Title

AUTONOMOUS VEHICLE INTERACTION WITH PEDESTRIANS IN URBAN ENVIRONMENT: A GAME THEORETIC FRAMEWORK

Author
Rafee Musabbir, Sarder

Publication Date
2023

Peer reviewed|Thesis/dissertation

AUTONOMOUS VEHICLE INTERACTION WITH PEDESTRIANS IN URBAN ENVIRONMENT: A GAME THEORETIC FRAMEWORK

By

SARDER RAFEE MUSABBIR
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Prof. H. Michael Zhang, Chair

Prof. Yueyue Fan

Prof. Chen-Nee Chuah

Committee in Charge

2023

Acknowledgment

It is not always true that our preconceived destinations bring us happiness. The tumultuous journey that we take to reach our heights makes life more insightful and worthy. My travel path leading to the academic degree has made my life more meaningful and knowledgeable. Along the journey, I am grateful to have constant support and insightful guidance from my supervisor, Prof. Michael Zhang. His humbleness, virtue, and academic prudence inspired me to go beyond and explore. I want to thank the committee members Prof. Yueyue Fan and Prof. Chen-Nee Chuah for their valuable time and insights in completing the research. I am also indebted to my departmental advisor Associate Prof. Miguel Jaller, and my long-time mentor Prof Md. Hadiuzzaman for their timely advice, motivation, and professional reference.

I want to express my profound appreciation to my family and friends, especially my wife Nazneen and son Ahyan, for their endless sacrifice and loving support throughout my academic journey. In all my efforts and success, I am forever grateful to the almighty for enabling me to pursue this endeavor.

Abstract

In this research, we develop a framework for autonomous vehicles to interact safely with pedestrians in urban scenarios based on a game theoretic approach. The primary idea is to simplify the complex interaction and capture distinct behavioral features of pedestrians. The game setup incorporates a feedback control system with nonlinear dynamics. The proposed concept works to linearize the nonlinear part of the player dynamics in each iteration and uses the quadratic cost to formulate the interaction patterns at the urban streets. The algorithm takes inspiration from the iterative linear quadratic form and builds on the stochastic game to find Nash Equilibrium. We divide the dissertation research into three approaches. In the first approach, we derive the Algebraic Riccati equation for the stochastic non-zero-sum (NZS) game to capture the competitive and cooperative nature of the interaction between players. We explore distinct features of pedestrians and develop the interaction framework for AV to capture these behavioral attributes, including conservative, aggressive, and grouping. After tackling the noise in player dynamics in the first approach, we investigate the external noise during an interaction event and define a robust game framework. The game setup allows for a finite disturbance in the player cost function and noise in the system dynamics. The disturbance in the cost function tries to maximize the overall cost as an adversarial input, whereas the noise perturbs the system dynamics. In the final approach, we extend the interaction framework with a finite delay in the feedback control of the players to reflect real-world scenarios. The interaction game framework developed in this research could complement the decision-making system of autonomous driving in urban streets to ensure accurate motion planning and safe interaction with respect to pedestrians.

Contents

List of Figures

List of Tables

Chapter 1

Problem Statement

1.1 Introduction

In the 21st century, autonomous (self-driving or driver-less) vehicle technology offers a renaissance for the transportation system in terms of safety, efficiency, and accessibility by replacing the human driver [1]. The reliability and market presence of the technology depends on the safety design and adaptability to the dynamic driving scene on both highways and urban areas. Commercial entities (e.g. Tesla, Waymo, GM) have developed formal autonomy features for the highway and test the AV fleet in urban settings.

Autonomous vehicle (AV) navigation in urban areas is a critical research problem due to complex and dynamic driving requirements with other road users (pedestrians, bicyclists, and motorists) [2]. In urban areas, AVs will face unique situations, including erratic driving from motorists, the impulsive crossing of pedestrians, sharp maneuvers from bicyclists, stop-sign crossing, and yielding the right-of-way to pedestrians. Motorists address these situations carefully with cognitive and visual information. However, these scenarios remain a monumental task for AVs with object detection, route planning, motion control, and interactive decision-making. Notably, vulnerable road users (pedestrians, bicyclists) sometimes communicate with the motorist using visual cues (cooperation) or compete for the right-of-way, which adds to the AV interaction challenge.

AVs operate on various sensors, including radar, lidar, and cameras. The mechanism fuses data from the sources to generate real-world surroundings, which in turn helps the vehicle to move forward. The classical architecture of AV consists of multiple core components, including localization, mapping, motion control, and decision-making. These components work together to ensure safe motion planning and behavioral prediction in urban areas. However, the behavioral uncertainties involved with human decision-making remain the primary constraint for AV interaction. Considering the challenges associated with urban interaction, a complex decision-making framework for AV is essential. A framework that captures the type of interaction (competition or cooperation), communication, behavioral dynamics, and uncertainty will add to the existing autonomous driving system. The primary goal of this study is to develop a reliable and efficient framework with human-like driving and decision-making features to tackle the unpredictable behavior of other road users.

Game theory is widely used as a framework to generate complex decision-making problems where the collective behavior of players depends on their interactions in the surrounding environment [3]. Out of the variations, dynamic games are more robust and relatable to real-world scenarios with flexible mechanism design, differential state dynamics, and complex cost or reward structure. Moreover, based on the unique and flexible features, many control problems with multiple decision-makers use the game theory framework, such as robust control, disturbance attenuation, consensus problem, and graphical game. Thus, game theory provides a suitable modeling and computational framework to capture interactions among multiple players with different objectives.

We use the dynamic game theory to capture the interaction specifics between agents, including reasoning capability, information sharing, and conflicting or cooperating objectives. The nonzero-sum game structure is a perfect fit to describe such interactions between agents with different characteristics. The primary distinction between nonzero-sum and zero-sum games is the type of objective for each player. In the zero-sum game, the conflicting goal of players balances their payoffs, which in turn defines the competitive interaction. It implies that the summation of the players' payoff will result in zero. In contrast, the nonzero-sum or general sum game may reflect competition or cooperation, where each player's payoff or cost does not necessarily balance out the others. Notably, Nash equilibrium defines a possible solution for both game types, where no player can improve their performance cost or payoff by unilaterally changing their strategy.

Communication or information sharing is a critical component in the urban interaction scene for autonomous vehicles. In some instances of interaction, the AV will be successful in passing information to other road users. However, in most cases, AVs will likely resort to their array of sensors to collect road users' positions and predict intentions to interact safely. The dynamic game environment with a feedback strategy provides an opportunity to capture the influence of pedestrian motion on AV, leading to a realistic approximation of the interaction. Moreover, the game framework can incorporate heterogeneity in human behavior with individual preferences in action. For instance, pedestrians have varied walking speeds and movement directions over the crosswalk.

The control system is the core of any framework based on dynamic game theory. It regulates the flow of information from one component to another to generate the optimal strategy for the players. Stabilizing the control system ensures that the game setup will converge to equilibrium. Learning from the game samples also provides a unique opportunity to stabilize the solution and converge to the local or global Nash strategies.

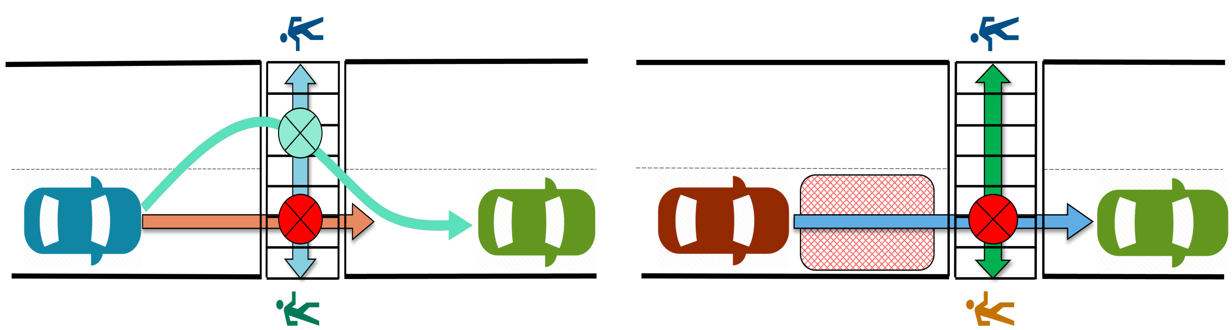
1.2 Problem Statement

Intersection navigation in the urban setting is a primary challenge for autonomous driving, especially in signalized scenarios with crosswalks. Presence of bicyclists and pedestrians while negotiating the turning movements makes it more difficult. For instance, when the vehicle is making a right turn on a red signal, the possibility of pedestrian collision arises as crosswalks parallel to the green traffic serve pedestrians simultaneously. As a result, pedestrians face a dilemma in deciding whether to yield or go faster as the turning traffic approaches the intersection stop line. In a typical scenario, the pedestrians cooperate or compete for the right-of-way. Pedestrians cooperate by communicating with motorists with visual cues and compete by crossing the intersection aggressively.

Conversely, a motorist can negotiate a turn at by yielding, communicating, or moving through the gap between the pedestrians crossing at the stop intersection. These actions remain a monumental task for AVs. AVs require a behavioral framework to navigate such urban scenarios and identify the pedestrian type (cooperating or competing) to ensure a safe interaction. This study focuses on the movement scenario for autonomous vehicles in the presence of pedestrians at the crosswalk. The real-time interaction scenarios describe the problem specifics in detail from the movement pattern of pedestrians and motorists at a signalized intersection (Laguna Beach Intersection, Orange County, California) [4].

The ***Conflict Zone*** is a roadway segment over the crosswalk where a pedestrian and vehicle may collide if one does not yield or swerve during a maneuver. Motorists generally focus on the signal head and moving traffic to make any maneuver without readily observing the pedestrian movement as per the pedestrian signal. Unlike motorists, AVs can detect the presence of pedestrians on the far side or near side of the crosswalk even before arriving at the intersection stop line. AVs can predict pedestrians' arrival time at the conflict zone and adjust speed accordingly to pass or yield the right-of-way if modeled practically for safe interaction.

The ***Gap Maneuver*** of AV is a way to balance yielding and passing maneuvers while interacting with pedestrians. The concept is similar to the pedestrian gap acceptance behavior, except vehicle length, mass, speed, and acceleration differ from pedestrians. If the predicted space and time headway is adequate between crossing pedestrians from the opposite sides of the crosswalk, the AV can maneuver through the gap/conflict zone to complete the turning maneuver. This action is practical for improved mobility or efficiency as the AV negotiates a passage with multiple or groups of pedestrians from both sides of the crosswalk.



(a) Gap Maneuver of AV (b) Pedestrian Dilemma while Crossing

Figure 1.1: AV Gap Maneuver & Pedestrian Dilemma

Pedestrian behavioral dynamics play a pivotal role in determining the safety issues that arise in the urban setting with the increasing movement of goods and people. Pedestrians are constantly at risk of sharing the road space with vehicular traffic at the signalized or unsignalized intersection. When the vehicle is making a right-turn maneuver on a red signal, the possibility of pedestrian collision arises as crosswalks parallel to the green traffic serve pedestrians simultaneously. Similarly, pedestrians crossing at the stop-sign or in the middle of street or waiting to cross during a left-turn traffic movement face unique situation. They face a dilemma in deciding whether to stop/yield or go faster as the turning traffic crosses the intersection stop-line. This study considers ***Pedestrian***

Dilemma as a behavorial feature where the crossing pedestrians either wait for the vehicle to cross or speed up to cross the roadway before the vehicle.

Based on the observed interaction between motorists and pedestrians, the following research goals develop a framework for AV in an urban setting.

Goal - 1: To investigate the movement pattern, behavioral dynamics, for pedestrians at the urban setting; and explore the effect of conflict zone, gap maneuver, and dilemma within the proposed interaction framework

Goal - 2: To design a control system for AV to tackle the uncertainty and variability of different types of road users during any form of interaction

Goal - 3: To incorporate the cooperation and competition in decision making of AV and ensure the safety, ease, and convenience of other road users, specifically for pedestrians

Goal - 4: To design and implement an information pattern for the control strategies in the game framework with respect to pedestrians during an interaction event to improve the decision making ability of AV

1.2.1 Research Questions

Before detailing the new methodology for the stochastic linear quadratic game, we provide an overview of the research questions that we tried to answer in this study and list our contribution as follows.

Why formulate Dynamic Games?

There are a few approaches to model the interactive behavior of AV in an urban setting, such as (a) scenario adaptive framework [5]; (b) behavioral prediction of other road users [6] ; (c) motion classifier and path prediction [7]; (d) micro-simulation using some fixed strategy [8]; (e) game theoretic highway driving. Notably, these models extract visual features of the surrounding environment and combine several models with learning to generate a safe behavior of AV in various scenarios. However, these models require extensive visual data combined with sensor data to develop a working model for AV. Interestingly, these models lack the concept of competitive and cooperative behavior and the stochastic nature of the interaction event.

The conservative design of the safety feature of AV translates to an always-yielding feature in an interaction event without exploring travel efficiency. These limitations suggest a new research direction incorporating intelligent control models to capture stochastic interaction events based on the pedestrian dilemma, conflict resolution, and gap acceptance behavior. On that note, dynamic games equipped with the differential equation of motion (nonlinear), cost function, and constraints for each player provide an adequate approach to modeling AVs interactive driving behavior (co-operative/competitive) in an urban setting. Moreover, the game is flexible enough to include the stochastic behavioral features of other road users (pedestrians) to develop a meaningful, realistic model.

What are the challenges in dynamic game formulation?

By taking the optimal control strategy of each player, the framework establishes an interactive game within the AV decision-making architecture. The computing complexity of the developed framework depends on the number of state variables inter-playing to describe the system's dynamics. Moreover, the induced non-linearity of player dynamics adds to the complexity and approximate solution approach.

Although most DG approaches describe and control the interaction between similar players, this study includes two types of players with different cost functions. Furthermore, the uncertain behavior of pedestrians gives rise to the stochastic form, which is challenging to solve. Finally, the players' communication strategy is feedback oriented and extended towards an asymmetric information pattern for realistic approximations. These considerations under the DG formulation increase the model complexity and computational cost.

How to formulate the game system to capture adequate real-world interactions?

The definition of the system is essential since the solution and verification of the control system vary. Control systems can be continuous, discrete, or Hybrid in design. Hybrid systems are a class of dynamical systems combining continuous-time dynamics with discrete events [9]. For this study, the AV and pedestrians reflect a continuous dynamic system transformed into discrete steps and linearize dynamics for numerical computation.

How to define the cost functions and constraints for AV as a safe and efficient agent or player?

This research aims to develop a safe and efficient interaction framework for AV from the system perspective. The safe interaction mechanism uses a circular bubble with a safe radius from other players and roadway boundaries. The speed and acceleration of AV use bounded constraints to introduce realistic behavior in an interaction event. Furthermore, the effect of cooperative or competitive behavior is more interesting to explore in the interaction event since it may indirectly affect the Av's yielding behavior or speed change. These constraints and behavioral features are part of the additive cost function of AV. The primary goal of the game formulation is to minimize the incurred cost.

What type of communication strategies are relevant to the framework?

Information distribution is essential to specify the details available to each player regarding the state and past actions of other players during the choice of control at each time step. Thus, the control strategy adopted by a player depends on the information available at each time. For this reason, different information structures may engender completely different game scenarios. Two information strategies are available for DG: (i) Open-Loop; and (ii) Feedback. Other problem-specific strategies (hierarchical play, delayed information) arise from the need to describe the problem in game form. For this study, the specific information structure involves delayed distribution, where the information will reach the interacting players at a finite delay. This will result in varying behavioral features for the active players compared to the perfect state information (feedback) scenario.

1.2.2 Contributions

This study uses game theoretic formulation to describe the competitive and cooperative nature of the interaction between AVs and pedestrians. It investigates pedestrians' aggressive or conservative walking behavior over the crosswalk. As described earlier, there is a transition in walking speed, suggesting that the pedestrians either walk fast or yield when the vehicle is adjacent to the crosswalk. Moreover, the study explores specific interaction events based on the stochastic nature of the players involved. The contributions are described in Chapters 3, 4, and 5. The contributions are listed as follows:

- We define a stochastic nonzero-sum dynamic linear quadratic game framework for AV interaction with pedestrians with Brownian motion noise on the players control. We incorporate the competitive and cooperative nature of the interaction framework. The framework is flexible for non-linear time-varying and noisy dynamics, and non-quadratic costs to incorporate the stochastic nature of the players. We derive and describe the model-based setting to compute the Nash Equilibrium with optimal feedback strategies and cost-to-go functions for the AV interaction framework. We provide numerical experimentation to illustrate the proposed AV interaction frameworks effectiveness in capturing and negotiating pedestrians behavior including conservative, aggressive, and grouping characteristics.

- We define a robust nonzero-sum dynamic linear quadratic game framework for AV interaction to allow for a finite disturbance in the player cost function and noise in the system dynamics. The disturbance in the cost function tries to maximize the overall cost as an adversarial input, whereas the system noise perturbs the system dynamics. We derive and describe the model-based setting to compute the Robust Nash Equilibrium feedback strategies and value functions for the AV interaction framework. We explore the proposed frameworks effectiveness in capturing and filtering uncertainty representations.

- We extend the interaction game framework with a finite delay in the feedback control of the players. We derive and describe the model-based setting to compute the Nash Equilibrium with delayed feedback strategies and HJB functions for the AV interaction framework. The delayed information pattern shows unique behavioral features compared to the regular feedback form.

1.2.3 Applications

The industry practice for AV involves learning from a large set of video and other sensor data. Such annotated data is combined into a neural network to generate potential motion planning and interaction strategies for urban scenarios. Unfortunately, the overall magnitude of the data and complexity of the urban scenarios with different road users (regular motorist, bicyclist, pedestrian) makes the interaction more challenging and tedious. However, in this study, we formulate a practical and safe motion control for AV by simplifying the complex interaction problem into a specific sub-problem, focusing on one approach and varying pedestrian behaviors. We validate the game model with real-world crossing and interaction time; and qualitative motion pattern matching with publicly available pedestrian datasets. The simplicity and flexibility of the developed stochastic game model dictate that it could fit into the decision-making framework of autonomous driving in urban scenes.

We develop a robust model on top of the stochastic game to tackle the time-varying external noise and introduce smooth motion planning for AV in the presence of noisy pedestrian motions. For instance, the flexibility of the robust game framework allows us to model the noise configuration for pedestrians and AV separately. This game formulation would aid in real-world motion control of AV that requires a smooth trajectory for ride comfort and safe interaction. In contrast, pedestrians' motion is always susceptible to environmental noise and individual crossing behavior. Thus, filtering the motion noise of AV for smooth trajectory and capturing the noisy behavior of pedestrians during an interaction event would aid the AV's decision-making capability in urban scenarios.

Finally, the proposed game model could be bolstered with learning mechanisms to understand scenarios from limited data and perform safely in real-world operating conditions. For instance, the game model could be structured with a modified actor-critic method to learn and update from running scenarios and perform offline simulations for random urban interactions with all road users. The flexibility of the game model already dictates that we could design customized cost functions and constraints for each type of player. The primary challenge would be the value function approximation for various players, including AV. However, the literature and our intuition suggest that we could bypass this limitation using neural network applications.

Chapter 2

Related Literature

2.1 Introduction

One of the significant challenges of autonomous vehicles (AVs) involves driving in an uncertain urban setting, emerging from behaviors of other road users (human drivers, pedestrians, cyclists) alongside varying road geometry and weather conditions. To drive in such an environment, AVs require interactive and communication features to understand and interact safely with other road users' intentions. Such interactions are critical, specifically between AV and pedestrians, the most vulnerable road user [10, 11]. Most of the safety literature on AV focuses on the vehicle-to-vehicle crash occurrence at the intersection. As such, the studies propose logical constraints as an optimization problem to solve the safety issue or design a conservative anti-crash or collision avoidance system [12, 13, 14]. Some of the approaches report on the safety issues using Naturalistic-Field Operational Test (N-FOT), an evaluation approach that tests prototype vehicles directly on public roads [15, 16, 17, 18] to explore the possible opportunities and ensure safety without considering the driving efficiency. The interaction between AVs and pedestrians warrants a research effort to describe their competitive or cooperative nature, dilemma during the crossing, and efficient driving system. This chapter reviews the interaction problem by decoupling the involved players (AV and Pedestrian) and exploring their behavioral dynamics individually and jointly based on different control and simulation approaches. The formal approach for this study is to design an interaction framework for AV to ensure an optimal trade-off between safety and mobility (travel efficiency).

2.2 Autonomous Vehicle Behavior

There are a few approaches to model the interactive behavior of AV in an urban setting, such as (a) scenario adaptive framework [5]; (b) behavioral prediction of other road users [6] ; (c) motion classifier and path prediction [7]; (d) micro-simulation using some fixed strategy [8]; (e) game theoretic highway driving. Notably, these models extract visual features of the surrounding environment and combine several models with learning to generate a safe behavior of AV in various scenarios. However, these models require extensive visual data combined with sensor data to develop a working model for AV. Interestingly, these models lack the concept of competitive and cooperative behavior and the stochastic nature of the interaction event.

Decision and control systems dictate the driving behavior and advanced safety features of autonomous vehicles. The control system design involves safety guarantees coupled with comfort and mobility (performance/ travel efficiency) in an uncertain urban setting, emerging from behaviors of other road users (drivers, pedestrians, cyclists) alongside varying road geometry and weather conditions. The conservative design of the safety feature of AV may translate to an always-yielding feature in an interaction event without exploring travel efficiency. One general approach in this domain is to utilize a hierarchical control structure with an upper and lower-level controller. The first one (outer loop) generates reference trajectories for the latter controller (inner loop) that governs vehicle dynamics based on the steering angle and acceleration/deceleration inputs required to track the reference trajectory [19]. For upper-level control (path planning, trajectory generation) of autonomous vehicles, the research approaches consist of; (a) decision trees [20, 21]; (b) partially observable Markov decision processes (POMDPs) [22], and multi-policy decision-making methods [23]. Notably, for the lower level controller (vehicle dynamics), model predictive control (MPC) can be listed as the generic approach [24, 25].

Geng et al. [5] introduced a scenario-adaptive approach for AV using prior driving knowledge (traffic rules, driving experience). First, the Hidden Markov Models (HMMs) learn the driving features. Then, the generated knowledge base specifies the model adaptation strategies and stores prior probabilities based on typical scenarios (road elements, traffic participants, and their inter-relationships). The approach test-bed includes one AV movement in Hefei city, China, for different road elements (route parts, segments, lanes, stop signs, markers, sidewalks, junctions, and traffic signs). Experimental results showed that the proposed approach extended the prediction time horizon by up to 56% (0.76 sec) on average. Furthermore, the precision improves by 26% for long-term predictions compared to the state-of-the-art driving behavior prediction models. However, although the ontology model augmented the scenario understanding ability based on semantic information, there is a need for specific scenario evaluation with interaction and decision models for pedestrians.

Chen et al. [26] proposed a new methodology for evaluating the safety and feasibility of the AVs driving strategy at un-signalized crossings. The study used Mobil-Eye sensors data installed on buses in Ann Arbor, Michigan, recording 2,973 passing events encountering pedestrians. A stochastic interaction model fitted with a bounded multivariate Gaussian mixture model (GMM) simulated and evaluated the movements of pedestrians. The pedestrians reacted to the oncoming vehicle and the passing strategies of automated vehicles when approaching un-signalized crossings. Notably, the AV performance results against the human drivers showed better efficiency and a lesser crash rate at an un-signalized crossing.

2.3 Pedestrian Behavior Dynamics

Pedestrian behavioral dynamics play a pivotal role in reflecting and determining the safety issues that arise in the urban setting with the ever-increasing movement of goods and people. Pedestrian crossing behavior is likely dependent on the surrounding environment and behavioral or decision-making process coupled with stochastic features. Previous studies on pedestrian behavior dynamics focused on movement and behavior in the regular vehicle environment. Building on the pedestrian dynamics concept, seven types of pedestrian models are available in the literature: (i) Cellular Automata (CA) models [27]; (ii) continuous force-based or social force based models [28]; (iii) microscopic models [8]; (iv) macroscopic pedestrian stream models [29]; (v) discrete choice model [30]; (vi) optimal control and game models [31, 32]; and (vii) lattice based models. A broad review of the crowd simulation models [33] grouped the models (microscopic and macroscopic) according to real-world applicability, precision, and computational burden.

For crosswalks, some approaches explored CA [34, 35, 36], gap acceptance behavior [37], car-following models [38]. However, none of these studies investigated the decision-making process of vehicle drivers and pedestrians during the interaction from a system perspective. To account for the decision process of vehicle drivers and pedestrians during the interaction at uncontrolled mid-block crosswalks, Chen et al. [39] used evolutionary game theory and cumulative prospect theory under bounded rationality and risk. The study used a CA-based model to reflect vehicle motion following the three-second rule. A modified set of pedestrian rules ensured the right moving preference and resolved the deadlock among mixed flows. The developed model was calibrated and validated using real data collected at Jianshe First Road in Wuhan, China. Results showed that the proposed model could simulate the interaction between vehicles and pedestrians. The proposed model also analyzed the effect of the interaction between vehicles and pedestrians on delays. The simulation results showed good matches with the actual observations.

[40] encoded the coupled nature of multi-pedestrian interaction using game theory and deep learning-based visual analysis to understand and predict pedestrian dynamics. The study used game theory to model the intertwined decision-making process of multiple pedestrians. Also, the visual classifiers learned mapping from pedestrian appearance for behavioral parameters. The authors argued for developing a predictive model for understanding the interaction between human and autonomous systems (e.g., autonomous cars, home robots, smart homes) that could preemptively respond to future human actions. The proposed model used video datasets consisting of multiple pedestrian interactions: (i) Zara Dataset [41]; (ii) Town Centre Dataset [42]; and (iii) LIDAR Trajectory Dataset. Experimental results showed that the model predicts and explains human interactions 25% better when compared to a state-of-the-art activity forecasting method.

To classify pedestrian behavior, Hoogendoorn and Bovy [43, 44] presented an integral theory based on three mutually dependent levels: (i) activity choice behavior and activity area choice, (ii) way-finding to reach activity areas and (iii) walking behavior. Given the activity set a pedestrian aims to perform, the theory asserts that pedestrians make a simultaneous path choice or activity schedule decision optimizing expected subjective utility, which reflects a trade-off between the utility of completing an activity and the cost of walking towards the activity areas. The walking cost consists of travel time, the discomfort of walking close to obstacles, and the stimulation of the environment. Notably, the stochastic pedestrian behavior predicts future conditions by assuming that the predicted routes are realizations of random processes [44]. The structure is modeled and operationalized as an optimal control process for pedestrian behavioral features. Later, applying this theoretical background is tested using a differential game framework, where the pedestrian is modeled as a closed feedback control system with necessary predictive and reactive capabilities [31]. However, for the differential game framework, the deterministic part is considered for the game solution strategy.

Avoidance dynamics between pedestrians where a novel modeling approach based on the Fokker Planck (FP) Nash game framework explains the causality of observed dynamics based on prescribed rules [32]. Furthermore, the study used a stochastic differential game framework, where the strategic game controls aim at avoidance maneuver by minimizing the cost functional and proved the existence of the Nash Equilibria solution. Finally, it used the experiments from cognitive psychology studies as a benchmark to draw a comparison with the resulting strategy for pedestrian motion [32]. However, the study do not consider other type of road users and focus only on the motion prediction of the pedestrians.

Cao et al.[45] reviewed the advances in modeling and controlling crowds in Cyber-Physical Systems (CPS) framework. They presented a schematic closed-loop framework for pedestrian evacuation. The study contribution enlists a fractional framework based on the physical part (mass pedestrian evacuation management of crowds) and cyber part (modeling and predicting crowds). The cyber part consists of ordinary differential equations (ODEs), partial differential equations (PDEs), and integral differential equations to describe the crowd of pedestrians using calculus of fractional order in micro-scale, macro-scale, and mesoscale, respectively. The physical part consists of pedestrian movement, smoothing fluids of crowds, and actions based on the game framework. Sensing and actuation carry out the connection or information transfer between these two parts. CCTV recorded the data on sensing (speed, density, flux, and formation patterns). Similarly, segways, cellphones, and other sensors guided the actuation task to calculate models, control the crowds, and predict the stampede. Although the framework promised many opportunities for modeling the dynamics on the micro-scale and macro-scale, rigorous testing and verification of the system is still in its early stage.

Antonini et al. [30] proposed a novel discrete choice framework for pedestrian dynamics, modeling the short-term behavior of individuals as a response to the presence of other pedestrians. The proposed model entails a physical space and a predictive model part. The physical space is a form of dynamic and individual-based spatial discretization. In comparison, the predictive part foresees a walking pedestrian's next step at a given time—two logit models (cross-nested, mixed nested) accumulated the difference in explaining pedestrian behavior. Notably, data from actual pedestrian movements from video sequences calibrated the models. The results indicated that the tendency to keep the current direction and go toward the final destination were decisions at higher levels in the individual decision process. The model captures the tendency of individuals to avoid crowded spatial positions. The authors argued that the discrete choice models' flexibility and dis-aggregate nature as a suitable fit to reflect the pedestrian behavior since the logit models capture the assumed inter-dependencies in the choice set. However, the problem framework is estimated for a specific scenario, thus requiring more data to generalize for many urban situations.

Ikeda et al. [46] developed a methodology to predict long-term pedestrian behavior (final goal/target and trajectory) to ensure adequate service from robots. The concept included a sub-goal (intermediate goals) based pedestrian behavior model and an algorithm that estimates sub-goals in an environment based on observed pedestrians' trajectories. These trajectories operated as sub-goal positions, and a probabilistic transition model between sub-goal locations predicted motion. The trajectory data set included a large shopping mall with several restaurants and shops, covering up to 860 square meters and 12003 observations in 7 hours. The chi-square method checked the prediction precision and compared it with other established methods (linear, pattern). The results indicate that the proposed methods outperformed other methods and generated 43% correct predictions.

Group dynamics while crossing is an essential consideration of our research problem. Moussad et al. [47] analyzed the motion of approximately 1500 pedestrians grouped under natural conditions to understand the group dynamics. The results showed that social interactions among group members generate specific walking patterns that influence crowd dynamics. At low density, group members walk side by side, forming a line perpendicular to the walking direction. As the density increases, however, the linear walking formation is bent forward, turning it into a V-like pattern. A well-designed model could describe these spatial patterns based on social communication between group members. The authors showed that the V-like walking pattern facilitates social interactions within the group. However, it reduced the flow because of its non-aerodynamic shape. Thus, when crowd density increased, the group organization resulted from a trade-off between walking faster and facilitating social exchange.

2.4 Interaction Between Pedestrian and AV

Researchers worked on pedestrian path prediction and motion classification to analyze the situation and possible collisions beforehand. Most of the studies involved Kalman Filtering [7], linear and multiple dynamical systems [48], Dynamical Bayesian Network [49], Gaussian Process Dynamical Models (GPDM) [50, 51]. Some also explored pedestrians' conforming behavior, such as head-turning for approaching right/left-turning vehicles when they cross intersections [52]. Moreover, others have [53], [54] analyzed the intersection layout and conditions surrounding pedestrians as variables of probabilistic functions. These functions represent pedestrians' crossing maneuvers at signalized intersections. For instance, Hashimoto et al. [55] described pedestrians' behavior based on the probabilistic Dynamical Bayesian Network (DBN). The model was constructed by imitating pedestrians' behavior and subjective assessment in real-world traffic scenarios at signalized crosswalks. It described the stochastic connections among external contexts, pedestrian behavior, and physical movement.

Millard-Ball et al. [56] explored a decision model for pedestrian crossing, including AV, where each crossing involves a subjective assessment of quick crossing or waiting and regular crossing or choosing a new route. Similarly, drivers must decide whether to yield or proceed at un-signalized crossings. The study explained this phenomenon using a crosswalk game of chicken between driver and pedestrian, where two drivers begin to drive head-on toward each other at high speed. The winner of the game is the driver who does not "chicken out" and swerve out of the way, whereas the worst possible outcome is that neither driver swerves and both vehicles collide. Thus, in equilibrium, one driver will chicken out, as losing the game is preferable to death. Though the game theoretic model is similar to that in [39] and [57], key differences include strategic incentives for pedestrian crossing and exclusion of bicycles, respectively. Based on a similar concept, Fox et al. [58] proposed a sequential chicken game to explain the interaction behavior at the un-signalized intersection. The simple model contained free parameters decrypting human preferences [58]. Later, Camaro et al. [59] extended the model to explore how such parameters can be fit to human data as a method of measuring behavior, combining empirical Game Theory and a Bayesian Gaussian Process analysis.

As part of the intelligent driving system, intention estimation algorithms gained interest to predict future actions of pedestrians [6], and drivers [60] for safe interaction. One classical approach towards intention estimation uses real-world data to predict pedestrian walking behavior. These models have two bases of estimation, one involving dynamic information such as the position and velocity of pedestrians [61]. The other involves the input of the contextual information of the scene, such as pedestrian signal state, individual pedestrian or group, and distance to the curb [62]. For example, Brouwer et al. [63] investigated the role of different types of information in collision estimation, considering the four factors: (i) dynamics (directions pedestrian can potentially move to and time to collision); (ii) physical elements (pedestrian's moving direction and distance to the car and velocity); (iii) awareness (in terms of head orientation towards the vehicle); and (iv) obstacles. The results indicated that physical elements and awareness are the best predictors of collisions in isolation. Moreover, combining the four factors provides the best possible estimation/prediction results. Considering all these details, Rasouli et al. [10] showcased a thorough review of the practical approaches to describe the interaction between pedestrians and AVs.

Intent communication of AV is part of the practical approaches to interaction with pedestrians. These communication mechanisms are investigated in several approaches, such as Vehicle to Vehicle (V2V) [64] and Vehicle to Pedestrian (V2P) [65] wireless communication mechanisms, and various visual intent displays such as LED lights [66] or projection [67]. For instance, Hussein et al. [65] proposed using a smartphone application that broadcasts the pedestrian's position and receives the location of nearby AVs based on V2P. The application then calculates and predicts the location and time of the collision, and if the pedestrian is in danger, sends a warning signal. Although these methods provide an ad-hoc solution to the communication mechanism, they lack the theoretical background of interaction and treat the problem dynamics as dealing with a rigid dynamic object rather than a social being [61].

Ningbo et al. [68] focused on the interactions at unmarked roadways and established a modified social force model, adding new force terms for pedestrian and vehicle models. In addition, they developed a decision-making model based on gap acceptance theory and conflict avoidance models to ensure safety and replicate pedestrians' perception of the traffic environment. The study site included one unmarked roadway near a bus station in Changchun City, China, with a width of 27 meters. The observation period of 2 hours continued for two consecutive working days, accumulating crossing trajectories of 1023 pedestrians and 517 vehicles with a total of 113,732 coordinates. The model was well-calibrated and validated using the observed trajectories and statistical means (RMSE). However, the interaction dynamics for turning traffic, crossing dilemma, and aggressive group behavior features require further exploration.

Schneemann [69] investigated the motorists' hand signals to express the intent to yield the right-of-way to pedestrians as part of autonomous driving functions. The study explored the possible approach of motorists to resolving an ambiguous situation when pedestrians do not provide any informal signals. The authors conducted a driving study with an instrumented Audi A7 in the inner city of Ingolstadt, Germany, with five crosswalks and two different speed limits (30 km/h and 50 km/h). Simple regression analysis is performed on the recorded dataset to identify the critical factors affecting drivers' and pedestrians' decision-making during an interaction event at the crosswalk.

2.4.1 Game Models in Autonomous Driving

With a motivation to establish an iterative differential game methodology in a linear quadratic format for an autonomous system, Fridovich-Keil et al. [70] explored the general-sum game as a framework to establish the multiagent interaction between players in a traffic scenario. The approach is closely related to the iterative linear quadratic regulator methodology and follows a feedback information system to update player control strategies over time. Furthermore, the game addresses convergence in local Nash Equilibria and is susceptible to arbitrary initial conditions.

Similarly, the ALGAMES uses the Lagrangian format of the cost or objective function to solve for the constrained linear quadratic game in open-loop form [71]. The game system uses primal-dual formulation and Khun-Katush-Tucker condition to solve the game in open-loop form and employs a receding horizon (Model Predictive Control) model to compare with the feedback game (iLQGames) [70]. The game results show more efficient performance due to the open-loop form and receding horizon window. However, the game system is susceptible to the prediction horizon window compared to the feedback game system.

Fiasc et al. [72] presented a hierarchical trajectory planning framework for mixed traffic scenarios on highways where an autonomous vehicle (AV) interacts with regular motorists. The study proposed a nonzero-sum game with a feedback information system following the Stackelberg equilibrium to explain the influence of an AV on motorists. Results from the game augmented the AV planning and prediction strategies in a low-level trajectory planner on highway scenarios. However, implementing the Stackelberg equilibrium (leader-follower) may only reflect cooperation in highway driving scenarios. Also, game planning is susceptible to the initial conditions where the leading vehicle (AV or motorist) may decide on a fixed outcome in different scenarios.

Liniger and Lygeros [73] proposed a non-cooperative (zero-sum) racing game between two autonomous cars to model the specific interactions in three directions. They modeled the first two approaches (sequential and cooperative) as bimatrix games where the collision constraints defined the interaction, and each player optimized their cost function. The third or final approach explored the blocking behavior by assigning greater rewards for staying ahead at the horizon's end. The games considered Stackelberg and Nash Equilibria using a sequential maximization approach. Similar to previous studies, the game includes a receding horizon component for online or real-time computation. The results indicated that the sequential game form was the most efficient blocking technique in a closed-loop setup. However, such cases neglected the follower and demonstrated the highest collision risk.

Others explored similar game theoretic planners for two-player [74] and multi-player [75] drone racing. The competitive game form included the goals and constraints of each player as shared knowledge without communicating the control strategies. Instead, the players interacted reactively to avoid collisions with each other. Finally, the authors explored Nash Equilibria as the game solution based on the premise that Stackelberg games may result in overly conservative action sequences. The planner approximates the Nash Equilibria using an iterative best response algorithm.

2.5 Research Gap Summary

The available models in the literature extract visual features of the surrounding environment and combine several models with learning to generate a safe and intelligent behavior of AV for interactive scenarios. As such, these models require an extensive/exhaustive amount of visual data combined with sensor data to develop a working model for AV. Interestingly, these models lack the concept of competitive and cooperative behavior and the stochastic nature of the interaction event. Moreover, the conservative design of the safety feature of AV may translate to an always-yielding feature in an interaction event without exploring travel efficiency. All these limitations point to a new research direction incorporating sensor-based control models of AV for stochastic interaction events, considering safety, pedestrian dilemma, conflict resolution, and gap acceptance behavior.

Chapter 3

Autonomous Vehicles Interaction with Pedestrians in Urban Streets

3.1 Introduction

AV driving action becomes more complex considering the movement patterns of pedestrians. Generally, advanced AV features anticipate pedestrian behavior using learned scenarios and real-time predictions [5], without considering the nature of interaction (competition or cooperation). Thus maintain a conservative safe distance from pedestrians during an interaction event. However, this study attempts to incorporate the nature of interaction within AV control system by developing an interaction framework. For that goal, the necessary assumptions of AV driving behavior are noted as follows.

The dynamic game consists of two parts: (a) game elements; and (b) algorithm design. The game elements describes players' behavioral features, cost components, and equilibrium concepts. On the other hand, the algorithm part builds on the local approximation, backward pass, control regularization, and line search for convergence. The study's uniqueness is combining two elements and developing a stochastic game theoretic framework for Autonomous vehicle interaction with pedestrians in urban streets. The stochastic component for the interaction framework builds on the control-dependent Brownian motion or Gaussian noise. The primary motivation for using control-dependent noise is to vary the control strategy of the linear quadratic game setting since noise on the state does not impact the dynamic system's optimal control.

3.2 Contributions

The primary contributions of this chapter are:

- We define a stochastic nonzero-sum dynamic linear quadratic game framework for AV interaction with pedestrians with Brownian motion noise on the players control. We incorporate the competitive and cooperative nature of the interaction framework.
- We establish the flexible interaction framework for non-linear time-varying and noisy dynamics, and non-quadratic costs to incorporate the stochastic nature of the players.
- We derive and describe the model-based setting to compute the Nash Equilibrium with optimal feedback strategies and cost-to-go functions for the AV interaction framework.
- We provide numerical experimentation to illustrate the proposed AV interaction frameworks effectiveness in capturing and negotiating pedestrians behavior including dilemma, conservative, aggressive, and grouping characteristics.

3.3 Game Elements

This section describes the multiplayer design setup for the interaction based linear quadratic dynamic game. This section provides an introduction to the optimal control framework and extends the framework to describe the behavior of the autonomous vehicles and pedestrians during an interaction event at the signalized intersection. Behavioral models of both AV and pedestrian are explored in this chapter, where the AV model incorporates deterministic feedback control approach and the pedestrian model involves stochastic feedback oriented approach. *Feedback* can be referred to as an environment where two (or more) dynamical systems are connected together such that each system influences the other and their dynamics are thus strongly coupled [76]. The cost function for both of these players differ significantly from each other based on the variability of utility, speed, acceleration, perception, and other behavioral dynamics.

The necessary assumptions of AV driving and pedestrian walking behavior are noted as follows.

Perception and prediction: The AV will predict other road users behavior, and uncertainty to navigate the urban scenarios using cooperative and competitive (non-cooperative) strategies. Unlike AV, pedestrians have limited predictive capability (audible and visual perception limitation) to assess the movement of other road users. They either cooperate or compete for the right-of-way with incoming vehicles at the crosswalk. Moreover, the group behavior on the crosswalk is vastly different (slow movement, aggressive) from an individual pedestrian.

Speed and safety preference: The AV will maintain a desired travel speed and comfort level with a specified acceleration-deceleration range in complement to the game strategy. On the other hand, pedestrians are likely to maintain a safe distance and comfortable pace over the crosswalk. However, pedestrians may compete for the right-of-way when the waiting time increases before crossing [4].

Performance feature: AV will try to minimize any incurring cost from driving at a desired speed and maintain safe distance with other road users. Similarly, pedestrians will likely minimize any cost by maintaining safe gap with respect to other road users and adjust preferable walking speed according to the presence of other road users.

Feedback Communication: Feedback information is the current value/observation of the state at every stage of the game. The AV behaves like a state-feedback controlled system [77]. Likewise, pedestrians walking behavior based on the current observation can be replicated through a feedback controller [44].

3.3.1 Player Model

In comparison to the pedestrians, AV continuously sense the surrounding environment to make informed decisions and predict into the near future for safe compatibility with other road users (e.g. motorists, pedestrians).

1. The AV behaves like a state-feedback controller. Using advanced onboard sensory application, it will react with the surrounding environment and adjust accordingly.

2. AV will predict and maintain a collision avoidance radius from other road users behavior to interact safely. Coupled with the state-feedback option, the AV will navigate more variant and challenging urban scenarios. Based on the road users behavior and its own set of parameters AV will formulate optimal strategies in two ways: (i) non-cooperative; (ii) cooperative.

3. AV will try to minimize any incurring cost from: (i) driving at a desired speed; (ii) smoothness and comfort by adjusting acceleration; (iii) optimal driving strategy considering the safe distance while interacting with pedestrians.

Inspired from the Hoogendron and Bovy's approach [44], the necessary assumptions of pedestrian walking behavior on the crosswalk are noted as follows.

1. Pedestrians adjust the walking behavior (routes, choices, etc.) according to the current observation of the surrounding environment, analogous to the operation of a feedback controller.

2. Pedestrians have limited predictive possibilities (audible and visual perception limitation) to assess the movement of other road users including other pedestrians or vehicles according to cooperative or non-cooperative strategy. This limitation is reflected by their actions over time and space, implying that they mainly consider pedestrians in their direct environment or area of influence and vision.

3. Pedestrians will be likely to minimize any cost from: (i) proximity cost or safe gap with respect to other road users; (ii) deviation from the initial planned or desired trajectory; (iii) interactive group behavior cost; (iv) control cost (slowing, running, side-stepping, etc.)

AV Dynamics

We model the AV dynamics using Kinematic Bicycle model [78] from the literature to simplify the vehicle dynamics without using the slip angle. The state dynamics of the vehicle, $\dot{z_v} = f(t, z_v, u)$

State:

$$z_v = [\quad X_v \quad Y_v \quad v_a \quad \phi_v \quad]^T \tag{3.3.1}$$

where, X_v, Y_v denotes the AV position in Cartesian co-ordinate, and v_a, ϕ_v denotes the speed and heading of AV, respectively. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{X_v} \\ \dot{Y_v} \\ \dot{v_a} \\ \dot{\phi_v} \end{bmatrix} = \begin{bmatrix} v_a \cdot cos\phi_v \\ v_a \cdot sin\phi_v \\ a_v \\ \frac{v_a}{L} \cdot tan\delta_v \end{bmatrix} \tag{3.3.2}$$

where the control variables are $u = (a_v, \delta_v)$.

Pedestrians Dynamics

We model the pedestrian dynamics using Unicycle model [78] from the literature to simplify the interaction problem. The state dynamics of the pedestrians , $\dot{z_p} = f(t, z_p, u)$ State:

$$z_p = [\quad X_p \quad Y_p \quad v_p \quad \theta_p \quad]^T \tag{3.3.3}$$

where, x_p, y_p denotes the pedestrian position in Cartesian co-ordinate, and v_p, θ_p denotes the speed and heading of the pedestrian. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{x_p} \\ \dot{y_p} \\ \dot{v_p} \\ \dot{\theta_p} \end{bmatrix} = \begin{bmatrix} v_p \cdot cos\theta_p \\ v_p \cdot sin\theta_p \\ a_p \\ \omega_p \end{bmatrix} \tag{3.3.4}$$

where the control variables are $u_t^p = (a_p, \omega_p)$.

Collision Avoidance

The collision avoidance is introduced in the system as a circle of radius r_c. The collision constraints are activated when each players move closer to each other. It is expressed in terms of relative position of the players. The collision avoidance is used a cost term for the player cost function.

$$\ell\{|p^i - p^j| < r_c\}(r_c - ||p^i - p^j||)^2 \tag{3.3.5}$$

The $\ell\cdot$ is an indicator function, which activates when the other road user is within the collision radius, otherwise it remains 0. Here, $p^i = (p_x^i, p_y^i)$ denotes the x and y position of the of the i-th player.

Travel Efficiency

The efficiency component ensures the maximum possible speed as per the speed limit. The term actually promotes the AV to behave as efficiently as possible given an interaction scenario in presence of other road users. The desired speed of the target vehicle (AV) is noted as $v_d(s)$, v_a is the current speed, and δ acts as the smoothing parameter for the output.

$$\delta_v(\|v_d(s) - v_a\|)^2 \tag{3.3.6}$$

Roadway Boundary

The roadway boundary is modeled to keep the vehicle within the designated roadway. It is also modeled as avoidance radius r_b where the distance between the vehicle and closest point on the boundary must remain larger than the radius value. Here b_p is defined as the distance to boundary.

$$\ell\{p^v > r_b\}(r_b - \|p^v - b_p\|)^2 \tag{3.3.7}$$

The $\ell\cdot$ is an indicator function, which activates when the distance between the vehicle and boundary is within the avoidance radius, otherwise it remains 0. Here, $p^v = (p_x^v, p_y^v)$ denotes the x and y position of the of the vehicle.

Goal

The goal or target is the final endpoint defined for each player in the game as follows,

$$\ell\{t > T - t_{target}\}(\|p^i - p^i_{target}\|)^2 \tag{3.3.8}$$

where, T is finite time horizon, p^i defines the player position, and p^i_{target} is the target or goal state for each player.

General Constraints

The general constraints include both state and control constraints. For state constraint, we can assign a maximum and minimum bound for speed of the AV. On the other hand we can also limit the maximum and minimum value for control components (acceleration and steering).

The speed bound is defined as follows:

$$\ell\{v_i > \bar{v}_i\}(v_i - \bar{v}_i)^2 + \ell\{v_i < \underline{v}_i\}(v_i - \underline{v}_i)^2 \tag{3.3.9}$$

where, $\bar{v}_i$ and $\underline{v}_i$ defines the maximum and minimum speed. The acceleration bound is defined as follows.

$$w_c\{\ell\{a_i > \bar{a}_i\}(a_i - \bar{a}_i)^2 + \ell\{a_i < \underline{a}_i\}(a_i - \underline{a}_i)^2\} \tag{3.3.10}$$

where, w_c is the assigned weight of the acceleration cost component, a_v is the acceleration of AV at time step t, and $(\bar{a}_i, \underline{a}_i)$ is the maximum or minimum acceleration. Based on the weight term the AV can be bound between a maximum and minimum acceleration.

The description and transformation of the cost function for each player is provided in the algorithm section below.

3.4 Algorithm Design

3.4.1 System Description

Consider the nonlinear dynamical system described by a stochastic differential equation,

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u})dt + \mathcal{F}(\mathbf{x}, \mathbf{u})d\omega \tag{3.4.1}$$

where, state $x \in \mathbb{R}^n$, control $u \in \mathbb{R}^m$ and standard Brownian Motion noise $\omega \in \mathbb{R}^p$. We define the running cost for each player $(i = 1, 2, .., N)$ as $L^i(x, u, t)$ and $h^i(x(T))$ as the terminal cost, where T denotes the final time of the finite horizon game. Here, $u(t)^i$ represents the control law for each player. The cost function is a collection of cost designed to accumulate over time with the game system initialized with each player's state x at time t, and controlled over the finite horizon T based on the control $u(t)^i$.

$$\mathcal{L}^i(x, u, t) = \mathbf{E}\Big[\mathbf{h}^i(x(T)) + \int_t^T L^i(\tau, x(\tau), u(\tau))d\tau\Big], \quad \forall i = 1, 2, .., N \tag{3.4.2}$$

The expectation arises due to the presence of stochastic process ω. In the game design, the noise is added to the control variables of each player, known as control-dependent noise. The premise of using noise is that the players do not fully know the control variables. For example, the AV can not fully access their control variable or the other players. In other words, AVs do not have complete knowledge of the control pattern of pedestrians. Likewise, pedestrians cannot assume the AV controls (acceleration and steering) while interacting at the crosswalk. Moreover, as we translate the game into the linear quadratic form, the control-dependent noise directly impacts the players' optimal control strategies, unlike state-dependent noise.

The objective of each player in the game is to find the control strategy that minimizes their respective cost function. Finding the optimum global strategy for each player may be challenging in stochastic game conditions. However, we can approximate the optimum local strategy for each player within a specific feasible condition. We will approximate the optimal control strategy in the vicinity of the optimal trajectory by applying the optimal control on the system dynamics.

We discretize the time horizon to formulate the system dynamics into several steps for the time-varying linear dynamics approach. Since the trajectory depends on the initial conditions, the optimal control approximations also depend on similar conditions.

3.4.2 Local Approximation

The locally optimized control law is constructed iteratively. Each iteration of the algorithm begins with a control sequence u_t and the corresponding trajectory x_t, obtained by applying u_t to the system dynamics (x_{t+1}) with initial states (x_0) following Euler's integration.

$$x_{t+1} = x_t + \Delta t \cdot f(x_t, u_t)$$

where time is discretized as $t = 1, ..., k$ with time step $\Delta t = T/(k-1)$. The time varying quantities such as controls for each player become $u_t^i \triangleq u^i((k-1)\Delta t)$.

We linearize the system dynamics and quadraticize the cost functions around the nominal trajectory $(\bar{x}, \bar{u})$ to obtain a discrete time linear system with quadratic cost. The linear approximation of the system dynamics transform the nonlinear system as follows,

$$x_{t+1} = \mathcal{A}_t x_t + \sum_{j=1}^{N} \mathcal{B}_t^j u_t^j + \mathcal{C}_{\mathbf{t}}(\omega_t), \quad \forall t = 0, ..., k-1 \tag{3.4.3}$$

where,

$$\mathcal{A}_t = \mathcal{I} + \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{x}}, \qquad \mathcal{B}_t = \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{u}}, \qquad \mathcal{C}_t = \sqrt{\Delta t} \cdot \frac{\partial \mathcal{F}}{\partial \mathbf{x}}, \quad \forall t = 0, ..., k-1$$

The quadratic transformation of the player costs results as follows.

$$\mathcal{J}_t^i = \frac{1}{2}\sum_{t=1}^{T}\Bigg[(x_t^T Q_t{}^i + 2q_t{}^i)x_t + \sum_{j=1}^{N}(u_t^{jT}\mathcal{R}_t^{ij}u_t^j + 2r_t^{ij}u_t^j)\Bigg], \qquad \forall i = 1, 2, ..N$$

where,

$$\begin{aligned} q_t &= \Delta t \cdot \frac{\partial \ell}{\partial \mathbf{x}}, \qquad Q_t = \Delta t \cdot \frac{\partial^2 \ell}{(\partial \mathbf{x})^2} \\ r_t &= \Delta t \cdot \frac{\partial \ell}{\partial \mathbf{u}}, \qquad \mathcal{R}_t = \Delta t \cdot \frac{\partial^2 \ell}{(\partial \mathbf{u})^2} \end{aligned} \tag{3.4.4}$$

The $\sqrt{\Delta t}$ term describes the linear growth of the covariance of the Brownian motion with time. The noise covariance is,

$$Cov\Big[\mathcal{C}_{\mathbf{t}}(u_t)\chi_{\mathbf{t}}\Big] = \sum_{j=1}^{N}(\mathcal{C}_{\mathbf{t}}{}^j u_t^j)(\mathcal{C}_{\mathbf{t}}{}^j u_t^j)^T \tag{3.4.5}$$

The noise model captures multiplicative noise to the control strategies of the players.

3.5 Computing Cost-to-go

The approximation of the optimal control strategy of players will be affine in form $u_t^{i*} = \pi_t^i(x_t) = K_t^i \cdot x_t + \alpha_t^i$, where K_t^i is the control gain. The feed-forward component α_t^i in the optimal control strategy arises from the transformation of the system and cost of players. The control strategy of players are approximately optimal because the game system has control constraints and non-convex costs.

We can design the control strategies u_t^i for the players for time steps $t, ..., (k-1)$. Then, the cost-to-go or value function $\mathcal{V}_t^i(x_t)$ for each player is well defined as the costs accumulates over time with initial state x_0 and time t_0, following the optimal strategies π_t^i for the rest of the time steps. The value function solution for the linear-quadratic game can be found in the quadratic form [3].

$$\mathcal{V}_t^i = \frac{1}{2} x_t^T \mathcal{S}_t^i x_t + \xi_t^{iT} x_t + n_t^i \tag{3.5.1}$$

with $\mathcal{S}_{T+1}^i = 0$, $\xi_{T+1}^i = 0$, $\eta_{T+1}^i = 0$ to be consistent with the final value condition. The linear part of the equation is represented by n_t^i. The Hamilton-Jacobi-Bellman equations for each player's value function,

$$\begin{aligned} \mathcal{V}_t^i(x_t) &= \ell_t^i(x_t, u_t) + E\Big[\mathcal{V}_{t+1}^i(x_{t+1})\Big] \\ &= \min_{u_t^i} \Bigg\{ \frac{1}{2}\Bigg((x_t^T Q_t^i + 2q_t^{iT}) x_t + \sum_{j=1}^{N} (u_t^T \mathcal{R}_t^{ij} + 2r_t^{ij}) u_t^j \Bigg) \\ &+ E\Bigg[\mathcal{V}_{t+1}^i\Big(\mathcal{A}_t x_t + \sum_{j=1}^{N} (\mathcal{B}_t^j u_t^j + \mathcal{C}_t^j \chi_t^j)\Big)\Bigg]\Bigg\} \end{aligned} \tag{3.5.2}$$

with the final value $\mathcal{V}_{T+1}^i(x_{t+1}) = 0$.

3.5.1 Backward Pass

The feedback control law is obtained by finding the minimizes of $\mathcal{V}_t^i(x_t)$. The assumption for this case relates to convexity and set the gradient to zero,

$$0 = \mathcal{R}_t^{ii} u_t^i + r_t^{iiT} + \mathcal{B}_t^{iT} \mathcal{S}_{t+1}^i \Bigg(\mathcal{A}_t x_t + \sum_{j=1}^{N} \Big(\mathcal{B}_t^j u_t^j + \mathcal{C}_t^j \chi_t^j \Big) \Bigg) + \mathcal{B}_{t+1}^{iT} \xi_{t+1}^i \tag{3.5.3}$$

The admissible optimal control u_t^{i*} based on the linear feedback is an affine function. The α_t^i term is added to the equation to capture the linear transformation of the nonlinear system.

$$u_t^{i*} = -\mathcal{K}_t^i x_t - \alpha_t^i \tag{3.5.4}$$

Inserting the optimal control strategy u_t^{i*} leads to,

$$0 = -\mathcal{R}_t^{ii}(\mathcal{K}_t^i x_t + \alpha_t^i) + r_t^{iiT} + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\Big(\mathcal{A}_t x_t - \sum_{j=1}^{N}\Big(\mathcal{B}_t^j(\mathcal{K}_t^j x_t + \alpha_t^j) + \mathcal{C}_t^j \chi_t^j\Big)\Big) + \mathcal{B}_{t+1}^{iT}\xi_{t+1}^i \quad (3.5.5)$$

Two similar systems of equations to find $\mathcal{K}_t^i$ and α_t^i,

$$\begin{aligned}
&\Big(\mathcal{R}_t^{ii} + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\mathcal{B}_t^i + \mathcal{C}_t^j\chi_t^j\Big)\mathcal{K}_t^i + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\sum_{j\neq i}\mathcal{B}_t^j\mathcal{K}_t^j = \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i A_t \\
&\Big(\mathcal{R}_t^{ii} + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\mathcal{B}_t^i + \mathcal{C}_t^j\chi_t^j\Big)\alpha_t^i + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\sum_{j\neq i}\mathcal{B}_t^j\alpha_t^j = \mathcal{B}_{t+1}^{jT}\xi_{t+1}^i + r_t^{ii}
\end{aligned} \quad (3.5.6)$$

We can define the value function approximation using $\mathcal{H}$, $\mathcal{M}$, $\mathcal{P}$

$$\begin{aligned}
\mathcal{H} &= \mathcal{R}_t^{ii} + \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\mathcal{B}_t^i + \mathcal{C}_t^j\chi_t^j \\
\mathcal{M} &= r_t^{ii} + \mathcal{B}_t^{iT}\xi_{t+1}^i \\
\mathcal{P} &= \mathcal{B}_t^{iT}\mathcal{S}_{t+1}^i\mathcal{A}_t
\end{aligned} \quad (3.5.7)$$

Now we derive the necessary expression for $\mathcal{S}_t^i$, ξ_t^i, and η_t^i.

$$\begin{aligned}
\mathcal{V}_t^i(x_t) = \Bigg\{&\frac{1}{2}\Bigg((x_t^T Q_t^i + 2q_t^{iT})x_t + \sum_{j=1}^{N}\big((\mathcal{K}_t^j x_t + \alpha_t^j)^T R_t^{ij} - 2r_t^{ijT}\big)\big(\mathcal{K}_t^j x_t + \alpha_t^j\big)\Bigg) \\
&+ \frac{1}{2}\Bigg(\Big(\mathcal{A}_t x_t - \sum_{j=1}^{N}\mathcal{B}_t^j(\mathcal{K}_t^j x_t + \alpha_t^j)\Big)^T\mathcal{S}_{t+1}^i + 2\xi_{T+1}^{iT}\Bigg) \\
&\times\Bigg(\mathcal{A}_t x_t - \sum_{j=1}^{N}\mathcal{B}_t^j(\mathcal{K}_t^j x_t + \alpha_t^j)\Bigg) + \eta_{t+1}^i\Bigg\}
\end{aligned}$$

To shorten the equation, we consider,

$$\mathcal{G}_t = \mathcal{A}_t - \sum_{j=1}^{N} \mathcal{B}_t^j \mathcal{K}_t^j$$
$$\beta_t = - \sum_{j=1}^{N} \mathcal{B}_t^j \alpha_t^j$$

The value function equation becomes,

$$\begin{aligned}
\mathcal{S}_t^i &= Q_t^i + \sum_{j=1}^{N} \mathcal{K}_t^{iT} R_t^{ij} \mathcal{K}_t^i + \mathcal{G}_t^T \mathcal{S}_{t+1}^i \mathcal{G}_t + \sum_{j=1}^{N} \mathcal{K}_t^{iT} \mathcal{C}_t^{jT} \mathcal{S}_{t+1}^i \mathcal{C}_t^j \mathcal{K}_t^i, \quad \mathcal{S}_{T+1} = 0 \\
\xi_t^i &= q_t^i + \sum_{j=1}^{N} (\mathcal{K}_t^{iT} R_t^{ij} \alpha_t^j - \mathcal{K}_t^{iT} r_t^{ij}) + \mathcal{G}_t^T (\xi_{t+1}^i + \mathcal{S}_{t+1}^i \beta_t) + \sum_{j=1}^{N} \mathcal{K}_t^{iT} \mathcal{C}_t^{jT} \mathcal{S}_{t+1}^i \mathcal{C}_t^j \alpha_t^j, \quad \xi_{t+1} = 0 \quad (3.5.8) \\
\eta_t^i &= \frac{1}{2} \left[\sum_{j=1}^{N} \left(\alpha_t^{jT} R_t^{ij} - 2 r_t^{ijT}\right) \alpha_t^j - \left(2\xi_{t+1}^i - \mathcal{S}_{t+1}^i \sum_{j=1}^{N} \mathcal{B}_t^j \alpha_t^j\right)^T \sum_{j=1}^{N} \mathcal{B}_t^j \alpha_t^j + \eta_{t+1}^i \right]
\end{aligned}$$

3.5.2 Regularization and Control

Regularization and control are two important parts of the proposed method. The combination of regularization and control ensures the stability and optimality of the proposed game system. For instance, regularization prevents the overfiting to the noise sequence and control stabilizes the system by reduce cost to ensure optimality. The cost components are added together to form the cost function for each player. The constraints are shifted into the cost or objective function, where the optimal condition tries to minimize the cost for each player. We also use weightage value on each of the cost components to control the constraints violation.

After deriving the dynamic programming solution for the game problem, the new control gains $\mathcal{K}_t$ and offset updates α_t as follows,

$$\begin{aligned} \mathcal{K}_t &= -(\mathcal{H}_t + \nu I)^{-1}\mathcal{M}_t \\ \alpha_t &= -(\mathcal{H}_t + \nu I)^{-1}\mathcal{P}_t \end{aligned} \tag{3.5.9}$$

where $\nu \geq 0$ is a regularization parameter to prevent $(\mathcal{H}_t + \nu I)$ from having negative eigenvalues. The ν is changed based on the performance result and updated across multiple iterations. This step helps the algorithm to reach fast convergence near a local minimum while ensuring the positive-definite property of $(\mathcal{H}_t + \nu I)$.

3.5.3 Line Search for Convergence

Line search is generally used to ascertain the convergence condition of an iterative type algorithm. In the proposed linear quadratic game with control dependent noise, it searches for the optimal strategy that minimizes the cost function associated with the players. More appropriately, the update in each iteration may lead to an increased cost or divergence in case the new trajectory is far from the valid region of local approximation. To avoid this incidence, the new nominal control trajectory is computed by backtracking line search with initial step of 0.5. The subsequent step is varied manually to test different approximation of the nominal trajectory. When the candidate trajectory results in a lower cost estimation or remains within the valid region of approximation, then this trajectory is accepted. Otherwise, the entire process is reiterated to find the appropriate trajectory where the control is within a threshold. After meeting the convergence criterion the proposed algorithm returns nominal trajectory and optimal feedback strategies.

3.5.4 Computational Complexity

The computational complexity of the stochastic game approach is comparable the iterative linear quadratic approach. The number of operations per iteration scales up with the number of players. At every iteration, the dynamic system with control-dependent noise is linearized. Given that the state dimension z_t^i of each player ($i = 1, .., N$) is greater than the control dimension u_t^i, linearization task computes $O(n^2)$ at each time interval, where n is the dimension. The quadratic transformation of the costs computes $O(Nn^2)$ at each time interval. However, the complexity is verified for specific deterministic cases. In the presence of Brownian motion noise with $\mu \geq 0.6$ and $\sigma \geq 1$ the computation complexity changes at every iteration of the proposed algorithm.

3.6 Numerical Results

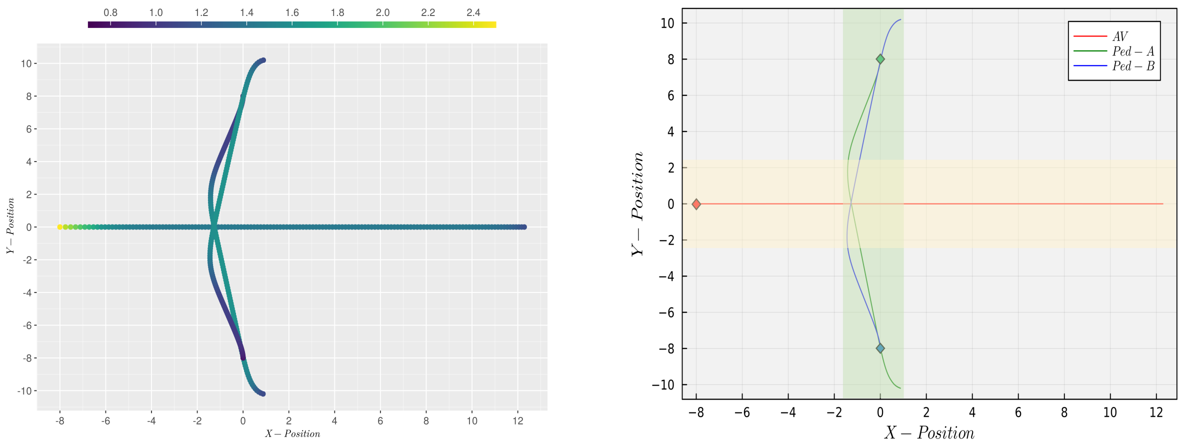
The numerical results section is divided into two parts. The first part describes the basic interaction of the players and show how the multiplicative noise effects the convergence with more players in the game. The second part identifies the aggressiveness and conservative behavior of pedestrians during crossing. The game parameters are presented in meters where applicable.

3.6.1 General 3-Player Game

The state trajectory plot highlight the trajectory of AV and interacting pedestrians.

Table 3.1: Game Parameters for 3-Player Interaction

Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-8, 0, 0, 2.5)	(0, 8, $-\pi/2$, 0.7)	(0, -8, $\pi/2$, 0.8)
Target State	(14, 0, 0, 3)	(1, -11, 0, 1.5)	(1, 11, 0, 1.5)
Brownian Motion Noise	0.1	0.1	0.1
Mean	1.2	1.2	1.2
Standard Deviation	1.3	1.3	1.3
Collision Radius (m)	2	1	1



(a) Player Trajectories Coupled with Speed (b) Player Trajectories

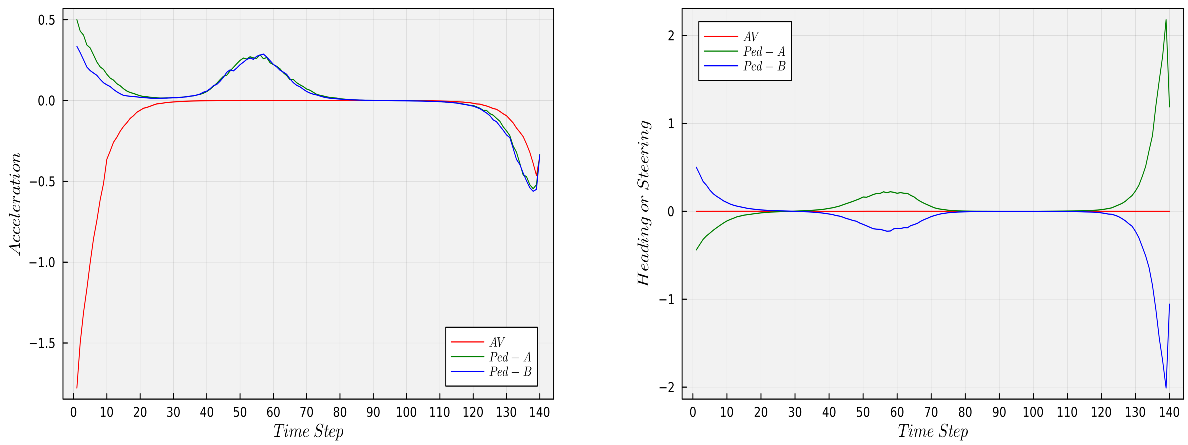
Figure 3.1: Three-Player Interaction Game Trajectory

The game trajectory describes the interaction between three players (one AV and two pedestrians over a crosswalk) (Fig 3.1). The interaction is sensitive to initial conditions as AV's higher speed and short distance from the crosswalk will likely encourage no major interaction event between AV and crossing pedestrians. Thus, the game is formulated in a way to promote an interaction between the players. The interaction dynamics of AV and pedestrians are governed by the transformed dynamics of the players and feedback strategy. The in-depth focus on the speed of the players in the figure gives an idea of the real-world interaction scenario. The control dependent noise is visible in the acceleration pattern of the players, specifically for pedestrians as they adjust their speed with respect to the initial condition and target position during crossing.

In Fig 3.3a the speed of the players at the endpoint of the game becomes identical as the players reach their destination within the game allocated time. The speed will not remain identical when the players can not reach their target state by the finite time horizon of the game.

Pedestrian Dilemma

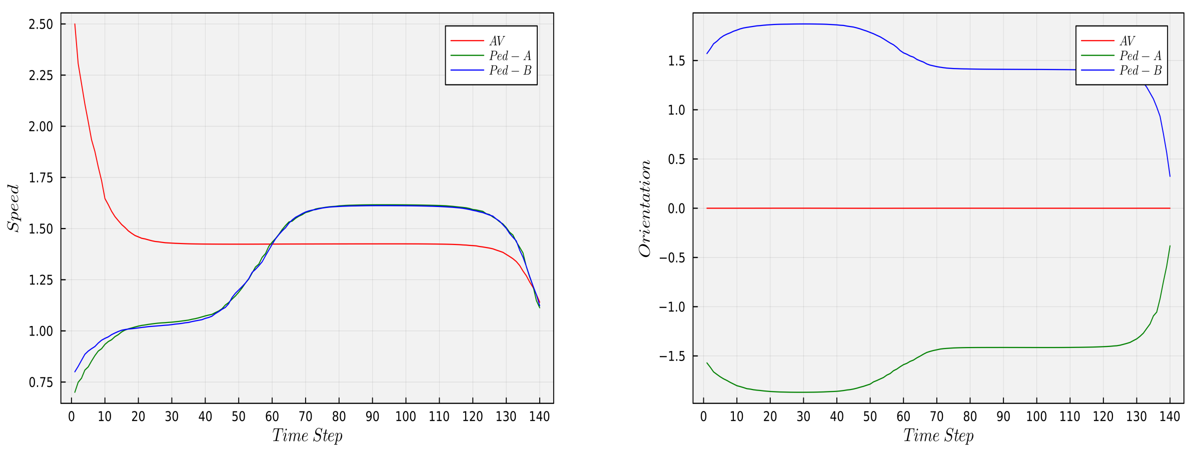
Pedestrian dilemma can be defined as a perception issue when the presence of vehicles results in a stage of confusion for the pedestrians either to walk fast to cross or slow down to let the vehicle



(a) Acceleration Control of Players

(b) Orientation Control of Players

Figure 3.2: Three-Player Interaction Game Controls



(a) Speed of Players

(b) Heading of Players

Figure 3.3: Three-Player Interaction Game States

pass in an interaction scenario. The primary reason for the existence of such dilemma zones is hinged on the perception level of pedestrians, since pedestrians attempt to assess/predict motorist behavior (aggressive or conservative) during crossing. Notably, the possibility of a dilemma zone is not limited to the intersection geometry, signal control setup, pedestrian volume, and the waiting time for turning vehicle; rather specific instances varying on the perception level of pedestrians and intent communication of motorist during the interaction. In case of vehicles approaching the crosswalk, the pedestrians predict the arrival of vehicle over the dilemma zone and decide to walk faster to cross or walk slowly to let the vehicle pass.

It is difficult to specify abstract point of the pedestrian dilemma as both AV and pedestrian know their respective location all the time. The stage of dilemma depends on the initial condition. The applied noise is control dependent and only affects in part on their speed progression. Given the premise a game with a small time horizon where the players are close to each other will reflect the pedestrian dilemma in the early stage of the crossing event. This implies that the pedestrian adjust their walking speed with respect to the incoming vehicle on the roadway.

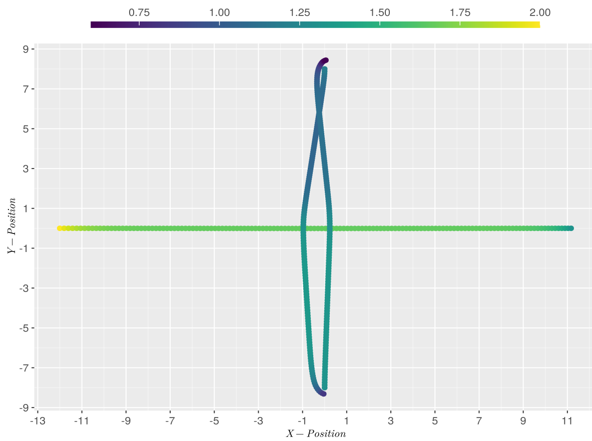
In the game environment the presence of dilemma can be traced from the crossing behavior of the pedestrians given they start from the same distance from both sides of the road (Fig 3.4). In this case the pedestrian from both sides of the crosswalk starts with the same initial condition. However, the Pedestrian-B crosses the roadway earlier than the AV and on the other hand Pedestrian-A crosses after the AV has passed. The speed plot of the players also reflect that Pedestrian-B speed up to cross the roadway and the AV slow down to facilitate the crossing. This concurs with the rationale that the pedestrians either speed up or walk slowly to facilitate the movement of the AV, when the AV is near the crosswalk. The game parameters are presented in meters where applicable.

Table 3.2: Game Parameters for Pedestrian Dilemma

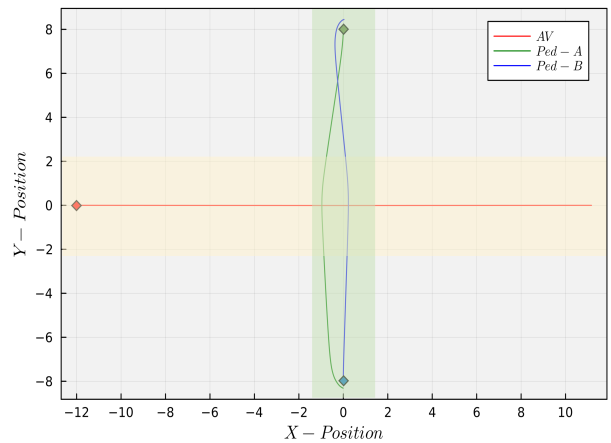
Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-12, 0, 0, 2)	(0, 8, $-\pi/2$, 1.2)	(0, -8, $\pi/2$, 1.3)
Target State	(14, 0, 0, 3.5)	(1, -11, 0, 1.5)	(1, 11, 0, 1.5)
Brownian Motion Noise	0.1	0.1	0.1
Mean	1.2	1.2	1.2
Standard Deviation	1.3	1.3	1.3
Collision Radius (m)	2	1	1

3.6.2 Safety Conservative Pedestrians

Safety conservative pedestrians interact with the AV starting from a safe speed and gradually accelerate once the AV has crossed the potential conflict point over the crosswalk (Fig 3.7). In the

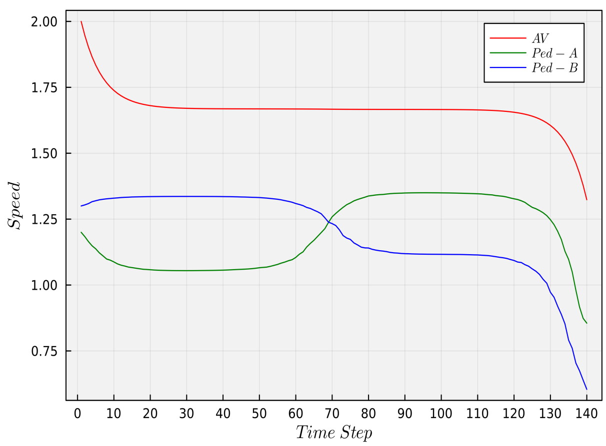


(a) Pedestrian Dilemma Trajectory with Speed

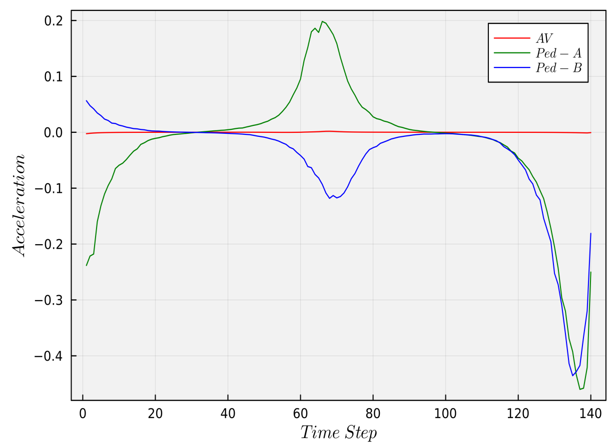


(b) Pedestrian Dilemma Trajectory

Figure 3.4: Three-Player Interaction Game with Pedestrian Dilemma



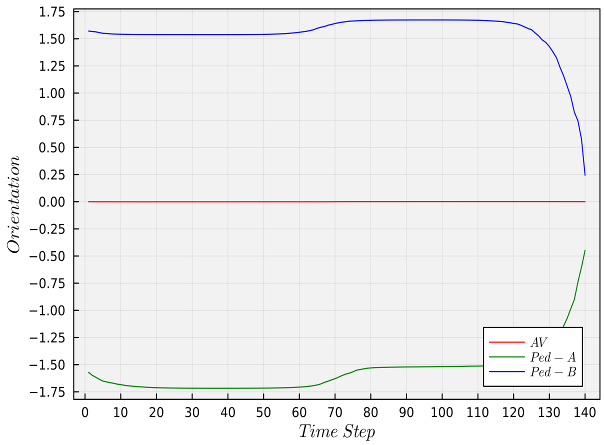
(a) Speed - Dilemma Condition



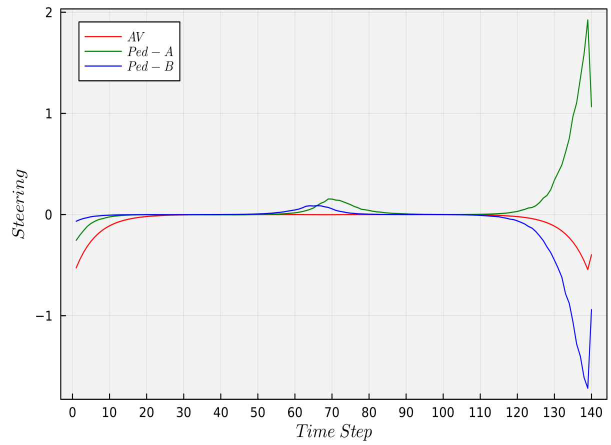
(b) Acceleration - Dilemma Condition

Figure 3.5: Three-Player Interaction Game with Pedestrian Dilemma

game scenario, the proactive sense of safety from pedestrians perspective is induced by maintaining a larger collision radius and assigning comparatively large weight on the collision cost component. The underlying behavior of the pedestrian remains same. However, due to the weighted cost component and preferred crossing speed, pedestrians behave more conservatively. They maintain a higher safe distance from other players including other pedestrians and AV.

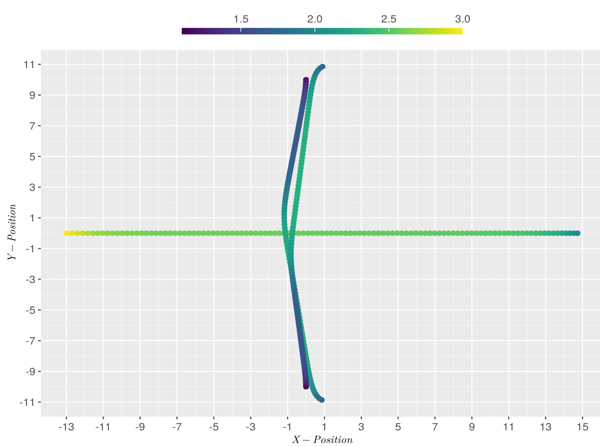


(a) Orientation - Dilemma Condition

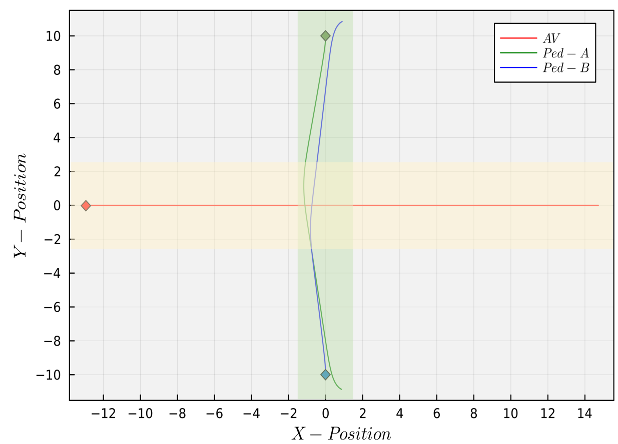


(b) Steering or Heading - Dilemma Condition

Figure 3.6: Three-Player Interaction Game with Pedestrian Dilemma



(a) Player Trajectories Coupled with Speed

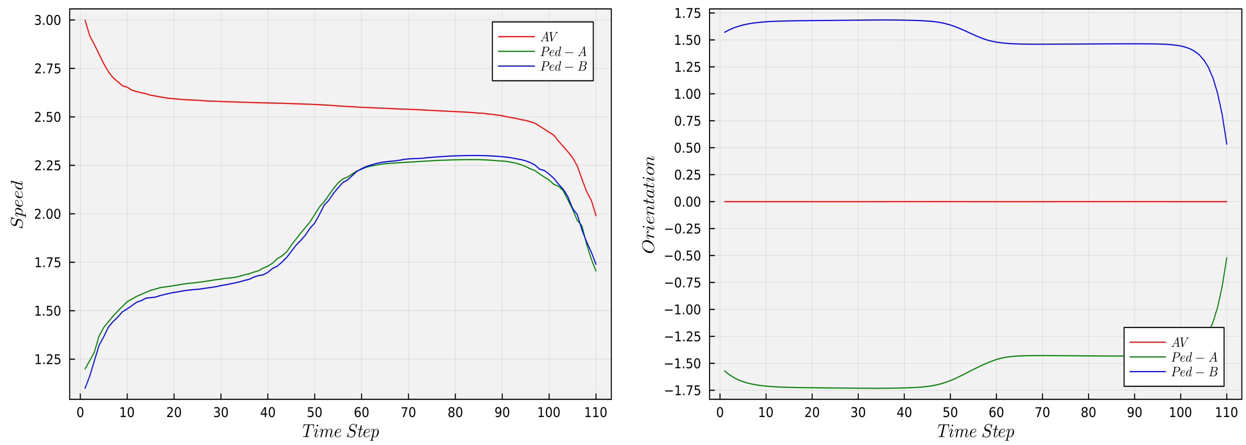


(b) Player Trajectories

Figure 3.7: 3-Player Game Trajectory with Safety Conservative Pedestrians

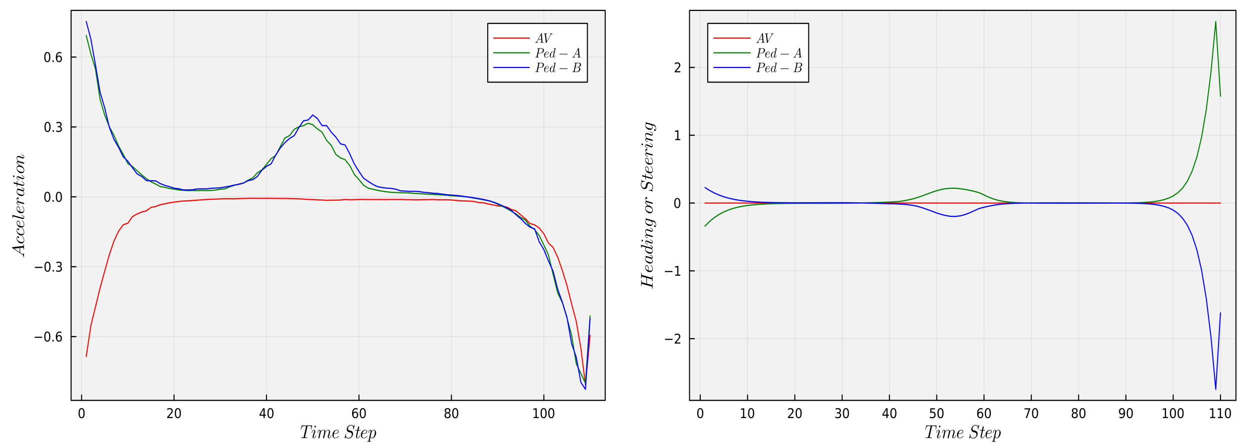
3.6.3 Aggressive Pedestrians

Aggressive pedestrians interact with the AV starting from a regular speed and accelerate before the AV reaches the potential conflict point over the crosswalk. In the game scenario, the aggressive pedestrian behavior is induced by maintaining a lower collision radius and assigning comparatively smaller weight on the collision cost component. The underlying behavior of the pedestrian remains same. However, due to the weighted cost component, preferred crossing speed, and assigned noise, pedestrians behave more aggressively. They maintain a lower safe distance from other pedestrians.



(a) Speed of Players (b) Heading of Players

Figure 3.8: 3-Player Game States with Safety Conservative Pedestrians



(a) Acceleration of Players (b) Steering or Movement Direction of Players

Figure 3.9: 3-Player Game Controls with Safety Conservative Pedestrians

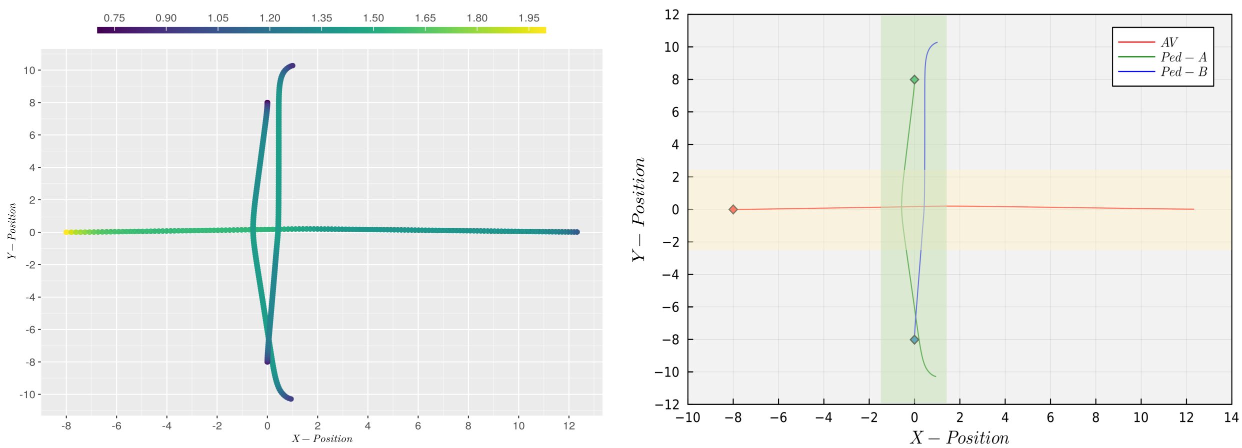
In this instances, AV adjust their speed with respect to the pedestrians and pass over the crosswalk once the pedestrians have crossed at a safe distance (Fig 3.10). The speed trajectory of the players show that the pedestrians speed up to cross the roadway from both direction.

3.6.4 Gap Maneuvering of AV

The gap behavior of AV represents a trade-off between the conservative safety concept and efficiency while interacting with pedestrians over a crosswalk in an urban setting. The concept is similar to

Table 3.3: Game Parameters for Aggressive Pedestrians

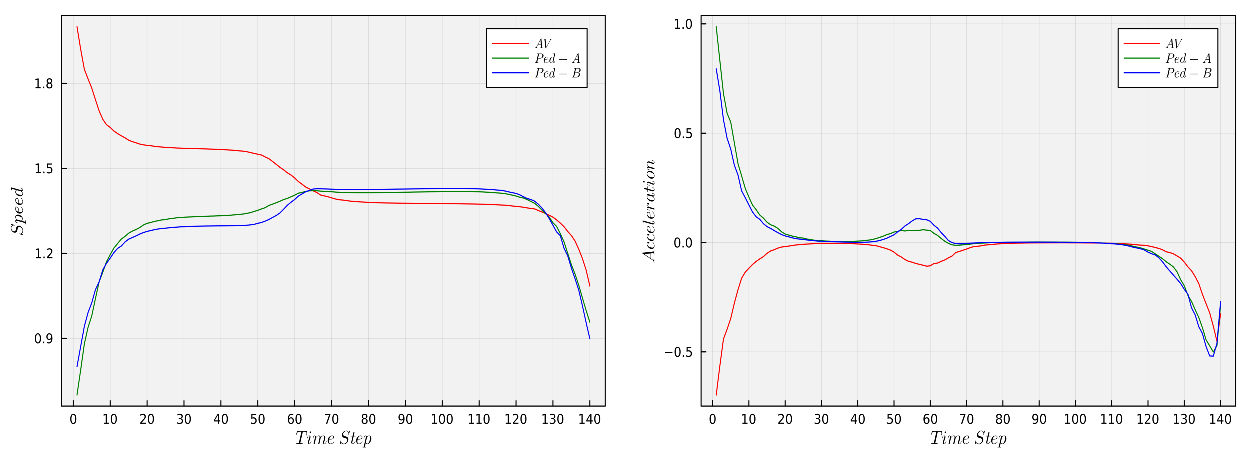
Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-8, 0, 0, 2)	(0, 8, -π/2, 0.7)	(0, -8, π/2, 0.8)
Target State	(14, 0, 0, 3.5)	(1, -11, 0, 1.5)	(1, 11, 0, 1.5)
Brownian Motion Noise	0.1	0.1	0.1
Mean	1.2	1.2	1.2
Standard Deviation	1.3	1.3	1.3
Collision Radius(m)	2	0.75	0.75



(a) Player Trajectories Coupled with Speed

(b) Player Trajectories

Figure 3.10: 3-Player Game Trajectory with Aggressive Pedestrians



(a) Speed of Players

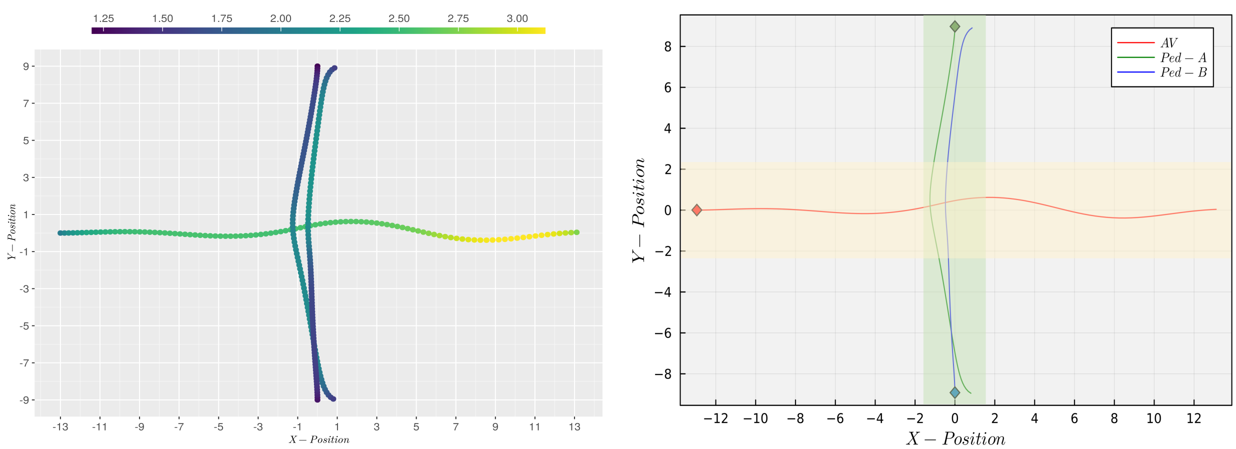
(b) Acceleration of Players

Figure 3.11: 3-Player Game Speed and Acceleration with Aggressive Pedestrians

the pedestrian gap acceptance behavior, except the length, mass, speed, and acceleration of the vehicle significantly differ than that of pedestrians. Based on the acceleration pattern coupled with the comfort and eco-driving component, the gap behavior of AV will vary for different scenario engendered from bi-directional pedestrians/groups. If the time headway between the bi-directional pedestrian movement is adequate, the AV will negotiate through the gap/conflict zone to complete the turning maneuver (Fig 3.12).

Table 3.4: Game Parameters for Gap Maneuvering of AV

Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-13, 0, 0, 2)	(0, 9, $-\pi/2$, 1.2)	(0, -9, $\pi/2$, 1.3)
Target State	(14, 0, 0, 4)	(1, -10, 0, 1.7)	(1, 10, 0, 1.7)
Brownian Motion Noise	0.1	0.1	0.1
Mean	1.2	1.2	1.2
Standard Deviation	1.3	1.3	1.3
Collision Radius (m)	1.5	0.85	0.85

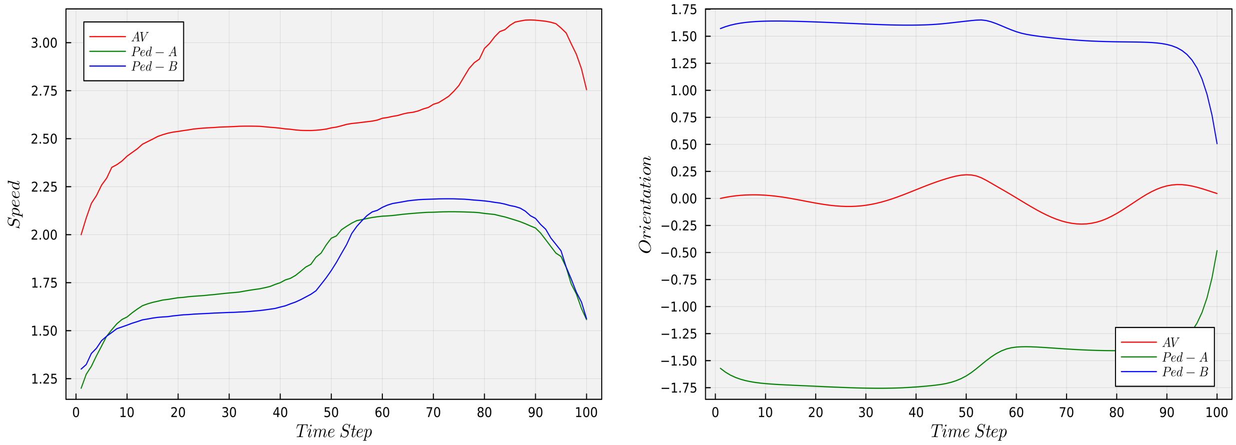


(a) Player Trajectories Coupled with Speed (b) Player Trajectories with Roadway Boundary

Figure 3.12: Game Trajectory of the Gap Maneuvering AV

3.6.5 Pedestrian Group Effect

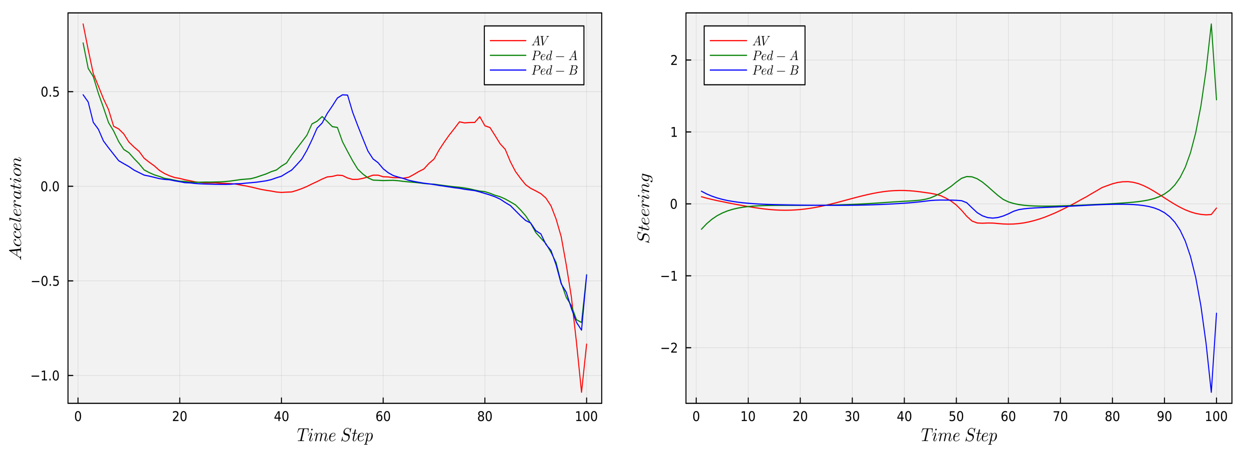
To test the group effect of the pedestrians over a crosswalk, we change the initial points from the pedestrians from one side of the crosswalk. Based on the same initial condition the pedestrian group may act differently. They can either cross before the AV or slow down and let the AV pass before crossing the roadway.



(a) Speed of Players

(b) Orientation of Players

Figure 3.13: Game States of the Gap Maneuvering AV



(a) Acceleration of Players

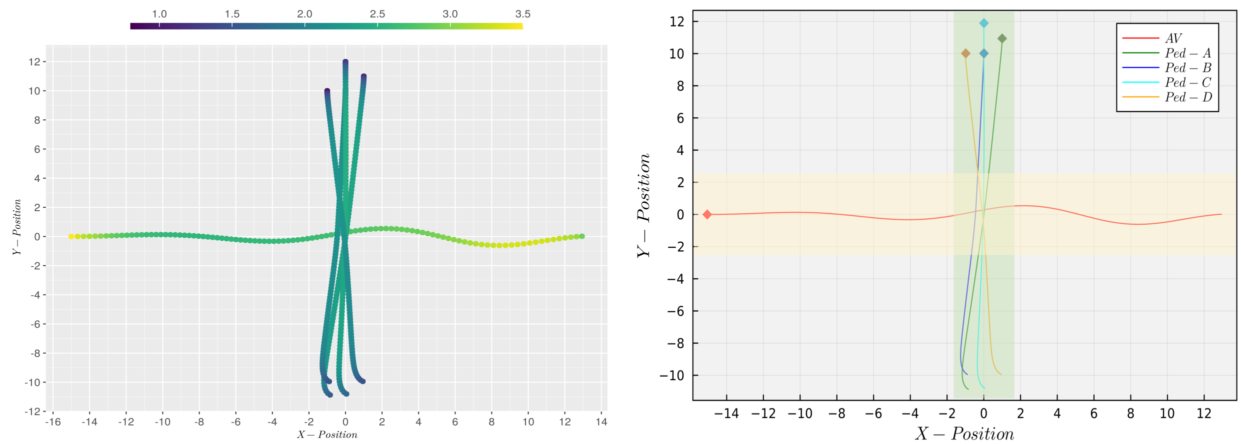
(b) Steering of Players

Figure 3.14: Game Controls of the Gap Maneuvering AV

Table 3.5: Game Parameters for Pedestrian Group Effect

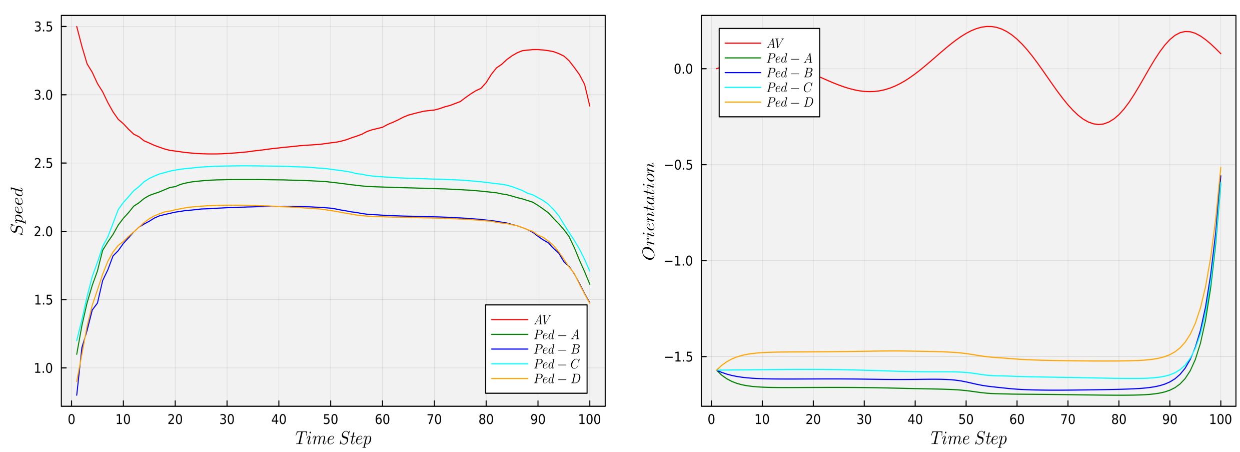
Parameters	AV	Pedestrian-A	Pedestrian-B	Pedestrian-C	Pedestrian-D
Initial State	(-15, 0, 0, 3.5)	(1, 11, $-\pi/2$, 1.2)	(0, 10, $-\pi/2$, 0.9)	(0, 12, $-\pi/2$, 1.2)	(-1, 10, $-\pi/2$, 0.9)
Target State	(14, 0, 0, 1.7)	(1, -10, 0, 1.7)	(1, 10, 0, 1.7)	(1, -10, 0, 1.7)	(1, 10, 0, 1.7)
Brownian Noise	0.1	0.1	0.1	0.1	0.1
Mean	1	1	1	1	1
Standard Deviation	1.2	1.2	1.2	1.2	1.2
Collision Radius (m)	1.5	0.75	0.75	0.75	0.75

The control trajectory defines the acceleration and steering for the AV and describes the pedestrian acceleration and heading for the proposed game framework. The pedestrian group cross the crosswalk first. Whereas, the AV reduces the speed and maneuvers its way ahead giving adequate safe space for the pedestrians to cross. The specific interaction happens in between the time $50-53$ seconds. The speed trajectory plot also shows that after the interaction time step the AV speeds up and the crossing pedestrians slow down to reach towards the target points (Fig 3.15).



(a) Player Trajectories Coupled with Speed (b) Player Trajectories

Figure 3.15: Game Trajectory with the Group of Aggressive Pedestrians

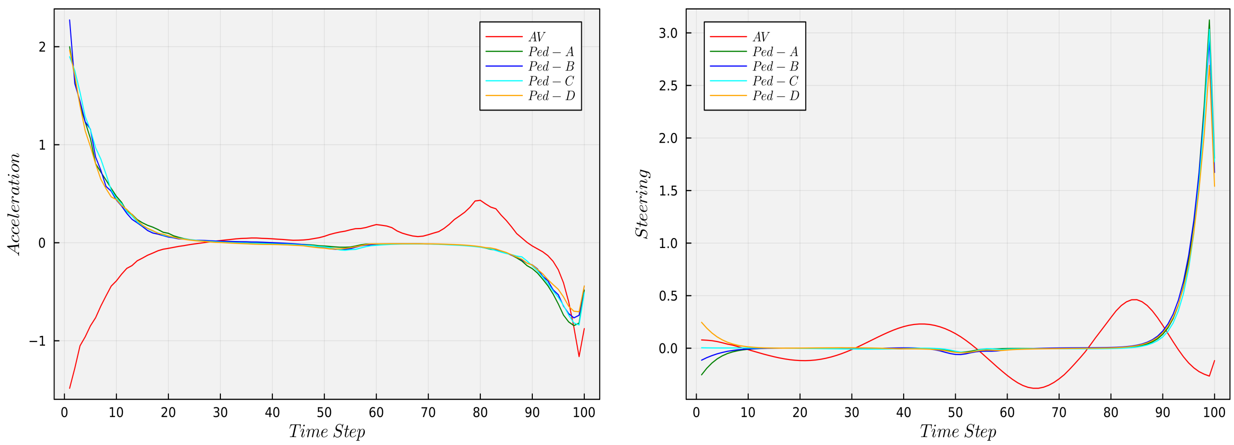


(a) Speed of Players (b) Orientation of Players

Figure 3.16: Game States of Aggressive Pedestrian Group

3.7 Quantitative and Qualitative Evaluation

A good measure for any model is to record and report the concerned agents or players' possible deviation or tracking error. This task may involve recording a trajectory from a real-world scenario



(a) Acceleration of Players

(b) Steering of Players

Figure 3.17: Game Controls of Aggressive Pedestrian Group

and building a model to replicate the trajectory with a possible bound for error. However, the real-world pedestrian reflects varying behavior that cannot be translated to one specific trajectory. For instance, our observation suggests that pedestrians can be conservative or aggressive during the crossing. Thus variation in the behavior can result in a very different set of trajectories. In our effort to evaluate the proposed game framework, we answer and rationalize two primary research queries as follows.

Why focus on the pedestrian motion for evaluating the proposed game framework?

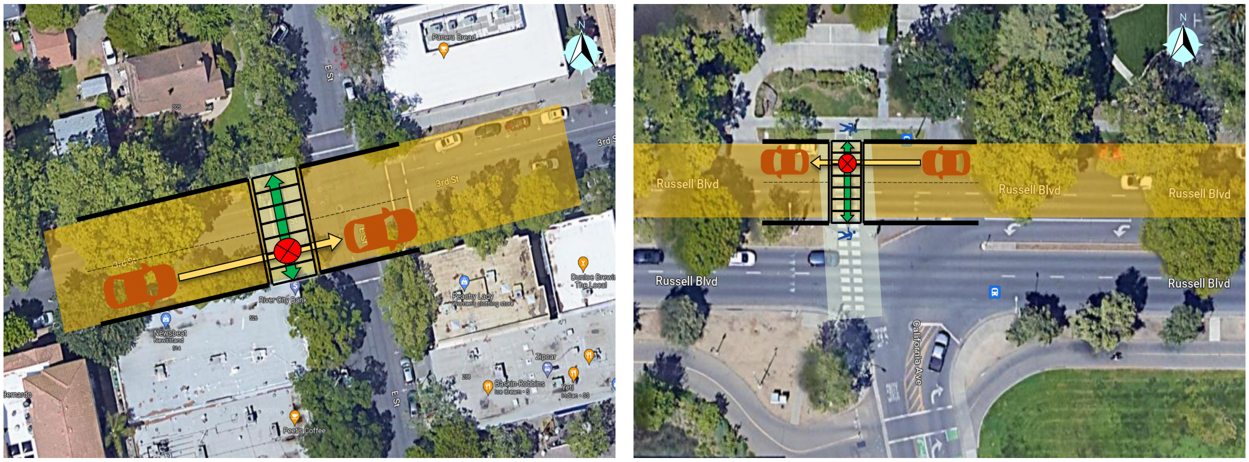
The proposed game framework's efficacy depends on accurately simulating the pedestrians' motion and behavior. In other words, the proposed game model for autonomous driving will be a practical feature when it can comprehend real-world pedestrian behavior during an interaction at crosswalks. Thus, we evaluate the model using pedestrian crossing time, interaction time, and motion patterns.

Why use the average crossing and interaction time of pedestrians for quantitative evaluation?

Detailed ground truth data is unavailable from an aerial perspective to analyze and test AV's motion planning and pedestrians' crossing behavior in a game framework. Moreover, the stochastic nature of the players embedded in the game makes it complex to set up an exact testing scenario. We tested several video cameras and drone data to extract meaningful information to evaluate the framework. However, the camera angle, occlusion, uncertainty of the interaction period, and low flying time of drones made it harder to extract any meaningful interaction information. Thus, we employ pedestrians' average crossing and interaction time as a quantitative reference to the real-world interaction scenario.

3.7.1 Quantitative Evaluation

This section reports the quantitative evaluation of the proposed interaction game framework for autonomous driving. The quantitative evaluation consists of two real-world scenarios: (1) a stop sign intersection; (2) a mid-block crosswalk. The stop sign intersection is downtown (3rd Street and E Street) (Fig 3.18), and the mid-block crosswalk is at the corner of Russell Boulevard and California Avenue in Davis, California (Fig 3.18). The preferred downtown location accommodates high pedestrian traffic on weekdays and weekends. The mid-block crosswalk has similar pedestrian traffic throughout the day and includes students and university staff nearby. For both scenarios, we count the crossing and possible interaction time of each pedestrian and record the type of behavior (cooperative, competitive) exhibited by them. The observation period was ten hours on different weekdays and weekends in May for multiple days to gather a representative sample of the crossing pedestrians. Notably, we only record the pedestrians crossing and interaction time over the crosswalk and their behavior with respect to the presence of a vehicle. The average crosswalk length is measured approximately 42 feet for the data collection locations.



(a) 3rd Street & E Street (b) Russell Blvd & California Avenue

Figure 3.18: Field Data Collection Locations at Davis, California

We measure the pedestrian crossing with two observers on each side of the crosswalk, one recording the crossing time and the other recording the probable pedestrian type and interaction time. The interaction time in the field is defined as the vehicle approaches the crosswalk and the pedestrians negotiate using a visual cue. For multiple pedestrians, we measure the incremental time each pedestrian took starting from the first crossing pedestrian. For instance, if the first pedestrian crosses the crosswalk in 10 seconds, we compute the other pedestrians crossing incrementally or add to the 10 seconds.

The cooperative behavior reflects the pedestrians crossing over the crosswalk while negotiating space with the vehicle using visual cues and waiting for the turn to cross. On the other hand, in the competitive case, the pedestrians compete for the crosswalk and walk fast to cross before the vehicle. This behavior-type assessment consists of real-world data from our study and the publicly available JAAD (Joint Attention in Autonomous Driving) dataset [79]. We explored the annotated JAAD dataset to understand pedestrian motion patterns and behavioral features in the context of autonomous driving in urban scenarios. This dataset comprises dashboard footage of the pedestrian-vehicle interaction in various scenarios, which is more suitable for understanding pedestrian behavior in urban streets. We observed the annotated video data to understand the pedestrian type and record similar behavior as cooperative or aggressive at the data collection sites. For computation, we divide the pedestrians into two behavioral groups and calculate the average crossing time concerning the number of pedestrians in each group.

Understanding the type and time of interaction is a complex and subjective task as it relies heavily on the observer's interpretation of the interaction event. Moreover, the game condition and real-world scenario may vary depending on the motorists' (drivers) and pedestrians' communication and behavior. In this case, we are documenting an interaction between a regular motorist and pedestrians in the real world. In contrast, the interaction between AV and pedestrians in urban scenarios may vary depending on the communication medium, pedestrian behavior, and driving style. For example, the AV can sense the presence of incoming pedestrians and may slowly approach the crosswalk to give enough time for pedestrians to cross and save energy without stopping. In some instances, it may speed up to cross before the pedestrian starts to walk near the crosswalk. These behaviors are typically absent in regular motorists as the sensor range is limited and subjective to the person driving the vehicle. Thus, traditional stop-interact-go situations may change with the adoption of AV in urban interaction scenarios. They are designed with a wide array of sensors, object tracking, motion planning, prediction, and other features to interact safely with other road users. Considering the uniqueness of the possible interaction scene and behavioral assumptions, the dynamic game framework is fitting for forming the interaction problem as a stochastic game.

The quantitative analysis table has multiple fields to compare real-world and game-simulated data (Table 3.6). First, we record the pedestrians for different scenarios from the study locations. Second, we record the type of behavior observed in the field. Third, we gather representative samples to compute the average crossing and interaction times for single, multiple, and group

Table 3.6: Quantitative Evaluation Summary

		Avg Crossing (s)				Avg Simulated (s)		Difference (s)	
Scenario	Sample	Mix	Con	Agr	Int	Crs	Int	Crs	Int
A (1)	16	10.331	10.750	9.075	2.994	10.251	3.127	0.080	0.133
B (2)	10	10.380	10.539	8.950	3.035	10.335	3.215	0.045	0.180
C (3)	7	10.519	10.506	9.200	3.048	10.563	3.154	0.144	0.106
D (4)	6	10.758	10.810	9.233	3.079	10.765	3.225	0.207	0.146
E (5)	5	10.972	11.068	9.133	3.092	11.285	3.286	0.413	0.194

*Avg - Average, s - Seconds, Mix - Mixed, Con - Conservative, Agr - Aggressive, Int - Interaction, Crs - Crossing.

interaction scenarios. The observers used stopwatches to record crossing time and interaction time. Notably, the interaction time was much shorter than the crossing time for all the scenarios.

To simulate the average crossing time for each representative scenario, we ran each game model 20 times. Each run produced a slightly different crossing time due to stochasticity. Afterward, we compute the average crossing and interaction time from the simulated trajectories and estimate the difference from real-world data.

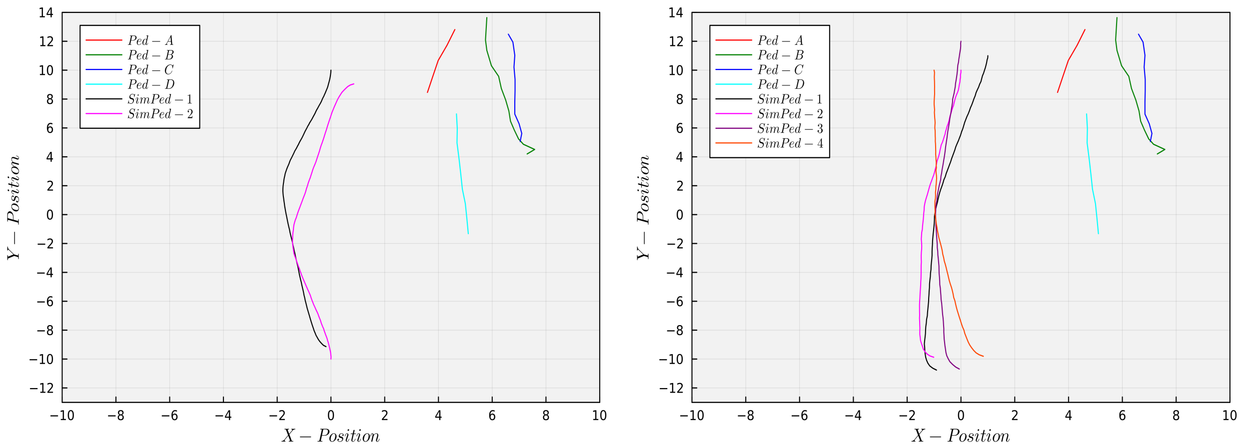
In the field, the crossing time of each group increases with the number of pedestrians as we add the crossing time incrementally after the first crossing pedestrian. In such cases, although the crossing time is similar for all the pedestrians in a group, the varying starting time of the pedestrians results in an incremental increase in the crossing time. This implies that pedestrians crossing last in a group starts later than the first crossing pedestrian and adds to the total time recorded in the field. However, in the game, we can pinpoint the exact crossing time of each pedestrian. Thus to draw an exact comparison, we also add a cumulative weightage to the game-simulated crossing time. The weightage is computed from the increased crossing time in the field data by adding pedestrians in scenarios A, B, C, D, and E. For instance, the increase in the average crossing time for a group of 4 pedestrians is about 4.13% from the average crossing time of one pedestrian in Scenario A. Then, to draw an exact comparison, we multiply our game-simulated average crossing time for four pedestrians with 1.0413, the cumulative weightage in this research.

The results indicate that the average crossing and interaction time for each simulated scenario closely matches real-world data. For instance, the difference between simulated and real-world crossing time for two pedestrians crossing from the opposite direction is 0.045 seconds. For the four pedestrians crossing scenario, the difference between average crossing time is about 0.207 seconds. The interaction time difference for multiple pedestrian scenarios (A, B, C, D, E) ranges from 0.133 to 0.194 seconds (Table 3.6).The maximum difference in average crossing time (0.413) is recorded for scenario E with 5 pedestrians crossing as a group. Notably, the aggressive behavior is recorded mostly from location 1 (3rd Street & E Street, Davis) because location 2 (Russell Blvd & California Ave) has a flashing light to warn the approaching vehicle near the mid-block crosswalk. Thus pedestrians feel safe and walk at a regular pace rather than walking fast to cross before the vehicle. The complete dataset recorded from the field is provided in the appendix.

3.7.2 Qualitative Evaluation

The qualitative evaluation visually compares pedestrian trajectories from public datasets with the game simulated trajectories. We explore publicly available datasets with recorded pedestrian features, including motion patterns. These datasets represent a collection of scenarios and annotations to explain pedestrian motion and intention. We investigate the motion pattern of pedestrians across the datasets and compare it with the game-simulated trajectories to determine the efficacy of the proposed framework.

We explored the annotated ETH ([80]) and Stanford drone dataset ([81]) to capture the motion pattern from real-world scenarios and compare it with simulated trajectories. The ETH dataset is a representative sample for pedestrian detection in urban scenarios, and the drone dataset records the students' motion patterns and interactions within the Stanford University campus. However, we could only use the ETH data for qualitative comparison as the drone data has detection and occlusion issues.



(a) Pedestrian Trajectories from 3-player Game (b) Pedestrian Trajectories from 5-player Game

Figure 3.19: Qualitative Evaluation of the Simulated Trajectories with respect to ETH data

The simulated pedestrian trajectories visually match the motion pattern of the real-world data with some differences (Fig 3.19). The visual difference with the real-world trajectories occurs as the pedestrians are not walking in a controlled crosswalk environment in the presence of a vehicle nearby. Moreover, the simulated trajectories have higher time resolution than the real-world data. The time interval between each data point may also contribute to some of the sharp edges in the pedestrian trajectories. Notably, the simulated pedestrian trajectories match the forecasted trajectories in other established studies in the pedestrian motion research domain ([82], [83]).

3.8 Conclusion

The proposed game simplifies a complex problem to describe specific features of the interaction scenario between pedestrians and AV in urban streets. The game investigates Brownian Motion noise on the control variable of the players and handles a specific noise. We use a game theoretic framework to detail pedestrian crossing behavior in various formats, including conservative, aggressive, and group. In turn, the pedestrians' behavioral features are part of the interaction framework for AV. The algorithm finds local approximations in a multiplayer stochastic dynamic or differential game setting and closely resembles the architecture of the iterative Linear Quadratic Regulator approach. The game solutions exhibit the competitive and cooperative behavior associated with the Nash Equilibrium for an interaction event.

Although the deterministic iterative game example is there for robotic movement following receding horizon methodology, only the proposed study investigates the specifics of a pedestrian crossing in urban streets considering several behavioral and movement features. The proposed interaction game model incorporates noise into the control variable to reflect the real-world stochastic condition. Our intention for the study is to capture some interesting observations from the pedestrian crossing behavior and design a game theoretic system for the AV driving system that accommodates such features.

It is difficult to compare the results to a real-time crossing event as the game's random noise and behavioral features will continue to change for each run of the proposed game framework. The error comparison between the real-time pedestrian trajectory (ground truth) will continue to differ as the same initial condition may change the game outcome due to noise on the control variables. However, applying noise only on pedestrian controls and testing the outcome of the interaction game is our future research direction. Notably, the qualitative comparison of other algorithms is difficult due to the uniqueness and setup of the equilibrium concepts, information-sharing framework, and other behavioral features that govern the players' costs. Computational cost can also be an issue, as the study explores the behavioral features of pedestrians rather than algorithmic efficiency. Considering all these limitations, we use a quantitative (average crossing time & interaction time) and qualitative metric (pedestrian motion pattern) to evaluate the proposed game model. The results described in the previous section indicate that the simulated pedestrian trajectory matches the real-world pedestrian movement regarding crossing time, interaction time, and motion pattern.

The proposed interaction framework based on dynamic games is sensitive to the initial conditions. This sensitivity is derived from the focused interaction between pedestrians and AV over the crosswalk. Although the game is set up in a finite horizon, the interaction happens only for a few seconds. The immediate information available to each player helps to identify a safe and optimal control strategy for any given interaction scenario. Notably, the feedback strategy influences the game setting, which means that every player knows each other state at all times.

The type of noise in the game system also plays a critical role in the robustness of the proposed framework. Although the developed game model can handle a specific noise range in dynamics, the model is not robust against external noise in urban settings. However, we can build a robust game system if we can assume the range of possible noise and insert it deterministically into the system. The following research approach will explore such robust feedback Nash Equilibrium for the proposed interaction game.

Chapter 4

Robust Game Theoretic Interaction Framework for Autonomous Vehicles

4.1 Introduction

Dynamic game theory provides a robust and rational framework for designing and understanding the strategic interactions among multiple decision-makers known as players. Game Theory is interesting because it can incorporate the players' behavior, access to information, and limited rationality in a dynamic environment. Moreover, it can accommodate uncertainties and external disturbances impacting players' strategies and cost function. Based on the game design, researchers have developed different ideas of equilibrium, including Nash, Pareto, and Stackleberg equilibrium. Notably, Nash equilibrium fits into the solution perspective of a non-cooperative differential game with a feedback strategy for the players. However, traditional Nash equilibrium does not accommodate the robustness of feedback strategies in a system with perturbations and facilitates uncertainties in the player cost function. The Robust Feedback Nash Equilibrium (RFNE) considers the robustness of strategies to deviations from equilibrium, feedback control mechanisms, and disturbances.

The algorithm part builds on the local approximation, backward pass, and line search for convergence. The study's uniqueness is combining two elements and developing a robust game theoretic framework for Autonomous vehicle interaction with pedestrians in urban streets. The stochastic component for the interaction framework builds on the system noise or the finite disturbance and the adversarial input in the player cost function.

In the previous chapter, we presented a stochastic linear quadratic game framework that can handle multiplicative noise on the control variables of the player. We want to further extend the game design following the safe systems approach. The concept of a safe system approach in transportation stems from the design perspective, where the roadways and vehicle features are designed considering human driving error. Following the safe system principle, we want to design a game framework that accommodates a range of system noise and incorporates the worst-case scenario for each player involved in the game. In short, the proposed framework in this chapter will accommodate certain disturbances and work towards mitigating the effect of such external uncertainties. The proposed method is partly analogous to the risk-sensitive iterative approach in the linear quadratic regulator literature domain.

This study assumes that the players can access their state information in a closed-loop system. We explore a robust feedback control strategy that follows the Nash Equilibrium, where the players consider the worst possible disturbance scenario. To solve such a game system, we introduce a robust form of the Hamilton-Jacobi-Bellman equation for each player, which computes the minimum of the controls and the maximum of the uncertainty. The flexibility of the proposed robust game allows the modeling of each type of player with different types of noise-based systems. Moreover, the robust game filters mitigate the external noise effects during an interaction event. As such, the robust game framework allows us to model the interaction with different noise scales for both players, given that we want a smooth trajectory for the AV and noisy motion patterns from pedestrians to reflect real-world scenarios.

4.1.1 Robust Control Literature

Modeling uncertainties are critical in developing a stable system for vehicle control, especially for autonomous driving. From an engineering perspective, the uncertainties in vehicle dynamics and surrounding environment noise can be time-varying, bounded, and follow certain distributions [84]. Robust control captures the model uncertainties. It is an essential feature in control theory that is applied to uncertain or noisy systems to capture worst-case scenarios in a system. It is also available in the game literature where the game design incorporates one or multiple control systems that must be robust for specific safety-critical applications. For instance, the altitude control of aerial vehicles [85], chemical reactions, and precision tracking systems [86] are examples of the robust game application.

Risk-sensitive or robust application for the zero-sum differential game with the infinite horizon is available in the literature. The benefit of studying the infinite horizon is that the final cost is not included in the cost function, and the overall cost is discounted over the horizon. In one approach, the authors established values and saddle-point equilibria for the formulated game with risk-sensitive payoff considering discounted and ergodic criteria. The study method focused on analytical expressions and proofs without numerical application [87]. Another study focused on the robust control of continuous-time singularly perturbed bi-linear quadratic systems with an additive disturbance. The uncertainty of the system acted as a player, transforming the control system into a two-person zero-sum game model. Finally, this study explored the saddle-point equilibrium for a chemical reactor model to verify the algorithm [88].

Regarding the non-zero-sum game form, an adaptive critic approach was proposed to approximate the online Nash equilibrium solution for the robust trajectory tracking control for continuous-time uncertain nonlinear systems. The proposed system incorporated a combination of the tracking error and the reference trajectory. It applied one critic neural network (NN) for individual players to approximate the coupled Hamilton–Jacobi–Bellman (HJB) equations and compute the feedback controls for the Nash equilibrium solution. The system was tweaked using the NN weight updates and verified by two numerical examples [86].

Engwerda presented the open-loop application of the robust Nash Equilibrium in a finite time setting and non-cooperative formulation [89, 90]. Later, others tried to expand the feedback game system with basic applications and simple numerical examples [91, 92]. Researchers also applied a variation of this method for robust event-triggered altitude control games for satellites traversing the earth's orbit. They studied the disturbance as a player in the game with a coordinated micro-satellite. The altitude regulation problem is formed into a robust differential game so each player (micro-satellite) can compute the worst control strategy to counter the time-varying disturbance with the finite distribution [85].

Robust control for autonomous driving applications is studied for lane change maneuvers in a hierarchical game framework for a highway driving scene [84]. The study developed a robust decision-making strategy to capture model mismatches in the prediction model. At the same instance, the model used the confidence of the driver's behavior to obtain less conservative actions for AV. Similarly, a recent study tested robots with a robust framework to understand pedestrian motion by combining robust control and intent-driven human modeling input to formulate a novel human motion predictor. The proposed approach tried to predict human states by trusting the intent-driven model to decide unlikely human actions and safeguard against all other possible actions in the form of a robust control predictor [93].

4.2 Contributions

The primary contributions of this chapter are:

- We define a robust nonzero-sum dynamic linear quadratic game framework for AV interaction to allow for a finite disturbance in the player cost function and noise in the system dynamics. The disturbance in the cost function tries to maximize the overall cost as an adversarial input, whereas the system noise perturbs the system dynamics.
- We derive and describe the model-based setting to compute the Robust Nash Equilibrium feedback strategies and value functions for the AV interaction framework.
- We provide numerical experimentation based on the AV interaction framework to illustrate the proposed algorithm's effectiveness in capturing and filtering uncertainty representations.

4.3 Game Elements

To understand the process and the robust game design lets start with a description of the dynamic system. We will delve into the specific AV and pedestrian dynamics game later in the chapter. We aim to define the general robust form of the non-cooperative differential game in this section by deriving inspiration from several studies ([94, 89, 91, 92, 85]). The general nonlinear dynamic system in continuous time is as follows,

$$x_{t+1} = f(x_t, u_t^1, .., u_t^N, \omega_t), \quad \forall t \in [t_0, T] \tag{4.3.1}$$

where $x_t \in \mathbb{R}$ represents state column vector of the game and $u_t^i \in \mathbb{R}$ reflects the control strategy at time t for each player i. Here i represents the number of players ($i = 1, 2, .., N$), $\hat{i}$ represents the combination of players other than i, and t_0 is the initial time. The noise or disturbance is reflected by ω_t. Here, ω is a finite unknown or external disturbance and square integrable ($\omega \in L^2[0,T]$). The general cost function or performance index for each players can be written as,

$$J^i(x_t, u_t, u_t^{\hat{i}}, \omega_t) = \sum_{t=0}^{T-1} c^i(x_t, u_t^i, u_t^{\hat{i}}, \omega_t) + h_T^i \tag{4.3.2}$$

where, c^i is the running cost and h^i is the terminal cost. Notably, the control region is a subset of $\mathbb{R}$. The formulation dictates that f, c^i, and h^i are such that for all $(u_t^i, u_t^{\hat{i}}, \omega_t) \in U_{adm}^i \times U_{adm}^{\hat{i}} \times L^2[0,T]$. Here U_{adm} represents all the admissible control strategies, U_{adm}^i control strategies for player i, and $U_{adm}^{\hat{i}}$ defines the control strategies of all other players except i. The theoretical premise is that the dynamics (x_{t+1}) admits an unique solution given that the cost function J_t^i is well defined.

Now, we explore the inclusion of worst case uncertainty from the i-th player's perspective. In this case the controls u^j are distributed according to the players other than the i-th player, where $j \in 1, ..., N$.

$$J^i(x_t, u_t^i, u_t^{\hat{i}}, \omega_t^{i*}) = \max_{\omega \in L^2[t_0,T]} J^i(x_t, u_t^i, u_t^{\hat{i}}, \omega_t) \tag{4.3.3}$$

Considering the premise of Nash Equilibrium and the existence of uncertainty function ω_t^{i*} for the i-th player, the control strategies $(u_{t,r}^1, u_{t,r}^2, ..., u_{t,r}^N)$ follow robust feedback equilibrium, where $(u_{t,r}^i, ..) \in U_{adm}$. In the form of inequality, this can be written as,

$$J^i(x_t, u_{t,r}^i, u_{t,r}^{\hat{i}}(t), \omega_i^*) \leq J^i(x_t, u_t^i, u_{t,r}^{\hat{i}}, \omega_t^{i*}) \tag{4.3.4}$$

Following these conditions it can be stated that $(u_t^{1,r}, u_t^{2,r}, ..., u_t^{N,r})$ is a vector of robust feedback Nash equilibrium strategies for all players. To understand and solve the robust interaction game we need to define the traditional value function and modify the contents according to the robust cost function from the previous section. Considering the control strategies $(u_t^1, u_t^2, ..., u_t^N)$ we can define the basic value function for player i is as follows,

$$V^i(x_t) = \min_{u^i \in U^i} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega^{i*}_{t,u^i,u_t^{\hat{i},r}}) \quad \forall i \in 1, 2, ..., N \tag{4.3.5}$$

Using the Bellman principle of optimality for the cost J^i, the value function definition can be defined as,

$$\begin{aligned} V^i(x_t) \leq J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega^{i*}_{t,u^i,u_t^{\hat{i},r}}) &= \max_{\omega \in L^2[t,T]} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega) \\ &= \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + \sum_{\hat{t}}^{T} J^i(x_t, u_t^{i,r}, u_t^{\hat{i},r}, \omega_t) \right\} \\ &= \max_{\omega \in L^2[t,T]} \left\{ \sum_{\hat{t}}^{T} J^i(x_t, u_t^{i,r}, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \end{aligned} \tag{4.3.6}$$

where, the control $u_{t,r}^{\hat{i}}$ represents the robust Nash strategies and $x_{\hat{t}}$ fulfills the general dynamics equation x_{t+1} for other players except i, and ω^* is the uncertainty function described earlier. Now taking the minimum over the control u_t^i yields to,

$$V^i(x_t) \leq \min_{u_t^i \in U^i} \max_{\omega \in L^2[t,T]} \left\{ \sum_{\hat{t}}^{T} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \tag{4.3.7}$$

If we consider that for any small deviation $\Delta_s > 0$, there is an admissible control $(u_t^{i,\Delta_s} \in U^i_{adm})$ with the following characteristics,

$$V^i(x_t) + \Delta_s \geq \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^{i,\Delta_s}, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \tag{4.3.8}$$

which dictates that $x_t^{\Delta_s}$ is the solution of the dynamics x_{t+1} with respect to the control u_t^{i,Δ_s} by maintaining the other players in a fixed position. Otherwise, if there exist a $\Delta_s > 0$ for any $u_t^i \in U_{adm}^i$ we would have the following inequality.

$$V^i(x_t) + \Delta_s < \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \tag{4.3.9}$$

Afterwards, if we consider the Bellman principle of optimality and control strategy $u_t^i = u_t^{i,r}$, the inequality arrives at the following contradiction.

$$\begin{aligned} V^i(x_t) + \Delta_s &< \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^{i,r}, u_t^{\hat{i},r}, \omega_t) + V_{\hat{t}}^i(x_{\hat{t}}) \right\} \\ &\leq \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^{i,r}, u_t^{\hat{i},r}, \omega_t) + \sum_{\hat{t}}^{T} J^i(x_t, u_t^{i,r}, u_t^{\hat{i},r}, \omega_t^*) \right\} \\ &= V^i(x_t) \end{aligned} \tag{4.3.10}$$

Therefore, from the previous inequality equation 4.3.8 we arrive,

$$V^i(x_t) + \Delta_s \geq \min_{u_t^i \in U^i} \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \tag{4.3.11}$$

Now, since the inequality 4.3.11 has $\Delta_s > 0$ and arbitrary, we can write,

$$V^i(x_t) \geq \min_{u_t^i \in U^i} \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \tag{4.3.12}$$

Considering the inequalities which follows the Nash equilibrium and Bellman principle of optimality, we can define the robust form of the dynamic programming equation as follows,

$$V^i(x_t) = \min_{u_t^i \in U^i} \max_{\omega \in L^2[t,T]} \left\{ \sum_{t}^{\hat{t}} J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + V^i(x_{\hat{t}}) \right\} \quad \forall \hat{t} \in [t, T] \tag{4.3.13}$$

The Hamilton-Jacobi-Bellman equation for player i's value function can defined by the following partial differential equation.

$$\mathcal{V}_t^i(x_t) = \min_{u_t^i} \max_{\omega_t} \left\{ J^i(x_t, u_t^i, u_t^{\hat{i},r}, \omega_t) + \mathcal{V}_{t+1}^i(x_{t+1}) \right\} \tag{4.3.14}$$

4.4 Robust Interaction Game

Now we shift our focus from the robust general game to the customized linear quadratic differential game between pedestrians and autonomous vehicles. In the game system, AV is modeled according to kinematic bicycle dynamics, and pedestrians are modeled on unicycle dynamics. The performance index or cost function of each player is designed to capture a specific noise or disturbance pattern that is bounded and time-varying. Thus the objective of each player lies in searching for a robust control strategy to realize the possible effect of the disturbance. In this section, we explore the unique formulation of the robust feedback Nash equilibrium strategies for each player of an interactive game with uncertainty. Notably, in an interaction event between pedestrians and AV, pedestrians are expected to behave randomly compared to AV. Thus, we need to design a system that is not overly aggressive or conservative from an AV perspective while interacting with pedestrians on urban streets. We can define this random disturbance using a time-varying and bounded noise system, which will inform the players over a know distribution. By formulation, players will be adept at handling such disturbances readily.

4.4.1 AV Dynamics

State Dynamics: Kinematic Bicycle model [78], $\dot{x_v} = f(t, x_v, u)$ State:

$$x_v = [\quad X_v \quad Y_v \quad v_a \quad \phi_v \quad]^T \tag{4.4.1}$$

where, X_v, Y_v denotes the AV position in Cartesian co-ordinate, and v_a, ϕ_v denotes the speed and heading of AV, respectively. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{X_v} \\ \dot{Y_v} \\ \dot{v_a} \\ \dot{\phi_v} \end{bmatrix} = \begin{bmatrix} v_a \cdot cos\phi_v \\ v_a \cdot sin\phi_v \\ a_v \\ \frac{v_a}{L} \cdot tan\delta_v \end{bmatrix} \tag{4.4.2}$$

where the control variables are $u_a = (a_v, \delta_v)$.

4.4.2 Pedestrians Dynamics

State Dynamics: Unicycle model [78], $\dot{x_p} = f(t, x_p, u)$ State:

$$z_p = [\quad X_p \quad Y_p \quad v_p \quad \theta_p \quad]^T \tag{4.4.3}$$

where, X_p, Y_p denotes the pedestrian position in Cartesian co-ordinate, and v_p, θ_p denotes the speed and heading of the pedestrian. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{X}_p \\ \dot{Y}_p \\ \dot{v}_p \\ \dot{\theta}_p \end{bmatrix} = \begin{bmatrix} v_p \cdot cos\theta_p \\ v_p \cdot sin\theta_p \\ a_p \\ \delta_p \end{bmatrix} \tag{4.4.4}$$

where the control variables are $u_p = (a_p, \delta_p)$. a_p is the acceleration, and δ_p is the change of heading per unit of time.

4.4.3 Game System Description

Consider the nonlinear dynamical system described by a noisy differential equation,

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u})dt + \mathcal{F}(\mathbf{x}, \mathbf{u})d\omega \tag{4.4.5}$$

where, state $x \in \mathbb{R}^n$, control $u \in \mathbb{R}^m$ and noise $\omega \in \mathbb{R}^q$. We define the running cost for each player $(i = 1, 2, .., N)$ as $\ell^i(x, u, t)$ and $h^i(x_T)$ as the terminal cost, where T denotes the final time of the finite horizon game. Here, u^i represents the control law for each player i. The cost function is a collection of cost designed to accumulate over time with the game system initialized with each player's state x at time t, and controlled over the finite horizon T based on the control u^i. With a little redundancy in notation we repeat the general form of the cost in continuous setting.

$$c^i(x, u, \omega, t) = \left[\mathbf{h}^i(x(T)) + \int_t^T \ell^i(\tau, x(\tau), u^i(\tau), u^{\hat{i}}(\tau), \omega)d(\tau)\right], \quad \forall i = 1, 2, .., N \tag{4.4.6}$$

Similar to the previous section, ω defines the disturbance in the system. In the game design, the noise is added to the state and control variables of each player, known as control-dependent noise. The premise of using a known noise is that the players do not fully know the state and control variables and within closed roadway environment the external disturbance can be defined with a time-varying known distribution. As noted in the literature application of robust control on AV and pedestrian are there for individual model purposes. This study attempts to apply the robust control on both players in a game environment. The previous chapter focused on the incomplete information access due to control dependent noise, which is also valid with distinction that this time we know the noise distribution over time for each player. Here, AV can not fully access their control variable of the other players. However, they tackle that lack of access to information with an adversarial input in the cost function that we will discuss later in this section. Moreover, as we translate the game into the linear quadratic form, the noise or disturbance is designed to directly impact the players' optimal control strategies.

Finding the optimum global strategy for each player may be challenging in the game conditions. However, we approximate the optimum local strategy for each player within a specific feasible condition. We approximate the optimal control strategy in the vicinity of the optimal trajectory by applying the optimal control on the system dynamics. We discretized the time horizon to formulate the system dynamics into several steps for the time-varying linear dynamics approach. Since the trajectory depends on the initial conditions, the optimal control approximations also depend on similar conditions.

4.4.4 Local Approximation

The locally optimized control law is constructed iteratively. Each iteration of the algorithm begins with a control sequence $u(t)$ and the corresponding trajectory x_t, obtained by applying u_t to the system dynamics (x_{t+1}) with initial states (x_0) following Euler's integration.

$$x_{t+1} = x_t + \Delta t.f(x_t, u_t)$$

where time is discretized as $t = 1, ..., k$ with time step $\Delta t = T/(k-1)$. The time varying quantities such as controls for each player become $u_t^i \triangleq u^i((k-1)\Delta t)$.

We linearize the system dynamics and quadraticize the cost functions around the nominal trajectory $(\bar{x}, \bar{u})$ to obtain a discrete time linear system with quadratic cost. The linear approximation of the system dynamics transform the nonlinear system as follows,

$$x_{t+1} = \mathcal{A}_t x_t + \sum_{j=1}^{N} \mathcal{B}_t^j u_t^j + \mathcal{D}_{\mathbf{t}}(\omega_t), \quad \forall t = 0, ..., k-1 \tag{4.4.7}$$

where,

$$\mathcal{A}_t = \mathcal{I} + \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{x}}, \qquad \mathcal{B}_t = \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{u}}, \qquad \mathcal{D}_t = \sqrt{\Delta t} \cdot \frac{\partial \mathcal{F}}{\partial \mathbf{x}}, \quad \forall t = 0, ..., k-1$$

The quadratic transformation of the player costs results as follows.

$$\mathcal{J}^i(x_t, u_t^i, u_t^{\hat{i}}, \omega_t) = \frac{1}{2} \sum_{t=1}^{\mathcal{T}} \left[(x_t^T {Q_t}^i + 2{q_t}^i) x_t + \sum_{j=1}^{N} (u_t^{jT} \mathcal{R}_t^{ij} u_t^j + 2 r_t^{ij} u_t^j) - \omega_t \mathcal{W}_t^i \omega_t \right], \quad \forall i = 1, 2, ..N \tag{4.4.8}$$

where, $\ell = \mathcal{J}$

$$\begin{aligned} q_t &= \Delta t \cdot \frac{\partial \ell}{\partial \mathbf{x}}, \qquad Q_t = \Delta t \cdot \frac{\partial^2 \ell}{(\partial \mathbf{x})^2} \\ r_t &= \Delta t \cdot \frac{\partial \ell}{\partial \mathbf{u}}, \qquad \mathcal{R}_t = \Delta t \cdot \frac{\partial^2 \ell}{(\partial \mathbf{u})^2} \end{aligned} \tag{4.4.9}$$

In the dynamics equation, $A_t \in \mathbb{R}^{n \times n}$, $B_t \in \mathbb{R}^{n \times u}$ and $D_t \in \mathbb{R}^{n \times r}$. The finite disturbance

ω enters the system through the matrix D_t. In the quadratic cost function x represents the state vector, u^i denotes control strategies of the player i, and $u^{\hat{i}}$ denotes control strategies of other players except i. ω_t denotes the predefined disturbance. The distinction in this case is that the objective of each player in the game is to find the robust control strategy that minimizes their respective cost while maximizing the adversarial input (disturbance) that is introduced in each players cost function. The first two terms in the cost function $x^T Q^i x$ and $u^{jT} R^{ij} u^j$ penalizes state and control deviation, respectively. Whereas the third term $(\omega_t \mathcal{W}_t^i \omega_t)$ works as an adversarial input to penalize the disturbance from player i's perspective.

Each player works to optimize their control strategy by optimizing their respective cost function independently and coordinates with other players to achieve the optimum interaction. The weighting cost matrices $\mathcal{Q}_t, \mathcal{R}_t$ are symmetric and semi-positive or positive definite. The matrix $\mathcal{W}_t$ is symmetric and positive definite for all players. This matrix also contains the disturbance vector or system noise in an indirect way to describe the model risk for each player in the game setting. The uncertainty setup is included in the framework in a way that players do not expect large deviation in dynamics when the uncertainty term is large enough. Moreover the control or input signal approaches to zero when player's assume a large value (worst-case scenario) of the weight matrix $\mathcal{W}$.

4.5 Robust Controller Design

The approximation of the optimal control strategy of players will be affine in form $u_t^{i*} = \pi_t^i(x_t) = -K_t^i \cdot x_t - \alpha_t^i$. The feed-forward component α_t^i in the optimal control strategy arises from the transformation of the system and cost of players. The control strategy of players are approximately optimal because the game system has control constraints and non-convex costs.

We can design the control strategies u_t^i for the players for time steps $t, ..., (k-1)$. Then, the cost-to-go $\mathcal{V}_t^i(x_t)$ for each player is well defined as the costs accumulates over time with initial state x_0 and time t_0, following the optimal strategies π_t^i for the rest of the time steps. The Hamilton-Jacobi-Bellman (HJB) equations for each player's value function derived in the game element section,

$$\begin{aligned}\mathcal{V}_t^i(x_t) &= \min_{u_t^i}\max_{\omega_t}\left\{J^i(x_t, u_t^i, u_t^{\hat{i}}, \omega_t) + \mathcal{V}_t^i(x_{t+1})\right\}\\ &= \min_{u_t^i}\max_{\omega_t}\left\{\frac{1}{2}\left((x_t^T Q_t^i + 2q_t^{iT})x_t + \sum_{j=1}^{N}(u_t^T\mathcal{R}_t^{ij} + 2r_t^{ij})u_t^j - 2\omega_t\mathcal{W}^i\omega_t\right) + \mathcal{V}_{t+1}^i(x_{t+1})\right\}\end{aligned} \tag{4.5.1}$$

where, $\ell^i = \mathcal{J}^i$ is the running cost and the final value $\mathcal{V}_{T+1}^i(x_{t+1}) = 0$. The value function solution for the linear-quadratic game can be found in the quadratic form [3].

$$\mathcal{V}_t^i = \frac{1}{2}x_t^T\mathcal{P}_t^i x_t + \mathcal{S}_t^T x_t + \eta_t^i \tag{4.5.2}$$

with $\mathcal{P}_{T+1}^i = 0$, $S_{T+1}^i = 0$, $\eta_{T+1}^i = 0$ to be consistent with the final value condition. Now considering the proposed solution of the $\mathcal{V}_t^i(x_t)$, the last term of the HJB equation 4.5.1 transforms into,

$$\mathcal{V}_{t+1}^i(x_{t+1}) = \frac{1}{2}x_{t+1}^T\mathcal{P}_{t+1}^i x_{t+1} + \mathcal{S}_{t+1}^{iT}x_{t+1} + \eta_{t+1}^i \tag{4.5.3}$$

Now, considering the game dynamics 4.4.7 the term can be written as follows,

$$\begin{aligned}\mathcal{V}_{t+1}^i(x_{t+1}) = \frac{1}{2}&\left(\mathcal{A}_t x_t + \sum_{j=1}^{N}\mathcal{B}_t^j u_t^j + \mathcal{D}_t(\omega_t)\right)^T\mathcal{P}_{t+1}^i\left(\mathcal{A}_t x_t + \sum_{j=1}^{N}\mathcal{B}_t^j u_t^j + \mathcal{D}_t(\omega_t)\right)\\ &+ \mathcal{S}_{t+1}^{iT}\left(\mathcal{A}_t x_t + \sum_{j=1}^{N}\mathcal{B}_t^j u_t^j + \mathcal{D}_t(\omega_t)\right) + \eta_{t+1}^i\end{aligned}$$

The feedback control law is obtained by finding the minimizes of $\mathcal{V}_t^i(x_t)$. The assumption for this case relates to convexity and set the gradient to zero,

$$0 = \mathcal{R}_t^{ii} u_t^i + r_t^{iiT} + \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i \left(\mathcal{A}_t x_t + \sum_{j=1}^{N} \mathcal{B}_t^j u_t^j + \mathcal{D}_t \omega_t \right) + \mathcal{B}_{t+1}^{iT} \mathcal{S}_{t+1}^i \tag{4.5.4}$$

The admissible optimal control u_t^{i*} based on the linear feedback is an affine function. The α_t^i term is added to the equation to capture the linear transformation of the nonlinear system.

$$u_t^{i*} = -\mathcal{K}_t^i x_t - \alpha_t^i \tag{4.5.5}$$

Inserting the optimal control strategy u_t^{i*} leads to,

$$0 = -\mathcal{R}_t^{ii} (\mathcal{K}_t^i x_t + \alpha_t^i) + r_t^{iiT} + \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i \left(\mathcal{A}_t x_t - \sum_{j=1}^{N} \mathcal{B}_t^j (\mathcal{K}_t^j x_t + \alpha_t^j) + \mathcal{D}_t^j \omega_t^j \right) + \mathcal{B}_{t+1}^{iT} \mathcal{S}_{t+1}^i \tag{4.5.6}$$

Two similar systems of equations to find $\mathcal{K}_t^i$ and α_t^i,

$$\begin{aligned} \left(\mathcal{R}_t^{ii} + \mathcal{B}_t^{iT} \mathcal{S}_{t+1}^i \mathcal{B}_t^i \right) \mathcal{K}_t^i + \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i \sum_{j \neq i} \mathcal{B}_t^j \mathcal{K}_t^j &= \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i A_t \\ \left(\mathcal{R}_t^{ii} + \mathcal{B}_t^{iT} \mathcal{S}_{t+1}^i \mathcal{B}_t^i \right) \alpha_t^i + \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i \sum_{j \neq i} \mathcal{B}_t^j \alpha_t^j &= \mathcal{B}_{t+1}^{jT} \mathcal{S}_{t+1}^i + r_t^{ii} + \mathcal{D}_t^i \omega_t \end{aligned} \tag{4.5.7}$$

Based on the similar systems of equations, we can compute the control strategy using linear equations $\alpha = \mathcal{H} \setminus \mathcal{M}\alpha$, $\mathcal{K} = \mathcal{H} \setminus \mathcal{M}\mathcal{K}$, where,

$$\begin{aligned} \mathcal{H} &= \mathcal{R}_t^{ii} + \mathcal{B}_t^{iT} \mathcal{P}_{t+1}^i \mathcal{B}_t^i \\ \mathcal{M}\alpha &= r_t^{ii} + \mathcal{B}_t^{iT} \mathcal{S}_{t+1}^i + \mathcal{D}_t^i \omega_t \\ \mathcal{M}\mathcal{K} &= \mathcal{B}_t^{iT} \mathcal{S}_{t+1}^i \mathcal{A}_t \end{aligned} \tag{4.5.8}$$

Similarly we can find an equation for the disturbance ω_t for player i by maximizing the $\mathcal{V}_t^i(x_t)$,

$$\omega_t^{i*} = (\mathcal{W}^i - \mathcal{D}_t^T P_{t+1}^i \mathcal{D}_t)^{-1} \Big\{ \mathcal{D}_t^T P_{t+1}^i (\mathcal{A}_t x_t - \sum_{j=1}^{N} \mathcal{B}_t^j u_t^j) + \mathcal{S}_t^{iT} \mathcal{D}_t \Big\} \tag{4.5.9}$$

4.5.1 Game Solution

Now we derive the necessary expression for $\mathcal{P}_t^i$, $\mathcal{S}_t^i$, and η_t^i of the proposed solution of HJB equation.

$$\begin{aligned} \mathcal{V}_t^i(x_t) = \frac{1}{2}\Bigg((x_t^T Q_t^i + 2q_t^{iT}) x_t + \sum_{j=1}^{N} \big((\mathcal{K}_t^j x_t + \alpha_t^j)^T R_t^{ij} - 2r_t^{ijT} \big) \big(\mathcal{K}_t^j x_t + \alpha_t^j \big) - \omega_t \mathcal{W}_t^i \omega_t \Bigg) \\ + \frac{1}{2}\Bigg(\Big(\mathcal{A}_t x_t - \sum_{j=1}^{N} \mathcal{B}_t^j (\mathcal{K}_t^j x_t + \alpha_t^j) + \mathcal{D}_t \omega_t \Big)^T \mathcal{P}_{t+1}^i + 2\mathcal{S}_{T+1}^{iT} \Bigg) \\ \times \Bigg(\mathcal{A}_t x_t - \sum_{j=1}^{N} \mathcal{B}_t^j (\mathcal{K}_t^j x_t + \alpha_t^j) + \mathcal{D}_t \omega_t \Bigg) + \eta_{t+1}^i \end{aligned} \tag{4.5.10}$$

To shorten the equation, we consider,

$$\begin{aligned} \mathcal{G}_t &= \mathcal{A}_t - \sum_{j=1}^{N} \mathcal{B}_t^j \mathcal{K}_t^j \\ \beta_t &= - \sum_{j=1}^{N} \mathcal{B}_t^j \alpha_t^j \\ \mu P &= (\mathcal{W}^i - \mathcal{D}_t^T P_{t+1}^i \mathcal{D}_t)^{-1} (\mathcal{D}_t^T P_{t+1}^i \mathcal{G}_t) \\ \mu S &= (\mathcal{W}^i - \mathcal{D}_t^T P_{t+1}^i \mathcal{D}_t)^{-1} (\mathcal{D}_t^T P_{t+1}^i \beta_t) \end{aligned}$$

The value function equation becomes,

$$\begin{aligned} \mathcal{P}_t^i &= Q_t^i + \sum_{j=1}^{N} \mathcal{K}_t^{iT} R_t^{ij} \mathcal{K}_t^i + \mathcal{G}_t^T \mathcal{P}_{t+1}^i \mathcal{G}_t + \mathcal{G}_t^{iT} \mathcal{D}_t^{jT} \mathcal{P}_{t+1}^i \mu P, \quad \mathcal{P}_{T+1} = 0 \\ \mathcal{S}_t^i &= q_t^i + \sum_{j=1}^{N} (\mathcal{K}_t^{iT} R_t^{ij} \alpha_t^j - \mathcal{K}_t^{iT} r_t^{ij}) + \mathcal{G}_t^T (\mathcal{S}_{t+1}^i + \mathcal{P}_{t+1}^i \beta_t) + \mathcal{G}_t^{iT} \mathcal{D}_t^{jT} \mathcal{P}_{t+1}^i \mu S, \quad \mathcal{S}_{t+1} = 0 \end{aligned} \tag{4.5.11}$$

4.6 Iterative Algorithm Overview

The iterative algorithm for the robust interaction framework in this study is inspired from several iterative control and game theoretic approaches [95]. The algorithmic steps are listed below.

- In the first step we initialize the player trajectory with initial trajectory vector x_t and simulate the trajectory with respect to the system disturbance and initial control strategies to arrive to the next operating point of the game.

- In the second step we proceed to linearize the nonlinear dynamics x_{t+1} an quadraticize the cost $\mathcal{J}^i$. We follow the procedure mentioned in the local approximation section and complete this step with linearized dynamics and quadratic costs. The finite disturbance ω is added to game system and the adversarial input $\omega_t \mathcal{W} \omega_t$ is included in the player cost function.

- In the third step we continue to the linear quadratic game formulation. Using the robust dynamic programming equation we solve for the robust feedback Nash equilibrium. We complete this step in backwards from final time horizon to initial time to compute the optimal control strategies for the players in the game. We use the HJB equation solution noted in the game solution section.

- In the fourth step we employ the combination of regularization and control to ensure the stability and optimality of the proposed game. For instance, regularization prevents the overfiting to the noise sequence and control stabilizes the system by reducing cost to ensure optimality. After deriving the dynamic programming solution for the game problem in the previous step, we get the new control gains $\mathcal{K}_t$ and offset update α_t.

- In the final step we perform line search, which is used to ascertain the convergence condition. In the proposed linear quadratic game with a known system noise or disturbance, it searches for the optimal strategy that minimizes the cost function associated with the players. More appropriately, the update in each iteration may lead to an increased cost or divergence in case the new trajectory is far from the valid region of local approximation. To avoid this incidence, the new nominal control trajectory is computed by backtracking line search with initial step of 0.5. The subsequent step is varied manually to test different approximations of the nominal trajectory. When the candidate trajectory results in a lower cost estimation or remains within the valid region of approximation, then this trajectory is accepted. Otherwise, the entire process is reiterated to find the appropriate trajectory where the control is within a threshold. After meeting the convergence criterion the proposed algorithm returns nominal trajectory and optimal feedback strategies. Otherwise it iterates over the first step and revisits the steps all over again until convergence.

4.7 Numerical Results

The player model assumptions are same as those in chapter-3. The cost components and constraints for both AV and pedestrians are also similar to that of Chapter-3. The numerical results section complements and extends the results obtained in Chapter-3. The iterative algorithm of this study is different in design, application, and follows the identical steps conforming to the risk sensitive iterative LQR design [95]. In the numerical result section, we show that the presence of a time-varying finite disturbance in the system can be captured and mitigated considering the adversarial input in the player cost function. We draw comparison between a stochastic linear quadratic game model that undergoes system disturbance without the adversarial input and a robust game model that addresses the disturbance with the adversarial input as a special player component in the cost function.

4.7.1 3-Player Robust Interaction Game

The three-player interaction game parameters are given in tabular format, including the time-varying finite disturbance, initial states, target states, and collision radius of players. The game parameters remain exactly the same for both type of games. The game parameters are in meters where applicable.

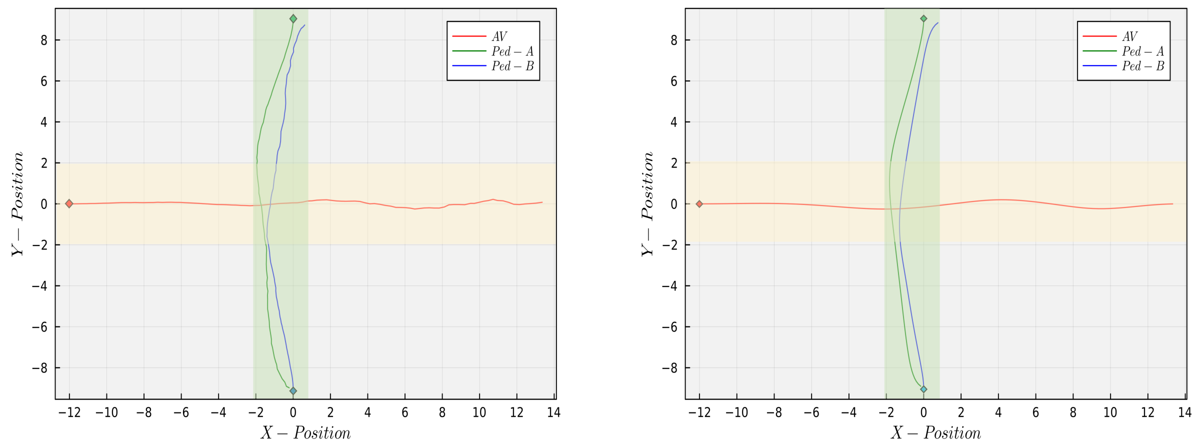
Table 4.1: Game Parameters for 3-Player Interaction

Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-12, 0, 0, 2.2)	(0, 9, $-\pi/2$, 1.2)	(0, -9, $\pi/2$, 1.2)
Target State	(15, 0, 0, 4)	(0, -10, 0, 1.5)	(1, 10, 0, 1.5)
Finite Noise	0.9sin(0.01t)exp0.01t	0.9sin(0.01t)exp0.01t	0.9sin(0.01t)exp0.01t
Collision Radius (m)	2.5	0.9	0.9

The game trajectory describes the interaction between three players (one AV and two pedestrians over a crosswalk). The interaction is sensitive to initial conditions as AV's higher speed and short distance from the crosswalk will likely encourage no major interaction event between AV and crossing pedestrians. Thus, the game is formulated in a way to promote an interaction between the players. The interaction dynamics of AV and pedestrians are governed by the transformed dynamics of the players and feedback strategy. The finite disturbance is visible in the state and control variable of the players, specifically for players in the regular stochastic game without the adversarial input.

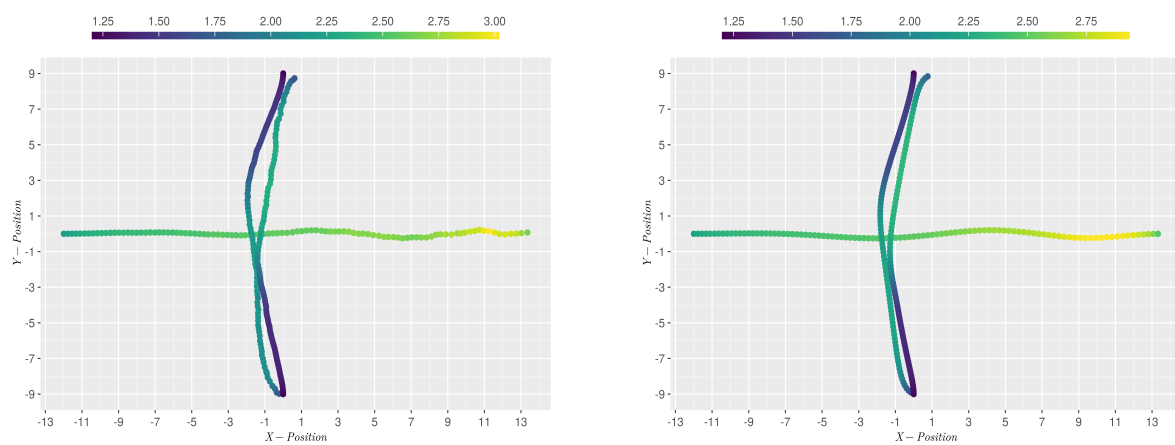
Exploring the Figure 4.1 shows that the stochastic AV interaction framework described in the previous chapter cannot handle the time-varying finite disturbance in the system dynamics. The adversarial input in the cost function and robust Nash feedback approximation captures the finite disturbance and filter it out from the player trajectory.

Similarly, the trajectory fused with the speed (Fig 4.2) shows that the robust game handles the finite disturbance well compared to the stochastic game formulation. The comparison is drawn



(a) Three Player Game without Adversarial Input (b) Three Player Robust Game with Adversarial Input

Figure 4.1: Three-Player Interaction Game with and without Adversarial Input

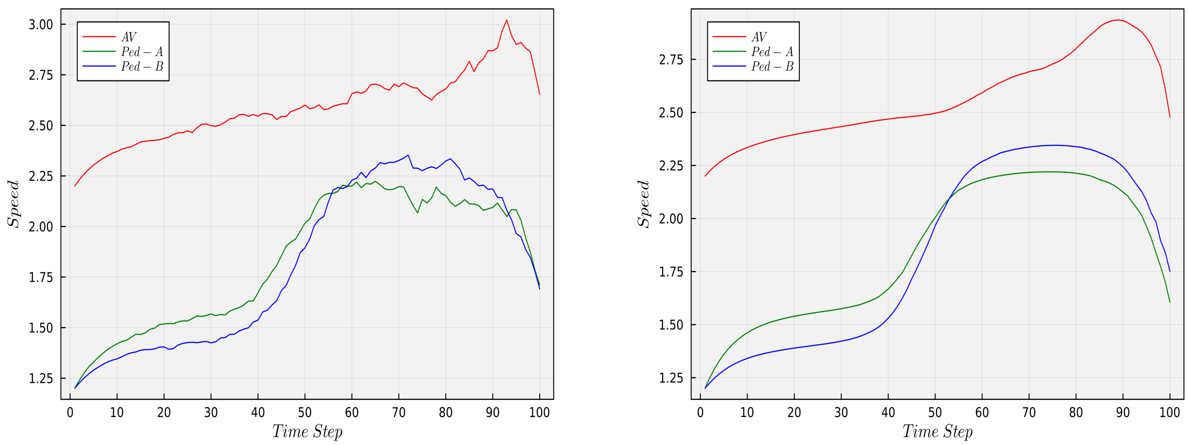


(a) Player Trajectory and speed without Adversarial Input (b) Player Robust Trajectory and speed in the presence Adversarial Input

Figure 4.2: Three-Player Interaction Game Trajectory and speed combination with and without Adversarial Input

with the same initial conditions for both type of games. With similar finite disturbance in the system the speed distribution for the players are almost identical for both cases. The primary goal in this case is to achieve a control strategy that captures the finite disturbance and adjust the game variables in the AV framework to interact safely.

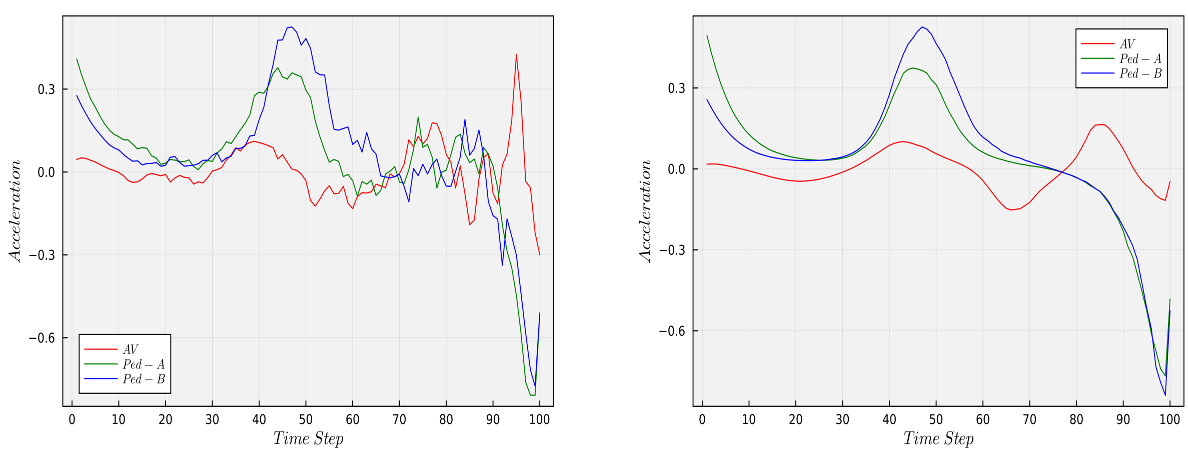
The individual speed distribution of the players over the finite horizon (Fig 4.3a) also shows significant impact of the finite disturbance throughout the interaction event. However the impact of



(a) Three-Player Speed Trajectory without Adversarial Input (b) Three-Player Robust Speed Trajectory with Adversarial Input

Figure 4.3: Three-Player Interaction Game Speed Trajectory with and without Adversarial Input

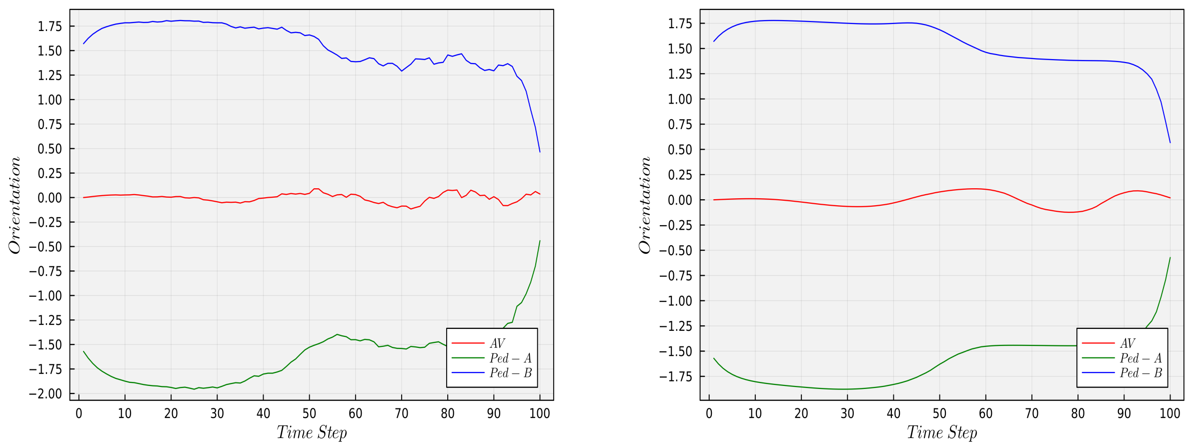
the disturbance is filtered in the robust game framework (Fig 4.3b). The robust game framework will aid AV in smooth motion planning and discard abrupt motion changes due external disturbance. This will in turn improve the ridership quality and trajectory prediction in the real-world AV decision making process.



(a) Three-Player Acceleration Trajectory without Adversarial Input (b) Three-Player Robust Acceleration Trajectory with Adversarial Input

Figure 4.4: Three-Player Interaction Game Acceleration Trajectory with and without Adversarial Input

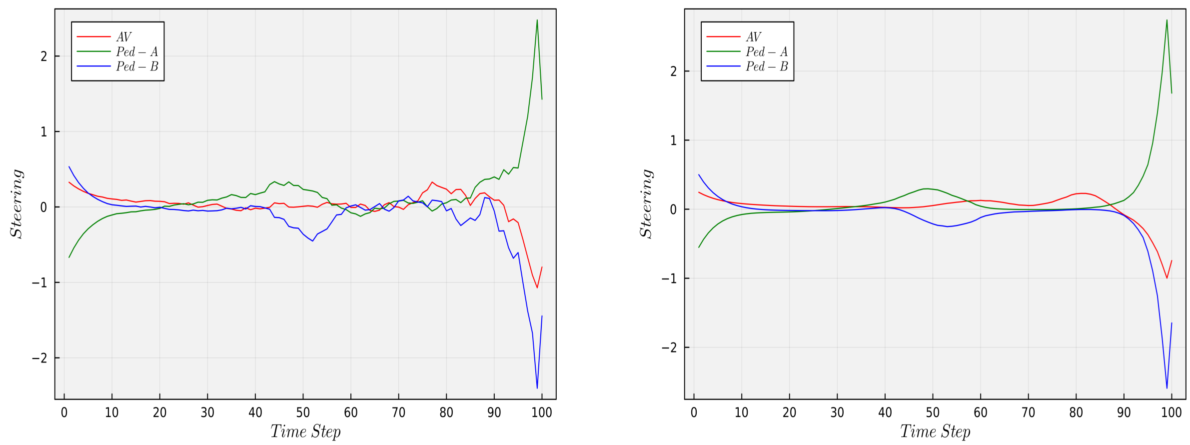
The acceleration distribution without adversarial input shows the maximum impact of the finite noise distribution (Fig 4.4a). It implies that in certain game conditions players behavior is affected by the finite disturbance and the previous stochastic framework is not adequate to capture and filter the disturbance. The proposed robust game captures the disturbance with the help of adversarial input and robust dynamic programming equation. Close inspection of the robust acceleration (Fig 4.4b) shows that there is some effect of noise near the final time as the disturbance is time-varying and reaches maximum at the end horizon. However, we are limited to the interaction event that occurs in between 4 to 7 seconds in the time horizon.



(a) Three-Player Orientation without Adversarial Input (b) Three-Player Robust Orientation with Adversarial Input

Figure 4.5: Three-Player Interaction Game Orientation with and without Adversarial Input

From the numerical results of two types of game with respect to orientation (Fig 4.5) and steering control (Fig 4.6), we can see that the robust formulation captures the finite time-varying disturbance and performs a smoothing or filtering type of effect on the game variables. Notably, the robust game formulation is flexible to explore all the behavioral conditions discussed in Chapter-3 to verify the robustness of the new formulation for the AV interaction framework. Moreover, we can extend the applicability of the robust control to only on AV with random noise on the pedestrians dynamics to make the AV interaction framework adept to urban setting.



(a) Three-Player Steering Control without Adversarial Input

(b) Three-Player Robust Steering Control with Adversarial Input

Figure 4.6: Three-Player Interaction Game Steering Control with and without Adversarial Input

4.8 Quantitative and Qualitative Evaluation

The quantitative and qualitative evaluation in this section follows the same methodology detailed in Chapter-3.

4.8.1 Quantitative Evaluation

Table 4.2: Quantitative Evaluation Summary for Robust Game

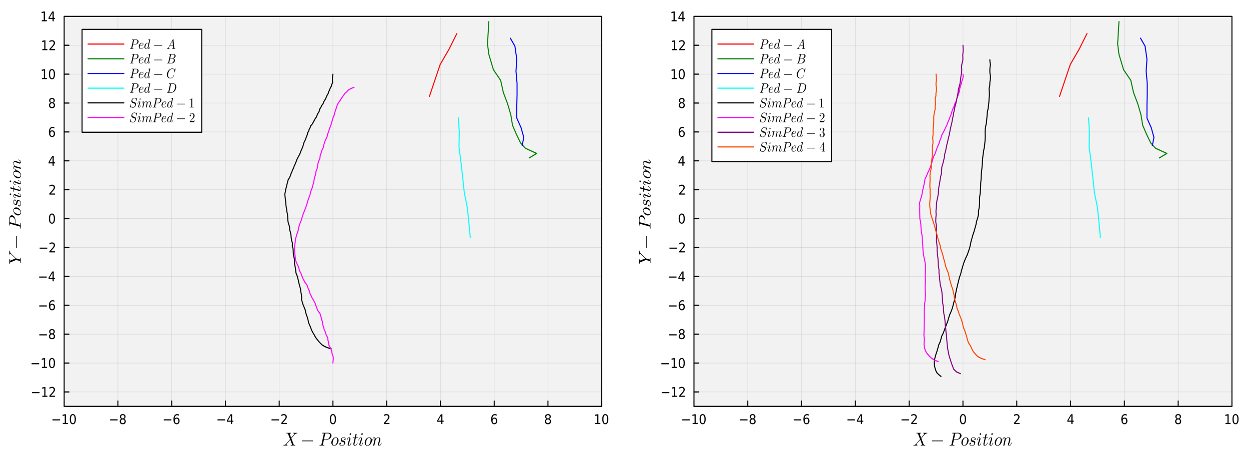
		Avg Crossing (s)				Avg Simulated (s)		Difference (s)	
Scenario	Sample	Mix	Con	Agr	Int	Crs	Int	Crs	Int
A (1)	16	10.331	10.750	9.075	2.994	10.372	3.252	0.041	0.258
B (2)	10	10.380	10.539	8.950	3.035	10.475	3.112	0.095	0.077
C (3)	7	10.519	10.506	9.200	3.048	10.621	3.311	0.102	0.263
D (4)	6	10.758	10.810	9.233	3.079	11.225	3.321	0.467	0.242
E (5)	5	10.972	11.068	9.133	3.092	11.595	3.251	0.623	0.159

*Avg - Average, s - Seconds, Mix - Mixed, Con - Conservative, Agr - Aggressive, Int - Interaction, Crs - Crossing.

The results indicate that the average crossing and interaction time for each simulated scenario closely matches real-world data. For instance, the difference between simulated and real-world crossing time for two pedestrians crossing from the opposite direction is 0.095 seconds. For the four pedestrians crossing scenario, the difference between average crossing time is about 0.467 seconds. The interaction time difference for multiple pedestrian scenarios (A, B, C, D, E) ranges from 0.258 to 0.263 seconds (Table 4.2). The maximum difference in average crossing time (0.623) is recorded for scenario E with 5 pedestrians crossing as a group. The average interaction time remains identical across the scenarios.

4.8.2 Qualitative Evaluation

We evaluate the robust game using qualitative pedestrian trajectories. The motion pattern of real-world pedestrians relates to the simulated trajectories with noise. The robust game filters the noisy trajectory and generates a smooth trajectory for AV and pedestrians. The flexibility of the smoothing effect is controlled through the weightage of the adversarial input into the system. This implies that the AV can fine-tune the robust control to adjust to real-world pedestrian motion patterns and interact safely. The flexibility of the robust model developed in this chapter will aid in smoothing AV trajectory while maintaining the noise element in the pedestrians' motion to reflect real-world conditions.



(a) Pedestrian Trajectories from 3-player Robust Game

(b) Pedestrian Trajectories from 5-player Robust Game

Figure 4.7: Qualitative Evaluation of the Simulated Trajectories with respect to ETH data

The simulated pedestrian trajectories visually match the motion pattern of the real-world data with some differences (Fig 4.7). The visual difference with the real-world trajectories occurs as the pedestrians are not walking in a controlled crosswalk environment in the presence of a vehicle nearby. The flexibility of the robust game framework allows us to model the noise configuration for pedestrians and AV separately. This game formulation would aid in real-world motion control of AV that requires a smooth trajectory for ride comfort and safe interaction. In contrast, pedestrians' motion is always susceptible to environmental noise and individual crossing behavior. Thus, filtering the motion noise of AV for smooth trajectory and capturing the noisy behavior of pedestrians during an interaction event would aid the AV's decision-making capability in urban scenarios.

4.9 Conclusion

The proposed game describes a specific noise structure to handle stochastic features of the interaction scenario between pedestrians and AV in urban streets. Based on our understanding of the stochastic noise process in the previous chapter, the proposed robust game in this chapter investigates a novel formulation using an adversarial input in the cost function in the form of a finite disturbance. The concept of designing the AV interaction framework stems from the safe systems approach in transportation, where the goal of designing any particular strategy or infrastructure incorporates the possibility of driving or human error. The game framework also establishes the robust dynamic programming equation with the help of the optimality principle, which implies that each player tries to minimize the core cost components related to state and control. Whereas the adversarial input always tries to maximize the cost components.

We use a novel game theoretic approach to detail the AV interaction against finite disturbance and adversarial input. The game framework is flexible because we can customize different types of finite disturbance on the system and develop an adequate adversarial input for each player's cost function to design a robust framework for different scenarios. From an algorithmic perspective, the interaction framework for AV finds local approximations in a multiplayer noisy dynamic game setting. It resembles the architecture of the risk-sensitive iterative Linear Quadratic Regulator setting. The game solution preserves the competitive and cooperative behavior associated with the Nash Equilibrium and extends the solution concept for time-varying finite disturbance instead of the previous control-dependent noise. Although the time-varying disturbance is maximum near the end of the finite horizon, our goal is to explore the impact of the disturbance on the players during the actual interaction and find possible ways to mitigate the noisy state and control trajectories.

The benefit of the proposed robust game framework is two folds. Firstly, the interaction framework of AV will be ready to capture such external disturbances and plan accordingly in the presence of other players. Secondly, the game can be formulated in a way that only one player (AV) may emphasize the robust control, whereas the other players (pedestrians) may operate on random stochastic noise distribution. The smooth trajectory will aid AV motion control, planning, and rider comfort in a real-world interaction. The scenarios we explored in the previous chapter for different types of pedestrian crossing behavior can also be demonstrated with the robust game framework with finite disturbance and adversarial input.

Our investigation into the system mechanism shows that the robust game is sensitive to initial conditions and communication strategies. For this chapter, we proposed a robust differential game that can handle external noise as an adversarial input to the system. We evaluate the robust game using quantitative and qualitative methods. The quantitative method follow the field data from Chapter-3 and draws a difference between actual and simulated crossing and interaction time. The results described in the previous section show that the robust framework closely matches the real-world data. Similarly, the visual comparison of motion pattern of real-world pedestrians relates to the simulated trajectories with noise.

The robust game in this chapter is set up in a finite horizon and follows a feedback control strategy. The feedback control implies that immediate information is available to each player, which helps to identify a safe and optimal strategy for any given interaction scenario. However, in real-world scenarios, each interacting player can access limited or delayed information about other players' states. For instance, pedestrians are limited by peripheral vision and other constraints, whereas vehicle control may vary with the ever-changing urban scene. Thus, a strategy with delayed information available to the players will reflect the interaction event more accurately. In our following research approach, we will explore the delayed feedback strategy for pedestrians and AV with the goal of updating the current framework.

Chapter 5

Delayed Feedback Interaction Framework for Autonomous Vehicles and Pedestrians

5.1 Introduction

In this chapter, we explore a delayed feedback information system for the players based on the stochastic linear quadratic game proposed in chapter-3. The proposed interaction framework accommodates the Brownian Motion noise on the control of the players and a finite delay in the game system. This study assumes that the players can access their state information in a system delayed by a time step. To solve such a game system, we introduce a novel form of the Hamilton-Jacobi-Bellman equation for each player, which computes the delayed feedback at each time step and solves the game for Nash equilibrium.

The delayed feedback game formulation includes the general elements, including (a) game elements; (b) controller design; and (c) algorithm design. The game elements define players' behavioral features and cost components, drawing parallels with those described in Chapter-3. The controller design is the core component of the study, as all the derivation and equilibrium concepts related to the time-delayed feedback game stem from this section. In contrast, the algorithm design enforces local approximation, backward pass, control regularization, and line search for convergence conditions similar to that of Chapter-3. The study's uniqueness is combining all the described elements and developing a stochastic game theoretic framework for Autonomous vehicle interaction with pedestrians in urban streets with delayed communication patterns. The stochastic component for the interaction framework builds on the control-dependent Brownian motion or Gaussian noise. The primary motivation for using control-dependent noise remains the same as in Chapter-3, given that the game formulation comprises a linear quadratic structure.

5.2 Contributions

The primary contributions of this chapter are:

- We define a stochastic nonzero-sum dynamic linear quadratic game framework for AV interaction with pedestrians with a finite delay in the feedback control of the players.
- We derive and describe the model-based setting to compute the Nash Equilibrium with delayed feedback strategies and HJB functions for the AV interaction framework.
- We provide numerical experimentation based on the AV interaction framework to illustrate the proposed algorithm's effectiveness in capturing the associated change in feedback strategies.

5.3 Proposed Interaction Game

Now we shift our focus from the robust general game to the customized linear quadratic differential game between pedestrians and autonomous vehicles. In the game system, AV is modeled according to kinematic bicycle dynamics, and pedestrians are modeled on unicycle dynamics. The performance index or cost function of each player is designed to capture a specific noise or disturbance pattern that is bounded and time-varying. Thus the objective of each player lies in searching for a optimal control strategy to realize the possible effect of the delay. In this section, we explore the unique formulation of the delayed feedback Nash equilibrium strategies for each player of an interactive game with stochastic control dependent noise. Notably, in an interaction event between pedestrians and AV, pedestrians are expected to behave randomly compared to AV. Thus, we need to design a system that is not overly aggressive or conservative from an AV perspective while interacting with pedestrians on urban streets.

5.3.1 AV Dynamics

State Dynamics: Kinematic Bicycle model [78], $\dot{x_v} = f(t, x_v, u)$ State:

$$x_v = \begin{bmatrix} X_v & Y_v & v_a & \phi_v \end{bmatrix}^T \tag{5.3.1}$$

where, X_v, Y_v denotes the AV position in Cartesian co-ordinate, and v_a, ϕ_v denotes the speed and heading of AV, respectively. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{X}_v \\ \dot{Y}_v \\ \dot{v_a} \\ \dot{\phi}_v \end{bmatrix} = \begin{bmatrix} v_a \cdot cos\phi_v \\ v_a \cdot sin\phi_v \\ a_v \\ \frac{v_a}{L} \cdot tan\delta_v \end{bmatrix} \tag{5.3.2}$$

where the control variables are $u_a = (a_v, \delta_v)$.

5.3.2 Pedestrians Dynamics

State Dynamics: Unicycle model [78], $\dot{x_p} = f(t, x_p, u)$ State:

$$z_p = [\quad X_p \quad Y_p \quad v_p \quad \theta_p \quad]^T \tag{5.3.3}$$

where, X_p, Y_p denotes the pedestrian position in Cartesian co-ordinate, and v_p, θ_p denotes the speed and heading of the pedestrian. The state dynamics can be written as follows,

$$\begin{bmatrix} \dot{X_p} \\ \dot{Y_p} \\ \dot{v_p} \\ \dot{\theta_p} \end{bmatrix} = \begin{bmatrix} v_p \cdot cos\theta_p \\ v_p \cdot sin\theta_p \\ a_p \\ \delta_p \end{bmatrix} \tag{5.3.4}$$

where the control variables are $u_p = (a_p, \delta_p)$. a_p is the acceleration, and δ_p is the change of heading per unit of time.

5.3.3 Game System Description

Consider a non-linear dynamical system with time-delay described by a stochastic differential equation,

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{u}(t-\tau))dt + \mathcal{F}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{u}(t-\tau))d\omega(t) \tag{5.3.5}$$

where, state $x \in \mathbb{R}^n$, control $u \in \mathbb{R}^m$, Brownian motion $\omega(t) \in \mathbb{R}^q$ and τ is a fixed delay. The initial conditions for state and control are $x(0) = x_0$ and $u(s) = 0$ for any $s \in [-\tau, 0]$. f and F defines the continuous dynamics function and drift dynamics of the system. We define the running cost for each player $(i = 1, 2, .., N)$ as $\ell^i(x, u, t)$ and $h^i(x_T)$ as the terminal cost, where T denotes the final time of the finite horizon game. Here, u^i represents the control law for each player. The cost function is a collection of cost designed to accumulate over time with the game system initialized with each player's state x at time t, and controlled over the finite horizon T based on the control u^i. With a little redundancy in notation we repeat the general form of the cost in continuous setting.

$$\mathcal{J}^i(x,u,t) = \mathbf{E}\Big[\mathbf{h}^i(x(T)) + \int_t^T \ell^i(s,x(s),u(s))ds\Big], \quad \forall i = 1,2,..,N \tag{5.3.6}$$

The cost function $J(x,u,t)$ represents the total cost accumulated from initial time t to final time T with the state system (x) initialized at the time instance t and control input $u = u(t,x)$. The expectation arises due to the presence of stochastic process ω. In the game design, the noise is added to the system variables of each player, known as control-dependent noise. The premise of using noise is that the players do not fully know the control variables. For example, the AV can not fully access their control variable or the other players. In other words, AVs do not have complete knowledge of the control pattern of pedestrians. Likewise, pedestrians cannot assume the AV controls (acceleration and steering) while interacting at the crosswalk. Moreover, as we translate the game into the linear quadratic form, the control-dependent noise directly impacts the players' optimal control strategies, unlike state-dependent noise.

5.3.4 Local Approximation

The locally optimized control law is constructed iteratively. The time is discretized with time step (sample time) $\Delta t = T/(K) = \tau/l$. The time varying quantities such as controls for each player become $u(t)^i \triangleq u^i(k\Delta t) = u_k$ with $k\Delta t \leq t < (k+1)\Delta t$. Each iteration of the algorithm begins with an initial control sequence $\bar{u}(t)$ and the corresponding trajectory $\bar{x}(t)$, obtained by applying $\bar{u}(t)$ to the system dynamics $(\dot{x}(t) = f(x(t),u(t),u(t-\tau)))$ with initial states $(x(0) = x_0)$ following

Euler's integration.

$$\bar{x}_{k+1} = \bar{x}_k + \Delta t.f(\bar{x}_k, \bar{u}_k, \bar{u}_{k-1})$$

We linearize the system dynamics around $(\bar{x}, \bar{u})$ to describe a discrete time linear system governed by the state and control deviations such as $\delta x_k = x_k - \bar{x}_k$ and $\delta u_k = u_k - \bar{u}_k$, respectively. Following the deviation quantities, the approximation of the nonlinear continuous time state dynamics equation becomes,

$$\delta x_{t+1} = \mathcal{A}_t \delta x_t + \sum_{j=1}^{N} (\mathcal{B}_t^{0j} \delta u_t^j + \mathcal{B}_t^{1j} \delta u_{t-1}^j) + \mathcal{C}_{\mathbf{t}}(\delta u_t, \delta u_{t-1}) \chi_{\mathbf{t}}, \quad \forall t = 0, 1, ..., k-1$$

with $\delta x_0 = 0$ and $\delta u_{-1} = \delta u_{-2} = ... = \delta u_{-l} = 0$. Now with a slight abuse of notation we can define $\delta x = x$ and $\delta u = u$ for easier representation. Thus the dynamics equation becomes,

$$x_{t+1} = \mathcal{A}_t x_t + \sum_{j=1}^{N} (\mathcal{B}_t^{1j} u_t^{0j} + \mathcal{B}_t^{1j} u_{t-1}^j) + \mathcal{C}_{\mathbf{t}}(u_t, u_{t-1}) \chi_{\mathbf{t}}, \quad \forall t = 0, 1, ..., k-1 \tag{5.3.7}$$

where,

$$\mathcal{A}_t = \mathcal{I} + \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{x}}, \quad \mathcal{B}_t^l = \Delta t \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{u}_l}, \quad \mathcal{C}_t = \sqrt{\Delta t} \cdot \frac{\partial \mathcal{F}}{\partial \mathbf{u}_l}, \quad \forall t = 0, ..., k-1 \quad \& l = 0, 1, ..$$

Now, we can explain each of these terms from the linearization perspective around a nominal trajectory $(\bar{x}, \bar{u})$ as follows,

$$\begin{aligned} \mathbf{f} &= f(\bar{x}_t, \bar{u}_t, \bar{u}_{t-1}), \qquad \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \frac{\partial f(\bar{x}_t, \bar{u}_t, \bar{u}_{t-1})}{\partial \mathbf{x}}, \qquad \mathcal{F} = F(\bar{u}_t, \bar{u}_{t-1}), \\ \frac{\partial \mathbf{f}}{\partial \mathbf{u}_0} &= \frac{\partial f(\bar{x}_t, \bar{u}_t, \bar{u}_{t-1})}{\partial u_t}, \qquad \frac{\partial \mathbf{f}}{\partial \mathbf{u}_1} = \frac{\partial f(\bar{x}_t, \bar{u}_t, \bar{u}_{t-1})}{\partial u_{t-1}}, \\ \frac{\partial \mathcal{F}}{\partial \mathbf{u}_0} &= \frac{\partial F(\bar{u}_t, \bar{u}_{t-1})}{\partial u_t}, \qquad \frac{\partial \mathcal{F}}{\partial \mathbf{u}_1} = \frac{\partial F(\bar{u}_t, \bar{u}_{t-1})}{\partial u_{t-1}} \forall t = 0, ..., k-1 \& l = 0, 1, .. \end{aligned}$$

The noise model ξ_t $N(0, I_p)$ captures multiplicative noise to the control strategies of the players for $t = 0, 1, ..., k-1$. The $\sqrt{\Delta t}$ term describes the linear growth of the covariance of the Brownian motion with time. The noise covariance is,

$$Cov\Big[\mathcal{C}_{\mathbf{t}}(u_t, u_{t-1})\chi_{\mathbf{t}}\Big] = \sum_{j=1}^{N}(\mathcal{C}_{\mathbf{t}}{}^{0j}u_t^j + \mathcal{C}_{\mathbf{t}}{}^{1j}u_{t-1}^j)(\mathcal{C}_{\mathbf{t}}{}^{0j}u_t^j + \mathcal{C}_{\mathbf{t}}{}^{1j}u_{t-1}^j)^T \qquad l = 0, 1, ..p \tag{5.3.8}$$

Now, we quadraticize the cost functions around the nominal trajectory $(\bar{x}, \bar{u})$ to obtain a discrete time linear system with quadratic cost. The quadratic transformation of the general player costs results as follows.

$$J_t^i = \frac{1}{2}\sum_{t=1}^{T}\left[(x_t^T {Q_t}^i + 2{q_t}^i)x_t + \sum_{j=1}^{N}(u_t^{jT}\mathcal{R}_t^{ij}u_t^j + 2r_t^{ij}u_t^j)\right], \qquad \forall i = 1, 2, ..N \tag{5.3.9}$$

where, i is the number of players in the game, and

$$\begin{aligned} q_t &= \Delta t \cdot \frac{\partial J}{\partial \mathbf{x}}, \qquad Q_t = \Delta t \cdot \frac{\partial^2 J}{(\partial \mathbf{x})^2} \\ r_t &= \Delta t \cdot \frac{\partial J}{\partial \mathbf{u}}, \qquad \mathcal{R}_t = \Delta t \cdot \frac{\partial^2 J}{(\partial \mathbf{u})^2} \end{aligned} \tag{5.3.10}$$

5.4 Computing Cost-to-go

The approximation of the optimal control strategy of players will be affine in form $u_t^{i*} = \pi_t^i(x_t) = K_t^i \cdot x_t + \alpha_t^i$. The feed-forward component α_t^i in the optimal control strategy arises from the transformation of the system and cost of players. The control strategy of players are approximately optimal because the game system has control constraints and non-convex costs. The actual format of the feedback will change based on the delay equation. We define the formal optimal control equation in the later part of the chapter. The general quadratic cost function for the players is as follows,

$$J_t^i = \frac{1}{2}\sum_{t=1}^{T}\Bigg[(x_t^T {Q_t}^i + 2{q_t}^i)x_t + \sum_{j=1}^{N}(u_t^{jT}\mathcal{R}_t^{ij}u_t^j + 2r_t^{ij}u_t^j)\Bigg], \qquad \forall i = 1, 2, ..N$$

We can design the control strategies u_t^i for the players for time steps $t, ..., (k-1)$. Then, the cost-to-go or value function $\mathcal{V}_t^i(x_t)$ for each player is well defined as the costs accumulates over time with initial state x_0 and time t_0, following the optimal strategies π_t^i for the rest of the time steps. The value function solution for the linear-quadratic game can be found in the quadratic form [3].

$$\mathcal{V}_t^i = \frac{1}{2}x_t^T\mathcal{P}_t^i x_t + \mathcal{S}_t^T x_t + \eta_t^i \tag{5.4.1}$$

with $\mathcal{P}_{T+1}^i = 0$, $S_{T+1}^i = 0$, $\eta_{T+1}^i = 0$ to be consistent with the final value condition. The Hamilton-Jacobi-Bellman (HJB) equations for each player's value function,

$$\begin{aligned}\mathcal{V}_t^i(x_t) &= \ell_t^i(x_t, u_t) + E\Big[\mathcal{V}_{t+1}^i(x_{t+1})\Big]\\ &= \min_{u_t^i}\Bigg\{\frac{1}{2}\Bigg((x_t^T Q_t^i + 2q_t^{iT})x_t + \sum_{j=1}^{N}(u_t^T\mathcal{R}_t^{ij} + 2r_t^{ij})u_t^j\Bigg)\\ &+ E\Bigg[\mathcal{V}_{t+1}^i\Big(\mathcal{A}_t x_t + \sum_{j=1}^{N}(\mathcal{B}_t^j u_t^j + \mathcal{C}_t^j\chi_t^j)\Big)\Bigg]\Bigg\}\end{aligned} \tag{5.4.2}$$

with $\ell_t = J_t$ and the final value $\mathcal{V}_{T+1}^i(x_{t+1}) = 0$.

Alternatively, for the delayed system of a player,

$$\mathcal{V}_{t-1}^i(x_{t-1}) = J_{t-1}^i(x_{t-1}, u_{t-1}) + E\Big[\mathcal{V}_t^i(x_t)\Big] \tag{5.4.3}$$

We can write the equation as follows,

$$\begin{aligned}\mathcal{V}_{t-1} = J_{t-1}(x_{t-1}, u_{t-1}) + E\Big[&\eta_t + (\mathcal{A}_{t-1}x_{t-1} + \mathcal{B}^0_{t-1}u_{t-1} \\ &+ \mathcal{B}^1_{t-1}u_{t-1} + \mathcal{C}_{t-1}(u_{t-1}, u_{t-1-l})\xi_{t-1})\mathcal{S}_t \\ &+ (\mathcal{A}_{t-1}x_{t-1} + \mathcal{B}^0_{t-1}u_{t-1} + \mathcal{B}^1_{t-1}u_{t-1})^T\mathcal{P}_t(\mathcal{A}_{t-1}x_{t-1} + \mathcal{B}^0_{t-1}u_{t-1} + \mathcal{B}^1_{t-1}u_{t-1}) \\ &+ (\mathcal{C}_{t-1}(u_{t-1}, u_{t-1-l})\xi_{t-1})^T\mathcal{P}_t(\mathcal{C}_{t-1}(u_{t-1}, u_{t-1-l})\xi_{t-1})\Big]\end{aligned} \tag{5.4.4}$$

Now, after evaluating the expected term and using the fact that trace(UV) = trace(VU),

$$\sum_{s=1}^{p}(\mathcal{C}^0_{s,(t-1)}u_{t-1} + \mathcal{C}^1_{s,(t-1)}u_{t-1-l})^T\mathcal{P}_t(\mathcal{C}^0_{s,(t-1)}u_{t-1} + \mathcal{C}^1_{s,(t-1)}u_{t-1-l})$$

5.5 Controller Design

In this section, we focus on the approximation and the delayed system to design the optimal controller. In this case, we explore the optimal control u^{i*}_{t-1} for player i by minimizing the cost function J_t, when the optimal strategies $u^*_0, u^*_1, ..., u^*_{t-2}$ are given.

Considering the dynamic system with stochastic state and quadratic costs, the optimal control strategy $u_0, u_1, ..., u_{k-1}$ for player i satisfies the following,

$$u^{i*}_t = \pi^i_t(x_t) = \mathcal{K}^i_t x_t + \alpha^i_t + \sum_{s=0}^{min\{(K-t),l\}-1} M^{is}_t u_{k+s-l}, \quad \forall t = 0, 1, ..., K-1 \tag{5.5.1}$$

The minimum of the cost-to-go is equivalent to,

$$\mathcal{V}^i_t = \frac{1}{2}x^T_t\mathcal{P}^i_t x_t + \mathcal{S}^T_t x_t + \eta^i_t + \sum_{j=1}^{N}(\sum_{s=0}^{min\{(K-t),l\}-1} u^{jT}_{k+s-l}\mathcal{R}^{sj}_t u^j_{k+s-l} + \sum_{s=0}^{min\{(K-t),l\}-1} u^{jT}_{k+s-l}r^{js}_t) \tag{5.5.2}$$

The feedback control law is obtained by finding the minimizes of $\mathcal{V}^i_{t-1}(x_{t-1})$. Now we want to estimate the coefficients of the costs for a player as follows,

$$\mathcal{H}_{t-1} = (\mathcal{B}^0_{t-1})^T \mathcal{P}_t \mathcal{B}^0_{t-1} + \sum_{s=1}^{p} (\mathcal{C}^0_{s(t-1)})^T \mathcal{P}_t \mathcal{C}^0_{s(t-1)}, \qquad K \leq t \leq (K-l)$$

$$= (\mathcal{B}^0_{t-1})^T \mathcal{P}_t \mathcal{B}^0_{t-1} + \sum_{s=1}^{p} (\mathcal{C}^0_{s(t-1)})^T \mathcal{P}_t \mathcal{C}^0_{s(t-1)} + \mathcal{R}_t^{(l-1)(l-1)} + (\mathcal{B}^0_{t-1})^T \mathcal{R}_t^{l-1} + (\mathcal{R}_t^{l-1})^T \mathcal{B}^0_{t-1}, 0 \leq t \leq (K-l-1)$$

$$\mathcal{G}_{t-1} = -(\mathcal{H}_{t-1})^{-1} \alpha_{t-1} + \mathcal{K}_{t-1} x_{t-1} + \mathcal{M}^0_{t-1} u_{k-l-1}, \qquad K \leq t \leq (K-l) \tag{5.5.3}$$

$$\begin{aligned} \alpha_{t-1} &= -(\mathcal{H}_{t-1})^{-1}((\mathcal{B}^0_{t-1})^T \mathcal{S}_t), \qquad K \leq t \leq (K-l) \\ &= -(\mathcal{H}_{t-1})^{-1}((\mathcal{B}^0_{t-1})^T \mathcal{S}_t + r_t^{l-1}), \qquad 0 \leq t \leq (K-l-1) \end{aligned} \tag{5.5.4}$$

$$\begin{aligned} \mathcal{K}_{t-1} &= -(\mathcal{H}_{t-1})^{-1}((\mathcal{B}^0_{t-1})^T \mathcal{P}_t \mathcal{A}_{t-1}), \qquad K \leq t \leq (K-l) \\ &= -(\mathcal{H}_{t-1})^{-1}((\mathcal{P}_t \mathcal{B}^0_{t-1}) + \mathcal{R}_t^{l-1})^T \mathcal{A}_{t-1}, \qquad 0 \leq t \leq (K-l-1) \end{aligned} \tag{5.5.5}$$

$$\mathcal{M}^0_{t-1} = -(\mathcal{H}_{t-1})^{-1} [(\mathcal{B}^0_{t-1})^T \mathcal{P}_t \mathcal{B}^l_{t-1} + \sum_{s=1}^{p} (\mathcal{C}^0_{s(t-1)})^T \mathcal{P}_t \mathcal{C}^l_{s(t-1)}], \qquad K \leq t \leq (K-l)$$

$$= -(\mathcal{H}_{t-1})^{-1} [(\mathcal{B}^0_{t-1})^T \mathcal{P}_t \mathcal{B}^l_{t-1} + \sum_{s=1}^{p} (\mathcal{C}^0_{s(t-1)})^T \mathcal{P}_t \mathcal{C}^l_{s(t-1)} + (\mathcal{R}_t^{l-1})^T \mathcal{B}^1_{t-1}], \quad 0 \leq t \leq (K-l-1)$$

$$\mathcal{M}^s_{t-1} = -(\mathcal{H}_{t-1})^{-1}(\mathcal{B}^0_{t-1})^T\mathcal{R}^{s-1}_t, \quad K \leq t \leq (K-l), \quad s = 1,2,..,k-t$$
$$= -(\mathcal{H}_{t-1})^{-1}[(\mathcal{B}^0_{t-1})^T\mathcal{R}^{s-1}_t + \mathcal{R}^{(l-1)(s-1)}_t], \quad 0 \leq t \leq (K-l-1), \quad s = 1,2,..,l-1 \tag{5.5.6}$$

Considering the coefficients, the cost function components can be written as follows,

$$r^0_{t-1} = (\mathcal{B}^1_{t-1})^T\mathcal{S}_t - (\mathcal{M}^0_{t-1})^T\mathcal{H}_{t-1}\alpha_{t-1}$$
$$r^s_{t-1} = r^{s-1}_t - (\mathcal{M}^s_{k-1})^T\mathcal{H}\alpha_{t-1}, \quad s = 1,2,...,mink-t,l-1$$
$$\mathcal{R}^0_{t-1} = (\mathcal{B}^1_{t-1})^T\mathcal{P}_t(\mathcal{A}_{t-1}) - (\mathcal{M}^0_{t-1})^T\mathcal{H}_{t-1}\mathcal{K}_{t-1} \tag{5.5.7}$$
$$\mathcal{R}^s_{t-1} = (\mathcal{A}^1_{t-1})^T\mathcal{R}^{s-1}_t - (\mathcal{K}_{t-1})^T\mathcal{H}_{t-1}\mathcal{M}^s_{t-1}, \quad s = 1,2,...,mink-t,l-1$$

Following the cost representations, the optimal control for player i can be written as follows,

$$u_{t-1} = -(\mathcal{H}_{t-1})^{-1}\mathcal{G}_{t-1} = \alpha_{t-1} + \mathcal{K}_{t-1}x_{t-1} + \mathcal{M}^0_{t-1}u_{t-l-1} \tag{5.5.8}$$

The value function for player i becomes,

$$\mathcal{P}_{t-1} = Q_{t-1} + (\mathcal{A}_{t-1})^T\mathcal{P}_t\mathcal{A}_{t-1} - (\mathcal{K}_{t-1})^T\mathcal{H}_{t-1}\mathcal{K}_{t-1}$$
$$\mathcal{S}_{t-1} = q_{t-1} + \mathcal{S}^T_t\mathcal{A}_{t-1} - (\alpha_{t-1})^T\mathcal{H}_{t-1}\mathcal{K}_{t-1} \tag{5.5.9}$$
$$\eta_{t-1} = \eta_t - \alpha^T_{t-1}\mathcal{H}_{t-1}\alpha_{t-1}$$

5.6 Numerical Results

The algorithm structure in this chapter is analogous to that of Chapter-3 with minor variations to accommodate delayed feedback control. The numerical results are generated for two types of interaction game. One is the regular stochastic game with the control dependent noise and the other follows a feedback delay for the interacting players. In the second type of game formulation, we try to analyze the implication of delayed communication for both players in an interaction event. Notably the effect of delayed feedback is tested with more players in the game environment. We explore the three player and four player interaction game with the delayed feedback. We consider two types of initial conditions for the player in two instances. Based on the setup, the time when players get their information could just be one or more time step later than the current or running time interval.

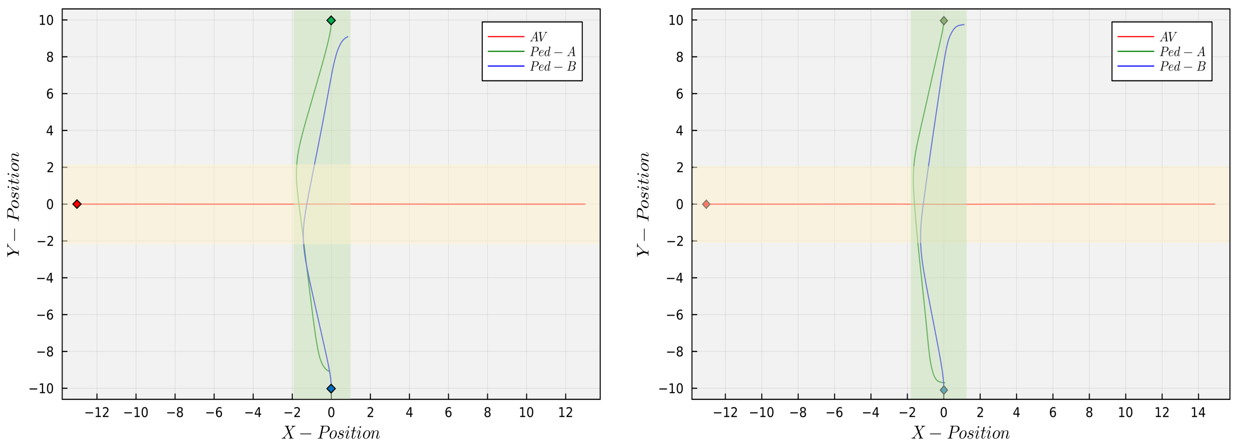
5.6.1 3-Player Interaction Game

In the three player interaction game, we consider one AV and two pedestrians. Here we assume one AV traveling on the roadway and interacting with two pedestrians crossing from opposite directions over the crosswalk.

Table 5.1: Game Parameters for 3-Player Game

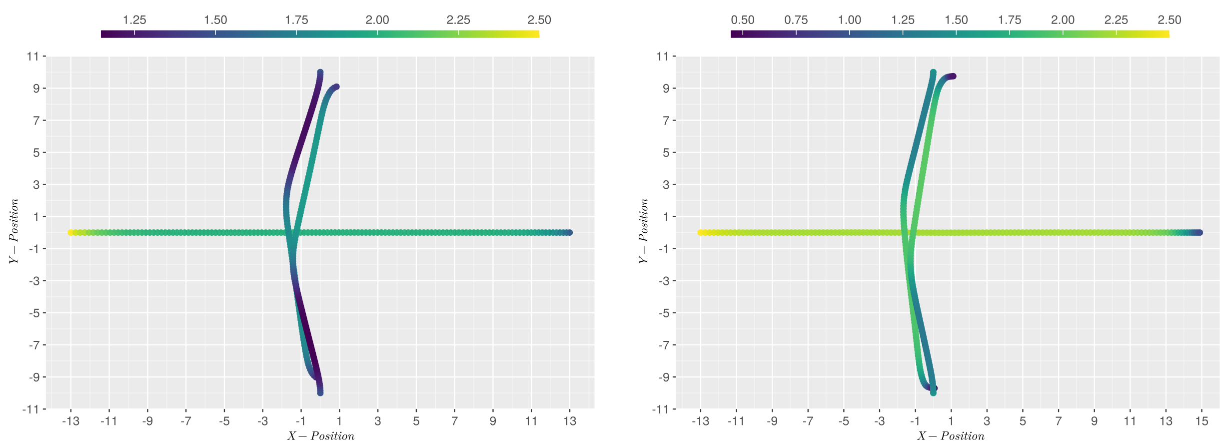
Parameters	Autonomous Vehicle	Pedestrian-A	Pedestrian-B
Initial State	(-13, 0, 0, 2.5)	(0, 9, $-\pi/2$, 1.2)	(0, -9, $\pi/2$, 1.2)
Brownian Motion Noise	0.1	0.1	0.1
Mean	1.2	1.2	1.2
Standard Deviation	1.3	1.3	1.3
Collision Radius (m)	2	0.8	0.8
Delay (s)	0.1	0.1	0.1

The game trajectory describes the interaction between three players. The interaction is sensitive to initial conditions as AV's higher speed and short distance from the crosswalk will likely encourage no major interaction event between AV and crossing pedestrians. Thus, the game is formulated in a way to promote an interaction between the players. The interaction dynamics of AV and pedestrians are governed by the transformed dynamics of the players and delayed feedback strategy. The difference in trajectory with respect to the regular game with similar initial condition is visible (Fig 5.1). The players with delayed feedback information travels further compared to the regular game with same initial condition and time horizon. From the conceptual perspective the decision making period for the players is delayed by a fixed interval or time step.



(a) Player Trajectory in a Regular Game (b) Player Trajectory with Delayed Feedback

Figure 5.1: Three-Player Interaction Game Trajectory

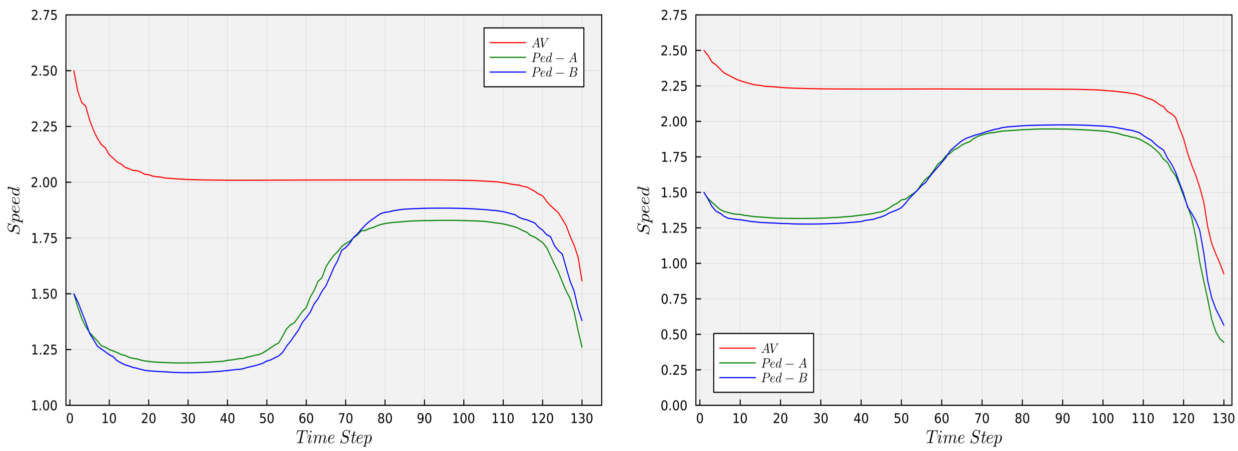


(a) Player Trajectory with Speed in a Regular Game (b) Player Trajectory with Speed in Delayed Feedback

Figure 5.2: Three-player Game Trajectory with Speed

In both type of three-player games the pedestrians interact with the AV in a safe and conservative way. In the game scenario, the proactive sense of safety from pedestrians perspective is induced by maintaining a larger collision radius and assigning comparatively large weight on the collision cost component. The underlying behavior of the pedestrian remains same. However, due to the weighted cost component and preferred crossing speed, pedestrians behave more conservatively. They maintain a higher safe distance from other players including other pedestrians and AV. With the delayed feedback the players adjust their control one step after the current time step. The pedestrians start from a slow walking speed and gradually accelerate once the AV has crossed the potential conflict point over the crosswalk (Fig 5.1). However, the pedestrians in the delayed feedback game has much higher walking speed compared to the regular game scenario (Fig 5.2). This shift in behavior is derived from the delay in the communication between players.

The speed profile of the players also show that the players interacting with a communication delay has a higher speed compared to the regular stochastic interaction scenario 5.3.



(a) Player Speed in a Regular Game (b) Player Speed Trajectory with Delayed Feedback

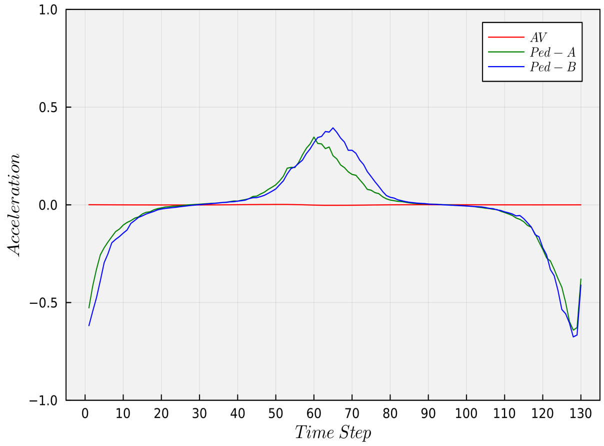
Figure 5.3: Three-Player Game Speed Trajectory

The control dependent noise is visible in the acceleration pattern of the players, specifically for pedestrians as they adjust their speed with respect to the initial condition during crossing. The pattern is different for pedestrians in a delayed feedback game.

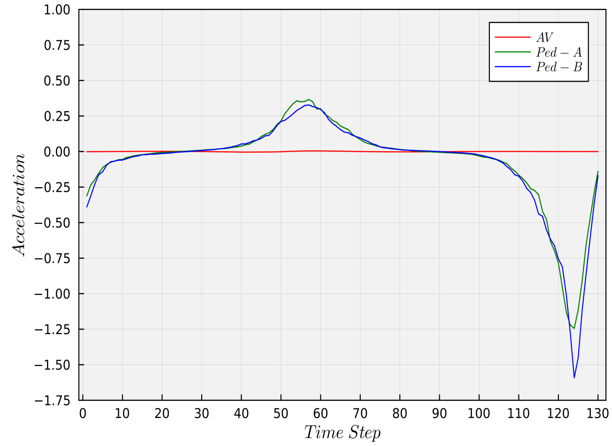
5.6.2 4-Player Interaction Game

In this four player interaction game, one AV interacts with three other pedestrians.

In this case, both the pedestrian cross over the crosswalk and negotiate the space with the nearest AV. The pedestrians cross after the AV has passed over the crosswalk (Figure 5.7).

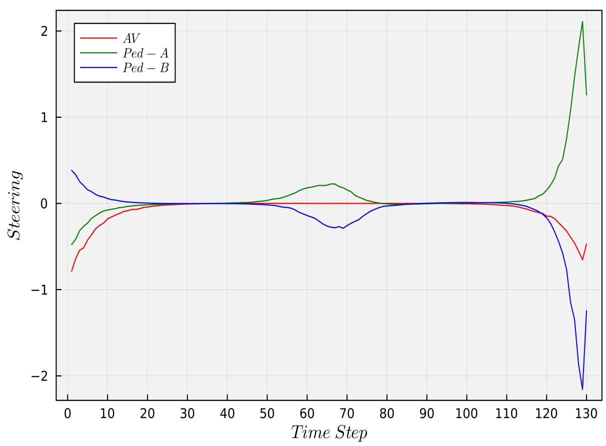


(a) Player Acceleration in a Regular Game

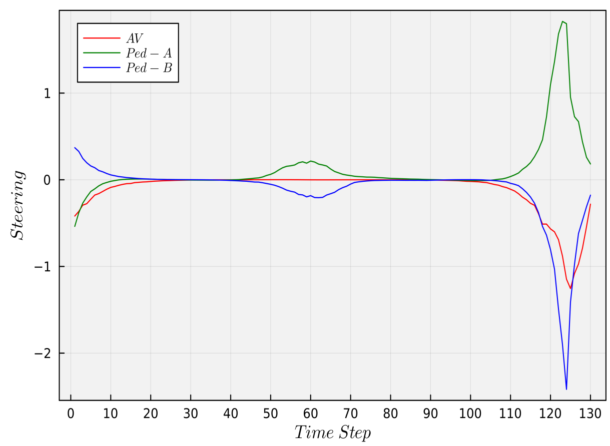


(b) Player Acceleration with Delayed Feedback

Figure 5.4: Three-Player Game Acceleration Trajectory



(a) Player Steering in a Regular Game



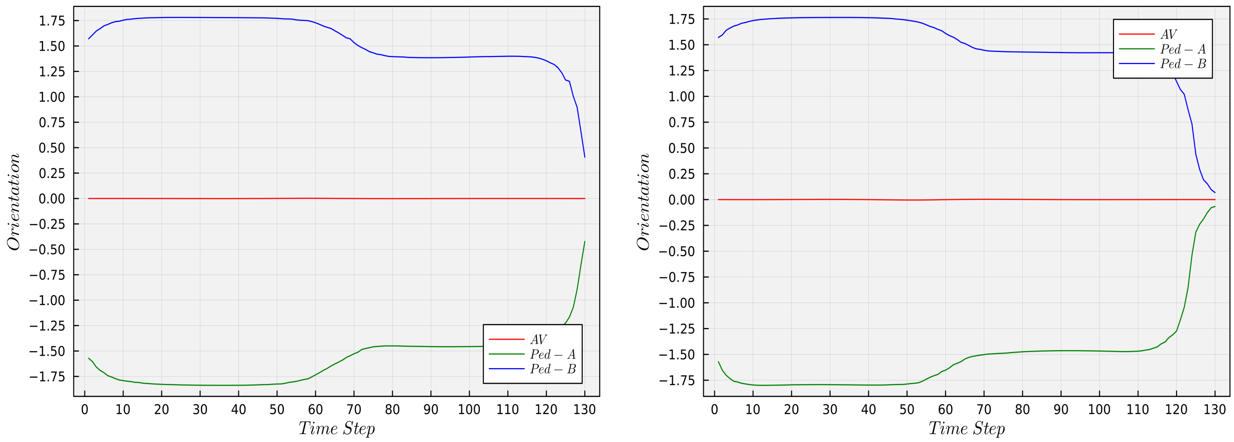
(b) Player Steering with Delayed Feedback

Figure 5.5: Three-Player Game Steering Trajectory

Table 5.2: Game Parameters for 4-Player Game

Parameters	AV-A	Pedestrian-A	Pedestrian-B	Pedestrian-C
Initial State	(-15, 0, 0, 2.5)	(1, 10, $-\pi/2$, 1.2)	(0, 10, $-\pi/2$, 1.2)	(-1, 10, $-\pi/2$, 1.2)
Brownian Noise	0.1	0.1	0.1	0.1
Mean	1	1	1	1
Standard Deviation	1.2	1.2	1.2	1.2
Collision Radius (m)	1.75	0.75	0.75	0.75
Delay (s)	0.1	0.1	0.1	0.1

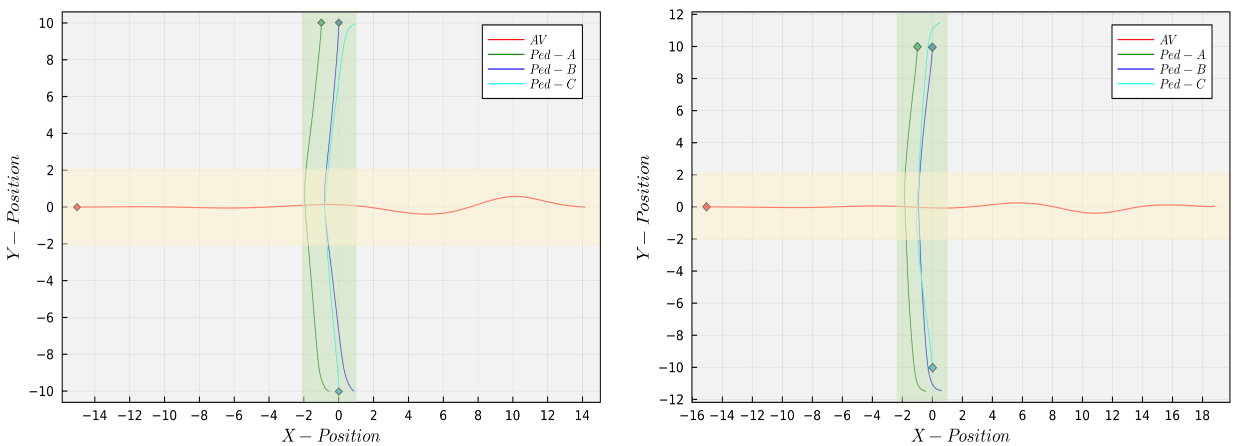
The control dependent noise is visible in the acceleration pattern of the players (Figure 5.10), specifically for the pedestrians as they adjust their speed with respect to the initial condition. The



(a) Player Orientation in a Regular Game

(b) Player Orientation with Delayed Feedback

Figure 5.6: Three-Player Game Orientation Trajectory



(a) Player Trajectory in a Regular Game

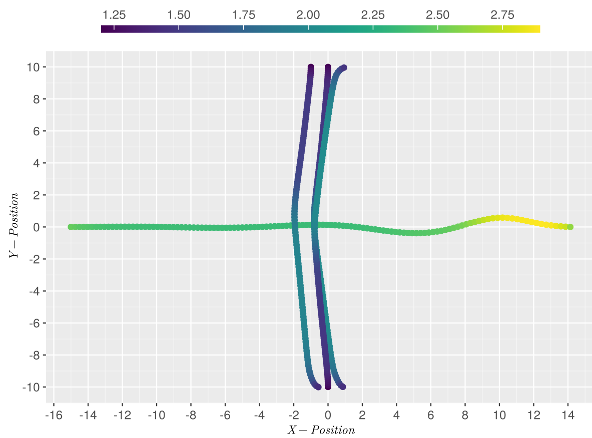
(b) Player Trajectory with Delayed Feedback

Figure 5.7: Four-Player Interaction Game Trajectory

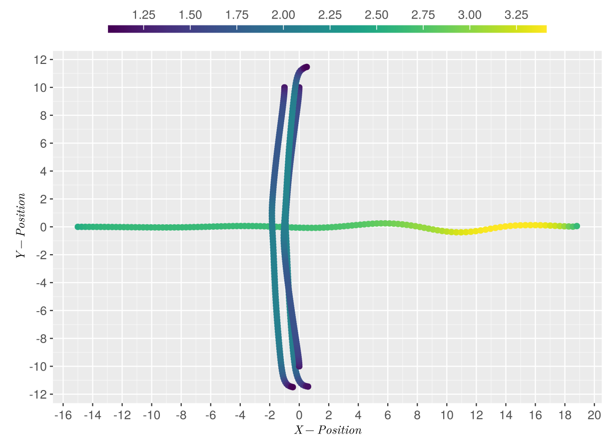
delay in the feedback control just shifts the control system of the interacting players, keeping all the other features same.

5.7 Quantitative and Qualitative Evaluation

The quantitative and qualitative evaluation in this section follows the same methodology detailed in Chapter-3.

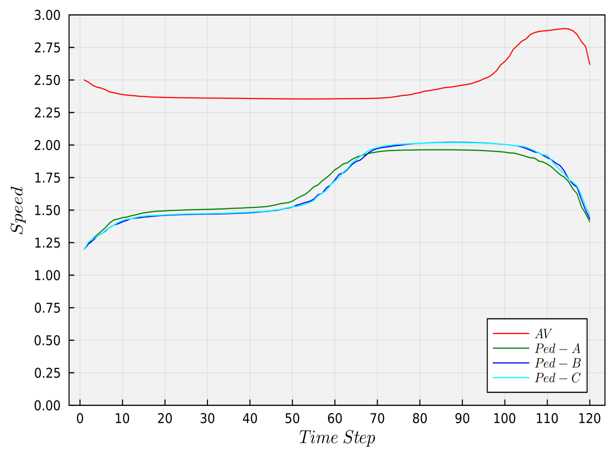


(a) Player Trajectory with Speed in a Regular Game

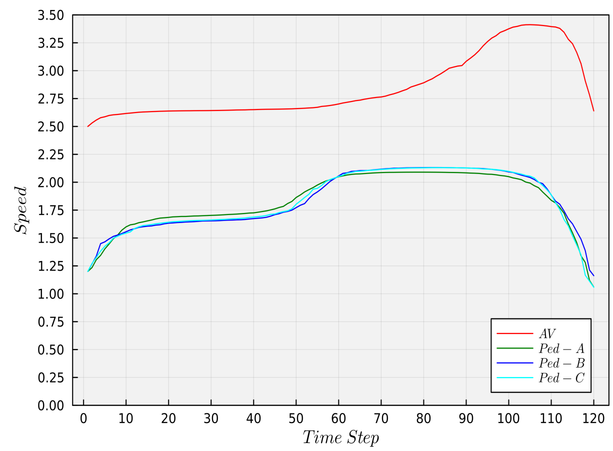


(b) Player Trajectory with Speed in Delayed Feedback

Figure 5.8: Four-Player Interaction Game Trajectory with Speed



(a) Player Speed in a Regular Game



(b) Player Speed with Delayed Feedback

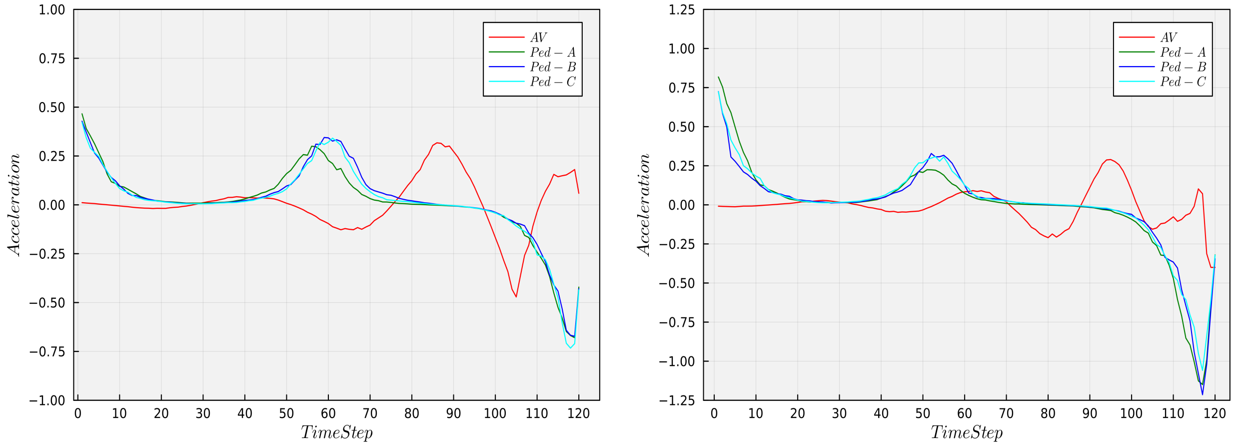
Figure 5.9: Player Speed Trajectory

5.7.1 Quantitative Evaluation

Table 5.3: Quantitative Evaluation Summary for Delayed Feedback Game

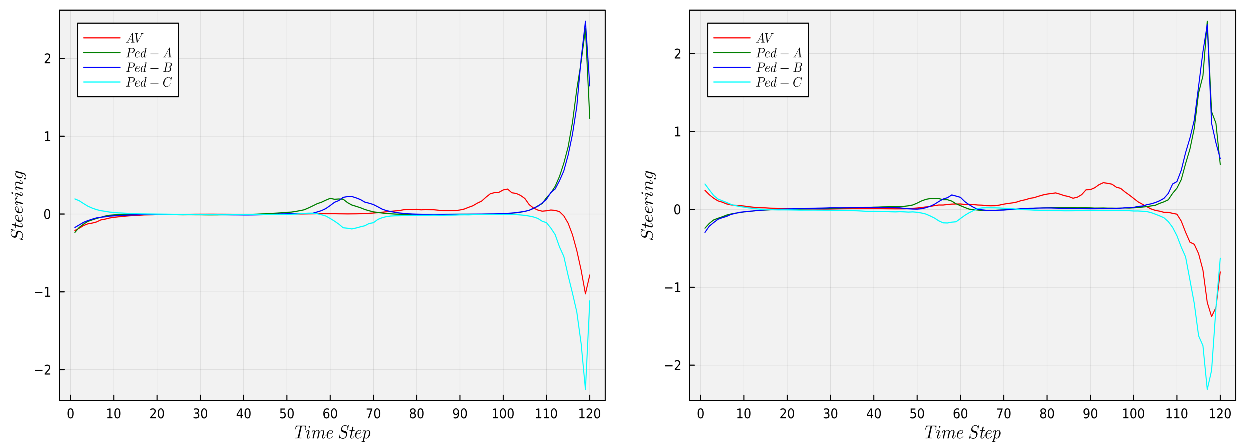
		Avg Crossing (s)				Avg Simulated (s)		Difference (s)	
Scenario	Sample	Mix	Con	Agr	Int	Crs	Int	Crs	Int
A (1)	16	10.331	10.750	9.075	2.994	10.412	3.524	0.081	0.530
B (2)	10	10.380	10.539	8.950	3.035	10.625	3.824	0.245	0.789
C (3)	7	10.519	10.506	9.200	3.048	10.824	3.925	0.305	0.877
D (4)	6	10.758	10.810	9.233	3.079	11.561	3.815	0.803	0.736
E (5)	5	10.972	11.068	9.133	3.092	11.859	3.719	0.887	0.627

*Avg - Average, s - Seconds, Mix - Mixed, Con - Conservative, Agr - Aggressive, Int - Interaction, Crs - Crossing.



(a) Player Acceleration in a Regular Game (b) Player Acceleration with Delayed Feedback

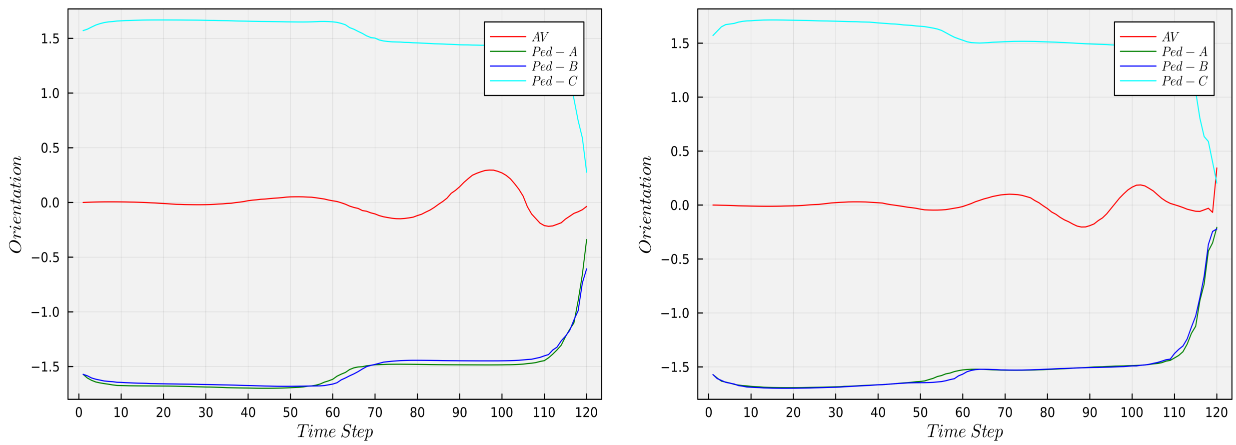
Figure 5.10: Player Acceleration Trajectories



(a) Player Steering in a Regular Game (b) Player Steering with Delayed Feedback

Figure 5.11: Player Steering Trajectories

The results indicate that the average crossing and interaction time for each simulated scenario are relatable to the real-world data. For instance, the difference between simulated and real-world crossing time for two pedestrians crossing from the opposite direction is 0.245 seconds. For the four pedestrians crossing scenario, the difference between average crossing time is about 0.803 seconds. The interaction time difference for multiple pedestrian scenarios (A, B, C, D, E) ranges from 0.530 to 0.877 seconds (Table 5.3). The maximum difference in average crossing time (0.887) is recorded for scenario E with 5 pedestrians crossing as a group. The average interaction time



(a) Player Orientation in a Regular Game (b) Player Orientation with Delayed Feedback

Figure 5.12: Player Orientation Trajectories

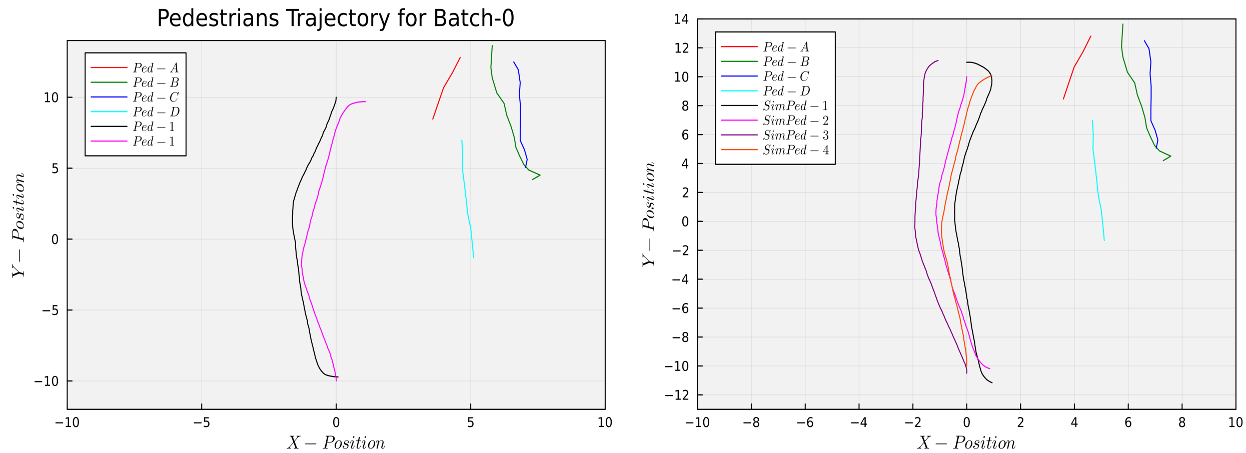
remains identical across the scenarios.

We can see a subtle increase in the difference between real-world and simulated trajectories. This effect exists because the feedback introduced in both players' control directly impacts the crossing and interaction time. The finite delay in communication strategy increases the crossing and interaction time by a small magnitude compared to our previous research results in Chapters 3 and 4. This increment may result from the increase in delay from both players' controls instead of a mixed strategy. It will be interesting for future research to test a hypothesis indicating a possible balance case for mixed strategy effect. For instance, we could check if the increase from the feedback delay is mitigated when AV has regular feedback and pedestrians have delayed information.

5.7.2 Qualitative Evaluation

We evaluate the delayed feedback game using qualitative pedestrian trajectories. The motion pattern of real-world pedestrians relates to the simulated trajectories with noise. The finite time delay introduced in the stochastic game system affects the interaction and crossing time, indicating that the game is sensitive to the control strategy followed by the players.

The simulated pedestrian trajectories visually match the motion pattern of the real-world data



(a) Pedestrian Trajectories from 3-player Delayed Feedback Game (b) Pedestrian Trajectories from 5-player Delayed Feedback Game

Figure 5.13: Qualitative Evaluation of the Simulated Trajectories with respect to ETH data

with some differences (Fig 5.13). The negligible visual difference with the real-world trajectories occurs as the pedestrians are not walking in a controlled crosswalk environment in the presence of a vehicle nearby. The effect of delayed feedback is not much significant from the qualitative comparison as the pedestrian motion pattern remains identical.

5.8 Conclusion

The proposed game investigates the specifics of a pedestrian crossing in urban streets, considering delayed information patterns for both types of players (AV and pedestrians). The game incorporates noise into the control variable to reflect the real-world stochastic condition and accommodates feedback delay as part of the communication strategy. Notably, the feedback delay introduces some behavior shifts in the player compared to the regular stochastic game introduced in Chapter 3. In Chapter 5, we augment the previous interaction game framework with delayed feedback control features to accommodate the players' communication strategy variation, which may arise in a real-world interaction. This addition ensures that the proposed game framework is flexible enough to facilitate a different communication strategy with relevant changes within the game formulation.

The game proposes a finite feedback delay in communication between the players with a noise on the controls. This research uses the game-theoretic framework to detail pedestrian crossing behavior and explore the change in player response with the finite delay in the feedback control strategy. The numerical results show a marked effect of the finite delay on the player control strategy compared to the stochastic game proposed in Chapter-3. The proposed game admits Nash's solution since players follow the same delay. Notably, a higher delay or delay greater than multiple time steps moves away from the Nash equilibrium due to the game formulation and constraints.

We used the same quantitative and qualitative metrics from Chapter-3 to evaluate the proposed game with feedback delay. The quantitative results indicate that the average interaction time of pedestrians has a marked difference of 0.5 to 0.9 seconds approximately between real-world and game-simulated trajectories. This difference in behavior results from the delayed communication between the players, which indicates that the players take more time to decide and negotiate space during an interaction over the crosswalk. Notably, the difference in average crossing time between the simulated and real-world pedestrians remains almost the same as in Chapter-3.

The game framework with delayed feedback shows that the interaction time is affected by the finite delay in communication between the players. This implies that type of communication used to relay the control strategies between players is critical for the interaction framework of AV in urban settings. For future research direction, we can look towards mixed strategies for different types of players. For instance, we can model the AV with a regular feedback control strategy and design the pedestrian model with delayed feedback. Although the current game framework has the flexibility to incorporate changes in the mechanism design or formulation, Nash equilibrium based on the premise of mixed feedback control strategy remains an open question. Moreover, the current research trend has shifted towards using machine learning methods to obtain classical features in a sizable macroscopic version known as a mean-field game.

Chapter 6

Conclusion and Future Work

6.1 Conclusions

This dissertation presents a stochastic game framework to capture pedestrian behavior for autonomous driving in urban settings. The game development is divided into three parts (Chapters 3, 4, and 5) to complement the proposed interaction framework. The first part (Chapter 3) is the primary component that captures the pedestrians' behavior to improve autonomous driving in urban settings. The game focuses on the interaction scenarios specific to pedestrians over the crosswalk in the presence of an AV. Based on differential game theory, we propose a novel approach to capture the pedestrians' stochastic behaviors and describe the nature of the interaction (cooperative and competitive) with AV. We model AV and pedestrians using the kinematic bicycle and unicycle models. Then, we linearize the nonlinear dynamics of the players (pedestrians and AV). We use Brownian motion noise in the state dynamics equation to reflect the stochasticity in the players' behavior. Specifically, we test the players' behavior with noise on the control variables. Player-specific behavior and control features in a linear quadratic optimal control formulation inspired the control-dependent noise setup. The cost functions for each type of player are designed based on collision, acceleration, and boundary cost. The constraints on the players' state and control are included as a penalty in the combined cost function. The proposed game mechanism can also handle different crossing behavior of pedestrians, including conservative, aggressive, and group. The interaction game admits the Nash equilibrium solution in a feedback form. The solution algorithm closely resembles the architecture of the iterative Linear Quadratic Regulator.

Although the developed game model in Chapter 3 can handle a specific noise in dynamics, the model is not robust against external noise in urban settings. To tackle this limitation, we can build a robust game assuming the range of possible noise and insert it deterministically into the system. Thus, in Chapter 4, we test the robustness of the proposed interaction game using an adversarial input.

We evaluate the robust game from qualitative pedestrian trajectories. The motion pattern of real-world pedestrians matches the simulated trajectory with noise. The robust game filters the noise and generates a smooth trajectory for AV and pedestrians. The flexibility of the smoothing effect is controlled through the weightage of the adversarial input into the system. This implies that the AV can fine-tune the robust control to adjust to real-world pedestrian motion patterns and interact safely. The flexibility of the robust model developed in this chapter will aid in smoothing AV trajectory while maintaining the noise element in the pedestrians' motion to reflect real-world conditions.

The robust game is set up in a finite horizon and follows a feedback control strategy. The feedback control implies that immediate information is available to each player, which helps to identify a safe and optimal strategy for any given interaction scenario. However, in real-world scenarios, each interacting player can access limited or delayed information about other players' states. For instance, pedestrians are limited by peripheral vision and other constraints, whereas vehicle control may vary with the ever-changing urban scene. Thus, a strategy with delayed information available to the players will reflect the interaction event more accurately. Therefore, in Chapter 5, we explore the delayed feedback strategy for pedestrians and AV to update the current framework. We used the stochastic game framework in Chapter 3 to detail pedestrian crossing behavior and explore the change in player response with the finite delay in the feedback control strategy.

The numerical results show a marked effect of the finite delay on the player control strategy compared to the stochastic game proposed in Chapter-3. The proposed game admits Nash's solution since players follow the same delay. We used the same quantitative and qualitative metrics from Chapter-3 to evaluate the proposed game with feedback delay. The quantitative results indicate that the average interaction time of pedestrians has a marked difference between real-world and game-simulated trajectories, resulting from the delayed communication between the players. Moreover, the difference in average crossing time between the simulated and real-world pedestrians remains almost the same as in Chapter-3. The results indicate that the players with delayed information take more time to decide and negotiate space during an interaction over the crosswalk.

6.2 Limitations

The limitations of this research are two folds. One relates to the formulation difficulty, and the other relates to the availability of the model-specific validation data. For instance, we measured the data for the pedestrian crossing considering the time interval to cross the approaching vehicle over the crosswalk from one specific approach of the intersection. However, the vehicle arriving from the opposite side may influence the pedestrian's behavior. In short, the crossing pedestrian's behavior may change midway on the crosswalk, given a vehicle waiting or approaching the intersection from the other side. The proposed models cannot capture this distinction that may arise in real-world interaction scenarios. For this research, we focus our model on explaining pedestrian behavior and possible interaction pattern for AV from one approach.

Understanding the type and time of interaction is a complex and subjective task in data collection as it relies heavily on the observer's interpretation of the interaction event. Moreover, the game condition and real-world scenario may vary depending on the motorists' (drivers) and pedestrians' communication and behavior. In this case, we are documenting an interaction between a regular motorist and pedestrians in the real world. In contrast, the interaction between AV and pedestrians in urban scenarios may vary depending on the communication medium, pedestrian behavior, and driving style. For example, the AV can sense the presence of incoming pedestrians and may slowly approach the crosswalk to give enough time for pedestrians to cross and save energy without stopping. Sometimes, it may speed up to cross before the pedestrian starts to walk near the crosswalk. These behaviors are typically absent in regular motorists as the sensor range is limited and subjective to the person driving the vehicle. Thus, traditional stop-interact-go situations may change with the adoption of AV in urban interaction scenarios.

This research proposes a decision framework for autonomous driving to capture pedestrian-specific behavior and safely interact in urban scenarios. However, it only proposes a model to solve a sub-problem of the complex interaction issues arising in real-world interaction scenarios, where the AV will interact with multiple types of road users, including bicyclists and other motorists.

In terms of robust formulation in Chapter-4, this research captures the external noise as an adversarial input and mitigates the noise effect on the players' trajectory. However, drawing an exact range of noise for urban scenarios is complex. The more accurate, robust interaction framework for AV driving should include a different set of noise for different types of road users in urban scenarios.

The communication pattern or control strategy for players plays a critical role in the game formulation. In Ideal conditions, the pedestrians should have access to small information compared to the AV equipped with a wide array of sensors. Thus, the formulation would be more realistic for the proposed interaction framework when each player has access to a different type of information.

A possible validation method that we may use is a tracking error for the real-world trajectory of a pedestrian or vehicle. Using the vehicle trajectory would be practical as the vehicle motion model is more reliable in tracking, whereas pedestrian motion is unpredictable. However, using pedestrian motion tracking will be unrealistic in some instances as we are developing a model to handle challenging interaction scenarios in urban streets. In other words, we can feed the model with different pedestrian trajectories to develop an expected range of pedestrian behavior in urban streets.

For this type of research, we should allocate a funding stream to collect relevant data from urban streets and merge it with other public data sources to develop a representative sample. Then, after collecting, cleaning, and annotating the dataset, we could employ inverse optimal control to learn from the pedestrian-specific trajectories. Finally, we can design the interaction game with AV using the learning parameters from the real-world pedestrian model.

6.3 Future Work and Open Questions

In this game formulation, we use simple kinematic bicycle dynamics and unicycle dynamics for the pedestrian model to simplify the complex interaction scenario. In future research, we can add slip angle to vehicle dynamics and other motion equations to accurately reflect real-world AV motion control. Similarly, we can introduce complex behavioral features (reactive and proactive) and additional motion states for pedestrians.

Next, we can build on the research by introducing more complex urban scenarios to negotiate right-turn and left-turn in stop sign intersections. Moreover, the interesting behavior in the presence of pedestrians at the roundabout can also be explored by extending the game mechanism design or formulation.

Noise is also a critical component of the system. We can test the flexibility using different types of noise in the system and check the variability of the output and stability of the game system.

We can conduct a detailed sensitivity analysis listing the range of pedestrian motion patterns by controlling the weightage of the adversarial input in the robust game framework.

Finally, communication pattern or control strategy specific to each player is a critical part of the game framework. We have seen from the results in Chapter 5 that the player behavior and game equilibrium are influenced by the type of strategy the players follow. Therefore, in ideal research conditions, the game should incorporate a mixed strategy for the players, indicating that each player will access the state information at different intervals to reflect more accurate real-world interaction scenarios. In literature, this type of game is called a mixed strategy game. However, the Nash feedback solution of this type of game remains an open question.

Appendices

Appendix-A

Table 6.1: Recorded Field Data for Scenario A

Scenario	Location	Sample No	Pedestrian No	Pedestrian Type	Crossing Time	Interaction Time
A	1	1	1	Conservative	10	3
A	1	2	1	Aggressive	9	2.8
A	1	3	1	Conservative	10.4	2.6
A	1	4	1	Aggressive	9.3	2.8
A	1	5	1	Aggressive	9.1	3.1
A	1	6	1	Conservative	10.2	2.7
A	1	7	1	Conservative	10	3.1
A	1	8	1	Conservative	10.4	3.2
A	1	9	1	Conservative	11	3.5
A	1	10	1	Conservative	11.5	3
A	1	11	1	Aggressive	8.9	3.2
A	2	12	1	Conservative	11	2.9
A	2	13	1	Conservative	11.2	3
A	2	14	1	Conservative	12	3
A	2	15	1	Conservative	10.5	3.2
A	2	16	1	Conservative	10.8	2.8

*Sample size - 16, No. of Pedestrians in Each Sample - 1

Table 6.2: Recorded Field Data for Scenario B

Scenario	Location	Sample No	Pedestrian No	Pedestrian Type	Crossing Time	Interaction Time
B	1	1	1	Conservative	10.2	3.1
B	1	1	2	Conservative	10.1	2.9
B	1	2	1	Conservative	10.5	3
B	1	2	2	Aggressive	9.1	2.8
B	1	3	1	Conservative	10.8	3.1
B	1	3	2	Conservative	11	3.2
B	1	4	1	Conservative	10.1	3
B	1	4	2	Conservative	10.5	3
B	1	5	1	Conservative	10.6	3.1
B	1	5	2	Conservative	11	2.9
B	1	6	1	Aggressive	8.8	3
B	1	6	2	Conservative	10.2	2.9
B	2	7	1	Conservative	10.3	3
B	2	7	2	Conservative	10.5	3.3
B	2	8	1	Conservative	10	3.1
B	2	8	2	Conservative	10.2	2.8
B	2	9	1	Conservative	10.3	3
B	2	9	2	Conservative	11.2	3.4
B	2	10	1	Conservative	10.2	3
B	2	10	2	Conservative	12	3.1

*Sample size - 10, No. of Pedestrians in Each Sample - 2

Table 6.3: Recorded Field Data for Scenario C

Scenario	Location	Sample No	Pedestrian No	Pedestrian Type	Crossing Time	Interaction Time
C	1	1	1	Conservative	10.2	3
C	1	1	2	Conservative	10.6	2.8
C	1	1	3	Conservative	11	3
C	1	2	1	Conservative	10.4	3.2
C	1	2	2	Aggressive	9.2	2.6
C	1	2	3	Conservative	10.9	3
C	1	3	1	Conservative	10.8	3.1
C	1	3	2	Conservative	11.3	3.2
C	1	3	3	Conservative	11.8	3.4
C	1	4	1	Conservative	10.1	3
C	1	4	2	Conservative	10.7	2.9
C	1	4	3	Conservative	11.2	3
C	1	5	1	Aggressive	9.4	3.1
C	1	5	2	Conservative	10.2	3.2
C	1	5	3	Conservative	10.8	3
C	1	6	1	Conservative	10.3	3
C	1	6	2	Aggressive	9	3.2
C	1	6	3	Conservative	10.7	3.3
C	2	7	1	Conservative	10.2	3
C	2	7	2	Conservative	10.6	2.9
C	2	7	3	Conservative	11.5	3.1

*Sample size - 7, No. of Pedestrians in Each Sample - 3

Table 6.4: Recorded Field Data for Scenario D

Scenario	Location	Sample No	Pedestrian No	Pedestrian Type	Crossing Time	Interaction Time
D	1	1	1	Conservative	10.2	3
D	1	1	2	Conservative	10.5	2.8
D	1	1	3	Conservative	11	3.2
D	1	1	4	Conservative	11.2	3
D	1	2	1	Conservative	10.3	3.3
D	1	2	2	Aggressive	9.1	3
D	1	2	3	Conservative	10.6	3.1
D	1	2	4	Conservative	11.5	2.9
D	1	3	1	Conservative	10.5	3
D	1	3	2	Conservative	10.8	3
D	1	3	3	Conservative	11.2	3.2
D	1	3	4	Conservative	12	3.2
D	1	4	1	Conservative	10.5	3.4
D	1	4	2	Aggressive	9.1	3
D	1	4	3	Aggressive	9.5	3.1
D	1	4	4	Conservative	11.5	3
D	1	5	1	Conservative	10.8	3
D	1	5	2	Conservative	11.2	2.8
D	1	5	3	Conservative	11	3.2
D	1	5	4	Conservative	12.3	3
D	1	6	1	Conservative	10.3	3
D	1	6	2	Conservative	10.6	3.5
D	1	6	3	Conservative	11	3
D	1	6	4	Conservative	11.5	3.2

*Sample size - 6, No. of Pedestrians in Each Sample - 4

Table 6.5: Recorded Field Data for Scenario E

Scenario	Location	Sample No	Pedestrian No	Pedestrian Type	Crossing Time	Interaction Time
E	1	1	1	Conservative	10.2	3.1
E	1	1	2	Conservative	10.6	3
E	1	1	3	Conservative	10.8	3.3
E	1	1	4	Conservative	11.5	2.8
E	1	1	5	Conservative	12.1	3
E	1	2	1	Conservative	10.5	3.5
E	1	2	2	Conservative	11.5	3
E	1	2	3	Conservative	12	3
E	1	2	4	Aggressive	9.2	3.2
E	1	2	5	Conservative	12.4	3.3
E	1	3	1	Conservative	10	3
E	1	3	2	Conservative	10.5	3.1
E	1	3	3	Conservative	10.9	3
E	1	3	4	Conservative	11.5	2.9
E	1	3	5	Aggressive	8.9	3.4
E	1	4	1	Conservative	10.3	3
E	1	4	2	Conservative	10.8	3.3
E	1	4	3	Conservative	11.3	3
E	1	4	4	Conservative	11.8	3
E	1	4	5	Conservative	12.5	3.2
E	2	5	1	Conservative	10.5	3.2
E	2	5	2	Conservative	11.2	3
E	2	5	3	Conservative	11.8	3.1
E	2	5	4	Conservative	12.2	2.9
E	2	5	5	Aggressive	9.3	3

*Sample size - 5, No. of Pedestrians in Each Sample - 5

References

[1] J. M. Anderson, N. Kalra, K. D. Stanley, P. Sorensen, C. Samaras, and O. Oluwatola, "Autonomous vehicle technology: A guide for policymakers. santa monica, ca: Rand corporation," 2016.

[2] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 187–210, 2018.

[3] T. Basar and G. Zaccour, *Handbook of Dynamic Game Theory.* Springer, 2018.

[4] S. R. Musabbir and M. Zhang, "Safety effects of the yellow light border (ypb) pedestrian signal: An evaluation or an empirical study on crossing behavior at a signalized intersection with high-volume pedestrians," tech. rep., 2020.

[5] X. Geng, H. Liang, B. Yu, P. Zhao, L. He, and R. Huang, "A scenario-adaptive driving behavior prediction approach to urban autonomous driving," *Applied Sciences*, vol. 7, no. 4, p. 426, 2017.

[6] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Are they going to cross? a benchmark dataset and baseline for pedestrian crosswalk behavior," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 206–213, 2017.

[7] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking system with kalman filtering," in *IEEE Intelligent Vehicles Symposium, 2004*, pp. 584–589, IEEE, 2004.

[8] M. Liu, W. Zeng, P. Chen, and X. Wu, "A microscopic simulation model for pedestrian-pedestrian and pedestrian-vehicle interactions at crosswalks," *PLoS one*, vol. 12, no. 7, p. e0180992, 2017.

[9] J. Park, *Safe controller design for intelligent transportation system applications using reachability analysis.* PhD thesis, The Ohio State University, 2013.

[10] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[11] I. Wolf, "The interaction between humans and autonomous agents," in *Autonomous driving*, pp. 103–124, Springer, 2016.

[12] A. A. Hassan and H. A. Rakha, "A fully-distributed heuristic algorithm for control of autonomous vehicle movements at isolated intersections," *International Journal of Transportation Science and Technology*, vol. 3, no. 4, pp. 297–309, 2014.

[13] I. H. Zohdy and H. A. Rakha, "Intersection management via vehicle connectivity: The intersection cooperative adaptive cruise control system concept," *Journal of Intelligent Transportation Systems*, vol. 20, no. 1, pp. 17–32, 2016.

[14] J. Lee and B. Park, "Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, pp. 81–90, March 2012.

[15] D. Zhao, H. Lam, H. Peng, S. Bao, D. J. LeBlanc, K. Nobukawa, and C. S. Pan, "Accelerated evaluation of automated vehicles safety in lane-change scenarios based on importance sampling techniques," *IEEE transactions on intelligent transportation systems*, vol. 18, no. 3, pp. 595–607, 2016.

[16] M. Aust, "Evaluation process for active safety functions: Addressing key challenges in functional, formative evaluation of advanced driver assistance systems," *Chalmers University of Technology*, pp. 1243–1262, 2012.

[17] M. Akamatsu, P. Green, and K. Bengler, "Automotive technology and human factors research: Past, present, and future," *International journal of vehicular technology*, vol. 2013, 2013.

[18] T. Gordon, H. Sardar, D. Blower, M. Ljung Aust, Z. Bareket, M. Barnes, A. Blankespoor, I. Isaksson-Hellman, J. Ivarsson, B. Juhas, *et al.*, "Advanced crash avoidance technologies (acat) program-final report of the volvo-ford-umtri project: Safety impact methodology for lane departure warning-method development and estimation of benefits," tech. rep., United States. National Highway Traffic Safety Administration, 2010.

[19] A. Carvalho, S. Lefévre, G. Schildbach, J. Kong, and F. Borrelli, "Automated driving: The role of forecasts and uncertainty—a control perspective," *European Journal of Control*, vol. 24, pp. 14–32, 2015.

[20] I. Miller, M. Campbell, D. Huttenlocher, F.-R. Kline, A. Nathan, S. Lupashin, J. Catlin, B. Schimpf, P. Moran, N. Zych, *et al.*, "Team cornell's skynet: Robust perception and planning in an urban environment," *Journal of Field Robotics*, vol. 25, no. 8, pp. 493–527, 2008.

[21] L. Claussmann, A. Carvalho, and G. Schildbach, "A path planner for autonomous driving on highways using a human mimicry approach with binary decision diagrams," in *2015 European Control Conference (ECC)*, pp. 2976–2982, IEEE, 2015.

[22] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic decision-making under uncertainty for autonomous driving using continuous pomdps," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 392–399, IEEE, 2014.

[23] E. Galceran, A. G. Cunningham, R. M. Eustice, and E. Olson, "Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction: Theory and experiment," *Autonomous Robots*, vol. 41, no. 6, pp. 1367–1382, 2017.

[24] P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, and D. Hrovat, "Predictive active steering control for autonomous vehicle systems," *IEEE Transactions on control systems technology*, vol. 15, no. 3, pp. 566–580, 2007.

[25] A. Carvalho, Y. Gao, A. Gray, H. E. Tseng, and F. Borrelli, "Predictive control of an autonomous ground vehicle using an iterative linearization approach," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 2335–2340, IEEE, 2013.

[26] B. Chen, D. Zhao, and H. Peng, "Evaluation of automated vehicles encountering pedestrians at unsignalized crossings," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1679–1685, IEEE, 2017.

[27] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz, "Simulation of pedestrian dynamics using a two-dimensional cellular automaton," *Physica A: Statistical Mechanics and its Applications*, vol. 295, no. 3-4, pp. 507–525, 2001.

[28] M. Chraibi, A. Seyfried, A. Schadschneider, and W. Mackens, "Quantitative description of pedestrian dynamics with a force-based model," in *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, vol. 3, pp. 583–586, IEEE, 2009.

[29] X. Tao, Z. Peng, S. Wong, S. Chi-Wang, and Z. Meng-Ping, "A macroscopic approach to the lane formation phenomenon in pedestrian counterflow," *Chinese Physics Letters*, vol. 28, no. 10, p. 108901, 2011.

[30] G. Antonini, M. Bierlaire, and M. Weber, "Discrete choice models of pedestrian walking behavior," *Transportation Research Part B: Methodological*, vol. 40, no. 8, pp. 667–687, 2006.

[31] S. Hoogendoorn and P. HL Bovy, "Simulation of pedestrian flows by optimal control and differential games," *Optimal Control Applications and Methods*, vol. 24, no. 3, pp. 153–172, 2003.

[32] S. Roy, A. Borzì, and A. Habbal, "Pedestrian motion modelled by fokker–planck nash games," *Royal Society open science*, vol. 4, no. 9, p. 170648, 2017.

[33] D. C. Duives, W. Daamen, and S. P. Hoogendoorn, "State-of-the-art crowd motion simulation models," *Transportation research part C: emerging technologies*, vol. 37, pp. 193–209, 2013.

[34] J. Zhang, H. Wang, and P. Li, "Cellular automata modeling of pedestrian's crossing dynamics," *Journal of Zhejiang University-SCIENCE A*, vol. 5, no. 7, pp. 835–840, 2004.

[35] Z. Sun, B. Jia, and X.-G. Li, "The study of the interference between pedestrians and vehicles based on cellular automaton model," 2012.

[36] X. Xin, N. Jia, L. Zheng, and S. Ma, "Power-law in pedestrian crossing flow under the interference of vehicles at an un-signalized midblock crosswalk," *Physica A: Statistical Mechanics and its Applications*, vol. 406, pp. 287–297, 2014.

[37] Q. Chao, Z. Deng, and X. Jin, "Vehicle–pedestrian interaction for mixed traffic simulation," *Computer Animation and Virtual Worlds*, vol. 26, no. 3-4, pp. 405–412, 2015.

[38] S. Jin, X. Qu, C. Xu, and D.-H. Wang, "Dynamic characteristics of traffic flow with consideration of pedestrians' road-crossing behavior," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 18, pp. 3881–3890, 2013.

[39] P. Chen, C. Wu, and S. Zhu, "Interaction between vehicles and pedestrians at uncontrolled mid-block crosswalks," *Safety science*, vol. 82, pp. 68–76, 2016.

[40] W.-C. Ma, D.-A. Huang, N. Lee, and K. M. Kitani, "Forecasting interactive dynamics of pedestrians with fictitious play," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 774–782, 2017.

[41] A. Lerner, Y. Chrysanthou, and D. Lischinski, "Crowds by example," in *Computer graphics forum*, vol. 26, pp. 655–664, Wiley Online Library, 2007.

[42] B. Benfold and I. Reid, "Stable multi-target tracking in real-time surveillance video," in *CVPR 2011*, pp. 3457–3464, IEEE, 2011.

[43] S. Hoogendoorn and P. Bovy, "Pedestrian travel behavior in walking areas by subjective utility optimization," in *Transportation Research Annual Meeting 2002*, 2002.

[44] S. P. Hoogendoorn and P. H. Bovy, "Pedestrian route-choice and activity scheduling theory and models," *Transportation Research Part B: Methodological*, vol. 38, no. 2, pp. 169–190, 2004.

[45] K. Cao, Y. Chen, D. Stuart, and D. Yue, "Cyber-physical modeling and control of crowd of pedestrians: a review and new framework," *IEEE/CAA Journal of Automatica Sinica*, vol. 2, no. 3, pp. 334–344, 2015.

[46] T. Ikeda, Y. Chigodo, D. Rea, F. Zanlungo, M. Shiomi, and T. Kanda, "Modeling and prediction of pedestrian behavior based on the sub-goal concept," *Robotics*, vol. 10, 2013.

[47] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, “The walking behaviour of pedestrian social groups and its impact on crowd dynamics,” *PloS one*, vol. 5, no. 4, p. e10047, 2010.

[48] N. Schneider and D. M. Gavrila, “Pedestrian path prediction with recursive bayesian filters: A comparative study,” in *German Conference on Pattern Recognition*, pp. 174–183, Springer, 2013.

[49] J. F. Kooij, N. Schneider, and D. M. Gavrila, “Analysis of pedestrian dynamics from a vehicle perspective,” in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pp. 1445–1450, IEEE, 2014.

[50] C. G. Keller and D. M. Gavrila, “Will the pedestrian cross? a study on pedestrian path prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 494–506, 2013.

[51] R. Quintero, I. Parra, D. F. Llorca, and M. Sotelo, “Pedestrian path prediction based on body language and action classification,” in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 679–684, IEEE, 2014.

[52] H. Hamaoka, T. Hagiwara, M. Tada, and K. Munehiro, “A study on the behavior of pedestrians when confirming approach of right/left-turning vehicle while crossing a crosswalk,” *Journal of the Eastern Asia Society for Transportation Studies*, vol. 10, pp. 2109–2122, 2013.

[53] M. Iryo-Asano, W. K. Alhajyaseen, and H. Nakamura, “Analysis and modeling of pedestrian crossing behavior during the pedestrian flashing green interval,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 958–969, 2014.

[54] X. Zhang, P. Chen, H. Nakamura, and M. Asano, “Modeling pedestrian walking speed at signalized crosswalks considering crosswalk length and signal timing,” in *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 9, pp. 1–15, 2013.

[55] Y. Hashimoto, G. Yanlei, L.-T. Hsu, and K. Shunsuke, “A probabilistic model for the estimation of pedestrian crossing behavior at signalized intersections,” in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 1520–1526, IEEE, 2015.

[56] A. Millard-Ball, “Pedestrians, autonomous vehicles, and cities,” *Journal of Planning Education and Research*, vol. 38, no. 1, pp. 6–12, 2018.

[57] T. Bjørnskau, “The zebra crossing game–using game theory to explain a discrepancy between road user behaviour and traffic rules,” *Safety science*, vol. 92, pp. 298–301, 2017.

[58] C. Fox, F. Camara, G. Markkula, R. Romano, R. Madigan, N. Merat, *et al.*, “When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions,” 2018.

[59] F. Camara, R. Romano, G. Markkula, R. Madigan, N. Merat, and C. Fox, "Empirical game theory of pedestrian interaction for autonomous vehicles," in *Proceedings of Measuring Behavior 2018*, Manchester Metropolitan University, 2018.

[60] P. Molchanov, S. Gupta, K. Kim, and K. Pulli, "Multi-sensor system for driver's hand-gesture recognition," in *2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*, vol. 1, pp. 1–8, IEEE, 2015.

[61] A. T. Schulz and R. Stiefelhagen, "A controlled interactive multiple model filter for combined pedestrian intention recognition and path prediction," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 173–178, IEEE, 2015.

[62] Y. Hashimoto, Y. Gu, L.-T. Hsu, and S. Kamijo, "Probability estimation for pedestrian crossing intention at signalized crosswalks," in *2015 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pp. 114–119, IEEE, 2015.

[63] N. Brouwer, H. Kloeden, and C. Stiller, "Comparison and evaluation of pedestrian motion models for vehicle safety systems," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2207–2212, Nov 2016.

[64] L. Hobert, A. Festag, I. Llatser, L. Altomare, F. Visintainer, and A. Kovacs, "Enhancements of v2x communication in support of cooperative autonomous driving," *IEEE communications magazine*, vol. 53, no. 12, pp. 64–70, 2015.

[65] A. Hussein, F. García, J. M. Armingol, and C. Olaverri-Monreal, "P2v and v2p communication for pedestrian warning on the basis of autonomous vehicles," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2034–2039, IEEE, 2016.

[66] T. Lagstrom and V. M. Lundgren, "Avip-autonomous vehicles interaction with pedestrians," *Master of Science Thesis, Chalmers University of Technology*, 2015.

[67] T. T. Nguyen, K. Holländer, M. Hoggenmueller, C. Parker, and M. Tomitsch, "Designing for projection-based communication between autonomous vehicles and pedestrians," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 284–294, ACM, 2019.

[68] C. Ningbo, W. Wei, Q. Zhaowei, Z. Liying, and B. Qiaowen, "Simulation of pedestrian crossing behaviors at unmarked roadways based on social force model," *Discrete Dynamics in Nature and Society*, vol. 2017, 2017.

[69] F. Schneemann and I. Gohl, "Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments," in *2016 IEEE intelligent vehicles symposium (IV)*, pp. 38–43, IEEE, 2016.

[70] D. Fridovich-Keil, E. Ratner, L. Peters, A. D. Dragan, and C. J. Tomlin, "Efficient iterative linear-quadratic approximations for nonlinear multi-player general-sum differential games," in *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 1475–1481, IEEE, 2020.

[71] S. L. Cleac'h, M. Schwager, and Z. Manchester, "Algames: A fast solver for constrained dynamic games," *arXiv preprint arXiv:1910.09713*, 2019.

[72] J. F. Fisac, E. Bronstein, E. Stefansson, D. Sadigh, S. S. Sastry, and A. D. Dragan, "Hierarchical game-theoretic planning for autonomous vehicles," in *2019 International conference on robotics and automation (ICRA)*, pp. 9590–9596, IEEE, 2019.

[73] A. Liniger and J. Lygeros, "A noncooperative game approach to autonomous racing," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 3, pp. 884–897, 2019.

[74] R. Spica, E. Cristofalo, Z. Wang, E. Montijano, and M. Schwager, "A real-time game theoretic planner for autonomous two-player drone racing," *IEEE Transactions on Robotics*, vol. 36, no. 5, pp. 1389–1403, 2020.

[75] Z. Wang, R. Spica, and M. Schwager, "Game theoretic motion planning for multi-robot racing," in *Distributed Autonomous Robotic Systems*, pp. 225–238, Springer, 2019.

[76] K. J. Aström and R. M. Murray, *Feedback systems: an introduction for scientists and engineers.* Princeton university press, 2010.

[77] S. P. Hoogendoorn and P. Bovy, "Generic driving behavior modeling by differential game theory," in *Traffic and Granular Flow'07*, pp. 321–331, Springer, 2009.

[78] A. Boyali, S. Mita, and V. John, "A tutorial on autonomous vehicle steering controller design, simulation and implementation," *arXiv preprint arXiv:1803.03758*, 2018.

[79] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "It's not all about size: On the role of data properties in pedestrian detection," in *ECCVW*, 2018.

[80] A. Ess, B. Leibe, K. Schindler, , and L. van Gool, "A mobile vision system for robust multi-person tracking," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08)*, IEEE Press, June 2008.

[81] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, "Learning social etiquette: Human trajectory understanding in crowded scenes," in *European Conference on Computer Vision*, 2016.

[82] M. Lisotto, P. Coscia, and L. Ballan, "Social and scene-aware trajectory prediction in crowded spaces," *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 2567–2574, 2019.

[83] C. Anderson, X. Du, R. Vasudevan, and M. Johnson-Roberson, "Stochastic sampling simulation for pedestrian trajectory prediction," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4236–4243, 2019.

[84] G. S. Sankar and K. Han, "Adaptive robust game-theoretic decision making strategy for autonomous vehicles in highway," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14484–14493, 2020.

[85] Y. Chai, J. Luo, N. Han, and J. Sun, "Robust event-triggered game-based attitude control for on-orbit assembly," *Aerospace Science and Technology*, vol. 103, p. 105894, 2020.

[86] C. Qin, Z. Shang, Z. Zhang, D. Zhang, and J. Zhang, "Robust tracking control for non-zero-sum games of continuous-time uncertain nonlinear systems," *Mathematics*, vol. 10, no. 11, p. 1904, 2022.

[87] A. Basu and M. K. Ghosh, "Zero-sum risk-sensitive stochastic differential games," *Mathematics of Operations Research*, vol. 37, no. 3, pp. 437–449, 2012.

[88] N. Bin, C. Zhang, H. Zhu, and Q. Liu, "Game approach for robust control strategy to follow the production in the singularly perturbed bilinear dynamic input-output systems," *Mathematical Problems in Engineering*, vol. 2017, 2017.

[89] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.

[90] J. C. Engwerda, "Robust open-loop nash equilibria in the noncooperative lq game revisited," *Optimal Control Applications and Methods*, vol. 38, no. 5, pp. 795–813, 2017.

[91] N. de la Cruz and M. Jimenez-Lizarraga, "Finite time robust feedback nash equilibrium for linear quadratic games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 11794–11799, 2017.

[92] M. Jiménez-Lizárraga, S. V. Rodríguez-Sánchez, N. de la Cruz, and C. E. Villarreal, "Robust dynamic programming in n players uncertain differential games," *Informatica*, vol. 31, no. 4, pp. 769–791, 2020.

[93] A. Bajcsy, S. Bansal, E. Ratner, C. J. Tomlin, and A. D. Dragan, "A robust control framework for human motion prediction," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 24–31, 2021.

[94] G. Jank and G. Kun, "Optimal control of disturbed linear-quadratic differential games," *European Journal of Control*, vol. 8, no. 2, pp. 152–162, 2002.

[95] H. Nishimura, N. Mehr, A. Gaidon, and M. Schwager, "Rat ilqr: A risk auto-tuning controller to optimally account for stochastic model mismatch," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 763–770, 2021.