THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Towards Safe and Efficient, Interactive Planning with Learning-based Optimal Control

Nominal, Stochastic, and Distributionally Robust Approaches, with Applications to Autonomous Heavy Vehicle Combinations

ERIK BÖRVE

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, 2026

Towards Safe and Efficient,
Interactive Planning with Learning-based Optimal Control

Nominal, Stochastic, and Distributionally Robust Approaches,
with Applications to Autonomous Heavy Vehicle Combinations

Erik Börve

Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000

Cover:
Examples of interactive predictions and plans in various traffic scenarios.

Printed by Chalmers Digital Printing
Gothenburg, Sweden, April 2026

Towards Safe and Efficient,
Interactive Planning with Learning-based Optimal Control
Nominal, Stochastic, and Distributionally Robust Approaches,
with Applications to Autonomous Heavy Vehicle Combinations
Erik Börve
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Planning has emerged as a fundamental component in modern autonomous systems (AS), spanning a wide range of applications from manufacturing and robot manipulation, to the focus of this thesis: autonomous vehicles. A key remaining challenge for these systems is operating in environments shared with other actors, such as, humans or other robotic systems. To find a plan that is both safe and efficient, the AS needs to predict the plan of other actors. Crucially, this includes how the AS and other actors influence each other, or in other words, how they interact.

This thesis investigates the intersection of learning- and optimization-based control approaches. In particular, we investigate optimal control over finite, receding horizons, commonly referred to as Model Predictive Control. Three Model Predictive Control problems are investigated, each with a fundamentally different treatment of the predicted plan of other actors: Nominal, considering only the most probable future outcome; Stochastic, considering a distribution of the future outcomes, and Distributionally Robust, considering a set of distributions of the future outcomes.

The results provides steps towards interactive planning with guarantees on an optimal tradeoff between safety, and efficiency. In particular this includes, theoretical, algorithmical and applied developments, that introduce a close coupling between the planner for the AS and the predicted plan of other actors. Simulation studies show that the AS can obtain human-like reasoning, exploiting interaction for necessary performance increases, while acting cautiously when faced with uncertainties to not jeopardize safety.

Keywords: Interactive Planning, Learning-based Model Predictive Control, Stochastic Optimization, Distributionally Robust Optimization.

To my Family and Friends from Mölnlycke.

List of Publications

This thesis is based on the following publications:

[A] **E. Börve**, N. Murgovski, L. Laine, "Interaction-aware trajectory prediction and planning in dense highway traffic using distributed model predictive control". Published in *2023 62nd IEEE Conference on Decision and Control (CDC).*

[B] **E. Börve**, N. Murgovski, L. Laine, "Tight Collision Avoidance for Stochastic Optimal Control: with Applications in Learning-based, Interactive Motion Planning". Submitted for publication as journal article. 2025.

[C] **E. Börve**, N. Murgovski, M. H. Chehregani L. Laine, "Interactive Trajectory Planning with Learning-based Distributionally Robust Model Predictive Control and Markov Systems". Submitted for publication as conference paper. 2025.

Other publications by the author, not included in this thesis, are:

[D] J. Dong, **E. Börve**, M. Rafayelyan, M. Unser, "Asymptotic stability in reservoir computing". *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022.

[E] E. Dimou, **E. Börve**, A. Kanellopoulo, N. Murgovski, "Safe and Efficient Optimization-Based Trajectory Planning using Conformal Prediction". Submitted for publication as journal article. 2025.

In addition, the research presented in this thesis has led to the following patent-related publications:

[F] **E. Börve**, L. Laine, S. Börjesson, N. Murgovski, "System and method for controlling an autonomous vehicle in a multi-lane driving environment". *Patent Application*, 2024.

[G] D. Pathare, L. Laine, M. Hagir Chehreghani, **E.Börve**, S. Börjesson, "Method and computer system for multi-level control of motion actuators in an autonomous vehicle". *Patent*, 2025.

[H] **E. Börve**, L. Laine, D. Pathare, S. Börjesson, A. Eriksson, “Computer System and Method for Lane-change Assistance”. *Patent Application*, 2025.

[I] **E. Börve**, L. Laine, D. Pathare, S. Börjesson, A. Eriksson, “Computer System and Method for Overtaking Assistance”. *Patent Application*, 2025.

Contents

Acknowledgments

Fist and foremost I would like to extend my sincerest gratitude to my supervisor Nikolce Murgovski. You offer thoughtful guidance, listen with care, and have given me a solid footing to help me develop in research and life. I would equally like to extend gratitude to my industrial supervisor Leo Laine, for being consistently supportive, and driving many fruitful collaborations between our research ideas and the industry practice. I sincerely thank my co-supervisor Morteza Haghir Chehreghani for many insightful thoughts and feedback during our recurring discussions. My thanks also go to Deepthi, Stefan and Markus for valuable conversations and nice collaborations at the intersections of our research interests. With extension, I would also like to thank all collaborators, colleagues, students, managers, and friends that I have met during my time at Volvo, all the way from Simulation Dragons at EMOB, to Efficent Driving at SEDS. I am really enjoying the experience of pursuing a PhD in industry, and you have been a key part of this!

I would also like to extend a huge shout-out to all the very nice people that I have met through Chalmers and E2 during my PhD studies. Whether it be, swimming, cooking, sailing, gyming, pottery, watching movies, singing commercials, eating pizza, going for coffee, going for beer(s), talking nonsense in our office and the lunchroom, or even climbing, your company has made the last years a really nice experience. I would also like to thank Yoko for keeping me energized and nourished with an amazing beef noodle soup. I would similarly like to extend a huge thanks my family and friends from Mölnlycke. Having people I can always count on for fun in good times and support in bad times really means a lot to me.

I would finally like to express my sincere thanks to Volvo Group and Chalmers University of Technology for facilitating my PhD studies. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Erik Börve, Göteborg, April 2026

Acronyms

MPC:	Model Predictive Control
LB-MPC:	Learning-based Model Predictive Control
RMPC:	Robust Model Predictive Control
SMPC:	Stochastic Model Predictive Control
DR-MPC:	Distributionally Robust Model Predictive Control
OCP:	Optimal Control Problem
DRO:	Distributionally Robust Optimization
ML:	Machine Learning
DML:	Deep Machine Learning
AS:	Autonomous System
AV:	Autonomous Vehicle
HDV:	Human Driven Vehicle
HVC:	Heavy Vehicle Combination

Part I

Overview

CHAPTER 1

Introduction

1.1 Safe and Efficient Interactive Planning

Planning provides an overarching framework for regulation of movement in mechanical systems, integrating principles from many fields, such as mechanical, electrical, control and computer science. Through the use of motion control, it represents not only the coordination of actuators, sensors, and controllers, but also the deliberate shaping of motion into a meaningful and intelligent behavior. This technology forms the foundation of modern automation and robotics, enabling complex tasks such as precision manufacturing, robotic manipulation, and, of particular interest to this work, autonomous navigation. A common denominator in these applications is the requirement on the mechanical system to move through the environment to a goal or reference, while respecting its, and the environments' physical constraints. Further, tasks in modern automation often occur within dynamic environments, in turn introducing a broad range of additional challenges. The mechanical systems, often referred to as *agents*, must continuously respond to, e.g., variable external disturbances, and time-varying conditions, requiring controllers to leverage real-time sensory feedback and adaptive algorithms to maintain the desired contextual behavior.

Classical feedback control methods for autonomous systems (AS) directly leverage the most recent measurement of itself and the environment to produce a dynamically feasible control action for the current time. However, such approaches have demonstrated limited capability in addressing complex physical constraints, such as those related to environmental geometry or dynamics. This has led to a widespread adoption of planning algorithms for AS applications. On a high level, planning algorithms can be divided into *Motion planning* and *Trajectory planning*. These divisions are closely related, and occasionally even used interchangeably in academic works. Following [1], motion planning algorithms utilize a defined goal and measurements of the environment to produce a path through the environment that connects the AS and the goal. Trajectory planning algorithms on the other hand, utilize such a path to produce another that is dynamically feasible. In this thesis we will discuss algorithms that jointly tackle the objectives of motion planning and trajectory planning. Our algorithms will utilize a defined goal and AS state with measurements of the environment to produce dynamically feasible paths from the AS, through the environment, and to the goal. To avoid confusion between the two terms, we will simply refer to our methods as *planners* that aim to solve *planning* problems.

The methods of this thesis further aims to find efficient solutions with guarantees on safety. *Efficiency* if often defined in terms of how well the AS is able to complete the task, utilizing application-specific metrics, e.g., completion time or expended energy. *Safety* on the other hand, lacks a unified definition. In some settings, safe planning is synonymous with absolute constraints on violating the imposed system limitations. This approach has a solid theoretical foundation, but may fail when the constraints are effected by large uncertainty. Such uncertainties are particularly prevalent in dynamic environments that include exogenous agents, e.g., other mechanical systems or humans. Many works argue that it is impossible to ensure constraint satisfaction with absolute certainty under these conditions. The act of moving in an uncertain environment has an inherent risk that is only eliminated by not moving in the first place [2]. This highlights the fact that safety and efficiency can represent competing objectives. A planner that prioritizes caution may provide strong safety guarantees at a loss of operational efficiency, whereas a more bold planner may improve performance while jeopardizing safety. Hence, we treat safety and efficiency by aiming for a balanced trade-off, attaining high efficiency while imposing a high probability of satisfying the safety constraints.

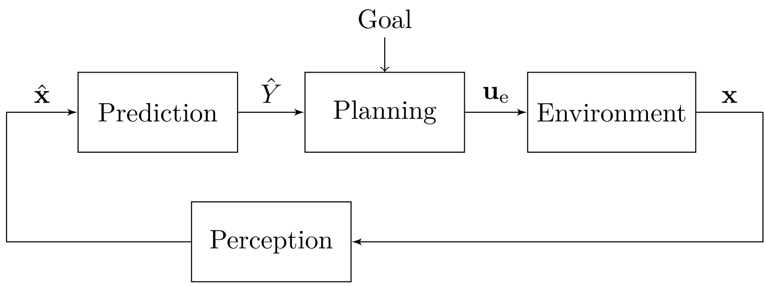


Figure 1.1: Common architecture for planning in dynamic environments.

Modern architectures for planning in dynamic environments are often divided into three subcomponents: *Perception* responsible for measuring the state of the environment, *Prediction* responsible for predicting the near future of the dynamic environment, and *Planning* responsible for producing the plan for the controllable mechanical system, often referred to as the *ego-agent*. A typical architecture based on this subdivision is displayed in Figure 1.1. This figure includes: Environment prediction $\hat{Y}$; Ego-agent control action $\mathbf{u}_\mathrm{e}$; Environment state $\mathbf{x}$, and environment state estimation $\hat{\mathbf{x}}$. The prediction and perception modules can be just as crucial as the planner. A poor prediction of the future environment state based on a poor estimate of the current environment state tends to produce infeasible plans.

The architecture of Figure 1.1 is prevalent in successful planning applications, however, it suffers from a key flaw: Decoupling prediction and planning. Consequentially, the plan of the ego-agent is assumed to not influence the dynamics of the environment. This can be particularly detrimental for applications in environments that include exogenous agents, where the plan of each agent may be contingent on the plans of others. This bi-directional influence between different agents is often referred to as an *interaction*, and is a setting that has received vast attention in recent years. A frequently appearing term in the context of interactive dynamic environments is *interaction awareness*. In past years, this term has been applied broadly to describe varying degrees to which interactions are considered by the planner. We believe this definition to have settled at describing a setting where the AS considers interactions in it's predictor, while assuming that the planner does not influence the prediction. This assumption can be validated by, e.g., considering a highly passive AS that interferes minimally with the plans of other agents. Indeed, this interpretation of interaction awareness can be handled by the architecture displayed in Figure 1.1. Importantly, this is not how interactions are treated in this thesis. Our aim is to actively consider how the ego-agent plan influences the plans of surrounding agents, requiring a coupling between the predictor and planner. Figure 1.2 displays the important adaptation of the architec-

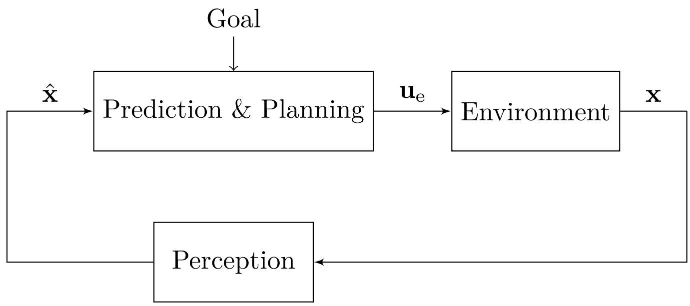


Figure 1.2: Architecture for interactive planning in dynamic environments.

ture, introducing a bi-directional coupling between prediction and planning. Hence, the planners discussed in this thesis are referred to as fully *interactive*, as they not only understand, but also can leverage the interaction to benefit their own plan. This allows our planners to utilize the safety-efficiency trade-off to the full extent, e.g., by avoiding intentionally passive behavior when it is excessively conservative. Simultaneously, this close coupling introduces large practical and theoretical challenges. The primary aim of this thesis is to address these challenges. While consideration of the Perception module may be of interest, this thesis will primarily aim to tackle the integration of Prediction and Planning assuming that a sufficiently accurate estimate of the environment state $\hat{\mathbf{x}} \approx \mathbf{x}$ is available.

1.2 Learning-based Model Predictive Control

A powerful framework for realizing safe and efficient planning objectives is *optimal control*. The solution of a predictive, optimal control problem (OCP) can produce plans that are dynamically feasible, respect environmental and physical constraints, and optimize some performance metrics. Indeed, this closely follows our outlined goals, making it a particularly compelling planning algorithm for our setting. These methods typically rely on model assumptions for the ego-agent and environment dynamics, but can provide strong guarantees on constraint satisfaction and optimality in many cases. In this thesis, we primarily discuss optimal control over a finite, receding horizon, popularly referred to as *Model Predictive Control* (MPC). As we consider potentially non-linear models, the MPC problems lack explicit solutions and need to be solved in real-time. This control architecture introduces requirements on the computational effort, in turn limiting the complexity and families of the models considered.

Learning-based Model Predictive Control (LB-MPC) further considers a setting where some models are partially or completely inferred from data. Typically, learning is used to capture external disturbances or to represent system dynamics that are not fully known a priori. In our setting, the most significant sources of uncertainty and incomplete information arise from the environment dynamics. Although the ego-agent dynamics are generally known or accurately modeled, the dynamics of exogenous agents are most often only partially characterized. Further, the decision-making process of such agents, e.g., humans or human-operated systems, is incredibly complex. Two different humans can make two drastically different decisions under the same external conditions. In recent years, state-of-the-art approaches to human motion and behavior prediction have increasingly relied on black-box machine learning (ML) models, such as large neural networks. While such prediction models typically achieve high accuracy on many widely used benchmarks, closely-coupling these models with MPC-based planning remains challenging. First, the complexity of such ML models can render real-time solution of the resulting MPC problem intractable, in particular when gradient-based optimization solvers are employed. Second, MPCs that account for modeling uncertainty are often restricted to considering certain families of distributions. This in turn places restrictions on the design of the ML predictor. Finally, inferring models from a finite number of samples implies that there will be errors in the predictions. To maintain the guarantees that are provided by the MPC these errors should be accounted for. This leads to three distinct research questions for this thesis: How can -

RQ1 - learning-based prediction methods be coupled with MPC methods to achieve interactive planning?

RQ2 - the inherent uncertainty of exogenous agents be incorporated into learning-based and stochastic MPC methods for interactive planning?

RQ3 - prediction errors arising from learning-based models inferred from a finite number of samples be accounted for in MPC for interactive planning?

1.3 Contributions

In this thesis, we propose three approaches to interactive planning, by introducing a close-coupling between the two modules. Hence, all three papers provide solutions towards our first research question. In particular, Paper A provides a method for coupling general black-box motion predictors with MPC for interactive planning through a distributed optimization approach. The method relies on an iterative scheme between predictor and planner, which terminates at a given tolerance or when convergence stops. To tackle the second research question, Papers B and C model decision uncertainty by moving over to the gray-box setting, treating humans as Markov Decision Processes with a finite number of decisions. In Paper B, we treat a stochastic LB-MPC problem that utilizes the distribution of interactive decisions to optimize the expected performance, subject to constraints on the probability of collisions. In particular, we propose a range of tight constraint reformulations that allow the ego-agent to successfully complete challenging maneuvers in crowded environments, which often coincide with emerging interactions. Finally, Paper C directly tackles the third research question, considering the same interactive planning setting as Paper B, while levering Distributionally Robust Optimization (DRO). We propose theoretical results for the learning method to allow for a learning-based Distributionally Robust MPC (DR-MPC) that utilize estimates of the size and probability of prediction errors, while actively accounting for them in the interactive plan.

The methods in this thesis are applicable for any autonomous system in an environment with exogenous agents. The demonstrated solutions are tailored for applications for autonomous vehicles (AVs), in particular, for autonomous heavy-vehicle-combinations (HVCs). In a similar fashion, the uncertain exogenous agents that we treat are primarily human driven vehicles (HDVs).

1.4 Outline

The thesis is structured as follows. Part I presents a broad background and overview of the accompanying papers. This includes: a general introduction in Chapter 1; an overview of different predictors, environments, and scenarios in Chapter 2; an overview of MPC theory in the deterministic, stochastic, and distributionally robust settings in Chapter 3; a summary of the accompanying papers in Chapter 4; and finally, a summary and outlook on future work in Chapter 5. Part 2 contains the accompanying papers.

CHAPTER 2

Interactive Prediction

The following chapter introduces the prediction problem in interactive environments and scenarios. We present the general problem setting in Section 2.1 and further introduce solutions using rule-based methods in Sections 2.2 and learning-based methods in Section 2.3.

2.1 Problem Formulation

Environment and Scenarios

In this thesis, we aim to treat environments with exogenous agents, of which the future behavior is inherently uncertain to the ego-agent. This is a particularly relevant setting when considering planning problems in environments with humans, or human-operated systems, such as HDVs. Indeed, two different humans can make two drastically different decisions under the same external conditions. In turn, this difference can depend on a vast amount of different factors, e.g., mental state, personality, and prior experiences. For example, an HDV may behave cautiously in an intersection if the human overheard an accident on the radio, but the same human may behave aggressively

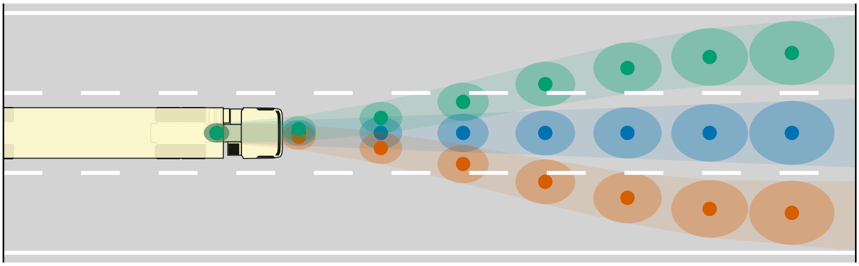
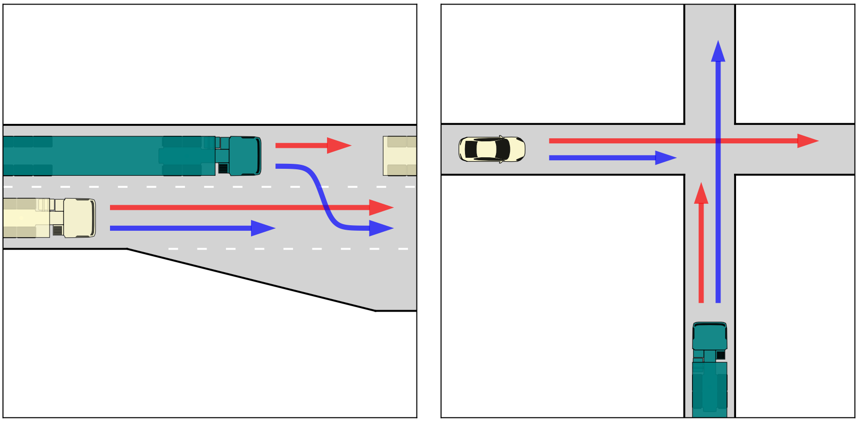


Figure 2.1: Conceptual prediction uncertainty for a highway setting.

in the same intersection if their favorite football team recently lost a match. Clearly it is ridiculous to assume that an AS could have perfect information of such factors, making the human future plan inherently uncertain to the ego-agent. Not only are high-level decisions of humans uncertain, finer details such as how much the steering wheel is turned or how heavily the gas pedal is pressed introduce similarly inherent uncertainties. Outside of the actions of the human, its dynamics may contain further uncertainties. While modern perception system often can classify the agent type and associated dynamics equations with high confidence, e.g., bipedal, bicycle, or vehicle, it may be more difficult to identify its' parameters, e.g., leg length, wheel base length, or tire rolling resistance.

The uncertainty corresponding to high-level decisions often amounts to multiple distinct trajectories, often referred to as multimodal uncertainty. Uncertainties that stem from finer details in the dynamics and decision-making typically cause a spread around each distinct mode. Figure 2.1 provides a conceptual visualization of these uncertainties in a highway environment.

In interactive environments, human actions are further affected by the actions of other humans and the autonomous system (AS). A common denominator in such environments is that multiple agents share the same resources, e.g., space on a road. Hence, it becomes particularly interesting to investigate specific scenarios where the AS and other humans are competing for said resource. Figure 2.2 displays two such scenarios. The arrows indicate two examples of compatible plans for the respective agent in corresponding color. In both scenarios, the autonomous HVC and the human driver must coordinate their use of the shared space. In Figure 2.2a the ego-vehicle aims to quickly change lane to meet an upcoming highway exit, while the human



(a) Highway driving environment. **(b)** One-way intersection environment.

Figure 2.2: Different compatible modes for interactive driving scenarios.

aims to keep its velocity in the current lane. In Figure 2.2b, both vehicles prefer to cross the intersection first, to avoid decreasing their velocity. Hence, in both scenarios, each vehicle prefers to utilize the resource prior to the other, causing a conflict which is resolved through interaction. All these factors have made trajectory prediction a notoriously difficult problem, which to this day receives great attention in both research and practical applications.

Predictor Interface

The goal of the interactive predictor module is to capture the uncertainties using relevant observations from the perception module and produce beliefs of the human plans for the near future. In a general sense, we may describe the interactive predictor as a parameterized function $\Phi_\theta \colon \mathcal{X} \mapsto \mathcal{P}$ that maps from input features $X \subseteq \mathcal{X}$ to a distribution $P(\mathbf{Y} = Y|\mathbf{X} = X) \subseteq \mathcal{P}$ describing the inherently uncertain human future $\mathbf{Y}$. The choice of parameters $\theta \subseteq \Theta$ can be inferred from expert knowledge, laws and data, varying based on Φ and the application. Regardless of the practical setting, the prediction problem aims to find the Φ_θ, which can describe the future of the environment with high accuracy. In practice, accuracy and computational effort may be a trade-off, varying with the application.

Considering a setting with one ego-agent e and $i = 1, \ldots, n_\mathrm{h}$ human agents, the input features typically consist of a collection of current, and m past observations of the agents' state $\mathbf{x} = [\mathbf{x}_\mathrm{e}^\top, \mathbf{x}_1^\top, \ldots, \mathbf{x}_{N_\mathrm{h}}^\top]^\top \in \mathbb{R}^{(n_\mathrm{h}+1)n_x}$. In the interactive setting, it becomes crucial to further consider the n future states of the ego-agent. That is, at discrete time k the input features become $X(k) = \{X_\mathrm{past}(k), X_\mathrm{e,plan}(k)\}$ where $X_\mathrm{past}(k) = [\mathbf{x}(k), \mathbf{x}(k-1), \ldots, \mathbf{x}(k-m)]$, $X_\mathrm{e,plan}(k) = [\mathbf{x}_\mathrm{e}(k+1), \ldots, \mathbf{x}_\mathrm{e}(k+n)]$. The output distribution attempts to estimate the probability of the environment state for n future time-steps, $\hat{P}(\mathbf{Y} = Y(k)|\mathbf{X} = X(k))$ where $Y(k) = [\hat{\mathbf{x}}(k+1), \ldots, \hat{\mathbf{x}}(k+n)]$ and $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^\top, \ldots, \hat{\mathbf{x}}_{n_\mathrm{h}}^\top]^\top \in \mathbb{R}^{n_\mathrm{h} n_x}$. It is similarly common to design predictors to only provide a nominal trajectory $\hat{Y}$, corresponding to the most likely future according to $\hat{P}$.

Importantly, the future ego-agent state is present in the input as it is necessary to predict the human interaction, but not in the output as it is determined by the planner. Each of the interactive predictors discussed in further detail in this chapter all consider a unique Φ_θ with varying X, $\hat{Y}$, and $\hat{P}$, but the general framework remains the same. Figure 2.3 displays a schematic view of the communication between the predictor and planner. The planner provides the future ego-plan, and the predictor provides the environment prediction, with a circular dependence.

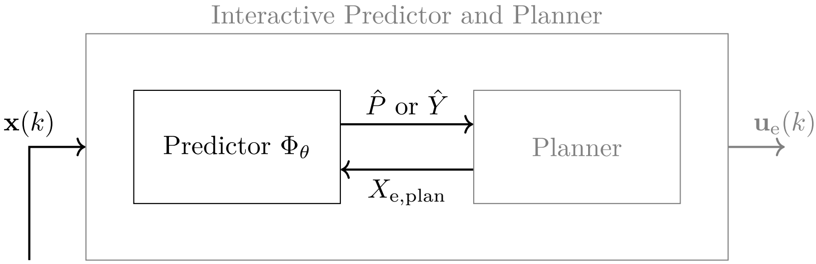


Figure 2.3: Interface for interactive predictor coupled with planner.

Indeed, the problem consists of finding a Φ_θ that attains the most accurate description of exogenous agents. However, the interactive prediction and planning setting restricts the flexibility of the method choice, as the predictor must be compatible with the planner. There are further high restrictions on computational effort, as the predictions need to be done in real time, and potentially multiple times per sample due to the circular dependence. Integration of prediction and planning through MPC is introduced in Chapter 3 and is discussed in detail in Papers A,B, and C.

2.2 Rule-based Models

A popular approach for the design of Φ_θ is to encode domain knowledge in the form of explicit rules, such as traffic regulations, right-of-way conventions, kinematic and dynamic constraints, and predefined interaction heuristics. Such rule-based methods have historically played a central role in behavior prediction, particularly in structured environments such as road networks, pedestrian walkways, and controlled robotic settings. By constraining predictions to physically feasible and legally compliant behaviors, rule-based models offer a high degree of transparency and interpretability, which is especially valuable in safety-critical applications like autonomous driving and advanced driver-assistance systems. Moreover, their relatively low computational requirements and predictable failure modes can make them more suitable for real-time deployment and for use in combination with MPC-based planners.

For interactive trajectory prediction applications, rule-based models are often integrated with dynamic models and applied step-wise over the prediction horizon. The rule-based model $\phi_\theta \colon \mathbf{x} \mapsto \mathbf{u}$ takes the state of the environment $\hat{\mathbf{x}}(k)$ and predicts control actions $\mathbf{u}(k)$, which then propagate through the dynamics $f(\hat{\mathbf{x}}(k), \mathbf{u}(k))$ to produce $\hat{\mathbf{x}}(k+1)$. The procedure is then repeated at step $k+1$ combining $\hat{\mathbf{x}}(k+1)$ with the ego-agent plan $\mathbf{x}_\mathrm{e}(k+1)$, continuing sequentially until step n. All together, the combination of ϕ_θ and f forms the predictor Φ_θ with input features $X(k) = \{\mathbf{x}(k), X_{\mathrm{e,plan}}(k)\}$. In principle, Φ_θ can be non-linear, and not continuously differentiable with respect to X.

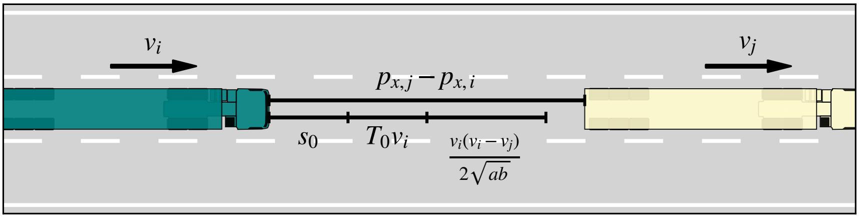
Traffic Models

Rule-based traffic models are of particular interest in this thesis. These models characterize the motion and interactions of road users through rules grounded in traffic regulations, human driving heuristics, and physical constraints. Through parameterized control laws, these models typically specify how the agents should behave to complete different driving objectives such as, maintaining a safe following distance, yielding the right of way, and adjusting speed to road geometry. One of the most widely used rule-based traffic models is the Intelligent Driver Model (IDM)[3], designed to describe longitudinal vehicle motion in traffic flow. The IDM expresses a vehicles acceleration as a continuous function by combining two intuitive driving objectives: maintaining a desired speed in free-flow conditions and preserving a safe gap when following another vehicle. For a vehicle with index i and its leading vehicle with index j we may express its acceleration as follows,

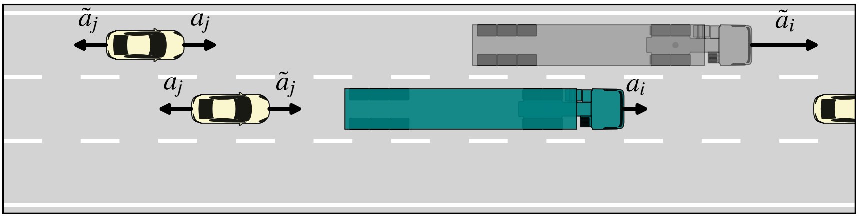
$$\dot{v}_i = a\left(1 - \left(\frac{v_i}{v_{\text{ref}}}\right)^{\delta} - \left(\frac{s^\star(v_i, v_j)}{p_{x,j} - p_{x,i}}\right)^2\right)$$
$$s^\star(v_i, v_j) = s_0 + T_0 v_i + \frac{v_i(v_i - v_j)}{2\sqrt{ab}}.$$

Here, a represents the maximum comfortable acceleration and determines how quickly a vehicle accelerates when unconstrained. In free-flow conditions, the vehicle speed v_i is regulated toward the desired reference speed v_{ref}, with the exponent δ controlling how smoothly acceleration decreases as this speed is approached. Interactions with a leading vehicle are governed by the ratio between the desired dynamic gap $s^\star$ and the actual gap $p_{x,j} - p_{x,i}$. The desired gap $s^\star$ encodes the spacing a driver aims to maintain and consists of three components: a minimum standstill distance s_0, a speed-dependent term determined by the desired time headway T_0, and a relative-velocity term that accounts for approaching a slower leader. This last term depends both on the acceleration parameter a and the comfortable deceleration b, reflecting the driver's anticipation of required braking effort. Together, parameters s_0, T_0, a, and b characterize individual driving styles, ranging from cautious to aggressive behavior. A conceptual visualization of the states and parameters is displayed in Figure 2.4a. Although the IDM is typically applied to longitudinal trajectory prediction in car-following scenarios, numerous extensions have been proposed to capture more complex traffic dynamics. Some particularly interesting examples include, highway merging [4], intersections [5], and probabilistic interpretations [6].

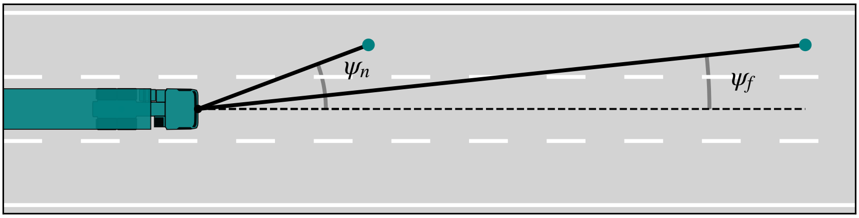
The MOBIL (Minimizing Overall Braking Induced by Lane changes) model is a widely used rule-based framework for modeling lane-changing behavior in traffic [7]. It complements longitudinal car-following models like the IDM by providing a mechanism to decide when a vehicle should change lanes based on safety and incentive criteria. MOBIL evaluates a potential lane change by balancing the expected acceleration gain for the lane-changing vehicle against



(a) Vehicle-following scenario using IDM.



(b) Highway lane-change scenario evaluation using MOBIL.



(c) 2-point visual steering model for highway lane change.

Figure 2.4: Conceptual visualizations of rule-based traffic models.

the negative impact it would have on surrounding vehicles, particularly those in the target lane. A key feature of the model is the inclusion of a safety criterion, ensuring that the lane change does not force other vehicles to decelerate beyond a safe threshold, and an incentive criterion, which determines whether the maneuver provides a net benefit considering both the driver's own advantage and a weighted factor for the surrounding traffic. For a vehicle index i and two vehicles following vehicles, one in the current lane j and one in the target lane k, the incentive criterion is formulated as,

$$(\tilde{a}_i - a_i) + p\,(\tilde{a}_j - a_j + \tilde{a}_k - a_k) \geq \Delta a_{\text{th}}.$$

Here, a and $\tilde{a}$ denote the current and expected acceleration of the respective vehicles, most often calculated with the IDM. The politeness factor p weighs how much the lane changing driver considers its impact on the surrounding vehicles, and Δa_{th} defines a minimum threshold of net acceleration gain required to justify the lane change. The safety criterion is typically expressed as a hard constraint on the deceleration of other vehicles, i.e., $\tilde{a}_j \geq b_{\max}$ and $\tilde{a}_k \geq b_{\max}$, for some $b_{\max} < 0$. This combination allows MOBIL to reproduce lane-change decisions while maintaining overall traffic stability. A conceptual visualization of a lane-change scenario is displayed in Figure 2.4b. Lastly, models such as the IDM and MOBIL are often complimented with a steering model. One such example is the two-point visual steering model [8]. It assumes that a driver tracks two points on the future path, typically a near and a far point along the lane, and continuously adjusts steering to minimize deviations from the intended trajectory. The near point is often chosen a few meters ahead of the vehicle to capture immediate lane alignment, while the far point is selected tens of meters ahead to anticipate long term effects. In the case of [8], the design is based on a PI-controller that aims to minimize the angle between the vehicle heading and the respective points as,

$$\dot{\delta} = k_f \dot{\psi}_f + k_n \dot{\psi}_n + k_I \psi_n.$$

The steering rate, $\dot{\delta}$, depends on the angular velocities of both points $\dot{\psi}_f$ and $\dot{\psi}_n$, together with the angular position of the far point ψ_n. The far point guides predictive, anticipatory steering, scaled by the gain k_f, while the near point provides fast, reactive corrections through k_n and k_I. Together, these terms capture how drivers combine immediate and future visual cues to produce smooth, controlled steering. A conceptual visualization of a lane-change scenario with corresponding angles, is displayed in Figure 2.4c.

Together, these models constitute the parameterized function $\phi_\theta(\mathbf{x}(k))$ that produces the control actions for each vehicle and time step $\mathbf{u}_i(k)$. The dynamics are often chosen based on the vehicle type, assumed available from the perception module. A common choice is the kinematic bicycle model, see e.g., [9]. For a passenger vehicle with state $\mathbf{x}_i = [p_x, p_y, v, \psi]^\top$ and control

$\mathbf{u}_i = [a, \delta]^\top$, the continuous time dynamics become,

$$\dot{\mathbf{x}}_i = \left[v\cos(\psi_1 + \beta),\ v\sin(\psi_1 + \beta),\ a,\ v\frac{\sin\beta}{\ell/2} \right]^\top$$

where $\beta = \arctan(\tan(\delta)/2)$. Further, p_x, p_y describes the position of the center of gravity, v describes its velocity, ψ is the angle to the reference frame, and ℓ is the length of the vehicle. The state dynamics are then discretized to form f, where the discretization method is chosen to obtain a suitable trade-off between computational effort and accuracy [10].

In Papers A, B, and C, these three rule-based traffic models are combined with kinematic models and used to varying extent for both simulation and prediction.

2.3 Learning-based Models

Rule-based trajectory prediction methods provide transparent and structured modeling of agent behavior, but also exhibit limitations when applied to more complex environments. Their reliance on handcrafted rules can make it difficult to account for rare events and subtle aspects of human decision-making, such as social cues or context-dependent behavior. As scenario complexity increases, maintaining and extending rule sets becomes progressively more challenging, often requiring simplifying assumptions to preserve tractability. These factors have driven a growing interest in learning-based approaches that instead aim to infer behavioral patterns from data.

In a general sense, the supervised learning process assumes a general structure of Φ_θ and learns the parameters θ by matching the prediction $\hat{Y}$ to labeled measurements Y. In recent years, the most popular approach is to utilize deep machine learning (DML). In this setting Φ_θ, often some constellation of a deep neural network, is treated as a black-box, producing $\hat{Y}, \hat{P}$ from X, without any discernible connection to e.g., physics or geometry. These methods often obtain state-of-the-art performance on benchmark datasets and have started to show impressive generalization capabilities in practical applications [11]. Some notable architectures include: Trajectron++ [12], and Scene Transformer [13]. This thesis does not focus on further developing ML methods for prediction, but rather on applying and adapting existing methods for interactive prediction and planning. Paper A in particular presents a method for interactive prediction and planning that is compatible with black-box predictors.

Although DML demonstrates remarkable performance across complex high-dimensional tasks, its reliance on highly parameterized NNs often limits transparency and human interpretability. In safety-critical applications, such as planning with exogenous agents, these attributes are not only desirable but essential. Without the possibility of further analysis, unexplainable errors could carry severe human and societal consequences [14]. In contrast, Interpretable Machine Learning (IML) methods are designed to make their decision-making processes transparent and understandable to humans. Unlike involved black-box models, these approaches allow practitioners to directly examine how input features contribute to predictions. Some classical examples include linear and logistic regression, where model coefficients quantify the influence of each variable; decision trees, which represent decisions as a sequence of human-readable rules; and adaptive rule-based models, which express predictions through structured, often simplified relationships between features and outcomes [15]. In this sense, IML models can serve as a middle ground between rule-based models and DML models, offering greater flexibility and expressive power than rule-based approaches while maintaining more transparency and interpretability than DML models. In Papers B and C, we shift our focus to IML models by representing the decision-making of exogenous agents as Markov processes which are combined with rule-based models.

Markov Processes

Among approaches to trajectory prediction in machine learning, structured probabilistic models stand out for their ability to represent assumptions about dynamics explicitly. One such framework is the Markov process, which provides a mathematical description of stochastic systems that evolve over time. A Markov process consists of a set of states and a transition mechanism that specifies the probability of moving from one state to another. Its defining characteristic is the Markov property, according to which the future evolution of the system depends only on the present state and not on the full sequence of past states. The behavior of the process is fully characterized by its transition probabilities, which determine how the distribution propagates across future time steps. The probability of a future state is commonly decomposed as follows,

$$\mathbf{p}(\mathbf{x}_{k+1}) = \mathbf{p}_{\mathbf{x}}(\mathbf{x}_{k+1}|\mathbf{u}_k, \mathbf{x}_k)\mathbf{p}_{\mathbf{u}}(\mathbf{u}_k|\mathbf{x}_k)\mathbf{p}(\mathbf{x}_k).$$

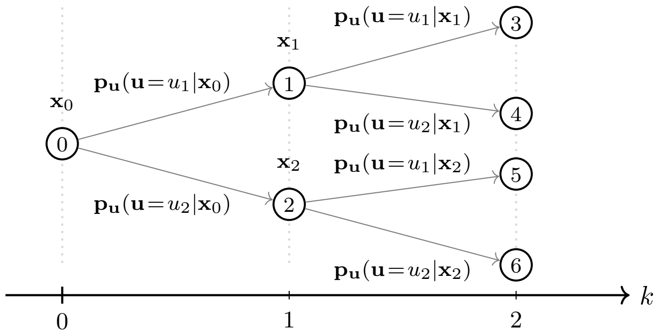


Figure 2.5: Multimodal Markov process evolving over a prediction horizon.

Here, $\mathbf{p}_\mathbf{x}$ describes the distribution of the dynamics and $\mathbf{p}_\mathbf{u}$ describes the distribution of the actions, often referred to as the policy. In a learning-based setting, it is common to consider a parameterized function describing the policy, i.e., $\mathbf{p}(\mathbf{u}|\mathbf{x}) = \phi_\theta(\mathbf{x})$, where the parameters θ are to be inferred from data. While applicable both in the DML and IML setting, this is particularly well suited to IML because of the structured representation of the dynamics through the Markov property. A particularly appealing property of Markov processes for interactive prediction is the ability to explicitly encode multimodality in the transition mechanism. Prior work has identified that Markov Jump systems are well suited towards this end [16]. Figure 2.5 demonstrates such a Markov Process with two possible modalities where $\mathbf{p}_\mathbf{u}$ uniquely determines the transition probability (dynamics $\mathbf{p}_\mathbf{x}$ assumed certain). Sequentially expanding the possible modalities over the prediction horizon builds a tree structure of possible futures.

A similar approach is utilized in Papers B and C, where the action probability is modeled by combining an interpretable $\phi_\theta(\mathbf{x})$ with rule-based traffic models.

Guarantees for Learning Schemes

The vast amount of learning schemes measures success in terms of how close $\hat{Y}$ is to Y on average. However, this notion of performance, typically captured through empirical risk on training or test data, provides only a partial view of reliability. The average error does not quantify how much performance may degrade on unseen data, nor does it characterize the probability of rare but consequential failures. In many planning applications, for example those that utilize MPC, this limitation becomes critical. Prediction errors directly influence constraint satisfaction, recursive feasibility, and stability. Not only is a good average behavior desirable, but deviations and tail events should be quantified and accounted for to ensure sufficiently safe operations.

In recent years, much research has been done towards combining MPC with learning systems that provide heuristic uncertainty quantification, such as Gaussian Process Regression [17], Monte Carlo dropout [18], or deep ensemble estimates [19]. These methods have been proven to provide useful approximates of predictive uncertainty, showing impressive performance also in practical applications. However, since these methods do not consider finite-sample estimation errors, they cannot establish rigorous performance guarantees.

In more recent years, Conformal prediction (CP) [20], has provided a complementary approach. Under the assumption of exchangeability, conformal methods construct prediction sets that achieve distribution-free coverage in finite samples. Crucially, they can be layered on top of arbitrary predictors, converting residual information into uncertainty sets. For MPC, such sets can be propagated through constraints to obtain probabilistic feasibility guarantees with explicit confidence levels [21] [22]. While CP-based MPC approaches can provide valid guarantees in many settings, they have failed to do so in interactive environments, as the process of interaction introduces distribution shifts that violate the exchangability assumption. Further, while the guarantees are valid, the methods instead struggle to be useful. Efficiency may be heavily impaired, or feasibility may be lost altogether, as the provided safety guarantees can become excessively conservative.

In Paper C we instead connect MPC with Statistical learning theory in the Markov process setting, to obtain an MPC that can account for finite-sample estimation errors without heavily impairing performance or jeopardizing feasibility, all while avoiding assumptions that hinder interactive planning.

CHAPTER 3

Interactive Planning

To complete the interactive prediction and planning architecture in Figure 2.3, we will now introduce three different planning approaches, each as Model Predictive Control (MPC) problems. Section 3.1 introduces the fundamental setting, considering a nominal prediction of exogenous agents $\hat{X}$. Section 3.2 instead describes the stochastic Model Predictive Control (SMPC) setting, considering an estimated distribution of the future state of the exogenous agents $\hat{P}$. Finally, Section 3.3 continues with the Distributionally Robust

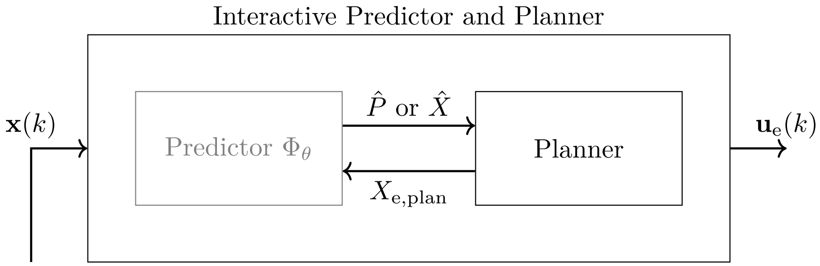


Figure 3.1: Interface for interactive predictor coupled with planner.

MPC (DR-MPC) setting, considering a set of distributions around the estimate. Adopting the same architecture as Figure 2.3, the MPC problem formulations will represent the complete interactive predictor and planner module, as show in Figure 3.1. The iterative communication between planning and prediction appears as the problem is being solved. In Paper A, this is done by leveraging distributed optimization. In Papers B and C, this occurs within the numerical optimization solver.

3.1 Nominal Model Predictive Control

To produce plans that are dynamically feasible, respect environmental constraints, and optimize performance metrics, our aim is to use optimal control over a finite receding horizon, commonly known as Model Predictive Control (MPC) [23]. As the ego-agent dynamics are described by nonlinear, continuous-state models, an explicit optimal solution does not generally exist, and it becomes suitable to employ gradient-based nonlinear optimization solvers [24]. In particular, we focus on Newton-based methods that exploit second-order information to achieve faster local convergence. Common tools tailored to formulate and solve MPC problems include, e.g., [25], [26], and [27]. Such methods solve for the locally optimal control inputs $U_\mathrm{e} = [\mathbf{u}_\mathrm{e}(0), \dots, \mathbf{u}_\mathrm{e}(n-1)]$ and corresponding states $X_\mathrm{e} = [\mathbf{x}_\mathrm{e}(0), \dots, \mathbf{x}_\mathrm{e}(n)]$, over the discrete-time horizon $k = 0, \dots, n$.

MPC problems typically include a few standard ingredients. The state and control are constrained to sets $\mathcal{X}_\mathrm{e} \subseteq \mathbb{R}^{n_{\mathrm{e},x}}$ and $\mathcal{U}_\mathrm{e} \subseteq \mathbb{R}^{n_{\mathrm{e},u}}$, respectively. To account for the unmodeled future over the infinite horizon, the terminal state is typically constrained separately to a terminal set $\mathcal{X}_{\mathrm{e},f} \subseteq \mathbb{R}^{n_{\mathrm{e},x}}$. It is similarly common to express these sets with vector-valued functions $g(\mathbf{x}_\mathrm{e}, \mathbf{u}_\mathrm{e}) \leq 0$ and $g_f(\mathbf{x}_\mathrm{e}) \leq 0$, respectively. The discrete-time dynamics of the ego-agent are modeled as a function $f \colon \mathbb{R}^{n_{\mathrm{e},x}} \times \mathbb{R}^{n_{\mathrm{e},u}} \mapsto \mathbb{R}^{n_{\mathrm{e},x}}$, describing how it evolves from the current state measurement $\mathbf{x}_\mathrm{e}(t)$, over the horizon, and to the terminal state $\mathbf{x}_\mathrm{e}(n)$. The final standard ingredients are the stage-cost $\ell \colon \mathbb{R}^{n_{\mathrm{e},x}} \times \mathbb{R}^{n_{\mathrm{e},u}} \mapsto \mathbb{R}$, and the terminal cost $\ell_f \colon \mathbb{R}^{n_{\mathrm{e},x}} \mapsto \mathbb{R}$. Both are intended to measure the performance of the controller, with ℓ_f additionally intended to measure performance in the unmodeled future over the infinite horizon [28].

In this thesis, we consider settings where the properties of the ego-agent, i.e., f, g, g_f, ℓ, ℓ_f, can be inferred with reasonable accuracy. Instead, as discussed in previous chapters, the main difficulties arise from operating in dynamic environments with interactive exogenous agents, such as HDVs. Indeed, the corresponding functions for the exogenous agents become much more difficult to obtain with the same accuracy. As outlined in Chapter 2, we consider $i = 1, \dots n_{\mathrm{h}}$ human agents whose future state prediction can be provided by a parameterized function Φ_θ. Within the MPC framework, the human state prediction is primarily incorporated to express the collision avoidance constraints. Collision avoidance typically amounts to ensuring that there is a strictly positive distance between the ego-agent and the human agent. How the distance is expressed may vary drastically depending on the geometry of both agents and the environment. In the most general sense, the distance between two agents represents the smallest norm difference between a pair of points $(\mathbf{p}_{\mathrm{e}}, \mathbf{p}_i)$ on the respective agents [29]. Considering that the ego-agent occupies a set $\mathcal{E} \subset \mathbb{R}^m$ and that a human occupies another set $\mathcal{H}_i \subset \mathbb{R}^m$ in m spatial dimensions, we may express a distance function as,

$$\mathrm{dist}(\mathbf{x}_{\mathrm{e}}, \mathbf{x}_i) = \min_{\mathbf{p}_{\mathrm{e}}, \mathbf{p}_i} \{||\mathbf{p}_{\mathrm{e}} - \mathbf{p}_i|| : \mathbf{p}_{\mathrm{e}} \in \mathcal{E}, \mathbf{p}_i \in \mathcal{H}_i\}.$$

A conceptual visualization of this function is provided in Figure 3.2. Here the function describes the distance between the nonconvex space occupied by an HVC and the space occupied by two HDVs, respectively. In practice, it is common to introduce an additional safety margin d_{safe} e.g., as to: (i) Increase robustness against unexpected errors, and (ii) Avoid causing potential discomfort or distress to human agents and technical issues for nearby robotic agents with implementations of similar safety margins [30].

The final and most crucial fact for interactive planning is that the prediction of human agents depends on the states of the ego-agent X_{e}. Hence, the prediction depends on the MPC optimization variables and the predictor must be directly incorporated into the MPC problem. In the deterministic setting, the MPC only considers the nominal prediction of human agents. Indeed, this lacks consideration of the many inherent uncertainties that persist. In practice, such safety concerns are typically mitigated indirectly, e.g., by heuristically increasing safety margins. As in Chapter 2 the predictor Φ_θ takes past observations X_{past} and future ego-state X_{e}, which here become optimization variables.

To neatly group the indices in the coming problem formulations, we consider $\mathcal{I}_{a:b} = \{a, a+1, \dots, b\}$ as the set of incrementally increasing integers from a,

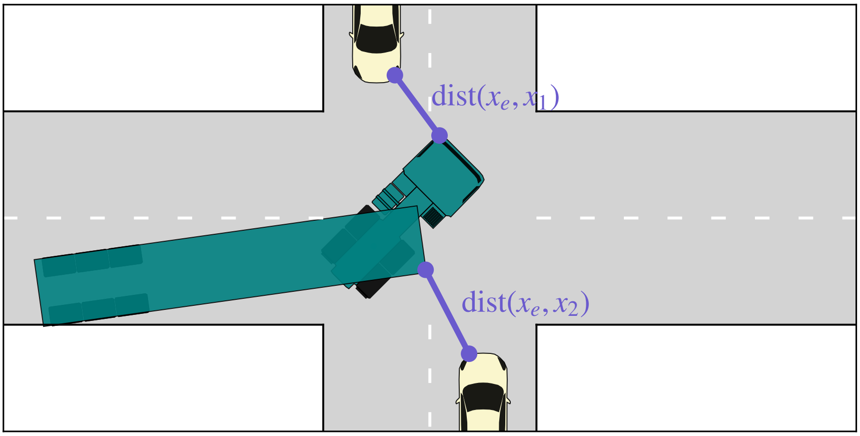


Figure 3.2: Distance between the ego-vehicle (teal), and two HDVs (yellow).

until b. We may now construct the deterministic MPC problem, considering a nominal prediction of the interactive human agents as follows,

$$\min_{X_e, U_e} \ell_f\big(\mathbf{x}_e(n)\big) + \sum_{k=0}^{n-1} \ell\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big) \tag{3.1a}$$

$$\text{s.t., } \mathbf{x}_e(k+1) = f\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big),\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.1b}$$

$$g\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big) \leq 0,\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.1c}$$

$$\text{dist}\big(\mathbf{x}_e(k), \hat{\mathbf{x}}_i(k)\big) \geq d_{\text{safe}},\ \forall k \in \mathcal{I}_{0:n},\ \forall i \in \mathcal{I}_{1:n_h} \tag{3.1d}$$

$$\hat{X} = \Phi_{\theta,i}\big(X_{\text{past}}, X_e\big) \tag{3.1e}$$

$$g_f\big(\mathbf{x}_e(n)\big) \leq 0,\ \mathbf{x}_e(0) = \mathbf{x}_e(t),\ \mathbf{x}_i(0) = \mathbf{x}_i(t),\ \forall i \in \mathcal{I}_{1:n_h}. \tag{3.1f}$$

Here, $\mathcal{I}_{0:n-1}$ emphasizes that the dynamics (3.1b), together with the state and control constraints (3.1c) are repeated for $k = 0, \dots, n-1$. Similarly, $\mathcal{I}_{0:n}$ and $\mathcal{I}_{1:n_\mathrm{h}}$ emphasize that the collision avoidance constraints (3.1d) are repeated for all $k = 0, \dots, n$ and for all humans $i = 1, \dots, n_\mathrm{h}$. Finally, (3.1f) describes the boundary conditions, where t describes the real-world continuous time.

To obtain feasible solutions of (3.1) with tractable computational effort using gradient-based optimization solvers, a number of requirements must be satisfied. Of foremost importance, all functions must be continuously differentiable [31]. This requirement excludes direct integration of many formulations such as: switching logic, absolute value functions, max/min operators, or piecewise-defined functions [32]. Such elements can create non-differentiable points where gradients are undefined or discontinuous, typically causing solver software to interrupt immediately. A significant direction in MPC research is devoted to reformulating or approximating these functions so that they become continuously differentiable, allowing the usage of general-purpose solvers [33], [34]. However, ensuring differentiability alone does not guarantee computational tractability. The effort required to compute first- and second-order sensitivities increases with the algebraic and compositional complexity of the model equations. In practical MPC implementations, these derivatives must be evaluated repeatedly within each optimization step and at every sampling instant. Consequently, moderately more complex gradient expressions that cause a moderate increase in the per-iteration cost, can significantly enlarge the cost of the complete nonlinear program [35].

In our setting, challenges with both these requirements emerge through the predictor Φ_θ. Indeed, large neural network-based black-box models typically achieve the best empirical prediction performance on popular datasets [36], [37]. However, many commonly used neural network architectures employ activation functions such as ReLU, sign, or max operators, which are not continuously differentiable. Although these functions are well suited for training via stochastic gradient methods, they violate the smoothness assumptions imposed by the gradient-based nonlinear programming solvers. Further, embedding a large neural network within the prediction model introduces a deep composition of nonlinear mappings whose derivatives must be evaluated in the process of solving the MPC problem. Backpropagation sensitivities of the optimization variables through many large network layers at every solver iteration can therefore add an intractable amount of computational effort [38], [35].

In Paper A, we address this challenge by employing distributed MPC techniques to allow the black-box predictor and the MPC-based planner to be separate modules. The scheme iteratively communicates the nominal HDV predictions and ego-vehicle plans between the two modules, allowing for interactive planning with black-box predictors.

Similar challenges also emerge through the collision avoidance constraints. As the general formulation of the minimum distance function introduces a min-operation, direct integration does not yield a continuously differentiable optimization problem [39]. In Paper A we propose an outer-approximation of the HDVs that constrain the ego-vehicle point mass, and consequently remove the min-operator. In Paper B, we extend existing methods for reformulating the min-operation and constrain the exact distance between the nonconvex HVC and HDVs, as it is displayed in Figure 3.2.

3.2 Stochastic Model Predictive Control

In the previous section, we established the standard construction of MPC problems together with the additional elements that are necessary for interactive planning. However, the key flaw with the nominal design is the fact that the predictor is assumed to represent a perfect, deterministic model of humans. While the resulting MPC may have some success in practice given sufficient heuristic safety margins, it fails to treat the key issue: the inherent uncertainty in the state of the human agents. To achieve this, we need to leverage tools from Stochastic Model Predictive Control (SMPC) [40], [41].

In the stochastic setting, the state trajectory of human agents receives a fundamentally different treatment. As a replacement for the point-wise interactive prediction $\hat{X}$ of all human states over the horizon, it instead considers the trajectories of all humans as a sequence of random variables $\mathbf{X}$. Crucially for interactive planning, the ego-agent needs to consider how the distribution of the human state is dependent on the ego-agents own plan. This implies that interactive planners must consider the conditional distribution $P(\mathbf{X} \,|\, \mathbf{X}_{\text{past}}, \mathbf{X}_{\text{e}})$, and consequently that the predictor must provide an estimate of said distribution as $\hat{P} = \Phi_\theta(X_{\text{past}}, X_{\text{e}})$ [42]. Figure 3.3 displays a nominal state prediction $\hat{X}_i$ and an estimated state distribution $\hat{P}_i$ for an HDV at an intersection. The nominal prediction represents what the predictor believes to be the most probable future, which clearly fails to capture

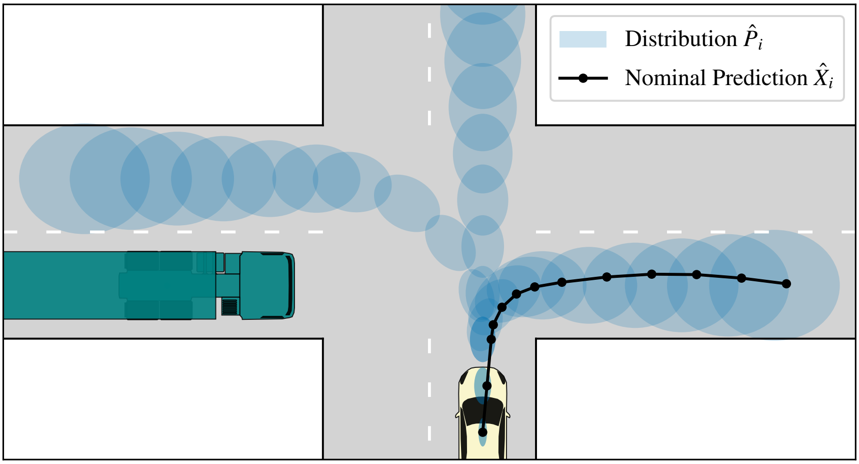


Figure 3.3: Nominal and stochastic trajectory predictions for an HDV at an intersection.

the spread and multimodal nature of the humans behavior. A planner that only relies on the nominal prediction in this setting could be at high risk of violating collision avoidance constraints.

With a distribution of the inherently uncertain and interactive environment, we may now design an SMPC to integrate it explicitly. As in the nominal setting, we consider that the properties of the ego-agent, i.e., f, g, g_f, ℓ, ℓ_f, can be inferred with reasonable accuracy. Hence, the optimizer treats them as deterministic mappings. The primary difference compared to the nominal MPC setting appears in the collision avoidance constraints. Not only should the ego-agent plan to avoid the most probable trajectory of the human, but it should avoid all regions where the human is likely to be. This is exactly what chance constraints aim to achieve [43]. Rather than constraining the distance between the ego- and human agents to a sufficient value, it constrains the probability that the distance between the two agents is sufficient. Describing the state of each human agent over the horizon as a sequence of random variables $\mathbf{X} = [\mathbf{X}_{1,1}, \ldots, \mathbf{X}_{n_\mathrm{h},1}, \mathbf{X}_{1,2}, \ldots, \mathbf{X}_{i,k}, \ldots, \mathbf{X}_{n_\mathrm{h},n}]$ with joint distribution P, we may express the chance constrained collision avoidance as,

$$\mathbb{P}_{\mathbf{X}\sim P}\left[\mathrm{dist}(\mathbf{x}_\mathrm{e}(k), \mathbf{x}_i(k)) \geq d_\mathrm{safe}, \forall i \in \mathcal{I}_{1:n_\mathrm{h}}\ \forall k \in \mathcal{I}_{0:n}\right] \geq 1 - \varepsilon$$

where $\varepsilon \in [0,1]$ is a threshold probability.

In addition to modifying the collision avoidance constraints, the presence of uncertainty may also influence the formulation of the cost function. As the human state is modeled by random variables, any cost term with a direct dependence on the random variables should be treated accordingly. The most common approach is to express the MPC objective as minimizing the expected value of the cost function, subject to the distribution of the human states. Assuming a similar structure to (3.1) but adapting the cost, collision avoidance, and predictor for the stochastic setting, finally yields the complete SMPC problem as,

$$\min_{X_\mathrm{e}, U_\mathrm{e}} \mathbb{E}_{\mathbf{X}\sim\hat{P}} \left[\ell_f\big(\mathbf{x}_\mathrm{e}(n)\big) + \sum_{k=0}^{n-1} \ell\big(\mathbf{x}_\mathrm{e}(k), \mathbf{u}_\mathrm{e}(k)\big) \right] \tag{3.2a}$$

$$\text{s.t., } \mathbf{x}_\mathrm{e}(k+1) = f\big(\mathbf{x}_\mathrm{e}(k), \mathbf{u}_\mathrm{e}(k)\big),\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.2b}$$

$$g\big(\mathbf{x}_\mathrm{e}(k), \mathbf{u}_\mathrm{e}(k)\big) \le 0,\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.2c}$$

$$\mathbb{P}_{\mathbf{X}\sim\hat{P}} \left[\mathrm{dist}(\mathbf{x}_\mathrm{e}(k), \mathbf{x}_i(k)) \ge d_\mathrm{safe}, \forall i \in \mathcal{I}_{1:n_\mathrm{h}}\ \forall k \in \mathcal{I}_{0:n}\right] \ge 1 - \varepsilon \tag{3.2d}$$

$$\hat{P} = \Phi_\theta\big(X_\mathrm{past}, X_\mathrm{e}\big) \tag{3.2e}$$

$$g_f\big(\mathbf{x}_\mathrm{e}(n)\big) \le 0,\ \mathbf{x}_\mathrm{e}(0) = \mathbf{x}_\mathrm{e}(t),\ \mathbf{x}_i(0) = \mathbf{x}_i(t),\ \forall i \in \mathcal{I}_{1:n_\mathrm{h}} \tag{3.2f}$$

where again the dynamics (3.2b), together with the state and control constraints (3.2c) are repeated for $\mathcal{I}_{0:n-1} = \{0, \dots, n-1\}$. The collision avoidance constraints (3.2d) are enforced jointly over the full trajectory $k = 0, \dots, n$ of all humans $i = 1, \dots, n_\mathrm{h}$ jointly. As before, (3.2f) describes the boundary conditions, where t describes the real-world continuous time.

As in the case of (3.1), all functions must be continuously differentiable and of tractable complexity to allow the use of gradient-based optimization solvers. Furthermore, in the stochastic setting, it becomes necessary to derive closed-form expressions for the probability measure and expectation with respect to the distribution of the human states. Exact reformulations of these expressions are only available for specific distributions, e.g., Gaussian or categorical. Further, as $\mathbf{X}$ represents a sequence of random variables that evolve over time, each $\mathbf{X}_{i,k}$ is dependent on all prior $k-1, \dots, 0$ through the dynamics, altering their marginal distributions into more involved expressions. In practice, this often implies that practitioners have to leverage approximations, or relax guarantees with additional assumptions, e.g., removing temporal dependence

[44]. For chance-constraint approximations, obtaining valid guarantees often implies making conservative outer-approximations [45], [46]. In the best case, this restricts the planners ability to accurately balance safety and efficiency; in the worst case, this heavily restricts the planners feasible space and may even produce infeasible problems. For relaxation of guarantees, this may, e.g., imply that the probabilistic constraints only hold marginally, and not over multiple time steps, limiting the statements that can be made regarding the success of the planning scheme [47].

In paper B we address both these points in a setting where the HDVs are treated as a Markov system with categorically distributed decisions. We provide tight approximations for the marginal chance constraints and further provide applications to different probability measures, such as the joint setting of (3.2d).

3.3 Distributionally Robust Model Predictive Control

As described in the prior section, learning-based SMPC accounts for the inherent uncertainty by assuming that the probability distribution has been estimated correctly from the available data. In the vast majority of real-world applications, this assumption does not hold, as the distribution governing the behavior of exogenous agents is estimated from finite data. The learned distribution will, almost surely, deviate from the true underlying distribution, which in turn degrades the safety guarantees [48], [49]. Without explicitly quantifying the discrepancy between the estimated and the true distribution, there is no guarantee that the learned model provides useful information for the interactive planner. A poorly trained model may, e.g., incorrectly assign a low probability to a safety-critical event, causing the estimated probability to fall below the threshold value ε. As a result, the SMPC may exclude the event from consideration, potentially leading to less safe solutions.

Learning-based Distributionally Robust Model Predictive (DR-MPC) leverages techniques from Distributionally Robust Optimization (DRO) to addresses this limitation by explicitly accounting for uncertainty in the estimation of the probability distribution itself [16]. The central idea in learning-based DRO is to replace the assumption of a single known distribution with an ambiguity set, defined as a collection of probability distributions that are deemed plausible given the available data. Instead of optimizing the expected cost with respect to a single learned distribution, the controller optimizes against the worst-case distribution within this ambiguity set. Similarly, the chance constraint considers the worst-case distribution, maximizing the probability of constraint violations. This approach provides robustness against distributional misspecification due to the finite sample, while maintaining a probabilistic description of uncertainty [50].

To ease notation in this section, we consider a single human agent index h with corresponding trajectory distribution $P_{\mathrm{h}}(\mathbf{X}_{\mathrm{h}} \,|\, \mathbf{X}_{\mathrm{past}}, \mathbf{X}_{\mathrm{e}})$, conditioned on the ego-agent trajectory and past observation. As in prior sections the predictor needs to provide an estimate of said distribution as $\hat{P}_{\mathrm{h}} = \Phi_\theta(X_{\mathrm{past}}, X_{\mathrm{e}})$, which crucially incorporates interaction through the dependence on X_{e}. The DR-MPC additionally considers an ambiguity set $\mathcal{P}_\theta$ around the nominal distribution, typically defined using a statistical distance or divergence measure such as Wasserstein distance [51], or ϕ-divergences [52]. The ambiguity set can be interpreted as a confidence region in the space of probability distributions which is designed such that it should include the true distribution with high probability. Learning-based DR-MPC thus naturally arises when considering ambiguity sets that are constructed with respect to the uncertainty when fitting the parameters θ based on a finite sample. Given a distribution estimate $\hat{P}_{\mathrm{h}}$ based on m samples and a divergence d, we may consider a divergence-based ambiguity set as follows,

$$\mathcal{P}_{\alpha,m} = \left\{ P : d(P, \hat{P}_{\mathrm{h}}) \leq r_{\alpha,m} \right\}.$$

Here, $r_{\alpha,m} \in \mathbb{R}$ is a radius and $\alpha \in [0,1]$ is a probability level. A key property of the ambiguity set is that the radius is chosen such that,

$$\mathbb{P}\left[P_{\mathrm{h}}(\mathbf{X}_{\mathrm{h}} \,|\, \mathbf{X}_{\mathrm{past}}, \mathbf{X}_{\mathrm{e}}) \in \mathcal{P}_{\alpha,m}\right] \geq 1 - \alpha.$$

Hence, with probability at least $1-\alpha$, the true distribution is contained in the ambiguity set $\mathcal{P}_{\alpha,m}$. The radius of the ambiguity set typically depends both on the sample size m and the confidence level $1-\alpha$. In particular, the radius decreases as the number of samples increases, reflecting improved statistical accuracy. Conversely, it increases as the confidence requirement becomes more stringent, i.e., as $1-\alpha$ increases.

We may now adapt the SMPC problem (3.2) to the Distributionally Robust setting by accounting for the ambiguity in the neighborhood of the distribution estimate. Indeed, if chance constraints are formulated with respect to

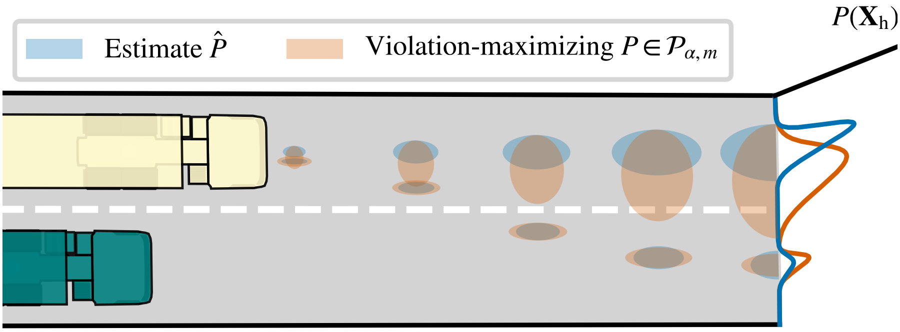


Figure 3.4: Stochastic trajectory prediction, and possible worst-case distribution in its ambiguity set that maximizes the probability of constraint violations.

the worst-case distribution in the ambiguity set, i.e., the distribution that maximizes the probability of constraint violations, then with confidence level $1-\alpha$, the true distribution will not lead to a larger probability of violations. Hence, we may formulate the ambiguous chance constraints as,

$$\max_{P \in \mathcal{P}_{\alpha,m}} \mathbb{P}_{\mathbf{X}_\mathrm{h} \sim P} \left[\mathrm{dist}\big(\mathbf{x}_\mathrm{e}(k), \mathbf{x}_\mathrm{h}(k)\big) < d_\mathrm{safe}, \forall k \in \mathcal{I}_{0:n}\right] \leq \varepsilon,$$

i.e., the probability of constraint violation for any distribution in $\mathcal{P}_{\alpha,m}$ must be smaller than or equal to ϵ. A conceptual visualization of an estimated trajectory distribution and the corresponding distribution in $\mathcal{P}_{\alpha,m}$ that maximizes constraint violations is displayed in Figure 3.4. Here, P_h describes two modes, one to keep the current lane and one to change in front of the ego-vehicle. To maximize the probability of violations, the density in the lane-keeping mode shifts towards the lane margins, and the density in the lane-changing mode is increased. This corresponds to the worst-case ground-truth distribution that we consider plausible, with probability $1-\alpha$.

The objective of (3.2) receives a similar treatment to the chance constraints. As before, we may consider the expected cost as an objective. Different from before, we may take the expectation with respect to the worst-case distribution inside of the ambiguity set. This is achieved in a similar fashion to the ambiguous chance constraints by introducing an additional max-operation over distributions in the ambiguity set. As a result, the observed average system cost will not exceed the cost anticipated by the optimizer, with a probability of $1-\alpha$.

Replacing the expected cost, and chance constraints by their ambiguous counterparts, and otherwise replicating (3.1), (3.2) yields the learning-based DR-MPC problem as follows,

$$\min_{X_e, U_e} \max_{P \in \mathcal{P}_{\alpha,m}} \mathbb{E}_{\mathbf{X}_h \sim P} \left[\ell_f\big(\mathbf{x}_e(n)\big) + \sum_{k=0}^{n-1} \ell\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big) \right] \tag{3.3a}$$

$$\text{s.t.,}\ \mathbf{x}_e(k+1) = f\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big),\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.3b}$$

$$g\big(\mathbf{x}_e(k), \mathbf{u}_e(k)\big) \leq 0,\ \forall k \in \mathcal{I}_{0:n-1} \tag{3.3c}$$

$$\max_{P \in \mathcal{P}_{\alpha,m}} \mathbb{P}_{\mathbf{X}_h \sim P} \left[\text{dist}\big(\mathbf{x}_e(k), \mathbf{x}_h(k)\big) < d_{\text{safe}},\ \forall k \in \mathcal{I}_{0:n} \right] \leq \varepsilon \tag{3.3d}$$

$$\hat{P}_h = \Phi_\theta\big(X_{\text{past}}, X_e\big) \tag{3.3e}$$

$$\mathcal{P}_{\alpha,m} = \left\{ P : d(P, \hat{P}_h) \leq r_{\alpha,m} \right\} \tag{3.3f}$$

$$g_f\big(\mathbf{x}_e(n)\big) \leq 0,\ \mathbf{x}_e(0) = \mathbf{x}_e(t),\ \mathbf{x}_h(0) = \mathbf{x}_h(t) \tag{3.3g}$$

where again $\mathcal{I}_{0:n-1} = \{0, \ldots, n-1\}$ for (3.3b) and (3.3c). As in (3.2) the chance constraints (3.3d) are expressed for the full human state trajectory, and (3.3g) describes the boundary conditions, with t representing the real-world continuous time.

As in the case of (3.1), all functions must be continuously differentiable and of tractable complexity to allow the use of gradient-based optimization solvers. As in the case of (3.2), we require analytical expressions for the expected cost and the probability of constraint violations, which, in turn, might require conservative approximations. The DR-MPC problem presents two further challenges: (i) We need to express the ambiguity set, i.e., find a suitable divergence d, and a radius $r_{\alpha,m}$, such that the set contains the true distribution with $1-\alpha$. (ii) We need to find a tractable, continuously differentiable reformulation of the additional max-operations with respect to distributions in the ambiguity sets. Indeed, if these two conditions are met, a solution of (3.3) would provide an interactive plan, accounting for the inherent uncertainty in human agents, and the estimation errors in the learning-based predictor, leading to valid safety guarantees.

The requirements on the human state distribution are challenging to fulfill already in the stochastic setting of (3.2). Conditions (i,ii) can be challenging to fulfill on their own for simple distribution, but become further complex when they need to hold simultaneously, and even more so in an interactive planning application. On its own, condition (i) is commonly analyzed by leveraging statistical inequalities together with assuming uniform convergence of the learning-scheme [53]. Informally, this implies that the learner uniformly approaches the true distribution as the number of samples tends towards infinity. In the MPC-setting, condition (ii) may be fulfilled by leveraging techniques from duality theory. Given that the max-operations in the cost and constraints are conic programs for each k, one may equivalently express them as continuously differentiable min-operation [43]. One may further consider the distribution of a sequence of random variables by considering a nested reformulation [54].

In the learning-based DR-MPC setting, conditions (i,ii) have been achieved in some prior work. One such setting is where HDVs are treated as Markov systems with categorically distributed decisions [55]. By modeling the Markov system with a state-independent transition kernel, learned online through a sample average approximation, they are able to find an ambiguity set and a tractable reformulation of the DR-MPC problem. However, as the transition kernel needs to be state-independent, the HDV model is not expressive enough to describe interaction between the HDV and the ego-vehicle. In Paper C, we extend prior work to allow for transition kernels that depend on the state of the HDV and ego-vehicle. This includes deriving a valid ambiguity set such that (i) holds and applying the nest reformulation to solve (ii). The result is a learning-based DR-MPC capable of generating interactive plans, accounting for inherent uncertainty in human behavior, and handling estimation errors arising from the finite sample size of the underlying distribution.

CHAPTER 4

Summary of included papers

This chapter provides a summary of the included papers.

4.1 Paper A

E. Börve, N. Murgovski, L. Laine
Interaction-aware trajectory prediction and planning in dense highway traffic using distributed model predictive control
Published in IEEE Conference on Decision and Control (CDC),
vol. 62, pp. 6124–6129, Dec. 2023.

This paper presents a novel framework that couples trajectory prediction and planning in multi-agent environments. Instead of directly integrating the predictor into the MPC problem formulation as in (3.1), the predictor is integrated through a distributed model predictive control scheme. Predictions of exogenous agents and plans for the ego-agent are communicated between the two modules iteratively, producing a result where predictions and plans agree. A demonstration of the framework is presented in simulation, employing a trajectory planner using non-linear model predictive control. Performance and convergence of the framework is further analyzed subject to different prediction errors. The results indicate that the obtained locally optimal solutions are improved, compared with decoupled prediction and planning.

Contributions: EB contributed with the idea, implementation, and main writing of the manuscript. NM and LL contributed with supervision and review of the manuscript.

4.2 Paper B

E. Börve, N. Murgovski, L. Laine
Tight Collision Avoidance for Stochastic Optimal Control: with Applications in Learning-based, Interactive Motion Planning
Submitted for publication as journal article. 2025 .

Trajectory planning in dense, interactive traffic scenarios presents significant challenges for autonomous vehicles, primarily due to the uncertainty of human driver behavior and the non-convex nature of collision avoidance constraints. This paper introduces a stochastic optimal control framework to address these issues simultaneously, without excessively conservative approximations. We opt to model human driver decisions as a Markov Decision Process and propose a method for handling collision avoidance between non-convex vehicle shapes by imposing a positive distance constraint between compact sets. In this framework, we investigate three alternative chance constraint formulations. To ensure computational tractability, we introduce tight, continuously differentiable reformulations of both the non-convex distance constraints and the chance constraints. The efficacy of our approach is demonstrated through simulation studies of two challenging interactive scenarios: an unregulated intersection crossing and a highway lane change in dense traffic.

Contributions: EB contributed with the idea, proofs, implementation, and main writing of the manuscript. NM contributed with guidance and ideas for the mathematical proofs. NM and LL contributed with supervision and review of the manuscript.

4.3 Paper C

E. Börve, N. Murgovski, M. H. Chehregani L. Laine
Interactive Trajectory Planning with Learning-based Distributionally Robust Model Predictive Control and Markov Systems
Submitted for publication as conference paper. 2025.

This work investigates interactive trajectory planning subject to uncertainty in the decisions of surrounding agents. To control the ego-agent we first learned the decision distribution and solved a Stochastic Model Predictive Control problem. To account for errors in the learned distribution we showed that it is possible to utilize Probably Approximately Correct learning in combination with Distributionally Robust (DR) optimization, to obtain a solution that accounts for the errors induced by the learning model. The results indicate that our PAC learning-based DR-MPC framework provides a method to interpolate between a robust MPC, and an omnipotent SMPC, based on the available number of samples.

Contributions: EB contributed with the idea, proofs, implementation, and main writing of the manuscript. NM and MHC contributed with guidance and ideas for the mathematical proofs. NM, LL and MHC contributed with supervision and review of the manuscript.

CHAPTER 5

Concluding Remarks

This thesis has investigated optimization-based approaches to interactive planning, with different learning-based modeling approaches for exogenous agents. We conclude the thesis by connecting the work to the research questions and discuss future work that we believe to be particularly promising.

5.1 Answering the Research Questions

The contributions towards the research questions are summarized in Table 5.1. This also includes the different models that may be considered for the exogenous agents, spanning from general black-box models, to Markov Processes with categorically distributed decisions, commonly referred to as Markov Jump Systems.

RQ1: *How can learning-based prediction methods be coupled with MPC methods to achieve interactive planning?*

While the planner and predictor have flexibility in their individual design, the key enabling factor for interactive planning is the close coupling between the two modules. To handle interactions, each agent needs to model how its

Table 5.1: Overview of research contributions by paper.

	RQ1	RQ2	RQ3	Exogenous Agent Model
Paper A	✓			Black-box
Paper B	✓	✓		Markov Jump Systems
Paper C	✓	✓	✓	Markov Jump Systems

own plan affects others and how others affect its own. Given that the agent wishes to optimize its own performance, the ideal interactive plan is fundamentally given by the solution of (3.1), (3.2), or (3.3). Whether each respective formulation is suitable depends on how the other agents are modeled, under what assumptions, and with what methods. To ensure compatibility with explicit MPC solvers, it is essential to preserve continuous differentiability while maintaining computational tractability. If one wishes to leverage black-box predictors, Paper A proposes a method for coupling the two modules using distributed MPC. In a general sense, it displays that distributed optimization schemes may be a promising alternative to increase flexibility in the choice of the prediction model. On the other hand, rigorously establishing guarantees on e.g., asymptotic convergence, constraint satisfaction, uncertainty quantification, and, recursive feasibility remains as open questions. Instead, the Markov system model of Papers B and C was a more promising approach to this end.

RQ2: *How can the inherent uncertainty of exogenous agents be incorporated into learning-based and stochastic MPC methods for interactive planning?* Papers B and C treat the decision making process of the exogenous agents as a Markov system with state-dependent and categorically distributed decisions. Compatibility with the optimal control framework was instead obtained by considering simpler, and continuously differentiable learning-models. Regardless, the papers show that it is possible to model interactive behavior between two agents by considering a probability distribution of different higher-level decisions. The obtained solutions show that the ego-agent can adapt its own state to alter the probability of different decisions for the exogenous agents, increasing the probability of favorable outcomes. Similarly, the chance-constraints allow the ego-agent to quantify safety, which in turn can be used to obtain a desirable balance between safety and efficiency. These findings highlight the critical role of accurate collision probability modeling, as conservative or imprecise approaches can significantly hinder performance. Due to the complexity of human decision making, learning the distribution directly from data appears practically promising. However, this inevitably introduces estimation errors into the model, which should be explicitly accounted for. If one aims to establish rigorous guarantees it becomes further necessary to treat the uncertainties that appear in the distribution estimate, due to the finite sample.

RQ3: *How can prediction errors arising from learning-based models inferred from a finite number of samples be accounted for in MPC for interactive planning?*

Papers A and B demonstrate the advantages of incorporating learning-based models to infer the behavior of exogenous agents from data. However, this practical utility is accompanied by estimation errors that are inherently introduced by the learning methodology. The ideal learning-based MPC should account for the induced prediction errors a priori. Paper C presents a methodology to systematically account for these estimation errors by leveraging distributionally robust optimization. The construction of ambiguity sets through statistical learning theory establishes a connection to the classical DRO setting, thereby facilitating the use of established solution methods. Crucially, Paper C extends the framework to state-dependent conditional distributions, allowing for interactive planning. As in Papers A and B, this implies that the ego-agent is able to optimize its own state to alter the probability of different decision for the exogenous agents. Unlike in Papers A and B, the probability that is considered is the one inside the ambiguity set that maximizes the cost in the objective and the probability of collisions in the constraints. As the number of samples increases, the ambiguity set shrinks, and the solutions approach the result we expect when leveraging the ground-truth distribution.

5.2 Future Work

Building on the contributions of Papers A–C, several directions for future research emerge, spanning theoretical developments, algorithmic enhancements, and practical applications.

For the framework introduced in Paper A, an important avenue lies in establishing stronger convergence guarantees for the proposed iterative schemes, particularly under more realistic operating conditions. In addition, integrating these methods with real-time iteration solvers and more sophisticated distributed optimization approaches such as, alternating direction method of multipliers, could significantly enhance real-time applicability. Such approaches could speed up each iteration and decrease the total number of necessary steps for convergence, reducing the overall solution time. Further, directly applying the framework to state-of-the-art trajectory predictors could improve its flexibility and relevance across diverse practical driving scenarios.

In Paper B, improving computational efficiency remains a central challenge. Future work could explore more intelligent branching strategies and scenario reduction techniques to reduce combinatorial complexity and enable practical deployment. It is moreover interesting to investigate extensions of the methodology to more realistic environments, including richer decision distributions and dynamics. In this context, the development and integration of learning models trained on naturalistic driving data represent a promising direction to enhance real-world performance.

For Paper C there is interesting theoretical and practical work. On the theoretical side, it is of particular interest to formally extend the framework to more general classes of probably-approximately-correct-learnable models, thereby broadening its theoretical scope and applicability. One could also establish MPC guarantees, such as stability and recursive feasibility. For practical applications, it is further interesting to reduce the conservativeness of the proposed methods. This includes refining the construction of ambiguity sets and risk measures to better balance safety and efficiency.

Finally, across all contributions, a critical step toward real-world impact lies in bridging the gap between methodological advances and application-driven performance metrics. Future work should focus on validating the proposed approaches in practical settings closely aligned with key business objectives, such as minimizing the total cost of operation including, e.g., energy consumption, driver compensation, and delivery times.

References

[1] S. M. LaValle, *Planning algorithms*. Cambridge university press, 2006.

[2] A. Hevelke and J. Nida-Rümelin, "Responsibility for crashes of autonomous vehicles: An ethical analysis," *Science and engineering ethics*, vol. 21, no. 3, pp. 619–630, 2015.

[3] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical review E*, vol. 62, no. 2, p. 1805, 2000.

[4] B. Brito, A. Agarwal, and J. Alonso-Mora, "Learning interaction-aware guidance for trajectory optimization in dense traffic scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18 808–18 821, 2022.

[5] A. Ghosh and T. Parisini, "Traffic control in a mixed autonomy scenario at urban intersections: An optimal control approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 17 325–17 341, 2022.

[6] E. Ward, N. Evestedt, D. Axehill, and J. Folkesson, "Probabilistic model for interaction aware planning in merge scenarios," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 2, pp. 133–146, 2017.

[7] A. Kesting, M. Treiber, and D. Helbing, "General lane-changing model mobil for car-following models," *Transportation Research Record*, vol. 1999, no. 1, pp. 86–94, 2007.

[8] D. D. Salvucci and R. Gray, "A two-point visual control model of steering," *Perception*, vol. 33, no. 10, pp. 1233–1248, 2004.

[9] C. Altafini, "Some properties of the general n-trailer," *International Journal of Control*, vol. 74, no. 4, pp. 409–424, 2001.

[10] J. C. Butcher, *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[11] Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, and H. Chen, "A survey on trajectory-prediction methods for autonomous driving," *IEEE transactions on intelligent vehicles*, vol. 7, no. 3, pp. 652–674, 2022.

[12] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone, "Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data," in *European Conference on Computer Vision*, Springer, 2020, pp. 683–700.

[13] J. Ngiam et al., "Scene transformer: A unified architecture for predicting multiple agent trajectories," *arXiv preprint arXiv:2106.08417*, 2021.

[14] *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Official Journal of the European Union, OJ L, 12.7.2024, 2024. [Online]. Available: `https://op.europa.eu/en/publication-detail/-/publication/d79f3e5d-41bc-11f0-b9f2-01aa75ed71a1/language-en`.

[15] C. M. Bishop, *Pattern recognition and machine learning*. Springer, vol. 4.

[16] M. Schuurmans and P. Patrinos, "Learning-based distributionally robust model predictive control of markovian switching systems with guaranteed stability and recursive feasibility," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 4287–4292.

[17] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using gaussian process regression," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2736–2743, 2019.

[18] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*, PMLR, 2016, pp. 1050–1059.

[19] C.-J. Hoel, K. Wolff, and L. Laine, "Ensemble quantile networks: Uncertainty-aware reinforcement learning with applications in autonomous driving," *IEEE Transactions on intelligent transportation systems*, vol. 24, no. 6, pp. 6030–6041, 2023.

[20] G. Shafer and V. Vovk, "A tutorial on conformal prediction.," *Journal of machine learning research*, vol. 9, no. 3, 2008.

[21] L. Lindemann, M. Cleaveland, G. Shim, and G. J. Pappas, "Safe planning in dynamic environments using conformal prediction," *IEEE Robotics and Automation Letters*, 2023.

[22] C. Stamouli, L. Lindemann, and G. Pappas, "Recursively feasible shrinking-horizon mpc in dynamic environments with conformal prediction guarantees," in *6th Annual Learning for Dynamics & Control Conference*, PMLR, 2024, pp. 1330–1342.

[23] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design.* Nob Hill Publishing Madison, WI, 2017, vol. 2.

[24] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[25] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, pp. 1–36, 2019.

[26] R. Verschueren et al., "Acados—a modular open-source framework for fast embedded optimal control," *Mathematical Programming Computation*, vol. 14, no. 1, pp. 147–183, 2022.

[27] "Forces nlp: An efficient implementation of interior-point methods for multistage nonlinear nonconvex programs," *International Journal of Control*, vol. 93, no. 1, pp. 13–29, 2020.

[28] B. Kouvaritakis and M. Cannon, *Non-linear Predictive Control: theory and practice.* Iet, 2001.

[29] X. Zhang, A. Liniger, and F. Borrelli, "Optimization-based collision avoidance," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 972–983, 2020.

[30] S. H. Nair, E. H. Tseng, and F. Borrelli, "Collision avoidance for dynamic obstacles with uncertain predictions using model predictive control," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, IEEE, 2022, pp. 5267–5272.

[31] M. Zanon, A. Boccia, V. G. S. Palma, S. Parenti, and I. Xausa, "Direct optimal control and model predictive control," in *Optimal control: novel directions and applications*, Springer, 2017, pp. 263–382.

[32] S. P. Boyd and L. Vandenberghe, "Convex optimization," in Cambridge university press, 2004, ch. 6, pp. 291–345.

[33] N. Murgovski, L. Johannesson, X. Hu, B. Egardt, and J. Sjöberg, "Convex relaxations in the optimal control of electrified vehicles," in *2015 American control conference (ACC)*, IEEE, 2015, pp. 2292–2298.

[34] N. Murgovski and J. Sjöberg, "Predictive cruise control with autonomous overtaking," in *2015 54th IEEE Conference on Decision and Control (CDC)*, IEEE, 2015, pp. 644–649.

[35] S. Gros, M. Zanon, R. Quirynen, A. Bemporad, and M. Diehl, "From linear to nonlinear mpc: Bridging the gap via the real-time iteration," *International Journal of Control*, vol. 93, no. 1, pp. 62–80, 2020.

[36] P. Sun et al., "Scalability in perception for autonomous driving: Waymo open dataset," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2446–2454.

[37] B. Wilson et al., "Argoverse 2: Next generation datasets for self-driving perception and forecasting," *arXiv preprint arXiv:2301.00493*, 2023.

[38] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, `http://www.deeplearningbook.org`.

[39] J. Fan, N. Murgovski, and J. Liang, "Efficient collision avoidance for autonomous vehicles in polygonal domains," *arXiv preprint arXiv:2308.09103*, 2023.

[40] A. Prékopa, *Stochastic programming*. Springer Science & Business Media, 2013, vol. 324.

[41] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Systems Magazine*, vol. 36, no. 6, pp. 30–44, 2016.

[42] Y. Chen, U. Rosolia, W. Ubellacker, N. Csomay-Shanklin, and A. D. Ames, "Interactive multi-modal motion planning with branch model predictive control," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5365–5372, 2022.

[43] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[44] L. Blackmore, M. Ono, and B. C. Williams, "Chance-constrained optimal path planning with obstacles," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1080–1094, 2011.

[45] R. Wang, M. Schuurmans, and P. Patrinos, "Interaction-aware model predictive control for autonomous driving," in *2023 European Control Conference (ECC)*, IEEE, 2023, pp. 1–6.

[46] A. Nemirovski, "On safe tractable approximations of chance constraints," *European Journal of Operational Research*, vol. 219, no. 3, pp. 707–718, 2012.

[47] M. Ono, M. Pavone, Y. Kuwata, and J. Balaram, "Chance-constrained dynamic programming with application to risk-aware robotic space exploration," *Autonomous robots*, vol. 39, no. 4, pp. 555–571, 2015.

[48] P. L. Bartlett, S. Boucheron, and G. Lugosi, "Model selection and error estimation," *Machine Learning*, vol. 48, no. 1, pp. 85–113, 2002.

[49] L. Devroye, L. Györfi, and G. Lugosi, *A probabilistic theory of pattern recognition*. Springer Science & Business Media, 2013, vol. 31.

[50] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," *arXiv preprint arXiv:1908.05659*, 2019.

[51] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1, pp. 115–166, 2018.

[52] G. Bayraksan and D. K. Love, "Data-driven stochastic programming using phi-divergences," in *The operations research revolution*, Informs, 2015, pp. 1–19.

[53] G. Pflug and D. Wozabal, "Ambiguity in portfolio selection," *Quantitative Finance*, vol. 7, no. 4, pp. 435–442, 2007.

[54] P. Sopasakis, M. Schuurmans, and P. Patrinos, "Risk-averse risk-constrained optimal control," in *2019 18th European Control Conference (ECC)*, IEEE, 2019, pp. 375–380.

[55] M. Schuurmans and P. Patrinos, "A general framework for learning-based distributionally robust mpc of markov jump systems," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 2950–2965, 2023.