



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## A Two-Stage Optimization-based Motion Planner for Safe Urban Driving

**Citation for published version:**

Eiras, F, Hawasly, M, Albrecht, SV & Ramamoorthy, S 2021, 'A Two-Stage Optimization-based Motion Planner for Safe Urban Driving', *IEEE Transactions on Robotics*, vol. Early Access. <https://doi.org/10.1109/TRO.2021.3088009>

**Digital Object Identifier (DOI):**

[10.1109/TRO.2021.3088009](https://doi.org/10.1109/TRO.2021.3088009)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

IEEE Transactions on Robotics

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# A Two-Stage Optimization-based Motion Planner for Safe Urban Driving

Francisco Eiras<sup>1,2</sup>, Majd Hawasly<sup>1</sup>, Stefano V. Albrecht<sup>1,3</sup>, Subramanian Ramamoorthy<sup>1,3</sup>

**Abstract**—Recent road trials have shown that guaranteeing the safety of driving decisions is essential for the wider adoption of autonomous vehicle technology. One promising direction is to pose safety requirements as planning constraints in nonlinear, non-convex optimization problems of motion synthesis. However, many implementations of this approach are limited by uncertain convergence and local optimality of the solutions achieved, affecting overall robustness. To improve upon these issues, we propose a novel two-stage optimization framework: in the first stage, we find a solution to a Mixed-Integer Linear Programming (MILP) formulation of the motion synthesis problem, the output of which initializes a second Nonlinear Programming (NLP) stage. The MILP stage enforces hard constraints of safety and road rule compliance generating a solution in the right subspace, while the NLP stage refines the solution within the safety bounds for feasibility and smoothness. We demonstrate the effectiveness of our framework via simulated experiments of complex urban driving scenarios, outperforming a state-of-the-art baseline in metrics of convergence, comfort and progress.

**Index Terms**—Motion planning, optimization, model predictive control (MPC), autonomous urban driving

## I. INTRODUCTION

WITH the rapid advancement of autonomous driving, it is becoming increasingly clear that guaranteeing safety across diverse driving scenarios is essential to the wider adoption of the technology [1], [2]. Performing safe, real-time planning in systems that must robustly achieve tight integration with scene perception and behavior prediction has continued to be an open challenge for the community [3], [4]. This challenge is enhanced by the complex environments and decision-making that arises in an urban, residential driving setting [5]. Fig. 1 shows an example of such environments.

An enticing approach to this problem, motivated by the ability to collect driving data from millions of miles driven by sensorized vehicles, is to exploit machine learning methods, such as in a deep imitation learning paradigm [6], [7]. While these methods have been shown to be successful in contained environments, it has been difficult to provide safety guarantees when policies are learned in this fashion, particularly with noisy training data and fault-susceptible real-time perception. Initial works aimed at providing robustness guarantees of neural networks, while encouraging, remain limited in scope in light of the scale of the models required in practical systems [8], [9].

The multi-dimensional requirements associated with planning in autonomous vehicles are inherently hierarchical in

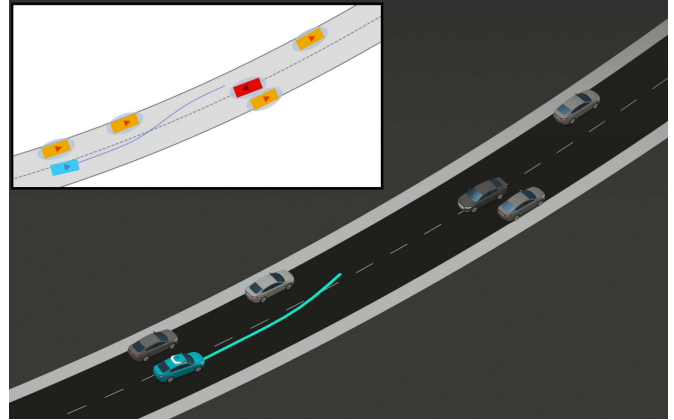


Fig. 1: *Urban driving*: simulation view of our planner overtaking static vehicles while handling an oncoming vehicle, with the planner’s view in the inset. A video of this and other planning situations are available in the supplementary material.

nature. The core concern of *avoiding collisions* with other road users and obstacles, in the interest of passenger safety, is a hard requirement. Other secondary concerns – such as continued progress towards a destination, comfort or power management – imply softer requirements. Some planning methods, including many reinforcement learning formulations and some forms of unconstrained trajectory optimization, expect the hierarchy to be resolved implicitly through the design of the optimization objective [10], [11]. However, such approaches tend to be brittle in enforcing the aforementioned safety-first hierarchy [12].

A promising approach to such prioritization is seen in the formal methods literature, where logical specifications capture safety guarantees explicitly. Examples include optimizing the robustness signal of Signal Temporal Logic formulas in a receding horizon framework [13], [14], and synthesis of policies that maximize the probability of satisfying objectives given as Linear Temporal Logic predicates [15]. Despite their high quality output (and recent runtime improvements [16]) many of these tools lead to performance bottlenecks that are prohibitive in real-time settings [17], [18].

A more scalable technique to achieve prioritization is to adopt a constrained optimization approach, in which inviolable rules of the road [19] are encoded as hard (in)equality constraints, while secondary objectives are satisfied in a soft manner as dictated by weights in a cost function [20]–[23]. In one approach, Schwarting *et al.* [24] formulated planning as a nonlinear constrained optimization problem in a Model Predictive Control (MPC) scheme. While this formulation efficiently

<sup>1</sup>FiveAI Ltd, United Kingdom, {first.last}@five.ai

<sup>2</sup>Dept. of Engineering Science, University of Oxford, United Kingdom

<sup>3</sup>School of Informatics, University of Edinburgh, United Kingdom

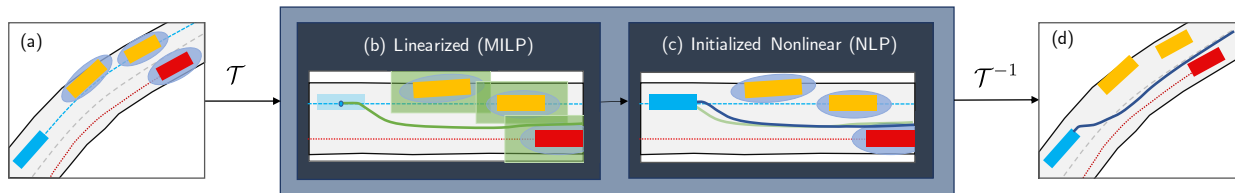


Fig. 2: A Two-Stage Optimization Planner Architecture: (a) from an initial scene - comprising driveable surface limits, static vehicles (yellow), moving vehicles with predicted trajectories (red), and a reference path to follow (light blue) - a transform  $\mathcal{T}$  yields the input to the planner in the reference path-based coordinate frame. The MILP stage (b) solves a linearized version of the planning problem, which initializes (c) the nonlinear, kinematically-feasible NLP stage. Then,  $\mathcal{T}^{-1}$  transforms the output back to a trajectory in the world coordinate frame (d).

produces safe and smooth trajectories when it converges, the authors note some challenges with this approach including the uncertain convergence and/or locally-optimal solutions, even when using state-of-the-art solvers [24]. Due to the fact that nonlinear, non-convex constrained optimization methods are not guaranteed to converge even if feasible solutions exist [20], [25], [26], their safe deployment necessitates the implementation of an additional *fallback mechanism* (e.g. emergency braking [24]), which might fall short in achieving the standard set by the safety hard constraints.

Our goal is to improve the convergence and solution quality of existing nonlinear constrained optimization methods, thus reducing the reliance on the potentially unsafe fallback mechanism and hence improving the safety of the overall system, while yielding lower-cost plans. We achieve this via an informed initialization step. Specifically, we pose the planning problem in terms of a first stage modeled as a Mixed-Integer Linear Program (MILP), whose linearized, non-kinematically feasible output initializes a Nonlinear Programming (NLP) problem, as shown in Fig. 2. The informed initialization offered by the first stage is an approximate solution that, theoretically, is globally  $\varepsilon$ -optimal<sup>1</sup> up to a user-defined receding horizon [28]. This initialization better enables the subsequent nonlinear, non-convex constrained optimization stage to converge to a safe, smooth and feasible trajectory of higher quality in terms of satisfying soft constraints. While this way of improving convergence and achieving higher quality solutions might compromise runtime efficiency, it suits applications that can tolerate longer planning times in return for higher-quality plans — such as in applications of high-fidelity simulation or urban driving.

**Contributions** Our contributions are twofold:

- 1) we formulate motion planning for urban driving in a novel two-stage constrained optimization framework, and
- 2) we evaluate the framework thoroughly in a range of simulated complex urban driving situations to demonstrate empirically that our approach leads to a higher convergence rate and to lower cost solutions when compared with the alternative methods considered.

<sup>1</sup>In practice, modern solver implementations might introduce further limitations to the theoretical guarantee [27].

## II. PRELIMINARIES AND PROBLEM STATEMENT

### A. Notation and Definitions

We use the shorthand  $k \equiv t_k = t_0 + k\Delta t$ , where  $t_0$  is current time and  $\Delta t$  is the timestep. We assume the planning temporal horizon to be  $\tau = N\Delta t$ , defined over  $N$  discrete steps. For any variable  $r_j$  with  $j \in \mathbb{N}^+$ , we use the shorthand notation  $r_{i:e} \equiv (r_i, \dots, r_e)$ . Vectors are in boldface.

We refer to the vehicle we are planning for as the *ego-vehicle* and consider its state at time  $k$  to be given by  $\mathbf{Z}_k = (X_k, Y_k, \Phi_k, V_k) \in \mathcal{W}_{\mathcal{Z}}$ , where  $(X_k, Y_k)$  is position,  $\Phi_k$  is heading, and  $V_k$  is speed, all in a global coordinate frame  $\mathcal{W}$ . The initial ego state at time 0 is the input  $\mathbf{Z}_0$ , while the output of planning is the vector of discrete ego states over the planning horizon,  $\mathbf{Z}_{1:N} \in \mathcal{W}_{\mathcal{Z}}^N$ .

Each other traffic participant (such as other vehicles, pedestrians and cyclists) is represented by a mean pose at time  $k$ ,  $\mathbf{O}_k^i = (X_k^i, Y_k^i, \Phi_k^i) \in \mathcal{W}_{\mathcal{O}}$ ,  $i \in \{1, \dots, n\}$ , and a Gaussian distribution for position uncertainty with mean  $(X_k^i, Y_k^i)$  and covariance  $\mathbf{\Gamma}_k^i$ . The predicted poses and covariances of all traffic participants over the planning horizon ( $\mathbf{O}_{0:N}^{1:n}, \mathbf{\Gamma}_{0:N}^{1:n}$ ) are also inputs to the planning problem.

We define the planning objective to be continuous progress along a differentiable and bounded two-dimensional reference path  $\mathcal{P}_{\text{ref}}$ . The reference path, of length  $|\mathcal{P}_{\text{ref}}|$ , is parameterized by the longitudinal distance from its start  $\lambda \in [0, |\mathcal{P}_{\text{ref}}|]$ , with the points  $(X^{\mathcal{P}_{\text{ref}}}(\lambda), Y^{\mathcal{P}_{\text{ref}}}(\lambda))$ .

To simplify the planning problem, we transform the global coordinate frame  $\mathcal{W}$  to a  $\mathcal{P}_{\text{ref}}$ -based representation  $\mathcal{W}_r$  under the invertible transform  $\mathcal{T}$ , as presented in the next section. Fig. 2 illustrates the steps to transform a problem input ( $\mathbf{Z}_0, \mathbf{O}_{0:N}^{1:n}, \mathbf{\Gamma}_{0:N}^{1:n}, \mathcal{P}_{\text{ref}}$ ) to the desired planning output  $\mathbf{Z}_{1:N}$  using  $\mathcal{T}$ ,  $\mathcal{T}^{-1}$  and our planner.

### B. Reference Path Representation

Given a reference path  $\mathcal{P}_{\text{ref}}$ , we can define its tangential and normal vectors in the global coordinate frame as:

$$\mathbf{t}_\lambda = \begin{bmatrix} \frac{\partial X^{\mathcal{P}_{\text{ref}}}(\lambda)}{\partial \lambda} \\ \frac{\partial Y^{\mathcal{P}_{\text{ref}}}(\lambda)}{\partial \lambda} \end{bmatrix}, \quad \mathbf{n}_\lambda = \begin{bmatrix} -\frac{\partial Y^{\mathcal{P}_{\text{ref}}}(\lambda)}{\partial \lambda} \\ \frac{\partial X^{\mathcal{P}_{\text{ref}}}(\lambda)}{\partial \lambda} \end{bmatrix} \quad (1)$$

The invertible transform  $\mathcal{T}$  operates on the input *poses*, *velocities* and *covariance matrices* as follows.<sup>2</sup>

<sup>2</sup>The derivation of  $\mathcal{T}^{-1}$  is similar to  $\mathcal{T}$ , and thus is omitted for brevity.

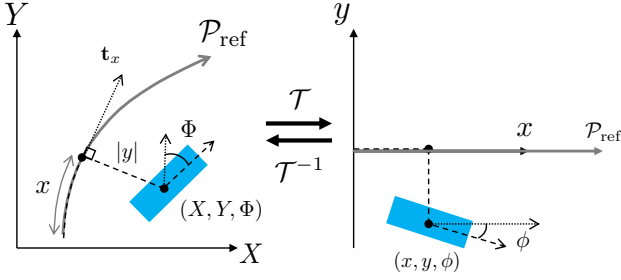


Fig. 3: Visual representation of the pose transform of  $\mathcal{T}$  from the global coordinate frame (left) to the reference path based coordinate frame (right), and the inverse transform  $\mathcal{T}^{-1}$ .

1) *Pose transform*:  $\mathcal{T}$  maps the pose  $(X, Y, \Phi)$  in the global coordinate frame  $\mathcal{W}$  to a pose  $(x, y, \phi)$  in the reference path frame  $\mathcal{W}_r$ , see Fig. 3.

- $x = \text{proj}_{\mathcal{P}_{\text{ref}}}[X \ Y]$  is the distance  $\lambda$  from the beginning of  $\mathcal{P}_{\text{ref}}$  to the projection of  $[X \ Y]$  into it, defined as:

$$x = \underset{\lambda}{\text{argmin}} \quad (X - X^{\mathcal{P}_{\text{ref}}}(\lambda))^2 + (Y - Y^{\mathcal{P}_{\text{ref}}}(\lambda))^2.$$

Due to the nature of the optimization, no closed-form solution can be obtained for  $x$ .

- $y = \frac{1}{\|\mathbf{n}_x\|} \mathbf{n}_x^\top \cdot \hat{\mathbf{y}}$ , where  $\mathbf{n}_x$  is the normal vector of the reference path at  $\lambda = x$  as in (1), and  $\hat{\mathbf{y}} = \begin{bmatrix} X - X^{\mathcal{P}_{\text{ref}}}(x) \\ Y - Y^{\mathcal{P}_{\text{ref}}}(x) \end{bmatrix}$ .
- $\phi = \angle \mathbf{t}_x - \Phi$ , where:

$$\angle \mathbf{t}_x = \arctan \left( \left. \frac{\partial Y^{\mathcal{P}_{\text{ref}}}(\lambda)}{\partial X^{\mathcal{P}_{\text{ref}}}(\lambda)} \right|_{\lambda=x} \right). \quad (2)$$

2) *Speed transform*:  $\mathcal{T}$  is a spatial transformation, so speeds are invariant:  $v = \mathcal{T}(V) = V$ .

3) *Covariance transform*: for a traffic participant with pose  $\mathbf{O}$ , covariance  $\mathbf{\Gamma}$ , and transformed pose  $\mathcal{T}(\mathbf{O}) = [x \ y \ \phi]^\top$ , the transformed covariance matrix is given by:

$$\mathbf{\Sigma} = \mathcal{T}(\mathbf{\Gamma}) = R(\angle \mathbf{t}_x - \phi) \mathbf{\Gamma} R(\angle \mathbf{t}_x - \phi)^\top \quad (3)$$

where  $\angle \mathbf{t}_x$  is as defined in (2), and  $R(\varphi) \in SO(2)$  is the rotation matrix for angle  $\varphi$ .

### C. Problem Statement

The planning problem is defined in the reference path coordinate frame  $\mathcal{W}_r$ . The state of the ego-vehicle at time  $k$  is given by  $\mathbf{z}_k = (x_k, y_k, \phi_k, v_k) \in \mathcal{Z}$  where  $(x_k, y_k) \in \mathbb{R}^2$  is the position,  $\phi_k \in \mathbb{R}$  is its heading, and  $v_k \in \mathbb{R}$  is its speed. The evolution of the state is given by a discrete general dynamical system:

$$\mathbf{z}_{k+1} = f_{\Delta t}(\mathbf{z}_k, \mathbf{u}_k), \quad (4)$$

where  $f_{\Delta t}$  is a discrete, nonlinear function parameterized by  $\Delta t$ , and  $\mathbf{u}_k = (a_k, \delta_k) \in \mathcal{U}$  is the acceleration and steering angle controls applied at time  $k$ . We consider the ego-vehicle to be a rigid body occupying an area  $S_e \subset \mathbb{R}^2$  relative to its center, and denote the area occupied by the ego-vehicle at state  $\mathbf{z}_k$  by  $\mathcal{S}(\mathbf{z}_k) \subset \mathbb{R}^2$ .

For other traffic participants  $i \in \{1, \dots, n\}$ , pose at time  $k$  is given by  $\mathbf{o}_k^i = (x_k^i, y_k^i, \phi_k^i) \in \mathcal{O}$  and position covariance by  $\mathbf{\Sigma}_k^i$ . Following the definition from [24], we denote the area

each traffic participant occupies with probability higher than  $p_\epsilon$  by  $\mathcal{S}^i(\mathbf{o}_k^i, \mathbf{\Sigma}_k^i, p_\epsilon) \subset \mathbb{R}^2$ .

We define the driveable surface area  $\mathcal{B} \subset \mathbb{R}^2$  to be the area in which it is safe for the ego-vehicle to drive in the reference path coordinate frame, and the unsafe area  $\mathcal{B}_{\text{out}} = \mathbb{R}^2 \setminus \mathcal{B}$ . With a cost function  $J(\mathbf{z}_{0:N}, \mathbf{u}_{0:N-1})$  defined over the ego-vehicle's positions and controls, we can now pose the planning problem.

**Problem 1 (Motion Synthesis).** Given an initial ego state  $\mathbf{z}_0$ , and trajectories of other traffic participants  $(\mathbf{o}_{0:N}^{1:n}, \mathbf{\Sigma}_{0:N}^{1:n})$  over the horizon  $N$ , compute:

$$\begin{aligned} & \underset{\mathbf{z}_{1:N}, \mathbf{u}_{0:N-1}}{\text{argmin}} \quad J(\mathbf{z}_{0:N}, \mathbf{u}_{0:N-1}) \\ & \text{s.t.} \quad \forall k \in \{0, \dots, N\} : \\ & \quad \mathbf{z}_{k+1} = f_{\Delta t}(\mathbf{z}_k, \mathbf{u}_k) \\ & \quad \mathcal{S}(\mathbf{z}_k) \cap \mathcal{B}_{\text{out}} = \emptyset \\ & \quad \mathcal{S}(\mathbf{z}_k) \cap \left[ \bigcup_{i \in \{1, \dots, n\}} \mathcal{S}^i(\mathbf{o}_k^i, \mathbf{\Sigma}_k^i, p_\epsilon) \right] = \emptyset \end{aligned}$$

### III. NONLINEAR PROGRAMMING FORMULATION

In this section, we describe a specific solution method to Problem 1, posing it as an NLP problem comprising: (1) *kinematic vehicle model* constraints on the model transitions and allowed controls; (2) *driveable area collision avoidance* constraints; (3) *traffic participants' collision avoidance* constraints; and (4) a *multi-objective cost function* over soft constraints. We will discuss these in order.

1) *Vehicle Model*: We consider a discrete kinematic bicycle model based on the center of the vehicle. Modeling the ego-vehicle as a rectangle with an inter-axle distance  $L$ , we write:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \phi_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \\ \phi_k \\ v_k \end{bmatrix} + \begin{bmatrix} v_k \cos(\phi_k + \delta_k) \\ v_k \sin(\phi_k + \delta_k) \\ \frac{2v_k}{L} \sin(\delta_k) \\ a_k \end{bmatrix} \Delta t \quad (5)$$

We additionally limit the maximum allowed steering  $|\delta_k| \leq \delta_{\text{max}}$  and acceleration  $a_{\text{min}} \leq a_k \leq a_{\text{max}}$ , as well as maximum jerk  $|a_{k+1} - a_k| \leq \dot{a}_{\text{max}} \Delta t$  and angular jerk  $|\delta_{k+1} - \delta_k| \leq \dot{\delta}_{\text{max}} \Delta t$ , to stay within the operational domain and for passenger comfort<sup>3</sup>. We also constrain speed,  $0 \leq v_{\text{min}} \leq v_k \leq v_{\text{max}}$ , to maintain forward motion below the set speed limit.

2) *Collision Avoidance with driveable surface boundaries*: The limits of the driveable surface can include both road limits and roadworks. The driveable surface constraint can be written as  $\mathcal{S}(\mathbf{z}_k) \cap \mathcal{B}_{\text{out}} = \emptyset$ . To define an adequate representation of the area of the ego-vehicle  $\mathcal{S}(\mathbf{z}_k)$  at state  $\mathbf{z}_k$ , we relax the definition of the boundary of the ego-vehicle to only considering its corners. For a rectangular vehicle of width  $w$  and length  $l$ , the positions of the corners at  $\mathbf{z}_k$  are:

$$\mathbf{c}^\alpha(\mathbf{z}_k) = \begin{bmatrix} x_k^\alpha \\ y_k^\alpha \end{bmatrix} = R(\phi_k) \left( \alpha^\top \circ \begin{bmatrix} w/2 \\ l/2 \end{bmatrix} \right) + \begin{bmatrix} x_k \\ y_k \end{bmatrix} \quad (6)$$

<sup>3</sup>The use of the same positive and negative jerk magnitude mirrors that of previous works [24], [29] and was observed to work in practice, however the framework can be trivially extended to use different jerk bounds.

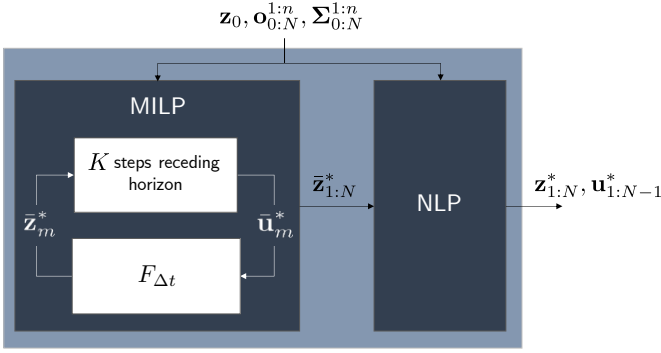


Fig. 4: *Detailed architecture*: the first stage (left) solves a receding horizon formulation of the MILP problem; the second stage (right) uses the solution of the first stage as initialization to solve the full nonlinear, non-convex, constrained optimization problem.

where  $\alpha \in \mathcal{A} = \{[1 \ 1], [-1 \ 1], [-1 \ -1], [1 \ -1]\}$ ,  $R(\phi_k) \in SO(2)$  is the rotation matrix for heading  $\phi_k$ , and  $\circ$  is the element-wise product.

With this, the constraint can be reduced to:

$$\forall \alpha \in \mathcal{A} : \mathbf{c}^\alpha(\mathbf{z}_k) \in \mathcal{B}. \quad (7)$$

We assume the driveable surface  $\mathcal{B}$  is described by two continuous and smooth functions of  $x$ ; the left border  $b_l(x)$  and the right border  $b_r(x)$ , such that  $\forall x : b_l(x) > b_r(x)$ . These boundaries impose a constraint on  $y$ , the lateral path deviation at position  $x$ . Then, (7) can be reduced to the set of constraints

$$\forall \alpha \in \mathcal{A} : b_l(x_k^\alpha) \geq y_k^\alpha \geq b_r(x_k^\alpha), \quad (8)$$

with  $[x_k^\alpha \ y_k^\alpha]^\top$  defined as in (6).

3) *Collision Avoidance with other traffic participants*: We will refer to the traffic participants as vehicles in this section, but other participants (like cyclists and pedestrians) could be handled similarly. We model the area of other vehicles  $\mathcal{S}^i(\mathbf{o}_k^i, \Sigma_k^i, p_\epsilon)$  in a similar fashion to [24]. Assuming  $a_{\text{shape}}$ ,  $b_{\text{shape}}$  to be the parameters of an ellipse  $\mathcal{L}$  that conservatively inscribes the vehicle's shape with  $\Sigma_k^i$  axis-aligned to the vehicle's axis, we write:

$$\mathcal{S}^i \subset \mathcal{L}(a_{\Sigma_k^i} + a_{\text{shape}}, b_{\Sigma_k^i} + b_{\text{shape}}) = \mathcal{L}(a_k^i, b_k^i) \quad (9)$$

where  $\mathcal{L}(a_k^i, b_k^i)$  inscribes vehicle  $i$  at time  $k$  up to an uncertainty  $p_\epsilon$ , see [24] for more details. Thus, we write the constraints:

$$g^{i,\alpha}(\mathbf{z}_k) > 1, \quad \forall \alpha \in \mathcal{A}, \quad (10)$$

$$g^{i,\alpha}(\mathbf{z}_k) = \begin{bmatrix} x_k^\alpha - x_k^i \\ y_k^\alpha - y_k^i \end{bmatrix}^\top R(\phi_k^i)^\top \begin{bmatrix} (a_k^i)^{-2} & 0 \\ 0 & (b_k^i)^{-2} \end{bmatrix} R(\phi_k^i) \begin{bmatrix} x_k^\alpha - x_k^i \\ y_k^\alpha - y_k^i \end{bmatrix}$$

4) *Cost function*: A multi-objective cost function to be minimized is defined over the set of soft constraints  $\mathcal{I}$  on the ego-vehicle's states and controls. For a soft constraint function  $\theta_\iota(\mathbf{z}, \mathbf{u})$  and weight  $\omega_\iota \in \mathbb{R}^+$  reflecting its relative importance, the cost function can be defined as:

$$J(\mathbf{z}_{0:N}, \mathbf{u}_{0:N-1}) = \sum_{k=0}^N \sum_{\iota \in \mathcal{I}} \omega_\iota \theta_\iota(\mathbf{z}_k, \mathbf{u}_k) \quad (11)$$

Soft constraints target various objectives, including:

- **Progress** towards a longitudinal goal  $x_g$ :  $\theta_x = (x - x_g)^2$ , a target speed  $v_g$ :  $\theta_v = (v - v_g)^2$ , or minimizing lateral deviation from the reference path,  $\theta_y = y^2$ .
- **Passenger comfort** defined as minimization of the norm of acceleration,  $\theta_a = a^2$ , and steering,  $\theta_\delta = \delta^2$ .

With the constraints and the cost function defined, we can now formulate the optimization as a constrained nonlinear program.

**Problem 2** (Nonlinear Programming Problem). Given an initial ego-vehicle state  $\mathbf{z}_0$ , trajectories of traffic participants  $(\mathbf{o}_{0:N}^{1:n}, \Sigma_{0:N}^{1:n})$  and soft constraints  $\mathcal{I}$ , compute:

$$\begin{aligned} \underset{\mathbf{z}_{1:N}, \mathbf{u}_{0:N-1}}{\text{argmin}} \quad & J(\mathbf{z}_{0:N}, \mathbf{u}_{0:N-1}) \\ \text{s.t.} \quad & \forall k \in \{0, \dots, N\} : \\ & \mathbf{z}_{k+1} = f_{\Delta t}(\mathbf{z}_k, \mathbf{u}_k) \\ & |\delta_k| \leq \delta_{\max} \\ & a_{\min} \leq a_k \leq a_{\max} \\ & |a_{k+1} - a_k| \leq \dot{a}_{\max} \Delta t \\ & |\delta_{k+1} - \delta_k| \leq \dot{\delta}_{\max} \Delta t \\ & v_{\min} \leq v_k \leq v_{\max} \\ & b_l(x_k^\alpha) \geq y_k^\alpha \geq b_r(x_k^\alpha), \alpha \in \mathcal{A} \\ & g^{i,\alpha}(\mathbf{z}_k) > 1, i \in \{1, \dots, n\}, \alpha \in \mathcal{A} \end{aligned}$$

Due to the nonlinearity of  $J$ ,  $f_{\Delta t}$ ,  $b_l$ ,  $b_r$  and  $g^{i,\alpha}$ , this is a nonlinear, non-convex, constrained optimization problem with equality and inequality constraints. While it is appealing to attempt to solve the problem directly or using a receding horizon formulation as in [24], there are two major challenges:

- **Uncertain convergence**: solvers for this type of problems are generally slow for large instances, especially when initialization is not carefully considered [25]. While advances in efficient primal-dual interior point solvers have mitigated this issue to a certain degree [30]–[32], the convergence to a solution is uncertain [24]–[26].
- **Local optima**: nonlinear constrained optimization solvers tend to be local in nature, finding solutions that are close to the initial point and, possibly, far from the desired global optimum [24], [25], [28].

#### IV. TWO-STAGE OPTIMIZATION

To mitigate the aforementioned issues of NLP, we propose the two-stage optimization framework presented in Fig. 4. The main motivation behind the architecture is to improve the quality of the initial solution provided to the NLP solver, leading to faster and more reliable convergence overall. The informed initial solution is generated through a precursory optimization procedure.

In our two-stage framework, the first stage solves a *linearized* version of Problem 2 formulated as a Mixed-Integer Linear Program (details in Sec. IV-A) in a finite, receding horizon manner. Even though this only gives MILP optimality guarantees for each individual step of the receding horizon (see

the discussion in Sec. IV-A5), it does act as a proxy toward reaching the global optimum of the full linearized problem. The receding horizon process avoids the high runtime that would otherwise be incurred if the MILP problem is attempted with the full  $N$  steps directly.

With the output of the MILP optimization taken as informed initialization, the second stage solves Problem 2. If the representations in the linearized and nonlinear problems are similar, this initialization should improve convergence, speed and the quality of the final solution.

### A. Mixed-Integer Linear Programming Formulation

To reformulate Problem 2 as a MILP problem we (1) consider a linear vehicle model with kinematic feasibility constraints; (2) devise an approach to collision avoidance which maintains the mixed-integer linearity of the model; and (3) adapt the soft constraints to a MILP cost function. We make use of the nonlinear operators  $|\cdot|$  and  $\max(\cdot)$  which we enforce through auxiliary binary variables under the *big-M* formulation [33], [34].

1) *Linear Vehicle Model and Kinematic Feasibility*: The nonlinear kinematic bicycle model presented in Sec. III-1 can be linearized around a point using a series expansion, but this approximation is only valid around the point and yields higher errors as the distance to the point increases. To avoid this issue, we consider a linear, nonholonomic vehicle model instead. We define the state of this linear vehicle model at time  $k$  as  $\bar{\mathbf{z}}_k = [x_k \ y_k \ v_k^x \ v_k^y] \in \mathcal{Z}^M$ , where

$$v_k^x = v_k \cos(\phi_k), \quad v_k^y = v_k \sin(\phi_k) \quad (12)$$

with controls  $\bar{\mathbf{u}}_k = [a_k^x \ a_k^y] \in \mathcal{U}^M$ , and consider a linear vehicle dynamics model defined as:

$$\bar{\mathbf{z}}_{k+1} = F_{\Delta t}(\bar{\mathbf{z}}_k, \bar{\mathbf{u}}_k) = A_d \bar{\mathbf{z}}_k + B_d \bar{\mathbf{u}}_k \quad (13)$$

where  $F_{\Delta t}$  corresponds to a zero-order hold discretization of the continuous state-space system:

$$\dot{\bar{\mathbf{z}}} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \bar{\mathbf{z}}^\top + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \bar{\mathbf{u}}^\top \quad (14)$$

This nonholonomic model is quite simplistic when compared with the kinematic bicycle model in (5), so in order to approximate kinematic feasibility<sup>4</sup>, we add the constraint:

$$v^x \geq \rho |v^y| \quad (15)$$

for a constant  $\rho \in \mathbb{R}^+$ . Assuming forward motion, that is  $\phi_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ , this constraint dictates that any movement in the  $y$  direction requires motion in the  $x$  direction as well.

We also enforce the same constraints as in the nonlinear model, in particular input bound constraints,  $a_{\min}^x \leq a_k^x \leq a_{\max}^x$  and  $a_{\min}^y \leq a_k^y \leq a_{\max}^y$ ; jerk constraints,  $|a_{k+1}^x - a_k^x| < \Delta a_{\max}^x \Delta t$  and  $|a_{k+1}^y - a_k^y| < \Delta a_{\max}^y \Delta t$ ; and speed constraints  $v_{\min}^x \leq v_k^x \leq v_{\max}^x$  and  $v_{\min}^y \leq v_k^y \leq v_{\max}^y$ , with  $v_{\min}^x \geq 0$  to guarantee forward motion.

<sup>4</sup>It should be noted that this model is not kinematically feasible with regard to the bicycle model, only approximately so.

2) *Collision Avoidance*: A key difference of the linear formulation to the nonlinear case with respect to collision avoidance constraints is the lack of explicit representation of orientation in the linearized state  $\bar{\mathbf{z}}$ . A linear approximation of the ego-vehicle's corners would thus induce large errors in the model. We instead consider the ego-vehicle to be a point mass and augment the area it occupies to linearized approximations of the limits of the driveable surface and traffic participants.

For the driveable surface, we formulate the constraint:

$$d + b_l^M(x_k) \geq y_k \geq b_r^M(x_k) - d \quad (16)$$

where  $d : \mathcal{S}_e \rightarrow \mathbb{R}$  is a function of the shape of the ego-vehicle, and  $b_l^M(x)$  and  $b_r^M(x)$  are piecewise-linear functions for the left and right road boundaries such that  $\forall x : b_l^M(x) > b_r^M(x)$ . For a rectangular vehicle with width  $w$  and length  $l$ , the most restrictive approximation of the shape of the ego-vehicle is given by  $d = \sqrt{(w^2 + l^2)}/2$ , limiting the driveable surface to  $\mathcal{B} = \mathbb{R}^2 \setminus (\mathcal{B}_{out} \oplus \mathcal{S}_e)$  where  $\oplus$  is the Minkowski-sum operator. For practical purposes, we consider  $d = w/2$ , which is exact when  $\phi = 0$ .

For traffic participants, we inscribe the ellipses  $\mathcal{L}(a_k^i, b_k^i)$  defined in (9) with axis-aligned rectangles, augmented with the shape of the ego-vehicle. With  $d^x$  and  $d^y$  being functions of the size of the ego-vehicle with respect to its center in the  $x$  and  $y$  direction, respectively, we define rectangles  $\mathcal{R}_k^i$  with the limits:

$$\begin{aligned} x_{k,\min}^i &= \left[ \min_x \mathcal{L}(a_k^i, b_k^i) \right] - d^x \\ x_{k,\max}^i &= \left[ \max_x \mathcal{L}(a_k^i, b_k^i) \right] + d^x \\ y_{k,\min}^i &= \left[ \min_y \mathcal{L}(a_k^i, b_k^i) \right] - d^y \\ y_{k,\max}^i &= \left[ \max_y \mathcal{L}(a_k^i, b_k^i) \right] + d^y. \end{aligned} \quad (17)$$

It should be noted that  $x_{k,\min}^i, x_{k,\max}^i, y_{k,\min}^i, y_{k,\max}^i$  can be computed from  $\mathcal{L}(a_k^i, b_k^i)$ ,  $d_x$  and  $d_y$  in closed form. Then, the collision avoidance constraint is the logical implication:

$$(x_{k,\min}^i \leq x \leq x_{k,\max}^i \wedge y \geq y_{k,\min}^i) \Rightarrow y \geq y_{k,\max}^i \quad (18)$$

which reads *if the ego's position is aligned with a vehicle along  $x$ , then it must be outside the vehicle's borders in  $y$* . Using the *big-M* formulation [33], [34] with a sufficiently large  $M \in \mathbb{R}^+$ , we get the corresponding mixed-integer constraint:

$$\begin{aligned} y_{k,\max}^i - M \mu_k^i &\leq y_k \quad \text{where} \\ \mu_k^i &= \max(x_{k,\min}^i - x_k, 0) + \max(x_k - x_{k,\max}^i, 0) \\ &\quad + \max(y_{k,\min}^i - y_k, 0) \end{aligned} \quad (19)$$

3) *Mixed-Integer Linear Cost Function*: It is imperative that the MILP cost function be similar to the one from the nonlinear stage in order to minimize the gap between the optima of the two stages. With the receding horizon formulation, the cost over each state is also defined independently in time, as, for  $k \in \{0, \dots, N\}$ :

$$J_c^{\mathcal{M},k}(\bar{\mathbf{z}}, \bar{\mathbf{u}}) = \sum_{\iota \in \mathcal{C}} \Omega_\iota \Theta_\iota(\bar{\mathbf{z}}, \bar{\mathbf{u}}), \quad (20)$$

for the set of soft constraints  $\mathcal{C}$ , where each constraint  $\iota$  defines a function to minimize,  $\Theta_\iota(\bar{z}, \bar{u})$ , and a weight  $\Omega_\iota \in \mathbb{R}^+$  to determine relative importance. Similar to the nonlinear stage, soft constraints target various objectives:

- **Progress** towards a longitudinal goal  $x_g$ ,  $\Theta_x = |x - x_g|$ ; a target speed  $v_g$ ,  $\Theta_v = |v - v_g|$ ; and lateral deviation from the reference path,  $\Theta_y = |y|$ .
- **Passenger comfort** as a minimization of the norm of lateral acceleration,  $\Theta_a = |a_y|$ .

4) *MILP Problem Definition*: With constraints  $\mathcal{C}$  and cost function  $J^{\mathcal{M},k}$ , we can now formulate the planning problem as a  $K$ -step receding horizon, Mixed-Integer Linear Program. That is, we obtain  $\bar{z}_{1:N}^*$  by solving  $N - K$  consecutive problems.

**Problem 3** (Receding Horizon MILP). Given an initial ego-vehicle state  $\bar{z}_0$ , trajectories of other traffic participants ( $\mathbf{o}_{0:N}^{1:n}, \Sigma_{0:N}^{1:n}$ ), a set of soft constraints  $\mathcal{C}$ , compute for planning iteration  $0 \leq m \leq N - K$  of the receding horizon:

$$\begin{aligned} \underset{\bar{z}_{m+1:m+K}, \bar{u}_{m:m+K-1}}{\operatorname{argmin}} \quad & \sum_{k=m}^{m+K} J_{\mathcal{C}}^{\mathcal{M},k}(\bar{z}_k, \bar{u}_k) \\ \text{s.t.} \quad & \forall k \in \{m, \dots, m+K\} : \\ & \bar{z}_{k+1} = F_{\Delta t}(\bar{z}_k, \bar{u}_k) \\ & v^x \geq \rho |v^y| \\ & a_{\min}^x \leq a_k^x \leq a_{\max}^x \\ & a_{\min}^y \leq a_k^y \leq a_{\max}^y \\ & |a_{k+1}^x - a_k^x| < \Delta a_{\max}^x \Delta t \\ & |a_{k+1}^y - a_k^y| < \Delta a_{\max}^y \Delta t \\ & v_{\min}^x \leq v_k^x \leq v_{\max}^x \\ & v_{\min}^y \leq v_k^y \leq v_{\max}^y \\ & d + b_l^{\mathcal{M}}(x_k) \geq y_k \geq b_r^{\mathcal{M}}(x_k) - d \\ & y_{k,\max}^i - M\mu_k^i \leq y_k, i \in \{1, \dots, n\} \end{aligned}$$

5) *On the optimality of the MILP stage*: The optimal solution  $\bar{z}_{1:N}^*$  of Problem 3 is used as an initialization to warm-start Problem 2 in order to mitigate the convergence and local optimality challenges discussed above.

A solution to this MILP problem can be obtained using Branch and Bound, a divide and conquer algorithm first introduced and applied to MILP by Land and Doig [35], proven to return the global  $\varepsilon$ -optimal solution [27], [28]. In practice, however, modern solvers (e.g. Gurobi or CPLEX) may fail to find that optimal solution due to rounding errors and built-in tolerances [27]. Moreover, the receding horizon formulation of Problem 3, introduced for the sake of computational tractability, results in suboptimality [36], [37]. Due to these factors, no strong theoretical guarantee can be given regarding the optimality of the MILP stage.

Our hypothesis, nevertheless, is that partial solutions that are close enough to the global optimum for each receding horizon iteration act as proxies towards a final, agglomerated solution that is, when compared with the global optimum, close enough to be useful in initializing the NLP stage in order to improve the quality of the final solution. Our experiments corroborate this hypothesis.

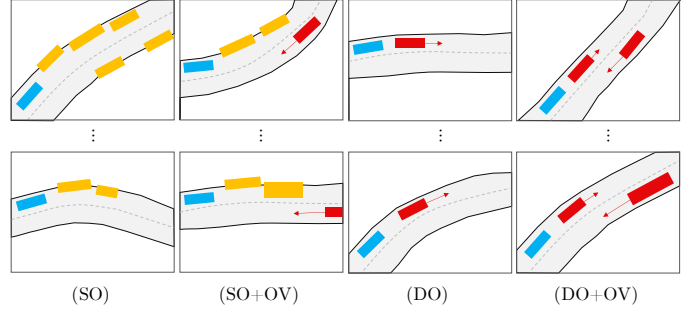


Fig. 5: *Urban Driving Scenario Classes*: with static vehicles in yellow and dynamic vehicles in red, two examples each of: (SO) static overtaking of parked vehicles; (SO+OV) static overtaking with an oncoming vehicle in the other lane; (DO) dynamic overtaking of a slow moving vehicle; (DO+OV) dynamic overtaking of a slow moving vehicle with an oncoming vehicle in the other lane. Parameters, such as the map layout and number of vehicles in an example, are randomly generated.

## V. EXPERIMENTAL RESULTS

In the following subsections, we show that:

- 1) for the NLP stage, our MILP formulation provides a *better initialization* when compared with simpler heuristics and ablations of Problem 3, leading to higher convergence (i.e., percentage of solved cases), faster NLP solving times, and lower-cost solutions, when compared with the NLP stage initialized by those methods (Sec. V-C); and
- 2) our two-stage method *outperforms*, in metrics of progress and passenger comfort, an alternative approach based on a Nonlinear Model Predictive Control (NMPC) scheme similar to [24] (Sec. V-D).

In our experiments, we used state-of-the-art off-the-shelf solvers for the two types of optimization problems, implementing the first stage (MILP; Problem 3) using Gurobi [38] and the second stage (NLP; Problem 2) using IPOPT [39]. Both solvers have a timeout of 25s, after which the optimization is stopped. We use  $N = 40$  and  $\Delta t = 0.2s$  for an output trajectory horizon of 8 seconds. Other parameters of the two problems are listed in Appendix A.

### A. Simulation Environment

Without loss of generality, we assumed left-hand traffic where drivers are expected to be on the left-hand side of the road. We use a proprietary simulator to run the experiments presented in this paper. It uses a fixed simulation frequency of 100Hz, and all agents are simulated using a kinematic bicycle model as described in Eq. 5. Perfect localization and full state observability at simulation time are assumed as sensor data for all agents. Agents also have access to a high definition map with road and lane boundary information, as well as markings, all represented as smooth piecewise linear functions. All non-ego agents implement an Intelligent Driver Model [40]-based decision-making process, allowing them to follow a lane, and perform adaptive cruise control [41] assuming a forward vehicle as well as basic overtaking maneuvers.

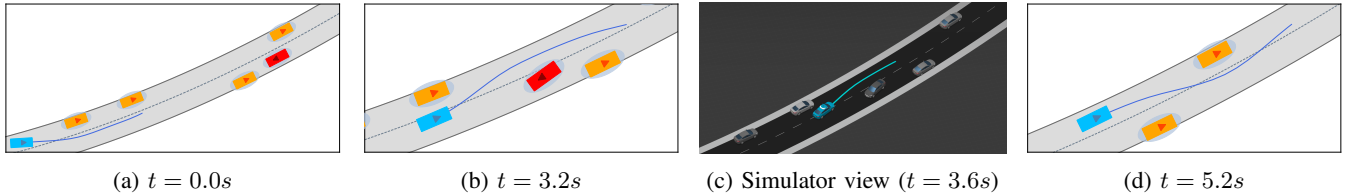


Fig. 6: Residential driving example: (a), (b) and (d) showing the planner’s view of a residential driving problem, with the ego-vehicle (blue), static obstacles (orange), oncoming vehicle (red) and ego’s plan (dark blue trace) at different times  $t$ ; (c) shows our simulator’s rendering of the situation at  $t = 3.6s$ .

|                   |               | (a) Converged (%) | (b) Per instance comparison                  |   | (c) Overall runtime |                                   |                                   |
|-------------------|---------------|-------------------|--|---|---------------------|-----------------------------------|-----------------------------------|
|                   |               |                   | $\Delta_{\text{OURS}}^{\text{NLP}}$ Cost (%) | $\Delta_{\text{OURS}}^{\text{NLP}}$ Runtime (%) | Init (s)            | NLP (s)                           | Total (s)                         |
| <b>Heuristics</b> | ZEROS         | 95.95             | 11.62  | 85.36   | 0.0                 | $0.71 \pm 0.56$                   | $0.71 \pm 0.56$                   |
|                   | CT. DEC       | 93.61             | 9.64   | 21.47   | 0.0                 | $0.44 \pm 0.37$                   | <b><math>0.44 \pm 0.37</math></b> |
|                   | CT. VEL       | 63.88             | 4.11   | 37.97   | 0.0                 | $0.63 \pm 0.89$                   | $0.63 \pm 0.89$                   |
|                   | CT. ACC       | 39.89             | -0.65  | 29.73   | 0.0                 | $0.68 \pm 0.64$                   | $0.68 \pm 0.64$                   |
| <b>MILP-based</b> | NoCOL + NoVEL | 37.49             | 0.27   | 37.66   | $0.14 \pm 0.08$     | $0.70 \pm 0.66$                   | $0.84 \pm 0.67$                   |
|                   | NoCOL         | 42.08             | -1.18  | 43.33   | $0.17 \pm 0.08$     | $0.65 \pm 0.48$                   | $0.83 \pm 0.49$                   |
|                   | NoVEL         | 93.22             | 3.60   | 15.70   | $0.71 \pm 0.35$     | $0.42 \pm 0.46$                   | $1.14 \pm 0.66$                   |
|                   | <b>OURS</b>   | <b>97.76</b>      | —  | —   | $0.55 \pm 0.37$     | <b><math>0.38 \pm 0.42</math></b> | $0.93 \pm 0.60$                   |

TABLE I: *Initialization Ablation*: (a) percentage of problems the NLP stage solves when initialized with the different methods; (b) average percent change in NLP cost and runtime in examples solved by our method and the alternative initialization (positive and higher is worse for the alternative method, better for our method); (c) running time per initialization method, and corresponding NLP stage and total runtime, in the respective subset of examples solved in the dataset by the method (mean  $\pm$  standard deviation).

The ego-vehicle has access to a route planner, which given a goal yields a reference path in the world coordinate frame, and a deterministic constant velocity and lane following predictor for other agents<sup>5</sup>. It plans synchronously at a frequency of 1Hz.

### B. Urban Driving Scenario Classes

Fig. 5 shows examples of the four classes of urban, two-lane driving scenarios we consider for the quantitative experiments in subsections V-C and V-D. Each of the classes presents a set of planning challenges associated with residential driving, among which: smooth overtaking of parked vehicles, negotiating an oncoming vehicle in the other lane, and the decision to overtake a slow-moving front vehicle.

Fig. 6 shows the solution to an example in which the ego-vehicle overtakes vehicles parked at the sides of a two-lane road while taking into account an oncoming vehicle that also has to overtake static vehicles on the other lane. These planning challenges are reflected in many driving situations. For example, we show in Appendix B another example of urban driving that requires a similar level of negotiation with other vehicles: an unprotected junction right turn. Videos of both situations are available in the supplementary material.

To showcase robustness, we perform our analysis over a procedurally generated dataset obtained by randomizing over relevant parameters in each class (e.g., lane widths, initial

speeds or number of vehicles). More details about the procedural generation process are presented in Appendix C.

### C. NLP Initialization Ablation

We considered four NLP stage heuristic initializations as alternatives to our MILP stage:

- ZEROS, in which all states and controls are initialized to zero;
- CT. VEL, in which the ego-vehicle is assumed to maintain a constant velocity throughout the solution;
- CT. ACC, in which the ego-vehicle maintains a constant acceleration of  $1ms^{-2}$  until the maximum speed is reached; and
- CT. DEC, in which the ego-vehicle maintains a constant negative acceleration of  $-1ms^{-2}$  until it stops.

For a more thorough analysis, we also considered three ablations of the MILP formulation presented in Problem 3:

- NoCOL + NoVEL, a similar initialization to the MILP formulation of Problem 3, excluding collision avoidance, speed limit and target speed constraints;
- NoCOL, a similar initialization to the MILP formulation of Problem 3, excluding the collision avoidance constraints with other road users;
- NoVEL, a similar initialization to the MILP formulation of Problem 3, excluding the speed limit and target speed constraints.

We generated a dataset of 4000 examples, 1000 examples per driving scenario class, and solved them using our method as well as with the NLP solver initialized by each of the

<sup>5</sup>While this is the case for the simulation environment in our experiments, we note that our method can include uncertainty-aware predictions following the modeling from [24].



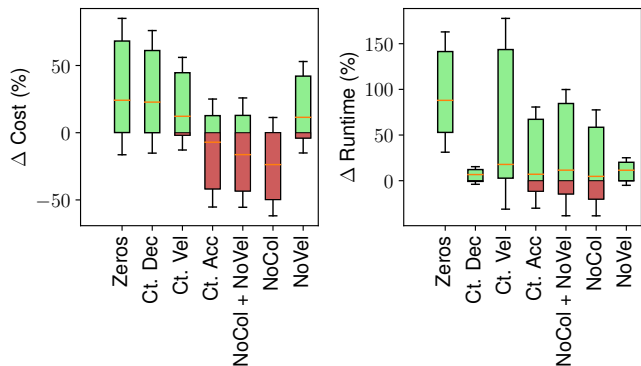


Fig. 7: *Initialization Ablation*: box plots of the percent change in NLP cost (left) and runtime (right) for each method when compared to ours, in examples where both converged but to different optima. Other methods having percent change values larger than zero,  $\Delta > 0$  (indicated by the green color), means better performance by our method.

alternative methods. All the NLP problems optimized the same constraints and cost function.

From Table I (a) we observe our MILP initialization leads to a higher percentage of solved examples than any other method considered. Table I (b) shows a case-by-case comparison between our MILP initialization and the other methods in terms of the average final NLP cost and runtime for the subset of examples solved by both, showing that our MILP initialization achieves better NLP runtimes compared to all the considered alternatives and better costs in most of the cases. To analyze that further, in Fig. 7 we present boxplots of the distribution of the per instance change in NLP cost and runtime of a method compared to ours in examples where both converged to *different* optima. In most cases, our MILP stage leads to NLP solutions that outperform the alternatives in terms of cost, and that converge faster in the second stage. While CT. ACC, NOCOL + NOVEL and NOCOL appear to find better solutions in terms of cost in a non-negligible percentage of the considered examples in Fig. 7, these approaches only converge to a valid solution in less than 43% of the dataset examples as Table I (a) shows and they are slower in general. In Fig. 8, we show a qualitative example from the dataset in which all initializations are successful and our MILP initialization outperforms the alternative heuristics in terms of NLP cost.

Table I (c) presents the running times of the initialization stage and the total runtime for the different methods we consider. Heuristic initializations have a negligible initialization time due to their simple closed-form formulations. Our MILP initialization leads to a lower NLP runtime than all other heuristic and ablation alternatives, indicating a higher-quality starting point that leads to quicker convergence. However, it does incur a higher total runtime than most due to the solving time in the first stage.

#### D. NMPC Baseline Comparison

We compared our method to an NMPC baseline which optimizes Problem 2 but in a receding horizon fashion similar to [24]. This baseline uses the same interior-point optimizer

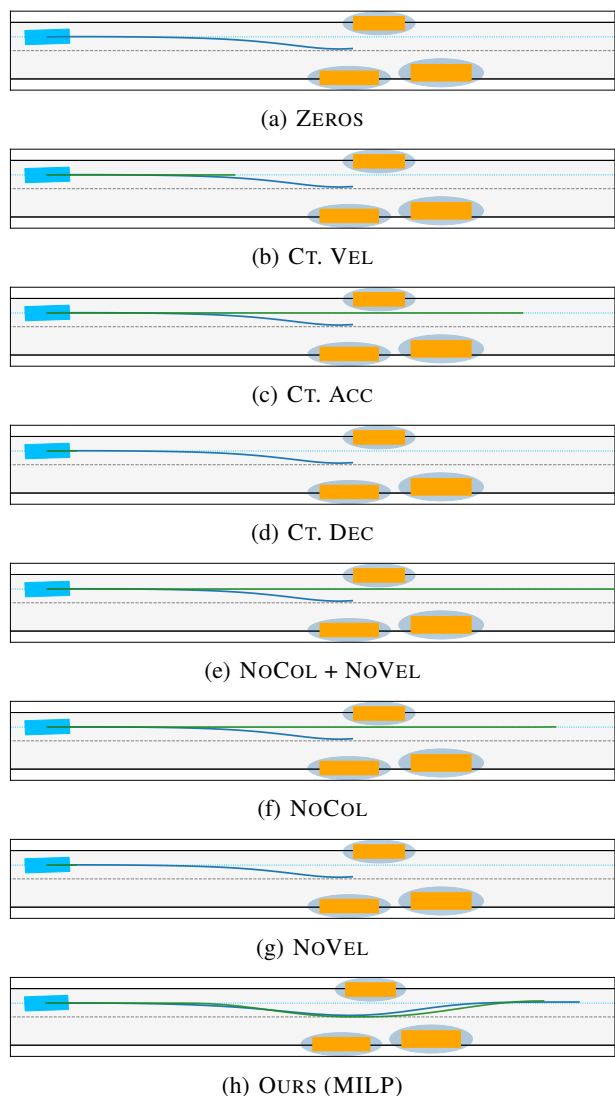


Fig. 8: *Initialization Ablation*: a qualitative example in the reference path coordinate frame of static overtaking from our dataset. The subfigures show the NLP output when initialized with (a)-(d) simple heuristics, (e)-(g) ablations of our MILP stage with subsets of constraints, and (h) our full MILP stage. The initialization trajectory is shown in green and the NLP stage output in dark blue.

as our NLP stage for each of the receding horizon iterations, initializing each iteration with the previous result shifted by one step.

To compare the merits of the two methods, we measured:

- *Solved*: percentage of dataset examples each method solved successfully (higher is better)
- *Progress at 8s* (P@8s): longitudinal distance covered along the reference path by the ego-vehicle within the first 8s of the example, computed over the set of solved examples (higher is better)
- *Speed tracking* ( $v$ ): average instantaneous speed in an example, computed over the set of solved examples (closer to the desired target of 8m/s is better)
- *Absolute jerk value* ( $|\dot{a}|$ ): average absolute value of lon-

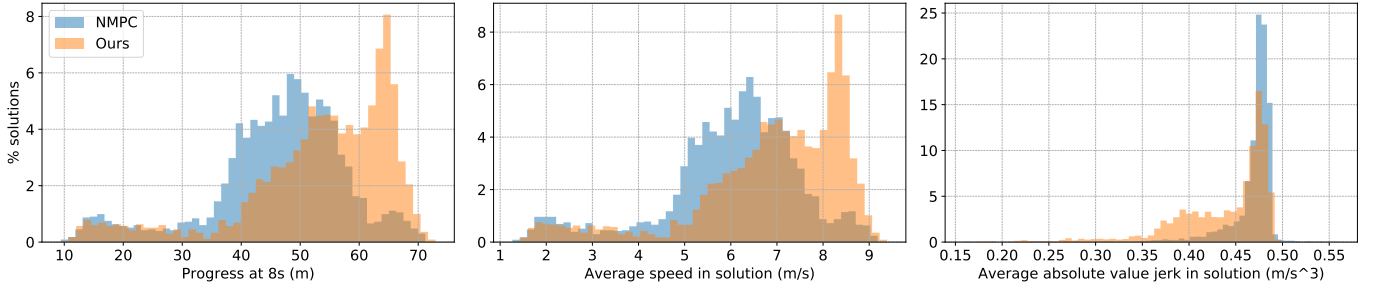


Fig. 9: *NMPC Comparison*: distributions of metric values for the subset of examples solved by both the NMPC (in blue) and our method (in orange). The metrics are: progress in meters after 8s (higher is better); average speed in an example (closer to the target of 8m/s is better); and average absolute value of jerk in an example (lower is better).

|  | NMPC        | OURS               |
|--|-------------|--------------------|
| Solved (%)                               | 87.79       | <b>98.32</b>       |
| Runtime (s)                              | <b>0.56</b> | 0.92               |
| Progress (P@8s in m)                     | 46.24±11.52 | <b>53.18±12.87</b> |
| Speed ( $v$ in m/s)                      | 6.00±1.49   | <b>6.88±1.65</b>   |
| Jerk ( $ \dot{a} $ in m/s <sup>3</sup> ) | 0.47±0.03   | <b>0.43±0.05</b>   |

TABLE II: *NMPC Comparison*: percentage of dataset examples solved, planning runtime, and metrics measured over the set of examples solved by both methods (mean  $\pm$  standard deviation) of progress after 8s (higher is better), average speed in an example (closer to 8m/s target is better), and average absolute value of jerk in an example (lower is better).

itudinal jerk in an example, computed over the set of solved examples (lower is better)

Absolute value of lateral jerk was not considered as part of our comparison since it can be misleading within the residential driving setup due to the fact that any overtaking maneuver will naturally incur higher lateral jerk values compared with situations when the planner does not perform an overtake, and thus it does not always reflect comfort. Longitudinal jerk, on the other hand, is invariant to the maneuver the planner is implementing, and can be considered a comfort metric.

We generated a new dataset of 4000 examples, 1000 per driving scenario class, and solved them using our method and the NMPC baseline.

Table II shows the percentage of examples solved by the methods, as well as the aforementioned metrics for the subset of examples solved by both. Our method outperforms the baseline in terms of the number of solved examples by over 10% (over 400 examples in total). On average, our method achieves significantly higher progress, better speed tracking across scenarios, and smaller longitudinal jerk values. This can also be seen from the distributions of the metric values over the set of examples solved by both methods presented in Fig. 9. With respect to runtime, our implementation of the NMPC baseline using IPOPT took on average 0.56s to generate a solution to each problem, making it nearly 64% faster than our method.

## VI. RELATED WORK

In this section we analyze works related to our method in the literature in terms of definitions of safety, vehicle dynamics

modeling, and other applications of constrained optimization.

### A. Safety in Planning for Autonomous Driving

Within the autonomous vehicles motion planning literature, an accepted definition of safety has been through the concept of the *inevitable collision set* (ICS): the set of states which will inevitably lead to a collision (sometimes also referred to as regions of inevitable collision or target set), where guaranteeing safety can be translated into computing and avoiding the ICS [42]–[47]. However, the computation of the ICS is intractable without some simplifying assumptions [48], which might lead to (i) undesirably-reactive behavior when the system approaches the boundary of the set and (ii) overly-conservative plans in interactive cases [24]. Similarly, reachability analysis methods, which compute the forward reachable set instead of the ICS for the sake of tractability, typically suffer from the same issues [49]–[51].

We follow the approach from [24], in which safe regions are defined by probabilistic constraints on the ego states in the plan. While the chance constraints introduce a conservative factor in the planner when compared to assuming deterministic future positions, the departure from the strict ICS analysis creates an important gap between the executed future states of other agents and the uncertainty-aware predictions. In [24] the authors address this through a high re-planning rate: the frequent update of the planned trajectory as the true states of other agents are observed induces a low uncertainty at the level of the implemented actions at a certain time.

It should be noted that the notions of safety mentioned above are based on the fundamental assumption that the vehicle dynamics model used is representative of the full set of differential constraints of the system, and thus the generated plans are both safe and feasible in the real world [24], [52]. This is not always the case for low-fidelity formulations that use approximate dynamics and operate under model mismatch in which a high frequency feedback controller tracks the generated plan [53]. Several works have shown that high-fidelity closed-form dynamics models can be used reliably in realistic simulation setups [24], [54] and in the real world [55], [56]. For all the experiments in this paper we used a kinematic bicycle model within the NLP problem (including in the NMPC implementation for Sec. V-D), yet the framework can be extended to higher fidelity models such as the one

considered in [24], possibly by considering a different linear vehicle model at the MILP stage (e.g. a first-order hold discretization as in [56]).

### B. Motion Planning via Constrained Optimization

Our framework fits within the constrained optimization literature [52] using a receding horizon [29], [57]. While some related work tackles restricted settings of the problem using linear, mixed-integer linear, or convex program formulations [58]–[60], we directly solve a warm-started, nonlinear problem, handling nearly arbitrary scenarios with complex combinations of dynamic and static vehicles over a long horizon (8s in the reported experiments). Rick *et al.* [20] formulate an NLP problem similar to ours, but with no uncertainty over the predicted positions of road users at the optimization level, and solve it within an NMPC loop initialized with the shifted previous solution. While our planning formulation accounts for state uncertainty, the task of state estimation/filtering [61] is outside the scope of this work.

The approach we take is closest to [24] in terms of the formulation of constraints (following from [21]–[23]), despite the fact that the problem in [24] refers to a corrective system in a semi-autonomous setting as opposed to the fully autonomous driving setting we tackle. Our work can be seen as a receding horizon approach, yet it differs from [24] in that the latter uses the shifted previous solution as the initialization for the NLP problem at each stage within an NMPC framework, whereas ours warm-starts each stage with a fresh MILP solution. Other works have used mixed-integer linear formulations for path planning [60], but, to the best of our knowledge, this is the first work within the autonomous driving literature that uses mixed-integer linear formulations to warm-start an NLP solver.

## VII. DISCUSSION

Safety is paramount for autonomous driving. At deployment time, optimization-based planners require fallback mechanisms to deal with infeasible problems so that a plan is generated at each planning step. This is exacerbated for nonlinear non-convex problem settings as convergence is uncertain, even for feasible problems [24], [25]. As such, a proxy for the deployed system safety could be seen in the percentage of solved problems. Our two-stage planner solves over 10% more examples than the NMPC baseline in our scenario classes, implying that the fallback mechanism would have to be called significantly less frequently than in the case of that baseline.

Furthermore, the presented results validate the claim that our framework with its informed initialization produces solutions of a higher quality than the ones produced by methods that rely on solving the NLP problem directly with simpler initializations or through an NMPC formulation that initializes it with a previous shifted solution.

The analysis of the initialization ablation presented in Table I and Fig. 7 might suggest only marginal improvements by our method in some cases when looked at with the runtime increase (e.g., our solution outperforms ZEROS by a margin of 2% in terms of convergence and 11.62% in terms of cost, but it runs on average 31% slower). However, taking the sequential

nature of the decision making problem into account – i.e., the fact that the planning problems need to be solved continually instead of the one-shot planning reported in Table I – means that these small differences will accumulate over longer runs. This phenomenon is highlighted, for example, by the receding horizon compounding in the NMPC comparison, as observed by the significant overall improvements in convergence and metrics in Table II and Fig. 9.

Generally, the non-convex nature of the NLP problem induces local optima that are highly influenced by the initialization, which, in the case of NMPC, is biased towards following a similar plan to the ones previously computed, and in the case of ZEROS, is biased towards plans closer to the origin in the solution space. The MILP initialization we use, on the other hand, enables the subsequent NLP to explore different parts of the solution space when needed.

Finally, the initialization ablation over the formulation of Problem 3 strongly justifies the inclusion of collision avoidance and speed related constraints in our method, the effects of which can be observed by the increased convergence. As expected, removing these constraints reduces the complexity of the problem, resulting in better initialization runtime. However, the deterioration in the quality of the initialization is also clear from the increased NLP runtime, yielding a comparable total runtime to our method.

## VIII. CONCLUSION

In this paper we introduce a two-stage optimization framework for autonomous driving which first solves a linearized version of the planning problem (formulated as a Mixed-Integer Linear Program in a receding horizon fashion) to warm-start the second nonlinear optimization stage. We show that our MILP initialization leads on average to a higher percentage of solved examples, lower costs, and better solving time for the NLP stage when compared with alternative initializations. Additionally, we show that the two-stage formulation solves more examples than an NMPC baseline, outperforming it in terms of progress and passenger comfort metrics.

By using generic, off-the-shelf solvers, our method trades off runtime for solution quality. One possible direction of future work is to investigate methods that can accelerate the solving time of the first stage, e.g. [62]–[64], in order to achieve faster runtime for real-world planning. This would also have the advantage of allowing for faster re-planning; bridging the safety gap between the uncertainty-aware predictions we use and the actual behavior of the other road users. Nonetheless, we have shown through our experiments that the current framework is still suitable for deployment in complex environments where the quality of the solution is extremely important.

## APPENDIX A OPTIMIZATION PARAMETERS

The parameters of Problem 2 and Problem 3 in the context of the results in Sec. V are defined in Table 1.

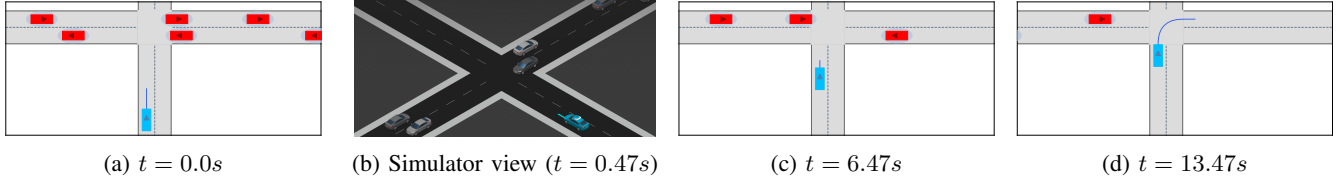


Fig. 10: Junction unprotected right turn example: (a), (c) and (d) showing the planner’s view of a junction handling problem, with the ego-vehicle (blue), dynamic vehicles (red) and ego’s plan (dark blue trace) at different times  $t$ ; (b) shows our simulator’s rendering of the situation at  $t = 0.47s$ .

| Parameter             | Stage | Value | Parameter           | Stage | Value  |
|-----------------------|-------|-------|---------------------|-------|--------|
| $L$                   | NLP   | 4.8   | $a_{\min}^x$        | MILP  | -3     |
| $\delta_{\max}$       | NLP   | 0.45  | $a_{\max}^x$        | MILP  | 3      |
| $a_{\min}$            | NLP   | -3    | $a_{\min}^y$        | MILP  | -0.5   |
| $a_{\max}$            | NLP   | 3     | $a_{\max}^y$        | MILP  | 0.5    |
| $\dot{a}_{\max}$      | NLP   | 0.5   | $\Delta a_{\max}^x$ | MILP  | 0.5    |
| $\dot{\delta}_{\max}$ | NLP   | 0.18  | $\Delta a_{\max}^y$ | MILP  | 0.1    |
| $\dot{a}_{\max}$      | NLP   | 0.5   | $v_{\min}^x$        | MILP  | 0      |
| $v_{\min}$            | NLP   | 0     | $v_{\max}^x$        | MILP  | 3      |
| $v_{\max}$            | NLP   | 10    | $v_{\min}^y$        | MILP  | -1     |
| $\omega_x$            | NLP   | 0.1   | $v_{\max}^y$        | MILP  | 1      |
| $\omega_v$            | NLP   | 2.5   | $\Omega_x$          | MILP  | 0.9    |
| $\omega_y$            | NLP   | 0.05  | $\Omega_v$          | MILP  | 0.5    |
| $\omega_a$            | NLP   | 1.0   | $\Omega_y$          | MILP  | 0.05   |
| $\omega_\delta$       | NLP   | 2.0   | $\Omega_a$          | MILP  | 0.4    |
| $\rho$                | MILP  | 1.5   | $M$                 | MILP  | $10^4$ |
| $d$                   | MILP  | 0.9   |                     |       |        |

TABLE 1: Parameters used in the MILP and NLP optimization stages

## APPENDIX B JUNCTION SCENARIO

Fig. 10 shows the ego-vehicle navigating a junction and performing an unprotected right turn.

## APPENDIX C GENERATION OF DATASET EXAMPLES

Here we give details of the procedural generation of scenarios presented in Sec. V-B.

| Parameter                 | Min                   | Max                   |
|---------------------------|-----------------------|-----------------------|
| Number of lanes           | 2                     | 2                     |
| Lane width (m)            | 3.5                   | 4.3                   |
| Ego initial $x$ (m)       | 0                     | 0                     |
| initial $y$ (m)           | $b_l(x) + 0.55 * 1.9$ | $b_r(x) - 0.55 * 1.9$ |
| initial $v$ ( $ms^{-1}$ ) | 0                     | 9.5                   |
| initial $\phi$ (rad)      | $-\pi/12$             | $+\pi/12$             |

TABLE 1: Common parameters

We assume the ego-vehicle has length  $4.8m$  and width  $1.9m$ , and that the scenario parameters are uniformly sampled from the ranges defined in Tables 1-5.

| Parameter                 | Min      | Max      |
|---------------------------|----------|----------|
| Number of static vehicles | 2        | 6        |
| Static vehicle $x$ (m)    | 0        | 80       |
| $y$ (m)                   | $b_l(x)$ | $b_r(x)$ |
| width (m)                 | 1.7      | 2.5      |
| length (m)                | 4.0      | 8.0      |

TABLE 2: Parameters of scenario SO

| Parameter                        | Min        | Max        |
|----------------------------------|------------|------------|
| Number of static vehicles        | 2          | 6          |
| Static vehicle $x$ (m)           | 0          | 80         |
| $y$ (m)                          | $b_l(x)$   | 0          |
| width (m)                        | 1.7        | 2.5        |
| length (m)                       | 4.0        | 8.0        |
| Oncoming vehicle initial $x$ (m) | 20         | 80         |
| initial $y$ (m)                  | $b_r(x)/2$ | $b_r(x)/2$ |
| initial $v$ ( $ms^{-1}$ )        | 1.0        | 8.5        |
| width (m)                        | 1.7        | 2.5        |
| length (m)                       | 4.0        | 8.0        |

TABLE 3: Parameters of scenario SO+OV

| Parameter                       | Min        | Max        |
|---------------------------------|------------|------------|
| Dynamic vehicle initial $x$ (m) | 20         | 80         |
| initial $y$ (m)                 | $b_l(x)/2$ | $b_l(x)/2$ |
| initial $v$ ( $ms^{-1}$ )       | 0.5        | 3.5        |
| width (m)                       | 1.7        | 2.5        |
| length (m)                      | 4.0        | 8.0        |

TABLE 4: Parameters of scenario DO

| Parameter                        | Min        | Max        |
|----------------------------------|------------|------------|
| Oncoming vehicle initial $x$ (m) | 20         | 80         |
| initial $y$ (m)                  | $b_r(x)/2$ | $b_r(x)/2$ |
| initial $v$ ( $ms^{-1}$ )        | 1.0        | 8.5        |
| width (m)                        | 1.7        | 2.5        |
| length (m)                       | 4.0        | 8.0        |
| Dynamic vehicle initial $x$ (m)  | 20         | 80         |
| initial $y$ (m)                  | $b_l(x)/2$ | $b_l(x)/2$ |
| initial $v$ ( $ms^{-1}$ )        | 0.5        | 3.5        |
| width (m)                        | 1.7        | 2.5        |
| length (m)                       | 4.0        | 8.0        |

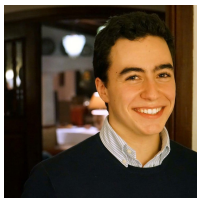
TABLE 5: Parameters of scenario DO+OV

## REFERENCES

- [1] N. Leveson, “Are you sure your software will not kill anyone?” *Communications of the ACM*, vol. 63, no. 2, pp. 25–28, 2020.

- [2] P. Koopman and M. Wagner, "Challenges in autonomous vehicle testing and validation," *SAE International Journal of Transportation Safety*, vol. 4, no. 1, pp. 15–24, 2016.
- [3] S. V. Albrecht and P. Stone, "Autonomous agents modelling other agents: A comprehensive survey and open problems," *Artificial Intelligence*, vol. 258, pp. 66–95, 2018.
- [4] S. V. Albrecht, C. Brewitt, J. Wilhelm, B. Gyevnar, F. Eiras, M. Dobre, and S. Ramamoorthy, "Interpretable goal-based prediction and planning for autonomous driving," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021.
- [5] M. Wulfmeier, D. Z. Wang, and I. Posner, "Watch this: Scalable cost-function learning for path planning in urban environments," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 2089–2095.
- [6] A. Bewley, J. Rigley, Y. Liu, J. Hawke, R. Shen, V.-D. Lam, and A. Kendall, "Learning to drive from simulation without real world labels," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 4818–4824.
- [7] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.
- [8] X. Huang, M. Kwiatkowska, S. Wang, and M. Wu, "Safety verification of deep neural networks," in *International Conference on Computer Aided Verification*. Springer, 2017, pp. 3–29.
- [9] E. Ayers, F. Eiras, M. Hawasly, and I. Whiteside, "PaRoT: A practical framework for robust deep neural network training," *arXiv preprint arXiv:2001.02152*, 2020.
- [10] C. You, J. Lu, D. Filev, and P. Tsotras, "Advanced planning for autonomous vehicles using reinforcement learning and deep inverse reinforcement learning," *Robotics and Autonomous Systems*, vol. 114, pp. 1–18, 2019.
- [11] M. Keller, F. Hoffmann, C. Hass, T. Bertram, and A. Seewald, "Planning of optimal collision avoidance trajectories with timed elastic bands," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 9822–9827, 2014.
- [12] D. Amodè, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in AI safety," *arXiv preprint arXiv:1606.06565*, 2016.
- [13] S. S. Farahani, V. Raman, and R. M. Murray, "Robust model predictive control for signal temporal logic synthesis," in *Conference on Analysis and Design of Hybrid Systems (ADHS)*. IFAC, 2015, pp. 323–328.
- [14] D. Sadigh and A. Kapoor, "Safe control under uncertainty with probabilistic signal temporal logic," in *Robotics: Science and Systems*, 2016.
- [15] M. Hasanbeig, Y. Kantaros, A. Abate, D. Kroening, G. J. Pappas, and I. Lee, "Reinforcement learning for temporal logic control synthesis with probabilistic satisfaction guarantees," *arXiv preprint arXiv:1909.05304*, 2019.
- [16] C. C. Constantinou and S. G. Loizou, "Automatic controller synthesis of motion-tasks with real-time objectives," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 403–408.
- [17] M. Luckcuck, M. Farrell, L. A. Dennis, C. Dixon, and M. Fisher, "Formal specification and verification of autonomous robotic systems: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 5, pp. 1–41, 2019.
- [18] S. A. Seshia, D. Sadigh, and S. S. Sastry, "Towards verified artificial intelligence," *arXiv preprint arXiv:1606.08514*, 2016.
- [19] D. S. Agency, *The Highway Code*. The Stationery Office, 2015.
- [20] M. Rick, J. Clemens, L. Sommer, A. Folkers, K. Schill, and C. Büskens, "Autonomous driving based on nonlinear model predictive control and multi-sensor fusion," *IFAC-PapersOnLine*, vol. 52, no. 8, pp. 182–187, 2019.
- [21] A. Liniger, A. Domahidi, and M. Morari, "Optimization-based autonomous racing of 1: 43 scale rc cars," *Optimal Control Applications and Methods*, vol. 36, no. 5, pp. 628–647, 2015.
- [22] D. Lam, C. Manzie, and M. C. Good, "Model predictive contouring control for biaxial systems," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 2, pp. 552–559, 2012.
- [23] T. Faulwasser, B. Kern, and R. Findeisen, "Model predictive path-following for constrained nonlinear systems," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*. IEEE, 2009, pp. 8642–8647.
- [24] W. Schwarting, J. Alonso-Mora, L. Paull, S. Karaman, and D. Rus, "Safe nonlinear trajectory generation for parallel autonomy with a dynamic vehicle model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 99, 2017.
- [25] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [26] P. E. Gill, W. Murray, and M. A. Saunders, "Snopt: An SQP algorithm for large-scale constrained optimization," *SIAM Review*, vol. 47, no. 1, pp. 99–131, 2005.
- [27] A. Neumaier, "Complete search in continuous global optimization and constraint satisfaction," *Acta Numerica*, vol. 13, pp. 271–369, 2004.
- [28] I. E. Grossmann, V. Voudouris, and O. Ghattas, "Mixed-integer linear programming reformulations for some nonlinear discrete design optimization problems," *Technical Report, Carnegie Mellon University*, 1991.
- [29] J. Wurts, J. L. Stein, and T. Ersal, "Collision imminent steering using nonlinear model predictive control," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 4772–4777.
- [30] A. Domahidi, A. U. Zraggen, M. N. Zeilinger, M. Morari, and C. N. Jones, "Efficient interior point methods for multistage problems arising in receding horizon control," in *51st IEEE Conference on Decision and Control*. IEEE, 2012, pp. 668–674.
- [31] R. J. Vanderbei and D. F. Shanno, "An interior-point algorithm for nonconvex nonlinear programming," *Computational Optimization and Applications*, vol. 13, no. 1-3, pp. 231–252, 1999.
- [32] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Review*, vol. 44, no. 4, pp. 525–597, 2002.
- [33] H. P. Williams, *Model building in mathematical programming*. John Wiley & Sons, 2013.
- [34] J. Omer and J.-L. Farges, "Hybridization of nonlinear and mixed-integer linear programming for aircraft separation with trajectory recovery," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1218–1230, 2013.
- [35] A. H. Land and A. G. Doig, "An automatic method for solving discrete programming problems," in *50 Years of Integer Programming 1958-2008*. Springer, 2010, pp. 105–132.
- [36] X. Geng and Y. Xi, "Suboptimality analysis of receding horizon predictive control with terminal constraints," *IFAC Proceedings Volumes*, vol. 32, no. 2, pp. 2984–2988, 1999.
- [37] L. Grune and A. Rantzer, "On the infinite horizon performance of receding horizon controllers," *IEEE Transactions on Automatic Control*, vol. 53, no. 9, pp. 2100–2111, 2008.
- [38] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2019. [Online]. Available: <http://www.gurobi.com>
- [39] 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.
- [40] 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.
- [41] L. Xiao and F. Gao, "A comprehensive review of the development of adaptive cruise control systems," *Vehicle system dynamics*, vol. 48, no. 10, pp. 1167–1192, 2010.
- [42] N. Chan, J. Kuffner, and M. Zucker, "Improved motion planning speed and safety using regions of inevitable collision," in *17th CISM-IFTOMM symposium on robot design, dynamics, and control*, 2008, pp. 103–114.
- [43] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games," *IEEE Transactions on automatic control*, vol. 50, no. 7, pp. 947–957, 2005.
- [44] A. Bautin, L. Martinez-Gomez, and T. Fraichard, "Inevitable collision states: a probabilistic perspective," in *2010 IEEE international conference on robotics and automation*. IEEE, 2010, pp. 4022–4027.
- [45] D. Althoff, M. Althoff, D. Wollherr, and M. Buss, "Probabilistic collision state checker for crowded environments," in *2010 IEEE international conference on robotics and automation*. IEEE, 2010, pp. 1492–1498.
- [46] T. Fraichard and H. Asama, "Inevitable collision states—a step towards safer robots?" *Advanced Robotics*, vol. 18, no. 10, pp. 1001–1024, 2004.
- [47] A. Blake, A. Bordallo, K. Brestnichki, M. Hawasly, S. V. Penkov, S. Ramamoorthy, and A. Silva, "Fpr—fast path risk algorithm to evaluate collision probability," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 1–7, 2019.
- [48] Y. Chen, H. Peng, and J. Grizzle, "Correct by construction design of autonomous vehicles through a barrier function method," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 4926–4931.
- [49] M. Althoff and J. M. Dolan, "Online verification of automated road vehicles using reachability analysis," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 903–918, 2014.
- [50] S. B. Liu, H. Roehm, C. Heinzemann, I. Lütkebohle, J. Oehlkerking, and M. Althoff, "Provably safe motion of mobile robots in human environments," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 1351–1357.

- [51] X. Wang, K. Leung, and M. Pavone, “Infusing reachability-based safety into planning and control for multi-agent interactions,” *arXiv preprint arXiv:2008.00067*, 2020.
- [52] B. Paden, M. Čáp, S. Z. Yong, D. Yershov, and E. Frazzoli, “A survey of motion planning and control techniques for self-driving urban vehicles,” *IEEE Transactions on intelligent vehicles*, vol. 1, no. 1, pp. 33–55, 2016.
- [53] S. Singh, M. Chen, S. L. Herbert, C. J. Tomlin, and M. Pavone, “Robust tracking with model mismatch for fast and safe planning: an sos optimization approach,” in *International Workshop on the Algorithmic Foundations of Robotics*. Springer, 2018, pp. 545–564.
- [54] H. Andersen, W. Schwarting, F. Naser, Y. H. Eng, M. H. Ang, D. Rus, and J. Alonso-Mora, “Trajectory optimization for autonomous overtaking with visibility maximization,” in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017, pp. 1–8.
- [55] R. Verschueren, S. De Bruyne, M. Zanon, J. V. Frasch, and M. Diehl, “Towards time-optimal race car driving using nonlinear mpc in real-time,” in *53rd IEEE conference on decision and control*. IEEE, 2014, pp. 2505–2510.
- [56] M. Brown, J. Funke, S. Erlien, and J. C. Gerdes, “Safe driving envelopes for path tracking in autonomous vehicles,” *Control Engineering Practice*, vol. 61, pp. 307–316, 2017.
- [57] V. A. Laurence, *Integrated motion planning and control for automated vehicles up to the limits of handling*. Stanford University, 2019.
- [58] P. F. Lima, M. Trincavelli, J. Mårtensson, and B. Wahlberg, “Clothoid-based model predictive control for autonomous driving,” in *2015 European Control Conference (ECC)*. IEEE, 2015, pp. 2983–2990.
- [59] M. Babu, Y. Oza, A. K. Singh, K. M. Krishna, and S. Medasani, “Model predictive control for autonomous driving based on time scaled collision cone,” in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 641–648.
- [60] T. Schouwenaars, B. De Moor, E. Feron, and J. How, “Mixed integer programming for multi-vehicle path planning,” in *2001 European control conference (ECC)*. IEEE, 2001, pp. 2603–2608.
- [61] S. V. Albrecht and S. Ramamoorthy, “Exploiting causality for selective belief filtering in dynamic Bayesian networks,” *Journal of Artificial Intelligence Research*, vol. 55, pp. 1135–1178, 2016.
- [62] D. Bertsimas and B. Stellato, “Online mixed-integer optimization in milliseconds,” *Journal of Optimization Theory and Applications*, vol. 183, no. 2, pp. 490–519, 2019.
- [63] P. Gupta, M. Gasse, E. Khalil, P. Mudigonda, A. Lodi, and Y. Bengio, “Hybrid models for learning to branch,” *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [64] V. Nair, S. Bartunov, F. Gimeno, I. von Glehn, P. Lichocki, I. Lobov, B. O’Donoghue, N. Sonnerat, C. Tjandraatmadja, P. Wang *et al.*, “Solving mixed integer programs using neural networks,” *arXiv preprint arXiv:2012.13349*, 2020.



**Francisco Eiras** received a BSc in Electrical and Computer Engineering from Técnico Lisbon in 2016, and an MSc in Computer Science from the University of Oxford in 2018. He worked as a Research Engineer within the Motion Planning and Prediction Applied Research team at Five for 2 years following his MSc, and is currently pursuing a PhD in Engineering Science at the Optimization for Vision and Learning (OVAL) group at the University of Oxford.



**Majd Hawasly** is a Senior Research Scientist in the Motion Planning and Prediction Applied Research team at Five. He received his PhD from the School of Informatics at the University of Edinburgh in 2014. Prior to his role at Five, he was a postdoctoral research fellow at the University of Leeds.



**Stefano V. Albrecht** is Assistant Professor in Artificial Intelligence in the School of Informatics, University of Edinburgh, where he leads the Autonomous Agents Research Group (<https://agents.inf.ed.ac.uk>). He is a Royal Society Industry Fellow working with UK-based company Five to develop AI technologies for autonomous vehicles. His research interests are in the areas of autonomous agents, multi-agent interaction, reinforcement learning, and game theory, with a focus on sequential decision making under uncertainty. Previously, he was a postdoctoral fellow at the University of Texas at Austin. He obtained PhD/MSc degrees from the University of Edinburgh and a BSc degree from Technical University of Darmstadt.



**Subramanian Ramamoorthy** is a Professor in the School of Informatics at the University of Edinburgh, where he holds the Personal Chair of Robot Learning and Autonomy. He is a Turing Fellow at the Alan Turing Institute, Executive Committee Member for the Edinburgh Centre for Robotics, and a Member of the UK Computing Research Committee. He received his PhD in Electrical and Computer Engineering from The University of Texas at Austin in 2007. He has been a Member of the Young Academy of Scotland at the Royal Society of Edinburgh, and has held Visiting Professor positions at the University of Rome “La Sapienza” and at Stanford University. His research investigates learning, adaptation, and control mechanisms that enable autonomous robots to cope with significant uncertainty and changes in tasks and environments. Between 2017–2020, he served as Vice-President – Prediction and Planning – at Five. He continues to be involved with the company as a Scientific Advisor.