



Degree Project in Technology

Second cycle, 30 credits

Vision-Based Localization at Intersections using Riccati Observers

ANNIKA WONG

Vision-Based Localization at Intersections using Riccati Observers

ANNIKA WONG

Master's Programme, Systems, Control and Robotics, 120 credits
Date: July 18, 2025

Supervisors: Zhiqi Tang, Frank J. Jiang

Examiner: Jonas Mårtensson

School of Electrical Engineering and Computer Science

Swedish title: Visionsbaserad lokalisering vid väggkorsningar med Riccati-observatörer

Abstract

Localization is critical for autonomous vehicles to reliably estimate their position and orientation within complex and dynamic environments. This thesis develops a lightweight, computationally efficient localization algorithm with guaranteed convergence, specifically tailored for intersection scenarios. Intersections pose significant challenges due to vehicles approaching from multiple directions and frequent occlusions from surrounding vehicles, complicating observability and localization accuracy.

The proposed bearing-based Riccati localization method integrates bearing measurements from known static landmarks and vehicle velocity measurements (linear and angular). The method is validated experimentally using a scaled vehicle platform (SVEA) equipped with a camera and ArUco markers as landmarks, simulating intersection conditions at scale. Scenarios examined include nominal driving conditions, aggressive maneuvers near operational limits, and degraded sensing conditions characterized by sensor inaccuracies and measurement noise.

The approach demonstrates high accuracy and robustness, maintaining reliable localization even during dynamic maneuvers and sensor disturbances. This thesis underscores the potential of Riccati localization as a practical and scalable solution for autonomous vehicle localization at intersections. Future work should involve formal analysis of tolerable delays to maintain system stability, exploration of cooperative multi-agent scenarios, and integration with advanced communication infrastructures such as 5G to further enhance real-time localization performance.

Keywords

Localization, Perspective-n-Point problem, Observability, Riccati observer

Sammanfattning

Lokalisering är avgörande för autonoma fordon för att pålitligt kunna uppskatta sin position och orientering i komplexa och dynamiska miljöer. Denna avhandling utvecklar en lättviktig och beräkningsmässigt effektiv lokaliseringsalgoritm med garanterad konvergens, specifikt anpassad för korsningsscenarier. Korsningar utgör betydande utmaningar på grund av fordon som närmar sig från flera riktningar och frekventa ocklusioner från omgivande fordon, vilket försvårar observerbarhet och lokaliseringsnoggrannhet.

Den föreslagna Riccati-baserade lokaliseringsmetoden utnyttjar bäringsmätningar från kända statistiska landmärken samt fordonets hastighetsmätningar (linjära och angulära). Metoden valideras experimentellt med hjälp av en skalerad fordonsplattform (SVEA) utrustad med kamera och ArUco-markörer som landmärken, vilket simulerar korsningsförhållanden i mindre skala. Undersökta scenarier inkluderar nominella körförhållanden, aggressiva manövrar nära fordonets dynamiska gränser samt försämrade sensorförhållanden präglade av mätfel och brus.

Metoden visar hög noggrannhet och robusthet, med pålitlig lokalisering även under dynamiska manövrar och vid sensorstörningar. Avhandlingen belyser potentialen hos Riccati-lokalisering som en praktisk och skalbar lösning för autonom fordonslokalisering i korsningar. Framtida arbete bör innefatta formell analys av tolerabla fördröjningar för att upprätthålla systemets stabilitet, utforskning av kooperativa multiagentlösningar samt integration med avancerad kommunikationsinfrastruktur såsom 5G för att ytterligare förbättra realtidslokaliseringsprestanda.

Nyckelord

Lokalisering, Perspective-n-Point problem, Observerbarhet, Riccati-observatör

Contents

1	Introduction	1
1.1	Overview	1
1.2	Structure of the thesis	2
2	Background	3
2.1	Perspective-n-Point (PnP) Problem	3
2.2	Bearing Measurement	4
2.3	Observability	5
2.3.1	Uniform Observability	5
2.3.2	Observability Grammian	5
2.4	Continuous Riccati Equation	6
2.5	Graph Theory	6
2.6	Related Work	7
3	Method	9
3.1	Scenarios and System Overview	9
3.2	Bearing-based Riccati Localization	10
3.2.1	Mathematical Models	11
3.2.1.1	Dynamic Model	11
3.2.1.2	Measurement Model	12
3.2.2	Observability Conditions	13
3.2.3	Algorithm Overview	13
4	Experiments	15
4.1	Hardware Setup	15
4.1.1	Small Vehicles for Autonomy (SVEA)	15
4.1.2	Intersection	16
4.2	Software Setup	17
4.2.1	Bearing-based Riccati localization	17
4.2.2	ArUco detection	18

4.2.3	Experimental Overview	18
4.2.4	Means of Evaluation	19
5	Results and Analysis	21
5.1	Evaluation Under Nominal Driving Conditions	21
5.2	Evaluation Under Stress-Tested Driving Conditions	23
5.3	Evaluation Under Degraded Sensing Conditions	25
6	Discussion	29
7	Conclusions and Future work	31
7.1	Conclusions	31
7.2	Future work	32
	References	33
A	Parameter Settings for each Experimental Scenario	37
A.1	Evaluation under Nominal Driving Conditions	37
A.2	Evaluation under Stress-Tested Driving Conditions	37
A.3	Evaluation under Degraded Sensing Conditions	38

Chapter 1

Introduction

1.1 Overview

Localization is a fundamental component of autonomous vehicle systems, enabling them to determine their position within a given environment accurately. A precise and robust localization framework underpins a wide range of autonomous driving functionalities, from basic motion planning to complex maneuvers in dynamic and uncertain traffic scenarios. Reliable pose estimation is essential for ensuring that autonomous vehicles can operate safely and efficiently without human intervention, thereby contributing to improved traffic safety, enhanced operational efficiency, and greater overall mobility.

To achieve accurate localization, autonomous vehicles typically use a combination of various sensors, including cameras, Light detection and ranging (LiDAR) [1, 2, 3], global navigation satellite system (GNSS) [4, 5], etc. Data from these sensors are often fused using various estimation and filtering algorithms, including various variations of Kalman Filter (KF) (Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF)) [6, 7], Simultaneous Localization and Mapping (SLAM) [8, 9, 10], and visual odometry techniques [11, 12]. Despite significant advances in sensor technology and localization algorithms, reliable localization remains challenging. Real-world environments are inherently dynamic and uncertain, and sensor measurements are subject to noise and drift. Such complexities hinder the accuracy and stability of localization, particularly in urban settings.

In relatively simpler scenarios, such as open highways or local streets, obtaining accurate localization is more straightforward. With predictable lane markings that provide visual cues, and stable GPS signals benefiting from minimal surrounding infrastructure, localization is relatively easier.

In contrast, intersections present a notably more challenging environment for localization, since vehicles approach from multiple directions, creating a dynamic and unpredictable environment. Moreover, in some cases, the dense urban infrastructure surrounding intersections creates an urban canyon that can degrade GPS signal accuracy and stability. Time is also a critical factor, as vehicles must rapidly interpret and react to complex interactions at intersections. Stopping mid-way would likely obstruct traffic and create hazardous situations.

The importance of intersections in urban transportation cannot be overstated. They serve as a vital node for connecting different routes, such as those linking rural areas to urban cities, and therefore improving the connectivity and accessibility within the city. Well-planned intersections also enhance traffic flow by offering alternatives, which is especially useful in alleviating congestion and improving resilience in case of accidents.

In this thesis, we focus on vehicle localization at intersections, under three conditions: (i) nominal driving under a typical configuration, (ii) aggressive driving near the vehicle's operational limits, and (iii) degraded sensing scenarios characterized by inaccurate or noisy measurements. The objective is to develop a lightweight and computationally efficient algorithm for localization while guaranteeing the convergence of vehicle position under certain conditions using bearing-based Riccati localization. The algorithm is validated using a 1/10th scale vehicle platform equipped with a camera, navigating a scaled intersection.

The contributions of this thesis are as follows:

- Proposed a lightweight with convergence-guaranteed bearing-based Riccati localization method for a vehicle at an intersection.
- Evaluated the performance of the localization method in real-world scenarios using a 1/10th scale vehicle platform.

1.2 Structure of the thesis

The thesis is structured as follows: Chapter 2 provides the theoretical foundations for bearing-based Riccati localization and includes a comprehensive literature review. Chapter 3 focuses on the explanation and derivation of bearing-based Riccati localization. Chapter 4 explains the details of the experimental setup. Chapter 5 presents the results of the experiments. Chapter 6 discusses the experimental results in depth. Chapter 7 summarizes the conclusion of the thesis and suggests directions for future work.

Chapter 2

Background

This chapter presents the theoretical foundation and context for bearing-based Riccati localization by introducing concepts including Perspective-n-Point (PnP) problem and observability etc. It ends with a summary of related work in localization, providing a basis for the experiments and analytical work presented in later chapters.

2.1 Perspective-n-Point (PnP) Problem

The Perspective- n -Point (PnP) problem involves estimating the pose, i.e., position and orientation of a calibrated camera given a set of n known 3D points in the environment and their corresponding 2D projections in the camera's image plane.

Let a set of n known 3D points in the inertial (world) frame be represented in homogeneous coordinates as $\mathcal{P}_i = [X_i \ Y_i \ Z_i \ 1]^\top$, and their corresponding 2D image point in homogeneous coordinates denoted by $p'_i = [u_i \ v_i \ 1]^\top$.

The objective is to estimate the camera pose in inertial (world) frame, consisting of a rotation matrix $R \in SO(3)$ and the translation vector $t \in \mathbb{R}^3$, which together define the transformation from the inertial frame to the camera frame.

By augmenting the extrinsic transformation into a single matrix $[R \ t] \in \mathbb{R}^{3 \times 4}$, the 3D-to-2D projection can be written as:

$$sp'_i = \mathbf{K} [R \ t] \mathcal{P}_i \quad (2.1)$$

, where s denotes the scaling factor, and \mathbf{K} represents the camera intrinsic

matrix, which encodes the intrinsic parameters of the camera, including focal lengths (f_x, f_y) , skew (γ) and principal points (u_0, v_0) .

The minimal case, known as the Perspective-Three-Point (P3P) problem, estimates the camera pose using only three known 3D-2D correspondences. This results in a system of algebraic equations with potentially multiple solutions. One of the earliest solutions to this problem was proposed by Grunert in 1841 [13], based on solving geometric constraints between rays and world points. However, this formulation may exhibit singularities under certain configurations, such as the "danger cylinder", which lead to unstable or ambiguous pose estimates.

When more than three correspondences are available (i.e., $(n \geq 4)$), the PnP problem can be solved using iterative optimization methods that minimize the reprojection error. Common algorithms include the Levenberg–Marquardt and Gauss–Newton methods [14]. Alternatively, non-iterative approaches such as EPnP and RPnP offer efficient closed-form solutions that scale better with the number of points.

2.2 Bearing Measurement

Bearing measurement refers to the direction of an object in space as inferred from 2D image data, specifically the direction from the camera's optical center to the object, irrespective of the object's distance. It captures directional information without providing depth. To compute this, the image coordinates are first converted into normalized image coordinates by removing the effect of the camera's intrinsic parameters. The normalized coordinates, which place the origin at the camera's principal point, are given by [15]:

$$u_n = \frac{(u_i - u_0)}{f_x} \quad (2.2)$$

$$v_n = \frac{(v_i - v_0)}{f_y} \quad (2.3)$$

The direction vector can be defined as $p' = [u_n \ v_n \ 1]^T$. The unit direction vector, used for bearing estimation, can be calculated as:

$$g' = \frac{p'}{\|p'\|} \quad (2.4)$$

2.3 Observability

Observability, a concept related to control theory, determines whether the internal states of a system can be inferred from its inputs and outputs [16]. It plays a crucial role in enabling accurate state estimation, which is essential for reliable operation and control.

Given the definition of a Linear Time-Varying (LTV) system as

$$\begin{cases} \dot{x} &= A(t)x + B(t)u \\ y &= C(t)x \end{cases} \quad (2.5)$$

where x denotes the system state, u denotes the control inputs, y denotes the measured output, and $A(t), B(t), C(t)$ denotes the time-varying system matrices.

The system is considered observable if, for any unknown initial state $x(0)$, there exists a finite time interval $t_l > 0$ such that the knowledge of the input u and the output y over $[0, t_l]$ suffices to determine uniquely the initial state $x(0)$ [16].

In the context of vehicle localization, observability guarantees that the vehicle's pose (position and orientation) can be deduced from the sensor data, such as camera images, and velocity measurements. Without sufficient observability, the localization system would not be able to reliably estimate the states of the vehicle, leading to navigation and control errors.

2.3.1 Uniform Observability

Uniform observability is a stronger claim of observability, meaning that the state of the system can be consistently estimated from the outputs over a given time interval. Specifically, a system is uniformly observable if the state $x(t)$ can be deduced from the information given by input $u(\cdot)$ and output $y(t)$ on the time interval $[t, t + \delta]$ [16].

2.3.2 Observability Grammian

Observability Grammian is a means to determine if the system is observable, which is defined as follows [16]:

$$W_Q^{A,C}(t, t + \delta) := \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(s, t) C^\top(s) Q(s) C(s) \Phi(s, t) ds \quad (2.6)$$

, where $\Phi(s, t)$ defined as the transition matrix of the system associated with $A(t)$.

If $A(t)$ and $C(t)$ are bounded and there exists a time interval $\delta > 0$ and such that $W_Q^{A,C}(t, t + \delta)$ is positive definite, then we can say that the system is uniformly observable.

Uniform observability is critical in time-varying estimation problems, where persistent and consistent reconstruction of the system state from sensor measurements is required to ensure stable and accurate navigation performance.

2.4 Continuous Riccati Equation

The Continuous Riccati Equation (CRE) is a fundamental tool in control theory that plays a crucial role in stability analysis, observer design, and optimal control.

Based on the linear time-varying (LTV) system described previously, CRE is expressed as [17]:

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)Q(t)C(t)P(t) + V(t) \quad (2.7)$$

where $P(t)$ is a symmetric positive definite matrix, $Q(t)$ is a non-negative semidefinite matrix, that is usually associated with the matrix that describes measurement noise, and $V(t)$ is a non-negative semidefinite matrix that is usually associated with the matrix that describes process noise.

The initial condition $P(0)$ should be symmetric and positive definite. The solution $P(t)$ from CRE is used in the calculation for the gain K , which minimizes the state estimation error:

$$K = kP(t)C^T(t)Q(t) \quad (2.8)$$

, where $k \geq 0.5$ is a bounded gain.

2.5 Graph Theory

Graph theory is a mathematical framework for representing the relationships and connections between components in a system. A graph is described as $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, consists of a set of vertices $\mathcal{V} = \{1, 2, \dots, n\}$ and a set of edges \mathcal{E} connecting pairs of vertices [18].

Graphs can be directed or undirected. An undirected graph has edges with no direction, representing a mutual relationship between vertices. A directed graph (digraph) has edges with specific directions, represented as ordered pairs of vertices, and it is defined as $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \dots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Each vertex i has a set of neighbors defined as $\mathcal{N}_i := \{j \in \mathcal{V} | i, j \in \mathcal{E}\}$.

A directed acyclic graph (DAG) is a type of directed graph that contains no cycles, meaning the path cannot start and end at the same vertex.

In this thesis, we consider vehicles under a directed leader-follower structure with $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ that is acyclic. \mathcal{V} consists of two types of vertices, with \mathcal{V}_f representing the vehicle vertex and \mathcal{V}_l representing the landmark vertex, i.e., $\mathcal{V} = \mathcal{V}_f \cup \mathcal{V}_l$. We assume that $\mathcal{V}_l = \{1, \dots, n_l\}$ and $\mathcal{V}_f = \{n_l + 1, \dots, n_f\}$, with $n = n_l + n_f$. For the landmark vertex, they have no neighbors, i.e., $\mathcal{N}_i = \emptyset$, $i \in \mathcal{V}_l$. For the vehicle vertex $i \in \mathcal{V}_f$, $\mathcal{N}_f \subseteq \{1, \dots, i - 1\}$.

2.6 Related Work

This section focuses on exploring localization methods that utilize the Perspective-n-Point (PnP) problem, visual-SLAM systems like ORB-SLAM, visual odometry, and those specialized methods at intersections.

The Perspective-n-Point (PnP) problem is a foundation of visual SLAM systems. PnP methods provide an efficient way to compute the position and orientation of the camera, relative to the known 3D points. One example of a visual SLAM system is ORB-SLAM [19, 20]. They first extract the feature points of the current frame using ORB, which is a fusion of FAST keypoint detector and BRIEF descriptor. The feature points are then compared and matched with the other frames to find correspondences. The pose of the camera is estimated using these matched points. RANSAC is employed to refine the pose estimation and handle outliers. PnP is also crucial for recovering the camera pose when tracking is lost, which happens in dynamic environments or during rapid camera motion.

While ORB-SLAM could provide robust pose estimation, they could be computationally demanding, limiting the performance of other crucial functionalities of the vehicles, including obstacle detection and avoidance.

Other than visual SLAM, visual odometry has become a widely used tool for motion estimation. [21] proposed a visual odometry system using monocular and stereo camera setups. For the monocular scheme, they make use of features in the frame and triangulate them into 3D points, before computing the pose of the camera using P3P and RANSAC for iterative

refinement. [12] proposed a visual odometry system specifically designed for car-like vehicles with Ackermann steering. A downward-facing camera captures ground features and calculates pixel displacement to estimate the vehicle's motion. [22] enhanced visual odometry by incorporating a Kalman filter to refine motion estimates based on feature tracking on the ground plane. Visual odometry methods are cost-effective, however, they could suffer from drift over time due to error accumulation, leading to long-term inaccuracies in localization.

Researchers have also looked into the issue of localization at intersections due to their complexity and crucial role in traffic environments. [23] proposed a method using traffic lights as fixed visual landmarks. The positions of traffic lights are stored in a high-precision map and are compared with the camera detection. This information is fused with the vehicle inertial navigation system (INS) and GPS measurements using an extended Kalman filter (EKF). This approach leverages static infrastructure to enhance localization accuracy. [24] uses a particle filter to estimate the vehicle's position and heading, with the update from image data and GPS. Unlike traditional methods that extract and match features, this approach predicts the full image of the scene and evaluates the coherency value derived from the structure tensor.

Chapter 3

Method

This chapter focuses on the mathematical formulation and derivation of the Riccati observer for vehicle localization. The experiments and implementation details will be addressed later in Chapter 4.

We begin with an overview of the problem settings and a description of the scenario. We then move on to the derivation of the Riccati observer.

3.1 Scenarios and System Overview

As mentioned previously, intersections present unique challenges for localization due to their dynamic and unpredictable natures, which causes the complexity in providing accurate pose estimation. The primary objective of this thesis is thus to develop a lightweight method capable of providing precise, real-time localization at an intersection, with guaranteed convergence to the actual vehicle pose.

The detailed flow of the system is shown in Figure 3.1. It starts with data acquisition from onboard sensors (actuators and a camera) and external data (HD map). With the velocity measurements estimated from the control signal issued to actuators and the pose estimation from the previous timestamp, the dynamic model predicts the pose (position and orientation) of the vehicle. Simultaneously, the camera image is processed to detect landmark bearings. This information, along with the landmark positions from the HD map and the prediction from the dynamic model, is fed into the measurement model to refine the pose estimation. The updated pose estimation is then iteratively fed back into the dynamic model for continuous real-time localization.

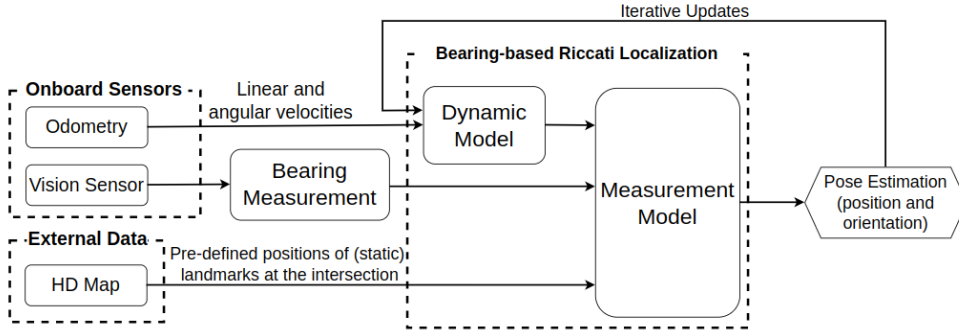


Figure 3.1: Overview of the system of the proposed bearing-based Riccati localization method

3.2 Bearing-based Riccati Localization

The bearing-based Riccati localization method is based on the Riccati observer proposed by Tarek Hamel and Claude Samson [25]. This approach addresses the non-stationary PnP problem, leveraging bearing measurements from known landmark positions and velocity measurements to estimate the position and orientation of a vehicle. Unlike the static PnP problem, this method uses vehicle motion as an additional source of information to enhance pose estimation.

In this scenario, we consider a vehicle at the intersection that is estimating its position and orientation, and n_l number of landmarks. We assume that the vehicle is equipped with an onboard camera and odometry to detect landmarks and measure its velocities.

We denote $\bar{p} \in \mathbb{R}^3$ as the position of the vehicle $i \in \mathcal{V}_f$, expressed in its body frame \mathcal{B} , with dynamics

$$\dot{\bar{p}} = -S(\omega)\bar{p} + \bar{v} \quad (3.1)$$

, where $\bar{v} \in \mathbb{R}^3$ is the linear velocity of the vehicle i expressed in \mathcal{B} . The body frame \mathcal{B} of the vehicle i with respect to the inertial frame \mathcal{F} is represented by the orientation matrix $R \in SO(3)$ with the following kinematic relations:

$$\dot{R} = RS(\omega) \quad (3.2)$$

, where ω is the angular velocity of vehicle i expressed in \mathcal{B} . We define the bearing vector of objects from the vehicle, measured from the vehicle's

onboard camera, expressed in the vehicle frame, as:

$$g_j := \frac{R^\top(p - z_j)}{|p - z_j|} \quad (3.3)$$

, where p is the position of the vehicle i in \mathcal{F} , and $z_j, j \in \mathcal{N}_i$ as the position of the landmarks expressed in \mathcal{F} .

We assume that there are at least three landmarks in the system with known positions $z_j, j \in \mathcal{V}_i$ in \mathcal{F} . To obtain the positions of landmarks, we assume that there are Intelligent Transport Systems (ITS) to facilitate vehicle-to-infrastructure (V2I) communication, which could provide the information in terms of a high-definition (HD) map, when the vehicle enters the intersection.

The vehicle i would measure the relative bearing g_j to its neighbor $j \in \mathcal{N}_i$, and is assumed to be well-defined $\forall t \geq 0$, i.e., the vehicle and landmarks are not collocated. With the onboard odometry, the vehicle can measure its own linear v and angular ω velocities. The velocity measurements are assumed to be bounded.

With the above setup and assumptions, we could design a bearing-based Riccati Localization approach to estimate the vehicle $i, i \in \mathcal{V}_i$'s pose (R, \bar{p}) .

3.2.1 Mathematical Models

3.2.1.1 Dynamic Model

Let (\hat{R}, \hat{p}) be the estimate of the vehicle pose (R, \bar{p}) . For the vehicle $i, i \in \mathcal{V}_f$, define the dynamics of (\hat{R}, \hat{p}) as,

$$\begin{aligned} \dot{\hat{R}} &= \hat{R}S(\hat{\omega}) \\ \dot{\hat{p}} &= -S(\omega)\hat{p} + \hat{v} \end{aligned} \quad (3.4)$$

, where $\hat{\omega}$ and \hat{v} are estimated angular and linear velocities to be designed later.

We define the orientation errors as $\tilde{R} := \hat{R}^\top R$. The unit quaternion associated with \tilde{R} is defined as $\tilde{\Lambda} := (\bar{\lambda}_0, \bar{\lambda})$ with $\bar{\lambda} \in \mathbb{R}^3$ denote the vector part of the quaternion, and $\bar{\lambda}_0 \in \mathbb{R}^1$ denote the scalar part of the quaternion. The corresponding Rodrigues formula relating $\tilde{\Lambda}$ to \tilde{R} is

$$\tilde{R} = I_3 + 2S(\bar{\lambda})(\bar{\lambda}_0 I_3 + S(\bar{\lambda})) \quad (3.5)$$

Define $\tilde{\omega} := \omega - \tilde{R}^\top \hat{\omega}$, we can deduce the following kinematic relations

$$\dot{\tilde{R}} = \tilde{R}S(\tilde{\omega}) \quad ; \quad \begin{cases} \dot{\bar{\lambda}}_0 = -0.5(\tilde{\omega})^\top \bar{\lambda} \\ \dot{\bar{\lambda}} = 0.5\tilde{\omega}\bar{\lambda}_0 + 0.5S(\tilde{\omega})\bar{\lambda} \end{cases}$$

and using the fact that $\bar{\lambda}_0 = 1 - |\bar{\lambda}|^2$, we have

$$\begin{cases} \tilde{R} = I_3 + 2S(\bar{\lambda}) + O(|\bar{\lambda}|^2) \\ 2\dot{\bar{\lambda}} = \tilde{\omega} - S(\omega)(2\bar{\lambda}) + O(|\bar{\lambda}||\tilde{\omega}|) + O(|\omega||\bar{\lambda}|^2) \end{cases} \quad (3.6)$$

Since $|\omega(t)|$ is bounded by assumption, we have

$$2\dot{\bar{\lambda}} = -S(\omega)(2\bar{\lambda}) + \tilde{\omega} + O(|\bar{\lambda}||\tilde{\omega}|) + O(|\bar{\lambda}|^2)$$

Under this parameterization, we can model the dynamics of the pose $(2\bar{\lambda}, \bar{p})$ as

$$\begin{cases} 2\dot{\bar{\lambda}} = -S(\omega)(2\bar{\lambda}) + \tilde{\omega} + O(|\bar{\lambda}||\tilde{\omega}|) + O(|\omega||\bar{\lambda}|^2) \\ \dot{\bar{p}} = -S(\omega)\bar{p} + \bar{v} \end{cases} \quad (3.7)$$

, where $S(x)$ denotes a skew-symmetric matrix associated with the cross-product in \mathbb{R}^3 .

3.2.1.2 Measurement Model

The measurement model considers the observation, i.e., the bearing measurements of the known landmarks detected by the vehicle's onboard camera.

The output function for each landmark, $y_j(\bar{\lambda}, \bar{p}, t)$ for $j \in \{1, 2, \dots, l\}$, is written as:

$$y_j(\bar{\lambda}, \bar{p}, t) = -\Pi_{g_j(t)}S(\hat{R}^\top z_i)(2\bar{\lambda}) + \Pi_{g_j(t)}\bar{p} + O(|\lambda|^2) \quad (3.8)$$

, where Π_x denotes an orthogonal projection operator in \mathbb{R}^3 that projects vectors onto the 2D subspace orthogonal to x and is defined as $\Pi_x := I_3 - xx^\top$. By defining $y := [y_1^\top, y_2^\top, \dots, y_l^\top]^\top$, we can obtain the output equation with

$C = [C_1 \ C_2]$:

$$C_1(x_1, \hat{x}_2, t) = \begin{bmatrix} -\Pi_{g_1(t)} S(\hat{R}^\top z_1) \\ \vdots \\ -\Pi_{g_l(t)} S(\hat{R}^\top z_l) \end{bmatrix} \quad C_2(x_1, t) = \begin{bmatrix} \Pi_{g_1(t)} \\ \vdots \\ \Pi_{g_l(t)} \end{bmatrix} \quad (3.9)$$

, $\{1, \dots, l\} \in \mathcal{N}_i$.

With the gain $K = kP(t)C^\top(t)Q(t)$ defined in equation (2.8), with $K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$ and $Q(t) = \text{diag}\{q_1 I_3, \dots, q_l I_3\}$. We design the estimated angular and linear velocities as

$$\begin{cases} \hat{\omega} = \omega + K_1(y - C_2 \hat{p}) \\ \hat{v} = \bar{v} + K_2(y - C_2 \hat{p}) \end{cases} \quad (3.10)$$

, where $(\hat{\omega}, \hat{v})$ is used to compute the estimated pose (\hat{R}, \hat{p}) from equation (3.4) via numerical integration.

3.2.2 Observability Conditions

The detailed proof can be found in the paper "Riccati Observers for the non-stationary PnP problem" by Tarek Hamel and Claude Samson [25].

To ensure uniform observability, a vehicle should, in general, maintain at least three non-collinear connections, such as detecting three landmarks distributed around the environment. Under these conditions, the danger cylinder configurations could arise. It is defined as the circle passing through the three landmarks, with an axis orthogonal to the plane containing the landmarks. Uniform observability is lost if the vehicle remains static on the danger cylinder, or when it drives along one of the three landmarks' axes that belong to the cylinder.

When more than three non-collinear landmarks are available, uniform observability is compromised if all landmarks and the vehicle lie on a circle, or when all landmarks are distributed on the circle and a straight line forming a degenerate horopter originating at the vehicle's pose.

3.2.3 Algorithm Overview

The proposed algorithm, detailed in Algorithm 1, builds upon the previously described models.

Algorithm 1 Bearing-based Riccati localization algorithm

```

1: while TRUE do
2:   Step 1: Prediction using Dynamics Model
3:    $\hat{\omega} \leftarrow \omega$ 
4:    $\hat{\dot{p}} \leftarrow -S(\omega)\hat{p} + \bar{v}$ 
5:   Step 2: Correction based on Observer
6:    $correction_{\hat{\omega}} \leftarrow 0$ 
7:    $correction_{\hat{\dot{p}}} \leftarrow 0$ 
8:   for each landmark  $z_i$  in  $z$  do
9:      $correction_{\hat{\omega}} \leftarrow correction_{\hat{\omega}} + q_i S(\hat{R}^\top z_i) \Pi_{g_i}(\hat{p} - \hat{R}^\top z_i)$ 
10:     $correction_{\hat{\dot{p}}} \leftarrow correction_{\hat{\dot{p}}} + q_i \Pi_{g_i}(\hat{p} - \hat{R}^\top z_i)$ 
11:   end for
12:    $correction_{\hat{\omega}} \leftarrow -kP \times correction_{\hat{\omega}}$ 
13:    $correction_{\hat{\dot{p}}} \leftarrow -kP \times correction_{\hat{\dot{p}}}$ 
14:   Step 3: Update Position and Orientation Estimation
15:    $\hat{\omega} \leftarrow \hat{\omega} + correction_{\hat{\omega}}$ 
16:    $\hat{\dot{p}} \leftarrow \hat{\dot{p}} + correction_{\hat{\dot{p}}}$ 
17:    $\hat{R} \leftarrow \hat{R}S(\hat{\omega})$ 
18:    $\dot{P} \leftarrow AP + PA^\top - PC^\top QCP + V$ 
19:    $\hat{R} \leftarrow$  compute  $\hat{R}$  using numerical integration of  $\dot{R}$ 
20:    $\hat{p} \leftarrow$  compute  $\hat{p}$  using numerical integration of  $\dot{p}$ 
21: end while

```

Chapter 4

Experiments

We conducted experiments for using a 1/10th-scale vehicle at a to-scale intersection. The experiments aim to evaluate the performance of Riccati localization in real-world environments under various conditions.

4.1 Hardware Setup

4.1.1 Small Vehicles for Autonomy (SVEA)

SVEA, in Figure 4.1, is a 1/10th-scale vehicle platform that had been previously developed at the Smart Mobility Lab (SML) [26]. It consists of an onboard NVIDIA Jetson TX2 computer with various onboard sensors, including an inertial measurement unit (IMU), a real-time kinematic (RTK) GPS, a camera, etc.

In the experiments, a Logitech 720p webcam is used as the primary vision sensor. Prior to data collection, intrinsic camera calibration is performed to ensure accurate bearing measurements. Details on the landmark detection process are provided later in this section. The vehicle's linear and angular velocities are estimated based on the control commands sent to the actuators, which are issued and recorded at a frequency of 50 Hz.

The SVEA platform was selected for experimentation due to its compact scale and sensor configuration, which collectively provide a representative proxy for a full-scale autonomous vehicle, without incurring the risks and costs associated with full-scale testing. Its onboard computing capabilities and full compatibility with the Robot Operating System (ROS), which is used to implement the proposed method, make it a suitable and practical testbed for evaluating the proposed Riccati observer in realistic intersection scenarios.



Figure 4.1: Small Vehicles for Autonomy (SVEA)

4.1.2 Intersection

A four-way intersection at the same scale as the SVEA, with one lane in each direction, was constructed at Kista Innovation Park (KIP), as shown in Figure 4.2. For easier visualization, some lane markings are added. A total of 32 10cm by 10cm ArUco markers are placed at the intersection as landmarks, with 16 inside the intersection and the rest along the edge of the intersection. To ensure observability is maintained effectively, the ArUco markers are double-sided, enabling detection from more angles. Additionally, the markers are positioned to avoid collinearity and ensure the vehicle can reliably detect a sufficient number of markers.



Figure 4.2: Setup of the Intersection

ArUco markers are chosen as landmarks due to their ease of detection and reliability. Each marker is embedded with a unique pattern, making it easier to identify with high accuracy and associate with the landmark position in the world frame. Additionally, the implementation is fairly simple thanks to the support from an open-source library, which reduces the complexity of the overall setup and allows us to focus on the implementation of Riccati localization. Lastly, ArUco markers provide precise measurements for both 2D and 3D poses, especially at close range, enabling the calculation of bearing measurement from the 2D poses and serving as a supplementary way to evaluate the performance of Riccati localization using 3D poses.

4.2 Software Setup

The software implementation is in Robot Operating System (ROS), which includes the bearing-based Riccati localization algorithm, sensors' measurements (camera, actuator, and RTK-GPS), and ArUco detection.

4.2.1 Bearing-based Riccati localization

The bearing-based Riccati localization algorithm has a list of configurable parameters:

Parameter	Description
$\hat{p}(0)$	Initial position estimation, $[x \ y \ z]^T$
$\bar{\Lambda}(0)$	Initial orientation estimation, $[qw \ qx \ qy \ qz]^T$
k	Gain
Q	Coefficient, $Q = \text{diag}\{q_1 I_3, \dots, q_t I_3\}$
V	Coefficient, $V = \text{diag}\{v_1 I_3, v_2 I_3\}$
$P(0)$	Initial Riccati Solution, i.e. $P(0) = \text{diag}\{P_1 I_3, P_2 I_3\}$

Table 4.1: Configurable parameters for bearing-based Riccati localization

The vehicle's onboard computer executes the bearing-based Riccati localization algorithm at 50 Hz. The algorithm first estimates the pose of the onboard camera, which is then transformed into the vehicle's pose using a known static transformation between the camera frame and the vehicle frame. For ease of interpretation and visualization, the vehicle's position \bar{p} , originally expressed in the body frame \mathcal{B} , is subsequently transformed into the inertial frame \mathcal{F} .

As discussed in Chapter 3, numerical integration is required to determine \hat{R} and \hat{p} . The Runge-Kutta 45 method (RK45) is selected due to its adaptive step-size control and error estimation during each step, which are essential for the algorithm to run more efficiently and maintain a decent level of accuracy.

4.2.2 ArUco detection

The OpenCV library is used to detect 4×4 ArUco markers at a frequency of 30 Hz. Since the detection results are published at a lower rate, we employ a zero-order hold (ZOH) scheme to maintain temporal consistency by holding the most recent detection result constant until a new measurement becomes available.

The OpenCV implementation provides both 3D positions, i.e., the distance between the ArUco marker and the camera, and the 2D position, i.e., the pixel position of the ArUco marker on the image plane during detection. The former is only used as a supplementary tool for evaluation, without being used in the algorithm, and the latter is then transformed into a bearing representation, using it as an input to the algorithm.

4.2.3 Experimental Overview

At the start of the experiment, the SVEA vehicle is randomly positioned at one of the intersection entrances, with only approximate knowledge of its initial location. The vehicle then proceeds forward toward the center of the intersection, where it either makes a turn or continues straight until it exits the intersection. This setup represents a typical intersection scenario, where the vehicle must navigate through uncertain initial conditions and make a directional decision based on its internal planning or control strategy. During this process, the vehicle runs both the RTK-GPS and the bearing-based Riccati localization algorithms concurrently, providing data for comparison purposes.

To comprehensively evaluate the performance of the bearing-based Riccati localization method, we consider three distinct operating conditions:

- (i) **Nominal driving under a typical configuration:** This serves as a baseline case that reflects typical operating conditions where sensor data is relatively clean and vehicle motion is within moderate bounds. It is used to validate the core accuracy and convergence behavior of the observer under standard circumstances.
- (ii) **Stress-tested driving near the operational limits:** This condition represents when the vehicle has quicker accelerations. This tests the

observer’s ability to maintain stability and accurate state estimation under fast-changing dynamics.

(iii) Degraded sensing conditions: This condition simulates real-world challenges such as sensor noises and reduced landmark visibility. Evaluating under such conditions is essential to demonstrate the observer’s robustness to imperfect measurements, which are unavoidable in partial deployments.

Due to computational limitations on the SVEA platform, all raw data were recorded during the experiment, and the bearing-based Riccati localization algorithm was executed offline for subsequent analysis.

4.2.4 Means of Evaluation

In order to evaluate the performance of bearing-based Riccati localization in an outdoor environment, where establishing an accurate ground truth is challenging, we decided to compare the performance against RTK-GPS. RTK-GPS is a satellite navigation system with real-time corrections from base stations that can achieve centimeter-level accuracy, which is significantly more accurate than standard GPS [27]. Given this level of precision, RTK-GPS serves as a practical benchmark for assessing the accuracy and reliability of the proposed localization approach.

However, it is important to acknowledge the limitations of RTK-GPS. Its performance can degrade in the presence of interference, adverse weather conditions, or restricted satellite visibility. In dense urban environments, the quality and stability of RTK-GPS could be compromised, as tall buildings and other infrastructure may obstruct line-of-sight to satellites or introduce multipath effects. Additionally, the accuracy of RTK-GPS is sensitive to satellite geometry and availability; poor configurations can significantly reduce its effectiveness.

Due to these limitations, a direct comparison between the performance of the RTK-GPS and bearing-based Riccati localization using numerical accuracy might not be entirely meaningful. As a complementary validation approach, we use captured camera images to assess localization accuracy through the visibility and placement of ArUco markers in the scene. This image-based verification provides qualitative insights into the accuracy of Riccati localization, particularly in situations where RTK-GPS data may be degraded or unreliable.

Chapter 5

Results and Analysis

This chapter presents the results of the experiments under various conditions, based on the setup described in Chapter 4. Each test case has a different parameter setting and is shown in Appendix A. The intersection setup is identical for all the test cases.

This chapter presents the results of the experiments conducted under a range of operating conditions, as introduced in Chapter 4. While the physical intersection setup remains consistent across all experiments to ensure a controlled evaluation environment, each test case is defined by a distinct set of parameter settings for the proposed localization approach. These parameter settings are summarized in Appendix A.

5.1 Evaluation Under Nominal Driving Conditions

This test case is designed to evaluate the performance of the bearing-based Riccati localization method under nominal operating conditions, where the vehicle is driven at approximately 50% of its maximum speed and does not approach its operational limits, providing a controlled setting for baseline analysis.

Table A.1 indicates the parameter setting for this test case, with both the initial pose and orientation estimates intentionally inaccurate to assess robustness.

Figure 5.1 presents the trajectories of the vehicle, as recorded by RTK-GPS, and the pose (position and orientation) estimated by bearing-based Riccati localization. The vehicle starts from the right entrance at roughly

(3.0, 0.5), turning left near the center before exiting the intersection at roughly $(-0.5, -2)$. A high degree of alignment between the two trajectories is observed, particularly during the turning phase. After completing the turn, however, a noticeable divergence occurs between the two estimated paths.

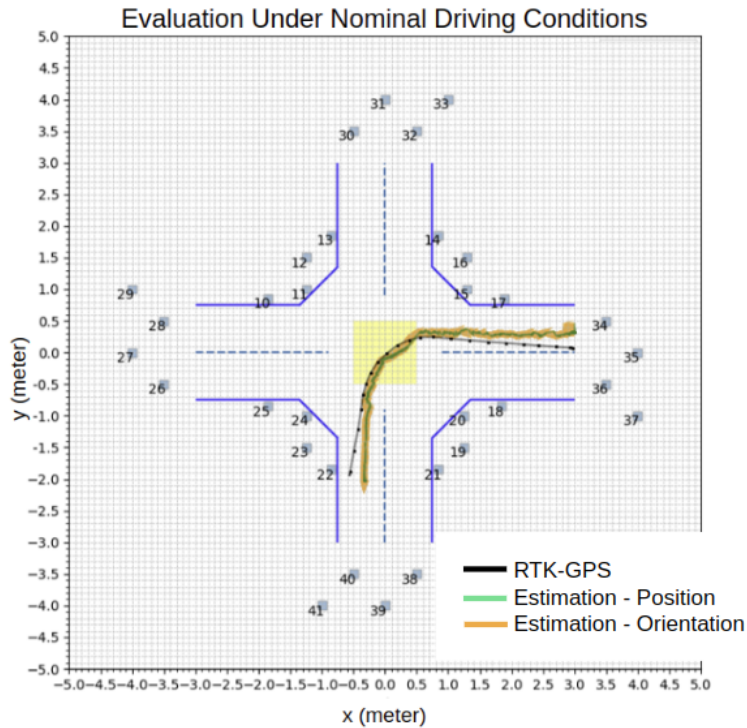


Figure 5.1: Vehicle trajectory comparison under the normal driving condition scenario: Estimation from bearing-based Riccati localization (green-orange) vs. RTK-GPS (black), with the square indicating the position of ArUco markers with IDs labelled

This divergence may stem from known limitations of RTK-GPS, such as multipath effects or transient signal degradation, rather than a failure in the proposed method. To investigate this, Figure 5.2, shows the images captured by the SVEA's onboard camera after the vehicle immediately after the turn, providing qualitative evidence to support bearing-based Riccati localization estimation. The vehicle's orientation and position align with the midpoint between ArUco markers #39 and #40. This visual evidence closely matches the pose estimated by the Riccati observer, suggesting that the deviation from the RTK-GPS trajectory is more likely due to inaccuracies in the GPS signal.

As mentioned previously, while RTK-GPS serves as a useful reference, it is important to acknowledge its limitations in outdoor environments where

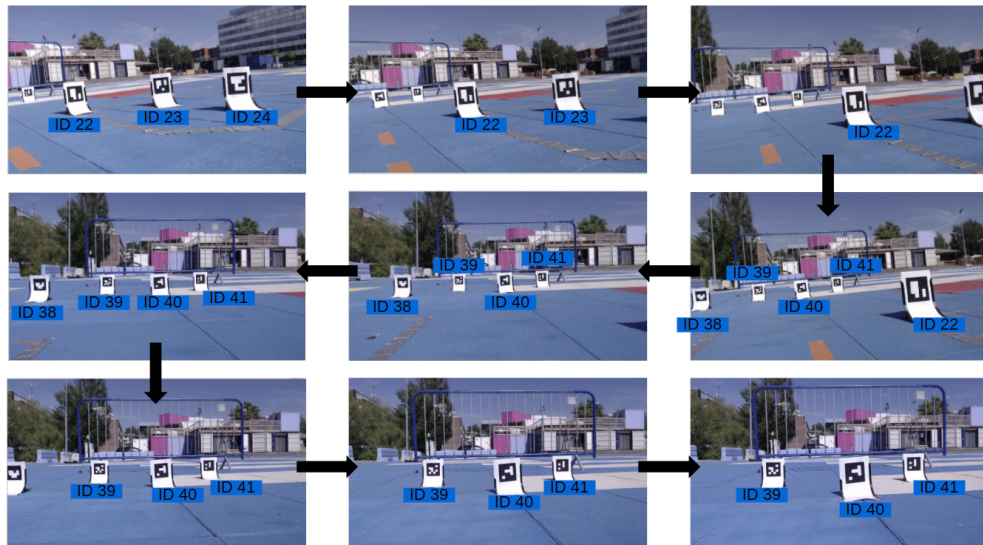


Figure 5.2: Sequential images captured by the SVEA's onboard camera after completing the turn in the nominal driving condition scenario, showing the visibility of ArUco markers

satellite visibility may be restricted or signal interference is present. In contrast, the Riccati localization method relies on onboard sensors, giving it a distinct advantage in scenarios where external localization aids are unreliable or unavailable.

These results highlight the robustness of the Riccati observer in maintaining accurate pose estimates through dynamic maneuvers and sensor imperfections. The combination of quantitative trajectory agreement and qualitative visual alignment supports its potential as an effective alternative solution to provide accurate localization, particularly in intersection scenarios where precision is critical and infrastructure uncertainty is high.

5.2 Evaluation Under Stress-Tested Driving Conditions

In this test case, we evaluate the performance of the Riccati localization method when the system operates near its dynamic limits, specifically by driving the vehicle at higher speeds.

Table A.2 indicates the parameter setting for this test case. Both the initial pose estimate and the initial orientation estimate are only rough estimations.

Figure 5.3 presents the vehicle's trajectories as recorded by RTK-GPS and

estimated by the Riccati localization method under higher-speed conditions. Compared to the nominal-speed case, more pronounced deviations are observed between the two trajectories, particularly during the turning phase.

Several factors contribute to this increased discrepancy. First, higher speeds reduce the effective sampling density of the onboard sensors. Since the vehicle covers more distance between consecutive sensor readings, the temporal resolution of both the motion and visual measurements decreases, which negatively affects the performance of the observer, especially during rapid maneuvers such as turns. Second, increased speed can lead to higher levels of sensor noise and motion blur, making ArUco marker detection less reliable. This results in fewer valid landmark observations per estimation cycle and, consequently, reduced system observability. The combination of degraded measurement quality and reduced observability delays the convergence of the pose estimate, leading to larger deviations from the reference trajectory.

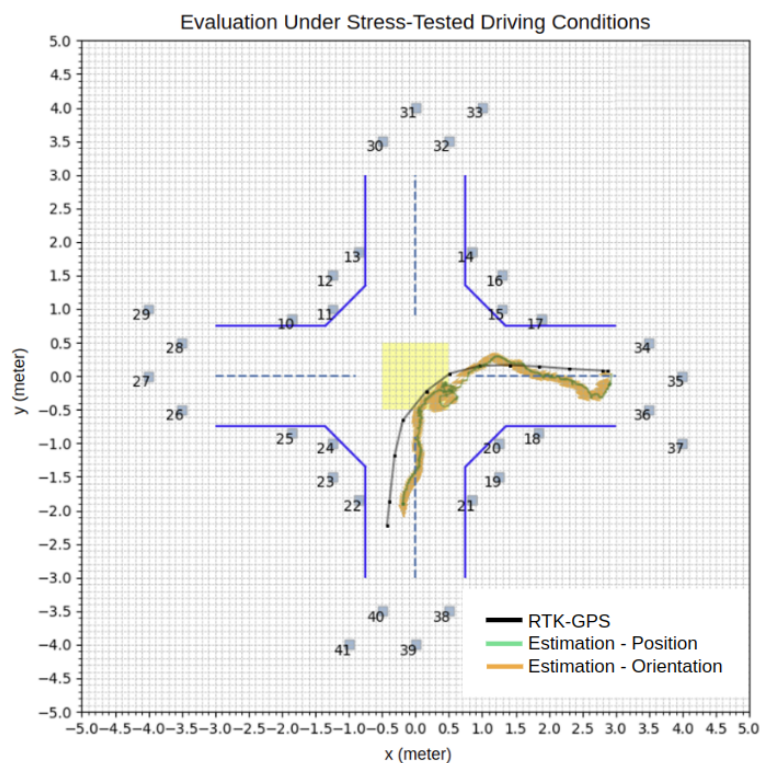


Figure 5.3: Vehicle trajectory comparison under the stress-tested driving condition: bearing-based Riccati localization (green-orange) vs. RTK-GPS (black), with ArUco marker IDs labeled on the figure

While the Riccati localization method demonstrates robustness under moderate-speed conditions, its performance degrades during rapid maneuvers, indicating a need for further improvement. This limitation is due to reduced sensor sampling density and degraded visual input at higher speeds. However, several strategies can be employed to mitigate these challenges. For instance, using a camera with a higher frame rate can improve the temporal resolution of visual measurements, allowing for more reliable landmark detection during fast motion. Additionally, integrating multiple cameras with overlapping fields of view can increase the likelihood of consistently detecting clear and trackable landmarks, thereby enhancing system observability and improving pose estimation accuracy under high-speed conditions.

5.3 Evaluation Under Degraded Sensing Conditions

The final test case evaluates the performance of the Riccati localization method in the presence of degraded sensor reliability. Specifically, the experiment introduces unexpected inaccuracies in angular velocity measurements to assess the observer's robustness when subjected to sensor bias or noise, which can impact estimation accuracy, especially during critical maneuvers.

Table A.3 indicates the parameter setting for this test case. Both the initial pose and orientation estimates are deliberately rough.

Figure 5.4 shows the trajectory estimated using Riccati localization. Figure 5.5 presents the linear and angular velocity readings from SVEA's onboard sensors, while Figure 5.6 provides synchronized camera images to validate the ground context during sensor anomalies.

By comparing the velocity plot and the camera images, a discrepancy is evident during the early phase of the experiment, where the vehicle is moving forward in a straight line. The angular velocity measurements exhibit an unexpected offset of approximately 0.3 rad/s, despite the vehicle driving in a straight line, as confirmed by visual evidence. This suggests a sensor bias or error.

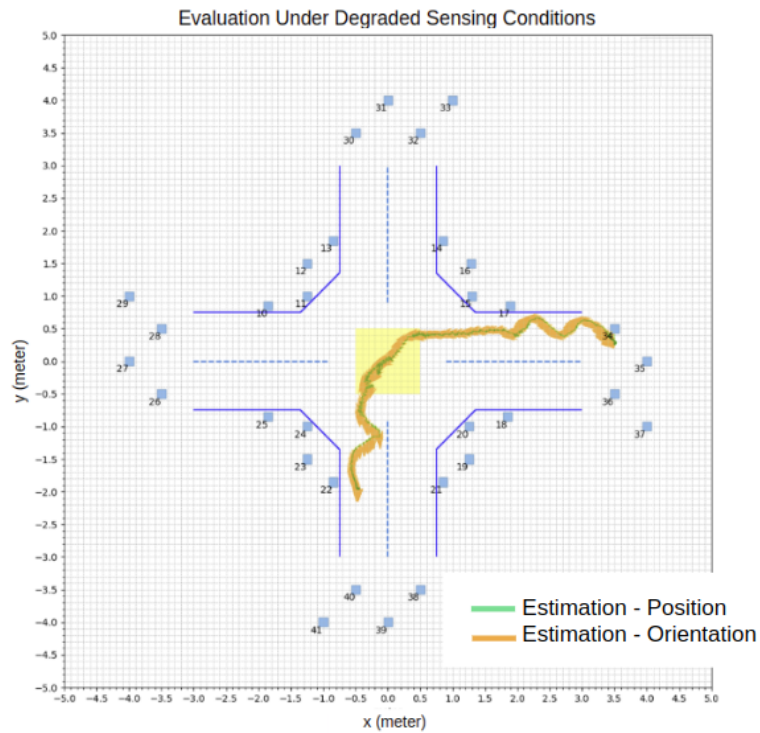


Figure 5.4: Vehicle trajectory under the degraded sensing condition scenario: bearing-based Riccati localization (green-orange), with ArUco marker IDs labeled on the figure

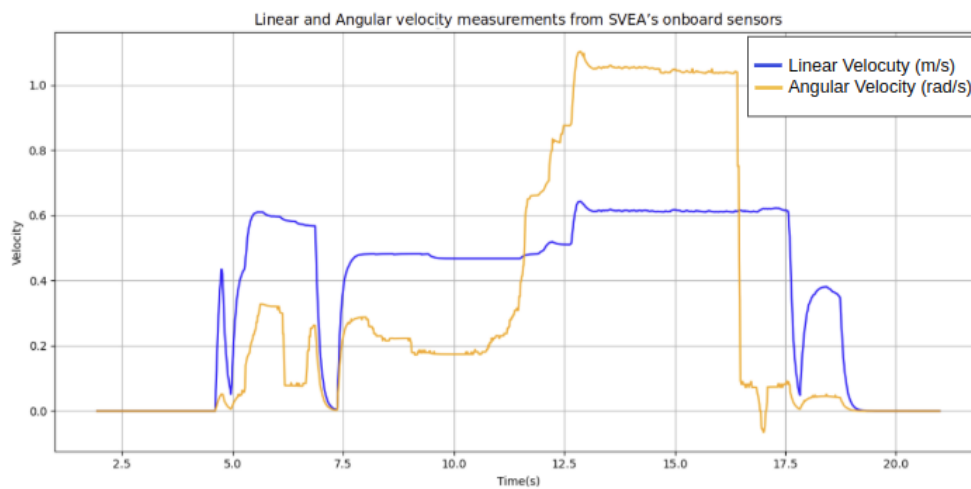


Figure 5.5: Linear and angular velocity measurements from SVEA's onboard sensors under the degraded sensing condition scenario

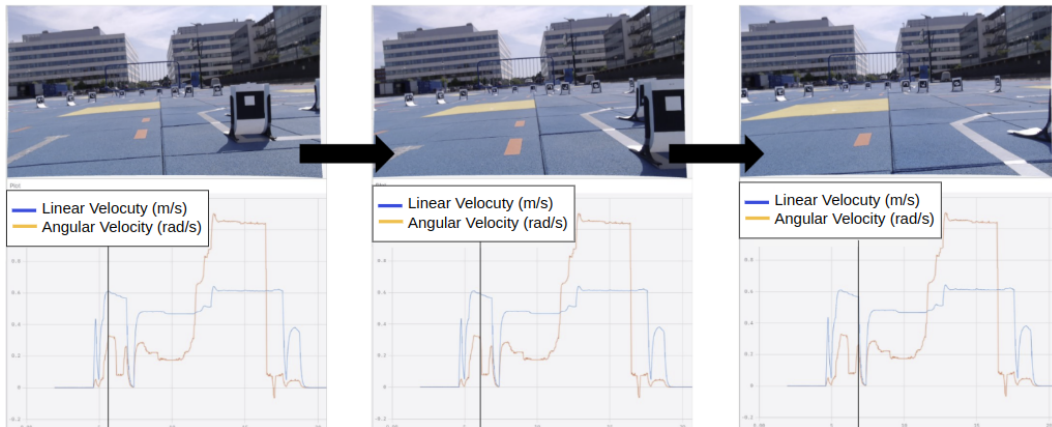


Figure 5.6: Images captured by the SVEA's onboard camera, highlighting the periods of unexpected inaccuracies in angular velocity measurements, as indicated by the black line, under the degraded sensing condition scenario

Despite this significant error in the early stage, the Riccati localization demonstrates a capacity to recover over time. As the vehicle continues through the intersection, the localization estimate stabilizes and closely aligns with the expected trajectory based on visual context and known ArUco marker placements. This suggests that the Riccati-based method is capable of mitigating the effects of sensor errors over time, which eventually stabilizes the estimation.

While the early deviation introduces some estimation drift, the observer's convergence behavior and overall trajectory remain within a practically acceptable range. These findings reinforce the robustness of the Riccati localization framework, particularly in scenarios where measurement quality is variable or intermittently degraded.

Chapter 6

Discussion

This chapter discusses the major findings gathered in the three experimental scenarios.

The results indicate that Riccati localization could achieve accurate localization, given a rough initial estimation of the vehicle's pose at the intersection's entrance. In real-world applications, such rough estimation could be provided by other more general localization methods, or by sensors such as RTK-GPS, which may not always guarantee high accuracy but can offer sufficiently reliable starting points.

In these experiments, localization was performed using a single forward-facing camera. While this configuration is generally sufficient under the controlled conditions of the study, characterized by a known environment and moderate vehicle dynamics, it may be less robust in real-world scenarios. Limitations such as a restricted field of view, susceptibility to motion blur at higher speeds, and greater sensitivity to camera calibration errors can all degrade system observability. These factors may lead to reduced landmark visibility, even failure to meet the observability conditions, ultimately impacting localization accuracy. To improve robustness in practical deployments, future implementations could benefit from using multiple cameras with overlapping fields of view or wide-angle lenses, reducing the likelihood of aligned landmarks and enhancing environmental coverage, thus maintaining observability under more diverse conditions.

Landmarks also play an important role in bearing-based Riccati localization. In the experiment, ArUco markers were used due to their simplicity in implementation and reliability. In real-world applications, these landmarks could include stationary features such as road signs, traffic lights, or corners of buildings, provided their coordinates are known relative to the intersection.

Chapter 7

Conclusions and Future work

7.1 Conclusions

This thesis has presented a lightweight and computationally efficient bearing-based Riccati localization approach specifically designed for vehicle localization at intersections. Leveraging bearing measurements from known landmarks and onboard vehicle velocity measurements (linear and angular), the proposed Riccati observer method offers accurate and reliable pose estimation, accompanied by a theoretical guarantee of convergence under clearly defined observability conditions.

Experiments conducted in a real-world, scaled-down intersection scenario using a 1/10-scale autonomous vehicle platform (SVEA) equipped with a single forward-facing camera and ArUco markers demonstrated the effectiveness and robustness of the proposed method. The evaluation considered different driving conditions, including nominal operation, higher-speed maneuvers, and scenarios involving sensor inaccuracies. Results consistently indicated that the bearing-based Riccati localization method achieves pose estimation accuracy comparable to, and in some cases exceeding, RTK-GPS, particularly in environments where RTK-GPS reliability is compromised.

The findings underscore the potential of the Riccati observer as a practical, scalable, and robust localization solution for intersection navigation, emphasizing its advantage in computational efficiency and resilience against sensor noise and transient errors.

7.2 Future work

Building upon the findings presented in this thesis, several promising directions for future research can be pursued to further advance and validate the Riccati localization method.

A key area for further exploration involves extending the methodology to multi-agent scenarios. In such cases, vehicles can act as landmarks for one another, sharing pose information collaboratively. This cooperative approach is particularly advantageous in busy intersection environments, where individual vehicles' fields of view may be significantly obstructed by others, potentially compromising observability and leading to unreliable localization estimates. Incorporating additional mobile landmarks through cooperation can enhance the overall localization performance and robustness, particularly at complex intersections.

Integration with advanced communication infrastructures, such as 5G networks, also offers considerable potential. Leveraging the high-speed and low-latency capabilities of 5G could facilitate real-time data sharing among vehicles and between vehicles and infrastructure. This integration would enable more efficient cooperative localization schemes and further improve localization accuracy and reliability in real-world conditions.

Additionally, conducting a formal stability analysis to determine the maximum tolerable communication delay for Riccati localization would be highly beneficial. Such an analysis would establish essential guidelines for system design under practical operating conditions. Exploring strategies for mitigating the adverse effects of communication delays, such as predictive filtering, would further enhance the robustness and effectiveness of the localization methodology.

References

- [1] K.-W. Kim and G.-I. Jee, “Free-resolution probability distributions map-based precise vehicle localization in urban areas,” *Sensors*, vol. 20, no. 4, p. 1220, 2020. [Page 1.]
- [2] Z. Ren and L. Wang, “Accurate real-time localization estimation in underground mine environments based on a distance-weight map (dwm),” *Sensors*, vol. 22, no. 4, p. 1463, 2022. [Page 1.]
- [3] N. Akai, L. Y. Morales, T. Yamaguchi, E. Takeuchi, Y. Yoshihara, H. Okuda, T. Suzuki, and Y. Ninomiya, “Autonomous driving based on accurate localization using multilayer lidar and dead reckoning,” in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017. doi: 10.1109/ITSC.2017.8317797 pp. 1–6. [Page 1.]
- [4] Y.-L. Hsueh and H.-C. Chen, “Map matching for low-sampling-rate gps trajectories by exploring real-time moving directions,” *Information Sciences*, vol. 433, pp. 55–69, 2018. [Page 1.]
- [5] M. Adjrad and P. D. Groves, “Enhancing least squares gnss positioning with 3d mapping without accurate prior knowledge,” *Navigation: Journal of The Institute of Navigation*, vol. 64, no. 1, pp. 75–91, 2017. [Page 1.]
- [6] K. Eckenhoff, P. Geneva, and G. Huang, “Sensor-failure-resilient multi-imu visual-inertial navigation,” in *2019 International Conference on Robotics and Automation (ICRA)*, 2019. doi: 10.1109/ICRA.2019.8794295 pp. 3542–3548. [Page 1.]
- [7] W. Liu, Y. Liu, and R. Bucknall, “A robust localization method for unmanned surface vehicle (usv) navigation using fuzzy adaptive kalman filtering,” *IEEE Access*, vol. 7, pp. 46 071–46 083, 2019. [Page 1.]

- [8] S. Fang, H. Li, and M. Yang, "Lidar slam based multivehicle cooperative localization using iterated split cif," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 21 137–21 147, 2022. [Page 1.]
- [9] H. Kim, K. Granström, L. Svensson, S. Kim, and H. Wymeersch, "Pmbm-based slam filters in 5g mmwave vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8646–8661, 2022. [Page 1.]
- [10] P. Schmuck and M. Chli, "Ccm-slam: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," *Journal of Field Robotics*, vol. 36, no. 4, pp. 763–781, 2019. [Page 1.]
- [11] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in *2008 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2008, pp. 3946–3952. [Page 1.]
- [12] N. Nourani-Vatani, J. Roberts, and M. V. Srinivasan, "Practical visual odometry for car-like vehicles," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 3551–3557. [Pages 1 and 8.]
- [13] J. A. Grunert, "Das pothenotische problem in erweiterter gestalt nebst bber seine anwendungen in der geodasie," *Grunerts Archiv fur Mathematik und Physik*, pp. 238–248, 1841. [Page 4.]
- [14] S. Pan and X. Wang, "A survey on perspective-n-point problem," in *2021 40th Chinese Control Conference (CCC)*, 2021. doi: 10.23919/CCC52363.2021.9549863 pp. 2396–2401. [Page 4.]
- [15] R. Hartley, *Multiple view geometry in computer vision*. Cambridge university press, 2003, vol. 665. [Page 4.]
- [16] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. USA: Oxford University Press, Inc., 1998. ISBN 0195117778 [Page 5.]
- [17] M. Salgado, R. Middleton, and G. C. Goodwin, "Connection between continuous and discrete riccati equations with applications to kalman filtering," in *IEE Proceedings D (Control Theory and Applications)*, vol. 135, no. 1. IET, 1988, pp. 28–34. [Page 6.]

- [18] R. Diestel, *Graph theory*. Springer (print edition); Reinhard Diestel (eBooks), 2024. [Page 6.]
- [19] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, “Orb-slam: A versatile and accurate monocular slam system,” *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015. doi: 10.1109/TRO.2015.2463671 [Page 7.]
- [20] R. Mur-Artal and J. D. Tardós, “Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras,” *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017. doi: 10.1109/TRO.2017.2705103 [Page 7.]
- [21] D. Nister, O. Naroditsky, and J. Bergen, “Visual odometry,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, 2004. doi: 10.1109/CVPR.2004.1315094 pp. I–I. [Page 7.]
- [22] D. V. Hamme, W. Goeman, P. Veelaert, and W. Philips, “Robust monocular visual odometry for road vehicles using uncertain perspective projection,” *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, pp. 1–21, 2015. [Page 8.]
- [23] C. Wang, H. Huang, Y. Ji, B. Wang, and M. Yang, “Vehicle localization at an intersection using a traffic light map,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1432–1441, 2019. doi: 10.1109/TITS.2018.2851788 [Page 8.]
- [24] N. Mattern and G. Wanielik, “Camera-based vehicle localization at intersections using detailed digital maps,” in *IEEE/ION Position, Location and Navigation Symposium*, 2010. doi: 10.1109/PLANS.2010.5507195 pp. 1100–1107. [Page 8.]
- [25] T. Hamel and C. Samson, “Riccati observers for the nonstationary pnp problem,” *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 726–741, 2018. doi: 10.1109/TAC.2017.2726179 [Pages 10 and 13.]
- [26] F. J. Jiang, M. Al-Janabi, T. Bolin, K. H. Johansson, and J. Mårtensson, “Svea: an experimental testbed for evaluating v2x use-cases,” in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022. doi: 10.1109/ITSC55140.2022.9922544 pp. 3484–3489. [Page 15.]

- [27] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, 2018. doi: 10.1109/JIOT.2018.2812300 [Page 19.]

Appendix A

Parameter Settings for each Experimental Scenario

A.1 Evaluation under Nominal Driving Conditions

The table [A.1](#) shows the parameters used for the nominal driving condition.

Parameter	Value
$\hat{p}(0)$	$[3 \ 0.5 \ 0]^T$
$\bar{\Lambda}(0)$	$[0 \ 0 \ 0 \ 1]^T$
k	80
Q	$q_j = 25, \forall j \in \mathcal{N}$
V	$diag\{I_3, 10I_3\}$
$P(0)$	$diag\{I_3, 100I_3\}$

Table A.1: Parameters configuration of bearing-based Riccati localization for the Nominal Driving Conditions

A.2 Evaluation under Stress-Tested Driving Conditions

The table [A.2](#) shows the parameters used for the stress-tested driving condition.

Parameter	Value
$\hat{p}(0)$	$[3 \ 0.5 \ 0]^T$
$\bar{\Lambda}(0)$	$[0 \ 0 \ 0 \ 1]^T$
k	300
Q	$q_j = 10, \forall j \in \mathcal{N}$
V	$diag\{100I_3, 100I_3\}$
$P(0)$	$diag\{I_3, 100I_3\}$

Table A.2: Parameters configuration of bearing-based Riccati localization for the stress-tested driving condition

A.3 Evaluation under Degraded Sensing Conditions

The table A.3 shows the parameters used for the degraded sensing condition.

Parameter	Value
$\hat{p}(0)$	$[3.5 \ 0.5 \ 0]^T$
$\bar{\Lambda}(0)$	$[0 \ 0 \ 1 \ 0]^T$
k	100
Q	$q_j = 10, \forall j \in \mathcal{N}$
V	$diag\{0.01I_3, 10I_3\}$
$P(0)$	$diag\{I_3, 100I_3\}$

Table A.3: Parameters configuration of bearing-based Riccati localization for the degraded sensing condition

TRITA-EECS-EX-2025:941
Stockholm, Sweden 2025

www.kth.se