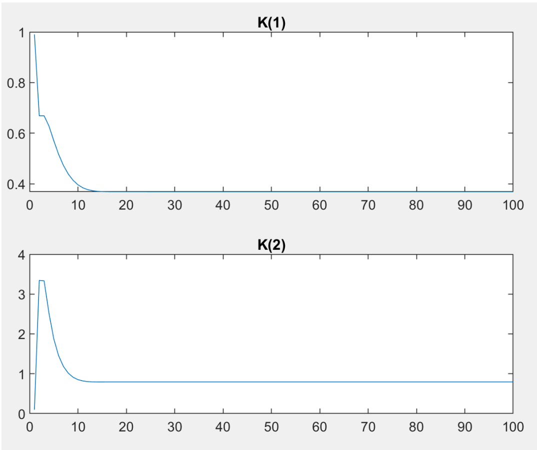


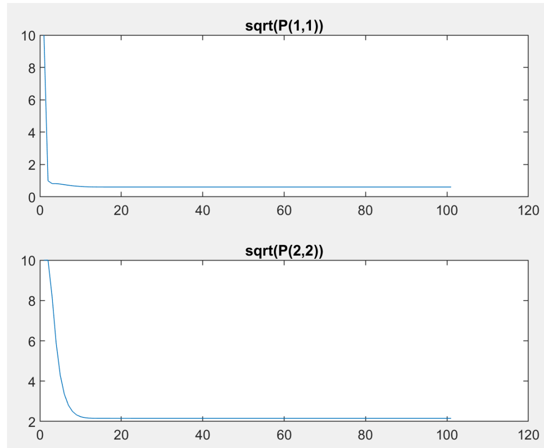
100x higher variance noise (Q)



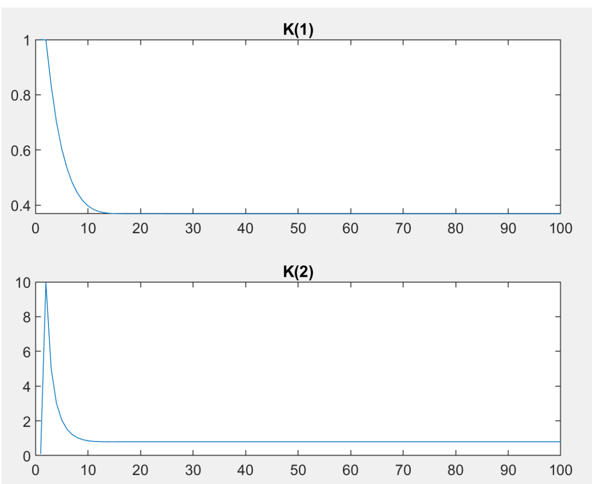


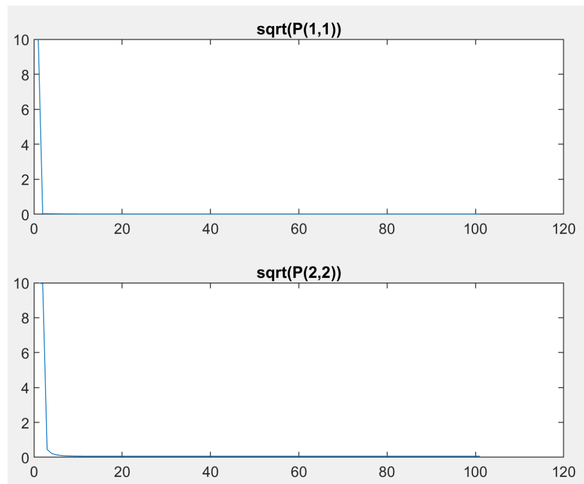
100 times R and Q





0.1 times R and Q





Question 4: How do the initial values for P and xhat affect the rate of convergence and the error of the estimates (try both much bigger and much smaller)?

Increasing the initial value for P has no impact on the rate of convergence and estimate error. However, decreasing the value of P makes the rate of convergence slower, but still does not change the error estimate. This makes sense, because it will take longer for the filter to handles a low initial covariance, but it should not affect the accuracy of the end result.

Changing the value of xhat does not change the rate of convergence or error of the estimate because the Kalman filter can correct for a poor initial state estimate using the measurement data.

Question 5: Which parts of (2) and (3) are responsible for prediction and update steps?
For both equations:

Prediction step

$$\int p(\mathbf{x_t}|\mathbf{u_t}, \mathbf{x_{t-1}})bel(\mathbf{x_{t-1}})\mathbf{dx_{t-1}}$$

Update step

$$\eta p(\mathbf{z_t}|\mathbf{x_t}, M)\overline{bel}(\mathbf{x_t})$$

Question 6: In the maximum likelihood data association, we assumed that the measurements are independent of each other. Is this a valid assumption? Explain why.

This is a valid assumption because we are using a white noise model for the measurement error of the white noise, so there is no correlation between two different measurements that can be discerned from the noise model . Also, the current measurement prediction does not directly depend on any previous measurements, it is predicted using the estimate of the current state.

Question 7: What are the bounds for δM in (8)? How does the choice of δM affect the outlier rejection process? What value do you suggest for λM when we have reliable measurements all arising from features in our map, that is all our measurements come from features on our map? What about a scenario with unreliable measurements with many arising from so called clutter or spurious measurements?

δM is a probability, so it has to be between 0 and 1 inclusive. If δM is higher, more measurements are accepted, and if it lower, less measurements are accepted. When the measurements are reliable, we want to keep most of them, which means we want most of the measurements to fall under the threshold. Therefore, λM should be large. The opposite is true if the measurements are unreliable- we want to keep less of them, so the threshold λM should be small.

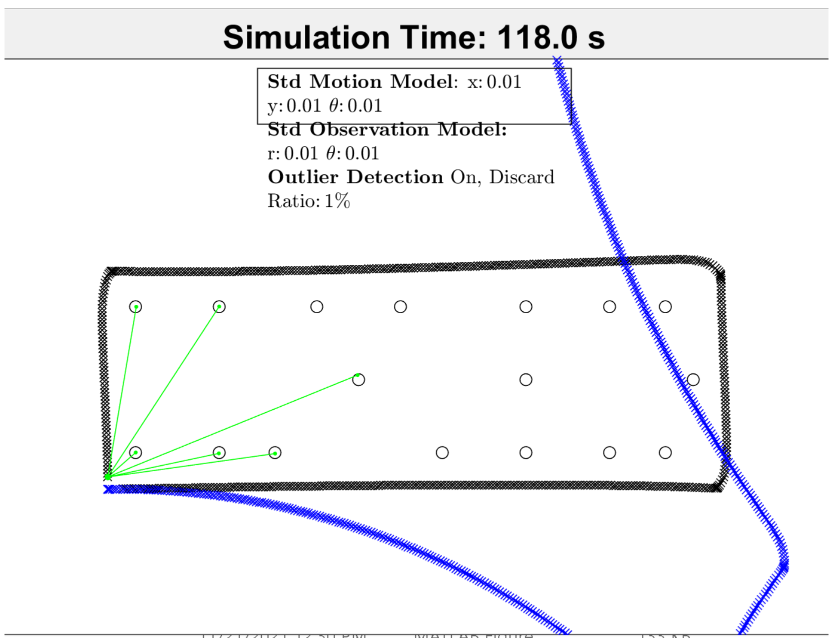
Question 8: Can you think of some down-sides of the sequential update approach(Alg 3)? Hint: How does the first [noisy] measurements affect the intermediate results?
If the first estimate is noisy (an outlier), and you perform an update with it, then the estimate may be poor and the uncertainty is artificially reduced. This can affect the intermediate results by incorrectly classifying them as outliers relative to the poor initial first measurement. This makes the sequential update process sensitive to error.

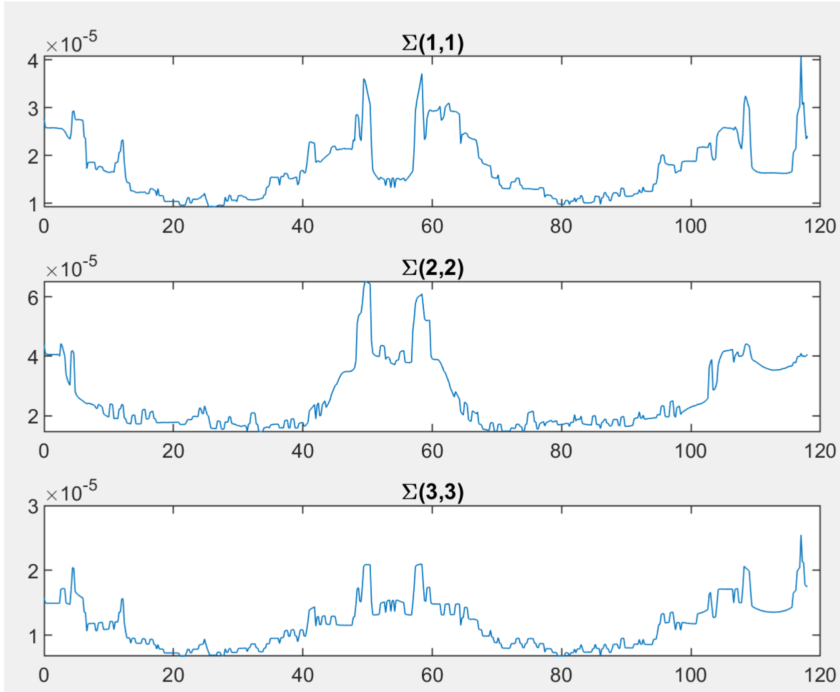
Question 9: How can you modify Alg 4 to avoid redundant re-computations?
Instead of iterating through every landmark in the map, there could be a way to potentially partition the map into sections, so then it could iterate just through the landmarks that correspond to the current section of the map that the robot is estimated to be located in. These

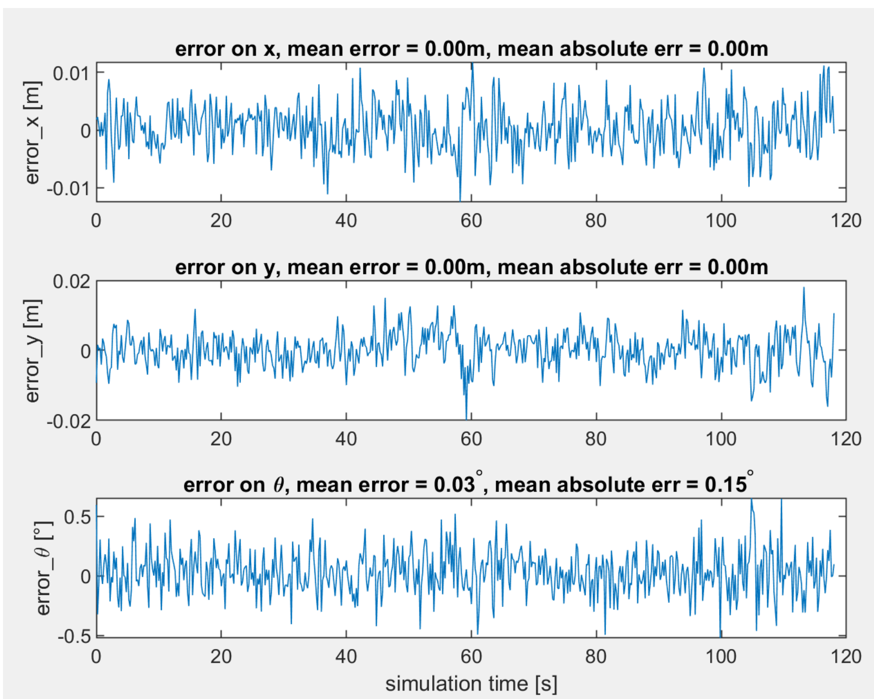
sections could be larger or smaller depending on the predicted error of the current state. In this way, only a subset of the landmarks need to be looked at instead of the whole map.

Question 10: What are the dimensions of v¯t and H¯ t in Alg 4? What were the corresponding dimensions in the sequential update algorithm? What does this tell you?
The dimensions of v¯t and H¯ t in Alg 4 are 2nx1 and 2nx3, respectively, while in the sequential update they were 2x1 and 2x3. This is because multiple measurements are being processed at the same time, so the order that the measurements arrive and/or a single noisy measurement does not affect the outcome as it does in the sequential update.

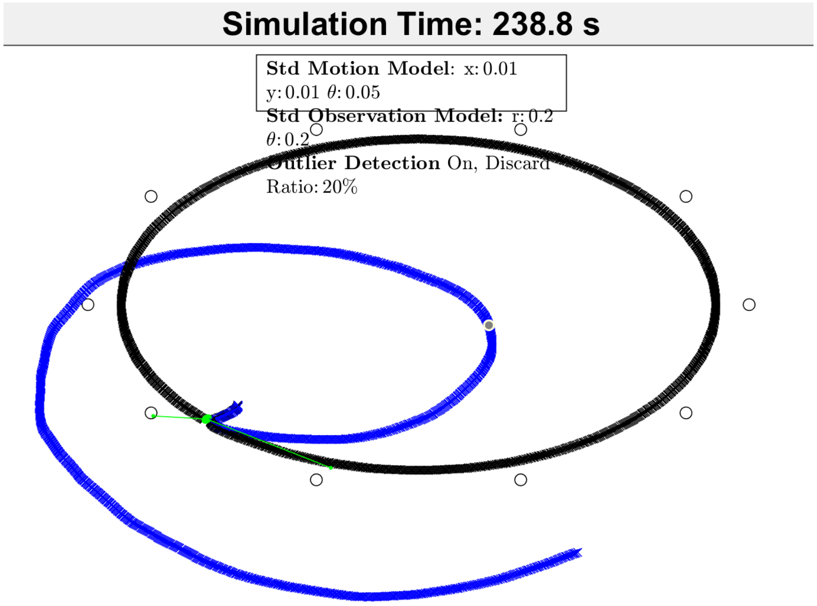
Dataset 1
The odometry estimate is not very accurate for the first two data sets, but the observation model is able to make up for the shortcoming and still produce an x,y, and theta within the specified ranges. Dataset 2 has a more challenging map with less landmarks to reference as the robot moves along the path and a very symmetric layout of landmarks, which can make it more difficult to identify individual ones. However, throwing out the outliers and modelling the noise correctly still allows a final pose within the desired error range.

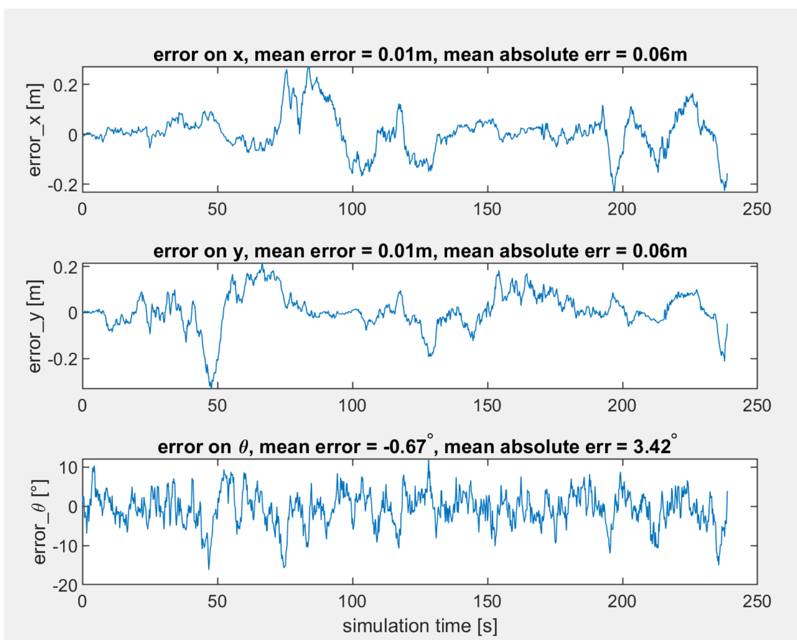


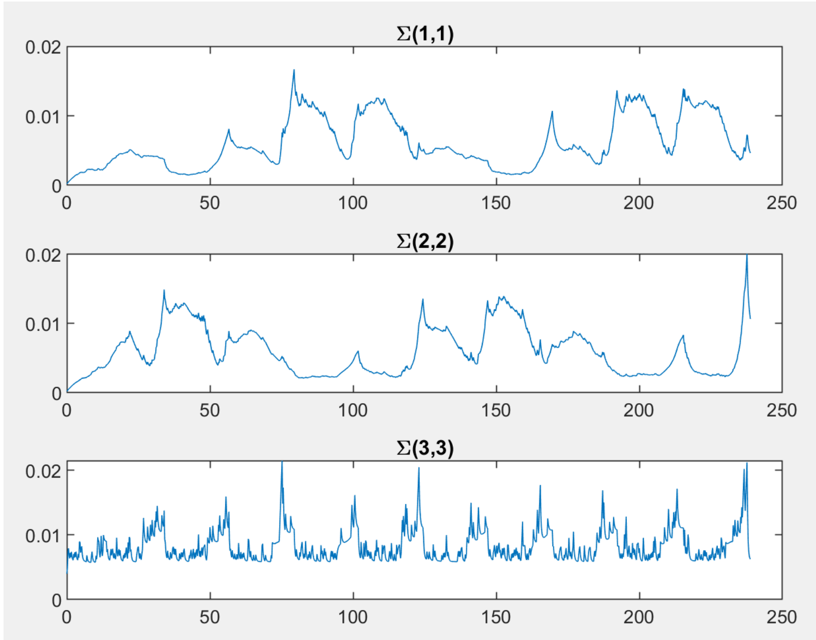

Σ(1,1)
×10^-5
Σ(2,2)
×10^-5
Σ(3,3)
×10^-5


error on x, mean error = 0.00m, mean absolute err = 0.00m
error_x [m]
error on y, mean error = 0.00m, mean absolute err = 0.00m
error_y [m]
error on θ, mean error = 0.03°, mean absolute err = 0.15°
error_θ [°]
simulation time [s]

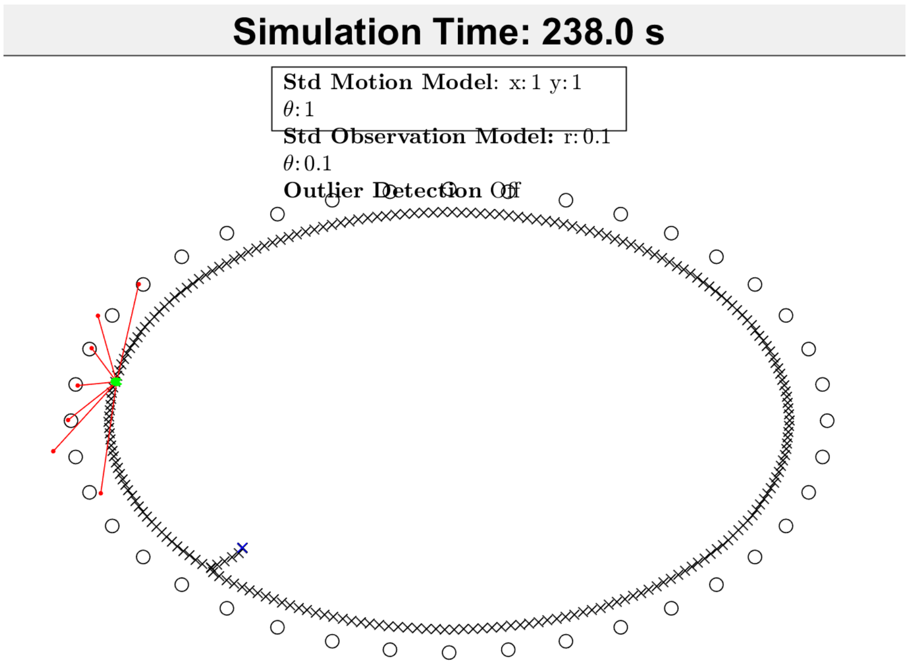
Dataset 2


Simulation Time: 238.8 s
Std Motion Model: x: 0.01
y: 0.01 θ: 0.05
Std Observation Model: r: 0.2
θ: 0.2
Outlier Detection On, Discard
Ratio: 20%

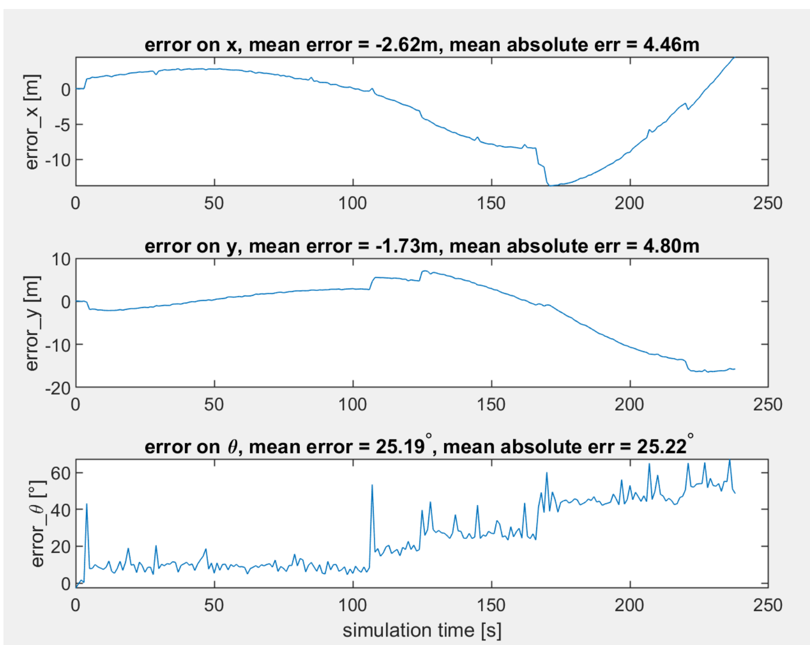

error on x, mean error = 0.01m, mean absolute err = 0.06m
error_x [m]
error on y, mean error = 0.01m, mean absolute err = 0.06m
error_y [m]
error on θ, mean error = -0.67°, mean absolute err = 3.42°
error_θ [°]
simulation time [s]



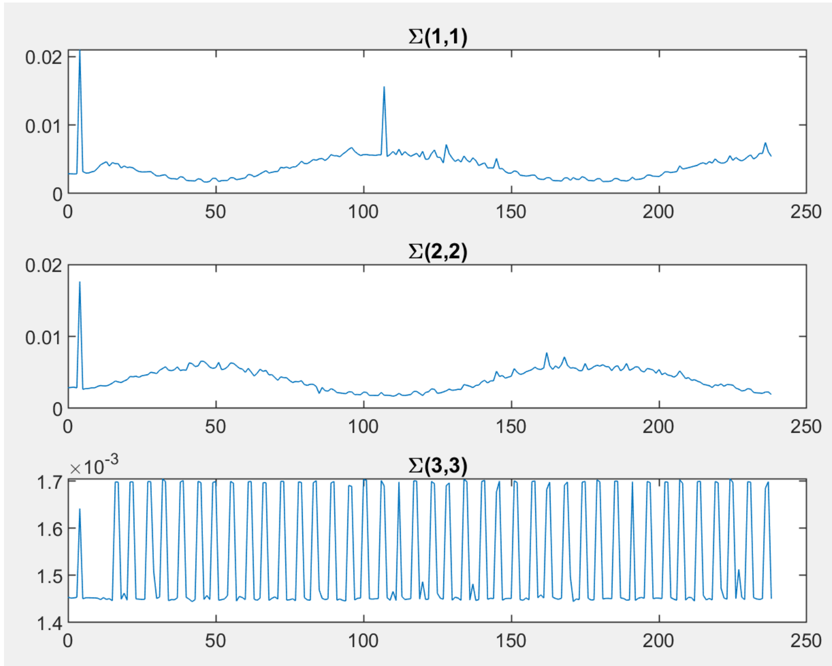
Dataset 3 Sequential



Simulation Time: 238.0 s
Std Motion Model: x: 1 y: 1
θ: 1
Std Observation Model: r: 0.1
θ: 0.1
Outlier Detection Off

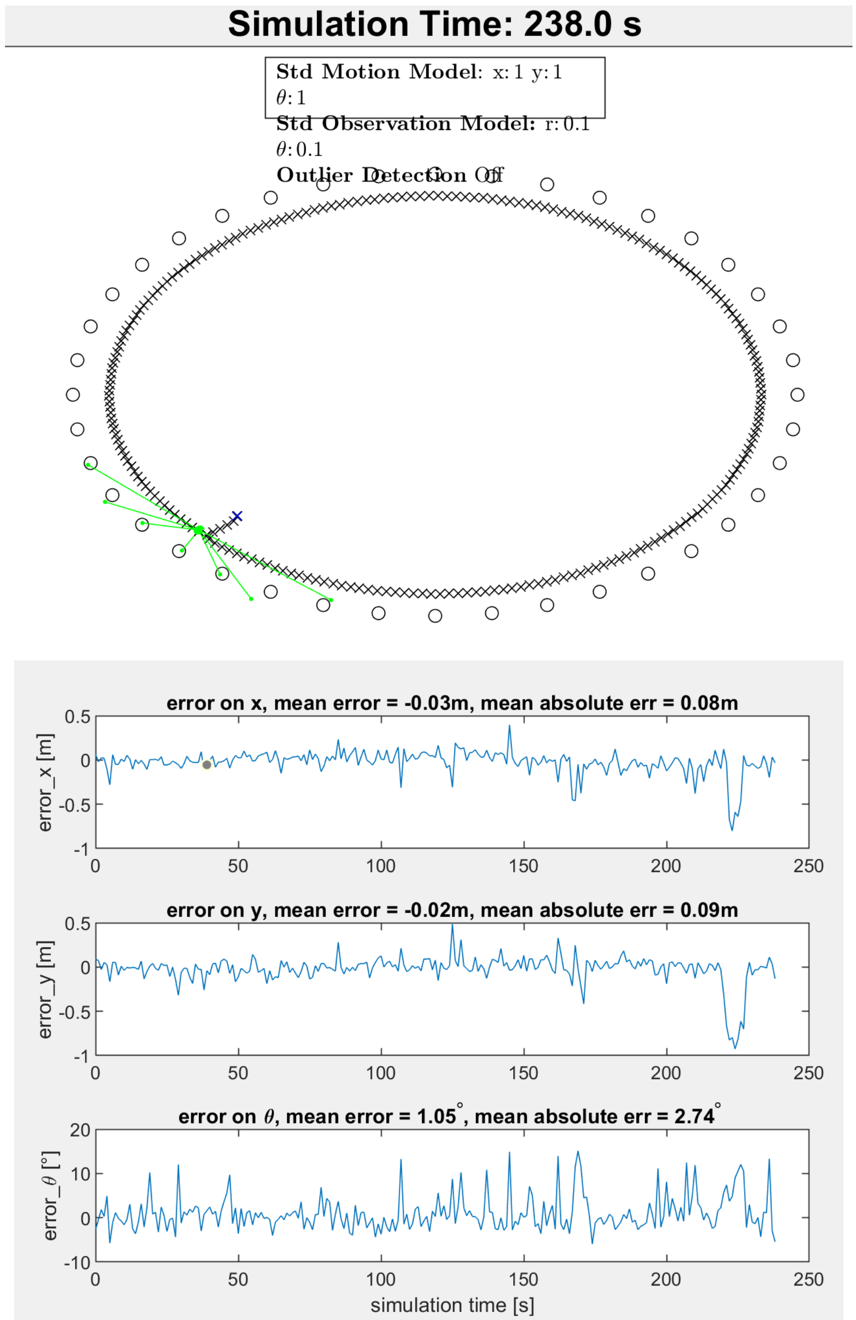


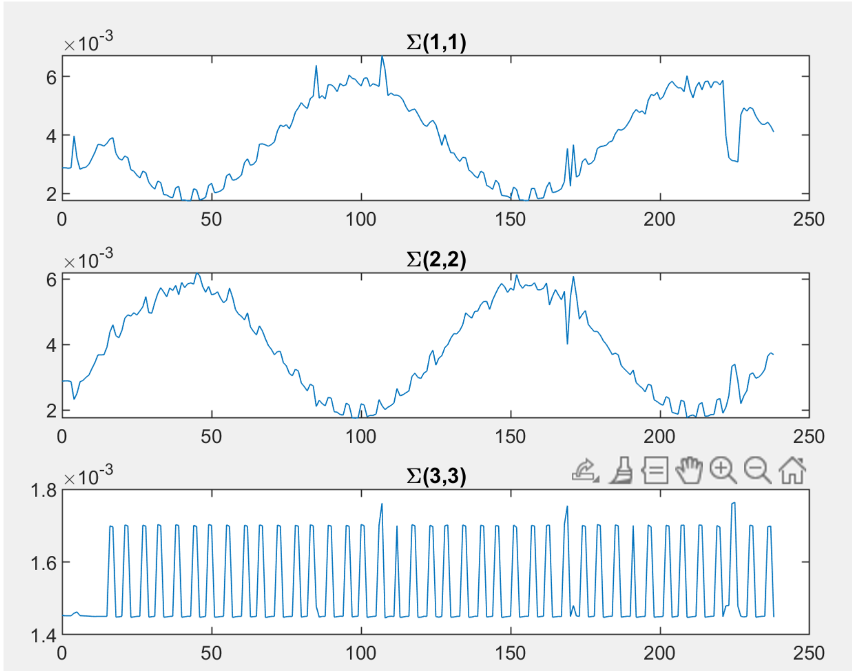
error on x, mean error = -2.62m, mean absolute err = 4.46m
error_x [m]
error on y, mean error = -1.73m, mean absolute err = 4.80m
error_y [m]
error on θ, mean error = 25.19°, mean absolute err = 25.22°
error_θ [°]
simulation time [s]



Σ(1,1)
0.02
0.01
0
0
50
100
150
200
250
Σ(2,2)
0.02
0.01
0
0
50
100
150
200
250
×10^-3
Σ(3,3)
1.7
1.6
1.5
1.4
0
50
100
150
200
250

Dataset 3 Batch





As expected, the batch update had a much smaller error and covariance than the sequential update. This is explained in Question 8 above.