UNIVERSITY OF WEST ATTICA
FACULTY OF ENGINEERING
Department of Electrical & Electronics Engineering
Department of Industrial Design & Production Engineering
MSc in Artificial Intelligence and Deep Learning

Master of Science Thesis

Development of estimation algorithms for navigation of unmanned aerial vehicles

Space to insert image, picture or diagram related to the thesis subject (optional)

Student: Kostakis Ilias-Odysseas
Registration Number: AIDL-0063

MSc Thesis Supervisor

Alexandridis Alexandros Professor

Insert supervisor name in latin characters, in the form Surname, name
Insert supervisor grade as: Professor, Assoc. Professor, Assist. Professor, Lecturer

ATHENS-EGALEO, February 2025

This MSc Thesis has been accepted, evaluated and graded by the following committee:

Supervisor	Member	Member
SIGNATURE	SIGNATURE	SIGNATURE
Surname, Name	Surname, Name	Surname, Name
Grade	Grade	Grade
Department or School	Department or School	Department or School
Institution	Institution	Institution

Declaration of the author of this MSc thesis

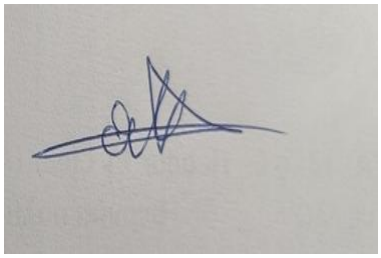
I, Ilias-Odysseas Kostakis, Pantelis with the following student registration number: mscaidl-0063, postgraduate student of the MSc program in "Artificial Intelligence and Deep Learning", which is organized by the Department of Electrical and Electronic Engineering and the Department of Industrial Design and Production Engineering of the Faculty of Engineering of the University of West Attica, hereby declare that:

I am the author of this MSc thesis and any help I may have received is clearly mentioned in the thesis. Additionally, all the sources I have used (e.g., to extract data, ideas, words or phrases) are cited with full reference to the corresponding authors, the publishing house or the journal; this also applies to the Internet sources that I have used. I also confirm that I have personally written this thesis and the intellectual property rights belong to myself and to the University of West Attica. This work has not been submitted for any other degree or professional qualification except as specified in it.

Any violations of my academic responsibilities, as stated above, constitutes substantial reason for the cancellation of the conferred MSc degree.

The author

Ilias-Odysseas Kostakis



{Dedication page (optional)}

{Acknowledgements page (optional)}

Abstract

Accurate attitude estimation and prediction are critical for the autonomous navigation and advanced control of unmanned aerial vehicles (UAVs). While classical sensor fusion algorithms effectively estimate current vehicle orientation, they are fundamentally reactive and cannot anticipate future states. Modern control strategies, such as Model Predictive Control (MPC), require multi-step ahead predictions of future attitude, creating a need for predictive models based on deep learning. This thesis develops a comprehensive framework for UAV attitude prediction centered on an attention-based Long Short-Term Memory (LSTM) network. The proposed architecture predicts future roll ϕ, pitch θ, and yaw ψ angles over a 10-timestep horizon using look-back windows of IMU measurements and control inputs. To establish a robust training dataset and baseline performance, four classical tilt estimation filters—the Complementary Filter, Mahony Filter, Explicit Complementary Filter, and Extended Kalman Filter (EKF)—were implemented in C++ and optimized using real flight data. By utilizing the EKF-estimated angles as training labels, the framework becomes self-contained and applicable to real-world scenarios where external ground-truth systems are unavailable. Experimental evaluation demonstrates that the attention-based LSTM not only achieves high estimation accuracy but also provides the essential multi-step prediction capability required for proactive MPC-based flight controllers, effectively bridging the gap between classical sensor fusion and modern deep learning for UAV navigation.

Keywords

UAVs, Complementary Filtering, EKF, Attitude estimation, Classic estimation methods, Neural Networks, Machine Learning, MPC, LSTM, RNN, Sensor Fusion

Table of Contents

List of Tables

List of figures

Acronym Index

IMU: Inertial Measurement Unit
EKF: Extended Kalman Filter
UAV: Unmanned Aerial Vehicle
RNN: Recurrent Neural Network
LLM: Large Language Model
LSTM: Long Term Short-Memory
MPC: Model Predictive Control

INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly known as drones, have emerged as transformative technologies across many sectors. From aerial photography, agriculture to package delivery, these autonomous platforms are reshaping how we approach tasks that were impossible in the past.

At the heart of every UAV lies a critical capability: the ability to know its own orientation in three-dimensional space. This ability, referred to as attitude, describes how the vehicle is tilted relative to the Earth's surface. Without accurate and continuous attitude information, a UAV cannot maintain stable flight, execute precise maneuvers, or respond appropriately to disturbances such as wind gusts. The attitude serves as the foundation upon which all flight control decisions are made.

Modern UAVs rely on Inertial Measurement Units (IMUs), compact electronic devices that combine multiple sensors to measure motion. A typical IMU contains gyroscopes, which measure rotational velocity, and accelerometers, which measure linear acceleration including gravity. These sensors are inexpensive, lightweight, and consume minimal power, making them ideal for small autonomous platforms. However, each sensor type has inherent limitations that prevent it from providing accurate attitude estimates on its own. Classical sensor fusion algorithms such as complementary filters and the Extended Kalman Filter (EKF) address this by combining gyroscope and accelerometer data to produce reliable estimates of the current tilt angles (ϕ, θ). While effective for real-time estimation these approaches are fundamentally reactive: they can estimate the current state and cannot anticipate future vehicle orientation.

However, modern advanced flight control strategies such as Model Predictive Control (MPC) require not just knowledge of the current attitude, but prediction of the future states over a finite horizon [12]. MPC computes optimal control actions by solving an optimization problem over predicted future trajectories, making the quality of the prediction critical to controller performance. This creates a need for algorithms that can provide accurate multi-step ahead attitude predictions.

Recent advances in deep learning have introduced new possibilities for time-series prediction in dynamical systems. Recurrent Neural Networks (RNNs) and in particular Long Short-Term Memory (LSTM) networks have demonstrated strong capabilities in learning temporal dependencies from sequential data [9]. In the field of UAV navigation, neural network-based approaches have been explored for attitude estimation as an alternative or complement to classical filters. Various techniques, algorithms and methods have been explored. Feedforward neural networks have been trained on IMU data and estimate orientation [10]. Also, RRN-based approaches have been used for inertial navigation [11]. However, the majority of these works focus on single-step estimation of the current state, rather than multi-step ahead prediction of future states. Furthermore, few works address the specific requirement of providing prediction horizons compatible with MPC-based flight controllers.

This thesis proposes an attention-based LSTM architecture for multi-step ahead attitude prediction that address this gap. Unlike existing neural networks approaches that estimate only the current state, the proposed method predicts the attitude angles (ϕ, θ, ψ) over a configurable future horizon of N timesteps ($k+1\ to\ k+N$), producing predictions directly suitable with MPC controllers. The attention mechanism enables the network to learn which timesteps within the input look-back window are most informative for the prediction task, improving prediction accuracy over standard LSTM architectures.

The subject of this thesis

This thesis addresses the problem of predicting future attitude states of unmanned aerial vehicles from sequential sensor and control data. Specifically, given a look-back window of K timesteps containing attitude angle estimates (ϕ, θ, ψ), raw gyroscope measurements, and control torques, the task is to predict the attitude angles for the next N timesteps into the future. On a real UAV platform, true attitude angles are not directly measurable by any single sensor — they must be estimated from raw inertial data using sensor fusion algorithms. This creates a two-stage problem: first, accurate attitude angles must be produced from raw IMU measurements; second, these estimated angles must be used to train a predictive model capable of forecasting future states. The quality of the first stage directly constrains the quality of the second, linking the two components into a single integrated pipeline. To address the first stage, four classical tilt estimation filters of increasing mathematical sophistication are implemented and comparatively evaluated: the Complementary Filter, the Mahony Passive Complementary Filter, the Explicit Complementary Filter, and the Extended Kalman Filter (EKF). The best-performing filter provides the attitude angle estimates that serve as training data for the second stage. For the prediction stage, an attention-based LSTM network is designed to learn temporal patterns across the input window and generate multi-step ahead forecasts. All components are implemented in C++ using the Eigen library for linear algebra operations and LibTorch (PyTorch C++ API) for the neural network, ensuring suitability for deployment on real-time embedded flight controllers.

Aim and objectives

The primary aim of this thesis is to develop a complete estimation-to-prediction pipeline for UAV attitude, from raw IMU data to multi-step ahead angle forecasts. The specific objectives are:

1. Implement and optimize four classic tilt estimation filters to produce accurate attitude angle estimates from raw 6-DOF IMU data, establishing a performance baseline.

2. Design and train an attention-based LSTM architecture that predicts the attitude angles (φ, θ, ψ) over a future horizon of N timesteps from look-back windows of sensor and control data.

3. Evaluate prediction accuracy per step (k+1 through k+N) and per angle using RMSE and MAE metrics, quantifying how performance evolves across the prediction horizon.

4. Implement the complete pipeline in C++ using Eigen and LibTorch, demonstrating feasibility for real-time embedded deployment.

Methodology

The methodology follows two phases.

Phase 1 — Tilt Estimation Filters. Four classical filters are implemented. Each filter processes 3-axis gyroscope and 3-axis accelerometer measurements collected from a quadcopter platform at 50 Hz (1409 samples) to estimate the roll (φ) and pitch (θ) angles. The filters represent a progression of mathematical approaches: weighted Euler-angle fusion (Complementary Filter), geometric error correction on SO(3) (Mahony Filter), bias-adaptive vector-based estimation (Explicit Complementary Filter), and probabilistic quaternion-based state estimation (Extended Kalman Filter). Each filter's hyperparameters are tuned via grid search, and performance is evaluated against ground truth measurements using RMSE and MAE. The EKF-estimated angles, having the lowest error, are used as the training labels for the LSTM.

Phase 2 — LSTM Multi-Step Ahead Prediction. An attention-based LSTM network processes look-back windows of K=10 timesteps, each containing 9 features: three attitude angles (from filter estimates), three gyroscope measurements, and three control torques from the PID controller. The network predicts the attitude angles (φ, θ, ψ) for the next N=10 timesteps. Training is performed on a separate flight dataset of 3397 samples at 33.3 Hz (Ts = 0.03 s) using MSE loss with the Adam optimizer. Evaluation metrics are computed per prediction step to assess how accuracy degrades over the prediction horizon.

Innovation

The key innovation of this thesis lies in the application of an attention-based LSTM network for multi-step ahead attitude prediction of UAVs. While existing neural network approaches in the literature primarily focus on single-step attitude estimation as replacement for classical filters [10][11], this work targets a different objective: predicting future attitude states over a configurable horizon of N timesteps.

This distinction is critical because multi-step ahead predictions are prerequisite for Model Predictive Control, which computes optimal control actions based on predicted future trajectories. By providing prediction at timesteps $k + 1\ to\ k + N$, the proposed LSTM architecture can serve as the prediction model with an MPC framework enabling proactive rather than reactive flight control.

A key aspect of the proposed methodology is that the LSTM is trained entirely on filter-estimated attitude angles rather than externally provided ground truth. The classical tilt estimation filters developed in this work process raw IMU data to produce the attitude estimates that form the LSTM's training dataset. This makes the approach self-contained and directly applicable to real-world scenarios where ground truth orientation is unavailable — the same filters that generate training data offline can provide real-time input to the LSTM during deployment. The incorporation of an attention mechanism further distinguishes this work from standard LSTM-based approaches. Rather than relying solely on the final hidden state, the

attention layer learns to weight the contributions of all timesteps in the look-back window, allowing the network to focus on the most informative inputs for each prediction.

Additionally, the complete pipeline is implemented entirely in C++ using LibTorch and Eigen, making it deployable on embedded systems. This is a practical requirement for real-world UAV applications that existing Python-based research implementations do not address.

Structure

Briefly outline the structure of the thesis, in chapters.

TODO After we finish writing just outline the structure ..

1 CHAPTER 1: Theoretical background

This chapter establishes the theoretical foundation necessary to understand the attitude estimation problem for UAVs. An introduction is being made to the attitude estimation in UAV, what sensors are using to measure motion and explore different mathematical representations of orientations. Later fundamental concepts of sensor fusion are discussed which motivate the classic filter algorithms developed in section 2. Finally neural network architectures are introduced for sequential learning, providing the theoretical background for the deep learning approach presented in Chapter 3.

1.1 The role of attitude in UAV flight

The ability of drones to operate autonomously has a critical dependency on their ability to determine and control their orientation in three-dimensional space.

The attitude of a UAV describes its rotation orientation relative to a fixed reference frame, typically defined with respect to the Earth's surface. In aviation, attitude is conventionally expressed using three angular coordinates known as Euler angles:

1) Roll (φ): Rotation around the longitudinal axis (forward-backward direction)
2) Pitch (θ): Rotation around the lateral axis (left-right direction)
3) Yaw (ψ): Rotation around the vertical axis (up-down direction)

These angles directly influence the vehicle's ability to maintain stable flight, execute precise maneuvers, and respond appropriately to environmental disturbances such as wind gusts. Without accurate attitude information a UAV can't, maintain level flight, execute controlled turns, provide stable platform for cameras and sensors, implement autonomous navigation algorithms.

Unlike position (which can be directly measured by GPS) or velocity (which can be derived from GPS updates), attitude cannot be directly measured by a single sensor. Instead, it must be estimated by fusing measurements from multiple sensors, each with different characteristics and limitations.

1.2 IMU and Sensors

An Inertial Measurement Unit (IMU) is a compact electronic device that combines multiple motion sensors to measure a vehicle's kinematic state. A typical 6-axis IMU contains an 3-axis gyroscope and 3-axis accelerometer.

1.2.1 Gyroscope

A gyroscope is an angular rate sensor that measures how fast an object rotates around one or more axes, and in UAVs it is one of core sensors for attitude estimations. In practice gyroscopes output angular velocity (e.g. in degrees or rad per second). Angular velocity is measured as $\omega \in R^3$ (rad/s). Each axis output corresponds to the rate of rotation about the x, y, z axes of the sensor frame.

Gyroscopes have key advantages and limitations. They provide very smooth short-term attitude information and are insensitive to linear accelerations, which makes them reliable during aggressive maneuvers where accelerometers are heavily corrupted by non-gravitational forces. However, their measurements contain bias and noise, so integration causes drift that grows over time, meaning gyro-only attitude estimates must be regularly corrected using other sensors such as accelerometers, magnetometers, or GNSS within a sensor fusion algorithm. [1]

1.2.2 Accelerometer

Accelerometers are fundamentally different from what their name suggests. They do not measure kinematic acceleration. Instead, they measure specific force, also called proper acceleration. It is measured as $a \in R^3$ (m/s^2). Each axis output corresponds to the specific force experienced along the x, y, z axes of the sensor frame, including gravity during static conditions.

Accelerometers excel at providing long-term absolute reference through the gravity vector, allowing computation of roll and pitch via inverse trigonometry when the vehicle is mostly stationary or slow moving. Accelerometer measures $a = [ax, ay, az]$ and we can calculate roll and pitch via these equations

$$roll = atan2(ay, az)$$

$$pitch = arsinc\left(\frac{ax}{g}\right)$$

1.3 Attitude Representation

Representing three-dimensional rotations mathematically is a non-trivial problem in attitude estimation. While the physical concept of orientation is intuitive, capturing it numerically requires careful consideration of mathematical properties, computational efficiency, and singularity avoidance. This section introduces three common representations used in attitude estimation algorithms, each with distinct advantages and limitations that influence their suitability for different filtering approaches.

1.3.1 Euler Angles

Euler angles are the most intuitive representation of orientation, describing attitude as a sequence of three rotations around specified axes. The aerospace convention, known as the ZYX (yaw-pitch-roll) sequence, applies rotations in the following order [2]:

1. Rotate by ψ (yaw) around the Z-axis (vertical)
2. Rotate by θ (pitch) around the new Y-axis
3. Rotate by φ (roll) around the new X-axis

It only has 3 parameters which it make's it more human interpretable and lesser memory footprint for computer calculations.

Euler angles suffer from non-uniqueness (multiple angle combinations can represent the same physical orientation) and require complex trigonometric transformations when computing angular velocity from angle rates

The Complementary Filter (Chapter 2) operates directly on roll and pitch Euler angles, taking advantage of their simplicity for the basic weighted fusion approach. Furthermore, all filters in this thesis output their final estimates as Euler angles for interpretability and comparison with ground truth measurements. For UAV flight control applications where extreme pitch angles (near ±90°) are rare, Euler angles provide an acceptable representation.

1.3.2 Rotation Matrices ($SO(3)$)

A rotation matrix $R \in SO(3)$ is a 3×3 orthonormal matrix that transforms vectors from one coordinate frame to another. The notation $SO(3)$ stands for "Special Orthogonal group in 3 dimensions," representing all possible rotations in three-dimensional space.

A rotation matrix can be constructed from Euler angles using the product of elementary rotation matrices [3]

Equation 1 Rotation matrix to euler angles: $\boldsymbol{R(\varphi, \theta, \psi) = Rz(\psi) \cdot Ry(\theta) \cdot Rx(\varphi)}$

Where Rx, Ry, Rz represent rotations around the X, Y, and Z axes respectively. The resulting 3×3 matrix contains nine elements, though only three are independent due to orthonormality constraints.

One drawback is redundancy: storing nine numbers to represent three degrees of freedom is inefficient. Also, they require periodic orthonormalization some common methods are Gram-Schmidt or singular value decomposition (SVD)[4].

The Mahony Passive Complementary Filter (Chapter 2) operates directly on rotation matrices, leveraging the geometric properties of SO(3) for error correction. The filter computes a cross-product-based correction term that exploits the manifold structure of rotation matrices, providing superior performance compared to Euler angle approaches. To maintain orthonormality, the Mahony filter implementation applies Gram-Schmidt orthonormalization after each prediction step, ensuring the rotation matrix remains valid throughout the flight.

1.3.3 Quaternions

A quaternion q is a four-parameter representation of rotation, consisting of one scalar and one vector component:

$q = [q_0, q_1, q_2, q_3]^T = [q_0, q_v]^T$

Where:

q_0 : Scalar part

$q_v = [q_1, q_2, q_3]^T$: Vector part

Like rotation matrices, quaternions require normalization after numerical integration to maintain the unit constraint. However, this is computationally cheaper than matrix orthonormalization (single division versus full matrix projection).

Extended Kalman Filter (EKF) uses quaternions for the state representation, with a 7-dimensional state vector that includes the quaternion and gyroscope bias estimates.

1.3.4 Sensor Fusion Fundamentals

Sensor fusion combines measurements from multiple sensors to produce estimates superior to any single sensor. For attitude estimation, gyroscopes and accelerometers exhibit complementary error characteristics: gyroscopes provide smooth, high-frequency measurements but suffer from integration drift, while accelerometers offer a long-term gravity reference but are noisy and corrupted by linear acceleration.

The fundamental challenge is designing algorithms that:

1. Trust gyroscope measurements for short-term dynamics (responsiveness)
2. Use accelerometer measurements for long-term drift correction (stability)
3. Reject accelerometer corruption during vehicle acceleration

Fusion strategies range from simple weighted averaging (complementary filters) to nonlinear filters operating on manifolds (Mahony filter on SO(3)) to probabilistic state estimation (Extended Kalman Filter). Performance is evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) against ground truth measurements.

1.4 Neural Networks for Time-Series Predictions

While classical filtering approaches (detailed in Chapter 2) rely on explicit mathematical models of sensor behavior, modern deep learning offers a data-driven paradigm. This section establishes the theoretical foundation for the LSTM-based predictor developed in Chapter 3, moving from basic sequential learning to the attention mechanisms that enable robust multi-step ahead prediction.

1.4.1 Time-Series Data and Sequential Learning

Attitude estimation is fundamentally a time-series problem. IMU measurements arrive sequentially at fixed intervals Ts, and the vehicle's state at time t is highly dependent on its state at t-1. Traditional feedforward networks process each sample independently, ignoring this temporal structure. For a UAV, where dynamics are continuous, the model must understand these dependencies to distinguish between sensor noise and actual motion.

1.4.2 Recurrent Neural Networks (RNN)

RNNs address the limitations of feedforward networks by maintaining a hidden state h_t that acts as a "memory" of previous inputs. At each timestep t, the RNN processes the current input x_t and the previous hidden state $h\_\{t-1\}$. While theoretically capable of learning long sequences, standard RNNs suffer from the vanishing gradient problem, where information from earlier timesteps is lost during training [6]. For attitude prediction requiring look-back windows of K >= 10 samples, this makes standard RNNs unreliable.

1.4.3 Long Short-Term Memory (LSTM)

The LSTM architecture, introduced by Hochreiter and Schmidhuber [7], solves the vanishing gradient problem through a complex gating mechanism. It maintains a cell state C_t that carries information across long intervals. The flow of information is controlled by three gates:

Forget Gate (f_t): Determines which historical data to discard
Input Gate (i_t): Updates the cell state with new information
Output Gate (o_t): Filter the cell state to product the hidden state h_t

For IMU data, LSTMs can learn to ignore high-frequency vibration noise (via the forget gate) while focusing on the persistent gravity vector (via the input gate).

1.4.4 Attention Mechanisms

Standard LSTMs use only the final hidden state h_k to generate predictions, which can create an information bottleneck. To improve prediction accuracy over a 10-step horizon (n=10) this incorporates an **Attention Mechanism**. This layer calculates a score for every hidden state in the look-back window, allowing the model to "attend" to specific past event (e.g., a sudden change in ω_gyro) that are critical for predicting the future state, The final context vector c is a weighted sum:

$$c = \sum_{i=1}^{T} a_i h_i$$

where a_i represents the learned importance of the $t - th$ timestep. The weights a_i are computed by passing each hidden state through a scoring function followed by a softmax normalization, ensuring they sum to one. This allows the network to dynamically focus on different parts of the input sequence depending on the prediction task.

1.4.5 Related Work

Recent literature has explored deep learning as an alternative or complement to classic filters for IMU-based attitude estimation. Liu et al. [10] demonstrated LSTM-based orientation estimation for UAVs, showing that recurrent networks can learn temporal patterns in IMU data. Weber et al. [13] proposed RIANN, a GRU-based network that takes raw 6-DOF IMU data and outputs quaternion orientation, consistently outperforming Mahony, Madgwick, and complementary filters across diverse motion types and sensor hardware without per-dataset tuning. Golroudbari and Sabour [14] developed hybrid CNN-LSTM architectures evaluated across seven public datasets totaling over 120 hours of IMU measurements, achieving state-of-the-art accuracy for real-time attitude estimation. Brossard et al. [15] took a different approach, training a convolutional network to denoise raw gyroscope signals which are then integrated through dead reckoning, achieving accuracy comparable to visual-inertial odometry systems. Cohen and Klein [11] similarly investigated hybrid models for sensor denoising.

However, a common limitation of all these approaches is that they focus exclusively on **single-step estimation** of the current attitude state. None of them address the problem of **multi-step ahead prediction**, where the model must forecast future attitude angles over a prediction horizon of N timesteps. This distinction is critical: while accurate current-state estimation is sufficient for reactive controllers, advanced control strategies such as Model Predictive Control (MPC) require predictions of future states ($k + 1$ through $k + N$) to compute optimal control actions [12].

2 CHAPTER 2: Tilt Estimation Filters

This chapter presents the implementation, optimization, and comparative analysis of four classical sensor fusion algorithms for UAV attitude estimation. Building upon the theoretical foundations established in Chapter 1, we now examine how different mathematical approaches translate into practical filtering solutions, each offering distinct trade-offs between computational complexity, estimation accuracy, and robustness to sensor noise.

The four algorithms implemented in this work represent a progression of increasing mathematical sophistication. The Complementary Filter provides a baseline approach through simple weighted fusion of gyroscope and accelerometer measurements in Euler angle space. The Mahony Passive Complementary Filter advances this concept by operating on the $SO(3)$ manifold using rotation matrices, incorporating geometric error correction through proportional feedback. The Explicit Complementary Filter extends the Mahony approach by adding integral feedback for gyroscope bias estimation, providing adaptive correction for sensor drift. Finally, the Extended Kalman Filter (EKF) implements a full probabilistic state estimation framework using quaternion representation, offering optimal estimation under the assumptions of Gaussian noise and linear measurement models.

All implementations are developed in C++ using the Eigen library for efficient matrix operations, with careful attention to computational efficiency suitable for real-time embedded deployment. Each filter processes 3-axis gyroscope and accelerometer measurements from a 6-DOF IMU operating at 50 Hz sampling rate. The algorithms are validated against ground truth measurements from a dataset containing 1409 samples collected during quadcopter flight operations, covering a range of dynamic conditions from hover to aggressive maneuvers.

For each filter, this chapter presents the mathematical formulation, implementation details, hyperparameter tuning methodology, and performance evaluation using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) metrics for both roll and pitch estimation. The analysis investigates how each algorithm handles the fundamental sensor fusion challenge: balancing short-term gyroscope accuracy against long-term accelerometer stability while rejecting accelerometer corruption during dynamic flight.

2.1 Quadcopter and Dataset

TODO

2.2 Complementary Filtering

The Complementary Filter is one of the most fundamental and widely used sensor fusion algorithms for attitude estimation in low-cost UAV applications due to its computational simplicity and effectiveness. It addresses the inherent limitations of individual inertial sensors by fusing their measurements in the frequency domain [8].

2.2.1 Algorithm Overview

The algorithm relies on the observation that gyroscope and accelerometer errors expose complementary spectral characteristics. Gyroscopes provide precise, low-noise measurements of angular velocity that are reliable in the high-frequency domain (short term) but suffer from low-frequency drift due to the integration of bias and noise over time. Conversely, accelerometers provide a stable absolute reference to the gravity vector that is reliable in the low-frequency domain (long term) but are subject to high-frequency noise and corruption from external linear accelerations (non-gravitational forces) during dynamic maneuvers.

The discrete-time implementation for the roll angle p is given by:

$$\varphi(t) = \alpha \cdot \left(\phi(t-1) + \omega_{gyro} \cdot dt\right) + (1-\alpha) \cdot \phi_{accel}$$

2.2.2 Parameters

The parameter α control the trust that is placed on each sensor. It takes values in the range of $(0,1)$ and acts as a weighting factor that determines how much the filter relies on gyroscope integration versus accelerometer measurements. When α is close to 1, the filter places greater trust on the gyroscope term. On the other side when close to 0 the filter relies on the accelerometer.

To determine the optimal value, a grid search was performed over $\alpha \in [0.01, 0.99]$ with a step size of 0.01. For each candidate value, the Complementary Filter was executed on the full dataset and the RMSE of the φ estimation against ground truth was computed. The search identified $\alpha = 0.79$ as the value minimizing the roll RMSE.

2.2.3 Experimental results

The Complementary Filter was implemented with a gain coefficient $a = 0.79$ and tested against the ground truth dataset. The results for the Roll and Pitch estimation are presented below.

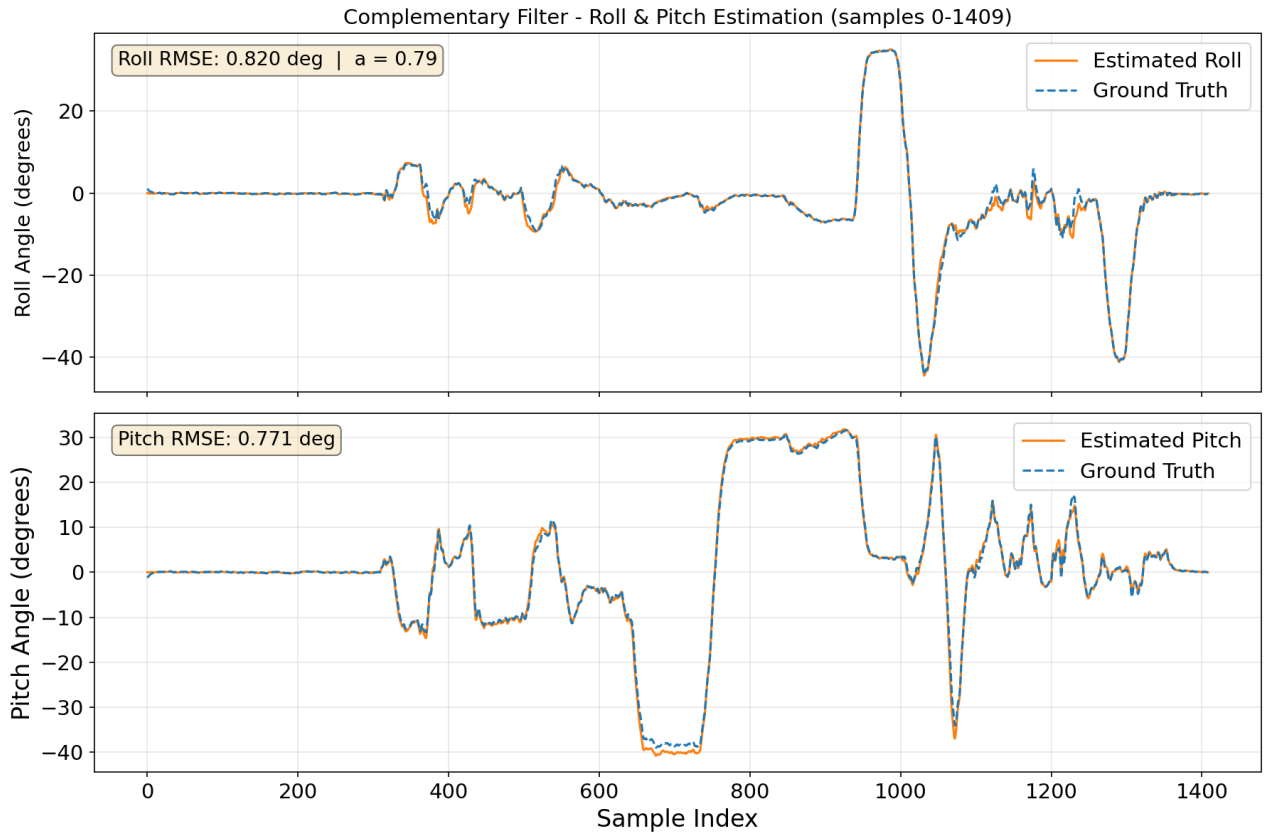


Figure 1 Complementary Filter Roll and Pitch Estimation

Figure 1 represents the global performance of the Complementary Filter over the entire flight duration (1409 samples). The estimated Roll and Pitch angles show a strong correlation with the ground truth data. The low RMSE values 0.820° for Roll and 0.771° for Pitch confirm that the filter successfully understands high-frequency gyroscope data with the low-frequency accelerometer data.

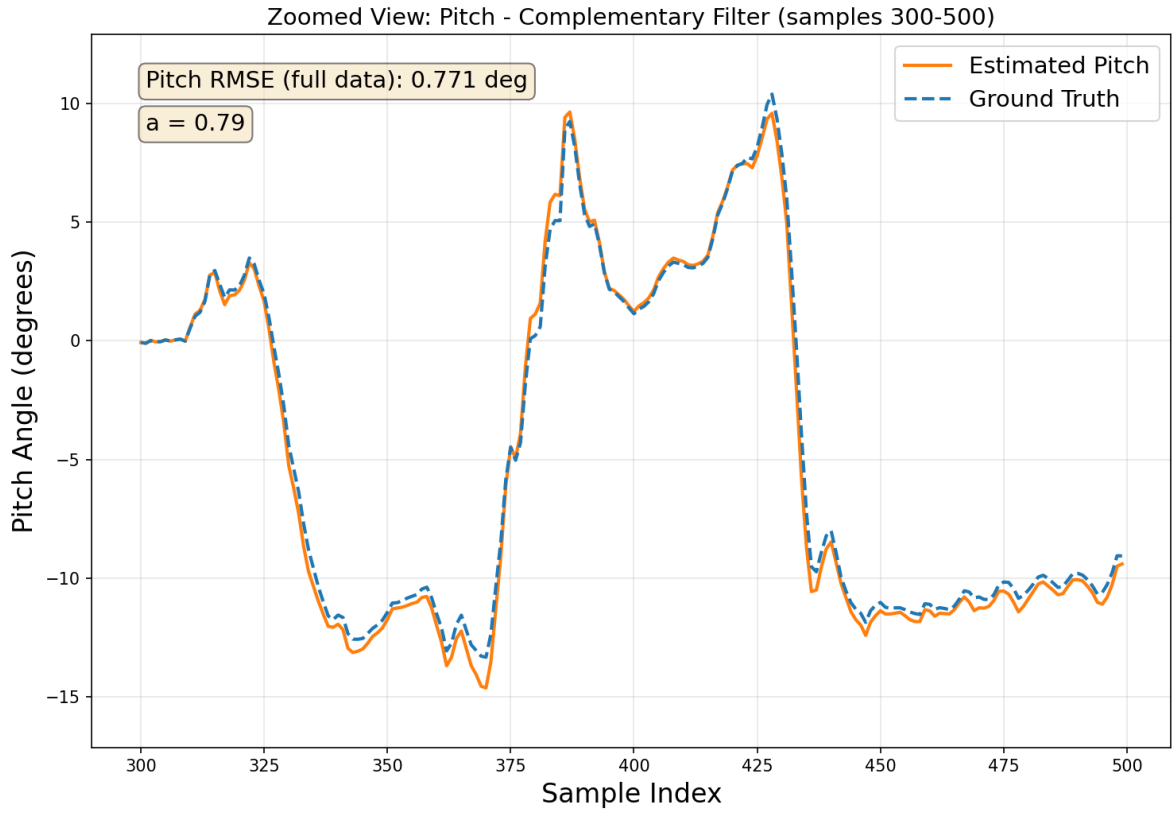


Figure 2 Complementary Filter Zoomed view (samples 300-500) Pitch Estimation

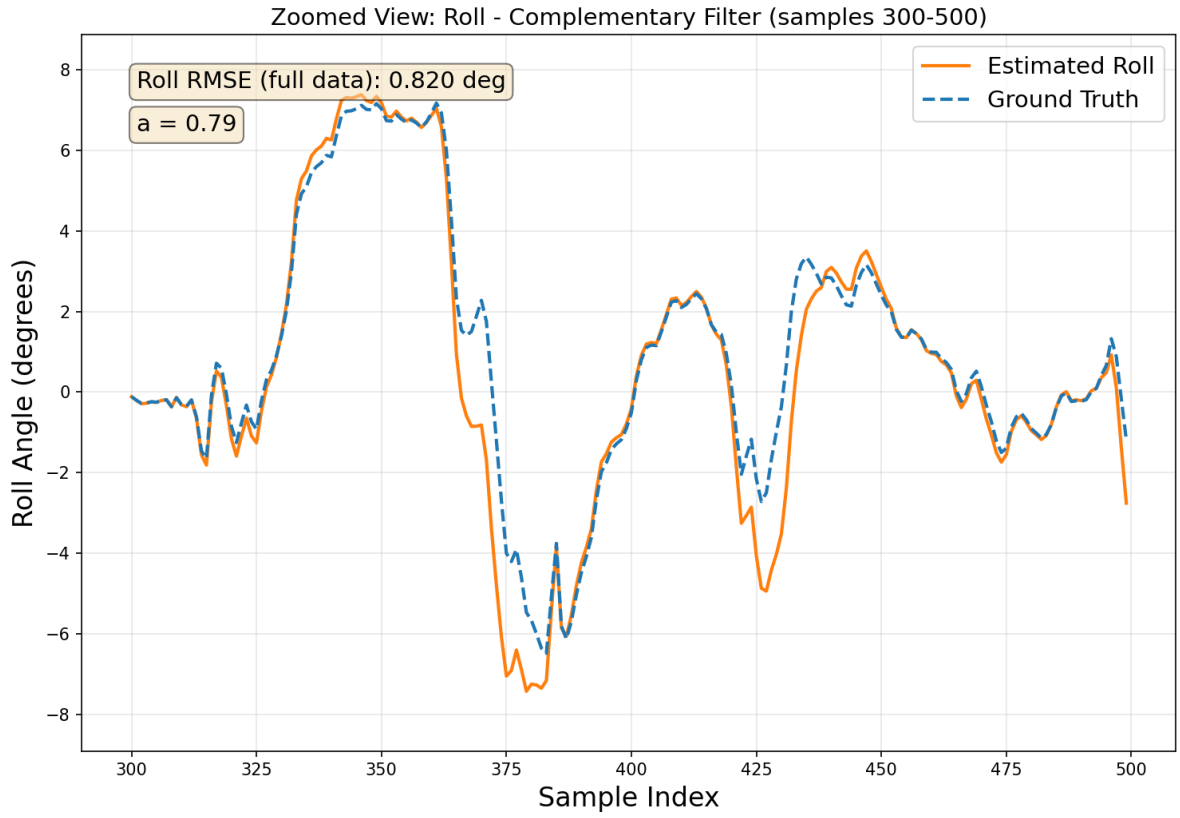


Figure 3 Complementary Filter Zoomed view (samples 300-500) Roll Estimation

Figure 2 provides a more detailed view of the Pitch estimation during a dynamic maneuvering phase (samples 300-500). It can be observed that the Pitch still follows close enough the truth data. Figure 3 illustrates the Roll estimation for the same interval. Generally, it performs almost as well as Pitch with a minor exception at around sample 375 which the filter fails to follow the truth data.

In conclusion, the Complementary Filter proves to be a robust and computationally efficient solution for general UAV attitude estimation, it successfully mitigates gyroscope drift and provides a smooth estimate suitable for flight control. The primary limitation is the reliance on a static α gain means it cannot adapt to various flight conditions that are encountered often in the real world.

2.3 Mahony Filter

The Mahony filter is a nonlinear complementary filter that operates directly on the Special Orthogonal group $SO(3)$, representing orientation using rotation matrices [4]. Unlike the standard Complementary Filter which operates on Euler angles, the Mahony filter avoids singularities (gimbal lock) and provides a more geometric approach to error correction.

The core principle of the Mahony filter is to estimate the rotation matrix $\hat{R}$ by comparing the measured direction of gravity (from the accelerometer) with the estimated direction of gravity. The error between these two vectors is used to generate a correction term that "steers" the gyroscope integration towards the true vertical.

The implementation done here is a "Passive" Mahony filter as described in [4].

2.3.1 Algorithm Overview

The Mahony filter maintains an estimated rotation matrix $\hat{R}$ that represents the current attitude. At each timestep, the filter compares the measured direction of gravity (derived from the accelerometer) with the direction predicted by the current estimate. The discrepancy between these two is used to generate a correction term that steers the gyroscope integration towards the true orientation.

A rotation matrix R_y is constructed from the accelerometer tilt angles, and the attitude error is computed as:

$$\tilde{R} = \hat{R}^T R_y$$

The correction vector is extracted from the anti-symmetric part of this error matrix:

$$\omega_{mes} = vec(\left(P_a(\tilde{R})\right), \quad P_a(\tilde{R}) = \frac{1}{2}(\tilde{R} - \tilde{R}^T)$$

The rotation matrix then updated using the correct angular velocity:

$$\dot{\hat{R}} = \hat{R}\left[\omega_y + k_p \omega_{mes}\right]_\times$$

After each discrete integration step, $\hat{R}$ is projected back onto SO(3) via SVD orthonormalization to prevent numerical drift.

2.3.2 Parameters

The Mahony filter has a single tunable hyperparameter: the proportional gain k_p. This parameter controls the accelerometer-based correction applied to the gyroscope integration at each timestep.

When k_p is small (close to zero), the filter relies almost entirely on the gyroscope for attitude estimation. The gyroscope provides accurate short-term measurements but accumulates drift over time. When k_p is large, the filter applies aggressive corrections based on the accelerometer measurement. While this eliminates gyroscope drift, it make the estimate highly sensitive to accelerometer noise and non-gravitational acceleration that occur during maneuvers [4].

The optimal value of k_p must balance these two effects: it should be large enough to prevent gyroscope drift within the operating timeframe, yet small enough to maintain smooth estimates during dynamic flight. Unlike the Complementary Filter's a parameter which directly weights two angle estimates, k_p operates in the angular velocity domain, scaling the correction signal that is added to the gyroscope before integration. This provides a more natural correction mechanism that respects the geometric structure of rotations.

To determine the optimal value, a grid search was performed over $k_p \in [1,100]$ with integer steps. For each candidate value, the Mahony filter was executed on the full dataset and the RMSE of both ϕ and θ was computed against ground truth. The optimization criterion was the average of the roll and pitch RMSE, ensuring that both angles are estimated well simultaneously. The search identified $k_p = 11$ as the optimal value.

2.3.3 Experimental Results

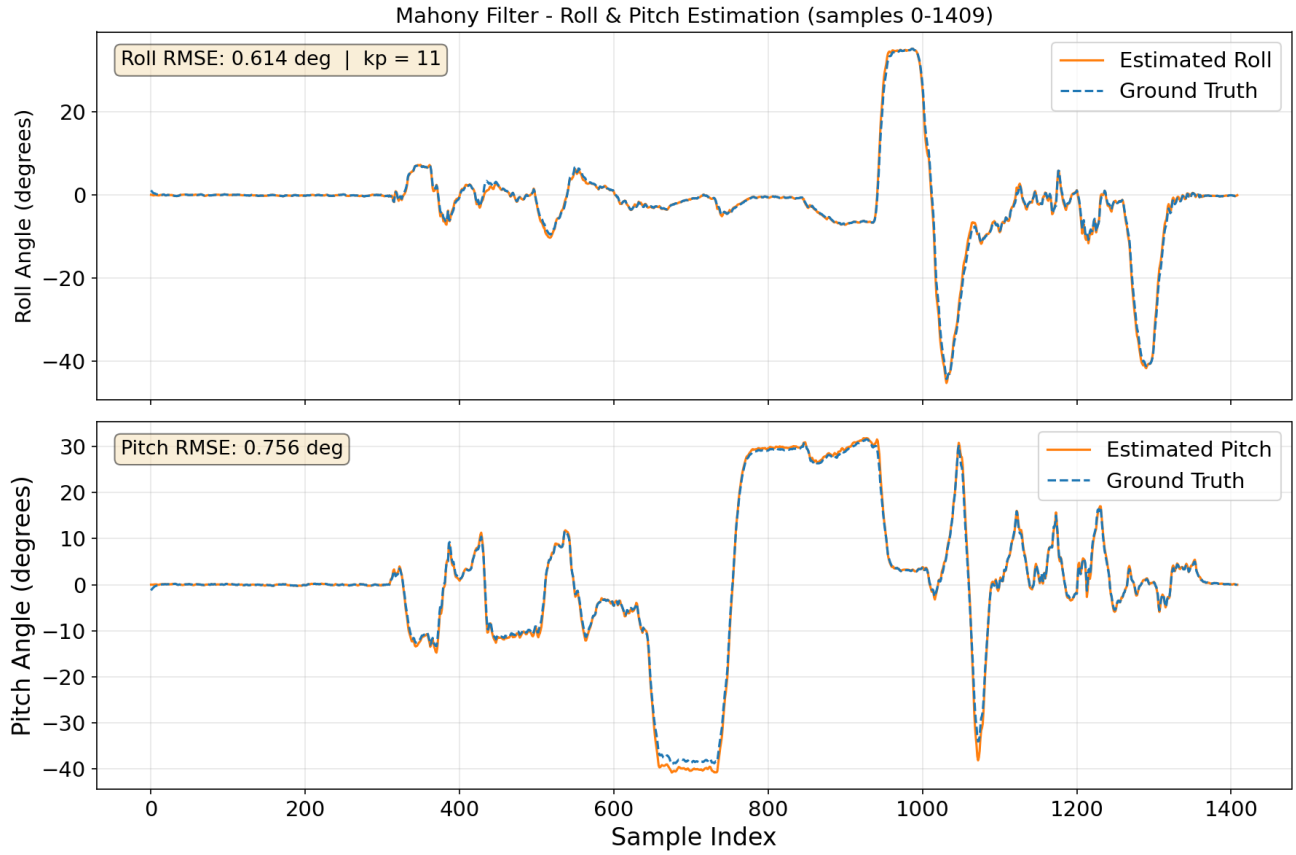


Figure 4 Mahony Filter Roll and Pitch Estimation

Figure 4 represents the global performance of the Mahony filter over the entire flight duration (1409) samples. The estimated ϕ and θ angles show a strong correlation with the ground truth data. The filter achieves an RMSE of 0.614 for ϕ and 0.756 θ, representing a significant improvement over the Complementary Filter. The improvement is particularly notable for the roll angle ϕ, where the filter reduces the RMSE 25% approximately.

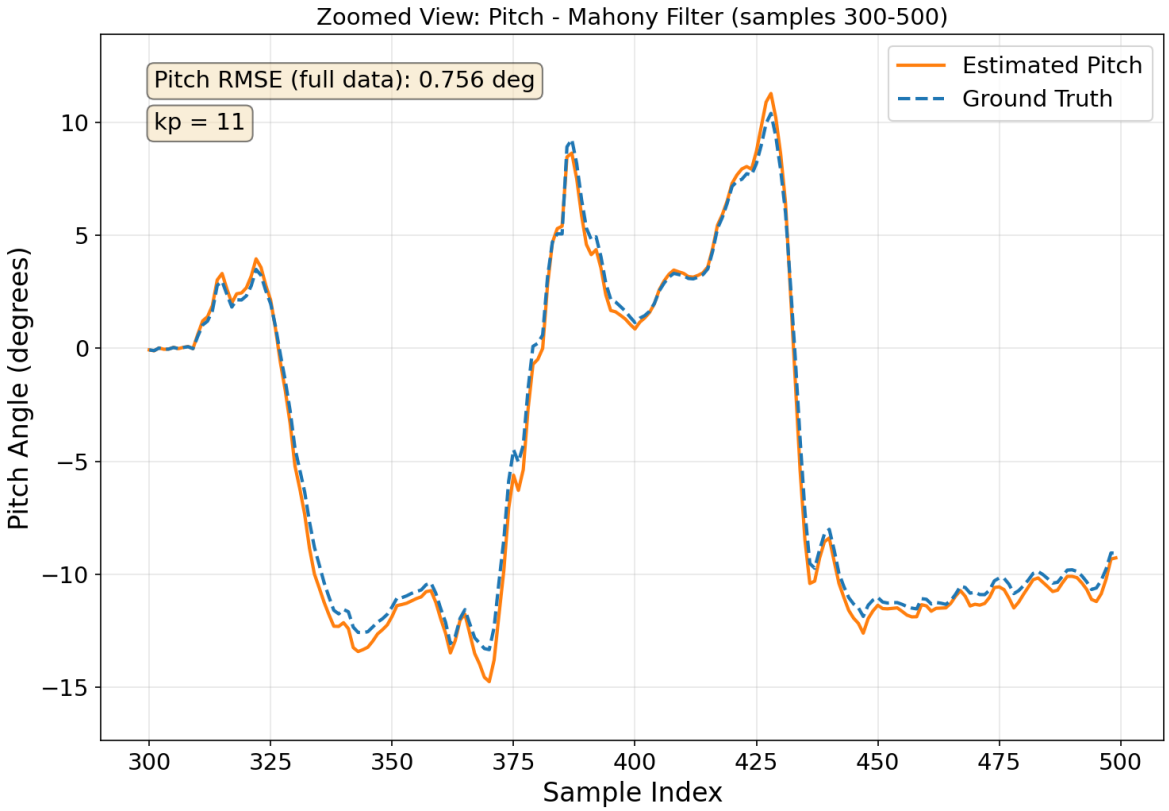


Figure 5 Mahony Filter Zoomed view (samples 300-500) Pitch Estimation

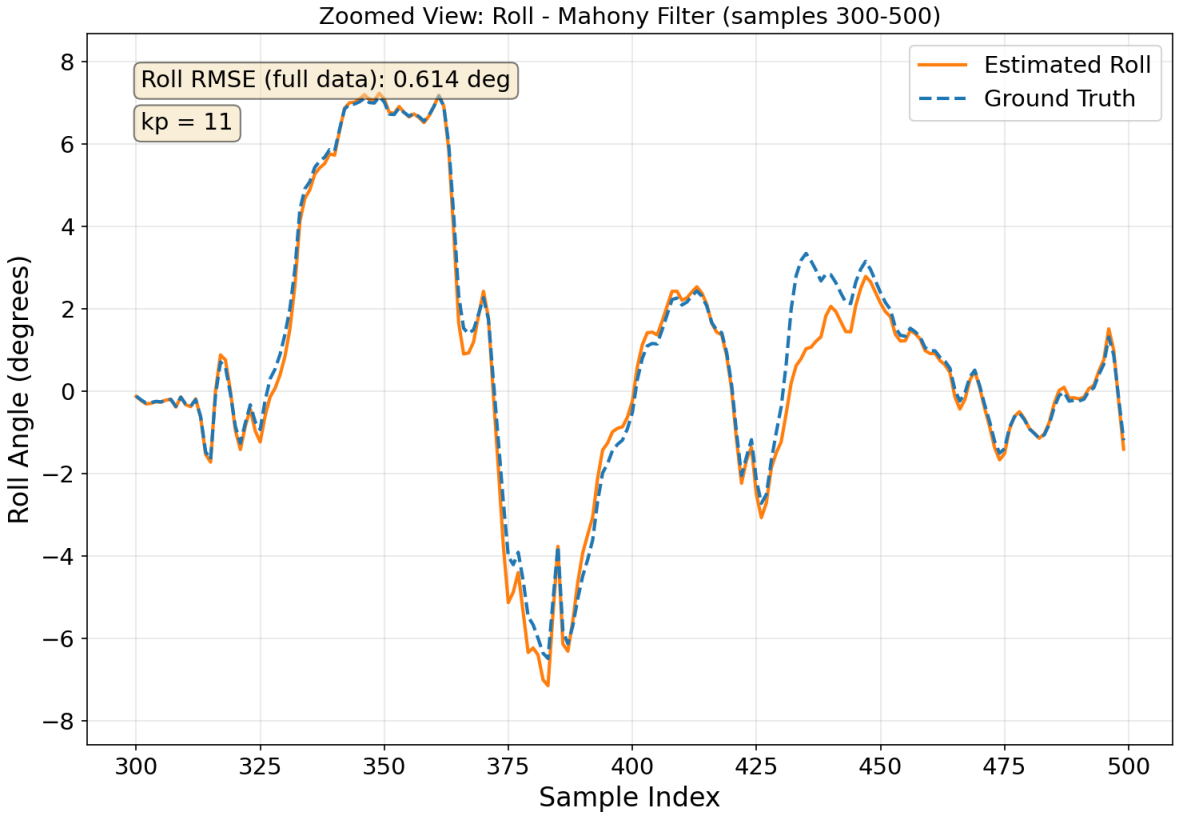


Figure 6 Mahony Filter Zoomed view (samples 300-500) Roll Estimation

Figure 5 and Figure 6 provide zoomed view of the attitude angles during a maneuvering phase (samples 300-500). Compared to the Complementary Filter, the Mahony filters the rapid angles changes more closely. The geometric error correction on $SO(3)$ allows the filter to handle large rotation more gracefully than the Euler angle-based approach of the Complementary Filter.

In conclusion, the Mahony Passive Complementary Filter demonstrates improved estimation accuracy over the basic Complementary Filter while maintaining computational efficiency suitable for real-time applications. Operating on $SO(3)$ with rotation matrices provides inherent robustness against gimbal lock and allows geometric error correction. The single hyperparameter k_p makes tuning straightforward, and the SVD-based orthonormalization ensures numerical stability during prolonged operation. However, the proportional-only correction means that persistent biases in the gyroscope may not be fully eliminated over very long time periods — a limitation addressed by the Explicit Complementary Filter in the following section.

2.4 Explicit Complementary Filtering

The Explicit Complementary Filter extends the Mahony approach by introducing two key modifications: it operates directly on vectorial sensor measurements without requiring intermediate attitude reconstruction, and it incorporates integral feedback for online gyroscope bias estimation. Both features originate from Section V of Mahony et al. [4], where the filter is derived as an "explicit" complement to the passive formulation presented in the previous section.

2.4.1 Algorithm Overview

The filter maintains two state variables: an estimated rotation matrix $\hat{R}$ in $SO(3)$ and an estimated gyroscope bias vector $\hat{b}$ in R^3. At each timestep, the correction term is computed as the cross product between the measured and estimated gravity directions:

$\omega_{mes} = u \times \hat{u}, \hat{u} = \hat{R}^T u_0$

where $u_o = [0,0,1]^T$ is the gravity reference direction in the internal frame, u is the normalized accelerometer reading and $\hat{u}$ is the filter's predicted gravity direction in the body frame. The cross product produces a correction vector that is zero when the two directions are aligned, and grows proportionally to the angular error between them.

The two remaining core filter equations are:

$$\dot{\hat{R}} = \hat{R}\left[\omega^y - \hat{b} + k_p \omega_{mes}\right]_\times$$

$$\dot{\hat{b}} = -k_i \omega_{mes}$$

The first equation propagates the rotation matrix using the gyroscope measurement ω^y, corrected for the current bias estimate $\hat{b}$ and augmented with proportional feedback scaled by k_p. The second equation drives the bias estimate toward the true gyroscope bias over time through integral feedback scaled by k_i. After each discrete integration step, $\hat{R}$ is projected back onto $SO(3)$ via SVD orthonormalization to prevent numerical drift.

2.4.2 Parameters

The filter has two tunable hyperparameters: the proportional gain k_p and the integral gain k_i.
The proportional gain controls the correction applied to the gyroscope integration at each timestep, serving the same role as it in the passive Mahony filter. When k_p is small the filter relies primary on the gyroscope integration. On the other side a large number implies aggressive correction to reduce drift but increase sensitivity to accelerometer noise.
The integral gain k_i dictates the adaptation speed of the bias estimator. While an aggressive (larger) value ensures rapid convergence toward the true gyroscope bias, it risks inducing instability and oscillatory behavior in the attitude estimate. Conversely, a conservative (smaller) value prioritizes estimation stability and noise rejection.

To determine the optimal values, a two-dimensional grid search was performed over k_p in [1.0, 15.0] with a step size of 0.5 and k_i in [0.05, 1.0] with a step size of 0.05, yielding 580 parameter combinations. For each pair, the filter was executed on the full dataset and the average RMSE of phi and theta was computed against ground truth. The search identified k_p = 11 and k_i = 0.05 as the optimal combination.

2.4.3 Experimental Results

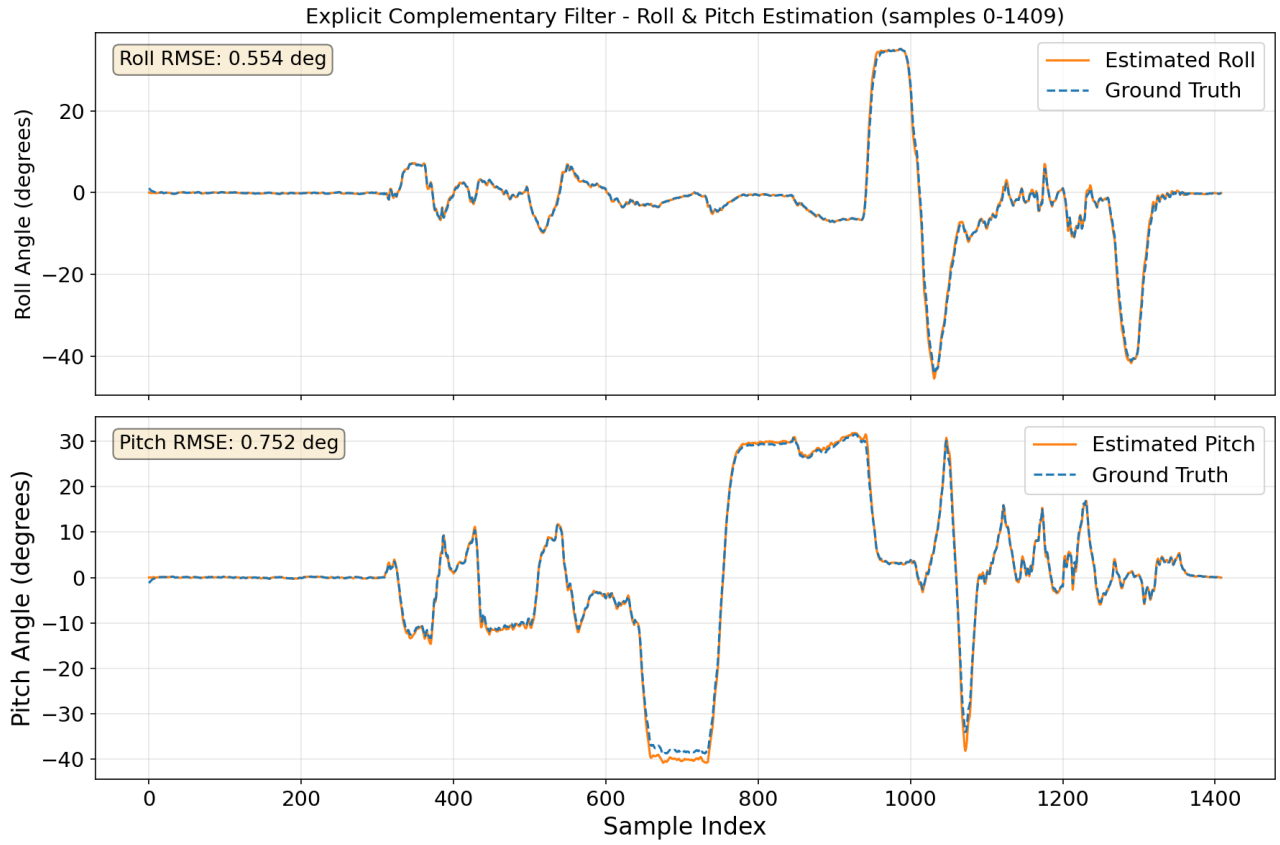


Figure 7 Explicit Complementary Filter Roll and Pitch Estimation

Figure 7 represents the global performance of the Explicit Complementary Filter over the entire flight duration. The filter achieves an RMSE of 0.554 degrees for ϕ and 0.752 degrees for θ. Compared to the "Passive" Mahony filter this one reduces the roll RMSE by an approximately 10% while maintaining comparable pitch performance.

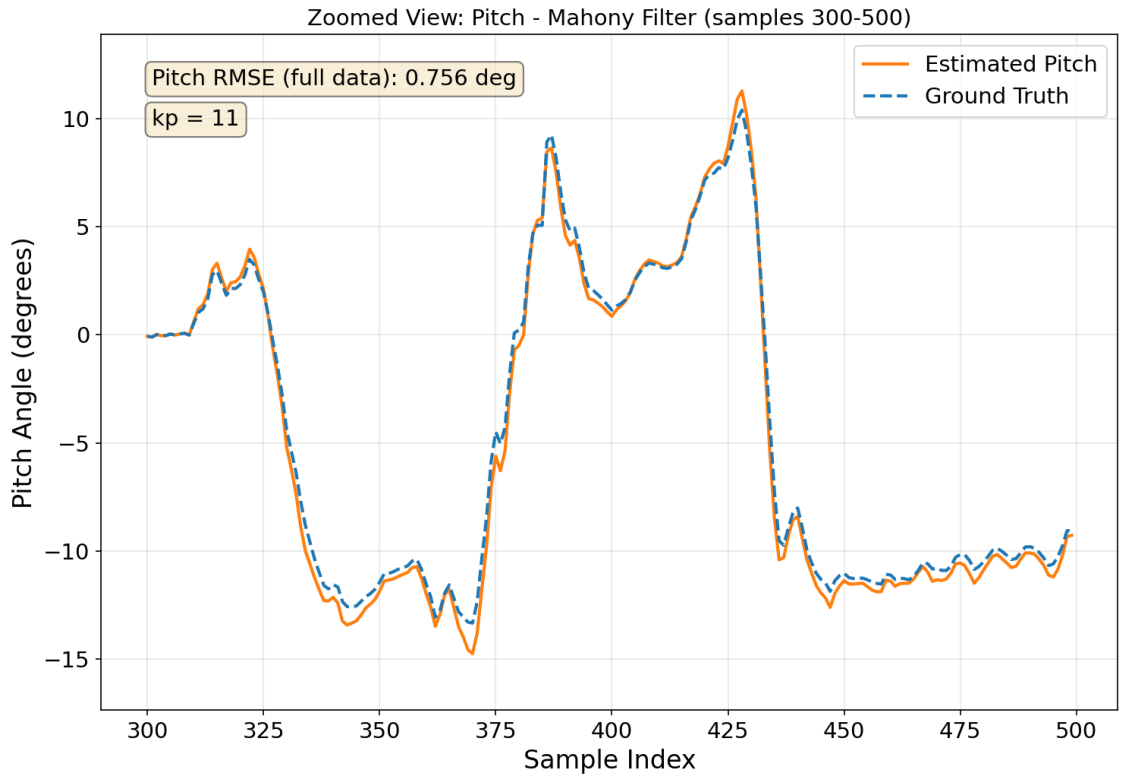


Figure 8 Explicit Complementary Zoomed view (samples 300-500) Pitch Estimation

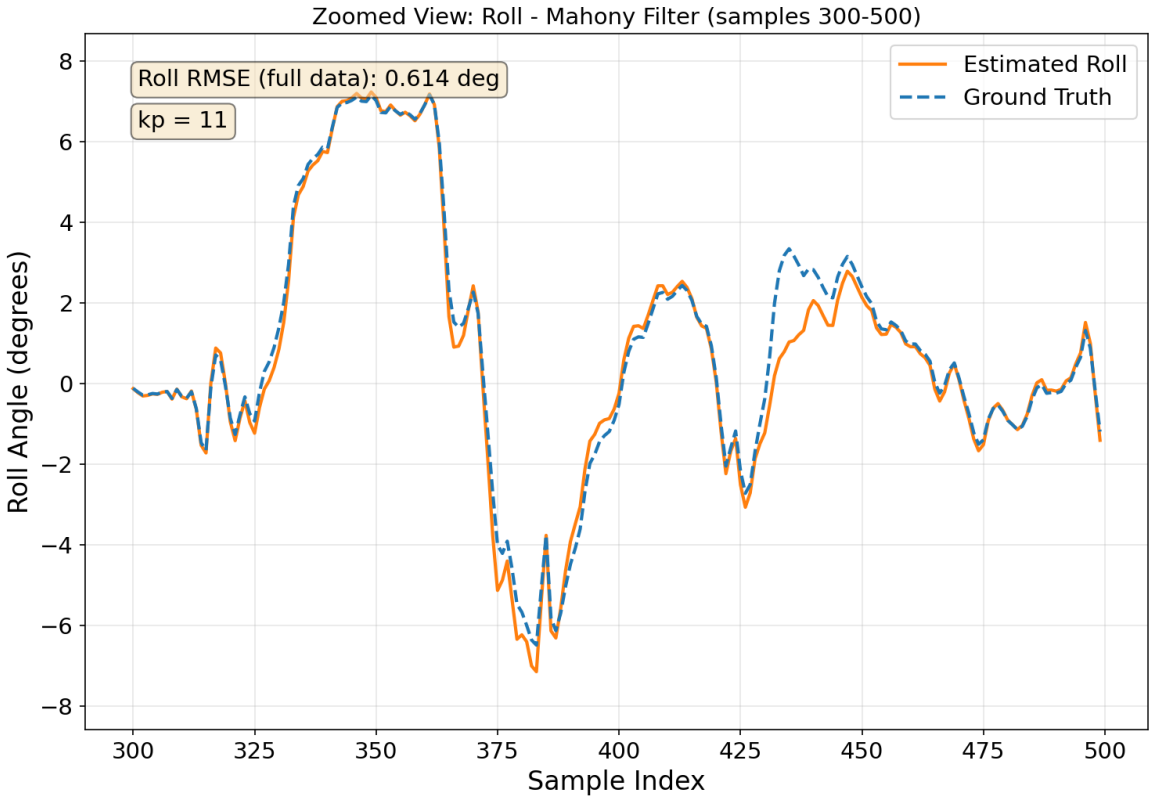


Figure 9 Explicit Complementary Zoomed view (samples 300-500) Roll Estimation

Figure 8 and Figure 9 provide a zoomed view of the attitude angles during a dynamic maneuvering phase. The estimates track the ground truth closely, with performance visually similar to the Mahony Filter. The improvement is most apparent in the roll angle, where the bias estimation mechanism provides a small but consistent reduction in tracking error.

In conclusion, the Explicit Complementary Filter demonstrates a modest improvement over the Passive Mahony Filter, primarily in roll estimation accuracy. Both filters share the same proportional correction mechanism operating on $SO(3)$, the same SVD-based orthonormalization, and the same optimal k_p value. The additional integral feedback for bias estimation provides a theoretical advantage that would become more pronounced in scenarios with significant gyroscope bias drift, which is not the case with the dataset. On the present dataset, where the gyroscope is well-calibrated, the improvement is limited to approximately 10% in roll RMSE. The filter's key architectural contribution, operating directly on vector measurements rather than requiring attitude reconstruction, reduces computational cost and eliminates potential numerical issues associated with the trigonometric extraction step, making it particularly suitable for resource-constrained embedded platforms.

2.5 Extended Kalman Filtering (EKF)

The Extended Kalman Filter is a probabilistic state estimation framework that extends the classical Kalman filter to nonlinear systems through the first-order linearization [16]. Unlike

the complementary filters presented in the previous sections, the EKF maintains an explicit uncertainty estimate in the form of a covariance matrix, and fuses sensor measurements in a statistically optimal manner under the assumption of Gaussian noise.

The EKF represents attitude using a unit quaternion rather than a rotation matrix. Quaternions avoid the singularities inherent to Euler angle representations (gimbal lock) and are computationally more efficient than rotation matrices for integration [17]. The state vector is:

$x = [q_0, q_1, q_2, q_3, b_x, b_y, b_z]^T \epsilon R^7$

The $q = [q_0, q_1, q_2, q_3]^T$ is the unit quaternion encoding the current attitude and $b = [b_x b_y b_Z]^T$ is the estimated gyroscope bias. The filter alternates between a prediction step driven by the gyroscope and an update step driven by the accelerometer.

2.5.1 Algorithm Overview

Prediction step: The quaternion is propagated forward using first-order quaternion kinematics. The bias-corrected angular velocity is $w = \omega_y - b$ and the quaternion update is $q_{k+1} = q_k + \frac{\Delta t}{2} \Xi(w) q_k$ where $\Xi(w)$ is the quaternion derivative matrix constructed from the bias corrected angular velocity w. The covariance is propagated using the linearized state transition Jacobian F:

$P_{k+1} = F\, P_k F^T$

Update step: The accelerometer provides a measurement of the gravity direction in the body frame. The expected measurement is computed by rotating the inertial gravity vector through the current quaternion estimate:

$h = R(q)^T g, \quad g = [0.0.1]^T$

The innovation $y = a_{norm} - h$, where a_{norm} is the normalized accelerometer reading, represents the discrepancy between the measured and predicted gravity directions. The Kalman gain and state update are then:

$K = PH^T(HPH^T + R)^{-1}$

$x \leftarrow x + Ky, \;\; P \leftarrow (I - KH)P$

2.5.2 Parameters

The EKF is governed by two noise covariance matrices: the process noise Q $(7x7)$ and the measurement noise R $(3x3)$[18]. These matrices encode the designer's trust in the process model versus the sensor measurements.

The process noise Q determines how much uncertainty is injected at each prediction step. A larger Q causes the filter to distrust its own propagation and rely more heavily on measurements, while a smaller Q makes the filter more confident in the gyroscope integration. Separate values are assigned to the quaternion components and the bias states, since these evolve at different timescales: the quaternion changes rapidly with motion, while the bias drifts slowly.

The measurement noise R encodes the expected noise level of the accelerometer. A larger R makes the filter discount accelerometer corrections, which is appropriate during dynamic maneuvers where non-gravitational accelerations contaminate the measurement. A smaller R increases responsiveness to accelerometer corrections at the cost of noise sensitivity.

For this implementation, the values were set to $Q_q = 0.001 * I_4$ (quaternion process noise), $Q_b = 0.0001 * I_3$ (bias process noise), and $R = 0.1 * I_3$ (accelerometer measurement noise). These values are consistent with standard practice for quaternion-based EKF attitude estimation [18].

2.5.3 Experimental Results

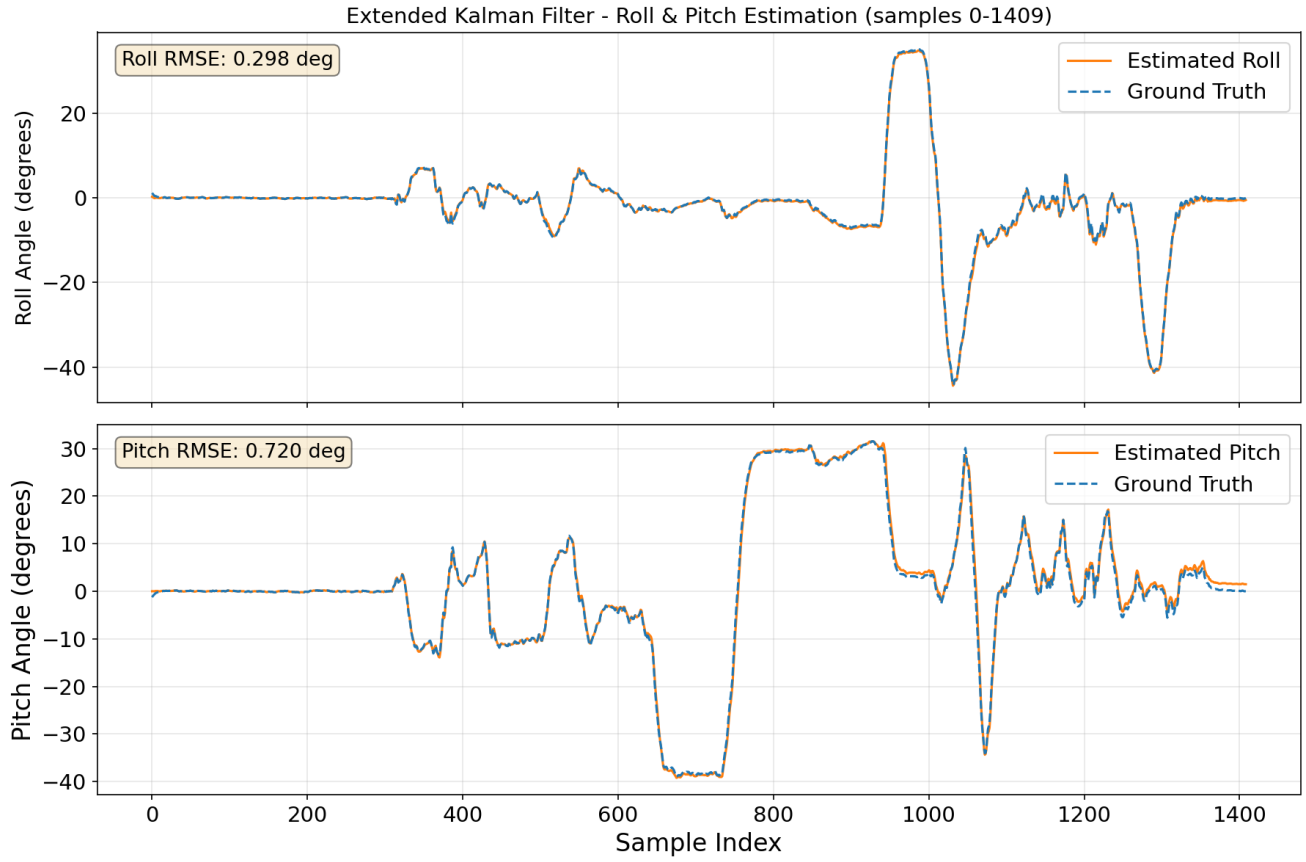


Figure 10 EKF Roll and Pitch Estimation

Figure 10 represents the global performance of the EKF over the entire flight duration. The filter achieves an RMSE of 0.298 degrees for ϕ and 0.720 for θ. The roll estimation represents that best performance among all four filters implemented in this chapter, with a 51% reduction in roll RMSE compared to the basic Complementary Filter and a 52% reduction compared to Mahony Filter.

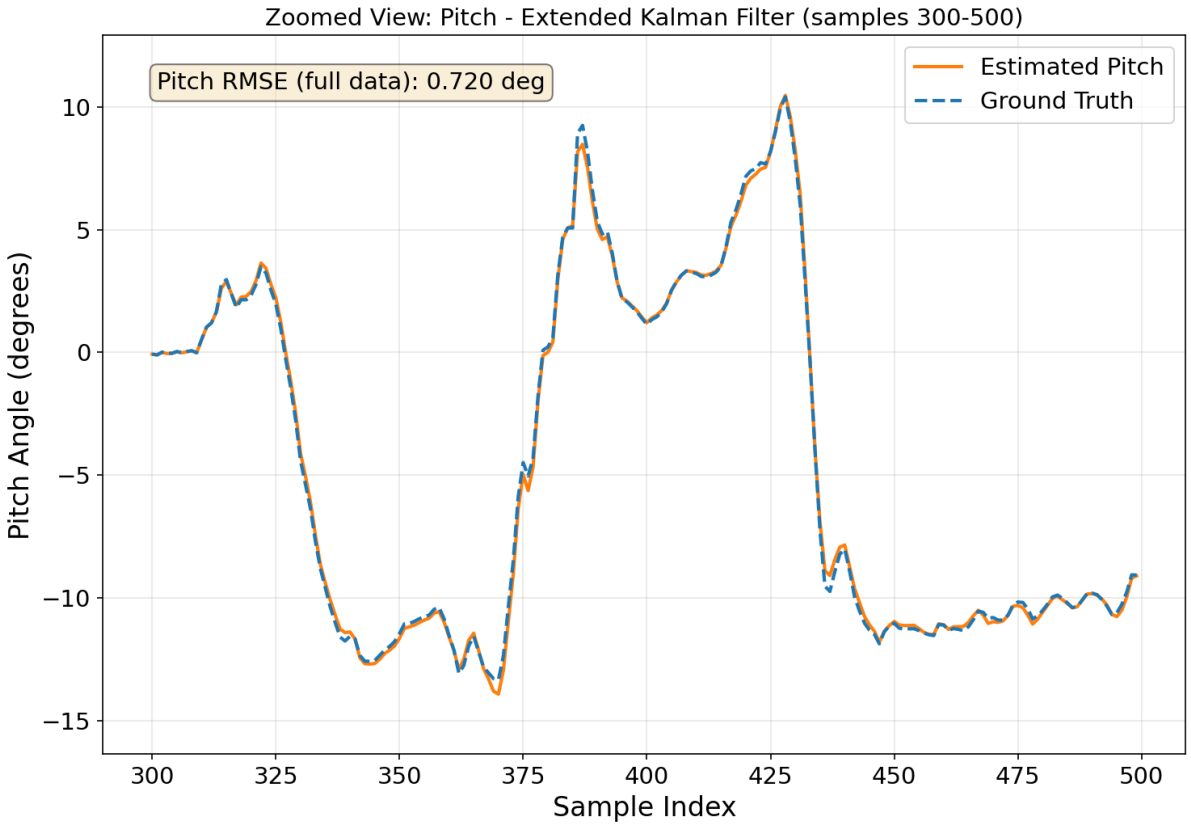


Figure 11 EKF Zoomed view (samples 300--500) Roll Estimation

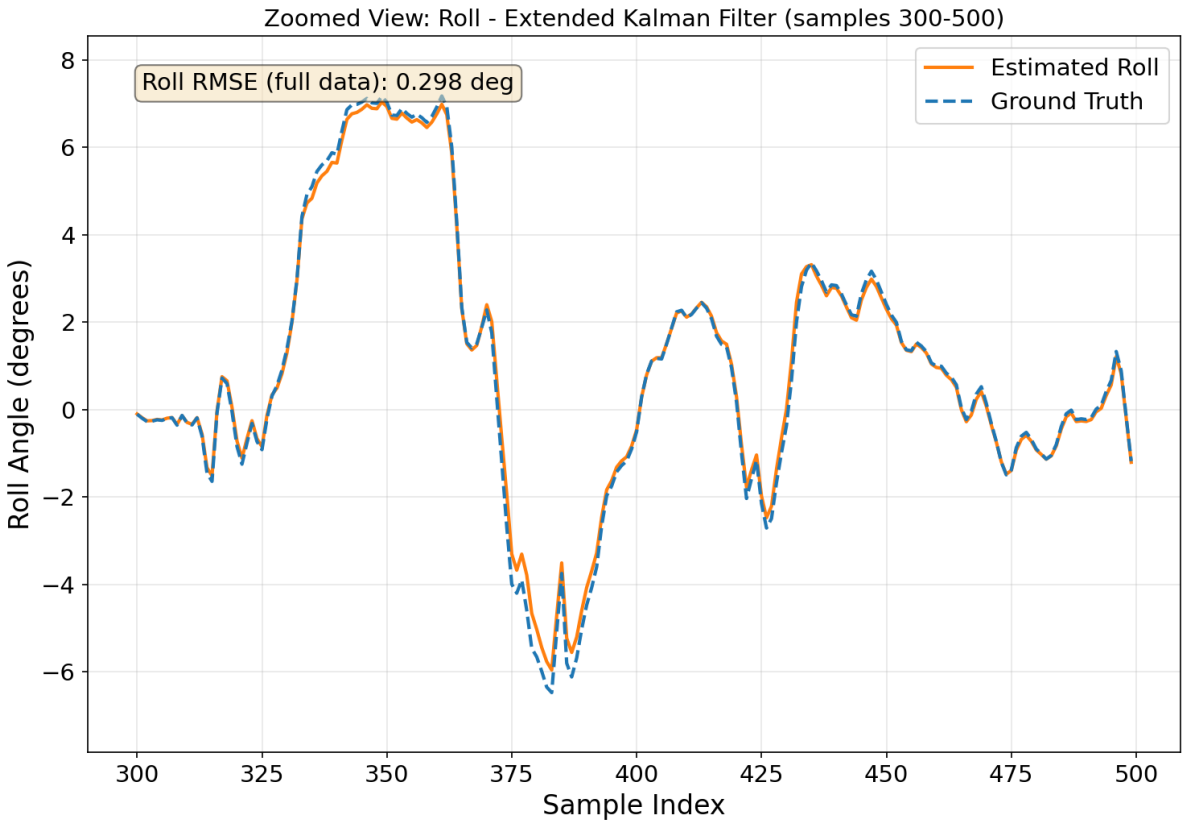


Figure 12 EKF Zoomed view (samples 300--500) Pitch Estimation

Figure **11** and Figure 12 provide a zoomed view during the dynamic maneuvering phase (samples 300-500). The EKF tracks the ground truth more closely than the complementary filters, particularly in the roll axis where the bias estimation and probabilistic weighting provide the most benefit.

In conclusion, the EKF achieves the highest roll estimation accuracy of the four filters, demonstrating the advantage of probabilistic state estimation with explicit uncertainty tracking. However, this performance comes at a higher computational cost: at each timestep the filter maintains and updates a 7x7 covariance matrix and computes a 3x7 Kalman gain, making it more demanding than the complementary filter approaches. The tuning of the Q and R matrices also requires more domain knowledge than the single-gain tuning of the Mahony and Explicit CF filters.

3 Chapter 3: Deep learning approach

TODO

Conclusions

Future Work

Bibliography – References – Online sources

1. Zhou, X., Chen, L., Sun, C., Jia, W., Yi, N., & Sun, W. (2025). Highly Accurate Attitude Estimation of Unmanned Aerial Vehicle Payloads Using Low-Cost MEMS. *Micromachines*, *16*(6), 632. https://doi.org/10.3390/mi16060632
2. Diebel, J. (2006). "Representing Attitude: Euler Angles, Unit Quaternions, and Rotation Vectors." Stanford University Technical Report
3. Craig, J. J. (2005). Introduction to Robotics: Mechanics and Control. 3rd Edition. Pearson Prentice Hall.
4. R. Mahony, Tarek Hamel, Jean-Michel Pflimlin. Nonlinear Complementary Filters on the Special Orthogonal Group. *IEEE Transactions on Automatic Control*, 2008, 53 (5), pp.1203-1217. ⟨10.1109/TAC.2008.923738⟩. ⟨hal-00488376⟩
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
6. Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157-166.
7. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780
8. Higgins, W. T. (1975). "A comparison of complementary and Kalman filtering." IEEE Transactions on Aerospace and Electronic Systems.
9. Gajamannage, K., Jayathilake, D. I., Park, Y., & Bollt, E. M. (2023). Recurrent neural networks for dynamical systems: Applications to ordinary differential equations, collective motion, and hydrological modeling. *Chaos (Woodbury, N.Y.)*, *33*(1), 013109. https://doi.org/10.1063/5.0088748
10. Liu, Y., Zhou, Y., & Li, X. (2018). Attitude Estimation of Unmanned Aerial Vehicle Based on LSTM Neural Network. 2018 International Joint Conference on Neural Networks (IJCNN), 1-6.
11. Cohen, N., & Klein, I. (2024). Inertial navigation meets deep learning: A survey of current trends and future directions. *Digital Signal Processing*, *157*, Article 104950. https://doi.org/10.1016/j.dsp.2024
12. Shi, X., Wang, J., Tang, L., & Liu, H. (2021). Receding horizon control of an unmanned quadrotor helicopter flying through a time-varying narrow aperture. *Aerospace Science and Technology*, *112*, Article 106602. https://doi.org/10.1016/j.ast.2021.106602
13. Weber, D., Guhmann, C., & Seel, T. (2021) "RIANN - A Robust Neural Network Outperforms Attitude Estimation Filters. AI (MDPI), 2(3), 444-463 .
14. Golroudbari, A.A., & Sabour, M.H. (2023). Generalizable End-to-End Deep Learning Frameworks for Real-Time Attitude Estimation Using 6DoF Inertial Measurement Units. Measurement, 217, 113105.
15. Brossard, M., Bonnabel, S., & Barrau, A. (2020). Denoising IMU gyroscopes with deep learning for open-loop attitude estimation. *IEEE Robotics and Automation Letters*, *5*(3), 4796-4803.
16. Welch & Bishop (1995) "An Introduction to the Kalman Filter."
17. Solà, J. (2017). *Quaternion kinematics for the error-state Kalman filter* (arXiv:1711.02508). arXiv. https://arxiv.org/abs/1711.02508

18. Sabatini A. M. (2006). Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing. *IEEE transactions on bio-medical engineering*, *53*(7), 1346–1356. https://doi.org/10.1109/TBME.2006.875664

Appendix A

……

Appendix B

……

Appendix C

……