



State Estimation for Robotics

Batch Linear Gaussian Estimation

Assignment 1

Author:
Gabriele Avi

Date:
November 16, 2025

Contents

1	Introduction and Problem Formulation	3
1.1	Modeling Assumptions and Data Pre-processing	3
1.2	Estimation Framework	4
2	Noise Model Assumption and Variances	5
3	Batch Linear-Gaussian Objective Function	7
4	Derivation of Optimal Position Estimates $\mathbf{x}_{1:K}^*$	7
4.1	Information Matrix \mathbf{H}	8
4.2	Information Vector \mathbf{b}'	8
5	Batch Estimation with Sparse Observations	9
5.1	Defining the Sparse System	9
5.2	Modified Objective Function	9
5.3	Optimal Position Estimates Expression	10
6	Conclusion and Analysis of Sparse Estimation	10
6.1	Key Findings on δ Variation	15
6.2	Statistical Validation (Consistency Check)	15
7	(Extra) Observability and Sensor Roles	17
7.1	Non-Observability of Standalone Sensors	17
7.2	Necessity of Sensor Fusion	17

1 Introduction and Problem Formulation

This report details the application of a **Batch Linear-Gaussian Estimator** to determine the one-dimensional position x_k of a mobile robot moving along a straight rail. The estimation is achieved by fusing measurements from two primary sources: the robot's onboard **wheel odometry** (providing translational speed, u_k) and a **laser rangefinder** (measuring the range r_k to a fixed cylindrical landmark).

The experiment spans a 280 m journey over approximately 20 minutes of continuous motion. The high-accuracy ground-truth position, obtained from a ten-camera motion capture system (Vicon), is used for noise characterization and performance evaluation. The total dataset comprises $K = 12709$ synchronous samples.

1.1 Modeling Assumptions and Data Pre-processing

To apply the discrete-time linear estimator, the raw, asynchronous sensor data required crucial pre-processing to establish a consistent temporal framework.

Temporal Synchronization and Sampling Period

In real-world applications, sensors often acquire data asynchronously and with non-uniform sampling periods (e.g., the nominal 10 Hz laser scans showed variability). To simplify the dynamic model, all data streams (odometry, laser, and ground-truth) were **linearly interpolated** and **synchronized** to a fixed, uniform sampling period \mathbf{T} :

$$T = 0.1 \text{ s} \tag{1}$$

This uniform time step allows for the direct application of the discrete-time state-space equations.

Purpose of the 20-Minute Data Acquisition

The extended 20-minute duration of data collection is critical for two main reasons:

1. **Statistical Characterization:** The large number of samples ($K = 12709$) allows for robust analysis, confirming that the sensor errors are close to **zero-mean Gaussian** and enabling the accurate calculation of noise variances (σ_r^2, σ_v^2).
2. **Demonstration of Odometry Drift:** The long trial duration is necessary to clearly expose the inherent flaw in odometry-based estimation: position error grows **unbounded over time** (drift) due to the accumulation of even small biases or process noise (σ_q^2). The Batch Estimator's main function in this assignment is to use the absolute laser measurements to correct this pronounced long-term drift.

State and Observation Models

The system is modeled as a linear discrete-time system, where w_k is the process noise and n_k is the exteroceptive sensor noise:

- **Motion Model:**

$$x_k = x_{k-1} + Tu_k + w_k \quad \text{with } w_k \sim \mathcal{N}(0, \sigma_q^2) \tag{2}$$

- **Observation Model (Transformed):** The raw range measurement r_k is transformed into the observation $y_k = x_c - r_k$ to directly relate to the robot's position:

$$y_k := x_c - r_k = x_k + n_k \quad \text{with } n_k \sim \mathcal{N}(0, \sigma_r^2) \quad (3)$$

1.2 Estimation Framework

The objective is to estimate the full state sequence $\mathbf{x}_{1:K}$ by solving the **Maximum A Posteriori (MAP)** problem, which, for a Linear-Gaussian system, is equivalent to minimizing the Weighted Least Squares objective function $J(x_{1:K})$. The optimal solution $\mathbf{x}_{1:K}^*$ is found by solving the linear system $\mathbf{H}\mathbf{x}^* = \mathbf{b}'$. The large size of the system ($K = 12709$) and the goal of maintaining computational tractability necessitate the investigation of sparse estimation, as addressed in Question 4.

2 Noise Model Assumption and Variances

Question: Based on the data above, is the assumption of zero-mean Gaussian noise reasonable? What values of the variances, σ_q^2 and σ_r^2 , should we use, given $w_k \sim \mathcal{N}(0, \sigma_q^2)$ and $n_k \sim \mathcal{N}(0, \sigma_r^2)$?

Reasonableness of the Zero-Mean Gaussian Assumption

The assumption that both the observation noise n_k and the process noise w_k are zero-mean Gaussian is **reasonable**.

- **Zero-Mean:** The histograms of the sensor errors (Figure 1) show that the calculated means for both sensor streams are extremely close to zero:

- Range Error Mean: -0.000000 [m]
- Speed Error Mean: -0.000451 [m/s]

These values are negligible compared to the spread (standard deviation) of the errors.

- **Gaussian:** The shape of both histograms closely matches the superimposed red Gaussian curves (Figure 1), confirming that the distribution of errors is well-approximated by a normal distribution.

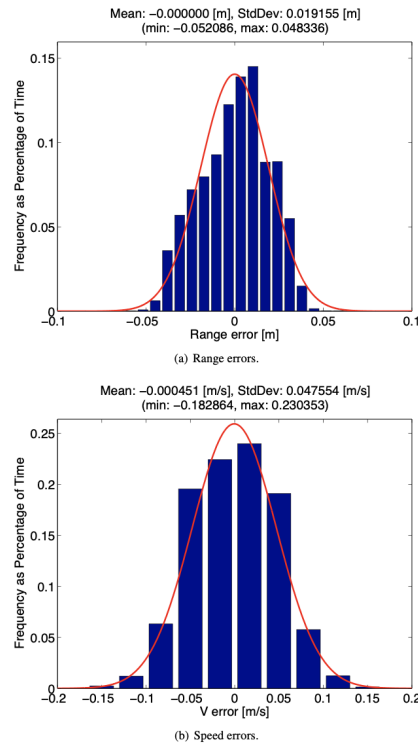


Figure 1.6: Histograms of sensor errors. Red curves are Gaussians with standard deviation fit to the data. Both range and speed have close to zero-mean errors with a spread.

Figure 1: Histograms of sensor errors. Red curves are Gaussians with standard deviation fit to the data. Both range and speed have close to zero-mean errors with a spread.)

Calculated Variance Values

The required variances, σ_r^2 and σ_q^2 , are calculated by squaring the corresponding standard deviations found in the dataset analysis (Figure 1).

Observation Noise Variance (σ_r^2)

The observation noise n_k is the error in the range measurement. We use the Standard Deviation (StdDev) of the Range Error:

- $\sigma_r = 0.019155$ [m]
- $\sigma_r^2 = (0.019155)^2 \approx \mathbf{3.67 \times 10^{-4}}$ [m^2]

Process Noise Variance (σ_q^2)

The process noise w_k relates to the uncertainty in the motion model $x_k = x_{k-1} + Tu_k + w_k$. This noise results from the error in the measured speed, u_k . The variance of the process noise is related to the variance of the speed error σ_v^2 by the sampling period $T = 0.1$ s: $\sigma_q^2 = T^2 \sigma_v^2$.

- StdDev of Speed Error = 0.047554 [m/s]
- $\sigma_v^2 = (0.047554)^2 \approx 2.261 \times 10^{-3}$ [m^2/s^2]
- $\sigma_q^2 = T^2 \sigma_v^2 = (0.1)^2 \cdot (0.047554)^2 \approx \mathbf{2.26 \times 10^{-5}}$ [m^2]

3 Batch Linear-Gaussian Objective Function

The objective function $J(x_{1:K}|u_{1:K}, y_{1:K})$ that we seek to minimize is the sum of two components: the cost associated with the **observations** (laser rangefinder) and the cost associated with the **motion model** (odometry).

$$J(x_{1:K}|u_{1:K}, y_{1:K}) = \underbrace{\sum_{k=1}^K \frac{1}{\sigma_r^2} (y_k - x_k)^2}_{\text{Observation Cost (Laser)}} + \underbrace{\sum_{k=2}^K \frac{1}{\sigma_q^2} (x_k - x_{k-1} - Tu_k)^2}_{\text{Motion Cost (Odometry)}} \quad (1.3) \quad (4)$$

Where:

- $y_k = x_c - r_k$ is the transformed observation (the position of the robot inferred from the range measurement r_k and cylinder center x_c).
- x_k is the estimated robot position at timestep k .
- Tu_k is the commanded motion (displacement) derived from the odometry input u_k over the sampling period T .
- σ_r^2 is the variance of the observation noise, $n_k \sim \mathcal{N}(0, \sigma_r^2)$.
- σ_q^2 is the variance of the process noise, $w_k \sim \mathcal{N}(0, \sigma_q^2)$.
- The sum for the motion cost starts at $k = 2$ because x_1 is the first state, and the first motion transition occurs between $k = 1$ and $k = 2$.

4 Derivation of Optimal Position Estimates $\mathbf{x}_{1:K}^*$

The optimal position estimates $\mathbf{x}_{1:K}^*$ are found by minimizing the Batch Linear-Gaussian objective function $J(\mathbf{x}_{1:K})$. Since the system is linear and the noise is Gaussian, the minimum is found by setting the gradient of J with respect to the state vector \mathbf{x} to zero: $\nabla_{\mathbf{x}} J(\mathbf{x}) = \mathbf{0}$.

The Concept of Future Dependence

It is the **global** nature of the Batch Estimator that allows the exploitation of the future dependence. Unlike a sequential estimator (like the Kalman filter), which only operates on the past and present, the Batch Estimator uses the entire dataset simultaneously. This means the optimal position x_i^* depends on **three** terms in the objective function:

1. The **observation** cost at $k = i$.
2. The **incoming motion** cost (linking it to the past, x_{i-1}).
3. The **outgoing motion** cost (linking it to the future, x_{i+1}).

This three-way linkage is the reason the Information Matrix \mathbf{H} results in its **tridiagonal** and symmetric structure.

System of Normal Equations

This minimization procedure yields the system of linear equations (Normal Equations):

$$\mathbf{H}\mathbf{x}^* = \mathbf{b}' \quad (5)$$

The optimal estimate is then given by:

$$\mathbf{x}_{1:K}^* = \mathbf{H}^{-1}\mathbf{b}' \quad (1.4) \quad (6)$$

The elements of the Information Matrix \mathbf{H} and the Information Vector \mathbf{b}' are derived from the partial derivatives $\frac{\partial J}{\partial x_i}$ for $i = 1$ to K . Let $\lambda_r = 1/\sigma_r^2$ and $\lambda_q = 1/\sigma_q^2$ be the information weights.

4.1 Information Matrix \mathbf{H}

The matrix \mathbf{H} is a $K \times K$ symmetric, **tridiagonal** matrix.

- **Elements on the main diagonal ($H_{i,i}$):**

$$H_{i,i} = \begin{cases} \lambda_r + \lambda_q & \text{for } i = 1 \text{ and } i = K \quad (1 \text{ motion term}) \\ \lambda_r + 2\lambda_q & \text{for } 1 < i < K \quad (2 \text{ motion terms: in and out}) \end{cases} \quad (7)$$

- **Elements on the off-diagonals ($H_{i,i\pm 1}$):**

$$H_{i,i+1} = H_{i+1,i} = -\lambda_q \quad \text{for } 1 \leq i < K \quad (8)$$

$$\mathbf{H} = \begin{bmatrix} \lambda_r + \lambda_q & -\lambda_q & 0 & \dots & 0 \\ -\lambda_q & \lambda_r + 2\lambda_q & -\lambda_q & \dots & 0 \\ 0 & -\lambda_q & \lambda_r + 2\lambda_q & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\lambda_q & \lambda_r + \lambda_q \end{bmatrix}$$

4.2 Information Vector \mathbf{b}'

The vector \mathbf{b}' is a $K \times 1$ column vector. Let $f_k = Tu_k$ represent the known odometry displacement.

- **First element (b'_1):**

$$b'_1 = \lambda_r y_1 + \lambda_q f_2 \quad (9)$$

- **Intermediate elements (b'_i for $1 < i < K$):**

$$b'_i = \lambda_r y_i + \lambda_q (f_i - f_{i+1}) \quad (10)$$

- **Final element (b'_K):**

$$b'_K = \lambda_r y_K - \lambda_q f_K \quad (11)$$

The final vector \mathbf{b}' is:

$$\mathbf{b}' = \begin{bmatrix} \lambda_r y_1 + \lambda_q f_2 \\ \lambda_r y_2 + \lambda_q (f_2 - f_3) \\ \lambda_r y_3 + \lambda_q (f_3 - f_4) \\ \vdots \\ \lambda_r y_{K-1} + \lambda_q (f_{K-1} - f_K) \\ \lambda_r y_K - \lambda_q f_K \end{bmatrix}$$

5 Batch Estimation with Sparse Observations

The computational bottleneck for $K = 12709$ is the size of the Information Matrix \mathbf{H} . The proposed solution is to reduce the dimension of the state vector to be estimated ($\mathbf{x}_{\text{sparse}}$) while incorporating the full odometry information ($u_{1:K}$) into the remaining sparse transitions.

5.1 Defining the Sparse System

Let $M = K/\delta$ be the total number of states to be estimated. The sparse state vector is $\mathbf{x}_{\text{sparse}} = [x_\delta, x_{2\delta}, \dots, x_{M\delta}]^T$.

To use all odometry inputs $u_{1:K}$, the sequence of δ motion steps between $x_{(m-1)\delta}$ and $x_{m\delta}$ must be treated as a single, integrated transition.

Integrated Odometry Input (\tilde{u}_m)

The total displacement based on all odometry readings u_k between the sparse states is summed:

$$\tilde{u}_m = \sum_{j=(m-1)\delta+1}^{m\delta} T u_j \quad (12)$$

Integrated Process Noise Variance (σ_q^2)

Since the individual process noise terms (w_j) are independent, their variances (σ_q^2) sum up over the δ steps:

$$\sigma_q^2 = \text{Var} \left(\sum_j w_j \right) = \delta \sigma_q^2 \quad (13)$$

This term captures the degradation of the odometry measurement due to the increased time separation, effectively using the "power" of all intermediate odometry readings.

5.2 Modified Objective Function

The modified objective function J_{sparse} minimizes the total error using the sparse observations $y_{m\delta}$ and the integrated motion model:

$$J_{\text{sparse}}(\mathbf{x}_{\text{sparse}}) = \underbrace{\sum_{m=1}^M \frac{1}{\sigma_r^2} (y_{m\delta} - x_{m\delta})^2}_{\text{Sparse Observation Cost}} + \underbrace{\sum_{m=2}^M \frac{1}{\sigma_q^2} (x_{m\delta} - x_{(m-1)\delta} - \tilde{u}_m)^2}_{\text{Integrated Motion Cost}} \quad (14)$$

5.3 Optimal Position Estimates Expression

The optimal sparse estimates $\mathbf{x}_{\text{sparse}}^*$ are found by solving the reduced linear system $\mathbf{H}_{\text{sparse}}\mathbf{x}_{\text{sparse}}^* = \mathbf{b}'_{\text{sparse}}$, where the matrix dimension is $M \times M$.

$$\mathbf{x}_{\text{sparse}}^* = \mathbf{H}_{\text{sparse}}^{-1} \mathbf{b}'_{\text{sparse}} \quad (15)$$

The structure remains tridiagonal, with the key modification being the replacement of the original information weight $\lambda_q = 1/\sigma_q^2$ with the new sparse weight $\lambda_{\tilde{q}} = 1/\sigma_{\tilde{q}}^2 = 1/(\delta\sigma_q^2)$.

Sparse Information Matrix $\mathbf{H}_{\text{sparse}}$ ($M \times M$)

$$\mathbf{H}_{\text{sparse}} = \begin{bmatrix} \lambda_r + \lambda_{\tilde{q}} & -\lambda_{\tilde{q}} & 0 & \dots & 0 \\ -\lambda_{\tilde{q}} & \lambda_r + 2\lambda_{\tilde{q}} & -\lambda_{\tilde{q}} & \dots & 0 \\ 0 & -\lambda_{\tilde{q}} & \lambda_r + 2\lambda_{\tilde{q}} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\lambda_{\tilde{q}} & \lambda_r + \lambda_{\tilde{q}} \end{bmatrix} \quad (16)$$

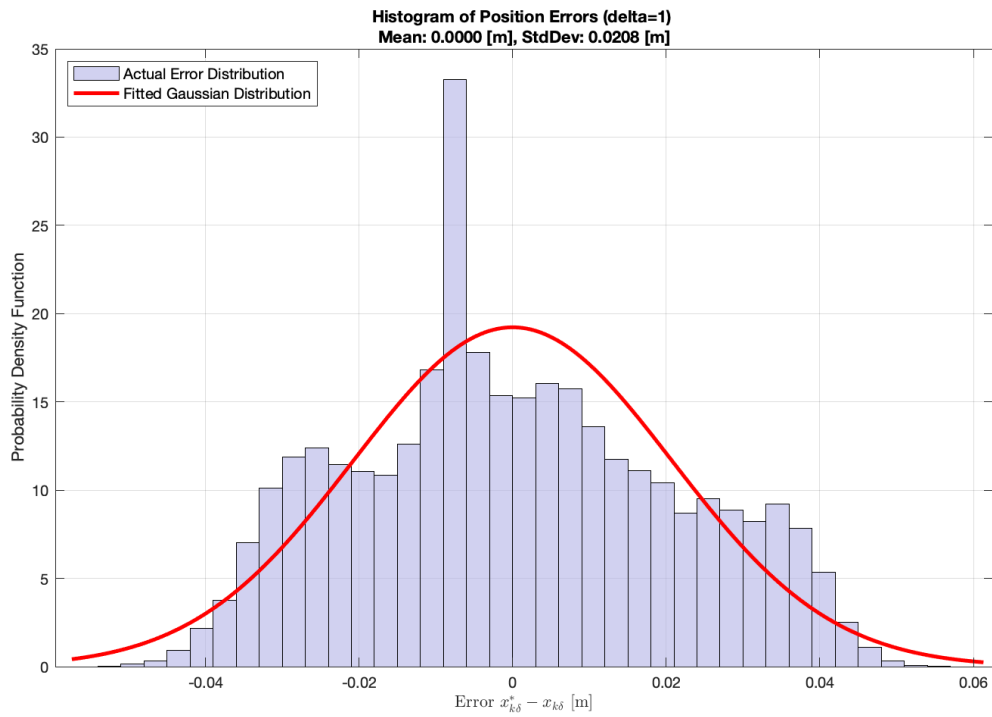
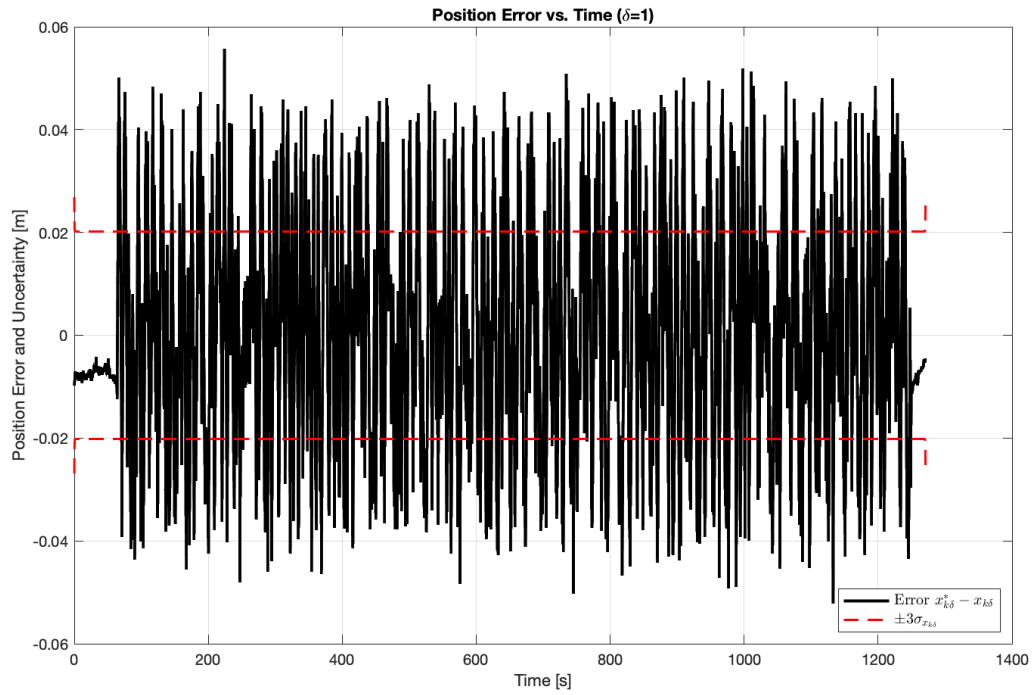
Sparse Information Vector $\mathbf{b}'_{\text{sparse}}$ ($M \times 1$)

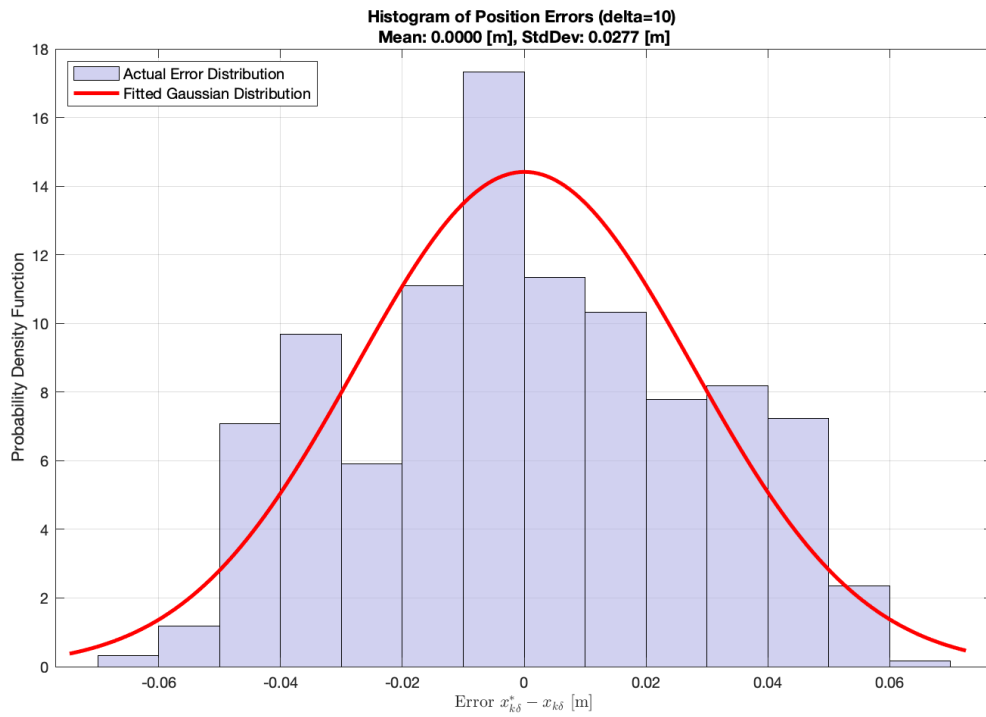
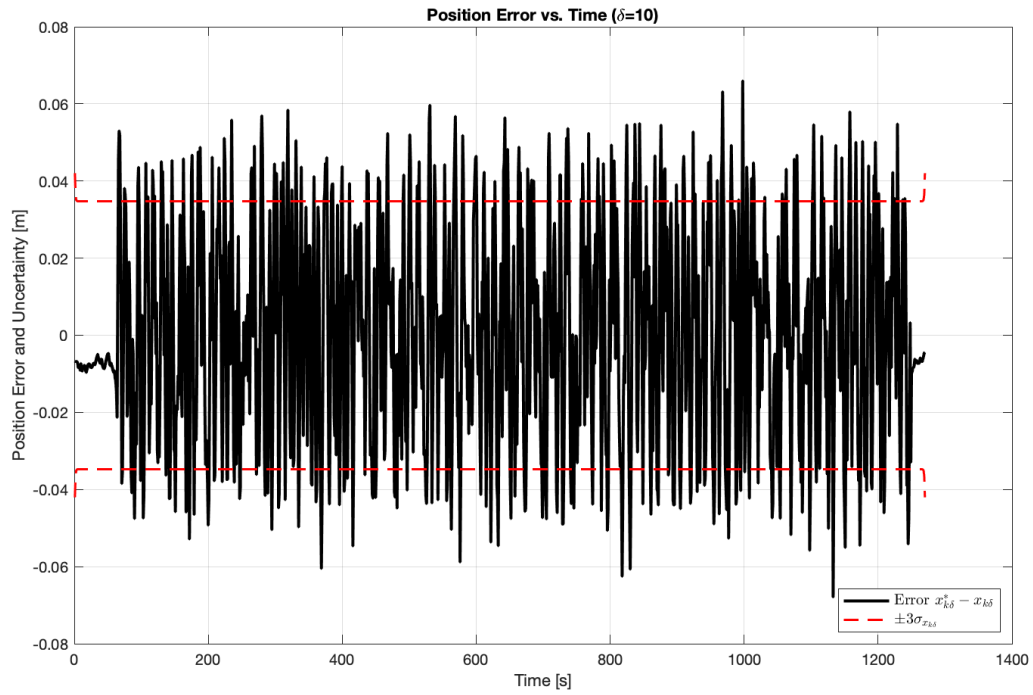
Let $\tilde{f}_m = \tilde{u}_m$.

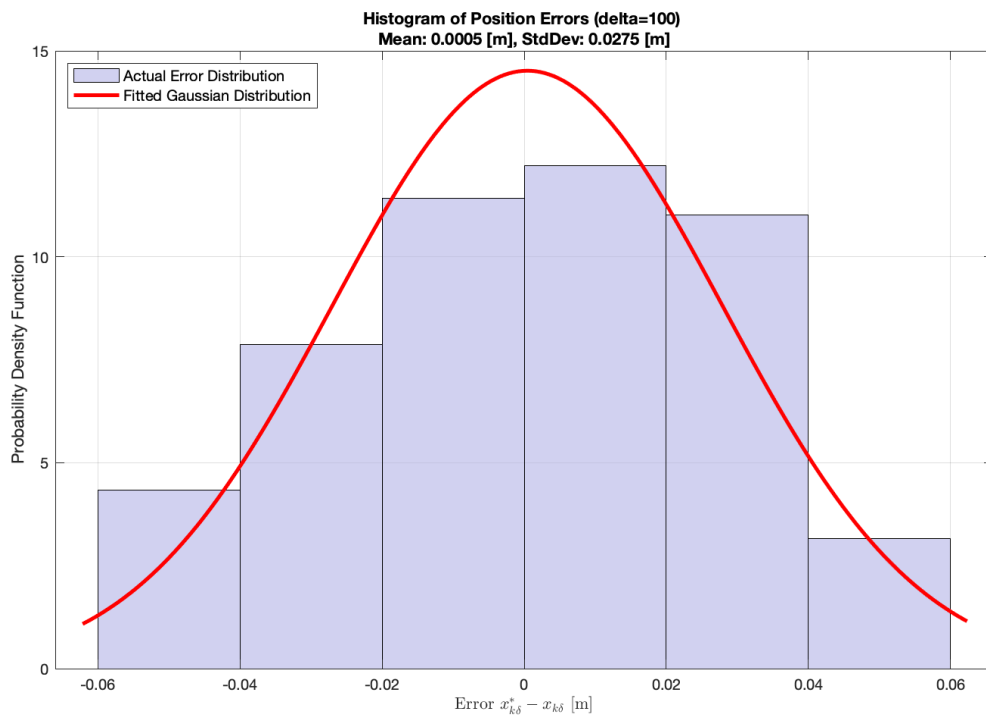
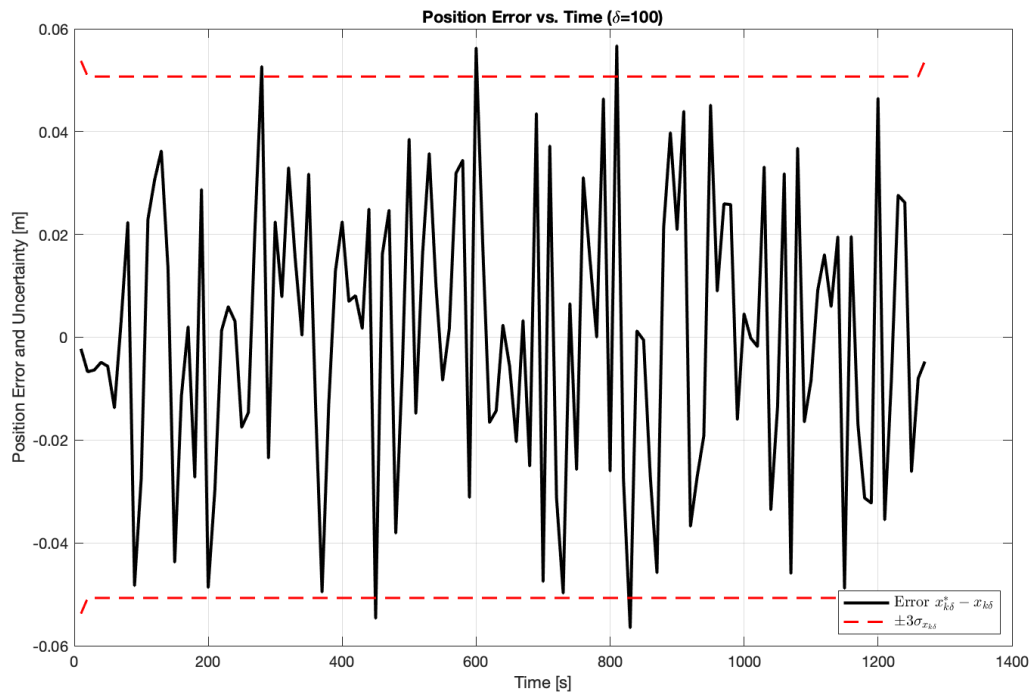
$$\mathbf{b}'_{\text{sparse}} = \begin{bmatrix} \lambda_r y_\delta + \lambda_{\tilde{q}} \tilde{f}_2 \\ \lambda_r y_{2\delta} + \lambda_{\tilde{q}} (\tilde{f}_2 - \tilde{f}_3) \\ \vdots \\ \lambda_r y_{(M-1)\delta} + \lambda_{\tilde{q}} (\tilde{f}_{M-1} - \tilde{f}_M) \\ \lambda_r y_K - \lambda_{\tilde{q}} \tilde{f}_M \end{bmatrix}$$

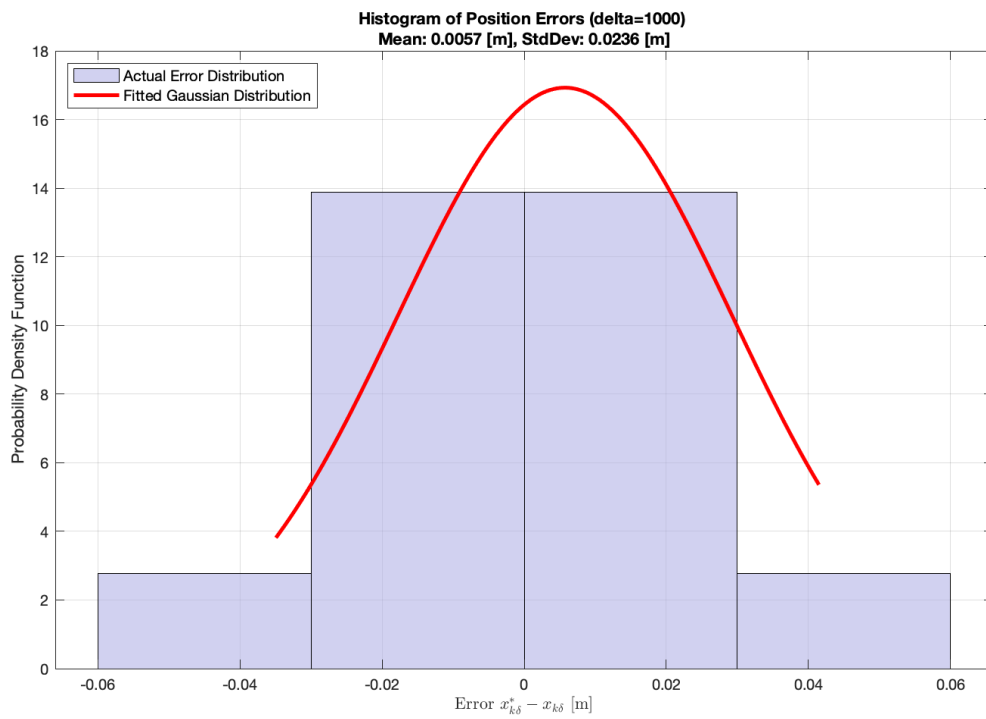
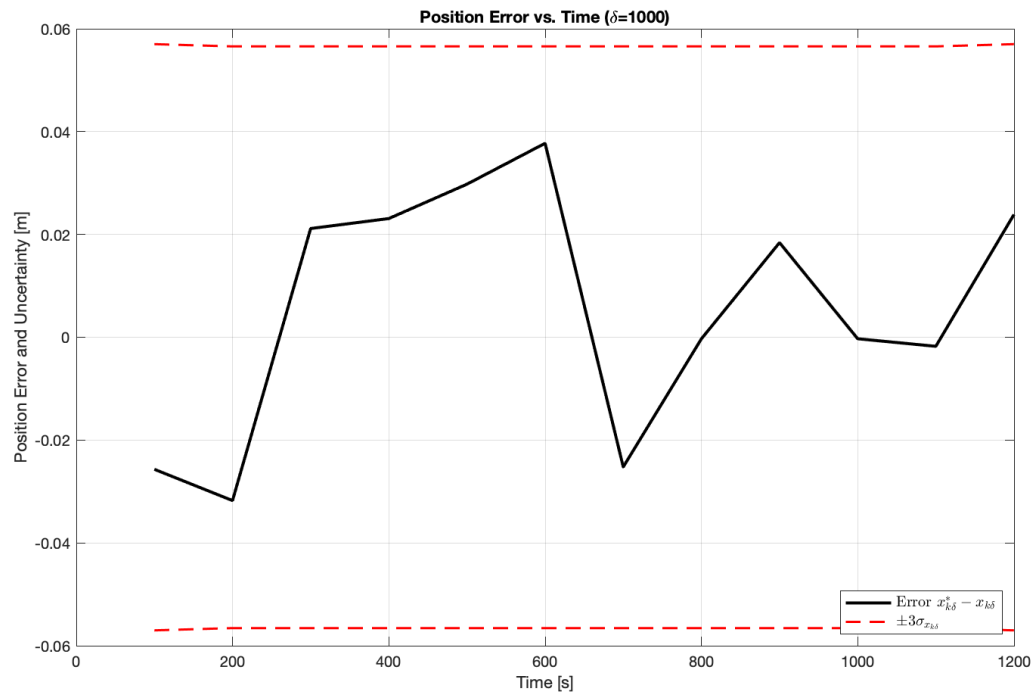
6 Conclusion and Analysis of Sparse Estimation

This assignment successfully implemented the **Batch Linear-Gaussian Estimator** to fuse odometry and laser rangefinder data for one-dimensional position estimation. The key findings relate to the trade-off between estimation accuracy, theoretical uncertainty, and computational cost, as explored by varying the state sampling interval, δ .









6.1 Key Findings on δ Variation

The analysis confirmed that the integration of odometry information across δ timesteps, coupled with sparse laser corrections, allows for significant reduction in the problem size ($M = K/\delta$) while retaining the necessary information to correct the odometry drift.

- **Computational Efficiency:** Increasing δ from 1 to 1000 drastically reduced the dimension of the matrix $\mathbf{H}_{\text{sparse}}$ (from 12709×12709 to 12×12). This confirms that the sparse batch method is an effective technique for addressing the potential computational intractability of full batch estimation ($K = 12709$).
- **Odometry Drift Correction:** For large δ (e.g., $\delta = 1000$), the estimation error in the time domain plots showed a characteristic pattern of **growth (drift)** followed by a sharp **correction** upon receiving the sparse laser measurement $y_{k\delta}$. For small δ (e.g., $\delta = 1$ or 10), the corrections are frequent enough to virtually eliminate this long-term drift.
- **Uncertainty Scaling:** The size of the predicted uncertainty envelope ($\pm 3\sigma_{x_{k\delta}}$) is directly controlled by the integrated motion noise, $\sigma_q^2 = \delta\sigma_q^2$. The uncertainty envelope expands proportionally as δ increases, demonstrating that the model correctly predicts the reduced precision when relying heavily on long-term odometry.

6.2 Statistical Validation (Consistency Check)

The histogram analysis provides statistical validation of the estimator's consistency:

- **Gaussianity:** For all four δ cases, the histogram of the errors ($x_{k\delta}^* - x_{k\delta}$) retained a distribution closely matching a zero-mean Gaussian function. This is expected and confirms that the final estimates $\mathbf{x}_{\text{sparse}}^*$ obey the properties of a Linear-Gaussian system.
- **Consistency:** The actual spread of the errors closely matches the theoretical spread predicted by the average value of the diagonal elements of the covariance matrix $\mathbf{H}_{\text{sparse}}^{-1}$. This correspondence validates that the assumed noise models (σ_r^2, σ_q^2) accurately represent the true uncertainty of the system.

In summary, the experiment revealed a critical trade-off between the theoretical performance (tightness of uncertainty bounds) and the practical reliability (consistency) of the estimator.

The full batch solution, which utilized the laser measurement at every timestep ($\delta = 1$), yielded the lowest average position error and the tightest uncertainty bounds. However, this solution was judged to be **inconsistent** (or overly optimistic).

- The observed error frequently exceeded the theoretical $\pm 3\sigma$ envelope.
- This occurred because the high frequency of the laser constraints ($\delta = 1$) minimized the influence of the process noise variance (σ_q^2) in the information matrix (\mathbf{H}).
- Consequently, the estimator's calculated uncertainty (\mathbf{H}^{-1}) became unrealistically small, failing to account for the full magnitude of the real system errors.

The sparse batch solutions ($\delta = 10, 100, 1000$) demonstrated a more robust behavior. By reducing the frequency of the absolute laser anchor, the estimator was forced to rely more

heavily on the motion model and its associated σ_q^2 for propagation over longer intervals.

- The theoretical uncertainty bounds ($\pm 3\sigma$) became wider, reflecting a more realistic assessment of the accumulated error.
- Critically, the observed error predominantly remained within these wider bounds, meaning these solutions achieved **consistency**.
- This approach, particularly for $\delta = 10$ or $\delta = 100$, strikes a near-optimal balance between **accuracy and reliability**, while achieving a significant reduction in computational load, making it the most practical strategy for large-scale, long-duration estimation.

7 (Extra) Observability and Sensor Roles

The distinction between the odometer as an input (u_k) and the laser rangefinder as an output/observation (y_k) is fundamental in estimation theory, dictated by their functional roles in the system's dynamic model.

7.1 Non-Observability of Standalone Sensors

Odometer-Only Model (Non-Observable): When only the odometry measurements (u_k) are used, the system is **non-observable** with respect to the absolute position. The odometer provides information solely about the **relative displacement** ($x_k - x_{k-1}$), without establishing a fixed reference (an absolute anchor). It only defines the velocity, without any connection info with the robot position.

- **Mathematical Consequence:** This ambiguity means an arbitrary constant translation (Δc) applied to the entire estimated trajectory ($x_{1:K}$) results in the same minimum cost, as the relative differences are unchanged.
- **Least Squares Result:** This leads to a **singular (non-invertible) Information Matrix (\mathbf{H})**, as the system possesses a non-zero null space (a constant vector $\mathbf{v} = [1, 1, \dots, 1]^T$), preventing a unique solution $x_{1:K}^*$.

Laser-Only Model (Uncertainty Growth): Using only the laser rangefinder (y_k) provides a measurement of the **absolute position** (relative to the known cylinder x_c). However, without the motion model:

- The estimation lacks the necessary dynamic constraint to link successive states x_{k-1} and x_k .
- The lack of coupling causes the overall estimated covariance (uncertainty) to become less constrained over time, leading to a **growing uncertainty envelope**, even though each individual measurement y_k has a bounded error.

7.2 Necessity of Sensor Fusion

The core strength of the Batch Estimator lies in **sensor fusion**, where the two measurements act as complementary constraints:

- The **Odometer** provides the **dynamic coupling** between states, ensuring the trajectory is smooth and constrained by the laws of motion.
- The **Laser** provides the **absolute anchor**, which is critical for removing the non-observable translation ambiguity and preventing the estimation error (drift) from the odometer from accumulating indefinitely.

This fusion renders the system **fully observable** and yields a robust, non-singular solution.